Exploring AI Applications in Web Interaction Mode Design: Chatbot, Accessibility, and Search Engines

Zikai Song[1*]

[1]School of Information and Optoelectronic Science and Engineering, South China Normal University, Guangzhou, 510006, China
*20213231040@m.scnu.cn

Abstract. The rapid development of the web has made it one of the most important elements of modern daily life. One recent trending approach to implementing the web is using artificial intelligence (AI) technologies. AI tools are becoming more and more reliable, and using them to develop the web has become a practical and efficient scheme. This article delves into several applications of AI in web interaction designs, including chatbots, web accessibility, and generative search engines. Web accessibility is a rather broad topic, and thus this article shall concentrate on three typical techniques, including text-to-speech (TTS), automatic speech recognition (ASR), and automatic image caption. Chatbots represent a noticeable application of AI technologies; chatbots can facilitate interactions between AI and human users across web platforms. Web accessibility technologies empowered by AI algorithms provide different approaches for disabled users to access websites and thereby enable their access to web content. Generative search engines take advantage of AI to improve search results. Despite these advancements, challenges persist in the implementation of AI-driven web interaction designs. Several issues about each of the three technologies are emphasized in this article. Addressing these challenges requires much effort. Possible future directions for addressing these challenges are suggested in this article. By exploring these dimensions and charting future directions, this study aims to provide a basic understanding of AI's role in shaping the digital landscape and potential approaches to future enhancements.

Keywords: Artificial intelligence, Chatbot, Web accessibility, Generative search engine

1 Introduction

In the 1950s, the notion of artificial intelligence was introduced [1], which marked the beginning of a transformative journey that continues to shape various facets of modern life. Over the decades, AI has evolved in many aspects, and its trajectory is marked with many remarkable achievements. Recent years have witnessed a surge in innovative advanced AI technologies, which make AI one of the most trending topics. Key achievements include breakthroughs in natural language processing (NLP) [2],

Y. Wang (ed.), *Proceedings of the 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024)*, Advances in Intelligent Systems Research 185,
https://doi.org/10.2991/978-94-6463-512-6_6

the advent of large language models (LLM), and the proliferation of Reinforcement Learning techniques, Transformer architecture, etc.

The efficacy of AI in diverse areas has been aptly demonstrated, and the impact of AI has been profound and multifaceted. More and more developers in various fields are integrating AI tools into their projects [3-4]. AI technologies greatly impact many fields, and web interaction design is one of those fields.

Generally, traditional implementations of web designs, including interaction designs, require programmers to code by hand. Hence, traditional approaches tend to be error-prone and inefficient [5]. AI technologies bring solutions to such difficulties. Therefore, the application of AI technologies in web interaction modes has become one of the trending approaches to implementing or improving web interaction, as AI tools have been applied to various aspects of web interaction mode designs.

This article concentrates on several AI applications in web interaction mode design, with a particular focus on chatbots, web accessibility, and generative search engines. Chatbots are computer programs that chat with human users in natural languages. Therefore, AI's ability to recognize natural language makes it a reasonable underlying technology with which to develop chatbots.

Similarly, web accessibility has benefitted from AI-driven advancements. Web developers have developed various advanced techniques to improve web accessibility. However, many traditional techniques are relatively time-consuming and difficult to ensure accuracy [6]. However, the application of AI technologies introduces innovative solutions to various accessibility features, thereby improving usability for individuals with diverse needs. Furthermore, generative search engines are improvements to traditional search engines. Traditional search engines are also error-prone and time-consuming to implement by hand, and they cannot always return satisfying content.

Although AI technologies bring solutions to the challenges of traditional web interaction, they also introduce new problems that would not otherwise exist. This article shall present and discuss several issues, such as the ability of chatbots to conduct analyses, hallucination, accuracy of the recognition of speech and image, correctness of the generated responses, etc.

