Comparative Evaluation of NRDC Web Portal Using Automated Tools

Jose T. Painumkal Sankeerth Ankam Sergiu M. Dascalu Frederick C. Harris, Jr
Department of Computer Science and Engineering, University of Nevada
Reno, Nevada, 89501, USA
{josepainumkal, sankeerth.ankam}@nevada.unr.edu {dascalus, fred.harris}@cse.unr.edu

Abstract

Websites play an important role in this competitive world. They are powerful tools for connecting business with the people, thus, creating a medium to deliver messages directly to the users. A good web interface must conform to various standards and guidelines, which determine the quality and effectiveness of the interface. To analyze the content and standards of a web interface, automated tools can be used. This paper explores various web-based automated tools that are currently available and discusses its capability in the evaluation of a web interface. Using automated tools, an experimental study was conducted on the web interface of Nevada Research Data Center (NRDC) to evaluate its design and usability standards. On the basis of the results from the tools, we performed a comparative evaluation of the NRDC's web interface with other similar portals on World Wide Web (WWW). The paper also includes the results of the comparative evaluation and gives recommendations on how to improve the standard and usability of the NRDC's web interface.

Keywords: Automated; Tools; Web; Compliance; Interface; Standards; Evaluation; Usability; Accessibility

1 Introduction

Websites are a powerful tool for connecting our business with the people. A good web interface must conform to various standards and guidelines stated by WWW. These standards ensure better quality and enhance the accessibility of the web interface. [1] (proceedings) The task of evaluating the standards of websites can be daunting given the quantity of sites being produced, the frequency of updates, and the sheer size of many sites. However, the advent of automated tools has simplified the evaluation process to a greater extent. These tools facilitate the inspection of several webpages at a time, and thereby save time and effort spent on the standardization of web pages. Automated tools can be used in measuring various quality standards of an interface like markups, readability, accessibility, mobile compatibility, page performance, etc. Thus, they help us in improving the overall quality of the final product and assure us a more useful and accessible web user interface.

The emphasis of this study is on assisting NRDC in finding the vital design flaws in their web portal for a better usability and user experience. [2] (proceedings) NRDC serves as a cyber-infrastructure hub providing data and computing resources for studying the effects of climate change in Nevada. It also provides a collection of software tools to handle a variety of services by utilizing the underlying architecture and web services of the portal. The web portal of the NRDC can be freely accessed by researchers, students, and organizations via its website.

The main goals of this study include:

1) Conducting research on the capabilities of various automated web-based tools.
2) Selection of appropriate tools for validating the conformance of NRDC's website to globally accepted standards.
3) Evaluating the design and usability standards of NRDC's interface using automated tools.
4) Interpreting the results obtained from the automated evaluation of the interface.
5) Comparing the results from automated tools with other similar portals to see how NRDC performs in terms of various standards.
6) Providing recommendations on improving the standards and usability of NRDC's interface

For comparative evaluation, we compared the interface of NRDC with web interfaces of Western Regional Climate Center (WRCC) and Consortium of Universities for the Advancement of Hydrologic Science, (CUAHSI).

The rest of the paper is organized as follows. Section 2 describes the details of the tools and metrics used for the evaluation. The design of the study is given in Section 3. The results of the tests are presented in Section 4. Section 5 describes the challenges and limitations faced during the tool-based evaluation of web interfaces. Recommendations for improving the usability of NRDC's web interface are described in Section 6. Section 7 gives the conclusion.

2 Tests and Tools Used

This section includes the details of the tools and metrics used in the evaluation of the interfaces. User testing is one of the primary testing processes involved in the evaluation of usability on any interface. Since our study was aimed at the automated evaluation of web-interfaces, the whole testing process was confined only to the tool-based assessment of usability standards. Choosing the right standards for the evaluation possess great importance, as it helps in figuring out the actual issues faced by the users of the interface. W3C standards define an Open Web Platform for application

development that ensure the web works equally well for everyone, regardless of their location, platform or technology. [3] (proceedings) They are widely recognized standards that are always maintained up to date with the latest innovations on the web and are relatively free from conflicts of interest. Due to these reasons, in our study, we had checked for conformance to W3C standards wherever they were applicable.

