MediMomo: Designing a Physical AI Platform for Content Accessibility in Healthcare

Philipp Thalhammer
LMU Munich
Munich, Germany
philipp.thalhammer@ifi.lmu.de

Alexander Wiethoff
LMU Munich
Munich, Germany
alexander.wiethoff@ifi.lmu.de

Figure 1: MediMomio from various angles ©Imago Design

Abstract

Interacting with artificial intelligence (AI) can be cumbersome for different user groups due to various limiting factors such as age, disability, or context. We propose a multimodal, barrier-free, and playful physical AI interface that can be used in the medical context to facilitate user engagement and playful interaction forms. Our prototype *MediMomo* is equipped with several input and output modalities that serve as an interaction framework and can be adapted to suit different use cases. In this paper we outline an exemplary use case to employ our prototype in a hospital environment to create a playful and engaging way for children to navigate the building and retrieve high-level medical information. We envision MediMomo to become a universal platform that can be used in different environments for different needs, enabling barrier-free content interactions.

CHI '25, Yokohama, Japan

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXX

CCS Concepts

• **Human-centered computing → Accessibility systems and tools**.

Keywords

Accessibility, Artificial Intelligence, Physical AI, Playful AI, Healthcare

ACM Reference Format:
Philipp Thalhammer and Alexander Wiethoff. 2025. MediMomo: Designing a Physical AI Platform for Content Accessibility in Healthcare. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 4 pages. https://doi.org/XXX

1 Introduction

Interacting with AI poses several challenges for people with difficulties interacting with technology, such as disabilities, fear, lack of personal mobile devices, or age as an excluding factor. Especially in the medical domain, this limitation restricts people such as children or the elderly from benefiting from the capabilities that Large Language Models (LLMs) such as ChatGPT[1] or DeepSeek[2] can provide. Our research focus, therefore, targets the design of barrier-free and accessible AI in the medical domain to overcome shortcomings and

[1]https://chat.openai.com
[2]https://www.deepseek.com

provide valuable in-context information such as, for example, medical knowledge and way-finding. To investigate this domain further, we have teamed up with a large German child clinic to provide first-hand information on the suitability of our approach. Our AI interface has the appearance of a monkey to promote higher user engagement, especially with children (see *Figure 1*). We introduce *MediMomo*: a physical AI interface that helps to build user engagement. Although we currently plan to verify this concept in the form of a system that helps users navigate hospitals, we envision the project to become a versatile platform that can be adapted for different accessibility needs in various environments. In this work, we analyze the relevant literature, describe the system architecture, and outline a study in a medical context. Our input in this CHI workshop is intended to share our research approach, discuss/adapt the potential study design, and invite others to collaborate with us in this emerging domain.

2 Related Work

2.1 AI in Healthcare

While AI in healthcare is often used for detection and classification purposes, such as the early detection of Alzheimer's disease [6] or cancer [10], it also has the potential to help humans in other domains that previously required a human agent. More specifically, the capability of generative AI to engage in full conversations could help mitigate the shortage of skilled workers in many developed countries: In their study about the use of AI chatbots to answer patients' medical questions, Ayers et al. [2] have shown that healthcare professionals preferred chatbot answers over human ones in 78.6% of cases. This does not mean that AI can replace human professionals, especially not in a high-stakes domain such as the medical field, however, it could be utilized to draft answers, which then would be reviewed by a human agent, transforming the workflow efficiently. Javaid et al. [5] proposed the usage of chatbots like ChatGPT to transcribe, translate, or summarize medical records and texts, to improve communication, and to enable medical staff to take better care of patients. Even though AI can potentially improve efficiency, it also comes with a set of challenges as identified by Sharma et al. [12]:

1. AI requires advanced technology (high-speed internet, computers, etc.)
2. Professionals need to maintain, develop, and operate the system
3. Language and cultural differences (Interfaces)
4. Data security and privacy
5. Financial resources

Especially, the privacy concern requires further investigation, as AI in healthcare would be given access to highly sensitive data, making locally hosted AI systems almost unavoidable. However, the potential benefits of AI in healthcare create a strong imperative for further research.

