

NEXT IAS

दैनिक संपादकीय विश्लेषण

विषय

भारत में AI सुरक्षा संस्थान की स्थापना:
आवश्यकताएँ और चिंताएँ

www.nextias.com

भारत में AI सुरक्षा संस्थान की स्थापना: आवश्यकताएँ और चिंताएँ

सन्दर्भ

- हाल ही में, केंद्रीय इलेक्ट्रॉनिक्स और सूचना प्रौद्योगिकी मंत्रालय (MeitY) ने IndiaAI मिशन के तहत एक AI सुरक्षा संस्थान स्थापित करने के लिए चर्चा शुरू की है, जिसका उद्देश्य घरेलू और वैश्विक AI सुरक्षा चुनौतियों का समाधान करना, भारत की अद्वितीय शक्तियों का लाभ उठाना एवं अंतर्राष्ट्रीय सहयोग को बढ़ावा देना है।

पृष्ठभूमि और तर्क: भारत का AI सुरक्षा संस्थान

- AI सेफ्टी इंस्टीट्यूट का विचार G20 और ग्लोबल पार्टनरशिप ऑन आर्टिफिशियल इंटेलिजेंस (GPAI) जैसे वैश्विक प्लेटफार्मों पर भारत की हालिया नेतृत्व भूमिकाओं के अनुरूप है।
- इसकी परिकल्पना पूर्वाग्रह, भेदभाव और AI सिस्टम की नैतिक तैनाती सहित AI सुरक्षा के आसपास बढ़ती चिंताओं को दूर करने के लिए की गई है।

उद्देश्य और लक्ष्य

- **घरेलू क्षमता बढ़ाना:** AI सिस्टम की तैनाती से पहले उसकी सुरक्षा का आकलन और सुनिश्चित करने के लिए भारत के अंदर मजबूत ढांचे एवं क्षमताओं का निर्माण करना।
 - इसमें सुरक्षित AI परिनियोजन के लिए रूपरेखा, दिशानिर्देश और मानक विकसित करना शामिल है।
 - भारत की विशिष्ट चुनौतियों और शक्तियों पर ध्यान केंद्रित करके, संस्थान यह सुनिश्चित कर सकता है कि AI प्रौद्योगिकियों को देश के अंदर जिम्मेदारी से और प्रभावी ढंग से तैनात किया जाए।
- सरकारी निकायों, उद्योग के खिलाड़ियों, शिक्षाविदों और नागरिक समाज सहित बहु-हितधारक सहयोग को बढ़ावा देना।
 - यह सुनिश्चित करता है कि AI सुरक्षा उपायों के विकास और कार्यान्वयन में विविध दृष्टिकोणों पर विचार किया जाए।
- **डेटा-संचालित निर्णय लेना:** AI नीतिगत निर्णयों को सूचित करने के लिए बड़ी मात्रा में डेटा का विश्लेषण कर सकता है, जिससे अधिक प्रभावी और लक्षित हस्तक्षेप हो सकते हैं।
 - यह स्वास्थ्य देखभाल, शिक्षा और सामाजिक कल्याण जैसे क्षेत्रों में सहायता कर सकता है, जहां डेटा-संचालित अंतर्दृष्टि परिणामों में सुधार कर सकती है।
- **मानव-केंद्रित AI :** यह सुनिश्चित करना कि AI विकास और तैनाती मानवाधिकारों, नागरिक स्वतंत्रता एवं विकासशील देशों की समावेशी भागीदारी को प्राथमिकता देती है।
- **अंतर्राष्ट्रीय सहयोग:** वैश्विक AI शासन संवाद में विविध दृष्टिकोण लाने के लिए AI सुरक्षा पर बैलेंसली प्रक्रिया और ग्लोबल डिजिटल कॉम्पैक्ट जैसी वैश्विक पहलों के साथ जुड़ना।

सामरिक महत्व

- **नवाचार और सुरक्षा संतुलन:** संस्थान का लक्ष्य नवाचार को बढ़ावा देने और यह सुनिश्चित करने के बीच संतुलन बनाना है कि AI सिस्टम सुरक्षित एवं नैतिक हैं।

- इसमें अत्यधिक निर्देशात्मक नियमों से बचना शामिल है जो मजबूत सुरक्षा उपायों को सुनिश्चित करते हुए नवाचार को बाधित कर सकते हैं।
- **वैश्विक उदाहरणों से सीख:** यूरोपीय संघ के AI कार्यालय और चीन की एल्गोरिदम रजिस्ट्री जैसे अन्य देशों के दृष्टिकोण से सीख लेते हुए, भारत का लक्ष्य एक लचीला लेकिन प्रभावी शासन मॉडल बनाना है।
- **वैश्विक नेतृत्व:** इस संस्थान की स्थापना करके, भारत वैश्विक AI शासन ढांचे को आकार देने में महत्वपूर्ण भूमिका निभा सकता है, यह सुनिश्चित करते हुए कि विकासशील देशों की आवाज़ सुनी जाए।

