The "playing list" metaphor for web content access: Towards increasing mobility and interaction quality

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

[1] Institute of Computer Science, Foundation for Research and Technology – Hellas

Heraklion, Crete, GR-70013 Greece

Tel.: +30.2810.391741, Fax: +30.2810.391740

{mourouzi, partarak, cdoulger, cs}@ics.forth.gr

[2] Computer Science Department, University of Crete

Abstract

This paper presents *webLection*, a prototype tool that aims at increasing the uptake of the Web by a variety of potential users. *WebLection* implements an innovative approach, in which Web content is automatically rendered into standard audio formats, and, ultimately, manipulated just like music employing the *playing list* behaviour as an intuitive interaction metaphor. Although *webLection* was primarily inspired from research for blind Web users, significant benefits emerge also for other target user groups. In particular, the proposed approach is considered to be useful for the typical, contemporary Internet user characterised by an increasing need for mobility and for interchanging between online and offline working modes; 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; people with disability, in addition to the blind, such as users with impaired memory who experience difficulties in exploration activities that involves remembering many combinations of actions and outcomes, and users with upper limp impairments who are finding user interfaces that require extensive mouse or keyboard usage prohibitive. 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 Web content available through its audio equivalent, and on the other hand to inspect the accessibility and usability of Web pages for users of voice browsers and screen readers.

Keywords

Hypermedia; mobility; accessibility; usability; non-visual interaction.

1. Introduction

The rapid spread of the Internet and the increasing deployment of the Web as a prevalent medium for remote human interactions, such as information, communication and collaboration, call for practical solutions towards further widening the potential user base and contexts of use of the Web. As a fact, nowadays, the Web has emerged into a most valuable gateway to accumulated knowledge, information, updating news, tools and services of all kinds, rendering indisputable the need for accessibility and high-quality interaction to anyone (Stephanidis et al., 1996), including blind users and people with visual impairments.

Currently, access to the Web by blind users is feasible mainly through special technologies, such as screen readers or voice browsers. This work proposes an alternative approach, in which Web content is (semi-) automatically rendered, stored and, ultimately, offered in audio format. 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 to the end-user. For example, it is argued that inexperienced blind users will find it easier and more satisfying to handle media players and audio files (or playing lists), rather than browsers, screen readers and Web pages.

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 his web mailer), loads them to his MP3 player (the same device he has bought 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. Furthermore, this approach offers the possibility to store the produced audio files and work offline, thus serving blind users also just like a printouts serves sighted users.

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, focus has been on two complementary directions:

(a) *Provide the functionality for rendering Web content into audio equivalent*. The motivation behind this direction is 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.

(b) *Integrate a voice player that allows hyperlinks-based, non visual interaction and navigation*. The motivation behind this direction relies on the fact that navigation through hyperlinks is more comprehensive than navigation based on the file hierarchy of Web pages (as dictated by their URLs), since hyperlinks, between them, control how the reader perceives, and/ or should perceive, the overall structure, and how the traffic flows between the different parts of the sites.

Although *webLection* was primarily inspired from research for blind users, significant benefits emerge for all, and the potential of increasing the uptake of the Web by a variety of potential users is offered. In particular, the proposed approach is considered to be useful for: the typical, contemporary Internet user characterised by an increasing need for mobility and for interchanging between online and offline working modes; 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; people with disability, in addition

to the blind, such as users with impaired memory who experience difficulties in exploration activities that involves remembering many combinations of actions and outcomes, users with upper limp impairments who are finding user interfaces that require extensive mouse or keyboard usage prohibitive. For instance, in the above use case example, a sighted user instead could 'read' the news and emails while driving to work. At last, but not least, webLection may additionally serve as a handy tool for Web developers, since it can be used to inspect the accessibility and usability of web pages for users of voice browsers and screen readers.

The next section of this paper 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. The paper concludes with an expert-based evaluation and discussion of the implemented features of the tool, leading to design recommendations for improving the tool and to an outline of the authors' plans regarding related future work.

