Towards Metrics for Web Accessibility Evaluation

Sinésio Teles de Lima
Mestrado em Gestão do Conhecimento e da Tecnologia da Informação
Universidade Católica de Brasília
SGAN 916 Norte Av. W5 - Sala A121 - Asa Norte - Brasília - DF - Brasil - CEP 70790-160
stlima@gmail.com
+55 61 3448-7137

Fernanda Lima
Universidade Católica de Brasília
ferlima@ucb.br
+55 61 3448-7121

Káthia Marçal de Oliveira
Universidade Católica de Brasília
kathia@ucb.br
+55 61 3448-7125

ABSTRACT
Nowadays, accessibility is a crucial factor for Web site development and use, and yet, people with visual disabilities face many accessibility barriers that hinder the adequate understanding of Web contents. In Brazil, the Federal Government published a law that formalizes the mandatory accessibility to Governmental Web sites' content. In this context, it is necessary to define ways to evaluate accessibility to guarantee the quality of these sites in this respect. Therefore, this article proposes a means of Web accessibility evaluation through metrics.

Author Keywords
Web accessibility, Metrics.

ACM Classification Keywords
K.5.2 [Legal Aspects of Computing]: Governmental Issues – Regulation; K.4.2 [Computers and Society]: Social Issues; D.2.8 [Metrics].

INTRODUCTION

Organizations use Web technology to make products and services available to potential clients. In this context, Web content accessibility is an important component of quality that must be observed among services offered by Web site providers. In Brazil, the Law Decree 5.296 explicitly states this importance, by compelling Governmental Web site contents to be in accordance with Web accessibility guidelines.

Therefore, it is crucial to define a means of evaluating accessibility in order to allow the elaboration of better software contracts between clients and providers. In this aspect, Service Level Agreements (SLAs) are used to establish minimum quality patterns pertaining to services supplied by providers to their clients [10]. The SLA's main component is a service catalog where one can define, among other things, measurable factors, or indicators that will permit visibility as to the specified service levels that were either violated or respected.

The present work proposes the use of software metrics to evaluate the accessibility that can be used to define service level indicators for Web accessibility SLAs. The following sections present Web accessibility concepts, our approach to Web accessibility evaluation and the work in progress.

WEB ACCESSIBILITY

Web accessibility means that people with visual, physical, speech, cognitive or neurological disabilities are given the opportunity to perceive, understand, navigate and interact with the Web [7]. According to [12], to perceive is "to become aware of something via the senses"; to understand is to grasp the meaning of something; and to interact is "to have and effect on each other".

According to W3C [1], evaluating Web sites for accessibility can be done using approaches such as: preliminary review, conformance evaluation using accessibility tools to determine if a Web site meets accessibility standards such as the Web Content Accessibility Guidelines (WCAG) and involving users in the evaluation.

The WCAG [4] proposes 14 general guidelines which are composed of checkpoints that explain specific accessibility aspects and techniques to use them; and of priorities that show how critical their implementation is. The current version of WCAG recommendations is 1.0. In the 2.0 WCAG Working Drafts [5], verification points and priorities have changed to 'success criteria' and 'conformance levels'. Furthermore, the guidelines were regrouped into four basic principles: "(i) Content

must be perceivable, (ii) Interface components in the content must be operable, (iii) Content and controls must be understandable, (iv) Content should be robust enough to work with current and future user agents (including assistive technologies)". Each principle contains general guidelines organized into levels and success criteria. The W3C Consortium describes the relationship between the two versions [3].

Other methods to evaluate Web accessibility can be found in the literature [6][11]. One of these methods, called participatory observation [11], proposes the evaluation of interactive sessions with disabled users and an observer, in an effort to identify barriers in Web access and strategies found by the user to try to overcome them.

