Enabling Accessibility and Enhancing Web Experience: Ordering Search Results According to User Needs

Alice Good and Jenny Jerrams-Smith

Department of Information Systems and Computer Applications
Burnaby Terrace, University of Portsmouth, UK
alice.good@port.ac.uk, jenny.jerrams-smith@port.ac.uk

Abstract. This paper presents the results of an exploratory study, which aimed to devise a means of improving web accessibility for users with impairments. It looks at how web pages can be rated for accessibility using specific algorithms. These algorithms analyse and select web pages according to user preferences. The study gathered user defined ratings of web page elements, for different categories of special needs' users that affect accessibility. The algorithms were then created using the results of these ratings. The paper also demonstrates the work in progress in relation to the extension of the Computer Aided Internet Navigation (CAIN) system, which aims to improve accessibility to web based information by re-ordering search results according to user needs.

Keywords: Accessibility; special needs; web navigation; algorithms.

1 Introduction

The purpose of this study is to focus upon addressing the issue of poor accessibility, which affects people with special needs. These needs explicitly relate to people with limitations that affect the ease in which they are able to access web based information, including older adults with age related impairments. The limitations considered here include all types of visual and motor restrictions as well as dyslexia. The scope of this study was focused on the more common impairments and hence other cognitive and language impairments will not be considered.

The strive towards accessibility has been a long-standing, major concern for representatives of special user groups and recent legislation has contributed to the importance of ensuring that these users are enabled and indeed have a right for equal access to information. Buzzwords such as 'Inclusive design' and 'universal access' are increasingly becoming omnipresent in the world of web design, along with numerous sets of guidelines and tools to assist and advise designers in ensuring that potential users are not excluded. As Nielson states: *"It has gotten much easier to advise people on making it possible for users with disabilities to use a website: just follow the official Web Accessibility Initiative Standard (WAI) from the World Wide Web Consortium (W3C)."* (Nielson, 1999, Jakob Nielson's Alertbox). The development of assistive and adaptive software have certainly provided better access to information for this user group though, in terms of equal access, there is no comparison to non-disabled users (Stephanidis, 2001; WAI, 2003; DRC, 2004). There

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 34–44, 2007.

is a general consensus that poor design and incorrect coding are key factors in the cause of problems related to accessibility (Nielson, W3C Schneiderman,).Much of the problem is also said to be a result of a lack of consideration to the entirety of user needs being incorporated in the design process; poor design then leads to poor usability (Nielson, 2003; Schneiderman,1999; Preece et al, 2000) The ever-increasing numbers of elderly people contribute to the scale of the problem given their susceptibility to similar disabilities such as reduced vision and mobility impairments. The problem of inaccessibility is not an isolated issue (Eklund & Ziegler 1996; Marshall 1995), other related or resulting factors include digital isolation, social exclusion and disempowerment (NTIA, 1999).

In addition to its renowned value to facilitate learning, the Internet is an expansive tool for reducing digital and social isolation and empowering individuals with disabilities. This then makes it all the more important to ensure these people have access to the information that they require. Difficulties in accessing information reduces the Internet's learning potential and prevents people from participating in the social, political and economic issues that are increasingly being enabled via Information and Communication Technologies. (ICTs) (Carver, 2000). This then increases the likelihood of social exclusion to the disadvantaged. (Schneiderman, 2000) Both disabled and elderly people commonly experience social isolation. This is then another area where the Internet could be particularly beneficial in assisting the maintaining of social contact. Then there is also the subject of disempowerment, the result of necessary reliance upon the assistance of others, that many people within this user group are subjected to. Those that are blind potentially have less need to be reliant upon others for assistance with tasks such as shopping, studying or working. People with mobility and hearing problems could access documentation online instead of facing difficult journeys and problems with communication. Furthermore, the ability to become more independent is enabled, thus increasing the individual's self esteem. Learning potential, inclusion and empowerment are all issues that could possibly go some way to compensate for the restrictions that are imposed by the effects that their impairments have upon their lives (Kobsa & Stephandis, 1998). In short, the improvement of accessibility will not only increase learning potential but is also likely to have a positive effect upon other aspects of their lives, which are affected by the nature of their disabilities. These problems and the issue of accessibility in general is also relevant to older adults.

