

## **1 Introduction**

Despite its relative youth, the World Wide Web has already revolutionized citizenship. However, in its current implementation, it falls short of fulfilling some of its users' needs and expectations. Recent efforts improving the Web's capabilities towards what is referred to as the 'Semantic Web' have demonstrated potential in overcoming these limitations and expanding the accessibility of structured, Web-based knowledge to a broader audience. The Semantic Web maintains the best characteristics of today's Web, such as the creation and adoption of open standards, more widely accessible information, and the encouragement of free dissemination and sharing of knowledge. In addition, it enables the use of intelligent agents to work autonomously and make trustworthy decisions on the user's behalf. We discuss how this transition is taking place, where it is headed, and the sustainable life changes that will be enabled once the Semantic Web vision is implemented.

In this paper, we provide a brief introduction to the standards, protocols, and languages that are an essential part of the Semantic Web effort. More importantly, we focus on the benefits that the widespread adoption of Semantic Web technologies may bring to humanity. These benefits range from timely access to scientific and technological knowledge to life-saving decision-making strategies.

This paper is structured as follows: Section 2 provides a brief historical account of the evolution of knowledge dissemination before the World Wide Web had been conceived. Section 3 examines selected aspects of today's Web, from its vision to the technologies that enable its implementation to some of the reasons behind its widespread use. Section 4 describes the Semantic Web effort through a futuristic scenario, which is then used to discuss the Semantic Web's vision, technologies, and impact on human life. Finally, Section 5 discusses the merits of this new approach and some of its dimensions and possible shortcomings.

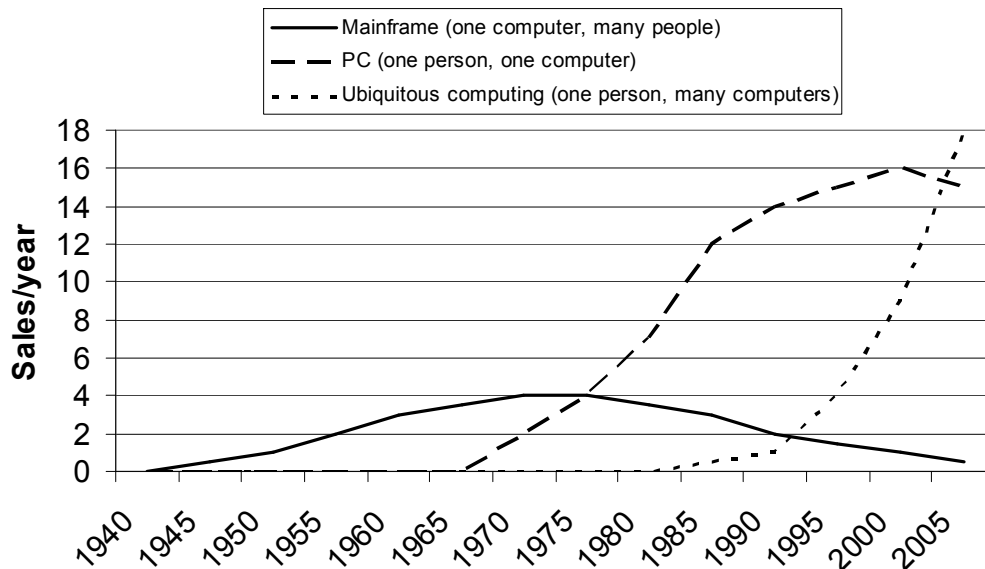
## **2 Acquisition and sharing of knowledge before the World Wide Web**

Prior to the spread of digital media and the popularization of personal computers accessing information was a far different process than it is today. Through ancient times and the Middle Ages the word-of-mouth was a slow, inaccurate way of transferring information between parties. A person had no choice but to trust their source. At best, hand-written documentation could be passed, increasing accuracy, but hardly changing efficiency.

Johann Gutenberg's invention of the printing press had far-reaching implications for the spread of information. The printing and distribution of larger and larger volumes became feasible. Besides helping to spread literacy the invention of the printing press has made books available to nearly everyone. Even today, books are the primary source of information for many people.

The late nineteenth and twentieth centuries have introduced innovations at a faster rate than ever before. The telephone promoted personal communication, while radio, cinema and television allowed information to be pushed to increasingly wider audiences. However, the retrieval of information, until the adoption of personal computers, was still a limited and largely local process that relied on local libraries, bookstores, and broadcast towers. Compared to today there was a narrow range of sources. Consumer interaction was limited – the end user had no control over this content, and the content was created by relatively few people. Then, querying was limited to an index in a book or the selection of a TV channel,

without a fraction of the flexibility we find today when using a Web-based library catalog or programming a digital video recorder (DVR) to record our favorite TV shows.



