Applications of Semantic Web techniques in Web 3.0

[1]Sushma Ravichandran

[1] Pursuing Bachelor of Engineering,
[1]Computer Science and Engineering,
[1]CMR Institute of Technology, Bangalore, India

***Abstract*— A computer can only understand data digital data in the form of 1s and 0s while the human mind can comprehend only in natural language. To implement a universal, common dialect for both machine and man, there are only two solutions: One, teach men to process in binary language or two, program a computer to grasp the natural language. Out of these two options, the former is infeasible because the average human brain is not so well developed as to perform such a task. Therefore, a rush to program a computer to understand natural language is under rigorous research. Because of the sheer availability of massive amount of data, this theory is first being implemented in large scale on The Internet. This requires the transformation of the currently used Web 2.0 to the new age Internet, Web 3.0 or Semantic Web as it is most commonly called. Although its development is in its primitive stages, the possible applications of Semantic Web are a popular area of research. This paper discusses ideas to incorporate the advantages of Semantic Web into the already existing system of different domains over The Internet.**

***IndexTerms*— Semantic Web, Web 3.0, Medical care, E Learning, Search Engine.**

I. INTRODUCTION

Semantic Web according to the World Wide Consortium is defined as a tool that "provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries". With the recent enthusiasm towards Artificial Intelligence, Semantic Web is the next step into machine learning. The goal of Semantic Web is to create a web of data inter-connected in a way whereby the computer can skillfully extract the necessary information and process it as required.

The Internet today is tailored for human use and not computer. This means that although a computer may watch a video, listen to music and read a web page, only a human will be able to comprehend the contents of the same. Similarly, a human user has to go through a number of old e-mails to answer questions like 'Will I be available on the coming Thursday between 4 and 6 for my dentist's appointment?' Semantic Web promises to not only answer this question but also arrange an appointment with the dentist after connecting with the doctor's availability calendar.

The Semantic Web would convert the present day document centric internet to a data centric Web of Data type of internet. Since research is in its primitive form, and as discussed in Paper 11 there are a number of problems while implementing the Semantic Web. Along with that, there are an equal number of applications of the same. Hence, we need to bring down the drawbacks until the variety of applications over weigh the shortcomings of Semantic Web.

The idea is to unite all of this to one common framework where the users are provided with not only correct but also only the necessary information.

Semantic Web uses a number of technologies such as Resource Description Framework (RDF), SPARQL Protocol and RDF Query Language (SPARQL), Web Ontology Language (OWL), and Simple Knowledge Organization System (SKOS). Each of these has been used separately with special emphasis on how it will be used for the advancement of services on Semantic Web.

II. MEDICAL SCIENCES

The most important application of Semantic Web will be its use in Medical Science. The biggest challenge when dealing with biomedical data is that the data needs to be integrated from various disparate sources.

For example when discoveries need to be made in the medical field, it would be helpful if all information about the interconnections between genes, proteins, vitamins, tissues, functions of organs, cells and chemical processes are all available and can be integrated [1]. Semantic web can make this possible. Another source of useful insight is provided by [2] which introduce the Dartgrid which is an application development framework together with a set of semantic tools to facilitate the integration of heterogeneous relational databases using semantic web technologies using Dynamic Semantic Query interface.

[3] shows how a neurology researcher, Ann, will be able to manage various tasks that deal with her research- all at once. When writing a research paper, she will have to explore a vast amount of data, summarize them manually, collect all the valid information for her research, re-enter data elsewhere and she might even have to correlate subjects that do not seem cogent at all. The paper then talks about a Semantically designed 'Paper Writing Workspace' where she is able to deal with this huge volume of data as well as manage certain other personal work and email notifications all at once. The papr also shows the various transitions made from the initial workspace to the final one where the final workspace has a number of layout specifications, menus, options and other buttons/keys to customize the window to her specific needs. This is the beauty of Semantic Web.

A patient's records can all be stored in a structured database. This would help detect any illness based on patient history. Medication can also be administered to the patients based on past records. Another important phase that Semantic web can be used in is Diagnosis. A user could report his/her ailments while semantic web could help determine the cause for these ailments based on patient history, history of other patients with similar ailments and a redefined database with symptoms and diseases. Clinithink is one example of an intelligent system that uses a similar kind of mechanism.
There are websites available[4] where different medical data are mapped together under a common interface. RDF, OWL and each of the sub parts of the Semantic Web will help achieve this goal of mapping data from different sources together. Combining metadata with Ontology, it creates a vocabulary to link different Medical data and collaborated them on one single website.
Imaging poses a problem of being translated and applied to non-imaging biomedical data. A medical image annotation technique as proposed by [5] where the different attributes of an image are observed (like size, density, length, breadth, etc.). This is then compared with the database of the imaging system itself. For example, different tissues of our body have different densities. So will the image of these tissues. This comparison helps us to semantically use images and use them effectively during medical research.
The recent surge in knowledge especially in the field of biomedicine is rapidly accelerating on the Web. This has developed the idea of e-Science which has lead to reconstructing the Internet Semantically in order to match the growing needs of bio medical researchers.

