डीपफेक: अवसर, खतरे और विनियिमन

यह एडिटोरियल 06/11/2023 को 'इंडियन एक्सप्रेस' में प्रकाशित "Rashmika Mandanna's deepfake: Regulate AI, don't ban it" लेख पर आधारित है। इसमें अभिनेत्री रश्मिका मंदाना की हाल में वायरल हुई 'डीपफेक' वीडियो के संदर्भ में चर्चा की गई है और ऐसी प्रौद्योगिकियों के विनियमन के लिये एक समग्र दृष्टिकोण की आवश्यकता पर विचार किया गया है।

प्रिलिम्स के लिये:

डीपफेक, आर्टफिशियिल इंटेलिजेंस (AI), आईटी अधिनियिम, 2000 और आईटी नियम, आईटी अधिनियिम की धारा 66D,

मेन्स के लिये:

डीपफेक: उपयोग, चुनौतियाँ, सरकार द्वारा निर्धारित नियम एवं आगे की राह

हाल ही में एक फैक्ट-चेकर वेबसाइट ने खुलासा किया कि लिफ्ट में प्रवेश करती अभिनेत्री रश्मिका की वायरल वीडियो वस्तुतः **'डीपफेक' (Deepfake)** है। इस वीडियो ने एक बहस छेड़ दी है जहाँ अन्य अभिनेताओं द्वारा डीपफेक वीडियो के कानूनी विनियिमन की माँग की जा रही है। इसकी प्रतिक्रिया में इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी (IT) राज्य मंत्री ने **आईटी अधिनियिम, 2000** के अंतर्गत मौजूद विनियिमनों का हवाला दिया है जो ऐसे वीडियो के प्रसार से निपट सकते हैं। हालाँकि, डीपफेक के विनियिमन के लिये एक समग्र दृष्टिकोण के तहत प्लेटफॉर्म और **कृत्रिम बुद्धिमत्ता (Artificial Inteliigence-AI)** विनियिमन के बीच की अंतःक्रिया और उभरती प्रौद्योगिकियों के लिये सुरक्षा उपायों को अधिक व्यापक रूप से शामिल करने के तरीकों पर ध्यान केंद्रित करना आवश्यक होगा।

डीपफेक क्या है?

- **डीपफेक (Deepfake)** शब्द सिंथेटिक मीडिया को संदर्भित करता है जहाँ किसी व्यक्ति की सदृश्ता को दूसरे व्यक्ति की सदृश्ता से बदलने के लिये डिजिटल रूप से हेरफेर किया जाता है।
- **डीपफेक** मशीन लर्निंग और AI के प्रभावशाली तकनीकों—जैसे कि डीप लर्निंग (deep learning) और जेनरेटिव एडवरसैरियल नेटवर्क (Generative Adversarial Networks- GANs) का उपयोग कर सृजित किये जाते हैं।
- डीपफेक तकनीक का उपयोग मनोरंजन, शिक्षा, कला और सक्रिय गतिविधियों (activism) जैसे विभिन्न उद्देश्यों के लिये किया जा सकता है।
  - हालाँकि, **यह गंभीर नैतिक और सामाजिक चुनौतियाँ भी उत्पन्न कर सकता है,** जैसे फेक न्यूज़ सृजित करना, भ्रामक सूचना का प्रसार करना, निजिता/गोपनीयता का उल्लंघन करना और किसी व्यक्ति की प्रतिष्ठा को हानि पहुँचाना।
  - इसका **उपयोग** नकली या **फेक वीडियो बनाने के लिये** किया जा सकता है; इसका उपयोग स्कैमर्स द्वारा मित्रों या प्रियजनों का रूप धारण कर लोगों से धोखाधड़ीपूर्ण तरीके से धन प्राप्त करने के लिये भी किया जा सकता है।

