# WABS: A WEB ACCESSIBILITY BARRIER SEVERITY METRIC

**Hayfa.Y.Abuaddous[1], Mohd Zalisham Jali[2] and Nurlida Basir[3]**

[1]*FST, Universiti Sains Islam Malaysia (USIM), Negeri Sembilan, Malaysia. haddose@yahoo.com*
[2]*iSER, Universiti Sains Islam Malaysia (USIM), Negeri Sembilan, Malaysia. zalisham@usim.edu.my*
[3]*iSER,Universiti Sains Islam Malaysia (USIM), Negeri Sembilan, Malaysia. nurlida@usim.edu.my*

**ABSTRACT.** A novel metric for quantitatively measuring the severity of websites barriers that limit the accessibility for disabled people is proposed. The metric is based on the Web Content Accessibility Guidelines (WCAG 2.0), which is the most adopted voluntary web accessibility standard internationally that can be tested automatically. The proposed metric is intended to rank the accessibility barriers based on their severity rather than the total conformance to priority levels. Our metric meets the requirements as a measurement for scientific research. An experiment is conducted to assess the results of our metric and to reveal the commonplace violations that persist in websites and affect disabled people interacting with the web.

**Keywords:** accessibility, barrier, metric, automated tools

## INTRODUCTION

The use of the Internet is rapidly spreading into most areas of human life. Therefore**,** it is imperative to have accessible web services for the majority of people including those with certain disabilities (permanent or temporary) in order to provide them with equal access. We can define web accessibility as making a website navigable and tractable by various user categories especially those who have disabilities and normally face obstacles when interacting with the web via electronic devices. According to World Wide Web Consortium (W3C,2005), web accessibility enables people with disabilities, i.e. blind, aged, to utilize the Internet in performing variety of tasks such as online purchasing and browsing.

Many research studies on accessibility have been conducted to evaluate the accessibility condition of public websites (Nizar et al., 2013; Lujan-Mora et al., 2014).Most of these studies focused on calculating barriers' frequencies and priority levels violated which led to inaccessible websites. However, there is an imminent need to look beyond the "Is the website accessible?" and "What are the violated priority levels?" questions and investigate the commonplace barriers (violations) spread across the web and contribute to accessibility issues.

In this article, we propose a novel metric for measuring the severity of accessibility barriers that affect disabled people. To be precise we give quantitative weights to accessibility barriers in order to rank them based on their influence on limiting accessibility. Measuring web accessibility barriers severity in precise and quantitative terms is important for many reasons. First, it would show whether barriers influence on accessibility differ. Second, in large-scale studies it would help accessibility evaluators in determining the influence of each barrier on accessibility rather than evaluating the absolute conformance of websites to guidelines. This will enable them to further investigate these types of barriers and spread awareness among web developers. Third, it would allow for a more precise and advanced statistical analysis for evaluating large-scale aggregate websites.

## WEB ACCESSIBILITY GUIDELINES

The World Wide Web Consortium (W3C, 2005) is an international consortium that aims to develop web standards. Its mission is pursued through making general guidelines that will lead to web standards. W3C's Web Accessibility Initiative (WAI), which is part of the W3C, focuses on enabling people with disabilities to create and interact with web content. The WAI has developed Web Content Accessibility Guidelines (WCAG 1.0 and WCAG 2.0), which describes how to make Web content and Web sites accessible. Although it is possible to conform either to WCAG 1.0 or to WCAG 2.0 (or both), the W3C recommends the use of WCAG 2.0 for the new and updated content. WCAG 2.0 is organized around four design principles that provide the foundation for web accessibility (perceivable, operable, understandable, and robust) (WCAG 2.0, 2008).  Each principle has guidelines and each guideline has testable success criteria (SCs) at levels A, AA, or AAA. SCs are the basis of determining the conformance of a level in WCAG 2.0. Table 1 demonstrates the WCAG 2.0 conformance levels.

