# Accessibility of machine learning and generative AI

## Abstract

TBA

## Status of this document

This is a draft collection of relevant information related to cross-disability accessibility guidance of how developments in machine learning and generative Artificial Intelligence (AI) clearly bears an impact on web accessibility standards and processes. Given the rapid changes in the consumption and development of AI design, this is intended to be a starting point to group the accessibility implications of machine learning and generative AI technologies.

Editors: Scott Hollier, (Janina, Jason, Raja)

For: RQTF

Date: 24 June 2024

Table of Contents

[Accessibility of machine learning and generative AI 1](#_Toc170205334)

[Abstract 1](#_Toc170205335)

[Status of this document 1](#_Toc170205336)

[1.0 Introduction 3](#_Toc170205337)

[1.1 Definitions 3](#_Toc170205338)

[1.1.1 Artificial intelligence 3](#_Toc170205339)

[1.1.2 Machine learning 3](#_Toc170205340)

[1.1.3 Generative AI 3](#_Toc170205341)

[1.2 Accessibility context 3](#_Toc170205342)

[2.0 AI accessibility use cases 4](#_Toc170205343)

[2.1 Relevance of current standards and guidance 4](#_Toc170205344)

[2.2 Alternative text for images 4](#_Toc170205345)

[2.3 Automatic Speech Recognition for captioning 6](#_Toc170205346)

[2.4 Plain language 7](#_Toc170205347)

[2.5 Plain Language detection 8](#_Toc170205348)

[2.5 Colour contrast 8](#_Toc170205349)

[2.7 Heading structure 8](#_Toc170205350)

[2.8 R adjustment of visual spacing 9](#_Toc170205351)

[2.9 Link purpose . 9](#_Toc170205352)

[2.10 Sign language 9](#_Toc170205353)

[3.0 Evaluation tools 9](#_Toc170205354)

[4.0 Accessibility Overlays 10](#_Toc170205355)

[Reference List 11](#_Toc170205356)

# 1.0 Introduction

## 1.1 Definitions

### 1.1.1 Artificial intelligence

Artificial Intelligence refers to a machine or computer program's ability to think and learn without explicit instructions. The terminology “AI” in this context is generally thought to have been established by John McCarthy in 1955 (Xu et al., 2021). Although AI is used in a variety of disciplines, it is currently most relevant in machine learning and consumer applications known as generative AI.

### 1.1.2 Machine learning

Machine learning represents a field of study within the AI domain that has a focus towards statistical algorithm - capable of learning from data and performing tasks without specific instructions. This form of AI tends to focus on determinations and predictions.

### 1.1.3 Generative AI

While generative AI is generally classified as a subset of machine learning, its difference lies within its specific focus on creating new data based on specific human creative inputs such as text, images and audio (Zhihan Lv, 2023).

## 1.2 Accessibility context

The rapid increase in the use of machine learning and the specific focus on generative AI represents potential benefits and challenges in the context of Web accessibility provision. As online generative AI platforms such as ‘ChatGPT’ continue to offer consumers the unrestricted ability to create text and images, including video and audio from a variety of inputs, it is important to consider how accessible these outputs will be presented, and if machine learning algorithm may address broader accessibility issues in everyday tasks. For example, a web browser can use machine learning and generative AI to update content in real-time. This could be done through means such as the addition of alternative text to images which do not have one, and recognising text that should be a heading but is not marked correctly.

These real-time changes would affect existing web standard such as the Web Content Accessibility Guidelines (WCAG), User Agent Accessibility Guidelines (UAAG) and the Authoring Tool Accessibility Guidelines (ATAG), assisting these standards in potentially addressing accessibility issues observed in various other accessibility related notes and resources. Furthermore, with the recent development by several manufacturers in launching computers and mobile devices with built-in Neural Processing Units (NPUs) to specifically support machine learning features on local devices, such as through the reduction of latency for live captioning. The potential for AI benefits within the web accessibility space could be profound.

In addition, machine learning also raises the prospect for existing tools to be used in identifying and remediating accessibility issues. For instance, the proficiency of automated testing tools and accessibility overlays could be significantly increased to provide greater accessibility support and guidance to relevant developers and designers. The WCAG Success Criteria could also be improved in the sense of how much can be checked using AI, while increasing the applicability of automated remediation.

However, the emerging nature of AI in its machine learning, and more specifically, its generative context, has made it difficult for users and developers alike to accurately perceive how effective automated accessibility tools are, and the degree of reliability they may offer in future directions. As such, this guidance is designed to provide some assistance in how machine learning is currently providing support, and its current level of effectiveness.

