Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 8, No. 4, August, 2024

Generative Artificial Intelligence and Web Accessibility: Towards an Inclusive and Sustainable Future

Patricia Acosta-Vargas [1*], Belén Salvador-Acosta [2], Sylvia Novillo-Villegas [1], Demetrios Sarantis [3], Luis Salvador-Ullauri [4]

[1] *Intelligent and Interactive Systems Laboratory, Universidad de Las Américas, Quito 170125, Ecuador.*

[2] *Facultad de Medicina, Universidad de Las Américas, Quito 170125, Ecuador.*

[3] *Operating Unit on Policy-Driven Electronic Governance, United Nations University, Guimaraes 4810445, Portugal.*

[4] *Software and Computing Systems, University of Alicante, 03690 Alicante, Spain.*

Abstract

This study examines the accessibility of Generative Artificial Intelligence (AI) tools for people with disabilities, using WCAG 2.2 success criteria as a reference. Significant accessibility issues were identified in the evaluated applications, highlighting barriers mainly affecting disabled users. Integrating accessibility considerations from the beginning of application development and adopting a proactive approach are emphasized. Although challenges are faced, such as the shortage of inclusive training data and opacity in AI decision-making, the need to continue addressing various aspects of accessibility in the field of generative AI tools is acknowledged. These efforts are based on regulatory compliance and ethical principles to ensure equal societal participation, regardless of individual abilities. The fundamental role of accessibility in realizing this vision is highlighted, aligning with the United Nations Sustainable Development Goals, particularly those related to equality, education, innovation, and inclusion. Improving accessibility meets regulatory requirements and contributes to a broader global agenda for a more equitable and sustainable future.

Keywords:

Generative Artificial Intelligence; Web Accessibility; WCAG 2.2; Sustainable Development Goals; Inclusive; Sustainable Future.

Article History:

Received:	06	April	2024
Revised:	10	July	2024
Accepted:	18	July	2024
Published:	01	August	2024

1- Introduction

Recent research has placed a spotlight on the impact of generative artificial intelligence (AI) on individuals with disabilities, acknowledging both its potential benefits and drawbacks. Despite recognizing the broad implications of generative AI within this demographic, a thorough exploration of its specific consequences remains largely untapped. In this context, Glazko et al. [1] embarked on a three-month autoethnographic study to delve into the utilization of generative AI for personal and professional purposes within a research team comprising individuals with and without disabilities. The study uncovered a spectrum of potential accessibility applications for generative AI while also flagging concerns regarding verifiability, training data, capability, and the validity of particular claims.

Generative AI encompasses computational techniques that generate novel and meaningful content from training data, spanning text, images, or audio. Prominent examples such as Dall-E 2, GPT-4, and Copilot have revolutionized how we collaborate and communicate, with applications ranging from artistic endeavors to intelligent question-and-answer

* **CONTACT**: patricia.acosta@udla.edu.ec

DOI: http://dx.doi.org/10.28991/ESJ-2024-08-04-021

systems. Within information systems, generative AI introduces opportunities and challenges, including implications for accessibility, as posited by Feuerriegel et al. [2].

Models like ChatGPT and Bing Chat have progressively integrated into our daily routines, potentially reshaping how we access, process, and produce information. However, Roegiest & Pinkosova [3] underscored concerns regarding the accessibility of such systems, particularly for individuals facing literacy challenges. The research underscores the imperative to investigate the readability and accessibility barriers posed by generative information systems, especially for those with literacy difficulties, as datasets used in training often produce content comprehensible only at a college level, thereby excluding a significant segment of the population. These findings emphasize the necessity of ensuring that advancements in this domain benefit all, with proposed enhancements to information search and retrieval mechanisms pivotal in bolstering accessibility and efficiency, consequently broadening impact and enhancing user engagement.

In a parallel narrative, Pascual & Osorio [4] contended that despite the United Nations lauding the internet as a catalyst for human rights and a cornerstone for achieving Sustainable Development Goals, the persisting digital divide complicates efforts to mitigate social inequalities. They identify low web accessibility as a critical factor hindering specific populations from accessing vast online resources. Proposing web accessibility as an underexplored facet of individual and societal inequality, they advocate for increased academic inquiry into how artificial intelligence can enhance web accessibility, positing that such improvements could bridge the digital divide and mitigate global inequities.

Furthering this discourse, Silva et al. [5] described the development and evaluation of a tool automating the generation of navigation aids for screen readers, leveraging topicalization and tagging algorithms to craft headings and internal links. They assert that such tools address the demand for navigational aid design insights for screen reader users, highlighting the potential integration of natural language processing techniques, including generative AI, in assistive technologies. However, they also acknowledge limitations that necessitate future research attention.

In essence, web accessibility is a cornerstone of inclusion and sustainable development. The abilities or limitations of users should not impede access to web-based information and services, with the Web Content Accessibility Guidelines (WCAG) 2.2 [6] providing a foundational framework to uphold this principle.

As generative AI rapidly evolves, revolutionizing technology interaction and unlocking new vistas in automatic content creation and design, attendant challenges in accessibility, particularly for individuals with disabilities, loom large. This study evaluates the accessibility of 50 generative AI tools through the lens of the WCAG 2.2 [6] framework.

By investigating the impact of generative AI on web accessibility and assessing alignment with WCAG 2.2, this research seeks to foster sustainable and inclusive development. The evaluation will encompass four key facets: Perceptibility, Operability, Understandability, and Robustness. The comprehensive assessment aims to pinpoint areas for enhancing generative AI tools' accessibility and furnish developers with actionable insights to create more inclusive technologies aligned with the Sustainable Development Goals (SDGs) [7].

The document unfolds as follows: Section 2 delves into Background and Related Work, exploring the history, current state, and accessibility considerations of generative AI applications while identifying research gaps. Section 3 elucidates the Methodology employed, detailing tool selection, evaluation criteria, data collection, analysis methods, and study limitations. Section 4 presents the Results and Analysis of accessibility assessments conducted on the selected generative AI tools, analyzing compliance with WCAG 2.2 and delineating strengths and weaknesses regarding accessibility and inclusion. Section 5 delves into the Discussion, examining implications for generative AI applications and accessibility and exploring challenges in ensuring inclusivity. Finally, Section 6 encapsulates Conclusions, Limitations, and Future Work, summarizing key findings, outlining contributions to accessibility and inclusion, and suggesting avenues for future research endeavors.

2- Literature Review: Generative AI, Web Accessibility, Sustainability

Generative AI is a transformative force, enriching individual and organizational efficiency while bolstering overall productivity. Its capacity to automate various tasks, from email responses to data analysis, endows entities with unprecedented speed and precision, optimizing operations and decision-making processes [8, 9]. Real-time data analysis capabilities empower swift and informed decision-making, offering a competitive edge in today's dynamic landscape [9].

However, applications leveraging generative artificial intelligence pose distinctive design challenges as they become increasingly ingrained in mainstream applications. With this integration comes a pressing need for guidelines to craft user experiences that foster practical and secure utilization [10], which delineates six principles to shape user experiences with generative artificial intelligence applications, address the nuanced dynamics of these interactions, and propose novel interpretations and extensions of existing design paradigms.

Moreover, Nasser et al. [11] presented an integrated approach harnessing emerging technologies to furnish practical and accessible solutions and intelligent transportation services tailored to individuals with visual impairments. Proposing a framework underpinned by artificial intelligence and machine learning algorithms, aiming to enhance system efficiency and enable expedited decision-making and more responsive services to meet the needs of visually impaired individuals.

Participatory methods, notably co-design, emerge as pivotal strategies to navigate the challenges inherent in digital government for individuals with disabilities, as argued by van Toorn [12]. Emphasizing inclusion as an ongoing process within the digital government context, the study underscores the conceptual value of continuous engagement, positioning disability as a realm ripe for exploration with generative AI technologies while acknowledging associated limitations. The study emphasizes the pivotal role individuals with disabilities play in shaping technological advancements.

Generative AI's profound impact on customer experience surfaces as a notable contribution, as elucidated by Haleem et al. [9] and Perez Medina et al. [13]. By facilitating personalized interactions and streamlining the shopping journey, generative AI enhances customer satisfaction, liberating time for individuals to focus on critical endeavors while fortifying security measures in real-time [14, 15]. These multifaceted capabilities optimize operational efficiency and decision-making and cultivate a secure environment conducive to business and individual prosperity.

Figure 1 depicts the trend of generative AI applications over the past 12 months, sourced from Google Trends [16]. ChatGPT emerges as the most utilized application, experiencing growth in January 2024, followed by Bing and Gemini, demonstrating continued ascendance since January 2023.

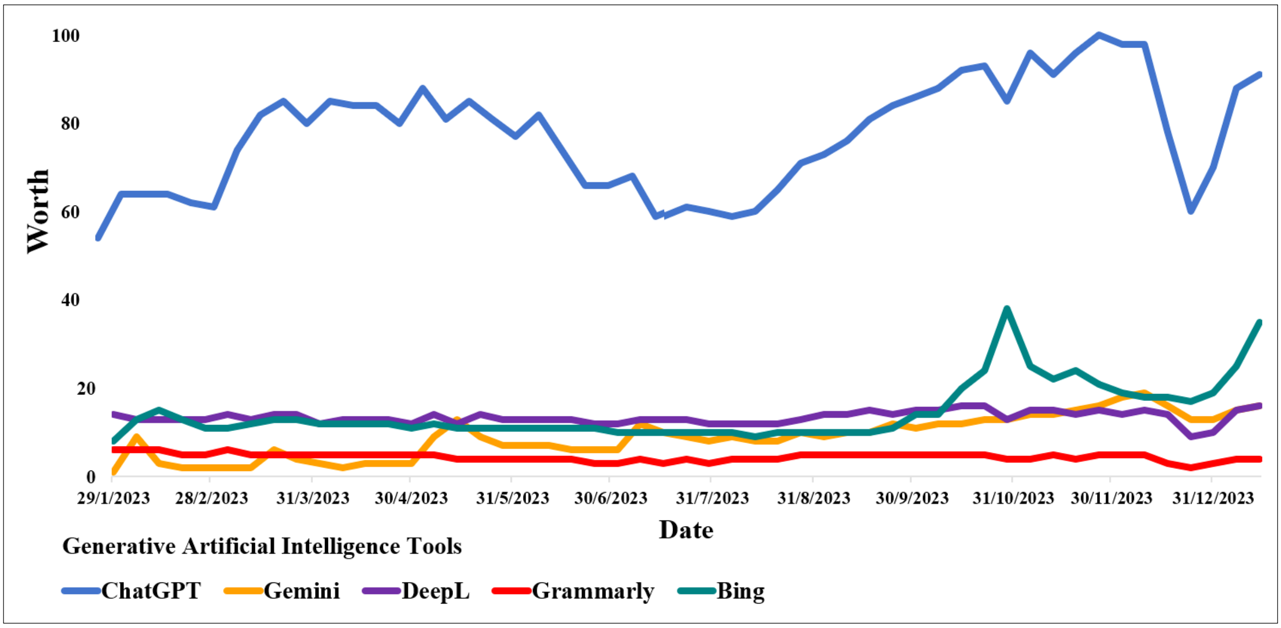

Figure 1. The trend of generative AI applications, data taken from Google Trends

Amidst this landscape, web accessibility emerges as a paramount concern, with more than 50 million accessibility errors detected across 1 million websites worldwide, as reported by the World Wide Web Consortium (W3C) [17, 18]. This alarming statistic underscores the imperative of adhering to guidelines such as WCAG 2.2 to ensure content accessibility. However, a significant proportion of websites fail to meet basic accessibility standards, posing barriers to individuals with disabilities. As home page complexity burgeons, strategies to mitigate accessibility impediments must be reevaluated, emphasizing prioritization of accessibility and adoption of streamlined development processes [19].

