



The Manager (Project Purchase)

Date: 13.10.2022

Open Tender Reference No: CS/MITES/049/2022/SPEECHDATA

GEM NAR ID: GEM/GARPTS/11102022/45578KHYKAFW

Due Date/Time: 02.11.2022@ 3:00 PM

Dear Sir/Madam,

On behalf of the Indian Institute of Technology Madras, offers are invited for

Speech Data Collection and Transcription: Major Indic languages

conforming to the specifications.

Tender Documents may be downloaded from Central Public Procurement Portal <https://etenders.gov.in/eprocure/app>. Aspiring Bidders who have not enrolled / registered in e-procurement should enroll / register before participating through the website <https://etenders.gov.in/eprocure/app>. The portal enrolment is free of cost. Bidders are advised to go through instructions provided at “**Help for contractors**”. [Special Instructions to the Contractors/Bidders for the e-submission of the bids online through this eProcurement Portal”]

Bidders can access tender documents on the website (For searching in the NIC site, kindly go to Tender Search option and type ‘IIT’. Thereafter, click on “GO” button to view all IIT Madras tenders). Select the appropriate tender and fill them with all relevant information and submit the completed tender document online on the website <https://etenders.gov.in/eprocure/app> as per the schedule attached.

1)	Pre-bid Meeting Details	:	NA
2)	ICSR Vendor Registration	:	<p><u>Vendor registration code.</u> Vendor registration with IC&SR (IITM) is mandatory for bidders to participate in tenders.</p> <p>** <u>For Vendor Registration & Guidelines, Please follow the website :</u> https://icandsr.iitm.ac.in/vendorportal; Helpdesk: vendorhelpdesk@icsrpis.iitm.ac.in</p>

No manual bids will be accepted. All tender documents including Technical and Financial bids should be submitted in the E-procurement portal.

Last date for receipt of tender	:	02.11.2022 @ 3:00 PM
Date & time of opening of tender	:	03.11.2022 @ 3:00 PM

3. Instructions to the Bidder:

<u>A)</u>	Searching for tender documents	:	<ul style="list-style-type: none"> • There are various search options built in the CPP Portal, to facilitate bidders to search active tenders by several parameters. These parameters could include Tender ID, organization name, location, date, value, etc. There is also an option of advanced search for tenders, wherein the bidders may combine a number of search parameters such as organization name, form of contract, location, date, other keywords etc. to search for a tender published on the CPP Portal. • Once the bidders have selected the tenders they are interested in, they may download the required documents / tender schedules. These tenders can be moved to the respective “My Tender” folder. This would enable the CPP Portal to intimate the bidders through SMS / email in case there is any corrigendum issued to the tender document. • The bidder should make a note of the unique Tender ID assigned to each tender, in case they want to obtain any clarification / help from the Helpdesk.
<u>B)</u>	Assistance to bidders	:	<ul style="list-style-type: none"> • Any queries relating to the tender document and the terms and conditions contained therein should be addressed to the Tender Inviting Authority for a tender or the relevant contact person indicated in the tender. • Any queries relating to the process of online bid submission or queries relating to CPP Portal in general may be directed to the 24x7 CPP Portal Helpdesk. The contact number for the helpdesk is [0120-4200462, 0120-4001002, 0120-4001005]
<u>C)</u>	Enrolment Process to Bidders	:	<p><u>REGISTRATION</u></p> <ul style="list-style-type: none"> • Bidders are required to enroll on the e-Procurement module of the Central Public Procurement Portal <u>URL:https://etenders.gov.in/eprocure/app</u> by clicking on “Online Bidder Enrolment”. Enrolment on the CPP Portal is free of charge. • As part of the enrolment process, the bidders will be required to choose a unique username and assign a password for their accounts. • Bidders are advised to register their valid email address and mobile numbers as part of the registration process. These would be used for any communication from the CPP Portal. • Upon enrolment, the bidders will be required to register their valid Digital Signature Certificate (Class II or Class III Certificates with signing key usage) issued by any Certifying Authority recognized by CCA India (e.g. Sify / TCS / nCode / eMudhra etc.) • Only one valid DSC should be registered by a bidder. Please note that the bidders are responsible to ensure that they do not lend their DSCs to others which may lead to misuse. • Bidder then may log in to the site through the secured log-in by entering their user ID / password and the password of the DSC / eToken.

			<ul style="list-style-type: none"> • Possession of a Valid Class II/III Digital Signature Certificate (DSC) in the form of smart card/e-token in the company's name is a prerequisite for registration and participating in the bid submission activities through https://etenders.gov.in/eprocure/app • Digital Signature Certificates can be obtained from the authorized certifying agencies, details of which are available in the web site https://etenders.gov.in/eprocure/app under the “Information about DSC”.
<u>D)</u>	Preparation of bids	:	<ul style="list-style-type: none"> • Bidder should take into account any corrigendum published on the tender document before submitting their bids. • Please go through the tender advertisement and the tender document carefully to understand the documents required to be submitted as part of the bid. Please note the number of covers in which the bid documents have to be submitted, the number of documents including the names and content of each of the document that need to be submitted. Any deviations from these may lead to rejection of the bid. • Bidder, in advance, should prepare the bid documents to be submitted as indicated in the tender document / schedule and generally shall be in PDF / XLS formats as the case may be. Bid documents may be scanned with 100 dpi with black and white option. • To avoid the time and effort required in uploading the same set of standard documents which are required to be submitted as a part of every bid, a provision of uploading such standard documents (e.g. PAN card copy, GSTIN Details, annual reports, auditor certificates etc.) has been provided to the bidders. Bidders can use “My Documents” area available to them to upload such documents. These documents may be directly submitted from the “My Documents” area while submitting a bid, and need not be uploaded again and again. This will lead to a reduction in the time required for bid submission process.
<u>E)</u>	Submission of bids	:	<ul style="list-style-type: none"> • Bidder should log into the site well in advance for bid submission so that he/she can upload the bid in time i.e. on or before the bid submission date and time. Bidder will be responsible for any delay due to other issues. • The bidder has to digitally sign and upload the required bid documents one by one as indicated in the tender document. • Bidder has to select the bid security declaration. Otherwise, the tender will be summarily rejected. • A standard BOQ format has been provided with the tender document to be filled by all the bidders. Bidders are requested to note that they should necessarily submit their financial bids in the format provided and no other format is acceptable. Bidders are required to download the BOQ file, open it and complete the detail with their respective financial quotes and other details (such as name of the bidder). If the BOQ file is found to be modified by the bidder, the bid will be rejected.

		<ul style="list-style-type: none"> • The server time (which is displayed on the bidders’ dashboard) will be considered as the standard time for referencing the deadlines for submission of the bids by the bidders, opening of bids etc. The bidders should follow this time during bid submission. • The Tender Inviting Authority (TIA) will not be held responsible for any sort of delay or the difficulties faced during the submission of bids online by the bidders due to local issues. • The uploaded tender documents become readable only after the tender opening by the authorized bid openers. • Upon the successful and timely submission of bids, the portal will give a successful bid submission message & a bid summary will be displayed with the bid no. and the date & time of submission of the bid with all other relevant details. • Kindly add scanned PDF of all relevant documents in a single PDF file of compliance sheet. • More information useful for submitting online bids on the CPP Portal may be obtained at: https://etenders.gov.in/eprocure/app. • All tender documents including pre-qualification bid, Technical Bid & Financial Bid should be submitted separately in online CPP portal as per the specified format only. Right is reserved to ignore any tender which fails to comply with the above instructions. No manual bid submission will be entertained.
F)	Marking on Pre-Qualification Bid	<ul style="list-style-type: none"> • The bidder eligibility criteria, technical specification and supply of item for this tender is given in Annexure A. • The Bidders shall go through the specification and submit the Pre-Qualification bid. • The Pre-qualification bid should be submitted in the proforma as per Annexure-B in pdf format only through online (e-tender). No manual submission of bid will be entertained. • The Pre-Qualification bid should have a page-wise heading as “Pre-qualification Bid” and page no. in all pages with seal and signature of authorized signatory. The total no. of pages should be mentioned at the last page of the documents
G)	Marking on Technical Bid	<ul style="list-style-type: none"> • The bidder eligibility criteria, technical specification and supply of item for this tender is given in Annexure A. • The Bidders shall go through the specification and submit the technical bid. • The Technical bid should be submitted in the proforma as per Annexure-C in pdf format only through online (e-tender). No manual submission of bid will be entertained. • The technical bid should have a page-wise heading as “Technical Bid” and page no. in all pages with seal and signature of authorized signatory. The total no. of pages should be mentioned at the last page of the documents. • The technical bid should consist of bidder eligibility criteria details and all technical details along with catalogue/ pamphlet which will give a detailed description of product with technical data sheet so

