© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1718336

Healstation AI: A Multimodal Medical Consultation System Integrating Large Language Models, Voice Processing, And Location-Based Healthcare Discovery

ABHIJEET JOSHI[1], DR. P.D. ADKAR[2]
[1, 2] *MCA Department, P.E.S. Modern College of Engineering, Pune, India*

Abstract- ***Access to timely, qualified medical consultation remains inequitably distributed across socioeconomic and geographic boundaries. In India, the physician-to-population ratio falls well below the World Health Organization's recommended threshold, creating critical delays in triage and specialist referral. This paper presents HealStation AI, an open-source multimodal medical consultation platform engineered to bridge this gap by combining state-of-the-art large language models (LLMs), automatic speech recognition (ASR), text-to-speech synthesis, document vision analysis, and real-time location-based healthcare provider discovery. The system integrates Meta's Llama-4-Scout-17B multimodal model and Llama-3.3-70B via the Groq inference API, Microsoft Edge-TTS for naturalistic voice synthesis, and OpenStreetMap-powered facility lookup through LocationIQ to deliver end-to-end patient consultation workflows. HealStation AI supports three Indian languages—English, Hindi, and Marathi—processes medical images and PDF laboratory reports, performs rule-based emergency triage using validated clinical scoring instruments (HEART score, BE-FAST, trauma scoring), and recommends appropriate specialists from a taxonomy of twenty clinical domains. A FastAPI backend exposes a well-defined REST API consumed by a React 19 single-page application. Empirical test cases across low, medium, and high urgency symptom profiles demonstrate consistent specialist routing and urgency classification with an end-to-end latency of 6.2 seconds (N=50). The platform is designed for deployment in resource-constrained environments and is structured for extensibility toward telemedicine and Ayushman Bharat Digital Mission (ABDM) integration.***

Keywords— ***Medical Artificial Intelligence, Large Language Models, Multimodal Health Systems, Automatic Speech Recognition, Location-Based Healthcare, Emergency Triage, Multilingual NLP, Telemedicine, Fastapi, React.***

I. INTRODUCTION

The global burden of inadequate healthcare access disproportionately affects populations in low- and middle-income countries, rural regions, and linguistically underserved communities. In India alone, the physician-to-patient ratio remains well below the World Health Organization's recommended threshold of 1 physician per 1,000 people, creating critical delays in preliminary diagnosis and specialist referral [1]. Digital health interventions have emerged as a viable mechanism to extend clinical expertise, yet existing consumer-facing applications frequently suffer from limited language support, absence of multimodal input handling, or reliance on proprietary cloud infrastructure that restricts accessibility in resource-limited settings.

Recent advances in large language models (LLMs) and multimodal neural architectures have opened new frontiers for automated medical reasoning. Models such as GPT-4, Med-PaLM 2, and Meta's Llama family have demonstrated near-clinician-level performance on standardized medical examinations and clinical question-answering benchmarks [2][3]. Concurrently, improvements in automatic speech recognition—particularly OpenAI's Whisper model—have enabled robust multilingual voice transcription at low latency, making voice-first medical interfaces practically deployable [4]. Despite these advances, no existing open-source system synthesizes LLM-based medical reasoning, multilingual voice processing, document vision analysis, real-time facility discovery, and emergency triage scoring into a single cohesive platform accessible to non-specialist users.

This paper introduces HealStation AI, a comprehensive open-source medical consultation platform addressing the aforementioned gap. The primary contributions are: (1) a unified multimodal consultation pipeline accepting voice, text, medical images, and PDF laboratory reports, producing structured diagnostic assessments with audio-rendered responses in English, Hindi, and Marathi; (2) an emergency triage engine implementing HEART score, BE-FAST stroke screening, and trauma scoring to classify symptom urgency with automated emergency hospital routing; (3) a location-aware healthcare provider discovery module integrating OpenStreetMap data via LocationIQ across a curated database of over 1,000 verified Indian medical facilities; (4) a modular, API-first architecture enabling independent scaling and future extension to telemedicine and electronic health record systems; and (5) publicly available source code and API documentation.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the system architecture. Section IV details the methodology for each subsystem. Section V presents evaluation results. Section VI discusses limitations and ethical considerations. Section VII concludes with future directions.