The goal of universal accessibility has been the driving force behind new techniques of information adaptation to cater for specific needs (Preece et al, 2001; Nielsen, 2001; Scheiderman, 2000). Although there is a wide variety of assistive technology to enable web pages to be made more accessible, this does rely upon correctly coded pages. For example, where pages do not have alternative text for images, screen readers cannot describe a picture to a blind user. In addition, there are elements of web pages that do not lend themselves to modification. For example, the layout of information can present problems for dyslexic users, whilst Web pages with interaction that requires precise mouse movements would present problems for motor impaired users. Users can then find themselves sifting through numerous pages to locate required content. This can be tiresome for even accomplished users but for people with special needs this task becomes even more arduous and frustrating, depending upon the extent of impairment. The content of a web page may match the user's goal but if it cannot be accessed as a result of poor design, then the content becomes useless. Selecting the presentation order of web pages according to specific needs could then be a possible alternative.

The specific requirements for the intended user group need to be identified as well as working out what adaptivity is needed. Elements of web pages that affect the accessibility for users with disabilities will need to be identified. This will help to understand how information can best be selected and ranked. To enable the pre-selecting of pages according to the user's specific requirements, some means of developing new attributes of web pages, to be stored in the web model, needs be explored. The web model will enable the process of automating the ratings of web pages, according to the specified attributes contained within the user model. For example, the *suitability* of web pages may be categorically rated as: very suitable, medium or unsuitable. Further explanation of this might be where a page's elements are all accessible, then it can be said to be suitable. A few non-accessible items may constitute medium, more than this would then deem the page as unsuitable.

Weightings of web page elements that affect accessibility for visually and motor impaired, as well as dyslexic users needs to be established. User defined ratings of these elements that affect accessibility would enable a set of user preferences to be created. Each category of impairments will have its own set of preferences, and an algorithm would then be used to rate web pages (Good, 2004). A default set of ratings for each category can then be adjusted according to user preference. The WAI gives a comprehensive description of which elements relate to which type of impairment and how they affect the user (W3C/WAI, 2004); some of which have up to nine specified elements. There are numerous elements that can affect a user with a particular impairment, however there appears to be little research conducted on the weightings of these elements even though they can differ in the extent to which they affect accessibility. It is the weightings of these elements that remain the focus of this study.

It is anticipated that to create an algorithm on the weightings of all elements that affect accessibility, would considerably increase the scope of this research. An important consideration has to be the time taken to analyse individual pages. The greater the number of elements, the longer the parsing time will take. In addition, to facilitate the selection of web pages based upon the entirety of specified elements would then significantly reduce the number of accessible pages and would impact upon the user's search criteria being fulfilled. It is also hoped that by limiting the algorithms to three elements, this will enable a preliminary analysis of the feasibility of selecting web pages according to user needs. The purpose of this study is then to gain user defined ratings of three elements for each category of users. Another reason for specifying that the users select only three is to reduce the likelihood of eliminating elements purely upon preference only, although it is hoped that further research will enable this facility to be implemented. It should be noted that this is an exploratory study and as such, the results will be used as an indication of preferences.

2 Method

The aim of this study was to gain user defined ratings of web page elements that affect accessibility. A user-centred approach was adopted for the study with a combined thinking aloud protocol and an observation study, with a follow-up questionnaire.

2.1 Procedure

Participants were asked to browse some information from the Web relating to news, health information and communication/recreational sites. They were given three tasks:

1. Find out what the current situation on SARS is (Severe Acute Respiratory Syndrome).
2. Find information on how to prevent asthma using the NHS web site.
3. Find a discussion group for gardening

Participants undertook various search related tasks. They were requested to talk through the process expressing their thoughts on the web pages they visited. In particular, they were to voice any problems relating to ease of usage and accessibility of web pages. Observations and comments were noted. After completing the task, participants were asked to specify which elements of the web sites they used affected the ease with which they were able to access the information. They were then asked to select the three most prominent elements and rate them between one and three with one being the element that made accessibility the most difficult.

