# Handbook of Research on Digital Libraries: Design, Development, and Impact

Yin-Leng Theng
*Nanyang Technological University, Singapore*

Schubert Foo
*Nanyang Technological University, Singapore*

Dion Goh
*Nanyang Technological University, Singapore*

Jin-Cheon Na
*Nanyang Technological University, Singapore*

Information Science REFERENCE

**INFORMATION SCIENCE REFERENCE**

Hershey · New York

| | |
|---|---|
| Director of Editorial Content: | Kristin Klinger |
| Director of Production: | Jennifer Neidig |
| Managing Editor: | Jamie Snavely |
| Assistant Managing Editor: | Carole Coulson |
| Typesetter: | Chris Hrobak |
| Copy Editors: | Ashlee Kunkel and Laura Kochanowski |
| Cover Design: | Lisa Tosheff |
| Printed at: | Yurchak Printing Inc. |

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

and in the United Kingdom by
Information Science Reference (an imprint of IGI Global)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 0609
Web site: http://www.eurospanbookstore.com



Library of Congress Cataloging-in-Publication Data

Handbook of research on digital libraries : design, development, and impact / Yin-Leng Theng ... [et al.], editors.
p. cm.
Includes bibliographical references and index.
Summary: "This book is an in-depth collection aimed at developers and scholars of research articles from the expanding field of digital libraries"--Provided by publisher.
ISBN 978-1-59904-879-6 -- ISBN 978-1-59904-880-2 (ebook)
1. Digital libraries. I. Theng, Yin-Leng, 1961-
ZA4080.H36 2009
025.00285--dc22
2008028570

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

# 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[1], YouTube[2], and Flickr[3]. 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 lead 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)[4] and resource description framework (RDF)[5] 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, these languages do not adequately provide a query mechanism to discover the complex and implicit relationships between the resources. Such complex relationships are called semantic associations (Aleman-Meza, Halaschek-Wiener, Sahoo, Sheth, & Arpinar, 2005; Sheth, Aleman-Meza, Arpinar, Bertram, Warke, Ramakrishnan, et al., 2005). The process of discovering semantic associations is also referred to as semantic analytics. This can be viewed as a special class of search applications which facilitates obtaining decidable knowledge from massive interconnected data resources. This process assists information analysis and provides new and unexpected insights to information search and retrieval (Aleman-Meza, Halaschek, Arpinar, & Sheth, 2003).

## RELATED WORK

The successful development of the Semantic Web applications depends on availability and adoption of ontologies and semantic data (Guha, 2003; Kiryakov, Popov, Terziev, Manov, & Ognyanoff, 2005; Shadbolt, Hall, & Berners-Lee, 2006). In the last few years different thesauruses, ontologies, and metadata structures have been introduced, such as friend of a friend (FOAF[6]) ontology, the GENE ontology[7], NCI meta thesaurus (Golbeck, Fragoso, Hartel, Hendler, Oberthaler, & Parsia, 2003), and Cyc ontology[8]. A number of works have been carried out to apply ontology-based information search and retrieval in various domains (e.g., Swoogle ontology search engine [Ding, 2004], TAP generic semantic search framework [Guha, 2003], semantic annotation in KIM (Kiryakov et al., 2005), semantics-enhanced multimedia presentation generation [Falkovych, Werner, & Nack, 2004; Rutledge, Alberink, Brussee, Pokraev, van Dieten, & Veenstra, 2003], and the semantic-based multimedia search engine Squiggle ]Celino, Valle, Cerzza, & Turati, 2006]). The following section reviews two representative domain independent semantic search infrastructures, that is, TAP (Guha, 2003) and KIM (Kiryakov et al., 2005). The next section describes related work to semantic association analysis, and then discusses the community-based approach to discover and analyze semantic associations.