OUR APPROACH TO EVALUATE WEB ACCESSIBILITY

In this research, the Goal-Question-Metric (GQM) approach [2] was used to define what was going to be evaluated. The main idea of GQM is that measurement should be goal-oriented. Initially, an explicit measurement goal is defined. Subsequently this goal is refined into several questions that break it down into its major components. Then, each question is refined into metrics that, when measured, will provide information to answer these questions. By answering the questions we will be able to analyze if the goal has been attained. The GQM goal for this work is:

To analyze Web sites,

for the purpose of evaluating,

with respect to accessibility,

from the view points of users with disabilities,

in the context of Brazilian Federal Government.

According to the definitions presented in the previous section, it was possible to identify that accessibility evaluation means, mainly, to evaluate the user's capacity to perceive, operate and understand Web site content. Furthermore, there is a concern, associated with accessibility, with respect to user capacity to reach the desired objective with an acceptable effort in a satisfactory manner. This aspect is explored by ISO/IEC 9126 [8] with respect to quality in use, i.e., the software product's capacity to guarantee that users will reach their specific goals with efficacy, productivity, security and satisfaction, in specific contexts of use. Based on these premises, the following questions were elaborated to evaluate accessibility:

Q1. What is the degree of perception with respect to Web site contents?

Q2. What is the degree of ease of operation of Web site contents?

Q3. What is the degree of understanding of Web site contents?

Q4. What is the user's efficacy in specific task execution while interacting with the Web sites?

Q5. What is the percentage of user productivity in specific task execution while interacting with the Web sites?

Q6. What is the degree of user satisfaction in specific task execution while interacting with the Web sites?

Based on the comparison between the two WCAG versions, metrics were defined to answer these questions (Table 1). The data was obtained in two ways: using a questionnaire filled out during participatory observation sessions and using an automatic tool called TAW [14]. The tool was chosen, among other reasons, because of its capacity to evaluate pages of arbitrary depth in a Web site and generating HTML reports with quantity of violations per page.

Three Brazilian Federal Government Web sites were chosen on which to apply these metrics (www.receita.fazenda.gov.br, www.previdencia.gov.br, and www.ibge.gov.br). The choices were based on the results of a Brazilian contest named iBEST Contest, where the best Brazilian sites, divided into different categories, receive prizes annually. The three sites above received the best prizes for the Government category.

One task was defined for each site. The first task was to find out whether a person is going to receive income tax returns, by filling out two fields: a number similar to a Social Security Number and a dynamically generated image shown as a capcha field (a code shown as a distorted image for security purposes). The second task was to identify the necessary documents to apply for a Social Security application. The third task was to find the Contact part of the site and identify the subjects that can be dealt with by email.

During participatory observation sessions, the tasks were executed by ten users with varying degrees of visual disability and different Web use expertise. All the sessions were conducted in environments with personal computers with keyboard, mouse, speakers, Web browsers (Internet Explorer or Firefox Mozilla) and screen reader software (Jaws or Virtual Vision).

Our approach was presented to the Brazilian scientific community in previous events related to multimedia and Web [9] and software quality [10]. These papers show some of the preliminary results. The current results (Table 1) show that the Web content levels of perception and understanding increase whenever the quantity of violations to the WCAG 1.0 checkpoints decreases. Analyzing the results of questions 1, 2 and 3, one can conclude that task 1 had the lowest degree of perception, operation and understanding, while task 3 obtained the best levels of the accessibility principles.