डीपफेक टेक्नोलॉजी के उपयोग

- **फिल्म डबिंग:** डीपफेक तकनीक का उपयोग विभिन्न भाषाएँ बोलने वाले अभिनेताओं के लिये यथार्थपरक लिपि-सिंकिंग (lip-syncing) के सृजन के लिये किया जा सकता है, जिससे **उक्त फिल्म वैश्विक दर्शकों के लिये अधिक अभिगम्य (accessible and immersive)** हो जाती है।
  - उदाहरण के लिये, मलेरिया के उन्मूलन का आह्वान करने के लिये एक याचिका शुरू करने के लिये एक वीडियो बनाया गया था, जहाँ डीपफेक तकनीक का उपयोग कर डेविड बेकहम, ह्यूज जैकमैन और बिल गेट्स जैसी मशहूर हस्तियों से विभिन्न भाषाओं में आह्वान कराया गया था।
- **शिक्षा:** डीपफेक तकनीक **कक्षा में ऐतिहासिक व्यक्तियों को जीवंत करने या विभिन्न परिदृश्यों के इंटरैक्टिव सिमुलेशन** बनाने के रूप में शिक्षकों के लिये आकर्षक पाठ प्रदान कर सकने में मदद कर सकती है।
  - उदाहरण के लिये, अब्राहम लिंकन के प्रसिद्ध गेटसिबर्ग संबोधन के डीपफेक वीडियो का उपयोग छात्रों को अमेरिकी गृहयुद्ध के बारे में शिक्षण प्रदान करने 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 साधन के रूप में किया जा सकता है।
  - उदाहरण के लिये, फ्लोरिडा में साल्वाडोर डाली के संग्रहालय को बढ़ावा देने के लिये डाली का एक डीपफेक वीडियो सृजित किया गया था,

जहाँ उन्होंने आगंतुकों के साथ संवाद किया और अपनी कृतियों पर टिप्पणी की।

- **स्वायत्तता और अभिव्यक्ति:** डीपफेक तकनीक लोगों को अपनी डिजिटल पहचान को नियंत्रित करने, अपनी गोपनीयता की रक्षा करने या विभिन्न तरीकों से अपनी पहचान व्यक्त करने के लिये सशक्त बना सकती है।
    - उदाहरण के लिये, **रिफेस (Reface) नामक एक डीपफेक ऐप** उपयोगकर्त्ताओं को मनोरंजन या वैयक्तिकरण के लिये वीडियो या जिफ (gifs) के रूप में मशहूर हस्तियों या चरित्रों के साथ अपना चेहरा बदलने की अनुमति देता है।
- **संदेश और उसकी पहुँच का विस्तार:** डीपफेक तकनीक **उन लोगों की आवाज़ और प्रभाव को बढ़ाने में मदद कर सकती है जिनके पास साझा करने के लिये महत्वपूर्ण संदेश हैं**, विशेष रूप से वे लोग जो भेदभाव, सेंसरशिप या हिंसा का सामना कर रहे हैं।
    - उदाहरण के लिये, सऊदी सरकार द्वारा हत्या करा दिये गए एक पत्रकार का डीपफेक वीडियो बनाया गया जहाँ उसने अपना अंतिम संदेश दिया और न्याय की गुहार लगाई।
- **डिजिटल पुनर्निर्माण और सार्वजनिक सुरक्षा:** डीपफेक तकनीक **गुम या क्षतिग्रस्त डिजिटल डेटा के पुनर्निर्माण** में मदद कर सकती है, जैसे पुरानी तस्वीरों या वीडियो का पुनर्निर्माण या निम्न गुणवत्ता वाले फुटेज को बेहतर बनाना।
    - यह आपातकालीन प्रतिक्रियाकर्त्ताओं, कानून प्रवर्तन एजेंसियों या सैन्य कर्मियों के लिये**यथार्थवादी प्रशिक्षण सामग्री का सृजन कर सार्वजनिक सुरक्षा में सुधार** लाने में भी मदद कर सकता है।
    - उदाहरण के लिये, स्कूल में गोलीचालन (शूटिंग) का एक डीपफेक वीडियो बनाया गया ताकि शिक्षकों को प्रशिक्षित किया जा सके कि वे ऐसे परिदृश्य में किस प्रकार प्रतिक्रिया दें।
- **नवाचार:** डीपफेक तकनीक **मनोरंजन, गेमिंग या मार्केटिंग जैसे विभिन्न क्षेत्रों और उद्योगों में नवाचार को बढ़ावा दे सकती है।** यह स्टोरीटेलिंग, इंटरैक्शन, डायग्नोसिस या प्रत्यायन/अनुनय (persuasion) के नए रूपों को सक्षम कर सकता है।
    - उदाहरण के लिये, सिंथेटिक मीडिया की क्षमता और समाज पर इसके प्रभाव को प्रदर्शित करने के लिये मार्क जुकरबर्ग का एक डीपफेक वीडियो बनाया गया था।

