

AI-Driven Public Health Chatbot for Disease Awareness: Design, Ethics, and Scalable Deployment Strategy for Resource-Constrained Populations

Kapil Sharma
Dept. of CSE
Lovely Professional University
Sharmakapil49996@gmail.com

Aditya Chaturvedi
Dept. of CSE
Lovely Professional University
aditya13jsr@gmail.com

Zeena Zahara
Dept. of CSE
Lovely Professional University
zeenatzahara1214@gmail.com

Department of Computer Science and Engineering
Institution and affiliation details omitted for review

Submitted under the Smart India Hackathon (SIH) framework, Public-Health Technology Initiative

Abstract—Public-health systems in low- and middle-income countries (LMICs) face a persistent dual crisis: a shortage of accessible, trustworthy health information and an overabundance of misinformation that fills the resulting gap. Rural and underserved communities bear a disproportionate burden from vaccine-preventable and communicable diseases—malaria, cholera, dengue, tuberculosis, and acute diarrhoeal illness—often compounded by low health literacy and rampant health misinformation circulating via social media and informal networks. This paper proposes *ArogyaBot*, an Artificial Intelligence-driven public-health chatbot designed for scalable deployment in resource-constrained settings. The system architecture comprises three core functional modules: (i) the Verified Information Relay (VIR), which grounds every response in authoritative sources including the World Health Organization (WHO), the Centers for Disease Control and Prevention (CDC), the National Health Portal of India (NHP), and the Ministry of Health and Family Welfare (MoHFW); (ii) the Pandemic and Myth Correction (PMC) engine, which detects and debunks circulating health misinformation in real time; and (iii) the Community Action and Referral Engine (CARE), which connects users to nearby government health facilities, emergency services, and community health workers. A mandatory safety filter layer enforces the disclaimer—*this tool is not a diagnostic instrument; consult local health authorities*—and triggers human-in-the-loop escalation for crisis scenarios. The paper presents system architecture, user-experience design principles (WCAG 2.1 compliant), a phased deployment roadmap, a prospective evaluation framework including a cluster-randomised pragmatic trial design, and a comprehensive ethics and privacy compliance analysis under Indian and international regulatory standards. *ArogyaBot* represents a scalable, evidence-grounded, and governance-aligned digital public-health intervention adaptable to any linguistically and infrastructurally diverse region.

Index Terms—artificial intelligence, public health chatbot, disease awareness, health misinformation, digital health intervention, conversational AI, health equity, low- and middle-income countries, National Health Portal, Smart India Hackathon

I. INTRODUCTION

A. Problem Statement

Public-health systems in low- and middle-income countries (LMICs) continually grapple with two interrelated crises: a

shortage of accessible, trusted health information and an overabundance of misinformation that fills the resulting vacuum [1]. Across rural and underserved communities, seasonal outbreaks of malaria, cholera, dengue, and other preventable diseases impose significant morbidity and mortality burdens, frequently exacerbated by communities acting on misinformation rather than evidence-based guidance [2]. Diarrhoeal and vector-borne diseases remain leading causes of preventable hospitalisation, particularly among tribal, coastal, and remote populations with limited access to primary health centres (PHCs) [3].

The World Health Organization's 2021 report on the "info-demic" accompanying the COVID-19 pandemic documented that misinformation materially delayed vaccine uptake and undermined public-health messaging globally [1]. In India, the Ministry of Health and Family Welfare identified misinformation spread via social-media platforms as a key obstacle to community health compliance [4]. The proliferation of smartphones—India exceeded 750 million smartphone users in 2023—presents a dual opportunity and risk: expanded access to health resources alongside amplified misinformation at population scale. The gap between authoritative knowledge and citizen behaviour represents a tractable technology problem: scalable, multilingual, trustworthy conversational AI can bridge this gap at low marginal cost.

