भारत सरकार

इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी मंत्रालय

राज्य सभा

अतारांकित प्रश्न संख्या 1497

जिसका उत्तर 31 जुलाई, 2026 को दिया जाना है।

09 श्रावण, 1948 (शक)

भाषिणी के अंतर्गत एआई आधारित भाषा अनुवाद उपकरणों का विकास

1497. श्री नरेश बंसल:
श्री मयंककुमार नायक:
डा. दिनेश शर्मा:
श्री अमर पाल मौर्य:
श्री मनन कुमार मिश्रः
डा. मेधा विश्राम कुलकर्णी:
श्री राजेश परमानंद शुक्रा:

क्या **इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी मंत्री** यह बताने की कृपा करेंगे किः

(क) भाषिणी के अंतर्गत भारतीय भाषाओं हेतु कृत्रिम बुद्धिमत्ता (एआई) आधारित भाषा अनुवाद उपकरणों के विकास में हुई प्रगति का ब्यौरा क्या है;
(ख) इस पहल के अंतर्गत अब तक कितनी भाषाओं और बोलियों को सम्मिलित किया गया है तथा इस हेतु कितने डेटासेट तैयार किए गए हैं;
(ग) क्या सरकार ने भाषिणी को सरकारी सेवा वितरण प्लेटफॉर्मों और ई-गवर्नेंस पोर्टलों के साथ एकीकृत करने हेतु कदम उठाए हैं, यदि हाँ, तो तत्संबंधी ब्यौरा क्या है; और
(घ) क्या सरकार ने डिजिटल सामग्री और सेवाओं को गैर-अंग्रेजी भाषी नागरिकों हेतु और अधिक सुलभ बनाने के लिए कोई कदम उठाए हैं, यदि हाँ, तो तत्संबंधी ब्यौरा क्या है?

उत्तर

इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी राज्य मंत्री (श्री जितिन प्रसाद)

(क) और **(ख):** सरकार इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी मंत्रालय (एमईआईटीवाई) के राष्ट्रीय भाषा अनुवाद मिशन (एनएलटीएम) के तहत भाषिणी (भारत के लिए भाषा इंटरफ़ेस) को लागू कर रही है ताकि एआई-संचालित बहुभाषी डिजिटल सार्वजनिक बुनियादी ढांचा (डीपीआई) विकसित किया जा सके और नागरिकों को उनकी पसंद की भाषा में डिजिटल सेवाओं तक पहुंचने में सक्षम बनाया जा सके।

भाषिणी के तहत, शासन, शिक्षा, स्वास्थ्य सेवा, कृषि, न्यायपालिका, वित्त और नागरिक सेवाओं जैसे क्षेत्रों में बहुभाषी संचार का समर्थन करने के लिए स्वदेशी एआई-आधारित भाषा प्रौद्योगिकियों का विकास किया गया है। यह प्लेटफ़ॉर्म सुरक्षित और स्केलेबल एपीआई के माध्यम से ऑटोमैटिक स्पीच रिकॉग्निशन (एएसआर), न्यूरल मशीन ट्रांसलेशन (एनएमटी), टेक्स्ट-टू-स्पीच (टीटीएस), ऑप्टिकल कैरेक्टर रिकॉग्निशन (ओसीआर), लिप्यंतरण, लैंग्वेज डिटेक्शन (टेक्स्ट और ऑडियो), बहुभाषी स्पीच सिस्टम और कन्वर्सेशनल एआई सहित एआई-आधारित भाषा सेवाएं प्रदान करता है।

वर्तमान में, यह प्लेटफॉर्म 36 से अधिक भारतीय पाठ भाषाओं, 23 भारतीय बोलचाल की भाषाओं और 35 अंतरराष्ट्रीय भाषाओं का समर्थन करता है, जो शैक्षणिक संस्थानों, अनुसंधान संगठनों, स्टार्टअप्स और उद्योग भागीदारों के सहयोग से विकसित 359 से अधिक एआई मॉडल द्वारा समर्थित है। प्लेटफॉर्म ने 8 अरब से अधिक एआई अनुमानों को संसाधित किया है, जिससे डिजिटल प्लेटफॉर्म पर प्रतिदिन लगभग 2 करोड़ बहुभाषी बातचीत संभव हो पाती है।

बहुभाषी एआई के विकास को समर्थन देने के लिए, भाषिणी ने भारत के सार्वजनिक भाषा संसाधन पारिस्थितिकी तंत्र की स्थापना की है जिसमें 298 बहुभाषी डेटासेट शामिल हैं जो 109 भाषाओं को कवर करते हैं, जिनमें 51 जनजातीय और कम संसाधन वाली भाषाएं शामिल हैं, और डेटा 165 जिलों और 31 राज्यों और केंद्र शासित प्रदेशों से एकत्र किया गया है। इन डेटासेट में समानांतर पाठ कॉर्पोरा, स्पीच डेटासेट, ओसीआर डेटासेट, लिप्यंतरण संसाधन, शब्दकोश और एआई मॉडल के प्रशिक्षण और सुधार के लिए उपयोग किए जाने वाले एनोटेट भाषा संसाधन शामिल हैं। भाषिणी समुदाये (सहयोगी) प्लेटफॉर्म के माध्यम से सरकारी एजेंसियों, शैक्षणिक संस्थानों, अनुसंधान संगठनों, स्टार्टअप और नागरिकों के सहयोग के माध्यम से डेटासेट निर्माण को सक्षम किया गया है, जिसने 10,000+ योगदानकर्ताओं को शामिल किया है।

