An Accessible Media Player as a User Agent for the Web

A. Mourouzis[1], N. Partarakis[1], C. Doulgeraki[1], C. Galanakis[1], and C. Stephanidis[1,2]

[1] Foundation for Research and Technology – Hellas (FORTH)
Institute of Computer Science
Heraklion, GR-70013, Greece
{mourouzi,partarak,cdoulger,galanakis,cs}@ics.forth.gr
[2] University of Crete, Department of Computer Science

Abstract. This paper presents *webLection*, a tool that aims at increasing the uptake of the Web by a variety of potential users. With *WebLection* Web content is locally rendered into standard audio formats, and, ultimately, manipulated just like music employing the *playing list* behaviour as an intuitive interaction metaphor. Although *webLection* originates from research for blind Web users, significant benefits emerge for other types users too, including Web developers, since it can be used, on the one hand, to easily transform and make Web content available through its audio equivalent, and on the other hand to inspect the accessibility and usability of Web pages for users of screen readers.

Keywords: Web accessibility, mobility; usability, non-visual interaction.

1 Introduction

Nowadays, access to the World Wide Web is critical in many aspects of life. Especially for blind people, the Web offers unprecedented opportunities for supporting independent living. In principle, an accessible Web can help blind users participate more actively in society and enjoy unique opportunities to information, education, employment, commerce, health care, recreation and more. Currently, access to the Web by blind users is feasible mainly through assistive technologies, such as screen readers or voice browsers. This work proposes an alternative approach, in which Web content is rendered into, and ultimately played back as, audio. In this way, blind users, who are reportedly keen users of media players, can access and navigate Web content employing a 'playing list' metaphor, just like when listening to music. It is argued that, from a human-computer interaction perspective, this approach shows a number of advantages. This builds on the fact that (inexperienced) blind users find it easier and more satisfying to handle media players and audio files (or playing lists), rather than browsers, screen readers and Web pages.

2 Concept Overview

A prototype research tool, named *webLection*, has been developed in order to experiment on the concepts described above. Towards the implementation of the proposed approach, efforts have been focused on the following directions:

K. Miesenberger et al. (Eds.): ICCHP 2008, LNCS 5105, pp. 474–481, 2008.

(a) *Provide functionality for rendering Web content into audio equivalent, with the objective* to support users in 'reading' Web content using any standard media player, as well as to support Web developers in easily adding to each produced Web page a link with the page's alternative in audio format (merely a printing facility for blind users).

(b) *Deliver the ability of hyperlinks-based, non visual interaction and navigation.* This choice is motivated by the fact that navigation through hyperlinks is more empowering than navigation based on the file hierarchy of Web pages (as dictated by their URLs), since hyperlinks, between them, control how the reader perceives the overall structure, and how the traffic flows between the different parts of t sites.

(c) *Integrate facilities for transforming textual content into audio equivalent*, thus supporting users in creating their own content and transforming it to forms compatible to any standard media player, and supporting developers in easily creating pre-rendered audio equivalents of textual content.

(d) Incorporate a *playing lists editor* that enables users to compile their own libraries by assembling pieces of information spread on various websites, or created by themselves as described in (c).

An indicative use case is that of a blind person who wakes up in the morning, and downloads her favourite newspaper (e.g., from a Web portal) and her emails (e.g., from her web mailer), loads them to an MP3 player (the same device used to listen to music), or to the playing list of an media player embedded in a cell phone, or even to the car stereo, and ultimately 'read' them on the way to work.

Although *webLection* was primarily targeted to blind users, significant benefits emerge for all users. In particular, the proposed approach is considered to be useful for (i) mobile users, as it supports interchanging between online and offline working modes; (ii) the very young, the elderly and the inexperienced Internet user, who often find it difficult to build effective conceptual models of the functionality of browsers or of the workings of the Web; and (iii) people with other disability, such as users with impaired memory who experience difficulties in exploration activities that involves remembering many combinations of actions and outcomes, or users with upper limp impairments who are finding user interfaces that require extensive mouse or keyboard usage prohibitive. Finally, *webLection* allows to assess, quickly and easily, the accessibility of Web content (e.g., a webpage), without the need, from the content developers side, to operate screen readers or to be able to understand and apply accessibility guidelines. Content developers need only to "play" a webpage under question and check, intuitively, whether the audio equivalent generated by webLection is comprehensive and efficient.

The next section provides a brief overview of related and background work in order to clarify the motivation and rationale behind the proposed approach. Then, the design and implementation of the prototype tool *webLection* are presented.

