Chapter 2
A Faceted Lightweight Ontology for Earthquake Engineering Research Projects and Experiments

Rashedul Hasan, Feroz Farazi, Oreste Salvatore Bursi and Md Shahin Reza

2.1 Introduction

The inventor of the Web, Tim Berners-Lee, envisioned a more organized, well connected and well integrated form of its data that are suitable for humans to read and for machines to understand (Berners-Lee 1999; Berners-Lee et al. 2001). This new form of the Web is called the Semantic Web. With the invention of the Semantic Web, computing paradigm is experiencing a shift from databases to Knowledge Bases (KB), where ontologies play a major role in enabling inferencing that can make hidden facts unconcealed to produce better results for users.

Moreover, KB-based systems provide a mechanism to manage information and semantics thereof that can make systems semantically interoperable and as such can exchange and share data between them. To overcome the interoperability issues and to exploit the benefits offered by the state of the art technologies, we moved to the KB-based system.

In fact, we have developed an ontology named as Earthquake Engineering Research Projects and Experiments (EERPE) using a faceted approach that gives emphasis on research project management and experiments. Following the validation of the ontology by a domain expert, it was published in the knowledge representation language RDF and integrated to the generic ontology WordNet (http://wordnet.

R. Hasan (✉) · O. S. Bursi · M. S. Reza
Department of Civil, Environmental and Mechanical Engineering,
University of Trento, Trento, Italy
e-mail: md.hasan@unitn.it

O. S. Bursi
e-mail: oreste.bursi@unitn.it

M. S. Reza
e-mail: Md.Shahin.Reza@unitn.it

F. Farazi
Department of Information Engineering and Computer Science,
University of Trento, Trento, Italy
e-mail: farazi@disi.unitn.it

F. Taucer, R. Apostolska (eds.), *Experimental Research in Earthquake Engineering,*
Geotechnical, Geological and Earthquake Engineering 35,
DOI 10.1007/978-3-319-10136-1_2

princeton.edu/). The experimental data coming from, inter alia, cyclic and pseudo-dynamic tests were also published in RDF. We used Jena (http://jena.apache.org/), OWLIM (http://www.onto-text.com/owlim) and Sesame (http://www.openrdf.org/) tools for publishing, storage and management, respectively. Finally, integrating the tools, ontologies and data, we developed a system to evaluate the effectiveness of the approach.

The rest of the paper is organized as follows. Section 2.2 depicts an ontology based information management approach. Section 2.3 describes the ontology development methodology that has been followed to build the ontology. Section 2.4 provides a brief description of the ontology representation languages RDF and OWL. In Sect. 2.5, we present existing ontology/thesaurus relevant for this work and as such worth discussing them. Section 2.6 provides the ontology integration. Section 2.7 presents the experimental set-up whilst Sect. 2.8 illustrates evaluation results that show the effectiveness of EERPE ontology. In Sect. 2.9 we briefly describe related work and, finally, in Sect. 2.10 we present conclusions.

2.2 Approach

Figure 2.1 describes an ontology based information management system development approach that involves standard three-tier architecture. KB works as a back-end of the system hosting ontologies represented in RDF while query processing, reasoning and inference mechanisms are incorporated in the business logic layer. User queries and corresponding results are shown in the User Interface (presentation) layer. However, for ontology development we follow the DERA methodology (Giunchiglia and Dutta 2011), for ontology representation in RDF we use Jena and for ontology integration we implemented a facet based algorithm (see Sect. 2.6).

