EVALUATION OF IMAGE ACCESSIBILITY FOR VISUALLY IMPAIRED USERS

Tahani Alahmadi¹, Steve Drew²

¹,²School of information and communication technology, Griffith University
²Tasmanian Institute for Learning and Teaching, University of Tasmania
¹College of Computer and Information Science, Princess Nourah Bint Abdulrahman University

Received: 2018-02-15 | Accepted: 2018-10-13 | Published: 2018-11-30

Abstract: The accessibility of website images is influenced by the availability and accuracy of descriptive text and its compatibility with the images’ complexity and purpose. Image accessibility evaluation cannot be fully affected through applying one method, and it can be enhanced by the inclusion of processes that consider the quality of descriptive text for images. The evaluation of descriptive text quality may initially involve human evaluation and then use of an automated evaluation tool to provide a counterpoint. In this paper, an analysis is presented of a dataset of 120 complex and informative images found on universities’ Web-based systems. This is supplemented with a detailed analysis of HTML image attributes and elements. Human and automated analyses of content are combined, and the information is integrated to inform the evaluation’s outcome. Our analysis illustrates a lack of accurate usage of HTML image attributes and elements, such as alt and longdesc. The findings provide insight into improving image accessibility by applying multiple evaluation methods and auto-generated descriptive text. This paper will be of interest to Web accessibility developers and researchers.

Keywords: image accessibility, descriptive text, alt text, visually impaired, human evaluation, automated tool evaluation.


125
Introduction

Accessibility evaluation is an important equity step in assessing the effectiveness and usefulness of online materials for users with disabilities. Pipino, Lee, and Wang (2002) considered accessibility a part of the data quality dimensions that they proposed. Using data quality assessments, the authors defined accessibility as “the extent to which data is available, or easily and quickly retrievable” (Pipino et al., 2002, p. 2). The ready availability of data to all users is the core dimension that affects data quality in any Web-based system. This quality encompasses the accessibility of media content, such as images and videos, and the availability and accuracy of text that describes images for visually impaired users. In providing descriptive text, the main considerations that developers should consider are the images’ complexity and purpose.

Accessibility is a complicated matter that involves the consideration of many aspects, including the features of systems, the characteristics of disabled user groups, the effects of embedded files, and the roles of assistive technologies. Considering these varied aspects, a multi-method evaluation scheme is well-matched to measuring accessibility and design development plans for specific Web-based systems such as university information systems. Aware that a single approach cannot accurately measure accessibility rate, many scholars (Biswas, Duarte, Langdon, Almeida & Jung, 2013; Gómez-Martínez et al., 2015; Sun & Strybel, 2017) have highlighted the vital contribution of combining methods to achieving favourable results. However, multi evaluation methods for image accessibility on university Web-based system has not been addressed in detail.

Using the above mentioned considerations as bases for evaluating large data sources on the World Wide Web may enable organisations to understand accessibility problems, develop image accessibility solutions, and improve the accessibility rates of current systems. To these ends, we conducted
human and automated evaluations to measure the accessibility of university Web-based systems. This study’s main aim is to highlight the importance of including human evaluation in image accessibility testing, as human evaluation is the only current means of measuring the accuracy of descriptive texts. Also, this study provides details on HTML image attributes and elements’ usage on university Web-based systems. From this analysis, elements of design for future accessibility-smart solutions can be used to create quality descriptive text with usable tools, even for complex images. A further aim is to highlight significant image accessibility barriers that prevent visually impaired users from receiving the same information from an image as their sighted peers.

The following sections discuss multi-method accessibility evaluation and present the methods adopted in this study and the findings that we derived.

Related Work

The literature was reviewed to identify the key issues related to evaluating image accessibility for visually impaired users. This section discusses multi-method accessibility evaluation (human and automated), with emphasis on the accessibility of images on Web-based university systems.

Visually impaired Characteristics

A sensory disability is defined as a disability that relates to one or more of the human senses, such as vision impairment, hearing impairment, or both (Oliver, 2017; World Health Organization, 2010). Vision-impaired individuals are the primary stakeholders in this study. Vision or visual impairment is a health condition of the eyes that cannot be corrected with standard solutions such as glasses. The World Health Organization (2010) defines three categories of vision impairment (severe, moderate, and mild impairment) and three categories of blindness based on visual acuity tests. Many people with disabilities who are blind have some vision (including those with light
sensitivity), very low or limited vision, or limited peripheral vision. Some visually impaired individuals have no light perception at all (World Health Organization, 2010). Understanding the characteristics of visually impaired users helps to determine the accessibility barriers as they interact with Web-based systems.

Heuristics ease the identification and prioritisation of characteristics for specific disabled groups. For visually impaired users, a missing text description of an image is a barrier (W3C, 2018). Moreover, visually impaired individuals use assistive software screen readers to interact with Web-based systems; thus, for example, if an image does not have descriptive text, they cannot access that image. Figure 1 illustrates the heuristically determined priority characteristics that should be applied for visually impaired users while they interact with images, video, voice and text. For example, images must be transferred to descriptive text; then, a screen reader can read it or print it as Braille code.