3 Background

Current approaches to the provision of accessibility to computer-based applications and services are mainly based on adaptations to existing systems or on "dedicated" developments targeted to specific user categories [1]. Along the same lines, attempts to provide Web accessibility are usually based on adaptations that can roughly be separated into three different levels: alternative access systems; the information content and structure; and the user interface [6]. For instance, access to the Web by blind users is mainly made possible through special technologies, such as screen readers (e.g., *JAWS*[1] and *HAL*[2]), or special Web browsers (e.g., the *BrookesTalk*[3] browser for visually impaired users, Zajicek, 2000; *Home Page Reader*[4], IBM's browser for users that are blind or have low vision). Following a more generic and systematic approach, AVANTI [5] was the first universally accessible web browser, addressing the requirements of able-bodied, blind and motor impaired users.

Analyses of blind users' interaction with such systems showed that they were unable to build appropriate conceptual models of the functionality of such systems, and they often become confused and frustrated (e.g., [8]). In addition, poorly developed conceptual models of the Web, as distinct from the browser or the screen reader, also form a major barrier to successful Web use for blind users. Sighted users rely on complex and contextual conceptual models and many visual clues to help them find information on the Web [9]. Furthermore, it is not equally easy for blind users to follow a hyperlink, reinforce their concepts, learn from their experience, and then return to their original position. Nonetheless, novice blind, but also many elderly, users find difficulty in remembering sequences of actions previously performed [7], and they often 'borrow' the model of the video recorder [8].

Under the light of the above, *webLection* is aimed at providing blind users with the desired freedom and ease of learn and use of the Web.

4 Rendering Web Content into Audio Equivalent

An independent mechanism has been prototyped to extract the content of webpages in a textual form and then transform it to one or more audio files of any desired format (e.g., mp3, wav). This way, developers for instance, may incorporate on their sites the produced "blind-friendly versions" of their webpages for end-user to be able to download them and listen to them through any standard media player. Furthermore, a text editor has been embedded in this mechanism to allow modifying and filtering the extracted content before producing the final audio artefact. Obviously, this is straight forward for static content, but an extension of the mechanism had to be implemented so that similar functionality could easily be integrated into more dynamic structures, such as portals, for producing the audio files at runtime.

The concept of transforming and storing an entire website in audio formats has many advantages, as it allows users to listen to Web content by using standard multimedia devices. However, the main limitation of this approach is that the transformed content is not fully equivalent as it looses the interactivity (navigation) provided by the actual Web interface through hyperlinks. To overcome this problem, the development of a specialized media player was considered.

[1] JAWS: http://www.freedimscientific.com/fs_products/software_jaws.asp

[2] HAL: http://www.dolphincomputeraccess.com/products/hal.htm

[3] BrooksTalk: http://www.brookes.ac.uk/speech

[4] Home Page Reader: http://www-306.ibm.com/able/solution_offerings/hpr.html

The *webLection* player, along with the aforementioned mechanisms and apart from the traditional features of media players (play audio, video, DVD, etc.), incorporates optional functionality for the transformation of Web content into its interactive equivalent in audio format. Overall, content can be saved either in xml files to be transformed in mp3 while browsing, or be exported directly to wav or mp3. Through this mechanism, users are able to download an entire site and browse it non-visually by following 'voice' hyperlinks. Additionally, standard player functions allow users to explore the saved content as ordinary video or audio files (e.g., pause audio stream, fast forward, next chapter etc). An important characteristic of the *webLection* player is that transformation of content occurs at the client computer, and therefore is not affected by the speed of Internet connections (the speed is vital in the case where the audio content is downloaded by the server hosting the Web application). webLection also supports multiple ways of interaction. Potential users of *webLection* are able to operate the application using standard software assistive technologies (screen readers, scanning software etc) or with the help of technologies embedded in webLection for non-visual interaction such as speech commands and synthetic voice output.

Finally, *webLection*, just like many media players, uses libraries. Typically, the library is designed to help users organize, or catalog, their music into categories such as genre, year, rating or other. In *webLection*, libraries work in a similar way, yet they take this notion a step by allowing the user not only to organise content but also to compile new artefacts by merging, modifying and editing existing content.