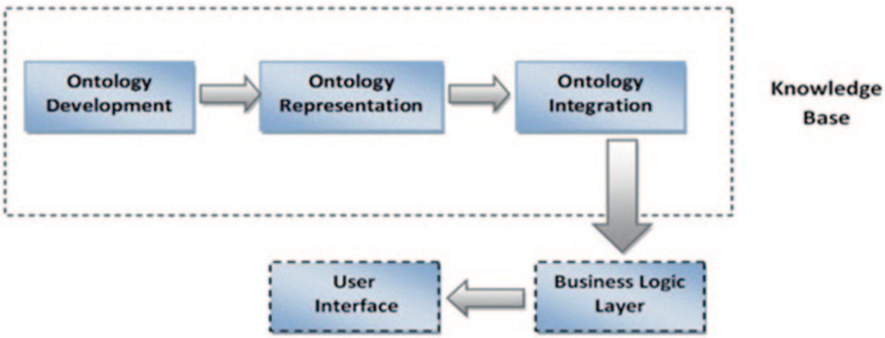

Fig. 2.1 Ontology based systems development approach

2.3 Ontology Development

We use the DERA methodology for ontology development. In fact it is known that it is extendable and scalable (Giunchiglia and Dutta 2011) and ontologies such as GeoWordNet were developed following this approach (Giunchiglia et al. 2010a).

DERA methodology allows for building domain specific ontologies. Domain is an area of knowledge in which users are interested in. For example, earthquake engineering, oceanography, mathematics and computer science can be considered as domains. In DERA, a domain is represented as a 3-tuple ***D*** **=<*E, R, A*>**, where ***E*** is a set of entity-classes that consists of concept sand entities; ***R*** is a set of relations that can be held between concepts and entities and ***A*** is a set of attributes of the entities.

In this three basic components concepts, relations and attributes are organized into facets; hence, the ontology is based on faceted methodology. Facet is a hierarchy of homogeneous concepts describing an aspect of a domain. Ranganathan SR, who was an Indian mathematician-librarian, was the first to introduce faceted approach capable of categorizing books in the libraries (Ranganathan 1967).

Note, however, that a domain can alternatively be called as domain ontology. Henceforth in this paper it will be referred to as domain ontology. Among the macro-steps to develop each component of a domain ontology, we used the following ones.

In the first step (identification) towards building an ontology, we identified the atomic concepts of terms collected from research papers, books, existing ontological resources and experts belonging to the Earthquake Engineering domain giving emphasis on research projects and experiments aspects. We found terms such as device, shaker, experiment, dynamic test, etc., and identified the atomic concept for each of them. We use WordNet as Knowledge Base. The term device has 5 different concepts in it. In our case, we selected the one that has the following description: "*device—(an instrumentality invented for a particular purpose)*". We have found 193 atomic concepts. In the second step (analysis) we analyzed the concepts, i.e. we studied their characteristics to understand the similarity and differences between them. Once the analysis was completed, in the third step (synthesis) we organized them into some facets according to their characteristics. For example, shaker is more specific than device, dynamic testing is more specific than experiment and we assigned the following relationship between them: shaker *IS_A* device and dynamic test *IS_A* experiment. In this way, we built 11 facets. Device and experiment facets are shown in Fig. 2.2 In the fourth step (standardization), we marked concepts with a preferred name in the cases of availability of synonymous terms. For example, while experiment and test are referred to the same concept, we assigned the former term as the preferred one. Finally, the ontology was validated by a domain expert.

Device

- Shaker
- Hammer
- Active Structural device
- Passive Structural device
  - Hydraulic damper
  - Electrical damper
  - MR damper
  - Friction damper

Experiment

- Static
  - Cyclic test
  - Monotonic test
- Dynamic
  - PSD (Pseudo-dynamic) test with substructuring
  - Shaking table test
  - Shaker-Based Test
  - Hammer-Based test

