DESIDOC Journal of Library & Information Technology, Vol. 31, No. 4, July 2011, pp. 217-225
© 2011, DESIDOC

Semantic Web Applications

Misha Mehra and Nishant Kumar

Defence Scientific Information & Documentation Centre, Delhi-110 054
E-mail: misha@desidoc.drdo.in; nishant@desidoc.drdo.in

ABSTRACT

With changing technology, the Internet has taken a pivotal role in all kinds of applications in our daily lives. To handle flood of information on the Internet, smarter Web technology is also required. This requirement has led to the advent of newer, smarter and better Web technology called 'Semantic Web'. Semantic Web is the next step in Web evolution. High usability of Semantic Web has found significant applications in the field of life sciences, crime investigation, scientific research, literary analysis, social networking, electronic commerce, knowledge management, digital libraries, defence, e-government, energy sector, financial services, healthcare, oil and gas industry, publishing, website back-ends, multimedia, etc. This paper discusses the most prominent areas for application of Semantic Web technology.

Keywords: World Wide Web, Semantic Web, Internet, knowledge management, social networking, electronic commerce, digital libraries

1. INTRODUCTION

Semantic Web is the next step in Web evolution. Semantic Web was conceived by Tim Berners Lee, the inventor of World Wide Web (WWW). He defined Semantic Web as Web of data that can be processed directly or indirectly by machines. Semantic Web is about reading, processing, transforming, and assembling the data repositories and documents available on the Internet into useful information. It relies on structured sets of information and inference rules that allow it to understand the relationship amongdifferent data resources[1]. It accesses and connects the information available on Websites, XML documents, and databases, and provides useful amount of data needed.

The data or information provided from semantics will take the future of Web to a new level, where there will be knowledge-based repositories having better understanding of Web content. This will facilitate accurate searching, accurate filtering, and categorisation of information and intelligent retrieval of data. The WWW contains boundless information in the form of documents. The websites represent text in natural languages and use graphics, images and videos. The information available in the Web documents is easy for people to read. However, this information cannot be perceived by machines the way it is perceived by humans. With the Semantic Web, this information can be made machine-readable. For example, if one wants to get some information from 10 different sites, one will have to go to all ten different sites and then copy the information one requires. Semantic Web will perform the same work intelligently, efficiently, and effectively. It will access the required information by connecting to all the websites in a network and will provide the consolidated data at the user's desktop. Due to these specialties of Semantic Web, major companies like Adobe, Oracle, IBM, HP, and Microsoft are developing and offering semantic tools[2]. Intensive research work is being carried out in this field.

2. APPLICATIONS OF SEMANTIC WEB: E-BUSINESS

According to Wikipedia, e-business may be defined as any business process that relies on an automated information system. E-business software solution allows the integration of intra- and inter-firm business processes and can be conducted using the Web, the Internet, extranet, etc. The application of Semantic Web in the field of e-business is wide and significant in terms of exchanging information between different business groups for mutual or collective purposes. It is believed to provide good semantic solutions when information

exchange is concerned. Semantic Web has found a prominent place in e-business in terms of searching of relevant data, exchange of information between different agents, filtering of relevant information useful for finding good business sites or analyzing new market trends, online advertisements, composition and integration of complex systems, multimedia collection, exchange of machine dialogue across the domains, virtual community and vocabulary flexibility, and standardisation[3].

One of the major limitations or drawbacks in e-business is the problem of interoperability between systems of two or more business partners (business-to-business). Interoperability is to be ensured for the business exchange to be effective and efficient between business companies. Extensible Markup Language (XML) has been used to provide the Web interoperability for the past few years. XML is capable of providing only syntactic interpretation, and not semantic interpretation. It does not understand the content and meaning of the messages being exchanged among different systems. Semantic Web solves the problem of interoperability using Web Ontology Language (OWL). OWL is a popular language used for representing ontologies on the Web. It is a World Wide Consortium (W3C) Standard and provides a strong ontology that is used as a standard. All the services on the Internet share the same standard for the interpretation of terms being exchanged between them. OWL facilitates greater interoperability of Web content. Moreover, the software used earlier for business purposes were hard-cored and always required direct human intervention by changing the code and then running the script. But with the use of Semantic Web tools and software, the pages can be dynamically changed and highly tailored and cost-effective results can be produced.