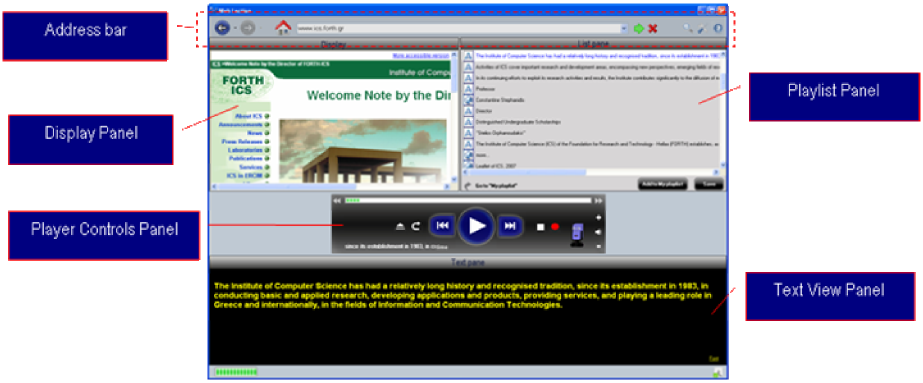
5 User Interface Design

In the design of *webLection* a user-centred approach was followed, based on iterative prototyping. Firstly, an initial interactive prototype was created, for feasibility purposes, and a cognitive walkthrough was conducted. Subsequently, a second interactive prototype was developed and evaluated by indicative user, confirming user interest and overall acceptability of the tool. Then, the final prototype was implemented.

Ultimately, the *webLection* player interface consists of:

- an *Address Bar*, which is used for address-based (and not hypelinks-based) navigation. In addition this bar provides access to the configuration editor of Weblection.
- a *Player Controls Panel*. This is the main component of webLection and has the same look-and-feel with a standard media player.
- a *Display Panel*, which is responsible for visualising content, including web pages, video, etc.
- a *Text View Panel* for visualising content in textual forms.
- a *Playlist Panel* which presents the underling decomposition of Web content into its primitive elements (to be further explained below).

All these, except the fist one, behave as windows that can be expanded (see Fig.1) or minimised (e.g., see Fig. 2) according to the needs of each individual user. For instance the configuration of the *webLection* presented in Fig.2 constitutes the default view for blind users, with the embedded text-to-speech facility being activated and input achieved through the use of keyboard shortcuts. Another example configuration is that with only the Display Panel being expanded, which has been identified as appropriate for web developers who wish to simulate easily how a website design would perform with users of assistive technologies such as screen readers. Similarly, the configuration in which the only expanded window is that of the Text View Panel is convenient for low vision users. As presented in Fig. 1, the page item in playback at given moment is concurrently presented as text employing a configurable setting (contrast, font type and size, etc.) to ensure high readability.



The Institute of Computer Science has had a relatively long history and recognised tradition, since its establishment in 1983, in conducting basic and applied research, developing applications and products, providing services, and playing a leading role in Greece and internationally, in the fields of Information and Communication Technologies.

Fig. 1. The *webLection* player (*fully expanded view*)

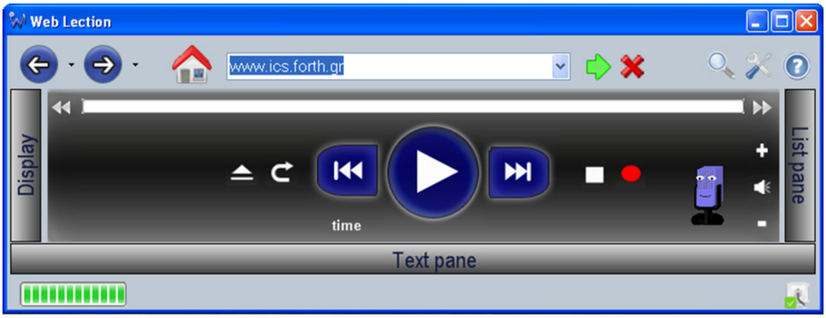
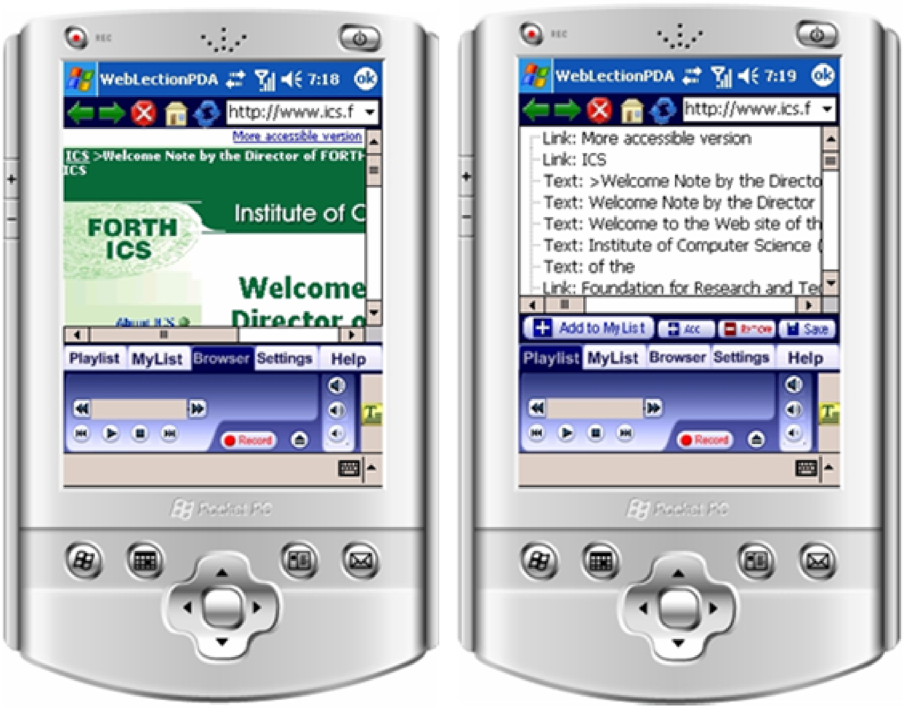


