Exploring the Semantic Web as a viable ontology for Knowledge Management

An undergraduate dissertation submitted in partial fulfilment of the requirements of IS335, the Information Technology Applications Module of BA (Hons) Computing and Information Systems.

Graham Robbins
February 2005

Approximate word count: 9,988

Abstract

The Semantic Web, a World Wide Web Consortium initiative, endeavours to seek out the hidden meaning behind data. Determining knowledge by forming associations between known facts, the Semantic Web's conclusions are derived through logical reasoning. Closely linked is the concept of Knowledge Management, a term coined for the systematic storage of an organisation's information. Following the growth of electronic media traditional methods of seeking data fall short of user needs. XML, RDF, OWL and other technologies are together realising the Semantic Web's potential. The impact on Knowledge Management is extensive, presenting the possibility of truly knowledge-centric organisations. The Semantic Web is a Herculean undertaking, yet the work of current academic and commercial projects offers viability to its argument. Given the nurture the Semantic Web richly deserves it may give rise to the renaissance of the knowledge age.

Where is the life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?

T.S. Elliot, The Rock (1934)

Contents

List of Figures

Introduction

Defining the Semantic Web is a greater challenge than comprehending its mechanics. A vital part of the World Wide Web Consortium's (W3C) long term strategy (Herman 2000) the Semantic Web is shaping the future Internet. The brainchild of Tim Berners-Lee, the W3C aim "to develop a software environment that permits each user to make the best use of the resources available on the Web" (Herman 2000).

The W3Cs goal is broad, and hardly unique. Rather, it is the sign of a maturing software industry (Antoniou 2004) whereby information use is evolving, "transforming the developed world toward a knowledge economy and, more broadly speaking, to a knowledge society" (Antoniou 2004, p. 1). Antoniou further states that computer systems have progressed from bespoke solutions to generic applications, computers are now perceived differently, "the purpose of computing is insight, not numbers" (Hamming 1962).

With the aim of realising a knowledge-centric society by means of the Semantic Web the first task is to appreciate what knowledge truly is. Subsequently analysing the Semantic Web Stack, see Figure i overleaf, lays bare the internal components. At its foundation shared technologies, like XML and Unicode, provide a platform to build Semantic Web technologies upon.

Further study of central topics, such as logic, reasoning and ontology, all of which have existed outside of computing for centuries, reveals why the Semantic Web is indeed required.

Rather than ascertaining exactly what the Semantic Web is, it is worthwhile determining what it is not. The Semantic Web is not a new Internet, it is an additional layer, where information is extended with well-defined meaning, and becomes machine-processable in a manner which is “ubiquitous and devastatingly powerful” (Shirky 2003). Thus attention will be given to how the Semantic Web can be integrated with the existing web and whether a complete conversion will ever be viable.

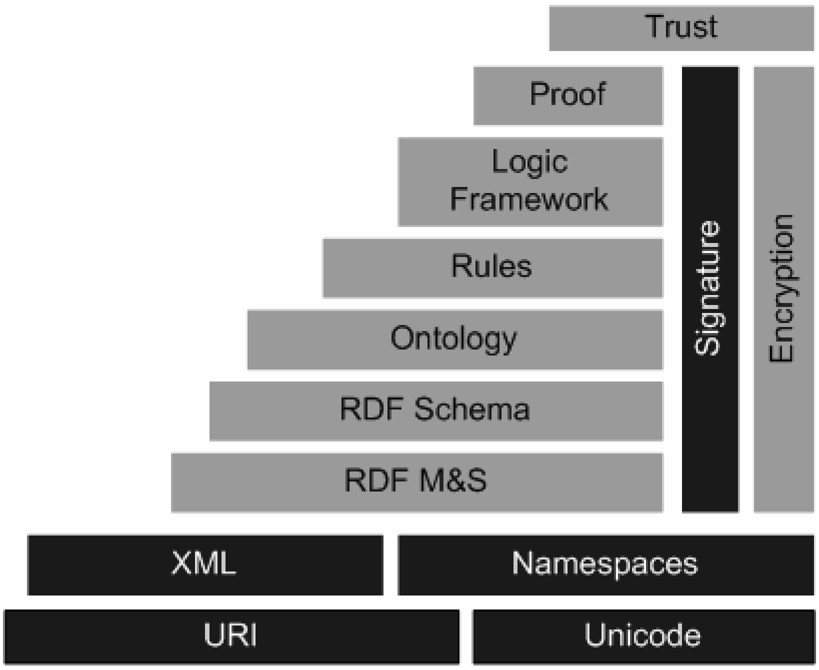

Figure i - The Semantic Web Stack
http://www.w3.org/2003/Talks/1022-Madrid-IH/32.html

The Semantic Web concerns extracting meaning from raw, unimpressive and ambiguous data. The driver is how to contrast two opposing worlds.

> “The philosophy was: what matters is in the connections. It isn’t the letters, it’s the way they are strung together into words. It isn’t the words it’s the way they are strung together into phrases. It isn’t the phrases it’s the way they are strung together in a document”.
>
> (Berners-Lee 1999, p. 13)

In his book “Weaving the Web”, Berners-Lee (1999) argued connections were key, the Semantic Web has well formed relations, not intelligence. Visualising the Web as associated objects can, as I shall establish, demonstrate fact.

Preceding the modern day Internet Berners-Lee presented a proposal concerning information management at CERN, now known as "European Organization for Nuclear Research". Berners-Lee demonstrated how a 'mesh' of interconnected documents could resolve data loss. The content of the document later evolved into the present web. However Berners-Lee's conclusion was rather different:

> "The aim would be to allow a place to be found for any information or reference which one felt was important, and a way of finding it afterwards. The result should be sufficiently attractive to use that it the information contained would grow past a critical threshold, so that the usefulness the scheme would in turn encourage its increased use."
>
> (Berners-Lee 1989)

The web is comparable to the original 'mesh' Berners-Lee envisioned. The seed for a Semantic Web had been sown, to finally achieve the goals in Berners-Lee's conclusion. Thus the driving question is can Semantic Web technologies deliver upon these goals? Moreover will their use result in a paradigm shift in knowledge management, verifying the viability of the Semantic Web initiative?

The Knowledge Gap

The need for information is great, but the loss of knowledge is equally massive, the philosopher Kant wrote, "If we only knew what we know, namely, in the use of certain words and concepts that are so subtle in application, we would be astonished at the treasures contained in our knowledge" (Kant 1800), from academic research to cinema time's information is the 'oil of civilisation' (Clegg 2001).

In his book, the "Professionals guide to mining the Internet", Brian Clegg (2001) outlines the differences between the Internet and libraries, the previous centres for knowledge. Primarily the Internet has no organisation or authority. Secondly the information is not always truthful. Gaps exist, culminating in areas of contention. Thirdly, users reduce their search to key terms, English verbosity is lost.

These key differences illustrate the current Web is not a salvation for information seekers. Antoniou (2000) identified problems with existing web technology; search engines suffer from high recall and low precision. Worthy results are interspersed with scores of others. Searches are sensitive to vocabulary, and only single web resources are returned.

> "The main obstacle to providing better support to web users is that, at present, the meaning of web content is not machine-processable".
>
> (Antoniou 2004)

The W3C aims to represent data in a manner which is meaningful to machines. This approach may require other changes, in late 2004 Pat Gelsinger, Intel's chief technology officer, said the Internet was stretched "to breaking point". Gelsigner advocated overlaying a new network upon the old, arguing better network security and integrity could be offered alongside data and information management, improving usability for all users (Anon 2004).

During the Web boom of the nineties the content available appeared to be huge, never ending streams of multimedia alongside office documents and emails (Oram 2002). Since the turn of the millennium the resources have grown further. Compounding this Paul Krill, during 'InfoWorld' in January 2000, discussed the importance of information overload. Krill argued corporate days are twenty-four hours long, data must provide a rapid response, and the propagation of the Internet has made matters far worse.

The result is a mixture of unrelated data needed in real time. The W3C hopes the Semantic Web can increase the quality of results, improving upon current information management which it describes as disconnected and vague (Daconta 2003). With user knowledge of the domain humans can conceptualise the meaning behind the data. Labelled by the authors as a trend, emerging in pilots, the Semantic Web is 'not a fad but a direction', the culmination of years of research.

> "The Semantic Web is different. It's a space of data. It's all the information which is now in databases, spreadsheets, and application-specific files, like calendar files or photo metadata. What's exciting about the Semantic Web is its potential for serendipity, the unplanned reuse of data. The effect will be even more powerful for the Semantic Web because you won't have to be a person following the links. A machine will be able to follow links."
>
> (Berners-Lee 2004, see Reinhardt 2004)

Clearly the Semantic Web is an ideal of the W3C, to fill the gaps in knowledge left by existing methods. To realise this, an understanding of the data must be sought, which is achievable by representing the data in a manner consistent with the knowledge it holds.

Foundation Stones

In Figure i, The Semantic Web stack shows four foundations. Each of these supports the later logical reasoning.

Unicode

Data is stored as binary, encoded for a particular characters language. The most simple to visualise is Morse Code a combination of just two symbols, dots and dashes, which can be transmitted by a number of mediums and understood at both ends. Traditionally ASCII was used, housing the entire Western Alphabet. However this is insufficient for XML, the Unicode Consortium strives to support all known world scripts, whether they be alphabetic, symbolic or ideographic – even this is not enough for all the known languages in the East. The Universal Character System (UCS) will in the future attempt to catalogue everything, storing approximately two billion symbols.

