

REQUEST FOR PROPOSAL (RFP)

Data Annotation for Foundational Numeracy Tool in UP

Closing Date for Submission

<August 21, 2026>

1. INTRODUCTION

Wadhwani AI is the flagship Artificial Intelligence (AI) initiative of Lord's Education and Health Society, 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, LEHS - Wadhwani 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 LEHS - Wadhwani AI to identify a dedicated partner to execute a large-scale handwriting data annotation 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, LEHS - Wadhwani 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 annotation 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 annotation program to label handwritten student responses from paper-based numeracy assessments. The annotated 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 annotated 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: Hindu Arabic numerals (0–9) and simple mathematical responses/symbols

Age group of students: 7 - 9 years old

Time for completion: 2 weeks

4. SCOPE OF WORK

Data Annotation: Handwritten Numeracy Assessment Worksheets

The project involves annotation of handwritten student responses from paper-based numeracy assessments. The annotated dataset will be used to train and evaluate OCR models that recognize handwritten Hindu Arabic numerals and simple mathematical responses from structured answer boxes.

The annotation partner will be responsible for transcribing student responses from worksheet images using their **own annotation tool** and adhering to a **custom annotation schema**.

A. Data Annotation Volume

The project requires annotation of approximately **9,000 worksheet page images**. Each worksheet page will be treated as a **single annotation unit**, containing multiple questions with predefined answer boxes.

The worksheets include foundational numeracy problems such as addition, subtraction, and fill-in-the-blank tasks, where students write responses in structured boxes.

The following table summarizes the expected annotation volume in 2 weeks:

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

B. Annotation Responsibilities and Training

The Partner must establish a robust annotation workflow to ensure accurate and consistent labeling of handwritten responses.

Annotation Task:

Annotators will transcribe handwritten student responses from answer boxes into a structured format as per the **custom schema provided by LEHS - Wadhvani AI**. The annotators will be annotating both the answer and the answer box in which the answer is written.

Tooling:

Annotation will be performed using the company's **own annotation tool**. The partner is responsible for ensuring that annotators are trained to use the tool effectively.

Training and Onboarding:

The Partner is responsible for training annotators on:

- Using their own tool
- Understanding worksheet structure and answer boxes
- Adhering to the annotation schema and quality standards (Such as transcription guidelines and handling of ambiguous responses)

LEHS - Wadhvani AI Support:

LEHS - Wadhvani AI will provide:

- Detailed annotation guidelines
- Initial training sessions and calibration exercises

C. Data Quality and Acceptance

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

Accuracy Requirement:

The partner must ensure a **minimum annotation accuracy of 95%**.

Quality Control Process:

The partner is expected to implement internal quality assurance processes, which may include:

- multi-level review of annotations
- sampling-based quality checks
- correction workflows for identified errors

Review Process:

LEHS - Wadhvani AI will review approximately **20% of annotated data** in batches.

Rework Clause:

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

D. Deliverables

The partner will deliver:

- Annotated worksheet data following the **custom schema defined by LEHS - Wadhvani AI**
- Complete annotation of all assigned worksheet images
- Data exported or submitted through the **provided annotation tool**

E. Data Security and Ownership

All worksheet images and annotations will reside on **LEHS - Wadhvani AI systems/cloud**.

The partner must ensure:

- no unauthorized copying or storage of data outside approved systems
- restricted access to annotation data to authorized personnel only

All annotated data will be the property of **LEHS - Wadhvani AI** and may be used for research, model development, and program implementation purposes.

F. Timeline

The expected timeline for completion of the annotation task is **2 weeks** from the date of project kickoff, subject to final agreement.

5. Sample Annotation Exercise (Mandatory for Evaluation)

To assess the capability and accuracy of the applying organizations, a mandatory sample annotation exercise is required as part of the submission. We have provided four (4) filled sample worksheets. Applicants must annotate these worksheets according to our provided guidelines and include the output with their proposal. The accuracy and adherence to guidelines in this sample annotation will form a critical basis for selection.

- Detailed Annotation Guidelines: Link:
https://drive.google.com/drive/folders/1_2oNwrJfZKS1E6s93o0yN8AY_GEWw5Jw?usp=s
haring
- Attachment 1: Link:
<https://drive.google.com/file/d/1yW9IyV3uHiaBcB8e5mk3UCyFeLEsxhMr/view?usp=sharing>
- Attachment 2: Link:
https://drive.google.com/file/d/1IoDphiJg_IHQ2TxVJXt__NKczF4C4yt8/view?usp=sharing
- Attachment 3: Link: <https://drive.google.com/file/d/1j4Ahb9-pg-GYu9z5-Pny-uRPxSwpPsCF/view?usp=sharing>
- Attachment 4: Link:
https://drive.google.com/file/d/1ItFoOoKR31_fc4AAU3DLe54GaKTJ7GiW/view?usp=sharing

Kindly submit the annotations in JSON format as given in the Detailed Annotation Guidelines along with the proposal. If JSON format is not possible, then try using a lossless format like UBJSON, Bencode, INI, etc

6. TIMELINES & ROUNDS

Sr. No.	Stage	Milestone Date
1	RFP Issue	
2	Q&A Date	
3	Submission Deadline	

7. DELIVERABLES

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

8. 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 annotators.
- Internal Quality Assurance (QA) mechanisms to ensure collected data meets required quality standards (e.g., all parts annotated, rules followed)

2. Financial Proposal

- Detailed cost breakdown.
- Applicable taxes and contingencies.

3. Organizational Profile

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

Submission Details

1. All proposals to this RFP must be received no later than 21st August 2026. The proposal should be submitted only through e-mail in PDF format addressed to the Procurement Team at: rfp.lehs@wadhwanai.org. The email's subject line must contain the reference number and title of the RFP: **Data Annotation for Foundational Numeracy Tool in UP**.
2. Any proposals received after the deadline prescribed in the timeline of this document are liable to be rejected.

9. EVALUATION MATRIX

Criteria	Weightage
Understanding of Objectives	5%
Proposed approach, methodology & workplan aligning with objectives	30%
Quality Assurance & Risk Mitigation	15%
Team Composition & their credentials	10%
Credentials of Organization in annotating photo/image-based data, handwriting data	15%
Sample annotation	20%
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 annotation, especially **image data** across a state
- Experience in training and operationalizing a cadre of on-ground data annotators

- 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

10. MANDATORY ENCLOSURES

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

- Company profile
- Relevant client references
- Work completion reports, purchase orders
- GST registration certificate.
- PAN & TAN registration.
- MSME Certificate, if any
- Audited Financials for the last three years
- 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