Ontology Based Semantic Annotation for Enhancing Mobility Support for Visually Impaired Web Users

Yeliz Yesilada, Simon Harper, Carole Goble and Robert Stevens
Information Management Group
Department of Computer Science
University of Manchester
Oxford Road
Manchester M13 9PL, UK
yesilady@cs.man.ac.uk

1. ABSTRACT

We have previously shown that the mobility, or ease of travel, of visually impaired Web users is reduced since most Web pages are usually designed for visual interaction[7]. Therefore, in a visually impaired person's environment objects that support travel are missing or inaccessible altogether. We aim to enhance the experience of visually impaired Web travellers by using annotation and *Semantic Web* technologies. In [17], we have proposed a semi-automated tool which encodes techniques for the support of travel upon the Web. The main goal of this tool is to analyse Web pages to identify objects that support mobility and travel, discover their roles, annotate them and transform pages based on these annotations to enhance the provided mobility support. This paper mainly presents the annotation part of the tool and provides some transformation examples which are based on the annotations. The main message of this paper is that visually impaired Web users could also benefit from the *Semantic Web* technologies and here we demonstrate a possible approach to achieve that.

2. INTRODUCTION

Our main goal is to improve the mobility of visually impaired Web users by providing tool support for the provision of mobility. The travel analysis framework which is the foundation for the tool is introduced in [17]. The aim of this tool is to analyse the travel support offered within a Web page and semi-automate the process of:

1. Extracting travel objects;
2. Discovering their roles;
3. Annotating the extracted objects;
4. Transforming the page with respect to annotations.

This paper presents the annotation part (3 above) of the tool[17] and discusses the associated challenges. The architecture of this tool is shown in Figure 1 and particularly, the focus of this paper is on the stage 3 of the architecture; in which Web pages are annotated with concepts from our authoring ontology and translated to mobility concepts which are derived using logical rules (heuristics).

Visually impaired people usually access Web pages either by using screen readers[10] or specialist browsers[3]. If the Web pages are properly designed and laid out in a linear fashion, these assistive technologies work satisfactorily. Particularly, nowadays some screen readers access the HTML source code rather than solely reading the *screen*, which enables them to provide better support. However, not many pages are properly designed; the focus is usually on the visual presentation which makes audio interaction almost impossible. Furthermore, chunking the page into several parts and presenting it in a nonlinear fashion is becoming popular which makes the provided functionalities of these assistive technologies insufficient. There are also guidelines to aid the designers in creating accessible pages[2], unfortunately, few designers follow these guidelines and therefore Web accessibility is still a problem.

The home page of 2^{nd} International Semantic Web Conference (ISWC2003) can be used to illustrate the problem (see Figure 7 part labelled as C). The page is visually laid out into two columns with the main content in the right column. Since the screen reader renders pages based on the tag order in the HTML code, visually impaired users have to *read* the entire left column in order to access the right column. The page is quite long and therefore it takes an unacceptable length of time to read the whole page. Accessibility, and in particular mobility, is not only about the provision of alternative text for images, but also about how easy it is for a traveller to complete a successful journey. For example, if the user wants to directly access the "register now" part of the page, the only way is to read almost the entire page (see Figure 7 part labelled as C). Therefore, the whole journey experience becomes frustrating and unsatisfactory. Further problems also exist when trying to gain an overview of the page. Some screen readers, for instance Jaws [10], provide information for overview when the user first accesses a page. This overview information includes, for example, the number of headings in the page based on the heading tags in the source code. However, if the page is not appropriately designed, such information could be misleading. For example, when ISWC2003 is accessed by Jaws, then it says "one heading". This information is misleading; although the page is quite long and heavy in the content, the user might expect to access a small page. This is a typical problem because the designers use visual effects to specify the headings in the page rather than the proper tags to specify the semantics of the structure.

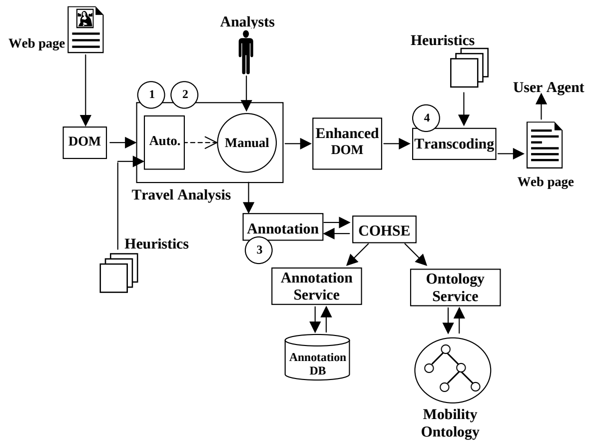

Figure 1: The basic architecture of the mobility support tool.

In fact, the page is semantically organized into chunks, but there is no mechanism for visually impaired users to access those chunks randomly or glance through the chunks. Whereas sighted users can change their focus easily and access the chunks randomly. Some screen readers provide a function for accessing the list of headings in the page based on the heading tags in the HTML source, which means that the page needs to be properly tagged. It allows users to scan the page rather than reading the entire page. These are crucial techniques for the mobility of the user, but very much dependent on proper HTML tagging. Our aim is to extend such techniques by using annotations and semantic Web technologies to improve the mobility experience of visually impaired Web travellers.