**Table 1. WCAG 2.0 Conformance Levels**

| **Conformance Level** | **A** | **AA** | **AAA** |
|---|---|---|---|
| **Explanation** | All SCs of level A are satisfied. This is the “minimum standard” which a website must meet to be considered accessible for any significant disability groups. | All SCs of Level A and AA are satisfied. This is a "professional practice standard”, which a website should meet to be accessible to a broad range of disability groups. | All SCs (at all conformance levels) are satisfied. This is a "gold standard" of maximum accessibility, which some websites may choose. |

Many issues have been raised about WCAG 2.0 such as the level of understanding of accessibility issues required when using them (Bittar et al., 2012).WCAG 2.0 documents are not facile to understand and require a certain level of technical knowledge of accessibility. Therefore, when developers or designers are required to implement accessibility, they do not always understand how to achieve the desired results. WCAG 2.0 documents are difficult to navigate and locate. All the documents related to this guideline exceed 450 pages with a few hundred navigation links on each page. As such, WCAG 2.0 could be unusable by real-world developers. Of a greater concern is that WCAG 2.0’s emphasis on perfect scores on all SCs to receive the level conformance logo since single SCs violation in a priority would render a website inaccessible. This makes it impossible for websites to achieve any acceptable level of conformance. WAI ignored that in the website development process, developers do not seek for perfection instead they aim for continuous, pragmatic improvement over versions. There are more than 60 SCs to follow within WCAG 2.0, which adds significant burdens to web developers to be successful. We believe that guidelines would be more usable and applicable if there is a quantitative metric to rank barriers (that corresponds with SCs violations) based on their commonalities and influence on the accessibility issue instead of focusing on all types of defects that may rarely occur. A quantitative metric can help reduce the time and effort required for the evaluation process. This stimulates interest in this research domain to further investigate the “What” question of “What is the influence of each barrier on websites accessibility?” and “Do all barriers contribute equally to the issue of accessibility?”

## WEB ACCESSIBILITY QUANTITATIVE METRIC

In this section, we cover the need for an accessibility metric as well as the importance of merging such metric with automatic evaluation. In addition, the requirements of the metric and the assumptions we made are described.

### The Need for an Accessibility Metric

Past studies proposed quantitative metrics to check the status of accessibility (Vigo et al., 2007; Brajnik and Lomuscio, 2007; Parmanto and Zeng, 2005). However, most of them need human judgment and expert review. Web Accessibility Barrier Severity (WABS) is different since it does not focus on measuring the accessibility of websites. We are more concerned about revealing the persistent barriers that limit the accessibility. A different measurement is needed for scientific exploration that takes into consideration: (1) The importance of the barrier to the other barriers that belong to the same priority level (2) The importance of the barrier to the website (3) The importance of the barrier to all the other barriers in the whole dataset (all websites).

### The Need for Automatic Evaluation

The proposed metric (WABS) is designed to work with an automated accessibility tool. It is calculated automatically from evaluation reports yielded by the evaluation tool. Automated testing can help in reducing the time and effort required for the evaluation process. Automated tools reduce the need for massive manual checking.Moreover, they allow large-scale assessment of aggregate Web sites.Our metric will utilize the results of the A-Checker tool since it covers the WCAG 2.0 guidelines and assists web developers to review and change violations in real time.

### Requirements and Assumptions

In this section, the requirements and assumptions for our metric are described to prove that it meets the measurements for scientific research.

**Requirement 1:** *Metric Results should be normalized.*

***Assumption1:*** The barriers must be measured in a quantitative score that provides a continuous range of values from perfectly major barriers to completely not important. A quantitative numerical score would allow assessment of barriers along groups of Web sites.

***Assumption 2:*** The metric values must have a large discriminating power beyond simply major to minor. A metric with good discriminating power would allow assessment of barriers affection on accessibility and show if there is significant difference between them. In order to rank barriers according to their severity; a weight with a positive value associated with each barrier is chosen so that values of the final metric are ranged 0 to 1. The closer the result of the metric is to 0 the less severe the barrier is and the closer to 1 is the more severe it is.

**Requirement 2:** *The metric should give one value for each barrier based on its influence on the priority class that it belongs to as well as to the whole dataset (the whole set of web page being assessed )*

***Assumption 1:*** Beside total frequency of errors for each barrier in the web page, the metric should also take into account the total number of times each barrier has been tested. The metric should not be based on the absolute frequency of errors found for each barrier but in the relative number of found errors in relation to the number of tested cases. That is, the ratio of errors and number of tested cases. In other words, the number of websites that contain the same type of barrier should be taken into account.

***Assumption 2:*** The metric should be scalable to conduct large-scale Web accessibility studies. Large-scale accessibility assessments require a metric that supports aggregation and allow for statistical analysis in order to measure if there is a significant difference between values.

