Abstract

Recently, the production and availability of multimedia Web content, as videos, have increased. In this scenario it is important to consider accessibility requirements so that any user can whelm the barriers to access content regardless of the limitations imposed by either deficiency or some temporary restriction. One of the main barriers found in the current players is the restriction to make videos accessible on the Web and little research on how to overcome those limitations has been conducted. This paper describes the three phases of User Centered Design, in which an evaluation with real users of an accessible video player is conducted. A video player, called Facilitas, is proposed in order to provide the rationale of how some of those barriers or limitations could be overcome. Its controls are new and different from the ones in other players. We have observed that by reducing accessibility barriers, the design process leads to an improved product in terms of usability. A user testing is described to explain which controls users frequently use to complete a task.

Selection and peer-review under responsibility of the Scientific Programme Committee of the 5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2013).

Keywords: Accessible Web Videos; Accessible Media Player; Facilitas.
1. Introduction

Recently, the production and availability of multimedia Web contents have increased [6], stimulating Web accessibility initiatives and attracting researchers from different fields due to the social inclusion contributions involved and challenges to make the huge amount of Web multimedia content accessible. Accessibility requirements must be considered so that any user can access content regardless of the limitations imposed by either deficiency or some temporary restriction [11].

The W3C (World Wide Web Consortium) has standardized a set of guidelines to make multimedia content accessible. Recommendations for making Web content more accessible are described in Web Content Accessibility Guidelines - WCAG. Guidelines for designing user agents to help disabled people are described in User Agent Accessibility Guidelines - UAAG. Guidelines to define how authoring tools should help Web developers produce Web content are accessible in Authoring Tool Accessibility Guidelines - ATAG. A problem to include those guidelines in Web content is that the designer needs to choose how they should be implemented, by, for example, considering the codification language or browser. When multimedia content, as video, is included in Web sites, the video access must be included in the code and the interaction with the users must be intuitive.

Videos have had a strong impact on people’s lives (for example, YouTube†, Facebook‡, Netflix§) and required special attention on their accessibility capacity. However, most of the assistive technologies in Web content concern text, indicating that more studies on other media types, such as audio and video are necessary. Some researchers have focused on video player functionalities [3, 4, 5, 7] and others on media player for disabled users [1, 9].

In [5] we can see a video player with accessibility resources regarding WCAG and UAAG guidelines. To the best of our knowledge, this is the first report in the literature on WCAG and UAAG video issues. However, the researchers in [5] do not conduct evaluations with real users in the HCI domain. We have taken the first steps to include user evaluation, following User Centered Design (UCD) and resources of multimedia content, as tags and search to improve media to be available in Semantic Web.

No research on the reduction of accessibility barriers to zero has been conducted, therefore we argue that a video player must be designed to minimize the accessibility barriers without disregarding usability. This paper describes the three phases of UCD, in which an evaluation with real users of an accessible video player is conducted. A video player, called Facilitas Player, is proposed in order to provide the rationale of how some of those barriers or limitation could be overcome. We propose and have tested new controls, such as tags and search (to cite some), which enable users to navigate through the video searching in closed caption or audio description. Such controls were designed to be compliant with W3C guidelines, especially UAAG. We have proved that, as Petrie concluded [9], if accessibility increases, usability also increases.

The paper is organized as follows: Section 2 discusses the related work and the W3C guidelines are presented in Section 3; Section 4 addresses the three phases of UCD process; finally, Section 5 concludes the paper and suggests some future works.

2. Related Work

According to González et.al 2011, some basic controls are necessary to obtain an accessible media player: play, stop, resize and volume. The authors describe some additional controls: enable/disable subtitles and audio description, search in the caption text, forward or delay seconds within a reproduction, change the size, font or color of the text, help documentation, among others. They also report the requirements concerning accessibility close to usability: maintain accessibility features configured by the users, enable/disable and adjust accessibility

† YouTube: www.youtube.com
‡ Facebook: www.facebook.com
§ NetFlix: https://signup.netflix.com/
features, provide information about keyboard shortcuts, enable navigation through the content, use keyboard to move through the menus, and enable users to set their preferences to configure the keyboard shortcuts.

Moreno et al. 2011 discussed accessibility for media players on the web, comparing YouTube, BBC iPlayer and CCPlayer, all of them developed with Flash technology. They are embedded in a web page and enable users to access the content without another application opening. To the best of our knowledge, CCPlayer was the most accessible player reported in the literature until [7] developed an accessible HTML5 Media Player using not only HTML5, but JavaScript and CSS as well. The player has the following controls: play, stop, rewind seconds, forward seconds, volume controller, audio controller, caption on/off, audio description on/off, help guide and select caption language, but it does not support search functionality (Fig. 1a).

