A Unified Web Evaluation Methodology Using WCAG

Eric Velleman[1], Christophe Strobbe[2], Johannes Koch[3], Carlos A. Velasco[3],
and Mikael Snaprud[4]

[1] Accessibility Foundation, Oudenoord 325, 3513EP Utrecht, Netherlands
e.velleman@bartimeus.nl
[2] Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Heverlee-Leuven, Belgium
{Christophe.Strobbe,Jan.Engelen}@esat.kuleuven.be
[3] Fraunhofer-Institut für Angewandte Informationstechnik (FIT), Schloss Birlinghoven,
D-53757 Sankt Augustin, Germany
{Johannes.Koch,Carlos.Velasco}@fit.fraunhofer.de
[4] Agder University College, Serviceboks 422, N-4604 Kristiansand, Norway
mikael.snaprud@hia.no

Abstract. Checking of web accessibility can be carried out in several ways along the same international standards and depending on the scale, the quality, availability of tools and the interpretation method that is followed. The Unified Web Evaluation Methodology is developed by European expert organizations and offers test descriptions to evaluate WCAG 1.0 conformance covering level AA, a clear sampling scheme, improved aggregations supporting confidence levels, Score cards and other instruments to help communicate the results of evaluations clearer and on more levels. The aim is to establish the UWEM as the basis for web accessibility evaluation, policy support and possible certification in Europe.

Keywords: web accessibility, conformance, certification, unified web evaluation methodology, uwem.

1 Introduction

Accessibility checks of webresources can be carried out in different ways even if the checks are based on the same internationally accepted guidelines. The Unified Web Evaluation Methodology (UWEM) is developed to ensure that large scale monitoring and local evaluation are compatible and coherent among themselves and with the Web Content Accessibility Guidelines from W3C/WAI [1]. The methodology is based on the Web Content Accessibility Guidelines of the World Wide Web Consortium (W3C) Web Accessibility Initiative (WAI), and will be synchronised with the foreseen migration from WCAG 1.0 to WCAG 2.0 [2]. Thus, the UWEM will sharpen the support for evaluating, certifying, and benchmarking web content in conformance with WCAG.

The Unified Web Evaluation Methodology offers test descriptions to evaluate WCAG 1.0 conformance covering level AA, a clear sampling scheme, improved

aggregations supporting confidence levels, Score cards and other instruments to help communicate the results of evaluations clearer and on more levels. The aim is to establish the UWEM as the basis for web accessibility evaluation, policy support and possible certification in Europe. In a recent communication on eAccessibility, the European Commission announces they will foster the use of three approaches not yet widely used in Europe: (i) accessibility requirements in public procurement, (ii) accessibility certification, and (iii) better use of existing legislation [3]. The UWEM aims to support these approaches.

2   The WAB Cluster

The UWEM has been produced from the Web Accessibility Benchmarking Cluster (WAB Cluster). The work of the Web Evaluation Benchmarking Cluster involves three European projects with a total of more than 21 partners from eight countries. They work together on harmonisation and evaluation to provide a European base for evaluation of Websites for accessibility. The projects participating in the WAB cluster are funded by the European Union in the second FP6 IST call (2003) of the eInclusion Strategic Objective. More information about the projects can be found on the project web sites:

http://www.wabcluster.org
http://bentoweb.org
http://www.eiao.net
http://www.support-eam.org

The Ben To Web project supports the European public and private sector by providing new software modules and methodologies for Web Accessibility, which are not analysed by existing tools due to their inherent complexity. BenToWeb also produces suites of test files for WCAG 2.0, which can be used for benchmarking accessibility evaluation and repair tools. More information: http://www.bentoweb.org

The EIAO project establishes the measurement machinery for a prototype EU Internet Accessibility Observatory. Frequently updated measurements will be available online from a data warehouse to enable large scale benchmarking, and provide a basis for policy-making, research and actions to improve the accessibility to Internet content. More information: http://www.eiao.net.