Fig. 2.2 Device and experiment facets

2.4 Ontology Representation

2.4.1 RDF

The Resource Description Framework (RDF) is a data model used to represent information about resources in the World Wide Web (WWW) and can be used to describe the relationships between concepts and entities. It is a framework to describe metadata on the web. Three types of things are in RDF: resources (entities or concepts) that exist in the real world, global names for resources (i.e. URIs) that identify entire web sites as well as web pages, and RDF statements (triples, or rows in a table) (Klyne 2004). Each triple includes a subject, an object and a predicate. RDF is designed to represent knowledge in a distributed way particularly concerned with meaning. The following RDF statements describe the resources 'Hammer' and 'Damper'.

```
<rdf:Descriptionrdf:about="http://earthquake.linkeddata.it/resource/Hammer">
        <rdfs:subClassOfrdf:resource="http://earthquake.linkeddata.it/resource/Device"/>
        <ontology:desciptionrdf:datatype="http://www.w3.org/2001/XMLSchema#string">A hand
        tool with a heavy rigid head and a handle; used to deliver an impulsive force by strik-
        ing</ontology:desciption>
        <rdf:typerdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>

  <rdf:Descriptionrdf:about="http://earthquake.linkeddata.it/resource/Damper">
        <rdfs:subClassOfrdf:resource="http://earthquake.linkeddata.it/resource/Device"/>
           <ontology:desciptionrdf:datatype="http://www.w3.org/2001/XMLSchema#string">A
        device that decreases the amplitude of electronic, mechanical, acoustical or aerodynamic
        oscillations</ontology:desciption>
        <rdf:typerdf:resource="http://www.w3.org/2002/07/owl#Class"/>
  </rdf:Description>
```

The above example represented relationship between 'Hammer' and 'Device' concepts; and the rdfs:subClassOf property is used to relate the former class to its more generic later class.

2.4.2 OWL

Web Ontology Language is designed to represent comparatively complex ontological relationships and to overcome some of the limitations of RDF such as representation of specific cardinality values and disjointness relationship between classes (Giunchiglia et al. 2010b). The language is characterized by formal semantics and RDF/XML based serializations for the web. As an ontology representation language, OWL is essentially concerned with defining terms that can be used in RDF documents, i.e., classes, properties and instances (Antoniou and van Harmelen 2004). It serves two purposes: first, it identifies current documents as an ontology and second it serves as a container metadata regarding the ontology. This language focuses on reasoning techniques, formal foundations and language extensions. OWL uses URI references as names and constructs these URI references in the same manner as that used by RDF. The W3C allows OWL specification includes the definition of three variants of OWL, with different levels of expressiveness. These are OWL Lite, OWL DL and OWL Full, ordered by increasing expressiveness.

2.5 Existing Ontology/Thesaurus

2.5.1 WordNet

WordNet (Miller et al. 1990) is an ontology that consists of more than 100 thousand concepts and 26 different kinds of relations e.g., hyponym, synonym, antonym, hypernyms and meronyms. It was created and is being maintained at the Cognitive Science Laboratory of Princeton University. The most obvious difference between WordNet and a standard dictionary is that its concepts are organized into hierarchies, like professor *IS_A* kind of person and person *IS_A* kind of living thing. It can be used for knowledge-based applications. It is a generic knowledge base and as such does not have good coverage for domain specific applications. It has been widely used for a number of different purposes in information systems including word sense disambiguation, information retrieval and automatic text summarization.

2.5.2 NEES Thesaurus

The Network for Earthquake Engineering Simulation (NEES) is one of the leading organizations for Earthquake Engineering in USA. They developed the earthquake engineering thesaurus; it is based on Narrower and Broader terms. It contains

NEES Earthquake Engineering Ontology by Broader Term		
Broader Term	Term	Narrower Term
AASHTO_2001	AASHTO_LRFD_Bridge_Design_Specifications	
Acceleration	Peak_Base_Acceleration	
Actuator	Dynamic_Actuator	
Actuator	Static_Actuator	
Axial_Load	Cyclic_Axial_Load	
Bearings	Preformed_Fabric_Pads	Cotton_Duck_Bearing_Pads

Fig. 2.3 NEES thesaurus

around 300 concepts and we have integrated in our ontology 75 concepts from NEES. Figure 2.3 depicts a small portion of NEES thesaurus.

