Screen Readers Cannot See

Ontology Based Semantic Annotation for Visually Impaired Web Travellers

Yeliz Yesilada, Simon Harper, Carole Goble, and Robert Stevens

Information Management Group
Department of Computer Science
University of Manchester
Manchester M13 9PL, UK
yesilady@cs.man.ac.uk

Abstract. Travelling upon the Web is difficult for visually impaired users since the Web pages are designed for visual interaction [6]. Visually impaired users usually use screen readers[1] to access the Web in audio. However, unlike sighted users, screen readers cannot see the implicit structural and navigational knowledge encoded within the visual presentation of Web pages. Therefore, in a visually impaired user's environment, objects that support travel are missing or inaccessible. Our approach to remedy this is to annotate pages with an ontology, the Travel Ontology, that aims to encapsulate rich structural and navigational knowledge about these objects. We use *Semantic Web* technologies to make such knowledge explicit and computationally accessible. Our semi-automated tool, *Dante* identifies travel objects on Web pages, annotates them appropriately with the Travel Ontology and uses this to transform the pages to enhance the travel support. Thus *Dante* uses the Travel Ontology to enhance the travel experience of visually impaired users. This paper introduces the Travel Ontology, the annotation pipeline used in the annotation part of *Dante* and some transformation scenarios to illustrate how the annotations are used to guide the transformation of Web pages.

1 Introduction

This paper introduces a semi-automated tool, *Dante*, for the support of travel and mobility for visually impaired Web users. The paper first presents an ontology, the Travel Ontology, and the annotation pipeline facilitated within *Dante*, and then discusses how the Travel Ontology is used to transform pages to enhance the travel experience of visually impaired Web users.

The visual navigational objects that support easy movement around Web pages, or *mobility*, are not appropriate and accessible to visually impaired Web users. These objects are crucial to confident, easy and accurate navigation, which we call *travel* [6]. In order to support mobility, these objects and their roles need to be identified, explicitly specified and presented in a way to fulfil their intended roles. The idea behind *Dante* is to analyse Web pages to extract such objects and annotate them with terms from the Travel

[1] Screen readers are special applications that vocalise the onscreen data. Pages are typically read from the top left to the bottom right, line by line, one word at a time [6].

N. Koch, P. Fraternali, and M. Wirsing (Eds.): ICWE 2004, LNCS 3140, pp. 445–458, 2004.
© Springer-Verlag Berlin Heidelberg 2004

Ontology that aims to encapsulate extensive knowledge about these objects. The Travel Ontology consists of several parts aiming to capture knowledge about how these objects are *presented* (their structural properties) and *used* (their role in supporting mobility) in a typical journey. These annotations, which are a way of associating extensive knowledge to these objects, can then guide the transformation of Web pages to enhance travel and mobility. For the annotation, we use *Semantic Web* technologies. However, unlike other examples[2], we are not annotating a Web page to convey the meaning, but rather to support mobility and convey information about the page itself. The architecture of *Dante*[3] is depicted in Figure 1 and its aim can be summarised as follows:

1. Identifying and extracting objects that support travel, travel objects, from the page;
2. Discovering their roles and structural properties;
3. Annotating the extracted objects by using the Travel Ontology;
4. Transforming the page with respect to these annotations.

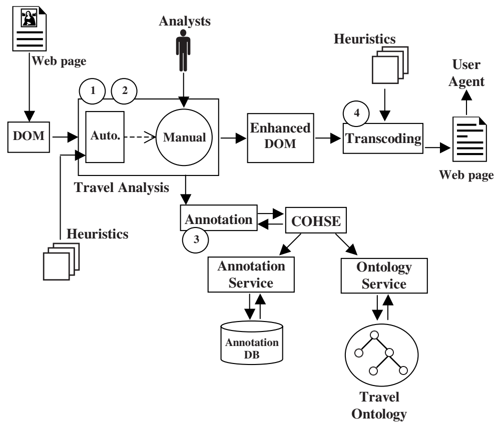

Fig. 1. The basic architecture of *Dante*.

