Web Access Technology for Visually Impaired Users

A White Paper
Introduction

The web currently holds over 500TB of data spread over some 50 million web sites and accessed by over 300 million people worldwide. 28 million people in the UK, 151 million in the USA and almost 1 billion people world-wide are using the Internet. And the web is an incredibly visual medium with pages built from an almost dizzying selection of colours, fonts, images, animations and even video.

However, the American Foundation for the Blind (AFB) estimates that there are 10 million Americans who are blind or have impaired vision. In the UK 395,000 people are registered as blind or partially sighted. The RNIB reckons that a further 750,000 people are eligible to register but haven't and a total of 2 million who have some sort of a sight problem. Worldwide it is estimated that 45 million people are blind, and 180 million people are visually impaired. The RNIB estimates that 1 in 12 of us will become blind or partially sighted by age 60, and 1 in 6 by age 75.

For blind and visually impaired people the web represents a real challenge. Being able to use the web has become such a key part of modern life that not to have access places them at a considerable disadvantage and builds a significant digital divide.

Whilst a number of technical solutions are already in use, all have their issues, and there may even be a fundamental flaw with the model on which they are based.

This paper examines the general issues around web accessibility for visually impaired users, looks at some of the existing technology solutions, presents a new technical approach, and finally looks at a fundamentally different model that could provide improved web accessibility in the future.

The Challenge

The fundamental challenge for visually impaired users is how to make use of a very graphic medium with little or no sight.
It should be noted that most blind users have perfectly good motor skills and so the use of a keyboard – particularly if marked in large letters or with braille labelling is not an issue. Even mice can be used if sufficient audio prompts are provided by the system. This means that whilst a lot of the focus of this paper is on text-to-speech systems, there is not an over-riding need to provide systems which use speech recognition to input commands.

A distinction also needs to be drawn between the needs of partially sighted and blind users (including those with no useful vision). For the former, solutions based around magnified text and improved contrast may be sufficient to make a web site usable. For the latter, the solution needs to use text-to-speech technology. It is the needs of completely blind users that are the focus of this paper.

Accessibility

Web accessibility has become a significant issue over the last few years, and several organisations have been involved in bringing it to the attention of web designers around the world.

The Web Accessibility Initiative is a global undertaking and a part of the World Wide Web Consortium. WAI has produced official guidelines for web accessibility which lay out the standards which a web site must meet in order to be deemed accessible (see left).

A number of software services are available to audit sites against this standard – the most well known one being Bobby. This will automatically review a web site and provide a report showing where it fails to meet the WAI standards.

The RNIB together with the Disability Rights Commission is currently working with the British Standards Institute in order to develop a Code of Practice for UK web developers.

In both the UK and US legislation is in place which places an obligation on some or all companies and organisations to make their services accessible. In the UK it is the Disability Discrimination Act (and the key undertaking came into force in October 1999), and in the US it is Section 508 of the Rehabilitation Act.
- Creating Accessible Web pages

Web pages can be made accessible in three broad ways.

- The web designer can use any conventional web design tool and manually ensure that the designed pages follow the accessibility guidelines

- The site can be designed from scratch on an accessibility aware content management system (such as Amexus or Libertas) which will enforce disability rules

- Almost any site can be served through a parser (such as WiderWeb) which reformats pages on-the-fly into an accessibility compliant format.

- Levels of Accessibility

Despite this activity progress in web accessibility has been slow. A study by DRC in 2004 of 1000 web sites found that 81% fail the most basic WAI accessibility category, only 19% were potentially level A compliant, 0.2% were AA compliant, and none achieved AAA! A survey in 2005 by Nomensa found that of 33 London local authority web sites, 85% have not met the target AA rating, only 6% are at AA, and 9% at AAA.

Current Approaches

There are three main approaches to converting web pages into speech.

- Screen Readers

Screen Readers are typically software programs installed on a user's PC which convert the text on screen into speech. They can read out details of screen icons and menus, the contents of any application windows and can allow the user to use almost any application on the PC, including a web browser. When combined with a speech recognition system they allow the user to have a completely voice-based computing system. For visually impaired users JAWS is the most well-known screen reader.

However, screen reader software needs to be installed on each PC that the user will access, and the installation process itself needs sighted help. The quality of voices shipped with most screen-readers also leaves much to be desired, and they typically have a very complex set of commands. There is no doubt that screen-readers are powerful tools – but that comes at

Top Accessibility Hints

How do you make a web site accessible? Here are some hints:

- Use the ALT tag to describe images.
- Use good contrast between text and background
- Use style sheets to control appearance
- Ensure that link text makes sense out of context
- Provide a text based site map
- Ensure that text size can be controlled by the browser
- Don’t use Frames
- Avoid absolute positioning
- Warn users about new windows and pop-ups
- Make sure that all multimedia elements are optional and provide a text alternative
- Warn users when linking to non-HTML resources
- Ensure that the page still makes sense if tables are flattened (i.e. the page is read row by row)
both a financial cost (JAWS costs $895), and at the cost of complexity.