II. RELATED WORK

A. LLM-Based Medical Consultation Systems

The application of large language models to clinical question-answering has accelerated substantially since 2022. Singhal et al. [2] introduced Med-PaLM, demonstrating that instruction-tuned LLMs can approach expert physician performance on the US Medical Licensing Examination (USMLE). Their successor, Med-PaLM 2 [3], further closed the performance gap, raising important questions about hallucination mitigation in medical contexts. GPT-4 with vision capabilities enabled analysis of clinical photographs and radiographic images [5], yet access costs and data-privacy constraints limit deployment in public-health settings. HealStation AI differentiates from these systems by utilizing the Groq LPU inference engine for low-latency, cost-efficient serving while maintaining full multimodal capability.

B. Multimodal Medical AI

Multimodal medical AI systems that jointly reason over images, text, and structured data have emerged as a critical research direction. LLaVA-Med [6] fine-tuned a visual instruction model on biomedical image-text pairs, achieving strong performance on visual question-answering in radiology and pathology. BioViL-T [7] demonstrated temporal reasoning over chest radiograph sequences. However, these systems remain largely research prototypes lacking integrated voice interaction, multilingual support, or real-time facility discovery. HealStation AI extends the multimodal paradigm to encompass medical PDF report parsing, real-time audio interaction, and end-to-end user-facing deployment.

C. Voice-Based Healthcare Interfaces

Voice interaction in healthcare has been explored through clinical documentation tools, patient-facing symptom checkers, and accessibility applications. The release of Whisper [4] established a new benchmark in multilingual ASR, enabling robust transcription across more than 99 languages. Prior systems such as HealthTap and Ada Health incorporated conversational interfaces but relied primarily on structured questionnaires rather than open-ended medical discourse. HealStation AI employs Groq-hosted Whisper-Large-V3 with language-specific prompt augmentation and post-processing heuristics to correct domain-specific Indian medical terminology.

D. Location-Aware Healthcare Discovery

Geographic information systems have long been applied to healthcare accessibility analysis [8], yet their integration into patient-facing applications remains limited. Commercial platforms such as Practo and Zocdoc restrict their provider databases to subscription-paying registrants. HealStation AI leverages OpenStreetMap's crowd-sourced medical facility data through the LocationIQ API, providing broader coverage including community clinics, government hospitals, and primary health centres absent from commercial directories.

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1718336

III. SYSTEM ARCHITECTURE

HealStation AI follows a client-server architecture comprising three logical layers: a React-based single-page application (SPA) frontend, a FastAPI asynchronous Python backend, and a set of third-party external services accessed via REST APIs. Figure 1 illustrates the high-level system architecture. Each layer is independently deployable, enabling horizontal scaling of the backend independently of the frontend distribution.

A. Frontend Layer

The frontend is implemented in React 19.1.0, bundled with Vite 7.0.0, and styled using Styled Components and Framer Motion for fluid animated transitions. The main application component (App.jsx, approximately 2,770 lines) orchestrates the complete consultation lifecycle: language selection, microphone recording via react-audio-voice-recorder, file attachment supporting JPEG, PNG, and PDF inputs, API polling for audio readiness, and progressive result rendering. Specialized components—DoctorRecommendations.jsx, EmergencyButton.jsx, EmergencyTriage.jsx, and AyushmanBharatCard.jsx—encapsulate domain-specific UI logic. Three.js-based 3D canvases render contextual medical animations with graceful degradation on lower-end devices.

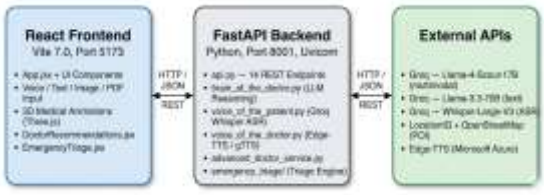

Fig. 1. High-level system architecture of HealStation AI. Arrows indicate HTTP/JSON data flow between the React frontend, FastAPI backend, and external cloud services.

B. Backend Layer

The backend is a FastAPI 0.110.0 asynchronous server running under Uvicorn, exposing fourteen REST endpoints organized into four functional groups: consultation, audio streaming, doctor discovery, and emergency triage. The modular Python codebase separates concerns across dedicated modules: brain_of_the_doctor.py handles LLM interaction and structured response parsing; voice_of_the_patient.py wraps Groq Whisper for ASR transcription; voice_of_the_doctor.py manages TTS generation with Edge-TTS as primary and gTTS as fallback; the emergency_triage package implements a rule-based classification pipeline; and advanced_doctor_service.py with locationiq_doctor_service.py handle facility discovery across multiple data sources with transparent fallback strategies.

