Next Generation Assisting Clinical Applications by using Semantic-aware Electronic Health Records

Pieterjan De Potter, Pedro Debevere, Erik Mannens, Rik Van de Walle
Ghent University - IBBT
Department of Electronics and Information Systems - Multimedia Lab
Gaston Crommenlaan 8 bus 201, B-9050 Ledeberg-Ghent, Belgium
{pieterjan.depotter;pedro.debevere;erik.mannens;rik.vandewalle}@ugent.be

Abstract

The health care sector is no longer imaginable without electronic health records. However, since the original idea of electronic health records was focused on data storage and not on data processing, a lot of current implementations do not take full advantage of the opportunities provided by computerization. This paper introduces the Patient Summary Ontology for the representation of electronic health records and demonstrates the possibility to create next generation assisting clinical applications based on these semantic-aware electronic health records. Also, an architecture to interoperate with electronic health records formatted using other standards is presented.

1 Introduction

The health care sector has been subject to computerization as any other sector. Paper records have been replaced by digital ones, clinical imaging is using digital assisting techniques and the resulting images are saved in digital formats, etc. The first electronic health record systems however targeted only the storage of health information and left too much room for interpretation to be processed by computer systems. These records can thus not be used by next generation assisting clinical applications such as medication conflict detection, chronic disease monitoring, etc.

The IBBT Share4Health project[1] aims at the development of a patient centric health care IT platform that addresses the needs of next generation clinical applications. The project focuses on collaboration between clinicians and assisting them in the decisions they take. One part of the project targets the creation of an ontology that models patient information and the definition of rules to identify clinical conflicts and to summarize relevant patient information. This paper aims at the introduction of the Patient Summary Ontology[2] for the representation of Electronic Health Records.

The remainder of this paper is organized as follows. In Section 2, related work in the area of medical ontologies is discussed. The Patient Summary Ontology is introduced in Section 3. In Section 4, the value of introducing semantic knowledge is illustrated by some use cases. An architecture that makes it possible to map existing health record systems to a semantic-aware system is treated in Section 5.

2 Related work

Semantic-aware representations of data are already used in the medical domain. Examples include the use of ontologies in adaptive questionnaires [5], the representation of the results of such a questionnaire [6], and the modeling of medical knowledge [9]. However, ontologies used in the medical domain model only knowledge from demarcated areas and are only used to solve a specific problem.

The Patient Summary Ontology aims to model the complete, dynamic medical history of a patient, including all procedures and actors involved. In this way, patient data can be used as input by a wide range of assisting clinical applications.

3 Patient Summary Ontology

To enable data processing in assisting clinical applications, the data, which is currently stored in Electronic Health Records (EHRs), needs to be remodeled first. Therefore, the Patient Summary Ontology (PSO) is introduced.

[1] `http://www.ibbt.be/en/project/share4health`

[2] The non-medical part of the Patient Summary Ontology can be retrieved from `http://eulersharp.sourceforge.net/2003/03swap`, the medical part can be found at `http://www.agfa.com/w3c/2009`.

In this section, first the preference for the OWL Full sublanguage is explained. Second, the fundamental principles of the PSO are discussed. Third, external ontologies from which concepts are used in the PSO are examined. Fourth, an overview of person related classes is given.

3.1 OWL Full

The ontology is created as an OWL Full [8] ontology since this introduces no limitations, making the creation a more natural process. The major drawback of this approach is that no reasoning engine will ever be able to provide a complete support for OWL Full [2], so there are still limitations introduced by the reasoning engine. However, it is possible to track these illegal statements and the supported statements are a superset of the statements allowed in OWL Lite and DL.

3.2 Fundamentals

The basic idea of the PSO is not to model the entire medical terminology, but only the general concepts. Details from specific medical domains (e.g., cardiology) can then be modeled in other ontologies, using the PSO as a foundation to build on. The PSO can thus be seen as a top-level ontology restricted to the health care domain.

The PSO consists of a collection of small ontologies, stored in different files. In this way, parts of the PSO can be efficiently reused in other ontologies by selecting the right files and the PSO can be easily extended for specific purposes.

