

REQUEST FOR PROPOSAL (RFP)

Pedagogical, Operational, and Program Partnership for Deploying Reading Assessment and Remediation Tool for the State of Tamil Nadu

Opening date of RFP: 17/10/2025

Closing Date of RFP: 29/10/2025

TABLE OF CONTENTS

I.	INTRODUCTION	2
II.	BACKGROUND	2
III.	PROJECT OBJECTIVES.....	3
IV.	SCOPE OF WORK.....	3
1.	Facilitate the onboarding of Wadhvani AI into the school ecosystem	4
2.	Content Creation for Data Collection: Create textual items for data collection of students' reading in Tamil and English for Grades 1-8.....	4
3.	Data Collection: Collect audio data of students reading texts in Tamil and English from Grade 1-8	5
4.	Data Annotation for ASR: Create timestamped human transcriptions as per the finalized protocol by Wadhvani AI for both Tamil and English.....	7
5.	Content Creation for Assessments for pilots, and census assessments for Grades 1-8 for both languages	10
6.	Data Annotation for Diagnostics Module and threshold setting	10
7.	Remediation Content Development and GenAI Output Validation.....	11
V.	TIMELINES & ROUNDS.....	12
VI.	DELIVERABLES.....	12
VII.	OUTCOMES	14
	Assessment:	14
VIII.	PROPOSAL SUBMISSION GUIDELINES.....	15
IX.	Evaluation CRITERIA FOR TECHNICAL PROPOSAL.....	17
X.	MANDATORY DISCLOSURES	18

I. 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) has been issued by Wadhvani AI, under the umbrella of LEHS, to find a dedicated partner to execute a critical project in the field of education in the State of Tamil Nadu.

II. 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:

- **Assessment:** The tool leverages highly contextualized and fine-tuned Automated Speech Recognition (ASR) models to analyze word and paragraph reading in different languages.
- **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.
- **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 the state of **Tamil Nadu**, covering two essential languages: **Tamil and English**. We are seeking an organization with a deep understanding of, and active engagement with, the Tamil Nadu state education system to engage in a comprehensive **Pedagogical, Operational, and Program Partnership**. This collaboration is essential to ensure the solution is academically rigorous, operationally sound, and seamlessly integrated into the state's educational ecosystem.

III. PROJECT OBJECTIVES

The overarching goal of this partnership is to successfully design, deploy, and integrate Wadhvani AI's Reading Assessment and Remediation Tool (RART) within the Tamil Nadu state education system. This collaboration aims to achieve universal attainment of foundational reading skills among primary and upper primary students.

The partner organization will be responsible for jointly achieving the following objectives:

1. **Establish Program Credibility & Reach:** Seamlessly onboard Wadhvani AI into the Tamil Nadu school ecosystem, securing the necessary approvals and organizational buy-in for successful large-scale deployment, and parental consent for each student assessed in compliance with DPDP ACT 2023.
2. **Ensure Pedagogical Rigor:** Validate and contextualize the RART's reading assessment, diagnostic, and remediation modules to align perfectly with the Tamil Nadu State Curriculum and the specific challenges of learning Tamil and English in the state.
3. **Guarantee Data Quality:** Execute a rigorous, large-scale data collection and annotation program to train and fine-tune the AI models, ensuring the tool's accuracy and diagnostic strength for Tamil and English.
4. **Drive Teacher Adoption & Usage:** Conduct effective teacher training and provide on-ground support to ensure educators can confidently utilize the RART's assessment, diagnostics, and Generative AI-powered remediation modules.
5. **Achieve Sustainable Scale:** Successfully pilot and deploy the RART across schools in alignment with the state's academic cycle, laying the groundwork for integration and long-term program sustainability.

Language: Tamil & English

Contract Duration: The initial contract duration will be **one year**, with potential for renewal based on satisfactory performance and mutual agreement.

