Software Development for Enhancing Accessibility and Fighting Info-exclusion

Protocols for Evaluation of Site Accessibility with the Participation of Blind Users

Abstract

User interfaces, through which users interact with the systems, should be designed according to the accessibility guidelines, with focus on usability. To this end, the designers of these interfaces should analyze whether their requirements meet the quality requirements, including their “user-friendliness” and “accessibility” features. The design of interfaces that are able to meet the needs of users with different characteristics is not a trivial task. It is paramount to involve the users and analyze how they interact with the system. Interface evaluation methods are absolutely necessary when gauging the system’s quality, since they allow the identification of many usability problems where the focus is on their accessibility. The present work aims to analyze two methods of observation involving people with totally impaired vision in order to develop a protocol with recommendations that can assist professionals in the identification of characteristics and problems, which can be solved or minimized during the interfaces’ evaluations, thus facilitating the process of system access focusing on users with totally impaired vision.

Keywords: Observation of users, Accessibility, Usability, Visual Impairment

1. Introduction

As user interfaces constitute a vehicle of communication with the Internet, through which a variety of information is transmitted to people spread out through the world [2], user interfaces should be designed according to the accessibility guidelines focusing on usability. To this end, one should pay special attention to the interface evaluation processes in order to analyze whether their requirements meet the quality criteria, including their “user-friendliness” and “accessibility” features. There is a body of recommendations and guidelines for designers in how to design accessible systems. In the case of already-existing systems, it is necessary that their interfaces be evaluated. Specific programs were developed to automatically evaluate the sites’ accessibility level. The site accessibility evaluation process is not restricted to the automatic procedure. It is also necessary to perform an evaluation by humans, specialists as well as impaired users.

In the design of a site that is universally accessible and is focused on usability it is paramount to observe and analyze the difficulties and skills of impaired users, as these difficulties and skills drive the mental model used by these users in their interactions with the system. This accessibility evaluation involving users should be the object of a case study, since the literature does not provide guidelines on how to conduct test with impaired users [12].

The observation methods should involve different kinds of users, including those with assorted limitations. By focusing only on people with totally impaired vision, the present work aims to analyze two methods of observation involving these users in order to develop a protocol with recommendations that can assist professionals in the identification of characteristics and problems, which can be solved or minimized during the interfaces’ evaluations, thus facilitating the process of system access focusing on users with totally impaired vision.

2. Evaluation of the Interfaces

The evaluation of interfaces consists of a systemic process of data collection with the intent of analyzing how the users utilize an artifact to execute their tasks in a computational environment [18]. When the evaluation methods of an interface are done in the presence of users they are referred as “observation methods or tests with users”. When the users are not involved, the evaluation is said to be done through “inspection methods or analytical or prognostic methods” [6].
2.1. Observation Methods with Users

During the interface evaluation it is important to involve the user, a better understanding of how this user creates his/her mental model of the computational system [7]. In observation methods involving users, the number of participants should be limited to five users of the same profile whereby these users use the system in similar conditions. This number was defined taking in account the recommendations proposed by Jakob Nielsen [13]. As the number of users increases, the information that is collected tends to repeat itself, providing a smaller amount of new information. According to research, five users are able to detect 85% of usability problems [13].

Observation methods of dealing with the context of the application’s use are performed in locations in which the user is accustomed (home, work, school). These observations allow the researcher to observe the participants in a familiar environment [17] [18]. However, interruptions caused by factors outside the evaluation process can lead to faults that are not a consequence of software problems, but are caused by the interruption of the task [5] [9].

In observation methods performed in a controlled context, the evaluation takes place in a monitored environment, such as labs. In these observations, the researcher is better able to control the variables influencing the evaluation [5] [9].

In the evaluation sessions, in the context of the user’s environment as well as in controlled spaces, the researcher should know the logical and strategic content that led the user to a decision during a task. For this purpose, it is necessary that the participants verbalize their thoughts during or after the evaluation session [5]. The simultaneous verbalization technique consists in asking the user to think aloud when using the interface [14]. In this type of verbalization users are stimulated to verbalize their thoughts during task execution. This technique shifts the user’s attention from task execution towards the explanation of the interaction, which can overload the user causing same to make errors in the interaction [5] [19]. In consecutive verbalization, the user is asked about his/her actions after task execution, in an interview. The researcher is responsible to remind the user about the interaction and ask the user to comment this situation. This solution increases the test duration. Moreover, the user may have forgotten the origin or cause of the problem [5] [19].

