Representing Clinical Knowledge in Oral Medicine Using Ontologies

Marie Gustafsson, Göran Falkman

School of Humanities and Informatics, University of Skövde, Skövde, Sweden

Abstract:

How can information technology be used to model and handle clinical knowledge in everyday work so that clinicians can more systematically learn from collected clinical data? Ontology is a crucial element in answering this question. Based on nearly ten years of clinical experience, fundamental requirements of an ontology for oral medicine are presented. The use of the proposed W3C standards RDF and OWL for the design and implementation of an ontology for oral medicine is then described and discussed. The reported work contributes to knowledge representation in oral medicine by presenting the pioneering work towards an ontology for oral medicine using RDF/OWL, thereby testing the latter on a new domain.

Keywords:
Oral Medicine; Dental Informatics; Information Dissemination; Evidence-Based Medicine; Internet; Knowledge Representation; Ontology; RDF; OWL

1. Introduction

The MedView computer system [1] is an implementation of evidence-based oral medicine based on the view that health care should be built on finding, validating and using the latest research results as a basis of clinical decisions. In order to provide support for subsequent analysis and validation, key issues in evidence-based medicine are the collection of the ‘right’ clinical data and the possibility of sharing the information generated thereby.

It has been argued that medical informatics is the bridge between the domain of medicine and practitioner psychology [2]. Ontologies [3] are a crucial element in building this bridge, in that ontologies try to capture the best existing medical knowledge, through the systematic and explicit representation of medical definitions, concepts and processes, and by providing methodologies and tools for constructing computer support systems that help clinicians to perform clinical tasks more effectively and efficiently. Over the past decade, there has been much research in using ontologies in different medical disciplines.

Recently, both the Resource Description Framework (RDF) [4] and the Web Ontology Language (OWL) [5] have become World Wide Web Consortium (W3C) recommendations for representing information, exchanging knowledge and for publishing and sharing ontologies. There has been much activity in developing applications and programming frameworks for creating and managing ontologies using RDF/OWL, although, so far, limited research connected to the medical domain has been reported. Despite these current developments, there is to our knowledge no previously reported work on ontology for oral medicine, and especially no work in which RDF/OWL have been used.
We therefore propose the use of ontologies in dealing with problems faced by a system for evidence-based oral medicine such as MedView. Further, we argue that adapting to standards such as RDF and OWL will increase possibilities of knowledge sharing, as well as giving the possibility of using external tools. The reported work contributes to knowledge representation in oral medicine by presenting the preliminary work towards an ontology for oral medicine using RDF/OWL, thereby testing the latter on a new domain.

1.1. MedView

MedView is the product of nearly ten years of cooperation between the Chalmers University of Technology, the Clinic of Oral Medicine at the Sahlgrenska Academy in Göteborg and the University of Skövde. The overall goal is to develop models, methods and tools to support clinicians in their daily work and research. A central question is how information technology can be used to model and handle clinical knowledge in everyday work so that clinicians can more systematically learn from collected clinical data.

The knowledge base built in the MedView project currently contains data from over 8000 clinical examinations, covering more than 2000 different cases. The main knowledge base is located at the Clinic of Oral Medicine at the Sahlgrenska Academy. Clinics within the Swedish Oral Medicine Network (SOMNET) and various non-Swedish clinics have local knowledge bases, which are regularly added to a central knowledge base, so that the entire amount of data collected can be accessed through one common knowledge base.

1.2. Overview

The article identifies some important problems concerning the representation of knowledge in oral medicine. Given these problems, we describe how a preliminary ontology for oral medicine can be designed and implemented, and investigate whether an approach using RDF and OWL might be fruitful in developing this ontology.

2. Knowledge Representation in Oral Medicine

Based on the experience gained during the first decade of the use of MedView and on recent interviews with domain experts and developers, the following requirements of an ontology for oral medicine have been identified:

- We must be able to utilise external sources of knowledge, e.g., general medical vocabularies and taxonomies of diseases and medications. Faster sharing of information is a prerequisite for effective evidence-based medicine [6].
- The relation between the conceptual models of fundamental clinical concepts, e.g., examination templates, lists of approved values for terms and groups of related terms, and their corresponding concrete entities must be formally examined.
- Relations and interactions between different entities of the ontology must be captured, e.g., that a certain answer to a specific question in an examination template triggers another question. By limiting the amount of questions to be answered, a potential barrier to clinicians entering the relevant information is diminished.
- A strong typing of elements is needed. We must be able to enforce that a given term only has values that are, e.g., numeric or a certain enumerated domain. The problem is amplified by the general problem in medicine of agreeing on canonical terms [7].
- We must be able to capture different kinds of meta-data, e.g., who is the creator of a specific examination template and what the purpose (scientific or clinical) is of the introduction of a specific examination template [8, 9].
• We need to differentiate between different ‘views’ of the underlying data, e.g., a patient, time or quantitative oriented view. The provision of different ‘views’ and the manipulation of these views are an intrinsic part of ontology management [10].