Generative AI applications are designed to adhere to WCAG 2.2 [6], striving for perceivability, operability, understandability, and robustness. These principles encompass efforts to ensure content accessibility for all users, regardless of disability or impairment [20, 21].

Integrating generative AI with WCAG 2.2 promises to bolster attaining the UN Sustainable Development Goals (SDGs) [3]. By promoting equal access to information and web services, generative AI can advance social participation and digital inclusion, which are crucial for fostering just and peaceful societies [22, 23]. Collaboration across sectors is imperative for developing and implementing generative AI solutions for web accessibility [24].

Nonetheless, integrating generative AI into web accessibility presents challenges, including algorithm biases and transparency issues that may adversely affect individuals with disabilities. Addressing these challenges necessitates regulatory frameworks ensuring ethical and responsible development and implementation of generative AI [25].

In conclusion, while the burgeoning volume of web content poses challenges to effective WCAG 2.2 implementation, generative AI emerges as a transformative tool poised to enhance web accessibility. Leveraging generative AI can drive a more inclusive and sustainable digital future, with collaboration and innovation serving as linchpins in this endeavor. This article aims to evaluate and explore the contributions of generative AI to web accessibility, shedding light on its impact on achieving the SDGs and offering recommendations for effective integration. Web accessibility is a fundamental right and a linchpin for sustainable development, and generative AI holds immense potential to advance this crucial agenda.

Generative AI applications [25] can assist developers in complying with both WCAG 2.1 [20] and WCAG 2.2 [6], thus representing a valuable tool for enhancing the productivity, innovation, and efficiency of any business. This article will evaluate the web accessibility of 50 generative AI applications. These applications not only provide an accessible interface for users to interact with web content but also can detect and correct content errors, personalize the user experience according to their preferences, and assist developers in optimizing web content for the best performance possible. In this accessibility analysis, we use the WAVE [26] with the plugin extension installed in the Google Chrome browser, which allows the evaluation of sites that require authentication.

Figure 2 depicts a graphical representation of a bibliographic neural network highlighting keywords linked to generative artificial intelligence, web accessibility, WCAG, inclusion, and sustainability.

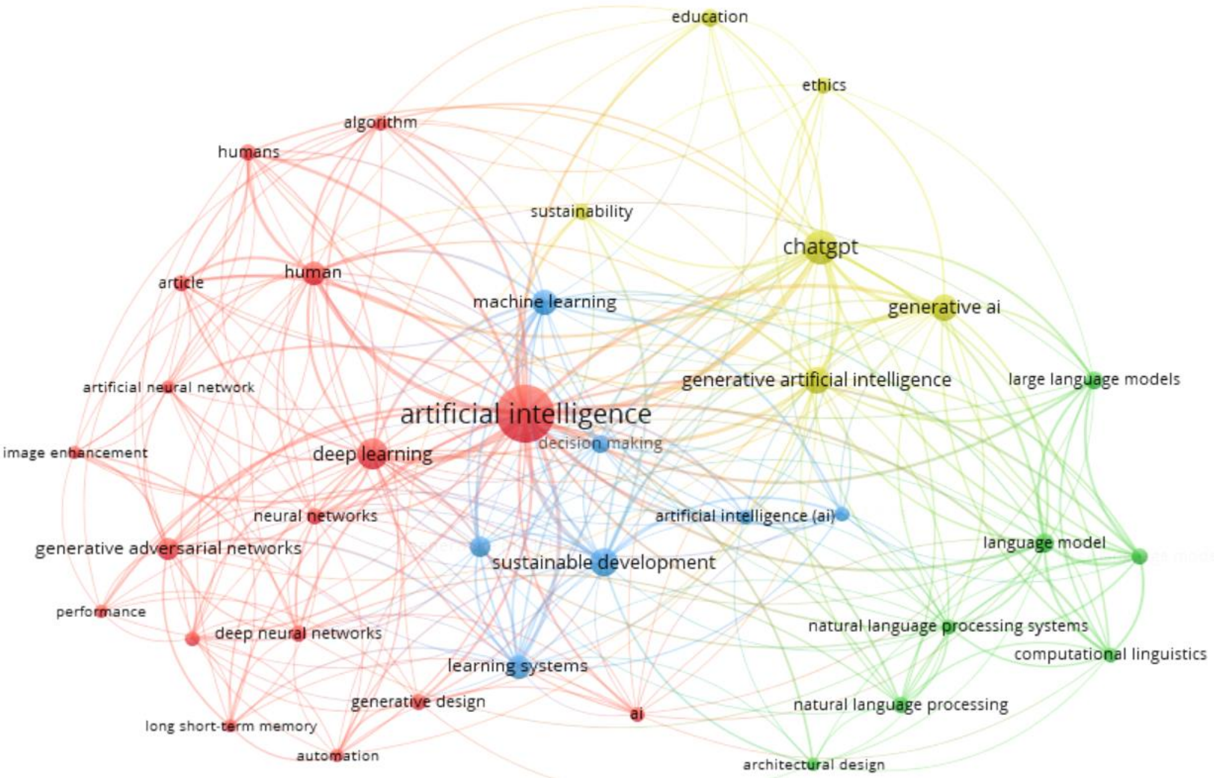

Figure 2. Presents a neural network of keywords organized into four groups

Figure 2 illustrates a framework of interconnected concepts derived from the analysis of 176 articles generated in VOSviewer [27] version 1.6.20, a tool that offers text-mining functionality to construct and visualize networks of co-occurring essential terms extracted from a set of scientific literature.

The figure presents 37 elements distributed across 4 groups interconnected by 301 links, indicating a robust interrelationship among the concepts. Each color in Figure 2 represents a group of elements or clusters, identified using the VOS algorithm to reveal strong relationships among elements. The data exhibits a well-defined cluster structure with related element groups.

The VOS algorithm [27] implemented in VosViewer is an ascending hierarchical clustering method based on similarities designed to identify groups in bibliometric networks, such as keywords, depending on the clusters' distances.

The similarity between each pair of clusters is calculated based on the strength of connections among the elements of each one; the co-occurrence between two elements defines the strength of the connection. Groups with the most significant similarities are merged into a larger one, incorporating elements from the original groups. This process repeats until a single group encompasses all network elements, representing the complete hierarchical clustering. The groups contain the following keywords:

Cluster 1: AI, algorithm, article, artificial intelligence, artificial neural network, automation, deep learning, deep neural networks, generative adversarial, generative design, human, humans, image enhancement, extended short-term, neural networks, performance.

Cluster 2: Architectural design, computational linguistics, language models, large language models, and natural language processes.

Cluster 3: Artificial intelligence (AI), decision making, generative model, learning systems, machine learning, sustainable development.

Cluster 4: ChatGPT, education, ethics, generative AI, generative artificial intelligence, sustainability.

From the analysis of the articles obtained from the search in Scopus, we observe that the previous study by Emenike & Emenike [28] argues that ChatGPT exemplifies how generative AI impacts academic integrity and assessment in chemistry education. Beyond teaching and learning, generative AI systems offer potential benefits and risks for students, teachers, and administrators, sparking discussions about leveraging capabilities while recognizing limitations and addressing equity and accessibility concerns.

The research argues that (Tiwary & Mahapatra [29]) millions of visually impaired people worldwide face online accessibility challenges, particularly in education and shopping. The study highlights issues with Indian university websites and proposes solutions, including AI-generated image descriptions, to ensure inclusivity and equal access for all users, regardless of their disabilities.

Draffan et al. [30] argues that automated web accessibility checkers have evolved to evaluate the barriers faced by users with disabilities, aligning with the WCAG proposed by the W3C. However, as web development progresses, questions arise about the reliability of these checks. The article explores integrating AI methods, such as image recognition and natural language processing, to improve accessibility assessment. Draffan et al. [30] argue that automated web accessibility checkers evaluate barriers for disabled users based on WCAG compliance. The increasing complexity of web code prompts updates to the guidelines, raising concerns about verifier reliability. The paper advocates the integration of AI, such as image recognition and natural language processing, to improve verifier accuracy and ensure effective barrier reduction strategies.

Acosta-Vargas et al. [21] argue that accessibility is critical to the success of e-commerce. One study evaluated 50 top-ranked e-commerce sites using modified WCAG-EM 1.0 and the WAVE tool. The errors found revealed that most contrast errors were common, emphasizing the need for improvement. Future work involves the development of AI-powered accessibility software.

Finally, Reddy [31] argue that generative AI offers transformative potential in clinical decision-making and patient care, particularly in healthcare. Despite ethical concerns, integrating generative AI requires meticulous planning and stakeholder engagement. Responsible governance ensures ethical and safe integration, improving the accessibility and quality of healthcare.

3- Material and Methods

Developing a methodology for integrating generative AI web accessibility involves a structured and collaborative approach. This section selects the automatic evaluation tool that helps the accessibility evaluation process. The researchers considered several parameters, such as accessibility principles, application of guidelines, level of accessibility, and features of the tools.

Based on previous studies by Acosta-Vargas et al. [32], we identified eight commonly used automated assessment tools, including AccessMonitor, Achecker, eXaminator, TAW [33], Tenon, WAVE [26, 34], Web Accessibility Checker, and Mauve++ Tool [32]. The tools were selected based on the ability to assess accessibility principles, accessibility level, and various features, such as load times, license, reporting, plugin, support and maintenance, functionality, portability, usability, connectivity, and security. The evaluation results concluded that the most suitable automatic tool to evaluate accessibility was WAVE; its advantages included that it was functional on the web with the browser extension and showed no problems with applications that required authentication. However, these tools are not definitive, and a more robust evaluation is required to assess accessibility comprehensively. Human evaluation is imperative to determine if an application meets accessibility standards, so manual review by accessibility experts with five years of experience is included in this research.

This research evaluates accessibility in 50 generative AI applications using a variation of the Website Accessibility Conformance Evaluation Methodology (WCAG-EM) 1.0 [35], comprising six phases described in Figure 3.

Phase 1: Selecting generative AI: At this stage, we select the generative AI technologies, which involve making strategic decisions about the model or approach we will use to meet the specific needs of the applications we are evaluating. During this process, we define fundamental requirements, such as the type of content to be generated, be it text, images, music, or other data.