2.2. Observation Methods with Visually-Impaired Users

The participation of user with limitations helps in understanding how they interact in the Web and how they use assistive technologies [1]. Through the observation of interaction strategies of different users in different contexts and using various assistive technologies, one can identify the difficulties faced by them [12], incorporating the experiences of these groups as users of the system [20].

Although the outcome of an evaluation of interfaces with the participation of users belonging to a single profile, such only blind persons or only deaf persons, produces significant results, the feedback from a single user profile should not be assumed to be equal for all. In the evaluation of interfaces, it is recommended to include users of different profiles (different visual impairments, audition impairments among others) [1]. However, it is not always possible to deal with all this diversity among users due to deadlines, limitations in budget and in the number of persons which would have to be involved [9]. For this reason the present research was limited to totally visually-impaired users. The user’s degree of experience in the web and with assistive technologies also influenced the evaluation results. Novice and experienced users can help accessibility assessments, if properly classified into different groups so that the results are significant to the experience of each group, including in this experience the use of computers, web browsing and the use of assistive technology [20].

For these reasons it is important to conduct a pilot test, a preliminary assessment that allows checking for possible problems and shortcomings in the evaluation of the system [1]. A pilot test helps the researcher to determine the following questions [5] [19]: whether the hardware and software configurations required for testing are compatible and how these settings work together; the possibility of video-recording the interaction; determining whether the location of the tests has barriers preventing access to the disabled; confirming that the strategy of observation and recording of the actions is correct; determining whether the participants have questions about the description of tasks, questionnaires or other procedures and if the estimated duration of the tests is appropriate.
2.3. Previous Works

This article has identified a number of proposals based on methods used in usability evaluation with the participation of users, in order to evaluate the accessibility of web pages. They are:

- A technique of qualitative and subjective evaluation of accessibility and usability [3]. This technique is based on the statements of participants and on the audio of screen readers used by visually impaired users to characterize the problems.
- Recommendations for accessibility evaluations with the methods utilized by persons involved in web projects [4]. This studied assessed the pros and cons of some accessibility evaluations methods. Afterwards, a list of recommendation for the evaluation of web accessibility for the visually impaired was developed.
- A guide of best practices of evaluation of web accessibility with the participation of visually impaired persons [8]. This guide recommends that before performing the evaluation with the users, other evaluations such as heuristic evaluation be performed. The guide also makes recommendations on how to evaluate accessibility with visually impaired persons.
- A formal evaluation of usability with focus on accessibility [9]. This proposal describes the necessary steps during a usability test with the participation of visually impaired persons in the development process of a system.

3. Research Method

The qualitative and exploratory research consisted of the five steps described below:

(a) Choice of the category of users: it was necessary to choose a specific category of users in order to understand the observation process of these persons. The choice was for totally visually-impaired persons.
(b) Study of tools and resources: it was obviously important to understand the principle of accessibility and its implications in the evaluation of site interfaces.
(c) Observation in the users’ context of use: five users were evaluated in their work environment, interacting with the sites through their computers.
(d) Observations in controlled environments: the interactions of five users were evaluated in a lab. The users taking part in this step were not the same as the ones taking place in the previous one.
(e) Analysis of the results: the data collected in steps (c) and (d) were analyzed.

3.1 Limitations

The first limitation was due to the fact that it was necessary to choose a category of visually impaired users for the study’s object. In reality, visual impairments are quite varied and it would have been interesting to evaluate people with residual vision. However, time limitations and the availability of volunteers shifted the present study’s focus to persons with total visual impairment. The fact that the evaluation was conducted only in a few sites is also a limitation, since the problems that were encountered may be due to accessibility errors of the site, thus interfering in the users observations, masking the results.

4. Results of Observations Involving Users with Total Visual Disability

The researched literature did not provide any information about the number of users that should be involved in accessibility observations [4]. However, since Nielsen [13] does not suggest involving more than five users in usability evaluations, the present work adopted this criterion to limit the number of participants. The outcome of the tests (in the context of use and in monitored environments) showed that, in fact, the last two participants in each type of observation added little to the results.

