Masaryk University
Faculty of Informatics


# Semantic Web

VB037 – Essay

**Jiří Kremser**

Brno, 5th of January, 2011

## Introduction

Natural language is amazing. Without any effort you can ask a stranger how to find the nearest bus stop; you can share your knowledge of music or martini making with your community of friends; you can go to the library, pick up a book, and learn from an author who lived hundreds of years ago. It is hard to imagine a better API for sharing knowledge. Even though sometimes the information is ambiguous, we can cope with it.

On the other hand, machines are not capable of such behaviour; they process programs in accordance with rules of logic. Logic itself is manipulation of symbolic structures on syntactic level. However, to deeply understand particular concept, not only the syntax is needed. Semantic Web is a set of methods to allow machines to understand the meaning – or "semantics" – of information on the World Wide Web (WWW). [Wik01]

> *"Syntax is not equivalent to nor sufficient for semantics."*
> —John R. Searle, 1980 [Dar98]

## History

The term "Semantic Web" was coined by Tim Berners-Lee, who was one of the co-founders of WWW in CERN. Apart from the(his) proposal of WWW, he is the director of World Wide Web Consortium (W3C) and his present initiative is concerned with the Linked Data.

The term "Semantic Web" denotes the idea of determining the meaning of semi-structured data, mostly web pages, and consecutive accessing the data by intelligent agents acting on behalf of the user. Since it is very difficult to make the present web pages machine-readable, the group of technologies and standards concerning the Semantic Web helps to achieve some subgoals of this main goal. The key standard is Resource Description Framework (RDF).

## RDF

This standard offers means for describing some concepts in terms of logic. The basic component of RDF is a triple *(predicate, subject, object)* in the sense that predicate holds for subject and object. In other words, if one wanted to convey the sentence "The Guardian costs 1£.", RDF triple would consist of *(costs, the Guardian, 50£)*. These simple statements are most often serialized into XML format and can be inserted to described data. For instance it can be included in the web page. The syntax of RDF assumes the knowledge of XML format – hence many web designers do not incorporate [SET09], unfortunately, such metadata into their web pages. The situation will hopefully have changed by the time the HTML5 will be finished.

## Microformats

What is the connection between Semantic Web and HTML5? The answer is microformats (abbreviated $\mu$F); with the new emerging standard HTML5 microformats are being developed. By using microformats it is much easier to anotate your data on web with additional semantic informations, furthermore there will be[1] support in all browsers for some particular types of data. For example, if a piece of the webpage is described by microformats, in this particular case by *hCallendar*, the browser recognizes the presence the callendar event and offers the user adding this event to his associated personal information manager. Another subsets of microformats describe another domains. In my opinion, microformat *hProduct* is very promising since it is used for describing arbitrary commercial goods, including price, standards for identifiers, URL, photo, brand, etc.

Once many products on the web are anotated by *hProduct*, various interesting application may be built on top of it. Gartner, Inc. pronounced following prognosis: "By 2013, mobile phones will overtake PCs as the most common Web access device worldwide." [Gar10] Keeping this prediction in mind imagine the synthesis of smartphones and the Semantic Web. User enters a bookstore, finds a book and then scan the barcode printed on the book with his smartphone. Smartphone is connected to internet, therefore it can use a service used for comparing products. If smartphone has a GPS sensor, it can offer to customer a contextual information: "The bookstore over the street has the same product for lower price."

1. Nowadays, there are plugins for all main browsers implementing this functionality.

## Conclusion

Label “Web 2.0” denotes the shift from static content to dynamically changing content when people create content on their blogs, Wikipedia, etc. There is no consensus on definition of the term “Web 3.0”, nevertheless,(?) scientists, all over the world, admit that the new version of web will contain many semantics elements. The internet community as a whole tends to find the two terms “Semantic Web” and “Web 3.0” to be at least synonymous in concept if not completely interchangeable.

I have to admin, I am looking forward for new technologies related with Semantic Web, but, at the same time, I am little bit worried about the security questions which come into play. The other side of the coin is that with linking data across the web, say on social networks, people reveal their personal data more and more.

### Word count

815

# References

[Dar98] Terry Dartnall. Why (A Kind of) AI Can't Be Done. In *Selected papers from the 11th Australian Joint Conference on Artificial Intelligence on Advanced Topics in Artificial Intelligence*, pages 1–13, London, UK, 1998. Springer-Verlag.

[Gar10] Gartner, Inc. Gartner Highlights Key Predictions for IT Organizations and Users in 2010 and Beyond, 2010. [Online; accessed 14-January-2011].

[SET09] T. Segaran, C. Evans, and J Taylor. *Programming the Semantic Web*. O'Reilly Media, 1005 Gravenstein Highway North, Sebastopol, USA, 2009.

[Wik01] Wikipedia. Semantic web — Wikipedia, the free encyclopedia, 2001. [Online; accessed 14-January-2011].