प्रोजेक्ट ग्लासविंग: AI से साइबर सुरक्षा में खोज से सुधार तक की दूरी

कृत्रिम बुद्धिमत्ता, साइबर सुरक्षा और महत्वपूर्ण अवसंरचना (सामान्य अध्ययन-III) | समाचार संदर्भ: जून 2026 में अंतरराष्ट्रीय विस्तार और भारत की साइबर एजेंसियों की भागीदारी संबंधी रिपोर्ट; अक्टूबर के बदलाव सहित। | प्रकाशित: 24 अक्टूबर 2026

खबरों में क्यों

जून 2026 में प्रोजेक्ट ग्लासविंग के विस्तार की घोषणा हुई। कंपनी Anthropic ने कहा कि वह लगभग 150 नए संगठनों तक अपनी साझेदारी बढ़ाएगी और पहुंच देने से पहले सुरक्षा संबंधी शर्तें जांचेगी। शुरुआती लगभग 50 भागीदारों के साथ प्रस्तावित समूह करीब 200 संगठनों का बनता था। इसे सभी संगठनों को तत्काल पहुंच मिल जाने का प्रमाण नहीं मानना चाहिए।

भारत में कुछ सरकारी साइबर संस्थाओं को पहुंच मिलने और बड़ी सूचना-प्रौद्योगिकी कंपनियों से बातचीत की खबर भी आई। यह जानकारी रिपोर्टों में अधिकारियों के हवाले से थी। सभी नामों की सार्वजनिक, स्वतंत्र आधिकारिक पुष्टि उपलब्ध होने का दावा यहां नहीं किया जा रहा है।

इस विषय को जून पर रोकना भी अधूरा होगा। 6 अक्टूबर को कंपनी ने ग्लासविंग और अपने साइबर सत्यापन कार्यक्रम को विस्तारित व्यवस्था में जोड़ने की घोषणा की। 8 अक्टूबर को खुले स्रोत वाले सॉफ्टवेयर के लिए स्वैच्छिक स्कैनर की घोषणा हुई। इसलिए शुरुआती सीमित साझेदारी को आज की स्थिर और अपरिवर्तित व्यवस्था के रूप में नहीं पढ़ना चाहिए।

पहले समझें: सुरक्षा भेद्यता क्या है?

सॉफ्टवेयर या उसकी व्यवस्था में ऐसी कमजोरी, जिसका दुरुपयोग करके कोई अनधिकृत काम किया जा सके, सुरक्षा भेद्यता कहलाती है। हर तकनीकी त्रुटि समान खतरा नहीं होती। यह देखना पड़ता है कि उसे वास्तव में इस्तेमाल किया जा सकता है या नहीं, उससे क्या नुकसान होगा और किन संस्करणों पर उसका असर है।

ग्लासविंग क्या करता था?

PCSMANTRA.COM

ग्लासविंग का शुरुआती उद्देश्य चुनिंदा संगठनों को उन्नत AI की सहायता से महत्वपूर्ण सॉफ्टवेयर में सुरक्षा कमजोरियां खोजने और सुधारने का अवसर देना था। इसमें Claude Mythos Preview नामक मॉडल का उपयोग बताया गया। यह सामान्य उपभोक्ता के लिए उपलब्ध किसी सार्वभौमिक सुरक्षा गारंटी का नाम नहीं था।

कंपनी का कहना था कि ऐसे मॉडल जटिल कोड की जांच करके कमजोरियों की पहचान में मदद कर सकते हैं। इस दावे का महत्व है, लेकिन कंपनी का प्रचार, परीक्षण का परिणाम और स्वतंत्र रूप से सिद्ध वास्तविक सुरक्षा अलग चीजें हैं। मॉडल के सुझाव की तकनीकी जांच और सुधार के बाद परीक्षण जरूरी रहता है।

AI की दोहरी उपयोगिता:

एक ही कमजोरी की जानकारी से सुरक्षा टीम उसे ठीक कर सकती है और हमलावर उसका दुरुपयोग कर सकता है। इसीलिए साइबर क्षेत्र में उपयोगिता और जोखिम साथ आते हैं। तेजी से खोजी गई कमजोरियां तभी सुरक्षा बढ़ाती हैं

जब उनकी सूचना सही लोगों तक सुरक्षित तरीके से पहुंचे और समय पर सुधार लागू हो।

यह समस्या केवल किसी एक मॉडल की नहीं है। बेहतर स्वचालित उपकरण सुरक्षा विश्लेषण का समय घटा सकते हैं, मगर गलत रिपोर्टों का बोझ भी बढ़ा सकते हैं। किसी मॉडल को शक्तिशाली बताना इस बात का प्रमाण नहीं कि वह हर हमले को रोक लेगा या सभी संगठनों को एक जैसी मदद देगा।

खोज से सुरक्षा तक:

पड़ाव	क्या करना होता है?	क्यों जरूरी है?
संभावित कमजोरी की पहचान	संदिग्ध कोड या व्यवहार की रिपोर्ट	यह केवल जांच की शुरुआत है
सत्यापन	देखना कि समस्या वास्तविक, नई और प्रभावशाली है या नहीं	गलत रिपोर्ट तथा दोहराव छांटना
सुधार तैयार करना	कोड या व्यवस्था बदलकर परीक्षण करना	सुझाव से नई खराबी न पैदा हो
सुरक्षित तैनाती	प्रभावित संस्थाओं में सुधार लागू करना	केवल नया संस्करण उपलब्ध होना पर्याप्त नहीं
निगरानी	परिणाम और संभावित दुरुपयोग पर नजर	पुराने या छूटे हुए सिस्टम पहचानना

उदाहरण के लिए, किसी बैंक के सॉफ्टवेयर में सुधार तैयार हो जाए लेकिन उसकी कुछ शाखाओं या जुड़े सेवा-प्रदाताओं में पुराना संस्करण चलता रहे, तो जोखिम बना रह सकता है। इसलिए खोजी गई कमजोरियों की संख्या और वास्तविक रूप से सुरक्षित हुए सिस्टमों की संख्या एक नहीं है।

जिम्मेदार सूचना क्यों?

यदि कमजोरी का पूरा विवरण सुधार से पहले सार्वजनिक हो जाए, तो हमलावरों को अवसर मिल सकता है। दूसरी ओर सूचना को अनिश्चित समय तक दबाए रखना भी उपयोगकर्ताओं के हित में नहीं हो सकता। जिम्मेदार खुलासा इन दोनों चिंताओं के बीच समन्वय की प्रक्रिया है।

CERT-In की नीति में रिपोर्ट की जांच, संबंधित उत्पाद-निर्माता और शोधकर्ता से समन्वय तथा उचित समय पर सूचना जारी करने की व्यवस्था दी गई है। नीति रिपोर्ट करने वाले को लागू कानूनों का पालन करने से छूट नहीं देती। किसी कमजोरी की खोज अपने-आप दूसरे व्यक्ति के सिस्टम में अनधिकृत प्रवेश की अनुमति नहीं बन जाती।

PCSMANTRA.COM

इसलिए परीक्षण का दायरा, अनुमति और संवेदनशील जानकारी की सुरक्षा स्पष्ट होनी चाहिए। यह लेख हमले की विधि नहीं, सुरक्षा-प्रशासन का ढांचा समझाता है।

भारत के लिए महत्व:

बैंकिंग, बिजली, दूरसंचार और स्वास्थ्य जैसी सेवाओं में डिजिटल व्यवधान का असर केवल डेटा तक सीमित नहीं रहता। भुगतान रुकना, सेवाएं बंद होना या संवेदनशील जानकारी का दुरुपयोग सार्वजनिक जीवन को प्रभावित कर सकता है। महत्वपूर्ण सूचना अवसंरचना की सुरक्षा के लिए भारत में NCIIPC राष्ट्रीय नोडल संस्था है। CERT-In की भूमिका साइबर घटनाओं और भेद्यता संबंधी समन्वय में महत्वपूर्ण है। दोनों को एक ही संस्था नहीं समझना चाहिए।

AI उपकरण अपनाने से पहले संस्थाओं को यह देखना होगा कि कौन-सा कोड या डेटा बाहर जाएगा, उसे कौन देख सकेगा और जानकारी का उपयोग किन शर्तों पर होगा। विदेशी या घरेलू सेवा होने से अकेले सुरक्षा का निर्णय नहीं हो जाता। वास्तविक अनुबंध, डेटा-सुरक्षा, पहुंच और तकनीकी नियंत्रण देखने पड़ते हैं।

अक्टूबर में क्या बदला?

6 अक्टूबर की कंपनी घोषणा में विस्तारित साइबर सत्यापन कार्यक्रम के अलग-अलग पहुंच स्तर बताए गए। रक्षा-संबंधी काम, अधिकृत सुरक्षा परीक्षण और विशेष रूप से संवेदनशील प्रणालियों के परीक्षण के लिए अलग शर्तें रखी गईं। पुराने ग्लासविंग भागीदारों को विशेष स्तर में स्थानांतरित करने की व्यवस्था बताई गई। यह घोषणा है, हर भारतीय संस्था की मंजूरी या वर्तमान उपयोग की अलग पुष्टि नहीं।

8 अक्टूबर को घोषित OSS Scanner खुले स्रोत वाली परियोजनाओं के लिए स्वैच्छिक सेवा है। कंपनी ने स्पष्ट किया कि इस रास्ते से मिलने वाली रिपोर्टें बिना मानव समीक्षा के मॉडल द्वारा तैयार हो सकती हैं और गलत या अमान्य भी हो सकती हैं। इसे मानव-सत्यापित समन्वित खुलासा प्रक्रिया के समान नहीं समझना चाहिए।

यही अंतर सार्वजनिक नीति के लिए महत्वपूर्ण है। तेज रिपोर्ट मिलना लाभ है, लेकिन उसे जांचने के लिए समय और विशेषज्ञ चाहिए। छोटी परियोजना को बड़ी संख्या में रिपोर्ट दे देना उसकी क्षमता बढ़ाए बिना उलटा दबाव डाल सकता है।

दावों का मूल्यांकन:

उम्मीदवार और सत्यापित कमजोरी: संभावित समस्या, मानव-सत्यापित समस्या, पहले से ज्ञात समस्या और दोहराई गई रिपोर्ट अलग श्रेणियां हैं। उन्हें जोड़कर उपलब्धि का बड़ा आंकड़ा बनाना गलत होगा।

गंभीरता और संदर्भ: परीक्षण में गंभीर दिखने वाली समस्या का वास्तविक असर सिस्टम की पहुंच और सुरक्षा व्यवस्था पर निर्भर कर सकता है। दूसरी ओर छोटी दिखने वाली कई कमजोरियां मिलकर बड़ा जोखिम भी बना सकती हैं।

खोज और सुधार: अधिक कमजोरियां मिलना बेहतर खोज का संकेत हो सकता है, पर इससे अपने-आप सभी सुधार लागू होने का निष्कर्ष नहीं निकलता। सुधार की गुणवत्ता, तैनाती का समय और सेवा की निरंतरता अलग मापदंड हैं।

कंपनी और स्वतंत्र प्रमाण: कंपनी के परीक्षण या भागीदारों के सर्वेक्षण को उसके स्रोत के साथ पढ़ना चाहिए। अलग नमूनों, अवधियों और परिभाषाओं वाले आंकड़ों की सीधी तुलना नहीं की जा सकती।

