Supporting Medical Doctors with AI-based Smart Wearable Assistance: Case Study of Diagnosis and Therapy Evaluation for Ward Visits

Agnese Augello[1], Giuseppe Caggianese[1], Silvia Franchini[1], Ignazio Infantino[1] and Luca Sabatucci[1,*]

[1]*Institute for High Performance Computing and Networking (ICAR) - National Research Council of Italy (CNR), Italy*

Abstract

The digitalization of healthcare requires innovative interfaces that support medical professionals without disrupting the delicate doctor-patient relationship. This paper presents a design-centric architectural study of a smart wearable assistance system developed within the SMASH (Smart Management and Assistance System for Healthcare) project. The architecture combines lightweight smart glasses with a tablet interface: while smart glasses provide immediate, hands-free access to essential clinical data via a see-through display, the tablet facilitates deep analysis of complex prognostic variations. The architecture focuses on interaction design and workflow integration, enabling clinicians to access situated information on the move while retaining the ability for deeper analysis when time permits. It also leverages AI-based services, including Large Language Models (LLM), speech recognition, and proximity sensors, to drive an adaptive eXtended Reality (XR) interface for hand-free clinical assistance during ward activities. Derived requirements and refined scenarios resulting from focus groups with healthcare professionals illustrate how the proposed interaction layer adapts to contextual conditions, manages uncertainty and noise, and exploits redundancy in sensing and interaction mechanisms, prioritizing how, when, and where information is presented rather than introducing new visualization paradigms.

Keywords

Healthcare, context-aware, XR, adaptive interfaces

1. Introduction

Hospitals represent highly dynamic and information-intensive environments, where clinicians must continuously coordinate patient data, clinical protocols, and time-critical decisions. Daily clinical activities are inherently mobile and situated: healthcare professionals move across corridors and rooms, interact with patients and colleagues, and must respond promptly to emerging conditions while maintaining high standards of safety, hygiene, and privacy. Among the activities carried out by doctors on a daily basis, daily ward rounds involve consulting various types of information while also maintaining empathetic communication with patients. The introduction of digital support in this type of task has the potential to enable professionals to apply protocols for planning, implementing, and executing healthcare strategies that are context-aware and profile-tailored, specializing in comprehensive interaction with the target user group. In this context, human-machine interfaces are of great importance for the acquisition and processing of real-time clinical information, as well as for facilitating simplified interactions between healthcare professionals and patients.

Augmented interaction technologies offer promising opportunities to bridge this gap by embedding digital information directly into the clinical environment and enabling hands-free or low-effort forms of interaction. In healthcare settings, smart glasses and augmented reality (AR) interfaces have been increasingly explored for a wide range of applications, including surgical support, remote assistance,

Joint Proceedings of the ACM Intelligent User Interfaces (IUI) Workshops 2026, March 23-26, 2026, Paphos, Cyprus

*Corresponding author.
agnese.augello@icar.cnr.it (A. Augello); giuseppe.caggianese@icar.cnr.it (G. Caggianese); silvia.franchini@icar.cnr.it (S. Franchini); ignazio.infantino@icar.cnr.it (I. Infantino); luca.sabatucci@icar.cnr.it (L. Sabatucci)
0000-0001-6463-9151 (A. Augello); 0000-0001-6607-6591 (G. Caggianese); 0000-0002-4845-5341 (S. Franchini); 0000-0002-0980-7391 (I. Infantino); 0000-0003-2852-9355 (L. Sabatucci)

training, and navigation. A growing body of work reports benefits, such as reduced workflow interruptions, improved access to contextual information, and better alignment with the mobile and situated nature of clinical activities [1, 2, 3].

Despite these results, the adoption of augmented interaction technologies in routine ward activities presents constraints related to usability, comfort, device intrusiveness, and integration with established clinical practice [4].Wearing head-mounted displays and simultaneously managing two different interactions: human-human and human-machine can lead to misunderstandings with co-workers [5, 6]. Moreover, the introduction of digital tools in medical visits often creates a physical and cognitive barrier; physicians forced to look at computer screens or handheld devices may lose eye contact and situational awareness [7] or, more generally, the use of such tools can become an impediment to natural face-to-face interactions [8].

This paper focuses on a design-centric architectural study of the system's interaction layer, which leverages eXtended Reality (XR) and Artificial Intelligence (AI) –including Large Language Models (LLM), speech recognition, and proximity sensing – to provide "invisible" support to the medical doctors. The necessity for such a system stems from the need to reduce doctors' cognitive load during rounds and to enable real-time visualization of prognostic data without interrupting the clinical interview. The main challenge addressed is the simultaneous interaction dilemma: the doctor must interact with the AI system to retrieve and analyze data while maintaining continuous conversation with the patient. This dual engagement creates potential interference between digital interaction and human communication, requiring careful design of when, how, and where information is presented.