Unicode was selected for XML documents as a means of universally storing characters for all humans, however only a subset is usually required.

URI

Objects, as shall be shown, can be represented by URIs. A simple premise for any Web user, however it must be noted that URIs are different from URLs, the latter points to a network resources, whereas the former refers to “anything that needs to be referred to in a statement” (Manola 2004). Thus URIs are perfect for the Semantic Web, linking conceptual objects.

XML

During the past decade computer science has been inundated by new acronyms, revolving around XML. XML aims to improve upon HTML through accuracy and flexibility (Herman 2000). XML includes metadata, any stored data, such as 'Adam Smith' is accompanied by 'name' or more verbose, 'employees individual name'. Therefore XML data is said to be self describing, by allowing fabricated names for metadata the author is able to explain their work as it is stored. XML is the "grease on the wheels of the information structure" (Ray 2001).

XML has numerous benefits,

- XML can retain organisation of data while stored
- XML is an open standard
- Unicode forms the core of XML
- Documents have a certain 'quality' that non-XML documents do not
- XML is easy for both humans and computers to interpret
- XML is easily styled and transformed

(Ray 2001)

Older editors such as TeX and LaTeX stored formatting alongside data, the resulting document was troublesome for computers attempting to distinguish data. With IBMs Generalised Mark-up Language (GML) generic styles were used. Developing into the Standard Generated Markup Language (SGML) which was flexible, but large.

The HTML Tim Berners-Lee worked upon solved a communication problem, growing from an in-house tool to a worldwide platform. However HTML was oversimplified, it sacrificed principles for ease of use and was wholly presentation orientated. Getting back to the ideals was to return to SGML, this was too big to handle, thus XML was born.

In the mid nineties the W3C set several goals for XML to achieve, before forming a recommendation. These concerned the nature of the mark-up, lack of ambiguity, error checking and separation of style and presentation. With the ability to describe any language "XML stands a good chance of becoming the lingua franca [standard] for computer communication" (Ray 2001).

Namespaces and Schemata

A term coined by Aristotle, 'schemata' are used throughout computing, for Aristotle they were patterns of valid syllogisms, organising elements of knowledge into larger forms, Kant commented "schema [can] … accommodate all possible aspects of a concept" (Kant 1787). For XML schemata constrains data through vocabulary and structure. Schemata relies heavily on namespaces to uniquely identify data structures. For example, company departments have employees, to indicate one particular department the employee metadata is linked to a namespace, such as 'accounts'.

Using XML any concept can be recorded in a machine-processable manner, Unicode allows worldwide participation while the use of unique URIs and Namespaces add scalability to the Semantic Web initiative.

From Aristotle to the Semantic Web

Syllogisms

Before analysing the Semantic Web's Logic Framework it is important to understand Knowledge Representation. Usually concerning concepts, contents and meaning, knowledge is incomplete without logic. Over seven millennia ago Socrates recognised humans "knew very little". Logic has been fundamental to the understanding of knowledge ever since. Studying the 'nature of knowledge and its justification' Plato extended these concepts, christening them epistemology. Later Aristotle initiated a paradigm shift, representing pure knowledge through logic (Sowa 1999).

Aristotle coined many everyday phrases. From Greek; category, metaphor and hypothesis, while from Latin, quantity, quality, genus, species, noun, verb, subject and predicate, amongst others (ibid). These words, alongside Aristotle's work, are fundamental to Semantic Web technology. Aristotle was the originator of predicate logic. Otherwise known as First Order Logic, this has become the cornerstone of Western civilisation through its application in science (ibid). Database management systems and artificial intelligence use predicate logic as a "discipline that studies the principles of reasoning" (Antoniou 2004), and the Semantic Web is no different.

Reasoning

Aristotle demonstrated logic, in a series of canonical syllogisms, could derive conclusions from reasoned fact. In the article "The Semantic Web syllogism and worldview" Carl Shirky (2003) explores their use and importance to the Semantic Web. A syllogism combines a major premise with a minor to derive a conclusion, thus;

Humans are mortal

Greeks are humans

Therefore Greeks are mortal

This is a reasonable conclusion. The two premises, presented as fact, are rationalised with logic. If Greeks are humans and humans are mortal then Greeks must be mortal. The axiom arrives at an a priori conclusion, based purely on what is generally understood. The Semantic Web, in turn, makes assertions of fact. Take for example;

Peter created the document in question

The document's author studied in London

Through syllogistic logic, one is able to ascertain the location and name of the author, thereby establishing truths from fact via association. Aristotelian syllogistic logic is enormously powerful, applied over thousands of years to discover and demonstrate reasoning, containing the knowledge of 'what should be so' (Sowa 1999). However Carl Shirky draws attention to their use in everyday life, which is limited. Life is not a combination of concatenating single fact premises.

Representing Logic

Nonetheless logic can be reduced to simple assertions, in the words of Conan Doyle (1892 see Doyle 2004, p.12) "when you have excluded the impossible, whatever remains, however improbable, must be the truth". Sherlock Holmes, as a detective, was referring to deductive reasoning. Still the same principles can be seen in syllogisms. Continuing this allegory, unlike logic, the Semantic Web is less interested in establishing strict truth than in making associations themselves, mirroring the ideals of Douglas Adam's Holistic detective's theory of the "fundamental interconnectedness of all things" (1988). The Semantic Web considers everything to be connected, via a series of pathways, if a premise can be proved as false it is as important, as one which is true.

This approach pictures the world as simplistic and reasonable. Labelled as data, the raw form is devoid of meaning. Once this data has been extrapolated by a knowledgeable source it can be renamed, information. This extrapolation is rarely based on deductive logic, thus is the world as simple as suggested?

As a consequence, Shirky (2003) believes, the "Semantic Web fails in one swoop of logic" it is unable to turn data into information without knowledge of the domain, which it lacks. Shirky's eagerness to signal the death of this emergent technology may have been premature. For instance he uses two syllogisms to demonstrate that reliable logic cannot always be deduced;

> The last Olympics were held in Athens
>
> People who live in Athens speak Greek
>
> Therefore the Olympics were held in Greek

The conclusion is sensible, and perhaps partly true, for Greek was the local language. The information, however, is insufficient to form a conclusion of the language used by athletes and spectators. Nevertheless it is presented as fact. In his second example Shirky is even more simplistic;

> All dogs have paws
>
> My cat has paws
>
> Therefore my dog is a cat

In philosophy this is identified as the 'fallacy of the undistributed middle' (Sowa 1999), what Shirky failed to appreciate in both examples is the misleading notion of mixing universal and individual logic. Namely the major premise, 'all dogs have paws' is universal to the canine species. Paws are a characteristic of canines. The minor premise, 'my cat has paws' is individual, relevant only to itself. This does not indicate individuals cannot inherit attributes from their universe, but logic cannot be deduced across the divide. It is conceivable that not all dogs are born with paws; their DNA is still that of a dog, the individual is independent of the universal whole.

Syllogistic fallacies, like those described, are driven by generalisations, and oversimplifications, thus;

> "Arguments … depend upon ambiguity. 'Those learn who know: for it is those who know their letters who learn the letters dictated to them'. For to 'learn' is ambiguous; it signifies both 'to understand' by the use of knowledge, and also 'to acquire knowledge'".
>
> (Aristotle 350_{BCE})

Consequently, statements cannot be manipulated scientifically for their meaning is not clear. Gottlob Frege held that predicate logic lays bare the 'misconceptions' that arise through the use of language (see Sowa 1999). However, John Sowa believed any factual information stated precisely can be represented in logic.

The Semantic Web is not immune from science, there are two opposing views as to how the task should be approached; top-down and visionary or bottom-up and functional. Shirky suggests a reliance on top-down ontology, empowering the argument for ultimate failure. The Semantic Web has set high goals for itself. It is unlikely it will succeed to a degree of forming a 'new world order' (van Dijck 2004) where all data is meaningful. There is no monolithic ontology to define the entire world, the point of ontology is to apply ideas to life, not alter that life. The bottom-up approach may be more accurate for the organisation, but leads to divergence, “We invent micro-worlds, where bottom up design leads to a different ontology result in each case” (Sowa 1999).

In his rebuttal of Shirky’s work “Themes and Metaphors in the Semantic Web discussion” Peter van Dijck (2003) concedes it is not feasible for all data to be in a Semantic Web. Even so, the needs of a companies internal data could be met by such a Web. The all encompassing mighty Web that Shirky advocates is not realistic, but idealistic. The vision may exist but the implementation is organic, the Semantic Web must grow.

In an additional and final denial to the claims of Carl Shirky, syllogisms are indeed used everyday in the implementations of databases. Databases store grouped simple statements which are analysed and interrogated by complex management systems to learn facts. Knowledge contained within a database would be better described as a series of axioms. Any statement in a database is assumed to be true, but it must be declared as only ‘contingently true’ (Sowa 1999) because it could be false. For example;

> Sarah is the mother of Rachel

Before Rachel was born that statement was not true, it was however absent as Rachael did not exist, thus no inconsistency occurred. The next statement;

> Every girl has a mother

Is declared as constraint truth, when Rachel’s existence is entered her mother’s existence is implied, regardless of name.