**Figure 1:** The continuing growth of ubiquitous computing (adapted from [21])

In the latter part of the twentieth century innovations in digital electronics made computers smaller (the first computers took up entire rooms), easier to use (a dedicated, highly-trained operator was no longer needed), and more affordable (cost continues to decrease dramatically). Computers, while essentially capable of performing simple operations, can perform them at amazing speeds. Modern computer science and computer engineering has enhanced the capabilities of these systems. The effects of digitized information have only begun to be explored, yet still are significant enough and hold the potential of and compare to the invention of the alphabet, language, and the printing press. Computers proliferate in our everyday lives and are embedded in a truly diverse variety of devices to the point where they are ubiquitous. Computers have evolved from mainframes (one computer, many users) to personal computers (one computer, one user), to an array of devices (one user, many computers) [21] (Figure 1). Today even devices such as coffee makers, watches, toasters and refrigerators often include microprocessors and can process information collected from local as well as remote sources. As this trend continues, the need to effectively organize information becomes more urgent. Providing access to structured knowledge is not only one of the goals of today's Semantic Web research, it has been anticipated in the past as well. Vannevar Bush's classical 1945 article [5] described a device, the *memex*, which would store the world's knowledge and would inspire the concept of hypertext that is fundamental to the Internet today.

The Internet has provided another breakthrough in recent decades. Information can now be shared globally, quickly, and inexpensively. Even before the World Wide Web, newsgroups, e-mail, FTP (File Transfer Protocol), and Telnet (a simple, bidirectional communications facility commonly used in command line sessions) emerged as viable communication media whose popularity has not been challenged to this date.

### 3 The World Wide Web

The World Wide Web – also referred to as *Web* – is among the best known components of the Internet. Hyperlinks are integral to the functionality of the Web because they provide a simple and effective way of connecting relevant areas among two or more documents. Moreover, creating and publishing contents on the Web has become increasingly easy, from design of customized pages to the use of HTML templates to the popularization of Web-based logs, known as *blogs*. In this Section we briefly discuss the Web's original vision, its technologies, impact, selected factors behind its huge success and popularity, and some of its limitations.

#### 3.1 Background

The original proposal of the Web, by Tim Berners-Lee [2] (Figure 2) suggested a document management framework in which hyperlinks would provide the connections between a document, its author and affiliation, related documents, related research groups and so on. Berners-Lee's vision called for descriptive hyperlinks that not only link two or more objects but qualify which type of relationship they have with each other. This functionality has not been implemented in today's Web and is part of the concepts that are only now being realized in the Semantic Web, as discussed in Section 4.2.

#### 3.2 Technologies

The popularity of the World Wide Web has been aided by the accessibilities of its technologies for content creation. Hyper Text Markup Language (HTML) is a markup language – derived from the syntax of SGML – that has been designed for the Web. The syntax of HTML is one of the keys to its ease-of-use (and, hence, its widespread adoption). Historically, HTML has accommodated incorrect syntax. Web browsers continue to render incorrect markup. This is a boon for novice users who may lack the familiarity necessary to write correct syntax.

HTML only provides facilities for the layout and formatting of items on a page. To add client-side interactivity JavaScript is typically used. JavaScript enables the content creator to access and manipulate the Document Object Model (DOM). The DOM is an object-oriented representation of an HTML page. A script written in JavaScript can change the appearance of a document (size, colors, etc.) and can also perform any calculations necessary. JavaScript's syntax is largely similar (but slightly different) to that of the many of the most popular high-level programming languages (e.g., C, C++, Java, PHP). Similarly to HTML, the cost (free) and availability of JavaScript (many contemporary browsers support it), in addition to its friendly syntax, have certainly helped its proliferation.

One major shortcoming with HTML pages is that the content is integrated with its formatting. This is intuitive because it simplifies the syntax needed to construct a single page. However, as design evolved and sites grew more complex it became clear that a new mechanism was needed to provide a better separation between presentation and content. Cascading Style Sheets (CSS) are typically separate documents that detail the appearance and position of items on a page. A single style sheet can be referenced by multiple pages. Not only does the use of CSS simplify the modification of a site's appearance, but – because CSS can be contained in a separate document which can be cached – it can save bandwidth as well. Today's Web pages are a combination of HTML, CSS, and JavaScript that is rendered by the client.



### 3.3 Impact, success factors, and limitations

Increasingly, routine human activities can be performed exclusively and efficiently over the Web. Solutions for routine query tasks such as looking up the definition of a word, checking television listings, or looking up addresses are commonplace on the Internet. More involved, user-specific queries can also be performed, including booking a flight or shopping online. The increasing preference towards using the Web is indicative of its success, and indications are that this shift of activities to the Internet will only continue.

Today, everything from authoritative encyclopedias and professional journalism to personal photographs and diaries exist in digital form. The ease of publishing has certainly provided the benefits of increased content. However, this enormous amount of information has created new challenges, particularly those of searching and organization. A digitized book can be quickly interpreted (by humans). It can spread over a network far faster than a printed copy can be shipped and can be (virtually) infinitely replicated at an insignificant cost. Moreover, there is more opportunity for false information to be disseminated as well. Without a semantic context it is difficult to establish credibility. Furthermore, the lack of semantics creates ambiguity and limits the user's ability to pose meaningful queries.

