Behaviour & Information Technology, Vol. 24, No. 6, November-December 2005, 407–417

Taylor & Francis
Taylor & Francis Group

A retrospective look at website accessibility over time

STEPHANIE HACKETT†*, BAMBANG PARMANTO† and XIAOMING ZENG‡

†University of Pittsburgh, Pittsburgh, PA 15260, USA, + 1 412-383-6650

‡East Carolina University, Greenville, NC, USA

Websites were retrospectively analysed to study the effects that technological advances in web design have had on accessibility for persons with disabilities. A random sample of general websites and a convenience sample of US government websites were studied and compared for the years 1997–2002. Web accessibility barrier (WAB) and complexity scores were calculated. Analysis of Variance (ANOVA) and Tukey's HSD were used to determine differences among years for general sites. Repeated measures of ANOVA were used to analyse trends in US government sites, and Pearson's correlation coefficient (r) was computed to evaluate the relationship between accessibility and complexity. Random websites become progressively inaccessible as complexity increases. US government websites remain accessible while increasing in complexity. Increasing complexity, often caused by adding complex components to a Webpage, doesn't have to contribute to increasing barriers to accessibility: US government websites remain accessible despite increasing complexity by limiting the number of scripts used in Webpage design.

1. Introduction

The world wide web (WWW) is an ever-popular way to share information. Constantly emerging technologies introduce new ways of presenting this information. These new technologies also put forth challenges to maintaining web accessibility for persons with disabilities, as many of these advances include multimedia components that present barriers to persons with varying disabilities. Utilising the Wayback Machine (InternetArchive 2003), a service from the Internet Archive and Alexa® Internet, websites were retrospectively evaluated to see how they have changed over time with respect to accessibility and the use of more complex webpage design components. The Wayback Machine allows one to look at websites as they existed at the time they were archived. The authors look at a sample of websites from 1997 to 2002.

2. Background

When the web first entered mainstream use it was primarily text-based. A blind person could access most of it easily through text-to-speech software. As webpage design has evolved, however, and web designers have started to include images, tables, animated Java applications, and streaming audio and video to organise information in more complex ways (Bucy *et al.* 1999, Amtmann *et al.* 2000, Heim 2000, USDOJ 2003a), the web has become saturated with obstacles for the blind user and for users with other disabilities.

Accessibility, when pertaining to a webpage, means that information has been made available for use by almost everyone, including persons with disabilities. This accessibility may be direct or through the use of assistive technologies. Web accessibility varies depending on the type of disability. Low-vision users might require a large font with a sharp contrast between the background and foreground colour, whereas colour-blind users may need to have colour-transmitted information translated into distinguishable shades of grey. Blind users may be accessing webpages using a screen reader, a type of assistive technology that translates text displayed on the computer screen into synthesised speech (FreedomScientific 2004, IBM 2004, WebAim 2004). Physically impaired

*Corresponding author. Email: srhst18@pitt.edu

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2005 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01449290500066661

users might need to navigate with a non-traditional input device.

Worldwide there are more than 300 million persons with a disability (WHO 2004), of which an estimated 180 million are visually disabled and 250 million have a disabling hearing impairment. According to the US Census Bureau (USCensus 2003), 19.3% of the US population have a long-lasting condition or disability. In the European Union (Eurostat 2001), 14.5% have a moderate or severe disability. Persons with disabilities have historically been segregated and denied opportunities that non-disabled persons take for granted. With regard to technology and the Internet this could be due, in part, to a lag in technological advances in assistive technology as compared to the advances in web application technology and design.

Many governments around the world have started to address this 'digital divide' issue. In the US, the federal US government acknowledges this with the Rehabilitation Act Amendments of 1998, which covers access to federally funded programmes and services. These amendments (USAccessBoard 2003a), known as Section 508, set standards requiring that all electronic and information technology (including the Internet and the web) developed or purchased by the US federal government be accessible by persons with disabilities. The UK's Disability Discrimination Act (DDA 2004) protects persons with disabilities against discrimination and led to the development of the Disability Rights Commission (DRC 2004), which plays a critical role in the advancement of disability rights in the UK. In 2004, the DRC investigated web accessibility in the UK and published a report (Disability Rights Commission 2004) on their findings along with recommendations for improving accessibility for persons with disabilities.

The World Wide Web Consortium (W3C), the standards-setting body for the web, has developed the Web Content Accessibility Guidelines (WCAG) (Chisolm *et al.* 2001). The WCAG is an internationally accepted standard consisting of 14 guidelines that provide specifications on how to develop accessible webpages. Each guideline includes a list of checkpoints, totalling 65, for evaluating webpages for their degree of accessibility to persons with visual, hearing, physical, or cognitive disabilities. The guidelines define accessibility at the level of web content, discussing accessibility issues and providing accessible design solutions. While not being an all-inclusive solution for every person with a disability, they address typical scenarios that may pose problems for users with certain disabilities (Chisolm *et al.* 2001). Version 1 of the guidelines was published in May 1999 and primarily addresses the needs of visually impaired users. Version 2, not yet finalised, will address more disabilities in depth.

The WCAG checkpoints are organised into levels of priority. Priority 1 checkpoints must be met to prevent lack of access for some groups of users. Priority 2 checkpoints should be met to prevent difficulties in access for some users, while Priority 3 checkpoints may be met to improve access to web documents (Chisolm *et al.* 2001). Web designers can use WCAG Conformance Logos to claim conformance to a specific level of the WCAG. Level 'A' implies that all Priority 1 checkpoints have been met, while 'AA' and 'AAA' imply that Priority 1 and 2 checkpoints and Priority 1, 2, and 3 checkpoints have been satisfied, respectively.

