आदि वाणी: AI से जनजातीय भाषाओं का डिजिटल संरक्षण - 9 भाषाएँ, 35 हज़ार+ डाउनलोड (PIB)



मौलिक प्रतीकात्मक चित्र: pcsmantra.com

खबरों में क्यों

जनजातीय गौरव पखवाड़ा (1-15 नवंबर 2026) और **जनजातीय गौरव दिवस** (15 नवंबर - भगवान बिरसा मुंडा जयंती) के अवसर पर जनजातीय मामले मंत्रालय ने 300+ परियोजनाएँ (228 उद्घाटन, 74 शिलान्यास) और दो राष्ट्रीय डिजिटल पहलें - **Jan-HIT पोर्टल** और **राष्ट्रीय FRA डिजिटल प्लेटफ़ॉर्म** - शुरू कीं। इसी संदर्भ में PIB की 15 नवंबर की रिपोर्ट **"आदि वाणी"** AI प्लेटफ़ॉर्म को उजागर करती है: 15 अक्टूबर 2026 तक यह **9 जनजातीय भाषाओं** में टेक्स्ट-अनुवाद देता है और **35,000+ Android डाउनलोड** दर्ज हैं।

आदि वाणी क्या है

- **परिचय:** जनजातीय मामले मंत्रालय का AI-आधारित अनुवाद और भाषा-प्लेटफ़ॉर्म - **1 सितंबर 2025** को लॉन्च; **IIT दिल्ली** के नेतृत्व में BITS पिलानी, IIIT हैदराबाद, IIIT नवा रायपुर और झारखंड, ओडिशा, MP, छत्तीसगढ़, मेघालय की **जनजातीय अनुसंधान संस्थाएँ (TRI)** सहित राष्ट्रीय कंसोर्टियम द्वारा विकसित - "जनजातीय समुदायों के साथ, जनजातीय समुदायों के लिए"।
- **भाषा-कवरेज:** शुरुआत **संताली, भिली, मुंडारी, गोंडी** से; 15 अक्टूबर 2026 तक **9 भाषाओं** में टेक्स्ट-अनुवाद; साथ में स्पीच-अनुवाद, **OCR**, वीडियो-सबटाइटलिंग और **300+ भाषा प्राइमर**।
- **पहुँच:** 35,000+ Android डाउनलोड, ~6,500 मासिक उपयोगकर्ता।
- **डेटा-निर्माण:** देशी वक्ता TRI और **AdiKosh** (शब्द-बैंक) के ज़रिए शब्द-वाक्य-आवाज़ देते हैं; भाषा-विशेषज्ञ जाँचते हैं; लक्ष्य - **प्रति भाषा ~1,00,000 वाक्य-युग्म और 100 घंटे वाणी**; 15 अक्टूबर 2026 तक **~13.9 लाख वाक्य (10 भाषाएँ)** और **~300 घंटे स्पीच डेटा** जमा।
- **आगे का रास्ता:** जनजातीय भाषाओं के कार्यकारी समूह (पहली बैठक 15 अक्टूबर 2026) ने आदि वाणी को **राष्ट्रीय जनजातीय भाषा भंडार (National Tribal Language Repository)** बनाने पर चर्चा की।

प्रमुख फ़ीचर

- टेक्स्ट-टू-टेक्स्ट अनुवाद (जनजातीय भाषाएँ ↔ हिंदी/अंग्रेज़ी)
- स्पीच-टू-टेक्स्ट (मौखिक भाषा का दस्तावेज़ीकरण)
- स्पीच-टू-स्पीच अनुवाद
- OCR - चित्र/स्कैन दस्तावेज़ों से टेक्स्ट निकालकर अनुवाद
- वीडियो सबटाइटलिंग और डिजिटल शब्दकोश/प्राइमर

क्यों महत्वपूर्ण है

- **GS-I/II (समाज/शासन):** जनसंख्या 2011 के अनुसार **~10.45 करोड़ अनुसूचित जनजाति (~8.6%)**; मातृभाषा में प्रारंभिक शिक्षा (NEP 2020 की भावना से सुसंगत), स्वास्थ्य-संचार और लोक-सेवाओं तक पहुँच।
- **ज्ञान-संरक्षण:** मौखिक परंपरा की भाषाएँ वन, कृषि, औषधि और जैव-विविधता के ज्ञान के वाहक - कई "कम-संसाधन" भाषाएँ डिजिटल तकनीक से वंचित रहीं।
- **सामुदायिक AI मॉडल:** डेटा-दान + विशेषज्ञ-सत्यापन + पुनः-प्रशिक्षण - भारत के लिए अनुकरणीय "लोक-भागीदारी AI" प्रयोग।

प्रिलिम्स परीक्षा अभ्यास प्रश्न

प्रश्न 1: आदि वाणी प्लेटफ़ॉर्म किस मंत्रालय की पहल है?