Semantic markup is not rendered by the browser and is, as a result, invisible to the user. The disparity between “contents that are appropriate for browser rendering” and “semantically-encoded contents” can be illustrated by a simple example of a document containing a job announcement (Figures 3 and 4). While the (HTML-formatted) contents of Figure 3 are appropriate to render the job announcement information with proper formatting in a browser window, tags such as `<h2>` or `<p>` do not convey any semantic information. Contrast that with an XML version of the same job announcement (Figure 4) and it is easy to see why XML is one of the fundamental building blocks of the Semantic Web. Tags such as `<skill>` and `<skill-list>` are almost self-explanatory, whereas `<title>` and `<description>` – once properly disambiguated by the use of namespaces – are extremely useful for human as well as computer consumption.

```
<html>
  <head>
    <title>Job Posting: Programmer</title>
  </head>
  <body>
    <h1>JOB POSTING</h1>
    <h2>Job Title: <i>Programmer</i></h2>
    <p><b>Job Description:</b> Our company is expanding! Duties will
include developing applications in PHP for e-commerce. </p>
    <p><b>Skills needed:</b> PHP, C++, and MySQL, good communication
skills</p>
  </body>
</html>
```

Figure 3: Job posting formatted in HTML.

Collaborative efforts have become one of the Web's driving forces, from which Wikipedia is one of the best-known examples [24]. It is an encyclopedia that can be edited by anyone, at any time, without any knowledge of specific syntax (such as HTML). An audit of changes is maintained, and the majority of information appears to be correct, but there is always the

potential for abuse. To date over a million articles have been written in the English language version alone. Users can only perform simple keyword queries and read the content of individual articles to extract information. No semantics are included, which precludes being able to aggregate information automatically across articles. Today such aggregation is performed manually by human operators, such as compiling and maintaining lists of the most populous cities or the GDP of nations. These limitations have been acknowledged and have become the motivations for the Semantic Wikipedia initiative [15].

```
<?xml version="1.0"?>
<job-posting>
  <title>Job Title: <emphasis>Programmer</emphasis></title>
  <description> Our company is expanding! Duties will include developing
applications in PHP for e-commerce.</description>
  <skill-list>
    <skill>PHP</skill>
    <skill>C++</skill>
    <skill>MySQL</skill>
    <skill>Good communication skills</skill>
  </skill-list>
</job-posting>
```

Figure 4: Job posting formatted in XML.

The increasing popularity of user-driven semantic annotation of content is also worth mentioning. Perhaps inspired by blogs, many news sites now provide the ability for anyone to comment on an article. Users can contribute opinions or provide links to relevant information, enhancing the content of the article by adding their own clarifications and semantics. This is evidence of the willingness and ability of users to include more semantics within content. Sites such as digg (a community-driven news site) [9], flickr (photo sharing) [22], and del.icio.us (a collaborative, categorized collection of web sites) [8] have gone beyond annotating existing content; the entire content of these sites is generated by users, typically with a particular user initiating a contribution (such as submitting a news article), enhancing its content (providing commentary and links to other content), and then evaluating the result (users can rate the quality of contributions). The style of these sites has become known as Web 2.0, a suggestive way to indicate the shift in design philosophy that is underway. However, there is a major limitation of Web 2.0 sites. In spite of the users' best intentions to add semantically meaningful knowledge, they often do so in a way that is loosely structured, known as "folksonomy", "as opposed to the classical 'taxonomy', which means a structuring of things into a hierarchy" [26]. Besides being unsuitable for machine consumption, the quality and expressiveness of this user-contributed content is questionable [26].

Content syndication is a recent example of a growing trend in today's Web: the adoption of RSS (Really Simple Syndication) providing real and direct benefits to today's Web users. The syntax of RSS is also XML-compliant. One of the advantages of RSS is that it allows many sources to be aggregated together, saving the user time. Furthermore, the presentation of the RSS feed is up to the user. For example, the user may choose to view feeds on a web page [12], in a standalone news reader [23], in a web-based aggregator [4], in a ticker [10], or inside their web browser [11]. Even the viewing device is not limited to a personal computer. RSS feeds can be viewed on mobile devices as well.

From a content production point of view RSS allows individuals to syndicate and distribute their own content (for example, many blog services automatically generate an RSS feed). The capabilities of RSS feeds are currently being extended beyond text. Podcasts are an example of RSS being used to distribute audio and video. RSS is one example of what we believe is a gradual transition to a more semantic Web.

In summary, while the technologies of the Web are expansive and work well for human consumption they were not designed with machine readability in mind. While tools exist to extract relevant information from a web page and recent user-initiated efforts show some advantages of adding semantics, there is a larger need of structuring knowledge for machine consumption which will not be solved without the introduction of a new way of structuring and processing information, the Semantic Web.

#### **4 The Semantic Web**

The Semantic Web can be defined as “an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation” [3]. Recent efforts within the computer science community have been focused on the creation and adoption of tools, standards, and methodologies that will foster computer-to-computer cooperation. This paper is mostly interested in the impact that the Semantic Web vision, once fully implemented, will have in human activities, particularly knowledge acquisition and dissemination, and the sustainable life changes that may ensue as a result of increased knowledge availability.

In this section we introduce an example scenario that demonstrates the potential of the Semantic Web. We then proceed to discuss the Semantic Web vision and how it more closely resembles the original vision for the World Wide Web. We also outline the key technologies that will make the Semantic Web concept a reality for the user. The wide adoption of Semantic Web tools and technologies and the corresponding human factors are then discussed. This section concludes with a brief analysis of the main challenges that lay ahead.