We propose adopting a dual-device strategy that combines lightweight smart glasses with a tablet-based interface. Smart glasses are used to support in-situ, hands-free micro-interactions, such as notifications, alerts, and short context-aware summaries, while the tablet provides a complementary space for more detailed exploration, analysis, and reporting when time and attention permit. This separation allows the system to accommodate the heterogeneous temporal and cognitive demands of clinical activities, ensuring immediate access to relevant information while preserving the possibility of deeper analytical exploration. The core contribution of this work is the interaction design strategy that manages the interference between the digital system and the human conversation, ensuring that the patient remains the focus of the interaction. The proposed architecture is grounded in constraints and requirements emerging from focus groups and requirements analysis activities within the research and development of the SMASH (Smart Management and Assistance System for Healthcare) initiative, which investigates digital proximity infrastructures, wearable technologies, and data-driven approaches to support context-aware healthcare services.

The remainder of the paper is organized as follows. Section 2 reviews related work on augmented solutions in healthcare environments. Section 3 presents the proposed augmented interaction solution, detailing its architecture and design rationale. Section 4 describes the use case and the reference scenarios analyzed to define the interaction modalities. Finally, Section 5 discusses design implications and outlines directions for future work in connection with the SMASH research programme.

2. Related Works

The integration of XR and AI in different domains has been widely explored in the last year; however, challenges regarding usability and social acceptance persist. From a usability perspective, old dilemmas have resurfaced, comparing user active interaction with a more passive approach that delegates action to agents [9, 10]. The Human-centered Artificial Intelligence (HCAI) is a growing viewpoint in which researchers are taking up the call to amplify, augment, empower, and enhance human performance, finding creative ways to support user self-efficacy, enable creativity, clarify responsibility, and facilitate social connectedness. Regarding the social acceptance of these tools and their application in sensitive domains, it is essential to consider the privacy of direct users as well as that of surrounding individuals, and to take into account the specific nature of some of the data that the system is allowed to process. In the healthcare domain, the social acceptance of AI-based tools is generally positive but tempered by concerns over trust, transparency, and the need for human oversight. Recent surveys have revealed varying levels of enthusiasm, with preferences for supportive rather than autonomous AI roles indicating that factors such as demographics, technical familiarity, and specific applications significantly influence attitudes [11]. Large-scale surveys indicate majority favorable views, such as 57.6% of over 13,000 respondents expressing positive attitudes toward AI in healthcare. However, trust is lower for specific aspects like treatment response accuracy, with only 41.8% confidence. Noninvasive uses, such as monitoring, rank highest in acceptance, while empathy-driven areas, such as psychotherapy, face skepticism [12]. To explore the acceptability and intention to use AI health assistants, another study found that technology attitudes, frequent device use, and AI knowledge drive higher acceptance, while women and less tech-savvy individuals show more caution. Social influence and perceived utility boost intentions to use AI health assistants, confirming that the perceived ”care” and utility of the assistant are paramount for its integration into the clinical routine [13].

Recent literature has begun to address these usability tensions through empirical studies of multimodal interactions and agent design in XR environments. Yoon et al. investigated preferential usability of intelligent agent-driven multimodal interfaces in XR, comparing independent task agents (ITAs)—embodied as floating orbs—and dependent task agents (DTAs) on panels using hand gestures, gaze, and voice on Meta Quest Pro headsets. Participants favored ITAs for higher hedonic quality (mean 1.81) and playfulness, highlighting embodied forms as superior for immersion and natural control, though both modalities excelled in pragmatic efficiency [14]. Zeng reviewed advances in AI-powered spatial intelligence for XR, emphasizing multimodal AI integration in hardware like vision transformers and software for adaptive, context-aware interactions. The work advocates HCAI-aligned designs that enhance user agency through hybrid inputs, leveraging digital twins and IoT to provide proactive assistance without diminishing self-efficacy [15]. In a comprehensive survey, researchers mapped large language model (LLM) integration in XR, categorizing applications from natural collaboration to generative content creation and underscoring usability gains from conversational agents over rigid menus. Similarly, an analysis of generative AI in XR examined interaction paradigms, finding hybrid active-passive systems most effective for creativity and responsibility clarification in domains like training and design [16].

From a perspective more closely related to interaction modalities, Wang et al. discuss the flexibility of gaze, gesture, and speech in AR systems, highlighting that while hands-free interaction is ideal for sterile environments, it can introduce social friction [17]. While the privacy and robustness issues due to the use of Voice User Interfaces (VUI) in shared spaces are considered in a recent review of Bai et al., which suggests that integrating gesture and speech can mitigate ambiguity, but in a clinical setting, minimizing conspicuous gestures is preferred to avoid alarming patients [18].

