REQUEST FOR PROPOSAL (RFP)
Stratified Audio Data Collection for Custom ASR Model in Telemedicine Platform

Closing Date for Submission

November 25, 2026, at 23:59 PM

I. INTRODUCTION

LEHS is a charitable organization in India whose purpose is to offer basic health and education for the poor. LEHS in furtherance of charitable objectives through its flagship programs Wadhwani AI which aims to build equitable and sustainable systems by making quality primary healthcare available and accessible to the underserved population and to bring the benefits of modern AI technology to underserved populations by building and deploying AI solutions for social impact across domains such as healthcare, agriculture, governance and education in India. LEHS aims to promote the integration of technologies, particularly in emerging domains like artificial intelligence and innovations into the Indian mainstream primary healthcare, education, and agriculture systems through a partnership with the State and National Government, apex institutions, international agencies, and private sector partners e.g. innovators, social enterprises and other ecosystem contributors in line with its stated objectives for the betterment of society particularly focusing on projects of national and social significance.

Wadhwani AI, a unit of LEHS, focuses on developing, deploying and evaluating artificial intelligence solutions to address critical social challenges in India, particularly in domains such as healthcare, agriculture, and education.
LEHS aims to promote the integration of technologies, particularly in emerging domains like artificial intelligence and innovations into the Indian mainstream primary healthcare, education, and agriculture systems through a partnership with the State and National Government, apex institutions, international agencies, and private sector partners e.g. innovators, social enterprises and other ecosystem contributors in line with its stated objectives for the betterment of society particularly focusing on projects of national and social significance. In line with its mission to support projects of national and social significance, LEHS also undertakes rigorous **monitoring, evaluation, and learning (MEL)** activities to assess the impact, usability, and scalability of different programmatic interventions.

II. BACKGROUND

Clinical documentation remains one of the most time-consuming tasks for frontline health workers. In India, Community Health Workers (CHOs)—most of whom are trained as nursing professionals or AYUSH practitioners—play a pivotal role in managing patient consultations at primary care levels. They are often responsible for collecting patient symptoms, summarizing case histories, and completing digital records. However, current systems rely primarily on manual typing or form-filling, which can be cumbersome, error-prone, and inconsistent.

To overcome these challenges, LEHS-AI is building an AI-enabled voice interface that allows CHOs to narrate symptoms and case summaries naturally, while the ASR engine accurately transcribes these into structured digital text. Off-the-shelf ASR engines, often trained on

Western datasets, perform poorly in Indian clinical settings due to varied accents, regional dialects, and context-specific medical terminology.

Therefore, LEHS-AI is investing in the creation of a domain-specific and India-centric audio corpus by collecting voice samples directly from CHOs as well as Medical officers (MO) working in the public health ecosystem. This corpus will represent natural clinical speech and cover the range of accent, pronunciation, and content variation required to train a high-quality ASR engine tailored for frontline clinical settings.

III. Project Objectives:

The key objective of this assignment is to collect high-quality voice recordings from CHOs and MOs, evenly split between English and Hindi. These recordings will form the training and testing datasets for a machine learning-based ASR model. The recordings must reflect real-world speech as observed during patient consultations in public health settings.

To ensure diversity, the dataset must be collected from CHOs and MOs with varied demographic and professional profiles, across urban, rural, and tribal locations. The final dataset must adequately reflect diversity in gender, age, region, qualification, dialect, and fluency level.

IV. SCOPE OF WORK

The selected agency is expected to:

- Establish partnerships with state health departments and relevant authorities.
- Identify appropriate health facilities for data collection (e.g., Health Sub-Centres, PHCs).
- Secure administrative approvals for audio recording activities.
- Ensure high-quality, timely, and complete data submissions with associated metadata.

All recordings must be conducted in real-life work settings (e.g., Health Sub-Centres, PHCs, CHC OPDs). The focus of this data collection activity is to build a foundational dataset for the development of ASR model, which includes both **Paragraph-Level Audio Data Collection** (using pre-designed scripts) and **Word-Level Audio Data Collection** (using clinical terms).

