WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Generative AI for Improved Web Accessibility

M. Shaun Daniel [1] **, *Dr. Shikha Tiwari*** [2] ***

[1]Postgraduate, Department of Computer Science & Engineering, Amity University Chhattisgarh, Raipur 493225, Chhattisgarh, India
[2] Faculty/Mentor, Department of Computer Science & Engineering, Amity University Chhattisgarh, Raipur 493225, Chhattisgarh, India

ABSTRACT :

Web accessibility is a critical aspect of ensuring equal access to digital information for all users, including those with disabilities. Despite significant advancements in assistive technologies, many web accessibility challenges remain unaddressed, particularly for individuals with visual, cognitive, and neurological impairments. This review paper explores the potential of generative artificial intelligence (AI) to enhance web accessibility.

Generative AI models, such as GPT-3 and DALL-E, have demonstrated remarkable capabilities in natural language processing and image generation, offering innovative solutions for accessibility. This paper reviews existing literature on web accessibility challenges and current assistive technologies, highlighting their limitations. It then examines the applications of generative AI in creating image descriptions, content summarization, and interactive question-answering systems. Furthermore, ethical and practical considerations in deploying these AI solutions are discussed.

By synthesizing current research and identifying future directions, this review aims to provide a comprehensive understanding of how generative AI can revolutionize web accessibility, paving the way for more inclusive digital experiences.

Keywords: Artificial Intelligence, Web Accessibility, Inclusivity, Disability, AI-driven Solutions, Digital Inclusion, Assistive Technology, User Experience

1. Introduction :

The internet has become an integral part of our daily lives, serving as a gateway to information, communication, and services. However, its benefits are not equally accessible to everyone. Millions of people worldwide face barriers when navigating websites, hindered by visual, cognitive, or motor impairments.[1] Traditional web design practices often fall short in accommodating these diverse needs, necessitating innovative solutions.
Approximately 1.3 billion people worldwide, or around 16% of the global population, live with a significant disability. In other words, one in six individuals faces challenges in daily life due to a disability. [1]

To bridge this gap, assistive technology are widely needed, but availability is limited, particularly in low- and middle-income nations. [3] Screen readers, for instance, leverage text-to-speech technology to audibly convey on-screen content, enabling users to comprehend text, navigate menus, and interact with user interface elements. Refreshable braille displays provide a tactile representation of web pages, allowing braille readers to access digital information seamlessly. Moreover, magnification tools and high-contrast modes offer enhanced visibility for individuals with low vision, ensuring a more inclusive digital experience.

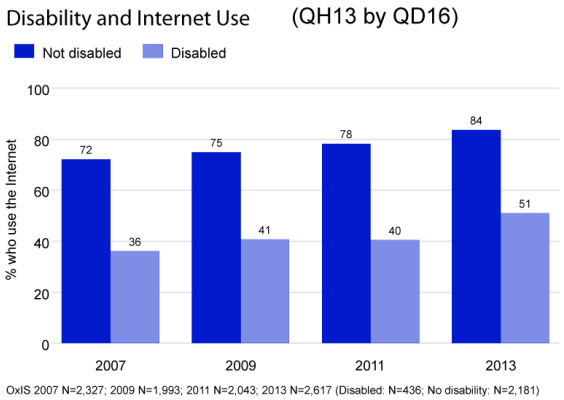

Fig. 1 - Trends in Disabled vs Not Disabled Internet Users [2]

This is where generative artificial intelligence (AI) emerges as a powerful solution. For visually impaired users, generative AI could automate the creation of accurate image descriptions and alternative text, providing crucial context and meaning that would otherwise be inaccessible through screen readers or braille displays alone. [4] These AI-generated descriptions could adapt to the user's preferences and level of detail required, enhancing their comprehension and engagement with image-heavy content such as e-commerce platforms, social media, and news websites.
Google's latest developments in generative AI for image captioning showcase the transformative potential of these technologies in enhancing web accessibility. [5] By leveraging these models, we can revolutionize how users with disabilities interact with and consume web content.

