Information Research - Vol. 30 No.iConf (2025)

Research on the influencing mechanism of blind or visually impaired persons' evaluation on generative AI in visual tasks

Huitong Chen, Yuting Pan, and Hui Yan

DOI: https://doi.org/10.47989/ir30iConf47110

Abstract

Introduction. Generative AI (GAI) has shown significant potential in assisting blind or visually impaired (BVI) Persons in visual tasks. However, existing evaluations of GAI tend to focus on technical performance, overlooking the specific usage contexts and experiences of BVI users.

Method. This study conducted action research and semi-structured interviews with 19 BVI persons, to explore their evaluations of GAI in visual tasks and the influencing mechanism of their evaluations.

Analysis. Following grounded theory, we identified 16 categories, and corresponding 5 core categories, as well as their relationships.

Results. The findings indicate that BVI persons primarily evaluate GAI based on three criteria: accessibility, credibility, and interactivity. Their evaluation is influenced by four main factors: system, information, BVI user, and context. Notably, both BVI user and contextual factors moderate the influence of the system and information on user evaluation.

Conclusions. This study develops a model that explains the influence mechanism behind the evaluation on GAI by BVI persons in visual tasks. It not only broadens the scope of human-AI interaction research by incorporating diverse user types and task contexts, but also provides an empirical foundation for developing human-centered GAI.

Introduction

There have been numerous assistive technologies designed for enhancing the accessibility of BVI persons. The booming development of GAI has ushered in new possibilities for advancements in assistive technology. In 2023, Be My Eyes announced the successful deployment of Be My AI, a visual assistance tool powered by OpenAI's GPT-4 vision model (Be My Eyes, 2023). Beyond the traditional way of resorting to volunteers, BVI persons can now utilize Be My AI to process visual information, which empowers them with greater independence, and improving well-being in individual and social terms (Bendel, 2024).

However, despite the proliferation of assistive technology for BVI persons in the market, their effectiveness and satisfaction in the actual lives of BVI persons remain limited (Khan & Khusro, 2021). While benchmarks in computer science effectively evaluate the performance of GAI in visual tasks, these tests tend to focus on the accuracy and efficiency of the models while ignoring the specific usage contexts and experiences of BVI persons. This disparity can lead to the development of technologies that are difficult to meet the real needs of BVI persons.

Therefore, this study examines the visual task encountered by BVI persons to investigate their evaluations on GAI in these tasks and explore the influencing mechanism of their evaluations. This not only provides a new theoretical perspective for research in human-intelligent interaction, but also introduces a BVI persons-centred approach to technology-driven model evaluation, offering valuable practical insights for the advancement of human-centred GAI.

Literature review

Evaluation of assistive technologies for BVI persons

Currently, an increasing number of assistive technologies based on artificial intelligence (AI) with computer vision (CV) are under development. There have been accumulated specific research findings about the evaluation of these technologies, which can be categorized into two orientations: technology-driven and user demand-driven. Among them, the former dominates the mainstream, focusing on the performance and practical implementation of assistive technologies.

Scholars in the field of computer science have constructed many benchmark datasets for vision-language models (Dai et al., 2023), and conducted the benchmark evaluation. Two key dimensions are involved in these evaluation metrics: automatic evaluation metrics, which principally include generated responses, Recall-Oriented Understudy for Gisting Evaluation (ROUGE), and BERT-based semantic textual similarity; human evaluation metrics which focus on the three main dimensions of correctness, actionability, and fluency (Yang et al., 2024; Zhao et al., 2024).

The user demand-driven evaluation approach focuses on the feedback and user experience from BVI persons. This feedback includes a rich set of metrics, including satisfaction (Rattanaphinyowanich & Nunta, 2021), accuracy, reliability, accessibility, privacy, security, compatibility, energy efficiency, usability (Bhagat et al., 2024), functionality, aesthetics characteristics and social acceptability (Hamilton et al., 2016; Phillips et al., 2018). In the field of library and information science, the information needs and information behavior of BVI persons have received particular attention (Berget & MacFarlane, 2020). Research has covered various topics, including the selection of information sources (Rahman et al., 2017; Chen et al., 2024), information-seeking behavior (Williamson et al., 2000), and interactions with information retrieval systems (Xie et al., 2021; Berget & MacFarlane, 2020). Among them, the accessibility and usability issues faced by BVI persons are the primary focus (Xie et al., 2020). These studies can be classified into two categories according to their research approaches: for one thing, examining established evaluation criteria, such as examining whether WCAG 2.2 effectively supports BVI persons' access to digital libraries from the perspectives of stakeholders including users, experts, and developers (Xie et al., 2022). For another, conducting empirical research on the challenges and perspectives

faced by BVI persons in interacting with information retrieval systems. The research findings show that while the library websites are accessible according to the extracted indices of W3C, empirical data from BVI users (i.e. successful task completion, working time, satisfaction level) suggests that the websites are not easy to use (Najafgholinejad, 2024).