*Figure 1. Heuristic priorities based on the characteristics of visually impaired users.*
Accessibility of University Websites

Accessibility evaluation is a vital equity step in assessing the effectiveness of online learning materials for students with disabilities. In an empirical study (Alahmadi & Drew, 2016), researchers assessed the websites of 60 top universities globally and in the Oceania and Arab region. They found 30,944 (37%) homepage errors in 180 evaluated pages. The study indicated no significant improvement in the accessibility of university websites between 2005 and 2015. Additionally, no significant difference in accessibility was found among top-ranking universities in developed or developing countries (Patra & Dash, 2017; Ringlaben, Bray & Packard, 2014; Zap & Montgomerie, 2013).

Educational Web-based information systems advance academic success among users with disabilities as long as the systems are designed for accessibility. Online courses provide enhanced solutions for students who experience barriers to attending traditional courses because of sensory or physical disabilities. Of all users with disabilities, visually impaired individuals are the most strongly affected by inaccessible educational systems (Paciello, 2000).

Fichten, Jorgensen, Havel and Barile (2006) demonstrated that most students with disabilities that they surveyed indicated that they need adaptive assistive technologies, such as screen readers and voice recognition software (VRS), to effectively interact with a university Web-based system. Visually impaired users typically rely on screen reader software (e.g., Jaws) based on text-to-speech techniques (TTSs), VRS (e.g., Dragon Naturally Speaking), and Braille note-taking devices and keyboards when interacting with university Web-based systems. A screen reader is characterised by a simple mechanism that scans a screen for text and then audibly reads the content for a user to hear. Screen readers offer accessibility solutions and provide visually impaired users a sense of independence, but similar to other programs, they also suffer from certain limitations. For example, screen
readers can only read text; they cannot read other media content, such as images or Flash animations. If a descriptive text for an image is not available or incorrect, then the screen reader cannot convey the image content to the user (Crow, 2008). Understanding current image accessibility problems may lead to better understanding of challenges among visually impaired users, help to develop a solution, and increase educators or developer awareness.

Current Web-based university systems can benefit from evaluating image accessibility with respect to visually impaired users’ characteristics and needs (Rodriguez-Ascaso, Boticario, Finat & Petrie, 2017). To ensure accessibility, developers should also take into account the requirements for descriptive text of images for visually impaired users to effectively access images as well as the possible impact of image accessibility on learning and study for visually impaired users when evaluating an entire Web-based university system.

Multi-Method Accessibility Evaluation

Using a multi-method approach to evaluation is the best way to measure accessibility and design development plans for Web-based systems, such as university websites, because accessibility is a complex, multi-faceted issue. A single method cannot guarantee improvement in accessibility rates, as indicated in many studies (Masri & Luján-Mora, 2011) that underscored the essentiality of combining approaches to achieve excellent results. Other studies (Gómez-Martínez et al., 2015) showed that using experimental methods and user-centred design tests is a unique direction in determining and rectifying the most critical problems faced by disabled users as they interact with Web-based systems.

Human assessment, which involves subjective and objective evaluation, is a consistent component of all accessibility evaluation methods. It enables efficient probing into a specific component of accessibility barriers in specific system functions. One way of obtaining reliable results is to gain an overview of the accessibility status of numerous Web-based systems through
manual evaluation by evaluators or users; however, this approach is often excessively time-consuming and costly (Bühler, Heck, Perlick, Nietzio, & Ulltveit-Moe, 2006). Human experts are highly accurate at evaluating accessibility and may use automated evaluation tools (AETs) only as supportive methods.

AETs present advantages in terms of productivity and rely on heuristics to detect guideline violations (Brajnik, 2008). The drawback of these tools is that many fail to effectively evaluate the accuracy of the correspondence between descriptive text for images and the images’ complexity and purpose. They are also unable to satisfy the mandatory requirements for Web 2.0 applications because they exhibit restricted crawling capabilities, some evaluate only static-generated HTML content, and they fail to verify dynamically created document object model elements that are critical to rich Internet applications (Velasco, Denev, Stegemann, & Mohamad, 2008; Watanabe, Fortes, & Dias, 2017). Human and AET evaluations are performed on the basis of accessibility standards.

Many accessibility standards, like WCAG 2.1 (W3C, 2018), BITV 1.0 (Bundesministerium, 2011), Stanca Act (Parliament, 2004), and Section 508 (U.S. Department of Justice, 2016), require descriptive text for nontext elements such as images. Table 1 provides a summary for the standard checkpoint/guideline numbers related to the criterion which “all image elements have an alt attribute”. Furthermore, Section 508 provides guidelines that require long descriptive text for complex images (Section 508[a]: Text Equivalents, Checkpoint ID 3) and state that all nondecorative images must have descriptive text (Section 508[a]: Text Equivalents, Checkpoint ID 4), essential images should not have spacer descriptive text (Section 508[a]: Text Equivalents, Checkpoint ID 5), and descriptive text for all images must contain all text in the image unless the image text is decorative or appears elsewhere in text in the web page (Section 508[a]: Text Equivalents, Checkpoint ID 11).
http://dx.doi.org/10.17411/jacces.v8i2.167