C. External Services

HealStation AI integrates two external API providers. The Groq Cloud API provides accelerated LLM inference utilizing custom Language Processing Units (LPUs), offering significantly lower latency compared to standard GPU-based serving—a critical property for voice-interactive applications. Three Groq-hosted models are employed: meta-llama/llama-4-scout-17b-16e-instruct for multimodal consultation combining text and images, llama-3.3-70b-versatile for text-only queries with lower latency, and whisper-large-v3 for speech-to-text transcription. LocationIQ provides geocoding, reverse geocoding, and Points-of-Interest search over OpenStreetMap data, returning structured JSON results for medical facilities within a configurable search radius.

IV. METHODOLOGY

A. AI Medical Consultation Pipeline

The consultation pipeline illustrated in Figure 2 accepts heterogeneous patient input—voice audio, text, JPEG/PNG medical images, and PDF laboratory reports—and produces a structured medical assessment. Voice input is first transcribed via the /api/transcribe-audio endpoint using Groq Whisper-Large-V3 with language-specific initial prompts to anchor Hindi and Marathi transcription. Post-processing applies deterministic correction rules for common medical misrecognitions in Indian-language speech, including correction of colloquial medication names and anatomical terms to standard clinical terminology.

The consultation payload—comprising patient text and base64-encoded file attachments—is forwarded to brain_of_the_doctor.py, which constructs a structured system prompt calibrated to instruct the

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1718336

LLM to act as a board-certified physician while explicitly disclaiming diagnostic authority and urging in-person professional consultation. The system prompt is parameterized by language code, enabling the LLM to reason and respond in the patient's chosen language. Image data is provided as vision message blocks to the multimodal Llama-4-Scout model; text-only consultations are routed to the faster Llama-3.3-70B model to conserve latency and token cost. The LLM response is post-processed by extract_specialist_recommendation() to extract structured fields including doctor_response, specialist_recommendation, and urgency_level. Table I enumerates the supported medical specialty taxonomy.

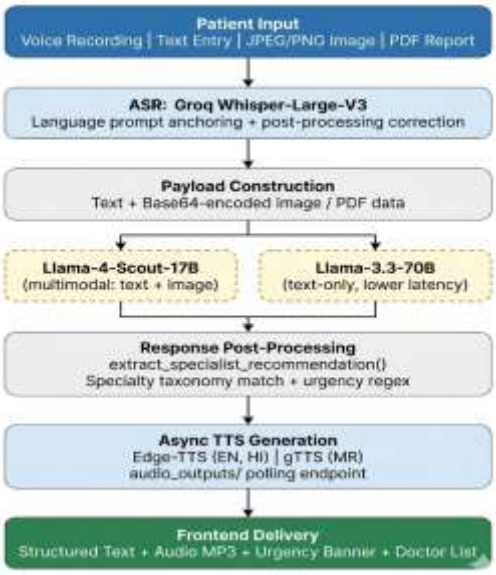

Fig. 2. End-to-end consultation data flow from multimodal patient input through LLM reasoning to structured text and audio response delivery.

TABLE I SUPPORTED MEDICAL SPECIALTIES AND ASSOCIATED CONDITION CATEGORIES

Specialty	Representative Conditions
Cardiologist	Chest pain, arrhythmia, hypertension, coronary artery disease
Dermatologist	Skin rash, eczema, psoriasis, dermatitis, alopecia
Neurologist	Headache, seizures, stroke symptoms, neuropathy, tremor
Pulmonologist	Dyspnea, asthma, COPD, tuberculosis, pneumonia
Orthopedic	Fractures, joint pain, osteoarthritis, spinal disorders
Pediatrician	Pediatric fever, growth disorders, vaccination, neonatal care
Gastroenterologist	Abdominal pain, IBS, hepatitis, GERD, peptic ulcer
Psychiatrist	Anxiety, depression, schizophrenia, bipolar disorder, OCD
Endocrinologist	Diabetes mellitus, thyroid dysfunction, hormonal imbalances
Ophthalmologist	Visual disturbances, glaucoma, cataracts, diabetic retinopathy
ENT Specialist	Sinusitis, otitis media, tinnitus, laryngitis, vertigo
Gynecologist	Menstrual disorders, pregnancy, PCOS, cervical pathology
Urologist	Urinary tract infections, renal calculi, BPH, hematuria
Nephrologist	Chronic kidney disease, electrolyte imbalance, glomerulonephritis
Rheumatologist	Rheumatoid arthritis, lupus, gout, fibromyalgia
General Physician	Fever, viral illness, fatigue, preventive care