(a) इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी मंत्रालय

(b) जनजातीय मामले मंत्रालय

(c) शिक्षा मंत्रालय

(d) संस्कृति मंत्रालय

उत्तर और व्याख्या

उत्तर: (b) जनजातीय मामले मंत्रालय का AI-आधारित अनुवाद प्लेटफ़ॉर्म; 1 सितंबर 2025 को लॉन्च; IIT दिल्ली नेतृत्व वाला कंसोर्टियम।

प्रश्न 2: आदि वाणी में "AdiKosh" क्या है?

(a) जनजातीय कलाओं का e-बाज़ार

(b) शब्द-बैंक - सत्यापित वक्ताओं द्वारा दर्ज शब्दकोश-प्रविष्टियाँ और बोले वाक्य

(c) जनजातीय छात्रवृत्ति पोर्टल

(d) वन-अधिकार दस्तावेज़-भंडार

उत्तर और व्याख्या

उत्तर: (b) AdiKosh आदि वाणी का वर्ड-बैंक है; इसमें दर्ज वाणी TRI द्वारा सत्यापित होकर AI मॉडलों के प्रशिक्षण में जाती है।

मेन्स परीक्षा अभ्यास प्रश्न

प्रश्न: "तकनीक भाषा-विविधता का बचावक बन भी सकती है और भस्मक भी - दिशा नीति तय करती है।" आदि वाणी के संदर्भ में टिप्पणी कीजिए। (250 शब्द)

रूपरेखा: भूमिका - कम-संसाधन भाषाओं की डिजिटल उपेक्षा; आदि वाणी मॉडल - सामुदायिक डेटा, TRI, सत्यापन; लाभ - मातृभाषा-शिक्षा, स्वास्थ्य-संचार, सेवा-पहुँच, सांस्कृतिक संरक्षण; जोखिम - डेटा-स्वामित्व, प्रतिनिधित्व-पूर्वाग्रह; निष्कर्ष - राष्ट्रीय जनजातीय भाषा-भंडार की दिशा सही कदम।

FAQ

संक्षिप्त उत्तर: आदि वाणी जनजातीय मामले मंत्रालय का AI प्लेटफ़ॉर्म है (1 सितंबर 2025 लॉन्च) - 9 जनजातीय भाषाओं में टेक्स्ट-अनुवाद, स्पीच/OCR/सबटाइटलिंग सुविधाएँ; 35,000+ डाउनलोड; ~13.9 लाख वाक्य और ~300 घंटे वाणी डेटा TRI/AdiKosh से; लक्ष्य - राष्ट्रीय जनजातीय भाषा भंडार।

प्रश्न: आदि वाणी की ख़ासियत क्या है?

यह जनजातीय भाषाओं के लिए AI-आधारित अनुवाद और डिजिटल-संरक्षण मंच है - टेक्स्ट/स्पीच अनुवाद, OCR, वीडियो सबटाइटलिंग और 300+ प्राइमर; डेटा देशी वक्ताओं से TRI/AdiKosh के माध्यम से, विशेषज्ञ-सत्यापन के साथ।

प्रश्न: अभी कितनी भाषाएँ और कितनी पहुँच?

15 अक्टूबर 2026 तक 9 भाषाओं में टेक्स्ट-अनुवाद (शुरुआत संताली, भिली, मुंडारी, गोंडी से); 35,000+ Android डाउनलोड और ~6,500 मासिक उपयोगकर्ता।

प्रश्न: जनजातीय भाषाओं का डिजिटल संरक्षण क्यों ज़रूरी?

ये भाषाएँ सदियों के लोक-ज्ञान (वन, कृषि, औषधि, जैव-विविधता) की वाहक हैं; डिजिटल उपलब्धता मातृभाषा-शिक्षा, स्वास्थ्य-सूचना और लोक-सेवाओं की पहुँच बढ़ाती है और भाषाओं के लोप को रोकती है।

स्रोत

- PIB (15 नवंबर 2026, आधिकारिक): Adi Vaani - Bringing India's Tribal Languages into the Digital Age