		that technical compliance can be verified.																							
H)	Marking on Price Bid	<ul style="list-style-type: none"> Financial bid (BoQ) should be submitted in the prescribed proforma format as per Annexure-D in xls format through e-tender only. No manual or other form of submission of Financial Bid will not be entertained 																							
4)	<p>EMD: Rs.10,00,000 /- (Rupees Ten Lakhs Only) has to be paid by means of bank transfer. Bank Account Details are given below. Proof of remittance has to be enclosed along with technical tender documents; else the bid will be rejected due to non-enclosure of EMD. This Earnest Money will be returned to the unsuccessful tenderers after the finalization of the tenders. EMD will not carry any interest. EMD will not be waived under any circumstances. EMD is exempted for Micro and Small Enterprises (MSE) as defined in MSE Procurement policy issued by the department of Micro, Small and Medium Enterprise (MSME). Relevant proof of registration of MSE with the competent authority should be submitted. For others, non-submission of EMD will lead to rejection of tender at the opening stage itself.</p> <table border="1"> <tr> <td>Institution Account Name</td> <td>The Registrar, Indian Institute of Technology - Madras</td> </tr> <tr> <td>Account No</td> <td>2722101003872</td> </tr> <tr> <td>Account Print Name</td> <td>The Registrar, IIT Madras</td> </tr> <tr> <td>IFSC CODE</td> <td>CNRB0002722</td> </tr> <tr> <td>Bank Name (in full)</td> <td>CANARA BANK</td> </tr> <tr> <td>Branch Name</td> <td>IIT -Madras Branch</td> </tr> <tr> <td rowspan="4">Complete Branch Address</td> <td>Canara Bank</td> </tr> <tr> <td>IIT -Madras Branch</td> </tr> <tr> <td>IIT -Madras Campus Post Office</td> </tr> <tr> <td>Sardar Patel Road</td> </tr> <tr> <td></td> <td>Guindy, Chennai - 600 036</td> </tr> <tr> <td>MICR No.</td> <td>600015085</td> </tr> <tr> <td>Account Type</td> <td>Savings Account</td> </tr> </table>		Institution Account Name	The Registrar, Indian Institute of Technology - Madras	Account No	2722101003872	Account Print Name	The Registrar, IIT Madras	IFSC CODE	CNRB0002722	Bank Name (in full)	CANARA BANK	Branch Name	IIT -Madras Branch	Complete Branch Address	Canara Bank	IIT -Madras Branch	IIT -Madras Campus Post Office	Sardar Patel Road		Guindy, Chennai - 600 036	MICR No.	600015085	Account Type	Savings Account
Institution Account Name	The Registrar, Indian Institute of Technology - Madras																								
Account No	2722101003872																								
Account Print Name	The Registrar, IIT Madras																								
IFSC CODE	CNRB0002722																								
Bank Name (in full)	CANARA BANK																								
Branch Name	IIT -Madras Branch																								
Complete Branch Address	Canara Bank																								
	IIT -Madras Branch																								
	IIT -Madras Campus Post Office																								
	Sardar Patel Road																								
	Guindy, Chennai - 600 036																								
MICR No.	600015085																								
Account Type	Savings Account																								
5)	<p>Performance Security: -</p> <p>The successful bidder should submit Performance Security for an amount of 3% of the value of the contract/supply. The Performance Security may be furnished in the form of an Account Payee DD, FD Receipt in the name of “The Registrar, IIT Madras” from any scheduled commercial bank or Bank Guarantee from any scheduled commercial bank in India. The performance security should be furnished within 14 days from the date of the purchase order.</p> <p>Performance Security in the form of Bank Guarantee: - In case the successful bidder wishes to submit Performance Security in the form of Bank Guarantee, the Bank Guarantee should be routed directly to IIT Madras from the Bank.</p> <p>The Bank Guarantee should remain valid for a period of sixty days beyond the date of completion of all contractual obligations of the supplier including the warranty obligations.</p>																								
6)	<p>Price:</p> <p>The offer/bid should be exclusive of taxes and duties, which will be paid by the purchaser as applicable. However, the percentage of tax & duties should be clearly indicated separately.</p>																								

7)	<p>Tenderer shall submit along with this tender:</p> <ul style="list-style-type: none"> (i) Name and full address of the Banker and their swift code and PAN No. and GSTIN number. (ii) GST registration proof showing registration number, area of registration etc. a) (iii) All of your future correspondences including Invoices should bear the GST No. and Area Code.
8)	<p>Terms of Delivery:</p> <p>The tenderer should clearly indicate the time of Delivery. In case there is any deviation in the delivery schedule, liquidated damages clause will be enforced or penalty for the delayed supply period will be levied.</p>
9)	<p>Period for which the offer will remain open:</p> <p>The Tender shall remain open for acceptance/validity till: 120 days from the date of opening of the tender. However, the day up to which the offer is to remain open being declared closed holiday for the Indian Institute of Technology Madras, the offer shall remain open for acceptance till the next working day.</p>
10)	<p>Risk Purchase Clause</p> <p>In the event of failure of provision of required services within the stipulated delivery schedule, the purchaser has all the right to get services from other sources on the total risk of the supplier under risk purchase clause.</p>
11)	<p>Payment:</p> <p>No Advance payment will be made for Indigenous purchase. Payment will be made 100% after the acceptance of completeness and accuracy of the said report by the Project Coordinator at IPM Cell, IITM.</p>
12)	<p>Acceptance and Rejection:</p> <p>Failure to comply with any of the instructions stated in this document or offering unsatisfactory explanations for non-compliance will likely to lead to rejection of offers.</p> <ul style="list-style-type: none"> (i) I.I.T. Madras has the right to accept the whole or any part of the Tender or portion of the quantity offered or reject it in full without assigning any reason.
13)	<p>Debarment from Bidding:</p> <p>In case of breach of Terms & Conditions, Bidder may be suspended from being eligible for bidding in any contract with the IIT Madras up to 2 Years [as per Rule 151(iii) of GFR] from the date of Tender.</p>
14)	<p>Disputes and Jurisdiction:</p> <p>Settlement of Disputes: Any dispute, controversy or claim arising out of or in connection with this PO including any question regarding its existence, validity, breach or termination, shall in the first instance be attempted to be resolved amicably by both the Parties. If attempts for such amicable resolution fails or no decision is reached within 30 days whichever is earlier, then such disputes shall be settled by arbitration in accordance with the Arbitration and Conciliation Act, 1996. Unless the Parties agree on a sole arbitrator, within 30 days from the receipt of a written request by one Party from the other Party to so agree, the arbitral panel shall comprise of three arbitrators. In that event, the supplier will nominate one arbitrator and the Project Coordinator of IITM shall nominate an arbitrator. The Dean IC&SR will nominate the Presiding Arbitrator of the arbitral tribunal. The arbitration proceeding shall be carried out in English language. The cost of arbitration and fees of the arbitrator(s) shall be shared equally by the Parties. The seat of arbitration shall be at IC&SR IIT Madras, Chennai.</p> <ul style="list-style-type: none"> a. The Applicable Law: The Purchase Order shall be construed, interpreted and governed by the Laws of India. Court at Chennai shall have exclusive jurisdiction subject to the arbitration