POSMANTRA.COM

आगे का रास्ता:

भारत के लिए प्राथमिकता किसी एक उत्पाद का नाम अपनाना नहीं, बल्कि सुरक्षा की पूरी श्रृंखला मजबूत करना है। सॉफ्टवेयर की सूची, जिम्मेदार टीम, अधिकृत परीक्षण, सुरक्षित रिपोर्टिंग, सुधार की जांच और समय पर तैनाती साथ चलें। छोटे सेवा-प्रदाताओं और खुले स्रोत वाली परियोजनाओं की क्षमता भी बढ़ानी होगी।

AI सुरक्षा विशेषज्ञ का सहायक बन सकता है। वह संस्थागत जवाबदेही, कानूनी अनुमति और तकनीकी सत्यापन का विकल्प नहीं है। ग्लासविंग की चर्चा का सबसे उपयोगी संदेश यही है कि कमजोरी खोजने की गति बढ़े तो उसे जिम्मेदारी से सुधारने की क्षमता भी उसी गति से बढ़नी चाहिए।

मुख्य परीक्षा अभ्यास:

प्रश्न: "AI की सहायता से सुरक्षा कमजोरियां तेजी से खोजना तभी उपयोगी है जब संस्थाओं की सत्यापन और सुधार लागू करने की क्षमता भी बढ़े।" महत्वपूर्ण सूचना अवसंरचना के संदर्भ में चर्चा कीजिए। (250 शब्द, 15 अंक)

परिचय: AI की दोहरी उपयोगिता और डिजिटल सेवाओं पर निर्भरता बताएं।

मुख्य भाग: पहचान, सत्यापन, जिम्मेदार खुलासा, सुधार और तैनाती समझाएं। डेटा-सुरक्षा, अधिकृत परीक्षण, छोटे रखरखावकर्ताओं की क्षमता और स्वतंत्र मूल्यांकन की चुनौतियां जोड़ें।

निष्कर्ष: उत्पाद-केंद्रित दृष्टिकोण के बजाय जवाबदेह संस्थागत सुरक्षा व्यवस्था पर जोर दें।

स्रोत और सीमाएँ

- इसी विषय का प्रारंभिक परीक्षा (तथ्य) लेख: प्रोजेक्ट ग्लासविंग और AI साइबर सुरक्षा — प्रमुख तथ्य
- 2 जून 2026: ग्लासविंग विस्तार की कंपनी घोषणा
- ग्लासविंग विस्तार और दावों पर स्वतंत्र रिपोर्ट
- जून 2026: भारतीय संस्थाओं की पहुंच पर रिपोर्ट
- CERT-In: जिम्मेदार भेद्यता खुलासा और समन्वय नीति
- NCIIPC की आधिकारिक भूमिका
- 6 अक्टूबर 2026: साइबर सत्यापन कार्यक्रम में बदलाव
- 8 अक्टूबर 2026: स्वैच्छिक OSS Scanner और उसकी सीमाएं
- **सीमाएँ:** जून का विस्तार और अक्टूबर के बाद के बदलाव (6 अक्टूबर CVP एकीकरण, 8 अक्टूबर OSS स्कैनर) अलग चरण के रूप में बताए गए हैं। किसी खोज को सत्यापित समस्या, जांचा गया पैच या तैनात सुधार नहीं माना गया। कंपनी के प्रदर्शन-आँकड़े और पूर्वानुमान नहीं दिए गए। भारत में पहुंच के बारे में बिना नाम अधिकारी के हवाले रिपोर्टिंग है; किसी एजेंसी की पुष्टि का दावा नहीं। CERT-In/NCIIPC को प्राथमिक संदर्भ के रूप में बताया गया है; रिपोर्टिंग परीक्षण की अनुमति नहीं बनाती। कोई शोषण-निर्देश या सार्वभौमिक AI-सुरक्षा दावा नहीं।

PCSMANTRA.COM