

PROJECT PROPOSAL: ADDING MEANING TO THE WORLD WIDE WEB

PROJECT OUTLINE

I propose to focus my project around the subject of the semantic web. The semantic web aims to give meaning to the World Wide Web, creating a universal medium that can be understood by both humans and machines. My rationale for this is that I have a strong interest in web standards and making my web sites accessible to as many people as possible. Therefore I would like to research more into this area to find out how the semantic web can help with this.

Currently, most web sites are written in a natural language that can only be truly understood in terms of context by humans. The semantic web seeks to create further understanding of such content. Its aim is to make text and data on web sites understandable to computers and software agents, which will allow them to find, integrate and aggregate data efficiently and effectively.

For example: an address of a business on a web site can easily be identified as such by a human, whereas a computer will have no ability to differentiate this from any other passage of text. The whole concept of the semantic web will rely heavily upon technologies such as Microformats, which add additional HTML mark-up around data and text, allowing them to be easily understood and recognised by computers. Once this data can be found and understood, it can be used in many ways e.g., enabling the address to be found by online business directories and giving people the ability to save the address to their address book in one click.

In my dissertation, I plan to cover the history of the semantic web, what it's trying to achieve, who's behind it and the criticism that it has received since its concept was first introduced.

I also plan to cover the enabling technologies such as: Microformats, RDF (Resource Description Framework), XML (Extensible Markup Language) and HTML (HyperText Markup Language) along with examples, their purpose and the positives and negatives of each.

METHODS

I plan to collect information using books, articles, references and web sites.

Reading books will help to provide me with a solid background to work from. This will help me gain a better understanding of the semantic web and also provide hints to what I need to research in a greater depth.

InfoTrac¹ will give me access to thousands of articles from publications around the world. This is a great resource and time saver because of the amount of full text articles stored in it, which are easily searchable to enable me to quickly find what I'm looking for. I have selected a few articles using the system, which I have discussed in the following section in more detail. They are also listed in the bibliography.

I found some additional information from an encyclopaedia listed on InfoTrac which supplied me with information about the internet's founder. He is also the man behind the vision of the semantic web.

In addition to the above, I discovered some digital web-only magazines publishing quality articles on the subject of web design. Within the content they have published, I found some

¹ <http://infotrac.galegroup.com/menu>

helpful articles relating to the semantic web and enabling technologies that will bring us closer to the vision.

I may also look into conducting a survey and/or interviews among web designers and developers to get a broader view of what they think about the semantic web. Asking questions such as what they like and don't like about it and how they think it will benefit them, their clients and end users.

LITERATURE REVIEW

The first source I consulted was a brief article from Computing magazine titled, 'What is the semantic web?' which gave a short explanation and definition of the semantic web along with an opinion from Richard Edwards, a senior research analyst. This could be useful in the introduction of my dissertation when explaining what the semantic web is. It may also be a good idea to look into Richard Edwards more and try and find out what he thinks of the semantic web in a greater depth.

An article from Digital Web Magazine², titled 'HTML5, XHTML2, and the Future of the Web' discusses future HTML specifications which are in development, along with the features and criticisms of each. They seem to be going towards a more semantic web in that they are adding new tags to the language to make the content more meaningful to computers. It also discusses the shortcomings of the current XHTML specification with the way it is processed in various browsers. This is a useful article in that provides information on one of the enabling technologies that will contribute towards a semantic web.

² <http://www.digital-web.com/>

Martin Kliehm's article 'Accessible Web 2.0 Applications with WAI-ARIA' from A List Apart³ gives advice on the problems and solutions of accessibility in web applications. For example: a container tag for bullet points can have a role of navigation attached to it. Screen readers can understand this and will know that it contains the navigation links for the web site. The article gives instructions on the available options to do this, examining the shortcomings of each. This could form part of the criticism of the semantic web in that it's not always easy to set it up. It can also have problems with browsers not fully supporting some of the approaches.

A blog post on SitePoint⁴ by Matthew Magain thinks we should recognise the internet as an ecosystem that is evolving and we should attach a version number to the internet that increments with each decade, e.g., Web 1.0, Web 2.0, Web 3.0, etc. He has linked to an interesting diagram created by Nova Spivack which shows where he predicts the web is heading over the next few decades. An interesting quote from the post is:

The ultimate goal should be the realization of a Semantic Web, where data is universally searchable, and understandable, by humans and machines.

This could be analysed in my dissertation in more depth. There are also some interesting comments to the post from web designers which could also be looked into.

The above concludes the sources I've already consulted. The following are sources which I have yet to consult in detail.