3 Results

The study yielded two types of data, qualitative and quantitative, which were analysed differently but will be used for comparison. The elements that were rated by users were analysed using a quantitative approach (Good, 2005). The thinking aloud protocol/observation method yielded qualitative data that was analysed using content analysis. It is the user defined ratings of elements that this study is predominantly concerned with and as such, the qualitative data will be used for reference only. It should be noted that there was little difference between the two sets of results. Statistical analysis was not deemed appropriate to analyse the results from this study because of the relatively small sample size and means of selection. It has previously been noted that this was an exploratory study and as such, the results will be used to give an indication of preferences. The results are therefore to be viewed as a pre-requisite for further research into user-defined ratings of elements that affect accessibility.

3.1 User Defined Elements

The participants were asked to list three elements and rate them in the order in which they affected web page accessibility. The elements were given scores according to their ranking. The first element was assigned a score of three, the second element received a score of two and the third element received a score of one.

Blind Participants. All six participants rated the lack of alternative text on graphics being the most significant accessibility issue for them. Three participants assigned frames as being the second element, with a rating of two that caused accessibility problems with the other three assigning this rating to pop-ups. The third element to cause accessibility issues was text embedded in pictures for three participants and pop-ups for the other three.

Other Visually Impaired Participants. There were fifteen participants who had varying degrees of visual impairment but were not medically defined as blind. Font size caused the most accessibility problems followed by low contrast between the font and the background colour, and with pop-ups being rated third. Other elements that caused problems were: patterned backgrounds; scrollbars; layout; embedded links and blue links. Only one element from the WAI list (W3C/WAI, 2004), was not specified and this was image text that cannot be re-wrapped. With the exception of font size, contrast between background and foreground, and layout, the remaining user defined elements do not feature in the list provided by the WAI. In terms of the second two elements: background colour and contrast between background and foreground, these two could be said to imply the same thing. Background colour can affect readability. Some justification for why the remaining elements were defined by the participants but were not specified by the WAI could be that these elements refer to the user's experience as opposed to accessibility. Results from the observational study did not highlight any additional elements.

Dyslexic Participants. Small font size and pages with large blocks of text were stated as being the elements that caused dyslexics the most problems. Participants rated these two elements equally. Web pages with white backgrounds were also problematic as was inconsistent layout, the lack of pictures and sentence length. The WAI have specified four additional elements that cause accessibility problems for the cognitive/language category. A likely explanation as to why they weren't rated by the participants is because the category includes a number of impairments including epilepsy and thus, video/audio frequencies triggering seizures would not be applicable for dyslexics. Conversely, there were a number of elements defined by the participants, which were not outlined as being problematic by the WAI. These include font size, too much text, background colour and sentence length. The observational study noted that when background colour was mentioned as being problematic, it was often because the colour was white. Other elements noted were lack of headings and very long web pages.

Motor Impaired Participants. There were ten motor impaired participants that took part in the study. The results show small links as affecting accessibility the most for this particular user group. Scroll bars came second with contrast between fore and background colour being third. All the elements defined by the participants were also outlined by the WAI. Conversely, the participants did not mention the following 3 elements as specified by the WAI: time limited response on web pages, no keyboard alternatives and forms that cannot be tabbed through. A likely explanation for this could be that the websites the participants visited did not feature these elements. The observational study highlighted the same elements specified by the participants with the exception of one, that being the lack of search boxes on large websites

4 Discussion

Where assistive and adaptive software have certainly provided better access to information for special needs users, in terms of equality of access, there is no comparison to non-disabled users. The problem itself is widely recognised with numerous interest groups researching new techniques and introducing legislation that will provide better access to future web-based information. However, despite the availability of standards and guidelines there is still a significant shortfall of compliance. It is this shortfall that prevents people with disabilities equal access of information. Where non-compliance to design standards prevails, there will always be users who are faced with barriers. Learning potential, inclusion and empowerment are all issues that are affected by inaccessible web pages. Selecting web pages according to their accessibility could help to reduce the amount of inaccessible information presented to the user and thus increase the value of the World Wide Web.