### Semantic-Enhanced Search

TAP is an infrastructure for sites to publish structured data, and for applications to consume this data. The main ontology in TAP is a broad knowledge base containing concepts such as people (e.g., musicians, athletes), organizations, places, and products, and a set of properties. TAP improves search results by utilizing semantic-enhanced and context-based approaches. It provides a simple mechanism to help the semantic search module to interpret the denotation of a query. This is important because one needs to deal with concepts ambiguity in the real world based on semantics rather than focusing only on the syntax. It also enhances the search results by considering search context and exploring closely related resources within a specified context.

KIM introduces a holistic architecture of semantic annotation, indexing, and retrieval of documents. Its aim is to provide fully automatic annotation methods using information extraction methods. The annotation framework is built upon a semantic repository which consists of two major components: a light-weight upper level ontology and a broadly-populated knowledge base. The ontology includes generic classes representing real world entities across various domains, such as people, location, organization, as well as their attributes and relationships. The advantage of using light-weight ontology is that it is easier to understand, build, verify, maintain, and get consensus upon. The entity annotation for an object in KIM includes both a reference to its most specific entity type in the ontology, and a reference to its entity description in the knowledge base. KIM improves search precision and recall by indexing and searching documents using the semantic annotation. Compared to traditional information search and retrieval approaches, semantic search and ontology-based information retrieval methods demonstrate several salient advantages, such as being able to incorporate query denotation, context-based exploration, enhanced data integration, query expansion, and consequently better recall and precision.

## Semantic Association Analysis

The conventional- and semantic-supported search approaches typically respond to user queries by returning a collection of links to various resources. Users have to verify each document to find out the information they need; in most cases the answer is a combination of information from different resources. Relations are at the heart of Semantic Web (Anyanwu, Maduko, & Sheth, 2005). Focusing on Semantic Web technologies, the emphasis of search will shift from searching for documents to finding facts and practical knowledge. Relation searching is a special class of search methods which is concerned with representing, discovering, and interpreting complex relationships or connections between resources. As the development of semantic-enhanced and community-based Web search continues, more ontologies are developed and used across the domains. This also leads to deployment and support of more semantic data in different systems and applications. One can expect that connections between entities on Semantic Web will become more complex and obscure. However, in some applications it is extremely important to have the ability to discover those distant and obscure connections between resources (Sheth et al., 2005). There are some existing work which have demonstrated utilizing semantic association analysis in different applications, for example, detection of conflict of interest (COI) (Aleman-Meza et al., 2005), detection of insider threat (Sheth et al., 2005), terrorism identification and flight security (Aleman-Meza et al., 2005; Sheth et al., 2005), and so forth. The semantic association discovery and analysis also plays a significant role in business intelligence, antimoney laundry, gene relationship discovery, medicine, and geographical systems.

Sheth et al. (2005) discuss an algorithm developed to process different kinds of semantic associations using graph traversal algorithms at the ontology level. The relationships between two entities in the results of a semantic query could be established through one or more semantic associations. In this case the semantic associations could be represented by a graph which shows the connection between entities. It is also important to process and prioritize the semantic association based on user preferences and the context of search. There are also ranking algorithms proposed based on different metrics to grade the semantic association (Anyanwu et al., 2005; Barnaghi & Kareem, 2007).

Baziz, Boughanem, Pasi, and Prade (2006) compare the classical keyword-based approach with the concept-based approach for information retrieval, and propose a fuzzy approach for query evaluation. The target documents and user queries are conceptually represented by means of weighted structures, and they are associated to an ontology. The documents and queries are compared with the fuzzy model based on the computation of degree of inclusion of features. The method holds its importance since the documents are retrieved based on conceptually-related documents, even if it does not contain weighted query keywords explicitly.

There are also ongoing researches to incorporate community interests and similarities amongst different groups and individuals to provide enhanced search results. The next section discusses the main issues of a community-based approach for a Web search.