In *Dante* (see Figure 1), the Travel Ontology is used as a controlled vocabulary to guide the transformation of Web pages. The COHSE[4] annotator is used to annotate pages with this ontology, the annotations are stored externally and accessed by the transformation component. The annotation process is encoded in a flexible annotation pipeline and the different parts of the ontology play an important role in this pipeline.

The rest of the paper is organised as follows: Section2 presents the motivation for our work. Section 3 explains the model of travel that is the foundation of *Dante* [17] and introduces the different parts of the ontology. Section 4 discusses how these different parts are used in *Dante*. Then, some example scenarios are explained in Section 5 that demonstrate how the annotated pages are used in the page transformation. Section 6 describes and discusses some related works. Finally, Section 7 provides a summary and offers some discussion.

---

[2] See `http://annotation.semanticweb.org`.

[3] [17] presents the travel analysis framework which is the foundation for *Dante*.

[4] Conceptual Open Hypermedia Service (COHSE) (`http://cohse.semanticweb.org`).

2 Motivation

Visually impaired people usually access Web pages either by using screen readers [9] or specialist browsers [3]. If the Web pages are properly designed and laid out in a linear fashion, these assistive technologies work satisfactorily. 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. Moreover the popular Web sites prove that the available guidelines for designers in creating accessible pages [1] are rarely followed.

The home page of the Mozilla Foundation can be used to illustrate the problem (see Fig. 3 part labelled as A). The page is visually laid out into two columns with the main content in the right column. Since most screen readers render pages based on following tags 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 "Documentation" part of the page, the only way is to read almost the entire page (see Fig. 3 part labelled as A). Therefore, the whole journey experience becomes frustrating and unsatisfactory. Further problems also exist when trying to gain an overview of the page[5].

As a summary, screen readers cannot see and understand the structural semantics implicitly encoded in the page so the mobility support is limited and fragile. Therefore, we need to make the implicit structural and navigational knowledge explicit and accessible to screen readers. The next section presents the Travel Ontology that aims to encode such knowledge.

3 The Travel Ontology

The Travel Ontology serves two purposes: (1) a representation of a shared conceptualisation of knowledge about the mobility of visually impaired people and structures widely supported by Web pages; (2) a controlled, shared vocabulary that can be communicated across applications. In the context of our tool, the ontology will be used as the controlled vocabulary to drive page transformations. Fundamentally, the ontology encodes three groups of concepts which will be presented in the following sections and which in summary hold information about:

– *Mobility* concepts: encapsulate the knowledge about the travel objects from real world mobility studies (how these objects are *used*). Objects can have a journey role which depends on the context of the journey being undertaken and can also have one or more environmental roles (Sect. 3.1);

[5] Please refer to [16] for detailed information.

- *Authoring* concepts: hold information about including hypermedia concepts and vocabularies used in previous work on transcoding and content management systems– encapsulate information about how the objects are *structured* and *presented* in Web pages (Sect. 3.2);
- The *context* of a journey: a Web journey can take place in different contexts [6] and concepts in this group provide contextual knowledge about a journey such as the purpose of the journey being undertaken (Sect. 3.3).

Table 1. The higher level concepts, their documentations and the number of the concepts in the each part of the Travel Ontology.

Part of the Ontology	**Concept**	**Documentation and Example Children**	**No. of Children**
Mobility Semantics:	EnvironmentalRole	Environmental features (elements) that are used or needed by travellers to complete a successful journey (e.g., WayPoint, TravelAssistant, etc.).	16
	JourneyRole	The role of an object in a particular context (e.g., Obstacle, Cue, OutOfView, etc.).	7
Authoring Semantics:	Atom	A coherent object that cannot be decomposed (e.g., Link, Headline, Caption, Footnote, Logo, etc.).	43
	Chunk	Several objects grouped together to form a coherent unit (e.g., Header, Footer, Section, Abstract, LinkMenu, SiteMap, etc.).	42
	Node	A composition of atom(s) and chunk(s) to form a meaningful group (meant to represent a Web page).	0
	Collection	A collection of nodes (meant to represent a Web site).	0
Context Semantics:	Purpose	Intention of either the user (e.g., Browsing, Scanning, etc.) or the object (e.g, AidsNavigation, AidsOrientation, etc.)	14

