*Online Journal of Applied Knowledge Management*
A Publication of the International Institute for Applied Knowledge Management

*Volume 3, Issue 1, 2015*

# The role of semantics in searching for information on the Web

**Ilona Pawełoszek,** *Czestochowa University of Technology, ipaweloszek@zim.pcz.pl*

**Jędrzej Wieczorkowski,** *Warsaw School of Economics, jedrzej.wieczorkowski@sgh.waw.pl*

## Abstract

*As the amount of data in the World Wide Web grows, the Internet becomes the biggest and often the primary information resource for many individuals and organizations. To make good use of the data, it is essential to provide effective and intelligent search capabilities. The users more and more often require direct and unambiguous answers, which are often not explicitly present in any document. The Web users also are not willing to learn complex query languages and need interfaces that will be easy to use and as close as possible to natural language. The aim of this paper is to illustrate the advantages that semantics brings to the users of contemporary Web search engines. The concepts of semantic Web and semantic search engines have been described along with some issues related to their development and usage, such as linked data, semantic query interfaces and the ways of publishing semantic data. The authors also explore the semantic features of contemporary search engines and indicate their future directions of development.*

**Keywords**: Semantic Web, search engine, linked data, semantic query interfaces, semantic data publishing

## Introduction

The World Wide Web has come a long way since its public introduction in April 1993. From that moment the WWW technology has been available for anyone to use on a royalty-free basis (History of the Web 2015). The number of Web pages sprung out rapidly and it is constantly growing. Also the internet services increase in number and sophistication. From once static pages, Internet sites have evolved and specialized in offering interactive and device-customized content. Despite the rapid development of the Web technologies, still the main purpose of the Web is to be exploited by human users, therefore the techniques of building websites mostly have been oriented on visual presentation of information and data.

More and more knowledge workers use the Web as a primary source of information in their daily tasks. A very important issue is that information in the Web documents often lacks structure and is many times rather complicated to read and understand. Data are often presented in tables interspersed with text in lengthy articles, whereas the user needs to interpret and extract the meaningful pieces of data. The internauts search for information, browse the content, download and make use of it in off-line and on-line applications. The use of search engines is the most popular way of accessing information on the Web. Since the very inception, the search engines have been single points of access to different kinds of Web content. The semantic Web technologies, which help to organize information in a meaningful way that enables reasoning as well by computers as by human users are promising and potentially important tools to make extensive use of Web resources.

The aim of this paper is to illustrate the advantages that semantics bring to the users of contemporary Web search engines. We also discuss some difficulties in applying the semantic approach. Section 1 introduces the reader to the idea of the semantic Web and semantic search. The discussion on semantic features already implemented in contemporary search engines on the example of google.pl has been presented in the section 2. In the section 3 the authors attempt to specify the categories of search- related problems that can be solved by exploiting semantics. Semantic Web as an emerging technology platform also creates challenges for companies trying to build their online presence and position their websites in the search engine results. The ways in which to present the company's data on the Semantic Web are also described in section 3. Finally, we conclude our considerations and follow that up with the future work that can be done to enhance semantic capabilities of today's Web applications.

## The semantic Web and semantic search – motivation and definition

In most Web search engines the list of retrieved documents is displayed after typing the keywords specifying the user's information-seeking goals. In the present scenario well-grounded keyword searching seems outdated and not responding to users' needs. The present generation search engines present a large amount of search results to the user in response to the query. Depending on the keywords and the way the query is posed some of the results are more or less irrelevant, and the user has an ordeal in sifting through the result sets to harvest some information of his interest (Mittal, Singh & Sachdeva 2011). Sometimes the answer to the user's query is not directly present on any Web document. In such a case it can only be inferred by analyzing and extracting data from many resources. This task, unfortunately so far must be done manually. The limitations of contemporary search engines have turned the researchers to seek for new alternatives which led to the emergence of the semantic approach to organize and search for information resources. Since the idea of semantic Web had been introduced in 2001 by Tim Berners-Lee (Berners-Lee, Hendler, & Lassila 2001). Many interesting and progressive technologies have emerged to pursue the vision of a Web that will not only simply contain data, but semantic information that is machine-processable in a meaningful way (Stoermer 2006). Semantics is the study of the meaning of words, phrases and sentences that are associated with the real-world objects and concepts. Semantic data structures are inspired by human natural language constructs in a way that they present knowledge as simple indicative sentences having subject, predicate and object.