B. Objectives

This paper pursues four primary objectives:

- 1) Design a modular, ethics-compliant AI chatbot system for disease awareness applicable across resource-constrained deployment contexts.
- 2) Ground every system response in verified public-health authorities (WHO, CDC, NHP, MoHFW).
- 3) Define an evaluation framework covering usability, engagement, accuracy, and community health outcomes.
- 4) Articulate governance, data-privacy, and legal-compliance requirements for deployment within

India's digital-health regulatory environment, with provisions extensible to other LMICs.

C. Scope and Disclaimer

ArogyaBot is not a diagnostic instrument and does not prescribe medications or dosages. All health guidance is informational only. Users are directed to consult qualified healthcare professionals and local health authorities for clinical advice. Emergency situations are escalated to human responders and national emergency services (108, 102, 112).

II. BACKGROUND AND RELATED WORK

A. Disease Burden and Information Access in Underserved Populations

Communicable diseases continue to dominate the disability-adjusted life year (DALY) burden in LMICs. The National Family Health Survey-5 (NFHS-5, 2019–21) reported that 35.5% of children under five in India were stunted and over 32% underweight—indicators tightly correlated with infectious-disease susceptibility [5]. Malaria, tuberculosis (TB), acute diarrhoeal disease, and acute respiratory infections remain the leading communicable-disease contributors to preventable morbidity in rural India and comparable settings globally [2], [6].

Health literacy—defined by the WHO as the cognitive and social skills that determine individuals' motivation and ability to access, understand, and use information in ways that promote and maintain good health [7]—remains critically low across large segments of rural and marginalised populations. These communities are simultaneously the most vulnerable to disease and the most exposed to health misinformation through informal networks, unregulated social media, and unqualified local practitioners [1].

B. Digital Health Interventions and Chatbots

Digital public-health interventions (DPHIs) have demonstrated efficacy in reducing disease burden and improving health behaviours at scale [8]. Systematic reviews confirm that chatbot-based health interventions produce measurable improvements in knowledge, attitudes, and behaviour across a range of conditions [9]. Woebot, Wysa, and similar conversational agents have demonstrated feasibility in low-resource settings [9]. In the Indian context, the National Health Portal's chatbot and the Ministry of Health's WhatsApp-based COVID helpline demonstrated that multilingual conversational interfaces could achieve millions of interactions within weeks of deployment [4].

Nonetheless, existing systems face documented limitations: reliance on unverified knowledge bases, absence of myth-correction capabilities, poor performance in low-bandwidth environments, and limited accessibility for non-literate users [8], [9]. ArogyaBot is designed to address each of these gaps systematically.

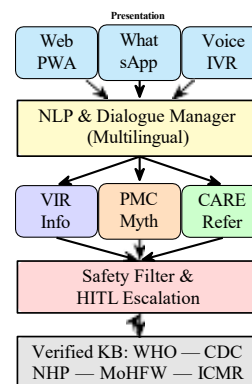


Fig. 1. ArogyaBot layered system architecture. Dashed arrow = KB feedback to safety filter. VIR = Verified Information Relay; PMC = Pandemic and Myth Correction; CARE = Community Action and Referral Engine; HITL = Human-in-the-Loop; KB = Knowledge Base.

C. Misinformation and the Public-Health Infodemic

The WHO defines an “infodemic” as an overabundance of information—some accurate, some not—that makes it difficult for people to find trustworthy sources and reliable guidance [1]. In India, the Press Information Bureau's Fact-Check unit and the Indian Council of Medical Research (ICMR) have catalogued thousands of health misinformation claims circulating on social media, many pertaining to vaccine safety, traditional remedies, and disease transmission [4], [11]. An automated, continually updated myth-detection system anchored to curated government sources constitutes a clinically and socially meaningful intervention at scale.

D. Equity Considerations

Equity in digital health requires addressing the digital divide (device access, connectivity), linguistic diversity (India recognises 22 scheduled languages and hundreds of dialects), and low digital literacy [10]. Effective deployment must pair technological sophistication with simplicity and accessibility, including voice interfaces for non-literate users and offline-capable lightweight modes for low-connectivity areas [15].