2.1 Research in E-business using Ontologies Provided by Semantic Web

Information exchange among business processes takes place through Web services. Web service is a self-describing, self-contained software system, which exposes its functions to consumers to support interoperable machine-to-machine interaction[4,5]. Web services for an organisation are registered in public registry like a catalogue of books. The three fundamental methods (Fig. 1) in a Web service system are: register, discover, and bind[6].

Figure 1 explains the working of Web services. During registry, Web services advertise their services in the public registry. Registry provides a database for the Web services. Once registry is done, service consumer or a requestor will send a request corresponding to a particular service. The registry will check in its database and will provide the relevant information of the most appropriate service. This procedure is called discover. At binding, the service consumer, upon receiving the relevant information, will directly bind with the Web service and data exchange will start taking place. In this way, exchange of data takes place among heterogeneous systems in a distributed network. In order to retrieve the most optimal Web services, registry function relies on standard ontology language like Simple Object Access Protocol (SOAP), Web Services Definition Language (WSDL) and Universal Description Discovery and Integration (UDDI), Web Ontology Language (OWL) and DARPA Agent Markup Language + Ontology Interchange Language (DAML+OIL). These ontologies are very popular as these are the fundamental building blocks in the development of e-business applications.

DAML+OIL is a semantic markup language for Web resources. It builds on earlier W3C standards such as Resource Description Framework (RDF) and RDF Schema, and extends these languages with richer modelling primitives[7]. The language has a clean and well-defined semantics. DALM-S uses DAML+OIL ontology language for the information exchange between different Web services across the Web. UDDI is a specification designed to allow businesses of all sizes to benefit in the new digital economy[8]. There is a UDDI registry, which is open to everybody. It is an Internet-wide registry of Web services. It allows companies to register their information like company name, specific service or type of service. This allows companies providing or needing Web services to discover each other and share the information over the Internet in a truly global manner. OWL is a semantic markup language for publishing and sharing ontologies on the WWW. OWL is developed as a vocabulary extension of the RDF. Researchers have used OWL-S for developing Web service models capable of exchanging information on the Web. OWL is very popular among the researchers due to its strong semantic interoperation.

DAML+OIL and UDDI have been used as ontology languages to provide architecture for the information exchange[5]. DAML-S defines DAML+OIL ontology for the description of Web services. The concept of process model allows interaction of both service consumer and Web service provider by specifying an interaction protocol where service consumer will know of what information to send and what information to receive. On the other hand, Web service consumer will also know about the status of the information received by Web service provider. The information being exchanged between both the parties is mapped into messages to be exchanged on the Internet by DAML grounding. Web architecture, using the process model and DALM-S grounding, consists of DAML parser, for transforming specification for process model and grounding into DALM-S descriptions; DAML processor for extracting

