परिप्रेक्ष्य: डीपफेक से निपटना

प्रिलिम्स के लिये:

डीपफेक टेक्नोलॉजी, डीप सिंथेसिस टेक्नोलॉजी, आर्टिफिशियल इंटेलिजेंस (AI), ब्लॉकचेन तकनीक, आईटी अधिनियम, 2000, आईटी नियम, 2021, आईटी अधिनियम, 2000 की धारा 66D

मेन्स के लिये:

डीपफेक टेक्नोलॉजी का प्रभाव, डीपफेक से निपटना, इसके दुरुपयोग से उत्पन्न नैतिक चिंताएँ

चर्चा में क्यों?

डीपफेक दुनिया भर में लोकतंत्र और सामाजिक संस्थानों के लिये एक गंभीर खतरा बनकर उभरा है। सोशल मीडिया प्लेटफॉर्मों के माध्यम से डीपफेक कॉन्टेंट के प्रसार ने इस चुनौती को बढ़ा दिया है।

- **इलेक्ट्रॉनिक्स और सूचना प्रौद्योगिकी मंत्रालय** ने समय-समय पर सोशल मीडिया प्लेटफॉर्मों को उचित निगरानी करने और डीपफेक के खिलाफ शीघ्र कार्रवाई करने की सलाह दी है।

डीपफेक क्या है?

- **परिचय:**
    - डीपफेक संश्लेषित मीडिया है, जो आमतौर पर किसी को धोखा देने या गुमराह करने के इरादे से दृश्य और ऑडियो सामग्री में परिवर्तन करने या निर्माण करने के लिये आर्टिफिशियल इंटेलिजेंस (AI) का उपयोग करता है।
    - जब वर्ष 2017 में एक Reddit सदस्य ने मशहूर हस्तियों की अश्लील फिल्में अपलोड कीं, तो पहली बार डीपफेक ने ध्यान आकर्षित किया। तब से घटनाओं की अन्य रिपोर्टें देखी गई हैं।
- **डीपफेक निर्माण:**
    - डीपफेक **जेनरेटिव एडवरसैरियल नेटवर्क (GANs)** नामक तकनीक का उपयोग करके बनाए जाते हैं, जिसमें दो प्रतिस्पर्द्धी तंत्रिका नेटवर्क शामिल होते हैं: **एक जनरेटर/उत्पादक और एक डिस्क्रिमीनेटर/विभेदक**।
        - जनरेटर **नकली छवियाँ या वीडियो बनाने की** कोशिश करता है **जो यथार्थवादी दिखते हैं, जबकि विभेदक असली और नकली के बीच अंतर करने** की कोशिश करता है।
        - **जनरेटर विभेदक की प्रतिक्रिया से सीखता है और अपने आउटपुट में सुधार करता है,** जब तक कि वह विभेदक को दुविधा की स्थिति में नहीं ला देता है।
    - **डीपफेक के लिये स्रोत और लक्षित व्यक्ति के फोटो या वीडियो जैसे वृहद् डेटा की आवश्यकता होती है** , जिसे अक्सर उनकी सहमति या जानकारी के बिना इंटरनेट या सोशल मीडिया से एकत्र किया जाता है।
    - डीपफेक **गहन संश्लेषण का एक हिस्सा** है, जो आभासी दृश्य बनाने हेतु पाठ, चित्र, ऑडियो और वीडियो उत्पन्न करने के लिये **गहन शिक्षण** और संवर्द्धित वास्तविकता (Augmented reality) सहित प्रौद्योगिकियों का उपयोग करता है।

डीपफेक टेक्नोलॉजी के क्या उपयोग हैं?