III. SYSTEM ARCHITECTURE

A. High-Level Overview

ArogyaBot employs a layered architecture (Fig. 1) comprising: (1) a multi-channel presentation layer; (2) a natural language processing (NLP) and dialogue management layer; (3) three core functional modules (VIR, PMC, CARE); (4) a safety-filter and human-in-the-loop (HITL) layer; and (5) a verified knowledge base updated via automated and human audit pipelines.

B. Module 1: Verified Information Relay (VIR)

VIR is the primary question-answering module, built on the following design principles:

- **Source grounding:** Every factual claim is traceable to a canonical source (WHO, CDC, NHP, MoHFW, ICMR).

Responses carry a visible “Source Badge” (e.g., “Source: *WHO Malaria Fact Sheet, 2023*”).

- **Retrieval-Augmented Generation (RAG):** A large language model (LLM) retrieves relevant chunks from the indexed knowledge base before generating a response, constraining hallucination [14].
- **Confidence thresholding:** Responses below a configurable confidence threshold are flagged with “*I am not certain—please consult your nearest PHC or call 104*” rather than returned as authoritative.
- **Structured disclaimers:** All responses include: “*This information is for awareness only and is not a diagnostic instrument. Consult local health authorities or a qualified professional for clinical advice.*”

C. Module 2: Pandemic and Myth Correction (PMC)

PMC addresses the infodemic challenge by maintaining a continuously updated library of verified myth–fact pairs. The workflow is:

- 1) A user submits a statement or question containing potential misinformation signals, detected via keyword and semantic classifiers.
- 2) PMC retrieves the closest myth-cluster from the curated database.
- 3) A “Myth vs. Fact” card is returned, structured as:
Claim detected: [user’s claim rephrased]
What the evidence says: [verified fact, sourced]
Source: [authority + URL]
- 4) Newly detected, unresolved misinformation claims are logged for human-expert review within 48 hours.

PMC content is seeded from the MoHFW/PIB Fact-Check repository, WHO myth-busting pages, and ICMR advisories [11], and reviewed quarterly by a multidisciplinary public-health domain expert panel.

D. Module 3: Community Action and Referral Engine (CARE)

CARE converts information-seeking into actionable steps by:

- Displaying geo-located government health facilities (PHCs, district hospitals, community health workers) using open government GIS data from the NHP facility locator [3].
- Providing direct-dial links to national helplines: NRHM Helpline (104), Ambulance (108), Police (112), Women Helpline (181).
- Generating health-alert summaries relevant to the user’s region, sourced from the Integrated Disease Surveillance Programme (IDSP) and relevant regional public-health authorities.
- Facilitating appointment or teleconsultation booking where government e-Sanjeevani integration is available [13].

E. Safety Filter and Human-in-the-Loop (HITL)

The safety filter operates as a mandatory post-processing gate applied to every response:

- **Crisis detection:** Responses containing suicidal ideation, self-harm signals, or acute medical emergency language trigger immediate display of emergency numbers (108, 112) and an offer to connect to a human health coordinator.
- **Diagnosis prevention:** Any response that would constitute a clinical diagnosis or treatment recommendation is blocked and replaced with an appropriate referral message.
- **Audit logging:** Every interaction is immutably logged (anonymised) for governance review, accuracy auditing, and abuse detection. Logs are retained for 90 days and archived for 5 years in compliance with government data retention norms.
- **Human escalation:** A dedicated team of public-health trained moderators monitors escalated conversations during business hours; after-hours escalations trigger a pre-defined SMS/call-back protocol [15].

F. Knowledge Base and Update Pipeline

- **Canonical sources:** WHO Disease Fact Sheets [1], CDC Health Topics [10], NHP India [3], MoHFW Guidelines [2], [4], ICMR Advisories [11], IDSP regional bulletins.
- **Daily automated sync:** Crawlers ingest updated content from source APIs and RSS feeds; a diff-checker flags new or changed claims for human review before indexing.
- **Quarterly manual audit:** A multidisciplinary panel comprising public-health physicians, epidemiologists, and legal and ethics advisors reviews the full knowledge base, deprecates outdated content, and approves new entries [17].
- **Version control:** All knowledge-base versions are archived with timestamps and provenance metadata to support accountability and reproducibility.