##### **4.1 Scenario**

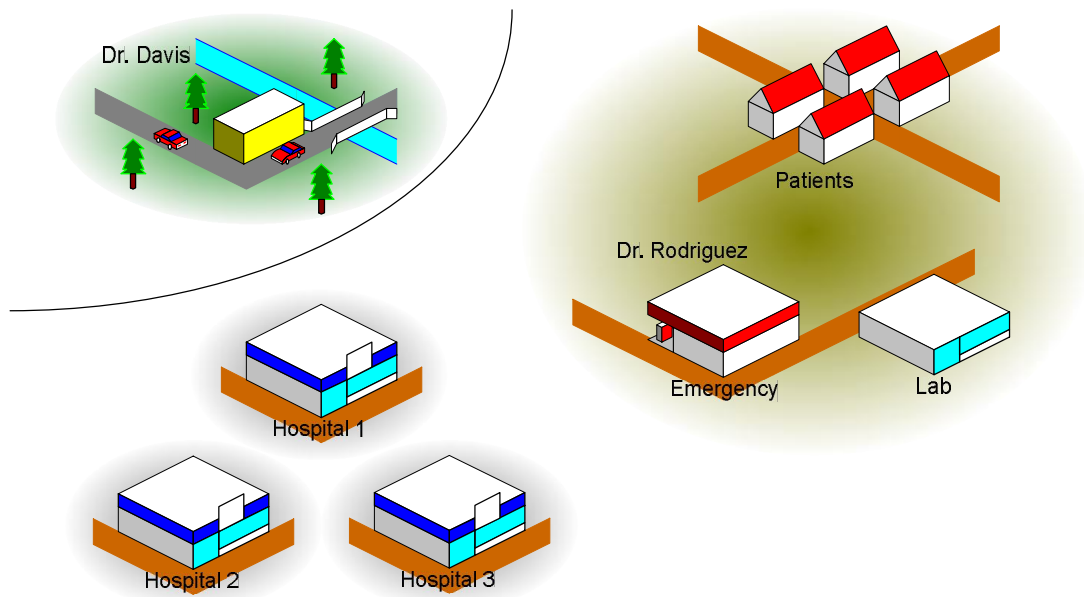
The following scenario (Figure 5), inspired by [3], illustrates the potential of Semantic Web tools and technologies applied to knowledge dissemination.

Dr. Rodriguez is a physician and chief resident in a medical school hospital in a remote location. In addition to teaching and supervising the work of his resident interns, he is also the doctor in charge of emergency care in this community.

Dr. Rodriguez is running the program in this location because of a higher incidence of a particular disease which is being researched by his remotely-located colleague, Dr. Davis. Patients come for regular checkups with Dr. Rodriguez’s team, where their status information is collected and sent to Dr. Davis’s team.

Dr. Rodriguez and Dr. Davis are part of a mutual collaboration where Dr. Davis benefits from Dr. Rodriguez’s personal supervision of the data collection and reciprocates by offering Dr. Rodriguez remote access to expert resources that would not be available onsite. The sum of teaching, supervision, and patient care activities makes the daily routine of Dr. Rodriguez quite busy, leaving little time for additional research. The ability to consult with world-

renowned experts on a variety of medical topics in a timely manner is one of the key aspects of Dr. Rodriguez's teaching activities and partly responsible for his ability to attract high-quality interns to that remote location.



**Figure 5:** A Semantic Web-enabled collaboration scenario.

In addition to Dr. Rodriguez's duties as chief resident and as Dr. Davis's correspondent, Dr. Rodriguez has to respond to the general medical needs of the community. His emergency medical care unit has a very limited number of beds and the minimum equipment necessary for immediate care in non-life-threatening situations. For more serious cases, Dr. Rodriguez needs on-the-fly time-sensitive guidance regarding where to direct patients who cannot be accommodated onsite. There are three hospitals that Dr. Rodriguez can send patients to, each of which has different capabilities and is located in a different direction. Coordination of time-sensitive, on-the-fly information combined with expert medical knowledge is absolutely needed.