The experimental study was carried out by focusing on the following metrics:

1) Compliance with W3C Hypertext Markup Language (HTML) standards
2) Compliance with W3C Cascading Style Sheets (CSS) standards
3) Compliance with W3C standards for mobile web viewers
4) Compliance with W3C WCAG
5) Readability of the text
6) Evaluation of page performance and web links.
7) Evaluation of web page links.

The first and second metric gives the minimum assurance that the web page will be displayed uniformly across different browsers and platforms. The third metric gives the extent to which the web page is compatible on mobile devices. The fourth metric indicates the level of accessibility of the web content. The fifth metric is a pointer to the understandability and readability of the content on the website. The sixth metric indicates the user's experience with the page on both mobile and desktop devices. The seventh metric indicates the performance of the links on the website, and alert developers on the issues faced by the users of the interface.

The details of the tools used for each test is described below.

Compliance to web standards is very important as it ensures the consistent handling of webpage across different platforms and web browsers. Invalid markups or styles could lead to unexpected errors and increases the risks of future browser incompatibility, poor SEO performance, cross-platform incompatibility, etc. For validating the compliance to HTML and CSS standards, we used the Validation Service offered by W3C.

Currently, web pages are designed in a more responsive way to support both mobile and desktop devices. To test the compliance of the page to mobile standards, we used W3C mobileOK Checker, a free service offered by W3C. This tool performs several tests based on [4] (online) Mobile Web Best Practices 1.0 standard and gives the level of mobile-friendliness of the page.

A web document has to adhere to various web accessibility standards to ensure accessibility to everyone. These may include the W3C's WCAG 1.0, WCAG 2.0, Section 508, the Stanca Act, BITV, RGAA or a combination of both. For this study, we evaluated the web page against the globally accepted WCAG 2.0 standards. An open source accessibility evaluation tool, [5] (proceedings) AChecker, was used to evaluate WCAG compliance.

Readability tests evaluate a web document based on two metrics; the first metric evaluates the easiness in reading contents on the webpage and the second metric determines the understandability of the content. For analyzing the easiness in reading the content, we used WAVE, a free web accessibility evaluation tool provided by WebAIM. For evaluating understandability of the content, we used a readability evaluation tool offered by Juicy Studio. This tool analyses the web content based on reading level algorithms (Gunning Fog, Flesch Reading Ease, and Flesch-Kincaid) and indicate us whether the web content is set at the right level for the intended audience.

Poor loading speed, unnecessary page redirects, broken or inactive page links affect the user experience on an interface, and could finally lead to the isolation of website from its users. To evaluate page loading speed, we used PageSpeed Insights, a free tool offered by Google. This tool evaluates the performance of a webpage on mobile as well as desktop devices, and provides recommendations for improving the user experience in Mobile and Desktop platforms. For evaluating page links, W3C Link Checker, a free tool offered by W3C was used. This tool can crawl over the links of a web interface and checks for issues in the links, unnecessary redirects & anchors in the page, or a whole website.

3 Design

The automated tools used in this study are stand-alone and work independently. Each tool is designed for evaluating a specific quality standard on the webpage, and gives recommendations on improving the compliance of the web page to that standard. Since the tools are independent in terms of functionality, the order at which the tests are carried out does not affect the capability of the tool in result generation. However, the order of the tests does affect the quality of the results generated from the tool. For example, it is always better to validate HTML before carrying out other tests. One reason for that is, HTML validation is widely regarded as a basic quality assurance standard for any webpage. Secondly, validating HTML could remove hidden and unidentified errors from the web page, and thereby improves the consistency of the page to other quality measures. Moreover, it was evident from the results of the tool-based evaluation that there exists an overlap between several guidelines of HTML and accessibility conformance check. For example, the guideline stating all image elements should have an alternate text attribute, is a common standard in both HTML as well as in accessibility validation. To confirm it, a sample of 10 URLs was randomly selected from the three web portals. The deciding factor in choosing the URLs was the occurrence of "missing

alternate text attribute" errors in the page. After evaluating the pages, it was noticed that proper markup validation could reduce accessibility issues by 45%. This is strictly dependent on what type of validation errors occur on the webpage. If the markup issues in the page overlap with the accessibility validation guidelines, then validating HTML could make a significant reduction in the accessibility issues on the page. Markup validation could even simplify link evaluation, and therefore should be considered as a recommended process before link evaluation. The proposed order for carrying out various validation tests on a webpage is shown in Figure 1.