# 2.0 AI accessibility use cases

## 2.1 Relevance of current standards and guidance

The W3C Accessibility Initiative (WAI) consists of three guidelines capable of assisting in the creation of accessible content that may potentially be supported by AI features. These guidelines include standards relating to web content, user agents and authoring tools.

The [Web Content Accessibility Guidelines (WCAG) 2.2](https://www.w3.org/TR/WCAG22/) features several guidelines and success criteria (SC) that are potentially being addressed, such as SC 1.1.1 with the provision of automated alternative text, and SC 1.2.4 with the provision of time-based media through automated live captioning.

In addition, the [User Agent Accessibility (UAAG) Guidelines 2.0](https://www.w3.org/TR/ATAG20/) could suggest that user agents, like web browsers with generative AI, could make real-time adjustments during browsing sessions, such as fixing color contrast issues and improving poor heading structures through the interpretation of codes.

Furthermore, the [Authoring Tool Accessibility Guidelines (ATAG) 2.0](https://www.w3.org/TR/UAAG20/) could also offer support, particularly in Part B, in which the creation of accessible content is of particular importance. For example, authoring tools that have automatically generated alternative text could support the creation of accessible content.

## 2.2 Alternative text for images

There currently exists a number of machine learning-based tools that have been integrated into popular social media platforms, alongside authoring tools, that are equipped to create an automated alternative text description based on machine learning algorithms that scan and determine the contents of visual materials, such as an image. Until recently, this automated process was considered to hold high inaccuracy to the point where its utility was questioned (Leotta et al., 2022). Recent developments have improved automated alternative text accuracy, but criticism persists due to limitations in providing detail and recognising the importance of relevant data.



Figure 1 – a coloured bar graph representing the favourite colour of children in a survey (Twinkl, n.d.)

A good example can be seen in a popular image used to illustrate data (Twinkl, n.d.). The image features a classic bar graph of responses from children clarifying their favourite colour, in which yellow has been found the achieve the highest result with 9 responses. While an appropriate alternative text for the image should endeavour to capture the significant points of the graph with detail, such as its drawn intention - the information organising its X and Y axis, as well as the resulting data, the automated alternative text simply describes this image as “a graph with different coloured bars”. While technically accurate, this information lacks depth to convey important technical details from the graph.



Figure 2 - An image provided by the James Webb Space Telescope

A second example is shown through an image from the James Webb Space Telescope. As all images publicly released include automated alternative text, the alternative text for Figure 2 was compared to that of a manually created alternative text. The former reads the description, “The image is divided horizontally by an undulating line between a cloudscape forming a nebula along the bottom portion and a comparatively clear upper portion”, while the latter states: “Speckled across both portions is a starfield, showing innumerable stars of many sizes. The smallest of these are small, distant, and faint points of light. The largest of these appear larger, closer, brighter, and more fully resolved with 8-point diffraction spikes. The upper portion of the image is bluish and has wispy translucent cloudlike streaks rising from the nebula below.”

Upon observation, the automated alternative text presents a simplified iteration of the image, using the brief narration, “a nebula in space with stars”. As such, once again, this comparison supports that, while automated alternative text provided by machine learning is representative of the image being studied and could assist in delivering a basic and minimised understanding of an image, it does not have the ability to incorporate the orientation of detail required to capture the essence of the image.

Although machine learning techniques embedded in authoring tools and other platforms may provide some information, generative AI platforms that are able to create images, videos and other visual media content based on text input tend not to provide automated alternative text. Hence, this would make it difficult for people who are blind or have low vision to attain a meaningful interpretation of these AI-generated outputs.

## 2.3 Automatic Speech Recognition for captioning

A common alternative accessibility support is the provision of live captioning through the features of Automated Speech Recognition (ASR) processes, whereby generative AI is employed to sample speech and convert them into captions as the same language, or subtitles in a different language.

Until recently, the accuracy and validity of ASR techniques in the same language were considered as ineffective practices as the translation quality was poor due to an extra machine learning step that attempted to convert ASR outputs further. Currently however, the provision of captions has become increasingly reliable, with an approximate success rate of 85%. Furthermore, its ability to translate in real time has improved substantially (Millett, 2021).

While ASR captions may provide merits when serving the role of a complementary tool alongside curated content, it is still generally perceived by the Deaf community as lacking in accuracy, thereby preventing ASR from being considered as a truly beneficial feature. Other issues associated with ASR may include delays in processing time, as most contemporary solutions require devices to be connected for the generative AI features to occur online, the lack of grammar and punctuation, the lack of block captioning associated with pre-recorded materials, and the necessity for good quality audio in a quiet environment in order for ASR to optimise its output.