IV. SCOPE OF WORK

This section outlines the specific responsibilities and deliverables expected from the selected program partners to ensure the successful implementation of the "Reading Assessment and Remediation" tool across the state of Tamil Nadu.

Summary of Activities

S No	Activity Description
1	Facilitate the onboarding of Wadhvani AI into the school ecosystem
2	Content Creation for Data Collection: Create textual items for data collection of students' reading in Tamil and English for Grades 1-8

3	Data Collection: Collect audio data of students reading texts in Tamil and English from Grade 1-8
4	Data Annotation for ASR: Create timestamped human transcriptions as per the finalized protocol by Wadhwani AI for both Tamil and English.
5	Content Creation for Assessments for pilots, and census assessments for Grades 1-8 for both languages
6	Data Annotation for Diagnostics Module and threshold setting
7	Remediation Content Development and GenAI Output Validation
8	Training of government schoolteachers to use the Wadhwani AI application
9	Integration of the Wadhwani AI application into the relevant state technology ecosystem

Each activity is detailed below:

1. FACILITATE THE ONBOARDING OF WADHWANI AI INTO THE SCHOOL ECOSYSTEM

- Facilitate the successful entry and integration of Wadhwani AI's Reading Assessment and Remediation Tool (RART) into the Tamil Nadu education system.
- Serve as the primary, high-level liaison between Wadhwani AI and all relevant state and local education authorities (e.g., SCERT, SSA, District Education Officials), school leadership, and teachers.
- Secure all necessary state-level permissions and clearances for data collection, teacher engagement, and pilot/assessment deployment.
- Assist in MoU signing with relevant state bodies, and in resolving data-related concerns.
- Proactively ensure the RART's features, deployment schedule, and data collection protocols are fully aligned with existing Tamil Nadu state education policies, foundational literacy programs, and administrative workflows.

2. CONTENT CREATION FOR DATA COLLECTION: CREATE TEXTUAL ITEMS FOR DATA COLLECTION OF STUDENTS' READING IN TAMIL AND ENGLISH FOR GRADES 1-8

Total Number of Items to be created: ~750-900 per language

- a. Create a corpus of textual assessment items in **Tamil and English** (Grades 1-8) for use in fine-tuning the Automated Speech Recognition (ASR) models.
- b. Create a total of **≈750–900 items per language** (Tamil and English). The overall corpus across both languages must cover **≈8,000–10,000 unique words** to ensure robust ASR model training.
- c. The items must cover four distinct types, contextualized for Tamil Nadu: **Sets of Letters, Sets of Words, Sets of Non-Words, and Comprehension items based on the reading paragraphs.**
- d. All item types must be leveled for each grade (1-8) to allow students to attempt reading them appropriately. Leveling parameters will be finalized in consultation with the Wadhvani AI team.
- e. Item-specific requirements:
 - Sets of Letters: Include individual letters only (no diacritic combinations in Tamil) and both capital/small letters in English.
 - Sets of Words: Must have a good balance of familiar and unfamiliar words
 - Sets of Non-Words: Must be generated using the phonotactic rules of each language (Tamil/English), and proper nouns must be excluded.
 - Paragraphs: Leveled paragraphs for each grade.
- f. Ensure the corpus includes a **balance of familiar and non-familiar words** for each grade/language to optimize ASR model exposure. Exact set composition will be finalized with the Wadhvani AI team.

3. DATA COLLECTION: COLLECT AUDIO DATA OF STUDENTS READING TEXTS IN TAMIL AND ENGLISH FROM GRADE 1-8

Total Number of Hours to be collected: ~400 hours for Tamil and ~130 hours for English

- a. Execute the end-to-end collection of student reading audio data in **Tamil and English** (Grades 1-8) using the Wadhvani AI-provided data collection tool (web link or PWA).
- b. Achieve a minimum collection of **≈400 hours for Tamil** and **≈130 hours for English.**
- c. Data will be collected through a data collection tool designed and provided by WIAI.
- d. Collect data distributed across the four item types (Letters, Words, Non-Words, Paragraphs) with an indicative breakup given below.

