

REQUEST FOR PROPOSAL (RFP)

Data Collection for Foundational Numeracy Tool in UP

Closing Date for Submission

<10th April 2026>

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 handwriting data collection project required for developing and contextualizing a Foundational Numeracy Tool for the state of Uttar Pradesh.

2. BACKGROUND

A significant number of students in Indian public schools lack foundational numeracy skills in the early grades, affecting their ability to progress in mathematics and other subjects. While initiatives such as the National Education Policy (NEP) 2020 and the NIPUN Bharat Mission have emphasized the importance of foundational learning, classrooms still lack scalable, data-driven tools to assess numeracy skills and identify specific learning gaps at the student level. Teachers often rely on manual grading of paper-based worksheets, which is time-consuming and limits the ability to conduct frequent and diagnostic assessments.

In response, Wadhvani AI is developing an AI-based system to automate the evaluation of paper-based numeracy assessments and provide teachers with actionable insights. The tool focuses on recognizing handwritten student responses to foundational numeracy problems such as addition, subtraction, and fill-in-the-blank number tasks. By digitizing worksheet responses using machine learning models, the system aims to reduce teacher grading effort while enabling more structured and frequent assessment of numeracy skills.

The system consists of three components:

- **Assessment**
- **Diagnostics**
- **Remediation**

To support the development and evaluation of the OCR models that power the assessment component, Wadhvani AI is collecting a dataset of handwritten student worksheets. We require an experienced partner to support the capturing of these worksheet images to generate high-quality labeled data for training and testing the OCR models.

3. PROJECT OBJECTIVES

The primary objective of this project is to execute a structured, large-scale data collection program to capture handwritten student responses from paper-based numeracy assessments. The collected dataset will be used to train and evaluate OCR models designed to recognize handwritten Arabic numerals and simple mathematical responses from student worksheets.

These worksheets contain foundational numeracy problems such as addition, subtraction, and fill-in-the-blank number tasks, where students write their answers in predefined boxes. The captured data will support the development of AI systems that automate worksheet evaluation and generate diagnostic insights to help identify foundational numeracy gaps among students.

Content Type: Handwritten responses to foundational numeracy questions
Script: Arabic numerals (0–9) and simple mathematical responses/symbols (+, -, <, >, =)

Age group of students: 7 - 9 years old

Time for completion : 3 weeks

4. SCOPE OF WORK

Data Collection: Handwritten Numeracy Assessment Worksheets

The project requires the collection of handwritten student worksheets containing responses to foundational numeracy questions. These worksheets will later be used to train and evaluate OCR models that recognize handwritten Arabic numerals and simple mathematical responses from structured answer boxes.

The data collection partner will be responsible for administering worksheets to students, capturing images of the completed worksheets, and ensuring that the collected images meet the quality standards defined by Wadhvani AI.

A. Data Collection Volume

The project requires collecting approximately **9,000 worksheet pages** from 3000 students completing foundational numeracy assessments using a data collection mobile application

developed by Wadhvani AI. Each worksheet page will be captured as a **full-page image**, preserving the layout and answer boxes required for OCR model development.

The worksheets will be administered primarily in partner schools where permissions are available. If sufficient data cannot be collected from these schools, additional data may be collected from other schools in Uttar Pradesh, for which Wadhvani AI will facilitate permissions and provide the mobile application for data collection.

Each worksheet contains structured answer boxes where students write responses to numeracy questions such as addition, subtraction, and simple fill-in-the-blank number tasks using **Arabic numerals**.

The following table summarizes the expected data collection volume expected in 3 weeks:

Item	Requirement
Total worksheet pages	9000
Unit	Full A4 size page
Content	Handwritten responses
Script	Arabic Numerals

B. Operational Responsibilities and Training

The Partner must establish a reliable operational process to administer the worksheets and capture high-quality images of the completed assessments.

Training and Personnel:

The Partner is responsible for recruiting and training data collectors to administer worksheets and capture images according to the data collection protocol. Training should include instructions on worksheet administration, image capture standards (preserving fiducial marker integrity, ensuring proper lighting, capturing full page clearly, maintaining stability and focus), consent collection procedures, and basic quality checks.

