The semantic web - a muggles* view

By Chris Gledhill, Managing Director, PDMS

For some time now Internet pioneers such as Tim Berners-Lee the inventor of the World Wide Web and founder of technical standards guardians The World Wide Web Consortium (W3C), have been working on the next generation of information sharing systems. They call this the Semantic Web, a system for capturing the underlying meaning of electronic data and thereby adding a whole new level of intelligence and order to the vast oceans of data sloshing around on the world's computer systems.

The current web was first developed by Berners-Lee whilst working at CERN (the high energy physics research facility) in Switzerland as a tool to publish the internal phone directory (allegedly). It became public property on the 30th of April 1993 when CERN announced that it would be free for all to use without any royalties.

Since that time the three Ws have, for most of us, become synonymous with the internet and represent the greatest ever collection of information, disinformation, fact and fiction ever assembled. The web is like a vast library in which the books are completely randomly arranged, each book has a predefined space on the shelf but there is absolutely no categorisation. Search engines such as Google act as librarians to help us find the resources we want, but in a most bizarre way. A web search engine is like some sort of autistic genius with an immense memory and perfect recall of the words in each book but absolutely no idea what they mean.

So instead of being able to ask Google to direct us to the section on advanced bungee jumping techniques or DIY space exploration, we have to think of some words that are likely to appear in the books we want and then wade through all of the irrelevant rubbish that happens to contain the same words in the hope of finding something useful.

The development of the web is (despite Microsoft's best efforts) a testament to the power of standards. The explosive growth of the web is based on the universal adoption of Berners-Lee's standard for URLs (uniform resource locators) or web addresses as we now think of them. Recently, the W3C completed work on the standards that will enable the Semantic Web to come into existence: the Resource Description Framework (RDF) and the Web Ontology Language (OWL). Together these standards create the foundation for a new, far smarter, network of information.

Fans of Harry Potter may be rather miffed to see the letters OWL which clearly stand for 'Ordinary Wizarding Level' hijacked in this way by the W3C. However the application of Ontology to computing also has some rather magical possibilities. In philosophy Ontology is a branch of metaphysics which seeks to find out what types of entities exist and to categorise them. In information science, by derivation, an ontology is a formal (and rigorous) description of all of the entities, relationships and rules within a particular domain of knowledge. The magical bit happens when you have created an ontology in a particular field. It then becomes possible

for computer software to make logical inferences from real world information to 'discover' new facts which have not been explicitly encoded.

In pharmacology an ontological schema could make the conection between the effects of a drug and the symptoms of an apparantly unrelated condition. For example, sildenafil citrate (aka Viagra), a drug which improves bloodflow, and impotence, a condition resulting from restricted bloodflow, even though sildenafil was originally developed as a heart medication.

So will the Semantic Web take over from the existing technicolor chaos of the three Ws? Or is it another compelling fantasy best enjoyed in the same spirit as J.K. Rowling's latest offering (she had better get on with book seven or there will be ructions in the Gledhill household). My view at the moment is a definite maybe. The wonderful thing about the existing web standards is that they do not embody any kind of opinion. Having a single standard for addresses is just common sense like not duplicating phone numbers or all driving on the same side of the road. The Semantic Web on the other hand requires us to agree (understand) and stick to predefined ontologies and I fear that this may be beyond the magical powers of even the greatest technical wizards.

Behind the scenes it is a different story. In specialist applications from counter terrorism to pharmaceutical research, the application of ontology and the potential of systems capable of inference is being taken very seriously. Similarly, the Semantic Web will be welcomed in any commercial field where the economic advantages of standardisation outweigh the benefits of confusing the opposition, punter or regulator. In other words, just like the real world, there will be many versions of the truth which will sit along side, rather than replace, the creative anarchy of the web as we know it now.

*Muggles - non magical folk

Tel: +44 (0)1624 823833 Fax: +44 (0)1624 825640 E-mail: enquiries@pdms.com