WebA: A Tool for the Assistance in Design and Evaluation of Websites

Luis Mena Tobar
(Laboratorio Aragonés de Usabilidad
Parque Tecnológico Walqa, Edificio 1 E-22197 Cuarte, Huesca, Spain
luis@menasl.com)

Pedro M. Latorre Andrés
(Departamento de Informática e Ingeniería de Sistemas
Centro Politécnico Superior
C/ María de Luna 1 E-50018 Zaragoza, Spain
platorre@unizar.es)

Elena Lafuente Lapena
(Laboratorio Aragonés de Usabilidad
Parque Tecnológico Walqa, Edificio 1 E-22197 Cuarte, Huesca, Spain
elafuente@laboratoriousabilidad.com)

Abstract: The development, analysis and follow-up of the processes designed to assure the usability and accessibility of websites is a tedious work for the moderator or evaluator. This is why there is a necessity for tools that can automate the processes. There are some tools that cover some of the technical aspects, but no tools have been identified that can tackle the process as a whole. For these reasons the Aragonese Usability Laboratory decided to develop WebA (Web Analysis), with the objective to have a complete application. This application is designed through modules with the objective of covering all of the evaluation and analysis phases, and it concludes with a process management module. For the analysis phase a Card Sorting module (open & closed) has been developed that uses the hierarchical and multidimensional cluster analysis, which allows a better information architecture. For the evaluation phases, modules have been developed that allow the semiautomatic evaluation of usability through user satisfaction tests based on Nielsen heuristics and ISO standards, the evaluation of accessibility through the verification of WCAG 1.0 guidelines. The application concludes with a process design and evaluation management module, and modules that automatically generate the reports of the analysis carried out.

Keywords: Usability Engineering, accessibility of websites, website usability evaluation, human computer interaction
Categories: H.1.2, H.5.2, H.5.3

1   Introduction

The measurement of the technical quality of computer application interfaces, and also that of websites in particular (correction, usability, etc.), is an expensive process that is usually approached with limited depth. The necessity to assure this quality has created specific norms for its verification (which are framed in the techniques of

evaluation of interfaces/users. [Dix et al. 04, Nielsen & Molich 90]) as well as laboratories specialized in the design and evaluation of usability.

This paper consists of an introduction in which the functionality of WebA and the initiative and activities of the Aragonese Laboratory of Usability (AraLUs) are shown. Subsequently, the evaluation methods and the verification of usability and accessibility and the systems that assist in the structuring of the information and their evaluation are reviewed in the available literature. Next, each of the available WebA modules is described in detail. Finally, conclusions are formulated and future work is planned.

1.1 WebA in short

WebA (from Web Analysis) is a tool created by the AraLUs that has evolved from a former application called WebA (Tools for the assistance in design and evaluation of websites) that had only a reduced pair of functionalities. WebA is evolving and adding new tools, and at the present time it is an application that

1) eases the evaluation of usability through the compliance of the ISO 9241-11 to 9241-17 norms and

2) support managing of tests of users' satisfaction,

3) allows the evaluation of accessibility through the verification of WAI standards (Web Accessibility Initiative),

4) has support for application design through the application of Card Sorting techniques (aCaSo, included in WebA) and

5) adds integrated management information tools. The processes can be carried out in situ or at a remote location.

1.2 The Aragonese Laboratory of Usability

The Aragonese Laboratory of Usability (AraLUs) [LAU] is an initiative of the General Directorate of Information Society Technologies of the Government of Aragon. Their activities are developed in the facilities of the Advanced Laboratory on Juridical and Managerial Applications in the Society of the Information [LABJE] (Technological Park Walqa [WALQA], Huesca), in the Faculties of Law and Economics and Business Administration and in the Higher Engineering School of the University of Zaragoza.

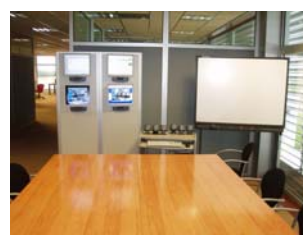*Figure 1: Observation room.*

Figure 2: Evaluation room.