## Social Network Analysis

A social network indicates the ways in which nodes (e.g., individual or organization) are connected through various relationships to each other. The social network analysis, in this context, is a way of processing and interpreting the nodes and relationships to realize mutual interests and connection between different groups and individuals in a community. Several studies have been carried out—mostly focusing on graph theory and statistical methods—to analyze relationships and connections in a social network. The Semantic Web technologies support social network analysis by providing explicit representation of the social network information (Ding, 2004). In recent years some social network studies have been conducted to adapt Semantic Web technologies in the social network research (Ding, 2004; Ding, Zhou, Finin, & Joshi, 2005; Matsuo, Hamasaki, Nakamura, Nishimura, Hasida, Takeda, et al., 2006; Mika, 2005). Flink (Mika, 2005) is one of the early works in this area that employs Semantic Web technologies for reasoning personal information extracted from heterogeneous sources, including Web pages, e-mails, publication archives, and FOAF profiles. It presents the professional work and social connectivity of researchers in the Semantic Web area. In a similar context, Ding et al. (2005) performed a network study based on the "*foaf:knows*" relation in a dataset constructed from online FOAF documents using Swoogle and other tools. The study primarily concentrates on basic graph features of the extracted social network. The authors state that the social network could be an implicit trust network to support applications such as knowledge outsourcing and online communities.

Matsuo et al. (2006) utilize extraction, analysis, and integration of multiple social networks from communities with similar characteristics. The work indicates the efficiency and significance of the research for locating experts and authorities, calculating trustworthiness of a person, detecting relevance and relations among people (e.g., COI detection), promoting communication, information exchange and discussion, ontology extraction by identifying communities, and so on.

In recent years, the Semantic Web community has been very active in promoting different applications of the proposed technologies in various application domains. The above discussed work outlines examples of several ongoing researches on semantic-enhanced and ontology-based information search and retrieval. They demonstrate the significance of semantic association analysis in different applications.

## PROCESSES AND COMPONENTS OF SEMANTIC ANALYTICS

The semantic association analysis consists of several key processes and components. We discuss ontology development, data set construction, semantic association discovery, semantic association ranking, results presentation, and performance evaluation in the following sections, respectively. However, there are also other important issues such as entity disambiguation, data set maintenance, and so forth. As our focus is to describe semantic association identification and interpretation, we will not discuss these issues in this chapter.

### Ontology Development

The creation of an **ontology** has been made easier because of the availability of some open source ontology creation tools like protégé,[9] ontology libraries such as DAML[10] and SchemaWeb,[11] and search engines (e.g., Swoogle[12]). For example, if one wants to create an ontology about countries, the creator does not need to create it from scratch, instead, the creator may find an existing ontology (or one in a similar context) through browsing the ontology library or using the ontology search engines. Further more, existing authoritative ontologies or vocabularies such as FOAF and Dublin Core[13] also contribute to the ontology engineering process. The adoption of existing vocabularies also promotes reuse and interoperability of an ontology.

## Data Set Construction

The data set, in some papers referred to as test bed (Aleman-Meza, Halaschek, Sheth, Arpinar, & Sannapareddy, 2004) or knowledge base (Guha, 2003), (Kiryakov et al., 2005), is in fact a collection of instances for a created ontology, or in Semantic Web jargon, the population of an ontology. The semantic association discovery is performed upon the test bed. The data usually come from different sources and are connected through relations defined in the ontology. Existing data sets for semantic analytics applications have some characteristics as summarized by Aleman-Meza et al. (2006). The data should be selected from highly reliable Web sites which provide data in structured, semistructured, or parse-able unstructured form, or with database backend. Structured data are preferred (i.e., RDF or OWL). Semistructured or parse-able unstructured data (i.e., XML) can be transformed to structured data using xPath or XSLT. Data with rich metadata and relations are preferred. For example, for a "computer scientist" class, the source also provides "address" and "country" attributes as well as some relations with other classes, such as "research area," "publication," and "organization." The data set should have rich relations and a large amount of instances which are highly connected.