Subsequently, we evaluate and select several generative AI models, highlighting some of the most recognized ones, such as GPT (Generative Pre-trained Transformer), VAE (Variational Autoencoder), and GAN (Generative Adversarial Network), each with its strengths and weaknesses. A crucial criterion in our selection process is the size of the generative AI model, which ranges from more compact models suitable for running on local devices to large-scale models requiring significant computational power in the cloud.

In the selection process, we prioritize those affordable generative AI technologies that do not demand excessive technical knowledge from the end user. We also carefully consider the model's performance evaluation, establishing clear metrics to measure success against the project objectives. In addition, we consider the model's ability to be upgraded and maintained over time.

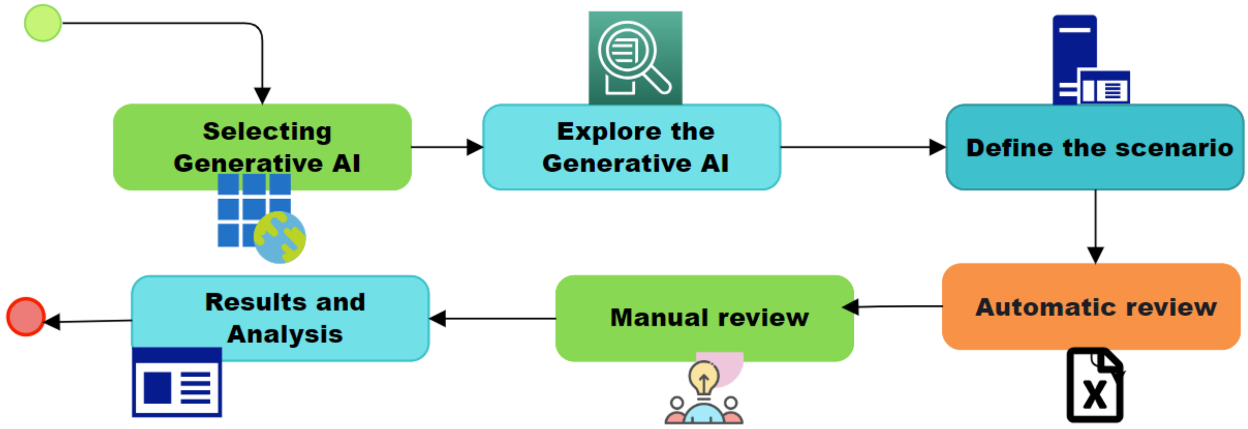

Figure 3. Combined approach to accessibility review

We took a random sample of 50 generative AI applications from the Synthesia site [36], which contains the most used applications arranged alphabetically by generative AI. Table 1 presents the tool name, the category to which it belongs, and the description. More details about the tools, the evaluation, and the accessibility analysis can be found in the Mendeley repository [37].

Table 1. A random sample of the 50 GAI applications

#	Category	Generative IA	Description
1	Image Generator Tool	Adobe Firefly	Realize an infinite range of image creations.
2	Image Generator Tool	AI Art	A text-to-image generator combined with a music generator.
3	Music Generator Tool	AIVA	AIVA is an AI music generator that creates background music for various use cases.
4	Character Generator Tool	Artflow.ai	This AI character generator allows it to create animated video stories by creating unique scenes, characters, and dialogs.
5	Design Tool	AutoDraw	This AI drawing generator suggests various shapes and symbols, can add text, change colors, and insert different shapes into the drawing.
6	Writing Tool	Bing	Bing is a web search engine from Microsoft that offers services such as image, video, map, and news search.
7	Music Generator Tool	Boomy	Online AI music generator that allows anyone to create their original music in seconds and share it with the world.
8	Character Generator Tool	Character.ai	Character.AI employs artificial intelligence to train realistic chatbots. Users chat, create characters, and enjoy ad-free messaging, including stories, academic help, and language learning.
9	Writing Tool	ChatGPT	It is the most popular AI tool and is helpful for content generation.
10	Writing Tool	ChatPDF	ChatPDF, with artificial intelligence such as ChatGPT and Claude, facilitates PDF conversation, summarization, and analysis in over 100 languages, providing versatility to users.
11	Writing Tool	Chatsonic	It is a chatbot that integrates with Google to enable tasks that involve using the latest information practically in real-time.
12	Video Generator Tool	Clipchamp	Clipchamp AI offers AI-powered video editing tools to enhance videos and automatically generate speech and subtitles, facilitating the efficient creation of engaging content.
13	Writing Tool	Copilot	Microsoft Copilot, the AI chat assistant, improves productivity in Microsoft 365, Word, Excel, PowerPoint, Outlook, and Teams, integrating into various applications.
14	Writing Tool	Copy.ai	Edits and generates documents, reports, and articles.
15	Image Generator Tool	Craiyon	Text-to-image AI generator.
16	Image Generator Tool	DALL-E 2	The most recognizable of AI image generators out there is undoubtedly DALL-E 2. This text-to-image system by Open AI can create realistic images and art based on text input.
17	Translator	Deep L	This tool is used to create high-quality and fast translations.
18	Image Generator Tool	DeepDream	Deep Dream Generator fuses art and artificial intelligence in an online platform that uses neural networks to create and share unique generative art.
19	Video Generator Tool	D-id	D-ID AI creates videos for digital people using artificial intelligence from a single image. Its application facilitates the rapid creation of AI videos for various applications.
20	Research tool	Elicit	AI tool for researchers based on GPT-3. People use it to find relevant papers and provide abstracts quickly.
21	Voice Generator Tool	FakeYou	Text-to-speech AI with a large number of voice choices.
22	Voice Generator Tool	Fliki.ai	Fliki AI, an online platform, uses artificial intelligence to create videos with realistic voices. It facilitates the efficient creation of customizable audiovisual content.
23	Design Tool	Fontjoy	Simple AI tool that mixes up fonts to help with the design.
24	Writing Tool	Frase.io	Frase.io with AI uses artificial intelligence to create and optimize SEO content with writing, text generation, and rewriting tools.
25	Presentation Tool	Gamma	Gamma App with AI uses artificial intelligence to create compelling presentations, documents, and web pages with templates and code-free editing.

26	Writing Tool	Gemini	Google AI's Gemini is a powerful language model that uses machine learning to generate text, translate, and answer questions.
27	Writing Tool	Grammarly	This tool checks spelling and syntax, improves clarity, and increases writing quality.
28	Image Generator Tool	Hotpot.ai	Hotpot AI uses artificial intelligence to generate images, art, and text. The AI art generator and other tools automate creative tasks and improve content quality.
29	Design Tool	Jitter	AI tools aid in creating animated designs in seconds; they can animate text and other design elements. It is perfect for animating interfaces or creating social media and other communication materials.
30	Design Tool	Looka	They need to design a logo and other branded materials; it helps them make a logo in seconds.
31	Design Tool	Magic Sketchpad	AI experiment that draws along with the final users.
32	Image Generator Tool	Midjourney	Midjourney, an AI platform, turns descriptions into images. It uses advanced machine learning, trained on diverse data, to create creative and technically sound images.
33	Bussines Tool	Namelix	A tool that aids in the search for a business name.
34	Writing Tool	Perplexity	Perplexity AI, based on ChatGPT, is a generative AI and conversational search engine that provides accurate answers to complex questions, making it easy to obtain information and suggestions.
35	Writing Tool	Plot Generator	AI generator of random plots and ideas for writing.
36	Writing Tool	Poe	"Poe," a fast chat application, provides instant responses and advanced conversations with artificial intelligence, including GPT-3.5, GPT-4, and Claude 2.
37	Design Tool	Quick, draw!	AI that recognizes a doodle drawn by the user.
38	Writing Tool	Quillbot	QuillBot, with AI, improves writing through suggestions when entering text. It offers rephrasing, article rewriting, and AI content generation.
39	Research tool	Rayyan	Rayyan is an online systematic review tool developed by Qatar University. It is used to help researchers systematically review the scientific literature in their field of study.
40	Voice Generator Tool	Resemble.AI	Clone voices and create synthesized voices
41	Bussines Tool	Salesforce Einstein	Integrated set of AI technologies intended for CRM
42	Design Tool	Sketch MetaDemoLab	Online tool that can bring the drawings (or those of your children) to life
43	Music Generator Tool	Suno	Suno AI crafts songs based on user prompts, utilizing artificial intelligence to generate music from scratch and streamlining the creative process.
44	Image Generator Tool	Supreme.ai	Sometimes, getting the right idea or finding a meme that stirs the right emotions is hard. This tool is the "Chief Meme Officer," and it is continuously updated with the latest meme trends.
45	Video Generator Tool	Synthesia	It is the absolute pioneer in AI video generation. It can choose between 65+ AI avatars (based on real actors) who can speak your words in 60+ languages.
46	Education Tool	Teach Anything	AI tool that offers a fast answer to any question.
47	Image Generator Tool	This Beach does not exist	Image AI generator of beach images.
48	Writing Tool	Tiny Stories	AI generation tool that aims to create stories for children.
49	Presentation Tool	Tome	Tome AI employs artificial intelligence to create prompts-based narratives, presentations, and storylines, facilitating creative editing and refinement.
50	Bussines Tool	Weblium	Aids with the construction of business websites. Offers more than 300 templates for mobile-friendly websites, each with a free domain that can be customized once the paid upgrade is added.

Phase 2: Explore the generative AI: In this second phase, we immerse ourselves in the detailed exploration of the selected generative AI applications. The main activity during this stage consists of experimenting with navigation and interaction on the first page of each application. This choice is based on the premise that if accessibility errors exist on the home page, they are likely to be reproduced on the rest of the pages [26].

During the exploration, we pay particular attention to the user experience when interacting with the functionalities of the first page of each application. We evaluate the navigation's fluidity, the prompts' clarity, and the interface's ease of understanding. In addition, we focus on identifying potential accessibility barriers that could affect users with disabilities.

If accessibility errors are detected on the home page, we believe these could indicate more generalized problems in the design and development of the application. This situation reinforces the importance of addressing accessibility issues early on, as correcting them at later stages of development can be more complex and costly.

Our main objective in this phase is to understand the usability and accessibility of the generative AI applications, laying the groundwork for more detailed evaluations later in the process. The information gathered during this stage allows us to make informed decisions about the suitability of each application to meet the established accessibility and usability requirements.

Phase 3: Define the scenario: In this stage of the evaluation process, we define a standardized test scenario to ensure that all tested applications perform the same functions and are subjected to similar conditions. The main goal is to ensure a comprehensive environment test covering all aspects of the applications, including input/output, user interface, security, usability, and performance.

The environment test is executed for different user scenarios and device types, thus ensuring that the application works correctly and context-independently. This approach is essential to obtain comparable and objective results between the tested applications.

As part of this phase, we pre-install Google Chrome browser version 110.0.5481.77 for 64-bit systems. In addition, we incorporate the WAVE extension provided by WebAIM, version 3.2.2.0, updated until December 9, 2023, within the Windows version 10 operating system environment.

We then perform the following actions in a standardized manner for all evaluated applications: (1) We open the Google Chrome browser. (2) We enter the specific URL of the application we are evaluating. (3) We load the main page of the application.

