AI-Assisted Voice Intake and Triage System

Patrick Keufack[1], Penelope Lopez[1],
Damian Archer[2], Yetunde Folajimi[1],
Salem Othman[1], Leonidas
Deligiannidis[1]

[1]School of Computing and Data
Science, Wentworth Institute of
Technology, Boston, MA 02115, USA
{keufackp, lopezp, folajimiy,
othmans1, deligiannidisl}@wit.edu

[2]Outer Cape Health Services,
Harwich Port, MA 02646, USA
darcher@outercape.org

***Abstract*—**Healthcare call centers frequently experience high call volumes, long hold times, and manual documentation burdens that can delay the identification of urgent patient needs. This paper presents ClearPath Health, an AI-assisted voice intake and triage prototype developed in collaboration with Outer Cape Health Services (OCHS), a community healthcare organization. The system uses a conversational voice interface and a Large Language Model (LLM)-based natural language understanding component to collect patient-reported symptoms and structure intake information. Final urgency classification is constrained by an OCHS three-tier clinical protocol that routes calls into emergency, same-day urgent, or non-urgent categories. Based on this classification, the system supports escalation, appointment routing, and structured call logging through a human-in-the-loop workflow. A controlled pilot study with 12 simulated patient-call scenarios demonstrated the feasibility of the prototype for voice-based intake, emergency escalation, and non-urgent scheduling support. The results should be interpreted as proof-of-concept evidence rather than clinical validation. Overall, the system illustrates how lightweight, protocol-constrained conversational AI may support call center staff by reducing routine documentation burden while preserving human oversight for clinical decision-making.

Keywords—AI in Healthcare, Voice Assistants, Medical Triage, Natural Language Processing, Human-in-the-Loop Systems, Call Center Automation

I. INTRODUCTION

Healthcare call centers serve as an important access point for patients seeking medical care, especially in community health settings where phone-based communication remains a major channel for care coordination. For patients who prefer direct conversation, have limited engagement with digital portals, or need help explaining symptoms, the phone remains an important connection to clinical staff. However, high call volumes can create long hold times, delayed message routing, and additional documentation burden for call center representatives and nurses. Traditional Interactive Voice Response (IVR) systems provide basic call routing through fixed menu options, such as “press 1 for pharmacy” or “press 2 for appointments.” While these systems can reduce some administrative workload, they are limited in their ability to collect structured clinical information, interpret patient-described symptoms, or identify high-risk cases dynamically. Prior work on symptom checkers and digital triage tools has shown that automated triage systems require careful evaluation because accuracy and safety can vary across systems and use cases [8]. Recent advances in conversational agents, natural language processing, and Large Language Models (LLMs) have created new opportunities for AI-assisted patient intake. Conversational agents have been studied for patient education, health behavior support, and clinical documentation assistance [6], [7], [13]. More recent work on LLMs in medicine has also shown their potential for clinical language understanding, while emphasizing the need for careful evaluation, human review, and safety controls before use in real healthcare workflows [12]. Healthcare applications of AI also require attention to privacy, transparency, accountability, and human oversight. The World Health Organization emphasizes that AI systems used in health should be designed with ethical governance, patient safety, and appropriate regulatory considerations in mind [10], [11]. For this reason, ClearPath Health is not designed to replace nurses or clinical staff. Instead, it supports a human-in-the-loop workflow by organizing patient-reported information and routing cases according to a predefined clinical protocol. This paper presents ClearPath Health, an AI-assisted voice intake and triage prototype developed in collaboration with Outer Cape Health Services (OCHS). The system uses a conversational voice interface to collect patient-reported symptoms, extract structured intake information, and route calls according to a predefined three-tier clinical triage protocol provided by OCHS. The three categories include emergency cases requiring escalation, same-day urgent cases requiring prompt follow-up, and non-urgent cases appropriate for standard scheduling. The main contribution of this work is a lightweight, human-in-the-loop prototype that combines LLM-based natural language understanding with protocol-constrained triage logic. Rather than replacing call center staff or nurses, the system is designed to reduce routine documentation burden, support early identification of urgent cases, and provide structured call logs for review. The prototype is evaluated through a controlled pilot study using simulated patient-call scenarios, and the results are presented as proof-of-concept evidence for AI-assisted voice intake in a community healthcare environment.

