Chapter 8
Semantic Web, Ontology, and Linked Data

Anindya Basu
Maharani Kasiswari College, India

ABSTRACT

Enormous amount of information is being produced every day and get consumed according to the needs of human being. Semantic web and ontology represent information which are machine processable and understand the semantics present among the entities. Ontology can be represented as Knowledge Organization and data modelling tool. Librarians are designated as "Information Custodian" or "Knowledge Keepers". Implication and application of concepts in LIS can play big role in shaping knowledge-based services and mining and inferring them in better way in future. Ontology and semantic web is the future of LIS as speculated by several professionals and experts. This chapter would delineate a basic overview of Semantic Web, Ontology and linked data.

INTRODUCTION

One of the most important inventions in last century is the evolution of information space through World Wide Web and information sharing through the Internet. What we see internet today, is a much transformed and agglomerated version of interoperable protocols and storage space distributed globally. Intuitive and intelligent applications are abundant in web platform now-a-days. But the greater question is being imposed on us whether the search application through the Web is intelligent enough to quench the thirst of proper information at right time. Can it handle complex queries which demands to understand and connect heterogeneous entities along with its attributes? Interestingly, the retrieved results from search engines hit out too many non-relevant information. Users always perceive that, any search tool would give answers rather just giving a huge collection of web-sites. This phenomenon forces us to relook at the basic architecture of html web-pages which are basically connected through hyperlinks(HTTP). HTTP links act as a locator only just identifying the web-page meant for only information representation. HTML is very much oriented to representation of textual content but not suitable for understanding the relationships present among information or data entities. It is also hard to extract out structured content

from HTML tags unless compatible micro-formats are defined. Still, they are semantically very poor because of their inability to understand web semantics.

Most of the SEs(search engine) depend on keywords as form of query and proximity of the words and consequently give the results mainly due to matching of the keyword against its index and return the hits. Another important feature of traditional SEs (Search Engines) it facilitates navigation and browsing which is helpful for a user to locate the relevant objects or documents. This is merely an indication of the fact that search engines cannot understand the basic meaning of the terms and also the relevant semantics of the query terms. Query term expansion and prediction approach acted in this direction but they rely mostly on statistical measure which lacks semantics of natural languages and show too much heuristics. As the volume of web-pages and heterogeneity of information/data are growing up, data discovery is still a problem area for search engines. For example, no SE can search within a spreadsheet or give numerical results as answers corresponding to a query. but Binding, defining, qualifying and standardizing data element with proper metadata have been a very new initiative from the scientific community.

Enormous amount of information is produced everyday to cater the needs of human being. There are so many users who have different information needs and most of them are dependant on the web to fulfill their needs. Web is a big semi structured database which provides a vast amount of information. WWW constitutes of near to 12 billion web pages. Through the rapid growth of web, it has become an easy way to access information. But for the rapid increase of information, users are facing new challenges. So locating the precise information also raises a big challenging task. Moreover, most of the search engines primarily give priority to index huge amount of information-contained files or web-pages, so that, once searched it can give maximum number of hits. But, evaluating enormous links by several parameters is strictly a statistical procedure. Semantic sense and relevance measures on the search space are still grey areas of research and raises many open questions.. Gruber defined as 'ontology' as 'a specification of a conceptualization'. If two people agree to use the same ontology when communicating, then there should be no ambiguity in the communication. To enable this, an ontology codifies the semantics used to represent and reason with a body of knowledge." In this effect, knowledge formalism for any discipline/subject/domain may be one of the fundamental work area for inference at semantic level.

The term Semantic Web was coined by Tim Berners-Lee. It is an effort to enhance current web and provides better machine processable data, programs to help users to find their requirements. Semantic web is not new web rather extension of existing web which connects data in fully classified manner inheriting all kinds of semantic and syntactic perspectives. It is a collaborative application provided by World Wide Web Consortium (W3C). It provides a common framework that allows data to be shared and reused across applications, enterprise, and community boundaries. The main objectives of the Semantic Web is to make an interface where user can be able to find, share, cluster and combine information more easily and understand the underlying semantics of an entity. The supporting formats of such semantic data are RDF(Resource Description Framework) and OWL(Web Ontology Language).

The concept of Ontology is a branch of philosophy which deals with study of nature as being, embodiment of semantic connection between entities through attributes and properties. Scientific community takes ontology as a multi-faceted tool for execution of different set of services and operations. Application of ontology can be visible in Knowledge Management, Knowledge formalism and representation, data modeling, vocabulary control, rule based services etc. Library science has been using vocabulary controlled devices like glossary, subject heading and thesaurus since a long time ago. Ontology opens up new arena in the context of information organization through symbolic knowledge representation, formal logic and automated reasoning and from conceptual modeling of any domain knowledge. Librar-

Related Content

Implementing Geospatial Web Services for Cloud Computing

Gobe Hobona, Mike Jacksonand Suchith Anand (2011). *Geospatial Web Services: Advances in Information Interoperability (pp. 287-308).*

www.irma-international.org/chapter/implementing-geospatial-web-services-cloud/51491

Geospatial Service Composition in Grid Environments

Tino Fleurenand Paul Müller (2011). *Geospatial Web Services: Advances in Information Interoperability (pp. 332-363).*

www.irma-international.org/chapter/geospatial-service-composition-grid-environments/51493

Semantic Enrichment for Geospatial Information in a Tourism Recommender System

Joan de la Flor, Joan Borràs, David Isern, Aida Valls, Antonio Moreno, Antonio Russo, Yolanda Pérezand Salvador Anton-Clavé (2012). *Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications (pp. 133-155).*

www.irma-international.org/chapter/semantic-enrichment-geospatial-information-tourism/65112

Security in Internet of Things: Requirements, Challenges, and Open Issues

Said Ul Abrar, Kamran Ullah, Saleem Zahid, Mohib Ullah, Irfan Ullah Khanand Muhammad Inam Ul Haq (2023). *Protecting User Privacy in Web Search Utilization (pp. 215-233).*

www.irma-international.org/chapter/security-in-internet-of-things/322593

A Deep Q-Learning Network for Dynamic Constraint-Satisfied Service Composition

Xuezhi Yu, Chunyang Ye, Bingzhuo Li, Hui Zhouand Mengxing Huang (2020). *International Journal of Web Services Research (pp. 55-75).*

www.irma-international.org/article/a-deep-q-learning-network-for-dynamic-constraint-satisfied-service-composition/264156