Due to the space limitations we cannot explain all of the concepts in the ontology[6]. However Table 1 provides an overview by presenting higher level concepts and the number of their children. A description of early work on the ontology can also be found in [16]. The ontology has been created using OilEd[7] and OWL[8].

3.1 Mobility Semantics

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 the Web world, *a model of travel* is introduced [6] and

[6] For the complete ontology, please refer to `http://augmented.man.ac.uk/ontologies/TravelOntology.owl`.

[7] See `http://oiled.man.ac.uk/`.

[8] See `http://www.w3.org/TR/owl-ref/`.

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: These are points within a journey at which a decision may be made that directly facilitates the onward journey. Way points are also further classified and an example is "Decision point" which is the point in the journey where a traveller has to choose from different possible paths (e.g., link menu).

Orientation points: 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 information are orientation points (e.g., logo).

Travel assistants: Sighted or visually impaired travellers experience problems in orienting themselves from time to time in both familiar and unfamiliar environments where they use different strategies to re-orientate themselves. The objects that they use in these strategies are grouped as travel assistants [17] (e.g., site map).

Fundamentally, a traveller navigates and orientates by consulting, detecting and identifying 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 [6].

Based on the model of travel, this part of the ontology holds information about the *travel objects*. Objects might have a specific role in an environment as explained above (travel objects) and based on the context, they might have another journey role. Therefore, beside the travel objects we also have concepts that are about the journey role of the objects. An object can be either *obstacle* or *cue* depending on the context of the journey being undertaken. 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 [8]. 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.

Real world mobility studies also suggest that visually impaired people travel a journey in a different way to sighted people, using a number of different cues. For example, visually impaired people use simple information more frequently than complex information [6]. Knowledge of these differences and how visually impaired people travel provide a context for their travel on the Web and this part of the ontology aims to capture this knowledge. The encoded information in the ontology then could be used to provide better support for the provision of mobility.

3.2 Authoring Semantics

Authoring concepts hold information about the hypermedia concepts, vocabularies used in previous work on transcoding and content management systems. In this case, we do not consider the role(s) 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). These concepts are more to do with the specific structures that can be used to define the overall structure of a page including for example, sections, summaries, abstracts, footers, etc. These constructs are usually implicit in the visual

Table 2. Example travel objects extracted from the home page of the Mozilla Foundation and some examples of mapping authoring concepts to mobility concepts. The table should be read in conjunction with Fig. 3 part labelled as A. Please refer to [17] for further information about the mobility concepts.

No	Authoring Concepts	Documentation	is a kind of...	Some Inferred Mobility Concepts
1	Header	Is printed at the top of a page and can include a company logo, the page title, a link menu and etc.	Chunk	Way Edge
2	Logo	An emblem or a device used to identify the page or a site.	SpecialGraphic	Reference Point
3,21	LinkMenu	A list of links meant to represent a menu.	NavigationalList	Decision& Navigation Point
4	SearchEngine	Consists of an edit box and a button.	Chunk	Information Point
5	Label	An identifying marker attached to an object.	Atom	Identification Point
6	Footnote	A note attached to a part of a page.	Atom	WayPoint
7, 8, 11, 16	Section	A self-contained part of a page.	Chunk	Way Edge
8, 9, 12, 17	Heading	Indicates what the part of the page below is about.	Atom	Identification Point
13	Headline	Is the highlighted heading which identifies the most important part.	Heading	Identification Point
15	Chunk	Several objects grouped together to form a coherent unit.	AuthoringConcept	Way Point
18	NavigationalList	A list of links.	List	Decision Point
19	Footer	Is printed at the bottom of a page and can include copyright information, a list of links and etc.	Chunk	Way Edge
20	Copyright	Is a note about the copyright and is positioned at the bottom of a page.	Footnote	Way Point