Data Collection Protocol:

Data collectors will administer the worksheets to students and capture images of the completed pages using an android application developed by Wadhvani AI on mobile devices. The following is the acceptance criteria for the captured images:

- **Preserve marker integrity:** Do not fold, crease, tear, or cover the worksheet corners where markers are printed. Avoid writing, stamping, or placing stickers on markers.
- **Ensure proper lighting:** Use even, diffused lighting across the page. Avoid strong shadows, glare, or reflections that may reduce marker visibility.
- **Capture the full page clearly:** Ensure all four corners and markers are visible in the frame. Do not crop the image, and maintain a distance that keeps the entire page legible.
- **Maintain camera stability and focus:** Hold the camera steady to avoid motion blur and ensure the image is in focus before capturing.
- **Handle sheets correctly:** Place worksheets on a flat surface, avoid capturing while holding them in the air, and smooth out any folds or creases before taking the image.

Data collectors should **avoid instructing students to write their names or any personally identifiable information** on the worksheets wherever possible.

Consent and Data Privacy Compliance:

The Partner will be responsible for ensuring that **appropriate consent forms are obtained from schools and/or parents/guardians prior to data collection**. Signed consent forms must be **physically collected and submitted to the Wadhvani AI office** as part of the project documentation. The forms will be provided in a digital format by Wadhvani AI for partner organisation to save, get printed and signed.

The data collected through this project will be stored on Wadhvani AI systems and may be **shared with the respective schools participating in the study**, in accordance with applicable data governance policies and the **Digital Personal Data Protection (DPDP) Act, 2023**.

Wadhvani AI Support:

Wadhvani AI will provide the worksheet templates, android application for data collection and detailed guidelines for administering the assessments and capturing worksheet images. Wadhvani AI will also conduct initial alignment sessions to ensure that the partner understands the data collection requirements, consent procedures, and quality expectations.

C. Data Quality and Acceptance

All collected data will be subject to quality review by Wadhvani AI.

Image Quality:

Collected worksheet images must be clear, all corners should be visible, properly aligned, and free from excessive blur, shadows, or cropping that could interfere with OCR model development.

Review Process:

Wadhvani AI will review approximately **20% of the collected images** in batches of 300 to verify adherence to the data collection guidelines. Out of the 20% Wadhvani AI checks, at least 90% should match the image criteria. If not, the entire batch will be rejected.

Re-collection:

If a batch does not meet the required quality standards, the partner will be responsible for re-collecting the rejected data at no additional cost.

D. Data Ownership

All collected data will be stored on **Wadhvani AI provided clouds**. The data may be used for research and development of AI systems for educational assessment and may be shared with relevant stakeholders as part of program implementation.

5. DELIVERABLES

1. ~3000 students' data or 9000 pages of worksheet.
2. Data must pass the **WIAI review process** (20% recordings checked) as per the details above.

6. 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 hiring, training and capacity building of data collectors.
- Internal Quality Assurance (QA) mechanisms to ensure collected data meets required quality standards (e.g., all corners covered, clear picture).
- 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.

Submission Details

1. All proposals to this RFP must be received no later than <>. 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 Numeracy Tool in UP**.
2. Any proposals received by Wadhvani-AI after the deadline prescribed in the timeline of this document are liable to be rejected.

7. Evaluation Matrix

Criteria	Weightage
Understanding of Scope of Work and Objectives	10%
Detailed Operations/Implementation plan	30%
Team Composition & Capability	10%
Experience in collecting photo data, handwriting data preferred	30%
Data Quality Assurance & Risk Mitigation	15%
Data Policies	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 **image data** across a state
- Experience in training and operationalising 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 the use of images in training AI models and importance of good quality images
- 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 UP, and prior experience in collecting data through state's administrative mechanisms

MANDATORY DISCLOSURES

Please submit the following as enclosures or attachments with your quotation:

- Company profile
- Relevant client references
- Work completion reports, purchase orders
- GST registration certificate.
- PAN & TAN registration.
- MSME Certificate, if any
- Audited Financials or Income Tax Return with computation of Income for the last three years
- Cancelled cheque.

- Copy of registration documents/certificate and most recent renewal as a legal entity

This section is mandatory for consideration of the proposal.

Disclaimer: LEHS reserves the right to:

- Accept or reject any or all proposals
- Modify or cancel the RFP at any stage
- Seek clarifications without obligation to select any bidder

If bidders do not hear from LEHS within two weeks of RFP closure, the proposal may be considered unsuccessful.

