# Where will the semantic Web take us with regard to improved access to Web-based information?

**Margaret van Steenderen**
margsvs@ananzi.co.za

With the semantic Web aiming at allowing computers to interpret Web content meaningfully, the question is where this will take us with regard to improved access to Web-based information, specifically with search tools. The meaningful interpretation of Web content will rely on search tools being developed to assist with more precise searching, and Web sites providing accurate information – some through the use of Web servers leveraging the existing information. There are already some exciting developments, such as Intellidimension's RDF gateway, (http://www.intellidimension.com/) and some other interesting projects are also on the way.

One way of improving search capabilities will be the ability for sites to present information that is relevant so that a search program will no longer have to strip out irrelevant data, such as advertisements, to get the relevant data (Berners-Lee and Miller 2002). For example, on a private cat Web site, *The Sims*, which is a game, is mentioned in passing. By analysing the results of queries that search engines bring to the site, *The Sims* is a common query – but actually not relevant to the site at all. It is therefore useless to the searcher. Making the site 'semantically' friendly would mean that only the words that are applicable to the site would be searched. This will prevent users who are looking for something completely different to stumble across useless information.

With one of the aims of the semantic Web being to make the Web more machine readable, it can also be left to the search tools to sort out the data and relationships between these data from various sites and to return the results. So, another way the concept of the semantic Web offers better search queries is through people generating the relationships between data. Examples of these relationships are 'hasLocation, worksFor, isAuthorOf, hasSubjectOf and dependsOn', which now become explicit (Berners-Lee and Miller 2002). These relationships allow more complicated queries to be sorted by the search tool. Figure 1 (Berners-Lee, Jendler and Lassila 2001) illustrates how searches will have the potential to be more elaborate and precise. The search tool must be able to understand relationships such as 'child of' and 'lives in'.

**Figure 1 Potential of figures to be more elaborate and precise**
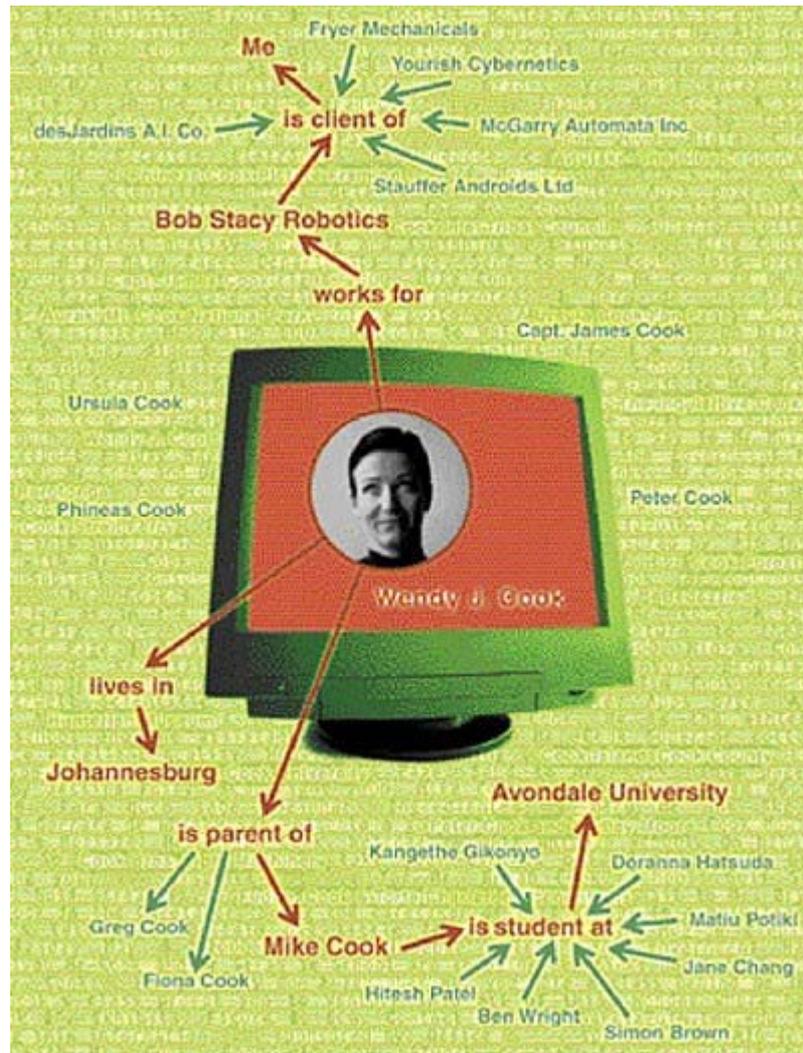
As Berners-Lee, Jendler and Lassila (2001) point out, the word cook would generally retrieve all sorts of irrelevant information, such as information on chefs, recipes, places, people and so on. However, with the collection of partially memorized information, the correct information can be found:

> 'Elaborate, precise automated searches will be possible when semantics are widespread on the Web. Here a search program correctly locates a person based on an assortment of partially remembered knowledge: her last name is "Cook," she works for a company on your client list and she has a son attending your alma mater, Avondale University. The correct combination of that information does not reside on a single Web page, but semantics make it easier for a program to discern the elements on various pages, understand relations such as "Mike Cook is a child of Wendy Cook" and piece them together reliably. More generally, semantics will enable complicated processes and transactions to be carried out automatically' (Berners-Lee, Jendler, and Lassila 2001).

The point of the semantic Web is to make searches more precise, to enable users to find exactly the information they are looking for, without having to sift through many irrelevant search results. Search tools can be improved though the right information being available to them, along with the correct relationships. This is what will enable machines to communicate with other machines more effectively, and in the long run lead to users trusting search results. The information can potentially have more meaning.

In the next issue, the functioning of the semantic Web is discussed, taking into account some projects that are already running, as well as putting together the big picture of the W3C's proposed standards (http://www.w3c.org/2001/sw/).

**References:**

Berners-Lee, T. and Miller, E. 2002. The semantic Web lifts off. [Online]. Available WWW: http://www.ercim.org/publication/Ercim_News/enw51/berners-lee.html (Accessed 14 February 2003).

Berners-Lee, T., Jendler, J. and Lassila, O. 2001. The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. [Online]. Available WWW: http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&pageNumber=1&catID=2 (Accessed 14 February 2003).

**About the author**

Margaret van Steenderen has been working in information technology since 1989, starting in support and training, and later moving on to system administration and system implementation. She has always had a special interest in the training side of information technology and currently teaches business information systems and end-user computing. In 1998, she studied part-time obtaining a BA Humanities in 2000 and Honors in Psychology in 2001 (both *cum laude*).

**Disclaimer**