Armed with an understanding of knowledge representation as associations between objects, the idea of a Semantic Web solidifies.

Logic Framework

Resource Description Framework: RDF

The Resource Description Framework (RDF) is one of two technologies vital to the success of the Semantic Web, the other is the Web Ontology Language. Working from the XML framework RDF is a Knowledge Representation Language making assertions about facts (Manola 2004). RDF identifies and associates subjects and objects with a predicate. The series of connections are made in ‘triples’, three-tuple associations, identical to Aristotelian syllogistic logic. The power behind RDF for the Semantic Web is drawn from reification, statements about concepts can be grouped together, to learn.

The first stage in forming meaningful information in RDF is by writing a simple, almost juvenile, description of the data to be stored: “There is a film titled ‘Monty Python’s Life Of Brian’ directed by ‘Terry Jones’ the Internet Movie Database (IMDB) stores film and director information at URI http://imdb.com/title/tt0079470/ and http://imdb.com/name/nm0001402/ respectively.”

To encode this within RDF the three-tuples are formed, as (subject, *predicate*, object).

The web resource in the subject represents a film stored on the IMDB web site.

(http://www.imdb.com/title/tt0079470/, *rdf:type*, imdb:film)

The qualified name (qname) rdf:type is the equivalent of “this object is a …”.

The following web resource represents the entity director of another web resource.

(http://www.imdb.com/name/nm0001402/, *dc:director*, http://www.imdb.com/title/tt0079470/)

The web resource has a fixed title.

(http://www.imdb.com/title/tt0079470/, *dc:title*, "Life of Brian")

When entities are standard the terminology used to define them can be shared, here "*dc*" is defined by the Dublin Core.

Amalgamated statements, such as these are machine interpretable, humans are also capable of making associations of fact. Extending the power of the Semantic Web, the Internet Movie Database is a Web resource. Visiting the URIs in a web browser displays HTML, however consider a Semantic Web where the machine could independently visit the URI and receive further triples. In our example the director is defined only by a URI, his name is unknown. However the additional triples received could contain extra detail; name, birth date and a photograph for example. Thus adding to the richness of the data the Semantic Web extracts.

Figure ii, overleaf, illustrates the second RDF statement as a graph, it is now easier to conceive of interconnected statements, all deriving meaning about one object.

The example given is an instance of the 'network effect' by identifying the objects in a generalised manner the types are shared, not defined for each document independently.

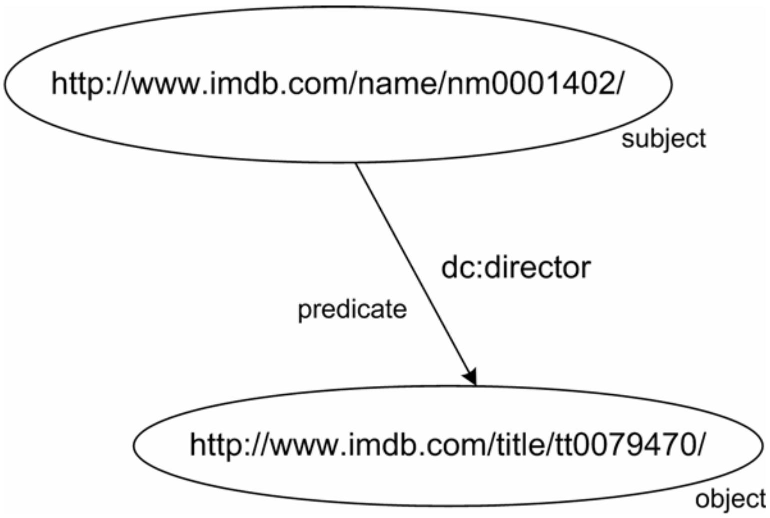

Figure ii - RDF Statements as a graph

RDF has not seen support in the commercial world like XML, Michael Daconta, Leo Obrst and Kevin Smith (2003) believe this is due to misconceptions. RDF is not XML, they are different but can work together. Secondly RDF is complex, it requires understanding of the domain and technology. Concepts must be mapped through reification to real world objects, and by following a chain of compounding assertions conflict and contradiction emerge.

RDF is however already used in collaborative endeavours. The Friend of a Friend (FOAF) community project, from Dan Brickley and Libby Miller, allows users to store their personal details as an RDF document. Specifying the vocabulary to allow the diction of ‘friend’.

RDF development has a long lead time, it must be cultivated to produce a useful result. In contrast XML benefits from a simple understanding, that better definition and interoperability are useful techniques, it requires little re-education and so can be implemented in the short to medium term.

RDF also contains schemata like XML, (perhaps the worst named of all the Semantic Web technologies) unlike XML schemata which constrain data, RDF schemata are a means of definition and liberation, where the authors own vocabulary can be created for the different classes of data to be stored. Through this non-contextual modelling a business document can be written, for example all businesses have a similar purchase order, any anomalies for particular organisations can be identified and added. The foundations of logic are the same.

RQL

The freedom that XML gives to the author is impressive, however the flexibility of recording the same data in several different manners is a hindrance to the Semantic Web. Take for example an XML cinema booking system, at its core it records the cinema complex, screen and times of their films. Another cinema implements a similar system. Written differently but achieving the same result, the XML could be quite different. XPath is capable of complex data mining, however it is unable to seek data, requiring prior knowledge of the XML structure. Antoniou and van Harmelen (2004), believe The RDF Query Language (RQL) is the answer, as it understands the model and adapts to the layout.

RDF has backing by the W3C as the principal Semantic Web technology. The data it stores, explained first by Aristotle, is capable of mining everything end-users require. However context is ambiguous, RDF has no comprehension of the world it describes, the Semantic Web Stack (see Figure i) transfers responsibility to ontology.

Ontology

The ontological spectrum

In order to understand ontology it is critical to understand its origins. In philosophy Ontology is the study of existence and being. David Koepsell (2000), from the Centre for Commercial Ontology Prospects, describes how ontology has been commandeered by computing science. "By prioritising items in a list, we are assigning relationships among the various things. Ontology can be relatively simple or it can be quite complicated". From here on for the purposes of computing we use the term ontology, while Ontology will be reserved for use by our philosophers.

Consequently, ontology is the conceptualisation of the world, logic cannot define something into existence, it must previously have existed, and ontology is the study of this, where vocabulary combines with meaning. "An ontology is an explicit and formal specification of a conceptualisation". (Studer 1998). Put simply ontology describes the real world, aiming to achieve a shared understanding of a domain. An ontology presents a list of concepts and the relationships between these concepts.

As ontology is a manner of demonstrating meaning, corresponding with the human understanding, it must be absent of ambiguity, representing the essence behind the knowledge, "It supplies the predicate of predicate logic" (Sowa 1999). Any representation can yield information, whether graphical or through RDF. The use of a Knowledge Representation Language demonstrates meaning behind the letters. Once interpretable by computers they interact at the human level. To achieve this, John Sowa (1999) claims we must observe. Observation and reasoning are inherently linked, reasoning is the act of making sense of observations made about the physical world, from which knowledge is gained.

Students of computing and logic are often taught ‘garbage in, garbage out’. This creates a flawed connection between data and processing. Data, unless syntactically incorrect, is important. In the Semantic Web approach all data could be perceived as garbage until it proves itself as reliable. Are customer name and addresses meaningful data? After verification that the customer does indeed live at the supplied address the data has a whole new meaning.

Natural language uses combinations of defined terms, which are ambiguous in nature. Although standards exist, and grammar rules are used, organisational language evolves. For example in computing the term ‘string’ is very different to a hardware store. It is not immediately obvious why semantics are easily understood by humans, “... the result of successful communication is at the same time a prerequisite for it” (Davies 2002) for humans have established over time an ever increasing complexity of language, it is self reliant. Ontologies for natural language have been written, WordNet calls itself the “online lexical reference system”, under the direction of George Miller (1995) has categorised English words into concepts and ‘synsets’.

The Semantic Web’s definitive goal is a situation where, given the same symbol based input, a machine and human would be capable of reaching the same conclusion. Thus data must be informative; “To describe a complex idea in a useful way is to use ontology” (Daconta 2003). For this, formal vocabulary must be presented in statements, later correlated into semantics.

For example mapping the statement ‘4 + 3’ as the addition of two integers. The three symbols provided must be matched with a pre-existing ontology about numbers and a conclusion drawn. Even with elaborate semantics the Semantic Web is not creating awareness or reflection from the computer program.

Terms are purely symbols devoid of meaning without association. For instance, the phrase 'R123ABC' has no meaning. However associated with the object 'automobile' it becomes apparent that the phrase is a vehicle registration number. Ontology captures the meaning that is useful for humans, after refining the original phrase is richer.

