EMusA: Enhancing Museum Accessibility through an Integrated Conversational AI Agent

Yael Avni[1], Juliana Najib Abushkara[1,†], Ahed Mohammad Zoabi[1,†], Lior Maman[1], Alexandra Danial-Saad[2], Tsvi Kuflik[1,*] and Oshrat Ayalon[1]

[1]*Information Systems Department, Faculty of Computer & Information Science, The University of Haifa, Israel*
[2]*Department of Occupational Therapy, Faculty of Social Welfare & Health Sciences, The University of Haifa, Israel*

Abstract

Ensuring accessible, meaningful, and autonomous museum experiences for blind and low-vision (BLV) visitors remains a persistent challenge. In response, cultural institutions increasingly seek to integrate digital and multimodal technologies to enhance accessibility and enrich museum experiences. This paper outlines the design of an interactive tangible system that will enable BLV visitors to explore 3D-printed replicas of museum artefacts and access contextual information through two complementary modalities: (1) structured, curator-approved pre-recorded audio descriptions (AD) triggered by a button press, and (2) open-ended voice queries answered by a dedicated conversational artificial intelligence (AI) agent. The system consists of a rotating wooden base containing four tactile artefacts, each associated with an individual sensor that automatically play an initial AD when lifted. Visitors may then request additional information at any point via an integrated button, which either delivers layered narrative AD or initiates a voice-based question for AI-generated responses. The forthcoming study will involve BLV participants and will investigate three core questions: the usability and clarity of the interaction sequence (tactile exploration, button activation, and voice recording), user preferences between structured narration and open conversational access, and participants' perceived autonomy and self-efficacy when engaging with the system. Data will be gathered through structured observations, pre- and post-interaction questionnaires, and qualitative interviews. This research aims to advance understanding of how tangible interfaces and conversational AI can support more inclusive, engaging, and user-directed museum experiences for BLV audiences.

Keywords

Conversational Agents, Natural Language Interaction, Multimodal Interfaces, Accessibility, User Evaluation, Cultural Heritage

1. Introduction

Museums serve as key cultural institutions, offering spaces for learning, reflection, and shared cultural engagement [1]. Ensuring that these environments are accessible to all visitors is therefore a central responsibility. Yet for blind and low-vision (BLV) individuals, who constitute an estimated 2.2 billion people worldwide [2], museum visits remain challenging, as exhibitions continue to rely heavily on visual modes of interpretation. This results in persistent barriers and low participation rates among BLV communities [3]. Over the past decade, research has increasingly focused on developing multimodal, user-centered solutions that support accessible museum experiences [4, 5]. A recent systematic review by Avni et al. [6] maps the landscape of technologies and frameworks developed to enhance the museum experience for BLV visitors, providing a foundation for the present work. Studies emphasize that effective interpretation for BLV visitors must integrate touch and audio as multisensory exploration fosters perception, memory, and spatial understanding [7]. Advances in digital fabrication have further enabled the creation of tactile replicas and bas-reliefs that support hands-on engagement with artworks and artefacts [8, 9, 10], while narrative "tactile kits" combining structured audio layers

Joint Proceedings of the ACM UMAP Workshops 2026, UMAP 2026, June 8–11, 2026, Gothenburg, Sweden
*Corresponding author.
†These authors contributed equally.
yavni03@campus.haifa.ac.il (Y. Avni); julabushkara@gmail.com (J. N. Abushkara); ahedzoabi493@gmail.com (A. M. Zoabi); liorm9607@gmail.com (L. Maman); asaad@univ.haifa.ac.il (A. Danial-Saad); tsvikak@is.haifa.ac.il (T. Kuflik); oayalon@is.haifa.ac.il (O. Ayalon)

with 3D models increasingly allow visitors to explore material culture at their own pace [11, 12]. Audio Description (AD) remains a cornerstone of accessibility, translating visual content into verbal form [13]. However, traditional AD, typically delivered through static, pre-recorded scripts, is time-consuming to produce, limited in personalization, and unable to address visitors' spontaneous questions or interpretive needs [14, 15, 16]. Digital platforms have since introduced layered AD and interactive tactile-audio systems [17, 18, 19], but recent advances in AI, particularly generative AI (GenAI), present new opportunities for scalable, adaptive, and conversational forms of access. AI-driven systems may generate draft descriptions, provide contextual information on demand, or support follow-up dialogue through conversational agents [20, 21, 22]. Nonetheless, these technologies also raise concerns related to accuracy, curatorial integrity, and ethical use [23, 24].

