EUTIC 2007

Enjeux et Usages des Technologies de l'Information et de la Communication

Médias et diffusion de l'information: vers une société ouverte

Colloque International 7, 8, 9 et 10 novembre 2007

Tome 1

Sous la direction de
Prof. Michel Meimaris et de
Dr. Dimitris Gouscos
Université d'Athènes, Laboratoire des Nouvelles Technologies

Challenges and Uses of Information and Communication Technologies

Media and information diffusion: towards an open society

International Colloquium 7, 8, 9 and 10 November 2007

Volume 1

Edited by
Prof. Michael Meimaris and
Dr. Dimitris Gouscos
University of Athens, New Technologies Laboratory

GUTENBERG - ATHENS

REFLECTIONS ON THE EVALUATION OF WEBSITES

Andreas P. Giannakoulopoulos,
New Technologies Laboratory, University of Athens
agsp96@media.uoa.gr

Abstract

The continually increasing rate of website development and the various different techniques employed with respect to web construction technologies have set the preconditions for the discussion which concerns the methods of measurement, evaluation and –under specific restrictions– certification of web-based applications. In this paper, certain questions are examined and theoretical reflections are exposed referring to the necessity of website evaluation and the appropriate methods and tools for such a task. In parallel, substantiated solutions to particular theoretical problems are sought. These problems lay within the core of the relevant discussion, the most prominent of which being the contradistinction between purely automated evaluation methods and the manual (subjective) ones.

The discussion is centered on the technological evaluation, in conjunction with the ergonomics of the constructions and the anticipated communicative results. The issues of the methods, the tools, the shape and quantitative expression of results and, sequentially, the matter of comparability among evaluated websites comprise both the main (technological) parameters and a proscriptive set of rules for the subject under discussion. Finally, the issue of website certification is raised, along with the accompanying exploration of the appropriate certifying institutions and the systems of evaluation.

Résumé

Le taux croissant de développement des sites Web et les diverses techniques utilisées a l' égard des technologies de construction web ont établie les conditions préalables pour cette discussion qui porte sur les méthodes de mesure, d'évaluation et – sous certaines restrictions spécifiques - de certification des applications web-basées. Dans ce contexte, nous allons examiner certaines questions et présenter des réflexions théoriques sur la nécessité d'évaluation des sites Web et sur les méthodes et outils appropriées dans ce but. Parallèlement, nous allons rechercher des solutions justifiées à des problèmes théoriques particuliers, dont le plus important est la distinction entre les méthodes d'évaluation purement automatisées et les manuelles (subjectives).

Plus precisement, nous allons nous centrer sur l'évaluation technologique, conjointement avec l'ergonomie des constructions et les résultats communicatifs prévus. Les questions des méthodes, des outils, de la forme et de l'expression quantitative des résultats et, par conséquent, de la question de la comparabilité parmi des sites Web évalués comportent les paramètres (technologiques) principaux et un ensemble des règles proscriptives liées au sujet de cette discussion. En conclusion, nous allons traiter la question de la certification des sites Web et identifier les établissements de certification et les systèmes de l'évaluation appropriés.

Keywords

usability, accessibility, web standards, evaluation methods, interface design

1. Introduction

Both the continually increasing rate of website development and the recent advances in the various different web construction technologies have set the preconditions for the discussion which concerns the methods of measurement, evaluation and –under specific restrictions– certification of web-based applications. This paper mainly examines the necessity, the methods, the tools and the objectives of website evaluation, from a theoretic and a technical perspective.

The main axes of the discussion have a follows: first the necessity and the rationale behind website evaluation is examined in order to set the basis for the following sections which concern the human participants of the overall task, namely the evaluators and beneficiaries. The next axis focuses on the technical aspects of the topic and in particular the methods and tools employed for the evaluation, which in turn leads to the fields of evaluation, which are organized under broad categories. Finally, through generalization, certain conclusions are drawn, including also a few basic notes on further steps to be followed in evaluation research, in accordance with the main scope of the discussion which is to examine in theoretical level the role and benefits of website evaluation towards a user-friendly and highly usable web.

