IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITP4735

How Accessible are the 50 Most Visited Web Sites?

Jack Cook, Ph.D., CFPIM, Wasner Martinez, David Messina, Melissa Piggott
Rochester Institute of Technology
College of Business, 105 Lomb Memorial Drive, Rochester, NY 14623-5608
Phone: (585) 414-7334, Fax: (585) 475-7055
JackCook@hotmail.com

INTRODUCTION

There are many identifiable tools that assist persons inflicted with some form of disabling circumstance. Many of these are easily recognizable such as ramps, elevators, Braille, and closed-captioning. These features have also been used and desired by everyone, especially families with strollers who can take advantage of elevators and curb cuts. The absence of these simple yet effective features results in a direct violation of governmental regulations as well as human rights. Would you deem it morally acceptable to restrict the hearing impaired from watching television because television networks and manufacturers refused to implement closed-captioning? Most people would say no. How about the World Wide Web? Should disabled persons not be able to utilize the resources of the Internet simply because they are disabled?

The purpose of this paper is to explore the accessibility of the web today. In order to accomplish this goal, the researchers replicated a study conducted by Terry Sullivan and Rebecca Matson which was reported in "*Barriers to Use: Usability and Content Accessibility on the Web's Most Popular Sites,*" published in 2000. Sullivan and Matson determined the fifty most popular sites through the use of www.Alexa.com. Accessibility of these sites may have either worsened or improved between the years 2000 and 2002, when our research was conducted. The goal of this project was to evaluate whether or not and to what degree this had occurred. This is significant because web accessibility is a growing issue in human rights and should be addressed as technology progresses.

The analysis portion of this project was completed using *Bobby. Bobby* tests web sites according to two criteria: (1) The W3C (World Wide Web Consortium) guidelines. This refers to fourteen guidelines outlined by the W3C in an attempt to evaluate Web content accessibility. See http://www.w3.org/TR/WCAG10/ for the fourteen guidelines. (2) The US section 508 guidelines. This refers to rules that pertain to Section 508 of the Rehabilitation Act of the ADA, which are closely knit with the web content guidelines proposed by the WAI (Web Accessibility Initiative).

According to the article, *Evaluating Web Resources for Disability Access* (Rowan, p. 81), some limitations of Bobby include: (a) The results may be overwhelming due to their length and complexity, (b) manual inspection may be required in order to separate actual problems from potential barriers, (c) may deem an acceptably accessible site as inaccessible, and (d) does not offer any recommendations for making a particular site more accessible.

The *HTML validation tool* was also utilized to see if the web pages conformed to HTML standards This is significant because "another way to ensure that a web site has optimal accessibility is to use HTML to convey meaning and not format or layout" (Casey, p. 23). Accessibility deals with both individuals that are disabled, as well as individuals that may possess some other restriction. Individuals may also have a temporary problem (such as a broken arm that restricts the use of a mouse) which simulates the experience of other disabilities.

LITERATURE REVIEW

What is Web Accessibility?

Web accessibility deals with the ability to access the World Wide Web (WWW). Prior to 1998, accessibility did not specifically apply to the virtual world under Section 508 of the Rehabilitation Act. However, during 1998, President Clinton expanded its scope, deeming that federal web sites should be fully accessible to disabled individuals.

How are web sites inaccessible to disabled individuals? The most commonly discussed issue stems from navigational problems encountered while attempting to access the web via voice-enabled browsers. These browsers not only interpret pages auditorily but also perform navigational commands via voice recognition. Problems associated with this form of assistive technology are more often linked to the design of the site because "if pages are poorly designed, no amount of adaptive technology or browser tweaking will help". (Larkin, p. 142) Aspects of web pages that are not friendly to assistive technologies include:

- Rollovers that do not contain additional code so that a mouse is not required.
- Frames and tables are not interpreted correctly because these browsers read information left to right, top to bottom.
- Links may not be descriptive of their proposed action, such as "click here".
- Graphical images that do not include the "alt" tag. An "alt" tag provides a textual description for a graphical image.
- Actions and information represented by animations that do not include descriptive text.
- Actions solely represented by images or colors.

The Continuing Need for Improving Web Accessibility

According to the research conducted, it appears that as technology advances, it is more difficult for assistive technologies to decipher web pages because of the increased graphical nature of web pages. A page can be graphically rich AND be accessible if designed properly. The world is becoming more dependent on virtual mediums such as the WWW. The dilemma is "as the online world grows more graphical it becomes less accessible to disabled users." (Heim, p. 1). If the WWW remain inaccessible to disabled individuals, these people will ultimately be out of the mainstream of information and left with fewer choices.