Since automated tools could not validate the entire website through auto crawling over the links, the compliance check was done by inputting URLs one by one to the tool. For websites comprising hundreds of webpages, it was highly impractical entering webpage URLs each time for checking the compliance to standards. Therefore, for the comparative evaluation, a sample set of URLs - comprising 51 webpages from NRDC, 36 webpages from WRCC and 36 webpages from CUAHSI, were collected from the websites through careful inspection on the usability and relevance of the web pages. For evaluating HTML compliance, CSS compliance, mobile compatibility, WCAG 2.0 compliance, readability and page performance, this sample set of URL's were used.

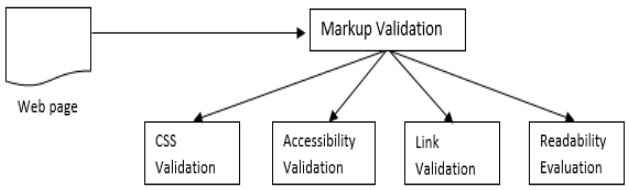

Figure 1: Proposed order for carrying out validation tests.

4 Results and Discussion

This section includes the results of the tool-based comparative evaluation of NRDC web portal with the interfaces of WRCC and CUAHSI. During this evaluation, various automated tools were used for measuring the usability and accessibility standards of the interface. The details of the test, metrics and interpretation of the results are given below.

4.1 Statistical Comparison

Statistical comparisons were performed to obtain an overview on the content and capabilities of the three different web portals, intended for serving somewhat identical goals. Since the websites were so large with hundreds of webpages, manual inspection was not a plausible approach for the evaluation. Therefore, website crawlers were used to crawl over website URLs and fetch key onsite elements to analyze from a SEO perspective. The results of the statistical comparison of three web portals using website crawlers are given in Table 1.

Table 1: Statistical comparison

Web interface	Page count	Depth (Click from start URI)	Size (in gigabytes)
NRDC	88	5	1.66
WRCC	944	4	21.99
CUAHSI	1106	6	21.36

Statistical comparisons give us valuable insights about the purpose and capabilities of different web interfaces. In Table 1, depth metric indicates how many clicks the user has to make from the home page to acquire the required information from the portal. The size field in the table indicates the total size of the html files in the website.

4.2 Compliance to HTML/CSS Standards

The results of the tests obtained while evaluating HTML and CSS compliance is shown in Table 2. The deciding factor in determining the HTML compliance of a web interface was the percentage of webpages that satisfies W3C Markup standards. If 70% of the evaluated web pages satisfy the validation standards, then the overall validity of the website was set as Y. The same criterion was followed for CSS validation. The compliance errors obtained during validation tests are shown in Table 3.

Table 2: HTML and CSS compliance of three web portals

Web interface	HTML valid? (Y/N)	Valid pages (%)	CSS valid? (Y/N)	Valid pages (%)
NRDC	Y	75	N	0.01
WRCC	N	0	Y	72.9
CUAHSI	NA	NA	N	0

Table 3: Validation error count in web portals

Web interface	Error count		Errors per page	
	HTML	CSS	HTML	CSS
NRDC	348	5866	6.82	115
WRCC	100	67	2.77	1.86
CUAHSI	NA	144	NA	4

Since more than 70% of the webpages were compliant to markup standards, the interface of NRDC was considered as HTML compliant. None of the pages in WRCC interface was HTML compliant. The HTML validator failed to

validate CUAHSI's interface, and hence marked as 'NA' in the table. Among the three interfaces, WRCC had the lowest CSS error count and, was the only CSS compliant interface.

4.3 Compliance to Mobile Standards

Table 4 illustrates the results of the evaluation on mobile compatibility of the interfaces. This test evaluates the compatibility of the web content on mobile devices based on W3C's Mobile Web Best Practices.