1. Paragraph-Level Audio Data Collection

- **Target Duration:** 100 hours per language (English and Hindi)
- **Geographical Coverage:**
  - English: Meghalaya, Mizoram, Manipur, Uttarakhand, Assam
  - Hindi: Uttar Pradesh, Madhya Pradesh, Assam, Arunachal Pradesh, Bihar
- **Distribution:** 20 hours per state
- **Content:** Scripted, paragraph-style summaries of chief patient complaints, read aloud by CHOs and MOs
- **Recording Format:** Paragraph-level narration using the WIAI-provided mobile tool

2. Clinical Vocabulary Word-Level Audio Data Collection

- **Total Word List:** ~5,000 clinical terms (provided by WIAI)
- **Speakers per Word:** Each term to be recorded by ~1,000 unique CHOs and MOs
- **Geographical Distribution:** Same set of states as above (~200 speakers per state)
- **Content:** Clean, isolated pronunciation of clinical terms
- **Recording Format:** Word-level recordings using the WIAI mobile application

Speaker Diversity Requirements

To ensure robustness of the ASR model across diverse user profiles, the agency must ensure the following representation in the speaker pool:

- **Gender Balance:** Approximately 50% female and 50% male speakers
- **Age Distribution:**
  - 22–30 years
  - 31–45 years
  - 46+ years
- **Qualification Diversity:**
  - BAMS / BHMS (AYUSH)
  - BSc Nursing / GNM
  - MBBS
  - Other CHO-eligible cadres
- **District Type Representation:** Proportional coverage across urban, rural, and tribal regions, aligned to workforce deployment patterns within each state

Metadata Requirements

Each audio file must be accompanied by standardized metadata fields, including:

- State
- District
- Health Facility Name
- Unique ID
- Gender
- Age Group
- Qualification
- Language

Tooling & Infrastructure Expectations

The selected agency **must have in-house or proven access to data collection and annotation tooling capabilities**, including:

- Tools for real-time audio quality monitoring
- Structured metadata tagging interfaces
- Annotation interfaces for timestamp labeling (if required in later phases)

Proposals should clearly outline methodology, team expertise, and pricing separately for each of the following scenarios:

- Audio sample collection only
- Audio sample annotation (including timestamp labeling)

End-to-end solutions covering both collection and annotation. Agencies may choose to apply for **only data collection**, **only annotation**, or **both collection and annotation**. However, **preference will be given to agencies demonstrating in-house capability and experience to deliver both components under a unified workflow**.

Agencies are requested to apply for complete **data collection and annotation**. Agencies that do not have in-house annotation capabilities may **partner with an annotation agency and submit a joint proposal as a consortium**. However, **preference will be given to proposals where both components are managed under a single unified workflow** to ensure accountability and minimize the risk of quality rejections and coordination delays.

The selected agency is expected to complete the audio data collection within 2 months and submit the data in 4 batches over intervals of 15 days each. Payment eligibility will be linked to the successful submission of each batch. The agency may choose to raise a single consolidated invoice after all 4 batches are delivered or submit invoices batch-wise after each delivery. Additionally, the agency will be required to participate in weekly progress reviews and monthly milestone checks to ensure timely completion and quality adherence.

Sample Script:
"A 40-year-old male presents with complaints of chronic back pain for the past six months. He describes the pain as dull and constant, worsened by prolonged standing. The patient also mentions occasional numbness in his legs. He has no history of back injuries but notes that his father had similar issues. Additional symptoms include difficulty sleeping and fatigue. He denies any recent trauma or changes in bowel or bladder function."

V. Timeline and Rounds

The timeline for the task of voice recording collection is 2 months.