The AraLUs develops projects of R&D in the field of evaluation of usability and accessibility of interfaces that are centred on the research of the consumer response to different stimuli and the application of different design techniques and evaluation of interfaces based as much on the user response as on the heuristic evaluation by experts. WebA is one of those R&D projects.

The Laboratory also develops activities with companies and institutions, helping to improve the quality of the product by means of introducing the verification of usability and accessibility in the development cycle of their computer applications and websites.

2   Methods for the Evaluation of Usability

Websites are the best interface for the integration of services by connecting people, forming the denominated social networks [Garton et al. 97]. The Internet has transformed into a key element, as much at a commercial level as at a social level. Hence, it is vital that the design of the sites is optimal from the point of view of the user: efficiency, effectiveness and satisfaction should be assured, and to achieve this depends to a great extent on the website's ease of use, or in other words, on its usability.

Different and varied methods of evaluating the usability of interfaces exist [Scholtz & Laskowski 98], classified according to different criteria. Considering Alan Dix's proposal [Dix et al. 04], a possible classification would be the following:

Evaluation through expert analysis
- Cognitive walkthrough
- Heuristic evaluation
- Review-based techniques
- Model-based evaluation

Evaluation through user participation
- Empirical Methods - Experimental evaluation
- Observational techniques
- Query and interview techniques
- Evaluation through monitoring physiological responses

From these methods, WebA has developed, by now, some Heuristic-based protocols (based on ISO norms) and query techniques (user satisfaction tests).

2.1   Automatic Evaluation of Usability

Unfortunately, not all the aspects of website usability can be evaluated in an automatic way [Balbo 95], since many of the points can only be validated and verified by expert professionals. Anyway, the automatic evaluation [Branjnik 00] of those points, where possible, based on the analysis of the website code (such as inadequate size of page, broken links, text format,...), involves a great help and facilitates the revision process vastly.

A diverse number of tools already exist in the market that can assist experts and automatically evaluate certain aspects of usability in a website. We will mention three examples:

- Morae [MORAE]: allows the management of usability tests.
- Noldus Observer [Noldus]: has some modules that give support for the user's observation and the execution of users' satisfaction test.
- W3C HTML Validator [W3C]: checks the validity of the HTML code

The great advantage of these tools is its speed, but its utility is limited since they do not detect the most important issue: the global questions of usability. Great part of these questions can only be detected by an expert human being, by means of techniques of heuristic evaluation.

2.2 Heuristic Evaluation

The methods of heuristic evaluation are based on the performance of a group of evaluators. These evaluators follow a group of heuristic rules to examine user's interfaces and judge their conformity according to the established principles.

There are several diverse aspects that should characterize a good user interface. As an example, it is necessary to mention the original list of ten rules developed by Nielsen and Mölich in 1990 [Nielsen & Molich 90] , as it is rewritten in Nielsen's website [UseIt]:

- Visibility of system status
- Match between system and the real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose, and recover from errors
- Help and documentation

This basic heuristics has been detailed and updated in many other publications, [Nielsen 94a, Nielsen 94b] but the basics remain the same. The heuristic evaluation detects most of the global questions of usability but it is time consuming and expensive.

2.3 Evaluation based on Standards

The design and evaluation of the interface can (and should) rely on guideliness and normalized standards, to assure the reliability of the process.

According to BSI [BSI], a standard is a published "specification that establishes a common language, and contains a technical specification or other precise criteria and is designed to be used consistently, as a rule, a guideline, or a definition". Standards are specifications of voluntary application, based on experience results and technological development. They are accepted by consent of interested parties.

Its application in the evaluation process consists in the tracking of each of the points that dictates the whole standard. This way it is made sure that each and every one of the outstanding aspects has been revised.

