An Adaptive Semantic Model to Enhance Web Accessibility to Visually Impaired Users

Short Paper

Tatiana Sorrentino[12], Alexandre Santos[2], Joaquim Macedo[2], Cláudia Ribeiro[1]

[1]IFRN, Rio Grande do Norte, Brazil
{tatiana.sorrentino,claudia.ribeiro}@ifrn.edu.br
[2]Centro Algoritmi, Universidade do Minho, Braga, Portugal, 4710-057
{alex,macedo}@di.uminho.pt

Abstract. Web has becoming an invaluable source of knowledge. However, visually impaired users have faced critical difficulties to access the services and data available on Web. To tackle this problem, this paper proposes a semantic model to improve accessibility to websites for users with visual disability. The model is made by components to identify and prioritize relevant information on web pages, converting pages elements into more understandable ones, by a strategy of semantically enrichment that results in an adapted page that meets the user's needs. The model is under development and has been partially implemented and validated in two different scenarios, one dealing with a site that portrays the Brazilian semiarid region and the other one regarding an adaptation of a social network page, replacing an image with an equivalent audio description. The experiment has shown the feasibility of semantic web technologies to improve accessibility.

Keywords: Web Accessibility Standards, Semantic Web, Ontology, Dynamic Adaptation

1 Introduction

Despite the Web evolution along the years, not all users can make use of the full potential it offers. People with special needs, for instance, use the Web and its related technologies at a lower level, when compared to the rest of the user population.

This work presents a proposal for a semantic adaptive model to improve web accessibility for the visually impaired people. In this sense, the model uses the semantic web as the main concept. This model aims to answer the following research question: "*How to enhance the understanding of the content available on Websites for the visually impaired users?*".

In general, websites present a large set of visual elements (videos, images, colors, animations and drawings). Thus, there is a great amount of elements that the screen reader cannot translate entirely. These elements have, in most cases, some components that cannot be interpreted by screen readers in a detailed way. In most cases, this happens because the reading performed by the screen reader software is based on syntactic elements, used in the construction of the web page. Large parts of information are not perceived by the software, as they are presented by non-textual elements that visually impaired people cannot understand due to their limitations.

In this context, the use of the semantic web brings a more effective improvement in the use of the Web. It brings added significance and stronger meaning to the elements and content of web pages. So, it would certainly increase the comprehension of this information by visually impaired users.

This paper is organized as follows: Section 2 introduces some related work; Section 3 describes the proposed model; Section 4 evaluates the experimental results obtained from two experimental scenarios, and finally, Section 5 presents some conclusions.

2 Related Work

Several studies suggest the use of technology to improve the accessibility of sites. Specifically about adaptations, Yesilada et al. [1], proposes a semi-automated tool called Dante, which aims to analyze web pages in order to extract travel objects, discover its profiles and describe them with a specific domain ontology about traveling. Based on the annotations taken, the pages are processed to increase support for mobility, the Wafa [2] ontology was proposed. Regarding the use of ontologies, the SEMA4A ontology, proposed by Malizia et al. [3] has as main objective the development of an ontology adapting notifications in emergency situations. In [4], is presented a ontology to describes the meaning of XHTML elements and associate this with annotations in CSS. [5] Proposes a system to help users in pervasive environments. Our model do the union of semantic and adaptations to improve accessibility.

3 Proposed Model

For the design of the model, a questionnaire was given to visually impaired users in order to identify the main difficulties they face when using the Web. According to their responses, twenty users answered, the adaptation of images and links was prioritized in our semantic model. These elements were identified as the most critical in the access and understanding of web pages.

3.1 Description of Components

As shown in Figure 1, the model is composed by an ontology called *VoidKB* and the components *Extractor*, *Organizer* and *Adapter*. Each of these components is now presented in the following subsections.

(a) *VoidKB* Ontology

The ontology, in the model @dapt++, aims to record additional information about the elements that are available on the web page. This information is used in adapting the pages, enriching semantically the elements, with the information recorded in the ontology. The construction of the ontology was initially based on the main concepts related with accessibility of websites. Then, the system identifies a set of elements whose adaptation enhances the understanding of visually impaired users.

