RESEARCH ON WEB ACCESSIBILITY IN HIGHER EDUCATION

Terry Thompson (tft@u.washington.edu)
Technology Specialist, Computing and Communication and DO-IT
University of Washington

Sheryl Burgstahler, Ph. D. (sherylb@u.washington.edu)
Director, DO-IT Program; Assistant Director - Information Systems,
Computing & Communications; Affiliate Associate Professor, College of Education
University of Washington

Dan Comden (danc@cac.washington.edu)
Adaptive Technology Consultant, University of Washington Adaptive Technology Lab;
Technology Coordinator, DO-IT Program
University of Washington

ABSTRACT

Several studies have used the automated tool Bobby in evaluating the accessibility of postsecondary educational institutions’ websites to people with disabilities. However, determining many aspects of web-content accessibility requires human judgment. With the present study, two web accessibility “experts” manually evaluated critical web pages of 102 public research universities using a 5-point rating scale that focuses on each site’s “functional accessibility,” i.e., whether all users can accomplish the perceived function of the site. The results of each evaluator approximate a normal distribution, and the evaluators’ results are positively correlated with one another (r = .597), suggesting that the procedure has inter-rater reliability. Also, the evaluators’ combined results are positively correlated with those obtained by using Bobby on the same sample (r = .595). Using these evaluations, the researchers were able to identify a distinct cluster of promising practices in web accessibility.

RESEARCH ON WEB ACCESSIBILITY IN HIGHER EDUCATION

Information technology, particularly the World Wide Web, is playing an increasingly integral role in higher education for delivery of academic, administrative, and student services. Many of these services are delivered in a way that is inaccessible to people with disabilities. The inaccessibility of campus web pages is especially significant because of the increasing importance of web resources for college studies and the increasing numbers of people with disabilities who are attending postsecondary institutions (Henderson, 2001). In addition, federal legislation requires that an institution’s programs and services be accessible to qualified individuals. Specifically, Section 504 of the Rehabilitation Act of 1973 mandates that such access is assured within institutions that receive federal funds. The Americans with Disabilities Act (ADA) of 1990 reinforces and extends Section 504 to public programs and services, regardless of whether or not these programs and services are federally funded. According to these laws, no otherwise qualified individuals with disabilities shall, solely by reason of their disabilities, be excluded from the participation in, be denied the benefits of, or be subjected to discrimination in these programs and services, unless it would pose an undue burden to do so. Although Section 504 and the ADA do not specifically address access to technology-based educational offerings and resources, the United States Department of Justice (Patrick, 1996) clarified that the ADA accessibility requirements apply to programs...
offered on the Internet.

More recently, amendments to Section 508 of the Rehabilitation Act require that electronic and information technologies that federal agencies procure, develop, maintain, and use are accessible to people with disabilities, both employees and members of the public, unless it would pose an undue burden to do so. In response to this law, the U.S. Architectural and Transportation Barriers Compliance Board (Access Board) developed accessibility standards http://www.access-board.gov/sec508/508standards.htm (Office of the Federal Register, 2000) to which federal agencies must comply. The standards apply to desktop and laptop computers, websites and other Internet resources, videotapes and multimedia products, software, telecommunications products, and other electronic and information technology. There are sixteen standards that specifically address accessibility of “web-based intranet and internet information and applications.” These standards are based in part on the Web Content Accessibility Guidelines 1.0 (WCAG), developed by the World Wide Web Consortium (W3C). The WCAG include 14 guidelines, plus a total of 65 checkpoints that provide implementation details for the guidelines. The 65 checkpoints are assigned priority levels 1 through 3, with Priority 1 checkpoints the most critical for accessibility. The Section 508 standards are similar, but not identical, to the WCAG Priority 1 checkpoints.

Although Section 508 directly applies to federal agencies, many questions and issues have been raised regarding its applicability to other entities, including public postsecondary institutions. Some public postsecondary institutions (e.g., California Community Colleges Chancellor's Office, 2001) consider the Section 508 regulations applicable to their institutions; others (e.g., California State University Office of General Counsel, 2001) do not. Some institutions have adopted the standard voluntarily. Regardless, the accessibility standards developed for the federal government can serve as one model as educational institutions develop their own guidelines for the design, purchase, and use of accessible websites and other information resources.