3. Accessibility barriers

Hypertext mark-up language (HTML) is the non-programmable technology used to create most websites. It can be augmented with other features (IBM 2003) and it is these features that add complexity to webpages and are often the barriers to accessibility. It is important to include text equivalents to non-text elements (such as images and multimedia objects) so that these textual descriptions can be rendered as speech or Braille output by assistive technologies, enabling access to the content regardless of disability or device constraints (USDOJ 2003b, W3C 2003).

Many of the design technologies that pose barriers to persons with disabilities are those that allow multimedia to be included in the website. Multimedia requires users to rely on more than one sense, usually vision and hearing, to fully comprehend the information being presented. This poses obvious barriers to someone with a disability affecting either of these senses. Additionally, some of these technologies require plug-ins, software add-ons that allow third-party companies to develop special content normally unavailable using HTML (USDOJ 2003b), and each plug-in may require a distinct strategy for accessibility (Foley and Regan 2003). Other technologies that tend to pose accessibility barriers include: Java applets, scripting languages and tables. Similarly, Java applets and scripting languages are used to create dynamic and interactive webpage effects or provide video and audio output. Tables pose special problems because of their multidimensional nature, with layout providing meaning to the content. Screen readers may skip empty cells in tables without warning (Amtmann *et al.* 2000) and information may not be reproduced in the intended order (Goodwin-Jones 2001, Pontelli *et al.* 2002, USAccessBoard 2003b).

4. Accessibility evaluation

The oldest and most well-known of the web accessibility evaluation tools is Bobby (Watchfire 2004), based on the WCAG and Section 508 guidelines (Watchfire 2004). Bobby provides an absolute rating of either 'approved' or 'not approved' based on automated and manually evaluated checkpoints. The web accessibility barrier

(WAB) score (Parmanto and Zeng 2005) provides a proxy of web accessibility that is derived by evaluating 25 of the WCAG checkpoints that can be automatically checked. By using checkpoints that allow for automated accessibility evaluation, a large number of websites can be consistently evaluated while minimising time and cost. The WAB score provides a continuous range of values, with a lower score indicating better conformance with WCAG. This measure is explained in greater detail in the Methods section.

Websites are constantly being altered to incorporate new features and technologies in webpage design, with the average life being approximately 100 days (Yaukey 2003). Using the Wayback Machine (InternetArchive 2003), a service from the Internet Archive and Alexa® Internet, one can compare archived versions of the website and see changes that have been made over time. An entire website is captured at a particular point in time by the Internet Archive's web crawlers and stored as it appeared at the point of capture. Any website that is available to the public has the potential to be archived in the Wayback Machine, but typically excluded websites are: those sites that are password-protected, blocked by the webmaster to be unavailable to web crawlers, or otherwise inaccessible to the Archive's automated systems (InternetArchive 2001).

4.1 *Related work*

A report funded by the PricewaterhouseCoopers Endowment for the Business of Government (Stowers 2002) found Section 508 regulations slow to take effect. Using Bobby, 148 major US federal websites were examined for accessibility. Although Section 508 required federal agencies to be in compliance as of June 2001, only 13.5% of the websites examined complied with the requirements a year after that deadline. An earlier study (West 2001) conducted by a Brown University researcher, also using Bobby, found that 37% of the US government websites were accessible.

A recently completed study (RINCE 2003) of accessibility of the 30 most popular French websites found that none of the 30 websites met conformance level 'A'. A similar study (McMullin 2002) conducted in Ireland found that 94% of the 159 websites tested failed to meet the minimum accessibility standard ('A'), and not one website met the guidelines of levels 'AA' and 'AAA'. Lazar *et al.* (2003) used Bobby to conduct a study on private and non-profit websites and found that 49 of 50 websites were found to have accessibility problems. Over time those websites became even less accessible (Lazar *et al.* 2004).

Parmanto and Zeng (2005) conducted an evaluation on a large sample of websites that considered themselves as accessible and self-rated (as per the icon present on their website) as 'A', 'AA', or 'AAA'. They found that even among websites that were self-declared as meeting 'AAA' conformance, only 8.81% were truly 'AAA'.

A usability study of blind web users was conducted by Zeng (2004) as part of his dissertation work. This study employed WCAG checkpoints to a Web Transcoder Gateway (WTG) server that complied with guidelines to remove barriers and transform original pages into accessible pages. Users were asked to access webpages and perform tasks on the original page and on the transformed page via the WTG server. This study found that users accessing the website via the WTG server found information more efficiently and with fewer errors than those accessing the information via the untransformed pages. Users also felt more satisfied, more confident and less frustrated while accessing pages using the WTG. This study shows that using WCAG to design accessible websites can make a difference for persons with visual disabilities.

5. Objectives and scope

The main objective of the study is to determine how changes in web design have affected accessibility over time. To analyse this, the authors utilise a proven metric (WAB score) to assess accessibility and employ the complexity algorithm to assess complexity of the website. Two categories of websites are compared: a random sample of general websites was selected for each of six years (1997–2002), with each year's sample being different; and compared to a convenience sample of US government websites, following the same websites through all the years.

6. Methods

6.1 *Selection of websites*

The unit of analysis in the study is the individual website. Because the number and distribution of websites are undeterminable due to the size and dynamics of the web, many probabilistic sampling methods are not applicable. An alternative sampling method was adopted for this study: a convenience sample was taken from a directory service provided by Alexa® Internet.

