

The Problem of Retrieving Clinical Guidelines in Self-Care

Juha Puustjärvi
Innomedi
Helsinki, Finland
email: juha.puustjarvi@cs.helsinki.fi

Leena Puustjärvi
The Pharmacy of Kaivopuisto
Helsinki, Finland
email: leena.puustjarvi@kolumbus.fi

Abstract— Self-care is what people do for themselves to establish and maintain health, prevent and deal with illness. A clinical guideline is a document with the aim of guiding decisions and criteria regarding diagnosis, management, and treatment in specific areas of healthcare. Clinical guidelines have been in use for thousands of years during the entire history of medicine. An interesting question arising from these notions is whether clinical guidelines can be used in self-care to ensure that the treatment is in line with the medical recommends. A problem, however, is that the retrieval methods of clinical guidelines are mainly aimed at healthcare personnel who are taught to retrieve relevant guidelines. In order to simplify patients' tasks in retrieving guidelines, we have designed an ontology (vocabulary) for annotating clinical guidelines. A patient can then query clinical guidelines by keyword expressions, which are intuitive and clear for patients. Technically annotations are presented according to the Semantic web technologies, and so they are machine understandable. Further, guidelines can be retrieved from a variety of sources within one query. In addition, the system can be easily extended such that guidelines and other relevant health-oriented information can be retrieved in a query.

Keywords – *Self-Care; Clinical Guidelines; Evidence-Based Medicine; Semantic Web; SPARQL*

I. INTRODUCTION

Nowadays, health care provision is moving towards self-care from the physician centric model where the treatment decisions are made almost exclusive by physicians. This is a challenging trend as it is estimated that 70% to 95% of all illnesses could be managed without the intervention of a doctor [1], i.e., by self-care [2].

To promote self-care some Personal Health Records (PHRs) already provide links to websites that provide information concerning patient's treatment and medication. However, many patients have regarded these sites to be overly commercial, and often patients cannot determine the source of the published medical information [3]. In addition, browsing these pages has turned out to be long lasting and frustrating as their provided medical information is not relevant, is overly specific for patients or is overly superficial [4].

After all, success in self-care requires patients to have relevant medical treatment information [5]. Traditionally, such information is provided for physicians in clinical guidelines, which are document with the aim of guiding decisions and criteria regarding diagnosis, management, and treatment in specific areas of healthcare [6].

Clinical guidelines are based on an examination of current evidence within the paradigm of *evidence-based medicine* (EBM) [7]. It emphasizes the use of evidence from well designed and conducted research in healthcare decision-making [8]. EBM is one of the most significant developments in the clinical use of information over the last decades [9]. Practicing EBM represents a thorough change in the way that physicians are taught to think and retrieve relevant guidelines [10].

There are a variety of ways how the digital volumes of guidelines can be organized. For example, guidelines may be grouped in an alphabetic order of the disorders they deal with, grouped by specialities, or grouped by body parts. Accessing guidelines from such digital volumes have proven to be easy for physicians and healthcare professionals as healthcare providers are obliged to know how to find relevant guidelines.

Instead searching relevant guidelines from such volumes is not necessary a trivial task for patients. This is regrettable as clinical guidelines are intended also for patients' use. Furthermore, clinical guidelines have special versions for different user groups.

Our argument is that clinical guideline oriented services should also provide specific retrieval methods for patients, as they are not taught to think in accordance with EBM. For example, a pregnant patient may be interested to know whether she has some risks in using pain drugs. In this case, finding relevant information from guidelines dealing with pregnancy and from guidelines dealing with pain drugs may be long lasting and frustrating process.

In order to develop appropriate retrieval methods for patients, we have annotated clinical guidelines by appropriate metadata items such as keywords and classification identifiers. A patient can then query clinical guidelines by Boolean expressions [11] comprising of operands and operations. The operands are the used keywords and the operands are typically "and", "or", and "not". The Boolean model is intuitive and clear. For example, the query "pregnancy and pain drugs" returns the guidelines (if any) that are annotated by these keywords.

Another problem concerning the retrieval of clinical guidelines is that they are fragmented. That is, there are many sites publishing guidelines, and therefore searching relevant guidelines from a variety of sites is often hard and long lasting. In order to avoid such weakness, a key point in our solution is that in one query patient can retrieve guidelines located in many web sites. This, however, requires that all the sites use the same vocabulary (ontology) and technology in annotating clinical guidelines.

Technically, in our solution, each site is assumed to provide a web service that supports SPARQL queries. SPARQL is a query language developed for Semantic web [12].