2.2 Physical Interfaces for AI Interaction

Bartneck et al. [4] introduce four key concepts for human-robot interaction: likeability, perceived intelligence, perceived safety, anthropomorphism, and animacy. Three of which (perceived animacy, perceived intelligence, and perceived anthropomorphism) Balakrishnan and Dwivedi [3] assigned to AI. While anthropomorphism and animacy may initially appear to reference the same attribute, "anthropomorphism refers to the attribution of a human form, human characteristics, or human behavior to nonhuman things such as robots, computers, and animals" [4] and animacy refers to how alive an object appears [4]. Research by Qiu and Benbasat [11] found that humanoid embodiment and human voice-based communication positively affect users' perception of social presence, heightening their trust in a system. An empirical study has shown "humanlike robots were judged as more competent than machinelike robots, and masculine robots produced higher levels of discomfort than feminine robots" [14], yet regarding to the uncanny valley theory from 1970 by Mori et al. [9], there exists a certain range in human-likeness that results in revulsion rather than empathy.

The 'T.A.I' [7] introduces a haptic component in addition to a text-based conversational AI agent capable of conveying emotions, which leads to a more ""creature"-like impression of the AI agent for the users" [7]. Similarly, 'MeBot' [1] uses limbs, a tiltable face, and a portable base to convey emotions through a video call heightening participants' psychological involvement, engagement, enjoyment, and willingness to cooperate.

2.3 Playful Technology in Healthcare

In their work about breastfeeding education, Tang et al. [15] found that even though their prototype succeeded in recreating common challenges in breastfeeding, users expected a more gamified approach, including some sort of quantification of their actions. However, they were able to demonstrate the viability of playful approaches in healthcare. Especially children respond well to gamified approaches in healthcare settings, as shown by Yawn et al. [16] with their asthma education program, reporting that "children were enthusiastic about the opportunity to play the game" [16]. The importance of playful activities for children in hospitals is also highlighted through a systematic review by Souza et al. [13], concluding that they "promote a less traumatic recovery, besides allowing for greater tranquility in the hospital environment" [13].

Malodia et al. [8] have shown that a higher playfulness in voice assistants leads to a higher likelihood of usage. We want to extend this concept into physical interfaces in the healthcare sector.

3 System

In this section, we provide an outline of the general technical interaction and implementation concept. We aim to allow multiple forms of input and output modalities to create various use applications where our prototype will serve as an AI interface gateway for our studies. MediMomo consists of a 3D-printed body and a detachable front plate for maintenance (see *Figure 2*).

All listed components are preliminary and may change in the final implementation. The system will be powered by a Raspberry Pi 5 (4GB RAM) and have the following input and output modalities:

(1) Touch display[3]
(2) Thermal printer[4]
(3) Speaker[5]
(4) Microphone[6]
(5) Camera[7]

Figure 2: Visualization of the 3D printed components of the physical AI interface

We plan to implement the system using Python and create different libraries to support the input and output modalities of the interface. Our system will support voice control and will be equipped with text-to-speech capabilities to communicate with potential users. A front-facing camera will be used to scan QR codes and detect the approximate age of the user and how many people are in front of our device. The front display might be used as both an input and output medium. We will also install a small-footprint thermal printer to produce physical paper artifacts that users can take with them. This physical token serves two purposes: (a) to remember certain information without requiring an additional mobile device (i.e., navigation instructions) and (b) as a reminder of the interaction experience and playful engagement artifact (e.g., collectible). The system's architecture is visualized in *Figure 3*.

4 Exemplary Use Case

To ensure our prototype aligns with the needs of our user group, we work with a hospital located in Germany to identify problems in the medical environment. Hospitals are often stressful environments, especially for families with children. One factor that can cause stress is finding the right place, as hospital buildings tend to be large structures that are difficult to navigate. Additionally, the medical sector constantly struggles with staff shortages, making it even harder for hospital visitors to access information. To help combat this issue, we want to employ physical AI interfaces that go beyond a simple information screen and offer a playful interaction to the user. To demonstrate our system, we plan to utilize it in a medical context to help users navigate a hospital environment. Users can interact with MediMomo via voice input and ask questions to receive directions to specific places in the hospital in a playful and engaging format (e.g. treasure-hunt). MediMomo responds using text-to-speech and also prints the directions using the included thermal printer (see *Figure 1*) to help users remember the instructions without using additional technical hardware (e.g, smartphone, tablet, watch, etc.). Further, MediMomo can answer simple medical questions in a way that is suitable for children and print a QR code. The code on the paper printout can then be scanned at a second MediMomo in the destination location (i.e., the responsible doctor's office) to act as an engaging scavenger hunt for children.