II. BACKGROUND AND RELATED WORK

Healthcare call centers play an important role in patient access, scheduling, symptom reporting, and communication between patients and clinical teams. In many healthcare settings, call center staff collect patient concerns, document symptoms, and route messages to nurses or providers. However, this process can become inefficient when call volumes are high, especially when staff must manually interpret unstructured patient descriptions and decide how quickly a message should be escalated. Traditional IVR systems are commonly used to route calls, but they rely on fixed menu structures and do not interpret patient symptoms. As a result, they are limited for clinical intake workflows where patients may describe symptoms in different ways, provide incomplete details, or require follow-up questions. This limitation creates an opportunity for conversational systems that can collect structured information through natural language interaction. Prior research on conversational agents in healthcare has explored their use in patient education, behavior change, and communication support [6], [7], [13]. These systems show that conversational interfaces can improve accessibility and engagement, particularly when users benefit from a natural dialogue rather than a form-based or menu-based interaction. However, healthcare conversational agents must be designed carefully because errors in interpretation can affect safety and care coordination. Digital symptom checkers and automated triage systems are also relevant to this work. Existing studies have shown that symptom checker performance can vary depending on the condition, input quality, and triage task [8]. This supports the need for careful evaluation and conservative design when using automated systems for healthcare triage. ClearPath Health addresses this concern by using AI for intake and language understanding, while constraining final classification through a predefined clinical triage protocol. Large Language Models (LLMs) have recently demonstrated strong capabilities in medical question answering and clinical language understanding [9], [12]. These capabilities make LLMs useful for extracting symptoms, identifying relevant details, and structuring patient-reported information. At the same time, LLMs may produce incorrect or unsupported outputs if used without safeguards. For this reason, the system presented in this paper does not rely on the LLM as an independent clinical decision-maker. Instead, the LLM supports natural language understanding, while the triage decision is guided by OCHS clinical rules. Healthcare AI systems also require attention to ethics, privacy, accountability, and human oversight. The World Health Organization emphasizes that AI systems in health should be governed by principles such as transparency, safety, responsibility, and protection of patient data [10], [11]. These concerns are especially important for voice-based systems because patient calls may contain sensitive health information. ClearPath Health follows a human-in-the-loop approach in which AI assists with intake, routing, and documentation, while clinical decision-making remains under human supervision. Compared with traditional IVR systems and general symptom checkers, ClearPath Health focuses on a narrower and more practical use case: AI-assisted voice intake for a community healthcare call center. The system is designed around OCHS workflow needs, including emergency escalation, same-day urgent routing, non-urgent scheduling support, and structured call logging. This makes the project less focused on autonomous diagnosis and more focused on improving the efficiency and consistency of patient intake.

III. PROJECT DESCRIPTION AND IMPLEMENTATION

The ClearPath Health system is designed as a modular AI-assisted voice intake and triage prototype that connects patient voice input to structured clinical routing actions. The system architecture, shown in Fig. 1, and the logical workflow, shown in Fig. 2, illustrate how unstructured patient speech is converted into structured intake data, triage classification, and appropriate follow-up actions. The system was developed around the workflow needs of Outer Cape Health Services (OCHS). During the design process, OCHS provided clinical triage guidance that separated patient calls into three urgency levels: Category 1 emergency cases, Category 2 same-day urgent cases, and Category 3 non-urgent cases. This protocol was used as the foundation for the system's classification and routing logic.

A. System Architecture

The architecture is divided into three primary layers: the Interface Layer, the Intelligence Layer, and the Action Layer.