<table>
<thead>
<tr>
<th>Standards</th>
<th>Guidelines</th>
<th>Checkpoint no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCAG 2.1</td>
<td>1.1 Text Alternatives</td>
<td>Success Criteria 1.1.1</td>
</tr>
<tr>
<td>BITV1.0</td>
<td>Group Level 1</td>
<td>Checkpoint 1.1</td>
</tr>
<tr>
<td>Section 508</td>
<td>A-text equivalents</td>
<td>Checkpoint ID 1</td>
</tr>
<tr>
<td>Stanca Act</td>
<td>Text Equivalents</td>
<td>Requirement 3</td>
</tr>
</tbody>
</table>

Providing descriptive text for media content improves accessibility (W3C, 2018), but this is effective only if the text is readily available and highly accurate. Alahmadi and Drew (2016, 2017a) found that failure to provide descriptive text for nontext elements, including images, is a serious accessibility error. This finding was confirmed by feedback from visually impaired users, who believed that such text is lacking from current Web-based systems (Alalhmadi, 2017a). Web localisers can bridge the knowledge gap and provide high-quality text alternatives when developers combine specialised and general Web accessibility evaluation tools (Vázquez, 2015). Splendiani and Ribera (2014) showed that a primary solution to image accessibility problems is the inclusion of alternative text through the use of decision trees. Multi-method evaluations of descriptive text for images are also expected to drive the discovery of website shortcomings that prevent the provision of accessible images.

**Image Accessibility: Descriptive Text for Visually Impaired Users**

Individuals describe objects through spoken, written, or typed language. A considerable amount of this language describes all objects in our lives, especially those that are visually based, such as images and videos. This language is likely a wealthy source of information about visual objects as...
well as methods for how individuals build natural language to describe visual objects (Kulkarni et al., 2013). Necessarily, then, a description of an image should contain a sufficient number of characters to highlight the principal image features. A complex image will require longer descriptions that reflect the main idea carried by the image. This requirement indicates a relationship between the number of characters in a description and the complexity of an image.

Descriptive text for images is necessary for visually impaired users (Connor, 2012). A simple textual description is not enough to convey the correct meaning of a graphic (Fitzpatrick, Godfrey, & Sorge, 2017). Automatically or human-generated descriptive text should lead to high-quality and accurate descriptions that reflect the key features of images. A deficiency in this regard diminishes the effectiveness of Web-based university systems (to which our model was applied). For instance, when an educator uploads a complex diagram, uses only two words to describe it, and neglects in-text explanations, visually impaired users will experience difficulty in understanding such important learning content.

Web-based university systems are characterised by a variety of images with equally varying purposes (Rice, 2012). An example is an image intended to deliver learning content. Each type of image needs a specific method of description, depending on the image’s purpose; there are complex or simple images, and some images are used as learning content, while other images are informative. HTML 5 (Connor, 2012; W3C, 2018) provides the necessary attributes and elements to add descriptive text for an image based on its purpose and type. Alt attributes are widely used to add alternative (description) text for nontext elements. In-text description is another method of describing an image with appropriate text in the paragraphs around the image on the web page. A null alt attribute adds a null value instead of text in the alt attribute. Table 2 provides a summary a summary of these attributes and its usages.
### Table 2. Summary of HTML Attributes and Elements to Describe Images (W3C, 2018).

<table>
<thead>
<tr>
<th>Image category</th>
<th>Example</th>
<th>Attribute/elements</th>
<th>Function</th>
<th>Number of characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex</td>
<td>Diagrams, graphs, maps, and</td>
<td>Longdesc attribute, area attributes, and figure elements</td>
<td>Add long text descriptions</td>
<td>More than 100</td>
</tr>
<tr>
<td>Simple</td>
<td>Informative image</td>
<td>Alt attribute</td>
<td>Add short text descriptions</td>
<td>Less than or equal 100</td>
</tr>
<tr>
<td>Simple</td>
<td>Decorative images</td>
<td>Null alt attribute value</td>
<td>Add a null value instead of a text description</td>
<td>Null value only</td>
</tr>
<tr>
<td>Simple</td>
<td>Functional images</td>
<td>Alt attribute</td>
<td>Add short text descriptions</td>
<td>Less than or equal 100</td>
</tr>
</tbody>
</table>

Moreover, diagrams, graphs, maps, and charts, which are considered complex images and used as learning content on Web-based university systems, necessitate long descriptions (more than 100 characters) that are placed under the images by using longdesc or area attributes, as well as figure elements. Also, MathML, for instance, is used to convert mathematical formulas into text in the absence of an in-text description or alt attributes (Connor, 2012; W3C, 2018). Text images, such as a scanned book chapter, require equivalent text files. Nonlearning images, such as those related to school administration, student accommodation, and alumni records, do not contain learning content. Examples include campus maps, images of boards of directors and related staff hierarchies, and diagrams of university...
pathways. Nonlearning images also need textual equivalents, but the impact of these images on visually impaired students is less than that of learning images.