YouTube is the most popular video search system on the web and provides some controls (Fig. 1b), like captions, screen resize, automatic transcription and subtitles, which could help accessibility issues. However, subtitles are difficult to operate in some browsers and are not accessible by the keyboard. Another problem is that screen reader tools for the visually impaired people cannot always distinguish the function of controls implemented in Flash, therefore some screen readers cannot access controls at all. Aiming to achieve an accessible Video Player, we have developed complementary technological solutions, provided in section 4.

Cheng et al. 2009 developed a video player that adjusts the current playback speed. Our video player has forward and rewind controls. Anthony et al. 2013 collected and analyzed 187 noncommercial videos uploaded to YouTube and coded them in a range of dimensions to characterize the interaction, the challenges encountered, and the adaptations adopted in daily use. The authors tested the videos with physically disabled users and showed that although many people with motor impairments find these devices empowering, accessibility issues still exist. Media repositories, as YouTube, not only enable users to upload their videos, but also encourage them to annotate the videos with descriptive words, called tags. Tags provide the description of video content and facilitate the categorization, sharing and search of videos. Li et al. 2011 claim even if the tags are provided for a whole video, they may describe only a small part of the video content. Therefore, when searching for video information via tags, users are often bewildered by the vast quantity of seemingly unrelated videos returned through video search engines and must browse through each video to find the interesting parts.

As can be seen, issues on Web video accessibility still need to be studied. The research found in the literature focuses on specific guidelines of W3C, emphasizing some functionalities different from the traditional ones. CCPlayer has search, but not audio description. Moreno's player has audio description, but not search. YouTube has automatic transcription and subtitles, but screen readers cannot access its functionalities at all.

3. Web Accessibility Guidelines

The WCAG documents explain how to make web content more accessible to disabled people [12]. Web "content" generally refers to the information on a web page or web applications, including natural information, such as text, images, and sounds, and code or markup that defines structure, presentation, etc. The ATAG documents define how authoring tools should help Web developers to produce Web content that is accessible and in conformance with WCAG. ATAG documents also explain how to make authoring tools accessible so that they can be used by people with disabilities [14]. The UAAG documents provide guidelines for designing user agents to reduce Web accessibility barriers for people with disabilities. User agents include web browsers, media players,
plug-ins and other types of software that help retrieve, render and interact with Web content [13]. A user agent that conforms to these guidelines will promote accessibility through the user agent interface and other internal facilities, including the ability to communicate with other technologies (especially assistive technologies). UAAG 2.0 also requires media players should provide documentation of accessibility features in an accessible manner and access to content through a variety of navigation mechanisms and implement interoperable interfaces for communication with other software [2].

As discussed in Section 2, related works have explored some guidelines, specially basic functionalities with another different functionality, but more functionalities still need to be tested. Some guidelines, as provide text configuration (Guideline 1.4), provide synthesized speech configuration (Guideline 1.6), configure and store preference settings (Guideline 2.7) and customize display of GUI (Guideline 2.8) have never been explored.

Most players also use Flash technology, however it is not accessible for assistive technologies in most cases. The only player developed with HTML5 technology is Moreno's player [7], but some functionalities have been added, because HTML5 provides only basic controls: play/pause, volume and full screen.

4. User Centered Design

The UCD process comprises three phases: design research, design, and design evaluation. During research, the designer’s purpose is, among other things, to assess the users and their needs. The second phase, based on findings from the design research, develops the user interface, the document, the information architecture, etc. Once a design has been drafted, the UCD practitioner evaluates it with users and revises it based on the results of the evaluation [15].

4.1. Phase 1 - Design Research

Our purpose is to assess users and their needs. We performed a preliminary user inquiry on how users watch videos on the Internet and which functionalities they use. There were 9 unpaid participants in this user inquiry: 5 males and 4 females, between 23 and 63 years old. Two participants only watched TV and 7 were computer-savvy users with experience in watching videos on the Internet. We asked the participants what types of video programs they would like to watch. The first five types were TV shows, films, video clips, documentaries, and cooking videos. We asked about the video players they usually used. The first five players were YouTube, Windows Media Player, Media Player Classic, VLC and a player from a news web site. Most of the participants used basic controls of media player: play/pause, resize, volume, caption and caption settings.