These steps are essential to establish a uniform and reproducible starting point for each application. In addition, using the WAVE extension provides us with specific tools to evaluate web accessibility, allowing us to identify potential problems and ensure that applications comply with accessibility standards. Phase 3 aims to ensure consistency in testing by providing a standardized framework that addresses all essential aspects of the applications and ensures a thorough and unbiased evaluation of their performance and accessibility.

Phase 4: Automatic review: The central purpose of this phase is to carry out an automatic review using the WAVE plugin, a tool that evaluates web accessibility. The actions deployed during this stage include running the WAVE plugin after completing the normalized activities of Phase 3. Before this process, a meticulous check is carried out to close all browser tabs, thus avoiding possible interference. Likewise, the stability of the Internet connection is ensured, guaranteeing a review without interruptions.

The WAVE scan is initiated to carry out the automatic review, allowing adequate time for a thorough evaluation. During this period, active monitoring is adopted, alerting users of any error messages or interruptions that might arise during the process. Particular attention is paid to carefully examining both the summary and the detailed information provided by WAVE. In this automated phase, observing the process to identify potential problems and understand their impact is crucial, paying particular attention to areas marked as critical or requiring immediate attention.

The implementation of this automated review is strategically positioned as an initial tool to detect accessibility issues efficiently. Its seamless integration into later phases of the assessment process results in the effective identification of accessibility challenges. As a final result, detailed information and a summary are obtained that serve as a solid basis for further evaluations and possible improvements to optimize the application's accessibility.

Phase 5: Manual review: Aims to enrich the accessibility evaluation through a detailed and contextualized review of the generative application of artificial intelligence. It begins with a page-by-page scan to check for consistency and coherence in the presentation of information. Active interaction with functions and elements allows for evaluating the system's response and ease of use, considering users' different needs. The quality of automatic descriptions of multimedia content is validated, and alternative navigation tests are carried out to identify possible obstacles. The interdisciplinary collaboration involves accessibility experts, users with disabilities, and user experience professionals, fostering a comprehensive and diverse review.

Detailed documentation of identified accessibility issues is prioritized based on their impact and urgency, facilitating improvement planning. Close collaboration with the development team allows problems to be addressed, with iterative testing to validate possible improvements. Generating a detailed report highlights identified issues, potential improvements, and recommended actions for future updates. Combining manual and automated reviews, this comprehensive approach ensures an accessible user experience and supports an ongoing commitment to improving web accessibility.

Special attention is given to the barriers related to errors and contrast errors, analyzing each element in close detail. The data classification is carried out according to the four fundamental principles of accessibility and the success criteria established by WCAG 2.2 at level AA.

Importantly, applications are evaluated against guidelines designed to make web content more accessible to people with disabilities. In this case, compliance with level AA is considered the minimum level of accessibility that must be achieved to ensure an accessible user experience for all people, regardless of their abilities or disabilities.

Automatic review with WAVE against the WCAG 2.2 guidelines at level AA provides an initial, quantifiable assessment of application accessibility, providing a starting point for the more detailed assessments that can be carried out later in the process.

The main steps of Phase 5 include exhaustive navigation, where we explore each page and function of the application in detail. We review the home page, the secondary pages, and specific features of the app.

In real-time interaction, we perform interactive actions within the application to evaluate the real-time response and user experience. This case included interacting with dynamic elements and using specific functions.

For multimedia content verification, we review and evaluate the accessibility of any multimedia content in the application, such as images, videos, or audio. We ensure that alternative descriptions will be provided and that content is accessible to people with sensory disabilities.

Keyboard and voice navigation testing: we check accessibility by using the keyboard only and, where possible, by voice navigation. This process ensures that the application is usable by people with motor or visual disabilities.

Contrast and readability validation: we evaluate the contrast between the text and the background as well as the overall readability of the interface. We ensure that the colors used meet accessibility guidelines and that the text is readable for all people.

Navigation consistency evaluation: we ensure that navigation within the application is consistent and predictable; consistency contributes to a smoother and less error-prone user experience.

Upon recording the findings, we document the findings in detail, indicating any problems or areas that require improvement. These logs provide valuable information to the development team and help guide necessary corrective actions.

The manual review complements the automated evaluation, providing a more complete view of the application's accessibility and allowing accurate identification of potential barriers for users with disabilities. In this phase, we consider 26 guidelines based on WCAG 2.2 applied in the manual review of generative AI by accessibility experts. The guidelines are summarized in Table 2.

Table 2. Guidelines based on WCAG 2.2 applied in the manual review of generative AI

Guideline	Barrier	WCAG (Principle)	Criteria of success	Criteria.Barrier	Level
G01	Content on mouse over focus	Perceptible	1.4.13	Content on mouse over focus	AA
G02	Easy-to-read font	Perceptible	1.1.1	Easy to read font	A
G03	Text alternatives	Perceptible	1.1.1	Text alternatives	A
G04	Subtitling	Perceptible	1.2.4	Subtitling	AA
G05	Automatic transcripts	Perceptible	1.2.5	Automatic transcripts	AA
G06	Sign language	Perceptible	1.2.6	Sign language	AAA
G07	Information and relationships	Perceptible	1.3.1	Information and relationships	A
G08	Sensory features	Perceptible	1.3.1	Sensory features	A
G09	Adjust screen settings	Perceptible	1.3.4	Adjust screen settings	AA
G10	Interface rearrangement	Perceptible	1.3.5	Interface rearrangement	AA
G11	Use of color	Perceptible	1.4.1	Use of color	A
G12	Contrast without text	Perceptible	1.4.11	Contrast without text	AA
G13	Well-spaced elements	Perceptible	1.4.12	Well-spaced elements	A
G14	Good audio techniques	Perceptible	1.4.2	Good audio techniques	A
G15	Minimal contrast	Perceptible	1.4.3	Minimal contrast	AA
G16	Images as sharp as possible	Perceptible	1.4.5	Images as sharp as possible	AA
G17	Enhanced contrast	Perceptible	1.4.6	Enhanced contrast	AAA
G18	Visual presentation	Perceptible	1.4.8	Visual presentation	AAA
G19	Accessible keyboard	Operable	2.1.1	Accessible keyboard	A
G20	Pause, stop, hide	Perceptible	2.2.2	Pause, stop, hide	A
G21	Language	Understandable	3.1.1	Language	A
G22	Consistent navigation	Understandable	3.2.3	Consistent navigation	AA
G23	Labels or instructions	Understandable	3.3.2	Labels or instructions	A
G24	Help	Understandable	3.3.5	Help	AAA
G25	Screen reader support	Robust	4.1.2	Screen reader support	A
G26	Status messages	Robust	4.1.3	Status messages	AA

Phase 6: Analysis and results: In this final stage of the accessibility assessment process, we analyze the data collected during the previous phases and present the results meaningfully. This analysis provides valuable information to the development team and is a basis for making informed decisions about improvements and fixes.

The data collected during the automatic and manual review phases are grouped and organized in this phase. This process includes information on accessibility errors, usability issues, and other relevant aspects identified during the evaluation. In phase 6 of this process, we systematically record the evaluation data in a spreadsheet. We use Microsoft Excel 365 to consolidate and analyze the collected information in this study. During the review, we examine the values associated with errors, contrast errors, alerts, functions, structural elements, and ARIA (Accessible Rich Internet Applications).

We then rank the identified issues based on relevance, prioritizing the most significant impact on accessibility and user experience. Additionally, we create a detailed report presenting the evaluation's results. This report includes specific findings, recommendations to improve accessibility, and, if necessary, suggestions for future evaluations. This phase encourages collaboration and ensures a common understanding of challenges and opportunities for improvement.

4- Results

In this section, we present the evaluation results of 50 generative AI applications, which were classified and tabulated according to the four accessibility principles of WCAG 2.2. The information was summarized using Microsoft Excel with dynamic tables and graphs, making it easy to identify patterns and trends. Dynamic reports simplified data exploration and enabled deeper analysis.

To calculate total accessibility using a combined method that integrates automatic and manual review, we introduce Equation 1:

$$Total\ Accessibility = Wa * Automatic\ Score + Wm * Manual\ Score \tag{1}$$

where Total Accessibility represents the overall accessibility score. Automatic scoring is derived from automatic review using the WAVE tool. Manual scoring is obtained through manual review by accessibility experts. *Wa* and *Wm* represent the weights assigned to automatic and manual reviews.

This formula combines automatic and manual scoring in a weighted manner, allowing for adjustments in the relative importance of each review method. Provides flexibility to tailor the assessment process to specific needs, allowing for greater emphasis on automated or manual review based on accessibility analysis requirements.

Table 3 presents the evaluation of the 50 generative AI applications, analyzing the sample available in the Mendeley data repository [37]. Provides a comprehensive analysis of these applications using the automated WAVE tool based on WCAG 2.2. The table highlights the diversity in general errors, contrasts, alerts, accessibility features, and structural elements, underscoring the variability in attention to accessibility principles in these applications.

The data in the table shows the evaluation of 50 generative AI applications using the WAVE tool to analyze accessibility. Each row represents a specific generative AI application, with metrics such as general errors, contrast errors, accessibility alerts, accessibility features, structural elements, and ARIA (Accessible Rich Internet Applications Suite) usage. Some critical observations from the data include the following: First, the variety in the results, where a wide range of results is shown between the different applications. Some have low scores on specific accessibility metrics, while others have high scores.

Second, contrast errors and general errors: Some applications show significant contrast errors, affecting readability for people with visual impairments. Additionally, general errors can indicate areas where accessibility needs to be improved.

Third, diversity in accessibility features varies in the number of accessibilities features they offer; some have a more complete set of features, which can improve the user experience for a broader audience.

Fourth, the use of ARIA and structural elements: The proper use of ARIA and the presence of structural elements can significantly improve the accessibility of an application. Some apps show strengths in this regard, while others may need improvement. The data provides valuable insights into accessibility in generative AI applications, which can help developers identify areas for improvement and work toward a more inclusive experience for all users.

Notable differences between the applications are identified by looking at the general errors; ChatGPT, Copilot, and Clipchamp show low scores, suggesting better consideration of accessibility in their design. On the other hand, other applications present higher numbers, indicating possible areas of improvement in attention to accessibility.