More image types are available under W3C classifications (W3C, 2018). Examples include images used to label information, such as telephone icons and file formats (Figure 3), which usually require only one to two words of description. Other examples include images used to supplement information, such as a picture of a set of books (Figure 4) placed next to a textual announcement of exam periods, or images reflecting emotion, such as those featuring triumphant student faces. These images need short text descriptions, probably around 10 words (W3C, 2018). Decorative images, such as a partial rendering of a page design or text link, provide appeal to a web page. These images can be described using a null alt attribute value. Finally, functional images, such as logos and icons, require descriptive text that accurately conveys the function represented by each image (around five words; W3C, 2018).

Figure 2. Example of a mathematical formula image.

\[ E(r) = 4\varepsilon\left(\frac{\sigma^{12}}{r} - \frac{\sigma^{6}}{r}\right) \]

Figure 3. Example of an informative image conveying file format.
Research Methods

This study has a combined quantitative and qualitative design; the methods adopted were experimental strategies that involved human and automated evaluations (Creswell, 2013). This section explains the image accessibility checkpoints and rules that were formulated, provides an overview of human expert- and AET-based evaluations, and describes the data sampling and collection methods used in this study.

Sampling Method

Examining all web pages against all evaluation criteria is generally impractical (Nietzio, Strobbe, & Velleman, 2008). In this research, many foundational steps were implemented before sample pages from the evaluated systems were chosen. The first step was defining the evaluation goals, and the second was determining the system’s features and functions. The third step involved highlighting the characteristics and needs of the target disabled users, and the fourth entailed determining the types and effects of content found in the selected systems. Finally, the types of pages that affected the accessibility of the Web-based systems to the target groups were determined and prioritised to formulate solutions. After the
foundational process, a sample of pages from the selected websites was evaluated. Uniform random sampling is necessary for replicable evaluation that enables synchronous or asynchronous comparisons. The sampling process is usually based on an ad hoc procedure, such as page type selection, random walk, and uniform random sampling (Brajnik & Lomuscio, 2007). The choice of sampling method affects the metric design, which should consider the size and complexity of a website during evaluation (Parmanto & Zeng, 2005). If a system is considerably large and complex, the system is highly likely to receive a low accessibility score.

The complexity of Web-based university systems, which contain thousands of pages that comprise many images, can decrease their accessibility. To address this issue, we evaluated both complex and simple images; usually, learning images are complex, and informative images are simple. We also formulated evaluation rules (Section 3.3) to guarantee the optimal judgement of whether an image is accessible or inaccessible.

In this study, we evaluated 120 web pages that included 120 images. In our main research project, we categorised web pages into four categories—video, image, document, and general web pages—based on a published evaluation model (Alahmadi & Drew, 2017b). We evaluated 265 document web pages to examine accessibility problems in all of the document files, 120 web pages that included 120 videos, and 1,000 general web pages to test all general accessibility problems. A total of 1,505 web pages were evaluated in our main project. Based on our assumptions, around 12% of web pages have image accessibility problems. This assumption came from using the Google Search Console tool, the Google search engine, sitemaps to generate the number of all web pages in one Web-based university system, the number of images (excluding decorative or functional images) published on the same Web-based university system, and the number of document files and videos. We found that in the chosen Web-based university systems, 12% of all web pages had images.
Dataset

The examined dataset contained 120 evaluated images published on 120 web pages of 64 Web-based university systems (including 38 Australian universities and 26 Saudi universities). This study was conducted in Australia and supported by the Saudi Ministry of Education. In 2016, among these universities, 9 ranked at the top 100 in the world, 14 ranked above the top 500 in the world, and 41 placed below the top 500, as determined from QS University Rankings (Dobrota, Bulajic, Bornmann, & Jeremic, 2016). The main language used on the web pages was English. The web pages, which contained complex or simple images, were randomly selected for the evaluations. Of the images examined, 37 were embedded on LMSs, such as course content web pages, and 83 were embedded on university web pages, such as online help and library pages; 66 were learning images, and 54 were non-learning images; 92 were considered complex, and 28 were regarded as simple (according to W3C definitions). We excluded decorative or functional images from the evaluations.

Image Accessibility Checkpoints and Rules

In our study, the learning images examined were published on web pages that delivered learning materials, such as course content and library pages. These images constituted a crucial part of the web pages’ content. The absence of descriptive text for such images means that part of the learning materials is also missing, thereby affecting the performance of visually impaired students/users. Most of the images are graphs, diagrams, and charts, which are regarded as complex images (W3C, 2018). As previously stated, complex images may require descriptions that are longer than 100 characters; such descriptions can be provided through the use of HTML5 attributes or in-text explanations (i.e., the text surrounding an image on a web page). Nonlearning materials that provide general information to students/users (e.g., administration and alumni web pages) are as important...
as learning images, except that they do not directly affect the achievements of visually impaired student/users in courses.