डीपफेक प्रौद्योगिकी से संबद्ध चुनौतियाँ

- **झूठी सूचना का प्रसार:** डीपफेक का उपयोग जानबूझकर झूठी जानकारी या गलत सूचना के प्रसार के लिये किया जा सकता है, जो महत्वपूर्ण विषयों के बारे में भ्रम पैदा कर सकता है।
    - उदाहरण के लिये, राजनेताओं या मशहूर हस्तियों के डीपफेक वीडियो का इस्तेमाल जनता की राय को प्रभावित करने या चुनावों को प्रभावित करने के लिये किया जा सकता है।
- **उत्पीड़न और धमकी:** डीपफेक को लोगों को परेशान करने, भयभीत करने, नीचा दिखाने और कमज़ोर करने के लिये डिज़ाइन किया जा सकता है।
    - उदाहरण के लिये, डीपफेक तकनीक अन्य अनैतिक कार्यों को बढ़ावा दे सकती है, जैसे कि रिवेंज पोर्न (revenge porn) बनाना, जिससे महिलाएँ असंगत रूप से हानि उठाती हैं।
        - डीपफेक पोर्न पीड़ितों की निजता एवं सहमति का भी उल्लंघन कर सकता है और मनोवैज्ञानिक संकट एवं आघात (trauma) का कारण बन सकता है।
    - डीपफेक प्रौद्योगिकी का उपयोग ब्लैकमेल या फिरौती की सामग्री के निर्माण के लिये किया जा सकता है, जैसे किसी व्यक्ति द्वारा कोई अपराध करने, प्रेम प्रसंग रखने या खतरे में होने के नकली वीडियो बनाना।
        - उदाहरण के लिये, एक राजनेता का डीपफेक वीडियो बनाया गया और इसे सार्वजनिक नहीं करने के बदले धन की मांग की गई।
- **झूठे साक्ष्य गढ़ना:** डीपफेक का उपयोग झूठे साक्ष्य गढ़ने के लिये किया जा सकता है, जिसका उपयोग फिर जनता को धोखा देने या राज्य की सुरक्षा को हानि पहुँचाने के लिये किया जा सकता है। डीपफेक साक्ष्य का उपयोग कानूनी कार्यवाही या जाँच में हेरफेर करने के लिये भी किया जा सकता है।
    - उदाहरण के लिये, डीपफेक ऑडियो या वीडियो का उपयोग किसी की पहचान या आवाज़ का प्रतिरूपण करने और झूठे दावे करने या आरोप लगाने के लिये किया जा सकता है।
- **प्रतिष्ठा धूमिल करना:** डीपफेक का उपयोग किसी व्यक्ति की ऐसी छवि बनाने के लिये किया जा सकता है जैसा वह नहीं है या किसी को कुछ ऐसा कहते या करते हुए दिखाने के लिये जैसा उसने कभी नहीं किया या किसी व्यक्ति की आवाज़ को ऑडियो फ़ाइल में संश्लेषित करने के लिये किया जा सकता है, जिसका उपयोग फिर उसकी प्रतिष्ठा को धूमिल करने के लिये किया जा सकता है।
    - उदाहरण के लिये, डीपफेक मीडिया का उपयोग किसी व्यक्ति या संगठन की विश्वसनीयता या विश्वस्तता को हानि पहुँचाने और प्रतिष्ठा संबंधी या वित्तीय नुकसान करने के लिये किया जा सकता है।
- **वित्तीय धोखाधड़ी:** डीपफेक तकनीक का उपयोग अधिकारियों, कर्मचारियों या ग्राहकों का रूप धारण करने और उन्हें संवेदनशील जानकारी प्रकट करने, धन हस्तांतरित करने या गलत निर्णय लेने के लिये भ्रमित करने हेतु किया जा सकता है।
    - उदाहरण के लिये, एक CEO के डीपफेक ऑडियो का इस्तेमाल एक कर्मचारी को भ्रमित करने और धोखाधड़ीपूर्ण खाते में 2,43,000 अमेरिकी डॉलर हस्तांतरित करने के लिये किया गया था।