However, that is not to discredit the fact that some of these issues are beginning to be corrected. With the introduction of NPUs into consumer devices, ASR captions can be rendered on devices rather than online. Early testing suggests that this significantly improves the speed of captions being presented, as well as provides enhanced auto-correction features in an instance where ASR misunderstands a word or phrase.

While these advancements do not resolve all issues of ASR, and the notion remains that this process is still not adequately accessible as a standalone feature, they demonstrate that current generative AI processes are capable of potentially contributing to accessibility improvements.

ASR technologies may also be applied to audio-only material in which pre-recorded content could be provided to a generative AI process to output a transcript. While the removal of real-time ASR may increase reliability to a point where some specialist writing tools could be applicable to utilise this function, issues of accuracy, spelling, grammar, punctuation and errors persist.

## 2.4 Plain language

Current demonstrations of generative AI often employ examples of language being translated or simplified to showcase the potential of these mechanisms. While the ability for popular AI platforms to convert language based on specific word count, rhyming pattern and subject matter are often viewed as successful outcomes, converting complex language content to the level of lower secondary reading abilities often require human intervention.

Key concepts including the need for common words, the definition of words, and the removal of double negatives are exemplary demonstrations of how generative AI may achieve some positive effects. However, other plain language concepts such as the conversion of text into literal language, understanding different grammar tenses, and addressing text with nested clauses highlight some current limitations of AI. For example, a translation of the popular poem, “Mary Had a Little Lamb”, into a non-English language, and then back to English observes the error of “it’s fleece was white as snow” being converted to “it snowed sheep hair”. While the other lines in the poem were largely accurately translated, the issue of literal language and the shortcoming of generative AI in understanding regional contexts and language structures remains.

 ----

##

## 2.5 Automated Language detection

Within the WCAG 2.2 standard, Guideline 3.1 explains the necessity to globally define language on a page, as well as when there is a change of language in parts. This would then allow assistive technologies to understand the language change and adjust its language selection accordingly. A potential benefit to machine learning could be the improved ability in identifying the type of language used. This would then enable assistive technologies such as screen readers to select that language if supported.

Although there are few examples in a web context where the language of content is defined by machine learning processes, rather than coded directly, the ability for assistive technologies to immediately comprehend the language being presented could provide a considerable benefit.

## 2.6 Colour contrast

Another area where machine learning could identify an issue lies in the remediation of colour contrast issues as they are recognised. Automated testing tools at present are capable of detecting some colour contrast issues (Almeida and Duarte, 2020), hence, the possibility of real-time remediation could be likely in the case that a detected text colour contrast is below the 4.5:1 colour contrast ratio for foreground, background and text colour, or the 3:1 colour contrast ratio for user interface elements. With generative AI processes, these elements could be made more contrasting through colour adjustments, while largely preserving the colours intended by the original author.

## 2.7 Heading structure

Currently, the use of elements such as bold text to look like a heading without the provision of a programmatically determined heading structure is a significant accessibility issue that remains in web content. With a generative AI feature, web pages or app screens can visually identify headings and their nested content, utilising current technologies. While there is presently no awareness of its potential implementation, the remediation of heading structure represents a significant opportunity to provide improvements in readability and navigation should generative AI be capable of effectively addressing this issue.

## 2.8 Adjustment of visual spacing

At present, some web browsers offer the proficiency to structurally reorder web content to optimise the aspect of readability. As enhancements in machine learning continues, it is likely that such features will continue to improve and confront accessibility concerns relating to text, word, and line spacing for the purpose of better supporting people with a cognitive and print disability.

## 2.9 Link purpose Presently, a widespread issue existing within online channels may be attributed to the use of non-descriptive links such as “click here” or “read more”, as identified in WCAG 2.2 SC 2.4.4 Link Purpose. While it is currently necessary to create links with a descriptive text to remediate the issue, generative AI represents an opportunity for issues to be addressed in real-time. This could be accomplished through the employment of AI following a non-descriptive link to its source, and distinguishing its content before resolving the link for users in a manner that is more indicative of what the link represents.

## 2.10 Sign language

The utility of sign language in relation to WCAG presently stands as a Level AAA success criterion that is not as widely implemented within policy-related legislative frameworks, with compliancy up to Level AA appearing to be the most adopted. Thus, support for sign language is limited within time-based media such as online videos.

However, thanks to generative AI, it is now feasible to provide language translation services from text or symbol-based languages to sign language. Some websites and movie apps already offer a limited version of this capability. On the other hand, while this observation shows promise for the future use of sign language, the effectiveness of generative AI in translating sign language, coupled with the diverse localised variations in sign language across countries, poses challenges for providing a fully effective automated solution at this time.