Research on human-AI interaction experience

With the development of AI technology, the focus of HCI research is shifting from human interaction with non-AI computing systems to human interaction with AI systems, thus giving rise to the cutting-edge topic of human-AI interaction (HAII) (Jiang et al., 2024). Researchers have conducted extensive and thorough explorations around human interactions with AI systems such as chatbots, voice assistants, virtual humans, autonomous vehicles, etc., with application scenarios covering shopping, healthcare, and transport sector, and more. As the two main elements in human-intelligent interaction, user characteristics and AI characteristics are the focus in the research of human-intelligent interaction experience. Among them, the inherent characteristics of individual users, such as gender identity, political ideology (Molina & Sundar, 2024), social role, user autonomy (Huh et al., 2023), motivation and social presence (Shao & Kwon, 2021), health condition (Esmaeilzadeh et al., 2021) can affect users' perception and evaluation of AI, as well as their willingness to adopt. Users' evaluation or perception of AI can be expressed in terms of ease of use (Loske & Klumpp, 2021), anthropomorphism (Pelau et al., 2021), etc. Characteristics of AI such as effects of explanation and synchronization (Fan et al., 2022), roles (Liao & Sundar, 2021) and communication models (Lew & Walther, 2023) also affect the experience of human-AI interaction. Furthermore, few studies have explored the role of task as a component in human-computer interaction. For example, people are less supportive of AI and its creators when AI performs highly (vs low) hedonic tasks (Yanit et al., 2023).

In summary, there is a relative scarcity of studies focusing on BVI users' evaluation on assistive technologies based on AI with CV. Research in the HAII domain has focused on the influence of both humans and AI on the human-AI interaction experience, and has conducted preliminary exploration into the mechanisms between each element. However, there is still a relative lack of research on the interaction between BVI persons and AI, especially on HAII in visual tasks. To fill this research gap, this paper aims to address two specific research questions: ***What are the evaluation criteria of BVI persons towards GAI in visual tasks? What are the underlying mechanisms influencing their evaluation?***

Research methodology

Data collection

With the help of the Capital Library and China Association of persons with visual disabilities, we recruited 19 participants, and their characteristics are summarized in Table 1. Among them, minor participants will take part in the interviews accompanied by their parents. We ensured that all the participants gave their voluntary consent based on full understanding.

For the 12 participants who had never used AI before, we conducted one-on-one '*Be My Eyes*' operation training for them at the Capital Library during March 2024. In the training, we collected behavioral data through participatory observation. After the training, we conducted interviews to gather participants' perceptions and evaluations of Gen AI, as well as the factors influencing their views. For the 7 participants who have experience in using GAI, we employed a semi-structured interview approach for data collection, questions centred around: visual challenges in daily life; experiences using GAI for visual tasks; evaluations of GAI and the influencing factors.

Categories	Subcategories	Quantity	Categories	Subcategories	Quantity
Gender	Male	9	Blindness Severity	Totally Blind	12
	Female	10		Low Vision	7
Age Group	9-18 Years	7	Experience in Using GAI	Never Use	12
	19-60 Years	5		Occasional Use	3
	41-40 Years	2		Regular Use	4
	Over 60 Years	5			

Table 1. Basic characteristics of interviewees

Data analysis

This study follows the principle and requirement of grounded theory. Data collection and analysis were alternately reciprocal and continuously comparative. Data were organized and coded immediately after each technical training or interview. The participant recruitment ended when we reached theoretical saturation. After three coding stage-open coding, axial coding, and selective coding (Strauss & Corbin, 1990), we identified 16 categories, and corresponding 5 core categories. The 5 core categories include system, information, BVI user, context and evaluation. Based on the coding results, a theoretical model is proposed (Figure 1).