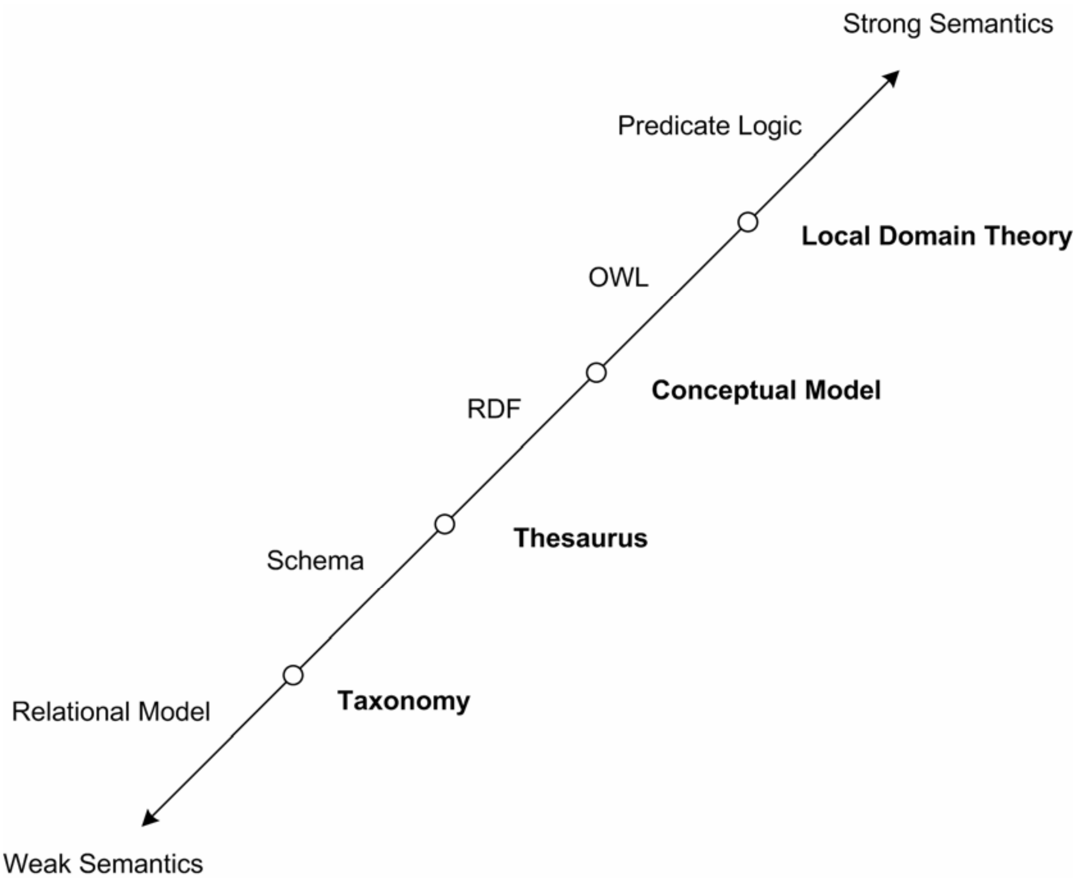

Figure iii - The Ontological Spectrum

Based on 'The Ontological Spectrum', The Semantic Web, P157

The ontology spectrum, Figure iii, demonstrates different levels of richness within Semantic Web technologies. Measured in terms of strength, semantic richness increases the higher up the chain. Simple meanings are said to be semantically weak, should it be strong then there is an apparent arbitrarily complex meaning.

Even if a semantic is rich there are distinctions, properties can change with time. For example, the Director General of the BBC is a definitive individual. To account for changing directors, ontology is forced to become a set of generic properties with particular semantics. Put simply, by using representational levels, further distinctions can be made. In our example, the first level of the Knowledge Representation Language maps to classes, in this instance to humans. The second level concerns the object, a person or role – Director General. Lastly at the third level the ontology object is instantiated as an individual – Mark Thompson.

Concepts can be divided into static and dynamic, for example 'human' is a static concept, one is either human or not, it does not change even post death. Whereas a 'person' has additional attributes, the individual is a teenager then an adult, or a client then a user, their role evolves. Consider, an adult, and a baby. Both are at different, definable, stages. Whereas a teenager and a young lady are conceivably the same individual.

Ambiguity does not rest with the tempory nature of some statements. For instance the maxim 'a happy child'. A child's emotive state is short lived and redundant as an indicator. However for a government agency this maybe the exact data required. Assessing a child's wellbeing over a period of time may hinge on accurately assessing their happiness. The child is an object, its pleasure an assertion. Can this be represented in a meaningful style?

Additionally can logic establish roles? Sowa (1999) suggests the terms "Loch Ness Monster" and "Cookie Monster" have two very different meanings. The former monster's name suggests a dwelling while the latter is a type of thief. The application of ontology is the only way to determine the context, and hence meaning, of any statement.

The question of how far ontology can take the Semantic Web was answered by the philosopher and mathematician Willard Van Orman Quine, who asked the ontological question, 'What is there' the answer Sowa suggests, is 'Everything'. In fact "to be is to be the value of a quantifiable variable". Our entire world can be reduced to objects and associations (Sowa 1999).

Application of ontology

The application of logic and ontology can be troublesome. Natural language can define empirical data, but only by sticking to formalisms. Language by definition is adaptable and expressive. This is the Semantic Web's greatest strength and weakness. Additionally all natural language requires background knowledge. The simplest of words contains a wealth of knowledge behind it. Interpretation is assumption based, a simple parser can understand syntax and a more intelligent version will comment on semantics, however a direct question will yield no answer, it requires outside knowledge. Alan Perlis (see Sowa 1999) writes, "It is not possible to translate informal specifications to formal specifications by any formed algorithm", answers are to be provided by humans, domain experts, in the form of further ontologies.

Take, for example, the following;

> Find me the scene from Monty Python where it is asked 'What have the Romans done for us?'

This requires an understanding of the term Monty Python and examination of the films and television scripts for the possible phrase. Should the knowledge representation system have such vast domain knowledge that it finds the reference in the "Life of Brian" and hence displays scene ten, "The

Commandos", relies upon interconnectivity of domain knowledge, ontologies support this by defining the connections and expanding the search vocabulary.

Knowledge Representation

The Principles of Knowledge Representation were explored by Randall Davis, Howard Schrobe and Peter Szolovits (1993), concluding that knowledge representation must fulfil several criteria. A knowledge representation system contains surrogates, which are objects, events and relationships represented purely by symbols and externally modelled. Second the knowledge representation system must fulfil ontological commitments, where the scope of existence, the domain, is addressed as the 'ontological commitment of the designer'. The need for reasoning must be addressed, the 'fragmentary theory of intelligent reasoning' looks at forming a knowledge based on understanding behaviour. All of this knowledge must be represented in an efficient medium for computation, computer encoding and human readable forms all reduce the computational time. Lastly a knowledge representation system must act as a medium of human expression, facilitating the knowledge engineers and domain experts working on representing knowledge.

Levels of representation

To encapsulate the discussions so far, the Semantic Web must apply representation at different levels (Brachman 1979) the first is implementation, where data structures are formed. Next logic displays the symbolic predicates and variables in a meaningful way. Then epistemological study defines the concepts in question through subtypes by applying inheritance between elements. Extended into the world of concepts, the semantic relations are defined clearly, lastly, linguistics must be resolved, understanding the natural language of the user.

The meaning triangle, see Figure iv, demonstrates these levels. The concept of a human is formed, connected to both the object – a person, and the symbol of a human, their name.

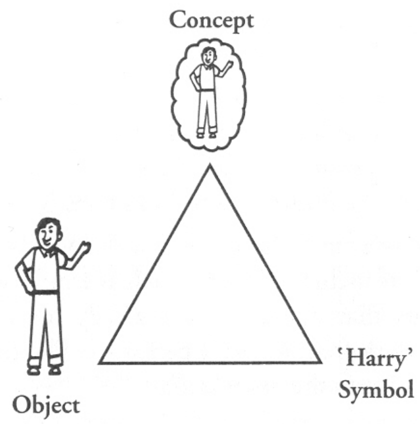

Figure iv - Meaning Triangle

"Knowledge Representation" p193

Meta level symbols

Charles Sanders Peirce was the first logician to recognise that by objectifying each element in a process a better understanding could be achieved (Sowa 1999). Events and elements which participate in a reaction are distinct. Peirce (see Sowa 1999), like others, ignored the tense of language as irrelevant. However earlier, in the fourth century BCE, Diodorus Cronus stated a sentence can be true in the past tense that is not true in the present, thus;

> "Alma had three husbands, Gusta, Walter and Franz"

Thus suggesting a distribution of several, sequential, marriages. Here we witness a distribution, where only one of the possible results is current. Unproved hypothesis and fiction create additional issues, for example, Sherlock Holmes is a fiction. However it is set in a real location and written as if it is fact. Logic and ontology together need to set context, allowing the Semantic Web to adapt to cope with such possibilities.

Consider for example, two seemingly similar sentences with opposing meanings;

> John sang strangely
>
> Strangely John sang

Although identical in lettering and punctuation the meaning is totally different. Robert Moore (1995) identified this as a major issue for knowledge representation. While the former suggests that John was unable to sing as the listener had anticipated, the second suggests John's timing was inappropriate or unusual. However what if the contrary is true. Two statements with different semantics mean the same, thus;

> It is true that Matthew is tall
>
> Matthew is tall

Arguably the language of the first premise is overly verbose, but both are accurate. Any knowledge representation system must be able to make that determination. Similarly the western philosopher Heraclitus (c. 500_{BCE}) wrote, "One cannot step twice into the same river". A river is in an ever changing state of flux. Can the Semantic Web appreciate the complexity of these examples? Likewise can ontologies supply the assistance to do so?