Against this backdrop, the present study will investigate BLV visitors prefer to access supplementary information about tactile museum artifacts. For the purpose of the study, "EMusA" — a multimodal interaction prototype was developed. It combines tactile and two different audio-based interaction methods. The study will compare two fundamentally different interaction paradigms: a closed, curator-approved modality based on pre-recorded extended audio (hereafter referred to as narrative-based), and an open, conversational modality in which visitors ask free-form spoken questions processed by a generative AI system (hereafter referred to as a conversational AI system). Through this comparison, the study will examine differences in usability, interpretive autonomy, and overall user preference. The findings are expected to contribute to ongoing efforts to design inclusive, trustworthy, and multimodal museum technologies. This paper presents the system design and the planned evaluation protocol; the user study is currently ongoing and is reported here as forthcoming work to invite early feedback from the community.

2. EMusA Development

The system was developed through a co-design process involving the local museum, members of the BLV community, and graduate students from the Information Systems and Occupational Therapy departments at the University of Haifa. Guided by participatory and user-centered design principles, the project was situated within a multidisciplinary course that treated the museum as a "living laboratory" for accessibility innovation. Stakeholders contributed at each stage, from conceptual framing to prototyping and content development, ensuring that the resulting interface aligned with both curatorial standards and the experiential needs of BLV visitors. The prototype centers on the exhibition theme *Women's Status in Antiquity* and consists of a rotating wooden stand containing four 3D-printed tactile replicas drawn from the museum's archaeological collection (Figure 1). Each replica is housed in an individual compartment and can be lifted for hands-on exploration. A microswitch beneath each compartment detects object removal and triggers a pre-recorded AD, providing an immediate, stable baseline of accessible information. The kit supports two distinct modes of interaction, which the study will compare empirically. (1) The first mode: *Narrative Mode* — offers structured, curator-approved ADs. For two of the artifacts, users may receive additional layers of information by pressing a physical button on the front panel; each press reveals a new thematic layer, such as historical context, material culture associations, or interpretive insights. This linear, progressively layered design reflects established best practices in AD and emphasizes accuracy, coherence, and curatorial fidelity. (2) The second mode supports a more open-ended interaction paradigm using a *conversational AI system.* For the remaining two artifacts, pressing the same button activates a speech-to-text prompt that invites visitors to ask a free-form question. The query is processed by a generative AI model trained on museum-approved textual sources relating to women's roles and material culture in antiquity. The system then returns a spoken, contextually relevant response. Unlike the fixed layers of Narrative Mode, this interaction offers adaptive, visitor-driven exploration and allows BLV users to follow their own curiosities or seek clarification in real time.

Designing EMusA as an intelligent conversational museum guide required orchestrating multiple AI and system platforms within a unified interaction loop. The system combines OpenAI Whisper for

real-time speech recognition, GPT-4o for dialogue generation and reasoning, and multilingual text-to-speech services (ElevenLabs and Narakeet), while continuously integrating asynchronous sensor events from an ESP32 microcontroller. These components operate across heterogeneous layers — including embedded firmware, a FastAPI-based backend, and a React web client — each introducing distinct timing, state, and reliability constraints. Maintaining coherent, artifact-specific dialogue therefore depends on tightly coupling physical interaction events with conversational context management and language-aware response generation. This cross-platform coordination illustrates the complexity of deploying conversational AI as an intelligent user interface embedded in a physically situated, accessibility-critical museum environment.