At least in part because of the passage of Section 508 and its resultant standards for technology accessibility, postsecondary institutions are becoming increasingly aware of their legal obligations to provide access to programs and services offered on the Internet. This in turn has led to an increased effort among higher education entities to address their web accessibility problems. In doing so, postsecondary institutions are taking a variety of approaches. Some campuses are developing web accessibility policies, which reside at various levels within the institutional policy structure. Some are integrating web accessibility into existing mainstream web and/or information technology (IT) policies. Some institutions are providing specialized training on web accessibility to faculty and staff who design web pages; some campuses are integrating accessibility training into their existing mainstream training options; some are doing both. Some institutions are using automated tools to evaluate the accessibility of web content; some perform these tests only when requested by faculty, staff, or students; others are systematically doing so, using software that spiders their entire web server.

Many published studies have compared the accessibility of select web pages at institutions of higher education. Schmetzke has completed eight web accessibility studies (Schmetzke, 2002a), the three most recent being evaluations of the University of Wisconsin’s 13 four-year campuses (Schmetzke, 2002b), the 56 North American colleges that offer ALA-accredited programs in library and information science (Schmetzke, 2002c), and the home pages of

Each of these studies used the web accessibility evaluation tool Bobby™, developed originally by the Center for Applied Special Technology (CAST) and now owned by Watchfire. Bobby automatically evaluates the accessibility of web pages on a number of objective measures. However, many of the authors noted above report the shortcomings of this tool and of automated evaluation tools in general. As the World Wide Web Consortium (W3C) points out, “Some of the web-content accessibility checkpoints cannot be checked successfully by software algorithms alone. There will still be a dependence on the user's ability to exercise human judgment to determine conformance to the guidelines.” For example, Bobby might check to see if an ALT attribute is provided for a graphic element, but it cannot tell if the text will be useful to a person who is blind and using speech or Braille output technology. In addition, current automated tools were originally developed for HTML and are unable to handle the increasing variety of scripting languages and techniques used to develop web pages. Pages may receive false positive or false negative ratings simply because the automated tool has ignored or misinterpreted the code. Another shortcoming of automated tools is their inability to take into account the "severity" of an error. For example, if Site A is missing ALT tags on its spacer images and Site B is missing ALT tags on its menu buttons, both sites are rated identically, when in fact Site B clearly has the more serious accessibility problem. Cooper (2001) notes that Bobby's fully automatic checking of accessibility covers 27% of points relevant to accessibility; the remaining 73% must be verified by manual checks.

As Brajnik (2001) points out, “Determining what to measure is a difficult decision: often we focus on attributes that are convenient or easy to measure rather than those that are needed.” Since those web features that are measurable by Bobby are easy and convenient, the current body of web accessibility research has focused primarily on these variables.

Some studies have attempted to address Bobby’s shortcomings by combining use of Bobby with a manual process, often using assistive technology. Rowan, Gregor, Sloan, & Booth (2000) propose a standard accessibility evaluation method that combines the use of Bobby, a W3C HTML Validation Tool, and a thorough manual evaluation that includes both general and detailed inspections, use of assistive technologies, and a usability evaluation. Johns (2002) evaluated iPac 2.0, a web-based library catalog, using Bobby, as well as the W3C HTML Validator, a screen reader, screen magnification software, and a color blindness
simulator. McCord, Frederiksen, & Campbell (2002) also used Bobby, as well as a screen reader and speech recognition software, in their evaluation of selected web-based health information resources. Also, Erickson (2002) supplemented his Bobby evaluation of 10 job boards and 31 corporate e-recruiting websites with a “simulated application process” in which each of the sites in the sample was navigated using a screen reader and keyboard.