One of the goals of the Semantic Web vision is to make knowledge accessible not only to humans but also to agents. In this framework, our goal is to make the role of the objects that support travel and mobility explicitly interpretable by agents rather than just being visually interpretable by humans. Therefore, it is necessary to associate metadata with the objects (machine-readable vs machine-understandable). One mechanism for associating such metadata is annotation.

Our aim is to annotate objects with semantic metadata that provides some indication of their role. This could be considered as another dimension of *semantic annotation* since in this case the annotation is not about the meaning of the resources, but it is rather about the role of the resources. This could be still considered as a further step along the way from simple textual annotation, as the intention within the Semantic Web context is that this information will be accessible to agents. The aim is still making the implicit knowledge explicit and making it accessible not only to humans but also to software agents.

The notion of travel and mobility on the Web is introduced to improve the accessibility of Web pages for visually impaired and other travellers by drawing an analogy between virtual travel and travel in the physical world [8]. Travel is defined as the confident navigation and orientation with purpose, ease and accuracy within an environment. Therefore, travel is not only about Web navigation but it is about the whole journey experience. Mobility is the ease and confidence at which travel can be accomplished. Environment is the context in which the traveller journeys through and includes the way the landscape is rendered and perceived [7]. Travel objects are environmental elements that are used during a journey; in the Web, they are supplied by the page design and the browser. Knowledge of how visually impaired people actually travel gives a context for their travel on the Web[7]. The ontology that we are proposing to use to annotate pages, aims to encapsulate that knowledge. This ontology could be considered as an assistive mechanism for applying physical travelling metaphors to movement around the Web. Our domain of interest is the mobility of visually impaired users and this ontology will be used as a controlled vocabulary for the transformation part of the tool. The COHSE[1] annotator[5] will be used to annotate pages with this ontology. The annotations will be stored externally and accessed by the transformation part of the tool. We also present some example transformation heuristics to illustrate how we will use the annotations.

The ontology that we are proposing to use to annotate pages consists of three parts. The first part encapsulates the knowledge about the travel objects from real world mobility studies– *mobility* concepts. The second part holds information about including hypermedia concepts and vocabularies used in previous work on transcoding– *authoring* concepts. The last part holds information about the context of a journey. The annotation process is encoded in an annotation pipeline. The first two parts of the ontology play important role in this annotation pipeline. Authoring concepts can capture the knowledge about how the objects are *presented* in the environment and mobility concepts can capture the knowledge about how the authoring concepts are *used* in a journey. Therefore, the combination of these two parts of the ontology, could provide extensive knowledge to perform the transformations of the Web pages to ease the travel.

The paper is **structured** as follows: **Section 3** explains the model of travel that is the backbone of the mobility support tool[17] and then introduces the ontology that encapsulates the knowledge provided by this model and by the mobility of visually impaired people. **Section 4** introduces our annotation pipeline. **Section 5** describes COHSE annotator and how the mobility ontology is used to annotate pages. Some example scenarios are explained in **Section 6** that demonstrate how the annotated pages will be used in the transformation stage of the mobility support tool. **Section 7** presents and discusses some related works. Finally, **Section 8** offers some conclusions.

3. MOBILITY ONTOLOGY

Travel in the Web world is likened to travel in the physical world [7]. There has been extensive work undertaken in the mobility of visually impaired people in the physical world, which can be transferred to the Web world. In order to transfer and adapt real world metaphors to Web world, a model of travel is introduced[7] and extended in[17]. In order to complete a successful journey, travellers use or may need to use *travel objects*. These objects are mainly grouped into three broad categories: way points, orientation points and travel assistants. Way points are points within a journey at which a decision may be made that directly facilitates onward journey. Knowledge about orientation suggests that a person needs information about location, distance and direction in order to be oriented in a journey and the objects that provide such an information are categorised as orientation points. Sighted or visually impaired travellers experience problems in orienting themselves from time to time in an unfamiliar or familiar environments, they use different strategies to re-orientate themselves. The objects that they use in these strategies are grouped as travel assistants[17]. Fundamentally, a traveller navigates and orientates by consulting, detecting

[1]The Conceptual Open Hypermedia Project (COHSE) (`http://cohse.semanticweb.org`), particularly, Mozilla plug-in version of the COHSE annotator is used.

and identifying all these travel objects. Consultation, detection and identification are accomplished through the mobility instruments of in-journey guidance, previews, probes and feedbacks. These components form the model of travel[7].

Real world mobility studies also suggest that visually impaired people travel a journey in a different way, using a number of different cues, to sighted people. For example, visually impaired people use simple information more frequently than complex information (for detailed information, please refer to [7]). Knowledge of these differences and how visually impaired people travel provide a context for their travel on the Web. Here, we introduce an ontology which aims to capture the knowledge about the mobility of visually impaired people. The encoded information in the ontology then could be used to provide better support for the provision of mobility.

The mobility ontology serves two purposes: a representation of a shared conceptualisation of knowledge about the mobility of visually impaired people, and a controlled, shared vocabulary that can be communicated across applications. In the context of the semi-automated tool that we aim to provide for the travel support, the ontology will be used as the controlled vocabulary to drive the transformations. Fundamentally, the mobility ontology encodes three groups of concepts which hold information about:

- The context of a journey;
- The role of the travel objects: objects can have a journey role which depends on the context of the undertaken journey and can also have one (or more) environmental role(s) (see Figure 3);
- The authoring concepts encapsulate the knowledge from hypermedia design and previous work on transcoding (see Figure 2).