- **फिल्म डबिंग:**
    - डीपफेक तकनीक के उपयोग से विभिन्न भाषाएँ बोलने वाले **अभिनेता सटीक लिप सिंकिंग कर** सकते हैं, **जिससे दुनिया भर के दर्शकों के लिये फिल्म की पहुँच और गहन गुणवत्ता बढ़ जाती है।**
        - उदाहरणतः मलेरिया को समाप्त करने के लिये एक याचिका शुरू करने हेतु एक वीडियो बनाया गया था, जहाँ **डेविड बेकहम, ह्यू जैकमैन और बिल गेट्स** जैसी मशहूर हस्तियों ने डीपफेक तकनीक का उपयोग करके विभिन्न भाषाओं में बात की ।

- **शिक्षा:**
  - शैक्षिक उद्देश्यों के लिये **ऐतिहासिक परिदृश्य का आवाज़ के साथ प्रदर्शन** जैसे **सकारात्मक विकास गहन शिक्षण प्रौद्योगिकी** द्वारा संभव बनाए जा सकते हैं।
    - उदाहरण के लिये **अब्राहम लिंकन के गेटसिबर्ग संबोधन का एक डीपफेक वीडियो का उपयोग** छात्रों को अमेरिकी गृहयुद्ध के बारे में जानकारी देने के लिये किया जा सकता है।
- **कला:**
  - डीपफेक तकनीक का उपयोग कलाकारों के लिये स्वयं को अभिव्यक्त करने, विभिन्न शैलियों के साथ प्रयोग करने या अन्य कलाकारों के साथ सहयोग करने के लिये एक रचनात्मक उपकरण के रूप में किया जा सकता है।
    - उदाहरण के लिये **फ्लोरिडा में अपने संग्रहालय को बढ़ावा देने के लिये साल्वाडोर डाली का एक डीपफेक वीडियो बनाया गया था**, जहाँ उन्होंने आगंतुकों के साथ बातचीत की और उनकी कलाकृतियों पर टिप्पणी की।
- **संदेश का प्रवर्द्धन और उसकी पहुँच:**
  - डीपफेक तकनीक उन लोगों की आवाज़ और प्रभाव को बढ़ाने में मदद कर सकती है जिनके पास साझा करने के लिये महत्त्वपूर्ण संदेश हैं, विशेषकर उन लोगों की जो भेदभाव, सेंसरशिप या हिंसा का सामना करते हैं।
    - उदाहरण के लिये **सऊदी अरब सरकार द्वारा मारे गए एक पत्रकार का एक डीपफेक वीडियो** अपना अंतिम संदेश देने और न्याय की गुहार लगाने के लिये **बनाया गया था**।
- **डिजिटल पुनर्निर्माण और सार्वजनिक सुरक्षा:**
  - डीपफेक तकनीक **खोए या क्षतिग्रस्त डिजिटल डेटा को फिर से बनाने में मदद कर सकती है**, जैसे- पुरानी तस्वीरें या वीडियो को पुनर्स्थापित करना, या कम गुणवत्ता वाले फुटेज को बढ़ाना।
  - यह आपातकालीन उत्तरदाताओं, कानून प्रवर्तन या सैन्य कर्मियों के लिये यथार्थवादी प्रशिक्षण सामग्री बनाकर सार्वजनिक सुरक्षा को बेहतर बनाने में भी मदद कर सकती है।
    - उदाहरण के लिये ऐसी स्थिति में कैसे प्रतिक्रिया देनी है, इसके बारे में शिक्षकों को प्रशिक्षित करने के लिये **एक स्कूल शूटिंग का डीपफेक वीडियो** बनाया गया था।
- **नवाचार:**
  - डीपफेक तकनीक मनोरंजन, गेमिंग या मार्केटिंग जैसे विभिन्न क्षेत्रों और उद्योगों में नवाचार को बढ़ावा दे सकती है। यह कहानी कहने, बातचीत, निदान या अनुनय के नए रूपों को सक्षम कर सकती है।
    - उदाहरण के लिये संश्लेषित मीडिया की क्षमता और समाज पर इसके प्रभाव को प्रदर्शित करने के लिये **मार्क जुकरबर्ग का एक डीपफेक वीडियो बनाया गया था**।

डीपफेक को लेकर वर्तमान चिंताएँ क्या हैं?