Three separate studies used exclusively manual processes in evaluating on-line library database services for accessibility. Axtell and Dixon (2002) manually checked one service for conformance to the Section 508 web accessibility standard. Byerley and Chambers (2002) documented the problems encountered in using screen readers with two products. Riley (2002) evaluated three products using a variety of assistive technologies. Horwath (2003) applied a unique method to her study: She administered a survey via email to 11 blind and visually impaired clients of the Canadian National Institute for the Blind Library for the Blind. In the survey, she instructed the subjects to perform specific tasks with various on-line library databases and asked them to rate the difficulty of the task. Subjects were also asked to rate the overall accessibility of each database.

In the present study, researchers have conducted a Bobby analysis and a proposed manual analysis. In contrast to previous studies, however, the present study has kept these methods completely separate in order to determine the correlation between results obtained by the two methods. An approach combining both manual and automated assessments may ultimately be the optimal solution. However, before assuming so, it is important to first establish that Bobby yields results that are generally consistent with those obtained manually by web accessibility experts. The bottom line with respect to web accessibility is whether an individual can perform a website’s intended function(s). As there will be varying degrees in the ease with which users can do so, such a measure does not lend itself to a binary “approved” or “not approved” rating. With this in mind, the evaluator of any web page should (a) identify its perceived intended function(s) and (b) rate the page on a scale that measures the ease with which any user, including a user with a disability, can perform the intended function(s).

Research Questions
The current study addresses the following research questions:

1. What is a reasonable system by which website accessibility could be manually evaluated? What is the correlation between the results of this system and those of Bobby?
2. Overall, how accessible are the key websites of postsecondary institutions with Carnegie classification “Doctoral/Research Universities–Extensive”? What is a reasonable method for choosing pages or sections of a site to test for accessibility?

METHOD

Participants
The sample for the study was public universities grouped under the Carnegie Classification “Doctoral/Research Universities–Extensive” (Carnegie Foundation, 2002). There are 102 universities that fall into this category. Choosing institutions from the same Carnegie classification assured that the sample was of a group of institutions that were homogeneous regarding key characteristics. The Carnegie system includes two classifications for
institutions granting doctoral degrees, intensive and extensive, which differ in the number of doctoral degrees offered. Institutions classified as extensive offer 50 or more doctoral degrees per year across at least 15 disciplines.

Universities in the sample host large quantities of web pages. There are, for example, over 200,000 pages on the primary web server at the University of Washington, with uncounted hundreds of thousands of additional pages on other campus servers. These pages receive over 2.5 million views each month. Within each institution, those individuals who seek to assure equal access to students with disabilities must make decisions regarding how to prioritize the daunting task of evaluating pages for accessibility and repairing those pages that are inaccessible. For example, responsible parties may decide to first address the accessibility of course pages in which students with disabilities are known to be enrolled; a second priority may be university pages that are systematically judged to be most critical. For the purposes of this study, the researchers sought to evaluate each university’s most critical websites, as these may indicate how completely the message of accessibility has permeated the institutional administration and information technology culture. The decision as to which university websites were “critical” was influenced primarily by traffic volume across a small sub-sample of universities from the research sample. The researchers’ primary resource was publicly available web traffic data for several universities in the sample. November 2001 data were used for comparison, as this was the only month that was consistently among the five busiest months for each institution. The following websites were consistently among the most requested sites for the sample universities. It can be reasonably assumed that these would be sites that would be frequently requested at most universities. The perceived function of each website is included in parentheses. Perceived function plays an important role in our research method as noted in the Methods section of this paper.

1. University home page (Read and navigate the entire page without missing important information.);
2. Search page (Search for any term and read the results.);
3. List of university colleges, departments, and/or degree programs (Read and navigate the entire page without missing important information.);
4. Campus directory of faculty, staff and/or students (Look up a name and read the results.);
5. Admissions home page (Read and navigate the entire page without missing important information.);
6. Course listings (Find a particular course listing and read all information about that course.);
7. Academic calendar (Read and navigate the entire calendar without missing important information.);
8. Employment home page (Read and navigate the entire page without missing important information.);
9. Job listings (Read and navigate the entire page without missing important information.);
10. Campus map (Locate a particular location on campus and identify how to get there.);
11. Library home page (Read and navigate the entire page without missing important information.).