- Point and Say

Several web based applications exist (such as BrowseAloud and ReadSpeaker) which work on a point and say model. Here the user selects a region of text in the web browser with the mouse, and the application then converts this to speech. Whilst ideal for users suffering from dyslexia or reading difficulties, or maybe even low vision, they are not ideal for blind users – although some do offer some mouse-free reading.

The technical approaches taken by such packages are quite different. Some (such as BrowseAloud) are downloadable applications or browser plug-ins, and the text-to-speech conversion takes place on the user's PC. However text-to-speech quality can be quite poor depending on the voices installed on the client PC, and the download and install process can take up to 5 minutes and needs sighted help – or a screen-reader!

Others (such as ReadSpeaker) take an ASP approach. The server takes the target segment of the web page as indicated by the user, converts it to audio, and then sends the audio file to the user. Being ASP based, these applications do not require a download or install, beyond that usually found already on PCs to play the sound files.

- Parsers

The third approach is that of the parser. A parser takes a whole web page, rather than a designated part, and reformats it (eg flattens tables) into a single simple text string ready to be passed to a text to speech system. To enable the user to move back and forward through the page the parser will often break the string into a series of blocks, and the user can then use keyboard keys to move back and forth as with a set of CD or VCR controls.

Text-to-Speech

Crucial to any system for blind users is the text-to-speech component. The conversion can be carried out either on the user's PC (as with BrowseAloud) or on a central server (such as ReadSpeaker and ReaderBot – our own solution).

The text-to-speech system will typically have one or more of the following features:
A range of voices, including different sexes to cater for users' tastes
- Support for a range of languages (e.g. an English TTS reading text written in French will be almost unintelligible – a French voice is needed instead)
- The ability to change the voice's pitch and/or speed
- An optimisation capability to ensure that the text is read well (e.g. handling numbers and telephone numbers, abbreviations, apostrophe'd words, non-phonetic words etc.)

There are also a number of approaches to TTS technology itself, from synthetic voice boxes to pre-recorded snippets of speech.

**Voice Recognition**

Although all of the assistive technology described is designed to work with keyboard input, some can also be controlled by speech – either through embedded functionality, or through the addition of another application such as Dragon Naturally Speaking. Such a system will then allow the complete web experience to be managed through voice alone.

**An ASP Approach**

Advanced Chatbot Solutions (ACS) is a Daden Limited technology venture working with chatbot, avatar and artificial intelligence technology. While the main focus of activity had been on developing talking chatbots using the Oddcast Sitepal avatar and TTS system, it became clear late in 2004 that the same technology could be used to deliver a new breed of virtual screenreader – the ReaderBot product.

Whilst accepting that the “power user” would probably want to use an application such as JAWS, discussions revealed that there was an opportunity for a web based screen reader for use either by web sites to make speech versions of their sites available (Site ReaderBot), or by individuals to surf the web (Personal ReaderBot).

The system works on the parsing approach and is delivered as an ASP, with the TTS taking place on the central Oddcast server. When the user clicks on a link to get a page as speech the ReaderBot server collects the target page, parses it, and breaks it down into a series of paragraph sized blocks. It also collects all the

---

**RSS**

Rich Site Summary (RSS) is an XML dialect for news feeds provided by web sites. It evolved out of the blogging movement to enable bloggers to inform others of their new posts, and to be able to syndicate their content out, and other people's content in.

RSS has a simple `<item>` format with no extraneous navigation information (and, until recently, no advertising), which makes it easy for machine parsing, such as for format translation for mobile or speech use.

For more information on RSS see:
- The RSS specification at: [http://blogs.law.harvard.edu/tech/rss](http://blogs.law.harvard.edu/tech/rss)
links on the page and reformats them to point via the ReaderBot server. The server then sends a new page to the user which contains the blocks, links and commands for the speech system – but which actually need contain no visible text at all. The user's browser then follows the commands on the page to send the first piece of text to the TTS server, which then sends the resultant audio back to the user. Although the process sounds complex it only takes a few seconds to start hearing the first audio, and subsequent blocks of audio come quicker still.

The system only requires that the user has Flash on their PC – which most PCs do anyway. Since Flash audio is used rather than streamed audio the system can work through almost every firewall.

The user navigates the page using simple single key commands - such as SPACE BAR to move onto the next block, L to access links and HOME to go to the start of the page.

ReaderBot also automatically detects skip to content, skip to navigation, and site-map land text-only links to ease navigation.

