Usability and Accessibility of Aging/Health-Related Web Sites

Panayiotis Zaphiris, Sri H. Kurniawan

Institute of Gerontology and Industrial & Manufacturing Eng.
Wayne State University
Detroit, MI 48202
+1 313 577 2297
p.zaphiris@wayne.edu, s.kurniawan@wayne.edu

ABSTRACT

This study is aimed at answering whether aging/health-related web sites of different domain extensions (i.e. .com, .edu, .gov and .org) differ in their accessibility and usability, and whether these two measures are correlated. The usability and accessibility of governmental, organizational, educational and commercial aging/health-related web sites were compared using two automatic evaluation tools: Bobby and LIFT. The governmental web site group has the highest compliance with Web site Content Accessibility Guide although only 52% got an approved status. The accessibility approval was found to correlate significantly with overall usability ratings for all groups, except the commercial web site group.

1. Introduction

Sullivan and Matson (2000) compared 50 most popular web sites in terms of their usability and content accessibility and found a marginal correlation (ρ=0.23) between manually analyzed content accessibility in conformance to the Priority 1 of the Web Content Accessibility Guidelines (WCAG) (Chrisholm, et al, 2001) and overall automated usability testing result provided by LIFT (http://www.usablenet.com/index.htm).

Older adults have some disadvantages in fully utilizing the Internet as their information resources. That is, older people had more trouble finding information in a Web site than younger people (Mead, et al, 1997). Very little research, however, focused on ensuring the accessibility and usability of Web health information for the elderly.

The present study aims to evaluate usability and accessibility aspects of web sites that might be of interest to older computer users. This study also extends Sullivan and Matson's (2000) study in two ways: by automating the content accessibility testing using Bobby (http://www.cast.org/bobby), which performs the test based on all priorities, and by performing group comparisons of commercial, educational, governmental and organizational web site groups in terms of their usability and content accessibility. Furthermore, this study aims to answer two research questions:

1. Do aging/health-related web sites of different natures (i.e. commercial, educational, governmental and organizational) differ in their accessibility and usability?
2. Is the result of accessibility evaluation of a particular group of web sites related to the result of its usability evaluation?

The two automatic evaluation tools used in this study are LIFT and Bobby. LIFT was chosen because it is the only automatic tool that performs usability evaluation. Bobby was one of the most widely used automatic accessibility evaluation tool.

LIFT provides a report of the number of catastrophic errors (errors that disable users to complete tasks), major errors (errors that cause users to face major impediments), minor errors (errors that are really a nuisance for users) and cosmetic errors (low priority materials). In addition, as a general rating, LIFT assigns a rating of excellent, good, fair or poor.

Bobby recommends effective Web page authoring for special Web browsers (e.g. the one which reads text out loud using a speech synthesizer for blind users). Bobby divides the accessibility errors into 4 sections: *Priority 1 Errors* (problems that seriously affect the page's usability by people with disabilities), *Priority 2 Errors* (secondary

access problems), *Priority 3 Errors* (third-tier access problems), *Browser Compatibility Errors* (HTML elements and element attributes that are used on the page which are not valid for particular browsers).

2. Methodology

The web sites analyzed in this study were collected using keyword search of "aging" from http://www.google.com search engine (all of them contain information about health). The web site's domain name extension (.com, .edu, .gov and .org) was used as a filter.

The analyzed web sites for each extension were limited to the sites listed in the first three pages (30 sites) of the google's search result because the majority of web users are not expected to go beyond 3 pages when looking for information (Zaphiris, 2000). After removing subsections of the same web sites and dead links, the numbers of analyzed sites for each extension ranged from 20-25.

3. Result and Discussions

Table 1 lists the mean and standard deviation of the usability and accessibility ratings for each group of web sites. Bobby's approval rating is converted into a binary variable with '0' representing 'Not Approved' and '1' representing 'Approved' status. The Usability rating is also converted into an ordinal scale with '1' representing 'Fair', '2' 'Good', and '3' 'Excellent'.

