The Semantic Web

Introduction

Undeniably Information Management is now a major issue of World Wide Web, as we can see, the problems is increasing on the web, e.g., the search engine only can search by the keyword, and it is returning a huge number of results which just including a small amount results meet user’s requirement. Users have to read through the results to find out the useful information.

The Semantic Web has been bring forward to address this shortcoming, As Berners-Lee said, the future of the web is Semantic Web, Current Web just a “container” of the information, it can’t manage the information, The Semantic Web is able to understanding the content of information, and it can classify or manage the information by the content, it is a real “information manager”.

The article “The Semantic Web” discusses the eventual use of the Semantic Web instead of the World Wide Web. Use a dummy story to show how the computer software agents carry out tasks for user.
Current World Wide Web & Semantic web

As the Current World Wide Web is based primarily on documents written in HTML, HTML has limited ability to classify the blocks of text on a page, apart from the roles they play in a typical document's organization and in the desired visual layout. For example, HTML is able to make assertion such as “This document’s title is ‘The Semantic Web’” But there is no way for HTML to understand the content, it is not able to say “This is an article”.

The author of this article, Berners-Lee, Tim thinks The Semantic Web can addresses this shortcoming by using the descriptive technologies RDF and OWL, and the data-centric, customizable markup language XML. These technologies are combined in order to provide descriptions that supplement or replace the content of Web documents. He thinks The Semantic Web will extend the ability of the World Wide Web through the use of Standards, markup languages and related processing, information is given well-defined meaning etc.

But I think to give “well-defined” meaning for each information on the Web, is Impossible to be done manually, like there are incalculability number of information on the web, and the number is increasing every hour, it may cost lots of human resource to do this job, there may be another way to achieve this objective is by using machine to do it, but there is another problem occur, is the machine able to define the meaning of an information, how long takes to make the machine able to understand the content of the information.

There is other main writer in the area, Clay Shirky’s, who write the article “Ontology is Overrated: Categories, Links, and Tags” He said “Ontology is overrated, many of the ways we're attempting to apply categorization to the electronic world are actually a bad fit, because we've adopted habits of mind that are left over from earlier strategies”

I do agree with him, poor classification may block the performance of semantic web, even in a nearly perfect categorization scheme, there maybe some other “pullback”, eg. Context errors, or the skill of the coder, or maybe the one who doing the classification misunderstanding the information etc.
Conclusion

This article gives us a great idea of what is semantic web, how it is works, and how is it eventually to be used. In which content on the Internet will be structured so that computer can carry out tasks for users by using software agents. That sounds we going to have an artificial intelligence Web in our world, but there are a lot of puzzle need to be solve, as the semantic web is based on machine and a set of operating rules, how we define these rules, and how can we make the machine to understand the contents

Reference

http://www.shirky.com/writings/ontology_overrated.html