S.No	Activity	Date
1	Release of the Request for Proposal (RFP)	18th November, 2026
2	Deadline for Expression of Interest	25th November, 2026
3	Query Addressal (Virtual Session)	30th November, 2026

4	Last date for submission of the complete proposal	2nd December, 2026
5	Notification to agencies - shortlisted in the technical round	7th December, 2026
6	Presentation and discussion on technical proposal	9th December, 2026
7	Notification to the agency that has been shortlisted	14th December, 2026

***Kindly share your expression of interest by 25th November 2026 and ensure that your contact details are correct so we can send all updates to interested parties via email.**

Note:
All necessary approvals, consents, and handling of Personally Identifiable Information (PII) (if applicable), related to mobile phones but not within the content of the recordings, must comply with applicable regulations and laws. It is the responsibility of the selected organization to secure the necessary approvals from the state and district for data collection. The organization is also responsible for ensuring strict adherence to data privacy and security measures throughout the project's duration.

VI. Required Deliverables

The selected agency will be responsible for delivering the following within the agreed project timelines and quality standards:

1. Audio Recordings

- **Paragraph-Level Audio Recordings:**
  - Minimum of **100 hours per language** (English and Hindi) of scripted paragraph-style patient complaint narration.
  - Data must be distributed across **5 states** (20 hours per state).
  - Recordings must be conducted using the **WIAI-provided mobile application**.

- **Word-Level Audio Recordings:**
  - Clear and isolated pronunciation of approximately **5,000 clinical terms**, each recorded by **~1,000 distinct CHOs and MOs**.
  - Equitable speaker participation across the same 5 states (~200 speakers per state).

Note: *The audio data collected by the selected agency shall not contain any personally identifiable information (PII), and the agency shall ensure appropriate measures are taken to prevent the collection or retention of such data.*

2. Speaker Diversity Documentation

- Evidence of adherence to demographic and professional diversity across:
  - Gender (approx. 50% male / 50% female)
  - Age groups (22–30, 31–45, 46+ years)

- Professional qualifications (e.g., BAMS, BHMS, BSc Nursing, GNM)
- Urban, rural, and tribal districts

3. Metadata

For each audio file submitted, the agency must include the following standardized metadata:

- State
- District
- Health Facility Name
- Name
- Gender
- Age Group
- Qualification
- Language

4. Tooling and Infrastructure

- Functional setup for **audio recording and quality monitoring**
- Tooling for **structured metadata tagging**
- Annotation interface for **timestamping** (if annotation is included in the scope)

5. Submission Schedule

- Deliverables must be submitted in **4 batches** over **15-day intervals**, with the **entire project completed within 2 months**.
- Audio files and metadata must be uploaded to a secure, mutually agreed platform within **48 hours of recording**.

6. Invoicing

- The agency may choose to:
  - Submit a **single consolidated invoice** upon completion of all four batches, or
  - Raise **batch-wise invoices** after each approved submission.

7. Review and Reporting

- Participation in **weekly review meetings** to report progress, challenges, and mitigation plans.
- Participation in **monthly milestone check-ins** to ensure quality and timeline adherence.

VII. Proposal Submission Guidelines

Your proposal must include:

1. **Technical Proposal**
   - Understanding of scope
   - Team composition and qualifications
   - Data collection plan

- Field management plan
- Quality assurance mechanisms
- Ethical safeguards

2. **Financial Proposal**
   - Detailed cost breakdown (by activity, training cost, number of personnel, travel, data collection days)
   - Applicable taxes and contingencies
   - Timeline-based payment schedule
   - Two separate final quotations must be submitted:
     1. Cost for collection and delivery of audio recordings only
     2. Cost for collection, delivery, and annotation (labeling) of audio recordings (e.g., timestamps of recording)
3. **Organizational Profile**
   - Relevant project experience
   - Sample of similar work (quant & qual tools/datasets)

VIII. Proposal Submission Guidelines

Who Can Apply

Non Profit Organizations, Business firms/ for Profit Organizations, Hospitals, Medical colleges, Public Health Organizations. The organization should have prior expertise in the thematic area of the work.