To formulate the list of websites for the random category, a list of the Top 500 ranked websites was obtained from www.alexa.com on 28 July 2003. Exclusion criteria for a website were: non-English as the primary language, no longer in service or could not be located, or contained explicit adult content. Of the Top 500, 221 (44%) had the potential to be included in the random population.

By completing a Wayback Machine search for each of the 221 websites, it was noted for which years the website is archived. The number of times a website is archived varies for each website and for each year, ranging from no archived instances in a particular year to several hundred. For a website to be included in the random category, it must have at least one archived instance during the period 1997–2002. From this information, lists were assembled for each year containing all websites that were archived in that year.

Using SPSS® for Windows, a random sample of 27 websites for each year was obtained from the aforementioned lists. Since many websites are short-lived, the authors felt that following the same sample of random websites through all years was not a true representation of the web and would bias the category towards websites that have longevity. An archived instance was then collected from the Internet archive for each website. For convenience, the first archived instance for each year was used. If the first instance was unable to be used (due to a web crawling exclusion or Internet archive error such as 'path index error', 'failed connection error', or 'file not found error'), the next archived instance for the year was used. Again for convenience, if an instance was unobtainable for a website, the website was replaced by the website that was next in the alphabetical list in the random population. Examples of websites in the random population include: www.aol.com, www.apple.com and www.matchmaker.com.

For purposes of comparison, a sample of US government websites was also obtained. The US government websites list was obtained from www.100topgovernmentsites.com (100 Top Government Sites 2003) on 15 August 2003. United States government websites included were limited to those ending in the postfix '.gov' and there were 32 unique '.gov' websites on this list. To obtain the sample of US government websites, only websites that were archived each year (1997–2002) were used. For convenience, the first archived instance for each year was used. If the first instance was unavailable, the next instance was used. If there was not an instance for a particular year, the website was not included in the study. By including only websites that had archived instances in each year, it is possible to analyse trends in the US government category. There were a total of 22 (69% of the unique list) US government websites included in the study. Websites in the US government category included: www.usda.gov, www.fbi.gov and www.cdc.gov.

6.2 *Measurement for evaluating accessibility*

The accessibility of each archived instance was measured using the web accessibility barrier score (Parmanto and Zeng 2005). Each measure is a site-based measure that includes the homepage and one level from the homepage (link from the home page to another page of the website). There are many previous studies on web accessibility. Almost all of these studies use only the homepage to evaluate website accessibility (Flowers *et al.* 1999, Davis 2002, Diaper and Worman 2003, Lazar *et al.* 2003). Jakob Nielsen (Nielsen 2000) argues that the homepage is the gateway to the site and therefore sets the tone of the entire website for the user. Evaluating web accessibility beyond the homepage requires extensive work. This study is one of the most extensive studies ever conducted on web accessibility. Future studies involving all levels of a website will be conducted. The best assessment of website accessibility is the average of the scores for all pages making up the website.

Researchers at the University of Pittsburgh developed the accessibility metric used in this study (Parmanto and Zeng 2005). The metric was developed with the intentions of overcoming the deficiencies of the current measurements used in web accessibility studies. The current rating system and the so called 'Bobby Approved' measurement reflect an absolute measure of accessibility: either the website conforms to all checkpoints or it is considered inaccessible. The new metric calculates a quantitative score that provides a continuous range of values ranging from perfectly accessible to completely inaccessible. This allows for assessment of changes in web accessibility over time and for comparison between websites or between groups of websites.

The metric (figure 1) is a proxy indicator of web accessibility and looks at 25 checkpoints that can be automatically evaluated, based on WCAG and Section 508 guidelines. Based on accepted standards, the metric measures accessibility for persons with disabilities.

This measure looks at the actual violations of the page and normalises them against the potential violations. For example, if the checkpoint looks at the number of images without alternative text, the number of violations would be the actual number of images without alternative text while the number of potential violations would be the number of images within the page. The measure utilises the checkpoint priorities in reverse. Priority 1 violations weigh three times

$$WABScore = \frac{\sum_{p}\sum_{v}(\frac{n_v}{N_v})(w_v)}{Np}$$

p : Total pages of a website
v : Total violations of a Web page
n_v : Number of violations
N_v : Number of potential violations
Wv : Weight of violations in inverse proportion to WCAG priority level.
Np : Total number of pages checked

Figure 1. The WAB formula.

more than a Priority 3 violation, since Priority 1 violations pose more difficulties in accessibility than Priority 3 violations. The WAB score for each website is the summed WAB score of the webpages normalised against the total number of pages. A higher WAB score means more accessibility barriers exist. A lower score means better conformance with WCAG guidelines. A score of zero denotes that the website does not violate any web accessibility guidelines and should not present any accessibility barriers to persons with disabilities.

A previous work (Parmanto and Zeng 2005) proposes the WAB as a novel metric for measuring content accessibility of the web for persons with disabilities. This study, which includes reliability and validity testing of the metric, found that the metric provides a good representation of a website's accessibility. The study calculated the scores of approximately 1 141 rated (accessible) websites with 'AAA', 'AA', and 'A' conformance levels; and 500 random non-rated (non-accessible) websites. Scores of the WAB metrics provide continuous 'degrees' of accessibility. The average scores of 'AAA', 'AA', 'A', and non-rated websites are 2.02, 2.74, 4.47 and 10.5, respectively. Further analysis revealed that a threshold exists between accessible and non-accessible websites. A threshold of 5.5 separates accessible websites from non-accessible ones with a high degree of accuracy (with area under the curve of 0.962 as measured using Receiver Operating Characteristics [ROC] Curve (Egan 1975)). A WAB score of 5.5 or less means a website has better conformance to the WCAG and contains few barriers to accessibility, whereas scores above 5.5 indicate increasingly more barriers to accessibility.