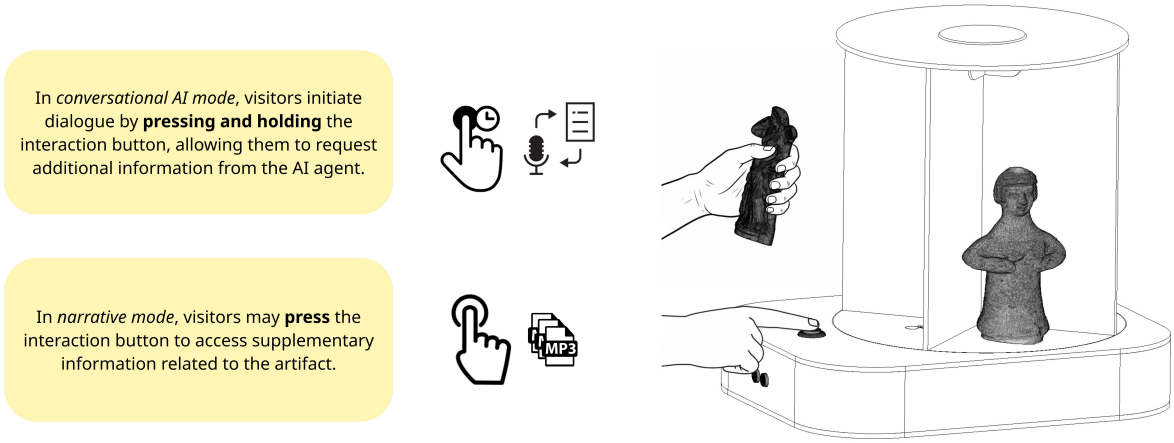

Figure 1: A dual-modality AD system. Upon selecting an artifact, the system assigns the corresponding compartment to one of two interaction modes: narrative playback or AI-based conversational interaction.

3. Study Design (Future Work)

The planned study will employ a mixed-methods, within-subjects design to examine how BLV museum visitors will prefer to access supplementary information about tactile exhibits using a multisensory interactive interface. The study will compare two modes of interaction: a closed Narrative Mode modality, and an open conversational AI modality. Approximately thirty BLV participants will be recruited to take part in this evaluation.

Before the interaction, participants will complete a demographic and background questionnaire that will record age, gender, vision profile, age of onset of blindness or visual impairment, and Braille literacy. The questionnaire will also assess prior familiarity with museum accessibility tools such as tactile models, raised-line images, embossed books, and braille displays, as well as their experience with screen readers, AD, voice assistants, and chatbot or AI-based systems. Participants will further rate their general confidence in using digital technologies, their comfort with voice-based interactions, and their expectations regarding clarity, accessibility, and autonomy during a museum visit. This pre-intervention phase will establish baseline characteristics and capture participants' initial expectations prior to encountering the system. Participants will be encouraged to explore the objects freely and to engage with the information pathways according to their preferences. To avoid response bias, they will receive a neutral explanation of the study's purpose without disclosing the specific research question being examined.

After completing the interaction, participants will fill out a detailed post-intervention questionnaire evaluating their experience with both interaction modalities. They will assess perceived usability, clarity of operation, ease of triggering information, and the extent to which the audio instructions were understandable. They will also report how effectively each modality supported their understanding

of the exhibits, whether they felt in control of the pace and depth of information, and whether the system addressed their personal needs as BLV visitors. Additional items will measure feelings of self-efficacy, including perceived independence during use, confidence in requesting additional information, and the likelihood of using similar technologies in future museum contexts. Participants will finally compare their actual experience with their initial expectations and provide qualitative feedback through open-ended questions regarding what they enjoyed, what they found challenging, which elements they would improve, and which type of interaction they preferred.

All interactions with EMusA will automatically logged, enabling qualitative analysis through the capture of time-stamped interaction events per artifact, user-initiated queries, system errors, and overall engagement patterns to support detailed examination of user behavior. The collected data will subsequently be analyzed using quantitative and qualitative approaches. Likert-scale responses will be examined to compare usability, comprehension, and self-efficacy across the two modalities, and engagement while open-ended responses will be thematically analyzed to identify recurring needs, preferences, and barriers. This study will allow us to characterize how BLV visitors engage with closed versus open information structures in a multisensory interactive museum environment and will inform the design of future accessible, conversational, and tangible museum interfaces.