- **प्रक्रियात्मक चिंताएँ:**
  - डीपफेक एक समस्या है क्योंकि इसका उपयोग विभिन्न दुर्भावनापूर्ण उद्देश्यों के लिये किया जा सकता है:
    - **प्रचार-प्रसार** और फर्ज़ी खबरें फैलाना;
    - **चुनावों** और जनमत को **प्रभावित करना**;
    - व्यक्तियों या संगठनों को ब्लैकमेल करना और **ज़बरन वसूली करना**;
    - मशहूर हस्तियों, राजनेताओं, कार्यकर्त्ताओं और पत्रकारों की प्रतिष्ठा एवं विश्वसनीयता को नुकसान पहुँचाना;
    - **बिना सहमति** और बदला लेने के लिये **अश्लील वीडियो बनाना**।
    - डीपफेक विभिन्न प्रकार से **नुकसान** पहुँचा सकता है, जैसे- **संस्थानों**, मीडिया और लोकतंत्र के प्रति **विश्वास को कम करना** तथा विधि के शासन एवं मानवाधिकारों को कमज़ोर करना।
    - डीपफेक तकनीक **व्यक्तियों की गोपनीयता, गरिमा और प्रतिष्ठा का उल्लंघन कर सकती है** तथा पीड़ितों, विशेषकर महिलाओं के मानसिक स्वास्थ्य और कल्याण को नुकसान पहुँचा सकती है, जो अक्सर इस तरह के दुर्भावनापूर्ण परिवर्तन का लक्ष्य होती हैं।
- **विधिक चिंताएँ:**
  - अमेरिका में डीपफेक वैधता जटिल है। **पीड़ित मानहानि का दावा कर सकते हैं, लेकिन सामग्री को हटाना सेंसरशिप के रूप में देखा जा सकता है**, यह संभावित रूप से पहले संशोधन का उल्लंघन है, जो धर्म, अभिव्यक्ति, सभा और याचिका की स्वतंत्रता की रक्षा करता है।
  - हालाँकि [भूल जाने का अधिकार](), उपयोगकर्त्ता को फेसबुक और गूगल जैसी उन कंपनियों से अनुरोध करने की अनुमति देता है, जिन्होंने उसका डेटा एकत्र कर इसे हटाने का अनुरोध किया है।
- **नैतिक चिंताएँ:**
  - डीपफेक का उपयोग **दुर्भावनापूर्ण उद्देश्यों के लिये किया जाता है, जिसमें बदला लेने के लिये पोर्न और चेहरे की पहचान प्रणालियों को हैक करना शामिल है।**
  - वे मीडिया में विश्वास को कम करते हैं और तथ्य एवं कल्पना के बीच की **रेखाओं को धुँधला करते हैं।**
  - **डीपफेक** द्वारा प्रचारित **फेक न्यूज़ को सच माना जा सकता है**, जिससे संभावित सामाजिक अशांति की स्थिति उत्पन्न हो सकती है।

डीपफेक से निपटने के लिये वैश्विक दृष्टिकोण क्या है?

- **भारत:**
  - भारत में ऐसे विशिष्ट कानून या नियम नहीं हैं जो डीपफेक तकनीक के उपयोग को प्रतिबंधित या विनियमित करते हों।
  - भारत ने ["एथिकल" AI]() उपकरणों के विस्तार पर एक वैश्विक ढाँचे का आह्वान किया है।