The evaluation is carried out by means of a team of experts in the topic (in this case usability) and in the standard to be applied. Each expert carries out a meticulous

inspection of the interface to check the compliance of each one of the points defined in the standard, and annotating the required observations. A main evaluator determines the team components and the work method, manages team answers and analyzes the results. Finally the definitive report is created.

The applicable standards that are directly related to the development of user interfaces are three: ISO 9241 [ISO 92], ISO 13407 [ISO 99] and ISO 14915 [ISO 02].

The standard ISO 9241 [ISO 92] "Ergonomic requirements for office work with visual display terminals (VDTs)" that is substituting the ISO 29241 of the same title, and has 17 parts which are the following.

1. General introduction
2. Guidance on task requirements
3. Visual display requirements
4. Keyboard requirements
5. Workstation layout and postural requirements
6. Guidance on the work environment
7. Requirements for display with reflections
8. Requirements for displayed colours
9. Requirements for non-keyboard input devices
10. Dialogue principles (Substituted by part 110)
11. Guidance on usability
12. Presentation of information
13. User guidance
14. Menu dialogues
15. Command dialogues
16. Direct manipulation dialogues
17. Form filling dialogues

Further standards in ISO 9241 are being added under the new title "Ergonomics of human-system interaction". These standards are now under development (most in a different final status of approval):

20. Accessibility guidelines for information/communication technology (ICT) equipment and services
110. Dialogue principles
151: Guidance on World Wide Web user interfaces
171: Guidance on software accessibility
300: Introduction to requirements and measurement techniques for electronic visual displays
302: Terminology for electronic visual displays
303: Requirements for electronic visual displays
304: User performance test methods
305: Optical laboratory test methods for electronic visual displays
306: Field assessment methods for electronic visual displays
307: Analysis and compliance test methods for electronic visual displays
308: Surface-conduction electron-emitter displays
400: Principles and requirements for physical input devices
410: Design criteria for physical input devices
920: Guidance on haptic and tactile interactions

The ISO 13407 standard [ISO 99] "Human-centred design processes for interactive systems" provides a general view of the design activities centred on the user, without going into technical details and methods. It is designed having in mind to be used by project directors and it is complementary with ISO 9241.

In this standard, besides indicating the general principles and planning the design process centred on the user, the activities to be carried out are detailed during the development of the project. These activities are grouped by four types:

- Understand and specify the context of use
- Specify user and organisational requirements and their context
- Produce design solutions
- Evaluate designs against user requirements

The ISO 14915 standard [ISO 02] "Software ergonomics for multimedia user interfaces" is composed of the following parts:

- Design principles and framework
- Multimedia navigation and control
- Media selection and combination

2.4 Evaluation of the Accessibility

Accessibility is "the group of characteristics that the environment, product or service should have to be usable in comfortable, secure and equal conditions for all persons, in particular, those that have some type of disability" [Plan].

The evaluation of the Accessibility is carried out based on the guidelines of Accessibility of Web content (WCAG 1.0) developed by the World Wide Web Consortium.

WCAG 1.0 is composed of 14 rules [WCAG] that constitute the general principles of accessible design. Each rule has one or more checkpoints that help to detect possible errors. Some checkpoints specify a priority level that may change under certain (indicated) conditions.

Each checkpoint has a priority level based on the checkpoint's impact on accessibility:

- Priority 1.- A Web content developer must satisfy this checkpoint. Otherwise, one or more groups will find it impossible to access information in the document. Satisfying this checkpoint is a basic requirement for some groups to be able to use Web documents.
- Priority 2.- A Web content developer should satisfy this checkpoint. Otherwise, one or more groups will find it difficult to access information in the document. Satisfying this checkpoint will remove significant barriers to accessing Web documents.
- Priority 3.- A Web content developer may address this checkpoint. Otherwise, one or more groups will find it somewhat difficult to access information in the document. Satisfying this checkpoint will improve access to Web documents.