*Figure 5: Main considerations when creating accurate descriptive text for an image (W3C, 2018).*

Image accessibility for visually impaired users necessitates accurate descriptive text that aligns with the images’ purpose and complexity (Connor, 2012; W3C, 2018; U.S. Department of Justice, 2016). For this reason, we formulated fundamental rules based on HTML5 attributes and element features as well as WCAG 2.1 and Section 508 standards that guide the evaluation of images embedded in educational Web-based systems. These rules are as follows:

- If an image is complex, then a long descriptive text (or in-text explanation) is required.
- If a descriptive text (or in-text explanation) is long, then the minimum number of characters required is >100 characters (W3C, 2018).
- If a long descriptive text (or in-text explanation) is used, then the accuracy of the description must be ensured.
- If an image is simple, then a short descriptive text is required.
- If a description is short, then the minimum number of characters required is ≤100 characters (W3C, 2018).
- If a short description is used, then the accuracy of the description must be ensured.
- If an image is decorative, then a null attribute can be used.
The rules above (W3C, 2018, 2018; U.S. Department of Justice, 2016) serve as the basic requirements for ensuring that any image is accessible to visually impaired users. Images that are intended to deliver learning content must also have accurate, meaningful, and high-quality text descriptions that are based on course outlines, resources, and strategies.

There are important considerations when evaluating the meaningfulness of descriptive text for an image:

- The descriptive text must describe an image in the form of complete sentences with accurate language, rather than unconnected words (Wu, Wieland, Farivar, & Schiller, 2017).
- It cannot contain acronyms or symbols without definitions (W3C, 2018).
- It must describe image features in the text similarly to human visual descriptions (Vedantam, Lawrence Zitnick, & Parikh, 2015).
- It must describe at least three main layers of statistical diagrams: a top-level summary, the major component layers, and single component explorations (Fitzpatrick et al., 2017).
- It must highlight most of the critical image factors: compositional, semantic, and context factors (Berg et al., 2012).
- It must describe all hierarchical chart components, cascading down from the top to the other components of the chart (W3C, 2018).
Evaluations by Human Experts and Automated Tools

The WCAG 2.1 guidelines (W3C, 2018) were created under the assumption that developers perform expert evaluations in the process of complying with the requirements of accessibility checkpoints. Consequently, the evaluation and accuracy of developed accessible web pages are directly associated with a developer’s level of experience. Understanding known and potential accessibility problems is expected to enable developers to create Web-based systems that are characterised by enhanced accessibility and data quality.

Bailey, Pearson, and Gkatzidou (2014) compared the reliability of accessibility evaluations carried out by novices versus experts. The authors found that expert evaluations were 76% reliable, whereas novice assessments were 65% reliable. The study partially supports the importance of expert evaluations in resolving the shortcomings of AETs. Expertise is accorded high priority in accessibility evaluations; it is paramount to the successful verification and application of WCAG-based techniques because the expert involvement ensures thorough knowledge of accessibility issues (Yesilada, Brajnik, & Harper, 2009). Nonetheless, an expert evaluation may be inaccurate or miss accessibility problems in web page analysis and thereby cause ambiguity in human evaluation (Brajnik, Yesilada, & Harper, 2010). It should, therefore, be supported by AETs to reduce the possibility of inaccuracies and lessen the time and effort involved in the evaluation. The use of AETs can be carried out as a second stage of the assessment.

In this study, we used the AChecker (AChecker Adaptive Technology Resource Centre, 2017) automated evaluation tool for many reasons. Firstly, we can extract the evaluation outcome as a PDF or CSV file to add to the research data as a reference. Also, we can check against many guidelines, such as WCAG 2, Section 508, BITV 1.0, and the Stanca Act. AChecker categorises the problems as known, likely, and potential problems. Finally, we can easily go to the checkpoints and the HTML line code that relate to image accessibility problems. In this study, we used AChecker against WCAG 2.0 standard level AAA to test all of the web pages that contained the
evaluated images and recorded the problems that were found. Also, we used AChecker as the second stage after human evaluation to validate the accuracy of the human decision and find the causes of any dissimilarities between the two methods. AChecker might provide false positive or false negative outcomes. However, this issue does not impact our study because human evaluation was the main evaluation method used on all of the images.

In image accessibility evaluations conducted by a human expert, the expert is obligated to ensure that the images and their purposes are fully accessible despite their complexity; all of the rules presented in Section 3.1 apply. The human expert not only examines the availability of descriptive text (or an in-text explanation) but also ensures the text’s quality and accuracy required by the image’s purpose and complexity. In this study, we evaluated each image on the basis of predetermined accessibility variables (Appendix A).

As shown in Appendix A, a number of known, likely, and potential accessibility problems were extracted using AChecker. These problems demonstrate the accessibility issues encountered on the web pages that contained the evaluated images. Appendix A also provides the HTML5 attributes (alt, longdesc, title, src, class, figure element, area) that are typically used as the bases in assessing image accessibility. The availability and accuracy of text descriptions are intended to be used as references in examining the quality of text descriptions and the number of words in such explanations. Complexity variables can be used to understand the purpose of an image, and the image category can be employed to determine whether an image is a learning or nonlearning image. Descriptive text and the words used in titles are designed to enable an analysis of the descriptive text’s quality through measurements of the words’ meaningfulness. Because we applied our method to university websites, an important requirement was to determine the web page type and system type and whether the images on the university websites included those from library web pages or, in particular, from LMSs.
Evaluation Process Flowchart

The evaluation process is based on testing and analysing HTML 5 code for image attributes and elements. Figures 6 through 9 illustrate examples of HTML 5 image HTML code.

