

Metrics for accessibility on the Vamolà project

Silvia Mirri
Department of Computer Science
University of Bologna
Via Mura Anteo Zamboni 7
40127 Bologna (BO), Italy
mirri@cs.unibo.it

Ludovico Antonio Muratori
Department of Computer Science
University of Bologna
Via Mura Anteo Zamboni 7
40127 Bologna (BO), Italy
muratori@criad.unibo.it

Marco Rocchetti
Department of Computer Science
University of Bologna
Via Mura Anteo Zamboni 7
40127 Bologna (BO), Italy
roccetti@cs.unibo.it

Paola Salomoni
Department of Computer Science
University of Bologna
Via Mura Anteo Zamboni 7
40127 Bologna (BO), Italy
salomoni@cs.unibo.it

ABSTRACT

How far are the Web sites from accessibility? Such a question implies a measure and, before, metrics or a set of metrics to quantitatively describe distances or closeness to one or more aspects of inclusion. Quantitative evaluations can effectively guide the improvement of sites quality and of skills which are necessary along the content authoring process. This paper details the hypothesized metrics and measures on the Vamolà project, which has been issued to integrate a monitor and a validator of the Web accessibility, according to the Italian Law. On Vamolà, the validation tool output can be effectively used in defining metrics inside the monitor application.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: Evaluation/methodology;

H.5.4 [Hypertext/Hypermedia]: User issues;

I.7.5 [Document Capture]: Document analysis;

K.4.2 [Social Issues]: Assistive technologies for persons with disabilities.

General Terms

Measurement, Human Factors, Verification.

Keywords

Accessibility, Monitoring accessibility, Evaluating accessibility, Metrics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

W4A2009 - Communication, April 20-21, 2009, Madrid, Spain.
Co-Located with the 18th International World Wide Web Conference.
Copyright 2009 ACM 978-1-60558-561-1...\$5.00.

1. INTRODUCTION

Italian regulations about Web accessibility state a set of mandatory requirements to be satisfied [5]. They meet a wide range of aspects about accessibility, also referring (and, indeed, restricting) the international guidelines [11] or some national existing rules such as the U.S. section 508 [7]. Every public institution is bound to follow such constraints both with automatic check procedures and with subjective, manual evaluation by accredited experts.

The Vamolà project is devised as a collaboration between the University of Bologna and The Emilia Romagna Region, gathering two integrated applications. The first one is a validator: its goal is getting together all the automatic procedures the law imposes and providing some guidelines to the subjective manual evaluation the regulations state as well. The second one is a monitor, whose aim is periodically querying the validation tool and reporting the accessibility of a given set of Web sites. Such a monitor utilizes messages and errors coming from the validator, in order to compute some sort of measures, on the strength of metrics. These metrics can, in turn, be overridden with extensions to the monitor itself. Simple formulas, as well as results of previous works, known in literature, have been taken into account during the design phase. Also the opportunity of expressing suitable parameters coming from expectations and preferences by a community of experts has been considered.

The goal of the whole project is aiding public institutions to effectively survey their Web contents and services and to point out a global, synthetic portrait of accessibility and its dynamics in time. Since a prototype of the monitor is being tested at present time, this paper details the features of Vamolà and its open trends about accessibility metrics and measures, without reporting results about its assessment, which is being done yet.

The remainder of this paper is organized as follows. Section 2 presents main work related to the definition and the implementation of metrics related to the quantitative evaluation of Web sites accessibility. In Section 3 we describe the main issues of the Vamolà project, illustrating some measures of accessibility,

provided in order to monitor and evaluate Web sites. Section 4 concludes the paper.

2. METRICS FOR MEASURING ACCESSIBILITY

Several works in the accessibility field are related to the definition and the proposal of quantitative metrics for measuring accessibility, totally based on WCAG 1.0 checkpoints [11]. In this Section we are going to present and briefly describe some of them.

Sullivan and Matson [9] proposed a proportion between potential points of failure and checked real errors on a single Web page and they called it “failure-rate”. The result range goes from 0 to 1.

$$failure_rate = \frac{real_errors}{potential_failures}$$

Such a rate does not take into account some fundamental issues, such as error impact and error nature (checkpoints could cause automatic errors, which can be detected by automatic validation tools, but also warnings and generic problems, which can be definitely identified only by a human operator). However, the authors were the first ones to propose the idea of continuous accessibility measurement, dealing with it in terms of a quantitative value, instead of the dichotomous accessible-inaccessible which was usually considered at that time. A numerical metric with continuous values would provide better discrimination power and promote a scientific approach to Web accessibility issues [6].