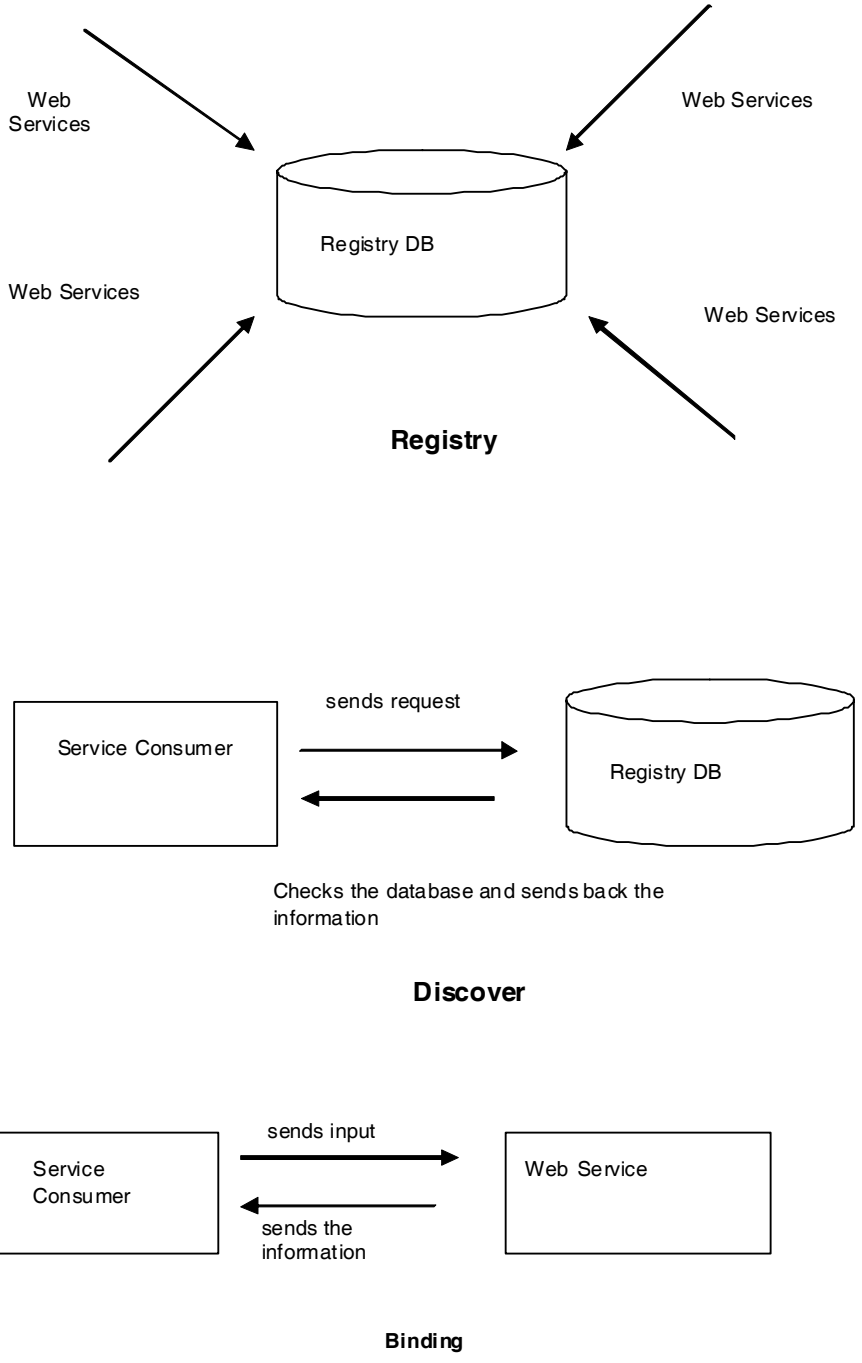

Figure 1. Web service system.

information protocol and composing message to be exchanged with the Web service provider; and Web service invocate for contacting other Web services.

One of the matching criteria in OWL-based Web services, where input and output of Web service is subsumed by input and output of service consumer[9]. This matching criterion has been further enhanced by putting constraints of subsuming the initial state of business process with the pre-conditions of potential Web service and also the after-effects of service execution with the final state of request[4].

The concept of Web service community was also used to provide a hierarchical classification of Web services sharing the common domain interests as subject clubs. Now the search engine matches request from the service consumers based on the representations of Web services in the subject clubs as well as the pre-conditions set in the matching algorithm.

3. APPLICATIONS OF SEMANTIC WEB: SOCIAL NETWORKING

Social networking has become an important part of the modern society and puts a strong impact on social, political, educational, professional, personal, and business life. It connects people across the world through social networking sites like Facebook, Orkut, MySpace, LinkedIn, etc. It allows information sharing on twitter, messaging through Yahoo Messenger, Google talk, content and ideas sharing through blogs, discussion forums, uploading and downloading of media, tagging, and folksonomies through wikis and podcasts, etc. The social networking have attracted millions of users across the globe and has become the most popular, convenient and cheaper mode of communication. A lot of social networking sites are coming into the business because of wealth being generated by these across the world.