We asked about problems they usually faced when they wanted to watch a video. Two participants reported problems with codec, two with volume control, and two with rewinding to some part to hear it again. Other problems were the difficulty in associating subtitle with video, errors with special characters in subtitles, speed of subtitles, difficulty in understanding subtitles, difficulty of keyboard navigation, lack of sound equalization, lack of shortcut keys for full screen, lack of close button and the automatic loading of the video.

We asked them about improvements to the video player. They answered they would like to edit the captions, access the keyboard, search a scene, select an interval, share some parts of the video, use voice recognition for controls, use annotation on the video because they usually needed to write the timestamp, read some description of the video to know the content and search by content and not by title.

Based on the difficulties and problems reported by the participants, we identified some accessibility problems. At first, we developed and tested some functionalities, as provide access to alternative content (Guideline 1.1), provide text configuration (Guideline 1.4), provide volume configuration (Guideline 1.5), ensure full keyboard access (Guideline 2.1) and provide text search (Guideline 2.4). Next, we tested tag control as a new functionality. All these functionalities are implemented in Facilitas player.
4.2. Phase 2 - Design

This phase involves brainstorming and conceptualizing and sketching initial drafts of the design based on findings from the design research. We have developed an accessible media player, Facilitas Player, using HTML5, JavaScript, jQuery, jQuery UI and CSS to provide functionalities to make videos accessible. The attributes are included in `<video>` tag. The player is made as a jQuery plugin, therefore, by simply calling $(selector).facilitasplayer(options) the Facilitas player will be loaded. Since it is a jQuery Plugin, its architecture is based on Implicit Invocation, i.e. after initialized, all controls and listeners are instantiated and the player is ready to receive event notifications from the video tag and/or from toolbar buttons. When an event notification is received, it invokes all procedures registered for that event. For instance, when the time changes, the video sends a notification announcing it. The player receives the notification and updates its interface, rounding the milliseconds to seconds and displaying the elapsed time of the video to the user.

The current controls include basic controls as play/pause, rewind, forward, volume controller and full screen and new controls as caption, search, tags, setting panel (text configuration) and light on/off. Some functionalities of accessibility also include highlight and keyboard access.

Some of the accessibility features of current media players are in conformance with the UAAG: G1.1: Alternative content, G1.3: Highlighting, G1.4: Text configuration, G1.5: Volume configuration, G1.8: Orientation in Viewports, G2.1: Keyboard access, G2.4: Search, G2.7: Preference settings and G2.8: Toolbar configuration. Based on the discussion in Section 2, we compared the features of players (Table 1).

<table>
<thead>
<tr>
<th>Controls</th>
<th>JW player</th>
<th>YouTube</th>
<th>BBC iPlayer</th>
<th>CCPlayer</th>
<th>Video Player[7]</th>
<th>Facilitas Player</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative content</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Highlighting</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Text configuration</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Volume configuration</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Orientation in Viewports</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Keyboard access</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Search</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>List with search result</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>Preference settings</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Toolbar configuration</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>Light</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>Tags</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Four of those functionalities are present only in Facilitas player: list with search result, toolbar configuration, light and tags.

Tag control enables the developer to add tags to the video creating links to divide videos into parts. Each tag is linked to a specific time in the video. Tags provide a short description of the video content and a long description when the tag is selected, facilitating the search of videos. For instance, in Fig. 2, the video has six tags. If we select the "Tip: dark chocolate" tag, the video skips to the third tag time and a long description appears.

The search control enables the search for a word or phrase that appears in the subtitle text. The player will show all results and when a result is selected, it skips to that point on the video. For instance, in Fig. 2, we searched for the word “butter”, returning a set of two entries. When a result is chosen, the player reproduces the selected part of the video.
Another functionality of Facilitas Player is the setting panel, which enables text configuration to change style, color and size in real time, as shown in Fig. 3. A control to move the Facilitas toolbar still in development, the toolbar will be docked at the top or bottom of Facilitas Player. Finally, a light functionality is represented as a lamp icon on the video (see Fig. 3) and is used to distinguish the video from the rest of the page.

4.3. Phase 3 - Design Evaluation

In the context of usability evaluation, Nielsen [8] proposed a set of 10 heuristics to guide experts to test a users interface. When accessibility is included in the test with users, the test can be affected by two heuristics: flexibility and efficiency of use (heuristic #7) and aesthetic and minimalist design (heuristic #8). Therefore it is important that the other usability heuristics are not damaged during the testing with real and common users. Following W3C guidelines does not necessarily mean to make content accessible to a wider range of people with disabilities [10].
Power et al. tested accessibility with blind users and the results suggest the need to research into the priority levels of the current WCAG 2.0 because they are difficult to measure. The priority levels for UAAG also need to be researched.