Figure 3 and 2 show the plain hierarchical representation of the concepts according to the groups above. These concepts mainly aim to provide a framework to encapsulate the information about a journey. Since we do not have the space to explain all the concepts in the ontology, please refer to `http://augmented.man.ac.uk/ontologies/TravelOntology.daml` for the complete ontology and detailed definitions. However, here we will introduce some of the important concepts from each group.

A Web journey can take place in different *contexts*. For example, the purpose of the journey can effect the way user travels through the Web; if the user is searching or browsing effects the objects they use in a Web page. Visually impaired Web users usually use assistive technologies such as screen readers with conventional browsers (e.g., Internet Explorer) to access pages and the assistive technologies can render the Web pages differently and provide a different environment which effects the undertaken journey. Therefore, all these aspects are important for the context of a journey.

The second part of the ontology holds information about the *travel objects*. Objects might have a specific role in an environment and based on the context, they might have another journey role. The journey role is context dependent, for example a graphic site map could be a cue to a sighted user but it could be an obstacle to a visually impaired user. An *obstacle* is an object that directly or indirectly obstructs the progress of a traveller to a specific destination and a *cue* is an object that orientates and encourages onward navigation[9]. The concepts under environmental role aim more to capture the knowledge of what kind of objects traveller use or may need to use to complete a successful journey. For example, decision and identification points are *way points*. *Decision points* are the choice points where alternative paths of travel are possible and identification points identify an object, a place or a person in the environment (for detailed explanations, please refer to [17]).

Authoring concepts hold information about the hypermedia concepts and previous work on transcoding. In this case, we do not consider the roles of the objects in the travel framework but we are more interested in how the objects are presented in the Web landscape. The Web landscape is defined as the combination of the page and the agent (e.g, browser and assistive technologies such as screen readers). For example, there are different ways of presenting boundaries between objects including distinctive colour, space and line and colour boundaries are usually invisible to visually impaired users. In the mobility framework, these objects can have different roles, for example boundaries are way edges.

Based on this ontology, there could be three different approaches for annotating Web pages, and in consequence, transforming them regarding the annotations for enhancing the provided travel support– Web pages could be annotated with a basket of ontologies. These approaches could be as follows:

1. We could use the *mobility journey roles* such as obstacle and cue (see Figure 3), and also the concepts that provide information about the context of the journey. In this case the contextual information is crucial because they determine the journey role of an object, such as the current travel purpose. If enough annotation could be provided to determine the journey role of an object, then we could try to turn obstacles into cues or remove them completely to provide a better journey support for visually impaired Web users.

2. We could use the *mobility environmental roles* (see Figure 3) to annotate the extracted travel objects such as identification point and decision point. These annotations would make the implicit environmental knowledge explicit and enable our agent to have the understanding of the objects in the environment. By the help of the transformation agent, objects will then be able to play their intended roles. Moreover, these annotations could be used to provide different views of a page and provide better support for the orientation and navigation, and thus for the mobility and travel.

3. Finally, we could use the *authoring concepts* (see Figure 2) and use a set of rules to map these authoring concepts to mobility environmental roles– knowing the authoring concept, the mobility environmental role could be inferred from that knowledge. Then based on the mobility environmental roles of the travel objects, we could transform the pages so that the objects can play their intended roles. Since the authoring concepts could be considered as an extension to structures supported by HTML, the rules would also address the basic HTML elements. Table 1 shows some example mappings; these mappings are turned into a set of heuristics. This approach is important for automating the annotation process. Automation could be done in two levels: first obtaining the physical role of the travel objects based on the HTML structural elements and then based on the authoring concepts, we could infer the environmental roles.

In this paper, we particularly explore the third approach – which is crucial for automating the annotation process – and provide further explanation in the next section. In the rest of the paper, we explain how we use the mobility ontology to annotate Web pages and how we transform the pages based on the provided annotations.

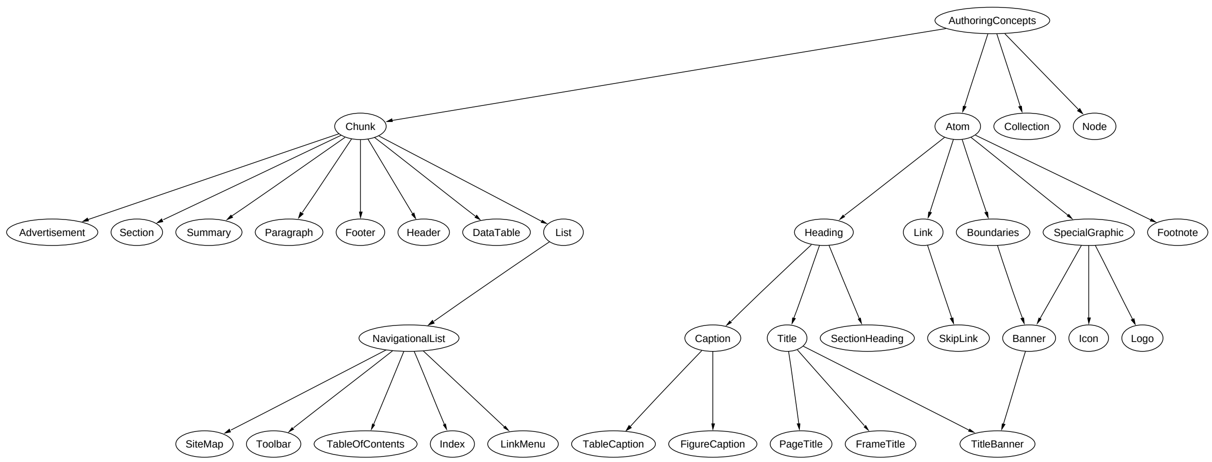

Figure 2: Hierarchical overview of the concepts encoding information about authoring concepts.