Social networking sites or services offer the basic features, but there are certain limitations in terms of connecting people and content in a meaningful way[10]. One of the limitations is the lack of interoperability among different social networking sites. Suppose a person having a profile account in one of the site wants to re-open the new profile account in some other site, and wants the same information in his previous account to be migrated into his new account, will not be able to do so because the provision is not available in social networking services. In such a scenario, the user has to re-enter the entire information, and update its information at two different sites, making the process more cumbersome. Second limitation is the lack of privacy. In case of certain centralised sites, the user does not have complete control over the information they provide[11].Thirdly, social networking sites do not provide relationships or connections between common profiles of people sharing common features liking visiting the same sites, or same likes, and hobbies. Semantic Web solves the problem of interoperability by providing globally accepted semantics to share information about people, their profiles, contents and connections through which they are interconnected. Security and privacy is being preserved as data is machine-readable.

Many projects have been initiated and various new applications and tools have been developed in the field of social networking. Search engines designed for public use are gradually assimilating the technologies of Semantic Web. Tumbup is a new search engine connected to Facebook, which analyses the activities of members of social network, like recommendations about places, products, etc. to produce more meaningful results[12]. Other search engines like Wolfram Alpha, True Knowledge, and Zoom are also integrated to the Intranet sites and corporate blogs. Facebook has introduced the Open Graph Protocol, a technology based on the Semantic Web, which allows third-party sites to interact with the social networking sites by sending and receiving information to Facebook. This new protocol integrates webpages into the social graph and sustains the social interactions between the visited Websites and the Facebook profile of a user[12]. Concept of Semantic Web has been used by many researchers and many projects have gained momentum in recent years. Some of prominent ones include Friend of a Friend (FOAF) and Semantically Interlinked Online Communities (SIOC).

3.1 Friend of a Friend

Friend of a friend (FOAF) consists of machine-readable pages used to describe people, their activities, and their relationships with the other people and objects. It connects social networking sites and the people it describes using a decentralised database. FOAF has got it applications in a number of services running on the Web. Some of the services are[13]:

- Google's Social Graph API indexes all the public FOAF data in the Web.
- Yahoo's Search Monkey also has support for FOAF.
- Garlik's QDos system includes a FOAF search facility, a FOAF builder tool, and various other services.
- The FOAF-search.net site offers search/browse for several million crawled FOAF file
- The Yandex search engine (wikipedia) has some support for FOAF.
- Export Address Book to FOAF is an AppleScript for Mac OS X that allows users to manage their FOAF profile from within the Address Book including uploading of the FOAF file to the Website using FTP.
- Morla is a RDF editor. There are templates for FOAF Person, FOAF Personal Profile Document, FOAF Image, etc.
- FOAFDrive is another FOAF XML file generator. It is also used for FOAF file import, APML, social networks connections etc.
- Elmo contains a set of tools for crawling, smushing and validating FOAF profiles
- MyFOAFexplorer Java 1.5 applet used to browse a scientific network based on FOAF
- FOAFMap display FOAF profiles on a Google Map
- FOAF-inator is a FOAF navigation tool developed in PHP

3.2 Semantically Interlinked Online Communities

Semantically Interlinked Online Communities (SIOC) allow the integration of online community information. It provides Semantic Web ontology for representing information in social Web like blogs, discussion forums, wikis, photo galleries, shared bookmarks, aggregated feeds, etc. SIOC has found significant place in the variety of commercial and open source applications and is used in combination with FOAF ontology for defining user profile and social networking information. Some of major application areas are[14]:

- WordPress SIOC exporter exports information about WordPress based blogs, blog users, posting and comments.
- OpenLink Data Spaces modules export SIOC meta data for various applications like ODS based blogs, Wiki, Bookmarks, AddressBook, Calendar, polls, photo Gallery, ODS-Feeds and Discussions.
- Talk Digger is a Web service for locating people and uses SIOC to export all its data in RDF/XML.
- SWAML is an exporter used for publishing mailing list archives into RDF format. It uses SIOC as the main ontology to represent a mailing list in RDF.
- Twitter2RDF is an RDF exporter for Twitter microblogs that uses SIOC (for the microblog entries) and FOAF (for describing the people).
- Gnizr, an open source social bookmarking and mashup application from ImageMatters exports saved bookmarks using SIOC.
- IkeWiki is a semantic wiki for knowledge engineering. IkeWiki attach discussions to wiki pages using the SIOC ontology.
- Engage, the community information application from Talis, is using SIOC, SKOS and FOAF.
- RDFa on Rails is a library for Ruby on Rails. It helps in publishing of RDFa data by providing set of helper methods to the rails application. SIOC terms are used to describe blog posts in this library.
- OpenQabal, an open source social networking and collaboration platform, also has SIOC support.
- The Seesmic micrologging service has decided to adopt the SIOC ontology as one of their open platform formats
- An RDF converter for IRC has been created that exports metadata in Turtle format, and SIOC is being used as one of the main representation formats.
- Faculty Academy's Fishtank project uses the RDF structure and searching power to fully utilise tags and feeds on blogs.
- BAETLE (Bug And Enhancement Tracking Language) uses SIOC for creating a software bug ontology that can be used by people to query for bugs.

4. APPLICATION OF SEMANTIC WEB: KNOWLEDGE MANAGEMENT

Knowledge management system (KMS) as a whole describes the creation of knowledge repositories, a method for knowledge access and sharing as well as communication through collaboration and also enhancing the knowledge environment and managing knowledge as an asset for an organisation[15]. Development of a good KMS is basically a collaborative effort. Figure 2 illustrates a traditional collaborative KMS.The traditional collaboration of KMS has undergone

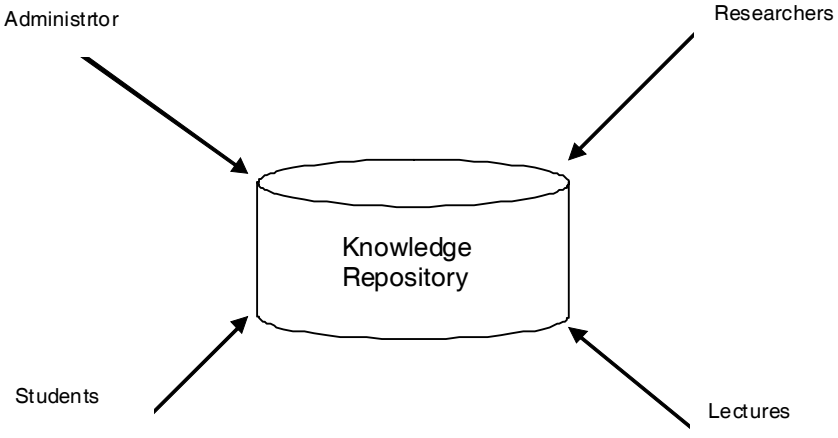

Figure 2. Knowledge management system.

a revolution with the advent of WWW. The WWW or Internet provides a knowledge repository with variety of information from various sources and also from geographically distant corners of the world. This makes the repository information rich like never before but this also brings the real challenges. These challenges, like overload of information, keyword searching being not appropriate, integration of information from heterogeneous sources and geographically-distributed Intranet problems, have been triggered by the Web. In fact, this enormous amount of data has made it increasingly difficult to search, access, present, and maintain the information required by a wide variety of users. This is because information content is mainly presented in a natural language. Thus, a wide gap has accured between the information available for tools aimed for knowledge extraction and the information maintained in human-readable form.

The most critical issue in intelligent knowledge management is how to represent and extract the semantic meaning from information contents. Researchers have tried to address this issue through various research areas including artificial intelligence, information retrieval, natural language processing, multimedia, knowledge management, etc. All these methods ask for a smarter Web to assist in knowledge acquisition, knowledge representation, knowledge sharing and distribution of human knowledge through the Web. This requirements have been answered by Semantic Web.