6.3 *Measurement for evaluating complexity*

Complexity of the websites was also examined for the purposes of explaining accessibility. Bucy *et al.* (1999) empirically evaluated the 'formal features' (the format structure, editing and digital effects, or the packaging of the information as opposed to the content) of 496 homepages. By manually analysing each homepage, Bucy *et al.* sought to answer the question of whether webpage complexity (as a result of the use of interactive capabilities) correlates with the popularity of the webpage. They specifically define complexity as any element that potentially increases the cognitive overload such as banners, homepage structure, and graphical, dynamic, asynchronous interactive and realtime interactive elements. Building on the Bucy study, we use the concept of different elements being more complex. The complexity score was designed to follow the sequence derived when one looks at a dot, a two-dimensional square and an enclosed three-dimensional cube: a dot is simpler than a square and a square is simpler than a cube. The complexity score (figure 2) is derived by automatically parsing the entire HTML document and assigning a value to each HTML tag (code used in HTML documents to indicate elements). By weighing certain HTML tags differently, the complexity score captures the fact that components of the webpage pose differing levels of barriers to accessibility for persons with disabilities. Object tags (e.g. < OBJECT > and < /OBJECT >), represented by the cube in the metaphor, are coded with a value of 100 units because they are the most complex elements. Script tags (e.g. < SCRIPT > and < /SCRIPT >), represented by the square in the metaphor, are coded with a value of 10 units because they are less complex than objects, yet more so than many other HTML tags. All other tags (e.g. < P >,< /P >, < TR >, < /TR >), represented by the dot in the metaphor, are coded with a value of one unit. A tag value starts at the opening angle, ' < ', and ends at the ending angle, ' > ', and only standard tag names are recognised, all other non-identifying modifiers (e.g. ID, VALIGN, etc.) are ignored. The tags unit values are summed and this summation represents the complexity score for the single page. The best assessment of website complexity for each website is the total of the scores of the pages evaluated for the website normalised by the number of pages analysed.

For illustration, the homepage and 13 second-level pages were evaluated for www.multimap.com (figure 3). The site contained a total of 495 single tags and 44 scripts and 0 objects. By applying the complexity algorithm and normalising for the number of pages analysed, the complexity score for Multimap.com is 66.79. Similarly, www.iVillage.-com was analysed. The homepage and 16 second-level pages were evaluated. There were a total of 4 066 single tags, 292 scripts and 7 objects. The complexity score for this site is 452.12.

$$\boxed{\mathit{Complexity} = \sum(\mathit{Tag}*1) + \sum(\mathit{Script}*10) + \sum(\mathit{Object}*100)}$$

Figure 2. Complexity score.

Complexity(Multimap) = 495 single tags*1 + 44 scripts*10 + 0 objects*100 = 935/14 pages = 66.79

Complexity(iVillage) = 4 066 single tags*1 + 292 scripts*10 + 7 objects*100 = 7 686/17 pages = 452.12

Figure 3. Complexity example.

6.4 *Statistical methods*

For the random category, Analysis of Variance (ANOVA) procedures were used to compare mean WAB scores and mean complexity scores of different years. When ANOVA showed statistical significance, Tukey's 'Honestly Significantly Different' (HSD) procedures were used to do pairwise comparisons among means of the years. Pearson correlation was performed to correlate WAB scores and complexity scores without regard to year. Repeated measures ANOVA, with time being the within subjects factor, was used to test for linear trends within the US government category, since the same sample of websites was followed through all six years. Pearson correlation was performed to correlate WAB scores and complexity scores without regard to year for US government websites.

7. Results

7.1 *Random websites*

A total of 162 archived instances of websites were analysed for the random category: 27 websites for each of six years (1997–2002). There were 105 unique websites, since it was possible for a website to be randomly selected in the sample for more than one year (i.e. www.circuitcity.com was in the 1997 sample and the 1999 sample). An alpha level of 0.05 was designated for analyses. Mean WAB scores increased each year (figure 4). ANOVA was computed comparing the mean WAB scores for each year. A significant difference was found among the years ($F[5,156] = 4.943$, $p < 0.001$). Tukey's HSD procedure was used to do pairwise comparisons among means of the years for random websites. This analysis revealed that websites selected for the year 2000 (m = 8.49, sd = 2.01) had significantly higher WAB scores than websites for the year 1997 (m = 6.52, sd = 2.50). Websites for the year 2001 (m = 8.82, sd = 2.37) also had significantly higher WAB scores than websites for the year 1997. Websites for the year 2002 (m = 9.32, sd = 2.75) had significantly higher WAB scores than websites for the years 1997 and 1998 (m = 7.21, sd = 2.72). Scores for year 1999 (m = 8.18, sd = 2.21) did not show a statistically significant difference with any other years.

As mentioned, our previous study (Parmanto and Zeng 2005) found the score that separates accessible websites from the inaccessible ones to be 5.5. We could extrapolate that the web, in general, was relatively accessible at the beginning of its history. By 1997, it was still not far from the accessible line. By 2001 and 2002, it has become inaccessible to persons with disabilities.

Complexity scores increased each year from 1997 to 2002 (figure 5) within the random websites. ANOVA was computed comparing the mean complexity scores for each year. A statistically significant difference was found among the years ($F[5, 156] = 13.587$, $p < 0.001$). Tukey's HSD analysis revealed that websites selected for the year 2002 had significantly higher complexity scores than websites for all other years and the year 2001 had significantly higher scores than 1997.