The Support EAM project has produced a CEN Workshop Agreement [4] on certification, many evaluation materials, tutorials and a curriculum. It has put down the basis for the use of the UWEM as a framework for a unified inspection process as part of a European certification scheme for Web accessibility based on interaction with users [5]. More information: http://www.support-eam.org

3   UWEM, Aims and Requirements

The Unified Web Evaluation Methodology provides an evaluation procedure offering a system of principles and practices for expert and automatic evaluation of Web accessibility for humans and machine interfaces. The Methodology is designed to be

conformant with WCAG 1.0 priority 1 and 2 checkpoints with regard to technical criteria.

The UWEM aims to increase the value of evaluations by basing them on a shared interpretation of WCAG 1.0 and a set of tests that are sufficiently robust to give stakeholders confidence in results. Web content producers may also wish to evaluate their own content and the UWEM aims to also be suitable to these users.

The methodology is designed to meet the following requirements:

- Technical conformance to existing Web Accessibility Initiative (WAI) Recommendations and Techniques documents;
- Tool and browser independence: questions and tests are given in a 'pure' form, making them as tool independent as possible;
- Unique interpretation: questions shall have only one way of being interpreted;
- Replicability: different Web accessibility evaluators who perform the same tests on the same site should get the same results within a given tolerance;
- Translatability: the methodology will address localization issues;
- Compliance with Regulation (EC) No 808/2004 of the European Parliament and of the Council of 21 April 2004 concerning Community statistics on the information society.

In the methodology information has also been included about:

- Scope and sampling;
- Reporting, interpretation and integration/aggregation of results.

The UWEM primarily covers methods to evaluate documents based on the following technologies:

- HTML 4.01
- XHTML 1.0 and 1.1
- CSS 2.x
- Other embedded objects in (X)HTML resources

4   Target Audience of the UWEM

The target audiences for the UWEM include, but are not limited to:

- Web accessibility benchmarking projects
- Possible Certification Authorities
- Web content producers wishing to evaluate their content
- Developers of Evaluation and Repair Tools
- Policy makers and Web site owners
- Other organizations evaluating Web sites

The European Commission, national governments and other organizations who wish to benchmark Web accessibility will be able to use the UWEM to carry out the

evaluations and compare their results in a meaningful way. For this reason, a score-card solution is presented.

The UWEM is an evaluation methodology and is not intended to provide information for Web content producers wishing to produce content conformant with WCAG 1.0. That information is provided in the WCAG 1.0 Techniques Documents [6] that are available through the W3C/WAI web site.

5   Automatic, Expert and User Testing

Web accessibility may be tested through automatic, expert and user testing. Some companies evaluating websites for accessibility on a commercial basis only use manual evaluation while others use only automatic testing or a combination of both [5]. Currently expert testing and automatic testing are supported in the UWEM. A section including testing protocols for specific web evaluation tests with people with disabilities has been delivered in an earlier stage. The testing protocols are used in the Ben-ToWeb project.
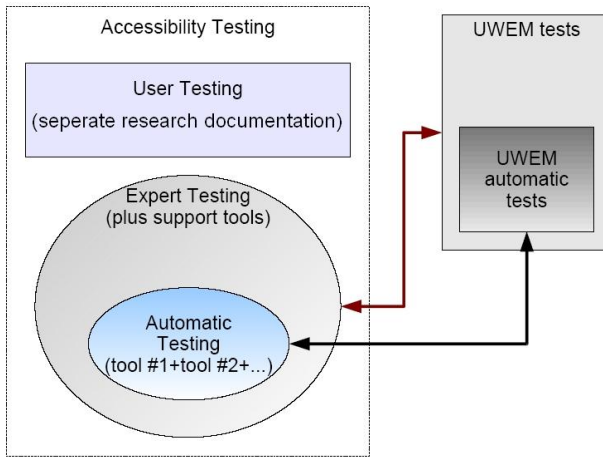

Fig. 1. The different evaluation methods and the coverage of the UWEM

Figure 1 describes three different evaluation methods. Currently two (automatic and expert) are covered in the UWEM. Automatic testing of all checkpoints for accessibility is not possible. If a checkpoint is fully automatable, this is indicated. Figure 1 shows that automatic tool #1 and #2 can only identify a subset of all possible accessibility barriers. Tools better support testing of very large numbers of Web resources within one or multiple Web sites. Some tools can also act as support systems in an expert evaluation process. The tools provide reliable results for a subset of tests and can not only speed up the process by performing some tasks automatically, but also, by providing hints about barrier locations, indicate areas the expert evaluators should focus on.