There are many advanced solutions that have been developed for many years, with the aim to help computers to interpret human language, and communicate in a human-like ways, these are artificial intelligence techniques such as natural language processing. However, they had not advanced to a level that could be usable for the machine interpretation of Web documents authored by humans. The original semantic approach envisioned by T. Berners-Lee and his coworkers was to keep the solutions simple enough so they can be used by webmasters without experience in knowledge engineering. Therefore the concept of the semantic web is based on a simple idea of annotating Web documents with Extensible Markup Language (XML) markup which is in many aspects similar to well-known and widely accepted Hypertext Markup Language (HTML) syntax.

The semantic annotations should correspond to some clearly defined and shared knowledge structure, which is called ontology.

Ontology is a specification designed to categorize and help explain the relationships between various concepts of in the given area of knowledge and research. The most widely accepted definition of ontology in the context of knowledge sharing and information science was proposed by T. Gruber (1995), he says that ontology is a specification of a conceptualization. The interpretation of this definition can be that

- Ontology provides a common understanding of a particular domain, or field, of study, and ensures a shared ground for those who study the domain.
- Ontology is useful for organizing concepts, information, and ideas, it helps to show the relations between concepts.
- Ontology can be formalized which means that it can be read and understood by computer applications.

The role of ontology in knowledge engineering is to enable construction of a domain model by describing a set of concepts and relations between them (Jurczyk & Pawełoszek 2015).

The language for annotating documents is Resource Description Framework (RDF) and for ontology definition RDF Schema (RDFS) or Web Ontology Language (OWL) are used. The resources on the Web, i.e., the Web pages, are crawled and looked for annotations done on them, if any. A semantic search engine uses ontologies to derive semantic associations among different words and concepts. These annotations are then compared to that ontology with which they have been tagged (Mittal, Singh & Sachdeva 2011).

According to a software company Cambridge Semantics Inc. headquartered in Boston, Massachusetts: The Semantic Web is a set of technologies for representing, storing, and querying information. Although these technologies can be used to store textual data, they typically are used to store smaller bits of data (Cambridge Semantics 2014). Essentially, the semantic Web focuses on pulling specific data (i.e., numbers, dates, locations) from multiple heterogeneous sources to answer directly the users' questions. Therefore the semantic Web is often described as a Web of Data. In different words: it works more like a huge database than a collection of textual documents. To make the Web of Data a reality, it is important to have the data on the Web available in a standard format, reachable and manageable by Semantic Web tools (W3C 2015).

The data extracted from Web documents, describe some real-world object, a person or a concept are generally referred to as entities. So called entity search fully reflects the idea of the semantic web, which is searching for the entities by asking about their attributes instead of searching for text describing the entity. T. Doszkocs (2012) explains the idea of semantic search as "a search or a question or an action that produces meaningful results, even when the retrieved items contain none of the query terms, or the search involves no query text at all". A good example is a question about a person's age. The user can specify a query as follows: "how old is person X". The answer may not be directly written in any document, although it can be deduced from the

data about the person's X Birthdate. The semantic search requires reasoning capabilities and knowledge base containing ontologies.

## Semantics in modern search engines

When studying the publications discussing the need for semantic Web and semantic search it is easy to find the opinions that using search engines is difficult, the search results are mostly irrelevant and the user has to browse many documents to find what he really needs. The common opinion is that the search engines are "unintelligent" and cannot understand the context of the user's query. The current generation of search engines is severely limited in its understanding of the user's intent and the Web content and consequently in matching the needs for information with the resources on the Web (Mika 2008). However, the study of the most popular search engines reveals that the aforementioned facts are not always so obvious. In the recent years the search engines are increasingly moving towards semantically enhanced results for user queries.