Table 4: Results of mobile compatibility check

Web interface	Compatible pages (%)	Critical issues	Errors per page
NRDC	42.3	114	2.23
WRCC	0	108	3
CUAHSI	0	144	4

42.3% of the webpages in NRDC's interface were mobile compatible, whereas none of the webpages in WRCC and CUASHI were compatible for mobile interfaces.

4.4 Compliance to WCAG

The web interfaces were evaluated to check conformance to WCAG 2.0 guidelines for measuring accessibility metric. [6] (online) WCAG 2.0 is a stable, referenceable technical standard, which has 12 guidelines that are organized under 4 principles: perceivable, operable, understandable, and robust. For each guideline, there are three levels of conformance in order to meet the needs of different groups and different situations: A (lowest), AA (mid-range), and AAA (highest). Conformance at a higher level indicates conformance at lower levels. For example, by conforming to AA, a Web page meets both the A and AA conformance levels. The results of the accessibility evaluation based on WCAG 2.0 compliance is shown in Table 5.

The results show that all the three interfaces had problems at all the three levels of accessibility. Among the three interfaces, NRDC had the highest number of errors in all levels of accessibility conformance.

Table 5: Accessibility evaluation of thee portals

Web interface	Accessibility errors per page		
	Level A	Level AA	Level AAA
NRDC	6.49	67.64	68.66
WRCC	6.13	7.36	7.38
CUAHSI	9.5	15.16	17

4.5 Evaluation of Readability

The results of the readability evaluation are illustrated in Table 6. For measuring readability, two aspects of reading were considered - understandability of the content, and easiness in reading the text on the web page.

For evaluating the easiness in reading, the webpages were evaluated for conformance to color contrast guidelines of WCAG 2.0. According to WCAG (Web Content Accessibility Guidelines), the text should have a minimum (4.5-to-1) color contrast difference with its background color. For this test, the web pages were evaluated individually. The total issues on the evaluated pages were used to determine the easiness in text reading on the interface.

The results show that all the three interfaces had color contrast issues in their webpages. WRCC had the least number of issues whereas; CUAHSI came last with highest number of issues. It was also observed that, the color contrast issues in the header and footer could make great impact on the overall readability of the interface. For example, in WRCC, there were 9 issues on the header and footer, which kept on occurring on all the web pages, resulting in 324 errors on evaluating 36 web pages. Therefore, fixing the issues in the header and footer alone can significantly enhance the compliance of the website to color contrast guidelines of WCAG.

Table 6: Readability metrics of three web portals

Web interface	Color contrast		Understandability		
	Valid pages (%)	Errors per page	Gunning Fog Index	Flesch Reading Ease	Flesch Kincaid Grade
NRDC	0.06	11.98	11.16	45.25	8.03
WRCC	0.08	9	10.52	44.05	7.7
CUAHSI	0.03	68.36	13.57	33.2	9.7

For evaluating the understandability of the content on the interface, readability metrics from several reading level algorithms were used. The reading level metrics used in this test includes Gunning-Fog Index, Flesch Reading Ease and Flesch-Kincaid grade. [7] (proceedings) The Gunning-Fog index and Flesch-Kincaid grade indicates a rough measure on how many years of schooling a person would take to understand the content. The lower the number, the more understandable the content will be to the readers. CUAHSI had the highest score for these two metrics, thereby indicating lower understandability of the content. The Flesch Reading Ease is another readability metric which uses a 100-point rating scale to measure the understandability of the content. For this metric, a higher score indicates easier to understand. On this scale also, CUASHI had the lowest content understandability.

After combining the results of understandability and easiness in reading, we observed that there was not much difference between the readability of the content in both WRCC and NRDC interfaces. Whereas the readability of CUAHSI interface could be improved further.

4.6 Evaluation of Page Performance

The results of the tests carried out for evaluating page performance of the web interfaces are shown in Table 7. The performance of the web page was evaluated based on [8] (online) ten speed rules and five usability rules specified by Google PageSpeed Insights. Based on the evaluation, the web page was assigned with an overall score ranging between 0 and 100. The web pages were evaluated individually and, the mean PageSpeed scores of the individual pages were analyzed to evaluate the overall performance of the website. The results of the evaluation on mobile and desktop devices are given in Table 7.

