



Accessible and Usable Web Content

Iker Hernandez Garcia

Supervisor: Morten Goodwin Olsen

Project report for Web mining in Autumn 2006

Adger University College
Faculty of Engineering and Science

Grimstad, November 25, 2006

Status: Final

Keywords: Wc3, Accessibility, Usability, WCAG

Abstract

We search for the connection between usability and accessibility. This project tries to find out how the adoption of w3c recommendations about the correct way to make web content accessible to people with disabilities can affect people with no disabilities. This information will be useful to obtain conclusions about if it's possible to get a friendly use webpage for both kinds of people at the same time. If we can find out this relationship we should have a powerful tool in the future for improve the quality of web content.

Contents

1	Introduction	3
1.1	Acknowledgements	3
2	Problem Description	4
3	Background	5
3.1	Literature Review	5
3.1.1	The WCAG recommendations	5
3.2	Choosing the users for the test	6
3.2.1	Requirements	6
3.2.2	Our Solution	7
3.3	Choosing the web pages for the test	7
3.3.1	Usability and Accessibility testing tools	8
3.4	The Questionnaire	8
4	Solution	9
4.1	First Stage	10
4.2	Second Stage	10
4.2.1	Rating Web Pages	10
4.2.2	User Testing Group	10
4.2.3	User Testing's Results	12
4.3	Third Stage	12
5	Discussion	14
5.1	The Confidence in the User Testing	14
5.2	The Confidence in the Automated Accessibility Rating	15
5.3	The Relationship is not Direct or Linear	15
6	Conclusion	17
7	References	18

List of Figures

1	Schema of the experiment	9
2	Age of Users	10
3	Country Of Users	11
4	Sex Of Users	11
5	Study Level Of Users	12
6	User Testing Results	12
7	Rating Average	13

List of Tables

1	World Internet Usage Statistics	7
2	World Internet Usage Statistics	7
3	Conversion Table	13
4	Accessibility Vs Usability table	13
5	Accessibility Vs Usability table (Results with no bias)	14

1 Introduction

The usability of the web content is not always a priority in the development, but is one of the most relevant things when you use a web page. Instead of usability, web developers make much more effort in other aspects of the web content like the appearance. Now days the importance of usability is growing up quickly, but of course we should never forget the accessibility to the web content of people with special needs. The idea of this project is to search if the adoption of the accessibility guidelines increases the usability of the web content or instead of that both things are not compatible.

Comparing the results of an usability user testing about some web pages with rated accessibility must be possible to find out the relationship between accessibility and usability, and get some information about how is it.

1.1 Acknowledgements

I would like to deeply thank the various people who, during the time in which this project was made, provided me with useful and helpful assistance. Without their care and consideration, this project would likely not have matured. Special thanks to my supervisor Morten Goodwin Olsen and all the participants of the user testing.

2 Problem Description

When web designers have to make web pages they must have under consideration many aspects. Most of times the most important aspect for them is the appearance of the web page, and they don't take so much care about accessibility, but this must change. Right now there are some recommendations made by the W3C in the WCAG to help web developers in their goal of improve the accessibility of web content. But there is no study about how the application of these recommendations affects the usability of the web pages, so this project complements the guideline giving more information to web developers. We will try to prove that the use of WCAG don't affect in bad way to the usability of the web content, and in the cases that it affect in substantial way we will give some new recommendations.

For this purpose we require two things, in one hand some control web pages, rated and measured by people with special needs. In the other hand one group of not disable people that will give their own opinion about the web pages. Comparing the results of the disable people with the non disable people, we will be able to find out the relationship between accessibility and usability. The accuracy of the results depends mostly of the number of people and the way to choose them. Of course we don't have to forget that choosing the correct web content and rate it in the best way is as important to get correct results as made good user testing.

This project can be a good tool to convince web developers that using WCAG recommendations they can achieve more than just improve the accessibility, they can made more friendly-use web pages for everybody.

3 Background

To start dealing with our Project, first of all we need a clear idea about what is accessibility and usability, for this goal nothing better to find definitions for both concepts. There are many definitions, many ways to say the same thing, these ones are easiest to understand.

- Accessibility in the context of a Web site is the degree to which that web site is usable by people with disabilities.
- Usability is the ease with which visitors are able to use a web site.