User testing is able to identify barriers that are not caught by other testing means, and is also capable of estimating the accessibility for tested scenarios. However, user testing is quite specialised, thus it is not generally suitable for conformance testing. The best approach to ensure both accessibility and the UWEM conformance is to use a combined approach encompassing all evaluation methods: automatic, expert evaluation and user testing of the Web site. Involvement of users in the evaluation of web content is described in the W3C/WAI evaluation suite (available on the WAI website).

6   Claiming Conformance

The purpose of the UWEM is to guarantee replicability of results. The unambiguous expression of the tested resources is of key importance for the aggregation and comparison of results. To claim conformance with the UWEM, it is minimally required that the resources sample and the scope for the evaluation are defined and that all resources in the sample should pass all applicable tests to the corresponding conformance level. The size and selection of samples is also described in the UWEM.

The conformance levels to the UWEM replicate those of WCAG10 [1]. The claims of accessibility conformance for manual evaluation according to the UWEM methodology must use the following form:

1. The UWEM version and its URI identifier, i.e.;
2. The URI to a document detailing the scope and the sample to which the claim refers;
3. The level of conformance.

Evaluation tools can also claim conformance to the UWEM 1.0. In that way, experts evaluating web sites according to the UWEM 1.0 will be able to rely on the results of the tool for the fully automatable tests of the methodology. To claim conformance to the UWEM 1.0, the tool MUST implement all fully automatable tests to the corresponding conformance level.

7   Tests and Structure of Tests for Conformance Evaluation

The UWEM offers tests for automatic and expert testing of Priority 1 and Priority 2 checkpoints of WCAG 1.0. It does not repeat information available in W3C documents. Instead, it provides pointers to the relevant W3C information if available and extends only information when necessary for the defined tests. Web content passes a checkpoint if it fails none of the applicable tests for that checkpoint. Web content fails a checkpoint if it fails any of the applicable tests for that checkpoint.

The structure of the tests is harmonized throughout the document and based on extensive evaluation experience by the partners in the projects:

1. Guideline: Quotation of the corresponding WCAG 1.0 guideline with pointers to additional clarifications on W3C website if available.

---

[1] http://www.w3.org/TR/WCAG10/#priorities

2. Checkpoint: Quotation of the corresponding WCAG 1.0 checkpoint. Pointers to additional clarifications on W3C website if available. For each checkpoint, a set of one or more tests is defined. If no automatic tests for a certain technology are defined, this means that there are no applicable tests for automated testing.
3. Technology specific tests: There are tests for (X)HTML resources and other technologies. Each test consists of:

- **Title and ID:** short descriptive title (informative) and unique identifier (normative);
- **Applicability criteria:** elements, attributes and combinations thereof used to determine the applicability of the test. Whenever possible, the criteria are presented as XPath expressions, otherwise a prose description is given;
- **Test procedure:** description in a tool-independent manner of the test procedure. The procedure may consist of multiple steps and is written so as to enable possible machine-testing;
- **Expected results:** statement defining the fail or pass conditions with regard to one or more steps in the test procedure. The elements or content specified in the accessibility criteria pass the test if the result is not FAIL;
- **Fully automatable:** statement whether the test procedure can be fully automated (yes/no).

8   Aggregation and Reporting of Test Results

The UWEM provides models for aggregation and reporting of test results including a User Centric Accessibility Barrier model, EARL schema, an Evaluation Report Template and a Scorecard that gives easy understanding and clear interpretation to the end-result. Following the European Commission regulation 808/2004 concerning community statistics for the information society, the UWEM provides a model for calculating the accessibility barrier probability for single Web pages and Web sites. A single value can then describe the probability of creating a barrier by violating any of the evaluation tests.