presentation of the page, however since they are not explicitly encoded in the underlying source code (e.g., HTML), they are inaccessible in any other form of interaction (e.g., audio interaction through screen readers). The aim here is to define a vocabulary that is already widely used between the designers but not formally explained and defined, that is to say we try to make the domain knowledge explicit. This part of the ontology could be considered as an extension to HTML– aims to provide a rich set of concepts that can be used to describe the overall structure of the pages so that they will be accessible in any from of interaction.

The home page of the Mozilla Foundation can be used to explain some particular concepts in this part of the ontology. Figure 3 (part labelled as A) shows some annotations that has been done by using authoring concepts and Table 2 provides documentation and hierarchical information about these concepts.

3.3 Contextual Semantics

The concepts in this part of the ontology aim to encode contextual information about a typical journey. They particularly address the purpose; this could range from the travellers' purpose (information seeking, surveying, orientation, navigation, browsing, scanning, etc.) to the travel object's intended purpose which is in fact the designer's purpose. One of the possible roles that this part of the ontology could fulfill would be to obtain enough knowledge about the traveller's purpose and to transform pages accordingly. For example, if the traveller wants to scan a page, we could try to provide an overview of the page or if he (she) wants to orientate himself (herself) in the environment (wants to learn where he (she) is in the environment) we could provide objects that support orientational information such as a title, logo, etc. Travel objects can also play different roles in different contexts, for example, for a visually impaired user, a graphic is an obstacle in the context of information searching but a cue in the context of orientation.

The main problem with this contextual information is that it is difficult to obtain. Typically, the traveller's purpose is not explicitly specified (or well-defined) and also the traveller can engage in many different purposes as they travel through the environment.

4 The Annotation Accumulation

This section explains how different parts of the ontology, particularly authoring and mobility concepts, are facilitated in *Dante*. We use a pipeline approach to maintain flexibility in the basket of possible annotation formats. The pipeline (see Fig. 2):

1. Receives inputs from many disparate sources and in many different forms (RDF[9], DC[10], RSS[11], etc.) including manual annotations done by using annotation tools such as COHSE.
2. Harmonizes these inputs into a canonical form based on a uniform ontological framework.
3. Recruits ontological annotations manually, semi and fully-automatically.
4. Translates between annotation vocabularies associated with authoring concepts and with mobility concepts in order to provide extensive knowledge about their roles in the travel framework.
5. Better realises – and simplifies – the complex transcoding activity associated with our final goal based on these now expansive annotations.

Figure 2 shows the annotation flow and relates the flow to the architecture of *Dante* which is illustrated in Fig. 1 (See Fig. 1 for the parts labelled as 3 and 4 on Fig. 2). Annotations can be received in different formats and translated into a canonical form, which we propose to use authoring concepts as explained in Sect. 3.2. Authoring concepts mainly provide information about how these objects are *presented* and *structured*. After we acquire authoring concepts, we use a set of rules[12] to translate authoring concepts to

[9] See `http://www.w3.org/RDF/`.
[10] See `http://dublincore.org/`.
[11] See `http://blogs.law.harvard.edu/tech/rss`.
[12] See the next page for an horn clause representation of an example mapping rule.

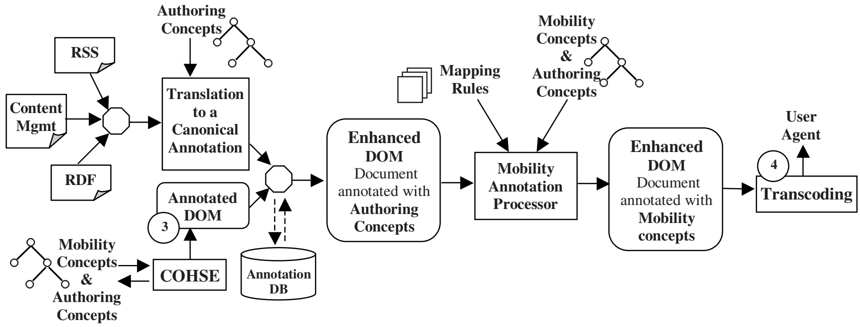

Fig. 2. The annotation pipeline (see Fig. 1 for the parts labelled as 3 and 4).