Tim Berners Lee has referred to Semantic Web as an extended Web of machine-readable information and automated services that extends far beyond current capabilities. Semantic Web is useful for KMS because this technology helps people to find and correlate the information they require irrespective of their sources and types. Semantic Web is relevant to knowledge management because it has the capacity to enhance the speed with which information can be synthesised manifolds. This is achieved by automating its aggregation and analysis. Most of the time information on the Web is presented in HTML format but the problem with it is that this format does not provide structure or metadata useful for effective management. Without structure, elements of content cannot be related to each other, and without metadata, the nature of the elements themselves cannot be known. Semantic Web is designed to provide these missing components. It can provide structure, (through the use of XML tags); metadata descriptors, (through RDF) and relationships (through Web Ontology Language). 'Ontology' is the key enabling power in realising the full potential of Semantic Web technology.

Ontology is not knowledge or information. It is in fact, meta-information which is information about information. In the context of the Semantic Web, using a special ontology language, the relationships between the various terms within the information can be encoded. Ontologies provide background information, which strengthens the description of the data and which helps in making the context of the information more explicit. Since ontologies are shared specifications, the same ontologies can be used for the annotation of multiple data sources, which include webpages, collections of XML documents, relational databases, etc.

The use of such shared terminologies enables interoperability between these data sources up to a certain extent. But, this does not solve the integration problem completely, because it is not possible for all individuals and organisations on the Semantic Web to use one common terminology or ontology. It is very much possible that different ontologies will appear and, to enable interoperation, mediation is required between these ontologies. Ontology mediation is necessary in semantic knowledge management for enabling sharing of data between heterogeneous knowledge bases and also to allow applications to reuse data from different knowledge bases. Another utility of ontology mediation appears in Semantic Web Services. In general, it is not necessary that the requester and the provider of a service use the same terminology in their communication, and thus, mediation is required to facilitate communication between the knowledge seekers and the knowledge providers.

4.1 Research in Semantic Web-based Knowledge Management System

4.1.1 Ontology-based Framework for Semantic Web Services for Knowledge Management Systems

Researchers have proposed a conceptual model of ontology-based framework for Semantic Web services for KMS. This framework has a layered structure. These layers are user interaction, interface, mediator, and ontology. All these four layers will communicate with other layers for specific roles and responsibilities of each layer; however each layer is designed to behave independently. This approach is based on 'Web service Modelling ontology' conceptual model which identifies four top level elements as the key aspects to define Semantic Web services: Ontologies, Web services, goals and mediators,[16] as shown in Fig. 3.

4.1.2 Components of Ontology-based Semantic Web Services Framework for KMS

Distributed KMS will enable different knowledge providers (individuals, groups, and organisations) to share

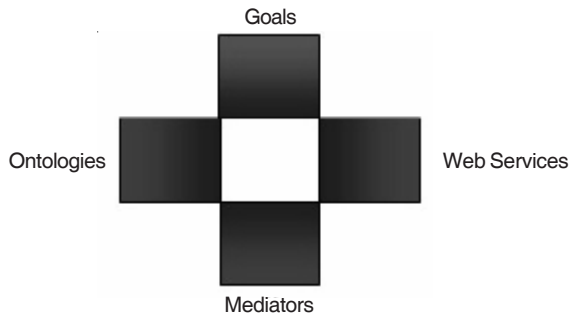

Figure 3. Ontology-based services framework for KMS.

the knowledge stored in various forms in distributed repositories within their organisations. The first layer, which is the user interaction layer of this framework, employs a Web services infrastructure and then enhances it ontologically so that lifecycle of a knowledge transaction is supported. Another layer, that is the interface layer specifies the functionality of the Web service through which it can communicate and cooperate with other Web services from different service providers.

Third layer, which decides the mediation, is concerned with handling heterogeneity, which is concerned with resolving the possible mismatches between resources that may be interoperable. The ontology layer is the important layer which provides support required by all the three components in mediator layer. It presents the semantic descriptions in three ontologies; user, domain and Web services ontologies.

4.2 On-to-Knowledge Project

On-to-Knowledge (OTK) project supports efficient and effective knowledge management by providing a tool environment powered by Semantic Web technology[17]. It focuses on acquiring, maintaining and accessing weakly structured information sources. The layered tool environment of OTK[17] is shown in Fig. 4.