Table 3. Evaluation of the 50 generative AI with WAVE

#	Generative IA	URL	Errors	Contrast Errors	Alerts	Features	Structural Elements	ARIA
1	Adobe Firefly	https://www.adobe.com/products/firefly.html	3	1	2	93	30	398
2	AI Art	https://boredhumans.com/text-to-image.php	6	0	7	1	6	4
3	AIVA	https://www.aiva.ai/	31	5	3	3	1	22
4	Artflow.ai	https://artflow.ai/	14	4	10	1	9	73
5	AutoDraw	https://www.autodraw.com/	19	13	1	1	13	0
6	Bing	https://www.bing.com/	4	0	4	2	20	155
7	Boomy	https://boomy.com/	8	14	7	14	64	26
8	Character.ai	https://beta.character.ai/	6	1	3	1	0	33
9	ChatGPT	https://chat.openai.com/chat	3	0	1	0	9	3
10	ChatPDF	https://www.chatpdf.com/	3	10	143	1	10	387
11	Chatsonic	https://app.writesonic.com/	5	8	21	20	7	34
12	Clipchamp	https://app.clipchamp	0	0	2	3	5	54
13	Copilot	https://copilot.microsoft.com/	3	0	1	0	11	28
14	Copy.ai	https://app.copy.ai/projects/1	6	8	7	3	30	91
15	Craiyon	https://www.craiyon.com/	4	5	1	31	15	9
16	DALL-E 2	https://openai.com/dall-e-2/	3	5	47	2	46	14
17	Deep L	https://www.deepl.com/	5	1	30	34	23	140
18	DeepDream	https://deepdreamgenerator.com/	5	10	2	0	10	7
19	D-id	https://www.d-id.com/	7	3	10	3	20	62
20	Elicit	https://elicit.org/	6	0	2	1	9	37
21	FakeYou	https://fakeyou.com/	10	6	12	4	8	111
22	Fliki.ai	https://app.fliki.ai/f	76	4	18	5	2	33
23	Fontjoy	https://fontjoy.com/	14	0	2	1	1	9
24	Frase.io	https://www.frase.io/	16	2	3	10	23	200
25	Gamma	https://gamma.app	51	84	71	2	13	567
26	Gemini	https://gemini.google.com/app	4	0	1	12	14	223
27	Grammarly	https://www.grammarly.com/	0	3	5	6	6	60
28	Hotpot.ai	https://hotpot.ai/logo-generator	30	74	25	27	32	0
29	Jitter	https://jitter.video/	8	3	3	1	7	14
30	Looka	https://looka.com/	11	30	24	11	5	13
31	Magic Sketchpad	https://magic-sketchpad.glitch.me/	25	7	1	0	4	0
32	Midjourney	https://www.midjourney.com/	1	6	22	5	22	522
33	Namelix	https://namelix.com/	1	5	2	9	10	0
34	Perplexity	https://www.perplexity.ai/	16	0	14	12	18	113
35	Plot Generator	https://www.plot-generator.org.uk/	68	4	21	2	17	4
36	Poe	https://poe.com/	20	8	15	7	12	12
37	Quick, draw!	https://quickdraw.withgoogle.com/	11	36	4	3	1	0
38	Quillbot	https://quillbot.com/	16	28	6	40	36	212
39	Rayyan	https://rayyan.ai	6	42	16	3	26	117
40	Resemble.AI	https://app.resemble.ai/	4	10	7	3	4	3
41	Salesforce Einstein	https://www.salesforce.com/products/einstein/	7	7	67	23	39	266
42	Sketch MetaDemoLab	https://sketch.metademolab.com/	0	1	2	34	3	1
43	Suno	https://www.suno.ai/	1	27	36	3	17	596
44	Supreme.ai	https://www.supermeme.ai/	1	0	3	8	11	41
45	Synthesia	https://www.synthesia.io/	6	2	11	18	10	93
46	Teach Anything	https://www.teach-anything.com/	1	2	1	5	8	15
47	This Beach does not exist	https://thisbeachdoesnotexist.com/	2	5	14	1	24	0
48	Tiny Stories	https://tinystorie.com/start/	5	0	2	1	0	0
49	Tome	https://tome.app/	28	0	34	2	26	161
50	Weblium	https://weblium.com/	9	30	9	74	74	49

Evaluating contrast errors reveals challenges in applications like Adobe Firefly, Gamma, and Gemini. These issues can affect readability for people with visual impairments, pointing to the need for adjustments in visual design to meet accessible standards. Accessibility alerts and features vary significantly; ChatGPT stands out with fewer alerts, suggesting a more robust design regarding accessible features. However, other applications can benefit from improvements to ensure a more inclusive experience. The presence of structural elements and the use of ARIA are critical aspects of accessibility. Apps like Tome, Sketch MetaDemoLab, and Looka show strengths, indicating a more robust approach toward creating accessible and structured interfaces.

This analysis highlights the diversity in the accessibility of generative artificial intelligence applications, highlighting specific areas requiring attention to move towards a more inclusive technology aligned with the principles of Sustainable Development Goals (SDGs).

Figure 4 presents the most significant errors highlighted in pink, contrast errors highlighted in yellow, alerts in gray, features in orange, structural errors in light blue, and ARIA errors in blue. Figure 4 provides a detailed look at the performance of various AI tools in terms of errors and specific accessibility features.

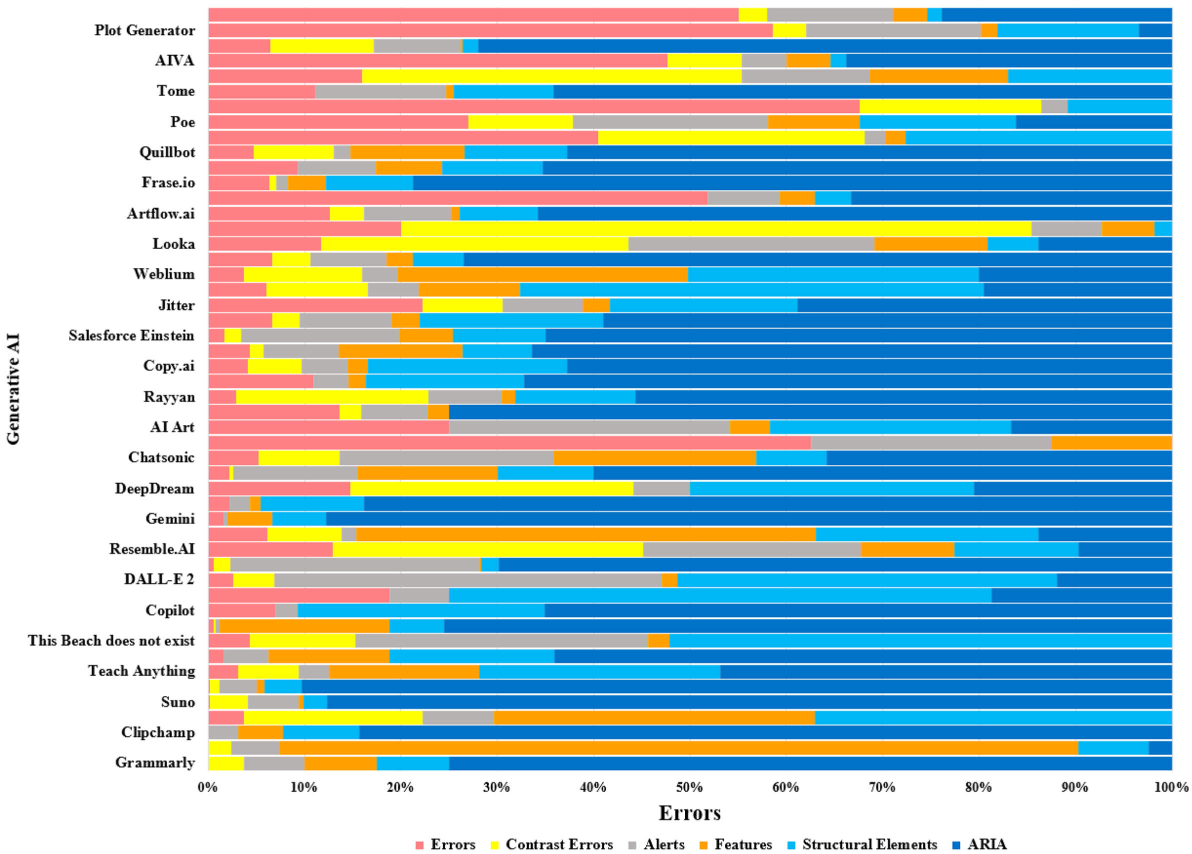

Figure 4. Overview of Significant Errors, Contrast Issues, Alerts, Features, Structural Errors, and ARIA Errors in Generative AI. Applications with Minimal Accessibility and Contrast Errors Highlighted

Grammarly stands out for its good performance, with zero general errors. However, it does have some issues with contrast errors, alerts, and accessibility features, suggesting areas for improvement to ensure a more inclusive experience.

Sketch MetaDemoLab shows solid performance with zero overall errors and low contrast errors. However, it scores higher in terms of accessibility features, pointing out possible improvements needed in this specific aspect.

Clipchamp, similar to Sketch MetaDemoLab, has zero overall errors but stands out with many contrast errors. It is essential to address these issues to improve visual accessibility.

Suno, unique in having only one general error, has many contrast errors and accessibility features. This situation indicates the need for adjustments to ensure a more equitable user experience.

Midjourney shows good overall performance with only one overall error. However, it has many contrast errors and accessibility features, suggesting specific areas for improvement.

These generative AI tools must address and correct the problems identified to move towards a more accessible technology aligned with the principles of the SDGs. This analysis allows developers and designers to focus on specific areas that require attention to improve the accessibility of their generative AI tools.

Table 4 provides a comprehensive breakdown of web accessibility following WCAG 2.2, categorizing criteria into Perceptible, Operable, Understandable, and Robust sections. Within the Perceptible category, emphasis is placed on legible font styles, utilization of text alternatives, and attention to visual presentation. The Operable category highlights keyboard accessibility and functionality for pausing, stopping, or concealing content. Understandable criteria assess the consistency of navigation elements and the presence of clear labeling and instructions. Lastly, robust evaluation focuses on supporting screen readers and handling status messages.

Table 4. Web Accessibility Guidelines (WCAG 2.2) Manual Assessment Summary: Perceptible, Understandable, and Robust Criteria

Guideline	WCAG 2.2	Criteria of success	Criteria.Barrier	Level	Total
G01	Perceptible	1.4.13	Content on mouse over focus	AA	2
G02	Perceptible	1.1.1	Easy to read font	A	50
G03	Perceptible	1.1.1	Text alternatives	A	3
G04	Perceptible	1.2.4	Subtitling	AA	4
G05	Perceptible	1.2.5	Automatic transcripts	AA	3
G06	Perceptible	1.2.6	Sign language	AAA	0
G07	Perceptible	1.3.1	Information and relationships	A	1
G08	Perceptible	1.3.1	Sensory features	A	2
G09	Perceptible	1.3.4	Adjust screen settings	AA	3
G10	Perceptible	1.3.5	Interface rearrangement	AA	3
G11	Perceptible	1.4.1	Use of color	A	50
G12	Perceptible	1.4.11	Contrast without text	AA	22
G13	Perceptible	1.4.12	Well-spaced elements	A	49
G14	Perceptible	1.4.2	Good audio techniques	A	3
G15	Perceptible	1.4.3	Minimal contrast	AA	22
G16	Perceptible	1.4.5	Images as sharp as possible	AA	22
G17	Perceptible	1.4.6	Enhanced contrast	AAA	22
G18	Perceptible	1.4.8	Visual presentation	AAA	22
G19	Operable	2.1.1	Accessible keyboard	A	22
G20	Perceptible	2.2.2	Pause, stop, hide	A	2
G21	Understandable	3.1.1	Language	A	28
G22	Understandable	3.2.3	Consistent navigation	AA	50
G23	Understandable	3.3.2	Labels or instructions	A	30
G24	Understandable	3.3.5	Help	AAA	50
G25	Robust	4.1.2	Screen reader support	A	29
G26	Robust	4.1.3	Status messages	AA	34

Perceptible (G01-G20): This category evaluates criteria related to content perception, such as font legibility, text alternatives, subtitles, automatic transcriptions, and visual presentation. Notably, compliance levels vary from AAA to AA, with some guidelines exhibiting more implementations than others. For instance, G02 (Easy-to-read fonts) and G11 (Use of color) demonstrate a robust implementation, each with 50 cases, while G06 (Sign language) lacks any recorded implementations in the sample.