Table 1. A Summary of the Results

	Blind	Short-sighted	Dyslexic	Motor
1	Images-no alt text	Font size	Font size	Small links
2	Frames	Contrast	Layout	Scroll bars
3	Tables	Pop-ups	Background colour	Contrast

This paper presented the results of an exploratory study to gain user defined ratings of elements that affect accessibility for users with special needs (Table 1). The study included participants with a wide range of sensory, physical and cognitive impairments. Participants were asked to specify and rate, with a scoring of one to three, the elements they felt affected accessibility of web pages. Many of the elements that were rated were ones that had been previously specified by the WAI. This is certainly true for all visually impaired participants. In terms of the top three rated elements (see Table 1), almost all dyslexic participants specified that font size was a problem and over half said that when the background colour was white, it made readability particularly difficult. These two elements were rated one and three respectively, however the WAI does not specify either of the elements as likely barriers. Those participants that experienced difficulty reading black text off a white background used coloured filters or wore precision tinted glasses, with individually chosen prisms and colour when reading web pages. Some of the elements that received lower weightings were also not ones that had been specified by The WAI. A possible explanation for this could be that they were elements that affected the user's experience as opposed to the effect they had upon accessibility. This would seem particularly relevant where over half of motor impaired participants specified that low contrast between foreground and background colours presented difficulties. A possible explanation as to why those elements that were outlined by the WAI were not defined by the participants could be because the web sites explored within the study did not feature them. Another reason could be that due to the small sample size and as such, the participants may not have experienced problems in relation to those elements.

5 Algorithms

The results from this study were used to create a set of algorithms, customised according to user needs. The purpose of the algorithms will be to analyse web pages and assess their suitability for users with specific needs. The algorithms calculate web page accessibility by looking for specific elements known to known to affect a particular type of user. The elements incorporated into the algorithm are the highest three rated by users from the reported study. The following sets of algorithms demonstrate how web page accessibility is calculated.

5.1 Algorithm for Short-Sighted Users

The following rule script shows how an algorithm makes decisions upon selecting the most accessible web pages for a short-sighted user:

```
If (user = ‘Short-sighted’)
{set page accessibility to 100
        {Check font size
                if (font size is fixed)
                        then reduce accessibility by 20
                if (font size of body text is less than 10pt)
                        then reduce accessibility by 20
                if (font size of body text is less that 12pt but greater than 10pt)
                        then reduce accessibility by 10         }
        { Check contrast between background and font colour
                if (contrast is less than 100)
                        then reduce accessibility by 50
                if (contrast is less than 200 but greater than 100)
                        then reduce accessibility by 40
                if (contrast is less than 300 but greater than 200)
                        then reduce accessibility by 30
                if (contrast is less than 400 but greater than 300)
                        then reduce accessibility by 20
                if (contrast is less than 500 but greater than 400)
                        then reduce accessibility by 10}}
```

5.2 Algorithm for Blind Users

```
If (user = ‘Blind’)
Total_area_ image = total area of images
na_image = non-accessible image
space_na_image = % of total images that non-accessible image(s) uses // non accessible image - no alt text
{       set page accessibility to 100
        Check instances na_image
                { Calculate space_na_image
                   total_area_image – total_area_na_image   //  total area of images minus total area of non accessible images
                        if  (space_na_image is greater than 75)
                                then reduce accessibility by 75
```

```
                    if  (space_na_image is greater than 50 but less than 75)
                              then reduce accessibility by 50
                    if  (space_na_image is greater than 25 but less than 50)
                              then reduce accessibility by 25}
      {Check usage of frames
                    if  (usage of frames)
                              then reduce accessibility by 15}
      Check usage of tables
                    if  (non-linear)
                              then reduce accessibility by 10}}
```

