Proceedings of the International MultiConference of Engineers and Computer Scientists 2012 Vol I,
IMECS 2012, March 14 - 16, 2012, Hong Kong

Integrating Knowledge Base Retrieval with Web Search using Semantic Roles

Pallavi Karanth and Kavi Mahesh

***Abstract*—An integrated search of the Web together with a local knowledge base is essential for specialized search applications in various domains for which rich local knowledge bases are available. Yet, there is no satisfactory solution to the problem either in research prototypes or commercial products that meets the requirements of high precision and effectiveness achievable only through the application of the semantic richness of the domain (as opposed to a general-purpose keyword-based search). We find three major problems in current solutions for such a domain-specific integrated search: (i) absence of links between conceptual elements of the knowledge base (i.e., the ontology) and textual elements in documents, (ii) drawbacks of keyword-based Web search engines such as lack of semantics and poor precision, (iii) the inadequacy of hyperlinks in terms of capturing the context and semantics of the links in addition to the lack of guidance for the user while navigating the Web together with a local knowledge repository. We present a solution to the above problems which caters to the needs of a knowledge worker in retrieving relevant information from both a local knowledge base and the Web using a richly interconnected link structure called *Superlinks* with well-defined *semantic roles*. We also illustrate a prototype implementation in the example domain of Indian medicinal herbs in Ayurveda.**

***Index Terms*—Knowledge base retrieval, Ontology, Search, Semantic role, Superlinks**

I. Introduction

A professional or a researcher in a domain of knowledge often faces the need to perform search in both a local, domain-specific knowledge repository and the World Wide Web. While the local knowledge base is likely to provide more relevant and domain-specific information, it is unlikely to be as comprehensive, current or varied as the Web. Despite the World Wide Web containing vast amounts of information, its usefulness depends on the "slippery slope of Semantics" [17]. The World Wide Web, and popular search engines used to retrieve information from it, although ideal for the information needs of ordinary users, fail to satisfy the needs of professional knowledge workers due to well-known problems of lack of precision and domain relevance. For instance, if an Ayurveda consultant needs to know about the medicinal benefits of the plant Arjuna, she can perform local search in an Ayurveda knowledge base. However, there is every chance that searching the Web can provide more recent or otherwise relevant additional information. It is necessary for the productivity of the knowledge worker to have a search engine that can concurrently search the Web and the local knowledge repository to provide the most relevant pointers to the search query.

Manuscript received December 30, 2011; revised January 17, 2012. This work is supported in part by a research grant from Visvesvaraya Technological University (VTU/Aca./2009-10/A-9/11626).

Pallavi Karanth is with P.E.S. Institute of Technology, 100 ft Ring Road, Banashankari 3rd Stage, Bangalore - 560085, India (e-mail: karanthpallavi@gmail.com).

Kavi Mahesh is with P.E.S. Institute of Technology, 100 ft Ring Road, Banashankari 3rd Stage, Bangalore - 560085, India (e-mail: drkavimahesh@gmail.com).

A knowledge worker often has the need to get not just generic descriptive content but highly relevant conceptual or technical information along with relevant explanations, examples, applications, images, videos and so on. Existing hyperlinks on the WWW together with methods such as the well-known Page Rank algorithm [16] are suitable for finding comprehensive but superficially relevant information about a given concept. The proliferation of search results and degree of freedom given to the knowledge worker in clicking on links to choose among them often leads to the user getting "lost in the hyperspace". On encountering a link, how does one decide if it is worth the distraction to follow the path that the link takes? Does the label appearing on the link tell us enough to decide? This problem could be called *informational myopia*. To overcome this, it is critical to have a linking structure that supports semantic and explicit interlinking of concepts to pieces of text which define, explain, illustrate or elaborate on the concepts [2]. The need of the hour is not only a search engine which offers the best information retrieval tools, but also an interactive, semantic browsing tool that can help knowledge workers navigate vast repositories of knowledge efficiently.