3.1 Literature Review

When we searched for literature about this topic, we didn't found very much, and this is the most important fact, there is lot of work to make in this field. Most of the works, are not based in any kind of user testing, they just compare the rules or recommendation for accessibility and usability, and get out some conclusions. Many of the papers and references used during this work were directly taken from EIAO¹, some of them were very useful sources² of information for the user testing. Anyway from these works we can as well get out some very important conclusions.

- Accessibility and usability can be improved at same time, their design philosophies are not incompatible.
- Is necessary to make a big effort in web accessibility and usability for improve quality, but this is not happening yet, because to several factors as the fast technological evolution.
- Quality is going to be in the near future the main goal of the designers.

3.1.1 The WCAG recommendations

The WCAG³ recommendations are the description of some techniques for authoring accessible content that apply across technologies. The recommendations are composed by several checkpoints. They explain how to develop accessible web content as how to test and measure the accessibility of web pages. The WCAG defines three levels of conformance and three different priority levels. Actually all these recommendations were taken from WCAG 1.0, we are waiting for WCAG 2.0 that will arrive soon.

Checkpoint priority levels:

¹European Internet Accessibility Observatory

²Specification and Implementation of the experimental environment, or Experimental results evaluation and analysis

³Web Content Accessibility Guidelines

- **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.⁴

Web page conformance levels:

- **Single-A:** All Priority 1 checkpoints are satisfied.
- **Double-A:** All Priority 1 and 2 checkpoints are satisfied.
- **Triple-A:** All Priority 1, 2, and 3 checkpoints are satisfied.

3.2 Choosing the users for the test

The population of interest in this study are all the internet users, that mean more than one billion people, of course is impossible make user testing to everyone. The idea is to take a sample group to characterize the billion users and making this is more art than science. We must think in witch characteristics of the subjects are relevant for our study, such as sex, nationality, study level, age. . .

The importance of the correct characterization is crucial to get correct results, making this section really important in the full experiment. We must be careful to don't get so complex sample impossible to manage it, and not so simple to get useless information.

3.2.1 Requirements

First of all we must think about the dimension of the sample group, as much bigger it is more accuracy in our results but more work to do. Usually this dimension should be between 0.0001/0.001⁵ percent of the population of interest, in our case using the lowest we have a sample group of 100.000 people.

Then we must choose the most relevant characteristics, in our case those are, in order of importance:

⁴Some checkpoints specify a priority level that may change under certain (indicated) conditions.

⁵Its not a fixed rate, just an aproximation taken from varius examples found in the web

	Country			Age		
	Occidental	Asiatic	Other	18-34	35-54	55+
Percentage	50	40	10	39	37	24

Table 1: World Internet Usage Statistics

	Studies Level			Sex	
	Univ. Student	Univ. Degree	Non-Univ.	Masculine	Feminine
Percentage	30	30	40	52	48

Table 2: World Internet Usage Statistics

1. Country: Or nationality, there are many differences in connection speeds and technological resources from our country to other as cultural and linguistic differences.
2. Age: The general computer knowledge level change a lot between people of different ages.
3. Studies level: For people with low studies profile can get more difficult to understand and get information from the web.
4. Sex: As long as the science knows, there are some differences between masculine and feminine brains, in relation to the search and organization objects in the space.

We search for the information⁶about those characteristics,we extrapolate and resume all this information, getting the results of table 1 and table 2. Of course the best would to get information from each country directly from their governments, and make the user testing in each country, but for obvious reasons that is not possible for our study.

3.2.2 Our Solution

As we now our resources are not enough to make a big user testing, so we are going to use just 10 people instead to the 100.000 people that we suppose to need. In the other hand we will try to make this group of people the most similar to our theory sample group.

3.3 Choosing the web pages for the test

We want to make a user testing of web content with different accessibility level. So we don't have many restrictions in this area, just found and rate different web pages, until we have a rated scale of 5 or 6 web pages. The number of web pages is limited by the resources we have for this project, but of course as in the

⁶www.internetworldstats.com/stats.htm and www.infoplease.com/ipa/A0908398.html

case of the choosing the users, the more web pages we use, the more accuracy we achieve.

Finally we decided to use some pages rated by the EIAO, from the complete list⁷ of pages, four were chosen. Because is time consuming for users to evaluate too many web pages, we will reduce this number to four pages. This time consuming makes more difficult to find participants for the users testing, which is the main reason to reduce the number of web sites to test. The selection criteria was the language and their rate, it was necessary to use English web pages of different rates. The selected pages are from England and Ireland:

1. <http://www.bankofengland.co.uk/>
2. <http://www.cso.ie/>
3. <http://www.taoiseach.gov.ie/>
4. <http://www.direct.gov.uk/>

3.3.1 Usability and Accessibility testing tools

The automated accessibility testing tools, have the ability to find out much faster that users the main accessibility issues of a web content, saving lot of time to developers. They are a very good help too for designers that are not familiar with the WCAG guidelines, telling them by easy and fast way if they are or not in the correct path.