It is indeed possible that the Semantic Web does not need the direct assistance of ontology to resolve this issue. By linking documents efficiently the author defines context. A proposition is only true in context, for example, using RDF notation;

TrueIn('Holmes is a detective', ContextOf('Sherlock Holmes stories'))

TrueIn('Holmes is a supreme Court Judge', ContextOf('US Legal History'))

Sowa (1999) declares these facts are no longer ambiguous, by declaring the context to which they belong. The notation additionally demonstrates the contextual declaration is disconnected, absent of any affect to the data, to which it belongs. Context is merely an association drawn between two entities.

Sowa (1999) cites the tale from Willard Van Orman Quine;

> "Take, for instance, the possible fat man in that doorway, and, again, the possible bald man in that doorway. Are they the same possible man, or two possible men? How do we decide? How many possible men are in that doorway? Are there more possible thin ones than fat ones? How many of them are alike? Or would their being alike make them one?"
>
> (Quine 1953)

Quine is raising the question, on occasion labelled as 'possibilia', there are no known answers here, and it is reasonable to conclude any object must exist, and thus in this case no man exists, bald or otherwise. This is a philosophical question, but no doubt one the Semantic Web will face.

Taxonomy

Reducing the complexity of any knowledge system is to categorise through a hierarchy. Each node reveals world entities, moving toward the root the subject generalises, while moving toward nodes the specialisation increases. This is known as taxonomy. Taxonomies, metadata schemata and ontologies are essential to establishing meaning. When taxonomy is semantically rich it specifies a distinguishing property for each subclass of taxonomy. For example 'manager' and 'employee' can be classes or attributes. Employee is semantically rich, the individual is employed, manager is not semantically useful, all managers are employees but only some employees are managers.

Why classify into taxonomies? Taxonomy allows one critical activity, browsing without a clear destination. The best taxonomy example is the Dewey Decimal System, with knowledge of only the subject it is possible to 'drill down' to the relevant items without prior knowledge.

Web Ontology Language: OWL

The second of the two technologies earlier suggested as vital to the Semantic Web is OWL, a manner of expressing ontologies. OWL is the ontology language for the Web, used to describe, in a defined way, a domain. The W3Cs goal is to permit reasoning of the described domain. Humans apply ontological reasoning to situations they are in, applying context to a discussion is essential to derive meaning. Although OWL documents are ontologies they are also RDF documents, utilising the benefits RDF brings.

To allow computerised reasoning OWL represents the objects humans already know about but computers can learn about, and apply reasoning.

Any reasoning language can achieve that end result, OWL however is intended for the Web, operating on a large scale, spread across platforms and boundaries. Designed by knowledge representation experts to become the formal knowledge representation language for the web, OWL represents an ambitious project for the W3C.

It is worth noting the acronym, OWL. For a technology dedicated to clarity it is surprising the W3C did not use WOL for Web Ontology Language. It is claimed that the developers, while selecting a name, wished to represent the great knowledge, and perhaps wisdom, held within; so played homage to the

Owl character, Wol, in A. A. Milne's Winnie the Pooh, thus reversing the acronym. (www.w3.org/2003/08/owlfaq, 2004)

OWL is an emergent technology, gaining recommendation early 2004, it therefore presents some issues, discussed by Antoniou and van Harmelen (2004). Sharing of work is so essential to the Semantic Web; however importing options are rather basic, limited to entire ontologies, instead of parts. Secondly, when a value is inherited it becomes the default if the level of specialisation is greater, this may not always be useful;

> "No consensus has been reached on the right formalisation for the Nonmonotonic behaviour."
>
> (Antoniou 2004).

OWL, in its processing, makes assumptions. Firstly it takes an Open-World stance to data, that is to say "a statement cannot be assumed true on the basis of a failure to prove it" (ibid). However the Closed-World assumption also has advantages, albeit in a limited capacity, where "a statement is true when its negation cannot be proved" (ibid). This constrains programmers. Secondly OWL assumes individuals with different names are different individuals. Likewise a database makes this logical assumption, however it is not always the case, many ontologies allow for individuals having aliases, Tom Baker, T Baker, Mr Baker are three different names, possibly the same individual.

The use of ontology is not widespread, the complex domain specific data is a hazard to its popularity. Ontological engineers must be domain experts. However, the benefits of ontological aided analysis are numerous. The Semantic Web demonstrates the capacity to store data, make assertions, and provide context. Attention now turns to the underlying knowledge to be represented.

The Knowledge Soup

In the article “Will the Semantic Web Change Education?” (Clark 2004) two notable differences between the printed world and the electronic were discussed. First, locating a literary source requires knowledge of the subject, suitable titles can be accessed via library catalogues or similar. A more effective method however is to follow up the references of a credible source, and in turn the references of the next source. The researcher quickly forms an idea of the subject by means of a variety of interconnected resources. Conversely electronic media may cite hyperlinks to supporting or opposing material in the body text, linking the researcher to a wealth of immediate knowledge.

The same idea of immediate expression through hyperlinks is part of the articles (ibid) second point, authors of journals and indeed books use footnotes and endnotes to add insight, often along a different tangent to the main text. They continue,

> “Thus, rather than creating concise or verbose linkage markers, scholarly discourse on the hypertextual Web is able to interleave and interweave such linkages within the main text itself. We can -- arbitrarily or elegantly -- make any text, within a scholarly hypertext, link to any other Web resource (or even to named parts, or fragments, of other Web resources). That difference in technology, which is admittedly quite subtle, calls forth and makes possible a change in the way that scholarly discursive practices are created and enacted.”
>
> (Clark 2004)

Altering the style of referencing sources affects the large amount of material re-published, or indeed first published, on the Internet.

Is knowledge power?

It is folly to say 'Knowledge is Power' as knowledge covers a wide area and numerous disciplines. Is knowledge just the recitation of fact, of data or of information. Is knowledge signified by better comprehension or understanding or via displaying an awareness of ones surrounds not held by the majority. Knowledge of an industry is different, it is focused and reasoned. This knowledge is correct or incorrect, meaningful or meaningless. Where one has a domain-specific knowledge base they do indeed have authority, as rule based knowledge, is power. The Semantic Web is distributed power. That is knowledge spread through the routers and connections of the Internet "knowledge becomes as distributed, dynamic and ubiquitous as the power flowing into the lamp by which you are reading these words" (Hendler 2000).

Managing the Knowledge Soup

Knowledge management has always been about acquiring, accessing and maintaining knowledge within an organisation (Antoniou 2004). However that knowledge has a weak structure. Searches are keyword based, agents extracting knowledge are relatively basic, maintenance is needed to resolve inconsistencies and outdated knowledge.

In the Semantic Web knowledge is organised into conceptual spaces, according to meaning. Aiming for automated inconsistency checks, query answering replaces traditional keyword searches while returning data over several documents and, lastly, enhanced security, even part of a document could be restricted from view.

> "It'll be 10 years before we have all the health information about drugs out there. I would expect then to be able to swipe a bar code on a package and get back any allergy information. Or in the supermarket, I should be able to swipe a bar code and find out from the Semantic Web whether it contains peanuts. It's so obviously valuable to have that information, so obviously in the public good, and it can obviously be connected with other information."
>
> (Berners-Lee 2004)

Berners-Lee describes a wealth of knowledge controlled centrally or voluntarily, it however has no focus. Scientists describe the confusion of knowledge as a knowledge soup. For example we are everyday provided with chunks of knowledge (Sowa 1999). Generalisations often omit the obvious, 'all birds fly' may seem true, but what about sleeping birds and Penguins? Statements omit the abnormal, 'I drive to work', does not consider a discharged battery. Similarly incomplete statements can neglect important detail: 'use the oil well', is the well capped or dry? Conflict may also appear, 'Richard Nixon is a Quaker and Republican' therefore is Nixon a pacifist or not? Equally unanticipated statements, is 'hair a body part', are extensions separate? (ibid).

These, according to John Sowa (1999), are not linguistic problems, each statement is valid, but represents inherent complexity in the system. Alfred North Whitehead said we must "discriminate exactly what we know already" to provide a structure for reasoning, which, Charles Sanders Peirce, labels a "solid foundation for great and weighty thought" (see Sowa 1999).

> "Knowledge representation is the application of logic and ontology to the task of constructing compatible models for some domain"
>
> (Sowa 1999)

In the words of Sowa, to represent knowledge is to use logic and ontology to construct models which symbolise the domain in question. If RDF claims to be able to define any object, then take for example the following diagram (Figure v). Assuming a user inputs the description, it is complex to interpret the three angles as representing the same object, can the Semantic Web achieve this?

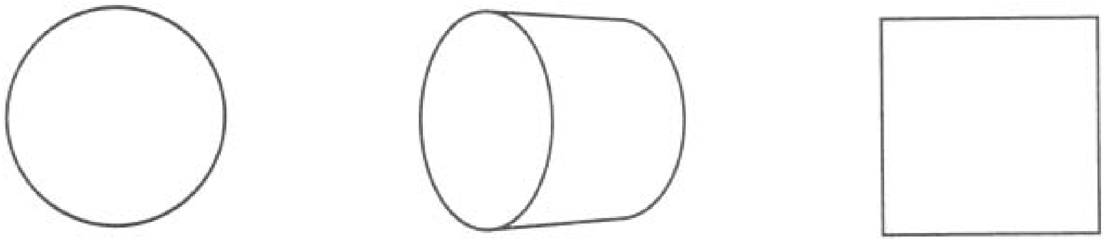

Figure v - Three views of an object, appearing both round and square

"Knowledge Representation" p. 335

Nonmonotonic reasoning

Taking a hypothetical axiom, of a car journey Sowa (1999) suggests that providing new data, such as snow on the road, does not block the original theory, instead it becomes more provable, that the journey was not possible or harder than anticipated. New data will always increase the accuracy of a reasoned conclusion.