A. Major Drawbacks of Keyword-Based Search Engines

Currently, there are a number of search engines for internet users. These engines differ in both the indexing methods and the search algorithm used for generating results. At present, keyword-based search engines rank web pages based on the relative importance of the web page on the World Wide Web as well as a lexical score [1]. Due to this method, a lot of redundant or irrelevant information is returned to the users as search results with an increase in the amount of time and effort spent by the user in obtaining optimal search results [3]. Also when the search query is ambiguous, having multiple senses, the sense relevant to the current context of the search is not taken into account while performing the search. Although page rank algorithms (e.g., Google's [16]) provide a significant improvement in the precision and ranking of search results, it would be more effective if the results are categorized in terms of the sense of the search term or the semantic roles that the result plays, such as whether it is a "definition," "example," "elaboration," and so on. Such descriptions in the links enable knowledge

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

workers to follow the results more judiciously thereby improving their productivity. Another kind of search engine, known as, metasearch engine gives a common interface to use the combined capability of several search engines, but suffers from the same lack of semantics in query processing.

B. Inadequacy of Hyperlinks

Although hypertext-based browsing offers hyperlinks to jump from one page to another, this may not be the most effective way to navigate in the knowledge space of a particular domain. Often, the structure of knowledge can be seen as a graph or interconnected network of concepts and relations among them. In order to browse such a knowledge space, there is a critical need to attach semantics to the links so that they capture the conceptual relations between the items being linked. The idea of attaching semantics to a link originated in the development of semantic networks, back in 1956 at the Cambridge Language Research Unit as an "interlingua" for machine translation of natural languages [15]. Though this key idea has been used in many projects like WordNet [3], it has not been adopted successfully to the present problem.

In hyperlinks, relationships between documents are modeled on a very low level of abstraction [16], preventing effective navigation through the hyperlinked document space. Hyperlinks are also limited by the amount of text to which they can point. In a single link, they cannot point to multiple documents or web pages or precisely to any internal structure within a document. In the context of knowledge browsing, this can be a major limitation as the very need is to point to the most appropriate part of the document which may be a single word, group of words, paragraph or other textual structure, in order to enable the knowledge worker to navigate to the right place on the right page.

The three critical problems outlined above – lack of semantics in linking, limitations of keyword-based search and the static, binary and imprecise nature of hyperlinks – accentuate the need for a semantics-based integrated search and browsing mechanism for knowledge workers to obtain relevant information efficiently from both domain-specific local knowledge repositories and the World Wide Web.

II. Previous Work

A. Art of Linking - Microformat, Microdata, RDFa, "n"-hrefs

Approaches to enable both machines and humans to understand a web page include embedded metadata formats like Microformats, Microdata and RDFa. Microformat could be seamlessly integrated with the existing HTML pages, but only with the significant disadvantage of no transformation to JSON and no DOM API support. RDFa provides a clear mapping to RDF and also expresses typed literal properties such as units of measurement and specific kinds of numbers in simple ways, but is rather complex to use. Microdata being a relatively new standard for embedding mark-up in web pages is more extensible as it allows new semantic vocabulary to be embedded in web pages. These formats provide for embedding metadata into existing web pages but they do not support inter linking of conceptual elements and their relationships with textual documents in a domain knowledge base.

Embedding *"n" - hrefs* from a source to a target in a web page [14] is proposed as a technique to specify author's intuition while creating a link in a web page. This strategy enables one to add multiple target destinations for a link in a page, thereby addressing the need to link the same source to many different targets by maintaining conformance to existing standards such as XHTML, Microdata, RDF and compatibility with navigation on existing browsers. This approach caters to solving the "unidirectional inadequacy" of hyperlinks, although it does not render capabilities to link ontological elements with the textual elements of web pages.

B. Federated Search and Desktop Search

Federated search is an information retrieval technology which allows simultaneous search of multiple searchable resources [6]. Federated search differs clearly from the current proposed approach in that the search results are an aggregation of results from multiple searchable online databases in contrast with an integration of semantically related results. Desktop Search engines, on the other hand, aim at searching the local computer which is neither semantically organized nor domain specific, by using a typical keyword matching approach. The primary difference between this and the current approach is that Integrated Search offers seamless integration of both keyword-based search results from the Web and the semantically rich search results from the local knowledge base in the domain of interest.