2. Background and Related Work

2.1 Brief reference to related features of the Internet and the Web

The Internet is a worldwide, publicly accessible collection of interconnected domestic, academic, business, and government networks, which together carry various information and services, such as electronic mail, online chat, file transfer, and the interlinked Web pages and other documents of the World Wide Web ("WWW" or simply the "Web"). The Web constitutes a collection of interconnected text documents, images, multimedia and many other items of information, referred to as resources. Resources are identified by short, unique, global identifiers called Uniform Resource Identifiers (URIs) so that each can be found, accessed and cross referenced in the simplest possible way. A Web page is a resource of information that is suitable for the World Wide Web and can be accessed through a Web browser. Web pages may be stored on a local computer or on a remote Web server. When Web pages are stored in a common directory of a Web server, they constitute a website. A website typically contains a group of Web pages that are linked together via hypertext links (hyperlinks), or have some other coherent method of navigation. Some websites require a subscription to access some or all of their contents. Examples of subscription sites include parts of many news sites, gaming sites, message boards, Web-based e-mail services, and sites providing real-time stock market data. Web pages may consist of files of static text within the Web server's file system (static Web pages), or the Web server may read files of computer code that instruct it how to construct the (X)HTML for each Web page when it is requested by a browser (dynamic Web pages).

A website can contain interactive elements and content which is able to be seen or heard by the user. These include, but are not limited to, menus, text entry fields, hyperlinks, text, graphics and images, audio and video, Java applets, and Macromedia Flash and Shockwave. A website can also contain content which is interpreted differently dependent upon the rendering browser and is typically not shown to the end user. These elements include, but are not limited to, scripts, meta tags, Cascading Style Sheets (CSS), and comments. When creating a web page, it is important to ensure it conforms to the World Wide Web Consortium[1] (W3C)

[1] W3C: http://www.w3.org/

standards for HTML, CSS, XML and guidelines for Web content accessibility[2]. The W3C standards and guidelines are in place to ensure that all browsers which conform to their standards can display identical content without any special consideration for proprietary rendering techniques. A properly coded web page is going to be accessible to many different browsers old and new alike, display resolutions, as well as to users with motor, audio or visual impairments who use alternative access systems.

2.2 Current approaches to Web accessibility for blind users

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 (Savidis & Stephanidis, 1995). Along the same lines, attempts to provide accessibility in the Web environment 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 (Treviranus & Serflek, 1996). For instance, access to the Web by blind users is currently made possible mainly through special technologies, such as screen readers (e.g., *JAWS*[3] and *HAL*[4]), or special Web browsers (e.g., the *BrookesTalk*[5] browser for visually impaired users, Zajicek, 2000; *Home Page Reader*[6], IBM's browser for users that are blind or have low vision). Following a more generic and systematic approach, AVANTI (Stephanidis et al., 1997) was the first universally accessible web browser, addressing the requirements of able-bodied, blind and motor impaired users. However, these solutions alone tie users to their PCs at home or their office, while increasing mobility creates the requirement for more freedom and the ability to travel around taking music, movies (with enhancements for the blind) and text to read on the way.

Furthermore, 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., Zajicek, 2000). 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 (Zajicek, Powell & Reeves, 1998). 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. Blind, but also many elderly, users find difficulty in remembering sequences of actions previously performed (Wilkniss et al., 1997), and they often 'borrow' the model of the video recorder expecting one press of a button to make everything happen (Zajicek, 2000).

2.3 WebLection vs. current approaches

Under the light of the above, *webLection* is aimed at providing blind users with the additional desired freedom and ease of learn and use of the Web. In fact, blind users are reportedly keen users of media players (software) and mp3 player (devices). Furthermore, as the proposed approach offers the possibility to store the audio equivalent of web content and work offline,

[2] W3C's guidelines for Web content accessibility: http://www.w3.org/WAI/intro/wcag.php
[3] JAWS: http://www.freedimscientific.com/fs_products/software_jaws.asp
[4] HAL: http://www.dolphincomputeraccess.com/products/hal.htm
[5] BrooksTalk: http://www.brookes.ac.uk/speech
[6] Home Page Reader: http://www-306.ibm.com/able/solution_offerings/hpr.html

even by means of traditional MP3 players, it could serve to blind users just like a printer friendly version serves to sighted users. Thus the employment of the media player behaviour and the 'playing list' metaphor is perceived as really promising in terms of its intuitiveness and usefulness for searching and navigating the Web.