Two additional sites were consistently among the most often visited but were excluded from the present research. The first, the entry page for the registration and records or enrollment system, was excluded because it generally has restricted password-protected access. The
second, the campus map page, was excluded because of differences across institutions in the perceived function for this site. Some institutions use campus maps to provide information about the buildings on campus and the divisions or departments that occupy them, whereas other institutions use campus maps exclusively to convey visual information about the campus layout.

In referring to each of these 10 categories of web content, the authors often use the term “site” rather “page” because evaluating the accessibility involved an evaluation of multiple pages. For example, evaluating the “search site” requires an evaluation of the page containing the search form, as well as the page that returns the search results. Both pages were considered and assigned one rating to the search site.

A total of 1013 sites from 102 institutions were included in the evaluation. Seven sites were excluded, because no site could be found that matched a particular category or because the site had restricted password-protected access. An evaluation for any one institution typically required fewer than 30 minutes.

**Procedures**

For the present study, the researchers developed a five-point scale to measure a site’s functionality. Note that the low end of the scale is 0, rather than 1, as the researchers in this study felt that a totally inaccessible site should receive a rating of 0. Points on the scale are not intended to be defined by specific guidelines or standards; some specific website characteristics are noted in parentheses after each rating as examples of types of issues that may warrant a rating, depending on context and overall site functionality. “Users” is defined as the majority of individuals, including those with disabilities using assistive technologies who access university websites.

4 = Accessible – Users can perform intended function(s) easily (Site must comply in principle with all Section 508 and WCAG Priority1 guidelines, including alternate text for all graphics, skip to main content link if needed, appropriate tags on forms and tables).

3 = Mostly accessible – Users can perform intended function(s), but site does not comply with all accessibility guidelines (e.g., missing ALT tags on “minor” images, missing <label> tags on “simple” forms, missing accessible table markup on “simple” two-dimensional tables, missing needed "skip navigation" links; these accessibility problems do not prevent most users from accessing the content).

2 = Fair – Accessibility problems exist, but skilled users can perform intended functions with patience.

1 = Mostly inaccessible – significant accessibility problems; skilled users can still perform many of the intended functions, but with considerable difficulty (e.g., complex tables or complex forms missing accessible markup).

0 = Inaccessible – not possible for some users to perform intended function(s) without assistance.

For each university in the current study, each site was evaluated on the above scale according to its perceived primary intended function(s).
After some training from the researchers, a student worker located and documented the URLs of critical sites for each of the universities in the sample. Evaluations were then conducted by two of the researchers. Each evaluator has at least 10 years of experience working in the technology accessibility field, and each is an expert on web accessibility, having given presentations and workshops on this topic locally, regionally, nationally, and internationally.

The two evaluators performed several initial evaluations together, using colleges and universities that were not included in the final sample. The purpose of these collaborative sessions was to assure that both evaluators had a similar understanding of the evaluation procedure and terminology and to identify and solve any initial problems with the evaluation method. Following the collaborative sessions, the evaluators worked independently to evaluate the study sample from October through December 2002. The universities were evaluated in alphabetical order, and the only communication between evaluators regarding specific sites was to (a) notify the other of a changed URL and (b) notify the other of current status. The latter communication was necessary in order for the evaluators to stay synchronized in their evaluations to assure they were evaluating the same site within a small time frame to minimize the chance that a site was updated or otherwise changed between the two evaluations.

The researchers adopted the following steps for evaluating the 10 sites at each institution:

1. Identify the perceived function of the site.
2. Using Internet Explorer, evaluate the site as follows:
   (a) Turn off images (using assigned hotkey), and make sure that all graphics have equivalent text.
   (b) If the page includes audio content, make sure this content is available through text equivalents.
   (c) Change the font size (larger and smaller) in the browser, and observe whether the page is still readable.
   (d) Set screen resolution to 640 x 480, or resize the browser window, and observe whether users can still access all information, with or without scrolling.
   (e) Examine how the site uses color, and ascertain that color is not used exclusively to convey information. Also examine the color contrast throughout the site. If inadequate contrast is suspected, print the page to ascertain whether all information is readable.
   (f) Tab through the links and form controls on a page, making sure that you can access all links and form controls without a mouse.
3. Use Jaws and observe whether the information available through a screen reader is equivalent to that available through a graphics-based browser. Note whether “skip navigation” links are present and functioning.