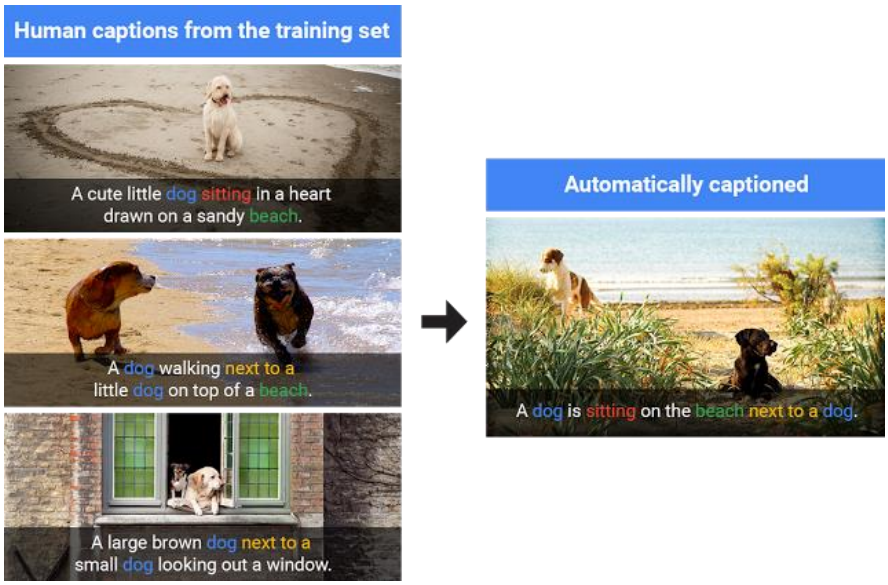

Fig. 2. Google's image captioning AI describing images [6]

2. Literature Review

This literature study goes into the many elements of web accessibility, putting light on the difficulties faced by visually impaired and cognitively/neurologically challenged users. It investigates the strengths and limitations of current assistive technology and provides an overview of generative AI models and their potential uses in improving accessibility. Understanding these important components lays the groundwork for designing inclusive solutions that remove barriers and unleash the full potential of the internet for all people, regardless of ability.

2.1 Web Accessibility Challenges for Visually Impaired and Cognitively/Neurologically Impaired Users

Visual Impairments: Visual impairments range from mild to severe, with each level presenting unique challenges for web accessibility. According to the World Health Organization (WHO), at least 2.2 billion people worldwide have a vision impairment, with at least 1 billion people having a moderate or severe distance vision impairment that could have been prevented or has yet to be addressed.[7]

Some examples of visual impairments are as follows :

- Blindness: Individuals who are blind or have complete vision loss rely heavily on screen readers and refreshable braille displays to access web content. Screen readers are assistive technologies that convert text, graphics, and other on-screen elements into synthesized speech or braille output. [8] While these tools are invaluable for blind users, they face several challenges in conveying complex layouts, images, and interactive elements effectively.

- Low Vision: People with low vision, or partial sight, often use magnification tools or high-contrast settings to enhance the visibility of web content. [9] However, over-reliance on magnification can affect the overall user experience, making it difficult to navigate and comprehend web pages with complex layouts. Additionally, high-contrast modes may not address all accessibility needs, as they do not necessarily improve the readability of text or the clarity of images.

Cognitive and Neurological Impairments: Cognitive and neurological impairments can significantly impact an individual's ability to comprehend, navigate, and interact with web content effectively. Some common conditions and their associated challenges include:

- **Autism Spectrum Disorder (ASD):** Individuals with ASD may face difficulties in processing information, understanding context, and navigating websites. Complex layouts, sensory overload, and the inability to interpret social cues on web pages can create significant barriers to accessibility [10].