Basic structure The basic structure of the ontology, as shown in Figure 1, is borrowed from the Health Level 7 Reference Information Model (HL7 RIM, [12]). HL7 RIM has four basic classes interacting with each other: Entity, Role, Participation, and Act.

In the PSO, the Entity class has been replaced by foaf:Agent[3] and the Act class has been renamed to Action, according to the naming in the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT, [13]). The participation class is replaced by the pso:actsIn property between the pso:Role and pso:Action classes since in the PSO, the information about this link is provided in the pso:Role and pso:Action classes.

This Agent - Role - Action model is used to allow flexibility: a person who is involved as a physician in a certain action, can be involved as a patient in another action (or even in the same action, in the case of self-examination). Since the semantic approach is about representing as much connections between objects from the real world as possible, we do not want to drop the knowledge that the physician in the first action and the patient in the second action are roles played by the same person.

[3]Since classes from different ontologies are used in the PSO, the "foaf" prefix will be used to indicate classes from the FOAF ontology, the "con" prefix will be used to indicate classes from the contact ontology and the "pso" ontology will be used to indicate classes from the Patient Summary Ontology.

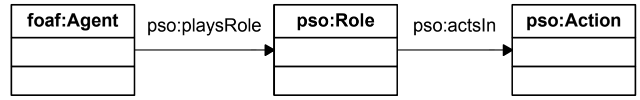

Figure 1. The basic structure of the Patient Summary Ontology

Terminology The terminology used in the PSO is mainly derived from SNOMED CT, a computer-processable structuralized clinical health terminology. It is a controlled terminology, manifested by inheritance relationships between the terms by the definition of "is a"-relations between them.

In SNOMED CT, each term has a unique name and a unique code attached. In the PSO, the unique name is used as a URI for a class, while the code is attached to this class to facilitate possible mappings to other systems or ontologies.

3.3 Integration of other ontologies

To facilitate possible future information integration from other sources, the PSO is built upon parts from other ontologies which are already widely used. In this section, these ontologies are briefly discussed.

FOAF The Friend of a Friend (FOAF, [7]) project is about describing people and the links between them. The FOAF formal vocabulary description has been developed for semantic web use. Although this specification still has a lot of unstable elements in it, its usage has grown with the expansion of social network communities [11].

In the PSO, the FOAF ontology is used to describe entities which are able to play a role. These entities are not limited to persons which can play roles of e.g., health care professionals or patients, but also include other actors such as for example viruses, organizations, or even machines. This interpretation leads to the need for an extension of the FOAF ontology, as illustrated in Figure 2.

Contact To represent the contact information, such as an address or a phone number, of a person who is playing a specific role, the contact [4] ontology is used by specifying the pso:PersonRole class as the intersection of the pso:Role and con:Person classes (see also figure 3). Also from the

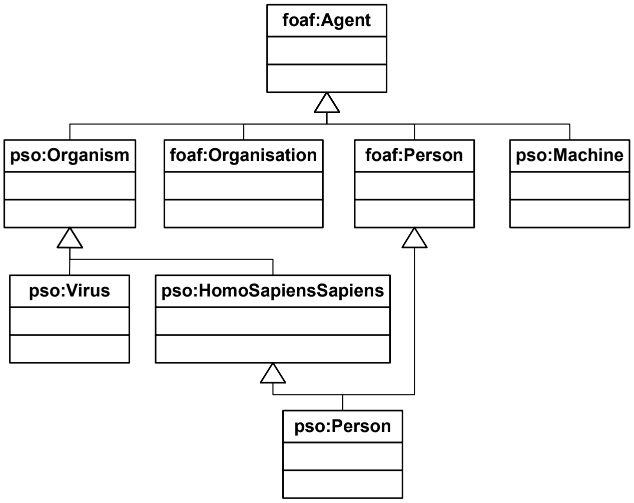

Figure 2. The extension of the FOAF ontology

contact ontology, con:ContactLocation is used for the specification of the seat of an organization in a specific role.