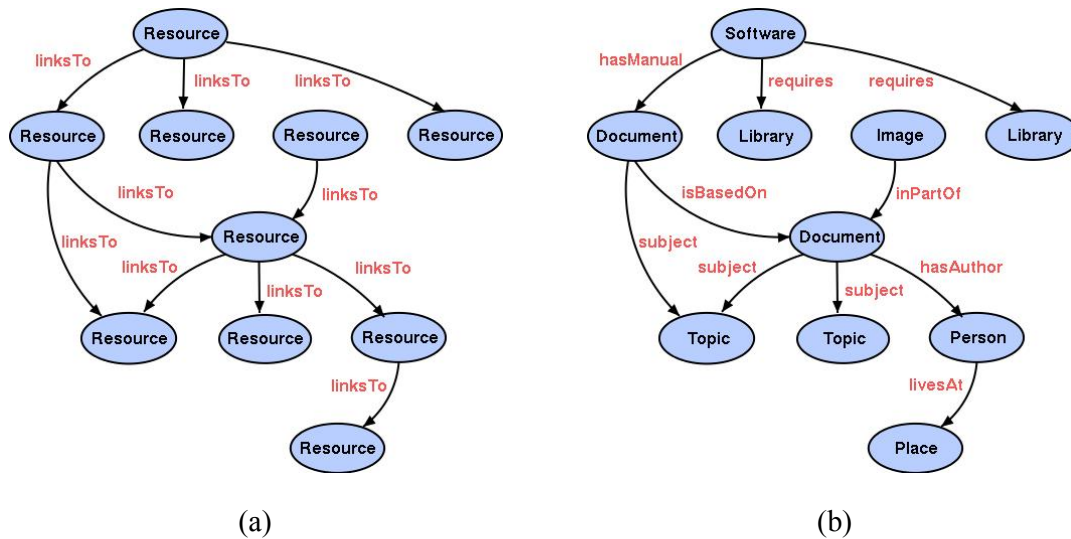
In a Semantic Web-enabled world, knowledge exchange between Dr. Rodriguez, Dr. Davis, and the general medical research community is facilitated by intelligent agents. A group of agents will be able to satisfy Dr. Rodriguez's needs, such as answering a difficult question from one of his students and compiling relevant literature, patient data, and research that provides a satisfactory answer with research from Dr. Davis. Intelligent agents would also assist Dr. Davis with collecting and analyzing the data provided by Dr. Rodriguez's weekly checkups.

For the time-sensitive, life-threatening situations, Dr. Rodriguez relies on another group of intelligent agents whose job is to combine the latest information on the candidate regional hospitals (e.g., bed availability, working state of equipment, blood supplies, staff shifts, etc.) and the roads that lead to them (e.g., traffic, weather and road conditions) with preexisting information (e.g., distances, alternate routes, contact person, etc.) and make a decision as to where to take a patient. In addition to synthesizing preexisting knowledge and current conditions, intelligent agents are able to make a decision and justify it. Such proof would consist of a brief report providing the evidence that was used to make the final decision.

## 4.2 Vision

The Semantic Web vision has been actively promoted since 2000 by the World Wide Web Consortium (W3C). Many research groups, both in academia and industry, have been developing and adopting Semantic Web-related tools, protocols, languages, and technologies over the past five years. Commercial product examples (see [7] for a complete list) include Adobe's Extensible Metadata Platform (XMP) – a labeling technology for embedding data about a file, known as *metadata*, into the file itself and Tidepool / Storymill [17] for image annotation and sharing. Research prototypes include the Magpie [14] browser plugin, which allows users to automatically highlight key items of interest (concepts linked to an underlying ontology) within any visited web page, and the *Friend of a Friend* (FOAF) project [18], focused on creating a Web of machine-readable homepages describing people, the links between them and the things they create and do.

An important part of the Semantic Web is its social component, which has been acknowledged since its very beginning and conceptual formulation. According to Berners-Lee, “the Semantic Web is a computer system, a distributed machine which should function so as to perform socially useful tasks [1]” Knowledge dissemination has also been among the Semantic Web's primary goals: “(...) if properly designed, the Semantic Web can assist the evolution of human knowledge as a whole [3].”



**Figure 6:** (a) Hyperlinks in today's Web; (b) Semantic Web links.

The Semantic Web concept is usually implemented through the use of intelligent agents equipped with instructions on a task to be achieved (e.g., finding the nearest hospital) and its constraints (e.g., equipment and blood availability). Agents use logical rules to select, compare, analyze and filter all the options, and eventually produce a working solution (in this case, a brief report indicating that the patient should be taken to a certain hospital, whose location is within reasonable distance, accessible through decongested roads and whose current equipment condition and supplies match the patient's needs).

The Semantic Web is – among other things – an attempt to restore meaning to hyperlinks, as envisioned by Berners-Lee original concept. In today's Web, links have been implemented as a way to indicate that a resource (e.g., Web page) is linked to another, but without indicating

the relationship between the two (Figure 6(a)). The Semantic Web provides means to make relationships explicit (Figure 6(b)) and uses these relationships to drive intelligent agents that use that information to satisfy users' needs (such as the selection of hospital in the example presented in Section 4.1).

### 4.3 Technologies

The Semantic Web is built upon languages, standards, and technologies, whose details are specified and maintained by Technical Committees under the W3C. Just as the current Web has become a huge success partially thanks to open standards, free and widely available tools, and highly collaborative standardization efforts, the Semantic Web is also mostly built on a free, non-proprietary framework.

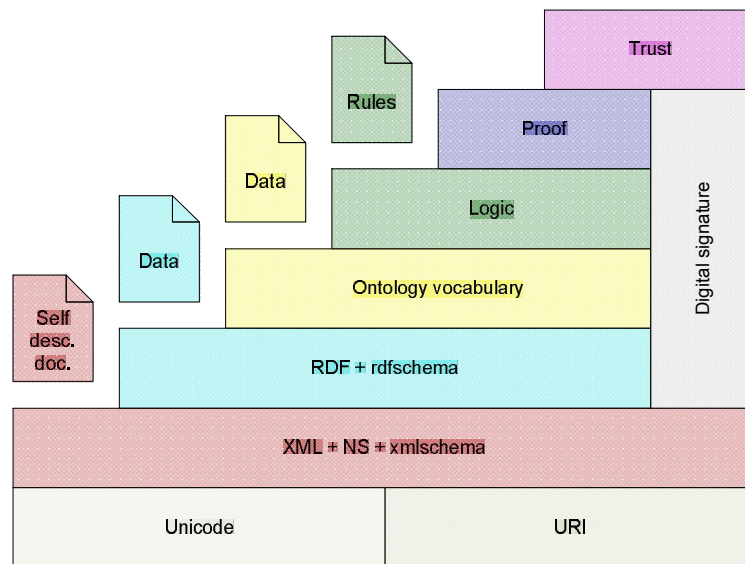


Figure 7: Semantic Web: layer cake (adapted from [13]).