5.3 Algorithm for Motor Restricted Users

```
   If (user = 'Motor-restricted')
   {      set page accessibility to 100
          {Check for a logical tab order through links
                    if (no logical tab order)
                              then check for keyboard shortcuts
                                        if (no keyboard shortcut)
                                                  then reduce accessibility by 60
             {Check scroll bar
                    if  (scroll bars)
                              then reduce accessibility by 30}}
```

5.4 Algorithm for Dyslexic Users

```
   If (user = 'Dyslexic')
   {      set page accessibility to 100
          {Check font size   // This would apply to main body text
                    if  (font size is fixed)
                              then reduce accessibility by 20
                    if (font size of body text is less than 10pt)
                              then reduce accessibility by 20
                    if (font size of body text is less that 12pt but greater than 10pt)
                              then reduce accessibility by 10
          Check for headings //  Checking for large blocks of text *}
          { Check background colour and font colour//  body text
                    if (background colour is white and font colour is black)
                              then reduce accessibility by 20
                    if(background colour is black and font colour is white)
                              then reduce accessibility by 20}}
```

6 How the System Works

The system will incorporate a user model and web model with new attributes that are particularly relevant to users with special needs. The information in the user model is used to determine which algorithm is used to rate the accessibility of a web page. The type of algorithm applied will depend upon whether the user is: blind; short-sighted; motor restricted or dyslexic. The user supplies this information. A database acts as the web model, which represents the accessibility of each page and includes attributes such as the URL and accessibility rating of web pages. Each web page is then analysed according to the constraints incorporated within the algorithm. When a user submits a query, the system refers to the user model to establish which algorithm is assigned. It then individually analyses the first 20 web pages that the search engine returns, and inputs the URL and the rating of accessibility it has given the page into the database. It is hoped that by restricting the rating of web pages to only 20, it is then less likely that accessibility will override content. Following this, the system then ranks the web pages according to the ratings applied and then presents the 20 pages as a list of URLs to the user, according to the ratings that have been assigned. Pages with high ratings are considered more accessible than those with low ratings.

6.1 How the System Works for a Specific User

Using short-sighted users as an example, the algorithm that would be applied will specifically look at font sizes and the contrast between font and background colour. Although the results indicate that users specified advertisement pop-ups as being the third element to affect accessibility, locating the presence of this element has proved to be problematic. This is because it is difficult to determine whether the HTML relating to the pop up explicitly relates to an advert or an instance of another window opening up. A manual check would look at embedded scripts, included scripts, and event handlers for "window.open()" methods, however this is beyond the scope of the system's capabilities. It is hoped that further research will allow for this feature to be incorporated.

Each web page is initially assigned a default rating of 100. The rating will decrease depending upon the extent to which affecting elements are present. For the short-sighted user, the algorithm will firstly check that web pages are not designed with a fixed font. The system considers this to be the most important feature when analyzing web pages for font size because most short-sighted users will need to increase the font. Pages that have a fixed font will then receive a higher penalty. The algorithm is then designed to look at the size of the font used for the body text. Only pages that use a font size below 12pt will be penalised. The next element that is checked will be colour contrast. According to the formula provided by the W3C, the range for colour difference needs to be above 500 if the text is to be readable:

> *Color difference is determined by the following formula: (maximum (Red value 1, Red value 2) - minimum (Red value 1, Red value 2)) + (maximum (Green value 1, Green value 2) - minimum (Green value 1, Green value 2)) + (maximum (Blue value 1, Blue value 2) - minimum (Blue value 1, Blue value 2))* (W3C, 2000)

The algorithm is designed to assign penalty points according to the extent of colour difference. Those pages with very low colour difference will receive a lower acessibility rating than those with a higher difference. These algorithms will be used to analyse and select web pages according to user preferences. An additional study will then evaluate the effectiveness of the presentation order of web pages. The study will also evaluate the results from the exploratory study for consistency. Further research will then be carried out to include all elements and then to develop more complex algorithms based upon the results of this.