नैतिक और सामाजिक निहितार्थ

- **नैतिक निरीक्षण और शासन:** संस्थान यह सुनिश्चित करने के लिए रूपरेखा और दिशानिर्देश विकसित कर सकता है कि AI सिस्टम को नैतिक रूप से डिजाइन एवं तैनात किया गया है। इसमें AI एल्गोरिदम में पूर्वाग्रह, भेदभाव और निष्पक्षता जैसे मुद्दों को संबोधित करना शामिल है।
- **गोपनीयता और डेटा सुरक्षा:** डेटा गोपनीयता और सुरक्षा के लिए मानक निर्धारित करके, संस्थान व्यक्तियों की व्यक्तिगत जानकारी को दुरुपयोग से बचाने में सहायता कर सकता है और यह सुनिश्चित कर सकता है कि AI सिस्टम डेटा सुरक्षा कानूनों का अनुपालन करते हैं।
- **पारदर्शिता और जवाबदेही:** संस्थान AI निर्णय लेने की प्रक्रियाओं में पारदर्शिता को बढ़ावा दे सकता है, जिससे यह समझना आसान हो जाएगा कि AI सिस्टम अपने निष्कर्षों तक कैसे पहुंचते हैं। इससे जनता का विश्वास बनाने और जवाबदेही सुनिश्चित करने में सहायता मिल सकती है।

उत्तरदायी AI विकास में वैश्विक नेता के रूप में भारत की स्थिति

- **रणनीतिक पहल और नीतियां:** नीति आयोग द्वारा शुरू की गई आर्टिफिशियल इंटेलिजेंस के लिए राष्ट्रीय रणनीति, AI विकास के लिए एक व्यापक रोडमैप की रूपरेखा तैयार करती है जो नैतिक विचारों, समावेशिता और पारदर्शिता को प्राथमिकता देती है।
 - इसका उद्देश्य सामाजिक और आर्थिक विकास के लिए AI का लाभ उठाना है, साथ ही यह सुनिश्चित करना है कि इसकी तैनाती नैतिक मानकों के अनुरूप हो।
- **नैतिक AI फ्रेमवर्क:** MeitY ने पूर्वाग्रह, निष्पक्षता और जवाबदेही जैसे मुद्दों को संबोधित करने के लिए AI के नैतिक उपयोग को बढ़ावा देने वाले दिशानिर्देश तैयार करने में महत्वपूर्ण भूमिका निभाई है।
 - ये ढाँचे यह सुनिश्चित करने के लिए डिज़ाइन किए गए हैं कि AI सिस्टम पारदर्शी, व्याख्या योग्य और भेदभाव को जन्म देने वाले पूर्वाग्रहों से मुक्त हों।
- **सार्वजनिक-निजी भागीदारी:** सामाजिक सशक्तिकरण के लिए जिम्मेदार AI (RAISE) शिखर सम्मेलन जैसी पहल नैतिक AI प्रथाओं पर चर्चा करने और उन्हें बढ़ावा देने के लिए विभिन्न क्षेत्रों के हितधारकों को एक साथ लाती है।
 - ये साझेदारियां AI समाधान विकसित करने के लिए महत्वपूर्ण हैं जो न केवल तकनीकी रूप से उन्नत हैं बल्कि सामाजिक रूप से भी जिम्मेदार हैं।
- **समावेशिता पर ध्यान:** डिजिटल विभाजन को समाप्त करने के लिए AI प्रौद्योगिकियों में कार्यबल को कुशल और पुनः कुशल बनाने के उद्देश्य से कार्यक्रम लागू किए जा रहे हैं।

- इसके अतिरिक्त, स्वास्थ्य देखभाल, कृषि और शिक्षा में चुनौतियों का समाधान करने के लिए AI समाधान विकसित किए जा रहे हैं, जिससे लाखों भारतीयों के जीवन की गुणवत्ता में सुधार होगा।
- **अनुसंधान और नवाचार:** IITs और भारतीय विज्ञान संस्थान (IISc) जैसे अग्रणी संस्थान AI अनुसंधान में सबसे आगे हैं, जो ऐसी प्रौद्योगिकियों को विकसित करने पर ध्यान केंद्रित कर रहे हैं जो नैतिक और समाज के लिए लाभदायक हों।
 - सरकार द्वारा वित्त पोषित अनुसंधान परियोजनाएं AI के सामाजिक प्रभावों की भी खोज कर रही हैं, यह सुनिश्चित करते हुए कि नैतिक विचारों को बुनियादी स्तर से AI विकास में एकीकृत किया गया है।