The Semantic Web 'layer cake' (Figure 7) provides a conceptual overview of how some of the most relevant technologies, languages, and standards are related. At the bottom of the cake there is a reference to two basic assumptions upon which the entire edifice is built: that characters can be encoded in a universally unambiguous way (Unicode) and that each and every resource accessible through the Web has a unique identifier (Uniform Resource Identifier, URI). The next layer highlights the eXtensible Markup Language (XML) – a key technology in which much of the Semantic Web is built – and the notion of *namespaces* (used to disambiguate tag names that may mean different things in different contexts, e.g., 'Editor' in a book context versus 'Editor' as in a software tool) and *schemas* (used to enforce a certain structure within the XML documents, e.g., that the 'Zip code' contents should only include digits). Immediately above the XML layer, RDF (Resource Description Framework) provides a data modeling framework for knowledge with a structure based on triples of subject, predicate, and object, whereas RDF schema (RDFS) allows simple ontology-modeling.

The ontology layer provides explicit ways of organizing contents in a user- or community-defined level of granularity and rigor and several languages (e.g. OWL – Web Ontology Language [27]) and tools (e.g. Protégé [28] and Swoop [29]) have been created for that purpose. The logic layer is where most of existing AI knowledge comes into play, with the

design of intelligent algorithms that will make inferences about facts encoded within the lower layers. In the case of our example scenario, relevant medical knowledge to Dr. Rodriguez's occasional questions as well as Dr. Davis's research would have been encoded in appropriate ontologies.

The proof layer is a self-policing mechanism to ensure that the conclusions reached by the logic layer are correct, analogous to a mathematical proof that shows in a convincing, step-by-step way how the agent went from input data to output results. The top layer reflects the fact that this entire concept is based on building a Web of trust, so that agents can rely on what they have been 'told' by other agents and that the ultimate human user – the ambulance driver, in the case of our example scenario – can safely drive to the indicated hospital and expect to find the resources to provide the medical care that the patient needed.

Much of the current research on Semantic Web topics focuses on the design and implementation of intelligent (ontology-driven) agents. An agent in this context can be defined as “programs that act as assistants to the user in the Web interface. They can track user interests, explore the Web proactively, learn through interacting with the user, provide personalized data and services, and much more [25]”.

In our example scenario, there are three distinct groups of agents: Group 1 includes agents that process Dr. Rodriguez's field data (raw data). In addition to usual tasks (such as statistical analysis of numerical results and updates on Dr. Davis's databases), these agents can also compare those results with comparable research being done by other groups, suggest alternative ways to parse and analyze the data, among other tasks. Group 2 contains agents in charge of processing Dr. Rodriguez's (and his students') research questions. Typical duties would include: accessing web-based repositories of medical research journals, summaries of relevant patient data at Dr. Davis's site, and performing the equivalent of getting a second opinion when needed. Finally, the third group of agents would deal with making decisions based on existing information (hospital distance, hospital facilities, etc.) and current information on time-sensitive aspects (traffic, hospital supplies, etc.) necessary to determine which hospital a patient should be sent to.

Work on expanding, refining, implementing, and applying these standards, languages, protocols, and technologies is ever-growing, as is the community of Semantic Web researchers. A large number of universities and companies are doing Semantic Web-related research, and many conferences and conference sessions are devoted to reporting the latest advances on the topic to the community of practitioners.

#### **4.4 Human aspects**

In this section we revisit the scenario of section 4.1 from the perspective of each of the key parties involved, namely: Dr. Rodriguez, Dr. Rodriguez's students, Dr. Rodriguez's patients, the hospitals, and Dr. Davis.

Dr. Rodriguez's students will have access to the latest research in response to their specific questions and needs, supplementing the accessibility to online repositories made possible by today's Web and enabling high-quality medical training in a remote location with rather limited facilities. The acquisition of this information will require modest time investment from Dr. Davis and will be trustable. The limited time of Dr. Rodriguez will be far less of a limiting factor.

Dr. Rodriguez will be able to accommodate very demanding and time-consuming tasks and yet ensure quality of education to his students and quality of care to his patients. His work as chief resident will be made easier by his collaboration with Dr. Davis. His need to be up-to-date with the latest medical literature will be satisfied in a more time-efficient way (e.g., by employing yet another group of agents in charge of collecting and digesting recent medical information relevant to Dr. Rodriguez's needs).