Depending on the conformance of checkpoints, a website can achieve one of the three following levels:

- Conformance Level "A": all Priority 1 checkpoints are satisfied;

- Conformance Level "Double-A": all Priority 1 and 2 checkpoints are satisfied;
- Conformance Level "Triple-A": all Priority 1, 2, and 3 checkpoints are satisfied;

ISO is also developing the 9241:20 standard, previously cited.

3   The structuring of information: Techniques of Card Sorting

The technique of Card Sorting [USANET] is a method whose objective is to find the latent structure between a group of ideas or non-prearranged questions. When a website is designed, titles for menus have to be perfectly understood by users, so understanding and grouping of these ideas or titles have to be tested in advance.

To apply this technique titles are identified and listed, and each of them is written in a card, be them real or "virtual", (stored in a file), depending on whether the process will be automated or not.

Once the cards have been done, the Card Sorting session can begin. User involvement is required. Each participant receives a pile of cards and he/she organizes them in groups according to some criteria. Depending on the method these groups have been previously named or can be named by the users at that moment, and the session evaluator could take them into account for the design of the final solution.

Later on, the obtained data are analyzed by means of statistical algorithms providing an arrangement based on the structures proposed by the different users. This process can be automated.

The Aragonese Laboratory of Usability has developed and integrated the application aCaSo into WebA that allows the automation of Card Sorting sessions.

4   The group of applications WebA (Web Analysis)

As mentioned before, the WebA tool is formed by a group of applications and assistance tools for the design and evaluation of websites. WebA consists at this time of the following modules:

- Evaluation based on ISO 9241-11 to 17 Norms.
  - ISO 9241-11: Guidance on usability
  - ISO 9241-12: Presentation of information.
  - ISO 9241-13: User guidance.
  - ISO 9241-14: Menu dialogues.
  - ISO 9241-15: Command dialogues
  - ISO 9241-16: Direct manipulation dialogues.
  - ISO 9241-17: Form filling dialogues.
- Evaluation of the Accessibility:
  The guidelines of accessibility are verified to the content in Web 1.0.
- Test of user satisfaction.
  Through this test the evaluator can identify data on what the user opinion is on structure, content, browsing, etc. of a website that can help him/her in the final evaluation of a site. In the AraLUs these types of sessions are carried

out in the evaluation room and are recorded by cameras, and while they are carried out, they are supervised from an observation room. The group of evaluators can get these recordings to assess the test answers.

- Evaluation session management module.
  WebA has four types of users: Administrator, Moderator, Evaluator and User, to carry out all kind of allowable tests.
- Generation of reports
  WebA allows the generation of reports with the results of the compliance with standards and the guidelines of Accessibility. Also, it allows the generation of some statistical graphs of the results of the evaluations based on standards.
- Forum for Evaluators and moderator.
- aCaSo (Applications for Card Sorting).
  The application allows users to carry out and manage Card Sorting sessions, apply the most appropriate analysis techniques and visualize results.

4.1 The Process of Evaluation of Websites in Detail

The evaluation process of a website is based either on the involvement of a group of experts that work individually and later combine their results (heuristic-based, standard-based) or else in the opinions of a group of users. The administrator chooses the methodology to apply and opens a new database to manage the process for the website to verify.

ISO standards and accessibility guidelines are structured in groups of recommendations that should be verified one by one. WebA automates the verification of those points whose fulfilment can be deduced through the webpage code, and shows the expert the results inside a form, allowing him/her to modify them if he/she considers it necessary in some cases. The recommendations that cannot be evaluated automatically are valued directly by the expert, who introduces the conformity grade of the website in regards to each norm or guideline, in accordance to the evaluation of usability or accessibility, in the corresponding forms.

4.1.1 Heuristic Evaluation

WebA shows the corresponding forms and it outlines, for each item, a conformity scale from 1 to 5 (1 being the most unfavorable value) stating whether automatic valuation is possible as well as if the expert should verify it,

Figure 3: Conformity scale.

The tool shows the link to the page that is being evaluated and the results are stored in a database. It is the evaluator who determines whether it is or it is not applicable and in what level. In the event of being precise he/she can make observations or comments that he/she considers convenient.