Contemporary search engines allow the user to specify the query and obtain better/more relevant results by offering many semantics-related options narrowing the results or changing their order. It is possible to specify a search query for text, images, videos and services such as maps, news, books etc. Users can also choose the desired language or domain (i.e. pcz.pl), however the latter requires knowledge how to use advanced options. The search engines offer many options and the possibility to use search operators to make the user's query more specific and to obtain better (more relevant) results. Moreover the advanced search options are quite easy to use (assuming that the user is aware of their existence).

Some of the contemporary search engines are context-sensitive. This means they already exploit some semantics. Very often it is easy to observe that search results are relevant to the user's location – for example, when the user localized in Katowice will type the query "restaurant" the first search results will probably be the links to the websites of the restaurants nearby the user's location.

The next "intelligent" feature of modern search engines is the awareness of synonyms and spelling mistakes. For example, if we type a query "information searching" (without quotes) the Google search engine also highlights the word "seeking" which is synonymous to "searching" (figure 1.).

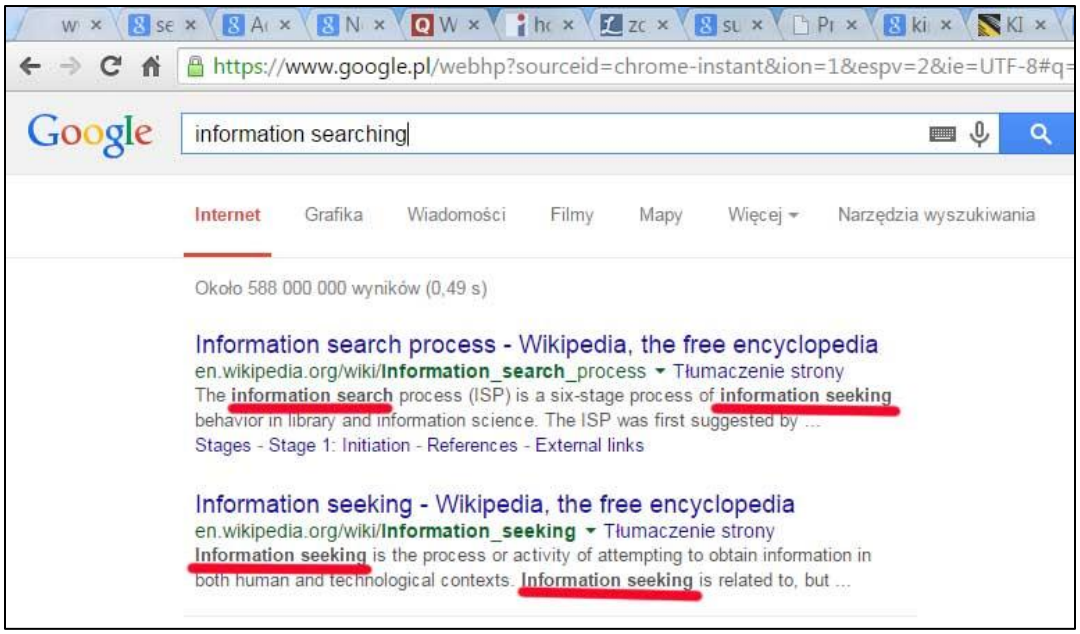


Figure 2. Example of the awareness of synonyms in Google search results

Source: Chrome browser and http://google.pl

While querying Google i.e. for the word "fox" it will return many results, but their sequence will be dependent on the language the user speaks. In case of Polish language the first search results will pertain to Fox TV channel, Fox Broadcasting Company. Although there is some inconsistence between text and image search results. The image search will return the pictures of fox – the animal. In many cases typing the foreign word or phrase will result in links to dictionaries and translators. The search engines can also serve as measurement and currency converters.

One of the significant changes in the ways of interaction with search interfaces is the conversational search presented by Google in its new Chrome browser. Conversational search is a new kind of philosophy for human-computer interaction. The principle behind it is that the user can speak a sentence into a device, and that device can respond with a full sentence (Techopedia 2015). The idea of conversational search is to make the human-computer interaction more natural and convenient for people. The feature is especially meant for mobile devices, while voice interface is more convenient for users while being on the go.