The rest of the paper is organized as follows. First, in Section II, we shortly present the architecture of the system that we are developing. Then, in Section III, we consider the Guideline ontology that we use in annotating guidelines. In Section IV, we give an example of a SPARQL query that exploits the Guideline ontology in searching guidelines from two data stores. In Section V, we consider the extension of the system by medicinal data. Finally, Section VI concludes the paper by considering our future research.

II. THE ARCHITECTURE OF THE GUIDELINE SERVER

In our developed architecture, users (patients) retrieve clinical guidelines through the Clinical Guideline Portal (Fig. 1). The portal provides a set of applications each providing a specific search method, such as searches based on keywords or body parts. These applications communicate with the SPARQL Processor, which access clinical guidelines through their Web Services.

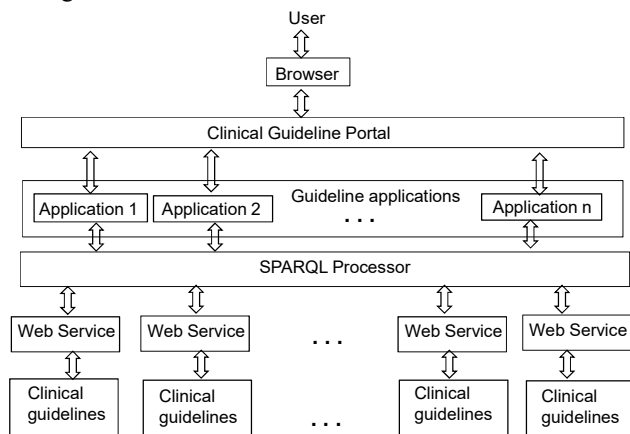


Figure 1. The architecture of the Clinical Guideline Retrieval system.

A useful feature of the SPARQL Processor is that it is capable for querying many RDF-formatted [13] data sources within one SPARQL query.

III. GUIDELINE ONTOLOGY

An ontology characterizes the meaning of concepts and their relationships [14]. Within computer science, an ontology is usually represented by using classes, properties, and their values as modeling primitives [15]. Hence an ontology provides a systematic way to standardize the used metadata items [16]. As an example consider our defined Guideline Ontology, which is graphically presented in Fig. 2.

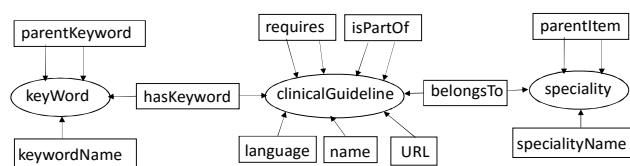


Figure 2. Guideline Ontology.

In this figure, ellipses represent classes and rectangles represent data type as well as object properties. These representations are modeling primitives in OWL [17]. Object properties (e.g., belongsTo) relate objects to other objects (or to itself, e.g., requires), and datatype properties (e.g., specialityName) relate objects to datatype values.

An instance of the Guideline Ontology is presented in RDF in Fig. 3. RDF is a key for representing machine understandable data. It is a data model with a variety of syntaxes for storing data files. By RDF we can express facts with tree-part statements called as triples. The subject identifies the thing being described, predicate is a property name, and object is property value. That is, each triple is like a little sentence that states a fact. However, RDF in itself does not bring interoperable semantics. It depends on the expression power of the used vocabulary. By a vocabulary we refer to a set of ontologies, which specifies the used terms and their semantics.

In order to illustrate the dependency of RDF and ontologies consider the RDF-description (a set of RDF-statements concerning the same subject) of Fig. 3. The RDF-description states, by using the Guideline Ontology as a vocabulary, that a guideline named *Pregnancy* has keywords *pregnancy* and *pain drugs*, and the guideline is a part of a guideline named *gynaecological diseases*. That is, pregnancy is a subclassification of gynecological diseases.

```
<rdf:RDF
  xmlns : rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns : no="http://www.helsinki.fi/GuidelineOntology#"
  <rdf:Description rdf:about="pregnancyGuideline">
    <rdf:type rdf:resource="&po;clinicalGuideline"/>
    <no : hasKeyword>pregnancy</no :hasKeyword>
    <no : hasKeyword>pian_drugs</no :hasKeyword>
    <no : isPartOf>gynaecological diseases</no:isPartOf>
  </rdf : Description>
</rdf:RDF>
```

Figure 3. An annotation of a guideline.

The Guideline Ontology enables to attach parenKeyword to each keyword meaning that keywords comprise taxonomies (hierarchies). The logic behind taxonomy is that when one goes up the taxonomy toward the root, the keywords become more general, and respectively when one goes down towards the leaves the keywords become more specialized [16]. We can also state this in a more formal way: depending on the direction of the link each link between a parent and a child node represents a subclassification relation or superclassification relation. For example, type 2 diabetes is a subclassification of diabetes, and diabetes is a superclassification of type 2 diabetes.