Figure 4: Form filling

4.1.2   Evaluation based on Standards

Also in this case WebA automates, as much as possible, the evaluation of usability by means of standards compliance. For those points whose examination cannot be automated it generates forms so that the expert can evaluate them. As much in the ISO usability norms as in accessibility guidelines some recommendation groups exist that need to be verified one by one.

The evaluation based on standards is based on ISO 9241-11 to 17, and it will be updated as soon as new ISO 9241 parts are published. Next we will briefly describe each one.

- ISO9241-11: 1998 Part 11: Guidance on usability.

This part measures the usability of an interactive application or website following some general recommendations about the ease of interface use.

Figure 5: Window with recommendations ISO 9241-11.

- ISO9241-12: 1998 Part 12: Presentation of Information

This Norm is in charge of the visual presentation of information by means of displays. It includes (in chapter 4) the design objectives that provide a high level path on information presentation. Chapters 5 through 7 provide recommendations for the design of each window or page and they are of application, in general, to all the dialogue techniques.

- ISO9241-13: 1998 Part 13: User guidance

This part of ISO 9241 norm treats the relative aspects of the guide or orientation to the users of interfaces for computer programs. The main objective of these orientations to the user is to assist in the user's interaction with the system,

- favoring the efficient employment of the system;
- avoiding the load of unnecessary mental workload;
- providing the users a means of management of errors;
- providing assistance to users with different levels of competency

Figure 6: Window with the recommendations ISO 9241-13.

- ISO9241-14: 1997 Part 14: Menu dialogues.

This part refers to the ergonomic design of dialogues by the means of menus. This group of norms provide recommendations for menu structuring used in user-computer dialogue to carry out tasks in offices through different techniques such as windows, panels, buttons, fields, etc….

This norm cannot be applied without a knowledge of the context design and user interface.

- ISO9241-15: 1997 Part 15: Command dialogues.

ISO 9241-15 Norm looks at the ergonomic design of the dialogues of command language type. In the dialogues of this type, the users introduce, by heart, expressions of complete or abbreviated commands, respecting the syntax of the language of commands, while the computer executes the actions corresponding to these parameterized commands.

- ISO9241-16: 1999 Part 16: Direct manipulation dialogues

Part 16 refer to the ergonomic design of direct manipulation dialogues, in which the user carries out the operations acting on objects presented on the screen, in a way similar to the manipulation of physical entities.

In practice the term direct manipulation is used as an alternative to graphic user interface (GUI). Nevertheless, in GUI other dialogue techniques are put into practice, such as dialogues by means of menus or by means of orders.

- ISO9241-17: 1998 Part 17: Form filling dialogues.

This part treats the relative aspects of communication by the means of dialogs of the website through filling out forms.
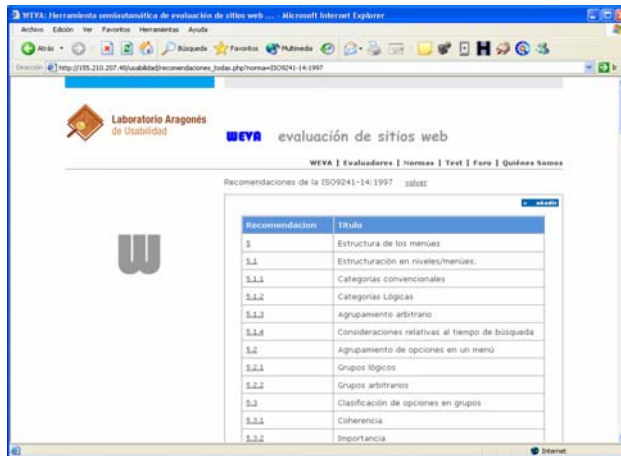
Figure 7: Window with the recommendations ISO 9241-14.

4.1.3 Evaluation of Accessibility

WebA evaluates in a semiautomatic way the accessibility of the websites according to the guidelines of content accessibility on the Web (WCAG 1.0) and it checks, or it allows checking, each verification point indicated by the guideline.