- **Interface Layer**: It uses Vapi as the voice interaction platform. This layer manages the conversation between the caller and the AI assistant, including speech input, transcription, and spoken responses. Instead of requiring patients to navigate a fixed IVR menu, the caller can describe symptoms in natural language.
- **Intelligence Layer**: This Layer serves as the core language understanding component of the system. It uses a Large Language Model (GPT-4.1 mini via Azure OpenAI) configured with a system prompt based on the OCHS three-tier triage protocol. The model supports entity extraction, including patient name, callback number, symptom description, duration, and severity. It also supports intent classification by identifying whether the call relates to an emergency, same-day urgent concern, or non-urgent scheduling need.
- **Action Layer**: It executes the appropriate workflow based on the triage category. Category 1 cases are routed for emergency escalation, Category 2 cases are routed for same-day follow-up, and Category 3 cases are routed for non-urgent scheduling. The system also records structured call data for later review through the dashboard.

B. Triage Logic and System Design

A common limitation of traditional voice assistants is their reliance on rigid keyword-based or "if-then" logic. ClearPath Health uses a hybrid AI-assisted approach in which the LLM interprets patient speech, while final triage behavior is constrained by the OCHS clinical protocol. This design prevents the system from acting as an independent diagnostic tool and keeps the focus on intake support and routing. The OCHS protocol defines three triage levels. Category 1 includes emergency symptoms that require immediate escalation, such as chest pain, severe difficulty breathing, stroke-like symptoms, loss of consciousness, uncontrolled bleeding, or other high-risk presentations. Category 2 includes same-day urgent concerns that require prompt follow-up, such as high fever, urinary symptoms with additional risk factors, abdominal pain with concerning symptoms, asthma or COPD exacerbation, or possible blood clot symptoms. Category 3 includes non-urgent concerns appropriate for standard scheduling, such as mild cough, rash, uncomplicated urinary symptoms, back pain without neurological symptoms, or other lower-acuity complaints. The system follows a safety-first design. When patient speech is incomplete, unclear, or interrupted, the assistant is designed to ask clarifying follow-up questions before assigning a category. If the caller cannot provide enough information or if the symptoms suggest possible risk, the system prioritizes escalation rather than under-classification. This supports a "do no harm" approach in which unnecessary escalation is considered safer than missing a critical condition.

C. Technical Implementation

The prototype was implemented using a combination of voice AI services, backend processing, structured data storage, and dashboard visualization.

1. **Natural Language Understanding (NLU)**:The system uses GPT-4.1 mini through Azure OpenAI as the primary language understanding component. The model is configured through prompt engineering rather than fine-tuning. The prompt includes the OCHS triage categories, required intake questions, escalation rules, and instructions for handling incomplete or ambiguous patient responses.
2. **Voice Interaction**: Vapi is used to manage the voice-based conversation. The assistant collects patient information, asks follow-up questions, and triggers tool calls based on the detected triage category. This allows the system to support a natural conversation while still following a structured workflow.
3. **Backend Integration**: A Python backend receives webhook data from Vapi and processes the structured call information. The backend stores call records, including patient details, symptoms, classification results, and actions taken.
4. **Database Logging**: SQLite is used to store call records locally during prototype testing. This improved the reliability of the system by creating a persistent record of each processed call instead of relying only on temporary call logs.
5. **Dashboard**: A Streamlit dashboard was developed for oversight and review. The dashboard displays call logs, triage categories, actions performed, and other structured information. This supports the human-in-the-loop design by allowing staff or reviewers to monitor system outputs.
6. **API Orchestration**: The system uses structured tool calls to connect triage classification with system actions. For example, emergency cases trigger escalation logic, while non-urgent cases can be routed toward scheduling support. Google Calendar API integration was used to support appointment-related workflows during prototype development [2].