Data Collection Volume

	Tamil	English
letter	44,000 char utterances	11,000 char utterances
	8,800 unique children	2,200 unique children
	~24 hours	~6 hours
word	22,000-word utterances	5,500-word utterances
	4,400 unique children	1,100 unique children
	~24 hours	~6 hours
non-word	22,000-word utterances	5,500-word utterances
	4,400 unique children	1,100 unique children
	~24 hours	~6 hours
paragraph	13,000 paragraph recordings	4,400 paragraph recordings
	13,000 unique children	4,400 unique children
	~330 hours	~110 hours

- e. Research, validate, and execute a data collection strategy that accounts for the dialectical diversity of Tamil Nadu.
 - Research and propose a structured categorization of districts into **District Groups**, where each group represents a unique linguistic accent or dialectical region of Tamil and English within the state. The proposed grouping must be verified and confirmed by the State Administration.
 - Distribute the total required data volume **equitably** across the validated District Groups. Based on the required volume, calculate the specific number of schools needed per district.
 - All data collection must occur exclusively in **Government Schools**. Within the finalized school list for a district, ensure an approximate split of **50% Urban areas** and **50% Rural areas**.
 - Strive for a **50:50 split between boys and girls** in the final data collected. Co-educational schools should be prioritized to facilitate this split.
- f. Establish a robust operational system for managing personnel, training, and ensuring the fidelity of all audio recordings.

- The Partner is solely responsible for **training all selected data collectors** on technical tool usage, data collection protocols, and quality standards (e.g., noise reduction, single speaker requirement, ethical consent).
 - Ensure that devices used by data collectors (e.g., mobile phones) have **microphones of sufficient quality** to accurately capture clear audio.
 - The partner needs to select and train reviewers who will drive a **robust, daily internal Quality Control (QC) process**. These reviewers will be given login access to the tool and must verify that each recording meets strict quality standards (e.g., no multiple speakers, minimal background noise) before being shared with Wadhvani AI.
- g. The Partner will provide Wadhvani AI with the finalized list of data collectors and mapped schools. Wadhvani AI will grant the necessary login access to the data collection tool.
- h. Wadhvani AI will review (either through a third party or by the internal WIAI team) - 10%-20% of collected recordings in batches and may **reject a batch** if it fails to pass the predetermined quality standards. The partner is responsible for re-collection of rejected data at no additional cost.
- i. All data collected will reside on WIAI servers with the intent of handing over the data once the assessments are completed. The partner is expected to resolve any concerns regarding Data Fiduciary responsibilities with the state.

4. DATA ANNOTATION FOR ASR: CREATE TIMESTAMPED HUMAN TRANSCRIPTIONS AS PER THE FINALIZED PROTOCOL BY WADHWANI AI FOR BOTH TAMIL AND ENGLISH.

Total Number of Hours to be annotated: ~360 hours for Tamil and ~115 hours for English

Data Annotation Volume

	Tamil	English
letter	40,000 char utterances	10,000 char utterances
	8,000 unique children	2,000 unique children
	~20 hours	~4-5 hours
word	20,000-word utterances	5,000-word utterances
	4,000 unique children	1,000 unique children

	~20 hours	~4-5 hours
non-word	20,000-word utterances	5,000-word utterances
	4,000 unique children	1,000 unique children
	~20 hours	~4-5 hours
paragraph	12,000 paragraph recordings	4,000 paragraph recordings
	12,000 unique children	4,000 unique children
	~300 hours	~100 hours