Figure 6. Example of an alt attribute.

```html
<img src="http://www.acu.edu.au/__data/assets/image/0006/1308975/Units_Results_LEO.png" alt="displays student portal and links to LED" class="img-responsive"/>
```

Figure 7. Example of image HTML code without an alt attribute.

```html
<img border="0" width="514" height="348" src="experiment_2_2016_files/image006.jpg" v:shapes="Picture_x0020_52"/>
```

Figure 8. Example of image HTML code with alt and title attributes.

```html
<img src="/sites/schools-engagement.cdu.edu.au/files/images/my_career_match_how_it_works.jpg" alt="How it Works" title="How it Works" width="633" height="127"/>
```

Figure 9. Example of image HTML code with an alt attribute including a NULL value.

```html
<img style="WIDTH: 584px; FLOAT: left; HEIGHT: 347px; MARGIN-LEFT: 10px; MARGIN-RIGHT:10px" alt="NULL" src="/sites/default/files/EBP%20diagram%20Nov%202013.png"/>
```

Figure 10 provides an overview of the human and AET evaluation process for one image published on one web page. Some essential variables were recorded before the evaluation process, such as the complexity level and image category.
Figure 10. Human and AET evaluation process for one image.

Findings of the Human Evaluations

After generating the dataset, we analysed the images on the basis of commonly used descriptive statistics. The accessibility problems discussed in this section cannot be identified by AETs. The human evaluation was directed towards the availability of HTML5 attributes and elements, with emphasis on alt and title attributes, in-text descriptions, and the accuracy of the descriptive text.

Availability of HTML5 Image Attributes and Elements

HTML5 image attributes and elements provide accessibility solutions (Connor, 2012). The more adequate the number of attributes and elements used, the clearer the information delivered by a screen reader to visually impaired users (W3C, 2018). As stated earlier, complex images need long descriptions. Our dataset comprised 92 complex images, for which the longdesc attribute was never used. It also contained numerous diagrams, charts, and maps, yet figure and area elements were also disregarded. The src attribute was used for 119 images, and the class attribute was used for 18 images (Figure 11). The title attribute is important because it shows users an image’s title before its description. Among the evaluated images, only 15 were given a title attribute. These findings indicated a lack of HTML5 attributes and elements that deliver significant information to visually impaired users and limit the number of accessibility problems encountered in screen readers.

Figure 11. Summary of uses of HTML 5 elements and attributes.
Availability of Descriptive Text (Alt Attribute and In-Text Descriptions)

The alt attribute was used for 99 of the images, suggesting that the developers or publishers were aware of the importance of using it when uploading an image. These images may pass the accessibility tests of AETs and may be considered accessible. In-text descriptions were used for 57 images, and in-text descriptions were used in conjunction with the alt attribute for 46 images. In-text descriptions without the alt attribute were used for 11 images, and the alt attribute without in-text descriptions was used for 53 images. Neither in-text descriptions nor the alt attribute was applied to 10 images (Table 3).

<table>
<thead>
<tr>
<th>ALT</th>
<th>NO</th>
<th>YES</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-text description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>11</td>
<td>46</td>
<td>57</td>
</tr>
<tr>
<td>No</td>
<td>10</td>
<td>53</td>
<td>63</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>99</td>
<td>120</td>
</tr>
</tbody>
</table>

As can be seen, most of the evaluated images were accorded an alt attribute and in-text descriptions. Such solutions do not suffer from availability issues. However, the correct usage of image accessibility solutions may influence accessibility but still generate an inaccessible image. To address this problem, we directed the human evaluation not only towards the availability of the alt attribute (as with automated tools) or in-text descriptions but also towards the quality and accuracy of the descriptive text.
The accuracy of the Descriptive Text

The accuracy of the descriptive text reflects its quality. The incorrect use of null values and the inadequacy of word counts influence descriptive text’s quality. Among the evaluated images, null values were used 38 times, but the combination of the alt attribute with null values should be used only for decorative images. Given that the developers or publishers of the examined websites used null values for 38 complex or informative images, we can say that these were inaccessible on the basis of W3C definitions and the rules formulated in this study. Table 4 shows the descriptive statistics for the word counts of the descriptive texts written with the alt attribute. The minimum number of words was 1, and the maximum was 316. The mean word count was only 4 words per image. The sum of all of the words in the descriptive texts for all 120 images was 514 words. Only one image had a 316-word descriptive text. Based on W3C definitions and the study’s rules, out of the 92 complex images evaluated in this research, only one was accessible.

