

# Methodology Draft:"Enhancing ePub Accessibility: Automating Compliance with WCAG 2.1 Standards"

Submitted on: 19-4-2024

Merijn van der Leek  
merijn.van.der.leek@student.uva.nl  
University of Amsterdam  
Amsterdam, The Netherlands

Dr. Yixian Shen  
M.y.shen@uva.nl  
University of Amsterdam  
Amsterdam, The Netherlands

## 1 INTRODUCTION

Electronic publications (ePubs) have significantly widened access to a large amount of literary content. However, the potential of ePubs is not fully realized for all users, especially those who rely on assistive technologies. This thesis addresses two large accessibility issues that impair the user experience for visually impaired individuals and those requiring language-specific support: the lack of descriptive alt text for images and the presence of incorrect or missing lang attributes. These deficiencies not only hinder the engagement of a significant user base but are also against the international accessibility standards stated in the Web Content Accessibility Guidelines (WCAG) 2.1 [1].

The importance of accessible digital content cannot be overstated, as it plays a crucial role in promoting inclusivity and equality. The World Wide Web Consortium (W3C), which develops the WCAG standards, advocates that accessible design enhances usability for all users, not just those with disabilities [1]. But unfortunately there is a pressing issue: a substantial portion of digital content remains inaccessible, alienating a large segment of potential users. In response, this project harnesses advanced machine learning techniques and open-source tools to automate and enhance ePub accessibility significantly. By focusing on generating precise alt text for images with AltText.ai and rectifying lang attributes, this methodology aims not only to mitigate immediate accessibility barriers but also to contribute substantially to the broader objective of making digital literacy resources universally accessible.

## 2 TOOL AND TECHNOLOGY SELECTION

**AltText.ai** was specifically chosen for its prowess in generating alt text due to its cutting-edge algorithms in deep learning, computer vision, and artificial intelligence, tailored specifically for automating alt text generation. This tool excels in producing high-quality, contextually aware captions, eliminating the need for human intervention and thereby reducing errors and enhancing efficiency [3].

**Langdetect** was selected for its simplicity and effectiveness, capable of distinguishing between 55 different languages with high precision. This tool is essential for ensuring that ePub content is accessible and comprehensible in users' native languages, adhering to international accessibility standards [2].

### 2.1 Practical Application and Output Examples

To illustrate the possibilities of this tool, we have applied AltText.ai to a randomly selected image from our dataset and generated the following descriptive alt text:

alt= Illustrated book cover titled "de letterruil" featuring a glowing light bulb inside a wire cage lamp on top of a scrolled paper, with a green vignette background.



Figure 1: The image used for alt text generation by AltText.ai.

This alt text successfully captures the essential elements of the image, providing a clear and concise description that can be used by screen-reading software to display the visual content to visually impaired users. The image, titled *de letterruil*, features a simplistic design which AltText.ai has accurately described, showcasing the model's ability to recognize and articulate the details in an image that are important to understand its context and content. In this case it even rightfully detected the image as book cover.

This example also serves as a benchmark for the quality of alt text we aim to produce for the entire dataset. By providing such

illustrative outputs, we can set expectations for the standard of alt text that would enable a rich reading experience for all users, regardless of their visual abilities.

The generated alt text is programmatically embedded into the ePub files using Python scripts that interface with the ePub structure. This automated process ensures that the alt text is associated with the correct images within the eBooks, following the rules for how HTML and ePub documents should be written.

### 3 DATA COLLECTION AND PREPARATION

Data collection was conducted with a lot of attention to ensure a diverse and comprehensive dataset. We accessed a variety of free eBook libraries that allow direct downloads, providing a broad spectrum of content. Additionally, we used tools such as a GitHub downloader for bulk acquisitions from Library Genesis, a platform known for its extensive repository of literature. This method enabled the efficient gathering of a large volume of texts, reducing manual effort while ensuring a rich collection of materials.

Moreover, we incorporated selections from Standard Ebooks, a source known for its high-quality, carefully curated content. This not only enhanced the quality of our dataset but also provided us with texts that are typically well-formatted and less prone to the common errors found in less careful prepared eBooks. The combination of these sources ensured that our dataset included a wide variety of genres, from children’s books that often contain vivid illustrations to complex cookbooks with detailed photographic content, and novels that offer a range of narrative-driven images. This diversity is crucial, as it allows AltText.ai to be tested across different types of visual content, simulating real-world applications where the context and subject matter of images vary significantly.