Hidden Knowledge

The knowledge soup shows information with hidden meaning. The Semantic Web however specialises in uncovering relationships, and hence knowledge. Organisation decisions could often be better informed, but the relationships are hidden in databases. In the months after the World Trade Centre attacks of 2001 the Federal Bureau of Investigation (FBI) director Robert Mueller called attention to the need of interoperability (Daconta 2003). Had the CIA and FBI systems worked together US Flight schools could have been easily cross-referenced with known terrorists. The American Global Intelligence Working Group (GIWG) was given the following remit:

> "To develop, build, and support the creation of the National Criminal Intelligence Sharing Plan, which provides law enforcement agencies with the ability to gather, analyze, protect, and share information and intelligence to identify, investigate, prevent, deter, and defeat criminal and terrorist activities, both domestically and internationally, as well as protect the security of our homeland and preserve the rights and freedoms of all Americans."
>
> (GIWG 2004)

The Semantic Web is clearly an appropriate answer, it can uncover implied and hidden relationships with ease, and thrives on disconnected, heterogeneous data (Daconta 2003). The Semantic Web changes an organisations outlook, their ethos of work must now be knowledge discovery, (Daconta 2003). Knowledge management is the dissemination of corporate knowledge through capture and cataloguing.

Realising the Web

Web Services

The process of building Web Services is intrinsically linked to the Semantic Web, both have the same goal of data harmonisation. However Alexander Nakhimovsky and Tom Myers (see Geroimenko 2003) point out that their orientation is different

> "… the Semantic Web thinks of the Web as repository of knowledge, while Web Services think of it as a marketplace. … While the Semantic Web's metaphor for the Web is an encyclopaedia, the Web services' is a telephone book, where you or your agent can call a number to obtain a service."
>
> (Nakhimovsky 2003 see Geroimenko 2003 p. 135).

Critically the Semantic Web remains an idealistic tool for knowledge, Nakhimovsky and Myers suggest Web Services are "closer to the metal and wire of the Web". Web Services are discovered, described and accessed directly via the SOAP protocol, "a lightweight protocol for exchange of information in a decentralised, distributed environment" (Mitra 2003).

Trust

The Semantic Web forms a matrix of associations by verifying the author, and hence trust. Non-repudiation, where legal proof of authorship is achievable, is needed. The Semantic Web relies upon XML signatures and XML encryption Documents, annotations or any other resources are signed by their author, and optionally encrypted.

Intended to use existing HTTP security, such as Secure Socket Layer (SSL) the need for purposeful XML security was realised with SAML. The Security

Assertion Markup Language is an emerging standard, protecting e-commerce over XML.

Nevertheless knowing the author does not result in absolute trust. In the article “Taking a Stand on the Semantic Web”, by Catherine Marshall (2003), she advocates a negative view of Semantic Web security. Describing a great potential for ‘poisoning’, through the fraudulent entering of metadata, Marshall suggests this issue alone is a stumbling block. No authority checks inputs, Marshall concludes “The Semantic Web: Unworkable, unnecessary, and unsafe at any speed”, a damning indictment. Even the most persuasive of proponents cannot help but agree the potential for misuse is vast.

Resources

Beginning to appear for academia are Semantic Web resources, IsaViz supported by the W3C creates RDF graphs. While work by the Israle Institue of Technology into automatic ontology building has resulted in ‘OntoBuilder’, an agent for building ontologies. In their paper, “The Use of Machine-Generated Ontologies in Dynamic Information Seeking”, the authors summaries their efforts:

> “... such a process would allow the creation of flexible tools to manage metadata, either as an aid to a designer or as an independent system (“smart agent”) for time critical missions. Also, an automatic reconciliation process would allow data management systems to use data even though it is originated for different ontologies”
>
> (Modica 2001)

RDF libraries and parsers for programming languages are appearing, ‘Drive’ (Singh 2003) is one such application, aimed at the Microsoft language C# for the .Net platform. Additionally the Redland RDF Application Framework

(Beckett 2005) is a University of Bristol project to provide RDF tools, with bindings for Java, Python and other popular languages. Such tools greatly simplify the process of creating meaningful data, ever-increasing the likelihood of realising the Semantic Web.

The Semantic Web Revolution

Changes in approach to knowledge

Paul Ford's (2002) futuristic article "August 2009: How Google beat Amazon and eBay to the Semantic Web" explores the potential future for this technology. Including what Ford calls the 'Google Marketplace Search' an interactive website passing simple RDF statements. Thus, entering 'sell:martin guitar' returns suitable buyers, and likewise for 'buy:'. Ford's ideas fall short of prophecy but his logic is sound. Google or any other inventor is capable of 'crawling' the Web for information and asserting links, RDF is the perfect storage mechanism. Ford continues, perhaps more far fetched predictions include future systems using the Semantic Web approach to verify age. The browser knows its users identity, by accessing the Social Security records or the local equivalent, age can be proved. The picture painted by Ford is of cultural change, the above is realised with a low margin for error, and thus its reliance for daily usage is feasible.

How fanciful are the notions put forward by Ford? Already sites like Alexander Loeser's (2002) "RooDolF" can interpret Google search results and translate into RDF. Using this technique another application is able to utilise the data more efficiently.

Richness and semantics can only proceed so far. It has previously been argued that untrue information to the Semantic Web is as meaningful as true, however what happens if data expressed as accurate has no basis in truth. In 2003 David Batty of the Guardian newspaper highlighted the governments Child Database disarray, one key issue was data:

> "A comparison of the data held on people who receive support from Hammersmith and Fulham social services, in west London, showed 48% of the department's records, covering nearly 55,000 people, did not match the local NHS database. About 3% of the records, 1,445, disagreed on whether a person was alive or dead, and 1% on gender."
>
> (Batty 2003)

This is a horrendous situation to be in, Hammersmith and Fulham are likely to be a good representation of the country, their database presents authoritative truths, but if such databases give conflicting data how can the Semantic Web adapt?

Military

One significant proponent of the Semantic Web is the military, due to their high use of information interchange. Operation Iraqi Freedom in 2003 saw the invasion and occupation of Iraq, Jim Hendler at the 2004 Semantic Technologies for e-Government Conference said the result of poor systems integration caused delays "The beginning of the Iraqi operation was postponed for weeks because information systems couldn't be made interoperable in the time required", this, Hendler suggested, could be solved by the Semantic Web.

The military involvement continues past recommendations. In early 2003 American company McDonald Bradley won a multimillion dollar contract to build a military Intranet. "The project is one of several meant to achieve the network-centric operations that the Defence Department is implementing" (Grimes 2004). The Intranet portal aims to be "an information tapestry" (John Stenbit see Grimes 2004).

> “McDonald Bradley will explore use of metadata tagging, data-content mark-up, taxonomies, ontologies and other extensible mark-up language and semantic Web techniques to help systems index and provide information.”
>
> (Jackson 2003)

Semandex is an example of the Semantic Web in military action, built for rapid information interchange Semandex was adopted by the American national defence and acted as a vital tool for Marines in combat (Gomezjurado 2004). Locating previously complex information like geospatial data could be reduced from six days to just one, it has also been implemented in intelligence gathering systems.

> “The Marines now have a system that can quickly find information they need and share it with others for analysis, instead of being swamped with thousands of search-engine hits”
>
> (Gomezjurado 2004).

The Semantic Web has been a huge advantage to the military, information can be gathered from a number of sources and cross-referenced in real time.

Changeover – now or never?

Explicit metadata is required, keyword matching can be confused, for example a standard web page can display the names of staff at an organisation. A keyword match will spot the names of each member, however it would be difficult without knowledge of the page design. With XML and metadata the staff, roles and other information is clearly defined, regardless of the organisation. Mark-up natural language analysis will write most mark-up, it will not be hand written like early HTML was.

The process of handling data in the Semantic Web will need to evolve, discovery and production must be changed. Once a document is written it should be checked for validity, sources verified and finally an XML schema applied. An XML signature can be used to assert authorship. RDF annotations give the data meaning, which are also signed. Next the document and annotations are mapped to an ontology, and ultimately published to a Web Service, and registered with the Web Service database. This makes the original meaningless document meaningful, trusted and easy to locate.

If this is performed at the time of document creation then later analysis is far more straightforward. The Semantic Web is a revolution, evolving the organisation toward a Knowledge-Centric outlook, the toughest part is changing the employers mindset.

Conclusion

The challenge facing the Semantic Web is not scientific, it is adoption of its techniques. The original web saw a similar growth in potential align with momentum for use.

The Semantic Web has been a phenomenal achievement, thus far, one of the biggest outstanding issues is how can the great number of ontologies be structured to realise the true Semantic Web vision. Few suggestions have been offered and like many Internet technologies may have to evolve. Commercial backing exists, Richard Hayes-Roth of Hewlett Packard stated “we expect the Semantic Web to be as big a revolution as the original web itself”.