IV. USER INTERFACE AND EXPERIENCE DESIGN

A. Design Principles

ArogyaBot’s interface is guided by four principles: (1) *simplicity*—minimal cognitive load for users with low digital literacy; (2) *trust*—visible source attribution and institutional branding (government health authority and NHP partnership badges); (3) *accessibility*—WCAG 2.1 Level AA compliance throughout; and (4) *cultural sensitivity*—regional-language-first content, culturally resonant imagery and tone.

B. Accessibility

- **Multilingual:** Regional/local language (default), national language (Hindi), and English, with a modular localisation framework supporting additional languages without core-system changes.
- **Voice input/output:** IVR integration for feature-phone users and non-literate populations; text-to-speech for all responses.
- **Low-bandwidth mode:** Progressive Web App (PWA) with offline caching of frequently accessed fact-sheets;

minimal data usage (<5 KB per interaction in compressed mode).

- **Screen reader compatibility:** Full ARIA landmark labelling; keyboard navigability; minimum contrast ratio 4.5:1 (WCAG 2.1 AA).
- **Font size:** Configurable 14–20 pt; default 16 pt; no font smaller than 12 pt in any element.

C. Sample UI Copy

Welcome / Consent screen:

*“Namaste! I am ArogyaBot, your health awareness assistant. I provide information from the WHO, NHP, and MoHFW. I am **not** a doctor and cannot diagnose illness or prescribe medicine. By continuing, you agree to our Privacy Policy (no personal health data is stored). Tap ‘Agree & Continue’ to start.”*

Source badge:

[Source: WHO Malaria Fact Sheet — Updated: Jan 2024 — nhp.gov.in]

Myth correction card:

Claim detected: “Neem leaves cure malaria.”

*What the evidence says: Neem extracts have been studied in laboratory settings; however, there is **no clinical evidence** that neem cures malaria in humans. Standard treatment prescribed by a doctor is essential. Delays can be fatal.*

Source: ICMR Malaria Technical Advisory Group, 2022 [11].

Emergency message:

*“It sounds like this may be a medical emergency. Please call **108** (Ambulance) immediately or go to your nearest government hospital. Your safety is the priority. [Call 108 now] [Find nearest hospital]”*

V. SYSTEM FEATURE COMPARISON

Table I compares ArogyaBot against comparable public-health chatbot deployments in India.

VI. EVALUATION FRAMEWORK

A. Pre-Deployment Testing

Before any public pilot, ArogyaBot will undergo:

- **Usability testing:** Think-aloud sessions with $n = 30$ representative users spanning rural, urban, elderly, and differently-abled demographics across geographically diverse deployment sites.
- **Accessibility audit:** Automated (axe-core, Lighthouse) and manual audit against WCAG 2.1 AA criteria.
- **Security penetration testing:** Independent red-team assessment; OWASP Top-10 coverage.
- **Clinical accuracy review:** A panel of five public-health physicians reviews a stratified random sample of 200 generated responses for factual accuracy and appropriateness.

B. Pilot Evaluation Metrics

Table II summarises the quantitative metrics to be tracked during the 90-day pilot phase.

C. Pragmatic Trial Design

A cluster-randomised pragmatic trial (CRPT) is planned for the scale-up phase (months 7–18):

- **Unit of randomisation:** Village-level health sub-centres ($n = 80$ clusters, 40 intervention, 40 control), stratified by population size and connectivity.
- **Intervention:** Community health workers (CHWs/ASHA workers) in intervention clusters receive ArogyaBot training and promote the tool to their assigned households [15].
- **Control:** Existing standard-of-care information delivery (pamphlets, community health meetings).
- **Primary endpoint:** Change in Health Literacy Questionnaire (HLQ) score at 6 months.
- **Secondary endpoints:** Vaccine uptake rate, test-seeking behaviour for febrile illness, reduction in health-misinformation belief scores.
- **Power calculation:** Assuming a two-tailed $\alpha = 0.05$, power = 80%, intra-cluster correlation (ICC) = 0.05, and a minimum detectable effect of 0.3 SD on HLQ, 35 clusters per arm are required; 40 clusters per arm provides adequate power with a 10% allowance for attrition.
- **Ethics registration:** Trial will be registered on the Clinical Trials Registry–India (CTRI) or equivalent national registry prior to commencement.