This kind of testing tools has become indispensable to test the accessibility of a web page. But the results they give to us must be validate always by user testing tools, in that way we can see how good is a testing tool, by comparing its results with the user testing results.

3.4 The Questionnaire

The questionnaire must be similar to the questionnaire made for people with special needs user testing, with some litle tasks, like search for some specific information, and some little questions about general appearance of the page.

The questionnaire⁸ is going to be send by e-mail to the different users, so we decide to ask them to fill in their personal data to know who has already send it and who don't. After get all the questionnaires we will delete the personal data page and just use the ID field. In the questionnaire there are four different tasks:

- Task1: Find out how many bank notes of 20 pounds were issued in 2005. I decide for this task because the information can be found in some tables, and is one of the accessibility weakness points of the page.

⁷The full list can be found at <http://eiao2.eiao.net:8080/eiaogui?>

⁸The complete questionnaire can be found in the annex.

- Task2: Get the number of births, boys and girls, from the year 1939 to 1945 in Ireland. To accomplish this task its necessary to use a data base, the database has poor accessibility because the form control are not labeled.
- Task3: Find out and download the constitution of Ireland. This task is a bit more difficult because is not so easy to know where suppose to be the constitution, you can find it by going to the youth zone or going out the main page. Then only accessibility problem of this web page is the caption of the tables.
- Task4: Find out if you need a visa or not to enter Great Britain as Au pair. Use your own nationality and more details, and write them down in the comment space. This page has not accessibility weakness, so we just try to make one difficult task asking for special au pair visa.

4 Solution

As it's explained in the background we expect to be able to find a relationship between accessibility and usability, with this goal we are going to made the following.

The experiment has 3 different stages, the first one is to seek out all the necessary elements for the experiment, it means to make a questionnaire and find the user group and web pages. In the second stage while the user testing is been doing, we have time to rate the web pages too. Finally in the stage 3 we compare the results from the user testing and the rating of web pages.

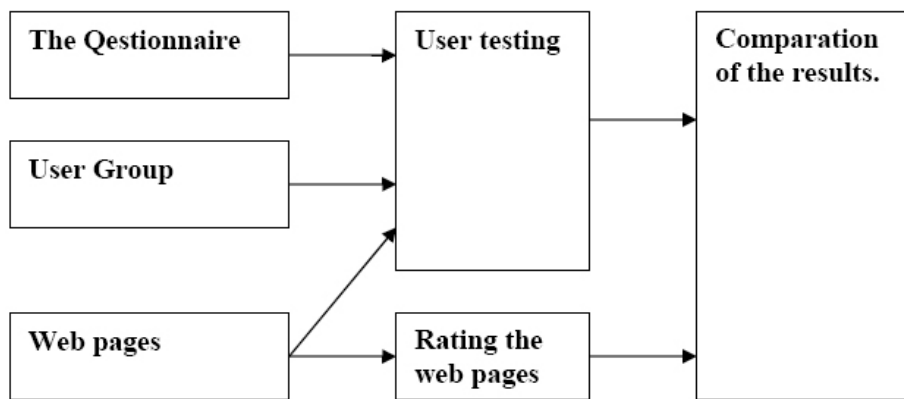


Figure 1: Schema of the experiment

4.1 First Stage

All the discussion about the requirements of the experiment are already reported in the background chapter.

4.2 Second Stage

4.2.1 Rating Web Pages

The pages selected were rated by automatic rating tool, this kind of tools just can find out the 30% of the accessibility troubles. These are EIAO results based on UWEM⁹ 0.5, since UWEM has changed a lot from 0.5 to 1.0 the accessibility results of these scores may also have changed. It means that the accuracy of this rating is not very high.

⁹Unified Web Evaluation Methodology

4.2.2 User Testing Group

The questionnaire was sent by e-mail to 15 candidates, from which just 9 send it back after the testing. The final composition of the user group was a bit different from the expected one, as we can see it in the Figures 2, 3, 4, 5.