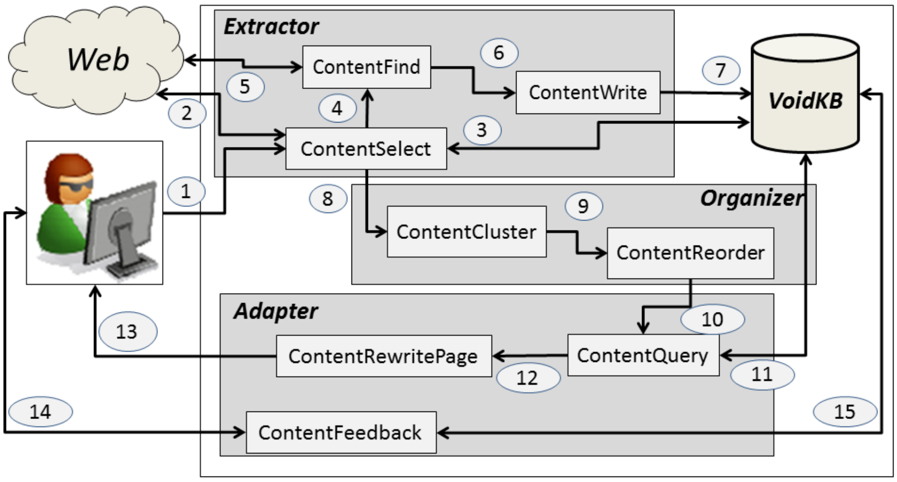

Fig. 1. Architecture of the model @dapt++

Figure 2 shows the classes of the ontology of the model @dapt++. In the ontology, it is possible to annotate descriptions that are used in the adaptation of sites elements, enriching the web page. It is noteworthy that the purpose of the model is to improve both links and images displayed on web pages. All of them are detailed below:

- Adaptation: It consists in the adaptations the system might perform according to the needs of the user.
- Application: The web system which is being used and needs adaptation.
- Context: The set of information about the conditions and applications being used. It encompasses environmental features and devices used to access and use the web system.
- Deficiency: A disability that the user presents.
- Element: This refers to the elements of a web page.
- Media: The forms of content delivery that are present on websites.
- Person: The user of the system.

In the ontology created by this model, the property *owl:sameAs* has been used. This property indicates that two resources are equivalent, that is, both resources represent the same thing. In this context, one may state that an image is equivalent to its audio description. Thus, in the adaptation of web pages, the image originally released on a page can be replaced with the corresponding audio description, increasing the understanding of the content for visually impaired users.

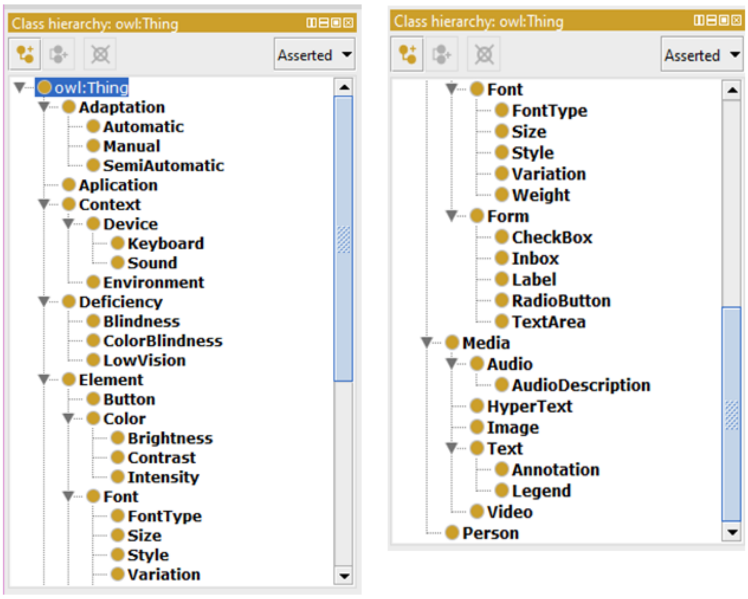

Fig. 2. Classes of the ontology *VoidKB*.

(b) *Extractor*

The *Extractor* component is proposed to perform an automated annotation of the ontology. Relevant information can be identified and extracted from the web page analysis and then used to search related information from other websites. This information can be used to semantically enrich the available elements. In many situations, it is available in an unstructured way and can be extracted automatically. An example of such information is the description of links.

The visually impaired user often do not notice whether the link will open in the same page or if he/she will be redirected to another page although this information is implicitly available in an unstructured way within the web page, but the users does not have access to it.