Conversational search exploits the natural language processing and speech recognition (Google's Voice Search) to formulate the query string. The generated query string is then analyzed to extract keywords. The keywords can be interpreted by semantic search mechanisms, which try to formulate the semantic query, usually in SPARQL (W3C 2008) language. Then, if possible, the full-sentence answer for the query is generated. The possibility of performing semantic search and answering with the full sentence is dependent on many conditions, the most important are: availability of semantic data sources for a given domain, the scope of semantic data sources and

underlying ontologies, and the accurateness of Automatic Speech Recognition (ASR) techniques. As the research shows the ASR engines may provide satisfying transcription (Silber-Varod, Geri 2014).

Another interesting concept (although not yet fully implemented) is called "previous query" – the idea is that the conversational search remembers the user's question and takes it into consideration giving the answers for the consecutive searches. For example, if someone searches for [Poland] and then [president] after that, the results should be altered to take the previous query into account. To some degree, it will be as if the second query was for [president of Poland].

## Semantic search problems and solutions

Semantic search problems can be regarded from the point of view of search engine developer, user and webmaster. From the point of view of application developer the basic practical problem in building semantic applications is how to extract entities from the text. If the entities are extracted the next step is to interpret the values of their attributes, compare them to the domain ontologies, infer the answer and display it to the user.

Extracting data from the webpage is not a trivial task. There are two approaches to building semantic applications. The first one assumes that the text contains markup explicitly identifying the entities and their attributes. In this case the search engine consumes structured data. The second approach (the implicit one) assumes that the search engine deals with unstructured text, therefore the advanced algorithms (stochastic, NLP) must be used to derive or infer the data.

Both the approaches have some advantages and pitfalls. The first - explicit approach requires from the webmasters the efforts of annotating webpages. The task of annotating is not so straightforward because it must be done manually, some programming environments do not have functions to support the user in adding annotations in the specific markup standard. The second issue is the decision about which ontology to choose as the reference model for annotations. The good thing about explicit approach is the fact that the annotations made by human are usually correct in the logical sense. Although the search engine based on this approach can only operate on the set of documents which had been annotated.

The second - implicit approach does not require annotated documents, therefore theoretically the search engine can work on any text documents found on the Web. The disadvantage of this approach is that it strongly depends on artificial intelligence algorithms and text mining which are not perfect and usually are restricted to some specific domain. Sometimes the context can be misunderstood and the entities or their attributes will not be properly identified. The implicit approach is often used in the cases such as: finding documents on a particular subject or domain, determining the sentiment of the documents (positive or negative context).

The third consideration of the semantic application developer is which ontologies to choose or if there are no suitable ontologies he may decide to build it on his own. The developers often look

for appropriate ontologies that can be integrated into their systems, rather than develop new ontologies from scratch. The choice of ontology may, have a major impact on the performance of the semantic application, including the quality of the results (Tan & Lambrix 2009). In case of building large ontology it is possible to integrate several existing ontologies describing portions of the large domain. One can also reuse a general ontology or taxonomy, such as the UNSPSC (United Nations Standard Products and Services Code) or PKWiU (Polish Classification of Goods and Services), and extend it to describe the domain of interest (Noy & McGuiness 2001).

The success of semantic applications strongly depends on its query interface. Users are not willing to get used to complicated kinds interaction. Therefore semantic search should not be much different than keyword search that the users are used to. Otherwise the search engine will not be used. On one hand the semantic Web query languages e.g. SPARQL (Prud'hommeaux & Seaborne 2008), SeRQL (Broekstra & Kampman 2004), RDQL (Seaborne 2004), are difficult to use, but on the other hand the strongly formalized graph-based approach represented by those languages allows for very precise query definition and addresses complex information needs. Among the aforementioned query languages SPARQL is the one that has been recognized as the de facto standard for the Semantic Web. The most user-friendly solution for query definition is undouble natural language, however it has unavoidably lower accuracy, as compared to systems with graph-based querying interfaces, which, in turn, are usually still too difficult for regular users (Styperek, Ciesielczyk & Szwabe 2014). Using semantic query languages poses two main challenges for the users, first of all they must be familiar with the language syntax, and the second important requirement is the knowledge of underlying ontologies. There are many propositions of semantic search engine interfaces. They usually present the tradeoff between the easiness of use and the semantic capabilities.