Although the availability of tools for annotations and RDF mark-up currently revolves around the academic world commercial applications will follow. Early successes will be seen in Intranets, where rules can be imposed. As with the Webs humble beginnings e-science may be the most loyal follower (Antoniou 2004). To science the Semantic Webs advantages are obvious, to e-commerce and the public it maybe more of a struggle. However Business-to-Business maybe seen as an additional driver, the economic potential of the Semantic Web is staggering.

Authors and technologies alike cite their work as being part of the “Semantic Web” without a clear definition it is difficult to disprove their claim. With all the hype and technological marketing surrounding the Semantic Web there is a high chance of confusion,

> "There are three reasons, or villains, that have kept the Semantic Web from succeeding. These include the fact that most documents are not tagged and are paper-based, and there are no tools to do the mark-ups; because there is a lack of understanding of technologies such as RDF; and due to corporate confusion."
>
> (William Rush 2002 see Krill 2002)

Rush continues, "The word ontology alone 'scares the heck out of people'", the issue with the Semantic Web is purely acceptance, the technology is there. At its core the Semantic Web is simple, to present data in a manner which is meaningful not meaningless. Additionally proponents such as Andy Oram (2002) are concerned many businesses will perform a cost/benefit analysis upon the Semantic Web and not add complex associations.

The Semantic Webs vision is to 'make the world make sense', by better utilising metadata, indeed metadata is the cornerstone of indicating meaning within a document. Semantic Web software agents, using metadata, will grow and develop, in fact only their appeal can improve their quality, as more are written their effectiveness will become apparent, leading to better programming and agents.

Armed with metadata, ontologies are key to understanding domains. Benefits include improved search results, following resource links, knowledge of alternative search criteria and small, defined, exact results, realising the needs of the web as identified by Brian Clegg (2001). Intelligence in the software implementation would allow the query's granularity (the level of detail returned) to fluctuate, automatically controlling the results.

OWL is a keystone of the Semantic Web, before the Semantic Web relied on older, less interconnected languages, such as DAML+OIL which left the Semantic Web isolated, an inaccessible utopia.

A complement to OWL is the more flexible Simple Knowledge Organisation System (SKOS), less restrictive than OWL it is targeted at organisations which cannot map directly to ontology.

Rooting itself squarely in knowledge representation OWL has potential, amalgamated with RDF they are a powerful combination.

> "In the Semantic Web vision, there is the expectation that hordes of developers, web masters, page authors, and even casual users will be creating and consuming Semantic Web data. Everyone will be hyperkrep hacker, able to casually create a mix of hypermedia and knowledge representation that fits in to the global hyperkrep system."
>
> (Clark 2004)

The Semantic Web relies on cooperation and understanding, knowledge representation engineers can be employed to study the domain in great detail, however using the triples system even a simple web master can make associations and hopefully comprehend the reason this needs to be done. Less 'hacker-like' tools than IsaViz will be integrated with office applications and operating systems.

Whenever a breakthrough is made, whether in science, technology or literature the end result must not just meet the level of the old but surpass it, in a novel manner. Tim Berners-Lee believes the Semantic Web is such a technology. "It will foster global collaborations among people with diverse cultural perspectives, so we have a better chance of finding the right solutions to the really big issues - like the environment and climate warming." Thus the impact of the Semantic Web will be greater than his original creation.

The Semantic Web is viable, it will be realised by the developers, the programmers, the web masters of the future. Users will experience a great push forward in information retrieval, but few will experience the Semantic Web direct. Semantically richer web pages result in software experiencing the Internet as a casual user would today, in turn knowledge management will reap the rewards sown by the Semantic Web. If "Knowledge management is a socio-technical discipline" (Davies 2002) it will require equal effort from society and technology, the latter is here, in the form of the Semantic Web, and is ready to meet society head on.

Acknowledgements

Thanks are owed to the librarians at the University of Brighton Aldrich Library and the web site team for providing access to a rich variety of sources. Thanks are also due to David Coutinho, Senior Lecturer, for planting a vision of interconnectivity devoid of presentation and comprising of pure meaning.

Bibliography

ADAMS, D., 1988. Dirk Gently's Holistic Detective Agency. Pan Macmillan

ALLEN, J., 2001. *Making a Semantic Web* [online]. Netcrucible. Available at http://www.netcrucible.com/semantic.html [Accessed 6 November 2004].

ARISTOTLE, 350BCE . *On Sophistical Refutations* [online]. Trans. PICKARD-CAMBRIDGE, W. A., 2004. Adelaide Library: University of Adelaide. Available at http://etext.library.adelaide.edu.au/a/aristotle/sophistical/ [Accessed 12 December 2004].

Anon, 2002. *The Next Weave of the Web?* [online]. Data Conversion Laboratory. Available at http://www.dclab.com/semantic_web.asp [Accessed 6 November 2004].

Anon, 2003. *Semantic Web - The Semantic Web Community Portal* [online]. Available at http://www.semanticweb.org/ [Accessed 1 September 2004].

Anon, 2004. *The Dublin Core Metadata Initiative* [online]. Available at http://dublincore.org/ [Accessed 12 December 2004].

Anon, 2004. *Intel sees big changes to the net* [online]. Broadcasting House: British Broadcasting Corporation. Available at http://news.bbc.co.uk/1/low/technology/3643902.stm [Accessed 3 October 2004].

Anon, 2004. *PlanetLab Consortium* [online]. PlanetLab. Available at http://www.planet-lab.org/consortium/ [Accessed 1 September 2004].

Anon, 2004. *Semandex Networks* [online]. Available at http://www.semandex.com/ [Accessed 12 December 2004].

Anon, 2005. *About Security Assertion Markup Language* [online]. XML Trust Centre. Available at http://www.xmltrustcenter.org/saml/ [Accessed 3 February 2005].

Anon, 2005. *SKOS* [online]. W3C. Available at http://www.w3.org/2004/02/skos/ [Accessed 2 February 2005].

ANDERSON, S., HODGES, J., MALER, E., 2005. *Defining and maintaining a standard, XML-based framework ...* [online]. OASIS Security Services. Available at http://www.oasis-open.org/committees/tc˙home.php?wg˙abbrev=security [Accessed 3 February 2005].

ANTONIOU, G., van HARMELEN, F., 2004. *A Semantic Web Primer*. Cambridge: MIT Press

BATTY, D., 2003. *Child protection scheme in disarray*. The Guardian, 24 October.

BECKETT, D., 2005. *Redland RDF Application Framework* [online]. University of Bristol. Available at http://librdf.org/ [Accessed 3 February 2005].

BERNERS-LEE, T., 1989. *Information Management: A Proposal* [online]. CERN. Available at http://www.w3.org/History/1989/proposal.html [Accessed 6 November 2004].

BERNERS-LEE, T., 1998. *Semantic Web Road map* [online]. W3C. Available at http://w3.org/DesignIssues/Semantic.html [Accessed 21 September 2004].

BERNERS-LEE, T., 1999. *Weaving the Web*. Harper: San Francisco

BERNERS-LEE, T., CONNOLLY, D., SWICK, R., 1999. *Web Architecture: Describing and Exchanging Data* [online]. W3C. Available at http://w3.org/1999/04/WebData [Accessed 6 November 2004].

BERNEWS-LEE, T., MILLER, E., 2002. *The Semantic Web lifts off* [online]. Ercim News. Available at http://www.ercim.org/publication/Ercim_News/enw51/berners-lee.html [Accessed 6 November 2004].

BRICKLEY, D., 1999. *Semantic Web History: Nodes and Arcs 1989-1999* [online]. W3C. Available at http://w3.org/1999/11/11-WWWProposal/ [Accessed 6 November 2004].

CLARK, K., PARSIA, B., HENDLER, J. 2004. *Will the Semantic Web Change Education?* [online]. Maryland Information and Network Dynamics Laboratory: University of Maryland. Available at http://www-jime.open.ac.uk/2004/3/clark-2004-3-disc-paper.html [Accessed 3 October 2004].

CLEGG, B., PAGE, K., 2001. *The Professional's Guide to Mining the Internet*. 2nd ed. Routledge: Falmer.

COHEN, F., 2003. *Debunking SAML myths and misunderstandings* [online]. IBM Developer Works. Available at http://www-106.ibm.com/developerworks/xml/library

CONNOLLY, T., BEGG., C, 2002. *Database Systems: A practical approach to Design, Implementation and Management*. 3rd ed. Addison Wesley.

DACONTA, M., OBRST, L., SMITH, K., 2003. *The Semantic Web: A Guide to the Future of XML, Web Services and Knowledge Management*. Indiana: John Wiley & Sons.

DANBRI, 2003. *RDFWeb and Friend of a Friend (FOAF)* [online]. RDFweb. Available at http://rdfweb.org/mt/foaflog/archives/000051.html [Accessed 3 February 2005].

DAVIES, J., FENSEL, D., and van HARMELEN, F., ed., 2002. *Towards the Semantic Web*. Chichester: Wiley.

DEAN, M., SCHREIBER, G., ed., 2004. *OWL Web Ontology Language Reference* [online]. W3C. Available at http://w3.org/TR/owl-ref/ [Accessed 6 November 2004].

DIENG-KUNTZ, R., 2002. *Corporate Semantic Web* [online]. ERCIM. Available at http://www.ercim.org/publication/Ercim_News/enw51/dieng.html [Accessed 12 December 2004].