For desktop devices, the interface of WRCC performed better than the other two interfaces, with a total page score of 83.64. For the mobile devices, NRDC's interface proved to be best among the three web portals.

Table 7: Page performance evaluation of three web portals

Web interface	Mobile	Desktop
	Speed score (100)	Speed score (100)
NRDC	62.08	66.4
WRCC	56.82	83.64
CUAHSI	48.91	59.11

4.7 Evaluation of Page Links

The results of the link evaluation are shown in Table 8. As part of the test, all the internal page links in the website were evaluated for generic internal server errors and client-side '404 Not Found' issues. Internal Server Error is a generic error message thrown from the server when an unexpected condition has encountered. “404 Not Found” is a client side error which occurs when the requested resource is not found by the server. The tests were done using web crawlers which could auto crawl over multiple webpages on the website.

No internal server errors were detected in WRCC and CUAHSI, whereas NRDC contains one link affected with internal server error. Client side errors were also minimal in all the three interfaces.

Table 8: Results of link evaluation in three portals

Web interface	500 Internal Server Error		404 Not Found Error	
	Error Count	Errors per page	Error Count	Errors per page
NRDC	1	0.01	4	0.045
WRCC	0	0	27	0.025
CUAHSI	0	0	44	0.039

5 Challenges and Limitations

Automated tools played an important role in simplifying the developer’s overhead while evaluating the conformance of web documents to various web standards. Even though the automated tools can ensure the conformity to a significant level, it does not completely eliminate the need for manual inspection. There are limitations for automated tools in the evaluation of web documents.

Most of the automated tools can process only one web page at a time and do not crawl over links on the webpage to process different webpages. Therefore, for evaluating a website, human intervention is still inevitable to manually input the webpages (via URI, file upload) to the tool, which makes the evaluation of large websites cumbersome and tedious. Automated tools may not generate accurate results at all times, or sometimes they may not even work at all. For example, the validation service of W3C could not validate the interface of CUAHSI, and ended up in throwing SSL handshake error. The evaluation of flash rich websites using automated tools are not advisable [9] (journal) as the results will not be accurate enough to interpret. Fortunately, none of the interfaces evaluated as part of this study had any flash-based content on it. In readability evaluation, the automated tools will not provide accurate results when the evaluation includes navigation items and other short items of web content. Therefore, readability evaluation tools are advisable only for websites rich in text content. In evaluating web accessibility, no tool can guarantee 100% assurance that the web content is accessible. Only a human can determine true accessibility.

Even with the advent of automated tools, validation errors are still very common in many of the well-established websites all over the world. There are several factors that retain validation errors in the websites. The validation of old websites proves to be tiresome and expensive, as it requires valuable man-hours and could even lead to major costly design changes. At some occasions, developer ignores validation errors due to time constrain. Ignoring validation errors could save time and effort during coding. However, in the long run, it turns out to be a costly mistake, as it affects code maintainability and would lead to serious compatibility issues. Sometimes, finding the actual cause for a validation error can be tricky and requires special skill. For example, validators may throw hundreds of errors on a webpage which is nowhere near valid. Mostly, the errors follow a cascading fashion, where one error leads to a series of errors. In such cases, finding the actual cause of the error is very crucial, and it could save a lot of time and effort.

6 Recommendations

After interpreting the results, it is observed that HTML and CSS errors were at an alarming rate in NRDC’s interface. HTML and CSS validation are two quality

measures that assure the consistency of the web page across multiple platforms. To improve the web standards of the interface, it is highly recommended to follow the markup/CSS guidelines from W3C.

Web accessibility ensures access of pages to everyone in the web. [10] (proceedings) It could also improve the search engine ranking of a website. To enhance web accessibility, the interface should conform to the guidelines of WCAG. Following WCAG guidelines, [11] (proceedings) not only benefits people with disabilities, but will also benefits others, for e.g. People using slow internet connection, people with changing abilities due to aging, etc.