- सूचना प्रौद्योगिकी अधिनियम, 2000 की धारा 67 और 67A जैसे मौजूदा कानूनों में ऐसे प्रावधान हैं जो डीपफेक के कुछ पहलुओं पर लागू हो सकते हैं, जैसे- मानहानि से संबंधित कॉन्टेंट को प्रकाशित करना।
  - भारतीय दंड संहिता (1860) की धारा 500 मानहानि के लिये सज़ा का प्रावधान करती है।
  - डिजिटल व्यक्तिगत डेटा संरक्षण अधिनियम, 2023 व्यक्तिगत डेटा के दुरुपयोग के खिलाफ कुछ सुरक्षा प्रदान करता है।
  - सूचना प्रौद्योगिकी नियम, 2021 दूसरों का प्रतिरूपण करने वाली सामग्री और कृत्रिम रूप से **रूपांतरित छवियों को 36 घंटों के अंदर** हटाने का आदेश देता है।
  - गोपनीयता, सामाजिक स्थिरता, **राष्ट्रीय सुरक्षा और** लोकतंत्र पर संभावित प्रभाव को देखते हुए भारत को विशेष रूप से डीपफेक को लक्षित करने के लिये एक व्यापक कानूनी ढाँचा विकसित करने की आवश्यकता है।
- **वैश्विक:**
  - हाल ही में दुनिया के पहले आर्टिफिशियल इंटेलिजेंस सुरक्षा शिखर सम्मेलन 2023 में अमेरिका, चीन और भारत सहित 28 प्रमुख देशों ने आर्टिफिशियल इंटेलिजेंस के संभावित जोखिमों को दूर करने के लिये वैश्विक कार्रवाई की आवश्यकता पर सहमति व्यक्त की।
    - शिखर सम्मेलन में **'ब्लेचली पार्क घोषणा'** में जान-बूझकर दुरुपयोग और AI प्रौद्योगिकियों पर नियंत्रण खोने के जोखिमों को स्वीकार किया गया।
  - **आर्टिफिशियल इंटेलिजेंस पर वैश्विक साझेदारी (GPAI)** शिखर सम्मेलन दिसंबर 2023 में नई दिल्ली में आयोजित किया गया था। शिखर सम्मेलन कृत्रिम बुद्धिमत्ता पर नई दिल्ली घोषणा को अपनाने के साथ संपन्न हुआ।
    - घोषणा ने GPAI सदस्यों के बीच **सुरक्षित, संरक्षित और भरोसेमंद AI** को आगे बढ़ाने और GPAI परियोजनाओं की स्थिरता का समर्थन करने की प्रतिबद्धता पर आम सहमति व्यक्त की।
- **यूरोपीय संघ:**
  - **दुष्प्रचार पर यूरोपीय संघ की आचार संहिता के तहत** तकनीकी कंपनियों को संहिता पर हस्ताक्षर करने के छह महीने के भीतर डीपफेक और नकली खातों की समस्या का समाधान करने की आवश्यकता होती है।
    - गैर-अनुपालन की स्थिति में तकनीकी कंपनियों को अपने वार्षिक वैश्विक कारोबार का **6% तक जुर्माना भरना पड़ सकता है।**
  - यूरोपीय संघ ने कृत्रिम बुद्धिमत्ता के उपयोग को विनियमित करने के लिये **दुनिया का पहला व्यापक कानून पारित किया** है। आर्टिफिशियल इंटेलिजेंस एक्ट (AI एक्ट) का उद्देश्य **आर्टिफिशियल इंटेलिजेंस के लिये एक सामान्य नियामक और कानूनी ढाँचा पेश करना है।**
    - मसौदा विनियमन का उद्देश्य यह सुनिश्चित करना है कि यूरोपीय बाज़ार में रखे गए और यूरोपीय संघ में उपयोग किये जाने वाले AI सिस्टम सुरक्षित हैं तथा मौलिक अधिकारों और यूरोपीय संघ के मूल्यों का सम्मान करते हैं।
- **संयुक्त राज्य अमेरिका:**
  - अमेरिका ने डीपफेक तकनीक का मुकाबला करने में **होमलैंड सिक्योरिटी विभाग** की सहायता के लिये **द्विदलीय डीपफेक टास्क फोर्स अधिनियम** पेश किया।
- **चीन:**
  - चीन ने **गहन संश्लेषण पर व्यापक विनियमन पेश किया**, जो वर्ष 2023 से प्रभावी होगा।
  - दुष्प्रचार पर अंकुश लगाने के उद्देश्य से विनियमन के लिये स्पष्ट लेबलिंग और डीपफेक सामग्री की ट्रेसबिलिटी की आवश्यकता होती है।
  - विनियम तथाकथित **"डीप सिंथेसिस टेक्नोलॉजी"** के प्रदाताओं और उपयोगकर्त्ताओं पर दायित्व निर्धारित करते हैं।
- **टेक कंपनियाँ:**
  - **मेटा और गूगल** जैसी बड़ी तकनीकी कंपनियों ने डीपफेक कॉन्टेंट की समस्या से निपटने के लिये उपायों की घोषणा की है।
    - हालाँकि उनके सिस्टम में अभी भी कमियाँ हैं,जो ऐसी सामग्री के प्रसार की अनुमति देती हैं।
  - **गूगल** ने वॉटरमार्किंग और मेटाडेटा सहित डीपफेक कॉन्टेंट की पहचान करने के लिये उपकरण पेश किये हैं।
  - **वॉटरमार्किंग जानकारी को कॉन्टेंट में शामिल करने से** यह संपादन के लिये प्रतिरोधी हो जाता है, जबकि मेटाडेटा मूल फाइलों को अतिरिक्त संदर्भ प्रदान करता है।