This article aims to present some of the current applications of AI technologies in chatbots, web accessibility, and generative search engines through a comprehensive review of existing literature. Web accessibility has many forms. Therefore, this article focuses on the following three kinds of web accessibility: TTS, ASR, and automatic image caption. More specifically, this review aims to provide brief introductions to each of the three technologies and discuss several typical challenges about them, as well as possible future directions to addressing these problems in hopes of inspiring further research and innovation, fostering the development of more inclusive, intuitive, and intelligent web interaction experiences.

2 AI-based Web Applications

Artificial intelligence stands as a prominent topic within computer science and has gained significant attention from several experts who have devoted themselves to this field. This section focuses on three typical applications of AI in web design: chatbots, web accessibility, and generative search engines.

2.1 AI-based Web Chatbot

The innovation of AI-based chatbots has had a large impact on many fields. In recent years, significant advancements have been witnessed in AI-based web chatbots. One notable example is ChatGPT, which has garnered considerable acclaim and popularity.

ChatGPT demonstrates proficiency in engaging with web users using natural language, which is made possible by NLP. Moreover, it possesses the ability to generate diverse content types, spanning from programs to movie scripts, owing to its advanced capabilities. Fig. 1 shows a piece of program generated by ChatGPT.

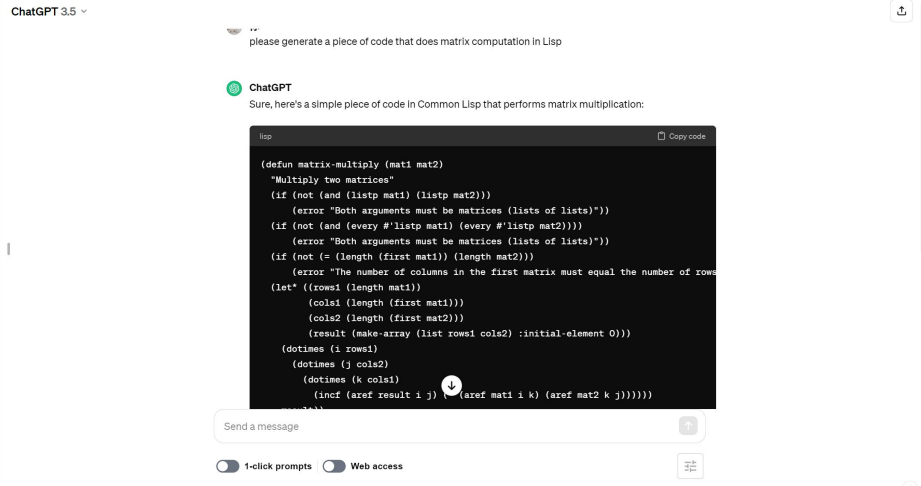

Fig. 1. Program computing matrix multiplication generated by ChatGPT（Picture credit: Original）

The utility of ChatGPT extends across various domains. For instance, recent research has applied ChatGPT in circuit theory teaching [7]. Other notable chatbots in existence include Bing, BERT, Erine Bot, Tencent Hunyuan, and RoBERTa. The advent of AI-based chatbots has profound influence.

2.2 AI-based Web Accessibility

Web accessibility serves as a methodology aimed at facilitating access to various web content for users with disabilities. Numerous standards have been proposed to ensure the creation of inclusive web environments. Recent standards include EN 301 549

V3.2.1 published by the European Telecommunications Standard Institute in 2021, and ISO/IEC 40500 published by ISO in 2019 [8].

As artificial intelligence continues to advance, using AI technologies becomes a natural choice to enhance web accessibility. It has been well established by a study that AI-based web accessibility technologies can effectively transform web content into various formats, thereby aiding impaired users in accessing web content. More explicitly, AI-based web accessibility technologies can now render web content accessible to users with cognitive impairments, hearing impairments, and other disabilities [8]. Fig. 2 shows Whisper's output, a speech recognition system developed by OpenAI.