NASA SWEET The NASA Semantic Web for Earth and Environmental Terminology (SWEET, [16]) ontologies provide a semantic framework for projects on the subject of Earth science. In the PSO, the SciUnits ontology from the second version of SWEET is used as a starting point for the modeling of units. Also, the "Country" Class from the geogBorder ontology is used in the definition of the nationality of a person.

3.4 Person overview

An overview of the pso:Person and pso:PersonRole related classes is given in Figure 3. The FOAF and contact ontologies are both involved in the definition of the classes pso:Person and pso:PersonRole. These ontologies are related since foaf:Person is defined as a subclass of con:Person in the FOAF ontology. To maintain the difference between the classes foaf:Agent and pso:Role, pso:Person and pso:PersonRole had to be defined as disjoint classes in the PSO.

4 Semantic knowledge: use cases

The process in which semantic knowledge is added to Electronic Health Records leads to some interesting insights. Even more interesting, though, are the applications that can be created using the semantic knowledge of the PSO. This section contains use cases to show the value of semantic modeling.

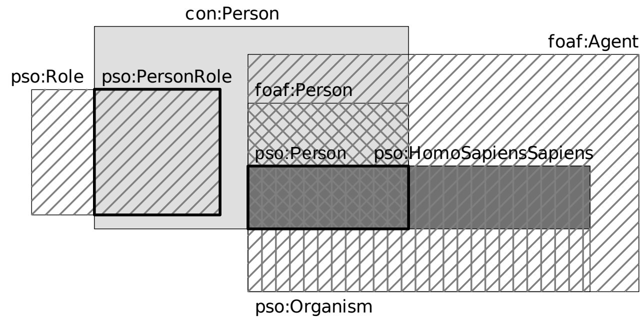

Figure 3. Person overview

4.1 Intelligent patient summary

The medical history of a patient can be very extensive and contain information that is irrelevant in some specific case. Also, medical histories typically keep privacy sensitive patient data. The information represented in a patient summary should thus depend on the situation and the role of a person requesting the data.

By the definition of rules and the usage of a reasoning engine to apply these rules, unwanted information can be filtered out. A first set of rules can be used to filter out data to which the requesting user has no right to view. A second set of rules can be used to filter out irrelevant results and maintain the relevant ones such as details about viral infections in the past, an overview of encounters were a headache was the primary symptom, etc.

4.2 Detection of medication conflicts

When a general practitioner prescribes medication to a patient, he has the time to check for possible conflicts with medication the patient is already taking. In the case of an emergency, one might not have the time to check all possible conflicts manually.

It is possible to describe the conflicts between different medications in rules using semantic concepts of the PSO. Using a reasoning engine to interpret these rules in combination with the medication record of a patient, it is possible to check for medication conflicts in a small time interval before administering medication.

Because the PSO contains knowledge about medication besides the medication records of patients, it is not only possible to describe relationships between particular medications. It is also easily possible to describe relationships between groups of medications, which then automatically apply to the medications in these groups, or to describe medications as being equivalent to each other, or interactions between components of medication.

In case a patient is unconscious, it might be important to have an idea of what prescription free medication (and possibly also illegal drugs) a patient might have taken and the directives of the patient to know which treatments are to be avoided. This makes it important for a patient to make sure his/her medical file is completely up to date and shows that the patients' involvement in EHRs is important.

4.3 Diagnosis support

It is possible to automate the diagnostic process by applying data mining techniques to data aggregated from a specific domain. The purpose of this automation is not to replace the physician, but to discover and report large deviations between the diagnosis made by the physician and the diagnosis predicted by the tool.

Because more information can be made accessible to the diagnosis support system, more factors can be taken into account making the diagnosis. Extra anonymized information sources that could be considered are for example the percentages of diagnosed diseases at other regional general practitioner practices, temporary epidemics at schools, etc.

4.4 Chronic disease monitoring

In case of some chronic diseases, the patient can measure some important indicators himself (e.g., weight, pulse rate) and feed them into a computer. This information can then be sent automatically to a central database.

By defining and applying disease specific rules to the submitted information, an alert with relevant information can be sent to the treating physician if a certain limit has been exceeded (e.g., too much increase in weight in a specified time interval). Also, simple but effective instructions can be given to the patient (e.g., rest, eat some sugar).

