

REQUEST FOR PROPOSAL (RFP)

Data Collection for Foundational Reading Tool in Chhattisgarh

Closing Date for Submission

<November 18, 2025>

1. INTRODUCTION

Wadhvani AI is the flagship Artificial Intelligence (AI) initiative of LEHS, a charitable organization dedicated to building equitable and sustainable systems that bring modern technology to underserved populations in India. Our core mission is to develop, deploy, and rigorously evaluate AI-powered solutions for social impact across domains, including healthcare, agriculture, governance, and education.

In alignment with our objective to support projects of national and social significance, Wadhvani AI promotes the integration of cutting-edge technologies like AI into the Indian mainstream systems. We execute this through strategic partnerships with State and National Governments, apex institutions, international agencies, and key ecosystem contributors.

This Request for Proposal (RFP) is issued by Wadhvani AI, under the umbrella of LEHS, to identify a dedicated partner to execute a large-scale voice data collection project required for developing and contextualizing a Foundational Reading Tool for the state of Chhattisgarh.

2. BACKGROUND

A significant number of students in Indian public schools (Grades 1–8) lack foundational reading skills, impacting their academic progress across all subjects. This challenge is compounded by the absence of scalable, data-driven tools for assessing reading in diverse languages and for identifying specific proficiency levels to enable targeted interventions.

In response, Wadhvani AI has developed and successfully deployed an **AI-powered tool** for Reading Assessment and Remediation designed to help teachers achieve universal attainment of foundational reading skills. The tool has three components:

1. **Assessment:** The tool leverages highly contextualized and fine-tuned Automated Speech Recognition (ASR) models to analyze word and paragraph reading in different languages.
2. **Diagnostics:** It generates key performance metrics, identifies specific reading miscues, and applies an intelligent layer to categorize students into reading levels. This provides teachers with actionable insights necessary for planning targeted classroom interventions.

3. **Remediation:** We plan to integrate a cutting-edge Generative AI module to produce practice worksheets tailored for each student cohort. This will provide teachers with customizable, context-specific remediation materials for use in the classroom.

To date, Wadhvani AI has successfully assessed over **7 million students** in two major states - Gujarat and Rajasthan, with this tool.

Wadhvani AI is now adapting and contextualizing the tool for deployment in **Chhattisgarh**, covering the essential languages: **Hindi** and **English**. We require an experienced partner to collect the necessary voice data to train the ASR models for the state's specific linguistic context.

3. PROJECT OBJECTIVES

The primary objective is to execute a rigorous, large-scale data collection program to collect voice data of students, reading different texts, in Hindi and English from the state of Chhattisgarh. This collected voice data will be used to train and contextualize the Automated Speech Recognition (ASR) models employed in the Foundational Reading Tool.

Language: Hindi & English

4. SCOPE OF WORK

Data Collection: Collect audio data of students reading texts in Hindi and English from Grade 1-8

Total Number of Hours to be collected: ~300 hours for Hindi and ~300 hours for English

A. Data Collection Volume

The project requires collecting a total of approximately 300 hours for Hindi and 300 hours for English from students across grades 1-8 in Government Schools in Chhattisgarh. This total volume must be distributed across four distinct item types using the Wadhvani AI-provided data collection tool (web link or PWA).

The total required hours is a cumulative target distributed across four item types. The following table illustrates the required volume of recordings and the minimum number of unique children contributing data to meet the target for each item type:

	Hindi	English
letter	20,000 char utterances	20,000 char utterances
	5000 unique children	5000 unique children
	~12 hours	~12 hours

word	10,000-word utterances	10,000-word utterances
	2000 unique children	2000 unique children
	~12 hours	~12 hours
non-word	10,000 word utterances	10,000 word utterances
	2000 unique children	2000 unique children
	~12 hours	~12 hours
paragraph	10,000 paragraph recordings	10,000 paragraph recordings
	10,000 unique children	10,000 unique children
	~250 hours	~250 hours

B. Operational Responsibilities and Training:

The Partner must establish a robust operational system for managing personnel, training, and ensuring the fidelity of all audio recordings.

- **Training and Personnel:** The Partner is solely responsible for training all selected data collectors on the technical tool usage, data collection protocols, and quality standards (e.g., noise reduction, single speaker requirement, ethical consent). The Partner will provide Wadhvani AI with the finalized list of data collectors and mapped schools. Wadhvani AI will provide the necessary login access.
- **Wadhvani AI Support:** Wadhvani AI will provide an online tool to collect the data on and a detailed document of the required data collection and will hold initial meetings to establish the required quality. Wadhvani AI will also review collected data, and the agency is responsible for modifying instructions given to data collectors accordingly.
- **Permissions:** Wadhvani AI has secured permissions with the state government of Chhattisgarh and will facilitate the finalization of the list of schools. The Partner is responsible for all on-ground logistical coordination with the finalized list of schools and ensuring smooth access for data collection teams.