To improve the mobile compatibility of the web pages: 1) Use legible font sizes. 2) Avoid links too close together. 3) Configure mobile viewport. 4) Size content to viewport. Avoid having content wider than the screen. 5) Refrain the use of incompatible plugins. To enhance the readability of a webpage, use short sentences made up of short words, and maintain the minimum color contrast difference (4.5 to 1) between the text and its background. To improve the page load speed, optimize the images and resources used in the web content and remove render blocking JavaScript/CSS. To improve link performance, ensure all the navigation links to the canonical version (with trailing slash) of the pages. Remove all obsolete inactive links from the interface and also make sure all the resources are named and linked correctly.

Other recommendations for improving the overall usability of the web interface are: 1) including sitemap in the web interface. Sitemap provides [12] (proceedings) users with the graphical representation of web pages on a website and present them with a complete alternative method of navigating the site, 2) providing a search option on each page of content rich websites, 3) providing an easy access to the home page of the interface by providing a link labeled “Home” on every webpage, 4) specifying appropriate width and height attributes for the images, which avoids unnecessary page text jumps while an image loads, and 5) using descriptive and meaningful unique title for each page.

7 Conclusion

In this paper, we conducted a study on the role of web-based automated tools in the evaluation of usability and design standards of a web interface. Using the automated tools, we evaluated the performance of the web interface of NRDC. With the results from the tools, we performed a comparative evaluation of NRDC’s interface with other two similar web portals on WWW. Based on interpreting the results from the tools, we provided NRDC with recommendations on improving the accessibility and user-experience of the interface.

References

[1] Brinck, Tom and Erik Hofer. 2002. "Automatically Evaluating The Usability Of Web Sites". CHI '02 Extended Abstracts On Human Factors In Computing Systems - CHI '02. doi:10.1145/506443.506652.

[2] Sergiu M. Dascalu Frederick C. Harris , Michael McMahon, Eric Fritzinger, Scotty Strachan and Richard Kelley.2014. "An Overview of the Nevada climate Change Portal". *Procs. Of the 7th Intl. Congress on Environ. Modelling & Software (iEMSS-2014)*, San Diego, CA, June 2014, vol.1,pp.75-82.

[3] Gibbs, Ivan, Sergiu M. Dascalu, and Frederick C. Harris. 2013. "Web Portal Usability Tests for the Nevada Climate Change Portal". 2013 *10th International Conference on Information Technology: New Generations*. doi:10.1109/itng.2013.14.

[4] Mobile Web Best Practices 1.0. W3.org, https://www.w3.org/TR/mobile-bp/., May 2016.

[5] Gay, Greg and Cindy Qi Li. 2010. "Achecker". *Proceedings of the 2010 International Cross Disciplinary Conference On Web Accessibility (W4A)* - W4A '10. doi:10.1145/1805986.1806019.

[6] "Web Content Accessibility Guidelines (WCAG) 2.0". W3.Org. https://www.w3.org/TR/WCAG20/. June 2016

[7] Goh, Ong Sing, Chun Che Fung, Arnold Depickere, and Kok Wai Wong. 2007. "Using Gunnnig-Fog Index To Assess Instant Messages Readability From Ecas". *Third International Conference On Natural Computation (ICNC 2007)* Vol V. doi:10.1109/icnc.2007.800.

[8] PageSpeed Insights Rules. Google Developers, https://developers.google.com/speed/docs/insights/rules #speed-rules., May 2016

[9] Youngblood, N. and Youngblood, S. User experience and accessibility: an analysis of county web portals. Journal of Usability Studies 9, 1 (2013), 25-41

[10] M. Ahmed Elgharabawy and M. A. Ayu, "Web content accessibility and its relation to Webometrics ranking and search engines optimization," *2011 International Conference on Research and Innovation in Information Systems*, Kuala Lumpur, 2011, pp. 1-6. doi: 10.1109/ICRIIS.2011.6125721

[11] Helen Petrie and Omar Kheir. 2007. The relationship between accessibility and usability of websites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '07). ACM, New York, NY, USA, 397-406.

[12] J. Manhas, "Design and development of automated tool to study sitemap as design issue in Websites," *Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on*, New Delhi, 2015, pp. 514-518.