

A method to automate the ranking of web pages according to user defined accessibility ratings

Alice Good

School of Computing
University of Portsmouth
PO1 3AE
UK

alice.good@port.ac.uk

Abstract. The premise of this research is to present the final results of an investigation, which looked at a means to automate the rating of web pages according to their accessibility to specific user groups. These groups include visual impairments, mobility restricted and dyslexia. This research identifies three integrated, user-centred studies that assisted in the development of this work. The research conducted for this project has collected data that will help to develop a better method for disabled users to search for and easily locate accessible web pages. It has investigated how web pages can be rated for accessibility using specific algorithms that have designed according to user defined ratings of accessibility. The results presented in this paper demonstrate that re-ordering search results, by ranking web pages according to user defined ratings, could provide a better user experience for people with disabilities.

Keywords: Accessibility; Algorithms; Disabilities; HCI

1 Introduction

Where assistive and adaptive software have certainly improved access to web-based information for users with disabilities, 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 accessibility could help to reduce the amount of inaccessible information presented to the user and thus increase the value of the World Wide Web.

Access to information is notoriously inadequate for disabled people; both current and earlier research confirms this [1,2,3,4]. In recent years, the medium has expanded to provide new services, such as the social networking sites including Facebook, Bebo and MySpace. However, according to AbilityNet, disabled users still find themselves excluded due to design problems [5]. While there are a number of contributing factors, the fundamental cause can be traced back to poor web design. Certainly, there is a marked failure to consider the needs of all users and incorporate the necessary adjustment into the design process. Consequently, poor design leads to poor accessibility [6,7]. With age comes increased disability: in the elderly, vision and mobility deteriorate with time. With an aging population on the increase, the severity of this problem continues to grow.

Population aging is on the increase and as such, there will be a greater percentage of people alive at this time who are likely to be experiencing age-related impairments than in the past. There is a definite need to 'bridge the gap' between elderly users and the Internet and with almost a fifth of the UK population aged over 65, it is a significant market to exclude [8]. Many of these people make use of the Internet in controlling their finances, keeping up to date with current affairs and staying in contact with family and friends. The Internet can then be said to be a valuable tool in reducing the risk of social exclusion, a concern that many of the elderly face. In considering that the elderly are very likely to develop age related impairments such as visual and/or mobility related disabilities, web accessibility then becomes an important consideration. The initiatives may be in place to encourage the elderly to get online but poor web accessibility is likely to affect not only the user's ability to access information but also create a less than favourable web experience. Better solutions are needed to improve accessibility for special user groups.

Much research has been dedicated to devising ways of modifying content according to users' needs. The focus has then stringently been placed upon the alteration of content, a process that retains a certain amount of infallibility due to its dependence upon correctly coded pages i.e. web pages that have been written with the relevant html, for example 'Alt-tags', so that screen readers are able to inform the user. However, correctly coded accessible web pages can be difficult to find. There needs to be a practical and effective way to guide users to accessible content. One possible solution could be adapting the order of search results according to user needs. This would then enable a means of retrieving accessible information, which is best suited to the users' needs and disabilities. Furthermore, such a method would contribute towards a better user experience too.

1.1 Inter-related studies

This research has featured a series of inter-related studies, which are outlined below:

1.1.1 Study One

In the initial requirements phase, user defined ratings of web page elements that were known to affect accessibility for specific user groups were obtained [9]. 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 particularly true for all visually impaired participants.

1.1.2 Study Two

These ratings formed the basis of a set of algorithms that were designed according to each user group, using the user defined elements of web pages that reduced accessibility outlined in study one (9). The objective of the algorithms was to analyse the accessibility of a web page, according to a specific user group. The algorithm would then assign a rating based upon the page's accessibility. Using these algorithms, users can enter a search query and Web results are then re-ordered according to how accessible they are to a particular user group.

1.1.3 Study Three

The focus of this paper is to present the final stage of this exploratory study, namely study three. The aim of this study was to evaluate the effectiveness of the algorithms in re-ordering search results according to user perceived accessibility in comparison with Google search engine. An empirical evaluation was conducted using the same participants from study one. The purpose was fundamentally to assess whether participants had a better Web experience using the Algorithm Based Search.