With the AI agents, these tools are no longer considered tools but more “companions” or persistent assistants. Recent studies on AI companionship highlight a complex landscape of benefits and harms, while some users report significant therapeutic benefits, reduced loneliness, and mental health improvements through always-available support [19], there is an emerging ”dark side” to these relationships. Research by Zhang et al. on large-scale user bases identifies a new category of relational harm, where algorithmic behaviors can diverge into harassment, verbal abuse, and privacy violations. Such harms not only affect the individual’s psychological well-being through emotional dependency and reality dissociation but may also erode their broader relational capacity and interpersonal connections. These findings underscore the critical need for ethical standards in AI design to prevent ”relational transgression” between the professional, the AI assistant, and the patient [20].

3. Smart Wearable Assistance System as support for ward visits

The proposed system is an edge computing solution designed to function robustly within hospital wards [21]. The context of use is daily visits to the ward, during which medical personnel collect information on the patient’s clinical status and analyze prognostic variations. The doctors are assisted by a context-aware multimodal system comprising an AR component, advanced multimedia visualization, artificial intelligence, and IoT devices, which becomes essential in supporting their daily screening. The challenge for these systems is therefore to create a seamless user experience that continuously supports the users without compromising ordinary activities or radically changing them. In designing a support service for doctors, we observed that no single device or interaction modality can adequately support the full range of activities performed by clinicians in hospital wards. Clinical work alternates between moments that require immediate, lightweight access to information and those that allow for more deliberate analysis and reflection. Designing a single interface to cover both extremes often leads to compromises that negatively affect usability and adoption. This led us to adopt a dual-device interaction strategy that combines lightweight smart glasses with a tablet-based interface. This approach enabled us to distribute interaction responsibilities across devices based on their strengths, while maintaining coherence at the system level.

3.1. Hardware Selection and Rationale

Based on a comparative analysis of current market devices, the Vuzix Shield [22] was selected as the primary wearable. Its lightweight, non-invasive design allows the doctor to maintain direct eye contact with the patient, reducing the impact on the patterns of non-verbal patient-professional communications [23]. The device utilizes microLED displays with high resolution but limited chromatic rendering (specifically, green monochrome). This technical constraint dictated the system design: the smart glasses display synthetic, high-contrast information (e.g., patient ID, text-based alerts, simple trend indicators), while a Tablet is used as a complementary device for visualizing complex chromatic data, such as ECG waveforms or detailed multi-variable graphs. Moreover, the selected wearable device incorporates a bone-conduction audio system that enables the physician to receive discreet auditory cues without occluding the ears, thereby preserving environmental awareness and minimizing isolation from the surrounding clinical context.

3.2. System Architecture

The augmented interaction infrastructure is designed to collect, manage, and visualize clinical and contextual data, and to present them in a contextualized form on mobile devices and smart glasses. The XR applications for wearable and mobile devices are focusing on the interaction layer that mediates between data sources, analytical services, and end-user devices.

The information conveyed through the XR application comes from the SMASH project platform that is in charge of providing contextual information related to the patient, such as medical records, and proximity information coming from patients' BLE (Bluetooth Low Energy) sensors, and also the access to other services essential to guarantee the user-device communication such as Automatic Speech Recognition (ASR), Large Language Model (LLM) and Text-to-Speech (TTS).

From an architectural perspective (see Figure 1), the entire system is organized around three main macro-components that communicate through secure REST-based APIs: (i) Knowledge/Context Logic, (ii) XR Applications, and (iii) Local server-based services. The Knowledge/Context Logic is responsible for data integration, contextual analysis, and orchestration of information to be delivered to edge devices. This component collects and processes heterogeneous data streams originating from multiple sources, including structured clinical data (e.g., JSON-based records, standards such as FHIR (Fast Healthcare Interoperability Resources) or DICOM (Digital Imaging and Communications in Medicine)), data from wearable sensors (e.g., BLE wristbands), and environmental inputs related to location and proximity. The primary role of this component is not only data aggregation, but also context interpretation: the system continuously evaluates contextual variables such as patient identification (e.g., via BLE proximity), clinical phase, user role, and location within the hospital. Based on this contextual analysis, the Knowledge/Context Logic dynamically selects, filters, and prioritises the information to be delivered to user devices. This mechanism is explicitly designed to reduce cognitive load by ensuring that only relevant and situation-appropriate information is presented on each device, adapting both content and level of detail to the current clinical scenario.