Before the start of the evaluation, the participants were informed about the objective of the research, and about the study and the procedures that would be conducted. The users were asked to authorize their participation in the
evaluations and to record their interactions, using audio recording resources and a software that makes a video of their interactions. This consent form, due to legal issues should be in a format that the user can read without anyone having to read it to him or her (or via braille screen reader). Once the term of consent in the research was accepted, a questionnaire was applied to identify the users’ profiles, and thereupon the observations were initiated whereby each participant was asked to perform a task at each time. The entire evaluation process was record in audio or video depending on the type of observation and on the notes taken by the researcher to help in tabulating the results.

The definition of the tasks was done by the researcher and was influenced by the analysis of other works, such as the comparative study of access methods of sites by visually-impaired persons [11]. The tasks were selected so that the site pages necessary for task execution presented different layouts and functionalities, according to the W3C recommendation [1].

The interaction with the visually-impaired user is done through audition. To help reading the tasks, a document was created and delivered to the participants. This document contained the tasks’ definition and a brief description of the research objectives, the adopted methods and support software (observations in the controlled environment). In this file, no paragraph, font or tabulation formatting tags were used, as they could generate unnecessary content to be parsed by the screen reader.

Before the evaluations were performed (user context and controlled environment), a pilot test appropriate to each type of observation was executed. During pilot test it was determined that it would be unfeasible to familiarize the users with the site during an initial navigation period, since the users were not interested in navigating without a defined objective, making little use of the available time to understand the site’s contents.

When entering data, visually-impaired users, mainly the ones with total impairment do not use the mouse, preferring the keyboard. The pilot test showed that the use of the keyboard to which the user is accustomed in essential for an interaction. A device adapted for blind people is not necessary, since the common keyboard is a facilitating solution able to be used by any blind person. But it is important that it is a keyboard familiar to the user.

After the initial phase, each participant was asked to access the browser and the screen reader which he/she was used to, in order to check whether they were correctly configured and working correctly [9] [10]. Although the pilot test participants were given a document with the task description, they preferred to have it read to them by the researcher. Thus, the researcher presented one task at a time. During task execution, the participant was allowed to request the researcher to orally repeat its description.

A form of simultaneous verbalization was utilized to allow the researcher to clear up doubts about the interaction strategies adopted by the participants. Simultaneous verbalization does not constitute a barrier in task execution, since the user was accustomed to pause the screen reader to interact with other people and, after the conversation, would turn to listen again to the screen reader. Only one user and the researcher were present in each session, observing problems, comments and behaviors.

4.1. Observation in the Context of Use with Totally Visually-Impaired Users

The method of observation in the context of the user with limitations, besides the advantage of being conducted in the users’ own environment, allowing easier coming and going by them, prevents participants to use assistive technologies with which they are not familiar. This situation can lead to the detection of problems of interaction with the site caused by the difficulty in using the assistive technology and not by the site’s lack of accessibility [9].

The evaluation was performed in each user’s work environment using computers and software to which they were accustomed. Five male users with college degrees took part in the observations (including the user who participated in the pilot test). All had more than one year’s experience in accessing the Internet. Four used the Internet daily while one used it at least three times a week. The users told of using the Internet at the work site and at home for the following purposes: accessing banks, news, e-mail, search sites and the Intranet of the organization they worked for. Four used their worksite’s Intranet to execute professional activities.

During the evaluation, screen readers belonging to the users were utilized. Three users utilized Jaws, one utilized Virtual Vision and one used DosVox/WebVox.

The definition of the proposed tasks consisted in the selection of activities that could be performed by the
participants in accessing the chosen sites, in the environments to which they were accustomed. Some task details differed from one site to the other, due to the form in which the information was available on each site. Another criterion in task definition was the limitation of each task to ten minutes, so that each user’s total time under observation was at most one hour. This limitation was necessary, firstly due to the fact that the observations were conducted at the worksite and also to avoid that fatigue would reduce the participants’ perception leading to unreliable results [5].

With the application of the pilot test, it was possible to observe peculiarities in the conduction of the tests, which resulted in some adjustments. The following main issues were identified:

- A need to reduce the number of proposed tasks so that each test does not run over the maximum duration of one hour.
- A need to adjust the tasks’ description so that they are better understood, matching the description as much as possible to the nomenclature used on each site.
- The unfeasibility of recording the user’s interaction on video since the tests were performed in a work environment where the installation of video-capturing software was not allowed. Due to this, the observed results were annotated instead of being voice-recorded.

After the initial phase, each participant was asked to access the browser and the screen reader which he/she was used to, in order to check whether they were correctly configured and working properly [9] [10]. Thus, the researcher presented one task at a time to the participants. During the task execution, the participant was allowed to request the researcher to orally repeat its description.