C. Current Approaches to Semantic Desktop Search

A combination of fact and document retrieval [10] is proposed as a new approach in a semantic desktop search engine which answers natural language queries through the techniques of semantic teleporting and spreading activation. It supports queries such as "Abstracts and Authors of Papers about Semantic Desktop Applications" to retrieve information or facts from Personal Information Models (PIMO) [10]. The strategy used here enables retrieval of facts to answer the queries in the form of graphs. Although the approach provides a method to exploit available semantics on the desktop, resembling a question answering system, it is inadequate for a typical inquisitive search by a professional or researcher who is looking for explanations, illustrations, definitions or elaborations in specific areas of interest.

Various approaches for answering natural language queries have been tried including pre-query disambiguation in Squiggle [11] which involves user interaction for resolving structural ambiguities of query terms and personalization of the retrieval process in Haystack [12] which considers information maximization to enable users to query their own "personal digital libraries". Nepomuk [13], a social semantic desktop provides a framework to annotate pieces of information on a desktop and enable enhanced semantic search on the desktop. However, these approaches do not provide an integration of domain specific, semantics-based retrieval from a knowledge base with regular Web search.

III. Proposed Solution

A. Knowledge Base Retrieval using Semantic Roles

Search is becoming the de facto means for finding information [4]. Although the use of either syntactic, keyword-based search engines or semantics-based search engines has many drawbacks, combining the two gives added advantages to the knowledge seeker. This paper presents an approach in which the knowledge worker browses or searches in a way that effectively reduces the *informational myopia* and prevents the *lost in hyperspace* syndromes.

The key idea is based on a recent research thrust where attempts have been made to encode the conceptual knowledge of a domain with well-structured links between concept ontologies and associated texts. The entire knowledge base is linked through special links called *Superlinks* [2] which embed semantics into the link structures, thereby overcoming the inadequacies of hyperlinks for effective knowledge browsing and efficient semantics-based search.

Conceptual knowledge of a domain is encoded as a set of documents with *Superlink* structures. Knowledge workers are able to browse through the repository of documents where semantics associated with multi-directional link structures make the experience of knowledge browsing rich and effective.

Web users tend to give simple search queries which result in poor precision. A way to improve the precision of search engines is to use more than one query term. In this regard, ongoing work suggests the idea of automatically adding *semantic roles* to the search query based on the context in which search is being performed [2]. The search query is pre-processed by the engine and the user can be involved in a dialog to resolve any ambiguity in the word senses of the query term.

The above method of inserting additional description to the query, based on a dialogue with the user helps the syntactic search engine to improve the relevance of its search results, thereby bringing it closer to producing the “most expected or appropriate” results.

The query can be expanded with *semantic roles* [2] (see below for an illustration of Superlinks and semantic roles in figures 1 and 2) and forwarded to a regular keyword based search engine. The integrated search results from both the local Knowledge base and web search provide the user with two options:

- The user can efficiently traverse the Superlinks to navigate in the local knowledge base; or
- The search interface also presents the results from the keyword-based search engine

1) Inter linking Concepts in Domain Ontology and Textual Elements using Semantic Link Structures

Performing a semantic role-based search involves the construction of a knowledge base in the domain (eg. Indian medicine – Ayurveda).

A novel strategy for constructing the ontology involves linking the textual basis for the terms, concepts, relations and attributes to the concepts in the ontology [2]. Such linking is extremely important to provide legitimacy to the entries in the ontology, apart from facilitating the development of the ontology itself and enabling semantic interpretation of the texts so linked. This new way of linking the concepts and the textual elements for effective knowledge browsing and search is termed as *Superlinking* in contrast to the normal hyperlinking and the links constructed are termed as *Superlinks* [2]. This methodology leads to a clear advantage – the knowledge workers are able to see in one Superlinked structure one or more texts that define the concept or relation, describe it, give an example for it, show its use, provide translations for it in different languages, or otherwise add to the information in the ontology in encyclopedic ways.