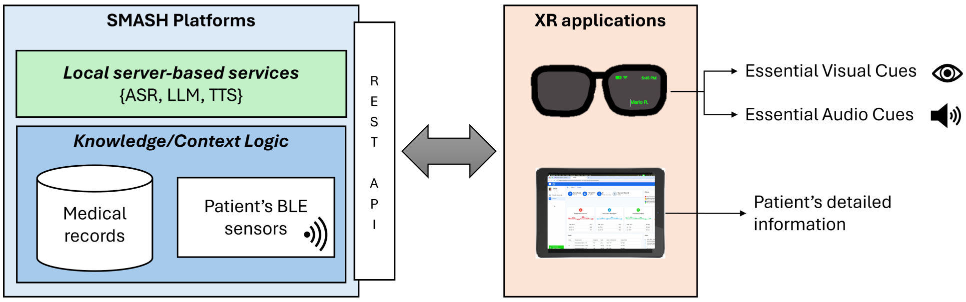

Figure 1: System Architecture

The XR Applications component is deployed directly on edge devices, including AR smart glasses and tablets, and is responsible for presenting information in a simple, concise, and context-aware manner. On AR smart glasses, the application is optimized for the limited field of view and supports rapid perception of essential information during mobility and bedside activities. Interaction with the system is multimodal, combining voice-based and touch-based input. In particular, the application exploits the directional microphones embedded in Vuzix Shield smart glasses to isolate the clinician's voice, preventing the capture of voices from patients or other healthcare staff. On AR smart glasses, the system displays essential information, including patient identification, current clinical status, high-priority alerts highlighted with graphical cues, and simple trend indicators for vital parameters. Visual elements are deliberately synthetic and high-contrast, relying on short textual labels, icons, symbols, and simple graphical indicators. Notifications may be accompanied by synchronized audio messages to reinforce critical information without increasing visual load. The tablet serves as a complementary edge device, providing access to more complex content, including detailed charts, historical reports, dynamic HTML or PDF documents, and, when necessary, pre-existing 3D models. This separation allows the system to align information richness with the temporal and cognitive demands of the clinical activity. The XR Applications layer is implemented using Android SDK and Vuzix SDK, ensuring compatibility with commercial smart glasses and established mobile development workflows. Advanced functionalities, such as speech processing, command interpretation, and the use of large language models, are delegated to services running on local or dedicated servers. These services are encapsulated in Docker containers and accessible exclusively via secure REST APIs. Whenever feasible, selected functionalities may also be executed directly on the tablet, provided that the computational requirements are compatible with the device's hardware resources. Voice data is processed in an ephemeral manner: audio streams are never persisted, stored, or transmitted to external services; instead, they are handled transiently and discarded immediately after processing. This design choice ensures a high level of privacy protection and compliance with healthcare data protection requirements.

4. Case Study: Designing the Doctor-System Interaction

The case study focuses on the design of the doctor's user interface. The main challenge addressed was the simultaneous interaction dilemma: the doctor must interact with the AI system to retrieve data while conversing with the patient. The general system requirements considered in the design are: (i) The system must assist the doctor during ward rounds without hindering doctor–patient communication, favouring hands-free and low cognitive load interactions, (ii) The system must distribute information between smart glasses and a tablet according to context, showing concise alerts and summaries on the glasses and detailed data on the tablet. (iii) Consider how the surrounding context may affect the interaction. For instance, the patient, unaware of the digital overlay, may interrupt the doctor or inadvertently trigger the system.

