

# Assignment 2 - Social Web

**Arthur-Ervin Avramiea**  
2517642  
a.e.avramiea@student.vu.nl

**Mihnea  
Dobrescu-Balaur**  
2549278  
mihnea@linux.com

**Kevin**

## 1. THE WEB AND SEMANTIC MARKUP

Much of a web developer's effort is geared towards providing the users with a rich, intuitive interface, that makes the information easy to find. However, most of the users' access to information on the web is nowadays mediated by search engines. The user enters a query and expects the search engine to provide him with websites that serve the content he is interested in. The search engine looks up for the string of text within the indexed database that resulted from web crawling. Nonetheless, textual representation of the information may differ between the different websites, and as such a lot of relevant data sources may not be identified. A solution to this issue is adding structure to the website content by annotating the text with metadata, in the lines of commonly accepted standards, which the search engine will be able to use to identify specific bits of data. This metadata tags the text that is to be presented to the user with its semantics, or meaning. In the end, the web developer's interest in making the information readable and easy to access for the user translates into an effort in making the information readable and easy to access for the search engines.

Several standards have been in use for the past years, among which RDFa, Microdata and Microformat[?]. Semantic markup technologies are adopted rapidly and on a wide scale, to the extent that, in 2013, 50% of the most popular websites according to Alexa, embed some form of structured data.

TOADD: Microformat description and example

TOADD: Microdata + schema.org description and example

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright ACM ...\$10.00

## 2. WEBSITES

### 2.1 Markup for personal webpage

The dynamics of human interaction has obviously changed with the advent of the social web. And in the hectic rhythm of our economy, we need to contact people, or find information about them more often. To be visible on the web, we need a form of presenting ourselves to the others, in the form of a social network account or personal website. It is useful, in this context, to structure the page(s) which contain contact details or other information which may be in the interest of somebody who searches for us, so that the search engine "knows" which information represents our telephone number, job position and company, etc. Not only does this help the search engine to point to such a page when somebody searches for a person, but it also helps the search engine website to provide with a summary of the individuals' information in the results, that is extracted by taking into account the specified metadata. This makes life easier for the search engine user - first of all, if the information which he looks for is already in the search engine, he does not need anymore to open the website and look for the information, second, because it allows the user to discriminate between different persons with the same name, before actually opening the web page.

Within these lines, we have chosen to annotate with semantic markup the personal website of Martin Fowler<sup>1</sup>, who has a diverse experience in the software industry, ranging from software engineer to training, authoring and consulting. The web page provides with rich, albeit unstructured information about him. We chose to use the Microdata standard with Schema.org due to the rich semantics of the available schemas, that are capable of representing a wide variety of aspects of a person. One of the facilities of Schema.org, of which we took advantage in our markup, is the possibility to create standalone items for organizations with which the person is affiliated, persons to which he is related, or items defining contact details, postal address, etc. This items than can be referred as properties of the person, and reused within other items, whenever that is needed.

---

<sup>1</sup><http://martinfowler.com/aboutMe.html>

Although Schema.org is sufficient for the details we have on Martin Fowler's page, if he would have also specified details about his education, skills, experience, scientific publications we would have better used microformat's hResume and hPerson schemas. These have provisions for most of the attributes that comprise a curriculum vitae. A disadvantage, however, with using microformat instead of schema.org, is that we lose the ability to reuse items that we have defined (for example to state that a person is now an alumni of an university, but they are also working at that university) .

## 2.2 Markup for product page

The increase in the number of online stores fosters competition between companies and allows the user with multiple choices. However, when searching a product across multiple websites, the user has to open the webpage of the on-line store to check for the product details and characteristics. In this sense, annotating the webpage of the product with semantic markup would allow the search engine to preview in the results a summary of the characteristics of the product, and thus reduce the time it takes to find the right product and make a choice on the provider.

Marktplaats.nl<sup>2</sup> is such a website, on which users and companies alike can advertise new or second hand products that they sell. We have chosen to markup the data using microformat, with the hProduct schema for describing the product, and hCard schema for describing the seller of the product.

Our choice of microformat was not as much determined by its advantages in this situation over other standards, as by our willingness to test it and see how it works. By comparison, Schema.org has some attributes of the product in the Product schema which are not present in microformat's hProduct, and may be useful for the products description. These include details about the physical properties of the object, such as color, weight, width, height; as well as aggregate ratings.

Although the product we have annotated is intended to be sold only one time (as it is only one second-hand bike available of that type), when a product is sold to multiple consumers by a company, or when multiple persons buy from a company, it is useful to have a review of the product or company. In this respect, although microformat provides with the hReview schema for annotating the reviews which belongs to the product, the Review schema from Schema.org is by far much more detailed.

Given these arguments, although for our product, the markup with microformat sufficed for describing the product and its seller, depending on the characteristics we would like to expose to the user, we would chose microdata with Schema.org instead.

<sup>2</sup><http://www.marktplaats.nl/>

## 3. PUBLISHER EVALUATIONS

## 4. CONCLUSION

## APPENDIX

### A. FIGURES