2) Semantic Structures in the Knowledge Base: Superlink

A Superlink [2] is a link among ‘*n*’ different points, which can be termed as an n-ary link. A point which is either a source or a target of the link can be a piece of text (e.g., a word, phrase, proper name, clause, sentence, paragraph, section or an entire document) or a structured element in a document (e.g., an entry in a table or an item in a list) or a concept, relation, attribute or attribute value in an ontology. Comparing the hyperlink to a Superlink, the hyperlink is a binary link. Even though the source can be any piece of text or structured element, the target can just be an entire document or an approximate location in it. Moreover, the hyperlink does not contain any semantics attached to the link between the source and target. A Superlink also specifies semantic roles and *xlink:show* and *xlink:actuate* attributes which specify to the browser how the link should be displayed and when the link should be activated. Using this, a Superlink can point exactly to the target location. The design of Superlinks provides two alternative ways of choosing the correct behavior by the browser on click: engaging the user in an interactive dialogue using the *semantic roles* (i.e., the typology of link semantics) to determine which particular target point(s) may be relevant to the user, or simultaneously taking the user to all the targets in the Superlink by visiting each one and presenting a *mash-up* of information gathered from all the targets.

3) Superlink Semantics

The *semantic role* attached to the Superlink specifies how the source and target of the link are related. The different roles are as shown below [2]:

- *Definition:* “TDef” for the link between concept and text and “CDef” for the inverse link gives the definition(s) of a conceptual element in the linked text(s).
- *Elaboration* or description: “TEl” for the link between concept and text and “CEl” for its inverse points to elaborate description(s).
- *Illustration:* “TIll” for the link between concept and text and “CIll” for its inverse represents examples illustrating a conceptual element in the linked text.
- *Analogy:* “TAna” for the link between concept and text and “CAna” for its inverse represents an analogy, metaphor or simile.
- *Generalization*: “TGen” for the link between concept and text and “CGen” for its inverse gives a generalization of the concept.

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

- *Translation:* "TTra" for the link between concept and text and "CTra" for its inverse represents a translation of the conceptual element in text from one language to another natural language.

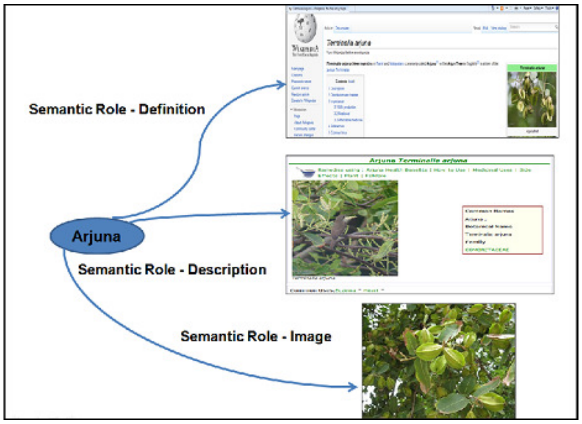

Figure 1: Pictorial Representation of a Superlink

A pictorial representation of a "Superlink" is shown in Fig.1 where "Arjuna" is a concept for which Superlinks are created. The first destination in the Superlink is a text file giving the concept's definition. The second target contains an elaboration or description of "Arjuna" and the third target is an image file showing a picture of the herb Terminalia Arjuna.

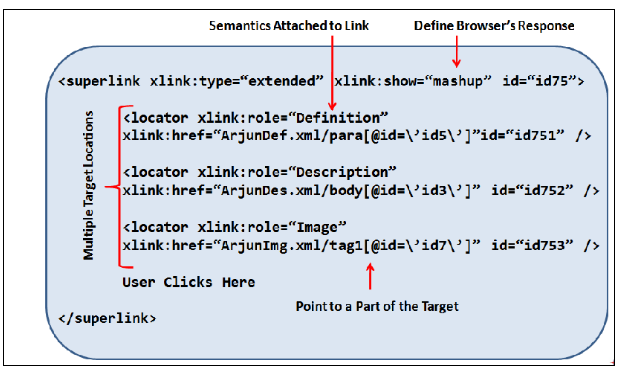

Figure 2: XML Representation of Superlink

The corresponding XML representation of the Superlink is shown in Figure 2. Each Superlink contains as many *locator* tags as the number of target destinations in the link. The Xlink attribute *xlink:role* specifies the semantic role played by the link in the context and *xlink:href* specifies the exact destination of the link in the target file which (in our ongoing work) will incorporate XPointer and XPath expressions to provide sufficient precision for specifying the targets of the link.

