Boletín de Ciberseguridad

Riesgo de vulnerabilidades IT y OT

Mayo - 2024

Sobre Axus

Axus es una empresa especializada en ciberseguridad, con presencia en el Perú y Latinoamérica. Asesora a clientes en la protección de su información e infraestructura tanto en ámbitos industriales (OT) como en ámbitos corporativos (IT) y viene registrando una creciente participación en el ámbito de Internet de las Cosas (IoT) y en la protección de entornos Cloud e Híbridos.

En base a su reputación como especialista con los principales fabricantes de la industria y su equipo profesional en permanente capacitación, Axus genera valor a través de una visión consultiva que busca definir soluciones a la medida de las necesidades de cada cliente.

Alertas de Ciberseguridad

A continuación, compartimos con ustedes algunas de las alertas más relevantes de mayo.

Alertas de Seguridad IT:

- Grave Brecha de Seguridad en Replicate: Modelos de IA de Clientes al Descubierto
- Cibercrimen y la Carrera por Dominar la IA: ¿Quién Lleva la Delantera?
- Impacto en Cadena: Filtraciones de Proveedores Dejan al Descubierto a Grandes Empresas Españolas

Alertas de Seguridad OT/ICS:

- Falla Crítica en Controlador Virtual de Honeywell Permite Ejecución Remota de Código
- Sistemas de Agua en EE.UU. Vulnerables: Autoridades Exigen Plan de Ciberseguridad Urgente
- Amenaza en Ascenso: Hacktivistas de Rusia Atacan Infraestructura Crítica

Grave Brecha de Seguridad en Replicate: Modelos de IA de Clientes al Descubierto

Tipo de Ataque: **Brecha de Seguridad.**

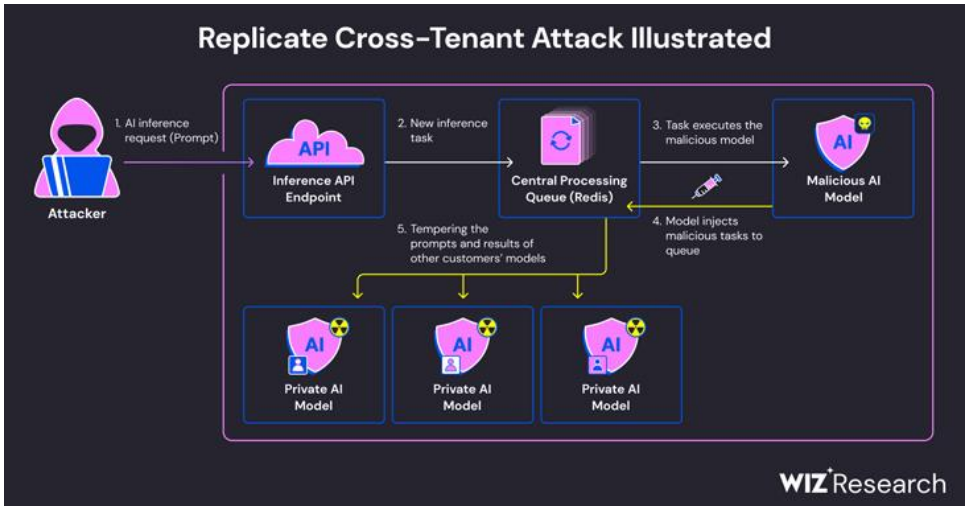
Medio de Propagación: **Internet**

1. **PRODUCTOS AFECTADOS:**
   - Servidores de Replicate
2. **RESUMEN:**

Investigadores de ciberseguridad han descubierto una grave vulnerabilidad en el proveedor de servicios de Inteligencia Artificial Replicate que podría haber permitido a atacantes acceder a modelos de IA privados y datos confidenciales al aprovechar una falla que permitía la ejecución de código malicioso a través de modelos IA.

3. **DETALLE:**

La vulnerabilidad residía en el hecho de que los modelos de IA se empaquetan típicamente en formatos que permiten la ejecución de código arbitrario. Los investigadores del cloud security firm Wiz crearon un contenedor malicioso utilizando la herramienta de código abierto Cog y lo subieron a la plataforma de Replicate, logrando así ejecutar código remoto con privilegios elevados en la infraestructura del servicio alojada en Google Cloud Platform.



Fuente: Wiz Research - The risk in malicious AI models: Wiz Research discovers critical vulnerability in AI-as-a-Service provider, Replicate.

Al aprovechar una instancia centralizada del servidor Redis utilizada para manejar múltiples solicitudes de clientes, los investigadores descubrieron que podrían manipular el proceso e insertar tareas maliciosas para impactar los resultados de los modelos de otros clientes de Replicate. Esto representa un riesgo significativo, ya que un atacante podría consultar modelos de IA privados, exponiendo conocimientos propietarios o datos sensibles utilizados en el entrenamiento de los modelos. Además, al interceptar las entradas (prompts), un atacante podría exponer información de identificación personal y otros datos confidenciales.