Additionally, one of the researchers conducted parallel evaluations using the Bobby web accessibility assessment tool. The researcher conducted a Bobby evaluation of a university immediately following completion of his manual evaluation of that same university. Bobby tests for compliance with various user-definable standards/guidelines. For the present study, Bobby was used to test for compliance with W3C Level A (compliance with all automatically measurable W3C Priority 1 checkpoints from WCAG 1.0). Bobby identifies the number of accessibility errors for each page assessed. For the present study, sites were assigned a rating of 0 if any accessibility problems were found and 1 if no accessibility problems were found. An overall Bobby Score was calculated for each university. This score
was a percentage of the university’s sites that received a rating of 1, i.e., were “Bobby approved at W3C Priority Level A.”

RESULTS

Overall Web Accessibility
Results for both evaluators approximate a normal distribution, with most sites falling within the 1–3 range. Across all websites at all 102 universities, the evaluators' combined mean rating was 2.26. The mean rating assigned by Evaluator 1 was 2.52, with a standard deviation of 1.17. The mean rating assigned by Evaluator 2 was 2.05, with a standard deviation of 1.00. A frequency distribution for each evaluator, showing frequency of each rating, is provided in Table 1.

<table>
<thead>
<tr>
<th>Evaluator 1</th>
<th>Rating</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>95</td>
<td>9.38</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>105</td>
<td>10.37</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>173</td>
<td>17.08</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>458</td>
<td>45.21</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>182</td>
<td>17.97</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1013</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluator 2</th>
<th>Rating</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>68</td>
<td>6.71</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>235</td>
<td>23.20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>332</td>
<td>32.77</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>336</td>
<td>33.17</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>42</td>
<td>4.15</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1013</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Inter-Rater Reliability
To measure the inter-rater reliability of the proposed functional accessibility evaluation method and accompanying 5-point scale, the researchers compared the raw evaluation data of the two evaluators. As seen in Table 1, Evaluator 1’s ratings had a higher variance, as he tended to be more liberal in assigning ratings of 0 and 4 than did Evaluator 2. However, the two evaluators were either in complete agreement or within one point of agreement on 845 of 1013 sites evaluated (83.42%). A summary of the rating differences between evaluators is
presented in Table 2. The Pearson Correlation Coefficient ($r$) for the two data sets is .597, significant at the 0.01 level (two-tailed).

Table 2
Frequency and percentage of all possible point differences between two evaluators on a 5-point scale

<table>
<thead>
<tr>
<th>Degrees of Difference</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>398</td>
<td>39.29</td>
</tr>
<tr>
<td>1</td>
<td>447</td>
<td>44.13</td>
</tr>
<tr>
<td>2</td>
<td>148</td>
<td>14.61</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>1.88</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.10</td>
</tr>
<tr>
<td>Total</td>
<td>1013</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Comparison Across Critical Site Categories
There was notable agreement in evaluator ratings when examined by critical site category (e.g., home page, search page). For each evaluator, mean ratings were calculated for each of the 10 critical site categories. Each evaluator’s mean scores were then ranked. There were slight differences in the evaluators’ exact rank values of site categories that ranked 2-9, although the same categories tended to fall either high or low in both evaluators’ rankings. Evaluators were in complete agreement on the two site categories which fell on each extreme of the ranked data set. Overall, the List of Colleges, Departments, and Programs sites tended to be the most accessible (mean of 3.0 for Evaluator 1, 2.54 for Evaluator 2), while the Course Listings/Timetable site tended to be the least accessible (mean of 1.84 for Evaluator 1, 1.33 for Evaluator 2). The full comparisons by critical site are provided in Table 3.