5 Interoperability with existing EHR systems

Multiple EHR systems are already in use, which makes interoperability an important issue. In this section, existing exchange standards are examined and the architecture of the part of the system taking care of the interoperability is introduced.

5.1 Exchange standards

In the domain of standardization of exchange documents for EHRs, two areas are to be considered:

- The structuring of documents for the exchange of clinical data. Several standards have been developed in this area: the Clinical Document Architecture (CDA, [10]), the Continuity of Care Record (CCR, [3]), Kind messages for electronic healthcare record (Kmehr-Bis, [1]), etc.
- The coding system used to represent concepts used by exchange formats. Examples are the International Statistical Classification of Diseases and Related Health Problems (ICD, [18]), Logical Observation Identifiers Names and Codes (LOINC, [15]), and the International Classification of Primary Care (ICPC, [17]).

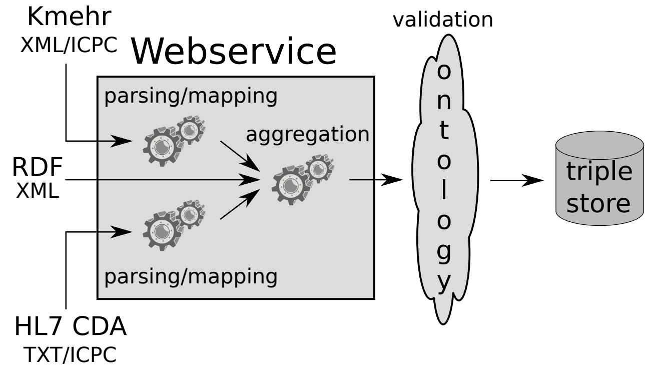

Figure 4. Input architecture

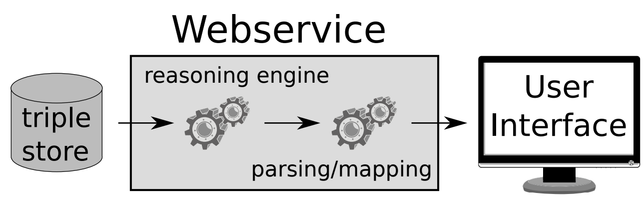

Figure 5. Output architecture

5.2 Architecture

To make it possible for the semantic enabled EHR system to exchange information with other systems, parsing of the exchange standards and mapping of the coding systems which are going to be used for the communication with these other systems are needed. After this transformation, the information is validated against the PSO. Information in RDF/XML [14], modeled using the PSO, can be directly validated against the PSO. After being validated, the information can be stored in a semantic database (a triple store). An overview of the input architecture is shown in Figure 4.

When a health care professional is asking for information (e.g., he/she would like to check for possible medication conflicts), the reasoning engine retrieves the information it needs from the database and presents the result to the end user. An overview of the output architecture is presented in Figure 5.

These transformation and presentation architectures together form a platform which can support next generation clinical assisting tools with patient data stored according to current EHR formats.

6 Conclusions and future work

The Patient Summary Ontology introduced in this paper has the potential to being used as the foundation of advanced clinical assisting tools. The included use cases demonstrate its usefulness. Also, the architecture presented in this paper shows that interoperability with other systems is possible.

Within the scope of the Share4Health project, a technical proof of concept for the detection of medication conflicts and disease diagnosis is being developed. Within the Congestive Heart Failure Project[4], another IBBT project, the usability of the system for monitoring chronic diseases is being tested.

A topic which certainly needs further investigation is how the patients' privacy can be guaranteed with the suggested system, and if another system is needed. Possible scenarios contain the application of privacy restrictions to entire EHRs or only to parts of them, or the use of rules to let the reasoning engine enforce the privacy restrictions.

7 Acknowledgments

The research activities as described in this paper were funded by Ghent University, the Interdisciplinary Institute for Broadband Technology (IBBT), the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), the Fund for Scientific Research-Flanders (FWO-Flanders), and the European Union.