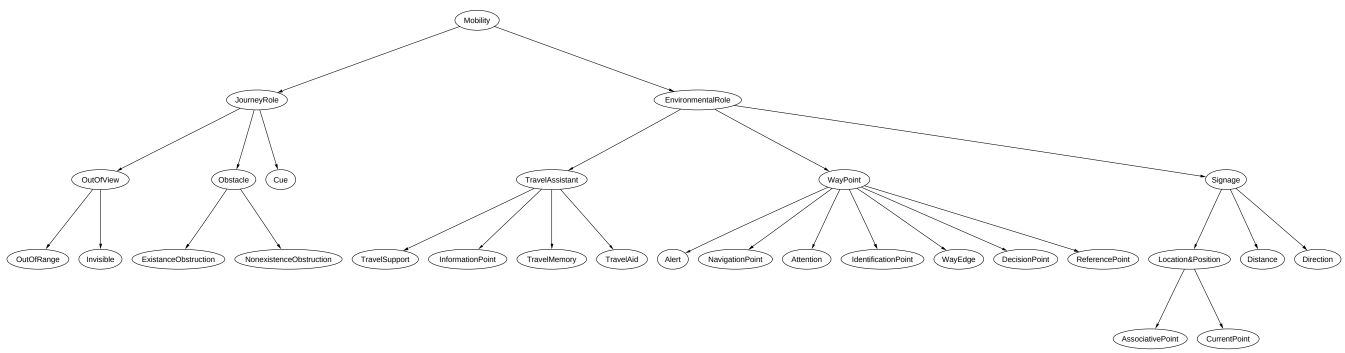

Figure 3: Hierarchical overview of the concepts encoding information about objects that support travel and mobility.

4. THE ANNOTATION PIPELINE

Our system uses a pipeline approach to maintain flexibility in the basket of possible annotation formats it can handle. By using a pipeline we provide for the possibility that annotations can arrive from many desperate sources and in many different forms (RDF, DC, RSS, CMS, etc.). The pipeline approach (see Figure 4) allows us to:

1. Receive different types of inputs from different sources.
2. Harmonize these inputs into a canonical form based on a uniform annotation framework.
3. Recruit annotations manually, semi-automatically, and fully-automatically.
4. Translate between annotation vocabularies associated with hypermedia concepts and 'expansive' annotation vocabularies associated with mobility concepts.
5. Better realise – and simplify – the complex transcoding activity associated with our final goal based on these now expansive mobility annotations.

Figure 4 shows the annotation flow and relates the flow to the architecture of our semi-automated annotation tool which is illustrated in Figure 1 (See Figure 1 for the parts labelled as 3 and 4 on Figure 4). Annotations are received in different formats and translated into a canonical form. To do this we define a simple vocabulary of authoring concepts; associated with previous work on transcoding and commonly used terms found in content management systems and site description languages. We also include an 'expansive' annotation vocabulary associated with mobility concepts. These vocabularies are encoded into an ontology (see previous section) which semantically links the authoring and mobility concepts. We can of course bypass this translation by using COHSE and our authoring, or mobility, concepts to directly annotate the page. Next we use a set of heuristic mapping rules in combination with the ontology to create an enhanced DOM annotated with mobility concepts. This new DOM is now in a suitable format for transcoding and the usually complex process of transcoding is dramatically simplified.
To demonstrate the necessity of our authoring annotation and the associated heuristic mapping rules we will focus on the more interesting part of the pipeline, the translation from authoring to mobility concepts. Table 1 shows some sample mappings from the examples in Section 6 (Figures 6 and 7). We can see that the first column shows an annotated authoring concept (in this case a Heading). The second column shows its physical characteristics taken from the DOM and we can see that in reality it is defined as bold text with a hyperlink and a different background colour to the default page colour. In effect this heading would be missed if it was not annotated as there is no way to distinguish it as a heading because the correct HTML element is not used. Figure 5 shows how this annotation along with a knowledge of the additional physical properties are used to automatically annotate the item with more