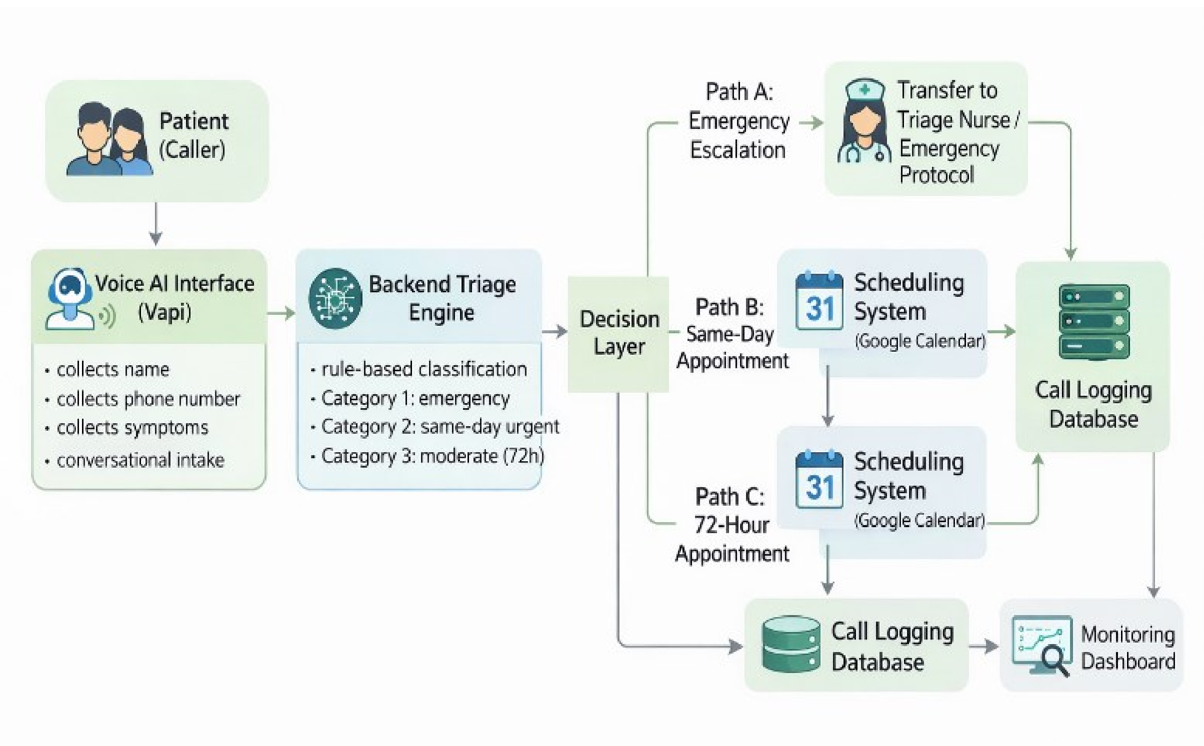

Fig. 1. System architecture of the AI-assisted medical triage system.