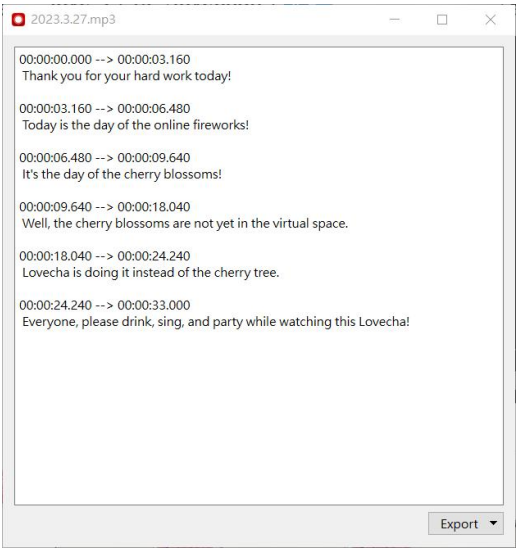

Fig. 2. Speech recognized by Whisper（Photo credit: Original）

2.3 AI-based Search Engines

Search engines play a vital role on the web, serving as the primary tool for users to navigate the internet. Therefore, the efficacy of search engines holds significant importance. Traditional search engines rely on crawling and ranking technologies to index and retrieve web pages relevant to a user's query. While traditional methods have been proven practical, there is room for improvement.

AI-based generative search engines have driven significant improvements. These engines offer a distinct advantage for their capacity to generate content in response to user queries rather than relying solely on pre-existing information like traditional search engines. Generative search engines are particularly valuable in situations where existing information fails to meet user requirements. Numerous generative search engines have emerged, including BingChat, Search Generative Experience (SGE), and NeevaAI. Some generative search engines are based on extensions to

language models; Fig. 3 shows one example, which is WebChatGPT, and, as suggested by the name, WebChatGPT is based on ChatGPT.

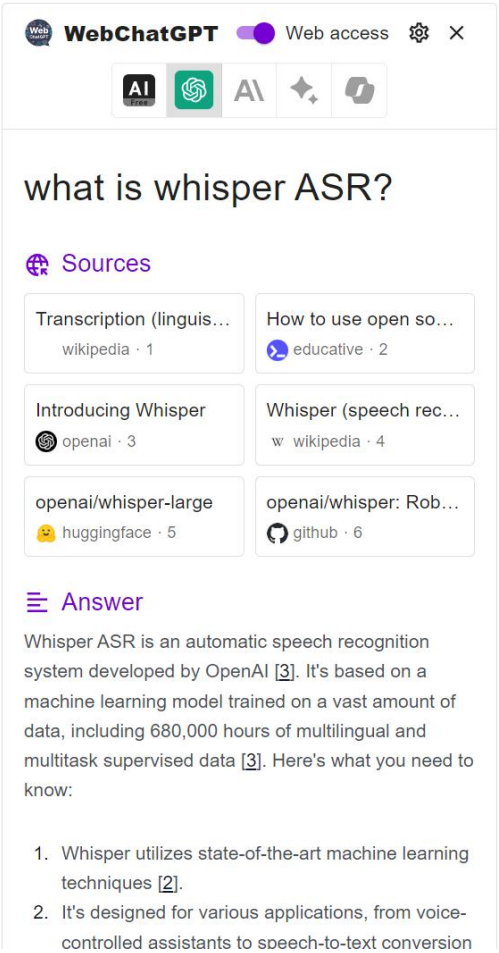

Fig. 3. Search results generated by WebChatGPT. last accessed 2024/04/09（Photo credit: Original）

3 Existing Issues and Possible Solutions

Despite the rapid advancement of artificial intelligence, the technology remains imperfect, and issues persist in its applications within web design. This section aims to highlight some of the challenges encountered in the three applications of artificial intelligence in web design discussed in the preceding sections. Additionally, possible future directions for each application are proposed in this paper.