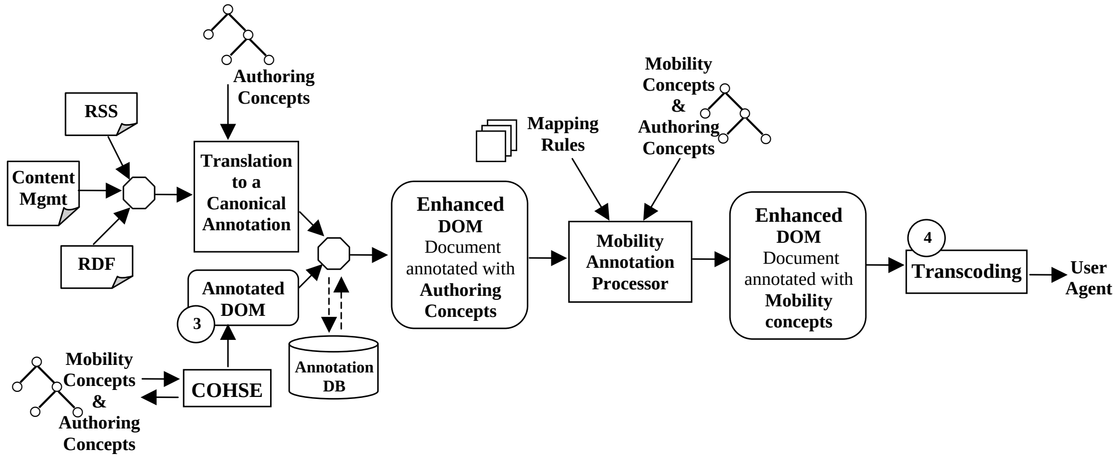

Figure 4: The annotation pipeline (see Figure 1 for the parts labelled as 3 and 4).

Authoring Concepts	Physical characteristics	Mobility Concepts
1. Heading	Bold text, colour boundary and link	Identification point, way edge, way point and navigation point
2. Heading	Bold text, colour boundary	Identification point, way point
3. Link	A text link	Travel memory , travel assistant , navigation and way point
4. Link	An image link	Navigation and way point
5. LinkMenu	A list of text links	Decision, navigation and way point, travel memory and travel assistant
6. LinkMenu	A list of image links	Decision, navigation and way point

Table 1: Some examples of mapping authoring concepts to mobility concepts. The table should be read in conjunction with the explanation in the text (see Section 4) and Figure 5 which describes the heuristic rules.

descriptive mobility concepts. The annotation and physical properties are passed through a set of horn clauses which translate the simple authoring concept to an expansive set of mobility concepts[2].

5. THE COHSE ANNOTATOR

We have used the browser plug-in[3] version of the COHSE annotator to annotate the Web pages with the mobility ontology. There are a number of annotation tools available including MnM and OntoAnnotate[4]. We used the COHSE annotator chiefly developed by our group which gives us an opportunity to develop the tools we need. Additionally, because of its compatibility with Mozilla which is our annotation delivery environment. We have implemented a prototype transformation tool as a plug-in to Mozilla, and using both plug-ins can create a single environment for authoring and publishing the annotations. In addition, the browser can take care of malformed HTML documents. By using a plug-in approach, the transformer, as well as the annotator can access the DOM object built by the browser and can base the transformations and annotations on that[5].

Fundamentally, COHSE aims to combine an open hypermedia architecture with ontological services in order to provide an architecture for Semantic Web[6]. The COHSE annotator facilitates the two supporting services of the COHSE agent: ontology service and annotation service. The annotator interacts with a Web browser and a collection ontologies, providing annotations which are then stored in an annotation service or RDF repository. Ontology service gives access to the ontologies that are used to annotate Web resources and the annotation service maintains mappings between resources and concepts.

The COHSE annotator works as follows: first the ontology is selected which is already uploaded to the ontology service, then a part of the document is highlighted and by selecting the concept, the annotation can be created and stored in the annotation service. It is easy and straightforward process. The COHSE annotator uses XPointer-like expressions to identify the region of the document. Work is underway to bring this into line with the W3C XPointer working draft[5]. Annotation service could be used to retrieve the stored annotations.

6. SOME ANNOTATION AND TRANSFORMATION SCENARIOS

This section introduces some transformation heuristics and their applications based on the external annotations provided. We have created a prototype browser plug-in into Mozilla to perform the transformations. Since the annotations done using the plug-in version of the COHSE annotator, the transformation prototype is also implemented using the same browser[6]. Annotations are stored externally based on the internal DOM tree of the browser, therefore, using the same DOM tree is important for consistency.

The transformation process also raised a number of issues concerning the usage of XPointer and external annotations. Since there

[2] An example of the menu translation is also provided.

[3] Versions have been implemented based on Mozilla and Internet Explorer and in this project we have used the Mozilla version.

[4] See `http://annotation.semanticweb.org/tools` for the list of available annotation tools.

[5] XML Pointer Language (XPointer) W3C Working Draft (http://www.w3.org/TR/xptr/).

[6] Mozilla (http://www.mozilla.org).