• The localisation of data has to be addressed rigorously. How can different language-based versions of the defined concepts, definitions and terms be provided? This is important since the transparent transition between language borders is a presumption of evidence-based medicine and knowledge sharing at a global level.

• An increasingly larger portion of medical data has its origin in images. The enormous amount of information obtainable from images is, however, difficult to grasp for the unaided human mind [7]. Thus, information contained in images, e.g., photos taken during the examination of patients, must be captured and represented.

3. Towards an Ontology for Oral Medicine

We will in the following introduce RDF and OWL and describe the design and the status of the implementation of the SOMWeb (Swedish Oral Medicine Web) ontology.

3.1. OWL and RDF

Essentially, RDF is a data-model, with subject-attribute-object triples, called statements, as its basic building blocks. Statements are graphs, where subjects and objects are nodes linked by attributes as the arcs. The abstract data model of RDF is most commonly described using XML (eXtensible Markup Language). Core concepts of RDF are resources and properties. Resources are the things we want to talk about, e.g., diagnoses, medications and allergies. Properties, a special kind of resources, describe relations between resources.

OWL is designed to be the standardised and broadly accepted language for describing ontologies, allowing users to write explicit, formal conceptualisations of domain models. OWL builds on RDF and uses RDF’s XML-based syntax. Some language elements of OWL are classes, properties and property restrictions. We can state that classes are equivalent, what a property's domain, range and inverse property are, and make restrictions on what values a certain property can take, among other things.

For constructing ontologies, an ontology editor can be used. Of the graphical editors, Protégé [11] is one of the more widely used. Among the plugins available for Protégé are tools for reasoning and visualisation, as well as an OWL-plugin [12]. Further, there are several query languages for RDF, which can be used for finding instances fulfilling criteria of interest, as well as programming frameworks to interact with RDF and OWL content.

3.2. Designing the Ontology

The different ‘views’ supported by the ontology should be chosen based on how well they support real-life clinical tasks [13]. Our MedView experience has shown that the concept of a ‘clinical examination’ is the natural starting point for an ontology for oral medicine. In this respect, our methodology for building the SOMWeb ontology can best be described as a ‘middle-out’ approach [8], based on real-life scenarios, e.g., an examination situation.

3.3. Implementing the Ontology

3.3.1. Examinations, Terms, Values and Classes

In the SOMWeb ontology, OWL classes are used to represent clinical terms, e.g., Allergy, and parts of examination templates, e.g., GeneralAnamnesis. Term values, e.g., Peanuts, are represented as RDF instances of these OWL classes (see upper part of Figure 1 below). A concrete examination record, created by the user filling in a form based on an examination template, is represented as an RDF document, part of which can be seen in the lower part of Figure 1 below.
OWL object properties, e.g., hasAllergy, are associated with terms, where the range gives the term the property it is related to, e.g., Allergy, and the domain gives the part of the examination it is connected with, e.g., GeneralAnamnesis (see Figure 2 below).

```
<rdf:Description rdf:ID="Peanuts">
  <rdf:type>rdf:Description rdf:ID="Peanuts"
  <rdf:type>"somwebOntology#Allergy"
  <rdfs:label xml:lang="en">Peanuts</rdfs:label>
  <rdfs:label xml:lang="sv">Jordnötter</rdfs:label>
  <dc:creator>"somwebPeople#Jontell"
</rdf:Description>

<owl:ObjectProperty rdf:ID="hasAllergy">
  <rdfs:comment xml:lang="en">What allergies do you have?</rdfs:comment>
  <rdfs:comment xml:lang="sv">Är du allergisk mot något?</rdfs:comment>
  <rdfs:range rdf:resource="somwebOntology#Allergy"/>
  <rdfs:domain rdf:resource="somwebOntology#GeneralAnamnesis"/>
</owl:ObjectProperty>
```

Figure 1 – RDF/XML for part of the values for the term Allergy and for part of the GeneralAnamnesis section of the examination record 32_041115

Figure 2 – OWL example for the object property hasAllergy

Classifying related terms is an important way for users to categorise different aspects of the data, e.g., different types of allergies. This can be accomplished using RDF/OWL at different levels. One way is to rigorously define subclasses, e.g., by defining FoodAllergy as a subclass of Allergy. We can also create descriptions of categories, e.g., for smoking habits, where an OWL restriction can be created for what it means to be a heavy smoker.

3.3.2. Using External Sources

An important contribution of using ontologies, and one of OWL’s central features, is the ability to reuse existing ontologies. Currently, we are looking into importing ontologies for countries, units, medications and diagnoses. The Dublin Core, which is an ontology defining a standard for information resource description, will also be used.