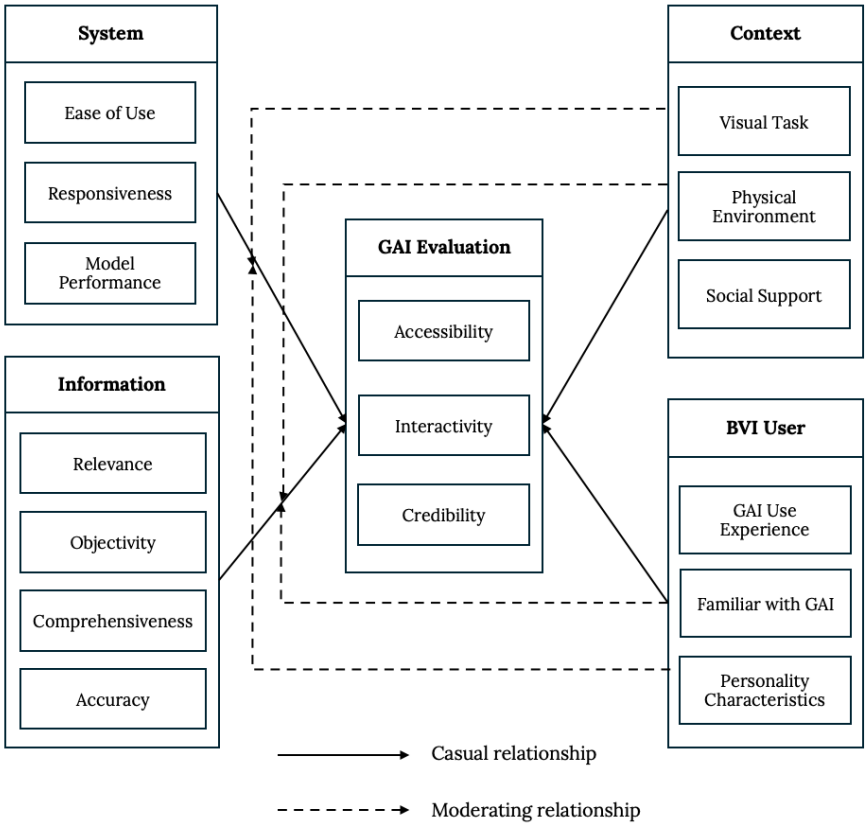

Figure 1. Model of the influence mechanism of BVI persons evaluation on GAI in visual tasks

Findings

Evaluation criteria of GAI by BVI persons

The research found that respondents' evaluation on GAI's performance in visual tasks primarily focused on accessibility, interactivity, and credibility. Specifically, this manifested as: (1) the degree of accessibility and operability of GAI at the physical level and its comprehensibility at the intellectual level. (2) The degree to which GAI can offer efficient and personalized information

delivery and feedback interface. (3) The degree to which GAI is honest or truthful, and can be trusted. This is specifically manifested in GAI's fairness and interpretability.

Influence of system quality on GAI evaluation

The system quality of GAI is the basic guarantee for providing visual information services to BVI persons. It is further categorized into ease of use, responsiveness and model performance.

Ease of use represents the technical threshold of GAI, which directly affects interviewees evaluation of accessibility. The accessibility of GAI increases as the download, registration, and operating procedures of GAI get simpler, and as the number of voice prompts it offers during the capture process increases. As Interviewee #01 noted, '*it is difficult for us to take clear photos and the operation is difficult for us.*'

Responsiveness refers to the efficiency and effectiveness of the system in responding to BVI persons requests, which greatly influences BVI persons evaluation of GAI's interactivity. During training, most interviewees had a particularly negative impression of the interactivity of tools like Be My Eyes due to their long processing time for images and frequent response failures. interviewee#11, who has used multiple GAI tools, had a better interactive experience in using the Luomo Toolbox, noted '*it is instantaneous and it only requires camera alignment to achieve immediate recognition.*'

Model performance refers to the capacity of the visual-verbal model embedded in the AI tool to process visual information. This ability influences BVI persons evaluation of GAI's credibility and interactivity through the quality of the generated information content.

Influence of information quality on GAI evaluation

The accuracy, objectivity, and comprehensiveness of information are essential for building trust in GAI among BVI persons. Additionally, the relevance of information influences their evaluation of the interactivity of GAI.

Accurate information is crucial for the safety and autonomy of BVI persons, as it forms the foundation for their reliance on GAI in decision-making. A model's ability to accurately identify and describe visual content directly correlates with the level of trust it can foster among BVI persons. Most interviewees express concern over the prevalent issue of AI hallucinations, hesitant to rely on GAI for some critical and extremely accurate processing of visual data.