## Semantic Association Discovery Algorithms

Semantic association discovery can be seen as a special class of semantic search aiming to find out complex relationships between entities. The problem can be generalized as enumerating all possible paths between any two nodes in a semantic graph. The search is performed using ontologies and semantic data sets. The structure of the ontology constrains the possible paths that one can take from one node to another. Typically, the structure of the ontology or relation between classes is simple; however, the relations between instances in the knowledge base (i.e., instances) might be very complicated depending on the connectedness of the graph.

RDF and OWL, which, in most cases, are the data model of the aforementioned data sets, have been visualized as a model of directed labeled graph and the problem of finding the semantic associations has been generalized as finding all paths between any given two entities in the graph (Aleman-Meza et al., 2004). A method for finding path associations using a recursive depth-first search is described by Sheth et al. (2005). The search is performed on the ontology, which can be viewed as a schema of the RDF instance data, to prune the search at the data level. The reason is that the number of entities in the schema is much less than in RDF data, thus the complexity of algorithm is reduced to a greater extent. The result of a search is a set of possible paths between C1 and C2 (both C1 and C2 are classes in an ontology) at the schema level and it is used to guide the search for paths between two entities e1 and e2 (e1 and e2 have "*rdf:type*" C1 and C2, respectively) at the data level. The search for join association is more complicated and is based on the path association algorithm. First it tries to find all paths between two classes and every other class at the schema level; the result is two sets of path. Then it compares every path in one path set and the other path set. If there is an intersection (e.g., joined node), then these two paths end at the same node in the schema. Finally, the result is used to perform a search at the data level.

## Semantic Association Ranking

A ranking mechanism is an important part of a search engine. A good ranking algorithm reflects the cognitive thought of human beings towards the ranking of real world objects according to their perceived importance. The PageRank algorithm (Anyanwub& Sheth, 2003) contributes to Google's success and it is one of the most important reasons that most people prefer to use it. Most of the current search engines rank documents based on vector space model and citation analysis (The PageRank algorithm can be seen as a variation of the citation-based approach) (Brin & Page, 1998). In semantic association analysis, an important task is incorporating the most meaningful associations out of all detected relations. However, new ranking algorithms need to be developed in order to utilize the advantages of Semantic Web technologies. SemRank (Anyanwu et al., 2005) describes a ranking algorithm which uses semantic and statistical metrics. The semantic metrics consists of context, subsumption, and trust. The value of trust is assigned according to the trustworthiness of source that made a statement. The statistical metrics includes rarity, popularity, and link length. The weight of a semantic association is the accumulated value of six factors. In a similar work, we have developed a relation robustness evaluation algorithm for semantic associations which computes the weight of an entity in the knowledge base with the queried entity using context, association length, and popularity metrics (Barnaghi & Kareem, 2007). It is worthwhile to note that the objects are being ranked in this algorithm rather than the semantic associations.

## Results Presentation

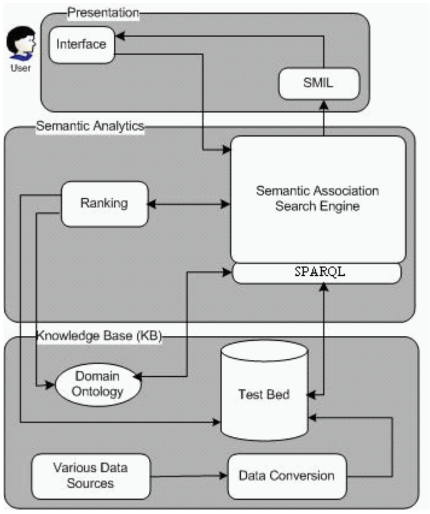
The presentation of the results is not the focus of this chapter, but for the sake of unification we briefly describe it here. In an existing work, the discovered semantic associations are presented as a list of property sequence with ranked value or as a 3D graph representation (Baclawski & Niu, 2005). The identified semantic associations could be presented to users in a meaningful way, which is able to help users understand the meaning of entities. We have implemented an automated hypermedia presentation generation engine, called MANA, to construct hypermedia presentations based on documents relating to entities in semantic associations (Deligiannidis, Sheth, & Aleman-Meza, 2006).