**Site and Personal ReaderBot**

Whereas most speech systems are designed for use either by web site owners, or by individual users on their own PCs, ReaderBot can be used in both modes.

Site ReaderBot can be licensed by a web site and configured to provide access to only their site. Audio announcements and a clear visual logo can indicate its presence, and the user just needs to press ALT-R to start the audio system.

Personal ReaderBot is designed for use a) by power users when away from their JAWS PC – e.g. at a colleague's desk or cyber-cafe, b) by more average blind users who needed occasional web access, and c) by clubs and libraries where any PC can provide access to the system without special software being installed.

For more information on ReaderBot please read the ReaderBot User Guide.

**Phone not Web**

An alternative approach to web based accessibility is to make use of the telephone. Just as a parser such as ReaderBot can take an HTML page and reformat it ready to be read by a web-based text-to-speech
system there is no reason why it can't reformat it ready to be used over the telephone.

What makes this possible, and affordable, is a protocol called Voice XML (VXML). As part of the XML family, VXML is almost as easy to write and deploy as HTML, and VXML hosting services exist which enable companies to offer VXML based telephone services but without having to invest in expensive telecommunications lines or IVR systems.

Whilst dedicated VXML based web to telephone systems have been around for a few years, ReaderBot is unique in being able to use the same parser for both telephone and web access – so helping to maintain a consistent user interface, and also ensuring that costs are only incurred once.

PhoneBot – the ReaderBot phone service, will take any ReaderBot enabled site or service and deploy it to the end of a phone line. The user then just dials the phone number (typically an 0870 number), and is greeted by a welcome message for the site. They can use the telephone keypad to navigate through the site in the same way that they would use the PC keyboard in the web version of ReaderBot. It is also far easier to implement voice commands on the phone system. Potentially this can bring the web to an even wider visually impaired audience, as users no longer even need a PC or Internet access - they just need a telephone.

Applications

Whilst the most common implementation of web access technologies will be in providing access to web based information, there are a wider array of web services and applications that could benefit from accessible access. Different access technologies will have different levels of success with each.

- **e-Commerce**: The ability to buy things over the Internet can be a significant boon to visually impaired users. Seeing as most e-commerce systems run in a secure environment, and that users may be reluctant to have their account details passed through a third party system, PC based accessibility solutions such as JAWS are probably the more preferred and viable route for e-commerce at the moment. However, combining audio browsing of the web site with a conventional telephone orderline, a visually impaired user can gain most of the convenience of on-line shopping.

- **Web Services**: Many applications on the web (e.g. Amazon, Google) are now making web services
available to their applications. This means that third party developers can develop bespoke interfaces to the systems for specific audiences. At ACS we have used the Amazon interface to produce a prototype “Amazon by phone” service, and used the Google interface to provide an accessibility specific search engine for ReaderBot. Using web service access to an application can provide a far more usable result than simple screen-scraping, but needs specific development for each system.

- **Community Systems:** Much of the value of the web lies in the person-to-person communications, whether it be by email, bulletin board, chat forum or web log (blog). Again a PC based screen-reader will enable access to these systems out of the box – although usability may be an issue unless an accessible aware application is being used (e.g. AccessibleMail). With all these systems, getting information out is a lot easier than putting information in. Our NewsBot product is designed to allow web log owners to use their RSS news feeds to produce a spoken version of their most recent postings. ACS also has phone and web audio based email readers. The way forward here to producing a good quality interface for visually impaired users is probably through the use of web services as they become available for each system to create a VI specific interface.

In addition to these conventional uses of the web and speech based technology, there are a number of other areas where development could be of significant benefit to visually impaired users.

- **Talking Newspapers:** Just putting an application like ReaderBot in front of a newspaper site will instantly create a talking newspaper – and in real-time, unlike taped talking books. The challenge, though, is in making the format portable – in the same way that a tape can be played from a Walkman. Two approaches seem possible. One is to use PhoneBot – so the user can dial in to the newspaper from anywhere – but this has both cost and coverage issues (e.g. on the tube). A second approach is to combine TTS with podcasting technology, so that set pages are converted to speech and then delivered to an MP3 based portable audio device.

- **Talking Books:** Just as ReaderBot type technology can convert newspapers to speech it can do the same with electronic books. Project Gutenberg (www.gutenberg.org) currently provides free access to over 15,000 books, mostly out of copyright classics. ReaderBot can instantly make these available to a blind user.
**Consumer Information:** New legislation is likely to force manufacturers to make all their consumer information available in an accessible format. Producing this as Braille is likely to be prohibitively expensive. With ReaderBot and PhoneBot such information can be published on the web using any content management system, and then accessed as speech over the web or phone for no extra effort.