The *Extractor* component has three main objectives. The first is to select the relevant information from the web page. This information will be handled by the other components of the architecture. In general, the web pages contain many elements that are purely decorative, which are not essential to understanding the page content. Not being essential, these elements may be ignored and this function is performed by the *ContentSelect* module, showed in Algorithm 1.

```
contentSelect(String uri, String user_profile)
{   String web_page := search_web_page(uri)
    Array adaptations_profile := query_ontology (us
```

```
 er_profile)
4    Array tag_elements := analyzes_web_page(web_page)
5    Repeat
6      Array keywords := select_keywords();
7      Array new_elements := remove_decorative_elements
(tags_elements);        }
8    Until tag_elements not null;
9 return keywords;}
```

Algorithm 1. *ContentSelect* Module

The algorithm used by the module *ContentSelect* receives two input parameters (line 1). The first is the URI of the web page chosen by the user. This module is responsible for processing this page and remove keywords about the main elements that the page contains. This information is passed to the *ContentFind* module to carry out an advanced search about page elements. The second parameter is the user profile which maintains information about how to transform the pages according to the user needs. The query on the ontology (line 3) is to consult the user's needs.

The second goal is to make an active search on the Web about the main content of the page. The *ContentFind* module (see Algorithm 2) has this feature.

```
1    contentFind(Array    keywords,    Array    adapta-
tions_profile,   Array ontology_data)
2 {String summary := busca_google_api(keywords);
3  String additional_data := busca_google_api(keywords,
ontology_data);
4  String other_information:= busca_swgoogle_api (key
 words, ontology_data);
5  Array data_valid := data_validation(summary,
additional_data, other_information);
6 return data_valid;}
```

Algorithm 2. *ContentFind* Module

This algorithm used by *ContentFind* module receives three input parameters (line 1). The first of these are the keywords about the elements contained in the original page chosen by the user, which are initially identified by *ContentSelect* module.

The second parameter is the information on the necessary adjustments according to the user profile, which are annotated in the ontology. This information represents the needs of the adaptation of the page, depending on the users requirements. For example, for a user with low vision, increasing the font size is a enough, but for a user with total blindness, it is not an appropriate adaptation. This information is configured on the ontology.

The third parameter regards the features of the ontology which are used in the search to better describe the elements available on the original page, for example, colors and shapes. Thus, the search (line 4) will be directed by the characteristics of the elements recorded in the ontology as well as the adjustments necessary to the user.

The third objective is to automate the annotation of information in the ontology. The data that will be annotated in the ontology results from active searches automatically performed. *ContentWrite* is the module responsible for this task. Its main interactions are with the user, the Web and the components *Organize* and the *VoidKB* ontology.

(c)*Organizer*.

The *Organizer* component is used to improve the organization of content within the page. From the analysis performed on the page, it is possible to identify related elements and group them in order to obtain an adapted page better organized in terms of content similarity.

One of objectives of the *Organizer* is to cluster the content available on a website. This clustering will be based on the semantic similarity of content. The output is a tree with the semantic structure of the web page. This structure will be passed to the module *ContentReorder*. This module is responsible for writing the web page according to the received tree structure. The *Organizer* interacts with *Extractor* and *Adapter* components.

(d) *Adapter*.

The *Adapter* component has as objectives to query the ontology, rewriting of the web page and get a user's feedback about the adaptations. In *VoidKB* ontology, the information on the page elements that should be adapted to improve accessibility is available. Thus, the *ContentQuery* module is responsible for query the ontology.

The *ContentRewritePage* module is responsible for writing the adapted web pages. It includes the adjustments to be made, enriching the pages with information which was added in the ontology.

The *ContentFeedback* module gathers feedback from the users about the quality of the web page adaptation which was done. This feedback can be collected with evaluation forms filled by the user about of adapted pages.

The main interactions of this component are with the *Organizer*, the *VoidKB* ontology and the end user. This service has been partly implemented.

(e)Interaction between Modules.

Figure 1 describes the operation of the system. In the following, a brief description of the interaction between modules:

1. The module *ContentSelect* receives the URL of the web page chosen by the user.
2. The module *ContentSelect* queries the ontology to access the user profile and information about the user's adaptation needs.
3. Furthermore, it selects important information from the content of the page.
4. The module *ContentFind* has the most important elements of the page from the *ContentSelect* as input. For each element, it searches related additional data on the

Web. This step may be skipped (or repeated) if the page was already adapted by the system.