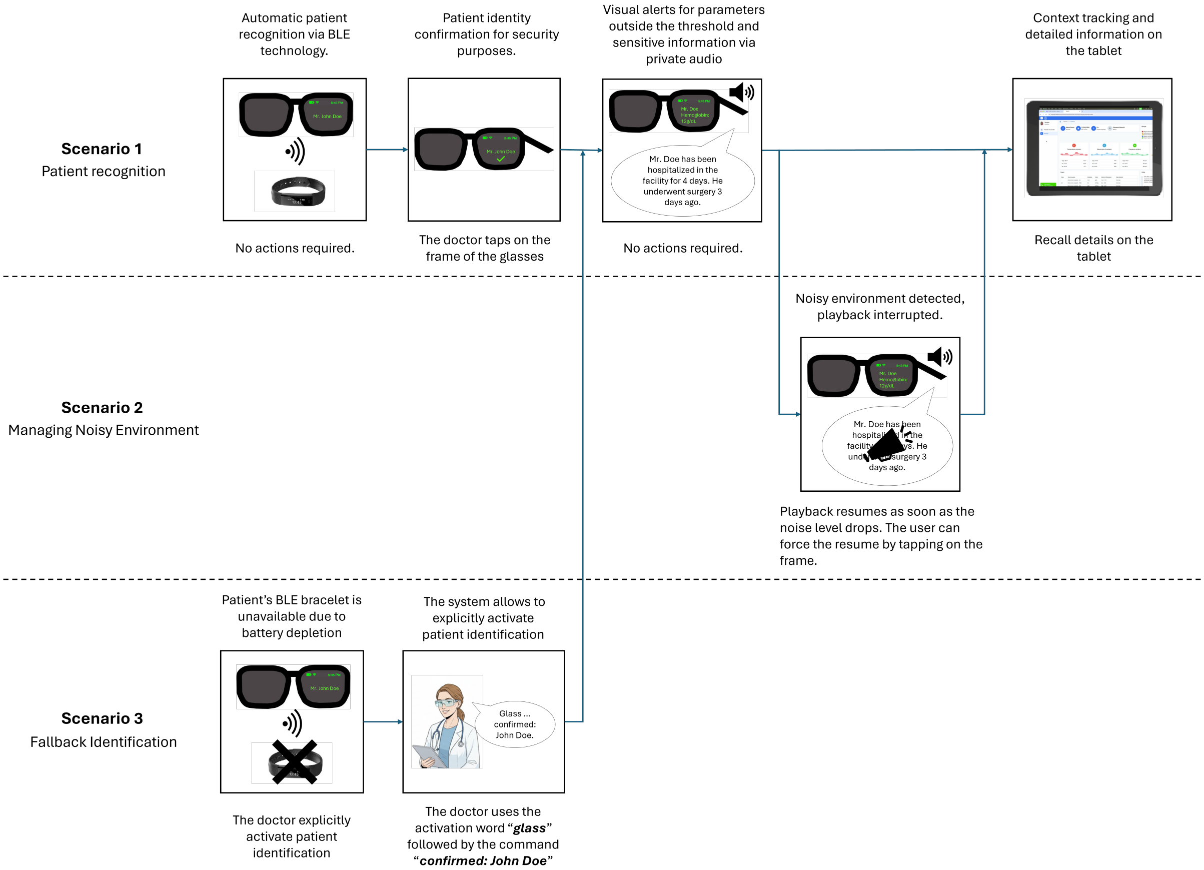

Figure 2: Operational Scenarios used to design an interaction layer adaptive with respect to the contextual conditions. The figure highlights the possible transitions among the scenarios and the clinician's interactions.

4.1. Operational Scenarios and Requirement Analysis

To evaluate the applicability of the proposed augmented interaction architecture in routine ward activities, we decomposed the ward round case study into three distinct operational scenarios. These scenarios are used to validate how the interaction layer adapts to contextual conditions, device constraints, and workflow contingencies, rather than to assess clinical outcomes. All scenarios take place in a post-operative ward and involve a clinician equipped with smart glasses and a tablet, interacting with hospitalized patients as they move through corridors and rooms. The scenarios emphasize hands-free interaction, context-aware information delivery, and the coordination between edge devices and define the requirements for the interaction design.

- SCENARIO 1 - Patient recognition: Upon approaching the patient or entering the room, the doctor is supported in the patient identification. The system provides immediate visual and audio feedback, displaying highlighted clinical alerts (e.g., critical values, allergies) and summarizing the patient's current state of health before the conversation begins, preparing the physician with essential context without manual queries. Patient information is sensitive data that must be delivered only to the doctor.
- SCENARIO 2 - Managing Noisy Environments: A typical ward environment is characterized by background noise, medical equipment alarms, or other running conversations. The system needs to differentiate the doctor's voice from the surrounding acoustic clutter. The goal is to ensure that voice commands are recognized accurately without false positives caused by environmental noise or the patient speaking.
- SCENARIO 3 - Fallback Identification: In the absence of an automatic patient identification, the system allows for a manual confirmation method. The doctor confirms the patient's identity by

voice, triggering the data retrieval process.

The scenarios reported in Figure 2 were developed following a User-Centred Design (UCD) approach [24, 25], adapted to the constraints of a technology-driven research project with limited access to end-users during the early design phases.

Requirements Elicitation and Visual Scenario Design. Following a series of preliminary meetings with technical partners and domain experts, we collected the main functional and interaction requirements, which we directly translated into visual scenarios [26]. These scenarios illustrate the interaction between clinicians, patients, and the technological infrastructure across different operational contexts.

Multidisciplinary Focus Group. The visual scenarios served [27] to facilitate discussion during a structured focus group involving a heterogeneous audience: software engineers, HCI researchers, and clinicians with experience in hospital ward management. The use of visual scenarios enabled participants from different disciplinary backgrounds to develop a shared understanding of the proposed interaction paradigms and to identify potential criticalities, contextual constraints, and alternative solutions. The visual representation proved particularly effective in bridging the communication gap between technical and medical perspectives, allowing clinicians to comment on workflow feasibility while enabling engineers to assess technical implementability.

