Generating Semantic Annotations during the Web Design Process

Sven Casteleyn[1], Peter Plessers[1], Olga De Troyer[1]

[1]Vrije Universiteit Brussel
Pleinlaan 2
1050 Elsene – Brussel
Belgium
+32 2 629 37 54
{Sven.Casteleyn | Peter.Plessers| Olga.DeTroyer}@vub.ac.be

ABSTRACT

An important milestone in the evolution of the Web is the Semantic Web: a Web in which the semantics of the available content and functionality is made explicit. Web design methods, originally aimed at offering the designer a well-structured, systematic approach to Web design, now face new challenge: to facilitate the (semi-) automatic generation of semantic annotations describing both content and structure of the website, thereby enabling the Semantic Web. In this short paper, we describe how the existing design method WSDM was adapted to use Semantic Web technology and how the use of Semantic Web technology was exploited to support the generation of semantically annotated websites.

Categories and Subject Descriptors

H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia – *architecture, navigation, theory*; I.6.5 [**Computing Methodologies**]: Model Development – *Modeling methodologies*

General Terms

Design, Human Factors

Keywords

Semantic Annotations, Structural Annotations, Web Design, Semantic Web, Web Engineering.

1. INTRODUCTION

Web design methods were conceived to help the Web designer in coping with the complexity of designing and creating websites. By offering abstraction mechanisms for different design concerns (i.e. most methods distinguish a data, navigation and presentation design) in combination with a systematic development approach, these design methods help the designer to create more usable websites.

The latest developments in the field of the Web are related to the vision of the *Semantic Web* [1]: *an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation*. Web design methods thus face a new challenge: to facilitate the generation of semantically annotated websites, websites in which

the semantics of the *content* is made explicit by means of annotations (with an ontology). For these annotations, existing ontologies can be used, or a suitable (domain) ontology can be created during design. Besides annotating content, it is also useful to annotate the website such that the semantics of its *structure* are made explicit. Dedicated ontologies describing the semantics of structural elements for a particular use (e.g. the WAfA ontology [6] is dedicated to assist visually impaired users) can be used to make the semantics of the different structural elements explicit. Existing Web design methods are thus faced with the challenge to generate these two types of semantic annotations: content-related and structural annotations. In this short paper, we briefly describe how WSDM, an existing design method, was adapted to use semantic web technology internally, and how this is exploited to generate both content-related and structural annotations.

2. WSDM & SEMANTIC ANNOTATIONS

WSDM is an existing design method [2] that consists of five design phases: mission statement specification, audience modeling, conceptual design, implementation design and implementation generation. With the emergence of the Semantic Web, WSDM was adapted to use semantic web technology, and to generate semantic annotations. More specifically, (1) an OWL ontology was specified to formally define the different WSDM design models, and (2) OWL is used as modeling language for the content of the website. The OWL ontology, which formally specifies the different WSDM design models, is called the WSDM Ontology. A full specification of the WSDM Ontology can be found at http://wise.vub.ac.be/ontologies/WSDMOntology.owl. In this short paper, we will only mention WSDM Ontology concepts where relevant for the generation of semantic annotations.

Support for content-related semantic annotations is rooted in WSDM's information modeling phase (part of conceptual design), which serves to model the information and functionality needed by the users. For this purpose, *object chunks* are created. OWL is used as modeling language for the object chunks. All object chunks together form what we call the application ontology (see also [3]). The application ontology can thus be incrementally constructed during the design, but WSDM also allows re-use of existing (domain) ontologies to model object chunks. In this case, object chunks are defined as views over the existing ontologies, which themselves are first aligned using a so called *reference ontology*. In this case, the reference ontology plays the role of application ontology.

Based on an explicit mapping between the application ontology and the actual datasource (e.g., a relational database), content-related annotations can be generated along with the actual content

of the website (see also [4]). Consider as an example an object chunk 'PersonChunk' that represents the address (hasStreet, hasCity and hasCountry) of a person. Using the aforementioned mapping between the application ontology and the datasource, this object chunk can be transformed into an executable query (in our example SQL) on the actual data source. This query looks as follows: SELECT ID, street, city, country FROM Person. The result of this query is a set of instances, in the form of a relational database table. The data of this result table is subsequently transformed into a set of (partial) instances of the application ontology with the requested property values (using the backward data source mapping). In the example, for one particular row this results in the following instantiation:

```
<Person rdf:ID="23">
  <hasStreet>aStreet</hasStreet>
  <hasCity>aCity</hasCity>
  <hasCountry>aCountry</hasCountry>
</Person>
```