The researcher recorded the time spent on the execution of each task [15]. Even though the time spent was recorded, this time was not the objective of this research. The purpose of timing the execution time of the tasks was to prevent a task from taking longer than anticipated (ten minutes), making the evaluation session time-consuming and tiresome. If the participant informed that the task was completed in the allotted time, the researcher verified if the task had really been performed [15] [16].

After the tasks had been completed, the participants were asked to report their impressions about the sites and the tasks. The reports of their experiences were not detailed, but all criticized the sites’ lack of accessibility and of the Internet in general.

Since the researcher would go to the user’s environment, the access to a pre-determined test site was not a problem. However, it turned out that it was impossible to use a single hardware and software environment for the test, since each user had his/her own configured environment to which said user was accustomed. The change to an environment with a single configuration could lead to results that would not only show the difficulties related to site accessibility, but also difficulties related to the lack of familiarity to a pre-determined environment.

The researcher faced a challenge in keeping the research time chart, considering that the users would participate on different dates, therefore having to schedule these sessions and proceed to the test sites. Oftentimes, the time the researcher spent in reaching the test site surpassed the time spent with the user during the accessibility evaluations.

4.2. Observation in Controlled Environments with Totally Visually-Impaired Users

A difference between observations in the user’s context and in controlled environments is the need to define the monitored environment. This means that an evaluation in a lab-like environment makes the preparation of an environment that reproduces the hardware and software used by the participants more difficult, besides causing some hardship when people with some kind of disability have to get to the site and return to their homes [5] [9]. On the other hand, the observation in monitored environments allows the use of tools and techniques that help the observations, for example: software that generates a video of the interaction that can be watched at a later time.

The monitored observations were performed during eight months, in a lab assembled in a University. In order to keep the same number of people of the sample used in the observations done in the users’ context, only five persons took part (including the user who participated in the pilot test), consisting of two males and three females.

All the participants were expert Internet users whose daily on-line time varied from 1 to 12 hours of interaction. They all reported that they used the Web to access e-mails, read news, etc.
Since the observations did not take place in a work environment, it was not necessary to limit the task execution time with the sessions’ duration being variable. To avoid fatigue, the participant was allowed to perform the test in as many sessions as necessary. The tests took approximately 36 hours. This was only possible due to the fact that the evaluations were conducted in a controlled environment.

The first stage of this observation consisted in defining an environment to be used, as it became evident that it was possible to use techniques supported by simulators of visual impairment. The tests were performed on a laptop using Mozilla Firefox 3.0 and the Jaws screen reader in its version 8.0. According to the participants, this software is the most popular among the blind, even though it is not free. Since the software is not free, it cannot be always used in observations in the user’s context, since many users prefer to use free software, like DosVox.

It was also found that it was feasible to record the users’ evaluations through software that makes videos of the interaction. Also used was Cantasia 5.1.0, screen-capturing software that records the user’s browsing throughout the site, in addition to the commands executed via keyboard which are read by the screen reader. Resources that recorded the screen reader’s output, the image of the user interacting with the system and his/her comments were also used.

The tasks were defined so as to simulate the daily use of systems by the participants, who could stop the session and resume it on another day.

The pilot test was performed with one user who was visually impaired with the objective of setting the path of the observations and checking the understanding to the tasks and questionnaire. It was possible to identify problems that could affect, directly or indirectly, the whole research. The main problems and their adjustments were:

- The need to configure the navigation system: several browsers were tried out, but Firefox had the best performance, even considering some crashes.
- Need to configure the hardware: The user executed the “keyboard recognition” of the laptop used in the tests, reporting that the placement of some keys did not conform to their disposition in said user’s computer. As a result it was necessary to use a keyboard in the familiar standard, connecting same to the equipment, in order to achieve a better performance of all the other participants.

After the necessary adjustments, the tests were performed with the other users. The first difficulty was to find totally visually impaired volunteers to take part in the observations, mainly due to the fact that these observations were to be done in a lab. Two participants reported that they had trouble getting to the locations and asked whether the observations could be performed at home, which, given the study’s nature, was not feasible.

Although the study’s objective was not to evaluate accessibility problems of the websites, but the method of observation, for purposes of information, when questioned about the main difficulties in general on the Web, all reported the same problems described by the users observed in their context of use.

