

# Chapter XIII

## Semantic Association Analysis in Ontology–Based Information Retrieval

**Payam M. Barnaghi**

*University of Nottingham, Malaysia*

**Wei Wang**

*University of Nottingham, Malaysia*

**Jayan C. Kurian**

*University of Nottingham, Malaysia*

### **ABSTRACT**

*The Semantic Web is an extension to the current Web in which information is provided in machine-processable format. It allows interoperable data representation and expression of meaningful relationships between the information resources. In other words, it is envisaged with the supremacy of deduction capabilities on the Web, that being one of the limitations of the current Web. In a Semantic Web framework, an ontology provides a knowledge sharing structure. The research on Semantic Web in the past few years has offered an opportunity for conventional information search and retrieval systems to migrate from keyword to semantics-based methods. The fundamental difference is that the Semantic Web is not a Web of interlinked documents; rather, it is a Web of relations between resources denoting real world objects, together with well-defined metadata attached to those resources. In this chapter, we first investigate various approaches towards ontology development, ontology population from heterogeneous data sources, semantic association discovery, semantic association ranking and presentation, and social network analysis, and then we present our methodology for an ontology-based information search and retrieval. In particular, we are interested in developing efficient algorithms to resolve the semantic association discovery and analysis issues.*

## INTRODUCTION

The current Web provides a universal platform to explore and contribute to the global information network. Undoubtedly, the Web has emerged as the world's major information resource with immediate accessibility in a world-wide scale. Currently, in most of the cases, in order to transform available information into meaningful knowledge, machines have to depend on the human inference ability (Craven, DiPasquo, Freitag, McCallum, Mitchell, Nigam, et al., 2000). Contemporary popular online search engines and information retrieval systems index and search the Web documents based on analysis of the document link structures and keywords. The keywords are often extracted from the documents according to the frequency of occurrence and considered as standalone entities without application contexts and other semantic relationships. This superficial understanding of content prevents retrieving implicitly-related information in most of the cases. It also in some cases returns irrelevant results to the user. In the context of multimedia Web, the current search systems are even more limited. Most of the multimedia search on the current Web relies on text explanations extracted from accompanying pages or tags provided by content authors. There are commercially successful Web sites for multimedia publishing, sharing, and retrieval on the Web such as MySpace<sup>1</sup>, YouTube<sup>2</sup>, and Flickr<sup>3</sup>. These Web sites have demonstrated a great achievement in acquiring millions of users to form communities and contribute to content generation; they also provide interfaces to search and view the published contents, but the search functions regularly rely on conventional methods and keyword matching mechanisms. As overwhelming information is published on the Web, new information search and retrieval methods are needed in order to enable users to find more relevant information based not only on keywords, but also context and preferences of each individual user. This has led to the

introduction of a new era for Web information search and retrieval, namely, community-based and semantic-enhanced search.

The emergent **Semantic Web** technologies provide the possibility to realize the vision of meaningful relations and structured data on the Web. As an extension to the current Web, the Semantic Web technologies enable computers and people to work in cooperation (Berners-Lee, Hendler, & Lassila, 2001). The Semantic Web focuses on publishing and retrieving machine-processable Web contents (Dayal, Kuno, & Wilkinson, 2003). In the Semantic Web framework, flexible and interoperable structures such as Web ontology language (OWL)<sup>4</sup> and resource description framework (RDF)<sup>5</sup> are used to represent resources. The relationships between entities in a particular domain can be explicitly expressed using an ontology (Chandrasekaran, Josephson, & Benjamins, 1999). To describe multimedia data and documents on the Semantic Web, the ontology concepts are required to be mapped to the metadata description structure, which is usually referred to as semantic annotation. The semantic-enhanced search focuses on utilizing the structured description and knowledge description ontologies to enhance the results of information search and retrieval process on the Web. The better the relationships are processed and analyzed, the more relevant context results are obtained to be shown to users.

A significant feature in the information search and retrieval systems developed based on Semantic Web technologies is the analysis of meaningful relationships between Web resources to provide enhanced search results. In order to do so, a semantic query framework is required to support this process. There are various semantic query languages such as SPARQL (2007), RQL (Karvounarakis, Alexaki, Christophides, Plexousakis, & Scholl, 2002), and SeRQL (Broekstra, Kampman, & Harmelen, 2002) that are able to query the Semantic Web data (i.e., RDF or OWL data) based on ontological concepts. However,

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/semantic-association-analysis-ontology-based/19876](http://www.igi-global.com/chapter/semantic-association-analysis-ontology-based/19876)

## Related Content

---

Digital Library Service System; Digital Library System; Document Model; Grid Computing; Grid Infrastructure; Institutional Repository; Service Oriented Architecture

Mohammed Nasser Al-Suqriand Esther O.A. Fatuyi (2009). *Handbook of Research on Digital Libraries: Design, Development, and Impact* (pp. 8-15).

[www.irma-international.org/chapter/digital-library-service-system-digital/19865](http://www.irma-international.org/chapter/digital-library-service-system-digital/19865)

Traffic-Signs Recognition System Based on FCM and Content-Based Image Retrieval

Yue Liand Wei Wang (2011). *International Journal of Digital Library Systems* (pp. 1-12).

[www.irma-international.org/article/traffic-signs-recognition-system-based/63646](http://www.irma-international.org/article/traffic-signs-recognition-system-based/63646)

Demographics, Socio-Economic and Cognitive Skills as Barriers to Information Seeking in a Digital Library Environment

Felicia Yusuf, Sola Owolabi, Ayooluwa Aregbesola, Sunday Oguntayo, Foluke Okochaand Toluwani Eyiolorunse (2016). *Information Seeking Behavior and Challenges in Digital Libraries* (pp. 179-202).

[www.irma-international.org/chapter/demographics-socio-economic-and-cognitive-skills-as-barriers-to-information-seeking-in-a-digital-library-environment/159598](http://www.irma-international.org/chapter/demographics-socio-economic-and-cognitive-skills-as-barriers-to-information-seeking-in-a-digital-library-environment/159598)

Use of Digital Resources Among Social Scientists: An Evaluative Study

S. Thanuskodi (2020). *Challenges and Opportunities of Open Educational Resources Management* (pp. 1-23).

[www.irma-international.org/chapter/use-of-digital-resources-among-social-scientists/258125](http://www.irma-international.org/chapter/use-of-digital-resources-among-social-scientists/258125)

Copyright Issues in the Context of the Digital Library

Pedro Pina (2016). *Information Seeking Behavior and Challenges in Digital Libraries* (pp. 70-83).

[www.irma-international.org/chapter/copyright-issues-in-the-context-of-the-digital-library/159593](http://www.irma-international.org/chapter/copyright-issues-in-the-context-of-the-digital-library/159593)