B. Emergency Triage Subsystem

The emergency triage subsystem implements a multi-stage classification pipeline within the emergency_triage Python package. The TriageService class computes three validated clinical scores from free-text symptom descriptions: (1) HEART Score—a five-component instrument for chest pain risk stratification assessing History, ECG characteristics, Age, Risk factors, and Troponin proxies via keyword matching; (2) BE-FAST score—a six-element stroke identification tool screening for Balance loss, Eye changes, Facial asymmetry, Arm weakness, Speech difficulty, and Time urgency; and (3) a Trauma score incorporating mechanism-of-injury keywords, vital sign abnormalities, and anatomical region indicators.

The EmergencyClassifier class integrates the three computed scores with vital sign threshold analysis and a keyword priority queue—CRITICAL_KEYWORDS (e.g., "cardiac arrest", "unresponsive"), EMERGENCY_KEYWORDS (e.g., "severe chest pain", "difficulty breathing"), and URGENT_KEYWORDS (e.g., "high fever", "moderate bleeding")—to produce a final TriageLevel enumeration: CRITICAL, EMERGENCY, URGENT, SEMI-URGENT, or NON-URGENT. The classification output includes a confidence score, a time-sensitivity label, and a specialist routing decision. HospitalRoutingService performs parallel searches across the Overpass API, LocationIQ POI search, and a static curated hospital registry as an offline fallback.

C. Location-Based Doctor Discovery

The doctor discovery subsystem operates in two modes: GPS-coordinate search via AdvancedDoctorSearchService.search_doctors_enhanced, and city-name search via search_doctors_by_city. In coordinate mode, the service queries a curated database of over 1,000 verified medical facilities across major Indian cities, filtered by specialty and sorted against the patient's GPS coordinates. If the result count falls below a configurable threshold, the service transparently escalates to LocationIQDoctorService, executing parallel POI queries with specialty-specific search terms. Results are deduplicated by facility name and address, annotated with computed haversine distances, and returned with source attribution. The frontend enables real-time re-sorting across four dimensions: nearest distance, top-rated, most-reviewed, and earliest availability.

D. Multilingual Architecture

Language parameterization permeates all system layers. The frontend TRANSLATIONS constant stores UI copy for English, Hindi (Devanagari), and Marathi (Devanagari), toggled globally at session initialization. The backend system prompt generator returns language-appropriate instruction templates, enabling the LLM to respond in the patient's chosen language. Whisper ASR uses language-coded API parameters with per-language correction dictionaries. TTS generation employs Edge-TTS voice identifiers en-US-GuyNeural for English and hi-IN-MadhurNeural for Hindi, falling back to gTTS for Marathi owing to the absence of a production-quality Edge-TTS Marathi voice at the time of implementation.

V. EVALUATION AND RESULTS

A. Functional Test Cases

Five representative consultation scenarios were designed to validate specialist routing and urgency classification across the full urgency spectrum. Each test case was executed five times with paraphrased symptom descriptions to assess LLM response consistency. Table II summarizes the evaluation results. The system achieved consistent urgency classification and specialist routing across all high- and medium-urgency cases. Minor variability was observed in TC-5, where the LLM alternated between Orthopedic and Rheumatology recommendations depending on symptom framing, reflecting genuine clinical ambiguity rather than a system defect. All CRITICAL-level cases (TC-4) were correctly classified across all five repeated trials without exception.

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1718336

TABLE II FUNCTIONAL EVALUATION: SYMPTOM INPUT, URGENCY CLASSIFICATION, AND SPECIALIST ROUTING

TC	Symptom Input	Exp. Urgency	Obs. Urgency	Exp. Specialist	Obs. Specialist	Consist.
TC-1	Severe chest pain, shortness of breath, sweating; age 52	HIGH	HIGH	Cardiologist	Cardiologist	5/5
TC-2	Persistent fever 3 days, productive cough, body aches	MEDIUM	MEDIUM	Gen. Physician	Gen. Physician	5/5
TC-3	Minor skin rash on left forearm; no pain; onset yesterday	LOW	LOW	Dermatologist	Dermatologist	5/5
TC-4	Sudden severe headache, facial drooping, arm weakness	HIGH (CRIT.)	HIGH (CRIT.)	Neurologist	Neurologist	5/5
TC-5	Bilateral knee pain, morning stiffness for 2 months	MEDIUM	MEDIUM	Orthopedic/Rheum.	Orthopedic	4/5

B. System Performance

Response latency was measured across 50 consultation requests under local network conditions. Table III summarizes mean latencies for each pipeline stage. Total end-to-end latency averaged 6.2 seconds under free-tier Groq API rate limits (30 requests per minute), with LLM inference accounting for the dominant contribution (3.1 s for multimodal, 1.8 s for text-only). TTS generation and audio file polling contributed an additional 1.5–2.0 seconds. These values are consistent with acceptable user experience thresholds for non-emergency medical consultation applications where asynchronous response delivery is the expected interaction model.