The tool carries out the verification according to the priorities established in the guidelines. This way it can determine the grade or level of conformity of the website that it is analyzing.
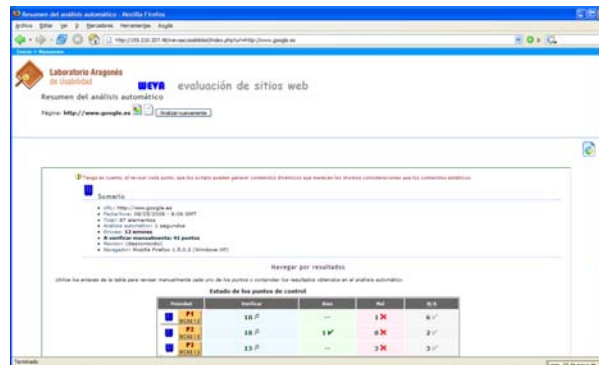

Figure 8: Window showing the summary of the analysis of the accessibility of a website.

The follow-up of guidelines during the evaluation process assures that all the guidelines have been revised, in an automatic way as well as by means of an expert. As always, the whole process cannot be automated due that some aspects of the guidelines exist that only an expert can verify.

4.1.4   aCaSo (Applications for Card Sorting)

The aCaSo tool, integrated in Weba, has as an objective to computerize Card Sorting sessions.

Figure 9: Management window for Card Sorting sessions.

This tool allows us:
- To acquire the data, to carry out the Card Sorting session and present the results fulfilling the requirements determined by the standards of Engineering Usability.
- To choose among open and closed clustering
- To use different analysis methods.
- To manage Card Sorting sessions.
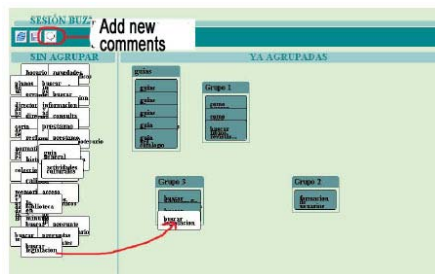- To carry out online evaluation sessions and analysis of results.

Figure 10: Grouping cards.   *Figure 11: Generating the dendrogram.*

4.1.5   User satisfaction tests

User satisfaction tests are carried out following the Nielsen general principles [UseIt] and the refined usability problems [Nielsen 94a]. Those tests are normally adapted to the specific application or website that is being analyzed.

4.1.6   Evaluation session management

The tool has the following levels of users:

ADMINISTRATOR

The Administrator is the user with the most privileges. He/she is in charge of :

- opening a new session or evaluation process
- adding / modifying / deleting moderators, evaluators and/or users.
- adding / modifying / deleting norms, recommendations and/or guidelines.
- administering the forum.
- adding / modifying / deleting evaluation groups.

Figure 12: Window showing the options for the Administrator.

MODERATOR

This user can create groups of evaluators and give evaluation guidelines, and to decide what norms or rules will be proven in the evaluation.

EVALUATOR

This user can enter and evaluate those websites for which he/she has been enabled (registered) as an evaluator, and generate reports and evaluation statistics.

Figure 13: Report of filling in the ISO 9241-11 norm.

USER

He/she can only enter the system to carry out the test of user's satisfaction on a website.

Figure 14: User´s satisfaction test window.

5   Conclusions and future work

Usability laboratories and the companies dedicated to development need tools that allow the realization and the management of the design process and evaluation, automating all those tasks that are possible.

As answer to this necessity, the Aragonese Laboratory of Usability has been developing a group of applications integrated in WebA.

WebA increases the reliability of the design process and evaluation of application interfaces and it reduces evaluation time, since it does not only facilitate the verification of the usability and accessibility of the site but rather gives assistance in the design phase of the interface to get an effective, efficient design and that satisfies the user.

In short, WebA allows the evaluation of usability by means of verification of ISO 9241-11 to ISO 9241-17 norms, guaranteeing the adjustment to these industrial standards and the evaluation of the accessibility based on the guidelines of accessibility to web content (WCAG).