डीपफेक के खतरे से निपटने के लिये क्या किया जाना चाहिये?

- **अन्य देशों से सीखना:**
  - डीपफेक के जीवन चक्र को तीन भागों में विभाजित किया जा सकता है- निर्माण, प्रसार और पता लगाना। AI विनियमन का उपयोग गैर-कानूनी या गैर-सहमति वाले डीपफेक के निर्माण में कमी लाने के लिये किया जा सकता है।
    - **चीन जैसे देश जिन तरीकों से इस तरह के विनियमन का रुख कर रहे हैं,** उनमें से एक यह है कि डीपफेक प्रौद्योगिकियों के प्रदाताओं को अपने वीडियो में मौजूद लोगों की सहमति प्राप्त करने, उपयोगकर्त्ताओं की पहचान सत्यापित करने और उन्हें सहारा देने की आवश्यकता है।
    - **डीपफेक से होने वाले नुकसान को रोकने के लिये कनाडाई दृष्टिकोण** में बड़े पैमाने पर जन जागरूकता अभियान और संभावित कानून शामिल हैं, जो दुर्भावनापूर्ण इरादे से डीपफेक बनाने और वितरित करने को अवैध बना देंगे।
- **सभी AI-जनरेटेड वीडियो में वॉटरमार्क जोड़ना:**
  - AI-जनरेटेड वीडियो में वॉटरमार्क जोड़ना प्रभावी पहचान और एट्रिब्यूशन के लिये आवश्यक है। वॉटरमार्क विभिन्न उद्देश्यों को पूरा करते हुए सामग्री की उत्पत्ति और स्वामित्व को प्रकट करते हैं। **वे सामग्री के निर्माता या स्रोत को, विशेषकर जब विभिन्न संदर्भों में साझा किया जाता है, स्पष्ट करने में सहायता करते हैं।**
    - दृश्यमान वॉटरमार्क अनधिकृत उपयोग के **खिलाफ एक निवारक के रूप में भी कार्य करते हैं**, जिससे यह स्पष्ट हो जाता है कि सामग्री को उसके स्रोत पर वापस खोजा जा सकता है।
    - इसके अलावा **वॉटरमार्क मूल निर्माता के अधिकारों का प्रमाण प्रदान करके जवाबदेही का समर्थन करते हैं**, AI-जनरेटेड सामग्री के लिये कॉपीराइट और बौद्धिक संपदा सुरक्षा के प्रवर्तन को सरल बनाते हैं।