Criteria for Application

Applicants must meet the following criteria to be considered for the audio data collection project:

- Inclusion of actual CHOs and MOs from the public healthcare system, or a combination of mimicked and actual personnel.
- Clear definition of the states where audio data collection will be conducted.
- Specification of the gender distribution within the sample.
- Detailed explanation of how necessary approvals will be obtained and managed.
- Identification of the recording locations (e.g., hospitals, public health facilities, etc.).
- Quantified and targeted sample size for CHOs and MOs.
- Documentation of past relevant experience in audio data collection.

Submission Details

- All responses to this RFP must be received no later than **25th November, 2026.** The proposal should be submitted only through e-mail in PDF format addressed Procurement Team in the below-given e-mail id: rfp.lehs@wadhwaniai.org. The email's subject line must contain the reference number and title of the RFP: **Proposal for Stratified Audio Data Collection – WIAI (Name of the Agency)**.

- Any proposals received by Wadhwani-AI after the deadline for submission of proposals prescribed in the timeline of this document are liable to be rejected.

IX. Evaluation Criteria for Technical Proposal

Criteria	Weightage
Experience in similar studies	20%
Field capacity and demonstrated ability to hire, train, mentor and monitor teams timely	20%
Team expertise	25%
Technical soundness of approach	25%
Data privacy and ethics compliance	10%

Only participants in RFP qualifying on technical evaluation will be eligible for financial evaluation. Financial proposals will be evaluated on criteria including total cost, cost breakdown, cost reasonableness, cost-effectiveness, payment terms, and financial stability of the participating organization.

X. Mandatory Enclosure

Please submit the following as enclosures or attachments with your quotation. Bidders must provide all the information requested above in section VIII. Quotations that do not provide the required information and certificates, or do not follow the submission requirements, may not be reviewed.

Company Profile, testimonials, work completion reports, purchase orders, and

reference (NGOs).

- NGOs need to provide a valid CSR-1 issued in favor of the NGO by the MCA.
- GST registration certificate.
- PAN & TAN registration.
- Audited Financials for the last three years.
- Income Tax Return for the last three years.
- Cancelled cheque.
- Copy of registration documents/certificate and the most recent renewal as a legal

entity.

- Preferably a Private Ltd company and has sound business records.
- Authorized signatory letter.
- Conflict of interest declaration.
- Contact person details.

Memorandum of Association & Articles of Association (or Trust Deed / Bye-laws as applicable).

- List of Governing Body Members/Board Members of the organization.
- 12A and 80G Registration Certificates and Revalidation Document.
- PF, ESI, and Professional Tax Registration Certificates (if applicable).
- Finance Policy, HR Policy, Anti-Sexual Harassment Policy, Procurement Policy, and any other relevant internal policies.
- List of Pending Litigation/Cases against the Entity (if any).
- Latest TDS Return filed (Form 26Q and 24Q).
- GSTR filings for the last 3 returns (if applicable).
- Sample Audited Utilization Certificate (UC) submitted to any donor for CSR Funding.
- Audit Report in Form 10B/BB for the last 3 years.
- Foreign Contribution (FC) Audited Financials for the last 3 years (if applicable).
- FCRA Returns (FC-4) for the last 3 years (if applicable).

All responses to this RFP must be received no later than **25th November, 2026.** The proposal should be submitted only through e-mail in PDF format addressed Procurement Team in the below-given e-mail id: rfp.lehs@wadhwaniai.org

For clarifications and queries, please direct all queries to rfp.lehs@wadhwaniai.org by 25th November, 2026.

Note: Only shortlisted vendors will be contacted for presentations or negotiations. If you do not hear from us within two weeks of submission, consider your proposal not selected. LEHS reserves the right to reject any or all proposals and to negotiate terms and conditions with the selected vendor.