## 3.0 Evaluation tools

At present, there are several automated testing tools available based on coding assessments within websites, apps and documents. Such tools are often employed to identify issues of non-compliance with the WCAG standard, additionally providing subsequent guidance in remediation. Some of these tools are free of charge but limited in functionality, while others provide enterprise-level remediation guidance, and is capable of monitoring web content in real time.

Although these tools are generally considered a helpful companion for remediation, they currently possess flaws; some notable issues include acquiring different results between the use of different tools (Ismailova & Inal, 2022), difficulty in determining if the tool has identified a specific issue, or simply noting that the issue requires review. These issues follow research that has constantly indicated automated tools bear the restricted capacity of executing low-level coverage in checking and recommending remediations, with the remaining requiring some form of intervention (Vigo et al., 2013).

While automated testing tools have been largely based on scripting language such as JavaScript to manually review and report on code, deep machine learning is likely to improve these tools to a considerable degree going forward, allowing the application of generative AI processes to more clearly identified issues. For instance, most tools may accurately determine if alternative text is available for images but cannot determine the effectiveness of present alternative text. With the inclusion of generative AI, improvements are likely to be offered in what automated tools can evaluate, such as the quality of captions, descriptive links and correcting other issues mentioned previously such as the use of language and headings. Although there is currently little evidence of generative AI featuring in such tools, it is very much possible that testing and evaluation of web content will improve along with the rest of the rapidly evolving generative AI content.

## 4.0 Accessibility Overlays

The rapid increase of accessibility overlays on websites has been viewed as rather controversial by people with disability. While these tools could be useful for individuals unfamiliar with assistive technologies that are built into computing and mobile devices, critics of overlays point to the tools being marketed as an accessibility solution, thus causing the code to interrupt the use of more developed assistive technologies such as screen readers (Morris, 2022). Furthermore, these overlay features carry the tendency to be limited in functionality as compared to tools installed in an operating system.

However, the promise of generative AI may be able to address the criticism that such tools lack functionality. An accessibility overlay capable of utilising generative AI functionality may be able to provide increased real-time support in overcoming accessibility issues or improving its interpretation of content, such as for images, language and page structure. Although these tools are currently promoted as a collection of accessibility features somewhat independent from the content, the applicability of an overlay that contributes accessibility improvements is similar to the use of AI chatbots and other prompting mechanisms, thereby suggesting this may prove to be another area where generative AI could introduce improvements.

## Reference List

Almeida, R., & Duarte, C. M. (2020). Analysis of automated contrast checking tools. *Proceedings of the 17th International Web for All Conference*, 1–4. https://doi.org/10.1145/3371300.3383348

Ismailova, R., & Inal, Y. (2022). Comparison of Online Accessibility Evaluation Tools: An Analysis of Tool Effectiveness. *IEEE Access*, *10*, 58233–58239. https://doi.org/10.1109/access.2022.3179375

Leotta, M., Mori, F., & Ribaudo, M. (2022). Evaluating the effectiveness of automatic image captioning for web accessibility. *Universal Access in the Information Society*, 1–21. https://doi.org/10.1007/s10209-022-00906-7

Millett, P. (2021). *Accuracy of Speech-to-Text Captioning for Students Who are Deaf or Hard of Hearing* (pp. 1–13). https://www.edaud.org/journal/2021/1-article-21.pdf

Morris, A. (2022, July 13). For Blind Internet Users, the Fix Can Be Worse Than the Flaws. *The New York Times*. https://www.nytimes.com/2022/07/13/technology/ai-web-accessibility.html

Twinkl. (n.d.). *What is Bar Chart?* [Image ]. Retrieved June 24, 2024, from https://www.twinkl.de/teaching-wiki/bar-chart

Vigo, M., Brown, J., & Conway, V. (2013). Benchmarking web accessibility evaluation tools. *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility - W4A ’13*, 1–10. https://doi.org/10.1145/2461121.2461124

Xu, Y., Wang, Q., An, Z., Wang, F., Zhang, L., Wu, Y., Dong, F., Qiu, C.-W., Liu, X., Qiu, J., Hua, K., Su, W., Xu, H., Han, Y., Cao, X., Liu, E., Fu, C., Yin, Z., Liu, M., & Roepman, R. (2021). Artificial Intelligence: a Powerful Paradigm for Scientific Research. *The Innovation*, *2*(4). Sciencedirect. https://doi.org/10.1016/j.xinn.2021.100179

Zhihan Lv. (2023). Generative Artificial Intelligence in the Metaverse Era. *Cognitive Robotics*, *3*, 208–217. https://doi.org/10.1016/j.cogr.2023.06.001