Table 4. Descriptive Statistics for Word Counts of the Descriptive Texts.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Sum</th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word count</td>
<td>1</td>
<td>316</td>
<td>514</td>
<td>4.28</td>
<td>28.843</td>
</tr>
</tbody>
</table>

A descriptive text with a low word count can be considered accessible if the words used to describe an image are highly meaningful. Correspondingly, we evaluated the meaningfulness of each image’s descriptive text with respect to the purpose and complexity. We also analysed all of the texts with the alt attribute, thus generating seven categories of descriptive text. The findings revealed that out of 120 images, only one was fully accessible in terms of meaningfulness with respect to the image’s purpose and complexity. Table 5 provides details on the text categories.

### Table 5. Descriptive Text Categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of descriptive texts in each category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete descriptive text</td>
<td>1</td>
<td>At the first level, we have Federation College and Secondary School. In Federation College at FAST, which . . . diploma.</td>
</tr>
<tr>
<td>Incomplete descriptive text</td>
<td>6</td>
<td>Blended learning model with the high-quality resources section highlighted . . .</td>
</tr>
<tr>
<td>Meaningless</td>
<td>15</td>
<td>Mapwagga</td>
</tr>
<tr>
<td>Link</td>
<td>8</td>
<td><a href="https://www.kfu.edu.sa/PW2D.jpg">https://www.kfu.edu.sa/PW2D.jpg</a></td>
</tr>
<tr>
<td>Title</td>
<td>17</td>
<td>A process flow chart</td>
</tr>
<tr>
<td>File name</td>
<td>15</td>
<td>Ann16-3.png</td>
</tr>
<tr>
<td>Symbol</td>
<td>1</td>
<td>. . .</td>
</tr>
</tbody>
</table>

As previously indicated, the title attribute was used for 15 images. We evaluated each text on the basis of the title attribute usage and found that in eight of the 15 images, the text used for the alt attribute as descriptive text was used for the title attribute as the image’s title. We also identified six meaningless title texts, one file name, one link, two null values, and only five complete title texts. Thus, the alt and title attributes were inaccurately used to deliver the images’ information.

In-text descriptions also function as accessibility solutions, with long descriptions usually accompanying complex images. We assessed each image and its corresponding in-text description to determine the text’s accuracy and completeness. The findings showed that 57 images had in-text descriptions. We found 14 complete in-text descriptions and 43 incomplete ones, indicating that the descriptions insufficiently described the images. The W3C (2017) recommendations indicated that the alt attribute should be employed with in-text descriptions to help visually impaired users understand those descriptions. However, this recommendation was not followed for the in-text descriptions accompanying the 14 images.

To sum up, out of the 120 images evaluated, 15 were accessible, 14 came with in-text descriptions, and one had the alt attribute (Figure 12).

**Figure 12. Image accessibility outcomes from human evaluations.**

Findings of the Automated Evaluation Tool

The AET revealed 5,641 known accessibility problems in the 120 evaluated web pages; the maximum number of problems on a web page was 520, and the minimum was 0. A total of 564 likely accessibility problems were identified, with the maximum being 97 and the minimum being 0. The
evaluation found 52,982 potential accessibility problems, with the maximum being 1,633 and the minimum being 0. AChecker also determined that all of the images suffered from accessibility problems, which all corresponded to violations of the WCAG 2.0 standards, particularly Guideline 1.1 (“Text Alternatives”) and Success Criterion 1.1.1, which require the provision of descriptive text for all non-text elements. The top nine image problems identified in the AET evaluation (AChecker) are listed in descending order below:

- Image elements required long descriptions.
- Image elements were missing the alt attribute.
- An alt text was not empty for an image that may have been decorative.
- An alt text did not convey the same information as what the image expressed.
- An embed element was missing a noembed element.
- An image had a title attribute, but the image may have been decorative.
- An image used for an input element was missing an alt text.
- An image used as an anchor was missing a valid alt text.
- An image’s alt text was lengthy.

None of the evaluated images passed the AET evaluation. Some images had known problems, some images had likely problems, and others had potential problems. Some evaluators considered an image accessible if it had likely or potential problems only. However, images with these problems should be checked by humans to determine if they are accessible. In this study, we checked all problem types, which resulted in zero accessible images for the AET. There were 61 images that did not have known problems but had likely or potential problems. Some evaluators may have considered these images accessible even though they contained serious accessibility problems. Overall, this finding highlights the importance of considering the
involvement of human experts in evaluating, designing evaluation rules, and using advanced evaluation methods, such as data mining.

Discussion

Evaluating image accessibility is an effective step that opens up opportunities to develop practical solutions to ensure equal image access by visually impaired users. Thus, multi-method evaluation of image accessibility provides a vital contribution to achieving favourable results. Human and automated evaluations can work together to assess image accessibility synergistically.

Human evaluation is an essential method for discovering the details of image accessibility problems. Usually, these problems cannot be identified by automated evaluation alone. After recognising whether an image is complex or simple and determining its purpose, the human evaluation process examines all HTML image attributes and elements. In our findings, these attributes and elements saw limited use. Many reasons can limit the use of HTML image attributes and elements. An author’s or developer’s knowledge regarding accessibility can affect the quality of the accessible image they create (Moreno, Castillo, Williams & Menez, 2015). Moreover, regulating Web accessibility is not an internationally recognised practice (Cleary & Maurer, 2017). Most organisations do not apply accessibility standards, and 75% of them do not enforce accessibility evaluations (Moreno et al., 2015). Organisations use various authoring or content tools (W3C, 2018). A noticeable shortcoming of most of these tools is that they do not facilitate the creation of accessible content and therefore do not provide intelligent features. Innovations like generating automated alternative text, text to speech (TTS), and speech to text (STT) may complete the vision of an adaptive and accessible Web-based system for all users.