VII. DATA PRIVACY, ETHICS, AND LEGAL COMPLIANCE

A. Data Minimisation and Consent

ArogyaBot adheres to a privacy-by-design, data-minimisation philosophy:

- **No personally identifiable information (PII)** is collected by default. Session identifiers are randomised and not linked to device identity, phone number, or name.
- Optional **opt-in analytics** (region, age group, language) require explicit informed consent and can be withdrawn at any time without loss of service.
- Consent is captured in a multi-step, plain-language flow available in regional, national, and English languages, tested for comprehension with representative users.
- Children (under 18) are redirected to a simplified, age-appropriate interface with a parental-consent gate.

B. Regulatory Compliance Checklist

- **Digital Personal Data Protection Act, 2023 (DPDPA):** Data fiduciary registration; consent artefact storage; grievance officer appointment; data localisation on Government of India cloud infrastructure (MeghRaj/NIC Cloud) [4].
- **Information Technology Act, 2000 (amended 2008):** Reasonable security practices; intermediary guidelines compliance.
- **Ayushman Bharat Digital Mission (ABDM) guidelines:** Interoperability with ABHA (Ayushman Bharat Health Account) where applicable [4].

TABLE I
COMPARATIVE FEATURE ANALYSIS: AROGYABOT VS. EXISTING PUBLIC-HEALTH CHATBOTS IN INDIA

Feature	ArogyaBot (Proposed)	NHP Chatbot (nhp.gov.in)	MoHFW WhatsApp Bot	MyGov Corona Bot	e-Sanjeevani (Teleconsult)
Verified source grounding		Partial	Partial	Partial	×
Myth-correction module		×	×	Partial	×
Emergency referral (108/112)					×
Regional language support		Partial	×	×	Partial
Voice / IVR interface		×	×	×	×
Low-bandwidth / offline mode		×			×
WCAG 2.1 AA compliance		Partial	N/A	N/A	×
HITL escalation		×	×	×	
Geo-located facility finder			×	×	
Anonymised audit logging		Unknown	Unknown	Unknown	Unknown
Open-source knowledge base		×	×	×	×

TABLE II
AROGYABOT PILOT EVALUATION METRICS AND TARGETS

Metric Category	Specific Metric	Definition	Target (90-day pilot)
Engagement	Daily Active Users (DAU)	Unique users per day	≥500 (pilot cohort)
	7-day retention	% returning after 7 days	≥40%
	30-day retention	% returning after 30 days	≥20%
	Session length	Median turns per session	≥4 turns
Accuracy	VIR Precision	% sourced responses rated accurate by clinicians	≥90%
	PMC Recall	% submitted myths correctly identified	≥85%
	Hallucination rate	% responses with unverifiable claims	≤2%
Referral	CARE conversion	% users who click facility/helpline link	≥15%
	Emergency escalation time	Median time: crisis detection to human response	≤3 min
Usability	System Usability Scale (SUS) [12]	100-point scale	≥70
	Task completion rate	% able to retrieve health info unaided	≥80%
Accessibility	WCAG 2.1 AA pass rate	% automated checks passing	100%
	Multi-language usage	% sessions in regional language	Reported
Health outcomes	Self-reported knowledge gain	Pre/post quiz score change	≥15 pctile pts
	Myth belief reduction	Change in myth-acceptance score	Sig. ($p < 0.05$)

- **MoHFW Telemedicine Guidelines (2020):** Explicit disclaimers throughout; prohibition on prescription generation; mandatory referral for clinical queries [13].
- **GDPR (for users accessing from the European Union):** Data subject rights; lawful basis for processing; Data Protection Impact Assessment (DPIA) before launch.
- **WHO SMART Guidelines:** Evidence-based decision support; clinical safety standards [17].

