GOVERNMENT OF INDIA
MINISTRY OF ELECTRONICS AND INFORMATION TECHNOLOGY
RAJYA SABHA
UNSTARRED QUESTION. NO. 3883
TO BE ANSWERED ON: 27.03.2026

EXPANSION OF INDIAN AI LANGUAGE ECOSYSTEM

3883. DR. PARMAR JASHVANTSINH SALAMSINH:
SHRI SHAMBHU SHARAN PATEL:
SHRI MAYANKKUMAR NAYAK:
DR. MEDHA VISHRAM KULKARNI:
SHRI JAGGESH:

Will the Minister of ELECTRONICS AND INFORMATION TECHNOLOGY be pleased to state:

(a) whether the Bhashini Samudaye workshop has successfully expanded India's language Artificial Intelligence (AI) ecosystem for administrative workflows;
(b) if so, the details thereof;
(c) the manner in which the Shrutlekh tool has improved real-time transcription and translation for governance use cases across 22 languages;
(d) the status of the 'Bhasha Daan' initiative in crowdsourcing diverse linguistic data for training Indian AI models; and
(e) the manner in which the deployment of voice-first AI tools has successfully improved service delivery for illiterate population in rural areas?

ANSWER

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

(a) to (e): Bhashini Samudaye is centralised digital platform bringing together individual contributors, startups, academic institutions, and research organisations to support the development of a multilingual AI ecosystem.

It facilitates access to language resources for training AI models and supports creation of datasets, domain-specific glossaries, and AI models. A structured onboarding mechanism has been developed under this initiative.

Over 10,000 contributors have been onboarded.

The initiative also supports participation in data creation and annotation tasks, contributing to ecosystem development and capacity building.

Shrutlekh is an AI-powered, real-time speech-to-text transcription and translation tool developed by Bhashini. It enables AI-based services across 22 Indian languages.

Shrutlekh has been deployed across 100+ sessions at major national and international platforms, including the India AI Impact Summit, World Summit on Disaster Management (Dehradun), Veer Bal Diwas (Puducherry), AI4Agri Summit 2026, RBI Hindi Seminar, and the TiE Global Summit (Rajasthan Chapter).
It has been used in major conferences like World Disaster Summit where multiple sessions were supported by real time transcription. Additionally the software has been integrated with the TAARA device being built by Niti Aayog.

Bhasha Daan initiative is an ongoing national effort to crowdsource diverse linguistic data at scale for strengthening Indian language AI models. It enables active participation from citizens, volunteers, and institutions in contributing text and speech data across multiple Indian languages.
It includes low-resource and underrepresented languages which help build an inclusive and representative language ecosystem.

The initiative has facilitated the creation of large-scale, high-quality datasets, including parallel corpora, speech datasets, and transliteration resources. These are being used to train and continuously improve AI models for translation, speech recognition, and language technologies.
The deployment of voice-first AI tools under BHASHINI has significantly improved the accessibility and delivery of digital services for non-literate and semi-literate populations, particularly in rural areas.

MahaVISTAAR-AI, an AI-powered mobile application developed by the Department of Agriculture, Government of Maharashtra, leverages BHASHINI-enabled multilingual capabilities to provide real-time advisory services to farmers.

The platform delivers personalized guidance on sowing, pest management, yield prediction, and climate-resilient practices, along with real-time updates on weather, market prices, and agromet advisories, through multilingual conversational interfaces.

By enabling interaction through speech in native languages, these solutions reduce dependence on text-based interfaces and allow citizens to access services in a more intuitive and inclusive manner.

****