Automated evaluation tools provide a list of all image barriers on a webpage. However, when comparing human and automated evaluation, we found that...
the evaluation outcomes of automated tools might negatively impact Web accessibility in two situations. Firstly, there is a high chance that a web page will be judged as having zero problems even with an inaccessible image uploaded on it. Secondly, there is a high chance that the opposite will occur, when an image is considered inaccessible even when it provides a quality in-text description. This situation leads to no accessible images being found by automated evaluation. In this study, the human evaluation results showed that 15 images were accessible, and their descriptive texts considered the images’ complexity and purpose.

Efficiently applying evaluations of HTML image attributes and elements will improve accessibility outcomes. A set of regulations or rules imposed by an organisation may be effective in generating developer and author awareness, resulting in practical improvements. It is vital that accessibility is considered part of the development of any Web-based system. As part of that consideration, adopting a multi-method evaluation process will improve the detection of image accessibility problems.

**Conclusion and Future Work**

The evaluation method developed in this study is applicable not only to university websites but also to other institutions using Web-based systems and organisations for which effective interaction between online platforms and disabled users is essential. The findings underscored the necessity of probing into images’ accessibility and ensuring that system modifications positively affect individual users. The human and automated evaluations trialled here provided insight into how image accessibility problems can be identified and understood. Human evaluation is essential, particularly in cases in which the quality of descriptive text needs to be tested.

The study methods and findings revealed a number of potentially productive directions for future work. We intend to evaluate image accessibility through data mining, with a particular focus on the use of classification algorithms,
to compare the results of human, AET, and mining-based evaluations. We plan to illuminate the outcomes of each method and determine how and why such outcomes vary across approaches. The rules developed in this study will be used in the data mining to classify each image case as accessible or inaccessible.

One-size-fits-all user interfaces and content can be a source of inequity, but methodical differentiation diminishes the likelihood that users with disabilities will benefit from image content (Gajos, 2014). The accessibility of images published on Web-based systems, especially university websites, should thus be given more attention. The availability and accuracy of descriptive texts and their compatibility with the image’s complexity and purpose should be ensured for all images because the presence of high-quality descriptive texts improves image accessibility (W3C, 2018). The findings derived from this work showed evidence of a lack of awareness by developers/authors, thus negatively affecting image accessibility on the evaluated sites. Apart from increasing developer/author awareness, adaptive approaches can be used to optimise accessibility to users with different disabilities. Employing adaptive content that is tailored to the abilities and characteristics of visually impaired users enhances accessibility when these individuals interact with a Web-based system (Stephanidis et al., 1998). A proper application of this principle is reflected in Wu et al.’s (2017) use of automatic alt-text (AAT) in Facebook. AAT is a technique that applies vision technologies to recognise faces, objects, and themes in images and thereby generates image alt-texts for screen reader users. AAT demonstrates that artificial intelligence techniques can be used to enhance the online experiences of visually impaired users (Wu et al., 2017).
References


Appendices

Appendix A

Table of Research Variables From the Dataset.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webpage type</td>
<td>{1, Home page} {2, Course content} {3, Administration} {4, Online help} . . .</td>
</tr>
<tr>
<td>System type</td>
<td>{1, University website} {2, LMS}</td>
</tr>
<tr>
<td>Image category</td>
<td>{1, Learning} {2, Non-learning}</td>
</tr>
<tr>
<td>Complexity level</td>
<td>{1, Complex} {2, Simple}</td>
</tr>
<tr>
<td>Alt-text availability</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>Alt text accuracy</td>
<td>{1, High} {2, Low}</td>
</tr>
<tr>
<td>In-text availability</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>In-text accuracy</td>
<td>{1, High} {2, Low}</td>
</tr>
<tr>
<td>Known problems</td>
<td>Total number of known problems</td>
</tr>
<tr>
<td>Likely problems</td>
<td>Total number of likely problems</td>
</tr>
<tr>
<td>Potential problems</td>
<td>Total number of potential problems</td>
</tr>
<tr>
<td>Alt</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>Alt text</td>
<td>Descriptive text (words)</td>
</tr>
<tr>
<td>Number of words in text</td>
<td>Total number of words</td>
</tr>
<tr>
<td>Longdesc</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>Title</td>
<td>{1, Yes} {2, No}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title text</td>
<td>Title text (words)</td>
</tr>
<tr>
<td>Src</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>Class</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>Figure element</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>Area</td>
<td>{1, Yes} {2, No}</td>
</tr>
<tr>
<td>Accessibility outcome</td>
<td>{1,Accessible} {2,Inaccessible}</td>
</tr>
</tbody>
</table>