	<p>clause.</p> <p>Any legal disputes arising out of any breach of contract pertaining to this tender shall be settled in the court of competent jurisdiction located within the city of Chennai in Tamil Nadu.</p>
15)	<p>Force Majeure: The Supplier shall not be liable for forfeiture of its performance security, liquidated damages or termination for default, if and to the extent that, it's delay in performance or other failure to perform its obligations under the Contract is the result of an event of Force Majeure.</p> <p>For purposes of this Clause, "Force Majeure" means an event beyond the control of the Supplier and not involving the Supplier's fault or negligence and not foreseeable. Such events may include, but are not limited to, acts of the Purchaser either in its sovereign or contractual capacity, wars or revolutions, fires, floods, epidemics, quarantine restrictions and freight embargoes.</p> <p>b. If a Force Majeure situation arises, the Supplier shall promptly notify the Purchaser in writing of such conditions and the cause thereof. Unless otherwise directed by the Purchaser in writing, the Supplier shall continue to perform its obligations under the Contract as far as is reasonably practical, and shall seek all reasonable alternative means for performance not prevented by the Force Majeure event.</p>
16)	<p>Preference to "class I Local Suppliers": preference will be given to "class 1 local suppliers" (subject to class -I local supplier's quoted price falling within the margin of purchase preference) as per public procurement (preference to make in India) order 2017 .O.M No P- 45021/2/2017 – pp(BE - 11) dt 04/06/2020 subject to the conditions that the "class 1 Local Supplier" should agree to supply goods / provide service at L1 rate and furnish a certificate with the technical bid document that the goods/service provided by them consists local content equal to or more than 50%.(certificate from Chartered Accountant in case value of contract exceeds Rs 10 crore).</p> <p>**Note: Local content percentage to be calculated in accordance with the definition provided at clause 2 of revised public procurement preference to Make in India Policy vide GoI Order no. P-45021/2/2017-PP (B.E.-II) dated 15.06.2017 (subsequently revised vide orders dated 28.05.2018, 29.05.2019and 04.06.2020) MOCI order No. 45021/2/2017-PP (BE II) Dt.16th September 2020 & P- 45021/102/2019-BE-II-Part(1) (E-50310) Dt.4th March 2021</p>

<p>17</p>	<p>Preparation of Tender: The bidders should submit the bids in three bid system as detailed below.</p> <p>Cover I _Pre-Qualification Bid</p> <p>The Pre- Qualification Bid for in the Bidder Eligibility criteria I & II along with all relevant proof needs to be submitted in the proforma given in Annexure – B.</p> <p>The proof attached should be indexed and page numbers of the attached proof should be clearly mentioned in the Pre-Qualification Bid format.</p> <p>Cover II _Technical Bid</p> <p>The technical bid should consist of technical specification compliance sheet as per Annexure-C.</p> <p>Cover III _Price Bid</p> <p>The price bid should be submitted in excel format (BoQ) as per the proforma (Annexure D) uploaded in the e-Tender web site. The Quoted price should be for the services offered and inclusive of all cost and statutory levies at IIT Madras. The vendor should Quote the cost on Per Hour Basis as per the BOQ in Annexure D.</p>
<p>18</p>	<p>Bidder Eligibility Criteria I:</p> <ul style="list-style-type: none"> a) As per the Government of India Order, only “Class - I Local Suppliers” and “Class - II Local Suppliers” can participate in this tender. Declaration to be submitted as per Annexure - E b) Bidder should confirm their acceptance that they comply with the provisions with report to “Guidelines for eligibility of a bidder from a country which shares a land border with India as detailed at Annexure-F. The bidder should submit Certificate for “Bidder from/ Not from Country sharing Land border with India & Registration of Bidder with Competent Authority” as per Order of DoE F.No.6/18/2019-PPD dated 23.07.2020 as mentioned. c) Proof of Submission of EMD. d) Consent Form for the Bidder as per Annexure -G e) Declaration for Blacklisting as per Annexure - H <p>Bidder Eligibility Criteria II:</p> <ul style="list-style-type: none"> a) The Bidder’s firm should have existed for a minimum of 4 years. (Enclose Company Registration Certificate) b) The bidder should have done at least 1000 hours of speech data Collection and at least 500 hours of transcription as scope of work with reputed organizations in both Academia and Technology Corporate organizations. PO copies or work completion certificates should be

	<p>submitted as proof of the same. The above-mentioned projects should have been done for at least 10 of constitutionally recognized Indian languages (at least 100 hours of speech data collection and 50 hours of speech transcription per language) and languages of preference are Tamil, Malayalam, Kannada, Telugu, Bengali, Marathi, Gujarati, Hindi, Odia, Urdu and Sindhi. PO copies or work completion certificates should be submitted as a proof of the same.)</p> <p>c) The Bidder should submit at least 3 customer references from their previous work with the customer satisfaction qualitatively as excellent or higher on the bidder's capability to deliver the agreed scope and timelines and adhering to the required quality norms. Proof to be submitted in formal documentation (Customer Satisfaction) and signed/ approved by the customer.</p>
<p>19)</p>	<p>Evaluation of Bids</p> <p>Bid evaluation will take place in three stages.</p> <p>Stage I Pre-Qualification Bid evaluation</p> <p>The Pre- Qualification Bid of the bidders will be evaluated and the bidders who have qualified all the criteria stipulated in bidder eligibility criteria I & II will alone be selected and their technical bids will be opened.</p> <p>Stage II Technical Bid evaluation</p> <ol style="list-style-type: none"> 1. All bidders who qualify in the Stage I should Provide a detailed approach of implementation for data collection in Indian languages. 2. In addition, they would transcribe 1 hour of data per language for 10 Indian languages. They will be evaluated for transcription correctness. Upon publishing the Pre-Qualification result IITM will forward the Data for transcription to bidders through email with time duration of 3 days for submission of transcript data. 3. Stage 1 Qualified bidders need to make a power point presentation before an EPC Committee. Date for presentation will be informed to bidders through email. <p>The technical submission, Data transcription and technical presentation together will be considered for technical bid evaluation.</p> <p>The total marks for technical bid will be for 100 marks based on the weightage of each category, out of which the bidders should score minimum 70 marks to qualify for opening of their financial bid. The Financial bids of those bidders who have scored less than 70 marks will not be opened.</p>

The detailed scoring system for arriving at the total technical score of the bidders who have qualified Stage I are as

S.No.	Category	Criteria	Maximum Score within the category	Category weightage
1.a)	Bidder Credentials	<p>Empanelment as a Bidder for Data Collection (Voice and/or Non-Voice) in English or one of the Indian languages. This work has been done earlier with IIT Madras or one of other IITs or IISc or IIITs or NITs or other Central educational and research institutions.</p> <p>Non-Voice Data collection (80%)</p> <p>Voice Data Collection (20%)</p> <p>Proof of Empanelment (Work order or work completion certificate attached from premier education institute)</p> <p>Proposed Scoring criteria</p> <ul style="list-style-type: none"> • At least 1000 hours of data collection in 10 Indian languages and at least 100 hours per language - 20 • 1000 hours of data collection in Indian languages and 100 hours per language – 15 • Less than 1000 hours of data collection, but data collection experience in 10 Indian languages and at least 75 hours in each Indian language – 10 • Else 0. 	20	30%
1.b)		Number of Current or ongoing Projects carried out or being carried out in Educational and Research Domain and/or Deep Technology Digital organizations in Data collection	20	

