



**INDIAN INSTITUTE OF TECHNOLOGY BOMBAY**  
**MATERIALS MANAGEMENT DIVISION**  
Powai, Mumbai 400076.

**PR. No. 1000027374**

**RFx. 6100001156**

**Translation Service from Hindi To Marathi.**

**Technical Bid**

**A) Scope of Work: -**

Center for Indian Language Technology (CFILT), IIT Bombay under the leadership of Prof. Pushpak Bhattacharyya is creating an Hindi - Marathi parallel corpus. This is being done as part of the NLTM-Vidyaapati project for Text to Text machine translation, awarded by Ministry of Electronics and Information Technology to various top educational institutions across India. This is a part of National Language Technology Mission applied to education domain.

The projects aims to design, develop and deploy Neural Machine Translation (NMT) system for Hindi to Marathi translation. **Around 20,000 Hindi sentences from different domains would be translated to Marathi language.** The direction of translation will be from Hindi to Marathi and Marathi to Hindi in the ratio of 80:20.

Large volumes of training data, also called parallel corpora, are needed to train the Machine Learning models used in MT. Such large parallel corpora is not available in most of the Indian languages and this has been a major barrier for the progress of MT technology in these languages. To address the data gap in the Marathi language, CFILT, IIT Bombay had set up the Hindi-Marathi parallel corpus creation challenge and had opened it to participants from across the country.

An Hindi-Marathi corpus consisting of around 2 lac sentence pairs would be needed to build a professional quality MT system that can be fed into NLTM Bhashini system. Start-ups would be engaged to help IIT-B in creation of this huge parallel corpora. IIT-B will provide text documents or web links to the participant having Hindi sentences and the participants are required to produce high-quality Marathi translations. Start-ups/bidders are required to execute the work within a period **of 2 months** from the date of receipt of the work order.

Item No	Item Code	Item Description	Quantity	Unit
10		Translation from Hindi to Marathi	1.000	Activ.unit
10.1		Translation from Hindi to Marathi	5000	Each

20		Translation from Hindi to Marathi	1.000	Activ.unit
20.1		Translation from Hindi to Marathi	5000	each
30		Translation from Hindi to Marathi	1.000	Activ.unit
30.1		Translation from Hindi to Marathi	5000	each
40		Translation from Hindi to Marathi	1.000	Activ.unit
40.1		Translation from Hindi to Marathi	5000	each

## **B) Eligibility Criteria for Service providers:**

1. The service provider should have an established firm for the minimum period of two year (Please attach Certificate of incorporation/Shop & Establishment License/ Deed of partnership/any Registration certificate).
2. The service provider should enclose copy the following:
  - a) GST certificate/ proprietorship certificate
  - b) PAN Card
3. The service provider should have an annual business turnover of minimum of Rs. 5 lakhs for F. Y. 2019-2020, 2020-2021 and 2021-2022
4. The service provider should submit the Income Tax Return copy of last 3 previous year.
5. The service provider should have similar experience in Hindi to Marathi (primary task) and/or having experience in other Indian language translation (added advantage) services for government offices/PSU/autonomous bodies/ Private organizations. They will be required to attach relevant documents as proofs.

## **C) Miscellaneous Instructions:**

1. Average word count in each sentence would be 15-20 words per sentences. Word count may increase or decrease as per the context of the sentence.
2. Sentences can be of any domain viz. administration, law, judiciary, education, culture or tourism.
3. The base for price/cost calculation will be the words in source sentences. Therefore, the price is to be quoted on per word basis.
4. Source sentence word count will be considered when processing the bill.

5. Payment to bidder/vendor will be as per IIT Bombay norms.

6. Sentences will be shared in word, excel or pdf form. Or, weblink will be shared. Bidders/translators are required to scrape the data from there.