The architecture presented in (Wang et al 2008) proposes a solution to translate keyword queries to formal queries. The Authors leverage terms extracted from Wikipedia to enrich literals described in the original RDF data. This way, users need not use keywords that exactly match the RDF data.

Another interesting proposition (Bäurle 2011) combines and extends established components of other search user interfaces - namely keyword search, facets, proposals and breadcrumbs. The user interface provides only a single input field as it is also common for current keyword search applications. When something is typed, it automatically shows different proposals for words, relations, entities, or semantic classes that are used to build the search query.

From the perspective of the webmaster a crucial problem is how join the semantic Web, this means publishing the data for more effective discovery, automation, integration, and reuse. It is especially important for internet marketers and the SEO (Search Engine Optimization) professionals. At present the most popular search engine is Google, and it also is making great efforts to develop its semantic capabilities. It is more and more often that after posing a question in the natural language one can get the exact answer. For example a question about the age of Polish president Bronislaw Komorowski results in very precise answer displayed along with some other information about the person (Figure 2). Moreover the semantic features are

dependent on language version of the user's operating system. The example on Figure 2 also illustrates another important concept strongly related to semantic Web – the Linked Data, which is a collection of interrelated datasets on the Web. The query to the semantic search engine displays desired entity and its attributes, but also proposes some other entities somehow related to the one the user is asking about. In the case of the query for Polish president, the search engine displays also two important Polish politicians, and the president's wife. The entities are connected with some mutual relationships of different kinds. The concept of Linked Data is about using the Web to connect related data (Linked Data 2015). The Linked data concept emphasizes that not only the Semantic Web does need access to data, but relationships among data should be made available.


Figure 2. The display of semantic search results

Source: Chrome browser and http://google.pl

The entity (a person or a business unit) must comply with some requirements to be found by semantic search engine. The requirements for being displayed in semantic search results are different than in case of traditional SEO. There are three basic methods to serve linked data on the Web:

- Google Knowledge Graph (),
- publishing RDF files,
- labeling HTML content with microformats, microdata or RDFa.

Google Knowledge Graph is the secret behind the entity-based search and the straightforward immediate answers to the user's question posed to Google Search engine. The Google Knowledge Graph, includes information about the entities along with their relationships. The Knowledge Graph takes information from trusted data sources such as Freebase (2015),

Wikipedia (2015) and Google+ (2015) account, so it is recommended to join the communities and fill one's profile.

However recently Google has been developing a new generation base for semantic search which is called Knowledge Vault (Dong et. al. 2014) . The difference between Google's existing Knowledge Graph and the Knowledge Vault is the way that facts are accumulated. The Knowledge Graph acquires information from trusted sources. The Knowledge Vault is an accumulation of facts from Web content (obtained via analysis of text, tabular data, page structure, and human annotations) with prior knowledge derived from existing knowledge repositories. The major difficulty is that, by its very nature, the Semantic Web is a large, uncensored system to which anyone may contribute. This raises the question of how much credence to give each source.

The next generation solution employs supervised machine learning methods for fusing distinct information sources and probabilistic inference methods that compute probabilities of fact correctness (Dong et al 2014).

Publishing RDF files on the Web requires the webmaster to provide some external URLs pointing to them, so that crawlers can discover the new added data. The RDF data source can be added to the ESW Wiki datasets list, also links can be added from one's Friend of a Friend (FOAF) profile.

The lightway semantisation for non-technical users can be achieved by adding so called microformats to the HTML websites. Microformats (2015) are XHTML tags for marking up people, organizations, events, locations, blog posts, products, reviews, resumes, recipes etc. A data with microformats can be consumed and used by search engines, browsers, and other sites. Examples of well-established microformats are:

- hCalendar – events,
- hCard - people, organizations, contacts,
- rel-license - licensed content,
- rel-nofollow - links in untrusted 3rd party content,
- rel-tag - tag posts and pages by subject,
- XFN - social relationships and rel-me links among profiles for the same person,
- XMDP - define a microformat vocabulary / profile,
- XOXO – outlines.