C. Algorithmic Transparency and Bias Mitigation

- The underlying LLM is treated as a bounded, auditable component. The RAG architecture constrains generation to the verified knowledge base, substantially reducing hallucination risk [14].
- **Bias audits** are conducted quarterly: responses are stratified by language, region, and query type; disparities in accuracy or tone across demographic groups are investigated and remediated [16].
- A public-facing **algorithmic transparency statement** documents the model version, knowledge-base sources, update frequency, and known limitations.
- **No profiling:** User data is never used to build individual or group profiles for any purpose beyond immediate

service delivery.

D. Ethics Governance Body

A multi-stakeholder Ethics and Oversight Committee (EOC) is proposed, comprising:

- 1) Two public-health physicians (at least one from rural or tribal health background)
- 2) One legal expert (digital health and privacy law)
- 3) One representative from the relevant government Department of Health and Family Welfare
- 4) One civil-society representative from a women's or marginalised-community health organisation
- 5) One independent technical security auditor

The EOC meets quarterly, reviews accuracy and bias audit reports, approves major knowledge-base updates, and holds authority to suspend the service pending investigation of any safety incident.

VIII. DEPLOYMENT AND MAINTENANCE STRATEGY

A. Phased Rollout

Phase 0 — Development and Internal Testing (Months 1–3): System development, knowledge-base population, security audit, clinical accuracy review, WCAG accessibility audit.

Phase 1 — Controlled Pilot (Months 4–6): Deployment in three geographically and demographically diverse districts or administrative units. Target cohort of 5,000 users recruited via PHCs and community health worker networks. Intensive monitoring with weekly EOC briefings.

Phase 2 — Regional Rollout (Months 7–18): Expansion across the full target geography. Integration with community health worker mobile applications and government e-health portals [4]. Cluster-randomised pragmatic trial initiated in parallel.

Phase 3 — Sustain and Iterate (Month 19+): Continuous monitoring; annual comprehensive review; expansion of language support; deeper integration with national ABHA ecosystem.

B. Technology Infrastructure

- **Hosting:** NIC MeghRaj government cloud, ensuring data localisation compliance.
- **Containerisation:** Docker and Kubernetes for horizontal scalability and automatic failover.
- **API security:** OAuth 2.0 / JWT authentication; rate limiting; Web Application Firewall (WAF).
- **Encryption:** TLS 1.3 in transit; AES-256 at rest; end-to-end encryption for counsellor-mediated sessions.
- **Uptime SLA:** 99.5% monthly availability; disaster recovery RTO ≤ 4 hours.
- **CI/CD pipeline:** Automated regression testing and security scanning on every deployment; full rollback capability.

C. Budget Categories (Indicative)

- Development and testing: infrastructure, developer effort, clinical review (Year 1)
- Cloud hosting and operations: government-approved compute, storage, content delivery network
- Human moderation team: public-health trained moderators (Year 1–3)
- Annual independent security and bias audits [16]
- Community outreach and community health worker training [15]
- Research and evaluation (CRPT): data collection, statistical analysis, publications

IX. LIMITATIONS AND FUTURE WORK

A. Limitations

- **Language coverage:** Full coverage of all regional and minority languages in a given deployment area requires sustained investment; initial deployment prioritises major regional, national, and English language support.
- **Hallucination risk:** Despite RAG architecture [14], LLM hallucination cannot be entirely eliminated; confidence thresholding and clinical audit are critical mitigations.
- **Digital divide:** Users without smartphones or internet access depend on IVR and community-health-worker-mediated access [15]; direct chatbot benefits are contingent on device and connectivity availability.