³ <http://www.alistapart.com/>

⁴ <http://www.sitepoint.com/>

The first book I found was called ‘Semantic Web: Concepts, Technologies and Applications (NASA Monographs in Systems and Software Engineering)’ which was published recently in January 2007. I looked at the synopsis found on Amazon.co.uk which describes it as “well-paced introduction to the Semantic Web”. It covers a range of topics about enabling technologies, both old and new, along with real world examples and how the technology can be extended into other areas, e.g., Geographic Information Sciences, Bioinformatics and Fine Arts. A particularly interesting point about the book is that it is from the book series ‘NASA Monographs in Systems and Software Engineering’. On the publishers web site⁵ it says:

A feature will be an emphasis on the relevance of the technologies described to NASA missions and projects. This will include prior work that has been applied, or is currently being applied, to NASA problems, as well as future work that may be appropriate to apply to issues facing NASA in future space exploration missions.

This sound like it could be fascinating to read and find out how the semantic web has been applied to problems faced by NASA.

The second book I thought may be useful is ‘Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential’ which has a foreword by Tim Berners-Lee, the inventor of the internet and the man behind the vision of the semantic web. From what I can gather from the synopsis, it covers a wide rang of topics from software agents that can negotiate and collect information to knowledge systems that enable machines to read Web pages and determine their reliability. I found the following line from the synopsis particularly interesting as it sums up what the semantic web is made up of.

⁵ <http://www.springer.com/west/home/computer?SGWID=4-146-69-173624766-0>

The truly interdisciplinary Semantic Web combines aspects of artificial intelligence, markup languages, natural language processing, information retrieval, knowledge representation, intelligent agents, and databases.

I found further information on Tim-Berners Lee in the Gale Encyclopaedia of E-Commerce using InfoTrac. It provides information on his background and other web related technologies that he's been involved in.

Business Week Online⁶ published an interview with him in April of this year. I have only briefly looked through it, but it seems to have lots of information about the Semantic Web, how it can be used and some of the problems that come with it. This could be very useful as it has come from the man who is behind the vision and I'm sure it will provide an interesting look into what he envisions for the web in the future and what he's working towards.

The final article I found on InfoTrac was from Computer Weekly⁷ titled 'Prepare now for Semantic Web, urges Berners-Lee.' which contains information from Tim Berners-Lee about the semantic web. This may be similar to the article on Business Week Online as they are both interviews of the same person. I will have to read each carefully and pick out interesting quotes which will be of use to me.

My final resource is a section titled 'Semantic Web' on the W3C (World Wide Web Consortium)'s web site⁸. It contains an introduction to the Semantic Web along with links to technical specifications, publications, articles, interviews and presentations. I think it is worth following some of the links from the page and looking at them in more detail as they are likely to provide some valuable information to help in my research.

⁶ <http://www.businessweek.com/>

⁷ <http://www.computerweekly.com/>

⁸ <http://www.w3.org/>

BIBLIOGRAPHY

SOURCES ALREADY CONSULTED

MAGAZINES & REFERENCE

Friedlos, D (2007) 'What is the semantic web?'. *Computing*, March 22. p. 14

WEB PAGES

Kliehm, M (2007) *Accessible Web 2.0 Applications with WAI-ARIA* [online]. Available at: <http://www.alistapart.com/articles/waiaria> [Accessed 22 April 2007]

Magain, M (2007) *Are You Ready For Web 3.0?* [online]. Available at: <http://www.sitepoint.com/blogs/2007/04/02/are-you-ready-for-web-30/> [Accessed 22 April 2007]

Anderson, D (2007) *HTML5, XHTML2, and the Future of the Web* [online]. Available at: http://www.digital-web.com/articles/html5_xhtml2_and_the_future_of_the_web/ [Accessed 22 April 2007]

SOURCES TO BE CONSULTED

BOOKS

Breitman, K. and Casanova, M. A. (2007) *Semantic Web: Concepts, Technologies and Applications (NASA Monographs in Systems and Software Engineering)*. Springer-Verlag London Ltd

Fensel, D. et al. (2005) *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. The MIT Press

MAGAZINES & REFERENCE

Anon. (2007) 'Berners-Lee, Timothy'. *Gale Encyclopaedia of E-Commerce*, 1. p. 57-58

Knights, M (2007) 'Prepare now for Semantic Web, urges Berners-Lee. (Tim Berners-Lee)'. *Computer Weekly*, March 20

Anon. (2007) 'Q&A with Tim Berners-Lee; The inventor of the Web explains how the new Semantic Web could have profound effects on the growth of knowledge and innovation. (CEO GUIDE TO TECHNOLOGY)'. *Business Week Online*, April 9

WEB PAGES

W3C (no date) *Semantic Web* [online]. Available at: <http://www.w3.org/2001/sw/> [Accessed 22 April 2007]

AREAS ON WHICH I NEED FURTHER ADVICE

I may need further advice on keeping a balance between technical and cultural information as the subject itself has a lot of technical information available due to the many technologies it involves. To combat this I will try to write about the effects that these technologies have/will have on end users and the history behind the semantic web.