Scorecards	RGB	Barrier probability Fs	Level of accessibility	Interpretation
A	204/236/255	Fs = 0.00%	Accessibility is fully supported	Full accessibility
B	153/204/255	0% < Fs <= 25%	Within levels of confidence no or very few accessibility barriers have been identified	Very good accessibility
C	0/153/204	25% < Fs <= 50%	A few accessibility barriers have been identified	Medium accessibility
D	51/102/153	0% < Fs <= 75%	Several accessibility barriers have been identified	Poor accessibility
E	0/51/102	75% < Fs <= 100%	Web Site is inaccessible	Very poor accessibility
n/a	255/255/255	Not measured	Not tested or not applicable	n/a

Fig. 2. UWEM barrier scorecard scores represented by letters and colours to indicate barrier probability. Note that the levels chosen are preliminary values that will be evaluated when real data measurements from the EIAO project are available.

The scorecards are based on this value that describes the probability of creating a barrier. The scorecards can be used to monitor the status and progress of single Web sites and groups of Web sites over time.

The scorecard gives a helicopter view of the status and progress, and allows to monitor a set of indicators. Figure 2 shows the colour and letter codes used for the UWEM barrier status score. The coloured codes were based on the shades of blue as this is a set of colours that ensures a clear distinction regardless of colour-vision impairment and on the black-and-white printout.

The Manual Evaluation Report Template included in the UWEM closely follows the W3C evaluation suite template and offers evaluators a clear template to report accessibility.

9   Conclusion and Future Work

The Unified Web Evaluation Methodology aims to increase the value of evaluations by a shared interpretation of WCAG 1.0 and a set of tests that are sufficiently robust to give stakeholders confidence in the results. It offers detailed test descriptions covering WCAG1.0 level AA, a clear sampling scheme, improved aggregations supporting confidence levels, Score cards and other instruments to help communicate the results of evaluations clearer and on more levels.

There are parts of the methodology where more information and research is needed. This requires more evaluation results than currently available. More detailed recommendations will be added in the next version of the UWEM and will be largely based on the results from experiments within the EIAO and BenToWeb projects. Also more public and user comments are necessary. A possible channel for that is provided through the forum on www.wabcluster.org/forum/. Based on these evaluations and experiments, a next version will be made available that includes support for WCAG2.0. This next version will also include more information on user testing, aggregation, reporting and sampling information and offer the possibility to choose the WCAG version.

The current version of the UWEM has been used extensively by different European organization for commercial testing of websites. Already a French and a Spanish translation exist (www.wabcluster.org).

This project has been undertaken in the framework of the project BenToWeb IST-2-004275-STP funded by the IST Programme of the European Commission and involves the partners working in the projects EIAO and Support EAM.

References

1. W3C Recommendation. Web Content Accessibility Guidelines 1.0 (WCAG 1.0). In: Chisholm, W., Vanderheiden, G., Jacobs, I(eds.) World Wide Web Consortium (MIT, INRIA, Keio) (May 5, 1999) available at http://www.w3.org/TR/WCAG10/
2. W3C Working Draft. Web Content Accessibility Guidelines 2.0 (WCAG 2.0). In: Caldwell, B., Chisholm, W., Vanderheiden, G., White, J. (eds.) World Wide Web Consortium (MIT, ERCIM, Keio) (2006) available at http://www.w3.org/TR/WCAG20/
3. European Commission: Communication COM, 425 - Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of Regions. Brussels (September 13, 2005) (2005) also available at: http://europa.eu.int/information_society/policy/accessibility/com_ea_2005/a_documents/cec_com_eacc_2005.html
4. CWA 15554, Specifications for a Web Accessibility Conformity Assessment Scheme and a Web Accessibility Quality Mark. CEN (2006), available at ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/WAC/CWA15554-00-2006-Jun.pdf
5. Support, EAM.: deliverable D3.2 Stakeholders' needs for a Quality Mark (April 8, 2005) also available at http://www.support-eam.org
6. W3C note. Techniques for Web Content Accessibility Guidelines 1.0. In: Chisholm, W., Vanderheiden, G., Jacobs, I. (eds.) W3C Note. World Wide Web Consortium (November 6, 2000) (1999) Available at: http://www.w3.org/TR/WAI-WEBCONTENT-TECHS/
7. Web Accessibility Benchmarking Cluster: D-WAB2 unified web evaluation methodology (UWEM 1.0) (2006) available from http://www.wabcluster.org/uwem/