2.6 Ontology Integration

Developed facets include concepts that were selected from NEES thesaurus to be incorporated into our ontology. This integration was accomplished in fact when we built the facets. In this Section, we describe how we integrated our developed ontology with Wordnet. Basically, we applied the semi-automatic ontology integration algorithm proposed in Farazi et al. (2011). In particular, we implemented the following macro steps:

- ***Facet concept identification:*** For each facet, the concept of its root node is manually mapped to WordNet, in the case of availability.
- ***Concept Identification:*** For each atomic concept *C* of the faceted ontology, it checks if the concept label is available in WordNet. In the case of availability, it retrieves all the concepts connected to it and maps with the one residing in the sub-tree rooted at the concept that corresponds to the facet root concept.
- ***Parent Identification:*** In the case of unavailability of a concept it tries to identify parent. For each multiword concept label it checks the presence of the header, and if it is found within the given facet, it identifies it as a parent. For instance, in WordNet it does not find hydraulic damper for which damper is the header and that is available there in the hierarchy of device facet. Therefore, it recognizes the damper with the description *"damper, muffler—(a device that decreases the amplitude of electronic, mechanical, acoustical, or aerodynamic oscillations)"* as the parent of the hydraulic damper.

2.7 Experimental Set-Up

Figure 2.4 represents the architecture of our KB-based information management system that uses semantic tools and technologies. We published our developed domain specific ontology into RDF using Jena, a Semantic Web tool for publishing

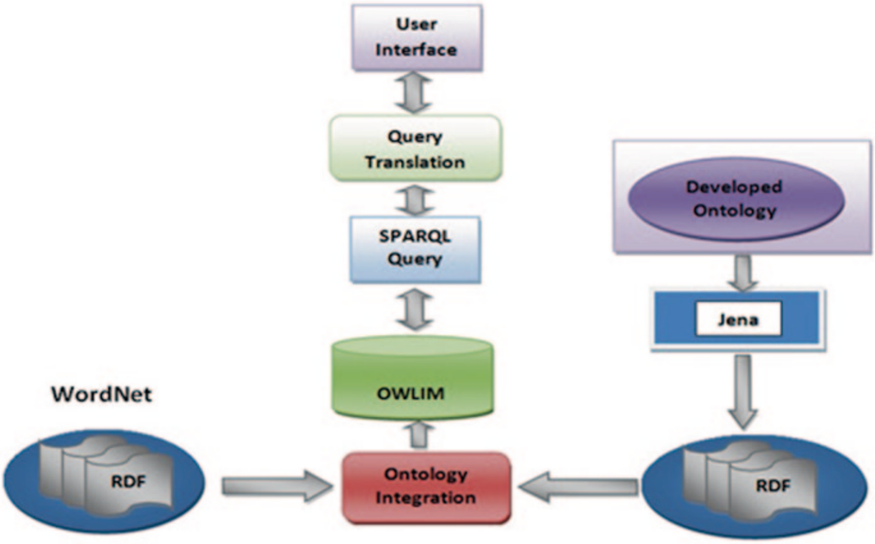

Fig. 2.4 KB-based system architecture

and managing ontologies and integrated it with WordNet RDF using the approach described in Sect. 2.6. To increase the coverage we integrated our developed ontology with WordNet. The outcome of the ontology integration was put in OWLIM triple store that was queried using Sesame API. For translating user natural language queries into SPARQL queries, we developed a few lines of code written in java. SPARQL is a query language specific to query RDF representations. It allows add, update and delete of RDF data.

2.8 Results

In this Section, we describe basically what advantages users can get with KB-based systems over traditional DB systems. In particular, we performed synonym search and more specific concept search.

Synonym Search When a concept is represented with two or more terms, they are essentially synonymous that are represented in RDF with *owl: equivalent Class*. For example, test and experiment represent the same concept and in the ontology they are encoded accordingly with equivalent relation. Therefore, user query for test can also return experiment, (Fig. 2.5) because they are semantically equivalent.

