

**GOVERNMENT OF INDIA  
MINISTRY OF SCIENCE AND TECHNOLOGY  
DEPARTMENT OF SCIENCE AND TECHNOLOGY  
LOK SABHA  
UNSTARRED QUESTION NO. 3827  
ANSWERED ON 18/12/2024**

**LAUNCH OF BHARATGEN**

**3827. SMT. SMITA UDAY WAGH:**

**Will the Minister of SCIENCE AND TECHNOLOGY विज्ञान और प्रौद्योगिकी मंत्री be pleased to state:**

- (a) the manner in which BharatGen ensure that its generative AI models adequately represent country's diverse linguistic landscape and what specific measures are being taken to include lesser-represented languages;**
- (b) the steps will BharatGen take to prioritize making its generative AI models accessible as a public good, particularly for marginalized and under represented communities in the country;**
- (c) the manner in which BharatGen contribute to the development of region-specific content, ensuring that local languages and dialects are effectively utilized in promoting cultural identity and regional development; and**
- (d) the partnerships or collaborations are being considered to enhance the inclusivity of BharatGen's models and ensure equitable technological access across different socio-economic groups in the country?**

**ANSWER**

**MINISTER OF STATE (INDEPENDENT CHARGE) OF THE  
MINISTRY OF SCIENCE AND TECHNOLOGY AND EARTH SCIENCES  
(DR. JITENDRA SINGH)**

विज्ञान और प्रौद्योगिकी तथा पृथ्वी विज्ञान मंत्रालय के राज्य मंत्री (स्वतंत्र प्रभार)  
(डॉ. जितेंद्र सिंह)

- (a) BharatGen is a multimodal multilingual large language model initiative, developing advanced generative AI models tailored to India's linguistic, cultural, and socio-economic diversity. To ensure that generative AI models adequately represent India's diverse linguistic landscape, BharatGen has launched an initiative called**

**“Bharat Data Sagar”, focusing on primary data collection. This data collection attempts to meet the requirement that training data is available for Indian languages that are lesser represented in data corpora.**

**(b) BharatGen is building partnerships with research groups across the country, to ensure that the generative AI models being developed can be extended by partners and made available to the larger research and non-academic community for further development and usage. BharatGen is also developing partnerships with the government, industry and start-ups for applications geared towards efficient administration and public at large, including marginalized and underrepresented communities in the country.**

**(c) In order to promote cultural identity and regional development, BharatGen provides technologies and tools that will support the development of region-specific content by seamlessly translating across local languages and dialects.**

**(d) BharatGen includes a consortium of top AI researchers across premier academic institutions in India that include IIT Bombay, IIIT Hyderabad, IIT Mandi, IIT Kanpur, IIT Hyderabad, IIM Indore, and IIT Madras. These research groups are partnering with the government, industry and startups to develop models, keeping in mind the linguistic and cultural diversity of India and inclusivity for citizens and to ensure equitable technological access across different socio-economic groups in the country.**

**\*\*\*\*\***