Metadata for Learning Resources: Technologies and Directions of the Semantic Web – A Brief Review

Ossi Nykänen

Digital Media Institute, Hypermedia Laboratory, Tampere University of Technology, Finland
Ossi.Nykanen@tut.fi

Abstract

The appearance of the Semantic Web technologies has a significant impact on the development of all semantically interoperable tools for the Web. This article provides an outline of the SW technologies and tools from the perspective of learning systems.

1. Introduction

The Semantic Web (SW) is both a technical framework and a vision of making semantically aware applications for the Web [15]. In brief, the SW is a set of universal, neutral standards and tools for publishing and processing (meta)data in applications [13, 12, 4].

The SW also establishes the technical foundation for the metadata of learning resources. In principle, learning systems simply introduce additional, domain-specific semantics to the standard (see e.g. [14]).

This article provides a brief overview of the SW technologies from the perspective of learning systems.

2. Models and Description Languages

The core of the Semantic Web is defined by a set of W3C recommendations. These basically establish the Resource Description Framework (RDF), Web Ontology Language (OWL), and the related RDF Schema Specification, and the RDF and OWL Semantics. The official status of the W3C technical reports and publications is available at [16].

2.1. RDF Data Model, Syntax, and Schemas

Conceptually, the Semantic Web is defined by the RDF data model (Resource Description Framework (RDF): Concepts and Abstract Syntax). RDF is an assertional language, suitable for making statements about, e.g., Web resources. RDF statements have a simple form associating a subject, a predicate, and an object, whose interpretation can be formalized.

A statement asserts that a given thing has a specific property. RDF statements assign universal properties, identified by URI references (URIrefs), for things, optionally identified with URI references. Using URIrefs for denotation is technical design decision, allowing the decentralized development of names.

Different kinds of vocabularies are designed for different kinds of purposes. By sharing the nodes of individual statements, assertions may be linked into uniform network structures, called RDF graph(s).

The syntax of the abstract is reinforced with the RDF/XML Syntax Specification (Revised) and the RDF Vocabulary Description Language 1.0: RDF Schema. In short, the syntax specification defines how RDF models are written to text files, accompanied with few primitive structure definitions (such as bags and lists).

RDF Schemas allow modeling the use of RDF vocabularies formally. The RDF Schema introduces two important concepts: Classes and properties (relations). Classes denote (hierarchical) sets of things while properties denote (hierarchical) sets of relations. In addition, RDF Schema supports the use of rich datatypes. In short, RDF Schemas provide a syntactic mean for controlling the usage of vocabularies (such as document type definitions allow validating structured documents). In addition, the schema information also establishes a basis for formal inference.

2.2. Ontology Framework and Reasoning

An ontology is a formal, explicit specification of a shared conceptualization of a particular domain [7]. Ontologies are a vehicle for communication, reflection, and entailment. In the computerized context, ontologies are used for management, reuse and organization of knowledge, e.g., for purposes of decision support.
The Web Ontology Language develops RDF Schemas by strengthening its description capabilities (OWL Web Ontology Language Guide; OWL Web Ontology Language Reference). Ontologies establish new class and property definitions, and allow declaring classes via set operations. Ontologies may also have names and descriptive labels and an ontology may import another ontology, subject to simple versioning, including an option for declaring backwards compatibility.

To manage the different needs of SW reasoning, OWL features three increasingly expressive, overlapping sublanguages: OWL Lite, OWL DL, and OWL Full. These differ in their language constructs, restrictions, and completeness. For instance, OWL Lite does not allow the Boolean combinations of classes.

The specifications of RDF Semantics and OWL Web Ontology Language Semantics and Abstract Syntax establish a formal interpretation of RDF. These define the valid inference rules which allows deducing (implicit) facts from RDF graphs.

3. Tools

From the developer's perspective the specifications are not enough: tools are needed. The widely adopted set of (non-standardized) tools includes interfaces for RDF data, query languages, and rule systems.

In principle, SW applications are simply programs manipulating RDF data. Low-level development takes place on the level of the RDF data model. Several class libraries, modules, and application frameworks exist, such as Jena (Java), RDFStore (Perl), RDFLib (Python), and Redland RDF (C). These allow managing RDF data using design patterns (such as iterators) close to the abstract data model [2,9,8,10].

A bit more abstract level of SW applications is based on query languages. Indeed, the uniform data model and the rigorous semantics establish a firm basis for accessing, merging, and even mining data via easy-to-learn query expressions. A typical example of an RDF query language is the SQL-like RDQL, implemented, e.g. in the Jena Toolbox [2]. An example of a more efficient query language is RQL [5], which as such also recognizes schema and ontology concepts.

Finally, RDF rule systems abstract the task of SW application programming a step further, by introducing a declarative language for expressing if-then rules or material implications, including logical and procedural rules (modifying the local RDF database). The best known SW rule system is probably the CWM application [11]. CWM allows implementing reasoning systems and integrating these with shell applications.

4. Discussion and Conclusions

In short, the Semantic Web allows publishing and processing metadata in the Web. In the context of learning applications this basically establishes the potential of integrating learning systems and data, searching and reusing content and tools, and automating some of the related, mundane tasks.

The universal metadata standards significantly shape the development of learning systems. It seems likely that semantic applications are expected either to be directly based on, or map onto the native SW concepts.

It is worth emphasizing that SW has the role of an enabling technology: the standardization of the content is still a major challenge, both socially and technically. To manage this complexity, several educational and other standardization forums exist (see, e.g., [3,6]).