A Pearson correlation coefficient was computed to evaluate the relationship between WAB scores and complexity scores without regard to year for the random category (see Random in figure 6). By graphing a box-plot, outliers (5) were identified and removed. A significant positive correlation was shown between the two ($r = 0.497$, $p < 0.01$): as complexity scores increase so do WAB scores. The original correlation prior to removing outliers was 0.516 ($p < 0.01$).

7.2 *US Government websites*

For comparison, US government websites were also examined as a group. As previously mentioned, US government websites are the only websites required to

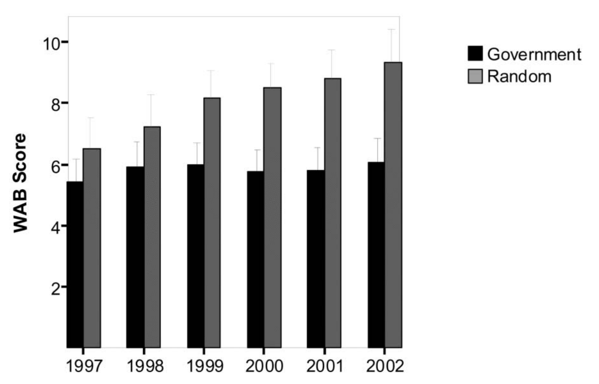

Figure 4. Comparison of mean WAB scores for US government and random websites.

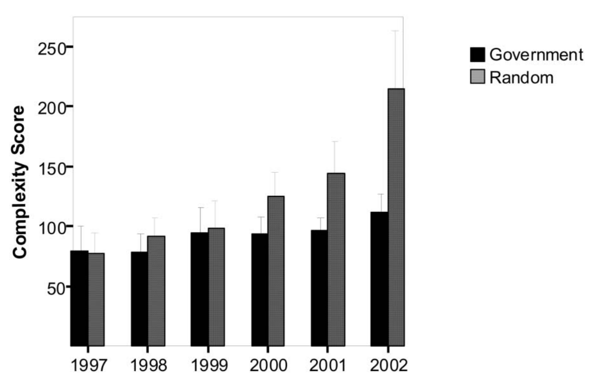

Figure 5. Comparison of mean complexity scores for US government and random websites.

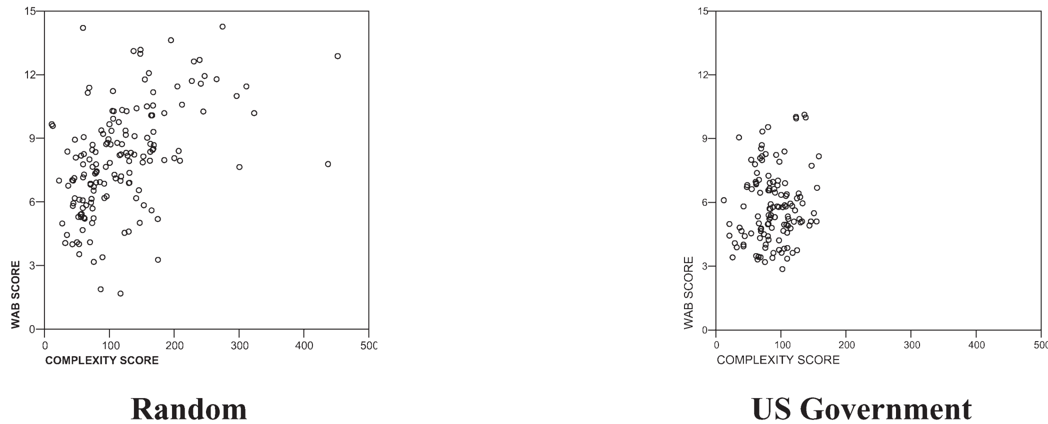

Figure 6. Scatter plot of correlation between WAB and complexity scores 1997 – 2002, outliers removed.

meet accessibility standards and therefore form a good basis for comparison. There were 22 US government websites studied and the same websites were evaluated for each year. Mean WAB scores remain fairly unchanged from 1997 to 2002 (figure 4); the scores of US government websites over the years are close to the accessible line. A repeated measure ANOVA was computed to test for a significant change in accessibility over time. Because of evidence that the assumption of sphericity was not met, results using the Greenhouse-Geiser correction were interpreted. The results were not significant ($F[5,105] = 1.148$). There was no linear trend for the WAB score across the years. Table 1 shows results of the 25 checkpoints comprising the WAB score for both US government and random websites. The number presented is the normalised rate of violation for each checkpoint.

The complexity of US government websites increases through time (figure 5). A repeated measure ANOVA was computed to test for a significant change in complexity over time. Because of evidence that the assumption of sphericity was not met, results using the Greenhouse-Geiser correction were interpreted. The results were significant ($F[5,105] = 3.758$, $p < 0.01$). There is also a significant linear trend in the data ($F[1,21] = 9.926$, $p < 0.005$), indicating that there is a tendency for complexity scores to increase each year.

A Pearson correlation coefficient was computed to evaluate the relationship between WAB scores and complexity scores of US government websites without regard to year (see US government in figure 6). By graphing a boxplot, outliers (6) were identified and removed. There was no relationship between the two ($r = 0.14$). The original correlation prior to removing outliers was 0.206.

Since the same US government websites were studied for all years, it was possible to differentiate those that are becoming more accessible from those that are not. The Federal Trade Commission's website (www.fcc.gov) has had an accessible website (WAB < 5.5) each year with the exception of 1997 despite having one of the most complex websites of the US government sample studied. Other US government sites that increased in complexity while remaining accessible include: Centers for Disease Control and Prevention (www.cdc.gov), Census Bureau (www.census.gov) and Department of Education (www.ed.gov), among others.