The entities are hyperlinked to those documents which are able to provide external explanations that help users to explore relevant information regarding a submitted query. Figure 1 shows different components and layers in an ontology-based information search, retrieval, and presentation.

## SEMANTIC SEARCH ARCHITECTURE

We have built a test bed upon which our semantic association discovery and ranking algorithms are evaluated. The system uses a domain ontology and semantic annotated resources in fine arts domain. When applied to fine arts (in particular, painting), analytical modules of the MANA are typically able to extract the name of an artwork, its creator, details on its features (i.e., style, period, materials), its image file, and possibly some information about other related documents. Paintings, as well as other objects, are associated with other

*Figure 1. Different components in a semantic-enhanced information search and retrieval system*



artworks in terms of creator, style, material, and so forth. Examples of information "triples" (in free-text form, to avoid syntax issues) are listed in the following:

*"Painter X Paints Painting Y".*
*"Painting Y's style is cubism".*
*"Painting Y is a contemporary artwork".*
*"Contemporary artworks have specific features".*
*"The features of contemporary art are described in Z".*

The logical reasoning could be outlined based on stated triples. For example:

*X painted Y, Y's style is cubism, Z describes cubism [implies that] "X's work style is described in Z".*

Other spatial and temporal similarities between entities could also be considered in a semantic association discovery. For example, two painters living in the same period of time and same geographical location would be related based on some queries and contexts. The domain ontology's hierarchical structure is shown in Figure 2.

The system uses an inference engine which is responsible for discovering the relationships between entities, and calculating and assigning weights to selected objects based on the proposed ranking mechanism. The details of the ranking mechanism are described by Barnaghi and Kareem (2007). The work introduces a knowledge-driven methodology which provides a multicriteria search method based on different attributes and ranking metrics. Figure 3 illustrates the results of a query (e.g., Cubism) and the associated ranking weights for the results obtained from the semantic association analysis.

## CONCLUSION

This chapter describes research and issues in the semantic analytics area and, in particular, discovery and interpretation of complex relations between entities in a knowledge base. In the Semantic Web, semantic analytics demonstrate significant importance in various application domains by enabling search mechanism to discover and process meaningful relations between infor-

*Figure 2. The domain ontology concepts*

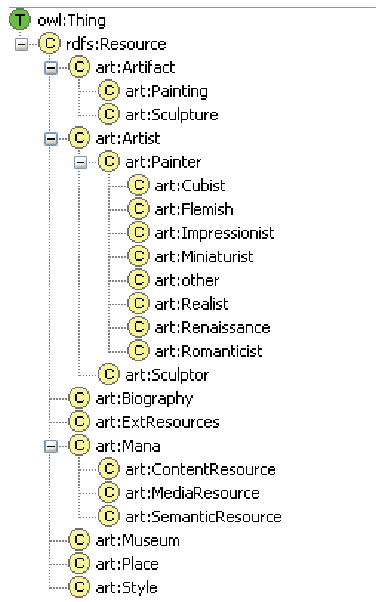
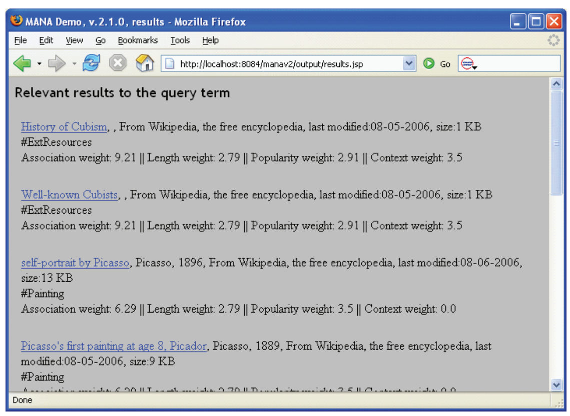



*Figure 3. Semantic search and ranking*



mation resources. In this chapter, we introduced the problem of semantic association identification, and we described the enhanced search methods based on semantic and community-based approaches. A prototype of semantic association-based information search and retrieval is also described through the chapter.