Dr. Davis's research will be carried out in a way that optimizes his time. The data collected at Dr. Rodriguez's site will be analyzed by intelligent agents, saving Dr. Davis and his students' time and providing more accurate results. Dr. Davis's data may also be made more accessible to the larger interested community.

The hospitals will benefit from being able to deliver better care as fewer resources (time and otherwise) are wasted and emergency teams are better prepared thanks to the advance notice triggered by the agents' search for patient accommodations.

Finally, the patients benefit in multiple ways. In addition to having the medical assistance of Dr. Rodriguez they have much better chances of being given the best care at the most appropriate location. Their health will be better-documented, increasing their own health awareness. Additionally, the discovery of trends in their health is also made more apparent to their caretakers, increasing the chances that a cure to common ailments is found. Ultimately, patients are given fairer access to health care.

This scenario is among many that demonstrate the potential of the Semantic Web to promote sustainable life change. Except of the emergency medical case component this scenario is applicable to other domains that share similar characteristics, such as agricultural research, marine biology, and geology.

#### **4.5 Promises, challenges, and expectations**

The dynamics of today's Web reflect the extremely short cycles between conception and implementation. Semantic Web proponents recognize that fact and believe that this is the best way to show its potential users that their efforts are not in vain. To that end, since 2003 there has been an annual *Semantic Web challenge* [6], whose general guideline is to encourage creation of fully functional end-user applications that integrate, combine and deduce information needed to assist users in performing tasks, with Semantic Web standards and technologies under the hood. Another intriguing challenge, initiated by Tim Bray [16], the owner of the RDF.net domain name, offers the domain to the first person or organization that presents an RDF-based app that "I actually want to use on a regular basis (at least once per day), and which has the potential to spread virally"[16]. Both challenges suggest that the Semantic Web community is aware that the transition – and the user efforts that are required to make it – will only be perceived as worthwhile if users get used to applications that they use on a regular basis and that help them achieve their goals more effectively than current Web-based solutions do.

Let us take the search task as an example. Users need to locate relevant information rapidly [20]. Even though the speed at which information can be accessed is ever increasing, due to a combination of faster and cheaper computers, higher bandwidth telecommunication infrastructure and fast search algorithms, finding some information quickly is by no means a

guarantee of its relevance. The problem is compounded when the aspect of trustworthiness is factored into the picture; after all, despite explicit attempts of telling sponsored advertisements from true search results it is difficult to be sure of whether the highest ranked results actually point to reliable information. There is no end to threats to trustworthiness today. Spam, hoaxes, phishing, pharming, and many other variants of electronic scams are all manifestations of such threats. If the Semantic Web can provide a more reliable way to find relevant information at least as quickly as today's search engines the end-user will benefit.

The materialization of the Semantic Web vision will not be without challenges. From ontology creation, adoption, mediation, and maintenance to the potentially huge efforts and cost associated with semantic annotation of legacy contents to successfully building and enforcing a Web of trust, the way ahead is paved with technical challenges for the computer science community. AI researchers, in particular, seem to be enjoying the challenge, understanding that it is a great opportunity to develop an AI-based solution that can live up to its expectations, overcoming the dreadful memories of underwhelming AI applications during the 1980s known as the AI Winter.

## **5 Concluding remarks**

This paper has presented an overview of the Semantic Web vision. We postulate that the Semantic Web will help implement a worldwide knowledge management environment that will bring many benefits to knowledge workers, regardless of their specific needs resulting in improved human conditions.

We have illustrated a detailed example to demonstrate the concept, vision, and applicability of the Semantic Web. The same methodology can be applied to diverse aspects of human life. Scenarios, such as the one described in Section 4.1 will lead to sustainable life change.

The fact that certain things can be done quite easily with today's Web does not mean that they are done ideally and cannot be improved. The human time spent on those activities and the inconsistent level of trust associated with the resulting information are two compelling reasons for improvement. We postulate that the Semantic Web is a framework for overcoming some of the limitations of today's Web and that its envisioned benefits reach far beyond providing machines with the ability to exchange meaningful marked-up data in an automatic and intelligent way. Human users can also benefit from the additional semantics – and the intelligent algorithms to process it – and should use it as a motivating factor for organizing knowledge into a Semantic Web-friendly format.

Many of the agents discussed in this paper could be implemented using today's methods of well-maintained databases and associated applications. In principle, this is partially true, but a number of factors, such as system compatibility and infrequent repository updates considerably reduce the reliability of these solutions.

Our lifestyles will have to change to accommodate the increased presence of agent-based decisions. It is very possible that imperfect implementations and malicious incidents will make certain agents underperform, thereby impacting the perception of the amount of trust they deserve. We expect that such obstacles will not completely stop the progress that can be made, similarly to how the discovery of a new computer virus does not stop people from sending e-mail.

Creating and maintaining ontologies is a critical component of the Semantic Web and entire books (see [19] for an example) have been written on this topic. Our personal view is that – with the best intentions in mind – the Semantic Web community has spent too much time developing competing and partially overlapping languages and technologies and debating which one should be considered the winner, which has cooled down some of the enthusiasm of potential early adopters. If the cost and size of ontology creation and maintenance is intimidating, the additional uncertainty of which language, technology, or format to use makes potential users even more reluctant in embracing the task.