a. Human Transcription

- Perform exact transcription of the audio recordings in **Tamil and English** to serve as the ground truth for ASR model training, for all four task types (sets of letters, sets of words, sets of non-words, paragraphs).
- Timestamping:
 - For paragraphs: the transcription must include **precise timestamps** for *each individual word*. Where words are broken down sub-lexically (e.g., sound-by-sound reading), timestamps must be provided for the **sub-lexical portions** rather than the word as a whole.
 - For Letters, Words and Non-words: A JSON with timestamps will be provided to the partner. Transcription for each time-stamped portion will have to be done by the partner.
- Transcription:
 - For Tamil: Annotators must transcribe **exactly what is heard**, including all reading miscues such as **mispronunciations, partially read words, incorrect readings, fillers and pauses**.
 - For English:
 - The partner will need to source **10-15 ideal readings of each data collection item** from each district to ensure dialectical diversity. This will serve as a reference for the annotators.
 - Each correctly pronounced English word, referenced through the ideal readings, needs to be transcribed in English.

- For English mispronounced words, the Partner must utilize a **phonetic script/system** selected by Wadhvani AI. The final transcript for English will be a combination of standard English text and this phonetic notation. WIAI will provide finalized instructions for annotations of English recordings.
 - A detailed transcription protocol will be provided by WIAI for each language.
- b. Labels
- In addition to the speech transcript, wherever appropriate, labels are to be added to the audio to denote the presence of Fillers, Noise, Long silences, Mumbling, Multiple Speakers, and accent-influenced pronunciation.
 - Whenever a word is read in a broken manner with pauses between sounds, a special label indicating an intra-word pause needs to be marked at all instances of such pauses.
 - Final labels to be added will be provided by WIAI.
- c. Quality Review
- An internal system of quality checks should be in place for every transcription, both for the quality of transcription and for the quality of labels.
 - The transcripts should be delivered to Wadhvani AI in batches of 20 hours each, and two rounds of quality check feedback might be given by Wadhvani AI on each batch. WIAI's review will include a review of transcriptions as well as a review of labels.
- d. Data Policy
- Wadhvani AI will share anonymized data, stripped of PII, uploaded in an AWS S3 bucket with the selected agency, from the data collected.
 - The agency needs to have processes in place to prevent the download of data by individual annotators
 - The Partner is required to sign an undertaking guarantee that this voice data will **not be used for any proprietary purposes** outside of this contractual agreement.
- e. Formatting & Delivery
- Delivering translations in a JSON format as specified by Wadhvani AI.

- The Partner is free to choose and utilize any annotation software, provided the software and its data storage methods **pass the data security and privacy policies of Wadhvani AI.**

f. Tool for Annotation

- WIAI does not have an internal tool for annotation. The partner is expected to have a tool for annotation or to procure a tool for this purpose that serves all the above-mentioned requirements.

5. CONTENT CREATION FOR ASSESSMENTS FOR PILOTS, AND CENSUS ASSESSMENTS FOR GRADES 1-8 FOR BOTH LANGUAGES

- a. Create a robust, grade-specific framework for writing test items that governs the structure, difficulty, and pedagogical focus of all assessments.
- b. The framework must explicitly take into account and align with a) Tamil Nadu State Pedagogy and Curriculum. b) Established Foundational Literacy Frameworks such as Early Grade Reading Assessment (EGRA) and Global Proficiency Framework for Reading (GPF-R).
- c. The final assessment framework must be mutually validated and formally approved by both the Wadhvani AI internal pedagogy team and the State Council of Educational Research and Training (SCERT).
- d. Design and write the required volume of assessment items in both Tamil and English for Grades 1–8, strictly adhering to the approved framework. The number of assessment items would be decided in consultation with the state government.
- e. A new, unique set of assessment items must be created for each distinct round of pilot or census assessment conducted throughout the project duration.

6. DATA ANNOTATION FOR DIAGNOSTICS MODULE AND THRESHOLD SETTING

- a. Engage a team of expert teachers to label a distinct subset of student reading recordings (~5000 recordings per language) with advanced diagnostic metrics to train the AI's intelligent layer for student cohorting.
- b. Apply detailed labels to each word in the designated audio subset to differentiate reading methods: **Lexical Reading:** Labeling the word if it is read as a whole unit. **Sub-Lexical Reading:** Labeling the word if it is decoded sound-by-sound or syllable-by-syllable.
- c. Annotate each recording for the presence and accuracy of reading pauses: Labeling instances of **correct pauses** (e.g., at punctuation marks). Labeling instances of **incorrect/inappropriate pauses** (e.g., mid-sentence, intra-word, or disruptive hesitations).