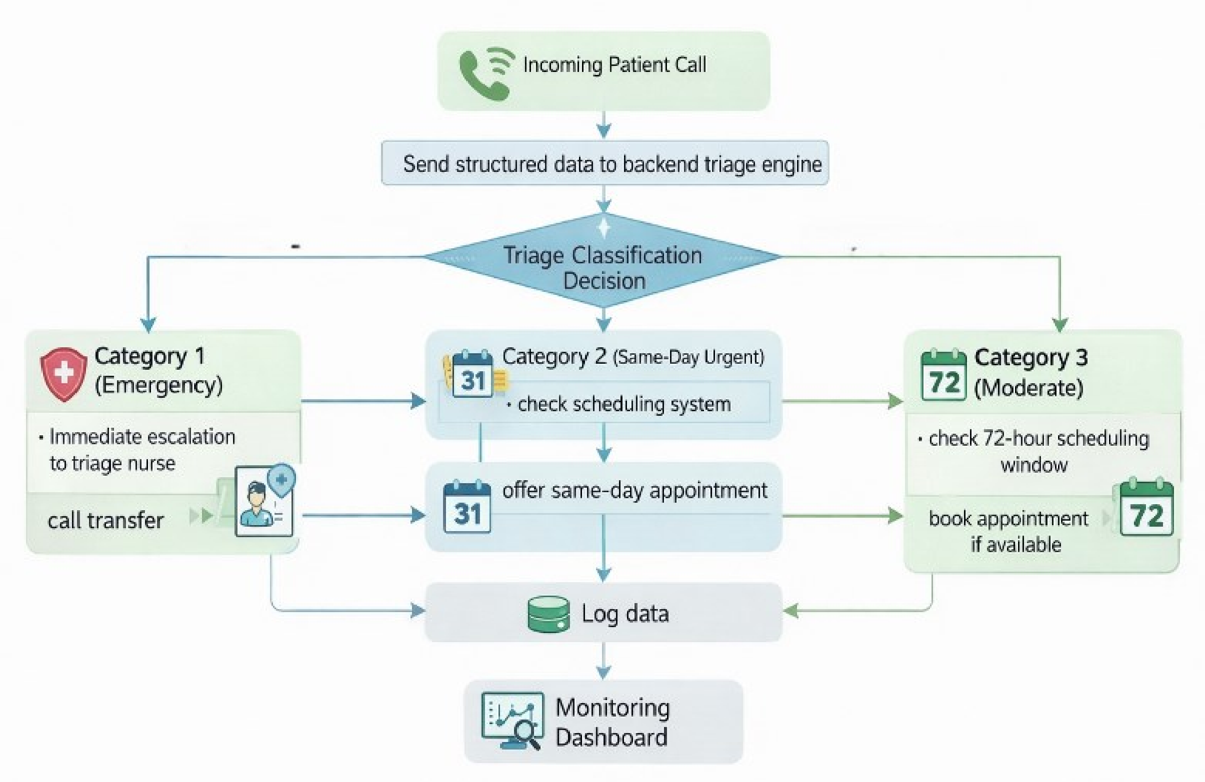

Fig. 2. Patient call workflow and triage decision process.

IV. Preliminary Results and Evaluation

To evaluate the prototype, we conducted a controlled pilot study using 12 simulated patient-call scenarios. The participant group included 5 native English speakers, 4 non-native English speakers with varying levels of fluency and accent, and 3 native Spanish speakers. Participants were recruited from the student population and the authors' extended network. The scenarios were designed to represent the three OCHS triage categories: Category 1 emergency cases, Category 2 same-day urgent cases, and Category 3 non-urgent cases. Participants were asked to either follow or paraphrase symptom scenarios during voice interactions with the system. Because this was a controlled proof-of-concept study, the results should not be interpreted as clinical validation.

A. Operational Evaluation Framework

Before analyzing system performance, we defined three evaluation criteria. These criteria were selected to measure whether the prototype could support structured voice intake and routing in a simulated call center workflow:

1. First, triage classification was evaluated by comparing the system's assigned category with the expected category defined by the OCHS clinical triage protocol. This measured whether the system routed emergency, same-day urgent, and non-urgent cases appropriately.
2. Second, entity extraction was evaluated by checking whether the system captured the required intake information, including patient name, callback number, symptom description, and other relevant details when provided by the caller.
3. Third, action execution was evaluated by checking whether the system triggered the appropriate workflow after classification, such as emergency escalation, same-day follow-up, or non-urgent scheduling.

All classifications were verified against the predefined OCHS clinical triage guidelines to ensure consistency in evaluation.