Parmanto et al. proposed an earlier version of their WAB (Web Accessibility Barrier) formula in [4] and a more mature one in [6]. As the previous one, also this metric takes into account only WCAG checkpoints. This formula considers the total pages of a Web site, automatically detectable accessibility errors and potential ones. The WAB score is calculated according to the following formula:

$$WAB_score = \frac{\sum_p \sum_v \left(\frac{n_v}{N_v} \right) w_v}{N_p}$$

where p is the total number of page of a Web site, v is the number of violations of a single Web page, n_v is the number of violations, N_v is the number of potential violations, W_v is the weight of violations in inverse proportion to WCAG priority level, N_p is the total number of pages checked. The number of violations of each checkpoint is a component of WAB score, hence a lower score means fewer accessibility barriers for people with disabilities, while a higher score indicates more barriers. A score of zero denotes that the Web site does not violate any Web accessibility guidelines and should have no accessibility barriers to people with disabilities.

Bühler et al. [3] proposed an approach in order to adapt measurement to different disabilities groups. A simplification of that model is the following:

$$A(u) = 1 - \prod (1 - R_b S_{ub})$$

Where R is the evaluation report and S is a severity value from 0 to 1 (for each barrier type b and user group u). This metric is

supposed to be integrated in the Web accessibility benchmarking framework defined in [8].

Finally, Vigo et al. [10] proposed a metric which is automatically calculated and faces the challenge of overcoming the drawbacks that the previous metrics contain. As an example, they cover all the aspects of Web accessibility, even if some of them are focused on specific user groups or limited to some guidelines. In fact, the authors tried to take advantage over the other approaches by adopting the good points of those metrics. In particular, they identified a set of facts, requirements and assumptions and defined a consequently algorithm to calculate a quantitative value. For instance, such a value should be normalized, should reflect the priority of an unfulfilled checkpoint, should consider in the same way errors and warnings (intended as manual checking). This algorithm takes into account also WCAG 2.0 principles and guidelines [12].

Since such formulas have been defined by international research groups, none of them is devoted to evaluate Italian law requirements, but they consider only international guidelines, (WCAG 1.0 and WCAG 2.0). Our validator and monitor are designed in order to evaluate Web sites according to our national requirements, by considering automatically detectable errors, warnings and suggestions of manual controls.

3. DESIGN ISSUES

Vamolà, which is the Italian acronym for “accessibility validator and monitor” has been designed as an integrated couple of applications. Users interested in validating and monitoring accessibility are typically public institutions, engaged in offering Web contents and services.

The validation tool is currently in the implementation phase and it will allow the automatic and semi-automatic accessibility evaluation of Web sites, according to the Italian Law [5]. This latter is attached by an evaluation procedure, organized into 22 requisites to be satisfied. Evaluation on Vamolà has been gained so as to let each single requirement to be faced. The validator is implemented as a three-version application in cooperation with the AChecker development team [2] from ATRC [1]:

- 1) a Web-centered application to evaluate a Web page, by submitting its URL or uploading its .html file; a prototype is under developing and evaluating phase;
- 2) a Web service (with the REST protocol) to be asked for by any application;
- 3) a module for the Apache Web server, to automatically generate reports for every requested Web page.

On its first form, the validator prototype lets the users to select the requirements they are going to evaluate on an URL or a file they can upload. On its second version the validator will be queried by means of an XML document describing the requirements to check out and their Web target. The third form of the validator will let Apache generate reports on every requested Web page, depending on a set of configuration parameters (typically, the law requirements to be controlled, the caching rules to avoid redundancy, the recipients of the reports). Apache module for the validator will provide its result through suitable emails, RSS feed or Web pages.

The Italian law allows the old sites not to be strictly XHTML 1.0 (but they have to avoid presentational tags and attributes, they have to let presentational and layout control to CSS files and to totally separate content and presentational aspects), while the new

ones have to be so. Hence, the prototype let the users validate a new or an old Web page, by setting up the appropriate controls. The difference between new and old Web pages is formally stated by the Italian Law [5].

Four classes of messages have been defined about:

- 0) errors automatically found by the validator (e.g., an `` without the `alt` attribute);
- 1) potential errors to be manually controlled (e.g., the `alt` value has to correspond to the actual content of the image);
- 2) potential errors to be found with heuristics (e.g., the `alt` value is parsed in order to look for a pre-defined set of string, for instance all the extension of image formats and words, such as "picture", "photo", "image"; all of these may represent a fake textual alternative; if one of them is found, the validator returns an alert message);
- 3) errors indirectly referring to the requirements or generally bound to the accessibility of the target.

The accent placed on messages is about errors and the way they can be corrected; users must always do controls (neither "should", nor "could"). Such a choice has been taken depending on the mandatory form of the Italian Law.

Finally, every message can be hidden on the report the validator provides, so as to avoid excessive verbosity (above all on 1, 2 and 3 types of messages), once a validation is repeated after corrections.