Taxonomies can be exploited in searching, if the used keywords return too many guidelines. For example, we could replace keyword *diabetes* by its subclassification *type 1 diabetes*, *type 2 diabetes*, or *gestational diabetes* (a condition in which women without previously diagnosed diabetes exhibit high blood glucose levels during pregnancy).

IV. QUERING GUIDELINES FROM MULTIPLE DATASTORES BY SPARQL

SPARQL is a query language that is able to retrieve and manipulate data stored in RDF-format. It was made a standard by the RDF Data Access Working Group of the World Wide Web Consortium, and is recognized as one of the key technologies of the semantic web [18].

A SPARQL query specifies the pieces of data that meets the stated requirements. The requirements are described with triple patterns, which are similar to RDF triples but they include variables to add flexibility in how they match against the queried data [18].

There is a variety of SPARQL technologies available for running queries against data both locally and remotely. For example, SPARQL provides two ways for querying remotely: using FROM keyword or using SERVICE keyword. In the former way the FROM keyword names a dataset to query that may be local or remote file. In the latter and our used way, instead of pointing at an RDF file somewhere, a SPARQL endpoint is pointed. An *endpoint* is a Web service that accepts SPARQL queries, runs the queries, and finally returns the result.

Further, Federated Queries in SPARQL allow searching multiple datasets with one query. For each dataset is created a subquery which access datasets by using SERVICE keywords. To illustrate this, consider the federated SPARQL query presented in Fig. 4. It accesses two datasets through SPARQL endpoints. The result of the query is the union of the results of the two subqueries. The query returns links (URLs) to those clinical guidelines that are annotated by the keywords *pain drug* and *pregnant*. The query is based on the ontology presented in Fig. 2 (the prefix *no* in the query refers to this ontology).

```

PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX no: <http://www.cs.helsinki.fi/registryOntology#>
SELECT ?URL
WHERE
{
  SERVICE <http://documentRegistry_A/sparql>
  { SELECT ?URL
    WHERE
    {
      ?no:name no:hasKeyword pain_drug ;
      ?no:name no:hasKeyword pregnant ;
      ?no:name no:URL ?URL .
    }
  }
  SERVICE <http://documentRegistry_B/sparql>
  { SELECT ?URL
    WHERE
    {
      ?no:name no:hasKeyword pain_drug ;
      ?no:name no:hasKeyword pregnant ;
      ?no:name no:URL ?URL .
    }
  }
}

```

Figure 4. A simple federated SPARQL query.

V. EXTENDING THE GUIDELINE SERVER

The Clinical Guideline Retrieval System presented in Section II is focused only on retrieving relevant guidelines. However there are also websites that have information about medication, and thus they can provide valuable supplementary information for self-care. For example, a patient may be interested to know more about a pain drug recommend in a clinical guideline.

A problem however is that most health-oriented web sites are developed only for human consumption, and so they as well as their metadata is not machine understandable. In order to transform their metadata in machine understandable format, we have developed an appropriate ontology (called Medicine Ontology) for annotating the documents that deal with medicine.

A portion of the Medicine Ontology is graphically presented in Fig. 5. In this graphical representation ellipses represent classes and rectangles represent data type and object properties. Object properties relate objects to other objects while data type properties relate objects to datatype values. Classes, data type properties and object properties are modelling primitives in OWL.

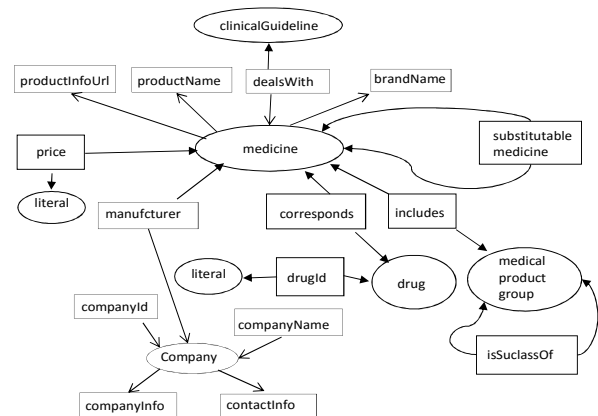


Figure 5. Graphical presentation of the Medicine Ontology.

