Generalized ontology modelling for integration of heterogeneous patient-individual data

Guido Schmidt[1], Keno März[1], Matthias Eisenmann[1], Maximilian Knoll[1], Mohammedreza Hafezi[4], Nassim Fard[4], Ali Majlesara[4], Sepehr Abbasi[4], Maria Maleshkova[2], Darko Katic[3], Stefanie Speidel[3] Arianeb Mehrabi[4], Lena Maier-Hein[1]

[1]Department of Medical and Biological Informatics, Computer assisted interventions, DKFZ, Heidelberg, Germany
[2]Institute of Applied Informatics and Formal Description Methods, KIT, Karlsruhe, Germany
[3]Institute for Anthropomatics, KIT, Karlsruhe, Germany
[4]Department of General, Visceral and Transplantation Surgery, University of Heidelberg, Heidelberg, Germany
{k.maerz|l.maier-hein}@dkfz.de

Abstract. In daily clinical practice vast amounts of patient-specific heterogeneous data including family history, medication, imaging data, diagnostic information etc. is accumulated. Due to the resulting information overload it remains a serious challenge for physicians to choose the best treatment for patients, especially in complicated cases. On the other hand, the immense capabilities of machines in terms of storage and processing power have not yet been fully exploited in this context. Hence, it is desirable to enrich data with semantics to make it understandable and enable accessibility of the data for intelligent information- and decision support systems. Ontological modelling has been successfully used to semantically annotate data in various domains, but lacks adaption to clinical workflows and applications. We attribute this to the high complexity of standard-conform, reusable ontological modelling. To overcome this issue, we introduce a new software framework for building domain-specific medical ontologies in an intuitive manner. In this context, we present a software framework that allows creating medical ontologies that can be linked to existing external ontologies, instantiation of patients with their individual data and building user interfaces (UI) adapted to the needs of clinicians. The framework was successfully applied to the domain of liver tumour treatment planning and will be made open source upon publication of this paper.

1 Introduction

Everyday clinical routine is dominated by vast amount of data generated by a lot of heterogeneous sources, including laboratory examinations, medical imaging and guidelines. Additionally, data is encoded in a heterogeneous manner, for instance in binary files, textual documents, images, tables etc [1]. Another challenge is the large amount of literature published on a daily basis. Klerings et al. [2] state that publishing rates of papers regarding treatment planning are too high to allow a single physician to incorporate them into their decision process in time. Thus, clinicians face information overload and may not be able to choose the optimal treatment for each individual patient based on latest scientific advice[3]. To solve these issues, intelligent systems like clinical decision support (CDS) systems could give treatment advice [4,5] by holistic processing of information from relevant data sources. Therefore, there is a high need of forming structured, semantically annotated data. Ontological modelling addresses this issue and has been applied successfully in other domains, e.g. for enriching hypertext with semantic meaning in the world wide web [6]. Attempting to transfer these concepts to the medical domain often leads to limited acceptance in clinical applications [7]. Approaches have been presented that try to adopt semantic modelling to the medical domain[8], but only focus on subsets of relevant data like the patient medical history. We attribute this to the fact that both the medical domain and the technical domain are highly complex. To create an ontology suited to the physician's needs, experts from both the technical and clinical domain need to work together [9,10], while communication overhead as well as the required understanding of basic ontology concepts form high barriers. Attempting to solve this issue, applications for ontology modelling such as Protégé [11] have been developed. However, they typically provide complex interfaces which require deeper understanding of ontology modelling. Other approaches discuss methods to display relevant information in form of a context-based electronic health record, but so not transferability and adaptability of the model[3]. To overcome these barriers, we propose simplified concepts for building an ontological patient model which integrates all the patient-specific heterogeneous data needed for further treatment planning. In this context, we present the first software framework which features a complete ontology modelling workflow from design to patient instantiation up to user interaction. The framework is scheduled to be open-sourced in June 2016 and is currently used to manage a set of more than 300 patient instances in the Heidelberg university hospital.

2 Materials and methods

Our software framework was designed to support both (1) ontology design as needed to develop a base ontology used by a backend and (2) instantiating patients through a UI frontend (see Fig. 3). The backend manages the ontology through an object oriented representation capable of processing and serializing Resource Description Framework (RDF) files using the Sesame library (`http://rdf4j.org/`) for Java. Interaction with the model is achieved with several services built with a Representational State Transfer (REST) architecture (see Fig. 2). Users can interact with the ontology over HTTP requests. The UI is implemented following concepts from the reactive programming paradigm which promotes building components that react to data flow or events. Our UI component library is built using react.js. This allows building independent, scalable components which can be nested or used standalone. With respect to the ontology modelling, we ensure a design understood by domain experts by reducing complex ontology concepts to main concepts from the domain. Our framework was developed with focus on the following aspects:

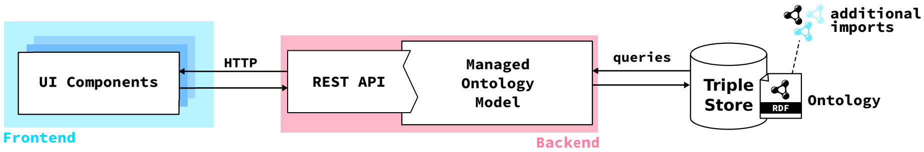

Fig. 1. Design of the proposed system as described in section 2.