B. System Performance Results

The results of the 12-scenario pilot study are summarized in Table I. Overall, the system demonstrated the ability to collect structured intake information, classify simulated symptom descriptions, and trigger the expected routing action in controlled testing conditions. The system correctly identified all simulated Category 1 emergency scenarios and routed them for escalation. It also successfully routed Category 3 non-urgent scenarios toward scheduling support. In same-day urgent scenarios, the system generally followed the expected routing behavior. One scenario involving asthma symptoms was classified at a higher urgency level than the target category. This was treated as a safety-pass outcome because the system escalated the case rather than under-classifying a potentially risky respiratory concern.

Table I: Performance Metrics Across 12 Simulated Pilot Scenarios

Language	Target	System output	Data Extraction Complete?	Result
English	Category 1 (Chest Pain)	Category 1	Yes	Success (Auto-Transfer)
English	Category 3 (Back Pain)	Category 3	Yes	Success (Scheduled)
Spanish	Category 2 (Fever/UTI)	Category 2	Yes	Success (Same-Day)
English	Category 1 (Stroke)	Category 1	Yes	Success (Auto-Transfer)
Spanish	Category 1 (Breathing)	Category 1	Yes	Success (Auto-Transfer)
English	Category 2 (Abdominal)	Category 2	Yes	Success (Same-Day)
English	Category 3 (Rash)	Category 3	Yes	Success (Scheduled)
English	Category 2 (Blood/Urine)	Category 2	Yes	Success (Same-Day)
English	Category 1 (Trauma)	Category 1	Yes	Success (Auto-Transfer)
English	Category 3 (Cough)	Category 3	Yes	Success (Scheduled)
English	Category 2 (Asthma)	Category 1 (Override)	Yes	Safety-Pass (Higher Triage)
Spanish	Category 2 (Flank Pain)	Category 2	Yes	Success (Same-Day)

C. Discussion of Results

The pilot results suggest that ClearPath Health can support structured voice-based intake and triage routing in controlled testing conditions. The system was able to process patient-reported symptoms, extract required intake information, and apply the OCHS triage protocol across emergency, urgent, and non-urgent scenarios. The Category 1 emergency cases are particularly important because missing these cases could create safety risks. In the pilot study, all simulated emergency cases were routed for escalation. The asthma scenario also demonstrated the safety-first design of the system. Although the target category was Category 2, the system classified the case as Category 1 due to respiratory risk. While this reduced category precision, it aligned with a conservative triage strategy where over-escalation is preferable to under-escalation. The multilingual scenarios also suggest that the system may support basic cross-language intake under controlled conditions. Spanish-language cases were successfully processed during testing. However, this result should be interpreted cautiously because the pilot sample was small and did not include enough participants to fully evaluate multilingual performance across accents, dialects, and mixed-language speech. Overall, the evaluation demonstrates proof-of-concept feasibility rather than real-world clinical validation. Larger studies with real call center conditions, broader patient populations, and direct comparison against existing IVR or manual workflows would be needed before deployment in a clinical environment.

V. Discussion and Limitation

The preliminary results suggest that ClearPath Health can support structured voice-based intake and triage routing in controlled testing conditions. The system demonstrated the ability to collect patient-reported symptoms, extract structured intake information, and route cases according to the predefined OCHS triage protocol. These findings suggest that conversational intake combined with protocol-based routing may offer potential advantages over traditional IVR workflows, especially in flexibility and completeness of data collection. However, several limitations should be considered. First, the evaluation was conducted as a small-scale pilot study using 12 simulated patient-call scenarios rather than real patient calls. As a result, the findings may not fully reflect real-world variability in patient speech, background noise, emotional distress, interruptions, or incomplete symptom descriptions. Second, speech-to-text and language model interpretation errors remain possible. Strong accents, unclear speech, mixed-language communication, or ambiguous symptom descriptions may affect the quality of extracted information. However, similar challenges are also present in human-led triage, where communication errors, incomplete descriptions, or misinterpretation can occur during patient interactions. Third, the system follows a safety-first design that may lead to over-classification of higher urgency categories. While this can increase the number of escalations, it reflects a deliberate “do no harm” approach, where missing a critical condition is considered more severe than triggering an unnecessary escalation. Fourth, the current evaluation did not include a direct baseline comparison with an existing IVR system or manual call center workflow. Therefore, claims about efficiency improvement remain preliminary. Future work should compare ClearPath Health against traditional IVR and human-only intake workflows using metrics such as call completion time, documentation completeness, escalation accuracy, and staff workload. Finally, real-world deployment would require stronger privacy, security, and system integration controls. Because patient calls may contain protected health information, future versions would need secure data transmission, access control, audit logging, and compliance with healthcare privacy requirements before use in a clinical environment [10], [11]. Integration with electronic health records would also require additional validation, governance, and clinical oversight.