		<p>for Natural language processing initiatives from 2019 onwards with size of data collection being at the size of 25% or higher of the size of this tender (approximate 2500 hours of voice data collection and 10K participants). Proof of work orders or Work in progress certificates from customers to be attached.</p> <p>Proposed scoring approach Greater than 25%: 17-20 Between 15 and 25% of the size of this tender: 12-16 Between 10 and 15% of the size of the tender: 8 – 11 Else 0</p>		
	1.c)	<p>24 hours of sample audio data collection in English and for at least 5 Indian languages for validation and rating for quality. 4 hours / Language</p> <p>The data would be uploaded by the bidder in cloud site to be provided by IITMadras and/or USB and to be delivered to the following address Nilekani Centre at AI4Bharat Department of Computer Science and Engineering, IIT Madras, PIN : 600036</p> <p>Proposed Scoring Approach Meets all the criteria specified in Key criteria section: 8 – 10 Meets 80% of the criteria specified in key criteria section: 6-8 Meets 60% of criteria specified in key criteria section: 4-6 Meets 50% of criteria specified in key criteria section: 1-3 Else 0.</p>	10	
	1d).	Empanelment as a Bidder for	25	

			<p>Transcription work done earlier with premier educational institutions in India (IITs, IISc, IIITs, NITs and other central government accredited premier institutes and universities).</p> <p>a) English (75% weightage) b) Indian Languages (25% Weightage)</p> <p>Proof of Empanelment (Work order or work completion certificate attached from premier education institute)</p> <p>Proposed Scoring criteria</p> <ul style="list-style-type: none"> • At least 5000 hours of voice data transcription done in English and at least 300 hours of data transcription done in at least in the 5 Indian languages. (Score 20 to 25) • At least 1000 to 5000 hours of voice data transcription in English and at least 100 hours of voice data transcription done in each of 5 Indian Languages (Score 15 to 20) • Less than 1000 hours of voice data transcription in English and at least 75 hours of voice data transcription each in 5 Indian languages (Score 		
--	--	--	---	--	--

			<p>10 to 15)</p> <ul style="list-style-type: none"> • Else 0 		
	1e).		<p>10 hours of audio transcription sample shared for each Indian language by the bidder along with audio sample – reviewed and rated on scale of 5. Rating is done based on key criteria outlined in this document.</p> <p>Proposed scoring approach</p> <ul style="list-style-type: none"> • Rating 5: 22-25 • Rating between 4 and 5: 17-22 • Rating between 3 and 4: 12-17 • Rating between 2 and 3: 8-12 • Else 0 	25	
	2a)	<p>Bidder– Project Management and Implementation capability</p>	<p>The organization has a certain proportion of the employees working out of rural or tier-2 or tier 3 towns workplaces/centers or do remote working with good knowledge of one or more Indian languages and with an ability to mobilize right participants for data collection. Proof submitted of offices in rural areas and employee strength in those offices</p> <p>Proposed Scoring Criteria</p>	10	50%

		<p>25% or higher number of employees work out of tier-2, tier-3 towns and /or rural location - 8- 10</p> <p>15 to 25% of employees working out of tier-2, tier-3 towns or rural areas: 6 to 10</p> <p>10 to 15% of employees working out of tier-2, tier-3 towns or rural areas: 4 to 6</p> <p>Less than 10% of employees working out of tier-2, tier-3 towns or rural areas: 1 - 3</p> <p>Else 0</p>		
	2b)	<p>Workflow / Process (Unit: Batch as defined in scope section of speech data collection, review, and approval).</p> <p>Bidder submits detailed workflow or process steps that would be followed in data collection (implementation phase)</p> <p>Proposed scoring criteria</p> <p>The workflow is in alignment with scope & key criteria defined in this document and is fairly detailed: 8 – 10</p> <p>The work flow has some alignment with scope and key criteria defined in this document and is to some extent detailed: 5-8</p> <p>The work flow is not very well aligned with scope and key criteria defined in this document and is not very detailed: 2-5</p>	10	

		Else 0	
2c)		<p>Team Structure (Team leads Profiles shared and committed to this project) with at least experience of 2 projects where they delivered on time and as per quality norms of data collection work in Indian Languages.</p> <p>The data collection team has an employee to consultant ratios of at least 70:30</p> <p>The bidder submits the team lead profiles along with official data that outlines their employee to consultant ratio is 70:30 or higher.</p> <p>Proposed Scoring criteria</p> <ol style="list-style-type: none"> 1. Team leads profiles shared and committed to project and employee to consultant ratio is 70: 30 or higher : 5 2. Else 0. 	10
2d)		<p>Schedule of the work and ability to deliver as per expected timelines (3 months for each language)</p> <p>Proposed Scoring Criteria</p> <p>The schedule is detailed enough and is realistic addressing risks /constraints: 8 – 10 (agile and balanced in terms of budget and bandwidth)</p> <p>The schedule is detailed to some extent only and ignores major risks / constraints : 5-8 (aggressive and may not factor in constraints in budget/bandwidth and associated risks)</p> <p>Else 0.</p>	10

	2e)		<p>Solution approach explores way to optimize overall costs of data collection while adhering to required quality norms. This specifically when the Bidder scope includes multiple Indian languages for which they would implement the data collection. Experience with using the Karya tool for data collection would be a big plus.</p> <p>Proposed Scoring Criteria</p> <ol style="list-style-type: none"> 1. Solution approach is detailed enough and demonstrates ability to optimize costs of data collection when done for multiple languages (8-10) 2. Solution approach has some detail and demonstrates only some confidence in optimizing the costs of data collection (5-8) 3. Else 0. 	10	
	2f)		<p>Workflow / Process (Unit: Transcribing, Reviewing, and editing of batch unit provided audio- or audio-visual content. Include solution approach to adhere to quality norms. Each batch is expected to be 200-400 hours per language.</p> <p>Proposed scoring criteria</p>	10	

		<p>The workflow is in alignment with scope & key criteria defined in this document for speech transcription and is fairly detailed: 8 – 10</p> <p>The work flow has some alignment with scope and key criteria defined in this document for speech transcription and is to some extent detailed: 6-8</p> <p>The work flow is not very well aligned with scope and key criteria defined in this document and is not very detailed: 1-5</p> <p>Else 0</p>		
	2g)	<p>The bidder would be given speech data samples to transcribe and submit the transcription. The transcription would be evaluated for adherence to key criteria mentioned in this document. This work would be at least 2 hours of speech data to be transcribed in 5 of the Indian languages in scope of this bid.</p> <p>Proposed Scoring Criteria:</p> <ol style="list-style-type: none"> 1. For all languages, transcription meets the desired criteria: 9-10 2. For 4 languages, transcription meets the desired criteria: 7-8 	10	

		<p>3. For 3 languages, transcription meets the desired criteria : 5-6</p> <p>4. Else 0.</p>	
2h)		<p>Workflow / Process (Unit: Transcribing, Reviewing, and editing of batch unit provided audio- or audio-visual content. Include solution approach to adhere to quality norms. Each batch is expected to be 200-400 hours per language.</p> <p>Proposed scoring criteria The workflow is in alignment with scope & key criteria defined in this document for speech transcription and is fairly detailed: 8 – 10</p> <p>The work flow has some alignment with scope and key criteria defined in this document for speech transcription and is to some extent detailed: 6-8</p> <p>The work flow is not very well aligned with scope and key criteria defined in this document and is not very detailed: 1-5</p> <p>Else 0</p>	10
2i)		<p>Team Structure (Team leads Profiles shared and committed to this project) with at least experience of 2 projects where they delivered on time and as per quality norms of</p>	5

		<p>transcription work, with size of content transcribed being 250hours or higher</p> <p>Their transcription team employee to consultant ratios is 70:30</p> <p>The bidder submits the team lead profiles along with official data that outlines their employee to consultant ratio is 70:30 or higher.</p> <p>Proposed scoring criteria</p> <ol style="list-style-type: none"> 1. The bidder submits the profiles of two team leads who have delivered transcription projects of size at least 250 hours each or higher and their employee to consultant ratio is 70: 30 or higher (8 to 10) 2. The bidder submits the profiles of two team leads who have delivered transcription projects of at least 150 to 250 hours each and their employee to consultant ratio is in range of 60:40 to 70:30 (5-8) 3. Else 0 		
	2j)	<p>Schedule of the work and ability to deliver as per expected timelines (3 to 5 months for each language) and overall timelines for all Indian languages in scope would be</p>	5	