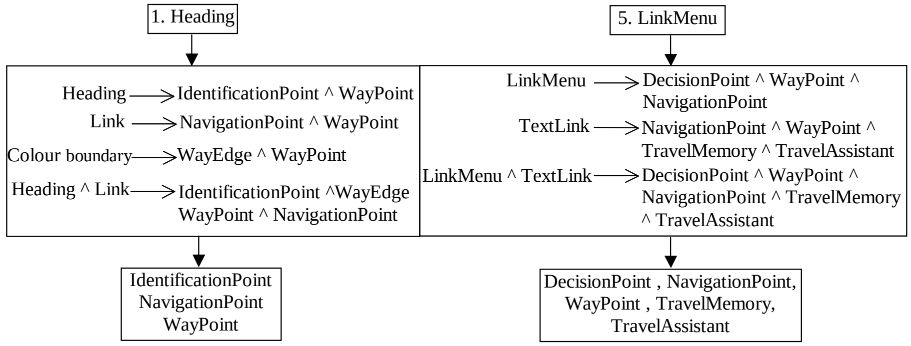

Figure 5: Some example rules for mapping Authoring concepts to Mobility concepts (also see Table 1).

are a number of transformation heuristics that we want to apply, applying one after the other could change the DOM tree and invalidate the existing XPointers in the annotation store. Therefore, in the current prototype, before the transformation process, we have included an intermediate stage to transform the external annotations to internal annotations by using the internal DOM tree of the browser. In this way, we are not actually modifying the original document, but we only add the external annotations to the internal DOM representation of the browser. This intermediate stage is hidden from the user. Since it is not a difficult process we have used an ontology as a controlled vocabulary. However, we still have the problem of dynamically changing pages. Some Web pages change their content and layout almost every day, therefore, even though the annotations are created and stored, they could be easily invalidated. Therefore, we envision incorporating the annotations and mobility ontology either with the content management systems or within the designing process. However, the annotations and the created prototype of the transformation tool could be considered as a *proof of concept*; we would like to demonstrate that the annotations and transformations could improve the mobility of visually impaired Web users.

We return to the home page of the 2^{nd} International Semantic Web Conference (ISWC2003) to demonstrate the implementations of some transformation heuristics based on our annotations (see Figure 7 part labelled as C). This is used because the design is simple, yet good enough to demonstrate some of the issues concerning the mobility support provided by the page. Figure 6 shows the annotations. The page is originally annotated with the simple concepts from the authoring concepts part of the mobility ontology (see Figure 2). Then the mobility environmental roles are inferred automatically from these annotations and the underlying source code, through the process encoded in the annotation pipeline which is explained in Section 4. Figure 6 shows some example translations from simple authoring concepts to an extended set of mobility environmental roles.

We propose to use the annotations to provide techniques for overview as it is explained in the introduction, mainly for enhancing the mobility of the user. For instance, we could use the annotated identification points to provide a kind of table of contents (TOC) (see Figure 6). The TOC could be considered as a way of providing the *bird's eye view* of the page. The annotated identification points can be considered to represent the chunks in the page. We add links from TOC to identification points and also back to the TOC. This could actually be considered as logical fragmentation of the page. Based on the identification points and way edges in the page, we logically fragment the page and allow user to have the preview of these logical fragments. These logical fragments aim to represent the implicit chunks within the page. This is a technique to improve the intra[7] mobility support (mobility support *within* the page), but once we improve the intra mobility support, this could effect the inter[7] (mobility support *between* the pages) and collection wide[7] mobility support (mobility support within the site). We could also physically fragment the page by creating separate pages based on the chunks in the page and allow the user to move from TOC to these pages and back. These two approaches have pros and cons. For example, in the logical fragmentation, the user can continue to read the next chunk without returning back to the TOC. However, the number of links in the page (from/ to TOC) might be too many and difficult for the user to manage them. The extra added links could increase the cognitive demand and maintenance problem. Fragmentation of the Web page is important for good mobility for visually impaired users. It is well known that visually impaired people orientate themselves frequently; returning back to the TOC could be used as an orientation technique, and their route is broken into a greater and more complex number of stages; moving from TOC to chunks and v.v. could increase the number of stages to access the chunks. Fragmentation divides the environment into more manageable and easy to travel units. Moreover, it makes the environment more regular, increases the information flow and supports granularity[7].

Skip links are popular for enhancing the navigation, and thus the mobility, support provided by the page for visually impaired users. They are mainly used at the top of the page to provide a link to the main content, so that the user does not have to *read* the information until the main content of the page. This is mainly for avoiding repetitions, so that whenever visually impaired users access the page, they do not have to read the information at the top to reach to the main content. Therefore, we have a set of heuristics concerning the addition of skip links and particularly deciding upon their targets. For example, if there is a decision point closer to the top of the page, then we add a skip link at the top of the page pointing the first element just after the decision point (see Figure 7 part labelled as A). This heuristic is derived by analysing a number of pages and observing that usually Web pages have a decision point on the left hand side and closer to the top of the page. Adding a skip link can be considered as a simpler version of creating a TOC.

We have some heuristics particularly concerning decision points.

[7] Please refer to the ontology at `http://augmented.man.ac.uk/ontologies/TravelOntology.daml` for detailed explanation.