- **Dyslexia:** Dyslexic users experience challenges with reading, spelling, and comprehension. Poorly structured content, unclear typography, and lack of alternative text for images can make it difficult for them to access and understand web content effectively [10].

- **Attention Deficit Hyperactivity Disorder (ADHD):** People with ADHD often struggle with maintaining focus and attention. Websites with distracting elements, long paragraphs, and cluttered designs can exacerbate these challenges, making it difficult for users with ADHD to effectively navigate and engage with web content [10].

- **Intellectual Disabilities:** Individuals with intellectual disabilities may have limitations in understanding complex instructions or abstract concepts. Content that requires high cognitive load, jargon-heavy language, and lack of clear navigation can create significant barriers to web accessibility for this user group [10]

2.2 Current Assistive Technologies and Their Limitations

Screen Readers and Refreshable Braille Displays:

- Strengths:
    - Screen readers provide auditory output, enabling blind users to access text content.
    - Braille displays offer tactile feedback for reading.

- Limitations:
    - Complex visual layouts (e.g., tables, graphs) pose challenges.
    - Interactive elements (buttons, forms) may not be conveyed effectively.

Magnification and High-Contrast Tools:

- Strengths:
    - Magnification helps users with low vision read content.
    - High-contrast modes enhance visibility.
- Limitations:
    - Over-reliance on magnification can affect overall user experience.
    - High contrast may not address all accessibility needs.

2.3 Applications for Web Accessibility:

Generative AI models can be leveraged in various ways to enhance web accessibility, particularly for users with visual or cognitive/neurological impairments. Some applications are as follows :

- **Image Captioning:** By utilizing models like DALL-E and Stable Diffusion, generative AI can automatically generate accurate and detailed descriptions of images on web pages, benefiting users who rely on screen readers or braille displays to access visual content. These descriptions can adapt to the user's preferences and level of detail required, providing crucial context and meaning that would otherwise be inaccessible.

- **Content Summarization:** GPT-3 and other language models can be used to summarize complex articles or web pages, making it easier for users with cognitive or neurological impairments to comprehend and navigate the content. These summaries can be tailored to the user's needs, presenting the most relevant information in a concise and accessible format.

Fig. 3. The 'Alttext' (in red) vs. meaningful interpretations from Google's LLM PaLi [11]

- **Question Answering:** Generative AI models can be integrated into web browsers and assistive technologies to provide concise answers to user queries, enhancing the interactive experience for users with disabilities. [12] By leveraging natural language processing capabilities, these models can understand and respond to questions related to web content, facilitating better comprehension and navigation.

3. Proposed Approach

3.1 Image Captioning and Description using Generative AI:

Methodology and Models: Generative AI models have revolutionized the field of image captioning and description, enabling the automatic generation of accurate and contextually relevant textual descriptions from visual input. Two of the most prominent models in this domain are DALL-E and Stable Diffusion.

- **DALL-E:** Developed by OpenAI, DALL-E is a cutting-edge generative AI model that can generate high-quality images from textual descriptions, as well as generate descriptive text from given images.
- **Stable Diffusion:** Stable Diffusion, developed by Stability AI, is another powerful generative model that has gained significant traction in the image captioning and description domain, Similar to DALL-E.

Training Data and Preprocessing: To train generative AI models like DALL-E and Stable Diffusion for image captioning tasks, large-scale datasets containing image-caption pairs are required. Two of the most commonly used datasets in this domain are:

- **COCO (Common Objects in Context):** The COCO dataset is a widely used benchmark in computer vision and natural language processing tasks, including image captioning. It contains over 300,000 images with multiple captions per image, providing a diverse and challenging dataset for training generative AI models. [13]
- **Conceptual Captions:** The Conceptual Captions dataset, introduced by Google AI, is another valuable resource for training image captioning models. It consists of over 3 million image-caption pairs, covering a wide range of concepts and scenarios, making it a comprehensive dataset for training models to generate accurate and diverse captions. [14]

Preprocessing techniques are crucial for preparing the data for model training. These techniques typically involve resizing and normalizing the images, tokenizing the caption text into sequences of words or subword units, and creating input-output pairs that can be fed into the generative AI model during training.