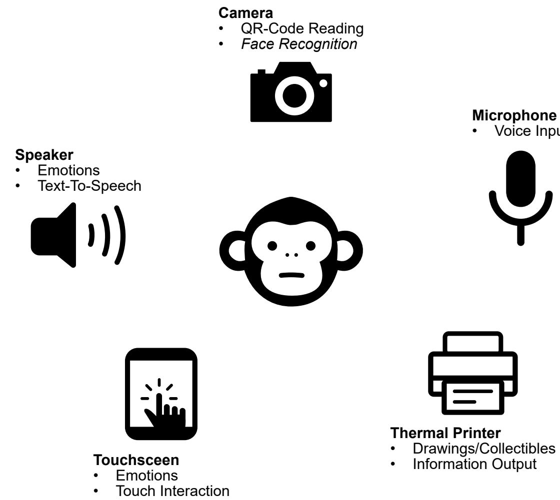

Figure 3: Schematic representation of the input and output modalities of the physical AI interface

5 Discussion and Conclusion

Considering the targeted implementation, we aim to investigate two research questions:

(1) How can physical artifacts foster the utilization of pre-trained LLMs?
(2) How can physical AI interfaces increase information accessibility in the medical domain?

We are aware that the *novelty effect* and the appealing physical design of our device might receive initial increased positive attention by users, which does not necessarily have a correlation to increased usability. However, we aim to reduce this biasing effect

[3]https://www.waveshare.com/product/displays/3.5inch-rpi-lcd-b.htm
[4]https://www.maximppl.com/PNP-500-micro-thermal-panel-printer.html
[5]https://www.berrybase.de/externer-usb-mini-lautsprecher-schwarz
[6]https://www.berrybase.de/usb-mini-mikrofon
[7]https://www.berrybase.de/raspberry-pi-camera-module-8mp-v2

through longer time spans of observation and exposure to different user groups (e.g., new vs. returning patients). In summary, we reflected on the design of MediMomo as a multipurpose physical AI input/output platform that can be used in various environments to foster user engagement and lower the participation barrier with LLMs in a medical context with a dedicated user group (i.e., children). We've outlined the concept of a physical AI interface with multi-modal capabilities to communicate with users and described an exemplary use case that will serve as a test-bed for our prototype in a real-world environment. By providing our setup, approach, and tools, we aim to spark discussions on the topic, raise awareness, and open our door to potential collaborators.

References

[1] Sigurdur O. Adalgeirsson and Cynthia Breazeal. 2010. MeBot: A Robotic Platform for Socially Embodied Presence. In *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction (HRI '10)*. IEEE Press, Osaka, Japan, 15–22.

[2] John W. Ayers, Adam Poliak, Mark Dredze, Eric C. Leas, Zechariah Zhu, Jessica B. Kelley, Dennis J. Faix, Aaron M. Goodman, Christopher A. Longhurst, Michael Hogarth, and Davey M. Smith. 2023. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Internal Medicine* 183, 6 (June 2023), 589–596. doi:10.1001/jamainternmed.2023.1838

[3] Janarthanan Balakrishnan and Yogesh K. Dwivedi. 2024. Conversational Commerce: Entering the next Stage of AI-powered Digital Assistants. *Annals of Operations Research* 333, 2 (Feb. 2024), 653–687. doi:10.1007/s10479-021-04049-5

[4] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics* 1, 1 (Jan. 2009), 71–81. doi:10.1007/s12369-008-0001-3

[5] Mohd Javaid, Abid Haleem, and Ravi Pratap Singh. 2023. ChatGPT for Healthcare Services: An Emerging Stage for an Innovative Perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations* 3, 1 (Feb. 2023), 100105. doi:10.1016/j.tbench.2023.100105