(ग) और **(घ):** भाषिणी को सार्वजनिक सेवाओं में बहुभाषी पहुंच, आवाज-सक्षम बातचीत और एआई-संचालित अनुवाद को सक्षम करने के लिए कई सरकारी डिजिटल प्लेटफार्मों और नागरिक सेवा पारिस्थितिकी तंत्र के साथ एकीकृत किया गया है।

भाषिणी भाषिणी ने 800 से ज़्यादा सरकारी वेबसाइटों और डिजिटल प्लेटफ़ॉर्म पर कई भाषाओं में सेवाएँ उपलब्ध कराई हैं, जिससे नागरिकों को अपनी पसंदीदा भाषा में सरकारी सेवाओं का लाभ लेने में मदद मिलती है।

प्रमुख तैनाती में शामिल हैं:

- बहुभाषी डिजिटल शिक्षा के लिए दीक्षा।
- बहुभाषी शिकायत निवारण के लिए रेल मदद।
- पंचायती राज शासन के लिए ई-ग्राम स्वराज।
- बहुभाषी शिकायत सहायता के लिए सीपीजीआरएएमएस समाधान दीदी।
- बहुभाषी कानूनी सहायता के लिए न्याय सेतु।
- बहुभाषी संसदीय और संस्थागत संचार के लिए सभासार।
- बहुभाषी सरकारी वर्कफ़्लोज़ के लिए eOffice.
- आरबीआई, एसबीआई, पीएफआरडीए, सीएजी और अन्य राष्ट्रीय संस्थानों के साथ एकीकरण।
- भारत विस्तार, महा विस्तार, अमूल एआई प्लेटफॉर्म सहित कृषि प्लेटफॉर्म बहुभाषी डिजिटल सेवाओं के माध्यम से गुजरात के 18,600 गांवों में 36 लाख से अधिक किसानों को लाभान्वित कर रहे हैं।
- सार्वजनिक वितरण प्रणाली पहल आशा, जो हर महीने 20 लाख से ज़्यादा लाभार्थियों के साथ कई भाषाओं में बातचीत करने में मदद करती है।

भाषिणी नागरिकों को एआई-संचालित बहुभाषी प्रौद्योगिकियों के माध्यम से उस भाषा में डिजिटल सेवाओं तक पहुंचने में सक्षम बनाती है जिसे वे सबसे अच्छी तरह समझते हैं।

प्रमुख पहलों में शामिल हैं:

- सरकारी वेबसाइटों और डिजिटल सेवाओं का एआई-संचालित अनुवाद।
- स्वचालित ऑडियो भाषा का पता लगाने, स्वचालित वाक् पहचान और टेक्स्ट-टू-स्पीच तकनीकों का उपयोग करके वॉयस-फर्स्ट इंटरफेस।
- नागरिक सेवाओं के लिए बहुभाषी संवादी एआई और चैटबॉट।
- वास्तविक समय भाषण प्रणाली और श्रुतलेख के माध्यम से लाइव ट्रांसक्रिप्शन, 180+ राष्ट्रीय कार्यक्रमों और सम्मेलनों में नियोजित किया गया है।
- भाषिणी मोबाइल ऐप, ट्रांसलेशन प्लग-इन, श्रुतलेख, वाणी-अनुवाद, लेखानुवाद, चित्रानुवाद, अभिलेख, पोडियम सॉल्यूशन और वोइसेरा सहित मोबाइल और वेब-आधारित समाधान उपलब्ध हैं।
- भीली जैसी भाषाओं को भाषिणी मोबाइल ऐप में शामिल करने के माध्यम से जनजातीय भाषाओं को शामिल किया गया है।
- भाषादान और भाषिणी समुदाय के माध्यम से नागरिक भागीदारी, भाषा डेटासेट और एआई संसाधनों के सहयोगात्मक निर्माण को सक्षम बनाती है।
- बहुभाषी डिजिटल समाधान बनाने में सरकारी विभागों, स्टार्टअप्स और डेवलपर्स का समर्थन करने वाले ओपन एपीआई।

ये एकीकरण नागरिकों को एआई-संचालित अनुवाद, स्पीच रिकग्निशन, वॉयस इंटरफेस और बहुभाषी संवादात्मक एआई के माध्यम से सरकारी जानकारी और सेवाओं तक पहुंचने में सक्षम बना रहे हैं, जिससे समावेशी डिजिटल शासन को मजबूत किया जा रहा है।

*****