The objectivity of the information described by the model also influences users' assessment of credibility. Interviewees #02 and #10 believe that AI is less susceptible to subjective feelings and cognitive limits when compared to humans, thus allowing for a more accurate and objective reflection of visual information. This perception leads them to view GAI as fair and inclusive.

Comprehensive information offers broad and complete coverage of the necessary or expected aspects of a problem or subject. Interviewees believe that the more detailed the GAI's description of visual information, the clearer their panoramic understanding of the environment, which in turn enhances their evaluation of the credibility of GAI. Interestingly, the comprehensive information provided by AIGC can create valuable Information Encountering. For instance, interviewee #10 casually used Be My Eyes to take a photo of the front and from AI's detailed description, he learned there was a bench ahead, allowing him to sit down and take a break. This unexpected help provided practical assistance and enhanced his trust in GAI.

The relevance of generated information content is manifested in its relevance to the situation and emotions of BVI persons. The greater the relevance of information, the better GAI can understand and respond to the personalized needs of BVI persons, resulting in higher user ratings of interactivity. For example, in a cooking scenario, interviewees only need GAI to quickly capture the

names and expiration dates of spices. However, the comprehensiveness of AI-generated information can make it difficult for BVI persons to quickly access the required information, thus reducing interaction efficiency. In reading scenarios, both interviewee#7 and interviewee#2 agreed that the GAI's voice '*sounds like a customer service, devoid of any emotional expression in its reading*'. This mechanized presentation provides a poor interactive experience for BVI persons.

Influence of user characteristics on GAI evaluation

User dimensions include their GAI use experience, familiarity with GAI, and personality characteristics. As users accumulate experience, BVI persons gradually become more familiar with the operation and functions of GAI, which in turn enhances their evaluation of accessibility and interactivity. In terms of personality, some interviewees consider themselves conservative and cautious, which leads to a natural distrust of new technologies such as GAI. On the other hand, interviewees with more positive and open-minded personalities tend to rate GAI with higher credibility.

Additionally, user dimensions can moderate the intensity and direction of the influence that system and information factors have on user evaluations. Users with prior experience, familiarity with GAI, and open-minded personality are more likely to be satisfied with the system and information provided by GAI, leading to more positive evaluations. For instance, experienced interviewees are generally accustomed to the response time of GAI, leading to higher evaluations of its interactivity.

Influence of contextual factors on GAI evaluation

The visual task, physical environment and social relationships experienced by BVI persons directly influence their assessment of GAI. Additionally, contextual factors can moderate the intensity and direction of the influence that system and information factors have on user evaluations.

In visual tasks that are urgent and important related to personal safety, such as reading hospital reports, medication information, or crossing the street, BVI persons may perceive that the responsiveness of GAI and the quality of the information it provides are insufficient to meet their need, leading to a decrease in their evaluation of its credibility and interactivity. However, in non-urgent and less important visual tasks, such as ordering food, checking nail designs, or picking up keys, BVI persons believe that using GAI reduces their reliance on and disruption to others, resulting in higher evaluations of its accessibility.

The physical environment refers to the objective conditions under which BVI persons encounter visual challenges, including sound environment, physical layouts, and network infrastructure. BVI persons primarily rely on auditory feedback to compensate for visual deficiencies. In noisy environment, they find GAI less accessible and interactive. The layout and design of the physical environment directly affect BVI persons' interactions with GAI. Many interviewees indicated that if commonly used items and devices in their home or workplaces are in fixed position, they would be able to take more focused photos, thus facilitating efficient information transfer and feedback with GAI. The quality of network infrastructure impacts the responsiveness of GAI systems, which in turn impacts users' perceptions of the accessibility and interactivity of Gen AI.

Social relationships, particularly strong ties, have a direct bearing on BVI persons evaluation of GAI's accessibility. Different parenting styles exert vastly different influences. For instance, the parents of interviewee#15 hold a very open-minded attitude towards AI and even suggested creating an AI chat group for her child after training, enabling interviewee#15 to successfully access and use GAI. In contrast, parents of interviewee#16 adopted strict control over their children, prohibiting interviewee#16 from accessing and using ICT. Consequently, interviewee#16 perceives restricted access to GAI.