4) Construction of the Knowledge Base of Ayurveda

The knowledge base of the domain "Ayurveda" was constructed from available HTML Web pages after transforming them to a suitable XML format. The Ayurveda ontology was constructed simultaneously by adding concepts of Ayurveda, basically consisting of traditional medicinal plants [7, 8, and 9]. The basic hierarchy in the ontology is as follows: Ayurveda is the root concept in the domain ontology. Medicinal Plants is a subclass of Ayurveda and have Chemical Composition, Medicinal Use and Distribution as properties.

The Superlinks are constructed in two primary ways:

- Linking the concepts, relation, attribute or attribute value in the Ayurveda Ontology to XML Web documents; or
- Linking two or more textual or structural elements in a list of the XML Web documents.

The XML documents so edited contain the semantic structures namely Superlinks built on top of the existing unstructured information from the Web. Thus the knowledge base is semantically constructed with semantic Superlinks attached to the concepts and related Web documents.

5) Semantics-based Knowledge Retrieval

As semantic search lends itself well to the kind of search normally done by a researcher where she is trying to locate a set of documents which together give her the required information, the main intent is to enable the user to obtain the semantic structures which most appropriately define/describe/elaborate/explain/illustrate the concept of interest. Semantic search in the chosen domain involves searching the most appropriate Superlinks in the knowledge base which give relevant information about the search query. The Superlinks contain locators which represent the targets of the link along with the associated semantic roles describing the exact relationship between the source concept or textual element and the target textual elements of the Web documents. Each of the locators embedded within the Superlink also contains a *rank* specifying the preference of the particular locator over other competing locators under the same parent Superlink. The search results contain relevant sets of Superlinks in the order of their ranks, with the most preferred locator at the top.

B. Integrating Knowledge Base Retrieval with Web Search

Although semantic search results from the knowledge base contain relevant results for the search query from the domain of Ayurveda, they are limited to the contents of the local knowledge base. Hence there is a need to integrate the semantically obtained search results with the syntactic, keyword-based search of the Web. The Web search is performed using a Web Search API such as "Google AJAX Search API" or "Yahoo Search Engine Service (B.O.S.S.)".

1) Pre-processing of the Search Query

The search query needs to undergo transformations in several steps in order to be able to retrieve relevant results from the Web. Query processing mainly involves a kind of shallow natural language processing by utilizing typical pre-processing steps such as stop word removal, stemming, word sense disambiguation and query expansion. The search query is expanded by appending the most frequently used semantic roles - definition, description and application - to the pre-processed query in cases where the input query given by the user does not contain any of the semantic roles. This strategy in query expansion helps not only increase the precision of Web search results but also in the seamless integration of the knowledge base and web search results.

2) Seamless Integration

The search query which is pre-processed and expanded is sent to the Web search engine using the API and results are obtained and integrated with semantic search results. The knowledge base retrieval results constitute sets of Superlinks in the order of their preference rankings while the syntactic search results have their own rankings as determined by the Web search engine. The two sets of results are seamlessly presented to the user based on the semantic role played by each result which specifies the relation between the search query and the result.

The search results containing both semantic and syntactic search results are presented with their URLs (and summaries where available) along with links to the user feedback system. The feedback form contains a simple set of fields for the user to enter the ranking for the link based on user's personal preference over other results for the query. The feedback is stored in a user profile base for future expansion of the local knowledge base and for refining of future search results.