RESEARCH

Prior to evaluating the web sites, it is imperative to discuss the priority levels created by the W3C. There are three levels and they each describe accessibility to a different degree.

- **Priority 1:** These checkpoints must be fulfilled. A failure to satisfy the Priority 1 guidelines implies that the page entails severe dilemmas with respect to accessibility for disabled individuals.

- **Priority 2:** These checkpoints should be fulfilled. A failure to pass this would deem the page as fairly accessible but would also highlight key access matters that should be addressed.
- **Priority 3:** These checkpoints may be fulfilled. "If you can pass all items in this section, including relevant User Checks, your page meets Conformance Level AAA for the Web Content Guidelines" (http://www.cast.org/bobby).

Sullivan and Matson's Research

Sullivan and Matson evaluated the 50 most frequently visited sites of the year 2000. **Methodology:** The scope of the project entailed the evaluation of eight guidelines that were titled under Priority 1, corresponding to guidelines 1.1, 1.2, 2.1, 6.1, 6.3, 7.1, 9.1, and 11.4. The researchers analyzed only the main page of the sites. This was because it would have been exhaustive to check each page within the sites and if the main pages were deemed inaccessible or poorly accessible, navigation through and onto other pages can be assumed to be similar. The researchers also evaluated the sites both automated and manually. The significance of this was to decipher between potential and actually areas of failure. Sites that contained text-only alternative home pages were evaluated solely on that text version. In order to display the results of the project more adequately, a four-tier model was used. These were ranked as follows:

- High Accessibility. This refers to sites that contained no perceived content access problems.
- Medium Accessibility (Mostly Accessibly). This refers to sites with few accessibility problems; five and less checkpoints identified.
- Medium Accessibility (Partly Accessibly). This refers to sites with many accessibility problems; between 5-10 checkpoints identified.
- Inaccessible. This refers to sites that contain major obstacles in the use of the page; more than 10 checkpoints identified.

Sampling: Sullivan and Matson took a sample of 50 web sites. These sites were significant because they represented elements of a purposive sample. That is to say, the sites were categorized as frequently visited. Therefore, they should be reasonably accessible to everyone.

Findings: Table 1 shows the results of their findings formatted in a four-tier table according to the rankings previously defined. The sites in the first tier are in alphabetical order because their accessibility ranking was identical. For tiers 2-4, the sites are listed in order of which was more accessible.

OUR RESEARCH

Methodology: Our research entailed an evaluation of the 16 checkpoints that are titled under Priority 1; as stated previously these guidelines must be satisfied in order to consider a site accessible. The purpose of focusing mainly on the Priority 1 guidelines is not only to more easily draw similarities to the research conducted by Sullivan and Matson but also to maintain consistency. Priority 1 guidelines also represent the minimum requirements to be *Bobby* approved. In an attempt to be as accurate as possible we ranked the sites' accessibility according to actual issues encountered by *Bobby*, which excludes user-checks. "User checks are triggered by something on the page; however, you need to determine whether they apply." (www.cast.org/bobby) These are usually checked manually because they represent potential, not actual violations. The significance of this is that the scope of the research solely deals with the automated evaluation of the web pages. The researchers furthered the previously conducted research by not only incorporating Web Content Accessibly Guidelines 1.0 but also the U.S. Section 508 Guidelines under *Bobby* and the *HTML validator* in order to be equipped to construct more detailed comparisons. As with the previous study, only the home page of each site was analyzed.

Table 1: Findings from Sullivan and Matson

Tier 1 Highly Accessible	Amazon, Gohip, Google, Goto, Hotbot, Microsoft, Monster, MSN, Snap
Tier 2 Mostly Accessible	AltaVista, Att.net, Excite, Icq, Tripod, Geocities, Lycos, Angelfire, Iwon, Yahoo, Infospace, Go
Tier 3 Partly Accessible	Dogpile, Looksmart, Preferences, Xoom, Bluemountain, Ebay, ZDNet, Netscape
Tier 4 Inaccessible	100free, Mp3, Homestead, Quicken, Ancestry, Webshots, Real, MSNBC, Freeserve, Cnet, About, Cnn, AOL, Hitbox, Askjeeves, Networksolutions, Ragingbull, Ign, Weather, Cdnow, This