Future work focuses on automated semantic annotation and constructing a social network and community information for academic publication archives.

## REFERENCES

Aleman-Meza, B., Halaschek, C., Arpinar, I. B., & Sheth, A. (2003). *Context-aware semantic association ranking* (Tech. Rep. 03-010). University of Georgia, LSDIS Lab, Computer Science.

Aleman-Meza, B., Halaschek-Wiener, C., Sahoo, S. S., Sheth, A., & Arpinar, I. B. (2005). *Template based semantic similarity for security applications* (Tech. Rep.). University of Georgia, LSDIS Lab, Computer Science Department.

Aleman-Meza, B., Halaschek, C., Sheth, A., Arpinar, I. B., & Sannapareddy, G. (2004). SWETO: Large-scale Semantic Web test-bed. In *Proceedings of the 16th International Conference on Software Engineering & Knowledge Engineering Workshop on Ontology in Action, Knowledge Systems Inst.* (pp. 490-493).

Anyanwu, K., Maduko, A., & Sheth, A. (2005). SemRank: Ranking complex relationship search results on the Semantic Web. In *Proceedings of the WWW 2005 Conference*.

Anyanwu, K., & Sheth, A. P. (2003). ρ-Queries: Enabling querying for semantic associations on the Semantic Web. In *Proceedings of the 12th International World Wide Web Conference*.

Baclawski, K., & Niu, T. (2005). *Ontologies for bioinformatics.* MIT Press.

Barnaghi, P. M., & Kareem, S. A. (2006). A flexible architecture for semantic annotation and automated multimedia presentation generation, In *Proceedings of the 1st International Workshop on Semantic-enhanced Multimedia Presentation Systems*.

Barnaghi, P. M., & Kareem, S. A. (2007). Relation robustness evaluation for the Semantic associations. *The Electronic Journal of Knowledge Management*, *5*(3), 265-272.

Baziz, M., Boughanem, M., Pasi, P., & Prade, H. (2006). A fuzzy logic approach to information retrieval using an ontology-based representation of documents. *Fuzzy logic and the Semantic Web.* Elsevier.

Berners-Lee., T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 35-35.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. In *Proceedings of the WWW 1998 Conference* (pp 107-117).

Broekstra, J., Kampman, A., & Harmelen, F. V. (2002). Sesame: An architecture for storing and querying RDF and RDF schema. In *Proceedings of the 1st International Semantic Web Conference* (LNCS 2342, pp. 54-68). Springer-Verlag.

Celino, I., Valle, E. D., Cerzza, D., & Turati, A. (2006). Squiggle: A semantic search engine for indexing and retrieval of multimedia content. In *Proceedings of the 1st International Workshop on Semantic-Enhanced Multimedia Presentation Systems.*

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are ontologies, and why do we need them? *IEEE Intelligent Systems*, 20-26.

Craven, M., DiPasquo, D., Freitag, D., McCallum, A., Mitchell, T., Nigam, K., et al. (2000). Learning to construct knowledge bases from the World Wide Web. *Artificial Intelligence,* 69-113.

Dayal, U., Kuno, H., & Wilkinson, K. (2003). Making the Semantic Web real. *IEEE Data Engineering Bulletin*, *26*(4), 4.

Deligiannidis, L., Sheth, A. P., & Aleman-Meza, B. (2006). Semantic analytics visualization. In *Proceedings of the Intelligence and Security Informatics, ISI-2006* (pp. 48-59).