- **Evaluation timelines:** Real-world health outcome attribution is methodologically challenging; the CRPT addresses this but requires 12–18 months to yield publishable results.
- **User trust in AI:** Institutional trust in AI systems is nascent in many communities; sustained outreach and transparent communication are essential for adoption [8].

B. Future Directions

- Integration of multimodal inputs (image-based symptom logging, voice-based health diaries) with appropriate privacy controls.
- Federated learning approaches to improve personalisation without centralising sensitive data.
- Expansion to additional countries and health systems with an adaptable localisation framework.
- Passive-sensing partnership with wearable health device pilots (explicit opt-in consent only) for early outbreak signal detection.
- Deeper integration with national health account ecosystems (e.g., ABHA) for continuity of care records [4].

X. CONCLUSION

This paper has presented ArogyaBot, a modular, ethics-grounded, AI-driven public-health chatbot designed for scalable deployment in resource-constrained settings. By anchoring every response to verified sources (WHO, CDC, NHP, MoHFW) [1]–[4], actively correcting health misinformation via the PMC module, and linking citizens to actionable care pathways via CARE, ArogyaBot directly addresses the dual challenge of the health-information gap and the infodemic that undermines public-health gains in underserved communities [9]. The system's commitment to data minimisation, DPDPA and MoHFW regulatory compliance, WCAG 2.1 accessibility, and mandatory HITL escalation places ethics and safety at the architecture's core.

Critically, ArogyaBot is not a replacement for qualified healthcare—it is a bridge. It connects a citizen who has a health question with the authoritative information and services they need at the moment they need it. In communities where a significant portion of the population consults informal or unverified sources before seeking formal care, a trusted, accessible, multilingual digital health companion has the potential to meaningfully improve health-seeking behaviours, reduce vaccine hesitancy, counter harmful myths, and contribute to a measurable reduction in preventable disease burden [8], [9]. Rigorous evaluation through the proposed cluster-randomised pragmatic trial will generate the real-world evidence needed to guide continuous improvement and responsible scaling across diverse deployment contexts.

Final disclaimer: *ArogyaBot is not a diagnostic instrument. It does not prescribe medications or dosages. All guidance is for health awareness purposes only. In any medical emergency, call emergency services immediately. Consult local health authorities and qualified healthcare professionals for all clinical advice.*