3.1 Analysis of ChatGPT: An Example

Since there are many chatbots, this section shall focus on only ChatGPT and the issues concerning it, as well as potential approaches to mitigate the issues. However, the problems discussed herein are not exclusive to ChatGPT alone, and the suggested approaches may apply to other language models.

Issues in ChatGPT. While ChatGPT demonstrates proficiency in understanding user queries presented in natural language and executing various tasks such as solving mathematical problems and writing code in different programming languages, studies have highlighted certain limitations. Specifically, ChatGPT encounters difficulties when tasks necessitate analytical reasoning. For instance, a study revealed that ChatGPT may provide incorrect answers to certain mathematical problems posed by users. Furthermore, ChatGPT showed subpar performance even in graduate-level mathematical exams [9].

Mathematics is not the only field in which ChatGPT has rather poor performance. Recent research revealed its shortcomings in medical domains. Although ChatGPT exhibits high accuracy in responding to general medical topics, it struggles to deliver satisfactory responses requiring analytical capabilities [10].

Possible Solutions. To address these challenges, employing unsupervised approaches in AI model development could offer a solution. Chatbots are commonly developed with supervised approaches, in which the models are trained with labeled datasets. A recent study highlighted the limitations of datasets, emphasizing their inability to accurately reflect real-world scenarios regardless of their size [11]. Therefore, unsupervised approaches may hold an advantage over supervised methods.

Also, different datasets may be of assistance. The lack of ability to perform analyses could be resolved by using specialized data which contains information on how humans analyze problems and solve them to train the models. In other words, data that record cognitive information that captures processes of how human minds analyze problems. However, acquiring such data requires significant advancements in cognitive science. Further exploration into the fundamentals of human cognition may be imperative.

3.2 Analysis of AI-based Web Accessibility

Numerous AI-based web accessibility technologies have been proposed and extensively utilized. Nevertheless, challenges persist in this domain. This section discusses deficiencies in three common AI-based web accessibility technologies and possible directions to address these challenges. Sections 3.2.1 and 3.2.2 shall specifically examine fundamental technologies such as TTS, ASR, and Automatic Image Captioning.

Typical AI-powered Web Accessibility Technologies. *TTS.* Text-to-speech technology employs AI to convert written text into spoken language. TTS engines are commonly integrated into web browsers or assistive technology tools to offer audio feedback for users with visual impairments. In recent years, prevalent implementations of TTS rely on deep learning (DL) [12].

ASR. AI-driven speech recognition technology has various forms of applications. One typical application is to employ ASR to recognize speech and translate it into forms accessible to users with impaired hearing, like text. ASR also allows users to navigate websites and interact with content using voice commands.

Automatic Image Captioning. Automatic image captioning can analyze visual content and generate descriptive text. Automatic image captioning can greatly assist people with disabilities by providing access to visual information that they may otherwise struggle to interpret. For individuals with visual impairments, image captions offer a way to understand the content of images they encounter online or in other media. DL also serves as a prevalent underlying technology in this domain.

Shortages Related to AI-powered Web Accessibility Technologies. *Difficulties in TTS.* A challenge encountered in TTS technology is the difficulty in consistently producing speech that sounds natural. Moreover, issues regarding accuracy and dataset limitations have been reported in certain TTS systems [12].

Future directions may include incorporating complementary techniques to build the models and augmenting datasets to improve overall TTS performance. Incorporating complementary techniques may involve integrating additional algorithms or methodologies, such as prosody modeling or linguistic rules, to enhance the accuracy of the resulting speech. These solutions may bridge the gap between synthetic and natural speech, ensuring better accessibility and user experience in TTS applications.

Difficulties in ASR. A recent study suggests that users with hearing disabilities have found ASR technologies to be limited in accuracy. The same study also revealed that ASR implementations are generally inaccurate. For instance, ASR struggles with recognizing pronunciation with accents [13].