Conclusion

This study delves into the evaluation criteria and influencing mechanisms of GAI from the perspective of BVI persons in visual tasks. The findings indicate that BVI persons' evaluations of GAI are primarily based on three criteria: accessibility, credibility, and interactivity. As traditional human-computer interaction transitions to human-AI interaction, these three evaluation criteria evolve to incorporate new dimensions. BVI persons' evaluation are directly influenced by four dimensions: system, information, user, and context. Furthermore, the user and context factors moderate the intensity and direction of the influence that system and information have on user evaluations. Specifically, ease of use and responsiveness of the system influence evaluations of accessibility and interactivity. The model performance influences evaluations of credibility and interactivity through the quality of the generated content. Additionally, user's experience and familiarity with GAI influence their evaluations of accessibility and interactivity. Visual tasks with high urgency and importance requirements tend to decrease the perceived credibility and interactivity of GAI. The physical environment impacts all three evaluation criteria, while social relationships primarily affect the evaluation of accessibility.

The main contributions of this study are as follows: It focuses on the interaction between BVI persons and GAI in visual task contexts, enriching the user types and diversifying task contexts in human-AI interaction research. Moreover, in the technology-driven domain of GAI evaluation studies, this study provides a comprehensive user-centred perspective based on grounded theory, shedding light on the factors and mechanisms that influence BVI persons' evaluations of GAI. However, the study has certain limitations. As an exploratory study, its primary focus is on identifying new variables and relationships. The stability of these relationships, however, requires further validation through larger sample sizes in future research.

Acknowledgements

The research is supported by National Social Science Foundation of China (Grant No. 23AZD093). This paper is one of the key projects of the National Social Science Fund of China's 'The Disruptive Impact of artificial intelligence on the National Economy and Research on Information Governance'. In addition, we would like to express our gratitude to Qihao Wang for inspiring the topic selection, as well as for the assistance provided during the data analysis process. We are also deeply thankful to the librarians of the Capital Library, Xuefeng Zhao, Qiangdong Li, Jing Su, Ruoxuan Yang, and Puhang Yu, for their unwavering support and assistance during our technical training sessions.

About the authors

Huitong Chen is a PhD student in Information Science at the School of Information Resource Management, Renmin University of China. Her research interests include the social impact of Artificial Intelligence, information behavior, and digital inclusion. She can be contacted at chenhuitong@ruc.edu.cn

Yuting Pan is a PhD student in Information Science at the School of Information Resource Management, Renmin University of China. Her research interests focus on community informatics. She can be contacted at sophie_pyt@163.com

Hui Yan is Professor and Doctoral Supervisor at the School of Information Resource Management, Renmin University of China. His research interests include the social impact of Artificial Intelligence, community informatics and digital inequality. He can be contacted at hyanpku@ruc.edu.cn

References

Be My Eyes (2023). Be My Eyes Integrates Be My AI™ into its First Contact Center with Stunning Results. https://www.bemyeyes.com/blog/introducing-microsofts-ai-powered-disability-answer-desk-on-be-my-eyes

Bendel, O. (2024). How Can Generative AI Enhance the Well-being of Blind? In Proceedings of the AAAI Symposium Series, 3(1), 340-347.

Berget, G., & MacFarlane, A. (2020). What is known about the impact of impairments on information seeking and searching? Journal of the Association for Information Science and Technology, 71(5), 596-611.

Bhagat, S., Joshi, P., Agarwal, A., & Gupta, S. (2024). Accessibility evaluation of major assistive mobile applications available for the visually impaired. arXiv preprint arXiv:2407.17496.

Chen, H., Yan, H., Wu, Z., & Zhao, X. (2024). Silicon-based Life or Carbon-based Life? An Exploratory Study on Visual Information Source Selection of Blind or Visually Impaired Persons. *Proceedings of the Association for Information Science and Technology*, 61(1), 493-498.

Dai, W., Li, J., Li, D., Tiong, A. meng huat, Zhao, J., wang, W., li, B., Fung, P., & Hoi, S. (2023). InstructBLIP: Towards General-Purpose Vision-Language Models with Instruction Tuning. ArXiv, 2305(06500v2). https://doi.org/10.48550/arXiv.2305.06500

Esmaeilzadeh, P., Mirzaei, T., & Dharanikota, S. (2021). Patients perceptions toward human–artificial intelligence interaction in health care: experimental study. Journal of medical Internet research, 23(11), e25856.

Fan, M., Yang, X., Yu, T., Liao, Q. V., & Zhao, J. (2022). Human-ai collaboration for UX evaluation: effects of explanation and synchronization. Proceedings of the ACM on Human-Computer Interaction, 6(CSCW1), 1-32.