Standard compliance To ensure standard compliance, the usage of a standardized upper ontology is crucial. The Open Biological and Biomedical Ontologies (OBO) Foundry (`http://www.obofoundry.org/`) defines principles and standards for working with ontologies. The OBO Foundry has established the Basic Formal Ontology (BFO) (`https://github.com/BFO-ontology/BFO`) [12] as a standard upper ontology for science, especially in the biomedical domain [13]. Because of its high abstraction of general entities and processes existing in nature, it suits the needs for modelling clinical objects and their relations. To represent measured values, the Information Artifact Ontology (IAO) (`https://github.com/information-artifact-ontology/IAO/`) provides concepts to extend the BFO. Our model is integrated with the BFO and the IAO to provide interoperability with other BFO-based ontologies. We abstract the clinical view of things by breaking them down to three major concepts: objects, qualities and measurements, as illustrated in Fig. 2.

Interoperability To ensure interoperability, our system provides interface components to map concepts to unique IDs of standardized clinical nomenclatures and other ontologies, such as SNOMED CT (`www.ihtsdo.org/snomed-ct`) or RadLex (`www.rsna.org/RadLex`) which define a standardized set of vocabulary to represent clinical terms.

Data entry & integrity To ensure data integrity, the possibility to instantiate, edit and change patient data is provided. Furthermore, a history of the measured values is stored using the concept of time-stamped measurements. UI components for data input, as well as browsing data history and marking obsolete values are also implemented.

Reduced complexity To reduce complexity, the system uses explicit mechanisms like hierarchical structures and modularization. Objects and qualities are additionally structured using an independent category tree, which allows reordering and sorting for e.g. workflow optimized views. The system is divided into backend and frontend as illustrated in Fig. 2. The backend provides an object-oriented wrapper of the ontology that guarantees correctness of the generated instances in regard to the ontology. The frontend UI is built using current web technologies and is constructed from structured, simple functional web components. These can be combined individually to build custom UIs.

Cognitive fit To avoid information overload, the UI is designed to suit human perception and cognition by implementing design recommendations as proposed by Horsky et. al [5].

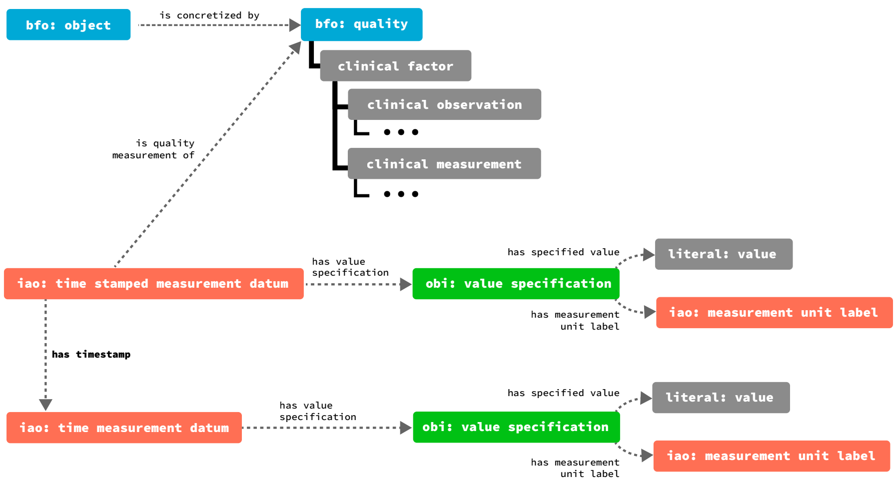

Fig. 2. Main concepts of our modelling approach: objects, qualities and measurements. Objects represent physical entities such as patients, organs, images etc. Qualities describe inherent attributes of these objects. The main parent quality class is a **clinical factor** which distinguishes between **clinical observations** and **clinical measurements**. **Qualities** can be measured, concretizing a quality with a value at a certain point in time. **Clinical measurements** describe numeric parameters, **clinical observations** describe observable attributes and can only assume true, false or left unmeasured. Each **quality** is concretized by **time stamped measurement dates** to describe its value at some point in time. Therefore **time stamped measurement dates** have a **time measurement datum**, which itself is a measurement of the time it was taken. Furthermore, each measurement datum has a list of **value specifications** from the Ontology for Biomedical Investigations (OBI) (`http://obi-ontology.org`) providing the possibility of storing values with different units. For instance a tumor (object) is concretized by its quality volume (**clinical measurement**) of 40 ml (**measurement datum**) on 2014-02-15 15:10:23 (**time stamped measurement datum**).

3 Results

Our concept was successfully implemented, is currently in use in the Heidelberg University Hospital. In the scope of the Collaborative Research Center 125: Cognition-guided surgery, the framework has been applied for liver tumour treatment planning: The resulting knowledge base consists of 2506 classes, 80 object classes, 1741 qualities, and a category tree of 685 categories. The example ontology was instantiated for over 300 liver cancer patients from the University Hospital of Heidelberg.