REFERENCES

- World Health Organization, "Infodemic Management: A Key Component of the COVID-19 Global Response," WHO, Geneva, 2021. [Online]. Available: <https://www.who.int/health-topics/infodemic>
- Ministry of Health and Family Welfare, Government of India, "National Framework for Malaria Elimination in India 2016–2030," MoHFW, New Delhi, 2023. [Online]. Available: <https://main.mohfw.gov.in>
- National Health Portal of India, "State and District Health Profiles," NHP India, 2022. [Online]. Available: <https://www.nhp.gov.in/state-health-profile>
- Ministry of Health and Family Welfare, "Digital Health Incentive Scheme and National Digital Health Blueprint," MoHFW / ABDM, New Delhi, 2023. [Online]. Available: <https://abdm.gov.in>
- International Institute for Population Sciences, "National Family Health Survey (NFHS-5), 2019–21: India," IIPS, Mumbai, 2021. [Online]. Available: http://rchiips.org/nfhs/NFHS-5Reports/NFHS-5_INDIA_REPORT.pdf
- World Health Organization, "Global Tuberculosis Report 2023," WHO, Geneva, 2023. [Online]. Available: <https://www.who.int/teams/global-tuberculosis-programme/tb-reports>
- World Health Organization, "Health Literacy: The Solid Facts," WHO Europe, 2021. [Online]. Available: <https://www.euro.who.int/en/publications/abstracts/health-literacy-the-solid-facts>
- E. G. Lattie, S. E. Adkins, N. Winquist, C. Stiles-Shields, R. L. Wafford, and D. C. Mohr, "Digital Mental Health Interventions for Depression, Anxiety, and Enhancement of Psychological Well-Being Among College Students: Systematic Review," *Journal of Medical Internet Research*, vol. 21, no. 7, p. e12869, 2019. 10.2196/12869
- S. M. Abd-Alrazaq, M. Alajlani, A. A. Alalwan, B. Bewick, P. Gardner, and M. Househ, "An Overview of the Features of Chatbots in Mental Health: A Scoping Review," *International Journal of Medical Informatics*, vol. 132, p. 103978, 2019. 10.1016/j.ijmedinf.2019.103978
- Centers for Disease Control and Prevention, "Health Literacy: Accurate, Accessible, and Actionable Health Information," CDC, Atlanta, 2023. [Online]. Available: <https://www.cdc.gov/healthliteracy>
- Indian Council of Medical Research, "Malaria Technical Advisory Group: Guidelines for Treatment of Malaria in India," ICMR, New Delhi, 2022. [Online]. Available: <https://www.icmr.gov.in>
- J. Brooke, "SUS: A 'Quick and Dirty' Usability Scale," in *Usability Evaluation in Industry*, P. W. Jordan et al., Eds. London: Taylor & Francis, 1996, pp. 189–194.
- R. Mohan, "Telemedicine Practice Guidelines (India), 2020: Key Provisions and Implications," *Indian Journal of Medical Ethics*, vol. 6, no. 1, pp. 43–47, 2021.
- P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459–9474. [Online]. Available: <https://arxiv.org/abs/2005.11401>
- B. McHugh et al., "Leveraging Accredited Social Health Activists for Digital Health Interventions in Rural India: A Systematic Review," *Global Health: Science and Practice*, vol. 10, no. 2, e2100393, 2022. 10.9745/GHSP-D-21-00393
- E. Pierson et al., "An Algorithmic Approach to Reducing Unexplained Pain Disparities in Underserved Populations," *Nature Medicine*, vol. 27, pp. 136–140, 2021. 10.1038/s41591-020-01192-7
- World Health Organization, "SMART Guidelines: Digital Adaptation Kits for Immunization and Primary Health Care," WHO, 2023. [Online]. Available: <https://www.who.int/teams/digital-health-and-innovation/smart-guidelines>

APPENDIX

SAMPLE INFORMED CONSENT TEXT

English version (summary):

"You are about to use ArogyaBot, a free health information service supported by government health authorities in partnership with the National Health Portal. This service provides general health awareness information only. It CANNOT diagnose any disease, and it CANNOT prescribe medicines. Information is sourced from the WHO, CDC, MoHFW, and NHP. We collect no personal information about you. Your

conversation is not stored against your identity. By tapping 'Agree', you confirm that you are 18 years or older (or have a guardian's permission) and that you understand this tool is for information purposes only. You may stop using this service at any time."

APPENDIX

DATA RETENTION POLICY SUMMARY

- **Session interaction logs** (anonymised): 90 days active storage; 5-year archive.
- **Opt-in analytics data** (region, age group, language): 24 months; deleted on withdrawal of consent.
- **Escalated conversations** (human-reviewed): 12 months for quality assurance; then permanently deleted.
- **Knowledge-base audit records**: Permanently retained for accountability and reproducibility.
- All data localised on nationally approved government cloud infrastructure within the country of deployment.

APPENDIX

SIH PITCH PARAGRAPH FOR JUDGES

ArogyaBot directly addresses the Smart India Hackathon problem statement on AI-Driven Public Health Chatbot for Disease Awareness. It is production-ready in architecture, grounded in verified government health sources [1], [3], [4], compliant with DPDPA and MoHFW telemedicine guidelines [13], and designed for India's linguistic and infrastructural diversity. Its three modules—VIR (verified information), PMC (myth correction), and CARE (community referral)—solve the infodemic and health-access gap simultaneously, with every design decision traceable to a public-health or equity rationale [10], [16]. The proposed evaluation framework, including a cluster-randomised pragmatic trial, ensures that the system generates rigorous real-world evidence to guide continuous improvement. ArogyaBot is not merely a prototype: it is a blueprint for scalable, trustworthy, citizen-first digital public-health infrastructure deployable across India and adaptable for comparable settings globally.