GOVERNMENT OF INDIA
MINISTRY OF ELECTRONICS AND INFORMATION TECHNOLOGY
RAJYA SABHA
UNSTARRED QUESTION NO. 1497
TO BE ANSWERED ON: 31.07.2026

DEVELOPMENT OF AI BASED LANGUAGE TRANSLATION TOOLS UNDER BHASHINI

1497. # SHRI NARESH BANSAL:
SHRI MAYANKKUMAR NAYAK:
DR. DINESH SHARMA:
SHRI AMAR PAL MAURYA:
SHRI MANAN KUMAR MISHRA:
DR. MEDHA VISHRAM KULKARNI:
SHRI RAJESH PARMANAND SHUKLA:

Will the Minister of Electronics and Information Technology be pleased to state:

(a) the details of the progress made under BHASHINI in the development of Artificial Intelligence (AI)-based language translation tools for Indian languages;
(b) the number of languages and dialects covered under the initiative so far and the number of datasets prepared therefor;
(c) whether Government has taken steps to integrate BHASHINI with Government service delivery platforms and e-Governance portals, if so, the details thereof; and
(d) whether Government has taken any steps to make digital content and services more accessible to non-English-speaking citizens, if so, the details thereof?

ANSWER

MINISTER OF STATE FOR ELECTRONICS AND INFORMATION TECHNOLOGY
(SHRI JITIN PRASADA)

(a) and (b): The Government is implementing BHASHINI (BHASHa INterface for India) under the National Language Translation Mission (NLTM) of the Ministry of Electronics and Information Technology (MeitY) to develop an AI-powered multilingual Digital Public Infrastructure (DPI) and enable citizens to access digital services in the language of their choice.

Under BHASHINI, indigenous AI-based language technologies have been developed to support multilingual communication across sectors such as governance, education, healthcare, agriculture, judiciary, finance and citizen services. The platform provides AI-based language services including Automatic Speech Recognition (ASR), Neural Machine Translation (NMT), Text-to-Speech (TTS), Optical Character Recognition (OCR), Transliteration, Language Detection (Text and Audio), Multilingual Speech Systems and Conversational AI through secure and scalable APIs.

At present, the platform supports 36+ Indian text languages, 23 Indian speech languages, and 35 international languages, backed by 359+ AI models developed

through collaborations with academic institutions, research organizations, startups and industry partners. The platform has processed over 8 billion AI inferences, enabling approximately 20 million multilingual interactions per day across digital platforms.

To support multilingual AI development, BHASHINI has established India's public language resource ecosystems comprising 298 multilingual datasets covering 109 languages, including 51 tribal and low-resource languages, with data collected across 165 districts and 31 States and Union Territories. These datasets include parallel text corpora, speech datasets, OCR datasets, transliteration resources, lexicons and annotated language resources used for training and improving AI models. Dataset creation has been enabled through collaborations with Government agencies, academic institutions, research organisations, startups and citizens through the BHASHINI Samudaye (Sahyogi) platform, which has onboarded 10,000+ contributors.

(c) and (d): BHASHINI has been integrated with multiple Government digital platforms and citizen service ecosystems to enable multilingual access, voice-enabled interaction and AI-powered translation across public services.

BHASHINI has enabled multilingual service delivery across 800+ Government websites and digital platforms, helping citizens interact with Government services in their preferred language.

Major deployments include:

- DIKSHA for multilingual digital education.
- Rail Madad for multilingual grievance redressal.
- eGram Swaraj for Panchayati Raj governance.
- CPGRAMS Samadhan Didi for multilingual grievance support.
- Nyaya Setu for multilingual legal assistance.
- SabhaSaar for multilingual parliamentary and institutional communication.
- eOffice for multilingual Government workflows.
- Integrations with RBI, SBI, PFRDA, CAG and other national institutions.
- Agriculture platforms including Bharat ViSTAAR, Maha ViSTAAR, Amul AI Platform, benefiting over 36 lakh farmers across 18,600 villages in Gujarat through multilingual digital services.
- Public Distribution System initiative ASHA, enabling multilingual interaction with over 20 lakh beneficiaries every month.

  BHASHINI also enables citizens to access digital services in the language they understand best through AI-powered multilingual technologies.

Key initiatives include:

- AI-powered translation of Government websites and digital services.
- Voice-first interfaces using Automatic Audio Language Detection, Automatic Speech Recognition and Text-to-Speech technologies.
- Multilingual conversational AI and chatbots for citizen services.
- Real-time speech systems and live transcription through Shrutlekh, deployed across 180+ national events and conferences.
- Mobile and web-based solutions including BHASHINI Mobile App, Translation Plug-in, Shrutlekh, VaaniAnuvaad, Lekhaanuvaad, Chitraanuvaad, Abhilekh, Podium Solution and VoicERA.

- Tribal language inclusion through integration of languages such as Bhili into the BHASHINI Mobile App.
- Citizen participation through BhashaDaan and BHASHINI Samudaye, enabling collaborative creation of language datasets and AI resources.
- Open APIs supporting Government departments, startups and developers in building multilingual digital solutions.

These integrations are enabling citizens to access Government information and services through AI-powered translation, speech recognition, voice interfaces and multilingual conversational AI, thereby strengthening inclusive digital governance.

****