Fig. 2. The *webLection* player (*fully minimised view*)

During the final iteration, a first attempt was conducted to provide the *webLection* functionality to users on the move, for example to user of PDA devices. This effort has led to the development the first interactive *webLection* prototype for PDAs presented in Fig. 3.



Link: More accessible version
Link: ICS
Text: >Welcome Note by the Directo
Text: Welcome Note by the Director
Text: Welcome to the Web site of th
Text: Institute of Computer Science (
Text: of the
Link: Foundation for Research and Te

Fig. 3. The *webLection* player implemented for PDAs: Having the Display Panel expanded (*left*) vs. having the Playlist Panel Expanded (*right*)

6 Implementation

For the implementation of the prototype, XML was used for storing the document transformed to interactive or non interactive form. For achieving maximum efficiency of data storage and retrieval, strongly typed datasets were created according to the specifications set by the appropriate XSD schemas [3]. These strongly typed datasets are used by the content retrieval API or the content transformation API in order to retrieve or save content. For presenting this information, the user interface layer uses the content rendering API. For web content, the AxWebBrowser control [1] is used for rendering web pages and for extracting the DOM document tree used for transforming html pages to their saved interactive representations. Additionally, the managed direct X Audio - Video playback[5] API together with the Microsoft Speech SDK [10] and the lame[6] wav to mp3 converter are used for transforming and rendering content. The processes involved in the underling player infrastructures for saving content on the one hand and rendering on the other are presented in Fig.4.

[5] http://msdn.microsoft.com/library/default.asp?url=/library/en-s/directx9_m/directx/ref/microsoft.directx.audiovideoplayback.asp

[6] http://lame.sourceforge.net/index.php

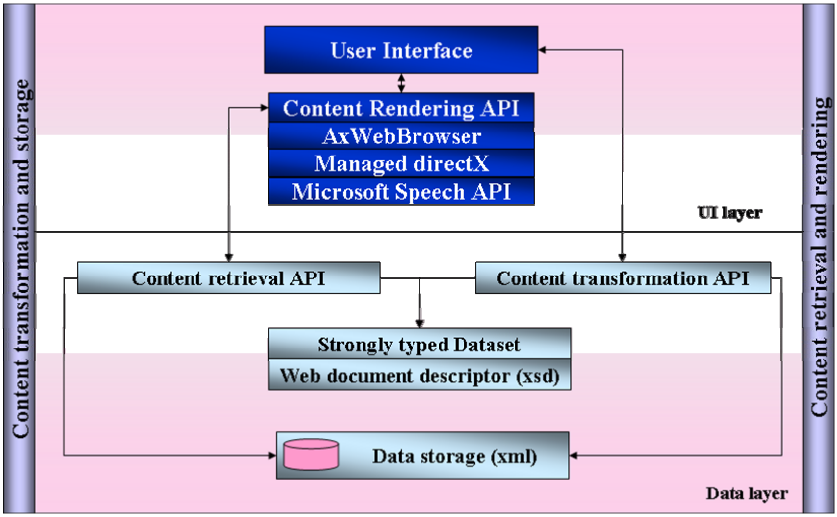


Fig. 4. webLection System architecture and implementation decisions

7 Conclusions and Future Work

This paper has presented an alternative approach for retrieving and reading Web documents. Further to the proposed approach, Web content is automatically rendered into standard audio formats), and manipulated through standard or special media players, just like music, by employing the playing list metaphor. A prototype tool, *webLection*, has been developed as a proof-of-concept in order to explore the involved concepts and design issues.