[6] Somayeh Khosroazad, Ali Abedi, and Marie J. Hayes. 2023. Sleep Signal Analysis for Early Detection of Alzheimer's Disease and Related Dementia (ADRD). *IEEE Journal of Biomedical and Health Informatics* 27, 5 (May 2023), 2264–2275. doi:10.1109/JBHI.2023.3235391

[7] Xin Liu and Kati London. 2016. T.A.I: A Tangible AI Interface to Enhance Human-Artificial Intelligence (AI) Communication Beyond the Screen. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. Association for Computing Machinery, New York, NY, USA, 281–285. doi:10.1145/2901790.2901896

[8] Suresh Malodia, Nazrul Islam, Puneet Kaur, and Amandeep Dhir. 2024. Why Do People Use Artificial Intelligence (AI)-Enabled Voice Assistants? *IEEE Transactions on Engineering Management* 71 (2024), 491–505. doi:10.1109/TEM.2021.3117884

[9] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley. *IEEE Robotics & Automation Magazine* 19, 2 (June 2012), 98–100. doi:10.1109/MRA.2012.2192811

[10] Annie Y. Ng, Cary J. G. Oberije, Éva Ambrózay, Endre Szabó, Orsolya Serfőző, Edit Karpati, Georgia Fox, Ben Glocker, Elizabeth A. Morris, Gábor Forrai, and Peter D. Kecskemethy. 2023. Prospective Implementation of AI-assisted Screen Reading to Improve Early Detection of Breast Cancer. *Nature Medicine* 29, 12 (Dec. 2023), 3044–3049. doi:10.1038/s41591-023-02625-9

[11] Lingyun Qiu and Izak Benbasat. 2009. Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems* 25, 4 (April 2009), 145–182. doi:10.2753/MIS0742-1222250405

[12] Sachin Sharma, Raj Rawal, and Dharmesh Shah. 2023. Addressing the Challenges of AI-based Telemedicine: Best Practices and Lessons Learned. *Journal of Education and Health Promotion* (Sept. 2023), 338. doi:10.4103/jehp.jehp_402_23

[13] Mariana Silva Souza, Daniel Lopes Araújo, Bruno Abilio da Silva Machado, Erik Bernardes Moreira Alves, Igor Mendes Mendonça, Cícero Santos Souza, José Eufrazino Júnior, José Marcos Fernandes Mascarenhas, Kayco Damasceno Pereira, Getulivan Alcantara de Melo, Maria Clara Souza Oliveira, Carla Michele Silva Ferreira, Danielle dos Santos Araujo, Leandro Luiz da Silva Loures, Miguel Felix de Souza Neto, Graziele Ferreira Nunes, Josué Brito Gondim, Millena Raimunda Martins de Almeida Carvalho, Marcos Van Basten do Nascimento Páiva, and Luana Nayra Coutinho de Meneses. 2021. Playful Strategies in the Pediatric Hospitalization Scenario. *Research, Society and Development* 10, 5 (April 2021), e6210514652–e6210514652. doi:10.33448/rsd-v10i5.14652

[14] Steven J. Stroessner and Jonathan Benitez. 2019. The Social Perception of Humanoid and Non-Humanoid Robots: Effects of Gendered and Machinelike Features. *International Journal of Social Robotics* 11, 2 (April 2019), 305–315. doi:10.1007/s12369-018-0502-7

[15] Kymeng Tang, Kathrin Gerling, and Luc Geurts. 2022. Challenges and Opportunities for Playful Technology in Health Prevention: Using Virtual Reality to Supplement Breastfeeding Education. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 406–409. doi:10.1109/VRW55335.2022.00088

[16] Barbara P. Yawn, Pamela J. Algatt-Bergstrom, Roy A. Yawn, Peter Wollan, Mark Greco, Marie Gleason, and Leona Markson. 2000. An In-School CD-ROM Asthma Education Program. *Journal of School Health* 70, 4 (2000), 153–159. doi:10.1111/j.1746-1561.2000.tb06462.x