			<p>for 15-18 months in total.</p> <p>Proposed Scoring Criteria The schedule is detailed enough and is realistic addressing risks /constraints: 8 – 10 (agile and balanced in terms of budget and bandwidth)</p> <p>The schedule is detailed to some extent only and ignores major risks / constraints: 1 – 8 (aggressive and may not factor in constraints in budget/bandwidth and associated risks)</p>		
	2k)		<p>Demonstration of technical tools used for transcription to improve accuracy and efficiency in the work done by the bidder earlier and how they propose to use them in improving agility and effectiveness of their proposed implementation approach</p> <p>Proposed scoring criteria The bidder demonstrates the use of tools and approach outlines with certain clarity on agility and effectiveness of transcription implementation : 4-5</p> <p>The bidder demonstrates the use of tools and approach and this outlines some clarity on agility and effectiveness of</p>	5	

		transcription implementation : 2-3 Else 0		
21)		Solution approach explores ways to optimize overall costs of transcription while adhering to required quality norms. This specifically when the Bidder bids for the scope that includes multiple Indian languages for which they would implement the transcription.	5	
3	Final Presentation (Data Collection and Transcription)	Bidder Presentation to committee – Approach and addressing key risks and dependencies 1. Why they need to be selected - USP 2. Their solution and delivery approach (Technical, PM) 3. Risks Identification and Mitigation 4. Value expected to be added to IIT Madras (Cost – Benefits)	100	20%
<p>Stage III: Price Bid Evaluation</p> <p>The price bid evaluation will be based on the rate quoted for each language by the bidder for the Collection of Speech Data Collection and Transcription: Major Indic languages . to arrival of Lowest Bid (L1) value in each language.</p>				
20)	<p>Selection of successful bidder and Award of Order</p> <p>The successful bidder will be selected based on assessment of skills, experience, and understanding/analysis of the project scope as per Technical Bid.</p>			

21)	<p>Apportion of work (Apportion by language) The L1 bidder for Each language will be awarded the contract.</p>															
22)	<p>Period of Service: The Collection of Speech Data in Indian Languages and Manually Transcription for building various Speech Technologies will be initially awarded for an estimated time 9575 – 11490 hours. Additional requirement may arise above 9575 – 11490 hours and the successful bidders shall extend the service on the same agreed rate.</p>															
23)	<p>Delivery and Payment Schedule:</p> <ul style="list-style-type: none"> • Delivery Schedule <ul style="list-style-type: none"> ○ Speech Data Collection <ul style="list-style-type: none"> ○ The speech data collection for each language except for Hindi, Telugu and Urdu should be completed in 3 months from date of start of work that is mutually agreed between IIT Madras and Bidder selected to deliver the work. <ul style="list-style-type: none"> ▪ 6-8 weeks: Planning and Preparation ▪ 4 weeks: Speech data collection for specific language. ○ For Hindi, Telugu and Urdu, the proposed timelines for data collection is 5-6 months from date of start of work that is mutually agreed between IIT Madras and Bidder selected to deliver the work ○ Speech Data Transcription <ul style="list-style-type: none"> ○ The transcription of speech data has to be completed at a rate of 100 hours of transcription completed, reviewed, and approved each month. ○ The indicative timelines for the work would be 3 to 6 months with languages such as Hindi and Urdu where data size / volume is higher expected to take 6 months. • Note: The indicative timelines for Speech data collection and Data transcription is expected to be around 3 to 5 quarters for all Indian languages in scope from Nov 2022 to January 2024 <p>The timeline for payment is as below.</p> <ol style="list-style-type: none"> 1. At the end of every month the data will be evaluated for the quality check by IITM. 2. The maximum acceptable data Errors is 5-7%. 3. No Payments will be made for the data with more than 15% Errors. <table border="1" data-bbox="261 1514 1476 1927"> <thead> <tr> <th data-bbox="261 1514 472 1612">Sl. No</th> <th data-bbox="472 1514 972 1612">Milestone</th> <th data-bbox="972 1514 1476 1612">Invoice (to be submitted by selected bidder)</th> </tr> </thead> <tbody> <tr> <td data-bbox="261 1612 472 1696">1</td> <td data-bbox="472 1612 972 1696">1 Month from Issue of PO</td> <td data-bbox="972 1612 1476 1696">15% of agreed Price with Bidder</td> </tr> <tr> <td data-bbox="261 1696 472 1787">2</td> <td data-bbox="472 1696 972 1787">40% of work is completed</td> <td data-bbox="972 1696 1476 1787">25% of agreed price with Bidder</td> </tr> <tr> <td data-bbox="261 1787 472 1864">3</td> <td data-bbox="472 1787 972 1864">70% of work is completed</td> <td data-bbox="972 1787 1476 1864">30% of agreed price with Bidder</td> </tr> <tr> <td data-bbox="261 1864 472 1927">4</td> <td data-bbox="472 1864 972 1927">100% of work is completed</td> <td data-bbox="972 1864 1476 1927">30% of agreed price with Bidder</td> </tr> </tbody> </table>	Sl. No	Milestone	Invoice (to be submitted by selected bidder)	1	1 Month from Issue of PO	15% of agreed Price with Bidder	2	40% of work is completed	25% of agreed price with Bidder	3	70% of work is completed	30% of agreed price with Bidder	4	100% of work is completed	30% of agreed price with Bidder
Sl. No	Milestone	Invoice (to be submitted by selected bidder)														
1	1 Month from Issue of PO	15% of agreed Price with Bidder														
2	40% of work is completed	25% of agreed price with Bidder														
3	70% of work is completed	30% of agreed price with Bidder														
4	100% of work is completed	30% of agreed price with Bidder														

	4. The total hours of work allocated, and completion will be inclusive of Data submitted with Errors.
24)	All information including selection and rejection of technical or financial bids of the prospective bidders will be communicated through e-Tender portal. In terms of Rule 173(iv) of General Financial Rule 2017, the bidder shall be at liberty to question the bidding conditions, bidding process and/or rejection of bids.
25)	The tenderer shall certify that the tender document submitted by him / her are of the same replica of the tender document as published by IIT Madras and no corrections, additions and alterations made to the same. If any deviation found in the same at any stage and date, the bid / contract will be rejected / terminated and actions will be initiated as per the terms and conditions of the contract.
26)	The bidders will not be entertained to participate in opening of Bids, since the tender is e-tender, the opening of the bids may be checked using the respective logins of the bidders.

ACKNOWLEDGEMENT

It is hereby acknowledged that I/We have gone through all the points listed under “Specification, Guidelines, Terms and Conditions” of tender document. I/We totally understand the terms and conditions and agree to abide by the same.

**SIGNATURE OF TENDERER ALONG WITH SEAL OF
THE COMPANY WITH DATE**

Scope of the Work.

Speech Data Collection and Transcription: Major Indic languages Tender No. CS/MITES/049/2022/SPEECHDATA

Deliverables

- The purpose of this tender is to identify bidders who will collect and deliver speech data from the field for the following Indian Languages. The scope also includes transcription of speech data collected from the field. The speech data volume collected is expected to be on an average, 25-30 hours per district (district is abbreviated as D below).
 - Tamil (38 D): Tamil 950 - 1140 hours
 - Malayalam (14D): Kerala 350-420 hours
 - Telugu (56D): AP and Telangana 1400 – 1680 hours
 - Kannada (31D): Karnataka 775 – 930 hours
 - Marathi (35D): Maharashtra 875 – 1050 hours
 - Bengali (23D): West Bengal 575-690 hours
 - Punjabi(23D): Orissa 575 – 690 hours
 - Gujarati (33D): Gujarat 825 – 990 hours
 - Hindi (at least 60D) 1500 – 1800 hours
 - Urdu (at least 60D) 1500 – 1800 hours
 - Sindhi – 250 – 300 hours

Total – 9575 – 11490 hours

1. **Bidders need to Mandatorily quote for all Districts in Each language .**
2. **Bidders need to collect Minimum 25 hrs and Maximum of 30 hours of Data per District in Each language**
 - a) **The minimum word vocabulary should be 400-550 words for 15 minutes of speech data collected from each participant.**
 - b) **Minimum No of speakers should be 10**
 - c) **And a Minimum time of 15 mins per speaker.**
 - d) **Data should be collected from both Male and Female equally**
3. **Agents should visit each district and collect data**
4. **Mode Of Delivery: In electronic form.**

Note: For certain languages like Hindi and Urdu, that is spoken across many states in India, the speech data collection and transcription would happen at least in 50 of districts across at least 3 states that has these languages as one of their primary languages. (for Eg, Hindi and Urdu – UP can be mandatory and other 2 states as optional)

Please note that implementation plan submitted by the bidder has to factor in a batch size for speech data collection and speech data transcription as per following guidelines.