5. The original elements and associated data are automatically recorded in the ontology by *ContentWrite* module.
6. The module *ContentWrite* updates the ontology, annotating the data received from *ContentFind* module.
7. The module *ContentSelect* simplifies the web page, maintaining only the important elements, and sends it to *ContentCluster*. This module is responsible for clustering the page elements, according with semantic similarity of their content. The output of the *ContentCluster* module is a tree structure with the content of the page semantically organized;
8. *ContentCluster* identifies the semantic similarity of the page contents and forwards the semantic structure of the page to *ContentReorder*. The resulting output of this step is a tree structure with semantically organized content;
9. *ContentReorder* reorganizes the elements following the semantic structure and forwards it semantically restructured to *ContentQuery*.
10. *ContentQuery* queries the ontology to find the annotations taken on the page content.
11. The module *ContentRewritePage* receives two pieces of information: 1) the reorganized page and 2) the semantically annotated content in the ontology about its elements. With this information, this module adds the ontology information on page, enriching it semantically. The result of this step is an adapted web page and better organized semantically;
12. The adapted web page is sent to the end user.
13. *ContentFeedback* sends a form to the user for adjustments on the page to be evaluated.
14. *ContentFeedback* collets the user's feedback about the adapted pages.
15. The results of these evaluations are recorded in Ontology *VoidKB*.

4 Experimental Results

To validate the model of accessibility @dapt++, two case studies were used. The first of them deals with the scenario of the Brazilian semiarid region. The second one is performed using the context of online social networks. In the context of semiarid region, a site in which provides the Technological Information Service in Agriculture (Infoteca-e) was chosen. The Infoteca-e site exhibits an image of the Brazilian flag that appears on the original page. The system searches a description that is equivalent to the flag of the Brazil in the ontology, and when it is found, the system automatically adds the description to the page replacing this image.

The other case study was done using the Facebook social network. In this experiment, a Facebook page was submitted to our accessibility system resulting in adapted pages.

That page contains an image of a palm tree created by artist named J.Alencar. Both the image and a video containing audio description of the artwork were recorded in the ontology. After recording all the elements in the ontology created, the page was submitted to *Adapter* service. As a result, a reformulated page was automatically obtained and in this adaptation, instead of the image on the original page, an audio description of the painting was included, giving to the user a detailed description.

5 Conclusions

The main objective of the @dapt++ model is to automatically perform adaptations in web pages. To fulfill this objective, some pages selected by visually impaired users were analyzed and the most relevant information was identified. This relevant information was then used as a basis to search additional and related information. The resulting information was annotated and represented by ontological elements. Besides, the page of components was analyzed and semantically grouped, according to content similarity criteria.

The resulting page is better organized and enriched with additional and relevant data automatically obtained from the ontology. Then, the adapted page is finally delivered to the user. It is important to notice that this adaptation process is in accordance with W3C recommendations.

Despite the fact that @dapt++ is an ongoing project, the results achieved so far, in some way, demonstrate its feasibility and relevance, specially, in the context of automated web accessibility tools. This is basically what motivates carrying out a new stage of the development of @dapt++, which regards the adding meaningful semantic structure on web pages.

Acknowledgements

This work has been partially supported by FCT - Fundação para a Ciência e Tecnologia, Portugal, in the scope of the project: UID/CEC/00319/2013.

References

[1] Y. Yesilada, R. Stevens, S. Harper, and C. Goble, "Evaluating DANTE," *ACM Transactions on Computer-Human Interaction*, vol. 14, no. 3, Sep. 2007.

[2] S. Harper and Y. Yesilada, "Web Authoring for Accessibility (WAfA)," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, no. 3, pp. 175–179, Sep. 2007.

[3] A. Malizia, T. Onorati, P. Diaz, I. Aedo, and F. Astorga-Paliza, "SEMA4A: An ontology for emergency notification systems accessibility," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3380–3391, Apr. 2010.

[4] S. Bechhofer, S. Harper, and D. Lunn, "SADIe: Semantic Annotation for Accessibility," *The Semantic Web-ISWC*, pp. 101–115, 2006.

[5] K. L. Skillen, L. Chen, C. D. Nugent, M. P. Donnelly, W. Burns, and I. Solheim, "Ontological user modelling and semantic rule-based reasoning for personalisation of Help-On-Demand services in pervasive environments," *Future Generation Computer Systems*, vol. 34, pp. 97–109, 2014.