From Table 1 it is apparent that governmental sites are in general the best in terms of accessibility and usability ratings compared to web sites of different extensions. The reason might be related to more strict enforcement by governmental agencies of the Section 255 of the Telecommunications Act to ensure that web sites maintained by the Federal governments are accessible and usable by most people (see Section 1.3 above). However, although the governmental sites are superior to other domain types, only half of the tested web sites (52%) were approved by Bobby.

Table 1 also shows high browser compatibility errors in all groups. One possible reason is that web site designers tend to rely on web design tools that are compatible with only one particular type of browser.

Figures 1 and 2 depicts the means of each web site group in terms of overall accessibility (Bobby's approval) and overall usability rating.

		.com N=23	.edu N=20	.gov N=21	.org N=25	Total N=89
Accessibility (Bobby)	Approval	0.13 (0.34)	0.25 (0.44)	0.52 (0.51)	0.24 (0.44)	0.28 (0.45)
	Priority 1	1.3 (0.93)	0.95 (0.69)	0.57 (0.75)	0.92 (0.64)	0.94 (0.79)
	Priority 2	3.39 (1.23)	2.2 (1.24)	2.67 (0.66)	2.8 (1.22)	2.79 (1.18)
	Priority 3	1.96 (0.21)	1.6 (0.68)	1.71 (0.56)	1.92 (0.4)	1.81 (0.5)
	Browser Error	13.13 (9.48)	7.75 (7.66)	6.81 (4.32)	9.16 (4.79)	9.31 (7.18)
Usability (LIFT)	Usability Rating	1.7 (0.63)	1.8 (0.7)	2 (0.63)	1.68 (0.56)	1.79 (0.63)
	Catastrophic	0.09 (0.29)	0	0	0.04 (0.2)	0.03 (0.18)
	Major	2.43 (1.5)	1.95 (1.23)	1.43 (1.12)	2.04 (1.4)	1.98 (1.36)
	Minor	2.91 (1.88)	3.25 (2.75)	3.33 (1.49)	3.64 (2.63)	3.29 (2.23)

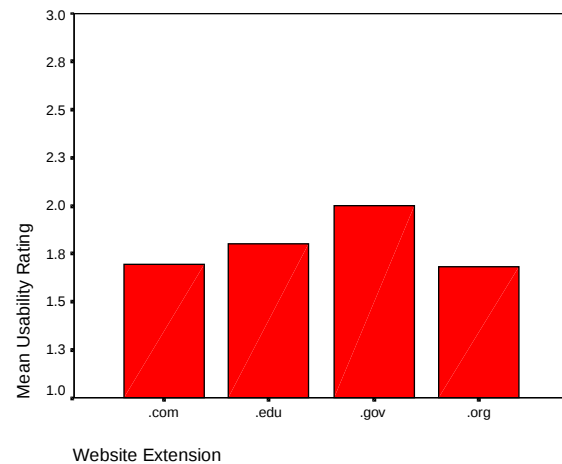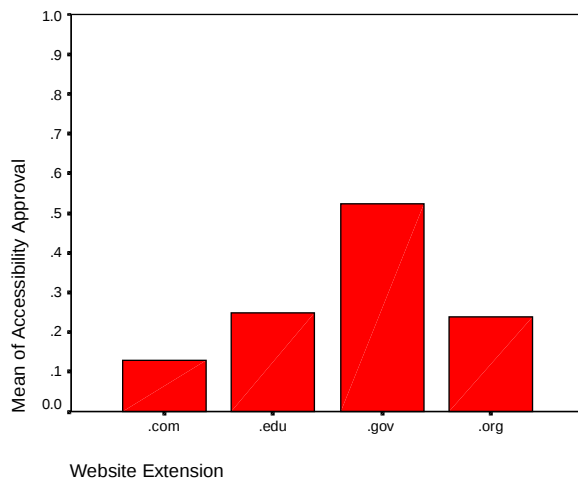Table 1: Mean Usability and Accessibility results (standard deviations in parenthesis)

Figure 1: The means of Bobby's Accessibility Approval by web site extension types