Note that the Medicine Ontology and the Guideline Ontology are overlapping in the sense that they both include the class *clinicalGuideline*. As a result we can use the Medicine Ontology in annotating clinical guidelines by medicinal data. For example, in Fig. 6, the annotation presented in Fig. 3 is extended by stating that the instance deals with Aspirin.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:no="http://www.helsinki.fi/GuidelineOntology#"
  xmlns:ro="http://www.helsinki.fi/MedicineOntology#"
  <rdf:Description rdf:about="pregnancyGuideline">
    <rdf:type rdf:resource="&po;clinicalGuideline"/>
    <no:hasKeyword>pregnancy</no:hasKeyword>
    <no:hasKeyword>pian_drugs</no:hasKeyword>
    <no:isPartOf>gynaecological diseases</no:isPartOf>
    <ro:dealsWith>Aspirin</ro:hasKeyword>
  </rdf:Description>
</rdf:RDF>

```

Figure 6. Using two ontologies in annotating a clinical guideline.

Similar to the extension on Medicine ontology we can also introduce other ontologies as well. For example, introducing an ontology for the terms of welfare might also be useful.

VI. CONCLUSION

Web-based e-health models and the Semantic Web support each other: Web-based health-oriented services provide a new paradigm for sharing health information while Semantic Web enables the presentation of information in a machine understandable form. As a result the use of Web-based e-health services is rapidly increasing. Yet, by using together a variety of services we can still achieve new services that would not be achievable by independently working systems. For example, we can achieve new outcomes by the interoperation of the systems developed for self-care, welfare, and smart homes. Technically the interoperation can be implemented by a SPARQL processor.

So far we have used these technologies in developing a patient-friendly retrieve method for clinical guidelines. We have also extended this solution by the services that provide information about medicines. In our future work we will consider how the information of clinical guidelines, medication, welfare and smart home can be used together in developing new services.

REFERENCES

- [1] "The Role of the Pharmacist in Self-Care and Self-Medication". Available at: <http://apps.who.int/medicinedocs/pdf/whozip32e/whozip32e.pdf>. Retrieved: 6, 2016.
- [2] "What is self-care?" ISF International Self-Care Foundation. Available at: <http://isfglobal.org/what-is-self-care/>. Retrieved: 7, 2016.
- [3] "Health Education in Self-Care: Possibilities and Limitations. Report of a Scientific Consultation". Geneva, Switzerland: World Health Organization; November 21–25, 1983.
- [4] S. Lowell, S. Levin, and L. Idler, "Self-Care in Health". Vol. 4: 181-201, 1983.
- [5] "Self-care in the Context of Primary Health Care". Report of the Regional Consultation Bangkok, Thailand, 7–9 January 2009. Available at: http://www.searo.who.int/entity/primary_health_care/documents/sea_hsd_320.pdf. Retrieved: 6, 2016.
- [6] L. Trevena, H Davey, A. Barratt, P. Butow, and P. Caldwell, "A systematic review on communicating with patients about evidence". *Journal of Evaluation in Clinical Practice*, 12,(1): 13-23, 2006.
- [7] J. Puustjärvi and L. Puustjärvi. "Practicing Evidence-Based Medicine in Developing Countries". In the proc. of the IST-Africa 2015 Conference & Exhibition. 2015.
- [8] N. Mulyar, M. van der Aalst, and M. Peleg, "A pattern-based analysis of clinical computer interpretable guideline modeling languages," *Journal of American Medical Information Association*, Vol. 14, no. 6, pp.781–787.2007.
- [9] M. Peleg, S. Tu, P. Ciccarese, and J. Fox, "Comparing computer-interpretable guideline models," A case-study approach. *Journal of American Medical Information Association*, 2003, pp.10–58.
- [10] T. Leaon, K. Kaiser, and S. Miksch, "Free and open source enabling technologies for patient-centric, guideline-based clinical decision support," A survey. *Methods of Information in Medicine*, pp. 74–86, 2007.
- [11] Baeza-Yates, R. and Ribeiro-Neto, B. 1999. "Modern Information Retrieval". Addison Wesley.
- [12] "SPARQL Query Language for RDF", Available at: <http://www.w3.org/TR/rdf-sparql-query/>. Retrieved: 2, 2016.
- [13] "RDF – Resource Description Language". Available at: <http://www.w3.org/RDF/>. Retrieved: 2, 2016.
- [14] T.R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing." Padua workshop on Formal Ontology, March 1993.
- [15] G. Antoniou and F. Harmelen, "A semantic web primer," The MIT Press. 2004.
- [16] M. Daconta, L. Obrst, and K. Smith, "The semantic web. :A Guide to the Future of XML, Web Services, and Knowledge Management," John Wiley & Sons. 2003
- [17] "OWL - WEB OntologyLanguage", Available at: <http://www.w3.org/TR/owl-features/> Retrieved: 2, 2016.
- [18] DuCharme. Learning SPARQL. O'Reilly Media.