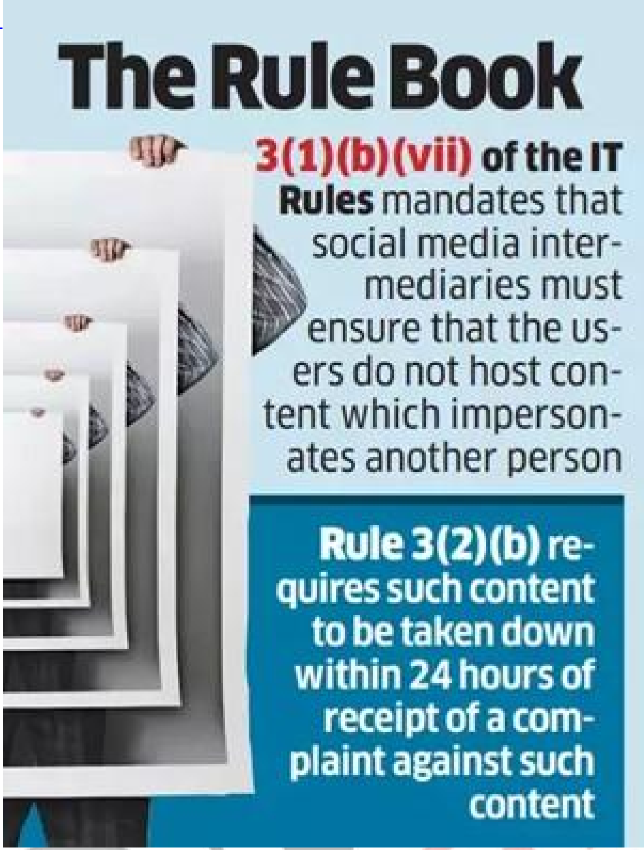
डीपफेक पर अंकुश लगाने के लिये सरकार द्वारा प्रवर्तित नियम

- **आईटी अधिनियम 2000 और आईटी नियम 2021:** आईटी अधिनियम और आईटी नियम दोनों में स्पष्ट निर्देश दिये गए हैं जो सोशल मीडिया मध्यस्थों पर ज़िम्मेदारी डालते हैं कि **वे यह सुनिश्चित करें कि इस तरह के डीपफेक वीडियो या फोटो को जल्द से जल्द हटा दिया जाएगा।** ऐसा न करने पर तीन वर्ष तक की कैद और 1 लाख रुपए के जुर्माने का प्रावधान है।
    - आईटी अधिनियम की धारा 66D: आईटी अधिनियम 2000 की धारा 66D में कहा गया है कि **जो कोई भी संचार उपकरण या कंप्यूटर संसाधन का उपयोग कर प्रतिरूपण के माध्यम से धोखाधड़ी (cheating by personating) करता है,** उसे तीन वर्ष तक की कैद और एक लाख रुपए तक के जुर्माने से दंडित किया जा सकता है।
    - नियम 3(1)(b)(vii): यह नियम कहता है कि सोशल मीडिया मध्यस्थों को यह **सुनिश्चित करना होगा कि उनके प्लेटफॉर्म के**

उपयोगकर्ता किसी भी ऐसी सामग्री को होस्ट न करें जो किसी अन्य व्यक्ति का प्रतिरूपण करती हो।

- **[नियम 3(2)(b)](#):** ऐसी किसी सामग्री के विरुद्ध शिकायत प्राप्त होने के 24 **घंटे के भीतर ऐसी सामग्री को हटाना** आवश्यक है।

[//](#)