Source: *Barrier to Use: Usability and Content Accessibility, p. 142*

This project will also rank the web sites in a similar four-tier model. However our technique of assigning the web sites with a specific rating will vary due to the manner in which we tested the sites. The sites will be ranked according to:

- Highly Accessible – site revealed no instances of Priority 1 guideline violations
- Mostly Accessible – site revealed only one instance of a Priority 1 guideline violation
- Partly Accessible – site revealed two instances of Priority 1 guideline violations
- Inaccessible – site revealed three or more instances of Priority 1 guideline violations

Sampling: Since this research was meant to be compared to the one previously conducted by Sullivan and Matson, we attempted to utilize their same sites to the best of our ability. Therefore the 50 most frequently visited sites of 2000, as obtained by www.Alexa.com, were again sampled. Due to the timely nature of this project a few of the web sites were no longer valid for analysis by 2002 — 3 sites were no longer in existence and 7 sites could not be successfully evaluated through *Bobby.* A possible reason is compatibility issues with JavaScript. In the cases mentioned above the researchers substituted those invalid sites with other sites that were relevant to the project scope. The substitute sites were frequently visited sites in their respective categories during the year 2002, according to www.100hotsites.com. It is important to note that www.Alexa.com temporarily does not carry a list pertaining to the top most frequently visited sites (see www.alexa.com/exec/faqsidos/help/index.html/index=3); hence the reason our research used www.100hotsites.com to find substitute sites. The list of sites broken out by category is:

Tier 1: Highly Accessible

FBI.Gov, Google, Microsoft, MSN

Tier 2: Mostly Accessible

AltaVista, AskJeeves, Att.Net, CDNow, CheapFares, Dogpile, Excite, Geocities, GoTo, HomeStead, Hotmail, ICQ, InfoSpace, Iwon, LookSmart, MTV, NetworkSolutions, Priceline, Quicken, Weather, WebShots, ZDNet,

Tier 3: Partly Accessible

100Free, About, Ameritrade, Ancestry, AngelFire, BlueMountain, BusinessNow, CNET, Ebay, Etrade, FreeServe, Go, GoHip, HitBox, HotBot, HotJobs, Tripod, Yahoo

Tier 4: Inaccessible

Amazon, EOnline, IGN, Lycos, MSNBC, Netscape

Findings & Interpretation: Generally, we encountered negative results with respect to the overall accessibility of the web's most frequently visited sites. The rankings of the fifty web sites were all placed in alphabetical order because no site has precedence over the other since our methodology was based on a specific number of errors that needed to occur in order to be categorized in a specific tier.

Overall no sites were considered to be *Bobby* approved. However 8% of the sites could be deemed as highly accessible since the analysis did

not reveal any instances of Priority 1 violations. This implies that amongst these 50 frequently visited sites most, that is 92%, possessed an actual significant barrier in respect to its accessibility for disabled individuals.

Eighty percent of the sites had medium accessibility, with 44 percent classified as mostly accessible, and 36 percent are partly accessible. The remaining 12 percent of the sites were considered inaccessible.

The sites were also evaluated according to the U.S Section 508 guidelines. Among these 50 sites it was found that 96% of the sites were not 508 approved, while 100% of the sites did not pass the *HTML validation*. The U.S Section 508 guidelines are less restrictive than those of the W3C. Hence, 4% of the sites that were not approved by *Bobby* were Section 508 approved.

Frequently violated guidelines were Priority Guidelines 1.1 and 12.1. The guidelines are described as follows:

- **Guideline 1.1**: "Provide a text equivalent for every non-text element" (Chisholm, pp. 39-40). 92% of the sites violated this guideline at least once. This problem can be fixed easily and with negligible amount of time.
- **Guideline 12.1**: "Title each frame to facilitate frame identification and navigation" (Chisholm, p. 46). 20% of the sites violated this guideline at least once.

COMPARISON

Table 2 compares our results with the previous study of Sullivan and Matson (2000). It depicts the specific guidelines that were violated, as well as their change, if any, over the past two years. Included are also the sites that we had to replace and their respective results.

From Table 2, nine sites were considered highly accessible during the year 2000; of these sites two were no longer valid for the evaluation. Three sites however remained highly accessible in 2002. On the other hand, four sites moved into a lower tier classification. Therefore, of the seven working sites, 57% of these became less accessible between 2000 and 2002. The researchers suspect that the magnitude of the decline in accessibility can be attributed to the move towards graphically rich pages. The accessibility trends of the highly accessible sites of 2000 are shown in Figure 1). This graph illustrates the specific number of sites (in their respective tier), whose ranking either changed or remained the same in comparison to the previous research conducted by Sullivan and Matson.

