

**Call for Research & Development Proposals under National Language Translation Mission
(NLTM)**

**Ministry of Electronics and IT (MeitY)
Government of India**

National Language Translation Mission (NLTM)

The government has been proactively promoting digitization to ensure that public services are made available electronically to citizens, even in the farthest corners of the country. The future lies in the uplifting of digital initiatives to national public digital platforms, which are simple, nimble and bring together existing systems, players and open the doors for innovation and value-added services roping in the public and private ecosystem players. Many such national public digital platforms viz. Aadhaar, UPI, GeM, GSTN, NDHM etc. are built using digital technologies including Artificial Intelligence, Machine Learning, Data Analytics and Automation. These platforms are built using open data, open source code and open APIs to encourage open and secure ecosystems, plus allowing the private sector and start-ups to innovate and bring value-added services to the citizens. This approach has ultimately made digital services more accessible. A similar approach has been adopted by the National Language Translation Mission (NLTM) which aims to build Speech to Speech Machine Translation and evolve a Unified Language Interface (ULI) for translation of Indian Languages in the next 3 years to bring together multiple efforts towards Indic Language Technologies.

The current approach aims to provide a free and open-source **National Public Digital Platform** to encourage open ecosystems, allow private sector innovation to provide improved services and products in the domain of language technologies, *in order to enable citizens to access digital services including educational content in their native language to further increase digital inclusion and accessibility*. The platform architecture consists of a Data & Model Repository layer which will be accessed through standard API - ULCA (Unified Language Contribution API), Foundation layer consisting of benchmarking tools, Leaderboards etc, Reference Application Layer and Ecosystem layer facilitating the startups/industry engagement. The equal participation of State Governments and other central ministries in the implementation of Language/State Missions is the key to strengthen the partnership based model to develop the platform and the platform acting as an orchestrator to bring together language technologies stakeholders for the aims of the mission.

For more details regarding National Language Translation Mission (NLTM), please refer to **NLTM White Paper at [Annexure](#)**.

1.0 Research & Development Proposals

Academia and research groups shall be responsible for carrying out research in the areas of language technologies and translation, covering a set of languages out of the current 22 scheduled languages. Since fundamental research with a long term horizon is important for the success of Mission, the research groups will be required to submit their proposals for the same, clearly outlining all objectives and deliverables. The projects will have short term (quarterly) and medium term (yearly) deliverables.

All research proposals shall be evaluated by an expert committee based on the research rubric in a fair and transparent manner.

Research proposals may often require data collection. The nature and amount of data to be collected should be justified taking into account the state of the art and available datasets on Bhashini platform. It is possible that the amount or nature of data to be collected changes due to technological trends. All data collected must be compliant with ULCA processes defined by the Data Management Unit under National Hub for Language Technology (NHLT), DIC to allow sharing and reuse of data with the ecosystem at the earliest possible following standard processes.

Models developed as part of the Government funded research projects must aim to improve open source baseline benchmark results that will be published on Bhashini platform and updated from time to time to reflect the state of the art.

Bhashini shall also allow submission of proposals for short term projects (upto 1 year), as the same may be required to make rapid progress on Bhashini. Central and state funded institutes' led consortia shall also be allowed to participate.

The output of all research proposals that are funded by the Government must be open sourced and ULCA compliant. It is further required that models developed by research must be made available as web (REST) service implementing the ULCA open API. Research proposals should include engineering resources to make this possible.

The process of providing grant for research would follow process mentioned below:

1.1 Initiation

Call for Proposals: A time window of 4 weeks shall be available to all institutions to submit **online** closed research and development proposals using MeitY's standard format to submit research proposals. This window would be opened 3 times yearly as per the budget availability.

1.2 Duration of project: The maximum duration of the project would be 3 years. However, based on the duration, the projects may be termed as:

1. Short Term Projects - up to 12 months
2. Medium Term Projects (1-2 years),
3. Long Term Project (3 years)

Note:

1. NLTM's Executive Committee at its discretion may recommend for increased/reduced duration of the project depending upon the nature of the research study as per research rubric.
2. Considering the disbursement schedule of NLTM, the disbursement may go beyond the duration of the project(s).

1.3 Eligibility:

Academic Institute, University, Research Organizations, registered/ accredited by a government can apply either alone, or in partnership with Startup(s)/Institute(s). The institution would be the primary applicant. The Project Investigator (PI) has to be a permanent faculty of the applicant entity. A co-PI of matching capability should also be identified for seamless shift of responsibility of the project in case of any unforeseen circumstances.