mobility concepts in order to accumulate enough knowledge about how these objects are *used* in a typical journey. We can of course bypass the translations by using COHSE [5] and our authoring, or mobility concepts to directly annotate the page.

The COHSE annotator uses XPointer[13] expressions to identify the region of the document and annotations are stored in an annotation service [5]. We have used the browser plug-in[14] version of the COHSE annotator to annotate Web pages. Although there are a number of available annotation tools including MnM and OntoAnnotate[15], we have preferred to use COHSE because of its compatibility with Mozilla which is also our annotation delivery environment. The prototype transformation part of *Dante* is implemented 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.

After annotating pages with authoring concepts, we use a set of heuristic mapping rules and the underlying HTML source code in combination with the ontology to create an enhanced DOM annotated with both authoring and mobility concepts. The mapping rules are encoded in JESS[16] which is implemented as a Java servlet. We use the internal DOM tree of Mozilla to obtain the properties of annotated authoring concepts and send these to JESS in order to infer the mobility concepts based on the mapping rules. In more recent work, we believe by evolving the ontology to extend the existing properties of concepts, we will be able to better exploit the reasoning mechanism of OWL to infer the relationship between these two ontologies. After we acquire the mobility concepts, we extend the internal DOM tree by using both annotated authoring and inferred mobility concepts. This new DOM is now in a suitable format for transcoding and the usually complex process of transcoding is dramatically simplified. Table 2 shows some sample

[13] See `http://www.w3.org/TR/xptr/`.
[14] It also has a proxy server version.
[15] See `http://annotation.semanticweb.org/tools` for the list of annotation tools.
[16] See `http://herzberg.ca.sandia.gov/jess/`.

mappings based on Fig. 3 (part labelled as A) and the following horn clause represents an example mapping rule:

$$\begin{array}{c} NavigationalList \\ \downarrow \\ NavigationalList \rightarrow DecisionPoint \wedge NavigationPoint \\ TextLink \rightarrow NavigationPoint \wedge TravelMemory \\ NavigationalList \wedge TextLink \rightarrow DecisionPoint \wedge NavigationPoint \wedge TravelMemory \\ \downarrow \\ DecisionPoint \wedge NavigationPoint \wedge TravelMemory \end{array} \quad (1)$$

This rule mainly applies to the objects that are annotated as a "NavigationalList" and all the links in the list are text. We confirm that the links are text and not images by checking the HTML source code (DOM tree). Therefore by using the provided annotations, the underlying source code and a set of rules, we can accumulate extensive knowledge about the role, structure and usage of the objects.

5 Using the Travel Ontology

We return to the home page of the Mozilla Foundation to demonstrate the implementations of some transformation heuristics based on our annotations (see Fig. 3). This page is used as an example since it is a typical corporation site and provides good demonstration of some of the issues concerning the mobility support provided by the page. Figure 3 (see part labelled as A) and Table 2 shows the annotations. The page is originally annotated with the authoring concepts (see Sect. 3.2), then the mobility concepts are inferred automatically from these annotations and the underlying source code[17]. The annotations are used to provide several techniques for enhancing provided mobility support. Essentially, the heuristics and transformations that we explain here are all simple but have high impact on the provided mobility support of the page and are good enough to illustrate how the annotations can drive the transformation of the pages.