3) User Profiling

Although Superlink structures are semantically constructed and contain rich links in the local knowledge base, they may not always render the most relevant results for a search query. Hence, the additional strategy of profiling the users is employed so that there is continuous improvement in the relevance of results and personalized preferences of the user. In such a personalized search service, an approach for building ontological user profiles by assigning interest scores to existing concepts in the domain ontology of Ayurveda has been developed. This approach successfully provides the user with personalized search results by retrieving only those results which are of high interest to the user. An ontological user profile is like an instance of the domain ontology in which concepts are annotated by interest scores given by the particular user upon sign in or sign up to the system. The rankings given by the user for the concepts in the ontology can also be seen as providing semantic evidence for solving the "cold start problem" in recommendation systems [5] apart from effective word sense disambiguation in search queries.

IV. Results and Discussion

A system for performing syntactic or regular keyword-based Web search and semantic retrieval from a local knowledge base using semantic structures called Superlinks was built in a knowledge base in the domain of Indian medicine - Ayurveda. The prototype system consists of several hundred entries for concepts in the domain ontology along with documents containing a few hundred superlinks interlinking the textual elements or concepts within the knowledge base and each Superlink containing on an average three to four target destinations. The search results from both the search processes (running as parallel threads) are integrated based on their semantic roles. As this is a proof of concept, the construction of a very large knowledge base requires the services of an Ayurvedic domain expert to build the domain ontology. Such a knowledge base with richly interconnected Superlinks and concepts of the domain can enable a quantitative evaluation of search quality.

A. Prototype System Capabilities

1) Word Sense Disambiguation based on Ontology

The Word Sense Disambiguation module takes the stemmed query term as input. For each stemmed query term, the word senses available are checked for ambiguity. The ambiguity can be resolved by the "user's ontological preferences" or the user's choice by involving the user in a dialog with the system.

2) Integrated Search Results

In Fig.3, integrated search results are shown with both Superlinks and hyperlinks from the knowledge base and Web. Each result URL contains feedback links for the user. The target file containing Superlink structures in the local knowledge base gets opened on navigating Superlinks.

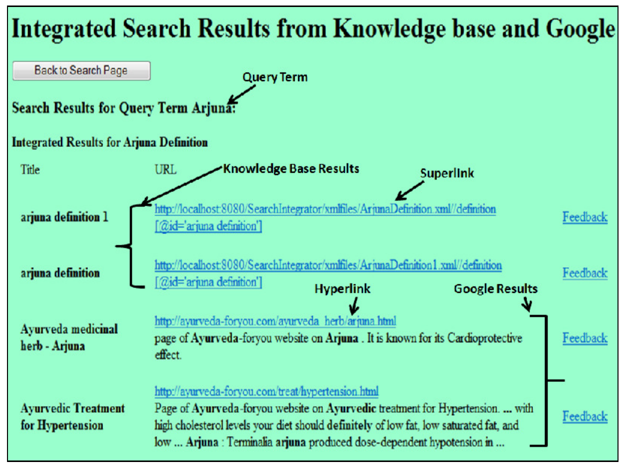

Figure 3: Integrated Search Results from Knowledge Base and Google

V. Conclusion and Future Work

Although there are several approaches to search including federated search, desktop search and the most popular Web search, all of them perform basically a keyword-based search providing a multitude of results that are in essence unidirectional hyperlinks devoid of any semantics or author intention. Results may or may not be relevant to a serious user with a genuine need for domain-specific knowledge with explanations, illustrations, definitions and elaborations on specific areas of interest. Moreover, although there exist tools for semantic annotation of texts and development of ontologies from text documents, they do not support concurrent development of hyperlinked texts and ontology concepts with interlinking mechanisms rich enough to capture the semantics of the domain. This paper demonstrated the idea of designing a vertical semantic service with the existing unstructured information of the World Wide Web and overcoming the inadequacies of hyperlinks by constructing new linking elements called Superlinks for linking both textual elements and concepts of ontology in the domain of Ayurveda. In order to resolve the Web search issue, a solution of integrating semantic search of the local knowledge base and syntactic search of the Web

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

was adopted to present the search results in a seamless fashion. The relevance of the results from the Web is improved by expanding the user queries with semantic roles and sending multiple, separate queries to the Web search API. The idea of getting user's feedback on the search results for future expansion of the knowledge base by the addition of Superlink structures and user profile refinement for obtaining more relevant results in the future were also presented.