III. E LEARNING

E-Learning is one domain that is seeing rapid developments every single day. One of the first requirements of an E-Learning system to flourish, an open intelligent educational environment must be created.
Sophisticated eLearning mechanisms are realised by summarising a framework for Semantic Web architecture. Also, a number of issues related to eLearning using semantic Web Technologies have been discussed. In conclusion, the paper discusses the model and lists all the future work that can be made in the model [6].
E-Learning should be personalised. The system should be capable of delivering results in response to each student's needs and tastes. These set of likes or dislikes should even be based on past records. Observations made during the course of a student's learning should all be made a note of and the system should be able to model a learner's current profile [7]. This way, it can even suggest possible courses for each student.
Three stages are identified when it comes to learning – pre-learning, learning and Post-Learning stages. Each of these stages is used in a semantically designed E Learning system so that the computer may learn about the interest of each of the learners. This concept is explained by [8] using services like Learning Object Metadata(LOM), Sharable Content Object Reference Model (SCORM), Learning Design and other pedagogy research in semantic e-learning
A mechanism to introduced a web based learning can also be introduced such as the one suggested by [9]. First, the contents of a learning system are divided into two categories- learning services and assessment services. A teacher can upload documents in many formats. Students' behaviour is continuously observed using the assessment services and the instructor searches for appropriate content for each student specific to his/her needs. These query results are stored in the student's database and may be used in the future. A structural metadata is created for the different subjects/ domains of interest for each student.
All the above mechanisms may improve the current education scenario. This would connect students from the most remote areas of a country and bring education to their door steps.

IV. SEARCH ENGINE

Today's search engines aim at assimilating the pages with the maximum number of 'hits' in the quickest possible time. What we do not know, is whether these 'hit' pages are relevant to the search request. We can take advantage of the Semantic Web technology to obtain cogent results depending on the request sent by each user.
One such search engine developed in 'Swoogle' [10][11][12]. Swoogle is among the first metadata and search engine for the semantic web. It uses a data centric architecture and is also extensible – meaning the different components can work independently and interact with one another through the database. Each sentence in a semantic document is tagged by RDF triplets. These triplets
To compute the importance of each Semantic document, Swoogle uses the idea of 'rank'. To be used in the search engine later. This way, pages based on 'rank' may be displayed on the output screen. The search engine 'Swoogle' proposed is semantically designed and incorporates three tasks-

- Finding appropriate ontology
- Finding instance data
- Characterizing the semantic web

These activities are accomplished by four separate components – Semantic Web Document (SWD) discovery, metadata creation, data analysis, and interface.[10] This can also be explained by four functions discovering, digesting, analysing and serving. In the first stage, the system updates itself constantly about the different SWDs available throughout the web. Next, at the semantic and syntactic level, metadata is created by caching snapshots of the objective details about the SWDs.
Recently, Swoogle has discovered 346,126 RDF documents with 65,747,150 triples. Although this number is still trivial in comparison with 8,058,044,651 web pages indexed by Google, it is a big number of semantic web researchers [11].

The next step to test the quality of the results is by using 'context' of the web pages. To achieve this, Web Of Belief (WOB) is created and then is modeled in all three different worlds – The agent World, the RDF graph world and The Web. The service component supports both human and software agents through conventional web interfaces.

V. STATISTICS

Machines find it difficult to comprehend the statistical data like graphs, tables and charts published by organizations on a daily basis. This is because organising and analysing these datasets is found to be extremely hard. A new method like Semantic Web shows promising results for the same.

Integrating data by linking them is found to be far more effective than a regular relational database. Scientists may put this linked data to further investigation and research. A few assessment tests for to help researchers consolidating information from many different sources have been defined in [13]. Moreover a number of statistical data may be compared at the same time.

One method for integrating scattered data is by using the Friend Of A Friend (FOAF) policy. Using this method, components can be easily interrelated and new components can be instantaneously detected and added to the network [14].

Linked Open Data (LOD) is a framework where a number of datasets are increasingly published. This mechanism has proven to not only to be able to scale up to larger domains but also operate on sparse datasets.[15]

Another example of an application to deal with semantic data is NESSTAR [16]. NESSTAR 'aims to streamline the process of finding, accessing and analysing statistical information'. NESSTAR aims at making statistical information available to machines and unlock them from being only human understandable. In NESSTAR, each data publisher uploads statistical data in his/her own server. This data is uploaded in the form of objects which are identified by NESSTAR and is semantically stored in a machine readable format.\

VI. CONCLUSION

Still in its rudimentary stage of research, there are a number of drawbacks that have to be overcome before Semantic Web is fully functional. Since the idea of Semantic Web is to make the machine intelligent enough to respond to human requests, all loopholes must be dealt with.

The first problem that we will face will be to determine which information on the Internet is truthful and which is not. There should be a mechanism to maintain trust or credibility of the web pages. Paper 4 suggests that we do this by first selecting a group of users that we trust. Each of these users is given the task of selecting another set of trustworthy pages/users. This way each page is given trust value and the credibility of the page can be thus determined.