2. Rationale, Beneficiaries and Evaluators

Primarily, the fundamental question with respect to website evaluation is "what do we need it for"? What are the benefits to the community of users and how do we gain them?

The actual phrasing of the above questions reveals the directions in which the answers have to be sought: the rationale behind the quest for website evaluation lies in the benefits that may be derived from such a task.

It is evidenced and acclaimed that website evaluation increases the acceptance and implementation of web standards, especially among the community of web developers (Zeldman, 2003). Compliance with the widely accepted web development standards, increases the operability, usability and accessibility of websites, which, in turn, leads to interoperable and device independent web constructions (Berners-Lee et al., 2006). Creating device independent web applications is a long awaited silent request of the community of users and is increasingly important in a world of mobile communications (Meimaris, 2007).

The same is true for content production which overcomes the limitations of particular displaying software packages and especially media players and browsers. Providing content that is highly accessible and usable further enhances the usefulness of the web and promotes communication and cooperation (Nielsen & Loranger, 2006). Website evaluation will support –by means of results and guidelines– content providers in order for their products to meet user-friendliness standards and therefore will be for the benefit of the community of users.

Finally, evaluated and accredited websites may serve as solutions-by-example to other developers, thus leading to increased competition and, in turn, state-of-the-art creations. It is well-known that the hypertext code is available and viewable via all browsers and web developers –as suggested even in textbooks– consult published solutions to particular design and development problems trying to improve their constructions (Berners-Lee & Fischetti, 1999). Good practices, characterized as such through evaluation, will spread among the community of developers resulting in a new circle of better designed websites, which will meet higher evaluation criteria and so on, in a evolutionary series of better techniques and practical solutions.

Secondly, provided that the necessity of website evaluation is well supported, the interrelated issues of the beneficiaries and evaluators need to be elaborated. We have to ask who is going to be the evaluator and for whom this evaluation takes place; who is capable and appropriate for the task of evaluation and who is interested in the results?

There are three main groups of people who benefit from the evaluation of websites. The first one is the broad category of end-users which actually contains each and everyone with access to the web, regardless of the level of technological literacy or experience (Giannakoulopoulos, 2007). The other two include website owners and web developers.

Website owners will benefit by being able to compare the quality of their websites related to the cost, as an indicator of return of investment. They will also be in the position to demand standard solutions for their web communication, which will, most probably, result in high performance websites.

Web developers, in my opinion, will benefit the most from the evaluation of their work. Shared techniques will emerge for known problems and their

motivation to build better sites will increase as a result of their will to be positively evaluated.

Hopefully, the final outcome of both the above processes will be user-friendly, top-quality websites, which obviously is for the benefit of the whole community of web users.

From this point onwards, the next issue regards the evaluation teams and in particular the academic, scientific and professional characteristics of the individuals and/or institutions that will undertake the role of evaluators and by which criteria are to be selected as appropriate and effective. There are two main categories of institutions capable for fulfilling the task and providing reliable results. Both of these categories share some key characteristics which offer them the required credibility: they are more or less independent from economic and political power and they are widely acclaimed in society.

The evaluating institutions may be independent ones, like international consortia, or scientific and academic ones, like university departments and research centers.In the former case, artists, designers and web developers, as well as experienced and non-experienced users, may take part in the process of evaluation. In the case of academic institutions the evaluating boards may consist of computer scientists, interface and usability experts, mass-media practitioners with academic experience and communication scientists.

In summary, up to that point and according to argumentation, website evaluation is consider to be desirable and useful, and it openly benefits website owners, web developers and the community of web users at large. It is also suggested that website evaluation be carried out by credible and independent from economic and political power organizations, preferably international consortia or academic institutions.

3. Methods and Tools

In practical level, the most important aspect of website evaluation is related to the available methods and tools. With respect to the methodology, the international discussion is already very active and is mainly focused on the contradistinction between the purely automated and the manual methods of evaluation. Furthermore, an interesting and unresolved yet issue concerns the fields of investigation: the actual aspects of both the developing processes and the final web product which have to be evaluated, especially with respect to usability and accessibility issues (Mueller, 2003).