5. Protocol for Conducting Observations with Users with Total Visual Impairment

The research was performed with ten visually-impaired users. Each method had five participants. The biggest difficulty in both methods was to find volunteers with total visual impairment to take part in the research. Based on these observations, it was possible to make recommendations for the conduction of observations involving visually-impaired users (Figure 1).

<table>
<thead>
<tr>
<th>Recommendations for observations in general</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limit the number of participants to five users of the same profile (similar deficiencies)</td>
</tr>
<tr>
<td>Make a pilot test (with only one participant) in order to check whether the procedures are adequate.</td>
</tr>
<tr>
<td>Inform the participants about the research objective and explain how the study and the procedures will be conducted.</td>
</tr>
<tr>
<td>Ask the users to sign a term of consent to take part in the test. This consent form, due to legal issues should be in a format that the user can read by him/herself (or via a Braille screen reader).</td>
</tr>
<tr>
<td>Apply a questionnaire to identify the users’ profiles.</td>
</tr>
<tr>
<td>Define the tasks according to some criterion, for example the W3C recommendation on feasibility evaluations [1].</td>
</tr>
<tr>
<td>Create a document containing the description of the tasks so that it can be read by a screen reader. Give preference to a text in a file, without paragraph, font, or tabulation formatting tags, in order not to generate unnecessary content to be parsed by the screen reader.</td>
</tr>
</tbody>
</table>
Recommendations for observations in general

Although the participants were given a document with the task description, they preferred to have it read to them by the researcher. The researcher should present one task at a time to the participants.

Allow the participant to request the researcher to orally repeat the task description.

Utilize the simultaneous verbalization technique whenever possible (as long as it does not disturb the user).

Observations in user contexts | Observations in controlled environments
--- | ---
It is not necessary to define the environment of the tests. | It is necessary to define the environment of the tests.
One should not count on the use of software that captures the screen and records the interactions on video. | One can (and should) utilize programs to record the interactions on video.
The observations should be concentrated in a short period and in a limited quantity. The time allotted for each test should be defined according to the environment’s conditions, such as work rules, if in such context. | The observations should be performed during a long period without limiting the number of tests. The user should be free to determine the time of each test.
The tasks need to be performed in a previously set time. | The tasks can be performed in a time not previously set.
The time spent by the researcher in reaching the test location should be taken in account. Sometimes this time is higher than that of the actual tests. | The researcher should not worry with the time spent to reach the lab.
The participants do not have trouble in reaching the test location, since they are already at the site. | Some participants may have trouble getting to the test location.
One can have more trust in the results. As the observations are usually conducted with assistive technologies running on computers and software configurations that the participants are already familiar with, the results represent only the difficulties related to accessibility. | One should be careful with the results. As the observations are usually conducted with assistive technologies running on computers and software configurations with which the participants are not familiar, the results represent difficulties related to the little experience of the participants with the preset environment.
One can use the user’s own keyboard. | It is necessary to configure the keyboard properly, and often connect a standard keyboard to computers used in the tests.
The observations are performed in different hardware and software environments. | It allows the use of a single hardware and software environment for the realization of the observations.

Fig. 1. Protocols for Conducting Observations with Users with Total Visual Impairment.

6. Final Considerations

The present study aimed to analyze the different observational methods involving visually impaired users: one performed in the user’s context, and another one in a controlled environment. Ten users were observed, five in each method. The biggest difficulty in both methods was to find volunteers with visual impairment to take part in the research.

User observation allowed the researcher to detect not only usability issues, with focus on accessibility, experienced by visually, but also peculiarities involved in the interaction, due to the characteristics of the observed users’ profiles (visual impairment).

As the tests were performed with visually-impaired users, it became possible to learn how they interact with the sites and to understand the difficulties they have during the interactions, observing the solutions that the users try to reach to attain their objectives as they navigate on the Web.

It was possible to verify that the main benefits of the observation with visually-impaired users were: learning a different kind of interaction; experience of the difficulties and solutions of each user in accessing the web and identification of problems not reported in the literature and/or by professionals who assess interfaces where these problems are related to the way users navigate.

The results were instrumental in creating a list of guidelines to help specialists and researchers in the conduction
of evaluations of accessibility for visually-impaired users, so as to contribute in the construction of a protocol with recommendations that could help the evaluators in identifying characteristics and problems, which could be solved or minimized in the evaluations. One hopes that these protocols provide a contribution so that the process of site accessibility focusing on users with totally visual impairment can be more systematic.

References