The US Senate website, however, is an example of a website that became less accessible with an increase in complexity. Upon further investigation, the increase in WAB score can be attributed to lack of alternative text for images and image map hot spots, and not providing links for client-side image maps elsewhere in the page, as well as a general increase in the use of tags.

The seven priority 1 checkpoints in the WAB score were further examined because they pose absolute barriers if violated. For both categories, the use of images has been increasing rapidly, more than doubling in US government websites from 2001 to 2002 and more than tripling in random websites during the same period. Applets and objects are not used frequently in either category, but also aren't accompanied by text equivalents describing the applet or object when they are used. Image-type buttons have begun to be used more often in both categories of websites. The increased presence of online forms may play a factor in this, as it has become good customer service to provide many services on the Internet that were previously handled in person or by telephone. In an effort to make pleasant looking, user-friendly forms, web designers use image-type buttons. Designers of US government websites are including alternative text for image-type buttons, with few violations. In random websites, however, the rate of not including alternative text for image-type buttons has been increasing. This indicates a lack of awareness that they should be accompanied by alternative text. The use of image maps is more common in random websites, with

Table 1. Checkpoints of the WAB score for random (R) and US Government (G) websites.

Checkpoint	Priority	Category	1997	1998	1999	2000	2001	2002
Provide alternative text for images	1	R	0.56	0.58	0.47	0.67	0.57	0.46
		G	0.34	0.23	0.31	0.24	0.18	0.08
Provide alternative text for applets	1	R	1	1	-	1	0.89	1
		G	1	1	0.94	0.75	0.29	0.70
Provide alternative content for objects	1	R	1	1	1	1	1	1
		G	-	-	-	-	1	1
Provide alternative text for all image-type buttons in forms	1	R	0.95	0.36	0.94	0.57	0.86	0.65
		G	0.50	0.50	0.10	0.03	0.03	0.02
Provide alternative text for all image map hot-spots (areas)	1	R	0.83	0.78	0.90	0.39	0.45	0.32
		G	0.82	0.65	0.67	0.37	0.48	0.26
Each frame must reference an HTML file	1	R	0.21	0.08	0.05	0.86	0.33	0.03
		G	0.14	0.04	0.02	0.16	0.47	0.20
Give each frame a title	1	R	1	1	1	1	1	1
		G	1	1	1	0.93	0.47	0.04
Use a public text identifier in a DOCTYPE statement	2	R	0.69	0.75	0.85	0.80	0.85	0.66
		G	0.68	0.69	0.74	0.69	0.69	0.59
Use relative sizing and positioning rather than absolute	2	R	0.09	0.09	0.07	0.15	0.12	0.17
		G	0.04	0.07	0.09	0.19	0.21	0.20
Nest headings properly	2	R	0.13	0.09	0.08	0.02	0.30	0.05
		G	0.28	0.08	0.15	0.06	0.07	0.09
Provide a NOFRAMES section when using frames	2	R	-	-	-	-	-	-
		G	-	-	-	-	-	-
Avoid blinking text created with the BLINK element	2	R	1	1	-	-	1	-
		G	1	1	1	1	1	1
Avoid scrolling text created with the MARQUEE element	2	R	-	1	-	1	1	-
		G	1	1	1	-	-	1
Do not cause a page to refresh automatically	2	R	1	-	-	-	-	-
		G	-	-	-	-	-	-
Do not cause a page to redirect to a new URL	2	R	-	-	-	-	-	-
		G	-	-	-	-	-	-
Make sure event handlers do not require use of a mouse	2	R	1	1	1	0.99	1	1
		G	1	1	1	1	1	0.98
Explicitly associate form controls and their labels with the LABEL element	2	R	0.84	0.88	0.92	0.85	0.83	0.82
		G	0.87	0.71	0.74	0.77	0.83	0.74
Create link phrases that make sense when read out of context	2	R	0.02	0.01	0.02	0.02	0.02	0.02
		G	0.05	0.04	0.05	0.03	0.04	0.06
Do not use the same link phrase more than once when the links point to different URLs	2	R	0.07	0.08	0.14	0.11	0.15	0.17
		G	0.27	0.22	0.22	0.27	0.18	0.19
Include a documents TITLE	2	R	< 0.01	< 0.01	0.02	< 0.01	< 0.01	0.02
		G	0.03	0	0.01	0.01	0.01	0.01
Client-side image map contains a link not presented elsewhere on the page	3	R	0.68	0.59	0.41	0.76	0.69	0.74
		G	0.74	0.64	0.63	0.61	0.69	0.66
Identify the language of the text	3	R	1	1	1	0.99	1	1
		G	1	0.99	0.99	0.93	0.88	0.71
Provide a summary for tables	3	R	1	1	1	0.99	1	1
		G	1	0.96	0.97	0.99	0.99	0.77
Include default, place-holding characters in edit boxes and text areas	3	R	0.55	0.78	0.40	0.55	0.40	0.44
		G	0.74	0.61	0.62	0.68	0.69	0.54
Separate adjacent links with more than white space	3	R	0.11	0.48	0.14	0.15	0.15	0.31
		G	0.27	0.19	0.26	0.32	0.24	0.37

usage substantially increasing in 2001 and 2002. The number of violations has also increased. The use of frames peaked around 1999–2001 for US government websites and 1999 for random websites.