Providing the Overview of the Page

We use the annotated headings[18] (identification points[19]) to provide a kind of table of contents (TOC) (see Fig. 3 part labelled as B). The TOC could be considered as a way of providing the *bird's eye view* (overview) of the page. The annotated chunks, sections and headings represent the fragments in the page. We add links from TOC to headings (identification points) and also back to the TOC to logically fragment the page. Based on the headings (identification points) and sections (way edges[20]) 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

[17] See the previous page (Sect. 4) for an horn clause representation of an example mapping rule.
[18] This type style indicates the concepts in the Travel Ontology.
[19] (A mobility concept) They identify an object, a place or a person in an environment [17].
[20] (A mobility concept) They are environmental elements that are linear (continuous) and act as boundaries between two areas [17].

the page. This is a technique to improve the intra (*within* the page) mobility support, but once we improve this, the inter (*between* the pages) and collection wide mobility supports (within the site) will also be improved.

We can 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. The extra added links can increase the cognitive demand.

Fragmentation of the Web page is important for good mobility for visually impaired users. Physical or logical 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 and consequently satisfies some of the mobility principles [6].

Enabling Movement to the Focus of the Page

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 unnecessary information and is mainly for avoiding repetitions. Therefore, we have a set of heuristics concerning the addition of skip links and particularly deciding upon their targets. The following two heuristics are examples for deciding upon a target for a skip link: (1) if an object is annotated as a headline then we could infer that the section after that is the most important part, therefore we provide a skip link to that object (see Fig. 3 part labelled as C); (2) if there is a decision point[21] 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. We have also assigned priorities to heuristics with the same purpose, in this case the first heuristic have higher priority than the second one. These heuristics are derived by analysing a number of pages and observing common patterns.

Structuring List of Links

The Mozilla home 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. Sighted users can change their focus easily and access the chunks randomly. Some screen readers provide a function for accessing the list of links in the page. It allows users to scan the page rather than reading the entire page. However this technique requires links to have proper labels so that they make sense when they are read out of context, but unfortunately many links are context dependent. For instance, in the home page of the Mozilla, there are links labelled as “More...” and many links repeated as “Windows”, “Linux” which provide links to different versions of a specific product, but if they are not read together with the product heading, it is almost impossible to understand where these links point to.

[21] See Sect. 3.1 for the definition.