Operable (G19): This criterion assesses the operability of content, particularly the accessibility of keyboard navigation. A compliance level of AA is observed, with 22 recorded cases.

Understandable (G21-G24): These criteria gauge content understandability, encompassing language specification, consistent navigation, and the provision of labels and instructions. Compliance levels range from AAA to A, with specific guidelines showing a higher implementation rate than others. For example, G22 (Consistent Navigation) and G24 (Help) are vigorously implemented, each with 50 cases.

Robust (G25-G26): Criteria related to content robustness are evaluated, including support for screen readers and status messages. Both guidelines maintain AA compliance levels, with 29 and 34 recorded cases, respectively.

This analysis underscores the significance of addressing all facets of web accessibility to ensure an inclusive user experience. Moreover, it offers a structured framework for identifying and prioritizing areas for enhancement, aligning with the commitment to SDGs and principles about equality and accessibility for all individuals.

To tackle the accessibility issues identified within the Perceptible, Operable, Understandable, and Robust categories, it's imperative to deploy a set of actions and strategies:

In Perceptibles (G01-G20), concerning the legibility of fonts, we recommend ensuring that the fonts used are legible and easy to read for users with visual disabilities. This study can be achieved by selecting appropriate fonts and ensuring that the size and contrast are appropriate.

It is also crucial to provide text alternatives for non-textual content, such as images, graphics, and multimedia so that users with visual disabilities can access information equivalently.

Regarding the problem of subtitles and transcriptions, we suggest including subtitles for multimedia content and automatic transcriptions to guarantee access to people with hearing disabilities.

About the problem of visual presentation, it is necessary to avoid using colors that make readability difficult and ensure that the visual elements are easily distinguishable.

Concerning the Operable principle (G19), Keyboard accessibility is essential. To solve this problem, ensuring that all content can be operated with the keyboard is essential, allowing users to navigate and interact without depending on the mouse.

Regarding the Understandable principle (G21-G24), it is essential to manage the language specification of the content to facilitate users' understanding. On the other hand, consistent navigation is vital throughout the website, making it easy for users to find and access content intuitively. Additionally, labels and instructions should be clear for forms and interactive elements, ensuring users understand how to interact with them.

Finally, the Robust principle (G25-G26) must be resolved by integrating support for screen readers and other assistive technologies used by visually impaired people. In addition, it is necessary to include clear and descriptive status messages to inform users about the status and actions taken in the user interface.

Addressing these issues directly contributes to SDG 10 (reduces inequality within and between countries) by promoting equal access to information and technology for everyone, regardless of their abilities or capabilities. They also support SDG 4 (quality education), facilitating access to online education for people with disabilities, and SDG 9 (industry, innovation, and infrastructure), promoting the creation of accessible technologies and the improvement of digital infrastructures.

Figure 5 presents a detailed analysis of accessibility errors based on WCAG principles, revealing various compliance levels and error distributions. Concerning the Perceptible principle, many errors stand out at level A, representing 31% of the total. These results point to specific areas that require improvement to ensure content is visible to a broader audience of users.

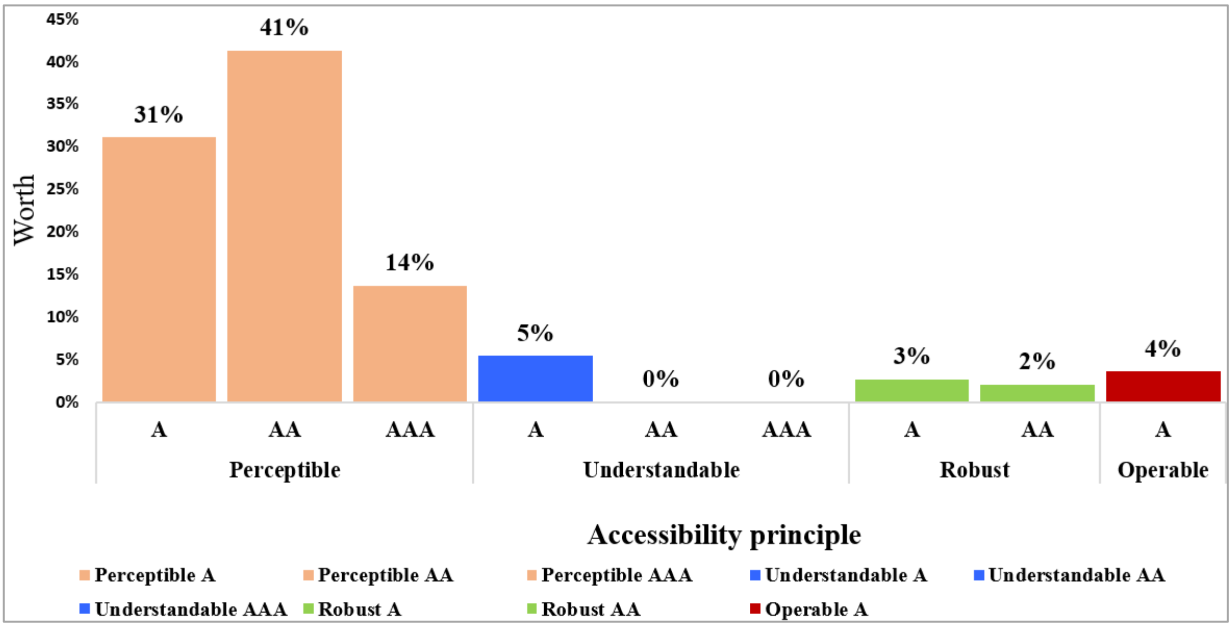

Figure 5. **Analysis of accessibility errors according to WCAG principles in generative AI**

The most considerable proportion of errors is found at level AA, at 41%, underlining the need to address more specific issues related to the perception of the content. Although smaller, the 14% errors at level AAA indicate the importance of improving more advanced aspects of perception.

Regarding the Understandable principle, errors are recorded at level A, but they represent a low percentage of 5%. This case suggests a solid initial understanding of the content, although there is room for improvement. No errors were identified at levels AA and AAA, which is positive and suggests a good understanding of the content.

In the Robust principle, a low level of errors is observed at levels A (3%) and AA (2%), indicating solid compliance with the robustness principles, ensuring that the content is compatible with different user agents and technologies of assistance.

Finally, the Operable principle reveals 4% errors at level A, pointing out areas of possible improvement to make the application more operable. The total number of errors provides a complete view, allowing specific focus areas to be identified to improve the system's overall accessibility. This analysis gives developers a clearer understanding of the distribution of errors based on accessibility principles, allowing them to prioritize and address specific areas to improve the user experience, aligning with the SDGs.

Figure 5 clearly shows the areas where web accessibility improvements are required, particularly addressing issues within the Perceptible category.

To solve this problem, we suggest AI-based content personalization for various needs. That is, integrating AI algorithms that analyze user behavior and browsing patterns. Based on that analysis, AI can personalize the content presentation by offering alternative text descriptions for images or adjusting color contrast for visually impaired users. This case would personalize the user experience while ensuring the content is perceptible to a broader range of users.

Figure 6 presents the rankings of the 10 most accessible generative AI tools, with separate evaluations across manual and automated evaluations. The analysis highlights the variability in accessibility performance among the top 10 generative AI tools; some show strong performance, while others have significant areas that need improvement. This study highlights the importance of prioritizing accessibility in developing these tools to ensure they are available and usable for all users, regardless of their abilities. Each application is assigned a total score that comprises its manual and automatic evaluations, revealing that:

Clipchamp shows excellent performance with total errors and contrast errors of 0.0%, indicating high quality in terms of accessibility. Supreme.ai and Sketch MetaDemoLab present similar accessibility problems, with a sum of errors and contrast errors of 3.8%. Although there is room for improvement, these results still indicate relatively strong performance compared to other tools. Copilot, Grammarly, ChatGPT, and Teach Anything have more accessibility issues, totaling 11.5% in errors and contrast errors. This analysis suggests that significant attention should be paid to improving the accessibility of these tools. Gemini, Bing, and Adobe Firefly are the three tools that show the worst performance in terms of accessibility, with 15.4% in errors and contrast errors.

This study indicates a notable number of issues that must be addressed to improve accessibility and ensure a fair user experience. To improve accessibility in generative AI tools, a context-sensitive wizard can be included and integrated within the generative AI tool. The assistant can analyze the user's input and preferences (screen reader usage, color blindness) to adjust the accessibility required by the user in real-time. These settings could offer alternative text descriptions for generated content, modify color contrast based on user settings, or provide keyboard shortcuts for commonly used functions. The assistant can also offer short, interactive tutorials to educate users about the accessibility features, promoting understanding and self-sufficiency. Another suggestion to improve accessibility in generative AI tools is to integrate accessibility testing functionalities directly within the generative AI tool. As users create content or use AI's generative features, the tool can automatically generate a preview that displays the content with different accessibility settings (high contrast mode, screen reader support). Users can then visually assess the accessibility of their creations and adjust parameters within the generative AI tool to ensure optimal visibility.

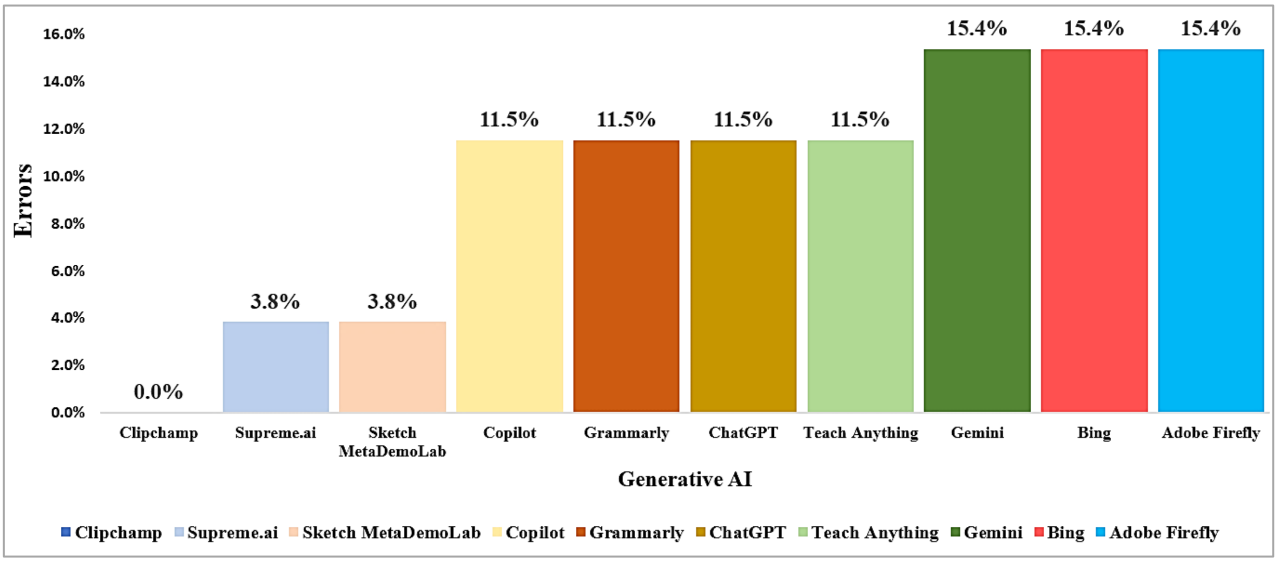

Figure 6. The ten most accessible generative AIs

5- Discussion

In the evaluation, while analyzing the accessibility of the applications, the success criteria frequently repeated were identified. These success criteria are guidelines established by W3C [38] to ensure that websites are accessible to people with disabilities.