The sites that were classified as mostly accessible in 2000 had similar trends to those that were highly accessible in that same year. All twelve of these sites were valid for the 2002 evaluation. Of these twelve sites, seven remained in the mostly accessibly bracket. Of the sites where accessibility was considered to have changed between 2000 and 2002, 100% fell into a lower tier of accessibility.

There were eight partly accessible sites of 2000; two of these were no longer valid for the 2002 appraisal. Two of these sites remained in the partly accessible classification while four shifted to other brackets. Three of these sites improved their accessibility. However, one site's research accessibility level slipped into a lower classification.

The benchmark project conducted in 2000 ranked 21 sites as inaccessible. Of these sites six were no longer valid for the present evaluation. Two sites remained inaccessible to disabled users, while thirteen

Table 2: Comparison of Sites from 2000 and 2002

Highly Accessible (Year 2000)	Number of Priority 1 Errors (Year 2002)	Improved, Worsened, or Remained the Same?
Amazon	3 (Guideline 1.1: 3 Instances)	Worsened
GoHip	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)	Worsened
Google	0	Remained the Same
GoTo	1 (Guideline 1.1: 1 Instance)	Worsened
Hotbot	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)	Worsened
Microsoft	0	Remained the Same
Monster	N/A	N/A
MSN	0	Remained the Same
Snap	N/A	N/A

**N/A means that the site no longer exists, or that there were errors when Bobby ran it.*

Mostly Accessible (Year 2000)	Number of Priority 1 Errors (Year 2002)	Improved, Worsened, or Remained the Same?
Altavista	1 (Guideline 1.1: 1 Instance)	Remained the Same
Att.net	1 (Guideline 1.1: 1 Instance)	Remained the Same
Excite	1 (Guideline 1.1: 1 Instance)	Remained the Same
ICQ	1 (Guideline 1.1: 1 Instance)	Remained the Same
Tripod	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)	Worsened
Geocities	1 (Guideline 1.1: 1 Instance)	Remained the Same
Lycos	3 (Guideline 1.1: 2 Instances Guideline 12.1: 1 Instance)	Worsened
Angelfire	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)	Worsened
IWON	1 (Guideline 1.1: 1 Instance)	Remained the Same
Yahoo	2 (Guideline 1.1: 2 Instances)	Worsened
Infospace	1 (Guideline 1.1: 1 Instance)	Remained the Same
Go	2 (Guideline 1.1: 2 Instance)	Worsened

**N/A means that the site no longer exists, or that there were errors when Bobby ran it*

Partly Accessible (Year 2000)	Number of Priority 1 Errors (Year 2002)	Improved, Worsened, or Remained the Same?
Dogpile	1 (Guideline 1.1: 1 Instance)	Improved
Looksmart	1 (Guideline 1.1: 1 Instance)	Improved
Preferences	N/A	N/A
Xoom	N/A	N/A
Blue Mountain	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)	Remained the Same
Ebay	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)	Remained the Same
ZDNet	1 (Guideline 1.1: 1 Instance)	Improved
Netscape	3 (Guideline 1.1: 2 Instances Guideline 12.1: 1 Instance)	Worsened

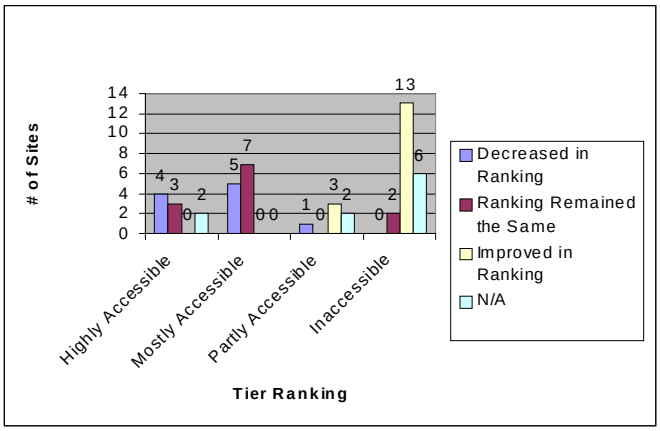
**N/A means that the site no longer exists, or that there were errors when Bobby ran it.*