- **Work Approach**

- **Speech Data Collection**

- Each district the speech data to be collected would be 25-30 hours and number of participants from whom the data will be collected would be at least 40 in each district.
 - The participant diversity would meet the following norms
 - Gender: 50% Male, 50% Female
 - Rural /Urban: As per 2011 census data at the state level
 - Age:
 - 18-30: 40%
 - 30-50: 35%
 - 50-60: 20%
 - 60+: 5%
 - Education Level
 - 0 – 5th Std: 10%
 - 6th to 12th Std: 10%
 - UG: 10%
 - Graduates and PG: 10%
 - Occupation
 - Regular: 25%
 - Gig work/Daily Wages: 25%
 - For each participant, the speech recording would be of around 15 minutes that will be organized into the following sections each around 5 minutes
 - Read Speech
 - Monologue
 - Dialog / Conversation

Note: IIT Madras would use the Karya App for speech recording and appropriate content/questions/prompts for use by participants in responding and recording of the 3 above mentioned sections of their speech. Preference will be given to bidders who have used the Karya app in the past. The Dialog / Conversation data will be collected over telephone channel (narrow band).

The Selected bidder should Mandatorily collect the Consent form from each participant prior to their participation in the data collection. The format for collection of Consent form from Participants will be shared to the Selected bidders at the time of award of Contract.

- **Speech Transcription**

- Listen to the audio-visual content and type contents verbatim and following the transcription guidelines provided.
 - Edit the typed content and make minor changes on a need basis for better understanding

- Review the transcribed text meets the guidelines provided by IIT Madras and also meets the quality norms prior to submission.

- **Key Criteria**

- **Speech Data Collection**

- The speech recording for each participant is around 15 minutes and meets quality norms.
 - Total number of samples collected for each language meets the diversity criteria prescribed.
 - Signed consent should be taken from all participants before collecting their voice sample with a clear indication that they give permission to IIT Madras and the sponsoring agency (Digital India Bhashini Mission, Ministry of Electronics and Information Technology, Government of India) under a permissible license such as the CC BY 4.0 license.
 - The speech recording meets the standards and specifications which will be specified in advance by IIT Madras (e.g., sampling rate 16KHz, .wav format, etc)
 - The recording would be reviewed and validated by the Bidder team post recording to ensure it meets quality norms prescribed and if not, re-record based on participant availability and convenience.
 - 20% of the total samples will be reviewed by IIT Madras team and it is expected that at least 95% of those samples adhere to the quality norms prescribed.
 - The quality norms are:
 - Sample recording quality should be similar to the quality of a reference sample provided by IIT Madras.
 - The speech should be clear and audible, despite any background noise.
 - For soft speakers, the audio volume must be adjusted to the right levels.
 - If the speaker speaks with a pronounced accent, the speech flow should be modulated in speed so that the words are clear and audible.

Note: The selected bidder's work would be internally reviewed by bidder's team and 20% of those samples by IIT Madras team and certified at batch level for each language, each batch would comprise of 200 to 400 participant samples and 50 to 100 hours of speech data collected across 5 districts.

- **Speech Data Transcription**

- The quality of transcription has an accuracy of 93-95% or higher
 - Adherence / conformance to guidelines is 93-95% or higher
 - A formal acceptance criteria would be discussed and agreed with the selected bidder once their bid is accepted by IIT Madras and a formal purchase order is initiated.
 - Note: If deliverables do not meet key criteria, rework must be done by the Bidder at no extra cost and as per revised schedule

PRE-QUALIFICATION BID PROFORMA
Tender No. EE/UME/40/2022/DATACOLLECT

Item Name: Speech Data Collection and Transcription: Major Indic languages

Bidder Eligibility Criteria:

I	Bidder Eligibility Criteria-I (Public Procurement – Preference to Make in India)	Class I / Class II	Local Content value	Reference, Page No.
a	Only 'Class-I local suppliers' and 'Class-II local suppliers', as defined under DIPP, MoCI Order No. P-45021/2/2017-PP (BE II) dated 16 th September 2020 and other subsequent orders issued therein. Declaration to be submitted as per Annexure - E			
b	Bidder should confirm their acceptance that they comply with the provisions with report to “Guidelines for eligibility of a bidder from a country which shares a land border with India as detailed at Annexure-F. The bidder should submit Certificate for “Bidder from/ Not from Country sharing Land border with India & Registration of Bidder with Competent Authority” as per Order of DoE F.No.6/18/2019-PPD dated 23.07.2020 as mentioned.			
c	Proof of Submission of EMD.			
d	Consent Form as per Annexure -G			
e	Declaration for Blacklisting as per Annexure - H			
II	Bidder Eligibility Criteria-II	Compliance (Yes/No)	Reference Page No.	Remarks, If any
a	The Bidder’s firm should have existed for a minimum of 4 years. (Enclose Company Registration Certificate)			
b	The bidder should have done at least 1000 hours of speech data Collection and at least 500 hours of transcription as scope of work with reputed organizations in both Academia and Technology Corporate organizations. PO copies or work completion certificates should be submitted as proof of the same. The above-mentioned projects should have been done for at least 10 of constitutionally recognized Indian languages (at least 100 hours of speech data collection and 50 hours of speech transcription per language) and languages of preference are Tamil, Malayalam, Kannada, Telugu, Bengali, Marathi, Gujarati, Hindi. Odia, Urdu and Sindhi. PO copies or work completion certificates should be submitted as a proof of the same.)			
c	The Bidder should submit at least 3 customer references from their previous work with the customer satisfaction qualitatively as excellent or higher on the bidder’s capability to deliver the agreed scope and timelines and adhering to the required quality norms. Proof to be submitted in formal documentation (Customer Satisfaction) and signed/ approved by the customer.			

TECHNICAL BID PROFORMA
Tender No. CS/MITES/049/2022/SPEECHDATA

Item Name: Speech Data Collection and Transcription: Major Indic languages

The Bidders should submit a detailed proposal for the scope of works of this tender covering the proposed,

S.No.	Category	Criteria	COMPLIED /NOT COMPLIED	Ref. Page No.
1.a)	Bidder Credentials	<p>Empanelment as a Bidder for Data Collection (Voice and/or Non-Voice) in English or one of the Indian languages. This work has been done earlier with IIT Madras or one of other IITs or IISc or IITs or NITs or other Central educational and research institutions.</p> <p>Non-Voice Data collection (80%) Voice Data Collection (20%)</p> <p>Proof of Empanelment (Work order or work completion certificate attached from premier education institute)</p> <p>Proposed Scoring criteria</p> <ul style="list-style-type: none"> • At least 1000 hours of data collection in 10 Indian languages and at least 100 hours per language - 20 • 1000 hours of data collection in Indian languages and 100 hours per language – 15 • Less than 1000 hours of data collection, but data collection experience in 10 Indian languages and at least 75 hours in each Indian language – 10 • Else 0. 		
1.b)		<p>Number of Current or ongoing Projects carried out or being carried out in Educational and Research Domain and/or Deep Technology Digital organizations in Data collection for Natural language processing initiatives from 2019 onwards with size of data collection being at the size of 25% or higher of the size of this tender (approximate 2500 hours of voice data collection and 10K participants). Proof of work orders or Work in progress certificates from customers to be attached.</p> <p>Proposed scoring approach Greater than 25%: 17-20 Between 15 and 25% of the size of this tender: 12-16 Between 10 and 15% of the size of the tender: 8 – 11 Else 0</p>		
1.c)		<p>24 hours of sample audio data collection in English and for at least 5 Indian languages for validation and rating</p>		