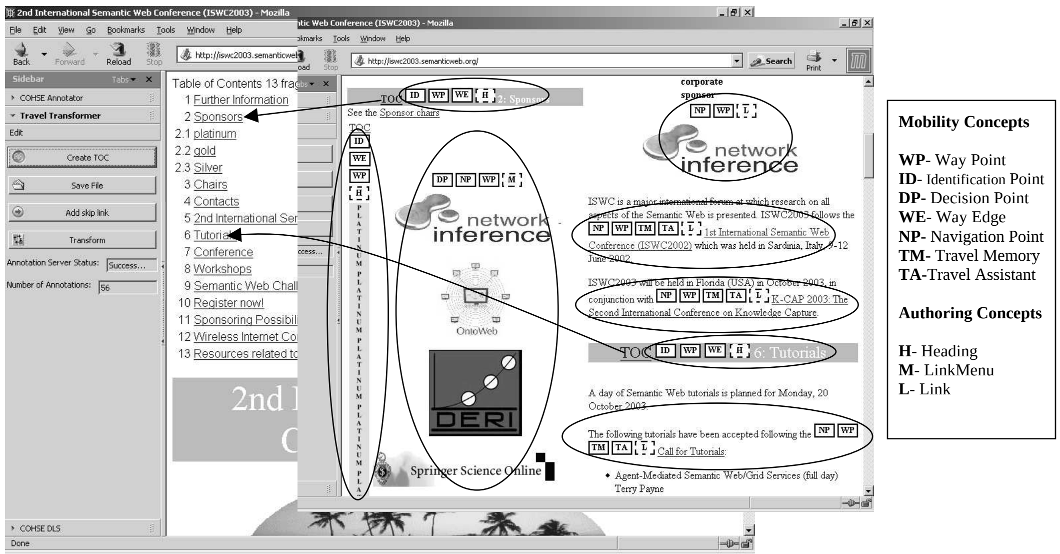

Figure 6: The home page of the 2^{nd} International Semantic Web Conference (12-July-2003) showing the annotations embedded into the page.

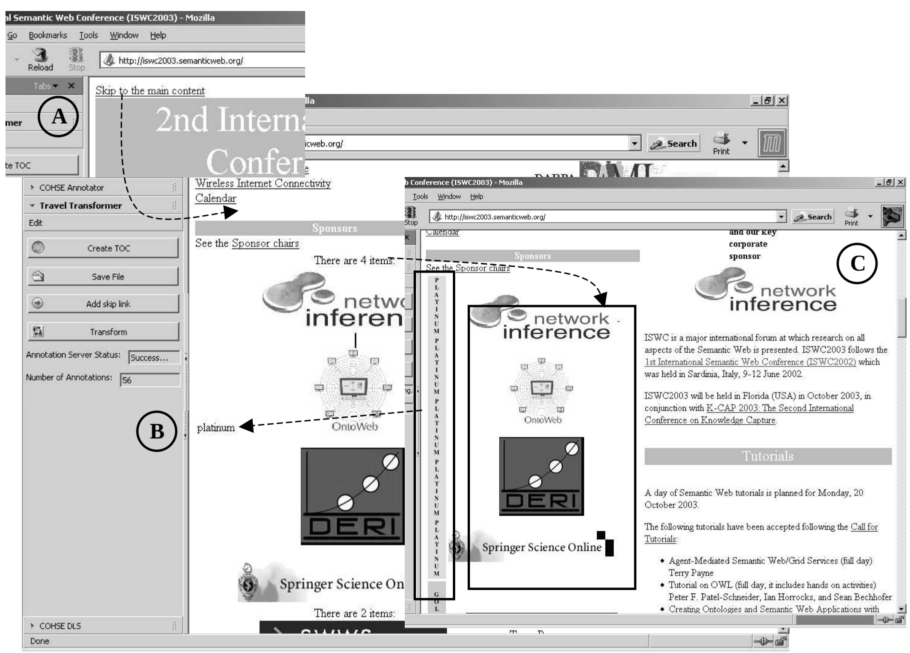

Figure 7: The home page of the 2^{nd} International Semantic Web Conference (12-July-2003) when transformed.

For instance, specifying the number of items explicitly just before the object is one of them (see Figure 7 part labelled as B). This could be important for the predictability and overview of the object. Based on this information, the user could have an idea about the size of the object and how long it will take to read the entire object. Sighted users can easily glance through the object and have an idea about the size, here we try to provide implicit information explicitly somehow to simulate the glancing activity. It could also be useful for the orientation of the user; when he (she) starts to access the items by knowing the size could help user to know where he (she) is.

There are also some heuristics generally applicable to all of the travel objects, for example checking the repetitions. We try to eliminate repetitions within specified objects, in order to reduce the cognitive overload and avoid the confusion the user might have when they have to decide which one to choose. Moreover, the aim is to try to make the page concise and task focused rather than verbose. For example, the sponsors section (see Figure 7 part labelled as B) is actually divided into three sections; platinum, gold and silver. However, the way that this part of the page is designed cause problems to visually impaired users because of several reasons. They are all images and only the images in the gold section have alternative tags, so the platinum and silver sections are invisible to the user. Additionally, images are repeated and the part with gold is rendered as “gold gold gold [network inference]”, which might be confusing to the user. Since they are annotated as identification points, the repeated images are removed and only one of them is left. We also have heuristics for missing Alt tags, such as using the file name of the image. Sometimes the file name might not appropriate but at least it gives an idea about the image. Figure 7 (part B) shows how the sponsors section is transformed.

Essentially, the heuristics and transformations that we have explained here are all simple but have high impact on the provided mobility support of the page and they illustrate how the annotations could derive the transformation of the pages.