VI. Conclusion

This paper presented ClearPath Health, an AI-assisted voice intake and triage prototype developed in collaboration with Outer Cape Health Services. The system combines a conversational voice interface, LLM-based natural language understanding, protocol-constrained triage logic, and structured call logging to support healthcare call center workflows. The prototype demonstrates the feasibility of using conversational AI to collect patient-reported symptoms, organize intake information, and route calls into emergency, same-day urgent, and non-urgent categories. The system is not intended to replace nurses or clinical staff. Instead, it supports a human-in-the-loop workflow by reducing routine documentation burden and helping prioritize cases that may require timely attention. Preliminary testing with 12 simulated patient-call scenarios suggests that the system can support structured intake and triage routing under controlled conditions. However, the results should be interpreted as proof-of-concept evidence rather than clinical validation. Future work will require larger-scale testing, direct comparison with existing IVR and manual workflows, stronger privacy and security controls, and closer integration with clinical systems. Overall, ClearPath Health illustrates how lightweight, safety-focused conversational AI tools may assist community healthcare organizations by improving patient intake workflows while preserving human oversight in clinical decision-making.

Acknowledgment

The authors would like to thank Outer Cape Health Services (OCHS) for their collaboration and for providing clinical insights that guided the development of this system.

References

[1] Vapi AI Voice Platform. [Online]. Available: https://vapi.ai[Accessed: March 7, 2026].

[2] Google Calendar API. [Online]. Available: https://developers.google.com/calendar[Accessed : March 7, 2026].

[3] A. Rajkomar, J. Dean, and I. Kohane, "Machine Learning in Medicine," *New England Journal of Medicine* , vol. 380, no. 14, p. 1347–1358, 2019.

[4] E. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine* , vol. 25, no. 1, pp. 44–56, 2019.

[5] Outer Cape Health Services (OCHS), "Clinical Triage and Appointment Scheduling Protocols," Internal Document, 2026.

[6] A. B. Kocaballi et al., "Envisioning an artificial intelligence documentation assistant for future primary care consultations," Journal of the American Medical Informatics Association, vol. 27, no. 11, pp. 1695–1704, 2020.

[7] S. Bickmore, D. Schulman, and T. Sidner, "Automated interventions for multiple health behaviors using conversational agents," Patient Education and Counseling, vol. 92, no. 2, pp. 142–148, 2013.

[8] J. Liang et al., "Evaluation and accuracy of symptom checkers: systematic review," BMJ Open, vol. 9, no. 12, 2019.

[9] OpenAI, "GPT-4 Technical Report," 2023. [Online]. Available : https://arxiv.org/abs/2303.08774.[Accessed: March 2026].

[10] World Health Organization, Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva, Switzerland: World Health Organization, 2021.

[11] World Health Organization, Regulatory Considerations on Artificial Intelligence for Health. Geneva, Switzerland: World Health Organization, 2023.

[12] K. Singhal et al., "Large language models encode clinical knowledge," Nature, vol. 620, pp. 172–180, 2023.

[13] A. Laranjo et al., "Conversational agents in healthcare: A systematic review," Journal of the American Medical Informatics Association, vol. 25, no. 9, pp. 1248–1258, 2018.