- It integrates the necessary participation of experts with automated evaluation.
- It facilitates the communication among the experts; it reduces time and cost for the realization of the corresponding meetings.
- It allows the evaluator to consult the previous evaluations, reducing work time.
- It automatically generates the report to present to the client.

The Laboratory continues to broaden the tool with new modules (log analysis, online tests of user response) [WebTrends, LogAnalysis]. It is also sought to generate a module for the validation of applications in mobile (WebA mobile) devices.

Acknowledgements

References

[Balbo 95] Balbo, S. (1995) Automatic evaluation of user interface usability: Dream or reality. In S. Balbo, Ed., Proceedings of the Queensland Computer-Human Interaction Symposium (Queensland, Australia, August). Bond University.

[Branjnik 00] Brajnik, G. (2000) Automatic web usability evaluation: Where is the limit? In Proceedings of the Sixth Conference on Human Factors & the Web (Auxtin, TX, June)

[BSI] BSI Bristish Standards group(http://www.bsi-global.com/en)

[Dix et al. 04]Dix, A., Finlay, J., Abowd, G., Beale, R. (2004) "Human-Computer Interaction". 3thd Edition,. Pearson Prentice Hall. ISBN 0130-461091

[Garton et al. 97] Garton, L., Haythornthwaite, C., Wellman, B. (1997). "Studying on line Socialnetworks", Journal of Computer Mediated Communication. Available and re-reviewed on April 2002.

[ISO 92] ISO 9241 Ergonomic requirements for office work with visual display terminals (VDTs). International Organization for Standardization, 1992-98

[ISO 99] ISO 13407 Human-centred design processes for interactive systems. International Organization for Standardization, 1999.

[ISO 02] ISO 14915 Software ergonomics for multimedia user interfaces. International Organization for Standardization, 2002-03.

[LABJE] Laboratorio Avanzado sobre Aplicaciones Jurídicas y Empresariales en la Sociedad de la Información (http://labje.unizar.es)

[LAU] Laboratorio Aragonés de Usabilidad (http://www.laboratoriousabilidad.org).

[LogAnalysis] Log Analysis (http://www.clicktracks.com)

[MORAE] Morae (http://www.morae.com).

[Nielsen & Molich 90] Nielsen, J., Molich, (1990) Heuristic evaluation of users interfaces. , Proc. ACM CHI'90 Conf. (Seattle, WA, 1-5 April), 249-256

[Nielsen 94a] Nielsen, J. (1994) Enhancing the explanatory power of usability heuristics. Proc. ACM CHI'94 Conf. (Boston, MA, April 24-28), 152-158.

[Nielsen 94b] Nielsen, J. (1994) Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, NY.

[Noldus] Noldus Observer (http://www.noldus.com).

[Plan] Spanish National Plan of Accessibility, 2004-2012.
http://usuarios.discapnet.es/disweb2000/lex/AccePlan2004-2012.pdf

[Scholtz & Laskowski 98] Scholtz J. and Laskowski, S. (1998) Developing usability tools and techniques for designing and testing web sites. In Proceedings of the Fourth Conference on Human Factors & the Web (Basking Ridge, NJ, June). Available at
http://www.research.att.com/conf/hfweb/proceedings/scholtz/index.html.

[USANET] http://www.usabilitynet.org/tools/cardsorting.htm

[WALQA] Parque Tecnológico Walqa (http://www.ptwalqa.com).

[WCAG] Web Content Accessibility Guidelines 1.0
(http://www.discapnet.es/web_accesible/wcag10/WAI-WEBCONTENT-19990505_es.html)

[W3C] W3C MarkUp Validation Service (http://validator.w3.org/).

[WebTrends] WebTrends Analytics (http://www.webtrends.com). [UseIt] Ten Usability Heuristics by Jakob Nielsen
(http://www.useit.com/papers/heuristic/heuristic_list.html)