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

- **अन्य देशों के अनुभव से सीखना:** डीपफेक के जीवनचक्र को तीन भागों में विभाजित किया जा सकता है -**निर्माण, प्रसार और इसका पता लगाना।** AI विनियमन का उपयोग गैरकानूनी या गैर-सहमति वाले डीपफेक के निर्माण के शमन के लिये किया जा सकता है।
  - **चीन जैसे देश** जिन तरीकों से इस तरह के विनियमन की ओर आगे बढ़ 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 साधन या अवलंब (recourse) प्रदान करने की आवश्यकता है।
  - डीपफेक से होने वाले नुकसान को रोकने के लिये **कनाडा का दृष्टिकोण** यह रहा है कि व्यापक जन जागरूकता अभियान चलाए जाएँ और ऐसे विधान बनाए जाएँ जो दुर्भावनापूर्ण इरादे से डीपफेक के सृजन एवं वितरण को अवैध बना दें।
- **सभी AI-जनित वीडियो में वॉटरमार्क जोड़ना:** AI-जनरेटेड वीडियो में वॉटरमार्क जोड़ना प्रभावी पहचान और श्रेय (attribution) के लिये आवश्यक है। वॉटरमार्क विभिन्न उद्देश्यों को पूरा करते हुए सामग्री के उद्गम और स्वामित्व को प्रकट करते हैं। वे सामग्री के निर्माता या स्रोत को स्पष्ट करके इसकी पहचान में सहायता करते हैं, विशेष रूप से जब इन्हें विभिन्न संदर्भों में साझा किया जाता है।
  - दृश्यमान वॉटरमार्क अनधिकृत उपयोग के विरुद्ध एक निवारक के रूप में भी कार्य करते हैं, जिससे यह स्पष्ट हो जाता है कि सामग्री के स्रोत या उद्गम का पता लगाया जा सकता है।
  - इसके अलावा, वॉटरमार्क मूल निर्माता के अधिकारों का प्रमाण प्रदान कर जवाबदेही का समर्थन करते हैं; इस प्रकार AI-जनित सामग्री के लिये कॉपीराइट और बौद्धिक संपदा सुरक्षा के प्रवर्तन को सरल बनाते हैं।
- **उपयोगकर्ताओं को अनुचित सामग्री अपलोड करने से रोकना:** ऑनलाइन प्लेटफॉर्म को उपयोगकर्ताओं को उनकी सामग्री नीतियों के बारे में शिक्षित और सूचित करने के लिये कदम उठाने चाहिए तथा उन्हें अनुचित सामग्री अपलोड करने से रोकने के लिये उपाय भी लागू करने चाहिए।
- **डीपफेक डिटेक्शन तकनीकों का विकास और सुधार:** इसमें अधिक परिष्कृत एल्गोरिदम का उपयोग करने के साथ ही नए तरीकों को विकसित करना शामिल हो सकता है जो डीपफेक के संदर्भ, मेटाडेटा या अन्य कारकों के आधार पर उनकी पहचान कर सकते हैं।
- **डिजिटल शासन और विधान को सुदृढ़ करना:** यह ऐसे स्पष्ट एवं संगत कानूनों एवं नीतियों को संलग्न कर सकता है जो डीपफेक के दुर्भावनापूर्ण उपयोग को परिभाषित एवं प्रतिबंधित करते हैं, साथ ही डिजिटल नुकसान के पीड़ितों और अपराधियों के लिये प्रभावी उपचार एवं प्रतिबंध प्रदान करते हैं।
- **मीडिया साक्षरता और जागरूकता बढ़ाना:** इसमें जनता और मीडिया को डीपफेक के अस्तित्व एवं संभावित प्रभाव के बारे में शिक्षित करना, साथ ही उन्हें संदिग्ध सामग्री को सत्यापित करने और रिपोर्टिंग करने के लिये कौशल एवं साधन प्रदान करना शामिल हो सकता है।
- **डीपफेक प्रौद्योगिकी के नैतिक और उत्तरदायी उपयोग को बढ़ावा देना:** इसमें डीपफेक प्रौद्योगिकी के सृजनकर्ताओं और उपयोगकर्ताओं के लिये आचार संहिता एवं मानकों की स्थापना और कार्यान्वयन के साथ ही इसके सकारात्मक एवं लाभकारी अनुप्रयोगों को प्रोत्साहित करना शामिल हो सकता है।

अभ्यास प्रश्न: डीपफेक प्रौद्योगिकी के संभावित उपयोग एवं खतरों की चर्चा कीजिये। इस संदर्भ में, उन उपायों पर भी विचार कीजिये जो सरकारें और प्रौद्योगिकी कंपनियाँ डीपफेक के नकारात्मक परिणामों के शमन के लिये अपना सकती हैं।

PDF Refernece URL: https://www.drishtiias.com/hindi/printpdf/deepfakes-opportunities,-threats,-and-regulation