C. Data Quality and Acceptance

- **All data collected is subject to Wadhvani AI's quality review** (either through a third party or the internal Wadhvani AI team).
- **Device Quality:** Ensure that devices used by data collectors (e.g., mobile phones) have microphones of sufficient quality to accurately capture clear audio.

- Review and Re-collection: Wadhvani AI will review 10%-20% of collected recordings in batches. If a batch fails to pass the predetermined quality standards, the partner is responsible for the re-collection of rejected data at no additional cost.
- Data Ownership: All data collected will reside on WIAI servers with the intent to handover the data to the state once the assessments are completed.

5. TIMELINES & ROUNDS

Sr. No.	Stage	Milestone Date
1	RFP Issue	November, 11 2025
2	Q&A Date	November 14, 2025
3	Submission Deadline	November 18, 2025

6. DELIVERABLES

1. ~300 hours of **Hindi** and ~300 hours of **English** collected as per the requirements listed above, through WIAI's tool.
2. Data must pass the **WIAI review process** (10% recordings checked) as per the details above.
3. All Data should be mandatorily collected between **5th December and 15th December**.

7. PROPOSAL SUBMISSION GUIDELINES

Your proposal must include:

1. Technical Proposal

- Detailed understanding of the scope, objectives, and deliverables.

- Team composition and qualifications of key personnel who will be leading the project.
- Detailed plan for the training and capacity building of data collectors.
- Internal Quality Assurance (QA) mechanisms to ensure collected data meets required quality standards (e.g., noise reduction, single speaker).
- Staffing Plan and Risk Management Plan.
- Sample of similar work conducted previously.

2. Financial Proposal

- Detailed cost breakdown.
- Applicable taxes and contingencies.

3. Organizational Profile

1. Relevant project experience, including technical and operational expertise.
2. Geographic coverage.
3. Staffing Plan.
4. Risk Management Plan.
5. Detailed Budget Format.
6. Sample of similar work.

Submission Details

1. All proposals to this RFP must be received no later than <November 18, 2025>. The proposal should be submitted only through e-mail in PDF format addressed to the Procurement Team at: rfp.lehs@wadhwaniai.org. The email's subject line must contain the reference number and title of the RFP: **Data Collection for Foundational Reading Tool in Chhattisgarh.**
2. Any proposals received by Wadhvani-AI after the deadline prescribed in the timeline of this document are liable to be rejected.

8. Evaluation Matrix

Criteria	Weightage
Understanding of Objectives	7%
Detailed Operations/Implementation plan	21%
Team Composition & Capability	7%
Experience in collecting speech audio data (students' reading) in Hindi and English	21%
Quality Assurance & Risk Mitigation	3.5%
Monitoring & Evaluation	7%
Data Policies	3.5%

Weighted Scoring Method.

70% weightage for the technical score and 30% weightage for the financial score.

Criteria for Application

Applicants must meet the following criteria

1. **Mandatory:**

- Demonstrated experience in data collection, especially **voice data** across a state.
- Experience in training and operationalizing a cadre of on-ground data collectors.
- Project management ability to coordinate diverse activities and deliver in stipulated timelines.
- Work experience of all key personnel involved must be included.
- Understanding of different accents present in the state; presence of natives from Chhattisgarh or linguists with an understanding of the linguistic landscape would be a plus.

- Robust monitoring, evaluation, and reporting capabilities.
- Financial stability: Annual corpus
- Demonstrate compliance with the provisions of the Digital Personal Data Protection (DPDP) Act, 2023, ensuring responsible collection, storage, and use of the personal data.

2. Preferred:

- Presence in Chhattisgarh, and prior experience in collecting data through state's administrative mechanisms

9. MANDATORY DISCLOSURES

Please submit the following as enclosures or attachments with your quotation. The agency must provide all the information requested above. 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, reference (NGOs).
- NGOs need to provide valid CSR-1 issued in favor of the NGO by the MCA. NGO's having valid FCRA registration shall disclose the same in their proposal and provide a copy of the registration certificate and relevant supporting documents along with the application.
- 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 most recent renewal as a legal entity.
- Preferably 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).

Note: Only Technically shortlisted Bidders will be eligible for the opening of the financial proposal. 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.

Bidders selected by the LEHS may be asked for their presentations after reviewing technical proposals and negotiations.

If bidders do not hear from LEHS within two weeks of closing of the RFP, consider your proposal not selected. LEHS reserves the right to reject any or all proposals at any given time without giving any justification or response to the bidders.