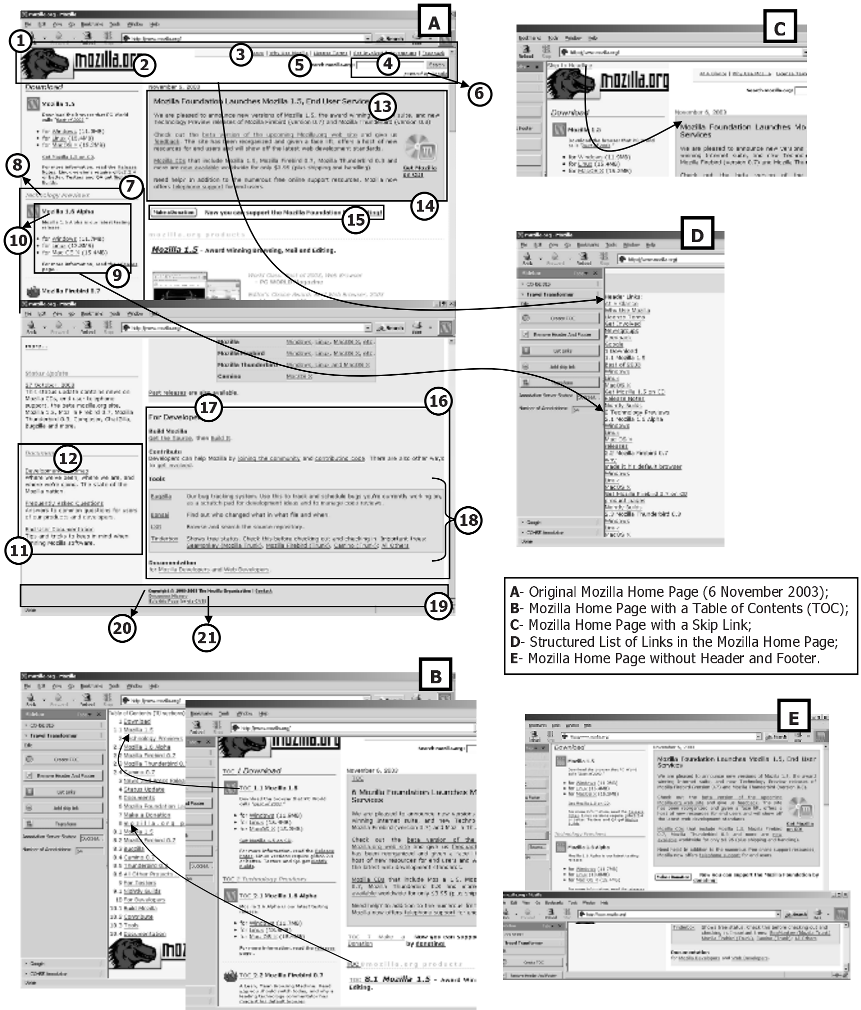

Fig. 3. The home page of the Mozilla Foundation with Transformations (06 November 2003). The part labelled as A should be read in conjunction with Table 2.

These are crucial techniques (e.g., providing a list of links and a list of headings) for the mobility of the user, but very much dependent on proper HTML design and tagging. These techniques can actually be improved by having the structural knowledge. For example, we can improve the provided list of links by putting links in a context but still keeping it short and concise for getting the overview. We propose to use the parts that are annotated as sections and chunks and by using the headings (identification points) in these parts, in order to provide a structure and context to links. This can be considered as grouping links (organizing) which is well-known to aid scanning and memorability of the links [14]. Figure 3 (see Part labelled as D) shows how we use annotations to structure the list of links in the Mozilla home page.

Eliminating Repetitions

There are also some transformations that are useful in case of accessing pages frequently. Some structures such as headers and footers can easily become repetitive and not quite useful if the page is accessed more than once. Sighted users tend not to read such constructs by skipping and directly focusing to the relevant part of the page. However, if you access a page with a screen reader, it is as if you have never been to that page and these constructs cannot be differentiated, therefore there is no supported function that allows you not *read* such parts of the page. Here we propose to remove header and footer in a page so that a shorter and concise page is provided to the user (see Fig. 2 part labelled as E). This technique is particularly useful if the page is accessed more than once or frequently.

6 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* [13], 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 [14], 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.

[4,15] 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*, [11] proposes a system to transcode an HTML document by fragmenting it into several documents. The transcoding is based on an external annotation framework. 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* [12]. 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 [10]. 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.

7 Summary and Discussion

We have first introduced the Travel Ontology[22]. that aims to encapsulate the knowledge from real world mobility studies, previous work on transcoding and information about hypermedia concepts. Then, we have described a possible annotation and transformation approach based on this ontology. In particular, an annotation pipeline is introduced which is the core of this approach. The annotation pipeline is used to annotate Web pages by using different parts of the ontology. Some annotation and transformation scenarios are also explained to illustrate the application and usage of this pipeline.

The annotation pipeline is promising and in this paper we demonstrate that once the annotations are acquired, even the simple transformations can have high impact on the mobility of the user. The future work and the issues concerning the pipeline has two fold: annotation acquisition and the page transformation, and here we discuss some of the issues and future work based on these two folds.

Concerning transformations, we are currently investigating the creation of transformation heuristics and mapping rules based on the type and functionality of the site. [2] proposes eight categories of sites (what a site is and not what is it about) based on their functionalities and we are working on analysing a number of Web sites from each category and detect common structural patterns within and between the categories. This study will provide us a well-defined set of heuristics for different types of sites.

The transformation process has also raised several issues concerning the usage of XPointer and external annotations. Since we want to apply a number of transformation heuristics, 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 and this intermediate stage is hidden from the user. This can be considered a partial solution because 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 the Travel Ontology either with the content management systems or within the designing process. However, the annotations and the created prototype can be considered as a *proof of concept*; our aim is to demonstrate that the annotations and transformations can improve the mobility of visually impaired Web users.