Due to the weighting system of the WAB metric, with Priority 1 checkpoints weighing more heavily, checkpoint violations causing the most accessibility barriers to persons with disabilities in US government websites are: not supplying alternative text for both images and image map hotspots. For random websites, it is these same checkpoints, as well as not providing alternative text for image-type buttons in forms. If designers made minimal effort to include alternative text for each image, image map hot spot and image-type button, their websites would be substantially more accessible. One particular 2002 website, consisting of the homepage and five second-level pages, was evaluated. The mean WAB score was 10.08. If the designer corrected violations for the most commonly violated checkpoints previously stated, the mean WAB score would decrease to 6.55. Since this website did not have any image-maps, alternative text would need to be supplied for only 28 images and eight image-type buttons.

The results of US government websites show that an increase in complexity does not necessarily translate into a decrease in accessibility. These results provide hope that if all websites take accessibility issues as seriously as US government websites do, the goal of universal web accessibility is one that is attainable.

8. Discussion

The findings from the ANOVA and Tukey's HSD analyses show that along with a statistically significant increase in accessibility barriers there has been a concurrent statistically significant increase in complexity in the random websites studied. As web designers have added increasingly complex components to the design of their webpages for the purposes of creating aesthetically appealing and interactive websites, they have also added barriers to accessibility for persons with disabilities.

The use of elements to make attractive looking websites that grab the attention of the consumer and hold their interest as they travel through the website has increased. As the use of more complex webpage components has increased, so has the number of guideline violations. Web designers in both categories studied have increased their use of tags in general in their design of webpages over the years, but this is occurring more so in the random category. Also seen in both categories is an increased use of scripts. While the use of scripts is increasingly common in US government webpage design, as well, it is seen with much greater frequency and is increasing much more rapidly in the random websites. Scripts appear to be the most distinguishable factor contributing to increased complexity scores of the random websites that is not seen in US government websites.

The number of links to other websites or to pages within the same site is shown to increase as well. Even though links themselves may not pose accessibility barriers as long as they are adequately descriptive, they have the potential to make the web less navigable to non-visually impaired users (Khan and Locatis 1998), which leads one to believe that an increased number of links could lead to additional navigation barriers to persons with disabilities, as well. One reason for this is that they may tire easily while listening to the screen reader read the many links out loud. Summaries are almost never provided for tables, making information contained therein virtually unavailable for users of screen readers.

Like websites in the random sample, the complexity scores of the US government websites included in the study also increased. Interestingly, these websites failed to have the concurrent increase in WAB scores. The WAB scores of US government websites remain consistently close to the accessible line. This is evidence that increasing complexity for purposes of aesthetically appealing or consumer-driven design does not necessarily equate to inaccessibility for persons with disabilities.

8.1 *Limitations*

The authors recognise limitations to this study. When using the Internet Archive, some websites in the random sample had to be replaced because they were not in the Wayback Machine. This included websites that are password protected, blocked by the webmaster to not be available for crawls, and websites that were inaccessible because of an Internet Archive error. The authors also recognise the statistically problematic issue that arises because the same random websites were not followed through the years. However, doing so would bias the random sample towards sites that have longevity, and thus not be a true representation of the web.

The complexity metric is an automatic subjective measure that was designed by the authors in an attempt to explain the increasing inaccessibility that is occurring on the web. This metric does not take into consideration the user's experience with the website although, by weighting the components differently, it does account for the varying degrees of accessibility posed by the components.

8.2 *Contributions*

Due to the flexibility of the WAB metric, it can be easily adapted to new guidelines and changes in checkpoints and priority levels. For example, future studies will include findings relating to cognitive disabilities once WCAG version 2 has been finalised. The metric can incorporate changes in user-agent technology in the same way: an individual may choose to look at only specific checkpoints or weigh checkpoints differently depending on the user-agent technology. The flexibility of the metric makes it possible for a webmaster to provide several WAB scores for a website depending on the user-agent technology accessing the website. A designer could also utilise the metric during the development or redesign of a website to ensure a particular WAB score.

Research studies like this one can provide valuable information that can lead to improvements in guidelines for accessible web design by identifying specific problems. These studies can also provide meaningful information to engineers of assistive technologies so that these technologies may become better suited to anticipate changes in the web, for example frames are not used with much regularity anymore and don't pose as much an issue as they have in the past.

Many web designers, especially of non-US government websites, may be inadvertently placing barriers to persons with disabilities simply by using the latest web design technologies. Although they may be providing an appealing site for some people, they are unaware of the implications such technology poses to other individuals. Guidelines and policies cannot only bring about adherence to better practises of webpage design; they can forge a more widespread awareness of the problem. By proving that these mandates do have an effect on accessibility, more industries will embrace policies similar to Section 508, making WCAG common practise and evolving the Internet into something that truly is accessible.

Acknowledgements

This research is supported in part by grants #42-60-I02013 from the National Telecommunications and Information Administration (NTIA) and #H133A021916 from the National Institute on Disability and Rehabilitation Research (NIDRR). The authors would like to thank Corinne Kirchner from the American Foundation for the Blind (AFB) for planting the idea of analysing web accessibility over time.

References

100 Top Government Sites, 2003, www.100topgovernmentsites.com.

Amtmann, D., Johnson, K. and Cook, D., 2000, Making web-based tables accessible for users of screen readers. *Library Hi Tech*, **2,** pp. 221–31.

Bucy, E.P., Potter, R.F., Lang, A. and Grabe, M.E., 1999, Formal features of cyberspace: Relationships between Webpage complexity and site traffic. *Journal of the American Society for Information Science*, **13,** pp. 1246–56.