Scenario Refinement and Semi-Formal Specification. The initial scenarios were revised to incorporate feedback received during the focus group, resolving ambiguities and addressing identified critical issues. The refined scenarios were then transformed into a set of semi-formal requirements and technical specifications which add up to the general system requirements reported at the start of the paragraph.

- SCENARIO 1 – Patient recognition:
  - The system must automatically recognize the patient near the bed using BLE sensors.
  - The system prompts the doctor to confirm the patient's identity with minimal interaction.
  - The system must display on the smart glasses the patient's identity and prioritized clinical alerts (critical values, allergies, recent exams) using synthetic, high-contrast visual cues.
  - The system must deliver sensitive information through a private audio channel audible only to the doctor.
  - The system must automatically load on the tablet a detailed view of the alerts (trends, reports, graphs) for in-depth inspection when the doctor picks it up.
- SCENARIO 2 - Managing Noisy Environments:
  - The system must distinguish the doctor's voice from background noise (alarms, other conversations, patient's voice), reducing false positives in voice commands.
  - The system must monitor ambient noise and, when audio feedback becomes unreliable, temporarily suspend voice messages while keeping visual alerts active with urgency coding.
  - The system must ensure that voice commands originate from the authorized doctor via speaker recognition, and ignore commands from the patient or other people.
- SCENARIO 3 - Fallback Identification:
  - The system must ensure uninterrupted use in the event of failure or unavailability of proximity sensors, with redundant mechanisms for patient identification and data access.
  - The system must support an alternative identification procedure based on a structured voice phrase from the doctor (e.g., "Confirm patient Mario Rossi").
- Cross-cutting privacy and security requirements that also involve the SMASH platform presented in Section 3.2:
  - The system must process voice commands using ASR and AI models running locally (edge or in local services), without storing or sending audio streams to external cloud services.
  - The system must expose its services (ASR, LLM, TTS, context orchestration) through secure APIs and comply with healthcare data protection and privacy regulations (e.g., GDPR).

4.2. Proposed Solutions

The first scenario focuses on the detection and communication of clinically relevant information during a routine bedside interaction, where the system facilitates early detection of critical conditions while enabling hands-free interaction and maintaining patient engagement. As the clinician approaches the patient's bed, patient identification is automatically performed via proximity-based BLE sensing, leveraging the patient's bracelet. To confirm the patient identity, enabling a double-check mechanism for safety, a minimal explicit action is required to the clinician that need to tap on the frame of the glasses. Once the patient is identified, the system retrieves contextual information related to the current clinical phase and recent examinations. A context-aware visual interface is designed to reduce clinician cognitive load; the visual interface on the glasses changes based on the detected context. When the analytics layer of the SMASH platform detects that a laboratory value exceeds a predefined threshold, this information is presented to the user through the smart glasses using a visual pre-attentive cue. The alert is deliberately synthetic and high-contrast, allowing rapid perception without interrupting the ongoing verbal interaction with the patient. Sensitive information is conveyed through a private audio channel using bone-conduction output, ensuring that only the clinician can access it. In parallel, the tablet automatically presents a more detailed visualization, including trend graphs of the relevant parameter over the previous days.

The scenario addresses the challenge of audio-based interaction in noisy ward environments, allowing the system to adapt the interaction modality to situational constraints, prioritizing comprehension and avoiding the risk of missed or misunderstood information. During bedside interaction, ambient noise generated by patients and surrounding activity interferes with the intelligibility of the audio feedback delivered through the smart glasses. In this scenario, the system continuously monitors environmental conditions and detects that the background noise level compromises the effectiveness of audio communication. As a result, the system temporarily suspends audio playback and waits for a more suitable moment to resume communication. Visual alerts remain active, with urgency levels encoded through graphical emphasis.

The third scenario explores system robustness in the presence of sensor failures, enabling resilience to component failures and supporting a redundant interaction mechanism to ensure continuity of care. In this scenario, the patient's BLE bracelet is unavailable due to battery depletion, preventing automatic proximity-based identification. To handle this situation, the system allows the clinician to explicitly activate patient identification through a voice command consisting of the activation word "glass" followed by the command "confirmed: John Doe". Voice-based confirmation enables the system to retrieve the correct patient context and continue the interaction without disrupting the workflow. The Vuzix Shield's integrated directional microphones are used for hardware-level beamforming. This physically isolates the wearer's audio source, significantly attenuating the patient's voice or background noise. The system is also calibrated to the specific timbre of the authorized doctor, so it ignores commands that do not match the doctor's voiceprint, preventing the patient from accidentally controlling the interface. For speech recognition, the system exploits lightweight models that run directly on the edge device or on local servers in the SMASH Platform. This ensures that voice data is ephemeral—processed instantly and never stored or transmitted to external clouds. Smart glasses provide confirmation feedback to the clinician, while alerts for abnormal parameters are displayed via compact visual indicators. When multiple parameters exceed thresholds, the smart glasses present a summarized alert with a visual marker indicating additional information, while the tablet simultaneously displays a consolidated overview with detailed graphs and values.