Acquiring means extracting semantic information from textual information, is done through text mining and extraction techniques. OntoExtract and OntoWrapper are the two tool supports for acquiring ontology extraction from text.

To explain the syntax and semantics of semi-structured information sources, RDF, XML, and OIL are used. OntoEdit is the tool support as ontology editor, and ontology storage and retrieval (Sesame), so as to enable automatic maintenance and view definitions of knowledge.

Accessing refers to accessing of information. It is supported by push-services and agent technology. RDFFerret is the tool support for ontology-based information navigation and querying, and Spectacle is the tool support for ontology-based visualisation of information. vAs a whole, the complete layered tool environment of Ontoknowledge operates like this: The unstructured and structured textual information sources from specified domains on the Web are extracted by OntoExtract and OntoWrapper. This extracted

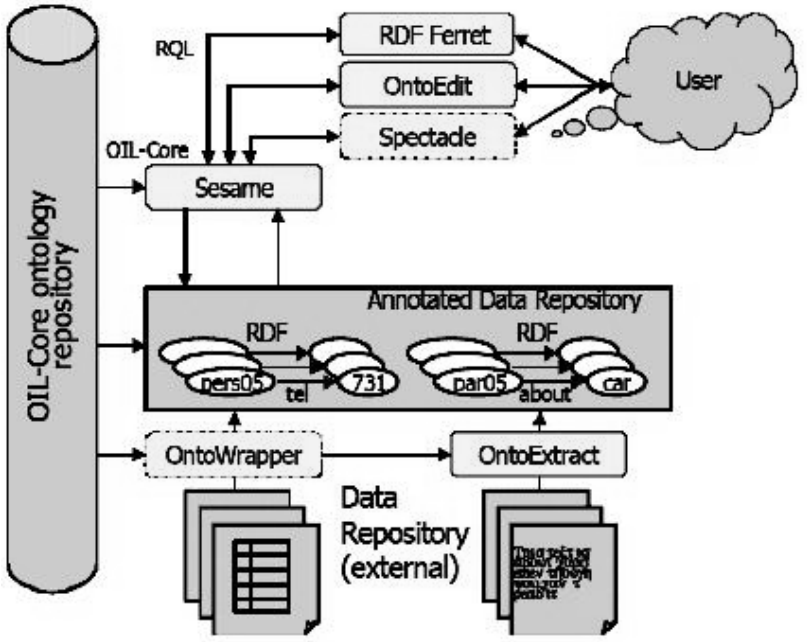

Figure 4. The layered tool environment of the on-to-knowledge.

information, now, is entered into the RDF database, where it can be edited with the OntoEdit tool. Finally, the RDF querying language reasoning engine can be used to query this database and delivers results to a user through RDF, which may be visualised by spectacle.

4.3 SWAP

To demonstrate the power of peer-to-peer (P2P) computing, and how the Semantic Web can actually be combined to support decentralised environments, where participants can maintain individual ideas of the world, by sharing knowledge in such a way that administration efforts are less but knowledge sharing and finding is easy, another IST project SWAP is undergoing. Key to the success of combining P2P solutions with Semantic Web technologies is the use of emergent semantics. Emergent semantics builds on lightweight and/or heavyweight ontologies that different individuals, departments or organisations have created. It basically works on the principle of ontology mediation as discussed above. There are certain challenges, which pop-up because of the major differences between ontologies in a P2P environment or because of differences in ontologies in an environment working on the client-server structure. The tasks of SWAP will mainly counter such challenges. Four of these can be categorised as (i) peer selection service, (ii) variation of ontologies, (iii) lack of ontological precision, and (iv) ontological drift.

5. CONCLUSION

Due to the huge collection of data, it has become difficult to manage and retrieve meaningful information from the Web. There is a need of Semantic Web when useful and meaningful data retrieval is required. Semantic Web has overcome the problem of interoperability by providing a standard format for data exchange using ontologies. It provides access to the information in heterogeneous and distributed environment using machine-readable semantics. This helps in development of KMS where input comes from distributed sources.