To tackle the accuracy issues in ASR, future directions could involve implementing more robust and inclusive training datasets that cover a diverse range of accents, dialects, and speech patterns. Also, advancements in machine learning (ML) techniques, particularly those exploiting deep learning and neural network architectures, could be applied to improving ASR accuracy by enhancing the model's ability to generalize various linguistic variations. ML could also be used to recognize less fluent speeches more accurately.

Furthermore, developing domain-specific ASR models and fine-tuning algorithms for specific applications or user groups could help mitigate biases and improve recognition accuracy in specialized contexts.

Difficulties in Automatic Image Caption. The following shortcomings of automatic image captioning have been emphasized [11]:
(1) Object hallucination: This occurs when the image captioning system generates descriptions for objects that are not in the image. It is often due to over-reliance on context or inaccurate object recognition.
(2) Illumination condition: Variations in lighting conditions can significantly impact the system's performance, like reduced contrast and shadows.
(3) Contextual understanding: This means comprehending the content of the given images, like the relationships between objects. However, models may struggle to grasp complex contextual relationships within images.
(4) Referring expression: Referring expressions pose a challenge as they require models to accurately identify and link objects in the image based on linguistic cues, enhancing caption quality by providing detailed descriptions for more accurate captions.

Future developments may benefit from contextual reasoning; techniques like graph-based representations or attention mechanisms could be applied to capture contextual dependencies and reduce object hallucination. Training on diverse datasets that cover a wide range of lighting scenarios could add robustness to lighting variations. Advanced techniques are also useful. Various advanced techniques are being developed, such as techniques to improve contrast, and restoring color, etc [11]. Unsupervised learning may ease the challenges related to contextual understanding and referring expressions.

3.3 Analysis of Generative Search Engines

Generative search engines can generate responses to the users. Thus, the correctness of the generated responses is crucial. Regretfully, the correctness remains to be improved. In this section, some typical difficulties encountered in generative search engines are presented and discussed.

Difficulties in Generative Search Engines. *Low Verifiability.* Some generative search engines include citations in their responses to enhance verifiability, as shown in Fig. 3. However, such generative search engines are lack of verifiability in a recent study. Although the generated responses are seemingly informative and natural, citations are far from accurate and supportive on average. In other words, the citations are not adequate, if at all, to support the response [14].

Hallucinatory Content. Hallucination poses a challenge for both image caption and generative search engines. In the context of generative search engines, hallucination refers to responses generated without proper consideration of meaning, resulting in improper and irrelevant content. This type of content does not have a direct basis but is made up. Hallucinatory content is not practical, even if it is coherently based on reliable sources.

Possible Solutions. Several techniques have been proposed to tackle these problems. Techniques like reinforcement learning are being developed to improve verifiability; also, new evaluation protocols aimed at achieving the same goal have been proposed [14]. Future developments may include improving mechanisms to allow users to provide feedback on the accuracy and supportiveness of citations.

Researchers are actively exploring advanced techniques to mitigate hallucination, such as Refusal-Aware Instruction Tuning. Future directions may also include improving evaluation protocols for the content of the generated responses. Also, employing unsupervised approaches to building the models may reduce hallucination.

4 Conclusion

This article presents three types of applications of AI tools in web interaction mode designs. The technologies include chatbot, web accessibility, and generative search engines. However, not all forms of web accessibility are covered; this article concentrates on only three forms: TTS, ASR, and automatic image caption, which can be used to ease web accessing for web users with issues like visual impairment, hearing impairment, and cognitive impairment.

Furthermore, the challenges of each of the three types of three types of web interaction technologies are discussed. Chatbots cannot perform complex and accurate analyses. As for web accessibility technologies: TTS is not able to produce natural speech; ASR suffers from low accuracy; automatic image caption suffers from object hallucination, poor illumination condition, limited contextual understanding, and limited ability to handle referring expressions. Generative search engines may generate content with low verifiability or even hallucinatory content.