The aforementioned contradistinction between the automated and the manual methods of evaluation is a key concept with respect to website evaluation due to the fact that it reflects a basic conceptual difference in the way the subject under discussion is approached: the more general diversity of objective and subjective criteria, the former being related to the automated methods, whereas the latter being related to the manual ones.

There are different tools for different aspects of website evaluation. Some of them, and especially the validators, are already in wide use. There is also an increasing interest in automatic accessibility assessment and certain tools have been developed for that task (Slatin & Rush, 2003). In general, the tools for automated web-based evaluation include markup and stylesheet validators, traffic and webpage structure analyzers, search engine optimizers and accessibility checkers.

The main objective regarding the software of the automated evaluation category is to integrate as many as possible of the above-mentioned tools in one-stop unified platforms.

Manual evaluation, on the other hand, is conducted by a board of evaluators, who independently or in collaboration evaluate websites according to their knowledge, expertise and experience. Therefore, both end-users and experts should be members of such committees. In that case, evaluation takes place in real-case testing environments and is site-specific and detailed (Krug, 2000). The fields of analysis include interface design, mainly employed by the so-called usability tests, content evaluation by methods of qualitative content analysis and aesthetics, which is commonly mentioned as the most intriguing and difficult field for evaluation, not only in the case of websites but also in almost every piece of work which involves artwork (Tidwell, 2005).

It is quite difficult, due to the strength of the arguments in favor of each category of techniques, to decide and suggest a certain methodology and approach for all cases. As a contribution to the discussion, the following conclusions for each method are put forward as topics for further investigation:

The automated evaluation method is objective and provides directly measurable results, but is inappropriate for content and artwork analysis which constitute a significant portion of the final product.

Accordingly the method of manual evaluation is detailed, site-specific and can provide deep and informative analysis, but is obviously prone to subjectivism which in the particular case under discussion is considered undesirable.

By trying to combine these methods, one may gain the best parts of both methods, but there is a critical point which may result in a major drawback: in order for the evaluation to be to the benefit of the community, the results must be comparable (Giannakoulopoulos, 2006). In the case of combined evaluation methodology, this is not an easy target to achieve.

4. Fields of Evaluation

With respect to the fields of evaluation, it will be useful to determine the particular web development areas which are appropriate for evaluation, in the sense that they consist autonomous entities with special features each, which allow for elaborate assessment. Getting into more details, we may organize the fields of evaluation into the technology related fields and the content and art ones.

The technological fields include the following areas of specialization:

- Web usability, concerning the interface and the user-friendliness of the websites (Nielsen and Loranger, 2006)
- Web accessibility, concerning the degree to which content is accessible by users with disabilities and the older population which is constantly increasing in numbers (World Wide Web consortium, 2007)
- Information architecture, regarding the structure and findability of information available via a particular website, with special emphasis on navigation and labeling (Wodtke, 2002)
- Web standards compliance, which is a purely technological field and is related to the validity of the website in terms of formal structure and presentation languages (Zeldman, 2003).

The content and art section includes all aspects of website development which may be labeled as "look, feel and content". It is quite difficult even to define this category and enumerate its items, let alone to evaluate it. But for the sake of categorization patterns it may be organized in the following four areas:

- The quality, richness and originality of content, including the use of language and the effective exploitation of hypertextuality
- The quality of aesthetics, artwork and design, and especially the usage of modern and less common features as parts of the interface design
- The effectiveness of various communicational issues, such as the frequency of updates, the effective use of multimedia and the successful employment of interactive communicational techniques
- Innovation and novelty which again is an obscure and undefined category, at least as the criteria to be employed for the assessment are concerned.

The above-mentioned categories form the basic areas in which website evaluation may take place. The list is obviously not exclusive, but rather indicates the fields in which evaluation may be more accurate and less biased, regardless of the specific methodology to be used.