Table 3
Mean Ratings and Rankings by Critical Site, by Evaluator

<table>
<thead>
<tr>
<th>Critical Site</th>
<th>Evaluator 1</th>
<th>Evaluator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Rating</td>
<td>Rank</td>
</tr>
<tr>
<td>List of colleges/depts.</td>
<td>3.00</td>
<td>1</td>
</tr>
<tr>
<td>Academic calendar</td>
<td>2.75</td>
<td>2</td>
</tr>
<tr>
<td>Library home page</td>
<td>2.75</td>
<td>2</td>
</tr>
<tr>
<td>University home page</td>
<td>2.61</td>
<td>4</td>
</tr>
<tr>
<td>Campus Directory</td>
<td>2.60</td>
<td>5</td>
</tr>
<tr>
<td>Search site</td>
<td>2.59</td>
<td>6</td>
</tr>
<tr>
<td>Job listings</td>
<td>2.55</td>
<td>7</td>
</tr>
<tr>
<td>Employment home page</td>
<td>2.40</td>
<td>8</td>
</tr>
</tbody>
</table>
Admissions home page | 2.09 | 9 | 1.98 | 8  
Course listings | 1.84 | 10 | 1.33 | 10

**Comparison of Universities**
In order to compare the accessibility of websites of universities in the sample, the two evaluators’ data sets were standardized (converted to z-scores), and a combined data set was derived by averaging each standardized value in the two original data sets. Then, an Accessible Web Index (AWI) was calculated for each university by calculating the mean combined standardized rating across all of the university’s critical sites. Finally, universities were ranked by AWI.

The University of Texas at Austin (UT) was the only university with an AWI greater than one standard deviation above the mean (1.19). There were no universities with an AWI less than one standard deviation below the mean. For both evaluators, the total raw score for UT was higher than the total raw score for any other institution.

**Comparison of Manual System with Bobby**
Data obtained using the manual evaluation was further compared to that obtained automatically using Bobby. Researchers observed that those universities with higher AWI ratings tended to have higher Bobby scores than those universities with lower AWI ratings. The most obvious example of this is that the only two universities to receive Bobby scores of 100 were ranked 1 and 5 based on their AWI ratings. Given the high manual rating of UT as reported in the preceding paragraph, it is interesting to note that UT was one of the two universities with “perfect” Bobby scores.

Statistically, the Pearson’s Correlation Coefficient (r) between AWI ratings and Bobby scores was .595, significant at the 0.01 level (two-tailed). This correlation is of comparable strength to that between the two manual evaluations, which argues that, despite its limitations, Bobby may be an adequate tool for gaining an overall measure of an institution’s accessibility.

There were, however, some important discrepancies observed for individual universities. For example, one university ranked in the top 30 on its AWI, but only 11% of its sites were Bobby-approved. This discrepancy was due, at least in part, to a situation where Bobby identified accessibility problems but was unable to evaluate the functional severity of these problems, which was judged to be minor in the manual evaluations. Conversely, there were several universities whose Bobby scores were above average for the sample but whose AWI rankings were in the lowest quartile. These discrepancies were most likely cases in which Bobby, due to its inherent limitations, was unable to automatically identify accessibility problems that were judged to present significant barriers when evaluated manually.

**The University of Texas at Austin : A Promising Practice**
As reported in the Results section, The University of Texas at Austin (UT) was the only university whose AWI was greater than one standard deviation above the mean (1.19). Also, UT received the highest total raw scores in the sample from both evaluators and was one of only two institutions to receive a “perfect score” of 100 using Bobby. Peripheral to the present study, the researchers sought to better understand the factors that contributed to UT’s exemplary performance. An in-depth discussion of UT’s historical efforts on web
accessibility, as well as a review of its current policies, procedures, and support strategies, is available at http://www.washington.edu/doit/Research/Web/texas.html.

Evaluation Tool and Procedure
The present research proposed an alternative approach by which websites can be evaluated, particularly when comparing websites for research purposes. The approach differs from those used in previous research studies in that it employs a manual evaluation procedure focusing on the sites’ functionality rather than on the automatic, inherently limited assessments conducting using Bobby and other automated tools. The proposed manual approach was shown to have inter-rater reliability. The strength of the inter-rater correlation and the strength of the correlation between the manual evaluation and that of the automated assessment tool Bobby were both about .6. From these results one can conclude that the overall results of tests for accessibility using Bobby and the proposed manual method are strongly correlated. They also suggest that improvements might be made to strengthen the reliability of the manual assessment procedure.