There are some obvious challenges with existing technology. For instance, hallucination affects both automatic image caption and generative search engines. More precisely, the output of an automatic image caption program may include a description of objects that are not actually in the image, and the content generated by generative search engines could include events or facts that are not in reality.

Approaches to solving these challenges are necessary for web interaction technologies to be reliable and practical. Aiming to provide some guidance, possible future research directions are shown in the article.

Web interaction mode design is a vast topic; any design must be sufficiently complicated to meet the requirements of users with diverse needs. Therefore, the coverage of this article is far from enough to fully demonstrate the significant impact brought to the field of web interaction mode designs by the innovations of AI tools. Also, the demonstrations of the challenges relating to each of the technologies described in the article are limited, as only a few challenges are mentioned.

In the future, a more detailed re-examination of network interaction pattern design will be conducted. Innovative application solutions of artificial intelligence in this field should be studied to deepen the understanding of related research.

References

1. Abou-Zahra, S., Brewer, J., Cooper, M.: Artificial Intelligence (AI) for Web Accessibility: Is Conformance Evaluation a Way Forward? In: Proceedings of the 15th International Web for All Conference, pp. 1-4. (2018).
2. Chen, Y., Wang, H., Yu, K., Zhou, R.: Artificial Intelligence Methods in Natural Language Processing: A Comprehensive Review. Highlights in Science, Engineering and Technology 85, 545-550 (2024).
3. Konya, A., Nematzadeh, P.: Recent applications of AI to environmental disciplines: A review. Science of The Total Environment 906, 167705 (2024).
4. Xu, R., Sun, Y., Ren, M., Guo, S., Pan, R., Lin, H., Sun, L., Han, X.: AI for social science and social science of AI: A survey. Information Processing & Management 61(3), 103665 (2024).
5. Muthazhagu, V. H., Surendiran, B.: Exploring the Role of AI in Web Design and Development: A Voyage through Automated Code Generation. In: 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), pp. 1-8. IEEE, 2024.
6. ARAa, J., Cecilia, S. L.: Artificial intelligence in web accessibility: potentials and possible challenges. In: Proceedings of IAC, 173 (2022).
7. Parra, J., Sanchis, P.: Development of a ChatGPT Web-Based Educational Chatbot for Circuit Theory Teaching. In: International Technology, Education and Development Conference2024 Proceedings, pp. 3062-3062. IATED, 2024.
8. Nacheva, R., Jansone, A.: Heuristic Evaluation of AI-Powered Web Accessibility Assistants. Baltic Journal of Modern Computing 11(4), 542-557 (2023).
9. Frieder, S., Pinchetti, L., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P., Berner, J.: Mathematical capabilities of ChatGPT. Advances in Neural Information Processing Systems 36 (2024).
10. Nastasi, A. J., Courtright, K. R., Halpern, S. D., Weissman, G. E.: Does ChatGPT provide appropriate and equitable medical advice? A vignette-based, clinical evaluation across care contexts. Preprint MedRxiv, 2023-02 (2023).
11. Ghandi, T., Pourreza, H., Mahyar, H.: Deep learning approaches on image captioning: A review. ACM Computing Surveys 56(3), 1-39 (2023).
12. Kumar, Y., Koul, A., Singh, C.: A deep learning approaches in text-to-speech system: a systematic review and recent research perspective. Multimedia Tools and Applications 82(10), 15171-15197 (2023).
13. Kuhn, K., Kersken, V., Reuter, B., Egger, N., Zimmermann, G.: Measuring the accuracy of automatic speech recognition solutions. ACM Transactions on Accessible Computing 16(4), 1-23 (2024).
14. Liu, N. F., Zhang, T., Liang, P.: Evaluating verifiability in generative search engines. Preprint arXiv: 2304.09848 (2023).