Evaluation Metrics: To assess the performance and quality of image captioning models, several evaluation metrics are commonly used. These metrics measure different aspects of the generated captions, such as their accuracy, relevance, and diversity. Some of the most widely used evaluation metrics include:

- **BLEU (Bilingual Evaluation Understudy):** BLEU is a widely used metric that measures the n-gram overlap between the generated captions and the reference (ground truth) captions. [15] It provides a quantitative measure of how similar the generated captions are to the reference captions in terms of word choice and word order.

- **METEOR (Metric for Evaluation of Translation with Explicit Ordering):** METEOR is another popular metric that considers not only the n-gram overlap but also takes into account synonyms and paraphrases when evaluating the quality of generated captions. This makes it more robust to variations in word choice and phrasing, providing a more comprehensive assessment of caption quality. [16]

Integration with Web Browsers and Mobile Apps: To leverage the power of generative AI-powered image captioning and description capabilities for enhancing web accessibility, seamless integration with web browsers and mobile applications is crucial. This integration can be facilitated through various approaches, such as:

- **Browser Extension:** A Browser extension can be developed that uses the generative AI models for image captioning. This extension would run in the background and automatically generate captions or descriptions for images displayed on web pages, providing crucial context and information for users relying on assistive technologies like screen readers or braille displays.

- **Mobile Apps:** Similar integration approaches can be applied to mobile applications, where image description features powered by generative AI models can be incorporated. This would allow visually impaired users to access detailed and accurate descriptions of images or visual content displayed within the app, enhancing their overall experience and comprehension.

To ensure a seamless user experience, the integration of these generative AI capabilities should be accompanied by customizable settings and preferences, allowing users to adjust the level of detail, verbosity, and formatting of the generated captions or descriptions according to their individual needs and preferences.

4. Conclusion

4.1 Summary of Key Points

This paper has provided a comprehensive review of the challenges faced by individuals with visual impairments and cognitive/neurological disabilities in accessing and navigating web content. Despite the advancements in assistive technologies, significant barriers persist, hindering equal opportunities and digital inclusion for these user groups.

The limitations of current assistive technologies, such as screen readers, refreshable braille displays, magnification tools, and high-contrast modes, were highlighted. While these tools have been instrumental in improving web accessibility, they still struggle to convey complex visual information, handle dynamic content, and provide seamless navigation experiences.

Generative artificial intelligence (AI) models, including DALL-E, Stable Diffusion, GPT-3, and others, were identified as promising solutions to enhance web accessibility. These models possess the capability to generate accurate image descriptions, summarize content, and engage in natural language interactions, addressing some of the key challenges faced by users with disabilities.

Potential approaches were discussed, such as integrating generative AI-powered image captioning, content summarization, and question-answering capabilities into web browsers and mobile applications. These integrations could facilitate more inclusive and comprehensible digital experiences for users with visual, cognitive, or neurological impairments.

4.2 Implications and Potential Impact

The integration of generative AI technologies into web accessibility efforts holds significant implications and potential impact. By leveraging the power of these models, we can bridge the accessibility gap and create a more inclusive online environment for individuals with disabilities.

For users with visual impairments, generative AI-powered image descriptions could provide crucial context and meaning, enabling them to better comprehend and engage with visual content that would otherwise be inaccessible through screen readers or braille displays alone.

Similarly, for users with cognitive or neurological impairments, content summarization and question-answering capabilities could enhance their understanding and navigation of complex web pages, articles, and digital information. Tailored summaries and natural language interactions could simplify the comprehension process and facilitate more effective engagement with online content.