1.4 Evaluation

The Research proposals will be approved by the executive committee based on rubric defined by Bhashini following below mentioned principles:

- Research proposals may define engineering costs related to any research deliverable (data, code, models, utilities, applications, workbench, pipelines etc.) as long as it is open source to allow the entire ecosystem to benefit. The research proposal must be relevant to Bhashini's goals.

- It must create a novel capability that has hitherto not been published or improve the baseline performance of existing capability. Multiple baseline shall be updated as state of the art evolves and new benchmarks are added at the platform by NHLT. **Till baseline of the NLTM Pilot Project and ULCA are published on the Bhashini platform, the research proposal can be submitted with an undertaking that the research project will follow NLTM approved baseline and improve upon that as part of the research project and explain in detail the baseline adopted in lieu of NLTM baseline.**
- Research proposals must include GPU cloud credits and not stand alone GPU hardware.
- The researchers must highlight their previous research contributions and experience in the field and engineering capability in the proposal.
- The research proposals submitted where the model(s) have been benchmarked will add more weightage to the proposal.
- Factors like data requirement, submitted model performance, future model performance, model size, compute , duration of research, cost, H-index of the team, past performance of projects etc will be considered during evaluation.
- The output of this research, including data, tools, utilities, pipeline, workbench, models shall be open source and made available for the entire community
- Data collection of types deviating from the already specified data types/specifications should be accompanied by the justification of the need for such data in general or for their specific project. Once approved by the expert group the data collection will be managed by MeitY following the standard procedure.
- Any data collected as part of the project will follow the ULCA process for collection and be compliant with the ULCA API (API should be extended if it does not support the data desired to be collected) and data collection requirements will be floated on GeM (or other e-Marketplace) by TDIL/MeitY.
- Any model submitted as part of the research must be packaged as a web service so that it can be automatically benchmarked. This approach will ensure near production quality deliverables. The engineering resources for this purpose may be hired through appropriate channels to allow for high quality resources at market costs.

The Proposal may also include applications where these models could be utilised by startups to aid the goals of Mission Bhashini.

1.5. Awarding

Based on the above factors, the Executive Committee (EC) will approve a research proposal. EC may also like to add specific terms and conditions for such funding, if required, over and above the standard terms and conditions of MeitY such as source code and data of all outcomes of the project will be made available in open source etc.

1.6. Funds Disbursement:

NLTM provides a contribution of up to 100% of project cost for research proposals. The support to the academia will be given as Grant-in-aid for the research. Contribution of funding from participating organisations/concerned State Governments/user agencies etc. will be a criteria in the Research Rubric for preferential treatment.

Instalment of the projects will be based on completion of clearly laid out milestones aligned with objectives of the mission on quarterly/6-monthly basis. In case of projects of duration 1 year or less, the milestones must be on quarterly basis, else they can be on quarterly/6-monthly basis. The fund release may also be on a pro-rata basis for partly completion of milestones and recommendation of the Executive Committee. The Milestones should also be correlated with projected benchmarks which could be verifiable through the Bhashini Platform.

1.7. Monitoring

The research progress would be monitored by PMU/EEU on a monthly/quarterly basis either through telecon/email/visits or any other suitable mechanism and a report will be required to be submitted to NLTM/TDIL/MeitY in a timely manner. EC will also hold quarterly Project Review Meetings.