Chisolm, W., Vanderheiden, G. and Jacobs, I., 2001, Web content accessibility guidelines. *Interactions*, **4,** p. 34.

Davis, J.J., 2002, Disenfranchising the disabled: The inaccessibility of internet-based health information. *Journal of Health Communication*, **4,** pp. 355–67.

DDA, 2004, Disability Discrimination Act 1995. www.disability.gov.uk/dda/ (accessed 13 August 2004).

Diaper, D. and Worman, L., 2003, Two falls out of three in the automated accessibility assessment of world wide web sites: A-prompt v. Bobby. In *People and Computers xvii*, Springer-Verlag, London, pp. 349–63.

Disability Rights Commission, 2004, The web access and inclusion for disabled people. www.drc-gb.org/publicationsandreports/2.pdf.

DRC, 2004, The Disability Rights Commission. www.drc-gb.org/ (accessed 13 August 2004).

Egan, J.P., 1975, *Signal detection theory and roc analysis* (New York: Academic).

Eurostat, 2001, Disability and social participation in Europe. http://europa.eu.int/.

Flowers, C.P., Bray, M. and Algozzine, R.F., 1999, Accessibility of special education program home pages. *Journal of Special Education Technology*, **2,** pp. 21–6.

Foley, A. and Regan, B., 2003, Best practices for web accessibility design and implementation. http://macromedia.com/resources/education/white-papers (accessed 21 June 2003).

Freedom Scientific, 2004, Jaws for Windows. www.freedomscientific.com (accessed 14 August 2004).

Goodwin-Jones, B., 2001, Emerging technologies – accessibility and web design why does it matter? *Language Learning and Technology*, **1,** pp. 11–19.

Heim, J., 2000, Locking out the disabled [web site accessibility]. *PC World (US Edn)*, **9,** pp. 181–5.

IBM, 2003, Development tools and technology. www.ibm.com (accessed 15 June 2003).

IBM, 2004, IBM accessibility center. www-3.ibm.com/able/solution_offerings/hpr.html (accessed 14 August 2004).

Internet Archive, 2001, The internet archive introduces a next generation library. www.archive.org/about/wb_press_kit.php.

Internet Archive, 2003, Wayback Machine. www.archive.org (accessed 7 August 2003).

Khan, K. and Locatis, C., 1998, Searching through cyberspace: The effects of link display and link density on information retrieval from hypertext on the world wide web. *Journal of the American Society for Information Science*, **2,** pp. 176–82.

Lazar, J., Beere, P., Greenidge, K. and Nagappa, Y., 2003, Web accessibility in the mid-Atlantic United States: A study of 50 home pages. *Universal Access in the Information Society Journal*, pp. 331–41.

Lazar, J., Dudley-Sponaugle, A. and Greenidge, K., 2004, Improving web accessibility: A study of webmaster perceptions. *Computers in Human Behavior*, **2004,** pp. 269–88.

McMullin, B., 2002, Users with disability need not apply? Web accessibility in Ireland. *First Monday*, **12,** www.firstmonday.org/issues/issue7_12/mcmullin/.

Nielsen, J., 2000, *Designing Web Usability* (Indianapolis, Indiana, US: New Riders Publishing).

Parmanto, B. and Zeng, X., 2005, Metric for web accessibility evaluation. *Journal of Society for Information Science and Technology* (in press).

Pontelli, E., Gillan, D., Gupta, G., Karshmer, A., Saad, E. and Xiong, W., 2002, Intelligent non-visual navigation of complex html structures. *Universal Access in the Information Society*, **1,** pp. 56–69.

RINCE, 2003, Des sites web francais de la vie quotidienne sont inaccessibles aus personnes handicapees. *Research Institute for Networks and Communications Engineering,* http://braillenet.org/eval_30sites_france_10seat2003.htm

Stowers, G.N.L., 2002, The state of federal websites. Pricewaterhouse-Coopers Endowment for the Business of Government, E-Government Series, www.businessofgovernment.org/pdfs/StowersReport0802.pdf

USAccessBoard, 2003a, The rehabilitation act amendments (section 508). www.access-board.gov/sec508/guide/act.htm (accessed 9 September 2003).

USAccessBoard, 2003b, Web-based intranet and internet information and applications. www.access-board.gov (accessed 9 September 2003).

USCensus, 2003, Disability status: 2000 www.census.gov/prod/2003pubs/c2kbr-17.pdf.

USDOJ, 2003, Federal agencies' Webpages. www.usdoj.gov/crt/508/report/web.htm (accessed 7 August 2003).

USDOJ, 2003, Section 508 self-evaluation Webpage accessibility questionnaire for component web contacts. www.usdoj.gov/crt/508/web.htm (accessed 7 August 2003).

W3C, 2003, Wai technical activity statement. www.w3.org/WAI/Technical/ (31 July 2003).

WATCHFIRE, 2004, Bobby. www.bobby.watchfire.com (accessed 14 August 2004).

WEBAIM, 2004, Screen reader simulation. www.webaim.org/simulations/screenreader (accessed 14 August 2004).

WEST, D.M., 2001, Wmrc global e-government survey. www.brown.edu/Departments/Taubman_Center/polreports/egovt01int.html.

WHO, 2004, World health organization. www.who.int.

YAUKEY, J., 2003, Archive site preserves earliest Webpages. www.gannettonline.com/e/trends/15000566.html (accessed 13 August 2003).

ZENG, X., 2004, Evaluation and enhancement of web content accessibility for persons with disabilities. Dissertation. University of Pittsburgh, Pittsburgh. http://etd.library.pitt.edu/ETD/available/etd-04192004-155229/.