Popular ontologies like OWL, DAML+OIL, FOAF, SIOC form the basic building blocks for the development of Semantic Web applications. Researchers have developed models using these ontologies. New ontologies have been developed by them to provide more flexibility and add more functionally to the tools and applications of Semantic Web. However the research is still in its preliminary stage and much more has to be achieved through continuous efforts.

REFERENCES

1. Semantic Web. http://www.altova.com/semantic_Web.html (accessed on 26 April 2011).
2. Herman, Ivan. What is being done today? http://www.w3.org/2009/Talks/1214-Darmstadt-IH/Applications.pdf (accessed on 26 April 2011).
3. Yuxiao, Zhao & Kristian, Sandahl. Potential advantages of Semantic Web for Internet commerce. *In* Proceedings of the International Conference on Enterprise Information Systems (ICEIS), Angers, France, 23-26 April 2003.
4. Huang, Xuejuan; Chen, Xinmeng; Xu, Jiazhen; Li, Hua & Liu, Jinshuo. Approach toward flexible semantic e-business in web service community. *In* Proceedings of the 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, SNPD, IEEE Computer, 2007. pp. 766-71.
5. Paolucci, Massimo; Sycara, Katia; Nishimura, Takuya & Srinivasan, Naveen. Toward a Semantic Web e-commerce. *In* Proceedings of 6th International Conference on Business Information Systems (BIS 2003), Colorado Springs, USA, 2003.
6. Bose, Ranjit. Applying Semantic Web services to web-based IS and applications. *In* Proceedings of the 5th International Conference on Information Technology, New Generations, IEEE Computer, 2008. pp. 1150-55.
7. "DAML+OIL Reference description. http://www.w3.org/TR/daml+oil-reference (accessed on 26 April 2011).
8. SOAP, UDDI and WSDL. http://www.enterprise-component.com/docs/ cxsLesson5.pdf (accessed on 26 April 2011).
9. Paolucci, M.; Awamura, T.; Payne, T.R. & Sycara, K. Semantic matching of web service capabilities. *In* Proceedings of 6th IEEE International Symposium on Wearable Computers (ISWC), Seathle, Washington, USA, 2002.
10. Breslin, John & Decker, Stefan. The future of social networks on the Internet: The need for semantics. *IEEE Internet Comput*., 2007, 86-90.
11. Mika, Peter. Social networks and Semantic Web. Springer, 2007. ISBN: 978-0-387-71000-6.
12. Semantic Web: The actual fields of applications. http://en.kioskea.net/faq/9880-semantic-Web-the-actual-fields-of-applications (accessed on 26 April 2011).
13. Tools. http://wiki.foaf-project.org/w/Tools (accessed on 26 April 2011).

14. SIOC applications. http://sioc-project.org/applications (accessed on 26 April 2011).

15. Abdullah, Rusli; Selamat, Mohd Hasan; Sahibudin, Shamsul & Alias, Rose Alinda. A framework for knowledge management system implementation in collaborative environment for higher learning institution. *J. Know. Manage*. *Practice*, 2005, **6**.

16. Che Cob, Zaihisma & Abdullah, Rusli. Ontology-based Semantic Web services: Framework for knowledge management system. *In I*nternational Symposium on Information Technology. ITSim 2008, Malaysia, 2008. pp. 1-8.

17. Fensel, Dieter; Harmelen, Frank Van; Klein, Michel & Akkermans, Hans. On-to-knowledge: Ontology-based tools for knowledge management. *In* Proceedings of the e-Business and eWork (EMMSEC) Conference, Madrid, Spain, 2000.

About the Authors

Ms Misha Mehra is a graduate in computer engineering from Jamia Millia Islamia. She is presently working as a Scientist at DESIDOC and handling software and Intranet Web services. She has also written technical papers on use of open source technology in library automation.

Shri Nishant Kumar is presently working as a Scientist at DESIDOC. He is an engineering graduate and has seven years of experience in the field of Internet and Intranet Web services. He has also worked in the field of data mining and machine learning. He has presented five research/technical papers in various conferences including two National conferences.