We performed an experiment to know the important functionalities of videos and those the participants chose to complete some tasks. The participants were provided with five videos using Google Chrome browser. They were instructed to choose two of them. They completed a series of tasks in which they had to answer three questions for each video. The first two tasks were questions about the video content, they had to show the scene where the answer could be found. They used some controls to find the scene: search, tag, rewind/forward and time bar. The third task referred to subtitle configuration and they used the settings panel to configure the color, size and font.

We used five videos with subtitles in Portuguese for the testing. Two of them have audio in English and three in Portuguese. For each video, we created between 2 and 7 tags (5 on average). The videos lasted 4 to 10 minutes (6 on average).

Table 2. Features of videos.

<table>
<thead>
<tr>
<th>Video type</th>
<th>Audio</th>
<th>Subtitle</th>
<th>#Tags</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooking video</td>
<td>Pt</td>
<td>Pt</td>
<td>7</td>
<td>4:25</td>
</tr>
<tr>
<td>Documentary</td>
<td>Pt</td>
<td>Pt</td>
<td>6</td>
<td>10:17</td>
</tr>
<tr>
<td>Comedy TV show</td>
<td>En</td>
<td>Pt</td>
<td>4</td>
<td>5:58</td>
</tr>
<tr>
<td>Terror film</td>
<td>En</td>
<td>Pt</td>
<td>2</td>
<td>4:29</td>
</tr>
<tr>
<td>Comedy Film</td>
<td>Pt</td>
<td>Pt</td>
<td>5</td>
<td>4:08</td>
</tr>
</tbody>
</table>

We used two applications of Morae software**, i.e. Recorder and Manager, to facilitate the research process and data analysis. Recorder enabled the capture of audio, video, user input and on-screen activity. Manager application enabled the analysis of the video records.

Ten people, six of them university students, participated in the experiment (See Table 3). Their age ranged from 23 to 63 years old. Four participants were female. Each participant spent approximately 30 minutes on the experiment. The participants chose two videos based on their preferences: 35% chose a cooking video, 30% chose a documentary, 20% chose a TV show, 10% chose a terror movie, and 5% chose a comedy movie. Participant 7 has a mild hearing impairment, but she did not use any hearing aid, three users (P1, P2 and P4) had myopia and wore glasses, which did not affect the interaction, and the others had no disability.

Table 3. Participant's Characteristics.

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>First Video</th>
<th>Second video</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>26</td>
<td>Masculine</td>
<td>Comedy TV show</td>
</tr>
<tr>
<td>P2</td>
<td>27</td>
<td>Feminine</td>
<td>Cooking video</td>
</tr>
<tr>
<td>P3</td>
<td>23</td>
<td>Feminine</td>
<td>Comedy TV show</td>
</tr>
<tr>
<td>P4</td>
<td>24</td>
<td>Masculine</td>
<td>Documentary</td>
</tr>
<tr>
<td>P5</td>
<td>32</td>
<td>Masculine</td>
<td>Documentary</td>
</tr>
<tr>
<td>P6</td>
<td>25</td>
<td>Masculine</td>
<td>Comedy TV show</td>
</tr>
<tr>
<td>P7</td>
<td>59</td>
<td>Feminine</td>
<td>Terror Film</td>
</tr>
<tr>
<td>P8</td>
<td>63</td>
<td>Masculine</td>
<td>Documentary</td>
</tr>
<tr>
<td>P9</td>
<td>23</td>
<td>Masculine</td>
<td>Cooking video</td>
</tr>
<tr>
<td>P10</td>
<td>23</td>
<td>Feminine</td>
<td>Cooking video</td>
</tr>
</tbody>
</table>

** Morae software - http://www.techsmith.com/morae.html
We analyzed the time it took each participant to complete a task. Figure 11 shows the time in seconds for the first two tasks. The tasks are labeled with a letter (P=participant, V=video, T=task) followed by a number.

In Fig. 4, tasks are filled with colors representing the type of video the participant chose. Red (P1V1, P3V1, P6V1, P10V2) represents a comedy TV show, purple (P1V2) represents a comedy movie, blue represents (P2V1, P3V2, P4V2, P5V2, P8V2, P9V1, P10V1) a cooking video, yellow (P2V2, P4V1, P5V1, P7V2, P8V1, P9V2) represents a documentary, and orange (P6V2, P7V1) represents a terror movie.