From this instance, the (regular) output is generated, in our case a string with the necessary span-tags:

```
<span id="1">aStreet</span> <span id="2">aCity
</span> <span id="3">aCountry</span>
```

Finally, we only have to link this generated HTML code ('page.html' below) and the instantiation of application ontology ('a' below) concepts together using XPointer expressions, e.g. for hasStreet as follows:

page.html#xpointer(id("1"))<=>a#xpointer(id("23")/hasStreet)

Support for structural semantic annotations is achieved by exploiting the semantics of the modelling concepts used in the different design models (and captured in the WSDM Ontology). More in particular, (instantiations of) navigation and presentation modelling concepts convey important semantic information about the structure of the website as a whole, and the internal structure and presentation of each webpage. To make this information explicit in a suitable form, a mapping between the WSDM Ontology concepts and an external ontology describing the semantic of structural elements is defined. Based on this mapping, structural annotation can be generated. As an example, consider the following example of a mapping rule between the WSDM Ontology and the WAfA ontology [5], which is dedicated to assist visually impaired users:

- **WAfA:Homepage:** *A single page that represents the homepage of a website.* The annotation can be created for a generated page that is the result of a wsdm:Page that contains an instance of wsdm:Root. The wsdm:Root concept is used in WSDM to indicate the starting point of the navigation structure of a website.
  ```
  wsdm:Page(?i) ∧ wsdm:hasNode(?i, ?x) ∧
  wsdm:Root(?x) ⇒  wafa:Homepage(?i)
  ```

Based on such mapping rules, content-related annotations can be generated along with the actual website. Consider as an example a simple homepage, containing the Person object chunk described before. At design level, this looks as follows:

```
<wsdm:Page rdf:id="page1">
 <wsdm:hasNode>
  <wsdm:Node rdf:id="node1">
   <wsdm:hasChunk rdf:resource="#PersonChunk"/>
  </wsdm:Node>
 </wsdm:hasNode>
</wsdm:Page>
```

This is transformed into an actual implementation:

```
<html id="page1">
<head></head>
<body>
<span id="1">aStreet</span> <span id="2">aCity
</span> <span id="3">aCountry</span>
</html>
```

And annotated using the aforementioned mapping rules as follows:

http://augmented.man.ac.uk/ontologies/ wafa.owl#Homepage
http://www.example.com/index.html
#xpointer(id("page1"))

A detailed description of structural semantic annotations for visually impaired users, including the generation process, can be found in [5].

3. CONCLUSION
In this short paper, we gave an overview of how the website design method WSDM has been adapted to suit the needs of the Semantic Web. The most important changes are the use of Semantic Web technology (OWL) to formally describe the different WSDM design models (captured by means of the WSDM Ontology), and to model the available data of the website. Next, we gave an overview of how the use of OWL facilitates the generation of two types of semantic annotations: content-related and structural annotations. The use of OWL for information and functionality modeling during the design process allows (semi-) automatically generation of content-related semantic annotations for the resulting website. Structural semantic annotations on the other hand are generated exploiting the semantics of the different design modeling concepts, formally captured in the WSDM Ontology.

4. REFERENCES
[1] Berners-Lee, T., Hendler, J., Lassila, O. The Semantic Web. In Scientific American (May 2001)

[2] De Troyer, O. and Leune, C. : WSDM: A User-Centered Design Method for Web Sites. In Proc. of the 7th WWW Conference, Elsevier, (1998) pp. 85-94

[3] De Troyer, O., Plessers, P., Casteleyn, S.: Conceptual View Integration for Audience Driven Web Design. In CD-ROM Proc. of the WWW2003 Conference, IW3C2, Budapest Hungary (2003)

[4] Plessers, P., De Troyer, O.: Annotation for the Semamtic Web during Website Development, In Proc. of the ICWE 2004 Conference, Munich Germany (2004) pp 349-353

[5] Plessers, P. Casteleyn, S., Yesilada, Y., De Troyer, O., Stevens, R., Harper, S., Goble, C. Accessibility: A Web Engineering Approach, In Proc. Of the 14th Int. WWW Conference (2005) pp 353-362

[6] Yesilada, Y., Harper, S., Goble, G., Stevens, R. Screen Readers Cannot See (Ontology Based Semantic Annotation for Visually Impaired Web Travellers). In Proc. of the ICWE 2004 Conference (2004) pp 445-458