The monitor is a Web-centered application: it will allow inserting into a database some characteristics of Web sites such as their referring institution with further information about their geographical location and administrative role. Authorized users can choose the time period, the set of requirements to be evaluated and the subset of inserted Web sites to be evaluated. The monitor will call the validator according to such parameters and will provide some in-time reports about the "current state" of them, their improvement and their dynamics. Thresholds on file-size differences have been chosen to state an evaluated Web page is the same as that one of the previous control (so as to avoid useless checks). Topology of evaluated sites is shown on an SVG map or on a table, according to the Italian regulations, into communes or municipalities, provinces and regions. The depth of evaluation, through the links of a target page, can be chosen as a configuration parameter too.

Reports about accessibility "snapshots" represent a base of knowledge to provide measures of accessibility. Type and frequency of errors point out the attention on peculiar aspects of accessibility. Accumulation of shortcomings, depending on the depth of an URL analysis can be taken into account to synthesize sites accessibility attention. The amount of changes along periodical controls can be, finally, an indicator of accessibility care.

3.1 Accessibility metrics on Vamolà monitor

The monitor of Vamolà is designed to provide some measures of accessibility, on the strength of some metrics that its administrators and users are able to set up. At their design phase, metrics definitions come from considerations about necessity to compare the accessibility among sites and evaluate them on an absolute scale, their dynamics in time and on the territory, and according to the users' expectations and preferences. A suitable

override system is planned to be implemented on Vamolà monitor, so as to allow the substitution or enhancement of applied metrics.

Since a metrics is a correspondence between tuples of items and positive or zero numbers, the first step toward a measure of accessibility, has to be establishing the weight (as a sort of referring importance number) of each item (which, generally, is not a number). Counting errors of each class, is a first, rough way to do such a weighting, which represents itself a metrics (and so, a measure) of accessibility. Referring to the above list, no doubt, the 0-type have to be considered as more relevant for accessibility, in comparison with the 1, 2, or 3-type which we can conceive only as potential mistakes. Multiplicative constants (e.g., on the [0,1] interval), applied on each amount of errors (gathered per class) express such an aspect on the sum of the four obtained numbers.

Moreover, clustering the Law requirements depending on a single aspect of disability they face to and, consequently, the messages they provide in case of error, represents a further way toward a measure of accessibility. Each cluster is a sub-tuple to be dealt with (by weighting its items), so as to provide a measure. As an example, let us consider all the messages about sight disabilities, from colour contrast to textual alternatives. They can be counted per type and string-pattern and weighted with some multiplicative constant to concur for a measure of sight-accessibility. Other criteria of clustering, such as the institution domain (geographically or administratively), the number of sites each institution masters, the size of the site (considered as the number of pages) or, finally, the amount of services provided (as forms, multimedia object, etc) have to be used to evaluate some measures and to be arranged onto a metric.

The final version of the monitor of Vamolà will let its administrators choose the measures to be shown and the parameters of each metrics (multipliers, sub-tuples, etc).

Parameters to be used in metrics could also be chosen from the expectations and preferences expressed by communities of authorized users. In a collaborative and shared way such communities could assign different weight and consider some particular aspects as prior or secondary and hence defining different metrics. Users' communities could be invited to submit questionnaires in which they state how much important they consider some aspects of accessibility in comparison with some other ones. Hence, such a sort of "majority report" could be synthesized as a set of parameters to build one or more customized metrics.

Let us consider each metrics as referring to one dimension of a sort of accessibility space. The differences (toward better or worst accessibility) can be evaluated as a single dimension segment length. The distance between two points of the accessibility space can, instead, represent a measure of the dynamics of accessibility in time. Indeed, the vector, together with such a distance exhaustively describes the sites dynamics. On Vamolà, an Euclidean metrics has to be chosen for the module of such a vector, while measures of each dimension will be derived from previous describe metrics and related considerations. Distance vector will be shown as a row of a table, where each column is the considered dimension.

A prototype of the monitor is already developed and Figure 1 depicts a screenshot of its interface (in Italian language). Vamolà has been used with about 800 web pages representing a number of institutions all over the Italian territory. A one-week interval control has been launched (for every requirement) on the sites, in

order to gather some preliminary results to analyzed and to test the prototype.