Second, in a world where everyone is taking active measure to showcase higher levels of anonymity, the Semantic Web would contrast this goal. Information from social networking sites like Facebook, MySpace, Twitter, etc is going to be added to the Semantic Web database through activities of users who have an account in these sites. Openness and transparency may become a forced option.

Third, mining only the essential details from the web is going to require the best natural language processing tools available. The currently available ones can extract useful information to an extent but they are not guaranteed to do the same once Semantic Web is implemented.

Fourth, even when referring to the same things by different users/websites, the vocabulary used may not be the same. There is a requirement of a tool that matches the vocabulary which if not done would be deemed incompatible and treated separately.

These are just some of the drawbacks which will definitely be highlighted when Semantic be will be put into use. All these deficiencies can be overcome by various technologies that are currently being researched on which will transform the structure of Web 2.0. In any case, as illustrated below, the applications of Semantic Web still overweigh these drawbacks.

REFERENCES

[1] Susie Stephens, David LaVigna, Mike DiLascio, Joanne Luciano, Web Semantics: Science, Services and Agents on the World Wide Web, Volume 4 Issue 3, September, 2006, Pages 216-221.

[2] Huajun Chen1, Yimin Wang2, Heng Wang1, Yuxin Mao1, Jinmin Tang1, Cunyin Zhou1, Ainin Yin3, and Zhaohui Wu, Towards a Semantic Web of Relational Databases: a Practical Semantic Toolkit and an In-Use Case from Traditional Chinese Medicine, The Semantic Web - ISWC 2006, Lecture Notes in Computer Science Volume 4273, 2006, pp 750-763.

[3] Karun Bakshi and David R. Karger, Semantic Web Applications.K. Elissa, "An Overview of Decision Theory,"

[4] A. A. Name, Towards a Semantic Medical Web: HealthCyberMap's Dublin Core Ontology in Protégé-2000, Kamel Boulos MN, Roudsari AV and Carson ER, Centre for Measurement and Information, School of Informatics, City University, London, UK.

[5] Daniel L. Rubin,1 Pattanasak Mongkolwat,2 Vladimir Kleper,2 Kaustubh Supekar,1 and David S. Channin2, Medical Imaging on the Semantic Web: Annotation and Image Markup, Issue No.01 - January/February (2009 vol.24, pp: 57-65.

[6] Khurram Naim Shamsi, Zafar Iqbal Khan, DEVELOPMENT OF AN E-LEARNING SYSTEM INCORPORATING SEMANTIC WEB, International Journal of Research in Computer Science, 2 (5): pp. 11-14, 2012

[7] Nicola Henze, Peter Dolog and Wolfgang Nejdl, Reasoning and Ontologies for Personalized E-Learning in the Semantic

Web, Educational Technology & Society, 7 (4), 82-97.
[8] Weihong Huang, David Webster, Dawn Wood and Tanko Ishaya, An intelligent semantic e-learning framework using context-aware Semantic Web technologies, British Journal of Educational Technology, Vol 37 No 3 2006, pp 351–373.
[9] Fayed Ghaleb, Sameh Daoud, Ahmad Hasna, Jihad M. ALJa'am, Samir A. El-Seoud, and Hosam El-Sofany, E-Learning Model Based On Semantic Web Technology, International Journal of Computing & Information Sciences, Vol. 4, No. 2, August 2006, Pages 63 – 71.
[10] Li Ding, Tim Finin, Anupam Joshi, Yun Peng, R. Scott Cost, Joel Sachs, Rong Pan, Pavan Reddivari, Vishal Doshi, Swoogle: A Semantic Web Search and Metadata Engine, Proceedings of the thirteenth ACM international conference and knowledge management, Pages 652- 659, ACM, New York, USA – 2004.
[11] Li Ding and Tim Finin, Boosting Semantic Web Data Access Using Swoogle, AAAI'05 Proceedings of the 20th national conference on Artificial intelligence, Volume 4, Pages 1604-1605.
[12] Tim Finin, Li Ding, Rong Pan, Anupam Joshi, Pranam Kolari, Akshay Java and Yun Peng, Swoogle: Searching for Knowledge on the Semantic Web, AAAI'05 Proceedings of the 20th national conference on Artificial intelligence.
[13] Benjamin Zapilko and Brigitte Mathiak, Defining and Executing Assessment Tests on Linked Data for Statistical Analysis, COLD – 2011.
[14] How the Semantic Web is Being Used: An Analysis of FOAF Documents, Li Ding, Lina Zhou, Tim Finin, Anupam Joshi, HICSS '05 Proceedings of the Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05) - Track 4 - Volume 04, Page 113.3, Washington, DC, USA -2005.
[15] Yi Huang, Volker Tresp, Maximilian Nickel, Achim Rettinger and Hans-Peter Kriegel, A Scalable Approach for Statistical Learning in Semantic Graphs, Semantic Web 5(1): 5-22 (2014).
[16] Pasqualino 'Titto' ASSINI (titto@nesstar.com) Nesstar Ltd - U.K., NESSTAR: A Semantic Web Application for Statistical Data and Metadata, International Workshop Real World RDF and Semantic Web Applications, 11th International World Wide Web Conference, (2002).