Fig. 4. Average Time for tasks

Tasks P1V1T2, P2V1T1 required more time than the other tasks, because in P1V1T2 the participant searched eight times with words that did not return the expected result. He used the progress bar to complete the task. P2 also said that she did not pay attention to the video because she was distracted using the video player controls. In P2V1T1 task, she used four tags and one search with an unsuccessful result; then she used the time bar to complete the task. After finding the answer she tested the search control using one word of the subtitle and checked if the search result was correct.

P1 spent less time on the second video than on the first he had watched. He paid attention and answered the question about the video content (P1V2T1) using his memory. In P1V2T2 he used one tag. Some tasks required more time than others independently of the participant's characteristics. For example, in the terror movie (see Fig. 4, P6V2 and P7V1) task 2 required more time than task 1. Both P6V2T1 and P7V1T1 used the progress bar twice to complete the task. P6V2T2 and P7V1T2 used search three times to complete the task. In P7V1T2 task, the participant wasted her time on misspelled words.

Two participants (P3 and P10) are women and chose a comedy TV show. The first task took longer than the second one. In P3V1T1 task, the participant used two tags and one search to complete it. In P10V2T1 task, she used the same word of P3 to search. In both P3V1T2 and P10V2T2 tasks, the participant used one search to complete them.

70% of the participants (P1, P2, P3, P5, P7, P8 and P10) spent more time on tasks for the first video than on those for the second video, i.e. they learned how to use the control functionalities in the first video.

To know what controls the participants used during a task, we analyzed each video and counted how many times they used each control until the task had been completed (Fig. 5a). During a task, the search control was used 53.6%, time bar 30.9%, tags 13.4% and forward 2.1%. We also counted how many times the participants used each control to complete the task (see Fig. 5b). For example, in V1T1, 6 tasks were completed using search control. The search control was used 62.5% to complete a task, time bar 30%, forward 5% and tags 2.5%.
Figs. 5a and 5b show that the search control was the most useful control, followed by the progress bar control. Although the participants used tags, only one task was completed using this control. The participants also used the forward control to complete two tasks. After having watched V1, P1 said "one thing that was different was the marking of tags. I had never seen this functionality before, but I was watching to see the tag functionality". After watching V2, P1 wanted to add new tags to the video and said "Adding tags to the video would be a good option. I would like to rename the first tag". P3 said that at the beginning she did not understand the tags, but then she realized that the tags were marked in the progress bar. P9 said that tags mean keywords and made a relation with the keywords used in YouTube. All participants, except P1 and P3, suggested that they could have had the option of adding some tags to the videos. P4 and P5 suggested deleting some tags.

Regarding the caption configuration task, we asked the participants to perform one task for video 1 and another for video 2. Three types of tasks were created: change font style to Verdana (red color in Fig. 6), change font color to yellow (yellow color in Fig. 6), and change font size to 20 (blue color in Fig. 6).

Four Participants (P2, P5, P7, P8 and P10) had difficulties in changing the subtitle color because in the configuration window there is a point to select color (see Fig. 3). By default the point was in the upper left corner (white color), so when the participant changed the color, the subtitle was always white, causing confusion. In all cases, that point was not observed by the participants. P2, P5 and P10 have computing skills, but P7 and P8 do not. When P7 learned how to change color, she began to test changing subtitle colors. A feature of the panel settings is to change style, color and size automatically without clicking on the button. P5 liked the color change feature and P10 said that the change size feature is really important. In fact this issue is the one directly related to the 8th usability principle mentioned by the users.
5. Conclusions and future work

We followed the three phases of UCD and tested Facilitas Player with real users, although the users are not disabled.

The experiment provided important results about new controls, tags and search, which will help to improve Facilitas Player. The tag control was not frequently used in the test cases, probably because it was different from other functionalities and participants had never seen it before. During the user test, we detected a misunderstanding among tags and keywords, therefore we proposed changing the name “tag” to “link” and added the phrase “Go to:” before the tags to orientate the participants.

On the other hand, the search control turned to be very useful, as it was used in 62.5% of the cases to successfully complete a task. This functionality was more intuitive than the tags. At the end of the user test, we confirmed Petrie’s conclusion [9]: the design regarding accessibility issues promotes better usability.

As future work, we propose adding some functionalities to Facilitas Player: documentation, language, on hover, preference settings and annotations. We also suggest testing it with older people to know what functionalities of video player are common for them.

Acknowledgements

The authors would like to acknowledge CAPES (PEC-PG) for the financial support provided to this project.

References

13. W3C, User Agent Accessibility Guidelines 2.0 (2013), http://www.w3.org/TR/UAAG20/