Microdata is an HTML specification to label content to describe a specific type of information, for example, reviews, persons, or events and their properties. For example, an event has the properties, such as: venue, starting time, name, and category.

RDFa is an extension to HTML5, designed for labeling content. It uses simple attributes in XHTML tags (often  or ) to assign brief and descriptive names to entities and properties. RDFa annotation tools are added to popular web editing tools For example, a plugin

AKSW (2015) added for TinyMCE (2015) editor which is a platform independent web based Javascript HTML WYSIWYG editor often implemented in Open Source Content Management Systems platforms.

All the aforementioned techniques allow the webmasters to markup things like People, Places, Events, Recipes and Reviews. Search Engines and Web Services use this markup to generate better search listings.

## Conclusions

Although there are many initiatives the Semantic Web has not reached its full potential yet. The real added value from the semantic Web has not yet been achieved due to many problems that must be faced before the intelligent Internet can be exploited as a reliable source of knowledge that can be processed in a fully automatic way. On the other hand the problems become a challenge for development of new business models (for example, providers of ontologies, trust services, semantic cloud services, semantic search interfaces and many more).

A very interesting future concept is the semantic personal data locker (Lockerproject 2015), which is a kind of the user's digital profile. The personal data locker contains the user's detailed data (such as preferences, professional and private schedules, memos, etc.) which can be used as the context parameters for the semantic applications. It would be for example possible to integrate the data from the user's shopping list with the information from the Web about sales and promotions and thus provide fully relevant and interesting search results. The idea of personal data locker assumes storing all the data in the cloud. Although the data are to be under total control of the user the concept still gives rise to many privacy concerns.

More and more companies and public institutions are discovering the power of the Internet in reaching their customers and expanding their marketing influence into previously unreachable localizations. While the Web increases in volume and heterogeneity, it becomes increasingly important resource of information and the area of business activity. Even the smallest organizations are quickly realizing the need for internet presence which is undeniable because individuals and businesses today predominantly make buying decisions based on what they discover online. The Semantic Web is often envisioned as the future of the Internet, therefore; it seems essential for the companies to keep an eye on the development of new technologies and adopt them to gain competitive advantage on the electronic market.

## References

Bäurle, F. (2011). Master's thesis, University of Freiburg, Retrieved January 19, 2015 from https://ad.informatik.uni-freiburg.de/files/ui_semantic_fulltext_search

Berners-Lee, T., Hendler, J. & Lassila, O. (2001). *The Semantic Web,* Scientific American, May 2001, (pp. 29-37)

Broekstra, J. & Kampman, A. (2004). SeRQL: An RDF Query and Transformation Language,*3rd International Semantic Web Conference*. Japan.

Cambridge Semantics (2014, November 17). Semantic Search and the Semantic Web, Retrieved January 19, 2015 from http://www.cambridgesemantics.com/semantic-university/semantic-search-and-semantic-web

Dong, X.L., Murphy, K., Gabrilovich, E., Heitz, G., Horn, W., Lao, N., Strohmann,T., Sun S.& Zhang W. (2014). Knowledge Vault: A Web-scale approach to probabilistic knowledge fusion, Retrieved January 19, 2015 from http://www.cs.ubc.ca/~murphyk/Papers/kv-kdd14.pdf

Doszkocs, T. (2012). Semantics for universal search and discovery, Retrieved January 19, 2015, from http://www.slideshare.net/jendre/semantics-for-universal-search-and-discover7

Gruber, T., 1995, Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal Human-Computer Studies Vol. 43, Issues 5-6, November, p.907-928, Retrieved March 29, 2015 from: http://tomgruber.org/writing/onto-design.pdf

History of the Web (n.d.). Retrieved January 19, 2015 from: http://webfoundation.org/about/vision/history-of-the-web/

Jurczyk, M. & Pawełoszek, I., The Concept of Semantic System for Supporting Planning of Innovation Processes, Polish Journal of Management Studies vol. 11/2015 (accepted for publication)

Linked Data (2015). Retrieved January 19, 2015 from http://linkeddata.org/

Lockerproject (2015). Retrieved January 19, 2015 from http://lockerproject.org/

Microformats (2015). Retrieved January 19, 2015 from http://microformats.org/

Mika, P. (2008). Microsearch: An Interface for Semantic Search, Proceedings of the Workshop on Semantic Search (SemSearch 2008) at the 5th European Semantic Web Conference ESWC 2008 volume 334 of CEUR Workshop Proceedings.