- d. Co-develop, validate, and apply a framework for annotating the **prosody and expressive quality** of the student reading.
 - **Framework Selection:** Explore and propose a suitable framework (e.g., adapting the NAEP framework or another validated standard) for assessing elements like pitch, stress, and volume.
 - **Annotation Application:** The expert team must label each recording based on this mutually agreed-upon framework.
- e. Implement rigorous quality checks to ensure the diagnostic labels are objective and consistent. Establish and maintain an active **Inter-Rater Reliability (IRR) system** across the expert team to mitigate subjectivity, especially for labels related to prosody and pauses. The IRR threshold must be mutually agreed upon with Wadhwani AI.
- f. **Collaborate closely with the Wadhwani AI Pedagogy Team** to jointly determine the **diagnostic metric thresholds** (e.g., error rates, pause frequency) that differentiate student proficiency levels. Map these finalized thresholds directly to the **four foundational reading cohorts** used by the tool to ensure actionable grouping of students for targeted remediation.
- g. Tool: Wadhwani AI will provide the tool for this labelling.

7. REMEDIATION CONTENT DEVELOPMENT AND GENAI OUTPUT VALIDATION

- a. For each of the four foundational reading cohorts, define and document the pedagogical requirements for the practice worksheets.
- b. Define specific worksheet templates that are pedagogically appropriate for the skills deficit of the corresponding cohort. This includes specifying the exercise type, format, and layout. Upto 5 templates to be defined for each cohort.
- c. Define the specific constraints and rules the GenAI model must adhere to when generating content (e.g., word length limits, number of items, target phonemes, grammar rules, context) for each worksheet template.
- d. Curate ≈ 50 high-quality examples for each defined worksheet template type. These examples will be used by Wadhwani AI to train and fine-tune the GenAI model's output.
- e. Write clear, practical Teacher Instructions and implementation guides to accompany each worksheet template. These instructions must detail the objective, suggested use, and expected teacher role.
- f. Establish and execute a robust human review process to validate all GenAI-generated worksheets before deployment.

- Review Rubric: Co-create a detailed Worksheet Evaluation Rubric with the Wadhvani AI team for human reviewers hired by the partner to assess the pedagogical, linguistic, and contextual quality of GenAI-generated content.
- Worksheet Approval: Partner to deploy a dedicated team of Expert Teachers or content creators to serve as human reviewers to approve all GenAI-generated worksheets. Every worksheet generated must be approved before being released to teachers.
- Annual Vetting Volume: The Partner is responsible for the quality assurance and approval of $\approx 5,000$ worksheets across all templates and cohorts within the first year of the contract.

g. Tool: Wadhvani AI will provide the tool for the worksheet review of 5000 worksheets.

8. Teacher Training: Train all government school language teachers, of the grades being assessed, about the AI application for carrying out on-ground deployment of the tool for data collection and assessment. Methodology for training teachers should be specified.

9. Deployment: Carry-out on-ground deployment of the tool for pilots, assessment, diagnostics, and remediation (or as the case may be) aligned with the state's academic cycle. The scope will include government liaison and co-ordination with state departments.

10. Monitoring and reporting of deployment: Facilitate evaluation via third party as and when required. The identification of a third party, and the cost for the evaluation will be borne by Wadhvani AI.

V. TIMELINES & ROUNDS

Sr. No.	Stage	Milestone Date
1	RFP Issue	October 17, 2025
2	Q&A Date	October 22, 2025
3	Submission Deadline	October 29, 2025

VI. DELIVERABLES

Activity-wise deliverables are listed below:

1. Facilitate the onboarding of Wadhvani AI into the school ecosystem

- a. MoU of Wadhvani AI with the state for the deployment of the RART.
- b. Training of all government teachers to use the application.