***Assumption 3:*** The priority of unfulfilled success criteria (barrier) should be reflected in the final result. Within WCAG 2.0 priority "A" SCs have more impact on the accessibility level of a web page than priority "AA" SCs and so on. No matter the value assigned for each priority, the value should reflect the difference between these priorities based on their importance. The unique restriction when selecting these weights is that 1 *>priorityA_weight>priorityAA_weight>priorityAAA_weight> 0*

In our study we adopted priorities weight suggested by (Vigo et. al., 2007), here (priority "A"=0.8, priority "AA"=0.16, priority "AAA"=0.04)

**Requirement 3:** *The measurement should be normative.*

The metric should be derived from standard guidelines of Web accessibility such as the WCAG. We use the WCAG 2.0 as the foundation for our metric since it is the latest one from WAI

**Requirement 4**: *Problems that need human judgment should not have influence on the final metric.*

Theoretically, the metric can be used to calculate scores based on SCs in WCAG 2.0 priorities. However, we focus only on barriers that can be checked automatically. Furthermore, when performing an automatic test, all web pages get the same report of generic problems that need manual review. Thus, a metric based on automatic evaluation should not take into account these generic problems.

## METRIC CALCULATIONS

Variables, constants and our final metric are discussed in this section. In order to demonstrate how we build the final metric, we break it down into sub equations. Equations cover the importance of a barrier to its priority class, its webpage and to the whole evaluation dataset.

### The Importance of the Barrier to the other Barriers that belong to the same Priority Level

$$\frac{\sqrt{\sum_{d=0}^{d=N} freq(bi,pc)^2}}{\sqrt{\sum_{d=0}^{d=N} b(pc)^2}} * Pc \qquad (1)$$

Where

*d*= Document (webpage) being tested

*N*= Total number of tested documents (webpages)

*bi*= barrier (violation) being checked

*Pc*= priority class weight which the tested barrier belongs to

Eq. (1) steps:

1. Divide the result of formula (1.1) over the result of formula (1.2)
2. Multiply the result of step 1 by the weight of the priority class which the barrier (bi) belongs to. Where (Priority "A"=0.8, priority "AA"=0.16, priority "AAA"=0.04)

$$\sqrt{\sum_{d=0}^{d=N} freq(bi,pc)^2} \quad (1.1)$$

Eq.(1.1) represents the length of a certain barrier that belongs to a specific priority level across the whole documents. For each webpage we count how many times a certain barrier (bi) appears in (i.e. frequency). After that we follow the steps below:

1. Calculate the square of (bi) frequency for each webpage
2. Find the summation of squares for step 1 across the whole dataset
3. Calculate the square root for step 2 final results

$$\sqrt{\sum_{d=0}^{d=N} b(pc)^2} \quad (1.2)$$

Eq.(1.2) calculates the length of the all barriers across the whole documents that belong to the same barrier's priority class. Note that this formula's result will be fixed for each set of barriers that belong to the same priority level. Steps are described hereunder:

1. For each webpage find the total number of barriers that belong to ( bi) priority class
2. Find the square of step 1
3. Repeat the previous steps for the whole dataset(all webpages)
4. Sum up the results of step 3 (the summation of barriers squares)
5. Find the square root for the final result (step 4)

**The Importance of the Barrier to the Whole Documents**

$$\frac{n(b)}{N} \quad (2)$$

Where
*n(b)*= the number of different documents (webpages) the barrier appears in
*N*= the total number of documents
This formula shows the ratio of documents that the barrier appears in to the total number of documents in the dataset. Steps are described below:

1. Calculate how many times a certain barrier (bi) appears in a different document.
2. Divide step 1 over the total number of documents (Fixed across the documents).

**The Importance of the Barrier to all the other Barriers in the Whole Dataset (the Whole Tested Webpages)**

$$\sum_{d=0}^{d=N} \sqrt{\sum_{b=0}^{b=t} b^2} \quad (3)$$

Eq. (3) is fixed for the all documents and calculated once. It calculates the total lengths of all barriers across the whole documents. The steps are as follow:

1. For each document find the square of all barriers that appear in it.
2. Sum up the result of step 1.
3. Find the square root of step 2
4. Repeat step 3 to all documents
5. Find the summation of step 4 results.

**The Final Metric**

Figure 1 represents the final metric (WABS) after aggregating the previous formulas. WABS is applied to each barrier separately to calculate its weight. Following, we can rank the violations severity once we find each barrier weight.

$$\frac{\sqrt{\sum_{d=0}^{d=N} freq(bi,pc)^2}}{\sqrt{\sum_{d=0}^{d=N} b(pc)^2}} * \frac{n(b)}{N} * \frac{Pc}{\sum_{d=0}^{d=N}\sqrt{\sum_{b=0}^{b=t} b^2}}$$