7. RELATED WORK

The goal of annotations for Web content transcoding is to provide better support either for audio rendering, and thus for visually impaired users, or for visual rendering in small screen devices. The problem of rendering Web pages in audio has some similarities to the problem of displaying Web pages on small-screen devices. For example, in both cases, only the small portion of the page is viewable at any point. However, there are major differences and requirements. Although the amount of information that could be accessed at once in a small-screen device is also limited, the interaction is still visual. The provided visual rendering is still *persistent*[14], screen acts as an external memory, as opposed to audio rendering which is *transient*. Additionally, compared to visual rendering, audio is less focused and more *serial* in nature[15], the user cannot easily and quickly shift the focus. It is then the aim of this section to discuss related work based on these two themes.

[4, 16] propose a proxy-based system to transcode Web pages based on the external annotations for *visually impaired users*. The main focus is on extracting visually fragmented groupings, their roles and importance. Eight different roles such as proper content, header and footer are proposed for annotation. These roles are mainly at abstract level and are not rich enough to fully annotate the page to enhance the mobility support. They do not support deep understanding and analysis of pages, in consequence the supported transcoding is constrained by these proposed roles.

For *small-screen devices*, [12] proposes a system to transcode an HTML document by fragmenting it into several documents. The transcoding is based on an external annotation framework[1]. Since the focus is the small-screen devices, physical and performance constraints of the devices need to be considered, such as screen size, memory size, and connection bandwidth. However, these are not the main requirements of the users accessing Web pages in audio and there are differences as explained above.

Another approach for content adaptation is page *clipping*[13]. The approach is annotating pages with elements such as keep (content should be preserved) and remove, and then at content delivery, filter the page based on these annotations. This approach is also used for converting HTML to VoiceXML[11]. This is simple and could be an efficient approach, however, our main goal is to identify the roles of the objects in a page and transform accordingly, rather than doing some kind of filtering.

8. SUMMARY

This paper has first presented an ontology that aims to encapsulate the knowledge from real world mobility studies, previous work on transcoding and information about hypermedia concepts. Then, it has discussed several possible annotation and transformation approaches based on this ontology. One of these strategies has been explored and presented in detail. In particular, an annotation pipeline is introduced which can be considered as the core of this approach. The annotation pipeline can be used to annotate Web pages by using different parts of the ontology. Some annotation and transformation scenarios are also explained here to illustrate the appllication and usage of this pipeline.

Our main goal is to improve the mobility support for visually impaired Web users and by using the proposed ontology and also the annotation pipeline, we expect to achieve our goal. The work presented here is still preliminary and much is to be done, in particular an evaluation of the annotation pipeline and the transformation process.

9. REFERENCES

[1] Annotation of web content for transcoding, 1999. W3C Note, `http://www.w3.org/TR/annot/`.

[2] Web content accessibility guidelines 1.0, 1999. `http://www.w3.org/TR/1999/WAI-WEBCONTENT/`.

[3] Chieko Asakawa and Takashi Itoh. User interface of a home page reader. In *Proceedings of the Third International ACM Conference on Assistive Technologies*, pages 149–156, 1998.

[4] Chieko Asakawa and Hironobu Takagi. Annotation-based transcoding for nonvisual web access. In *Proceedings of the Fourth International ACM Conference on Assistive Technologies*, pages 172–179. ACM Press, 2000.

[5] Sean Bechhofer and Carole Goble. Towards annotation using daml+oil. In *K-CAP 2001 workshop on Knowledge Markup and Semantic Annotation*, Victoria B.C, October 2001.

[6] Carole Goble, Sean Bechhofer, Leslie Carr, David De Roure, and Wendy Hall. Conceptual open hypermedia = the semantic web? In *SemWeb2001, The Second International Workshop on the Semantic Web*, Hong Kong, May 2001.

[7] Carole A. Goble, Simon Harper, and Robert Stevens. The travails of visually impaired web travellers. In *Proceedings of the Eleventh International ACM Conference on Hypertext and Hypermedia*, pages 1–10, 2000.

[8] Simon Harper. *Web Mobility for Visually Impaired Surfers*. PhD thesis, The University of Manchester, 2001.

[9] Simon Harper, Robert Stevens, and Carole Goble. Web mobility guidelines for visually impaired surfers. *Journal of Research and Practice in Information Technology Special Issue on HCI*, 33(2), July 2001.

[10] Henter-Joyce, Inc. *Jaws*. `http://www.hj.com`.

[11] Nichelle Hopson. Websphere transcoding publisher: Html-to-voicexml transcoder, 2002. IBM developerWorks.

[12] Masahiro Hori, Goh Kondoh, and Kouichi Ono. Annotation-based web content transcoding. In *Proceedings of the Nineth International World Wide Web Conference*, pages 197–211, 2000.

[13] Masahiro Hori, Kouichi Ono, Teruo Koyanagi, and Mari Abe1. Annotation by transformation for the automatic generation. In *Pervasive 2002*, pages 267–281, 2002.

[14] Jakob Nielsen. Voice interfaces: Assessing the potential. Alertbox, January 2003.

[15] Ian Pitt and Alistair Edwards. *Design of Speech-Based Devices - a Practical Guide*. Springer, London, UK, 2003.

[16] Hironobu Takagi and Chieko Asakawa. Transcoding proxy for nonvisual web access. In *Proceedings of the Fourth International ACM Conference on Assistive Technologies*, pages 164–171. ACM Press, 2000.

[17] Yeliz Yesilada, Robert Stevens, and Carole Goble. A foundation for tool based mobility support for visually impaired web users. In *Proceedings of the Twelfth International World Wide Web Conference*, 2003.