Si bien la vulnerabilidad fue reportada de manera responsable en enero de 2024 y ha sido solucionada, los investigadores advierten sobre el gran riesgo que representan los modelos IA maliciosos, especialmente para proveedores de IA como servicio, donde los atacantes podrían aprovecharlos para realizar ataques entre inquilinos y acceder potencialmente a millones de modelos y aplicaciones de IA privadas almacenadas en estas plataformas.

4. **RECOMENDACIONES: Validar**
   - No alimentar a los modelos de IA con data sensible o confidencial a menos que sea absolutamente necesario. Si es necesario alimentar modelos de IA con datos sensibles, realizar un análisis de riesgos para un adecuado tratamiento del riesgo.
5. **REFERENCIAS:**
   - https://thehackernews.com/2024/05/experts-find-flaw-in-replicate-ai.html
   - https://www.wiz.io/blog/wiz-research-discovers-critical-vulnerability-in-replicate

Cibercrimen y la Carrera por Dominar la IA: ¿Quién Lleva la Delantera?

1. RESUMEN:

De acuerdo con una actualización de Trend Micro presentada en la Conferencia RSA 2024, los ciberdelincuentes están rezagados en la adopción de la Inteligencia Artificial generativa (IA gen), utilizando principalmente productos de IA convencionales en lugar de desarrollar sus propios sistemas de IA. Sin embargo, se observa un crecimiento en servicios de "jailbreaking" que permiten a los criminales acceder de forma anónima a los Grandes Modelos de Lenguaje (LLM) existentes.

2. DETALLE

En su investigación de 2023 sobre el uso criminal de IA gen, Trend Micro ha encontrado solo un LLM desarrollado por delincuentes: WormGPT. En su lugar, hay una creciente incidencia, y por lo tanto un potencial uso, de servicios de "jailbreaking" como EscapeGPT, BlackHatGPT y LoopGPT, que permiten a los criminales acceder de forma anónima a LLMs populares como ChatGPT.

Además, hay un número creciente de servicios cuyo propósito es incierto, como FraudGPT, que solo mencionan supuestas capacidades sin proporcionar demostraciones o pruebas. Trend Micro los coloca en una categoría aparte denominada "posibles estafas".

En general, los criminales se están concentrando en el uso de productos de IA convencionales en lugar de desarrollar sus propios sistemas de IA. Esto se evidencia en el uso de IA dentro de otros servicios, como la herramienta de piratería Predator, que incluye una función de GPT utilizando ChatGPT para ayudar a los estafadores a crear texto.

También se observa un aumento en los servicios de deepfakes de imágenes, videos y voz, suficientemente buenos para engañar a personas sin un conocimiento íntimo del sujeto suplantado, y que se enfocan en eludir la verificación de conocimiento del cliente (KYC) para la creación de cuentas falsas.

A pesar de la actual falta de explotación criminal a gran escala de la IA gen, Trend Micro destaca indicios de que esto podría cambiar. Las principales prioridades de los delincuentes son aprender a utilizar la IA sin abandonar sus métodos existentes.

Esto explica la falta de urgencia, la escasez de LLMs desarrollados por criminales y el crecimiento de los servicios de "jailbreaking", que les permiten utilizar LLMs existentes de forma anónima.

Trend Micro prevé que estos servicios se volverán más sofisticados y adoptarán diferentes LLMs, ya que actualmente hay más de 6.700 LLMs alojados en HuggingFace. A medida que los criminales tengan acceso a LLMs más especializados, se espera que los utilicen para aplicaciones como deepfakes avanzados y evasión de verificación de usuarios.

Sin embargo, Trend Micro se mantiene cauteloso y no se une a los escenarios apocalípticos de IA. Hasta ahora, los criminales están rezagados en la adopción de IA gen debido a la renuencia a abandonar métodos exitosos existentes. En la carrera entre criminales y defensores, los defensores actualmente tienen la ventaja

.

3. **RECOMENDACIONES: Validar**
   - Verificar o fortalecer las defensas antimalware.
   - Verificar o fortalecer los mecanismos de recuperación como backups, especialmente si se tiene comunicación con la entidad afectada.
   - Fortalecer la concientización de los empleados, especialmente si se mantienen comunicación con la entidad afectada.
4. **REFERENCIAS:**
   - https://www.securityweek.com/criminal-use-of-ai-growing-but-lags-behind-defenders/

Impacto en Cadena: Filtraciones de Proveedores Dejan al Descubierto a Grandes Empresas Españolas

Tipo de Ataque: Brecha de datos

Medio de Propagación: Internet

1. RESUMEN:

En las últimas dos semanas, tres grandes empresas españolas del Ibex 35 - Banco Santander, Telefónica e Iberdrola - han sido víctimas de graves ciberataques que han resultado en el robo masivo de datos de millones de clientes y empleados. Estas filtraciones han generado alarma en el sector corporativo, ya que podrían formar parte de una campaña coordinada de ataques, poniendo en riesgo a otras compañías importantes.