4. Expected Outcomes

We expect that BLV participants will find the use of 3D-printed tactile artifacts engaging and experientially enriching, contributing positively to their overall museum experience. The combination of a physical interaction button and speech-based input is anticipated to be intuitive and comfortable, drawing on participants' prior familiarity with voice-controlled interfaces. We further hypothesize that participants with previous experience using conversational AI systems (e.g., smart voice assistants such as Alexa or Siri) will demonstrate more fluent and natural interaction patterns in the conversational AI mode compared to participants with limited prior exposure to such technologies. During AI-based interaction, we expect visitors to primarily ask artifact-related and contextually relevant questions, indicating appropriate use of the system rather than exploratory or off-task behavior. Finally, we anticipate variability in user preferences between interaction modes: while some participants may favor the clarity and structure of pre-recorded narrative descriptions, others are expected to prefer the flexibility of open-ended conversational interaction that allows deeper inquiry and personalized exploration. This contrast will provide insights into how different interpretive strategies support autonomy, engagement, and learning in accessible museum interfaces.

Acknowledgments

The authors would like to thank the staff members of the Hecht Museum for their support in the preparation of the content for the study.

Declaration on Generative AI

During the preparation of this work, the author(s) used Chat-GPT-4 in order to: Grammar and spelling check. After using this tool, the author(s) reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] G. Black, The Engaging Museum: Developing Museums for Visitor Involvement, Routledge, 2005.
[2] World Health Organization, Vision impairment and blindness, https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment, 2023. Accessed: 2025-04-17.

[3] R. Vaz, D. Freitas, A. Coelho, Perspectives of Visually Impaired Visitors on Museums: Towards an Integrative and Multisensory Framework to Enhance the Museum Experience, in: Proceedings of the 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, ACM, 2020, pp. 17–21. URL: https://dl.acm.org/doi/10.1145/3439231.3439272. doi:10.1145/3439231.3439272.

[4] A. Leahy, D. Ferri, Barriers to cultural participation by people with disabilities in Europe: A study across 28 countries, Disability & Society 39 (2024) 2465–2487. doi:10.1080/09687599.2023.2222898.

[5] S. Sylaiou, C. Fidas, Supporting people with visual impairments in cultural heritage: Survey and future research directions, International Journal of Human–Computer Interaction 40 (2024) 2195–2210. doi:10.1080/10447318.2022.2098930.

[6] Y. Avni, A. Danial-Saad, T. Kuflik, Design for inclusion: A systematic review of technologies and frameworks for enhancing the museum experience of blind and low-vision visitors, International Journal of Human–Computer Interaction (2026) 1–31. doi:10.1080/10447318.2026.2623227.

[7] E. L. Horton, R. Renganathan, B. N. Toth, A. J. Cohen, A. V. Bajcsy, A. Bateman, M. C. Jennings, A. Khattar, R. S. Kuo, F. A. Lee, M. K. Lim, L. W. Migasiuk, A. Zhang, O. K. Zhao, M. A. Oliveira, A review of principles in design and usability testing of tactile technology for individuals with visual impairments, Assistive Technology 29 (2017) 28–36. doi:10.1080/10400435.2016.1176083.

[8] M. Butler, L. Holloway, K. Marriott, A Closer Look: Multi-Sensory Accessible Art Translations (2019) 594–596. URL: https://dl.acm.org/doi/10.1145/3308561.3354617. doi:10.1145/3308561.3354617.

[9] K. Jabłońska, C. Lambert, N. Lister, T. Abram, C. Dalton, I. Frontani, Increasing Museum Accessibility for People with Visual Impairment Using Product Development and Additive Manufacturing, The International Journal of the Inclusive Museum 17 (2024) 169–191. URL: https://cgscholar.com/bookstore/works/increasing-museum-accessibility-for-people-with-visual-impairment-using-product-development-and-additive-manufacturing. doi:10.18848/1835-2014/CGP/v17i02/169-191.

[10] D. Reinhardt, L. Holloway, J. Thogersen, E. Guerry, C. A. C. Diaz, W. Havellas, P. Poronnik, The Museum of Touch: Tangible Models for Blind and Low Vision Audiences in Museums (2024) 135–155. URL: https://link.springer.com/10.1007/978-3-031-49511-3_8. doi:10.1007/978-3-031-49511-3_8.