1.2 Algorithm Based Search

The proposed system, known as Algorithm Based Search, will incorporate a user model specific to the disability group as well as a web model. 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. It is intended that the user supplies this information by specifying which user group best fits their needs: ; blind; short-sighted; motor restricted or dyslexic. 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. The purpose of the web model is to store the URL and the rating of the web page. 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. The algorithm 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. 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.

2 Method

A simulator was built using the algorithms created from user-defined ratings of accessibility that were obtained from the first study of this research [9]. Each disability user group had its own algorithm: blind; short sighted; motor restricted and dyslexic. The simulator was built in html. The simulator enabled a user to specify which user group they were in and the content of each of the two search queries (Scottish castles and Health Information). Twenty web pages were previously analysed according to the rating system incorporated within the algorithm. For example, for a short-sighted user, the user-defined rated elements from the first study that were found to affect accessibility were fixed font, low contrast between background and font colour. The simulator allowed the user to switch between the ordered set of results according to the applied algorithm and the set of results provided by Google. Each user set consisted of 10 search results for each query. There were two search queries and so each user set had two sets of results, one for 'Scottish castles' and one for 'Preventing asthma'. There was also the default set of search results provided by Google for each search query.

For this study, an empirical investigation was used to measure the effectiveness of the system to direct users to accessible pages. This system has been designed to improve the accessibility of search results for disabled users when compared to a traditional search engine. In this study, users were asked to rate the first ten pages of the search results presented only, as this amount was deemed a suitable sample. Google and most other on-line search engines limit the amount of results to 10 per page.

2.1 Participants

The second study included 31 participants. It is possible to obtain excellent results using small sample groups, as research undertaken by Nielsen and Landauer (1993) reveals. Sample groups consisting of no more than five people can yield meaningful data. Furthermore, frequent tests with smaller groups can produce better results than large-scale studies [10].

Participants ranged in age from 16-65; both males and females were included. Computer skills ranged from novice to intermediate. All users reported that they experienced impairments that affect the ease with which they were able to access web-based information.

2.2 Procedure

The participants were asked to perform two searches for specific information, using the Internet. They were asked to enter the search query regarding health related issue, firstly using the algorithm and then using Google search. Then users were asked to rate the pages, using a Likert-Scale, in terms of ease of access and suitability of content.

3 Results

Respondents were asked to rate web pages in terms of ease of access and precision of content using both the Algorithm Based Search system and Google. Results indicated statistically significant differences between respondents' perception of the ease of access and content of the webpage between Google search and Algorithm Based Search.

3.1 Comparison of responses

The Likert-Scale responses were converted onto a numerical scale with 'strongly disagree' responses assigned the minimum weight of 1 and 'strongly agree' response assigned the highest weight of 5. The differences in responses were calculated between the algorithm based search and Google, for each respondent and each web page. Differences in responses are recorded for each web page to analyse how the behaviour of the respondents change as they record their response for page 1- 10. A flat line above the zero mark will be the most ideal outcome as this would suggest that respondents consistently rated algorithm based search results better than Google results. Mean, Median and Standard deviation were calculated for difference in responses for each web page. Fig. 1 and 2 show the Mean, Median and Standard deviation of difference in responses for each web page.

3.1.1 Health Information Search

As fig. 1 shows, Mean, median and Standard Deviation were generally positive with page 8 and page 6 being the exception, where respondents felt that Google search results were better than code based search. Also, difference in responses was not consistent with 4 crests and 4 troughs. Respondents particularly reported negative responses for web page 6 while responses for web page 8 were marginally negative. However, the graph suggests that overall the respondents positively felt that the Algorithm Based Search results were better than Google search results in terms of ease of access and suitability of content