3. Design

3.1 Functionality for rendering Web content into audio equivalent

3.2.1 Static Web pages

A prototype mechanism has been designed to assist, on the one hand, end-users in downloading the audio alternatives of web pages and, on the other hand, developers in incorporating blind friendly versions of pages to their web sites. The mechanisms's main function is to extract the content of each page and transform it to one or more audio files in a desired light format (e.g., mp3, wav). Furthermore, this mechanism provides an embedded text editor for making the desired changes to the extracted page content in order to filter out content of low interest.

3.2.2 Dynamic Web pages

Obviously, this approach works fine with static content. An extension of the mechanism has been considered so that it can be easily integrated into a portal. In such cases, the audio files are automatically created by reproducing the *webLection* functionality. Additionally, users of the portal need to be able to download the content either of the entire portal or of a specific part. For example, a user may request the contents of the message board facility and the portal framework automatically creates and delivers the requested audio files.

3.2 Hyperlinks-based, non visual interaction and navigation

The concept of transforming (and downloading) an entire website in audio format has many advantages, as it allows users to browse Web content offline, by using standard multimedia devices. However, the main drawback of this approach is that the saved content looses the interactivity (navigation) provided by the actual Web interface through hypelinks. To overcome this problem, the development of a specialized media player was considered.

webLection, apart from the traditional features of media players (play audio, video, DVD, etc.), incorporates optional functionality for the transformation of a website into an interactive equivalent in audio format. More specifically, according to the functionality chosen, content can be saved either in xml files to be transform in mp3 while browsing or be exported directly to wav or mp3. Through this mechanism, users will be able to download an entire site and browse through non-visual interaction and by following 'voice' hyperlinks. Additionally, the standard player functions allow users to use the saved content as an ordinary video or audio file (e.g., pause audio stream, fast forward, next chapter etc). Another important characteristic of *webLection* is that the transformation of Web 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).

Concluding, a very important issue that needs to be investigated concerns the support for multiple ways of interaction. Potential users of the *webLection* are able to operate the application using standard software assistive technologies (screen readers, scanning software

etc) or with the help of embedded technologies for non-visual interaction (including speech commands to perform actions and speech output to identify available options).

3.3 Preliminary User Interface

Figure 1 presents the interface provided by *webLection* for rendering web content. This interface can be divided into the following main areas:

1. Content: This area is responsible for rendering visual content such as web pages, media files, etc.
2. Playlist: In this area the list of items representing the currently browsed web content or the list of media files to be rendered are displayed.
3. Track info: Provides information about the track currently browsed.
4. Functionality: From this area the traditional media player functionality is offered.
5. Web saving facilities: Facilities for saving sites in various interactive or non – interactive formats.