		<p>for quality. 4 hours / Language</p> <p>The data would be uploaded by the bidder in cloud site to be provided by IITMadras and/or USB and to be delivered to the following address Nilekani Centre at AI4Bharat Department of Computer Science and Engineering, IIT Madras, PIN : 600036</p> <p>Proposed Scoring Approach Meets all the criteria specified in Key criteria section: 8 – 10 Meets 80% of the criteria specified in key criteria section: 6-8 Meets 60% of criteria specified in key criteria section: 4-6 Meets 50% of criteria specified in key criteria section: 1-3 Else 0.</p>		
1d).		<p>Empanelment as a Bidder for Transcription work done earlier with premier educational institutions in India (IITs, IISc, IIITs, NITs and other central government accredited premier institutes and universities).</p> <p>c) English (75% weightage) d) Indian Languages (25% Weightage)</p> <p>Proof of Empanelment (Work order or work completion certificate attached from premier education institute)</p> <p>Proposed Scoring criteria</p> <ul style="list-style-type: none"> • At least 5000 hours of voice data transcription done in English and at least 300 hours of data transcription done in at least in the 5 Indian languages. (Score 20 to 25) • At least 1000 to 5000 hours of voice data transcription in English and at least 100 hours of voice data transcription done in each of 5 Indian Languages (Score 15 to 20) • Less than 1000 hours of voice data transcription in English and at least 75 hours of voice data transcription each in 5 Indian languages (Score 10 to 15) 		

		<ul style="list-style-type: none"> • Else 0 		
1e).		<p>10 hours of audio transcription sample shared for each Indian language by the bidder along with audio sample – reviewed and rated on scale of 5. Rating is done based on key criteria outlined in this document.</p> <p>Proposed scoring approach</p> <ul style="list-style-type: none"> • Rating 5: 22-25 • Rating between 4 and 5: 17-22 • Rating between 3 and 4: 12-17 • Rating between 2 and 3: 8-12 • Else 0 		
2a)	Bidder– Project Managem ent and Implementa tion capability	<p>The organization has a certain proportion of the employees working out of rural or tier-2 or tier 3 towns workplaces/centers or do remote working with good knowledge of one or more Indian languages and with an ability to mobilize right participants for data collection. Proof submitted of offices in rural areas and employee strength in those offices</p> <p>Proposed Scoring Criteria</p> <p>25% or higher number of employees work out of tier-2, tier-3 towns and /or rural location - 8- 10</p> <p>15 to 25% of employees working out of tier-2, tier-3 towns or rural areas: 6 to 10</p> <p>10 to 15% of employees working out of tier-2, tier-3 towns or rural areas: 4 to 6</p> <p>Less than 10% of employees working out of tier-2, tier-3 towns or rural areas: 1 - 3</p> <p>Else 0</p>		

2b)		<p>Workflow / Process (Unit: Batch as defined in scope section of speech data collection, review, and approval).</p> <p>Bidder submits detailed workflow or process steps that would be followed in data collection (implementation phase)</p> <p>Proposed scoring criteria</p> <p>The workflow is in alignment with scope & key criteria defined in this document and is fairly detailed: 8 – 10</p> <p>The work flow has some alignment with scope and key criteria defined in this document and is to some extent detailed: 5-8</p> <p>The work flow is not very well aligned with scope and key criteria defined in this document and is not very detailed: 2-5</p> <p>Else 0</p>		
2c)		<p>Team Structure (Team leads Profiles shared and committed to this project) with at least experience of 2 projects where they delivered on time and as per quality norms of data collection work in Indian Languages.</p> <p>The data collection team has an employee to consultant ratios of at least 70:30</p> <p>The bidder submits the team lead profiles along with official data that outlines their employee to consultant ratio is 70:30 or higher.</p> <p>Proposed Scoring criteria</p> <ol style="list-style-type: none"> 1. Team leads profiles shared and committed to project and employee to consultant ratio is 70: 30 or higher : 5 2. Else 0. 		
2d)		<p>Schedule of the work and ability to deliver as per expected timelines (3 months for each language)</p> <p>Proposed Scoring Criteria</p> <p>The schedule is detailed enough and is realistic addressing risks /constraints: 8 – 10 (agile and balanced in terms of budget and bandwidth)</p> <p>The schedule is detailed to some extent only and ignores major risks / constraints : 5-8 (aggressive and may not factor in constraints in budget/bandwidth and associated risks)</p> <p>Else 0.</p>		

2e)		<p>Solution approach explores way to optimize overall costs of data collection while adhering to required quality norms. This specifically when the Bidder scope includes multiple Indian languages for which they would implement the data collection. Experience with using the Karya tool for data collection would be a big plus.</p> <p>Proposed Scoring Criteria</p> <ol style="list-style-type: none"> 1. Solution approach is detailed enough and demonstrates ability to optimize costs of data collection when done for multiple languages (8-10) 2. Solution approach has some detail and demonstrates only some confidence in optimizing the costs of data collection (5-8) 3. Else 0. 		
2f)		<p>Workflow / Process (Unit: Transcribing, Reviewing, and editing of batch unit provided audio- or audio-visual content. Include solution approach to adhere to quality norms. Each batch is expected to be 200-400 hours per language.</p> <p>Proposed scoring criteria</p> <p>The workflow is in alignment with scope & key criteria defined in this document for speech transcription and is fairly detailed: 8 – 10</p> <p>The work flow has some alignment with scope and key criteria defined in this document for speech transcription and is to some extent detailed: 6-8</p> <p>The work flow is not very well aligned with scope and key criteria defined in this document and is not very detailed: 1-5</p> <p>Else 0</p>		

2g)		<p>The bidder would be given speech data samples to transcribe and submit the transcription. The transcription would be evaluated for adherence to key criteria mentioned in this document. This work would be at least 2 hours of speech data to be transcribed in 5 of the Indian languages in scope of this bid.</p> <p>Proposed Scoring Criteria:</p> <ol style="list-style-type: none"> 5. For all languages, transcription meets the desired criteria: 9-10 6. For 4 languages, transcription meets the desired criteria: 7-8 7. For 3 languages, transcription meets the desired criteria : 5-6 8. Else 0. 		
2h)		<p>Workflow / Process (Unit: Transcribing, Reviewing, and editing of batch unit provided audio- or audio-visual content. Include solution approach to adhere to quality norms. Each batch is expected to be 200-400 hours per language.</p> <p>Proposed scoring criteria</p> <p>The workflow is in alignment with scope & key criteria defined in this document for speech transcription and is fairly detailed: 8 – 10</p> <p>The work flow has some alignment with scope and key criteria defined in this document for speech transcription and is to some extent detailed: 6-8</p> <p>The work flow is not very well aligned with scope and key criteria defined in this document and is not very detailed: 1-5</p> <p>Else 0</p>		

2i)		<p>Team Structure (Team leads Profiles shared and committed to this project) with at least experience of 2 projects where they delivered on time and as per quality norms of transcription work, with size of content transcribed being 250hours or higher</p> <p>Their transcription team employee to consultant ratios is 70:30</p> <p>The bidder submits the team lead profiles along with official data that outlines their employee to consultant ratio is 70:30 or higher.</p> <p>Proposed scoring criteria</p> <ol style="list-style-type: none"> 1. The bidder submits the profiles of two team leads who have delivered transcription projects of size at least 250 hours each or higher and their employee to consultant ratio is 70: 30 or higher (8 to 10) 2. The bidder submits the profiles of two team leads who have delivered transcription projects of at least 150 to 250 hours each and their employee to consultant ratio is in range of 60:40 to 70:30 (5-8) 3. Else 0 		
2j)		<p>Schedule of the work and ability to deliver as per expected timelines (3 to 5 months for each language) and overall timelines for all Indian languages in scope would be for 15-18 months in total.</p> <p>Proposed Scoring Criteria</p> <p>The schedule is detailed enough and is realistic addressing risks /constraints: 8 – 10 (agile and balanced in terms of budget and bandwidth)</p> <p>The schedule is detailed to some extent only and ignores major risks / constraints: 1 – 8 (aggressive and may not factor in constraints in budget/bandwidth and associated risks)</p>		