C. Doctor Discovery Coverage

The doctor discovery module was tested across five major Indian cities, querying five distinct specialties per city (Table IV). The curated database contributed results for all metropolitan centres, while LocationIQ augmented results with community-level facilities. Delhi returned the highest average result count per specialty query (22.7 results), reflecting higher density of documented medical facilities in the national capital region. Chennai returned the lowest count (11.8), attributable to relative sparseness of OpenStreetMap medical POI data in southern metropolitan areas at the time of evaluation. All cities returned sufficient results (greater than 10 per specialty on average) to support meaningful specialist selection.

D. Multilingual ASR Accuracy

ASR accuracy was assessed using a reference set of 30 symptom utterances per language (90 total) recorded by native speakers. Transcription outputs were evaluated against reference transcripts using Character Error Rate (CER). English achieved a CER of 3.1%, Hindi 5.8%, and Marathi 9.4%. The higher Marathi CER reflects the relative scarcity of Marathi medical data in Whisper's training corpus; the post-processing correction dictionary reduced Marathi CER by approximately 2.3 percentage points from the raw transcription baseline. TTS intelligibility was rated qualitatively high for English and Hindi; Marathi gTTS output was rated acceptable but less natural than Edge-TTS voices, motivating future migration to a Marathi-capable neural TTS engine.

VI. DISCUSSION

A. Limitations

Several limitations constrain the current system. First, the absence of persistent patient data storage precludes consultation history retrieval for longitudinal monitoring. Second, LLM hallucination remains an intrinsic risk; the system may generate plausible but clinically incorrect information, particularly for rare conditions or atypical symptom presentations. The disclaimer modal and explicit system prompt language mitigate but cannot eliminate this risk. Third, free-tier API rate limits (Groq: 30 requests per minute; LocationIQ: 5,000

requests per day) restrict production scalability; high-traffic deployment requires paid-tier subscriptions and load-balancing infrastructure.

Fourth, the doctor discovery module relies on crowd-sourced OpenStreetMap data, exhibiting variable completeness across tier-2 and tier-3 Indian cities. Facilities may carry outdated contact information or missing entries. Fifth, the Marathi TTS fallback to gTTS produces lower naturalness compared to Edge-TTS voices, potentially reducing accessibility comfort for Marathi-speaking patients with lower digital literacy.

B. Ethical Considerations

AI-assisted medical consultation carries significant ethical responsibilities. HealStation AI is designed exclusively as an informational triage and referral tool, not as a diagnostic or prescriptive system. Every consultation response includes an explicit disclaimer urging in-person professional evaluation. Patient data (audio, images, documents) is not persisted beyond the active session; audio files are automatically purged after one hour. The system collects no personally identifiable information beyond optional age and gender fields supplied by the user.

Emergency triage outputs emphasize immediate contact with emergency services (108 in India) for CRITICAL and EMERGENCY classifications, prioritizing life safety over system engagement. Algorithmic fairness is a relevant concern, as the underlying LLM may reflect demographic biases from its pre-training corpus. Systematic bias evaluation across patient age groups, genders, and regional symptom presentations is recommended prior to any deployment in public health infrastructure.

VII. CONCLUSION AND FUTURE WORK

This paper presented HealStation AI, an open-source multimodal medical consultation system integrating large language models, speech processing, document vision analysis, emergency triage, and location-based healthcare discovery into a unified, language-aware platform. The system demonstrated consistent specialist routing and urgency classification across diverse symptom profiles, multilingual ASR performance approaching reference benchmarks, and broad doctor discovery coverage across major Indian metropolitan areas.

HealStation AI makes three primary contributions: (1) a demonstration that state-of-the-art LLM capabilities can be assembled into a coherent patient-facing application using exclusively free-tier APIs, substantially lowering the barrier for health-tech innovation in resource-constrained settings; (2) a validated multilingual voice interaction pipeline for Indian languages with documented ASR accuracy and post-processing correction strategies; and (3) an open reference architecture for LLM-powered emergency triage combining validated clinical scoring instruments with neural language understanding.