5. Conclusions

This paper presented an augmented interaction architecture designed to support context-aware clinical assistance in hospital wards and corridors. By combining lightweight smart glasses with a tablet-based interface, the proposed solution addresses the need for hands-free, situated access to information while preserving the ability to perform deeper analytical exploration when time and attention permit. The architecture emphasizes interaction design and workflow integration as key enablers for the effective adoption of augmented technologies in routine clinical practice. The case study illustrates how the interaction layer adapts to contextual conditions, manages uncertainty and noise, and supports redundancy in sensing and interaction mechanisms. Rather than introducing new visualization paradigms, the proposed approach focuses on structuring how, when, and where information is presented to clinicians, aligning augmented interaction with everyday ward activities.

Future work will focus on empirical validation in real hospital settings, including usability studies, assessment of cognitive load, and longitudinal observation of workflow integration. Further developments will explore tighter integration of spatial representations and 3D content where appropriate, as well as the extension of the interaction layer within the broader SMASH initiative to support additional clinical scenarios and organizational contexts. These steps will allow the proposed architecture to evolve from a validated design concept toward a deployable and scalable solution for augmented clinical interaction.

Acknowledgments

This research was funded by the Ministry of Enterprises and Made in Italy (MIMIT) – Fondo per la Crescita Sostenibile - Accordi per l'Innovazione – Project SMASH (Smart Management and Assistance System for Healthcare), D.M. 31/12/2021, grant number: F35017701-05X60, CUP B69I23003100007.

Declaration on Generative AI

During the preparation of this manuscript, the authors used generative AI tools solely to support grammar and spell checking and writing-style improvement, in accordance with the CEUR-WS activity taxonomy. The authors reviewed and edited all AI-assisted outputs and take full responsibility for the publication's content.

References

[1] B. Wang, Y. Zheng, X. Han, L. Kong, G. Xiao, Z. Xiao, S. Chen, A systematic literature review on integrating ai-powered smart glasses into digital health management for proactive healthcare solutions, npj Digital Medicine 8 (2025) 410.

[2] K. Apiratwarakul, L. W. Cheung, C. Pearkao, K. Ienghong, Smart glasses with augmented reality workflow; a modern tool for triage in mass casualty incidents, Archives of Academic Emergency Medicine 13 (2025) e53–e53.

[3] N. Zuidhof, O. Peters, P.-P. Verbeek, S. B. Allouch, Appropriation of smart glasses, European Journal of Workplace Innovation 10 (2025) 110–125.

[4] J. Sumner, H. W. Lim, A. Bundele, E. H. H. Chew, J. F. Chong, T. Koh, R. B. Sudin, A. W. Yip, Through the lens: A qualitative exploration of nurses' experiences of smart glasses in urgent care, Journal of clinical nursing 34 (2025) 948–958.

[5] A. Augello, G. Caggianese, L. Gallo, P. Neroni, Design strategies to enhance awareness in mr collaborative systems, in: L. T. De Paolis, P. Arpaia, M. Sacco (Eds.), Extended Reality, Springer Nature Switzerland, Cham, 2023, pp. 424–436.

[6] A. Augello, G. Caggianese, L. Casoria, L. Gallo, R. Montalti, P. Neroni, G. Rompianesi, R. Troisi, Augmenting collaborative practices: The hololiver case study, in: 2023 17th International Conference on Signal-Image Technology Internet-Based Systems (SITIS), 2023, pp. 234–241. doi:10.1109/SITIS61268.2023.00043.

[7] F. Marino, F. Alby, C. Zucchermaglio, M. Fatigante, Digital technology in medical visits: a critical review of its impact on doctor-patient communication 14 (2023) 1226225. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC10412815/. doi:10.3389/fpsyt.2023.1226225.

[8] L. Casoria, L. Gallo, G. Caggianese, Safeguarding face-to-face communication in augmented reality: An adaptive interface, in: 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 2022, pp. 127–132. doi:10.1109/MetroXRAINE54828.2022.9967661.

[9] B. Shneiderman, P. Maes, Direct manipulation vs. interface agents, Interactions 4 (1997) 42–61. URL: https://doi.org/10.1145/267505.267514. doi:10.1145/267505.267514.

[10] E. Horvitz, Principles of mixed-initiative user interfaces, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99, Association for Computing Machinery, New York, NY, USA, 1999, p. 159–166. URL: https://doi.org/10.1145/302979.303030. doi:10.1145/302979.303030.