Question	Metric	Value	Task 1	Task 2	Task 3
Q1	M1. Degree of content perception	From 0 (completely imperceptible) to 7 (completely perceptible)	2,9	5,5	6
Q1	M2. Quantity of violations to the perception principle in each task	Integer (X>=0)	15	59	22
Q2	M3. Degree of operation related to keyboard use	From 0 (completely inoperable) to 7 (completely operable)	5,4	5,8	6,2
Q2	M4. Degree of operation related to time for execution	From 0 (completely unsatisfactory) to 7 (completely satisfactory)	2,3	5,1	5,8
Q2	M5. Degree of operation related to navigation complexity	From 0 (very complex) to 7 (very simple)	5	5,8	6
Q2	M6. Degree of operation related to anchor existence	From 0 (very difficult) to 7 (very easy)	5,2	6	5,8
Q2	M7. Quantity of violations to the operation principle in each task	Integer (X>=0)	2	0	0
Q3	M8. Degree of understanding and comprehension	From 0 (completely not understood) to 7 (completely understood)	3,3	5,6	6
Q3	M9. Quantity of violations to the understanding principle in each task	Integer (X>=0)	5	4	0
Q4	M10. Percentage of Efficacy	Percentage = (number of users that concluded the task) / (number of users that executed the task)	10%	100 %	100 %
Q5	M11. Percentage of productivity (up to 10 min.)	Percentage = (number of users that concluded the task in 10 minutes or less) / (number of users that executed the task)	10%	80%	90%
Q6	M12. Degree of satisfaction	From 0 (completely dissatisfied) to 7 (completely satisfied)	1	5,6	5,7

Table 1. Metrics defined for Accessibility Evaluation and their application

WORK IN PROGRESS

Through the analysis of metrics and obtained data, it was possible to define a preliminary parallel that indicates the following: the lower the degree of perception, operation and understanding of web content, the lower will be the efficacy, the productivity, and the satisfaction of the users, during task execution with those contents.

The next steps aim to improve the analysis of the metrics, in order to produce indicators that can be used to compose a service catalog for a Web accessibility SLA. Accessibility SLAs will be useful by contributing to accessibility initiatives, in the future, as formal instruments, between Web content developers and clients.

REFERENCES

1. Abou-Zahra and EOWG. Evaluating Web Sites for Accessibility: Overview. http://www.w3.org/WAI/eval/Overview.html

2. Basili, V. R., Caldiera, G. and Rombach, H. D. The Goal Question Metric Paradigm. Encyclopedia of Software Engineering, v. 1, 528-32. John Wiley & Sons, 1994.

3. Candwell, B. et al. (Editors) Comparison of WCAG 1.0 checkpoints to WCAG 2.0. 2006. http://www.w3.org/TR/WCAG20/appendixD.html.

4. Chisholm, W., Vanderheiden, G. and Jacobs, I. (Editors) Web Content Accessibility Guidelines 1.0. 1999. http://www.w3.org/TR/WCAG10/.

5. Henry, S. L. (Editor) WCAG 2 FAQ. 2007. http://www.w3.org/WAI/WCAG20/wcag2faq.

6. Henry, S. L. Just Ask: Integrating Accessibility Throughout Design. http://www.w3.org/WAI/eval/users.html#ut-access.

7. Henry, S. L. and EOWG Group (Editors). Introduction to Web Accessibility. 2005 http://www.w3.org/WAI/intro/accessibility.php.

8. International Organization for Standardization. ISO/IEC 9126-1. Software Engineering - Product quality – Part 1: Quality model, 2001.

9. Lima, S. T.; Lima, F.; Oliveira, K M. Investigação da Acessibilidade em Sítios Web do Governo Brasileiro: Uma Abordagem para o Estabelecimento de SLA de Acessibilidade. Proc Brazilian WebMedia 2006 Symposium, SBC Press (2006).

10. Lima, S. T.; Lima, F.; Oliveira, K M. Avaliação da Acessibilidade de Sítios Web por meio de métricas de software. Proc Brazilian Software Quality 2007 Symposium, SBC Press (2007).

11. Melo, A.M., Baranauskas, M.C.C. and Bonilha, F.F.G. Avaliação de acessibilidade na Web com a participação do usuário: um estudo de caso. Proc Brazilian IHC Symposium, SBC Press (2004).

12. Merriam-Webster Dictionary http://www.mw.com/dictionary

13. Office for Government Commerce Itil. The Key to Manage IT Services: Service Delivery – Version 1.2; Crow, 2001.

14. Web Accessibility Test (TAW). 2006. http://www.tawdis.net/taw3/cms/en