7 Conclusion

Where many current methods tend to focus upon the adaptation of content, the proposed approach is to adapt the order of search results according to how accessible each page is to the user. Web pages will be assessed according to individual needs using user-defined ratings of elements that affect accessibility. The more accessible the web page is, the higher up the list of results it will feature. This is a far more simple approach than modifying content; a process that retains a certain amount of fallibility due to its dependence upon correctly coded pages. This proposed scheme would reduce the need to sift through numerous inaccessible pages to locate required content, which can be particularly arduous and frustrating task for an impaired user. This could then improve the effectiveness of the World Wide Web as a pedagogical tool for this particular user group.

Since this research commenced, Google has devised a means of selecting web pages specifically for visually impaired users. Their aim is to reduce the amount of inaccessible pages that visually impaired users are forced to sift through to find necessary information. Named 'Google Accessible Search' (Google, 2007), it searches for pages using its search algorithms and couples this by applying additional factors to rank the pages. This process involves looking at HTML markup. 'Google Accessible Search' assesses the overall simplicity of a page. The extent of imagery is examined as well as establishing how the page is affected by disallowing images. In addition, it will assess whether the page facilitates keyboard navigation (Google FAQ, 2006). The method that is proposed in this paper works on a similar principle, in that pages are ranked according to how accessible they are. Both systems use algorithms to rank pages according to factors of accessibility and whilst the factors are very similar for blind users, the proposed system here will also consider elements that specifically affect short sighted users such as contrast and fixed fonts. The other difference is that the proposed system is not limited to visually impaired and blind users but also considers dyslexic, and motor restricted users.

References

Deets, H.B: Aging and Technology: The Convergence of Two Revolutions. CyberPsychology & Behavior 2(6), 501–503 (1999)

Disability Rights Commission The Web: Access and Inclusion for Disabled People. London (2004)

Henze, N.: Open Adaptive Hypermedia: An Approach to Adaptive Information Presentation on the Web (2001)

Kobsa, A., Stephanidis, C.: Adaptable and Adaptive Information Access for all users, Including Disabled and Elderly People. In: Proceedings of the 2nd Workshop on Adaptive Hypertext and Hypermedia. Hypertext'98. Pittsburgh, USA (June 20-24, 1998)

Good, A.: Enabling Accessibility and Enhancing Web Experience: Ordering Search Results According to User Needs. In: Proceedings of HCI International 2005 (2005)

Good, A.: Addressing User Needs: Adapting Information Access For The Elderly. HCI 2004, HCI and the Older Population Workshop (2004)

Google Accessible Search FAQ. Google labs. (Retrieved January 11, 2007) from http://labs.google.com/accessible/faq.html

Nielson, J.: Beyond Accessibility: Treating Users with Disabilities as People. (Retrieved December 12, 2004) (2001) http://www.useit.com/alertbox/20011111.html

National Statistics Omnibus Survey Internet Access. Individuals and Households (2003)

Preece, J., Rogers, Y., Sharp, H.: Interaction Design. John Wiley & Sons, New York

Shneiderman, B.: Universal Usability. Communications of the ACM. vol. 43(5) (May 2000)

Skinner, B.F.: Teaching Machines. Science. vol. 128 (1958)

W3C/WAI Web Content Accessibility Guidelines 2.0. (Retrieved December 14, 2004) (2004) http://www.w3.org/TR/2004/WD-WCAG2020040311/

W3C/WAI How People With Disabilities Use the Web. (Retrieved November 18, 2004) (2004) http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/Overview.html

W3C Techniques For Accessibility for Evaluation and Repair Tools. (Retrieved March 28, 2005) (2000) http://www.w3.org/TR/AERT