References

[1] Kind messages for electronic healthcare record – Belgian implementation standard. Available from: `http://www.chu-charleroi.be/kmehr/htm/kmehr.htm` Accessed 2009 may 29.

[2] G. Antoniou, E. Franconi, and F. V. Harmelen. Introduction to semantic web ontology languages. In *Reasoning Web, Proceedings of the Summer School, Malta, 2005. Number 3564 in Lecture Notes in Computer Science*. Springer, 2005.

[3] ASTM Committee E31 on Healthcare Informatics. E2369-05, Standard Specification for Continuity of Care Record (CCR). Technical report, Am. Soc. Testing and Materials, 2005.

[4] T. Berners-Lee. Contact: Utility concepts for everyday life, 2008. Available from: `http://www.w3.org/2000/10/swap/pim/contact#` Accessed 2009 may 29.

[5] M.-M. Bouamrane, A. Rector, and M. Hurrell. Gathering precise patient medical history with an ontology-driven adaptive questionnaire. In *CBMS '08: Proceedings of the 2008 21st IEEE International Symposium on Computer-Based Medical Systems*, pages 539–541, Washington, DC, USA, 2008. IEEE Computer Society.

[6] M.-M. Bouamrane, A. Rector, and M. Hurrell. Using ontologies for an intelligent patient modelling, adaptation and management system. In *OTM '08: Proceedings of the OTM 2008 Confederated International Conferences, CoopIS, DOA, GADA, IS, and ODBASE 2008. Part II on On the Move to Meaningful Internet Systems*, pages 1458–1470, Berlin, Heidelberg, 2008. Springer-Verlag.

[7] D. Brickley and L. Miller. Friend of a Friend (FOAF) vocabulary, 2007. Available from: `http://xmlns.com/foaf/spec/` Accessed 2009 may 29.

[8] M. Dean and G. Schreiber. OWL web ontology language reference. W3C recommendation., W3C, Febuary 2004.

[9] R. Dieng-Kuntz, D. Minier, M. Rika, F. Corby, O. Corby, and L. Alamarguy. Building and using a medical ontology for knowledge management and cooperative work in a health care network. *Computers in Biology and Medicine*, 36(7):871–892, July 2006.

[10] R. H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron, and A. Shabo Shvo. HL7 Clinical Document Architecture, Release 2. *J Am Med Inform Assoc*, 13(1):30–39, 2006.

[11] J. Golbeck and M. Rothstein. Linking Social Networks on the Web with FOAF: A Semantic Web Case Study. In *Proceedings of the Twenty-Third Conference on Artificial Intelligence (AAAI'08)*, pages 1138–1143, 2008.

[12] Health Level 7 (HL7). HL7 Reference Information Model. Available from: `http://www.hl7.org/Library/data-model/RIM/modelpage_mem.htm` Accessed 2009 may 29.

[13] International Health Terminology Standards Development Organisation (IHTSDO). *SNOMED Clinical Terms User Guide*, January 2009 International Release edition, January 2009.

[14] G. Klyne and J. J. Carroll, editors. *Resource Description Framework (RDF): Concepts and Abstract Syntax*. W3C Recommendation. W3C, February 2004.

[15] C. McDonald, S. Huff, J. Suico, G. Hill, D. Leavelle, R. Aller, A. Forrey, K. Mercer, G. DeMoor, J. Hook, W. Williams, J. Case, P. Maloney, and Lab LOINC Developers. LOINC, a universal standard for identifying laboratory observations: A 5-year update. *Clinical Chemistry*, 49(4):624–633, April 2003.

[16] R. G. Raskin and M. J. Pan. Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). *Computers & Geosciences*, 31(9):1119–1125, 2005. Application of XML in the Geosciences.

[17] J.-K. Soler, I. Okkes, M. Wood, and H. Lamberts. The coming of age of ICPC: celebrating the 21st birthday of the International Classification of Primary Care. *Family Practice*, 25(4):312–317, 2008.

[18] World Health Organization. *International statistical classification of diseases and related health problems*, 10th revision, 2nd ed. edition, 2004.

[4] http://www.ibbt.be/en/project/chf