In the comprehensive accessibility analysis of 50 generative AI applications, WCAG 2.2 [20] was used as a reference framework. These guidelines, which focus on perceptible, operable, understandable, and robust principles, establish an international standard to improve not only web accessibility but also accessibility in the specific field of generative AI applications.

Glazko et al. [1] examined the impact of generative AI tools on people with disabilities. Researchers explored the accessibility and challenges of using generative AI. The results revealed diverse accessibility applications but raised concerns about verifiability and training data to maximize benefits and mitigate potential risks for people with disabilities. In our study regarding the perceptible principles [20], high compliance stands out, especially in the use of readable fonts (G02) and the implementation of alternative texts (G03), both achieving a rate of 100%. However, areas such as text-free contrast (G12) and audio techniques (G14) exhibit variations, pointing to specific opportunities for improvement. The Clipchamp app leads in perceptible accessibility with a notable 4%, followed closely by Copilot, ChatGPT, and Grammarly, hitting 8%.

In the area of operability [20], keyboard accessibility (G19) is highlighted with a compliance range of 30% to 60%, while the ability to pause, stop, or hide content (G20) shows a more modest compliance of 2%. These results highlight the need to optimize operational aspects in the evaluated applications to guarantee an inclusive experience.

Comprehensible principles [20] reveal a solid initial understanding of the content, with 28% and 30% compliance rates for criteria G21 and G23, respectively. Consistency in navigation (G22) stands out with a compliance of 50%. The presence of errors at level A is minimal, with 5%, evidencing an excellent initial understanding.

In terms of robustness [20], screen reader support (G25) shows 29% compliance, and status message handling (G26) reaches 34%. These results indicate a solid implementation of the robust principles, ensuring the content is compatible with various assistive technologies.

The ten most accessible tools stand out for their commitment to accessibility. Clipchamp leads with an outstanding 4%, closely followed by Copilot (8%), distinguished by a solid automatic evaluation. ChatGPT and Grammarly, at 28%, offer a notable balance of manual and automated assessments. Bing and Gemini, both with 30%, show consistent results in both evaluations, reaffirming their commitment to accessibility. These findings not only provide an accurate snapshot of the current state of accessibility in generative AI but also point out critical areas for future improvements.

To improve accessibility in generative artificial intelligence tools, it is suggested that artificial intelligence assistants be incorporated to personalize the user experience to achieve a higher level of accessibility. López-Gil & Pereira [39] investigated whether large language model (LLM) systems can effectively evaluate web accessibility criteria that traditionally require manual intervention. The results showed that the LLMs successfully detected accessibility issues that the automated tools missed, achieving a detection rate of 87.18%. They also suggested that LLMs can improve automated accessibility testing by capturing issues that current software misses. Future research should expand the scope of the evaluation.

This principle relates to the need for applications to be compatible with different user agents, including assistive technologies. The low number of issues in this area suggests that the applications were generally robust and capable of working with various user agents.

Accessibility is a critical aspect in the development of applications generative AI; here are some solutions and recommendations to improve the accessibility of these applications: User-Centered Design [9, 13]; generative AI applications must be designed considering all users' needs and capabilities, including those with disabilities. This process involves user testing with people with disabilities to identify and correct accessibility issues [13].

Othman et al. [40] advocate the effectiveness of large language models (LLMs), particularly ChatGPT, in automatically addressing web accessibility issues to comply with WCAG 2.1. Manual testing results highlighted the use of ChatGPT to improve web accessibility, offering valuable insights while striving to comply with WCAG and improve the inclusion of people with disabilities. We therefore suggest incorporating assistive technologies into generative AI applications, such as screen readers, to improve accessibility. Specific technologies can also be developed to improve the accessibility of generative AI, such as voice input systems and gesture recognition.

Generative AI applications must be transparent and explainable so that users with disabilities can understand how they work and how to make decisions based on their results. For example, if a chatbot is used to provide customer support, the explainability of its recommendations could be helpful for people with cognitive disabilities.

Generative AI applications must comply with accessibility regulations and standards, such as WCAG 2.1 and 2.2. to ensure that the application is accessible to the most significant number of users, including those with disabilities.

Generative AI application developers should be trained in WCAG and know the latest techniques and tools to improve the accessibility of their applications. Awareness of the importance of accessibility in the AI industry can also be promoted through education and outreach. Finally, inclusion and diversity should be considered from the beginning of the AI application development process [41]. Future studies can include a group of people with different types of disabilities to ensure that applications have a design that considers the needs of all users.

Solutions and recommendations aligned with the SDGs are proposed to strengthen the accessibility of applications with generative artificial intelligence. Firstly, it is crucial to adopt a user-centered design approach, where these applications are meticulously designed considering the needs and capabilities of all users, which is in line with SDG 10, which advocates reducing inequalities.

Second, generative AI applications can leverage existing assistive technologies, such as screen readers, and the proactive development of specific technologies, such as voice input systems and gesture recognition. This duality in the technological approach not only improves accessibility but also contributes to SDG 9, which seeks to foster innovation and sustainable infrastructure.

Transparency and explainability emerge as a third fundamental pillar. Generative AI must be transparent, allowing users, especially those with disabilities, to understand how they work and make informed decisions based on the results. This level of understanding contributes significantly to user autonomy and active participation, supporting SDG 16 on peace, justice, and strong institutions.

Fourthly, compliance with accessibility regulations and standards, such as WCAG 2.1 and 2.2, is unavoidable. This conformity not only promotes accessibility but also addresses SDG 4, ensuring quality and equitable education for all users, including those with disabilities.

Specialized training for developers constitutes the fifth essential point. These professionals should receive accessibility training and stay current on the latest techniques and tools, contributing to SDGs 4 and 9. Awareness of the importance of accessibility in the AI industry can be driven through educational programs and outreach campaigns, aligning with several SDGs, such as 4, 9, and 10.

Finally, inclusion and diversity must be considered from the initial phases of development, supporting SDG 10. The active participation of people with different types of disabilities in future studies will ensure that generative AI applications are designed to consider the diverse needs of users, thus contributing to the achievement of multiple SDGs. This inclusive approach is essential to build an equitable and sustainable technological environment in line with the comprehensive vision of the Sustainable Development Goals. Othman et al. [40] investigated using ChatGPT to automatically address web accessibility issues and ensure compliance with WCAG 2.1. The findings have significant implications for stakeholders striving to comply with WCAG and improve online inclusion for people with disabilities through more accessible platforms.

6- Conclusion

Ensuring the accessibility of generative AI tools for people with disabilities is paramount and identifying success criteria based on WCAG 2.2 is a crucial step toward this. The study's results highlight the presence of significant accessibility barriers in the evaluated applications, which particularly disadvantage users with disabilities. It is essential to emphasize accessibility considerations throughout the application design lifecycle, advocating a proactive approach rather than reactive measures. By prioritizing accessibility from the beginning of development, it is possible to align with WCAG standards, effectively addressing the multifaceted challenges inherent in the accessibility of generative AI tools. This proactive stance ensures regulatory compliance and upholds the ethical and moral imperatives of facilitating equitable social participation, regardless of ability.

In addition, improving the accessibility of generative AI tools contributes to broader global agendas, particularly aligned with the United Nations Sustainable Development Goals. The pursuit of accessibility aligns with equality, education, innovation, and inclusion goals, emphasizing its critical role in fostering a more equitable and sustainable future. Sometimes, collaborative efforts between generative AI developers, accessibility experts, and users with diverse abilities are crucial. These collaborations facilitate the identification of specific challenges and the design of inclusive solutions, improving the user experience for all. In addition, ongoing research efforts are essential to monitor best practices, adapt accessibility guidelines, and explore innovative approaches to integrate accessible principles into generative AI algorithms. At the same time, efforts to raise awareness and provide training in inclusive design and web accessibility are imperative, ensuring accessibility remains an integral part of education and professional development in artificial intelligence. Through these concerted efforts, we can aspire to a future where generative AI tools are accessible, contributing to a more inclusive and sustainable society.

6-1- Future Work

To move towards an inclusive and sustainable future, it is imperative to address the identified gaps in the accessibility of generative AI. First, close collaboration is required between generative AI developers, accessibility experts, and users with various abilities. This collaborative approach will identify specific challenges and design solutions that improve the user experience for everyone.

Second, it is essential to promote continued research in generative AI accessibility. The rapid evolution of this technology requires constant monitoring of best practices and adapting accessibility guidelines. Research should explore new ways to integrate accessible principles into generative AI algorithms, ensuring that technological innovation is aligned with inclusion.

Finally, future work should focus on awareness and training. It is necessary to raise awareness among developers, designers, and decision-makers about the importance of accessibility in generative AI. Training in inclusive design and web accessibility should be integral to education and professional development in artificial intelligence. This proactive approach will help cultivate a more equitable and sustainable technology ecosystem.

7- Declarations

7-1- Author Contributions

Conceptualization, P.A.-V., B.S.-A., and L.S.-U.; methodology, P.A.-V., and B.S.-A.; software, L.S.-U.; validation, P.A.-V., B.S.-A., and L.S.-U.; formal analysis, P.A.-V., B.S.-A., and L.S.-U.; investigation, P.A.-V., B.S.-A., S.N.-V., D.S., and L.S.-U.; resources, P.A.-V.; data curation, P.A.-V.; writing—original draft preparation, P.A.-V., B.S.-A., S.N.-V., D.S., and L.S.-U.; writing—review and editing, P.A.-V., B.S.-A., S.N.-V., D.S., and L.S.-U.; visualization, P.A.-V. and L.S.-U.; supervision, P.A.-V., and L.S.-U.; project administration, P.A.-V.; funding acquisition, P.A.-V. All authors have read and agreed to the published version of the manuscript.

7-2- Data Availability Statement

Publicly available datasets were analyzed in this study. This data can be found here: https://doi.org/10.17632/2kyksx8hty.2.