DOYLE, A., 2004. The Adventure of the Beryl Coronet [online]. Classic Literature Library. Available at http://sherlock-holmes.classic-literature.co.uk/the-adventure-of-the-beryl-coronet/ [Accessed 20 December 2004].

DODDS, L., 2004. *An Introduction to FOAF* [online]. XML.com. Available at http://www.xml.com/pub/a/2004/02/04/foaf.html [Accessed 3 February 2005].

DUMBILL, E., 2002. *Finding friends with XML and RDF* [online]. XMLhack.com. Available at http://www-106.ibm.com/developerworks/xml/library/x-foaf.html [Accessed 3 February 2005].

FENSEL, D., et al., ed., 2003. *Spinning the Semantic Web*. Cambridge: Massachusetts Institute of Technology.

FORD, P., 2002. August 2009: *How Google beat Amazon and Ebay to the Semantic Web* [online]. FTrain. Available at http://www.ftrain.com/google_takes_all.html [Accessed 25 September 2004].

FORD, P., 2004. *Uncle Sam's Semantic Web* [online]. O'Reilly Media, Inc. Available at http://www.xml.com/pub/a/2004/09/15/egov.html [Accessed 3 October 2004].

GEROIMENKO, V., CHEN, C., ed., 2003. *Visualizing the semantic Web: XML-based Internet and information visualization*. London: Springer.

GOMEZJURADO, J., 2004. *Marines in Iraq Find Battlespace Information Faster with Content-Based Network from Semandex* [online]. National Technology Alliance. Available at http://www.nta.org/docs/SemanticNetwork.doc [Accessed 20 November 2004].

GRANT, J., BECKETT, D., 2004, 2004. *RDF Test Cases* [online]. W3C. Available at http://w3.org/TR/rdf-testcases/ [Accessed 6 November 2004].

GRIMES, B., 2004. *McDonald Bradley to build net-centric intranet [online]. Washington Technology*. Available at http://www.mcdonaldbradley.com/news/articles/2004/washtechmbinetcentric.pdf [Accessed 6 November 2004].

HAMMING, R., 1962. *Numerical Methods for Scientists and Engineers*. New York: McGraw-Hill.

HANSON, E., 2001. *XML and the Semantic Web from 10,000 Feet* [online]. Aquameta. Available at http://www.aquameta.com/~eric/writings/semantic_web/ [Accessed 1 November 2004].

HEFLIN, J., HENDLER, J. 2001. *A Portrait of the Semantic Web in Action* [online]. IEEE Intelligent Systems. Available at http://www.cse.lehigh.edu/~heflin/pubs/ieee2000.pdf [Accessed 6 November 2004].

HEFLIN, J., 2004. *OWL Web Ontology Language Use Cases and Requirements* [online]. W3C. Available at http://w3.org/TR/webont-req/ [Accessed 6 November 2004].

HEMENWAY, K., 2002. *The Semantic Web: 1-2-3* [online]. Available at http://www.disobey.com/detergent/2002/sw123/ [Accessed 6 November 2004].

HENDLER, J., 2000. *Agent Mark-up Language*. Defence Advanced Research Projects Agency.

HENDLER, J., BERNERS-LEE, T., MILLER, E., 2002. *Integrating Applications on the Semantic Web* [online]. W3C. Available at http://w3.org/2002/07/swint [Accessed 6 November 2004].

HERACLITUS, c. 500BCE. *Fragment 41 from KAHM, C, 1979. The Art and Thought of Heraclitus*. Cambridge University Press: Cambridge, UK

HERMAN, I., 2000. *About the World Wide Web Consortium (W3C)* [online]. W3C. Available at http://www.w3.org/Consortium/#mission [Accessed 12 December 2004].

JACKSON, J., 2003. *McDonald wins DoD semantic Web work [online]. Washington Technology*. Available at http://www.washingtontechnology.com /news/1_1/daily˙news/20307-1.html [Accessed 5 November 2004].

JAKES, D., PENNINGTON, M., KNODLE, H., 2001. *Using the Internet to promote Inquiry-based learning* [online]. MyProjectPages. Available at http://www.myprojectpages.com/support/epaper/ibr.htm [Accessed 12 December 2004].

JOYCE, E., 2001. *Semantic Web Supporters/Detractors Sound Off* [online]. AtNewYork. Available at http://www.atnewyork.com/news/article.php/780491 [Accessed 5 November 2004].

KANT, I., 1787. *Critique of Pure Reason*. Trans. GUYER, P., WOOD, A. 1999. Cambridge University Press: Cambridge UK.

KANT, I., 1800. *Lectures on Logic*. Trans. YOUNG, M., GUYER, P., WOOD, A., 1992. Cambridge University Press: Cambridge UK.

KOEPSELL, D., 2000. *Centre for Commercial Ontology* [online]. Buffalo University. Available at http://www.acsu.buffalo.edu/˜koepsell/center.htm [Accessed 2 November 2004].

KOIVUNEN, M-R., MILLER, E., 2004. *W3C Semantic Web Activity* [online]. W3C. Available at http://w3.org/2001/12/semweb-fin/w3csw [Accessed 6 November 2004].

KRILL, P., 2002. *Exec touts Semantic Web* [online]. InfoWorld Media Group. Available at http://www.infoworld.com/articles/hn/xml /02/06/07/020607hnruh.html [Accessed 6 December 2004].

LOESER, A., 2002. *RooDolF* [online]. Available at http://nutria.cs.tu-berlin.de/roodolf/ [Accessed 12 December 2004].

MALLETT, S., 2004. *The Future of the Semantic Web is Here Today and is Evenly Distributed* [online]. O'Reilly Media. Available at http://www.oreillynet.com/pub/wlg/5607 [Accessed 6 November 2004].

MANOLA, F., MILLER, E., ed., 2004. *RDF Primer* [online]. W3C. Available at http://w3.org/TR/rdf-primer/ [Accessed 6 November 2004].

MARSHALL, C., 2003. *Taking a Stand on the Semantic Web* [online]. Texas A&M University. Available at http://www.csdl.tamu.edu/˜marshall/mc-semantic-web.html [Accessed 6 October 2004].

MATTHEWS, B., WILSON, M., BRICKLEY, D., 2002. *Semantic Web Advanced Development in Europe* [online]. Ercim News. Available at http://www.ercim.org/publication/Ercim˙News/enw51/matthews.html [Accessed 6 November 2004].

McDONALD BRADLEY, 2004. *The Semantic Web Foundations of Net-Centric Warfare* [online]. McDonald Bradley. Available at http://www.mcdonaldbradley.com/comps/white%20papers/The%20Semantic%20Web%20Foundations%20of%20Net-Centric%20Warfare.pdf [Accessed 6 November 2004].

MILES, A., 2005. *SKOS Core* [online]. Schema Web. Available at http://www.schemaweb.info /schema/SchemaDetails.aspx?id=101 [Accessed 3 February 2005].

MILLER, G., 1995. *Wordnet* [online]. Princeton University. Available at http://www.cogsci.princeton.edu/˜wn/ [Accessed 6 November 2004].

MITRA, N., 2003. *SOAP Version 1.2 Part 0: Primer* [online]. W3C. Available at http://www.w3.org/TR/2003/REC-soap12-part0-20030624/ [Accessed 12 December 2004].

MODICA, G., GAL, A., JAMIL, H., 2001. *The Use of Machine-Generated Ontologies in Dynamic Information Seeking* [online]. Israel Institute of Technology. Available at http://iew3.technion.ac.il/OntoBuilder/ [Accessed 12 December 2004].

MOORE, R., 1995. *Logic and Representation*. Center for the Study of Language and Information. Stanford.

MORVILLE, P., ROSENFELD, L., 2002. *Information Architecture for the World Wide Web*, 2nd ed. O'Reilly.

ORAM. A,. 2002. *The Semantic Web: It's Whom You Know* [online]. O'Reilly Media. Available at http://www.oreillynet.com/pub /a/network/2002/04/19/platform.html [Accessed 6 November 2004].

QUINE, W., 1953. *From a Logical Point of View*. Harper & Row: New York.

RAY, E., 2001. *Learning XML*. O'Reilly

REINHARDT, A., 2004. *The Web's Father Expects a Grandchild* [online]. Business Week. Available at http://www.businessweek.com/bwdaily/dnflash/oct2004/nf20041022_6972_db083.htm [Accessed 2 January 2005]

SHIRKY, C., 2003. *The Semantic Web, Syllogism, and Worldview* [online]. Shirky. Available at http://www.shirky.com/writings/semantic˙syllogism.html [Accessed 4 October 2004]

SINGH, R., 2003. *Drive* [online]. DriveRDF.org. Available at http://www.driverdf.org/ [Accessed 3 February 2005].

SOWA, J. 1999. *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Crawfordsville: Brooks Cole.

SWARTZ, A., 2002. *The Semantic Web (for Web Developers)* [online]. LogicError.com. Available at http://logicerror.com/semanticWeb-webdev [Accessed 12 December 2004].

van DIJCK, P., 22 November 2004. *Themes and metaphors in the semantic web discussion* [online]. Available at http://poorbuthappy.com/ease/semantic/ [Accessed 6 October 2004].

WILLIAMSON, B., MILLER, L., 2003. *The semantic web* [online]. Guardian Unlimited: E-learning. Available at http://education.guardian.co.uk /elearning/story/0,,981948,00.html [Accessed 5 November 2004].