The evaluators have identified possible reasons for the occasional differences in their ratings:

- Web fatigue–Evaluating many websites for accessibility can be a draining process. Although this evaluation was spread out over a three-month period, both evaluators still noticed changes in their rating tendencies after they had conducted several evaluations within a few hours, particularly if those evaluations included multiple inaccessible sites.
- Mood–Evaluators observed in the raw data a tendency to have “high days” and “low days”, specific days on which the majority of evaluations were either slightly higher or slightly lower than similar pages evaluated on other days.

It is recommended that future research be conducted in order to:

- Increase the inter-rater reliability of the 5-point scale for manual evaluation of accessibility/functionality of websites.
- Identify the optimal relationship between automatic and manual assessments in the evaluation process, i.e., how can the strengths of both approaches be fully utilized for maximum accuracy and efficiency?
- Explore the possibility of incorporating the concept of website functionality into new or existing web accessibility evaluation tools.
- Expand the scope of the present research to include educational entities in other Carnegie classifications, i.e., Master’s Colleges and Universities, Baccalaureate Colleges, Associate’s Colleges, Specialized Institutions, and Tribal Colleges and Universities.
- Track higher education web accessibility longitudinally, in order to identify patterns of improvement over time.
- Identify predictors of web accessibility in higher education entities. Possible predictors to explore might include institutional policies, procedures, and support strategies.

The State of Accessibility in Research Universities
In the present study most websites at the 102 major research universities were rated somewhere in between 4 (“accessible; student can perform intended function easily”) and 0 (“inaccessible; not possible for student to perform intended function without assistance”); scores of 0 and 4 were rare. These results suggest that for most sites it is possible for visitors with disabilities to access some or most web content at universities’ critical sites, although it
may be difficult and may require considerable patience and/or considerable expertise in using one’s assistive technology.

Researchers in this study observed that many sites could very easily improve their rating to 4 simply by practicing the most fundamental of web accessibility techniques, including the following:

- Add alternate text to images (using the HTML ALT attribute).
- Add a null ALT (ALT="") attribute to non–informative decorative or spacer images.
- Add a “Skip Navigation” link to the top of pages with extensive navigation.

Also, once accessibility is established, it is important to maintain it. Several universities, for example, had deployed a template that included a “Skip Navigation” link that was no longer functional. It was also noted that many sites that created a separate "text-only" version of a site or page did not keep it as up-to-date as the default page.

A few sites in the present study practiced particularly noteworthy techniques. Each of the sites described below includes a link to an archived page at the University of Washington. The researchers received permission from the authoring institutions to archive their exemplary pages, since web pages change frequently and any live URLs sited in the present context might otherwise have been modified or replaced by the date of this article’s publication.

The University of Kansas home page (http://www.washington.edu/doit/Research/Web/KU/) uses scalable fonts, which allows text size to be changed using browsers’ text resize functions. It also has a working “skip redundant navigation” link and an additional visible “toggle navigation” link, which turns all navigation links on or off. This latter feature is a nice example of universal design, as it benefits users with disabilities but also allows for a cleaner page for all users. All of these accessible features are built into the site template which is used on several other pages within the University of Kansas site. This is an excellent example of an automated tool such as Bobby being currently unable to evaluate the enhanced functionality of a page’s features.

The University of Kansas Admissions page (http://www.washington.edu/doit/Research/Web/KU/Admissions/), in addition to having a "skip to content" link, uses accessible form markup (i.e., HTML <label> elements are used to associate label text with its corresponding form fields). Without such markup, screen reader users can find it difficult to correctly identify the content that is being requested by form fields. Also, this page provides an excellent example of alternate text used effectively. Visually, the banner graphic on this page identifies the Office of Admissions and Scholarships, but it also includes a backdrop that uses a variety of colors and shows many smiling student faces to convey an atmosphere of enthusiastic learning. An ALT attribute that simply provides access to the text would not adequately convey the nonverbal message that is communicated to people who access this page visually. The ALT attribute used is "Graphical banner announces that this is the home of The University of Kansas's Office of Admissions and Scholarships.” In addition to providing access to the visual text, the action verb “announces” subtly conveys the enthusiasm depicted in the image.

