Lecture 16: Knowledge Sharing and Semantic Web

Dr. Roman V Belavkin

BIS4410

Contents

1 Introduction to Semantic Web

Historical Notes

- In 1989, Tim Berners-Lee proposed and created the first Internet Server and Client for sharing documents.

- It was based on the Hypertext Transfer Protocol (HTTP)

- Web documents are created and read by humans. They are not designed for machines.

- Ten years later Berners-Lee was promoting the idea of Semantic Web:

    ...a web of data that can be processed directly and indirectly by machines

Semantic Web

- *Semantic* means *meaning*.

- Computers can read a document (a Web page), but not really understand it.

- The idea of Semantic Web is that data can also be read and understood by computers and automated agents.

- Semantic Web is not yet a technology, but a vision of the future Web.

- A number of technologies already exists.

Main features of Semantic Web

- Meta-data (e.g. data about data, keywords)
- Better indexing and retrieving information (e.g. keywords, tags may not always help).
- Better annotations, comments (e.g. Wiki)
- Merging and sharing databases
- Better standards for storing data
- Automated services (e.g. making reservations, shopping)
- Resource discovery
- Intelligent agents

2   Semantic Web Components and Standards

Semantic Web Components and Standards

URI, URL : Uniform Resource Indicator (URI), Uniform Resource Locator (URL) for resources.

XML : Extensible Markup Language

XML Schema : A language defining structure of XML languages (extensions)

RDF : Resource Description Framework; language for describing information and metadata

RDF Schema : A vocabulary for describing properties and classes for specific RDF applications

OWL : Web Ontology Language; a family of languages for describing ontologies (concepts, relations between concepts)

Logic and Proof : Automatic reasoning used to establish consistency and correctness of resources, make automatic inference for applications.

Trust : Methods for authentication of data, services and resources.

Evidence of Semantic Web today

- There is an increasing use of meta-data, *tags* for indexing resources (e.g. YouTube, Last FM, WikiPedia).

- Search engines based on *semantic similarity*

- Linked data and databases (e.g. biological, medical and other natural science resources, financial data).

- Domain-ontologies and ontology learning.

- Annotation and user-editable contents (e.g. WikiPedia).

Summary

- Semantic Web is a development of Web technologies and standards to facilitate machine understanding of data, resources and automating applications.

- W3C promotes a set of standards (XML, RDF, OWL).

- Some technologies and standards become popular, and many new successful applications have emerged.

- Many ideas or components have not been or may never be implemented.

Additional Reading

1. Iba, Nemoto, Peters, and Gloor (2010)

    Analyzing the Creative Editing Behavior of Wikipedia Editors
    Through Dynamic Social Network Analysis

References

Iba, T., Nemoto, K., Peters, B., & Gloor, P. A. (2010). Analyzing the creative editing behavior of Wikipedia editors through dynamic social network analysis. *Procedia Social and Behavioral Sciences*, *2*, 64416456.