More Specific Concept Search In our ontology concept hierarchies are represented using *rdfs: subClassOf*. For example, hammer and damper are more specific concepts of device, hence, they are represented as follows: hammer *rdfs: subClassOf*device; and damper *rdfs: subClassOf* device. Moreover, hydraulic damper is more specific than damper and it is encoded as hydraulic damper *rdfs:subClassOf*

Fig. 2.5 Synonymous relationship

Fig. 2.6 Transitive relationship of device

damper. Note that *rdfs: subClassOf* is a transitive relation. Using OWL inference engine, we can utilize the power of transitivity and for a given concept we can retrieve all more specific concepts that are directly or indirectly connected by *rdfs:subclassof*. Therefore, a search for device retrieved all of its more specific concepts as shown in Fig. 2.6.

2.9 Related Work

NEES ontology has been developed in the domain of earthquake engineering. However, it is mainly a thesaurus encoding broader and narrower relations that cannot capture ontological details. For instance, it cannot be clarified in thesaurus whether a relation between two concepts is *IS_A* or *PART_OF*. As a result ontologies represented as thesaurus might lead to some unexpected results. DBPedia is an example that uses broader/narrower relations and ended up establishing connections between Telecommunication, and Flora and Fauna. In contrast, the ontology developed in this paper does not suffer from this issue; rather it provides better clarification because it exploits ontological relations.

2.10 Conclusion

In this paper, we provided a detailed description of the development of Earthquake Engineering Projects and Experiments ontology. We followed DERA methodology for building this domain specific ontology. We exploited an ontology integration algorithm that was employed to incorporate our ontology into WordNet. It helped to increase the coverage of the Knowledge Base. On top of the integrated ontology, i.e., put in an OWLIM store of an experimental setting, we experimented the semantic and ontological capabilities of the developed system and interesting results were found.

References

Antoniou, van Harmelen (2004) Grigoris Antoniou and Frank van Harmelen: a Semantic Web Primer. MIT Press

Berners-Lee T (1999) Weaving the web. Orion business books, London

Berners-Lee T, Hendler JA, Lassila O (2001) The semantic web. Sci Am J 284(5):34–43

Farazi Feroz et al (2011) A faceted ontology for a semantic geo-catalogue. The semanic web: research and applications. Springer, Berlin, pp 169–182

Giunchiglia F, Biswanath D (2011) DERA: a faceted knowledge organization framework

Giunchiglia F, Farazi F, Tanca L, de Virgilio R (2010a) The semantic web languages. Semantic web information management. Springer Verlag, Berlin

Giunchiglia F, Maltese V, Farazi F, Dutta B (2010b) GeoWordNet: a resource for geo-spatial applications. The semantic web: research and applications. Springer, Berlin, pp 121–136

Giunchiglia F et al (2012) A facet-based methodology for the construction of a large-scale geospatial ontology. J Data Semant 1(1):57–73

Grigoris A, van Harmelen F (2004) Semantic web primer. MIT Press, USA

Klyne G, Jeremy JC (2004) Resource description framework (RDF): concepts and abstract syntax. W3C Recommendation, 2004. World Wide Web Consortium. http://w3c.org/TR/rdf-concepts

Miller GA, Beckwith R, Fellbaum CD, Gross D, Miller K (1990) WordNet: an online lexical database. J Lexicograph 3(4):235–244

Ranganathan SR (1967) Prolegomena to library classification. Asia Publishing House, New York

http://www.springer.com/978-3-319-10135-4

Experimental Research in Earthquake Engineering
EU-SERIES Concluding Workshop
Taucer, F.; Apostolska, R. (Eds.)
2015, XXVII, 614 p. 463 illus., Hardcover
ISBN: 978-3-319-10135-4