प्रमुख चुनौतियाँ और संबंधित सुझाव

- **गोपनीयता संबंधी चिंताएँ:** शासन में AI के उपयोग में प्रायः बड़ी मात्रा में व्यक्तिगत डेटा का संग्रह और विश्लेषण शामिल होता है।
 - यह महत्वपूर्ण गोपनीयता संबंधी चिंताओं को जन्म देता है, क्योंकि नागरिकों को लग सकता है कि उनकी व्यक्तिगत जानकारी की निगरानी की जा रही है और उनकी सहमति के बिना इसका उपयोग किया जा रहा है।
 - AI के बढ़ते उपयोग के साथ, डेटा गोपनीयता और सुरक्षा सुनिश्चित करना सर्वोपरि हो गया है। व्यक्तिगत जानकारी की सुरक्षा के लिए कड़े डेटा संरक्षण कानून और प्रथाएं स्थापित करना एक महत्वपूर्ण चुनौती है।
- **समावेशिता और पहुंच:** AI-संचालित शासन के लाभों को समान रूप से वितरित नहीं किया जा सकता है, जिससे वर्तमान असमानताएं बढ़ सकती हैं। जिन लोगों के पास डिजिटल प्रौद्योगिकियों तक पहुंच या उनका उपयोग करने का कौशल नहीं है, वे पीछे रह सकते हैं, जिससे डिजिटल विभाजन और बढ़ जाएगा।
 - यह सुनिश्चित करना कि AI हाशिए पर रहने वाले समुदायों सहित समाज के सभी वर्गों को लाभान्वित करे, महत्वपूर्ण है। समावेशी और सुलभ AI समाधान विकसित करने से डिजिटल विभाजन को समाप्त करने में सहायता मिल सकती है।
- **संस्थागत क्षमता:** AI प्रणालियों का प्रभावी ढंग से मूल्यांकन और विनियमन करने के लिए आवश्यक संस्थागत क्षमता का निर्माण करना।
- **हितधारक जुड़ाव:** एक समग्र शासन ढांचा बनाने के लिए उद्योग विशेषज्ञों, नीति निर्माताओं और नागरिक समाज सहित विभिन्न हितधारकों की सक्रिय भागीदारी सुनिश्चित करना।
- **निर्देशात्मक नियामक नियंत्रण से बचना:** जबकि विनियमन आवश्यक है, AI सुरक्षा संस्थान को अत्यधिक निर्देशात्मक नियंत्रण से बचने की आवश्यकता है जो नवाचार को बाधित कर सकता है।
 - इसके बजाय, इसे एक सहायक वातावरण बनाने पर ध्यान केंद्रित करना चाहिए जो व्यवसायों, सरकारों और व्यापक पारिस्थितिकी तंत्र के बीच सक्रिय सूचना साझाकरण एवं सहयोग को प्रोत्साहित करता है।
- **नैतिक और नियामक ढाँचे:** व्यापक और अनुकूली नियामक ढाँचे विकसित करना जो तीव्र तकनीकी प्रगति के साथ सामंजस्य रख सकें, महत्वपूर्ण है।

निष्कर्ष और आगे की राह

- भारत के AI सुरक्षा संस्थान की स्थापना AI प्रौद्योगिकियों के सुरक्षित और जिम्मेदार उपयोग को सुनिश्चित करने की दिशा में एक महत्वपूर्ण कदम का प्रतिनिधित्व करती है।
- घरेलू क्षमता का निर्माण, बहु-हितधारक सहयोग को बढ़ावा देना, वैश्विक संवादों में शामिल होना और जोखिम मूल्यांकन एवं शमन पर ध्यान केंद्रित करके, संस्थान भारत तथा वैश्विक स्तर पर AI प्रशासन के भविष्य को आकार देने में महत्वपूर्ण भूमिका निभा सकता है।
- जैसे-जैसे भारत AI में अग्रणी के रूप में उभर रहा है, AI सुरक्षा संस्थान आगे आने वाली चुनौतियों और अवसरों को संबोधित करने में महत्वपूर्ण भूमिका निभाएगा।

[Source: TH](#)



दैनिक मुख्य परीक्षा अभ्यास प्रश्न

प्रश्न. प्रस्तावित भारत का AI सुरक्षा संस्थान कृत्रिम बुद्धिमत्ता के नैतिक और सामाजिक निहितार्थों को प्रभावी ढंग से कैसे संबोधित कर सकता है, संभावित जोखिमों को कम करते हुए उत्तरदायी और लाभकारी विकास सुनिश्चित कर सकता है?