Beyond improving accessibility, the integration of generative AI could also foster greater independence, educational opportunities, and social inclusion for individuals with disabilities. Furthermore, the development and adoption of these innovative solutions could drive broader awareness and

prioritization of web accessibility considerations within the technology industry and society at large. As we strive to create a more inclusive digital landscape, the potential impact extends beyond the immediate beneficiaries, fostering a more equitable and accessible digital future for all.

REFERENCES :

[1] "Disability," World Health Organization. Accessed: Jun. 09, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/disability-and-health

[2] "Almost half of people with disabilities don't use the Internet: but why? | Oxford Internet Surveys - OxIS." Accessed: Jun. 09, 2024. [Online]. Available: https://oxis.oii.ox.ac.uk/blog/almost-half-people-disabilities-dont-use-internet-why/

[3] V. Tangcharoensathien, W. Witthayapipopsakul, S. Viriyathorn, and W. Patcharanarumol, "Improving access to assistive technologies: challenges and solutions in low-and middle-income countries," WHO South East Asia J Public Health, vol. 7, no. 2, 2018.

[4] L. Henneborn, "Designing Generative AI to Work for People with Disabilities." Accessed: Jun. 09, 2024. [Online]. Available: https://hbr.org/2023/08/designing-generative-ai-to-work-for-people-with-disabilities

[5] E. Collins and D. Eck, "Google I/O 2024: Introducing Veo and Imagen 3 generative AI tools," Google Blogs. Accessed: Jun. 09, 2024. [Online]. Available: https://blog.google/technology/ai/google-generative-ai-veo-imagen-3/

[6] "Google's Image-Captioning AI Is Getting Scary Good." Accessed: Jun. 09, 2024. [Online]. Available: https://www.popularmechanics.com/technology/robots/a23019/google-ai-captioning/

[7] "Blindness and vision impairment." Accessed: Jun. 09, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

[8] A. Hill, "How to Make the World Accessible for the Visually Impaired." Accessed: Jun. 09, 2024. [Online]. Available: https://www.visioncenter.org/resources/visual-impairment-accessibility/

[9] "Accessibility Requirements for People with Low Vision." Accessed: Jun. 09, 2024. [Online]. Available: https://www.w3.org/TR/low-vision-needs/

[10] "Diverse Abilities and Barriers | Web Accessibility Initiative (WAI) | W3C." Accessed: Jun. 09, 2024. [Online]. Available: https://www.w3.org/WAI/people-use-web/abilities-barriers/

[11] E. Segalis, G. Research, D. Valevski, D. Lumen, Y. Matias, and Y. Leviathan, "A Picture is Worth a Thousand Words: Principled Recaptioning Improves Image Generation," 2023, Accessed: Jun. 09, 2024. [Online]. Available: https://www.w3.org/TR/2016/NOTE-WCAG20-TECHS-

[12] "Generative Question Answering over Documents with LLMs." Accessed: Jun. 09, 2024. [Online]. Available: https://nexocode.com/blog/posts/generative-question-answering-llms/#revolutionizing-information-retrieval-the-power-of-llms-in-question-answering

[13] "Introduction to the COCO Dataset - OpenCV." Accessed: Jun. 09, 2024. [Online]. Available: https://opencv.org/blog/introduction-to-the-coco-dataset/

[14] P. Sharma, N. Ding, S. Goodman, and R. Soricut, "Conceptual Captions: A Cleaned, Hypernymed, Image Alt-text Dataset For Automatic Image Captioning," pp. 2556–2565, Accessed: Jun. 09, 2024. [Online]. Available: https://en.wikipedia.org/wiki/Alt

[15] Y. Yi, H. Deng, and J. Hu, "Improving Image Captioning Evaluation by Considering Inter References Variance," pp. 985–994, Accessed: Jun. 09, 2024. [Online]. Available: https://github.com/ck0123/improved-bertscore-for-

[16] "METEOR - Wikipedia." Accessed: Jun. 09, 2024. [Online]. Available: https://en.wikipedia.org/wiki/METEOR