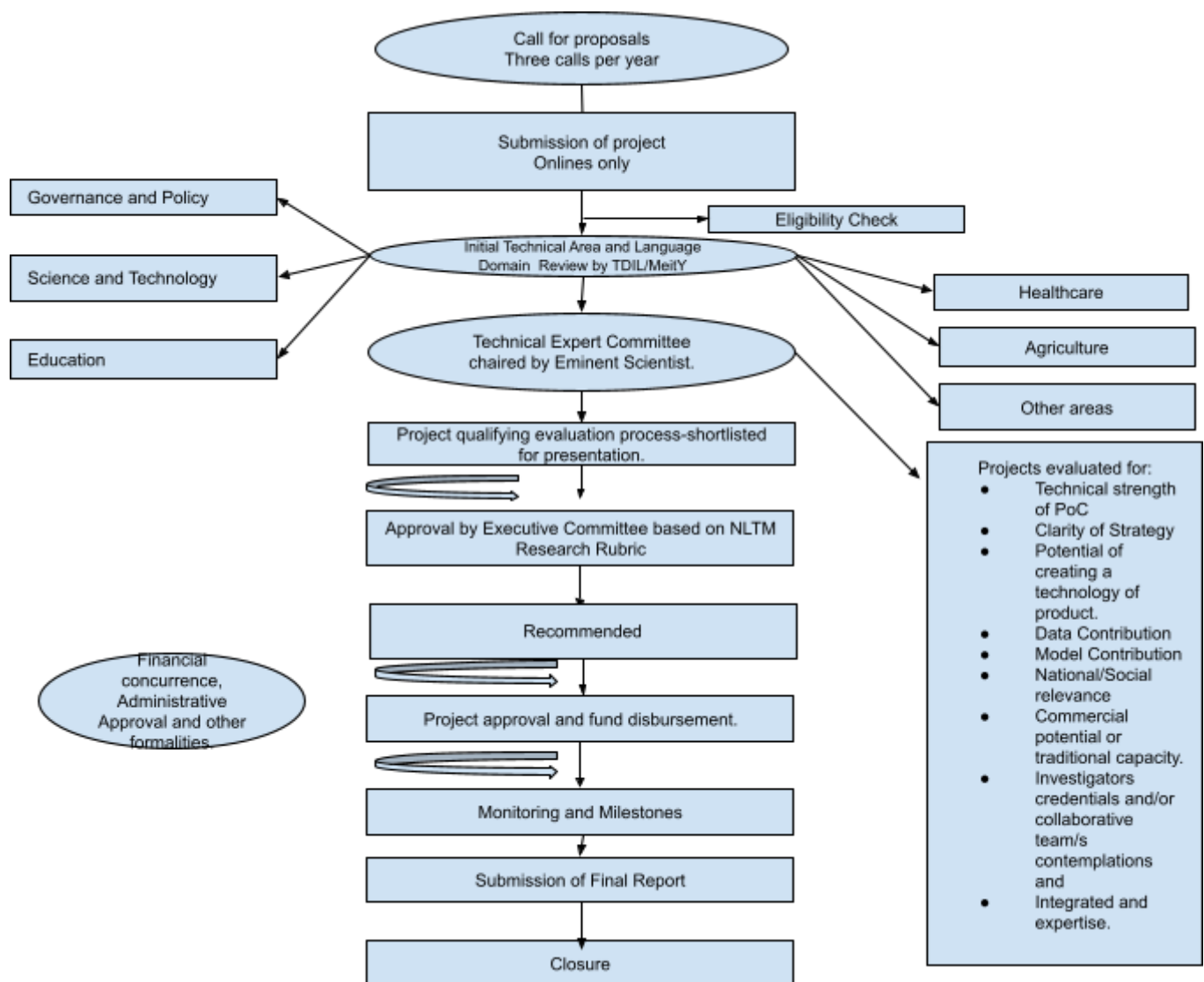
1.8 Research Rubric

The following parameters are indicative and Research Rubric will be finalized in consultation with the Domain Experts (Global as well as National)

Parameter	Remarks
Domain	Governance and Policy (Primary), Science and Technology, Education, Health, Agriculture etc.
Research Approach	Innovation and originality of approach
Research Potential	Likelihood of reaching high levels and attaining success
Data Contribution Size to NLTM in ULCA Format	
Minimum and Maximum Data Requirement with justification	
Current Model Performance and Technology Readiness Level (TRL)	
Future/Target Model Performance and Technology Readiness Level (TRL)	
Model Architecture and Size (Estimate)	Performance, Lightweight, Optimization, Innovative
Technology Delivery	Model as well as ancillary modules/subsystems
Model Compute Requirement	
Duration of the Research	
Cost	
Research Team's Performance	Publications, Prototypes Developed,

	Working Systems Delivered, Performance Levels
Past Performance of NLP/Speech Projects	<ol style="list-style-type: none"> 1. Business Adoption of Research 2. Usage of Research for social benefit 3. Impact on Research Community 4. TRL Level

1.9 Process Flow of Call for Research Proposal under NLTM



1.0 Proposal Submission

The proposal must be submitted in **standard format for submitting R&D proposals in MeitY** to nltn-research@meity.gov.in