Recent developments in search engine capabilities such as auto-completion of user queries and instant result pages can be integrated with the present system to further enhance the usability of the system and thereby the productivity of the knowledge worker. Search results may also list or highlight links which contain the same source obtained from both the local and global repositories. Capturing user's intentions during the search and navigation process to provide even higher relevance for future queries can also be an added benefit to the users. Supporting search queries in different languages can enhance the coverage of the system in terms of users from different cultural and linguistic backgrounds.

Some ontology and knowledge-base related extensions can be developed including the development of a sophisticated IDE for domain ontology development or by way of integration with existing ontology editors. Using the feedback from the user to enrich the knowledge base by the addition of new Superlink structures for the knowledge domain can help in expanding the local knowledge base. Using feedback from the user to periodically refine user preferences can help increase the relevance of search results. Extensions to browsers for supporting Superlinks in the form of plugins with new features like mash-ups or dialog-based retrieval can enable the user to seamlessly navigate the local knowledge base and the Web.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contribution of students of PESIT Nitish Shangari, Harshita Agarwal and Saurabh Shekhar towards the illustration of Superlinks in XML files.

REFERENCES

[1] Mark Levene, An Introduction to Search Engines and Web, 2nd Edition, Wiley Publishers, 2010
[2] Kavi Mahesh and Pallavi Karanth, Knowledge Browsing with Superlinks, In Proceedings of International Conference on Digital Libraries and Knowledge Organization, ICDK-2011, February 2011, Gurgaon, India
[3] Steve Lawrence and C. Lee Giles, Searching the Web: General and Scientific Information Access, Journal of the Institute of Image Electronics Engineers of Japan, ISSN - 0285-9831, 432-441(1999)
[4] B. Olstad and S. Seres, “A Roadmap to Search as a Strategic Enabler”, KM World, KMWorld Specialty Publisher Group, Camden, ME, USA, pp. 4-5, 2005
[5] S. Middleton, N. Shadbolt, and D. D. Roure, Capturing interest through inference and visualization: Ontological user profiling in recommender systems, Proceedings of the International Conference on Knowledge Capture, K-CAP 2003, pages 62–69, Sanibel Island, Florida, October 2003
[6] Federated Search Engines, Wikipedia, http://en.wikipedia.org/wiki/Federated_search
[7] Iswara Prasad A, Scientific Significance of Sacred Plants, National Institute of Vedic Sciences, first edition, January 2011
[8] M S Valiathan, The Legacy of Susruta, ISBN 978-81-7371-666-9 Universities Press, 2009
[9] Ravindra Sharma, Dictionary of Ayurveda, ISBN13 978-81-7035-315-7, Daya Publishing House, 2009
[10] Kinga Schumacher, Michael Sintek, Leo Sauermann, Combining Fact and Document Retrieval with Spreading Activation for Semantic Desktop Search, Proceedings of the 5th European Semantic Web Conference, LNCS 5021, pp. 569–583, 2008
[11] Irene Celino et al, Squiggle: a semantic search engine for indexing and retrieval of multimedia content, Proceedings of the 1st International Workshop on Semantic-enhanced Multimedia Presentation Systems, 2006
[12] Eytan Adar et al, Haystack - Per user Information Environments, Proceedings of the eighth international conference on Information and knowledge management, p.413-422, Kansas City, Missouri, United States, 1999
[13] Nepomuk - The Social Semantic Desktop, http://nepomuk.semanticdesktop.org/
[14] Eric van der Vlist, One Href is not Enough, We need n hrefs! Proceedings of Balisage - The Markup Conference , Montreal, Canada, 2011
[15] Interlingua, Wikipedia, http://en.wikipedia.org/wiki/Interlingua
[16] Larry Page et al, "PageRank: Bringing Order to the Web", Stanford Digital Library Technologies Project, August, 1997
[17] Peter Morville, Ambient Findability, ISBN 0596007655, O'reilly Publishers, 2005