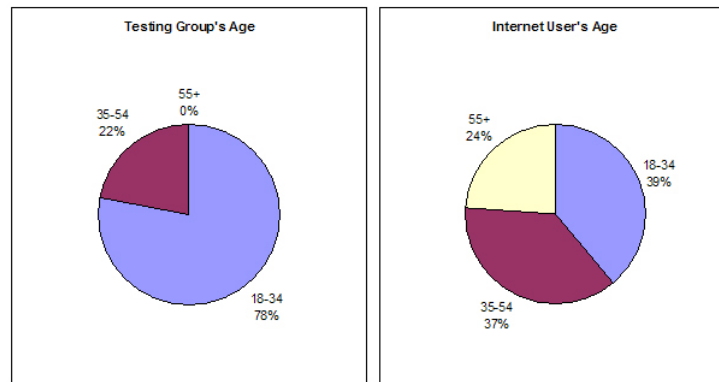


Figure 2: Age of Users

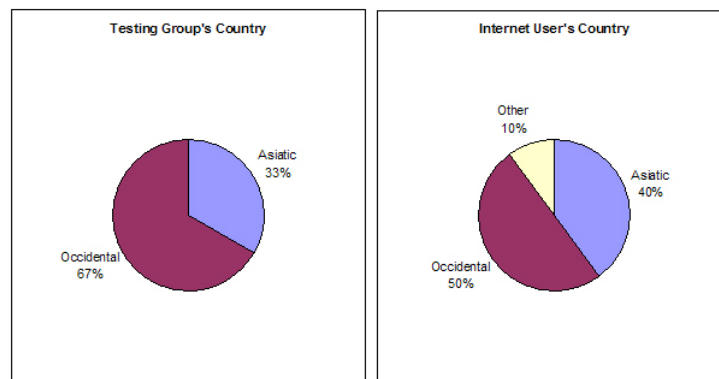


Figure 3: Country Of Users

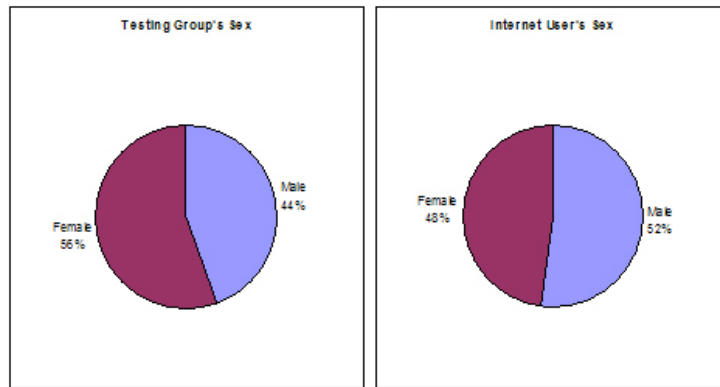


Figure 4: Sex Of Users

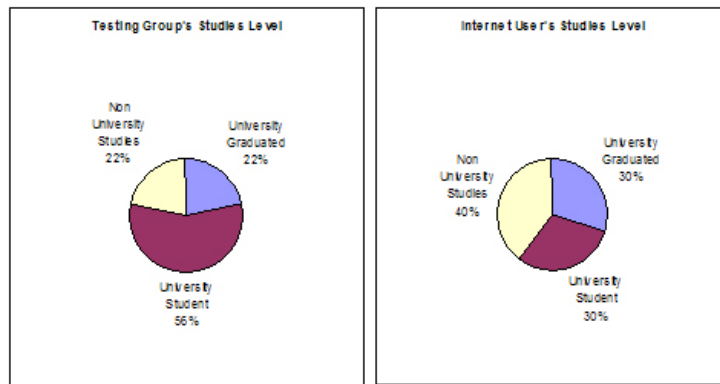


Figure 5: Study Level Of Users

4.2.3 User Testing's Results

The answered questionnaires can be found in the Annex II, and the results of the rating were used to create the Figure 6 and calculate the average showed in the Figure 7.