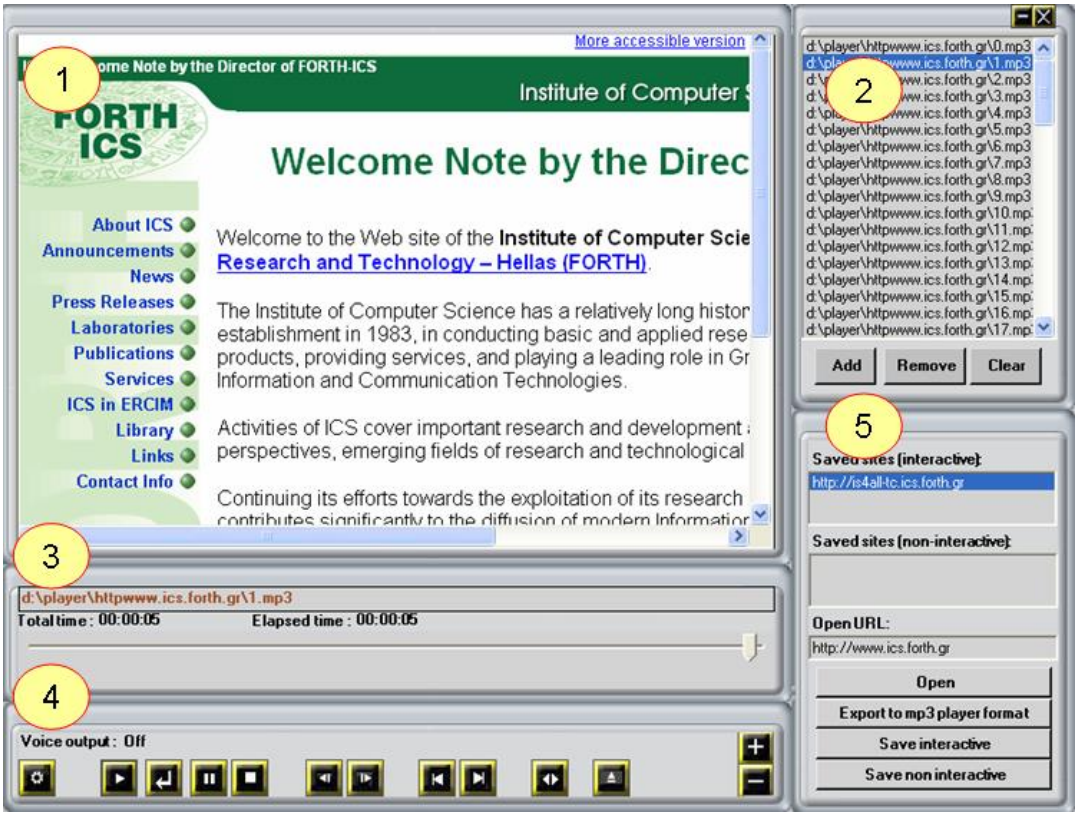

Figure 1. Screenshot from the first prototype of the webLection tool

4. 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 (Sillis, Ahmed, Dotthatcom.com, Boumphrey & Ortiz, 2002).

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 (Jones, 2004) 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[7] API together with the Microsoft Speech SDK (Zhuk, 2004) and the lame[8] 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 Figure 2.

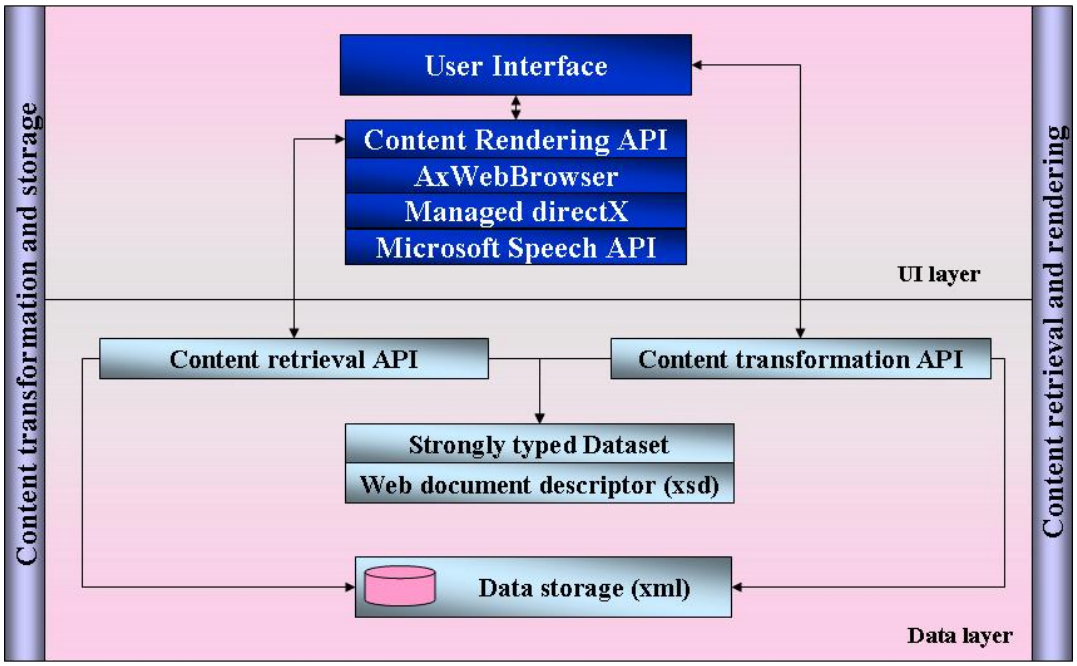

Figure 2. System architecture and implementation decisions underlying the webLection prototype