Hamilton-Fletcher, G., Obrist, M., Watten, P., Mengucci, M., & Ward, J. (2016). 'I Always Wanted to See the Night Sky' Blind User Preferences for Sensory Substitution Devices. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 2162-2174).

Huh, J., Whang, C., & Kim, H. Y. (2023). Building trust with voice assistants for apparel shopping: The effects of social role and user autonomy. Journal of Global Fashion Marketing, 14(1), 5-19.

Jiang, T., Sun, Z., Fu, S., & Lv, Y. (2024). Human-AI interaction research agenda: A user-centered perspective. Data and Information Management, 100078.

Khan, A., & Khusro, S. (2021). An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities. Universal Access in the Information Society, 20(2), 265-298.

Lew, Z., & Walther, J. B. (2023). Social scripts and expectancy violations: Evaluating communication with human or AI chatbot interactants. Media Psychology, 26(1), 1-16.

Liao, M., & Sundar, S. S. (2021). How should AI systems talk to users when collecting their personal information? Effects of role framing and self-referencing on human-AI interaction. In Proceedings of the 2021 CHI conference on human factors in computing systems (pp. 1-14).

Loske, D., & Klumpp, M. (2021). Intelligent and efficient? An empirical analysis of human–AI collaboration for truck drivers in retail logistics. The International Journal of Logistics Management, 32(4), 1356-1383.

Molina, M. D., & Sundar, S. S. (2024). Does distrust in humans predict greater trust in AI? Role of individual differences in user responses to content moderation. New Media & Society, 26(6), 3638-3656.

Najafgholinejad, A. (2024). Accessibility and usability of user interfaces of library information retrieval systems from the perspective of visually impaired users. Library and Information Sciences, 26(4), 83-110.

Pelau, C., Dabija, D. C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. Computers in Human Behavior, 122, 106855.

Phillips, M., & Proulx, M. J. (2018). Social interaction without vision: an assessment of assistive technology for the visually impaired. Technology & Innovation, 20(1-2), 85-93.

Rahman, A. L. A., Razali, T. R. a. T., Ghazali, A. M., & Kamarudin, M. H. (2017). BVIC's CIS in the technological environment. Advanced Science Letters, 23(1), 151-155.

Rattanaphinyowanich, T., & Nunta, S. (2021). Development of DAISY-WIBORD as computer assisted learning facilities for children with visual impairment. Journal of Physics, 1835(1), 012080.

Shao, C., & Kwon, K. H. (2021). Hello Alexa! Exploring effects of motivational factors and social presence on satisfaction with artificial intelligence-enabled gadgets. Human Behavior and Emerging Technologies, 3(5), 978-988.

Strauss, A., & Corbin, J. M. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Sage Publications, Inc.

Williamson, K., Schauder, D., & Bow, A. (2000). Information seeking by blind and sight impaired citizens: an ecological study. Information research, 5(4), 5-4.

Xie, I., Babu, R., Lee, T. H., Castillo, M. D., You, S., & Hanlon, A. M. (2020). Enhancing usability of digital libraries: Designing help features to support blind and visually impaired users. Information Processing & Management, 57(3), 102110.

Xie, I., Babu, R., Lee, H. S., Wang, S., & Lee, T. H. (2021). Orientation tactics and associated factors in the digital library environment: Comparison between blind and sighted users. Journal of the Association for Information Science and Technology, 72(8), 995-1010.

Xie, I., Babu, R., Wang, S., Lee, H. S., & Lee, T. H. (2022). Assessment of digital library design guidelines to support blind and visually impaired users: a study of key stakeholders' perspectives. The Electronic Library, 40(6), 646-661.

Yang, B., He, L., Liu, K., & Yan, Z. (2024). VIAssist: Adapting Multi-modal Large Language Models for Users with Visual Impairments. arXiv preprint arXiv:2404.02508.

Yanit, M., Yanit, M., & Wan, F. (2023). Right agent, wrong level of hedonism: How high (vs low) hedonic values in AI-performed tasks lead to decreased perceptions of humanlikeness, warmth, and less consumer support. Computers in Human Behavior, 147, 107870.

Zhao, Y., Zhang, Y., Xiang, R., Li, J., & Li, H. (2024). Vialm: A survey and benchmark of visually impaired assistance with large models. arXiv preprint arXiv:2402.01735.