Like the University of Kansas pages, the Campus Directory of the University of Texas at Austin (http://www.washington.edu/doit/Research/Web/Texas/Directory/) includes a "skip to
main content" link and accessible form markup. For the most part, it uses scalable fonts. It additionally uses HTML <H1> elements for section headings. Many sites rely merely on font size to visually suggest that text is a “heading” but don’t specifically identify it as such by appropriately using HTML. By using valid HTML as the University of Texas has done on this page, navigation is enhanced for site visitors with assistive technologies that support this markup. For example, some screen readers include functionality that allows users to jump between headings. Thus, screen reader users can get a better sense of how the page content is organized and can skip to sections that specifically interest them.

The University of Washington home page (http://www.washington.edu/doit/Research/Web/Washington/), like many pages, includes rotating images but it is unique in that the image’s corresponding alternate text also rotates. Most sites with rotating images use a single generic ALT attribute (e.g., ALT=”photo”). All characters on this page are actual text – rather than graphic depiction of text – formatted using cascading style sheets. Also, all text on this page is scalable so users can enlarge it as needed.

North Carolina State University includes “skip to content” links on many of their pages. The links at the top of the NC State Academic Programs page (http://www.washington.edu/doit/Research/Web/NCSU/) are noteworthy in that there are multiple links. In addition to "skip to main content,” additional links are provided to each of the main content headings, e.g., "skip to By College," "skip to Academic Resources,” etc.

The researchers were challenged to find a good example of an accessible site in the Course Listings category. As noted in the Results section and in Table 3, these sites were the least accessible in the present study. They typically provide an index or search interface whereby users can find relevant courses and can browse extensive database-driven detail about each course. The search forms are often highly complex, and few universities are using accessible form markup. Consequently, screen readers often read these forms incorrectly. The output is typically presented using a complex table with nested rows and/or columns. An accessible table is one in which a screen reader user can easily identify the column and row headers for a given table cell. Without this information, it’s difficult for blind users to track their position within the table, particularly tables with large quantities of information (e.g., long rows of fields containing data about class title, number, times, instructor, etc.). There are specific HTML techniques that allow screen readers to correctly read tables. However, few institutions are practicing these techniques on their course listings pages. In fact, many institutions present this content using <pre> elements, which provide a visual tabular layout by preserving spaces between table cells but do not provide any actual tabular structure, so screen readers simply read them as long lines of text, not as rows and columns of data.

One of the only institutions to provide an accessible Course Listings interface was Kent State, who did so via a separate text-only version of the content. This version includes accessible markup on its search form (http://www.washington.edu/doit/Research/Web/Kent/Search/) and presents a linear version of the results that is easily accessed by screen reader users (http://www.washington.edu/doit/Research/Web/Kent/Results/).

CONCLUSION
The present study was undertaken to create and test a manual system for evaluating websites for accessibility. The proposed manual system evaluates websites based on their primary function and addresses the limitations of automated accessibility testers such as Bobby. The present study also tested the accessibility of key websites of research universities and summarized successful policies and practices at the institution whose websites are most accessible. The results provide information that can help guide universities in their efforts to improve the accessibility of the websites on their campuses. Although few websites evaluated in this study were totally inaccessible, the opposite is also true: few were totally accessible. This confirms that continued effort is needed in order to educate administrators, faculty, and web designers about the need for web accessibility and the techniques for implementing it. As advocates continue their efforts to make the web in higher education more accessible, it is important to continue research that assesses progress toward this goal. More accurate and efficient methods for evaluating the accessibility of websites are critical to assuring that we have an accurate measure of web accessibility in higher education.

REFERENCES


Back to Table of Contents

Information Technology and Disabilities Homepage