[11] J. Kauttonen, R. Rousi, A. Alamäki, Trust and acceptance challenges in the adoption of ai applications in health care: Quantitative survey analysis, J Med Internet Res 27 (2025) e65567. URL: https://www.jmir.org/2025/1/e65567. doi:10.2196/65567.

[12] F. Busch, L. Hoffmann, L. Xu, L. J. Zhang, B. Hu, García-Juárez, Multinational attitudes toward AI in health care and diagnostics among hospital patients 8 (2025) e2514452. doi:10.1001/jamanetworkopen.2025.14452.

[13] J. Su, Y. Wang, H. Liu, Z. Zhang, Z. Wang, Z. Li, Investigating the factors influencing users' adoption of artificial intelligence health assistants based on an extended UTAUT model 15 (2025) 18215. URL: https://doi.org/10.1038/s41598-025-01897-0. doi:10.1038/s41598-025-01897-0.

[14] H. Yoon, H. Im, Y. Kim, Y. Song, T. Yi, Investigating preferential usability on intelligent agent-driven multimodal interaction within the extended reality environment, Archives of Design Research 37 (2024) 7–24.

[15] B. Zeng, Recent advances and future directions in extended reality (xr): Exploring ai-powered spatial intelligence, 2025. URL: https://arxiv.org/abs/2504.15970. arXiv:2504.15970.

[16] Y. Tang, J. Situ, A. Y. Cui, M. Wu, Y. Huang, Llm integration in extended reality: A comprehensive review of current trends, challenges, and future perspectives, in: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, CHI '25, Association for Computing Machinery, New York, NY, USA, 2025. URL: https://doi.org/10.1145/3706598.3714224. doi:10.1145/3706598.3714224.

[17] Z. Wang, H. Wang, H. Yu, F. Lu, Interaction with gaze, gesture, and speech in a flexibly configurable augmented reality system, IEEE Transactions on Human-Machine Systems 51 (2021) 524–534. doi:10.1109/THMS.2021.3097973.

[18] J. Bai, M. S. Sunar, N. Mohd Suaib, Augmented reality interaction: a comprehensive review of gesture and speech integration techniques 37 (2025) 11347–11377. URL: https://doi.org/10.1007/s00521-025-11190-w. doi:10.1007/s00521-025-11190-w.

[19] P. Pataranutaporn, S. Karny, C. Archiwaranguprok, C. Albrecht, A. R. Liu, P. Maes, "my boyfriend is ai": A computational analysis of human-ai companionship in reddit's ai community, 2025. URL: https://arxiv.org/abs/2509.11391. arXiv:2509.11391.

[20] R. Zhang, H. Li, H. Meng, J. Zhan, H. Gan, Y.-C. Lee, The dark side of ai companionship: A taxonomy of harmful algorithmic behaviors in human-ai relationships, in: Proc. of the 2025 CHI Conf. on Human Factors in Computing Systems, CHI '25, Association for Computing Machinery, New York, NY, USA, 2025. URL: https://doi.org/10.1145/3706598.3713429. doi:10.1145/3706598.3713429.

[21] A. Rancea, I. Anghel, T. Cioara, Edge computing in healthcare: Innovations, opportunities, and challenges, Future Internet 16 (2024). URL: https://www.mdpi.com/1999-5903/16/9/329. doi:10.3390/fi16090329.

[22] Vuzix Corporation, Vuzix shield smart glasses, https://www.vuzix.com/products/vuzix-shield-smart-glasses, 2022. Accessed: 15 Jan 2026.

[23] T. A. D'Agostino, C. L. Bylund, Nonverbal accommodation in health care communication, Health communication 29 (2014) 563–573.

[24] J. M. Carroll, M. B. Rosson, Getting around the task-artifact cycle: How to make claims and design by scenario, ACM Transactions on Information Systems (TOIS) 10 (1992) 181–212.

[25] C. Leonardi, L. Sabatucci, A. Susi, M. Zancanaro, Design as intercultural dialogue: coupling human-

centered design with requirement engineering methods, in: IFIP Conference on Human-Computer Interaction, Springer, 2011, pp. 485–502.
[26] C. Leonardi, L. Sabatucci, A. Susi, M. Zancanaro, Ahab's leg: exploring the issues of communicating semi-formal requirements to the final users, in: International Conference on Advanced Information Systems Engineering, Springer, 2010, pp. 455–469.
[27] S. L. Star, J. R. Griesemer, Institutional ecology,translations' and boundary objects: Amateurs and professionals in berkeley's museum of vertebrate zoology, 1907-39, Social studies of science 19 (1989) 387–420.