Figure 2: The means of LIFT's Usability Rating by web site extension types

The ANOVA follow-up analysis shows that the mean's differences between those groups of web sites are significant at $p<0.05$ for Bobby's approval ($F_{(3,85)}= 3.193$, $p=0.028$), Priority 1 errors ($F_{(3,85)}= 3.435$, $p=0.021$), Priority 2 errors ($F_{(3,85)}= 4.117$, $p=0.009$) and Browser compatibility errors ($F_{(3,85)}= 3.643$, $p=0.016$). On the other hand, all of the usability measures' group differences (LIFT) are not significant at $p<0.05$. Therefore, as measured by the automated evaluation tools used in the present study, the answer of the first research question is that web sites of different extensions differ significantly in their accessibility but not in their usability.

To answer the second research question, bivariate correlation of different usability and accessibility measures were observed. Observing all web sites (with all extensions), the accessibility approval correlates significantly with the overall usability rating ($\rho=0.531$, $p<0.01$). However, observing bivariate correlation of those measures within the same group, while the educational, governmental and organizational sites' usability and accessibility measures still correlate significantly, in the commercial group, that is not the case. LIFT's evaluation consists of two parts: accessibility issues and usability issues unrelated to accessibility. Since the commercial web site group is relatively high in usability rating yet low in accessibility approval, one can conclude that commercial sites perform high in usability issues unrelated to accessibility. In other words, commercial sites are user-friendly for people with no disability.

4. Conclusions

The analysis revealed that the governmental web site group has the highest compliance with Web site Content Accessibility Guide (WCAG) although only 52% got an approved status. The accessibility approval was found to correlate significantly with overall usability ratings for all groups, except the commercial web site group.

The present study brings about several implications for the practitioners. First, because some web sites' accessibility and usability measures are not predictive of each other, it opens a door into exploring the possibility of developing an integrated automated accessibility and usability evaluation tool. Second, the finding that most web sites did not receive the approved status from Bobby could be used to motivate web site designers to improve the accessibility and usability of web sites.

Further research could be conducted in several areas. First, in this study, only simple correlation and ANOVA were employed. Advanced statistical analysis such as structural equation modeling would be fruitful to explore the underlying relationship between different measures of usability and accessibility evaluation. Second, in the present study, the topic of interest is aging/health-related web sites. However, the methodology used in this paper could be applied in any area of interest (e.g. entertainment, e-commerce or, services).

Some limitations of using automatic evaluation tools need to be recognized:

1. There are important elements (such as the web navigation structure, the information's layout, the value of information, or various aesthetic aspects) which are not evaluated by the automatic tools.
2. The meaning/significance/appearance of graphics is not evaluated, only the inclusion of ALT tags are taken into consideration by Bobby and LIFT and only the number (higher number of graphics correlates to lower rating) of graphics is considered in LIFT.
3. Text-only web sites will get high ranking with both tools regardless of the quality of information or the readability of the fonts.

These limitations might imply that, although automatic evaluation tools provide a quick reference of the web site's accessibility and usability, formal usability evaluation involving user testing combined with a series of other non-empirical methods (such as cognitive walkthroughs or GOMS) still hold a major importance in the thoroughness of web site evaluation.

5. References

Mead, S. E., Spaulding, V. A., Sit, R. A., Meyer, B., & Walker, N. (1997). Effects of age and training on World Wide Web navigation strategies. *Proceedings of the Human Factors & Ergonomics Society 41st Annual Meeting* (pp. 152-156). Albuquerque, NM: HFES.

Sullivan, T. & Matson, R. (2000). Barriers to Use: Usability and Content Accessibility on the Web's Most Popular Sites. *Proceedings of the Conference of Universal Usability*. Arlington, VA: ACM.

Chrisholm, W., Vanderheiden, G., Jacobs, I. (2001) W3C Web Content Accessibility Guidelines (WCAG) [On-Line]. Available: http://www.w3.org/TR/WAI-WEBCONTENT

Zaphiris, P. (2000) Depth Vs Breadth in the Arrangement of Web Links", in *Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society*, HFES. San Diego, California, USA, August 2000.