**Home Automation:** Although some set-top boxes are beginning to offer accessibility features, providing speech versions of Electronic Programming Guide (EPG), sites could make TV listening a far easier process for blind users. The next challenge would be how to integrate EPG with VCR/DVR/PVR recording technology, and even with home media servers. Whilst this may be an area that is better served by PC based solutions (e.g. Windows Media Centre with speech), it could also be possible to implement a more centrally managed system which would minimise the problems of managing a system when it failed to perform as expected (e.g. the PC crashes).

**Games:** One of the biggest social changes of the last few years has been the rise of video games – but by their very nature they tend to exclude visually impaired users. Whilst one can't currently expect TTS technology to provide audio-described versions of Halo or The Sims, the technology used to provide a service such as ReaderBot can be used to either speech enable the many text based games found on the web (particularly the older MUDs and MOOs – Multi-User Dungeons and MUDs Object-Orientated!), or to develop audio specific games. In fact, ACS has already done some early work on a XML based language to create audio adventure games. And as The Sims shows, there is no reason why these Game packages shouldn't be used for more social and “serious gaming” applications.

**Chatbots:** ACS’ original, and still continuing, area of research is into chatbot based systems. These are computer programmes that allow you to have a typed or spoken conversation with a computer programme, and try to mimic the responses of a human. Apart from providing an alternative access technology to web information (see below), such chatbots could have real potential as personal assistants – and even virtual companions – for visually impaired users. A lot of research is being done in Japan on the role of robots in the social care of the elderly – and this again could have implications for VI users.

**Graphics:** All the standard web accessibility tools have the same approach to graphic and image content – they just read out the ALT tag. But wouldn't it be good if the application could actually describe the image? Whilst this may not be viable
for many years for your typical photo, it could be possible for charts, diagrams and even maps. The key technology here is SVG – Scalable Vector Graphics. Like all good modern technologies it is based on XML, and provides a way of creating an image using a set of “primitives” such as lines, squares, circles etc. Whilst a browser would render an SVG file as a conventional picture, it would be possible for a parser based system like ReaderBot to look into the SVG code and describe the image. We are currently working on a concept demonstrator around simple graphs to show the SVG to TTS technology in action.

The 1D-2D problem

Whilst we are proud of the capabilities of our ReaderBot application, and never cease to be amazed by the ability of the trained JAWS reader to get through web pages faster than we can, we do have a nagging concern that there is a key issue that isn't being addressed.

At its core it is this; the web is a two-dimensional medium, speech is one dimensional. All screenreader type applications, ReaderBot included, have to take the 2D information and try to present it as a linear stream of speech. The power of the 2D interface for a sighted person is that they can scan a whole load of information, and then pick out the bits that they are going to actually read – a 1D process. The visually impaired user does not have this luxury – they have to go through the page word by word – although applications do try and help by having skip forward/back and jump-to-content type links.

Perhaps we need a more fundamental review of how we provide visually impaired access. Two routes appear possible. One is to put some Artificial Intelligence into the screenreader that can do the “scanning” for the user – based on their likes and dislikes – and flag the items that are likely to be of interest. The second is to do away with the literal screen reading totally and instead focus on the chatbot approach described above – creating virtual assistants for each site who can just converse with the user in a one-dimensional way, and tell them what they want to know. At ACS we plan to investigate both avenues – whilst using ReaderBot to provide a service today until these newer approaches have matured.
The Future

Apart from the 2D-1D problem, we believe that this is an exciting time for accessibility technology, and indeed for visually impaired users. The Internet in all its forms is having a significant effect on the lives of sighted users, and for two long most visually impaired users have been on the wrong side of the digital divide. Modern web based access technology, in all its forms, is now helping not only to bridge the digital divide, but also to provide tools that can be used to develop dedicated services for visually impaired services in a very cost-effective manner.

In fact, the challenges of providing access for VI users are, in some areas, so great that the benefits of achieving them could be felt well beyond visually impaired users. We've all seen Star Wars and Star Trek, we all know that in the future we will talk to computers, and they will talk to us. Perhaps the future will come to visually impaired users first?

Conclusion

Hopefully this document has given you some idea of the current state-of-play of web accessibility technology, and in particular how our ReaderBot and PhoneBot technology fits in – and where things may go from here.

If you would like more information on the topics or technologies discussed, or to discuss our specific ReaderBot, PhoneBot and NewsBot solutions, then please contact us by any of the means detailed below. We would be happy to provide live demonstrations of either of the systems described, or of our wider web services, chatbot or VXML capability.

Who Are We?

Daden Limited is a consulting and venture business based around exploiting advanced technology. Our current venture business is Advanced Chatbot Solutions (www.chatbots.co.uk), focussing on speech, VI, avatar and chatbot technologies. We have been working with high technology for over 20 years. We have a deep understanding of the possibilities offered by new technologies, but with a good knowledge of the marketing and usability issues to help work out what might work, and sell, and what won't.