Although the prime target users of *webLection* are blind users, significant benefits may also emerge for other target user groups. In particular, the proposed approach is considered to be useful also for the typical, contemporary Internet user characterised by an increasing need for mobility and for interchanging between online and offline working modes. For example, a Web user can use the *webLection* tool in order to store locally (on his computer or on a standard mp3 player device) websites or individual Web pages as audio files and refer again to them once offline and on the move. However, this scenario restricts users into navigating through Web pages following their file system hierarchy predefined by the Web server. To address this limitation, and facilitate mobile, offline, hyperlink-based, and potentially non-visual navigation within a collection of Web pages, the *webLection* tool needs to be further enhanced in order to make it possible to install and operate it through PDAs and modern mobile phones. Furthermore, *webLection* can be particularly useful also for the very young, the elderly and the inexperienced Internet user, who often find it difficult to build effective conceptual models of the functionality of browsers or of the workings of the Web. In fact, *webLection* behaves like common players, rendering it easy to understand, learn and use. Users with impaired memory, who experience difficulties in exploration activities that involves remembering many combinations of actions and outcomes, are expected to find the use of *webLection* more friendly and less complex, due to its simplicity and its familiar and compact interface in the form of a common player console. Similarly, users with upper limp impairments who are finding user interfaces that require extensive mouse or keyboard usage prohibitive, will benefit from the configuration options of the system and minimise the amount of console's options down to the basics, and thereby minimise the amount of necessary key or special switches strokes required to read through a webpage or a Web site.

Finally, *webLection* may additionally serve as a handy tool for Web developers, since it can be used, on the one hand, to easily transform and make their artefacts also available through audio equivalents that can serve as 'printouts' for blind users, and on the other hand to inspect the accessibility and usability of Web pages for users of voice browsers and screen readers.

A potential direction for future work concerns the development of an improved version of webLection in order to take into account *User Agent Accessibility Guidelines* (UAAG 1.0) from W3C.

Acknowledgments. Part of the work reported in this paper has been carried out in the framework of the European Commission funded project ASK-IT[7].

References

1. Jones, A.: C# Programmer's Cookbook. Microsoft Press (2004)
2. Savidis, A., Stephanidis, C.: Developing Dual User Interfaces for Integrating Blind and Sighted Users: the HOMER UIMS. In: Proc. CHI 1995, Denver, USA, pp. 106–113 (1995)
3. Sills, A., Ahmed, M., Boumphrey, F., Ortiz, J.: XML.NET Develpoer's guide. Syngress Publishing Inc., Liberty (2002)
4. Stephanidis, C., Akoumianakis, D.: Usability Requirements for advanced IT products. In: Mital, A., Krueger, H., Menozzi, M., Fernandez, J.E. (eds.) Advances in Occupational Ergonomics and Safety, pp. 145–150. IOS Press, Amsterdam (1996)
5. Stephanidis, C., Paramythis, A., Karagiannidis, C., Savidis, A.: Supporting interface adaptation: the AVANTI Web-Browser. In: 3rd ERCIM Workshop on User Interfaces for All, Alsace, France (1997)
6. Treviranus, J., Serflek, C.: Alternative Access to the World Wide Web. University of Toronto (1996)
7. Wilkniss, S.M., Jones, M.G., Korol, D.L., Gold, P.E., Manning, C.A.: Age-related differences in an ecologically based study of route learning. Psychology and Ageing 12(2), 372–375 (1997)
8. Zajicek, M.: The construction of speech output to support elderly visually impaired people using the Internet. In: ICSLP 2000, vol. 1, pp. 150–153 (2000)
9. Zajicek, M., Powell, C., Reeves, C.: A Web Navigation Tool for the Blind. In: Proc. 3rd ACM/SIGAPH on Assistive Technologies, Los Angeles, USA (1998)
10. Zhuk, J.: Integration-Ready Architecture and Design: Software Engineering with XML, Java,NET, Wireless, Speech, and Knowledge Technologies. Press Syndicate of the University of Cambridge (2004)

[7] ASK-IT (Ambient Intelligence System of Agents for Knowledge based and Integrated Services for Mobility Impaired Users): http://www.ask-it.org