- c. Deploy Wadhvani AI's tool for assessment, diagnostics, and remediation in alignment with the state's academic cycle, and in collaboration with the state department of education. Training of Teachers: Pre-decided number of teachers trained.
- d. Integration of technology tools of Wadhvani AI into the state technology ecosystem through on-ground support, liaison, and successful demonstration of the implementation.

2. Content Creation for Data Collection: Create textual items for data collection of students reading in Tamil and English for Grades 1-8

- a. A CSV file for each language - Tamil and English with ~750-900 data collection items, divided across 4 different task types.
- b. Format of the CSV will be provided by WIAI.

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

- a. ~400 hours of Tamil and ~130 hours of English collected as per requirements listed above, through WIAI's tool.
- b. Data must pass WIAI review process (10% recordings checked) as per details above).

4. Data Annotation for ASR: Create timestamped human transcriptions as per the finalised protocol by Wadhvani AI for both Tamil and English.

- a. JSON files of transcriptions of recordings amounting to ~360 hours in Tamil and ~115 hours in English.
- b. Each recording to have a separate JSON file, format as specified by Wadhvani AI.

5. Content Creation for Assessments for pilots, and census assessments for Grade 1-8 for both languages.

- a. A CSV file for each language - Tamil and English with required number of assessment items, divided across multiple task types.
- b. Format of the csv will be provided by WIAI.
- c. A new csv will need to be created for every round of assessment.

6. Data Annotation for Diagnostics Module and threshold setting.

- a. ~5000 recordings labelled with metrics described above for each language.
- b. WIAI will provide the tool for this labelling, output file format will be defined as per this tool.

7. Remediation Content Development and GenAI Output Validation.

- a. A document of the worksheet framework, the format of the framework will be defined in collaboration with Wadhwani AI's pedagogy team. This document will also include teacher instructions to accompany each worksheet type.
- b. A compendium of 50 worksheets (machine readable PDF format) per worksheet type, uploaded in a folder. Folder paths will be provided by Wadhwani AI.
- c. A document of the worksheet review rubric, this rubric will be co-defined with the Wadhwani AI team.
- d. ~5000 worksheets reviewed by expert teachers employed by the partner. The review will be done on a tool provided by Wadhwani AI.

VII. OUTCOMES

ASSESSMENT:

- **At least 75% of teachers** report increased ability to identify and support students with oral reading difficulties through the platform's tools.
- **20% increase** in correct words per minute (CWPM) for at least **60% of students assessed**, within one academic year.

To effectively measure impact and ensure equitable outcomes, the program will incorporate the following:

Measurement Approach

- **Baseline Assessments:** Conduct initial assessments to establish a starting point for student performance and teacher capabilities.
- **Midline Assessment:** Implement mid-term assessments to monitor progress and identify areas for course correction.
- **Remediation:** Based on assessment results to address identified learning gaps and improve student outcomes. In case the State wants two assessment cycles - Baseline and Endline, then the Remediation is to be done after the Baseline.
- **Endline Assessments:** Perform final assessments to evaluate the overall impact and achievement of desired outcomes.

Equity Lens

The program will actively consider the following factors to address potential inequities:

- **Gender:** Analyze data and tailor interventions to ensure equitable outcomes for both male and female students.

- **Rural/Urban School Location:** Account for differences in infrastructure, internet access, and learning environments between rural and urban schools.
- **Multi-level, Multi-grade Classrooms:** Design strategies that are effective in diverse classroom settings.
- **Socio-economic Background of Students:** Consider the socio-economic context of students and adapt approaches to support all learners.

In addition to the above stated outcomes, the program aims to achieve significant system-level impact:

- System Level Outcomes:
 - State Adoption: Achieve formal adoption of the AI learning tools and methodologies by the state education department.
 - Cost per Child Reduction: Demonstrate a reduction in the cost per child for delivering quality education through the integrated AI solutions.
 - Policy Integration: Secure integration of the program's components into existing state policies, academic calendars, or specific state campaigns/initiatives.