**Figure 1. WABS Metric.**

## EXPERIMENTAL ANALYSIS OF THE PROPOSED METRIC

An experiment was conducted to demonstrate our findings using this metric. Homepages of 500 different Malaysian websites were evaluated. The homepages covered many sectors: public and private universities, education, arts and entertainments, recreations and sports, sciences and environments. Violations that can be evaluated automatically by A-Checker tool are analysed. Generic reports and errors that needed human review were ignored. Table 2 shows the list of barriers that were analysed along with their priority level and calculated weight ranked in a descending order. From the table below we note that the barriers weights fell within the range (0-1). Moreover, barriers weight differs even for barriers belong to the same priority level since each violation has a unique weight. One important finding is that barriers belong to priority "A" does not necessarily mean they are more severe than other priority levels. For example, the barrier (1.4.4) with a priority level "AA" scored higher weight than barrier (2.1.1) with priority level "A". In other words, barriers influence on accessibility is not only bounded to its importance to the priority level. Other factors affect the weight such as the importance of the barrier to other barriers within the same webpage and its importance to the whole evaluation set (the 500 webpages in our experiment).

**Table 2. The Ranked Barriers with Their Weights**

| **Barrier Id** | **Priority Level** | **Weight** |
|---|---|---|
| 1.1.1 | A | 0.00011 (most severe) |
| 1.4.4 | AA | 1.22E-05 |
| 1.4.6 | AAA | 4.40E-06 |
| 2.1.1 | A | 2.04E-06 |
| 3.3.2 | A | 1.78E-06 |
| 1.3.1 | A | 1.29E-06 |
| 2.4.4 | A | 1.05E-06 |
| 3.1.1 | A | 8.67E-07 |
| 2.4.6 | AA | 1.38E-07 |
| 4.1.1 | A | 1.29E-07 |
| 1.4.1 | A | 4.73E-08 |
| 2.2.2 | A | 1.73E-08 |
| 2.4.2 | A | 4.37E-09 |

| | | |
|---|---|---|
| 2.4.1 | A | 3.82E-09 |
| 3.2.2 | A | 1.67E-09 |
| 4.4.1 | A | 4.98E-10 |
| 3.3.1 | A | 3.15E-10 |
| 2.2.1 | A | 7.87E-11 (least severe) |

## CONCLUSIONS AND FUTURE WORKS

Web Accessibility Barrier Severity (WABS) is proposed. WABS requirements and assumptions provide evidence that is suitable to scientific research. The metric takes into account (1) The importance of the barrier to the other barriers that belong to the same priority level (2) The importance of the barrier to the webpage (3) The importance of the barrier to all the other barriers in the whole dataset. Results show that the importance of barriers differs and varies. Our future work includes evaluating a large set of websites. Furthermore, we want to implement a report generator system that enables web developers to statistically analyze web accessibility based on the WABS metric.

## REFRERENCES

A-Checker Tool. (2015). Retrieved April8, 2014 from http://achecker.ca/checker/index.php

Bittar, T., Amaral,L., Faria, F., Fortes, R. (2012). Supporting the Developer in an Accessible Edition of Web Communications: a Study of Five Desktop Tools. In *Proceedings of the Workshop on Information Systems and Design of Communication* (ISDOC '12). ACM, New York, NY, USA, pp.3-9.

Brajnik, G., Lomuscio, R. (2007). SAMBA: a semi-automatic method for measuring barriers of accessibility. In *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility* (Assets '07). ACM, New York, NY, USA, 43-50.

Lujan-Mora, S., Navarrete, R., Penafiel, M. (2014).Egovernment and web accessibility in South America. *First International Conference on eDemocracy&eGovernment (ICEDEG)*, 77-82, 24-25 April 2014. doi: 10.1109/ICEDEG.2014.6819953

Nizar, A.A., Obedidat, A., Abu-Addose, H.Y. (2013). Accessibility as an Indicator of Jordanian E-Government Website Quality. *Fourth International Conference on e-Learning "Best Practices in Management, Design and Development of e-Courses: Standards of Excellence and Creativity*, 156-160, 7-9 May 2013

Parmanto, B., Zeng, X. (2005). Metric for web accessibility evaluation. *J. Am. Soc. Inf. Sci. Technol.* 56, 13 (November 2005), 1394-1404. DOI=10.1002/asi.20233

Vigo, M., Arrue, M., Brajnik, G., Lomuscio, R., Abascal, J.(2007). Quantitative metrics for measuring web accessibility. In *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A)* (W4A '07).ACM, New York, NY, USA, 99-107.

W3C. (2005). Retrieved March 2, 2015 from http://www.w3.org/WAI/intro/accessibility.php

WCAG 2.0. (2008). Retrieved April5, 2014 from: http://www.w3.org/TR/WCAG20/