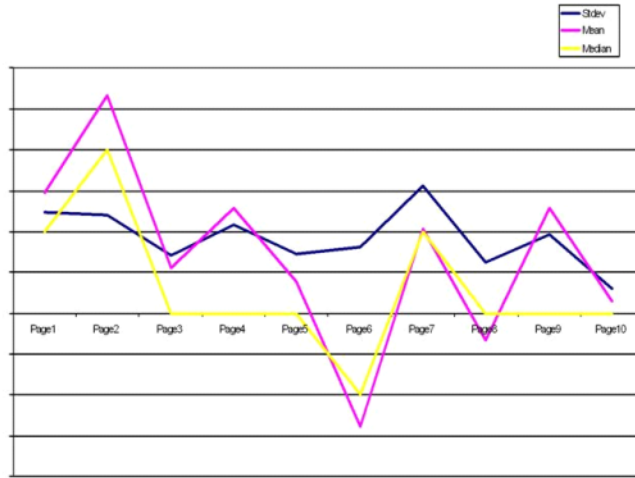


Figure 1: Health Information Search Statistics

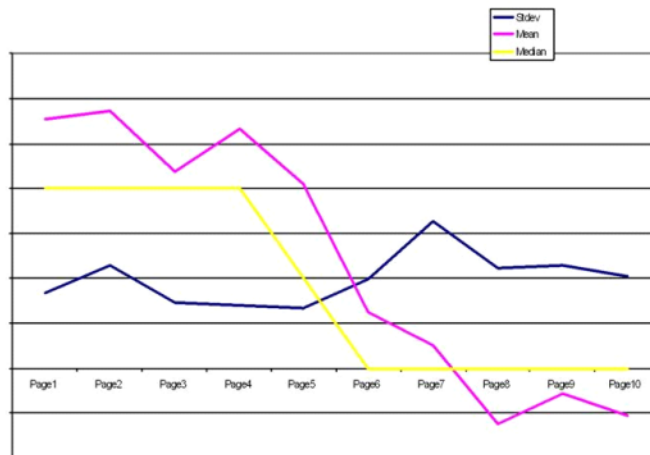


Figure 2: Scottish Castle Search Statistics

3.1.2 Scottish Castle Search

Fig 2 suggests that respondents felt positive about algorithm based results but their responses turned negative as they moved down the search results. This may be because the Scottish Castle related web pages were very graphical with many images. Obviously a high ratio of graphics to text would create accessibility issues for blind users. Because the algorithm checks for fonts and contrasts, the two attributes which do not affect images very much, the algorithm results down the order may have been less appealing to the respondents than Google results, however the respondents did report a significant and positive improvement in results with the algorithm.

3.2 Statistical Testing of Results

The Wilcoxon Signed Ranks Two-Tailed Test was also used to determine differences between users' responses on which search method was more convenient to them in terms of accessibility and in terms of finding the information that they were looking for. This non-parametric test for significance between items was considered most appropriate because the measure uses an ordinal scale. The Wilcoxon Signed Ranks Test approximates a z test. To further determine the magnitude of differences, all analyses were conducted at the $p < .01$ significance level.

We test the hypothesis:

H0: The median response is 3 i.e. neither agree nor disagree

H1: The median response is >3 i.e. the respondents felt a noticeable enhancement in results when searched using the code as compared to when searched using Google search engine.

Median	10
Mean	9.612903
Standard deviation	3.116381

Median	14
Mean	16.22581
Standard deviation	8.452651

Wilcoxon signed rank test was then performed using SPSS version 11.5.

3.2.1 Health Information Search

The results obtained for health related query were:

W- = 72.50, W+ = 227.50, N = 24, $p \leq 0.01346$

The huge difference between W+ and W- suggests that we can rule out the NULL hypothesis. This means that the users did experience a difference in browsing experience under the two settings- Google based and Algorithm Based Search. This conclusion was confirmed by the critical value comparison.

The selected (being smaller) value for the search is W-. Next, the critical value in the Wilcoxon signed rank test critical values table was found to be n=24 at 95% significance. This value is 91.

W- = 72.50 and
 $T_{\text{critical (95\% significance)}} = 91$

Thus $W(72.50) < T_{\text{critical (95\% significance)}}$

This means that we can reject the NULL hypothesis that respondents' responses for both the searches are identical. The higher value of W+ suggests that more respondents felt that their browsing experience was enhanced because of using the algorithm-based search. If the respondents felt a negative difference in browsing experience we would have obtained a higher value of W-.

In conclusion we can say that, at 5 percent significance level, the respondents felt that for health related query, algorithm based search results were better in terms of accessibility and content than the Google based search results.

3.2.2 Scottish Castle Search

The results obtained for Scottish Castle related query were:

W- = 109.50, W+ = 268.50, N = 27, $p \leq 0.02760$

Again, we witness a huge difference in W+ and W- values, which suggests that we can rule out the NULL hypothesis. This means that the users did experience a difference in browsing experience under the two settings- Google based and algorithm based. But this can only be confirmed by looking at the critical value comparison.

The selected (being smaller) value for the search is W-. Next, the researcher looked up for the critical value in the Wilcoxon signed rank test critical values table for n=27 at 95 percent significance. This value is 119.