Wikipedia is an excellent example of a desire and need to include semantic information with content. A project is currently underway to extend Wikipedia article with semantic meaning that can be queried by computers [15]. Wikipedia authors must already be familiar with a custom syntax to connect articles together with standard hyperlinks. The Semantic Wikipedia extends this syntax with attributes that contribute semantic knowledge.

The adoption and success of the Semantic Web faces many challenges. Though the vision is well-defined, there still lacks a realization of the "killer application". Standing technical challenges such as creating ontologies and annotating legacy content still remain. However, we believe that these challenges will be overcome. The model (mostly free, decentralized, highly collaborative, with just enough control) that has helped the Web grow to the point where it is now will continue to foster the adoption of technologies, paradigms, languages, and user practices that will enable broader access to and dissemination of knowledge.

In summary, we are optimistic about the potential of the Semantic Web to enhance the dissemination of knowledge, encourage collaboration, foster scientific progress, and expand learning opportunities for all users.

## References

1. Berners-Lee, Tim. "Axioms of Web Architecture." World Wide Web Consortium. 01 Dec 1999. World Wide Web Consortium. <<http://www.w3.org/DesignIssues/Interpretation.html>>.
2. Berners-Lee, Tim. "Information Management: A Proposal." Geneva, Switzerland: CERN, Mar 1989.
3. Berners-Lee, Tim., Hendler, J., & Lassila O. (May 2001). "The Semantic Web." *Scientific American*.
4. Bloglines. IAC Search & Media. <<http://bloglines.com>>
5. Bush, V., "As we may think." *The Atlantic Monthly*. 1945.
6. "Semantic Web Challenge." <<http://challenge.semanticweb.org>>.
7. "CommercialProducts - ESM Wiki." CommercialProducts. <<http://esw.w3.org/topic/CommercialProducts>>.
8. "del.icio.us." del.icio.us. <<http://del.icio.us>>.
9. "digg." digg. <<http://digg.com>>.
10. "ENewsBar: FREE scrolling news ticker. Also, import RSS and Atom feeds." ENewsBar. <<http://enewsbar.com>>.
11. "Firefox – Rediscover the Web. Mozilla Corporation." <<http://www.mozilla.org/firefox>>.
12. "Google." Google Inc. <<http://google.com>>.
13. Hendler, James. "Agents and the Semantic Web", *IEEE Intelligent Systems Journal*. Mar/Apr 2001.
14. Magpie - a tool for Semantic Web. Knowledge Media Institute. <<http://kmi.open.ac.uk/projects/magpie/main.html>>.
15. Max Völkel, Markus Krötzsch, Denny Vrandečić, Heiko Haller, Rudi Studer. "Semantic Wikipedia." *Proceedings of the 15th international conference on World Wide Web, WWW 2006, Edinburgh, Scotland*, May 23-26, 2006.
16. "ongoing – The RDF.net Challenge." The RDF.net Challenge. <<http://www.tbray.org/ongoing/When/200x/2003/05/21/RDFNet>>
17. "storymill – tidepool." Immuxa Corporation. <<http://storymill.com/tidepool/>>.
18. "The friend of a friend (foaf) project." The FOAF project. <<http://www.foaf-project.org>>.
19. "Towards the Semantic Web: Ontology-driven Knowledge Management", Edited by John Davies, Dieter Fensel & Frank van Harmelen. Wiley 2002.
20. Warren, Paul. "Knowledge Management and the Semantic Web: From Scenario to Technology," *IEEE Intelligent Systems*, Vol. 21, No. 1, 2006, pp. 53–59.
21. Weiser, Mark. "Ubiquitous Computing." Xerox PARC Sandbox Server. 17 Mar. 1996. Xerox Palo Alto Research Center. <<http://www.ubiq.com/hypertext/weiser/UbiHome.html>>.
22. "Welcome to Flickr - Photo Sharing." Yahoo Inc. <<http://flickr.com>>.
23. "What is RSS Feedreader? | Feedreader - Totally Free RSS / ATOM Newsreader / Aggregator." Feedreader. <<http://feedreader.com>>.
24. "Wikipedia." Wikimedia Foundation. <<http://wikipedia.org>>.
25. Frye, C., Plusch, M., and Lieberman, H. "Static and Dynamic Semantics of the Web", a chapter in Fensel, D. et al. (eds.) *Spinning the Semantic Web*, MIT Press (Cambridge, MA), 2003.
26. Best, David. "Web 2.0 Next Big Thing or Next Big Internet Bubble?," Technical Report, Technische Universiteit Eindhoven, <<http://page.mi.fu-berlin.de/~best/uni/WIS/Web2.pdf>>.
27. "OWL Web Ontology Language Overview", W3C Recommendation, February 10 2004, <<http://www.w3.org/TR/owl-features/>>.
28. "The Protégé Ontology Editor and Knowledge Acquisition System", Stanford Medical Informatics, <<http://protege.stanford.edu/>>.
29. A. Kalyanpur, B. Parsia, E. Sirin, B. Cuenca Grau, J. Hendler. "SWOOP: A Web Ontology Editing Browser" Elsevier's Journal of Web Semantics (JWS), Vol. 4, No. 2, 2005.