Several directions remain for future development. Integration with the Ayushman Bharat Digital Mission (ABDM) ecosystem would enable verified health record linkage. Extension to additional Indian languages—Tamil, Telugu, Bengali, Kannada—would expand accessibility to approximately 400 million additional native speakers. The addition of appointment scheduling, real-time telemedicine video consultation, and mobile applications via React Native would complete the clinical workflow from initial triage to care delivery. Systematic clinical validation studies comparing AI triage recommendations against attending physician decisions are necessary prior to deployment in public health infrastructure.

ACKNOWLEDGEMENTS

The author thanks the open-source communities behind Groq, Meta AI (Llama), OpenAI (Whisper), LocationIQ, FastAPI, and React for making their tools freely accessible. Special appreciation is extended to the contributors to OpenStreetMap whose crowd-sourced geographic data powers the healthcare facility discovery subsystem.

REFERENCES

[1] World Health Organization, “Health workforce: Doctor density per 10,000 population,” Global Health Observatory Data Repository, 2023.

[Online]. Available: https://www.who.int/data/gho

[2] K. Singhal, S. Azizi, T. Tu, et al., "Large language models encode clinical knowledge," Nature, vol. 620, pp. 172–180, 2023. doi:10.1038/s41586-023-06291-2

[3] K. Singhal, T. Tu, J. Gottweis, et al., "Towards expert-level medical question answering with large language models," arXiv:2305.09617, 2023.

[4] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in Proc. ICML, 2023.

[5] OpenAI, "GPT-4 Technical Report," arXiv:2303.08774, 2023.

[6] C. Li, C. Wong, S. Zhang, et al., "LLaVA-Med: Training a large language-and-vision assistant for biomedicine in one day," in Proc. NeurIPS Datasets and Benchmarks Track, 2023.

[7] K. Bannur, S. Hyland, Q. Liu, et al., "Learning to exploit temporal structure for biomedical vision-language processing," in Proc. CVPR, 2023.

[8] M. Guagliardo, "Theoretical grounds and measurement methods of geographic accessibility to health services," SESPAS Gac Sanit, vol. 19, no. 4, pp. 1–18, 2004.

[9] Meta AI, "Llama 4: The next generation of open large language models," Meta Blog, 2025. [Online]. Available: https://ai.meta.com/llama

[10] Groq Inc., "GroqCloud: LPU inference engine," Technical Documentation, 2024. [Online]. Available: https://console.groq.com/docs

[11] J. Lee, W. Yoon, S. Kim, et al., "BioBERT: A pre-trained biomedical language representation model for biomedical text mining," Bioinformatics, vol. 36, no. 4, pp. 1234–1240, 2020.

[12] LocationIQ, "Places API: Medical facilities search," API Documentation, 2024. [Online]. Available: https://locationiq.com/docs

[13] OpenStreetMap Foundation, "OpenStreetMap Wiki: Medical facilities tagging schema," 2024. [Online]. Available: https://wiki.openstreetmap.org/wiki/Healthcare

[14] V. Shah, V. Bhatt, S. Singh, and R. Kumar, "AI-based health chatbots and their role in the Indian healthcare ecosystem: A systematic review," Health Informatics J., vol. 28, no. 3, 2022. doi:10.1177/14604582221112321

[15] T. B. Brown, B. Mann, N. Ryder, et al., "Language models are few-shot learners," in Proc. NeurIPS, vol. 33, pp. 1877–1901, 2020.

[16] S. Sebastian and B. Balachandran, "Whisper-Hindi: Domain adaptation of automatic speech recognition for Hindi medical transcription," arXiv:2311.10240, 2023.

[17] G. Kaissis, M. Makowski, D. Rückert, and R. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," Nat. Mach. Intell., vol. 2, pp. 305–311, 2020.

[18] Ministry of Health and Family Welfare, India, "Ayushman Bharat Digital Mission: Technical specifications," NHA Report, 2023. [Online]. Available: https://abdm.gov.in

[19] S. Tipirneni and C. K. Reddy, "Self-supervised transformer for sparse and irregularly sampled multivariate clinical time-series," ACM Trans. Knowl. Discov. Data, vol. 16, no. 4, pp. 1–17, 2022.

[20] P. Rajpurkar, J. Irvin, K. Ball, et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv:1711.05225, 2017.