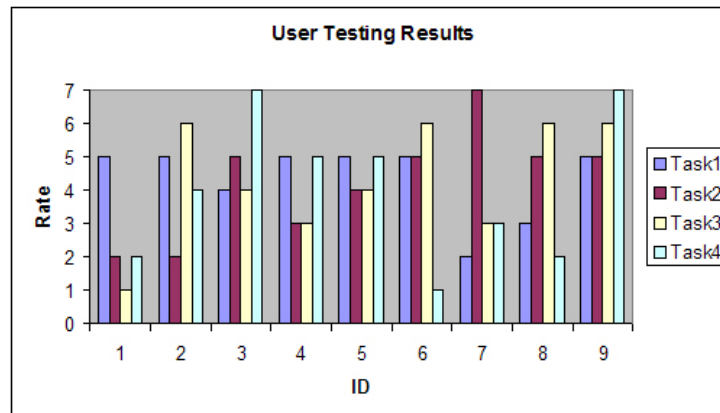


Figure 6: User Testing Results

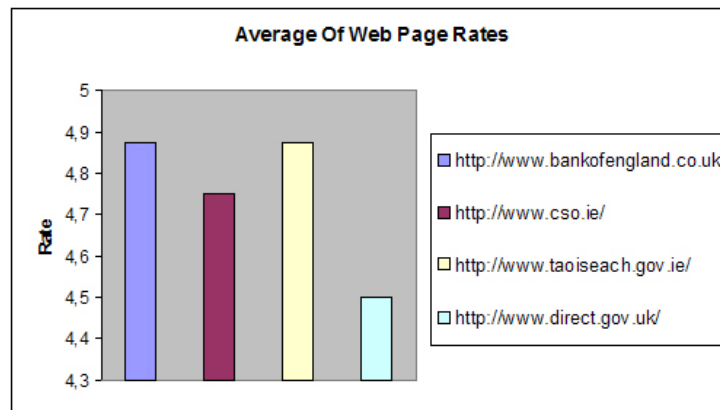


Figure 7: Rating Average

4.3 Third Stage

To compare the results of usability and accessibility we need equal scales for both, so we have to convert the usability results into the letter scale used in accessibility. To get this we calculate a step to create a conversion table, and

just convert the numeric values of usability that we have get in the user testing into letter values.

$$Step = \frac{(N_2 - N_1)}{L} \quad (1)$$

$N_1 = Lower\ Numeric\ Rate$

$N_2 = Maximun\ Numeric\ Rate$

$L = Number\ of\ letters$

Using this formula we can calculate the step:

$$Step = \frac{(7 - 1)}{5} = 1.2$$

With this step we create the table 3. Thanks to the table 3 we are finally able to convert and compare the results of usability and accessibility, in the table 4.

Letter	E	D	C	B	A
Numeric	1-2.2	2.2-3.4	3.4-4.6	4.6-5.8	5.8-7

Table 3: Conversion Table

	Task 1	Task 2	Task 3	Task 4
Accessibility	D	C	B	B
Usability	B	B	B	C

Table 4: Accessibility Vs Usability table

Looking to the results of the table 4 we see that it don't match our expectatives, in the next chapter we will discuss why .

5 Discussion

As we can see in the table 4 the results don't match the expectations, the usability user testing results are not equal to the accessibility results. In task1 and task2, we get better rating for the usability than for accessibility, instead of that the task4 has worst valuation of usability than accessibility. The only equal results are in the task3, that means that the solution match just 25% to our expectations.

At this point the question is if the main theory of a strong relationship between accessibility and usability is incorrect, or we mistake the way of probe it. The results shows differences between the rating, but those differences are small, several reasons can explain these gaps:

- The confidence in the user testing.
- The confidence in the automated accessibility rating.
- The relationship is not direct or linear.

5.1 The Confidence in the User Testing

Time and resource limitations force us to choose web pages from one list instead of create or rate our own web content, this make much more difficult to make an appropriate task for each web page. In task 4 we can see lot of bias, probably because to misunderstanding the language or because the participants are not very sure about what is a visa or au pair. We can see these biases in people with very bad rating of this web content. This problem have many different sources, first of all that many of the users didn't read the instruction set, or didn't pay enough attention to it, other possible source of the problem is the cultural differences, most of the Asiatic participants didn't know what au pair is.

The difficulty of the task has lot of importance in the achievement of the answer and people rate lower when are disappointed, we can see it clearly in the answered questionnaires of the Annex II. If we add to this factor that many of the participants didn't read the instruction and or for different reasons didn't pay enough attention to the user testing, we can not trust very much on these results. Using just the ratings of participants with no bias we found the table 5. As we can see we still have just one match now in the task 4, but the big gap between accessibility and usability of the task 1 has reduced. Is significant, but no determinant, because the results has varied a lot and the task 3 is almost a D instead of a C.

	Task 1	Task 2	Task 3	Task 4
Accessibility	D	C	B	B
Usability	C	B	C	B

Table 5: Accessibility Vs Usability table (Results with no bias)

Other important limitation that we have already mention about it in the background is the number of participants of the user testing, is too small to get an accurate results. And in this last result set of the table 5, we are using just four peoples results, so the results have really poor accuracy.