[11] Y. Avni, A. Danial-Saad, J. Sheidin, T. Kuflik, Enhancing museum accessibility for blind and low vision visitors through interactive multimodal tangible interfaces, International Journal of Human-Computer Studies 198 (2025) 103469. URL: https://linkinghub.elsevier.com/retrieve/pii/S1071581925000266. doi:10.1016/j.ijhcs.2025.103469.

[12] P. Pistofidis, F. Arnaoutoglou, G. Ioannakis, N. Michailidou, M. Karta, C. Kiourt, G. Pavlidis, S. G. Mouroutsos, D. Tsiafaki, A. Koutsoudis, Design and evaluation of smart-exhibit systems that enrich cultural heritage experiences for the visually impaired, Journal of Cultural Heritage 60 (2023) 1–11. URL: https://linkinghub.elsevier.com/retrieve/pii/S1296207423000055. doi:10.1016/j.culher.2023.01.004.

[13] C. Bartolini, Diversity in museums: The inclusive value of museum audio description, DIVE-IN – An International Journal on Diversity and Inclusion (2022) 107–138 Paginazione. URL: https://dive-in.unibo.it/article/view/15764. doi:10.6092/ISSN.2785-3233/15764.

[14] S. A. Doore, D. Istrati, C. Xu, Y. Qiu, A. Sarrazin, N. A. Giudice, Images, Words, and Imagination: Accessible Descriptions to Support Blind and Low Vision Art Exploration and Engagement, Journal of Imaging 10 (2024) 26. URL: https://www.mdpi.com/2313-433X/10/1/26. doi:10.3390/jimaging10010026.

[15] J. Snyder, The Visual Made Verbal: A Comprehensive Training Manual and Guide to the History and Applications of Audio Description, Æ Academic Publishing, 2020. arXiv:WHorEAAAQBAJ.

[16] R. S. Hutchinson, A. F. Eardley, Museum audio description: The problem of textual fidelity, Perspectives 27 (2019) 42–57. doi:10.1080/0907676X.2018.1473451.

[17] F. D'Agnano, C. Balletti, F. Guerra, P. Vernier, Tooteko: A case study of augmented reality for an

accessible cultural heritage. digitization, 3d printing and sensors for an audio-tactile experience., The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5/W4 (2015) 207–213. URL: https://isprs-archives.copernicus.org/articles/XL-5-W4/207/2015/. doi:10.5194/isprsarchives-XL-5-W4-207-2015.

[18] T. Graven, I. Emsley, N. Bird, S. Griffiths, Improved access to museum collections without vision: How museum visitors with very low or no vision perceive and process tactile–auditory pictures, British Journal of Visual Impairment 38 (2020) 79–103. URL: https://journals.sagepub.com/doi/10.1177/0264619619874833. doi:10.1177/0264619619874833.

[19] X. Wang, D. Crookes, S.-A. Harding, D. Johnston, Stories, journeys and smart maps: An approach to universal access, Universal Access in the Information Society 21 (2022) 419–435. URL: https://link.springer.com/10.1007/s10209-021-00832-0. doi:10.1007/s10209-021-00832-0.

[20] R. Dror, D. Hutchinson, M. Jones, V. Van Hyning, T. Kuflik, The Curator's Helper (2024) 496–504. URL: https://dl.acm.org/doi/10.1145/3631700.3664905. doi:10.1145/3631700.3664905.

[21] D. Fernández, R. B. Hagen, Guidelines for Creating Accessible Archaeological Museum Exhibitions for Visitors with Visual Impairments, 2024.

[22] S. Reinders, S. Ananthanarayan, M. Butler, K. Marriott, Designing Conversational Multimodal 3D Printed Models with People who are Blind (2023) 2172–2188. URL: https://dl.acm.org/doi/10.1145/3563657.3595989. doi:10.1145/3563657.3595989.

[23] J. Yan, Y. Wang, W. Liu, The Application and Challenges of Artificial Intelligence in Contemporary Art Curation, Journal of Education 60 (2025).

[24] C. Kist, Y. Moshfeghi, To bot or not? Conversational agents and authentic heritage experiences, in: Proceedings of the Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems, 2026, pp. 1–5. doi:10.1145/3772363.3798924.