5. Informal Evaluation and design recommendations

A preliminary heuristic evaluation has been conducted by an expert evaluator on the presented version of the *webLection* prototype. Below are the most important design issues and recommendations that emerged from the evaluation and which will be taken into account into the final version of the tool:

- Consider altering the order of functionality in conformance with the conventions followed by most media players.
- The buttons assigned to the functions for accessing player settings and switching to full screen mode do not clearly depict the underling functionality provided. In the case of full screen switching the positioning should also be changed.

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

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

- The player's functionality must be placed below the bar displaying tracks.
- Enable users to cancel functions especially in the cases of saving or exporting web sites.
- Inform users for the successful completion of any function
- Inform user about the specified sticky keys and if possible alter them in order to conform to commonly used conventions.
- Provide context sensitive help.
- Clearly mark the location of the main areas provided by the player such as the area displaying the current play-list.

6. Conclusions & 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 (by employing synthetic voice mechanisms), and, ultimately, manipulated through standard or special media players, just like music, by employing the playing list behaviour as an intuitive interaction metaphor. A prototype tool, *webLection*, has been developed as a proof-of-concept in order to explore the involved concepts and design issues. Expert-based evaluation of an early interactive prototype of *webLection* has brought forward some design issues and recommendations for further improving the usefulness and usability of the tool for all target users. 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 that can support light applications (such as media players).

Furthermore, *webLection* can be proved to 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.

Regarding future work, the development of an improved version has been planned in order to incorporate the design recommendations that emerged from the aforementioned preliminary inspection of the prototype. For the next version of the tool, the *User Agent Accessibility Guidelines* (UAAG 1.0[9]) from W3C will be also taken into account as they explain how to make user agents, including Web browsers and media players, accessible to people with various disabilities. Finally, upon the implementation of the improved version, user testing will be conducted with all the aforementioned target user groups.

Acknowledgments

Part of the work reported in this paper has been carried out in the framework of the European Commission funded project ASK-IT[10].

References

Jones A. (2004), C# Programmer's Cookbook, Microsoft Press

Savidis, A., and Stephanidis, C. (1995). *Developing Dual User Interfaces for Integrating Blind and Sighted Users: the HOMER UIMS.* In Proc. CHI '95, Denver, USA, pp 106-113.

Sills A., Ahmed M., Dotthatcom.com, Boumphrey F., Ortiz J. (2002) XML .NET Develpoer's guide, Syngress Publishing Inc. Liberty

Stephanidis, C., & Akoumianakis, D. (1996). *Usability Requirements for advanced IT products*. In Mital, A., Krueger, H., Menozzi, M. and Fernandez, J.E. (eds.). Advances in Occupational Ergonomics and Safety. IOS Press, pp 145-150.

Stephanidis, C., Paramythis, A., Karagiannidis, C., & Savidis, A. (1997) *Supporting interface adaptation: the AVANTI Web-Browser*. 3rd ERCIM Workshop on User Interfaces for All, Alsace, France.

Treviranus, J., & Serflek, C. (1996). *Alternative Access to the World Wide Web*. University of Toronto.

Wilkniss, S., M, Jones, M., G., Korol, D., L., Gold, P., E. & Manning, C., A. (1997) *Age-related differences in an ecologically based study of route learning*. Psychology and Ageing 12(2), pp 372-375.

Zajicek, M. (2000). *The construction of speech output to support elderly visually impaired people using the Internet*. In ICSLP 2000, Vol. 1, pp 150-153.

[9] UAAG 1.0: http://www.w3.org/TR/UAAG10/

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

Zajicek, M., Powell, C., & Reeves, C. (1998). A *Web Navigation Tool for the Blind*. In Proc. 3rd ACM/SIGAPH on Assistive Technologies, Los Angeles, USA.

Zhuk J. (2004), Integration-Ready Architecture and Design: Software Engineering with XML, Java, .NET, Wireless, Speech, and Knowledge Technologies, Press Syndicate of the University of Cambridge