5. Conclusions

In a field of research such as the one under discussion, in which on the one hand the interest is continuously growing and on the other there is little experience yet, it is normally expected that there is a wide area of further research topics, which need to be explored in order for stable conclusions to be drawn. In brief, the following subtopics are considered to be of high priority with respect to further research within the field of website evaluation:

First we need to reach a high level of consensus as far as the methodology that will be used for the task is concerned, including the tools, the scales and the

fields of evaluation, especially if measurable and comparable results are sought and expected.

We also need to develop one-stop platforms by the integration of automated tools and manual checks. Models of analysis as well as templates for final reports are also part of the integration process, which at last will overcome the problem of different methods and techniques currently used in website evaluation and resulting in the fragmentation of the field.

Finally, the evaluating institutions may provide quality assurance and conformation marks, much like the validity icons shown in figure 1, which are accredited by the World Wide Web consortium to valid markup and stylesheet code.

Fig. 1: The validity icons of the W3C for valid markup and stylesheets

Epigrammatically, the outcome of the theoretic argumentation supports the further exploitation of the field of website evaluation. In particular, the major conclusions, which may be drawn by generalization from the previous discussion, are the following:

- Website evaluation benefits web development as a process and especially website owners, content providers, web authors, developers, designers and, overall, the users of the web.
- The evaluating institutions must be independent, reputable, widely acclaimed in international level, available to undertake the task and willing to improve the field in terms of methodology and liability.
- It is of great importance that the evaluation methodology combines both automated tools and manual techniques in order for all aspects of web creation to be equally represented in the results of evaluation.
- The fields of investigation should cover the entirety of the development process and the final product, including all the technology related fields, the “look, feel and content” area, structural and navigational elements, the web interface design and the overall operability and usability of each website to be evaluated.

The main objective of the process of website evaluation is to obtain comparable results and that aim may be effectively reached only by employing standard, widely acceptable methods. The final outcome and the net profit for the web community at large will be a better, more accessible, user-friendlier and highly usable web.

Bibliographic References

1. Berners-Lee, T. & Fischetti, M. (1999) *Weaving the web:the original design and ultimate destiny of the World Wide Web by its inventor*. Harper Collins, New York.
2. Berners-Lee, T., Hall, W., Hendler, J.A., O'Hara, K., Shadbolt, N. & Weitzner D. (2006) *A framework for web science*. Now publishers, Hanover, MA.
3. Giannakoulopoulos, A.P. (2006) "Threefold web accessibility evaluation by the use of an integrated tool", in Freddy Malpica, Angel Oropeza, José V. Carrasquero & Paul Howell (eds.) *Proceedings of the 2nd International Conference on Social and Organizational Informatics and Cybernetics*, SOIC '06, July 20-23, 2006 – Orlando, Florida, USA (v. II, pp. 243-248).
4. Giannakoulopoulos, A.P. (2007) 'Web standards and consensus: ensuring the universality of Internet applications', *Communication Issues*, 5, pp. 33-45.
5. Krug, S. (2000): *Don't make me think: A common sense approach to web usability*. New Riders Publishing, Berkeley.
6. Meimaris, M. (2007) 'Communicational dimensions of the new technological environment', *Communication Issues*, 5, pp. 27-34.
7. Mueller, J.P. (2003) *Accessibility for everybody: understanding the Section 508 accessibility requirements.* Apress, Berkeley.
8. Nielsen, J. & Loranger, H. (2006) *Prioritizing web usability*. New Riders Publishing, Indianapolis.
9. Slatin, J. & Rush, S. (2003) *Maximum accessibility*. Addison-Wesley, Boston.
10. Tidwell, J. (2005) *Designing interfaces.* O'Reilly, Sebastopol.
11. Wodtke, C. (2002) *Information architecture: blueprints for the web*. New Riders Publishing, Indianapolis.
12. World Wide Web Consortium (2007) *Web accessibility initiative* [Internet]. Available from: <http://www.w3.org/WAI/> [Accessed 09 September 2007].
13. Zeldman, J. (2003) *Designing with web standards*. New Riders Publishing, Indianapolis.