- **उपयोगकर्त्ताओं को अनुचित सामग्री अपलोड करने से रोकना:**
    - ऑनलाइन प्लेटफॉर्म को उपयोगकर्त्ताओं को उनकी सामग्री संबंधी नीतियों के बारे में शिक्षित और सूचित करने के लिये कदम उठाने चाहिये तथा अनुचित सामग्री अपलोड करने से रोकने के लिये उपाय लागू करने चाहिये।
- **डीपफेक डिटेक्शन तकनीकों का विकास और सुधार:**
    - इसमें अधिक परिष्कृत एल्गोरिदम का उपयोग करने के साथ-साथ नए तरीकों को विकसित करना शामिल हो सकता है, जो उनके संदर्भ, मेटाडेटा या अन्य कारकों के आधार पर डीपफेक की पहचान कर सकते हैं।
- **डिजिटल प्रशासन और कानून को सुदृढ़ बनाना:**
    - इसमें **स्पष्ट और सुसंगत कानून एवं नीतियाँ बनाना** शामिल हो सकता है, जो डीपफेक के दुर्भावनापूर्ण उपयोग को परिभाषित और प्रतिबंधित करते हैं, साथ ही डिजिटल नुकसान के पीड़ितों और अपराधियों के लिये प्रभावी उपचार और प्रतिबंध प्रदान करते हैं।
- **मीडिया साक्षरता और जागरूकता बढ़ाना:**
    - इसमें **जनता और मीडिया को डीपफेक के अस्तित्व और संभावित प्रभाव के बारे में शिक्षित करना शामिल हो सकता है** । उपयोगकर्त्ताओं को संदिग्ध सामग्री को सत्यापित करने और रिपोर्ट करने के लिये कौशल एवं उपकरण प्रदान करना, जैसे:
        - **मीडिया में दृश्य और श्रव्य विसंगतियों पर नज़र रखना**।
        - मूल स्रोत या समान छवियों को खोजने के लिये **रिवर्स इमेज सर्च** का उपयोग करना।
        - छवियों या वीडियो की **गुणवत्ता, स्थिरता और प्रामाणिकता का विश्लेषण करने** के लिये **AI-आधारित टूल** का उपयोग करना।
        - मीडिया के स्रोत और अखंडता को सत्यापित करने के लिये **डिजिटल वॉटरमार्किंग** या [ब्लॉकचेन](ब्लॉकचेन) का उपयोग करना।
        - डीपफेक तकनीक और उसके निहितार्थों के बारे में **स्वयं को एवं दूसरों को शिक्षित करना**।
- **डीपफेक प्रौद्योगिकी के नैतिक और ज़िम्मेदारीपूर्ण उपयोग को बढ़ावा देना:**
    - इसमें डीपफेक तकनीक के रचनाकारों और उपयोगकर्त्ताओं के लिये आचार संहिता एवं मानकों को स्थापित करना तथा उन्हें लागू करना, साथ ही इसके सकारात्मक और लाभकारी अनुप्रयोगों को प्रोत्साहित करना शामिल हो सकता है।
- **ब्लॉकचेन-आधारित डीपफेक सत्यापन:**
    - डिजिटल मीडिया का कॉन्टेंट किसने बनाया, इसका अपरिवर्तनीय रिकॉर्ड बनाने और इसकी प्रामाणिकता को सत्यापित करने में पारदर्शिता सुनिश्चित करने के लिये **ब्लॉकचेन तकनीक** का उपयोग करना।
    - यह विकेंद्रीकृत दृष्टिकोण व्यक्तियों को मीडिया की उत्पत्ति और संशोधन के इतिहास का पता लगाने की अनुमति देता है, जिससे दुर्भावनापूर्ण डीपफेक के निर्माण एवं प्रसार को हतोत्साहित किया जा सकता है।

UPSC सिविल सेवा परीक्षा, विगत वर्ष के प्रश्न

??????????:

प्रश्न. विकास की वर्तमान स्थिति में कृत्रिम बुद्धिमत्ता (Artificial Intelligence) निम्नलिखित में से किस कार्य को प्रभावी रूप से कर सकती है?

1. औद्योगिक इकाइयों में विद्युत की खपत कम करना
2. सार्थक लघु कहानियों और गीतों की रचना
3. रोगों का निदान
4. टेक्स्ट से स्पीच (Text- to- Speech) में परिवर्तन
5. विद्युत ऊर्जा का बेतार संचरण

नीचे दिये गए कूट का प्रयोग कर सही उत्तर चुनिये-

(a) केवल 1, 2, 3 और 5
(b) केवल 1, 3 और 4
(c) केवल 2, 4 और 5
(d) 1, 2, 3, 4 और 5

उत्तर: (b)

?????:

प्रश्न. "चौथी औद्योगिक क्रांति (डिजिटल क्रांति) के प्रादुर्भाव ने ई-गवर्नेंस को सरकार का अविभाज्य अंग बनाने की पहल की है"। विवेचना कीजिये। (2020)

PDF Refernece URL: https://www.drishtiias.com/hindi/printpdf/perspective-combating-deepfakes