W- = 109.50 and
 $T_{\text{critical (95\% significance)}} = 119$

Thus $W < T_{\text{critical (95\% significance)}}$

This means that we can reject the NULL hypothesis that respondents' responses for both the searches are identical. The higher value of W+ suggests that more respondents felt that their browsing experience was enhanced because of using the Algorithm Based Search. If the respondents felt a negative difference in browsing experience, a higher value of W- would have been obtained.

In conclusion it can be said that, at 5 percent significance level, the respondents felt that for health related query, Algorithm Based Search results were better in terms of accessibility and content than the Google based search results.

4 Conclusions and Recommendations

The research has provided a means of rating Web content according to user needs. The results presented in this paper indicate that overall, participants had a better user experience when using the Algorithm Based Search. The algorithms developed will also go some way to improve overall accessibility by creating a filtering system that identifies and rates pages according to the user's accessibility needs. One of the objectives of this work was to develop these algorithms with the input of those directly affected by accessibility issues. This research is informed by a need to involve and engage with those who are directly affected by accessibility issues. It is believed that these individuals possess the knowledge and experience necessary to assist in addressing the problem of accessibility.

While the project has resulted in significant gains, there is considerable scope for future research. It is necessary to create algorithms for other impairments, such as auditory or cognitive impairments other than dyslexia. Additionally, specific visual impairments, apart from blindness and short-sightedness, also need to be addressed. Most importantly, algorithms need to be developed in order to meet the accessibility needs of users with multiple impairments. This is a particular challenge, as it will require reconciling the competing aims of different forms of impairment. In spite of the expected degree of difficulty, this research is essential. Multiple levels of disability affect many users. For example, the growing numbers of elderly users face a host of impairments as they age. As eyesight fails, arthritis sets in and other neurological conditions may develop.

4.1 Recommendations

As a result of the research undertaken for this project, the following recommendations for further research are suggested.

- Develop algorithms that allow for more elements of web pages known to affect accessibility to be included
- Enable user specific algorithms to have different level of searches according to extent of accessibility check required. For example – low, medium and high
- Develop algorithms that allow users to specify which elements they wish to look for.
- Perform study on larger group of users

- Conduct a comparative study between Google's accessible search engine and the algorithm-based search for users with visual impairments
- Develop additional algorithms need to be developed in order to address the needs of other users with impairments not covered by this research
- Develop algorithms to meet the needs of users with multiple impairments.
- Refine the simulator used to test the algorithms to ensure that it functions at a higher level.
- Implement system based upon simulator and test performance with live searches.

References

1. Lazzaro, J., Adaptive Technologies for Learning and Work Environments, American Library Association, USA (1993)
2. Webb, C., Ensuring the Internet Revolution reaches the blind (2001) <http://www.theage.com.au/news/2001/01/22/FFX5ORDL71C.html>
3. Hackett, S., Parmanto, B. & Zeng, X. Accessibility of Internet Websites Through Time, ACM SIGACCESS Accessibility and Computing Archive, no. 77-78, pp. 32-39. (2003)
4. Bevan, T. & Ahmed, A. An Investigation into Web Accessibility Standards as a Practical Study with Older and Disabled Users, ACM SIGACCESS Accessibility and Computing Archive, no. 88, pp. 9-14. (2007)
5. Lang, S., Social Networking Sites Exclude disabled users, says charity, The Guardian, 23 January, (2008)
6. Shneiderman, B. Leonardo's Laptop: Human Needs and the New Computing Technologies, Cambridge, Mass., MIT Press. (2002)
7. Preece, J., Rogers, Y., & Sharp, H. Interaction Design. New York, NY: John Wiley & Sons, Inc RBIB (2001)
8. Pinder A. Help to Bridge The Digital Divide. Computer Weekly, June 27. (2002)
9. Good, A., Jerrams-Smith, J., Enabling Accessibility and Enhancing Web Experience: Ordering Search Results According to User Needs. Lecture Notes in Computer Science: Universal Access in Human-Computer Interaction. Applications and Services. Springer Berlin / Heidelberg, Volume 4556/2007 (2007)
10. Nielsen, J. & Landauer, T. A Mathematical Model of the Finding of Usability Problems Proceedings of ACM INTERCHI '93 Conference, Amsterdam, The Netherlands, pp. 206-213. (1993)