Another possible solution to overcome the dynamically generated pages problem is to annotate pages automatically. Since the authoring concepts could be considered as an extension to structures supported by HTML, the translation rules that we have for mapping authoring concepts to mobility concepts could be extended to address HTML elements. This approach is important for automating the annotation process which could be done in two levels: first obtaining the properties of the travel objects based on the HTML structural elements and then based on the authoring concepts we can infer the mobility concepts. Therefore, we would have an automated process of annotating pages.

[22] For the complete ontology, please refer to `http://augmented.man.ac.uk/ontologies/TravelOntology.owl`.

Our main goal is to improve the mobility support for visually impaired Web users and using the proposed Travel Ontology and also the annotation pipeline lead us to achieve our goal. The work presented here is still continuing and there is still some work to be done, in particular an evaluation of the annotation accumulation process.

Acknowledgments. Yeliz Yesilada gratefully acknowledges the scholarships awarded her by ORS and the Department of Computer Science of the University of Manchester.

References

1. Web content accessibility guidelines 1.0, 1999. `http://www.w3.org/TR/1999/WAI-WEBCONTENT/`.
2. E. Amitay, D. Carmel, A. Darlow, R. Lempel, and A. Soffer. The connectivity sonar: detecting site functionality by structural patterns. In *Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*, pages 38–47, 2003.
3. C. Asakawa and T. Itoh. User interface of a home page reader. In *Proceedings of the Third International ACM Conference on Assistive Technologies*, pages 149–156, 1998.
4. C. Asakawa and H. Takagi. Annotation-based transcoding for nonvisual web access. In *Proceedings of the Fourth International ACM Conference on Assistive Technologies*, pages 172–179, 2000.
5. C. Goble, S. Bechhofer, L. Carr, D. D. Roure, and W. Hall. Conceptual open hypermedia = the semantic web? The Second International Workshop on the Semantic Web, 2001.
6. C. Goble, S. Harper, and R. Stevens. The travails of visually impaired web travellers. In *Proceedings of the Eleventh ACM on Hypertext and Hypermedia*, pages 1–10, 2000.
7. S. Harper. *Web Mobility for Visually Impaired Surfers*. PhD thesis, The University of Manchester, 2001.
8. S. Harper, R. Stevens, and C. Goble. Web mobility guidelines for visually impaired surfers. *Journal of Research and Practice in Information Technology Special Issue on HCI*, 33(2), July 2001.
9. Henter-Joyce, Inc. *Jaws*. `http://www.hj.com`.
10. N. Hopson. Websphere transcoding publisher: Html-to-voicexml transcoder, 2002. IBM developerWorks, `http://www7b.boulder.ibm.com/`.
11. M. Hori, G. Kondoh, and K. Ono. Annotation-based web content transcoding. In *Proceedings of the Ninth International World Wide Web Conference*, pages 197–211, 2000.
12. M. Hori, K. Ono, T. Koyanagi, and M. Abe1. Annotation by transformation for the automatic generation. In *Pervasive 2002*, pages 267–281, 2002.
13. J. Nielsen. Voice interfaces: Assessing the potential. Alertbox, January 2003.
14. I. Pitt and A. Edwards. *Design of Speech-Based Devices - a Practical Guide*. Springer, 2003.
15. H. Takagi and C. Asakawa. Transcoding proxy for nonvisual web access. In *Proceedings of the Fourth International ACM Conference on Assistive Technologies*, pages 164–171, 2000.
16. Y. Yesilada, S. Harper, C. Goble, and R. Stevens. Ontology based semantic annotation for enhancing mobility support for visually impaired web users. In *K-CAP 2003 Workshop on Knowledge Markup and Semantic Annotation*, 2003.
17. Y. Yesilada, R. Stevens, and C. Goble. A foundation for tool based mobility support for visually impaired web users. In *Proceedings of the Twelfth International Conference on World Wide Web*, pages 422–430, 2003.