Thanks to the observation made in our user testing we can improve the confidence in future user testing taking in consideration the following recommendations:

- Choose carefully the participants, and do the user testing under supervision. This supervision is not a requirement when we can trust that participants will pay attention to the user testing and will make it seriously.
- Make some web content for each task, instead of use generic web content, this should be very good but very time consuming. In case is too much work, our recommendation is to make more than one task for each web page, to evaluate more than one thing in each page. With this we should be able to don't get too difficult tasks that people will don't understand and rate lower the page than it should be.
- Use the maximum of participants, this is a main factor, but of course we have the resource limitations.
- Use numeric scale instead of letter scale.
- Discard results of participants with a certain percent of bias. A good rate can be no more of 25% of bias.

5.2 The Confidence in the Automated Accessibility Rating

As we have explain the web sites were rated by automatic rating tool, this tool is limited to automatic testing and because of this, only discover 30% of the accessibility barriers. That means that our confidence in the accessibility rating is not very high too. To improve it our recommendation is to make a new user testing to rate the accessibility of the web content. Of course in this case the challenge of the user testing is much higher because we need people with special needs, that means that we should need much more resources. The best way to do this user testing should be with the same tasks but of course probably this is not possible, anyway we should use the same rating scale and method in both user testing.

5.3 The Relationship is not Direct or Linear

The third possibility is that maybe the relationship between accessibility and usability is not as easy as we expect in a first moment, probably the bad accessibility will not be a usability problem until it reach some kind of minimum

level. It can happen the same in the other side, if we have very good accessibility or just good accessibility the result of usability could be the same. To find out this kind of relationship we should need much more resources and different approach to the problem. With our experiment we can not find how it works the relationship between usability and accessibility if is not linear. To achieve such task we would need get deeper in both things usability and accessibility and make test of each basic principle of both.

6 Conclusion

Our experiment try to find out some conclusions about the relationship between accessibility and usability, unfortunately this goal was not achieved. The limited resources or a mistake in the approach can explain this fail. By the other hand we get some very useful information and conclusion for future research on this field as for better way of make user testing. The results are not enough good to probe that a relationship exists between accessibility and usability but the logic and previous works that we have read in the literature review bring as to think that this kind of relationship exist. In our work the results has not the enough confidence for taking out conclusions, but we can see that accessibility and usability are not so different from each other.

In conclusions it looks like that exist a relationship between accessibility and usability but is a bit more complicated than expected in a first approach and the resources we should need to understand the relationship between them are quite more big than we have had for our project.

7 References

- Web Structure Mining
URL:<http://www.eiao.net/webmining/previousprojects/reportgroup7.pdf>
- The Web Content Accessibility Guidelines
URL:<http://www.w3.org/TR/WAI-WEBCONTENT/>
- Usability and Accessibility Comparison of Governmental, Organizational, Educational and Commercial Aging/Health-Related Web Sites
- Testing for conformance to web accessibility and usability standards, an explanatory guide
URL:http://www.theusabilitycompany.com/accessibility_testing_guide.pdf
- Usability and accessibility: best friends or worst enemies?
- The Impact of Accessibility and Usability on the Development of Web Applications
- Experimental results, evaluation and analysis: User Testing of WAM Relevance
URL:<http://www.eiao.net/publications/EIAO.D4.3.1.pdf>
- Specification and Implementation of the experimental environment;
URL:<http://www.eiao.net/publications/EIAO.D4.1.1.pdf>
- UWEM1.0; Unified Web Evaluation Methodology
URL:<http://www.wabcluster.org/uwem1/>
- EIAO D4.1.1; Specification and Implementation of the experimental environment
URL: <http://www.eiao.net/publications/>
- EIAO D4.3.2; Experimental results evaluation and analysis
URL:<http://www.eiao.net/publications/>
- EIAO; European Internet Accessibility Observatory
URL:<http://www.eiao.net>
- WCAG; Web Content Accessibility Guidelines
URL:<http://www.w3.org/TR/WAI-WEBCONTENT/>
- WCAG2.0; Web Content Accessibility Guidelines 2.0
URL:<http://www.w3.org/TR/WCAG20/>

- WAI; Web Accessibility Initiative
URL:<http://www.w3.org/WAI/>
- Relaxed; Relaxed the HTML validator
URL:<http://relaxed.sourceforge.net/>