Inaccessible (Year 2000)	Number of Priority 1 Errors (Year 2002)	Improved, Worsened, or Remained the Same?
100Free	2 (Guideline 1.1: 2 Instances)	Improved
Mp3	N/A	N/A
Homestead	1 (Guideline 1.1: 1 Instance)	Improved
Quicken	1 (Guideline 1.1: 1 Instance)	Improved
Ancestry	2 (Guideline 1.1: 2 Instances)	Improved
Webshots	1 (Guideline 1.1: 1 Instance)	Improved
Real	N/A	N/A
MSNBC	3 (Guideline 1.1: 2 Instances Guideline 12.1: 1 Instance)	Remained the Same
Freeserve	2 (Guideline 1.1: 2 Instances)	Improved
CNET	2 (Guideline 1.1: 2 Instances)	Improved
About	2 (Guideline 1.1: 2 Instances)	Improved
CNN	N/A	N/A
AOL	N/A	N/A
Hitbox	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)	Improved
Ask Jeeves	1 (Guideline 1.1: 1 Instance)	Improved
Network Solutions	1 (Guideline 1.1: 1 Instance)	Improved
Raging Bull	N/A	N/A
IGN	3 (Guideline 1.1: 2 Instances Guideline 12.1: 1 Instance)	Remained the Same
Weather	1 (Guideline 1.1: 1 Instance)	Improved
CDNow	1 (Guideline 1.1: 1 Instance)	Improved
This	N/A	N/A

**N/A means that the site no longer exists, or that there were errors when Bobby ran it*

10 Sites Added for Consistency	Number of Priority 1 Errors (Year 2002)
FBI.Gov	0
Priceline	1 (Guideline 1.1: 1 Instance)
MTV	1 (Guideline 1.1: 1 Instance)
Hotmail	1 (Guideline 1.1: 1 Instance)
CheapFares	1 (Guideline 1.1: 1 Instance)
HotJobs	2 (Guideline 1.1: 1 Instance Guideline 12.1: 1 Instance)
ETrade	2 (Guideline 1.1: 2 Instances)
BusinessNow	2 (Guideline 1.1: 2 Instances)
Ameritrade	2 (Guideline 1.1: 2 Instances)
EOnline	3 (Guideline 1.1: 3 Instances)

Figure 1: Accessibility Trends of the Highly Accessible Sites of 2000



sites moved into different brackets. These sites showed a general improvement concerning accessibility, revealing that 86% of the valid sites improved as opposed to remaining relatively constant. However, it is important to note that these sites still have a considerably high level of accessibility problems in 2002.

CONCLUSION

"One in five Americans have some disability; as the country ages, that percentage is expected to increase" (Heim, p. 182). This figure is shocking. Many of the sites that previously possessed satisfactory levels of accessibility worsened by 2002. Sites may need incentives in order to remain or become web accessible. The more publicity drawn to the issue of web accessibility and listings pertaining to 'accessible friendly' sites may curb this problem by generating some form of acknowledgment by web sites.

Many web sites appear to be improving accessibility. However, this improvement has not been significant. Communication via the Internet has expanded rapidly over the past few years. Therefore it would be safe to assume that the web's popularity has grown significantly, and one is expected to utilize various virtual mediums such as email. Barriers to access limit disabled individuals. This ultimately places these individuals in an overall disadvantageous position. Just as with buildings, all citizens should have access to the World Wide Web.

REFERENCES

Casey, C., (March 1999), Accessibility in the virtual library: Creating equal opportunity web sites. **Information Technology and Libraries, 18**(1), pp. 22-25.

Chisholm, W., Vanderheiden, M. G., Jacobs, I., (July/August 2001), **Web Content Accessibility Guidelines 1.0,** University of Wisconsin, Madison, pp. 35-53.

Heim, J., (Sept. 2000), Locking out the disabled, **PC World Magazine, 18**(9), pp. 181-185.

Larkin, M., (July 8, 2000), Web gears up for people with disabilities. **The Lancet, 356**(9224), pp.142.

Rowan, M., Gregor, P., Sloan, D., Booth, P., (2000), Evaluating web resources for disability access, **ACM**, pp. 80-84.

Sullivan, T., Matson, R., (2000), Barriers to use: Usability and content accessibility on the web's most popular sites, **ACM,** pp. 139-144.

http://www.100hotsites.com/
http://www.alexa.com/
http://www.cast.org/bobby
http://validator.w3.org/
http://www.w3c.org/