2k)		<p>Demonstration of technical tools used for transcription to improve accuracy and efficiency in the work done by the bidder earlier and how they propose to use them in improving agility and effectiveness of their proposed implementation approach</p> <p>Proposed scoring criteria The bidder demonstrates the use of tools and approach outlines with certain clarity on agility and effectiveness of transcription implementation : 4-5</p> <p>The bidder demonstrates the use of tools and approach and this outlines some clarity on agility and effectiveness of transcription implementation : 2-3</p> <p>Else 0</p>		
2l)		<p>Solution approach explores ways to optimize overall costs of transcription while adhering to required quality norms. This specifically when the Bidder bids for the scope that includes multiple Indian languages for which they would implement the transcription.</p>		
3	Final Presentation (Data Collection and Transcription)	<p>Bidder Presentation to committee – Approach and addressing key risks and dependencies</p> <ol style="list-style-type: none"> 5. Why they need to be selected - USP 6. Their solution and delivery approach (Technical, PM) 7. Risks Identification and Mitigation 8. Value expected to be added to IIT Madras (Cost – Benefits) 		

SIGNATURE OF BIDDER ALONG WITH SEAL OF THE COMPANY WITH DATE

FINANCIAL BID (PROFORMA) - BILL OF QUANTITIES (BOQ)

Item Name: Speech Data Collection and Transcription: Major Indic languages
Tender No. CS/MITES/049/2022/SPEECHDATA

It. No	Description of work	Quantity	Units	Basic Rate/Hour in INR	GST %	Total Amount with taxes in INR
1	Speech Data Collection and Transcription: Major Indic languages Listed below. Value to be quoted for per Hr basis for all districts in Each language					
	▪ Tamil (38 D): Tamil	1	Hrs.			
	▪ Malayalam (14D): Kerala	1	Hrs.			
	▪ Telugu (56D): AP and Telangana	1	Hrs.			
	▪ Kannada (31D): Karnataka	1	Hrs.			
	▪ Marathi (35D): Maharashtra	1	Hrs.			
	▪ Bengali (23D): West Bengal	1	Hrs.			
	▪ Punjabi (23 D): Punjab	1	Hrs.			
	▪ Gujarati (33D): Gujarat	1	Hrs.			
	▪ Hindi (at least 60D)	1	Hrs.			
	▪ Urdu (at least 60D)	1	Hrs.			
	▪ Sindhi – 250 – 300 hours	1	Hrs.			
	Grand Total					

Total Amount Rupees in words _____

Note: Financial bid (BoQ- Excel File) should be uploaded in the eProcurement portal, financial bid column only.

**FORMAT FOR AFFIDAVIT OF SELF-CERTIFICATION UNDER PREFERENCE TO MAKE IN INDIA –
PER ITEM**

Tender Reference Number:

Name of the item / Service:

Date: _____
I/We _____ S/o, D/o, W/o, _____
Resident of _____

Hereby solemnly affirm and declare as under:

That I will agree to abide by the terms and conditions of the Public Procurement (Preference to Make in India) Policy vide GoI Order no. P-45021/2/2017-PP (B.E.-II) dated 15.06.2017 (subsequently revised vide orders dated 28.05.2018, 29.05.2019 and 04.06.2020) MOCI order No. 45021/2/2017-PP (BE II) Dt.16th September 2020 & P- 45021/102/2019-BE-II-Part (1) (E-50310) Dt.4th March 2021 and any subsequent modifications/Amendments, if any and

That the local content for all inputs which constitute the said item/service/work has been verified by me and I am responsible for the correctness of the claims made therein.

Tick (✓) and Fill the Appropriate Category	
<input type="checkbox"/>	I/We _____ [name of the supplier] hereby confirm in respect of quoted items that Local Content is equal to or more than 50% and come under “Class-I Local Supplier” category.
<input type="checkbox"/>	I/We _____ [name of the supplier] hereby confirm in respect of quoted items that Local Content is equal to 20% but less than 50% and come under “Class-II Local Supplier” category.

- The details of the location (s) at which the local value addition is made and the proportionate value of local content in percentage

Address _____ Percentage of Local content: _____%

For and on behalf of (Name of firm/entity)

Authorized signatory (To be duly authorized by the Board of Directors)

<Insert Name, Designation and Contact No.>

[Note: In case of procurement for a value in excess of Rs. 10 Crores, the bidders shall provide this certificate from statutory auditor or cost auditor of the company (in the case of companies) or from a practicing cost accountant or practicing chartered accountant (in respect of suppliers other than companies) giving the percentage of local content.]

This letter should be on the letterhead of the quoting firm and should be signed by a competent authority. Non-submission of this will lead to Disqualification of bids.

(To be given on the letter head of the bidder)

Tender Ref. No. _____

Dated: _____

CERTIFICATE

(Bidders from India)

I have read the clause regarding restrictions on procurement from a bidder of a country which shares a land border with India and hereby certify that I am not from such a country.

OR (*whichever is applicable*)

(Bidders from Country which shares a land border with India)

I have read the clause regarding restrictions on procurement from a bidder of a country which shares a land border with India and hereby certify that I from _____ (Name of Country) and has been registered with the Competent Authority. I also certify that I fulfil all the requirements in this regard and is eligible to be considered. *(Copy/ evidence of valid registration by the Competent Authority is to be attached)*

Place:

Date:

Signature of the Tenderer
Name & Address of the
Tenderer with Office Stamp

(To be given on the letter head of the bidder)

Tender Ref. No. _____

Dated: _____

TO WHOMSOEVER IT MAY CONCERN

In response of IIT Madras’s Bid for the “Call For Collaborative Participation in Collection of Speech Data in Indian Languages and Manual Transcription for building various Speech Technologies”, I _____ (Representative of the Company/Organization) on behalf of _____ (Name of the company/organization), bearing CIN. _____ (preferred), having registered office at _____ do hereby give my unconditional irrevocable consent to IIT Madras, to collect, retain and use in various ways, the transcribed speech data and/or audio recorded on _____ (date of recording).

Further, in exchange of sufficient consideration, _____(name of the organization/company) do unconditionally irrevocably assigns all its rights and/or ownership in the transcribed speech data, recorded video and audio and provide a lifetime access to IIT Madras.

_____ (name of the company/organization) shall be personally responsible for the validity of the rights and ownership of the data collected and shall hold IITM harmless from any complaint, infringement proceedings, and/or any claim made by a third-party.

_____ (name of the company/organization) do understand that IIT Madras can lawfully use the transcribed speech data, recorded video and audio.

Signature:

Name:

Place :

Date:

Self-Declaration that the Service Provider has not been blacklisted

(To be given on the letter head of the bidder)

I

.....

..... S/o R/o

..... police station District

..... Director/ partner/ sole proprietor (Strike out whichever is not applicable) of

..... (Firm or Company) do hereby declare and solemnly affirm:

- i. That the Firm has not been Blacklisted or declared insolvent by any of the Union or State Government / Organization.
- ii. That none of the individual / firm / Company Blacklisted or any partners or shareholder thereof has any connection directly or indirectly with or has any subsistence interest in the deponent business / firm company.
- iii. That neither the Firm nor any of its partner has been involved / convicted in any criminal case / economic offence nor any criminal case / economic offence is pending against firm or any partner of the Firm before any Court of Law / Police.

Place:

Signature of the Tenderer

Date:

Name & Address of the Tenderer with Office Stamp