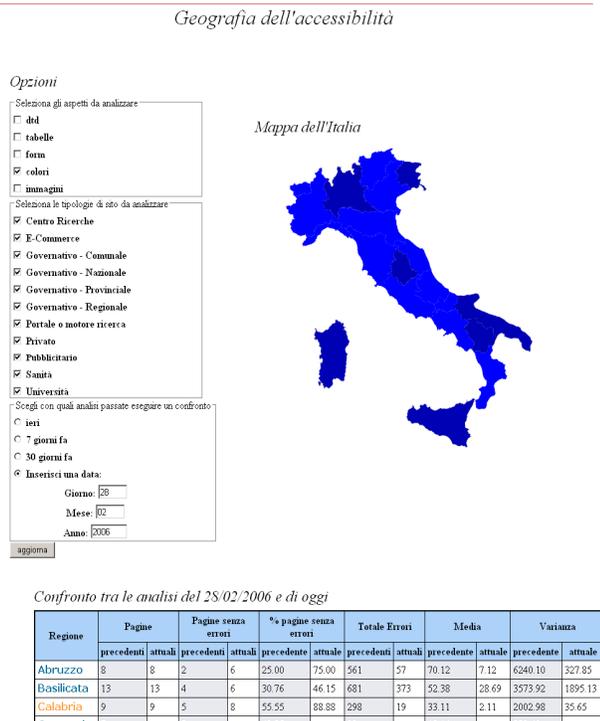


Figure 1. A screenshot of the monitor prototype interface (in Italian language).

The prototype interface is still under the design phase. A study is conducting in order to define adequate monitor and validator interfaces. In particular, the latter one should support end users in taking suitable actions to address the identified barriers.

4. CONCLUSIONS AND FUTURE WORKS

Whenever regulations can be backed up by deterministic constraints, such as measures or rigorous procedures to asses, the law is safer from elusion or violations. Vamolà has been designed to get together the Italian requirements about Web accessibility within an all-in-one application and to provide a synthesis of their application as a screening tool. Its aim is helping public institutions to face and (hopefully) to solve, accessibility issues and shortcomings. Metrics and measures on Vamolà monitor are thought to represent a further synthesis of results and evolution of Web sites in time. Besides some metrics, known in literature, other ones have been designed and have to be implemented. For all the parameters will be customizable, according to the expectations and preferences by a community of experts. At the present time, a prototype of the Vamolà monitor is being used to gauge all the parameters appearing on measures. Besides the

analysis of results from the prototype of the monitor, future work will deal with the developing phase and providing suitable, ad hoc guidelines, based on shortcomings clustering and measures values. Overridden enhancement and new metrics are also expected from a future community of developers on the Vamolà monitor.

5. ACKNOWLEDGEMENTS

The authors would like to thank Giovanni Grazia and Jacopo Deyla (from the Emilia Romagna Region), Marina Vriza and Ennio Paiella (from ASPHI) and Gregory R. Gay (from ATRC, Faculty of Information, University of Toronto).

6. REFERENCES

- [1] Adaptive Technology Resource Centre (ATRC), Faculty of Information, University of Toronto. Available from: <http://atrc.utoronto.ca/index.php>, 2009.
- [2] ATRC, AChecker. Available from: <http://www.atutor.ca/achecker/index.php>, 2009.
- [3] Bühler, C., Heck, H., Perlick, O., Nietzio, A., and Ullveit-Moe, N. Interpreting Results from Large Scale Automatic Evaluation of Web Accessibility. In K. Miesenberger et. al (Eds.). *Computers Helping People with Special Needs 2006*. Lecture Notes in Computer Science 4061. Springer, 184-191.
- [4] Hackett, S., Parmanto, B., and Zeng, X. Accessibility of Internet websites through time. In *Proceedings of 6th International ACM SIGACCESS Conference on Computers and Accessibility 2004*, pp. 32-39.
- [5] Italian parliament. Law nr. 4 – 01/09/2004. Official Journal nr. 13 – 01/17/2004, January 2004.
- [6] Parmanto, B. and Zeng, X. Metric for Web Accessibility Evaluation. *Journal of the American Society for Information Science and Technology*, 56(13):1394–1404, 2005.
- [7] U.S. Rehabilitation Act Amendments. Section 508. Available from: <http://www.webaim.org/standards/508/checklist>, 1998.
- [8] Snaprud, M.H., Ullveit-Moe, N., Pillai, A.B., and Olsen M.G. A Proposed Architecture for Large Scale Web Accessibility Assessment. In K. Miesenberger et al. (Eds.). *Computers Helping People with Special Needs 2006*. Lecture Notes in Computer Science 4061, pp. 234-241, Springer.
- [9] Sullivan, T. and Matson, R. Barriers to use: usability and content accessibility on the Web's most popular sites. In *Proceedings of ACM Conference on Universal Usability 2000*, pp. 139-144.
- [10] Vigo, M., Arrue, M., Brajnik, G., Lomuscio R. and Abascal, J. Quantitative Metrics for Measuring Web Accessibility. In *Proceedings of W4A2007* (Banff, Alberta, Canada, May 7-8, 2007) ACM Press, New York, NY, 2007, 99-107.
- [11] World Wide Web Consortium. Web Content Accessibility Guidelines (WCAG) 1.0. Available from: <http://www.w3.org/TR/WCAG10/>, 1999.
- [12] World Wide Web Consortium. Web Content Accessibility Guidelines (WCAG) 2.0. Available from: <http://www.w3.org/TR/WCAG20/>, 2008.