Semantic web

17/01/2008 21:00 by Sally Ahmed

Semantic Web

Introduction

Semantic web as defined by the creator of the web Tim Berners-Lee is *“a web of data, in some ways like a global database*” (Berners-Lee, 1998). To elaborate further Mr. Berners-Lee explains in an interview held by IDG Now, data is expressed on computers as associated files with applications that deal specifically with information, an example would be, data in calendars, bank systems, spreadsheets, and database application. Looking at a web page, data is not clearly defined and not associated with any of the applications usually on computers. Semantic web will allow data to interact and connect together; it will bring on a common data format for all applications, for databases and web pages alike (Moon, 1999). Semantic web is not to build an artificial intelligence system which allows computers to understand what humans write on web pages; on the contrary, it is an attempt to make web pages more understandable and well-defined to support automatic extraction of data from within web content (Berners-Lee, 1998).

Analysis

The emergence of the web and the way HTML took off was driven by how society’s needed to grew, from Internet chat to file transfer to high-end communities through blogs and wiki’s. HTML was not limited to web content, knowledge base and help files adapted the language as a format to document software applications and provide training material. The revolution of technologies on the Internet allowed companies like Google to index pages; a thought that was very far away, says Tim Berners-Lee in his lecture at MIT. Web services have evolved to pave the road for distributed information and modular programming allowing interoperability among sites. Through XML, data in one site can be used by another using the common protocols and standards supported by both (Berners-Lee, 1998). XML defines schemas that deal with fields of data, what is required is a system that can tell the computer what sort of information (data) it can derive from within a page (Moon, 1999). With Web 3.0 a site will provide data that can be navigated through and extracted from multiple sites, this is a result of the fact that semantic web data model is closely related to a relational database where records of data share common fields that connect them together (Berners-Lee, 1998).

The solution provided to support semantic web is in the form of metadata that describes the data contained on web pages. Resource Description Framework (RDF) is a base to manage metadata; it is the ground that computers can use to exchange and interact with applications on the Web (W3C RFC, 1999). The applications for RDF include digital libraries, online catalogs, and indexing systems that are usually associated with content and content relationships models deployed in most web pages. With RDF data within business portals will be analyzed and identified as resources, properties, or statements transparent to the domain, further more, the specifications will merge with other documents to comprise a framework of classes. Classes organized as a hierarchy comprises a schema that can be reusable as metadata definitions along side multiple platforms. Resources created in this hierarchy can be identified using a resource identifier (URI), which enables a document given to a machine with this identification to be recognized by the system and triggers it to dig and find similar data (Berners-Lee, 1998).

[-pagebreak-]

Implications

Semantic web can be the solution to overpower the limitations of current information management systems in finding and extracting data from unorganized resources. RDF is meant to describe any data regardless of its character, location, source, or type, the concept of URI is richer to uniquely identify any object on the web (Berners-Lee, 1998).

The pillars of Semantic web are standards and common protocols that are the bases for knowledge representation; HTML, RDF, the data language resource description web ontology language (OWL) that describes to the machine what is going on, in addition to RDF1 which is a query language to make inquiries among machines much easer, will all emerge and collaborate to bring in more to the web and more intelligent programs that will bring the Internet more closer (Cleave, 2004).

The current research and implementation of Abilene network and the Next Generation Internet (NGI) Internet 2 of high-performance backbone network linking major universities and research labs across the US, is a good foundation for what Semantic web can do, and represent the perfect platform for grid computing, digital libraries, virtual laboratories, and distance learning (Abilene, Internet2). Internet2 or I2 was developed by a group of universities in 1996 providing improved connectivity standards to reach 10gbps (gigabits per second). With more than 227 universities and libraries connected, network based applications and experimental programs can run on this network of high-bandwidth connection feeding on the latest technology of gigabit Ethernet and IP protocol version 6 (Reardon, 2004). Semantic web standards can be the base of material and data distributed on this network, providing the best test platform to explore the full potential and what can be achieved.

Conclusion

Tim Berners-Lee believes that with Web 3.0 we can succeed and fantastic things can happen, but the infrastructure need to be built, laws of privacy and security need to be revised and honored, further more, the web need to remain open for researchers to allow for continuous upgrade and development. Semantic web will kick off when individuals materialize the need to work on data processing, and think about collaborating their data, with company’s information and that of the government (Moon, 1999).

References: Berners-Lee, Tim. 1998. Semantic Web Road map.W3C team. (14 October 1998) http://www.w3.org/DesignIssues/Semantic (accessed 16 Jul 2007) Moon, Peter. 1999. The future of the Web as seen by its creator. IT World IDG Now (7 July 1999) http://www.itworld.com/Tech/4535/070709future/ (accessed 14 Jul 2007) W3C RFC. 1999. Resource Description Framework (RDF) Model and Syntax Specification. W3 Consortium (5 January 1999) http://www.w3.org/TR/PR-rdf-syntax/ (accessed 20 Jul 2007) Cleave, Kenith Van. 2004. Regis University Database Practicum Experience. Regis University. (14 November 2004) http://trackit.arn.regis.edu/dba/Thesis%2520Papers/ kvancleave_2004Bfinalreport_20041118.pdf (accessed 16 Jul 2007) Abilene, Internet2. http://abilene.internet2.edu/about/ (accessed 17 Jul 2007) Reardon, Marguerite. 2004. Internet2: 2004 and beyond. CNET, News (24 August 2004) http://news.com.com/2100-1034_3-5321053.html (accessed 19 Jul 2007)