Ding, L., Finin, T., & Joshi, A. (2004). Analyzing social networks on the Semantic Web. *IEEE Intelligent Systems*, *8*(6).

Ding, L., Finin, T., Joshi, A., Pan, R., Cost, R. S., Peng, Y., et al. (2004). Swoogle: A search and metadata engine for the Semantic Web. In *Proceedings the 13th ACM international Conference on Information and Knowledge Management* (pp. 652-659).

Ding, L., Zhou, L., Finin, T., & Joshi, A. (2005). How the Semantic Web is being used: An analysis of FOAF documents. In *Proceedings of the 38th International Conference on System Sciences* (pp. 113.3).

Falkovych, K., Werner, & Nack, F. (2004). Semantic-based support for the semi-automatic construction of multimedia presentations. In *Proceedings of the Interaction Design and the Semantic Web Workshop, the 13th World Wide Web Conference*.

Guha, R., & McCool, R., (2003). TAP: A Semantic Web test-bed. *Journal of Web Semantics*, *1*(1), 81-88.

Guha, R., McCool, R., & Miller, E. (2003). Semantic search. In *Proceedings of the WWW 2003 Conference*.

Golbeck, J., Fragoso, G., Hartel, F., Hendler, J., Oberthaler, J., & Parsia, B. (2003). The National Cancer Institute's thésaurus and ontology. *Journal of Web Semantics*, *1*(1).

Karvounarakis, G., Alexaki, S., Christophides, V., Plexousakis, D., & Scholl, M. (2002). RQL: A declarative query language for RDF. In *Proceedings of the 11th World Wide Web Conference* (pp. 592-603).

Kiryakov, A., Popov, B., Terziev, I., Manov, D., & Ognyanoff, D. (2005). Semantic annotation, indexing, and retrieval. *Elsevier's Journal of Web Semantics*, *2*(1).

Matsuo, Y., Hamasaki, M., Nakamura, Y., Nishimura, T., Hasida, K., Takeda, H., et al. (2006). *Spinning multiple social networks for Semantic Web.* American Association for Artificial Intelligence.

Mika, P. (2005). Flink: Semantic Web technology for the extraction and analysis of social networks. *Journal of Web Semantics Science*, *3*(2-3), 211-223.

Rutledge, L., Alberink, M., Brussee, R., Pokraev, S., van Dieten, W., & Veenstra, M. (2003). Finding the story: Broader applicability of semantics and discourse for hypermedia generation. In *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia* (pp. 67-76).

Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The Semantic Web revisited. *IEEE Intelligent Systems*.

Sheth, A., Aleman-Meza, B., Arpinar, I. B., Bertram, C., Warke, Y., Ramakrishnan, C., et al. (2005). Semantic association identification and knowledge discovery for national security applications. *Journal of Database Management on Database Technology*, *16*(1), 33-53.

SPARQL Query Language for RDF (2007, November 12). *W3C proposed recommendation*. Retrieved August 17, 2008, from http://www.w3.org/TR/rdf-sparql-query/

## KEY TERMS

**Information Search and Retrieval:** Finding out queried information and its descriptive details.

**Ontology:** Object description and relationship between objects in a domain.

**Semantic Association Analysis:** Discovering complex and meaningful relationship between objects.

**Semantic Web:** A Web of relations between resources together with well-defined metadata attached to those resources.

## ENDNOTES

1. http://www.myspace.com
2. http://www.youtube.com/
3. http://www.flickr.com/
4. http://www.w3.org/TR/owl-features/
5. www.w3.org/RDF/
6. http://xmlns.com/foaf/0.1/
7. http://www.geneontology.org/
8. http://www.opencyc.org/
9. http://protege.stanford.edu
10. http://www.daml.org/ontologies
11. http://www.schemaWeb.info
12. http://swoogle.umbc.edu
13. http://dublincore.org