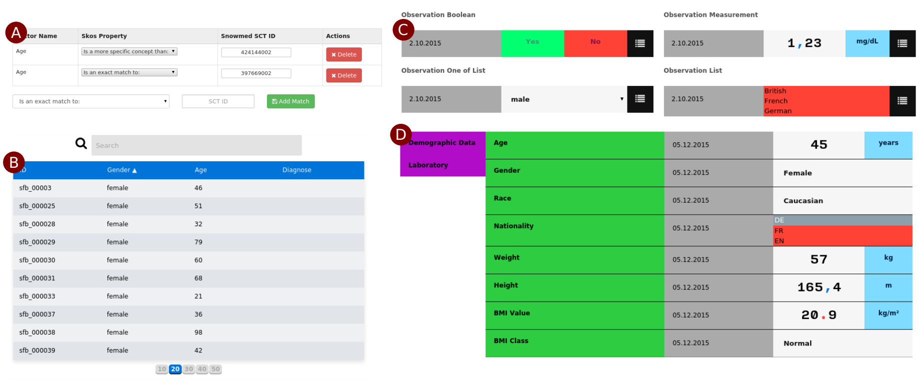

Fig. 3. Selection of exemplary UI components: A) Table for linking qualities to SNOMED CT identifiers, B) Sortable patient selection dialog, C) Different kinds of data input fields for observations, D) Top category flow showing all corresponding qualities for catgeory *Demographic Data*.

4 Discussion

In this paper, we present the first open-source framework that allows clinical domain experts to model and instantiate patients of their respective domain using streamlined ontological concepts in an intuitive and standard conform manner. The presented software components can either be downloaded as an open source package and used standalone or as a complete system. Our system potentially reduces the required amount of ontological modelling expertise and automatically generates a standard-compliant ontology, increasing reusability and extensibility. The created model can immediately be used to instantiate patients with minimal overhead and can be extended manually if required. The data is shareable with all projects adhering to the OBO Foundry standards. We furthermore plan an extensive evaluation of the interface and the interaction of physicians with the information system. As a long term goal we aim to extend the framework with further modules for modeling and decision support in the context of therapy planning. Finally, we intend to open development to the community as an open source project. To our knowledge, this is the first open-source project with a comparable functionality. The first version will be available for download June 2016.

Acknowledgements: This work was carried out with the support of the German Research Foundation (DFG) as part of project A02 of the SFB/TRR 125 Cognition-Guided Surgery, with additional support from project I01.

References

1. März K, Hafezi M, Weller T, Saffari A, Nolden M, et al NF. Toward knowledge-based liver surgery: holistic information processing for surgical decision support. Int J Comput Assist Radiol Surg. 2015;10(6):749–759.
2. Klerings I, Weinhandl AS, Thaler KJ. Information overload in healthcare: too much of a good thing? Z Evid Fortbild Qual Gesundhwes. 2015;109(45):285 – 290.
3. Hsu W, Taira RK, El-Saden S, Kangarloo H, Bui AAT. Context-Based Electronic Health Record: Toward Patient Specific Healthcare. IEEE Transactions on Information Technology in Biomedicine. 2012;16(2):228–234.
4. Schiff GD, Bates DW. Can electronic clinical documentation help prevent diagnostic errors? N Engl J Med. 2010;362(12):1066–1069.
5. Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: A targeted review of best practices for clinical prescribing interventions. J Biomed Inform. 2012;45(6):1202–1216.
6. Tim TBL, Hendler J, et al OL. The Semantic Web. Sci Am. 2001;284(5):28–37.
7. W Ceusters BS, Moor GD. Ontology-Based Integration of Medical Coding Systems and Electronic Patient Records. Med Inform (Lond). 2006; p. 447–475.
8. Bouamrane MM, Rector A, Hurrell M. Using Ontologies for an Intelligent Patient Modelling, Adaptation and Management System. On the Move to Meaningful Internet Systems. 2008;5332:1458–1470.
9. Garshasbi M, Asosheh A, Asadi H. Application and Effectiveness of Ontology on E-Health. In: 2014 7th International Symposium on Telecommunications (IST). IEEE; 2014. p. 544–549.
10. Almeida JS, Chen C, Gorlitsky R, Stanislaus R, de Sousa MA, et al PE. Data integration gets 'Sloppy'. Nat Biotechnol. 2006;24(9):1070–1071.
11. Noy NF, Sintek M, Decker S, Crubezy M, Fergerson R, a Musen M. Creating Semantic Web contents with Protege-2000. IEEE Intell Syst. 2001;16(2):60–71.
12. Smith B. Basic Concepts of Formal Ontology. In: Guarino N, editor. Formal Ontology in Information Systems. IOS Press; 1998. p. 19–28.
13. Spear AD. Ontology for the twenty first century: An introduction with recommendations. Institute for Formal Ontology and Medical Information Science, Saarbrücken, Germany. 2005;.