Mittal, H., Singh, J. & Sachdcva, J. (2011). ARAGOG Semantic Search Engine: Working, Implementation and Comparison with Keywoixi-Based Search Engines, Pred, A., Dietz, J.L.G., Liu, K., Filipc, J. (eds.) *IC3K 2009*. CCIS. vol. 128. (pp. 177-186). Springer, Heidelberg.

Noy, N.F. & McGuiness, D.L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. Technical Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880 Retreived January 19, 2015 from http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html

Prud'hommeaux, E. & Seaborne, A. (2008). SPARQL Query Language for RDF. Retrieved January 19, 2015 from http://www.w3.org/TR/rdf-sparql-query/

Seaborne, A. (2004). RDQL - a query language for RDF. W3C member submission, Retrieved January 19, 2015 from http://www.w3.org/Submission/RDQL/

Silber-Varod, V. & Geri, N. (2014). Can automatic speech recognition be satisficing for audio/video search? Keyword-focused analysis of Hebrew automatic and manual transcription, Online Journal of Applied Knowledge Management A Publication of the *International Institute for Applied Knowledge Management* Volume 2, Issue 1, 2014. Retrieved March 25, 2015 from: http://www.iiakm.org/ojakm/articles/2014/volume2_1/OJAKM_Volume2_1pp104-121.pdf

Singhal, A (2012). Introducing the Knowledge Graph: Things, Not Strings. Google Official Blog. Retrieved March 6, 2015 from: http://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html.

Stoermer, H. (2006). Introducing Context into Semantic Web Knowledge Bases, *Proceedings of the CAISE*06 Doctoral Consortium*, Luxembourg, 5th June 2006. Retrieved January 19, 2015, from http://ceur-ws.org/Vol-263/paper3.pdf

Styperek, A., Ciesielczyk, M. & Szwabe, A. (2014). Semantic search engine with an intuitive user interface, *Proceedings of the companion publication of the 23rd international conference on World Wide Web companion*, International World Wide Web Conferences Steering Committee (pp. 383-384).

Tan, H. & Lambrix, P. (2009). Selecting an Ontology for Biomedical Text Mining, *Proceedings of the Workshop on BioNLP*, pages 55–62, Boulder, Colorado, Association for Computational Linguistics.

Techopedia (2015). Conversational Search, Retrieved January 19, 2015, from http://www.techopedia.com/definition/29608/conversational-search.

W3C (2015). Linked Data, Retrieved January 19, 2015, from http://www.w3.org/standards/semanticweb/data

Wang, H., Zhang, K., Liu, Q., Tran, T. & Yu, Y. (2008). Q2Seinantic: A lightweight keyword interface to semantic search, Bechhofer, S., Hauswirlh, M., Hoffmann, J., Koubarakis, M. (eds.) *ESWC 2008*. LNCS, vol. 5021, (pp. 584-598). Springer. Heidelberg.

Dong. X., Gabrilovich E., G Heitz, W Horn, N Lao, K Murphy, T Strohmann, Sun S. & Zhang, W. (2014). Knowledge vault: a web-scale approach to probabilistic knowledge fusion. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining.

***Web References***

Freebase (2015) https://www.freebase.com/ Retrieved march 25, 2015

Wikipedia (2015) http://www.wikipedia.org/ Retrieved march 25, 2015

Google + (2015) https://plus.google.com/ Retrieved march 25, 2015

AKSW (2015) http://aksw.org/Projects/RDFaCE.html Retrieved march 25, 2015

TinyMCE (2015) http://www.tinymce.com/ Retrieved march 25, 2015

W3C (2008) SPARQL Query Language for RDF W3C Recommendation 15 January 2008 Retrieved march 25, 2015 from: http://www.w3.org/TR/rdf-sparql-query/