In total, approximately 300 eBooks were selected to form a comprehensive base for our statistical analysis, chosen to ensure a wide representation of linguistic diversity and graphical content. During the selection process, we actively screened for and removed any files that were corrupted or unreadable. These are common issues that can compromise the integrity of language detection and alt text generation. Additionally, any files that contained nonsensical strings of characters, which are irrelevant for language processing, were excluded to maintain the dataset’s quality.

### 4 BASELINE SETTING AND VALIDATION

Baseline settings were carefully chosen to establish a standard against which the enhancements brought by our project could be measured. For language detection, we set a straightforward baseline: all languages that were initially not detected were defaulted to English. This baseline serves as a control to highlight the effectiveness of Langdetect in accurately identifying a wide range of languages, demonstrating possible improvements over simplistic assumptions.

For alt text generation, we established a minimal baseline by setting all undetected alt texts to the phrase ‘Photo without alt text’. This baseline is critical for assessing the qualitative improvements achieved by AltText.ai. By comparing the output against this baseline, we can quantitatively measure the enhancement in quality and detail that our tool provides.

## 5 MODEL APPLICATION AND INTEGRATION

Integrating AltText.ai and Langdetect into our ePub processing workflow involved sophisticated software engineering efforts due to the complex nature of ePub files. These files are essentially compressed archives containing various types of data, including text, images, and metadata, each requiring different handling techniques. Our approach involved carefully configuring these tools to parse and extract the necessary data within this environment.

To navigate and decode the multi-layered structure of ePub files effectively, we utilized various Python scripts. These scripts leverage powerful libraries like BeautifulSoup for parsing HTML content and zip handling techniques for unpacking the compressed files. This setup allows us to accurately target and process the specific content within eBooks that is relevant for our language detection and alt text generation tasks, ensuring that our tools are applied efficiently and effectively.

## 6 CHALLENGES AND LIMITATIONS

The integration of advanced tools into a complex file format presented significant technical challenges, particularly in accurately parsing and manipulating the structures within ePub files. The decision has also been made to produce alt text exclusively in English for its practicality. However, future studies might benefit from generating alt text in the book’s language, leveraging the capabilities of our selected model. Furthermore, the project’s scope is currently limited to addressing only two specific WCAG errors, with numerous other accessibility concerns remaining unaddressed. This focus highlights the need for broader enhancements to fully realize the accessibility potential of eBooks.

## 7 FUTURE DIRECTIONS AND SCALABILITY

Looking forward, the project aims to expand its scope to address a more comprehensive range of WCAG/ACE errors, moving towards the goal of fully accessible eBooks. The scalability of the current approach is promising, validated by the diversity of the dataset and the adaptability of the implemented tools, which are designed to be applicable to a vast number of books. However, the financial implications of scaling, especially given the costs associated with services like AltText.ai, present a significant challenge that needs to be managed as the project evolves.

## 8 EVALUATION METRICS

The project employs a dual approach to evaluation, designed to assess both the technical efficacy and the practical utility of the enhancements we have implemented in ePub accessibility. This evaluation framework is crucial to ensure that our project’s outputs meet both the standards of technical accuracy and the functional needs of end-users.

**Quantitative Evaluation** For the quantitative aspect of our evaluation, we focus primarily on assessing the accuracy of the language detection tool, Langdetect. To do this effectively, we have compiled a test set of eBooks that already possess known language tags. These books serve as a benchmark, allowing us to measure the precision of Langdetect by comparing its language identification outcomes against the pre-established tags. This method provides a

clear, quantifiable metric to measure the accuracy of our language detection process. By using a test set with known variables, we can conduct an evaluation that gives us reliable data on the tool's performance.

**Qualitative Evaluation** On the qualitative side, we engage domain experts to review the alt text generated by AltText.ai. Their role is to evaluate the contextuality and relevance of the alt text in relation to the images they describe. This involves a detailed analysis of how effectively the alt text replaces the visual information with textual descriptions, ensuring that the content is both accurate and sufficiently descriptive to serve as a stand-alone representation of the visuals.

The experts review a range of outputs from AltText.ai, assessing the descriptive quality, clarity, and informativeness of the alt text. They provide feedback on how well the alt text would communicate the essential information of the images to a person who cannot see them. This feedback is crucial for refining the alt text generation process, guiding further improvements to ensure that the alt texts are not only technically accurate but also genuinely useful and engaging for all users.

## REFERENCES

- [1] World Wide Web Consortium (W3C), *Web Content Accessibility Guidelines (WCAG) 2.1*, 2021, <https://www.w3.org/TR/WCAG21/>.
- [2] Mimino666, *Langdetect: Language Detection Library for Python*, n.d., <https://pypi.org/project/langdetect/>.
- [3] AltText.ai, *About Us*, n.d., <https://alttext.ai/about>.