7-3- Funding

Universidad de Las Américas-Ecuador funded this research as part of internal project 489. A.XIV.24.

7-4- Institutional Review Board Statement

Not applicable.

7-5- Informed Consent Statement

Not applicable.

7-6- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

8- References

[1] Glazko, K. S., Yamagami, M., Desai, A., Mack, K. A., Potluri, V., Xu, X., & Mankoff, J. (2023). An Autoethnographic Case Study of Generative Artificial Intelligence's Utility for Accessibility. The 25th International ACM SIGACCESS Conference on Computers and Accessibility, 99, 1-8. doi:10.1145/3597638.3614548.

[2] Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2023). Generative AI. Business & Information Systems Engineering, 66(1), 111–126. doi:10.1007/s12599-023-00834-7.

[3] Roegiest, A., & Pinkosova, Z. (2024). Generative Information Systems Are Great If You Can Read. Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval, 165 - 177. doi:10.1145/3627508.3638345.

[4] Pascual, J. K., & Osorio, C. P. (2023). The Road to a More Accessible Web: Bridging the Inequality Gap with AI. Swiss Cognitive, Küsnacht, Switzerland. Available online: https://swisscognitive.ch/2023/10/10/the-road-to-a-more-accessible-web-bridging-the-inequality-gap-with-ai/ (accessed on May 2024).

[5] Silva, J. S. R., Cardoso, P. C. F., de Bettio, R. W., Tavares, D. C., Silva, C. A., Watanabe, W. M., & Freire, A. P. (2024). In-Page Navigation Aids for Screen-Reader Users with Automatic Topicalisation and Labelling. ACM Transactions on Accessible Computing, 12, 1-45. doi:10.1145/3649223.

[6] World Wide Web Consortium (W3C). (2023). Web Content Accessibility Guidelines (WCAG) 2.2. World Wide Web Consortium, Cambridge, United States. Available online: https://www.w3.org/TR/WCAG22/ (accessed on May 2024).

[7] López, J. L., Espinilla, M., & Verdejo, Á. (2023). Evaluation of the Impact of the Sustainable Development Goals on an Activity Recognition Platform for Healthcare Systems. Sensors, 23(7), 3563. doi:10.3390/s23073563.

[8] Sarker, I. H. (2022). AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. SN Computer Science, 3(2), 158. doi:10.1007/s42979-022-01043-x.

[9] Haleem, A., Javaid, M., Asim Qadri, M., Pratap Singh, R., & Suman, R. (2022). Artificial intelligence (AI) applications for marketing: A literature-based study. International Journal of Intelligent Networks, 3, 119–132. doi:10.1016/j.ijin.2022.08.005.

[10] Weisz, J. D., He, J., Muller, M., Hoefer, G., Miles, R., & Geyer, W. (2024). Design Principles for Generative AI Applications. Proceedings of the CHI Conference on Human Factors in Computing Systems, 1-22. doi:10.1145/3613904.3642466.

[11] Nasser, N., Ali, A. G., Karim, L., & Al-Helali, A. (2024). Advancing Vehicular Independence for the Visually Impaired Through Smart Technology. ScienceOpen Preprints, 1-12. doi:10.14293/PR2199.000771.v1.

[12] van Toorn, G. (2024). Inclusion interrupted: Lessons from the making of a digital assistant by and for people with disability. Government Information Quarterly, 41(1), 101900. doi:10.1016/j.giq.2023.101900.

[13] Perez Medina, J. L., Gonzalez, M., Pilco, H. M., Beatriz Jimenes Vargas, K., Acosta-Vargas, P., Sanchez-Gordon, S., Calle-Jimenez, T., Esparza, D., & Rybarczyk, Y. (2019). Usability study of a web-based platform for home motor rehabilitation. IEEE Access, 7, 7932–7947. doi:10.1109/ACCESS.2018.2889257.

[14] Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2021). Substantial capabilities of robotics in enhancing industry 4.0 implementation. Cognitive Robotics, 1, 58–75. doi:10.1016/j.cogr.2021.06.001.

[15] Alabdulatif, A., Khalil, I., & Saidur Rahman, M. (2022). Security of Blockchain and AI-Empowered Smart Healthcare: Application-Based Analysis. Applied Sciences (Switzerland), 12(21). doi:10.3390/app122111039.

[16] Google Trends. (2023). Trend in artificial intelligence-based applications in the last 12 months. Google Trends, the Mountain View, United States. Available online: https://trends.google.es/trends/explore?q=ChatGPT,Chatsonic,DeepL,Grammarly,Rayyan (accessed on May 2024).

[17] Fakrudeen, M. (2023). Accessibility of Multimedia Resources from Irish Universities for People with Epilepsy. Emerging Technologies in Computing, iCETiC 2022, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 463, Springer, Cham, Switzerland. doi:10.1007/978-3-031-25161-0_10.

[18] Acosta-Vargas, P., Zarate-Estrella, S., Mantilla-Vaca, F., Novillo-Villegas, S., Chimbo, C., & Luján-Mora, S. (2021). Towards accessibility and inclusion of native mobile applications available for ecuador in google play store. Sustainability (Switzerland), 13(20). doi:10.3390/su132011237.

[19] WebAIM (2023). Institute for Disability Research USU. The WebAIM Million 2023. Institute for Disability Research, Policy, and Practice, Logan, United States. Available online: https://webaim.org/projects/million/ (accessed on May 2024).

[20] W3C. (2018). Web Content Accessibility Guidelines (WCAG) 2.1 World Wide Web Consortium, Cambridge, United States. Available online: https://www.w3.org/TR/WCAG21/ (accessed on May 2024).

[21] Acosta-Vargas, P., Salvador-Acosta, B., Salvador-Ullauri, L., & Jadán-Guerrero, J. (2022). Accessibility challenges of e-commerce websites. PeerJ Computer Science, 8. doi:10.7717/PEERJ-CS.891.

[22] United Nations Economic and Social Commission for Asia and the Pacific. (2016). Reduce inequality within and among countries. Statistical Yearbook for Asia and the Pacific 2015, United Nations, New York, United States. doi:10.18356/7df2027f-en.

[23] United Nations. (2023). Promote peaceful and inclusive societies for sustainable development 2023. United Nations, New York, United States. Available online: https://www.globalgoals.org/goals/16-peace-justice-and-strong-institutions/ (accessed on May 2024).

[24] United Nations. (2023). Partnerships for the Goals 2023. United Nations, New York, United States. Available online: https://www.globalgoals.org/goals/17-partnerships-for-the-goals/ (accessed on May 2024).

[25] Barua, P. D., Vicnesh, J., Gururajan, R., Oh, S. L., Palmer, E., Azizan, M. M., Kadri, N. A., & Acharya, U. R. (2022). Artificial Intelligence Enabled Personalised Assistive Tools to Enhance Education of Children with Neurodevelopmental Disorders—A Review. International Journal of Environmental Research and Public Health, 19(3), 1192. doi:10.3390/ijerph19031192.

[26] Web Accessibility Evaluation Tool (WebAIM), Logan, United States. Available online: https://wave.webaim.org/ (accessed on May 2024).

[27] VOSviewer version 1.6.19 (2024). VOSviewer, Leiden University, Leiden, Netherlands. Available online: https://www.vosviewer.com/ (accessed on May 2024).

[28] Emenike, M. E., & Emenike, B. U. (2023). Was This Title Generated by ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators. Journal of Chemical Education, 100(4), 1413–1418. doi:10.1021/acs.jchemed.3c00063.

[29] Tiwary, T., & Mahapatra, R. P. (2022). Web Accessibility Challenges for Disabled and Generation of Alt Text for Images in Websites using Artificial Intelligence. 2022 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). doi:10.1109/icict55121.2022.10064545.

[30] Draffan, E. A., Ding, C., Wald, M., Everett, H., Barrett, J., Sasikant, A., ... & Newman, R. (2020). Can a Web Accessibility Checker Be Enhanced by the Use of AI?. Computers Helping People with Special Needs, ICCHP 2020, Lecture Notes in Computer Science, 12376, Springer, Cham, Switzerland. doi:10.1007/978-3-030-58796-3_9.

[31] Reddy, S. (2024). Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. Implementation Science, 19(1). doi:10.1186/s13012-024-01357-9.

[32] Acosta-Vargas, P., Novillo-Villegas, S., Salvador-Acosta, B., Calvopina, M., Kyriakidis, N., Ortiz-Prado, E., & Salvador-Ullauri, L. (2022). Accessibility Analysis of Worldwide COVID-19-Related Information Portals. International Journal of Environmental Research and Public Health, 19(19), 12102. doi:10.3390/ijerph191912102.

[33] Acosta-Vargas, P., Lujan-Mora, S., & Salvador-Ullauri, L. (2017). Quality evaluation of government websites. 2017 Fourth International Conference on EDemocracy & EGovernment (ICEDEG), 8-14. doi:10.1109/icedeg.2017.7962507.

[34] Ara, J., & Sik-Lanyi, C. (2022). Investigation of COVID-19 Vaccine Information Websites across Europe and Asia Using Automated Accessibility Protocols. International Journal of Environmental Research and Public Health, 19(5). doi:10.3390/ijerph19052867.

[35] World Wide Web Consortium (W3C). (2014).Website Accessibility Conformance Evaluation Methodology (WCAG-EM) 1.0. World Wide Web Consortium, Cambridge, United States. Available online: https://www.w3.org/TR/WCAG-EM/ (accessed on May 2024).

[36] Synthesia. (2023). 105 Fun AI Tools You've Never Heard of. Synthesia, London, United Kingdom Available online: https://www.synthesia.io/post/ai-tools (accessed on May 2024).

[37] Mendeley Data 2024;V2. (2024). Acosta-Vargas P. (Dataset) Generative AI tools. Mendeley Data 2024;V2. Elsevier, Amsterdam, Netherlands. Available online: https://data.mendeley.com/datasets/2kyksx8hty/2 (accessed on May 2024).

[38] World Wide Web Consortium (W3C). (2018). W3C Issues Improved Accessibility Guidance for Websites and Applications. World Wide Web Consortium, Cambridge, United States. Available online: https://www.w3.org/2018/06/pressrelease-wcag21.html.en (accessed on May 2024).

[39] López-Gil, J. M., & Pereira, J. (2024). Turning manual web accessibility success criteria into automatic: an LLM-based approach. Universal Access in the Information Society. doi:10.1007/s10209-024-01108-z.

[40] Othman, A., Dhouib, A., & Nasser Al Jabor, A. (2023). Fostering websites accessibility: A case study on the use of the Large Language Models ChatGPT for automatic remediation. Proceedings of the 16th International Conference on Pervasive Technologies Related to Assistive Environments, 707 - 713. doi:10.1145/3594806.3596542.

[41] Jiang, J., Ma, X., Ouyang, D., & Williams, R. O. (2022). Emerging Artificial Intelligence (AI) Technologies Used in the Development of Solid Dosage Forms. Pharmaceutics, 14(11), 2257. doi:10.3390/pharmaceutics14112257.