3.3.3. Typing of Elements

Several methods in RDF and OWL can be used in aiding the typing of terms and values. Apart from object properties, which relate objects to other objects, OWL also supports data type properties, which relate objects to data type values. A good candidate for a data type property is hasGoodHealth, taking a value of true or false. For OWL data type properties, XML Schema data types can be used, so hasGoodHealth would have range XMLSchema#boolean, as defined by W3C (see lower part of Figure 1).

On the issue of units, OWL and RDF provide good support. We have applied the Semantic Web Best Practices and Deployment Working Group’s working draft for defining n-ary relations\(^2\). For example, smoking habits could be represented as follows:

\(^2\) http://www.w3.org/TR/swbp-n-aryRelations/
<SmokingRelation rdf:ID="Relation_1">
  <smokingValue rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
    4
  </smokingValue>
  <hasTobaccoType rdf:resource="#Cigarette_Filter"/>
  <hasTimeUnit rdf:resource="timeOnt#Day"/>
</SmokingRelation>

This gives a separation of quantities and units, using fictive external time ontology. There are several initiatives for ontologies for units, which the SOMWeb ontology could use.

### 3.3.4. Meta-Data

For managing the multitude of terms, values and templates, using metadata is necessary. The SOMWeb ontology captures who created a term, value or template, and their reason for doing so. The Dublin Core includes elements such as creator, subject and description. In Figure 1, the creator element is used to describe who entered the value described.

### 3.3.5. Localisation of Data

We are especially interested in support for different languages. In the case of a value, e.g., Peanuts, and questions associated with the properties, e.g., hasAllergy, different labels can be specified for different languages using xml:lang (see Figure 1 and 2). Applications will use this to choose either the Swedish or the English labels and questions.

### 4. Discussion

Although ontologies may ease both knowledge sharing and the process of ontology construction, it is not a trivial process. Pinto and Martins [14] give an overview of methodologies for ontology construction and different processes of ontology reuse. Our approach is based on the observation that one success factor for bringing knowledge-based systems into daily clinical use is the careful planning of the introduction of the system in the clinical setting, in cooperation with end-users [15]. Right from the start, MedView was set up as a close collaboration between domain experts and computer scientists, in order to ensure that the developed systems were directly integrated into the clinical ‘workflow’ [17, 15, 9, 16]. Another key success factor is that the resulting system uses ‘real’ clinical data [13], a factor that is definitely met in MedView. Along the lines with previous research, we have started our work on constructing an ontology for oral medicine by focusing on a specific perspective of knowledge in oral medicine [9, 10], namely the concept of an ‘examination’.

Three requirements listed in Section 2 have not been thoroughly explored: In catering for perspectives other than ‘examination’, e.g., ‘patient’ and ‘time’, the possibility to focus on different parts of the graph structure of RDF and OWL could be used. Inverse properties could also be defined to aid in moving in both directions on the graph. For example, the hasAllergy property may need an inverse property allergyOf. For representing relations and interactions between different parts of the model, OWL rule languages, e.g., [18], could be used. When it comes to representing information in photos, there are several initiatives for annotating photos using RDF/OWL. Using one of these, information ranging from simple descriptors referring to the diagnosis and quality of the photo, to marking regions of interest could be added. Another aspect we wish to explore is user modelling, needed to support users of different familiarity with computers, different roles and working in different departments. This can be a way to cope with the problem of getting the end-user to feel comfortable with the application, meaning a greater chance of data being entered [19].

### 5. Conclusions and Future Work

We have identified requirements faced in representing knowledge in oral medicine. Key issues here are the adherence to proposed standards, the utilisation of external data sources, and the amplitude of conceptual entities that have to be formally represented. We have de-
scribed how these requirements can be fulfilled by constructing an ontology implemented in RDF and OWL. We have also begun identifying ontologies of interest for reuse.

During the initial implementation of this ontology, it has become apparent that RDF and OWL are applicable in fulfilling the requirements identified in Section 2. This work will be continued with further modelling of examinations and related concepts. Ontologies to be imported will be identified and hopefully successfully incorporated. As a case study, an existing MedView application will be adapted to using RDF and OWL, though the use of RDF and OWL will not be apparent to the end user. The new application will be tested as part of a more general evaluation of the RDF/OWL approach. Although reasoning is not our focus at present, advances on reasoning about instances may be relevant in the future.

As RDF and OWL are becoming widely used, we propose that these be used in the implementation of the SOMWeb ontology. By applying international standards, sharing knowledge and connecting ontologies should be achieved more easily. Further, the array of existing development tools will aid us in this implementation, and ensure that other ontology creators will have an equivalent ease of development.

6. References


Address for Correspondence
Marie Gustafsson, School of Humanities and Informatics, University of Skövde, PO Box 408
SE-541 28 Skövde, Sweden. Email: marie.gustafsson@his.se