VIII. PROPOSAL SUBMISSION GUIDELINES

Your proposal must include:

1. **Technical Proposal**
 - Understanding of scope and deliverables
 - Details of process and tools for each Activity
 - Team composition and qualifications
 - Quality assurance mechanisms
 - Data Policy
2. **Financial Proposal**
 - Detailed cost breakdown
 - Applicable taxes and contingencies
3. **Organizational Profile**
 - Relevant project experience including technical and operational expertise
 - Geographic coverage
 - Staffing Plan
 - Risk Management Plan
 - Detailed Budget Format
 - Sample of similar work

Proposals should include a timeline **ensuring full fund utilization by March 2026 and impact reporting through December 2027.**

Criteria for Application

Applicants must meet the following criteria:

1. **Mandatory:**

- a. Presence in Tamil Nadu with clear lines of communication with all relevant state education agencies and ability to collect audio data from across the state.
- b. Demonstrated experience in design, and implementation of programs in state school ecosystems through state-level partnerships.
- c. Experience in running scaled-up program i.e. more than 100000 students.
- d. Expertise in curriculum design, and capacity building. Educational content creation expertise for FLN for Tamil and English contextualised for Tamil Nadu to create data collection items, assessment items and worksheet templates.
- e. Native Tamil speakers in the team for participating and driving review processes for the 'data annotation for ASR' activities.
- f. Ability to engage expert teachers for Tamil and English languages to participate in annotation and review of 'data annotation for Diagnostic module' related activities and to review worksheets generated by genAI for remediation.
- g. Project management ability to coordinate diverse activities and deliver in stipulated timelines.
- h. Robust monitoring, evaluation, and reporting capabilities.
- i. Technology enabled, and ready in terms of applications, dashboards, integrations.
- j. Financial stability: Annual corpus.
- k. 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:**

- a. Prior experience in data annotation or experience in contracting other agencies for data annotation.
- b. Prior experience in creating assessment items to test foundational reading using specific criteria for Grades 1-8.
- c. Prior experience in co-creating AI tools for FLN or for reading.

Submission Details

- Both Technical and Financial Proposals must be submitted electronically to rfp.lehs@wadhwaniai.org **separately**. The email's subject line must contain the reference number and title of the RFP: **Pedagogical, Operational and Program Partnership for deploying Reading Assessment and Remediation Tool for the State of Tamil Nadu (Name of Agency)**.
- Financial Proposal must be password protected.

- All documents should be in PDF format.
- **Submission Deadline: [October 29, 2025 – 23:59 Hours]**
- Any proposals received 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

Evaluation Matrix

Criteria	%
Understanding of Objectives	5%
Technical soundness and methodology for each activity outlined in the Scope of Work)	30%
Existing relationships and work done with the State/Education Department	10%
Expertise in pedagogy & content creation	10%
Monitoring, Evaluation, and Learning (MEL)	10%
Quality Assurance & Risk Mitigation	5%
Team Composition & Capability	15%
Experience & Past Performance in Data Collection, Data Annotation: Experience in collecting speech audio data (students' reading) in Tamil and English with the State Administration. Experience in the creation of textual assessment items in Tamil and English in collaboration with SCERT. Experience in transcribing speech for training speech recognition models. Experience in transcribing speech with miscues, where miscues are transcribed verbatim. Technical soundness of approach (tools used to create timestamped transcripts, quality check process). Experience in transcribing Tamil and English speech and labeling audios according to speech patterns. Experience in developing and reviewing content for remediation programs for Tamil and English. Data policies in adherence with the legal requirements.	15%
Technical Total	100%
Weighted Scoring Method	
70% weightage for technical score and 30% weightage for financial score.	

X. MANDATORY DISCLOSURES

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, reference (NGOs).
- NGOs need to provide 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 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.