2. DETALLE:

El 14 de mayo, el Banco Santander informó a la Comisión Nacional del Mercado de Valores (CNMV) que sufrió un ciberataque y el robo de datos personales de clientes en España, Chile y Uruguay, así como de "todos los empleados y algunos exempleados", aunque no se proporcionó una cifra oficial de afectados.

El 28 de mayo, Telefónica anunció que estaba investigando un ciberataque que resultó en la filtración de información de 120.000 clientes.

Un día después, el 29 de mayo, Iberdrola aseguró haber sido víctima de otro ciberataque, en el que se robaron los datos de 850.000 clientes.

Estas graves intrusiones han puesto en alerta máxima a los directivos de tecnología de las empresas del Ibex 35, quienes temen ser los próximos objetivos y se preguntan si estos ataques forman parte de una misma campaña coordinada.

Investigaciones preliminares apuntan a que el centro de estos ataques podría ser el hackeo al proveedor de servicios en la nube Snowflake, cuyas credenciales robadas permitieron a los cibercriminales acceder a las bases de datos de cientos de empresas, incluidas Ticketmaster, Telefónica y Santander.

Además, se reveló que, en el caso de Iberdrola, el ciberataque se produjo a través de su proveedor de call center, Konecta, poniendo de manifiesto el riesgo que representan las cadenas de suministro y los proveedores externos con medidas de seguridad deficientes.

Los expertos advierten que estos ataques suelen ser facilitados por empleados descontentos que comparten credenciales a cambio de dinero, y que las empresas deben reforzar sus protocolos de seguridad, especialmente en lo que respecta a proveedores y cadenas de suministro.

Mientras tanto, los datos robados ya están circulando en la red oscura, con millones de registros personales, números de cuenta, tarjetas de crédito y listas de empleados siendo ofrecidos a precios asequibles para los cibercriminales, lo que podría alimentar una ola de ciberestafas y otros delitos.

3. RECOMENDACIONES

- Identificar la ubicación lógica y física de los activos a proteger y los riegos de ciberseguridad asociados.
- Implementar controles según el nivel de exposición al riesgo y el apetito al riesgo de la organización.
- Monitorear las herramientas de seguridad con personal especialista en detección y respuesta a incidentes.
- Tener planes de respuesta y recuperación probados, para una respuesta rápida y efectiva, de manera de minimizar el impacto a la organización.

4. REFERENCIAS:

- https://www.elconfidencial.com/tecnologia/2024-05-31/iberdrola-telefonica-santander-ciberataques-konecta-snowflake_3892772/

Falla Crítica en Controlador Virtual de Honeywell Permite Ejecución Remota de Código

Tipo de Ataque: Vulnerabilidades

Medio de Propagación: Internet

1. **PRODUCTOS AFECTADOS:**
   - Control Edge Virtual UOC de Honeywell
2. **RESUMEN:**

Investigadores de la empresa de ciberseguridad Claroty han descubierto vulnerabilidades críticas en el controlador de automatización industrial Control Edge Virtual UOC de Honeywell que podrían permitir a los atacantes ejecutar código remoto y tomar el control total del dispositivo.

3. **DETALLE:**

La firma de ciberseguridad Claroty, especializada en OT, IoT y dispositivos médicos, ha divulgado detalles sobre vulnerabilidades críticas descubiertas por sus expertos en el Controlador de Operaciones de la Unidad de Control Edge (UOC) de Honeywell.

Una de las fallas encontradas, identificada como CVE-2023-5389 y catalogada de "gravedad crítica", está vinculada a una función no documentada en el protocolo propietario EpicMo que permite a los atacantes escribir archivos en los controladores virtuales UOC.

Los investigadores de Claroty centraron su análisis en EpicMo, el protocolo utilizado para la comunicación entre servidores y controladores Honeywell Experion.

Esta vulnerabilidad podría permitir a un ciberdelincuente con acceso a la red OT de la entidad atacada ejecutar código remoto sin necesidad de autenticación, simplemente enviando paquetes maliciosos al controlador.

"Un ataque de este tipo podría llevarse a cabo de forma remota para modificar archivos, obteniendo control total del controlador y permitiendo la ejecución de código malicioso", explicaron desde Claroty.

Además, los expertos identificaron otra falla de gravedad media, CVE-2023-5390, una vulnerabilidad de trayectoria de ruta absoluta que potencialmente permitiría a un atacante leer archivos desde el controlador, exponiendo información limitada del dispositivo.

Las vulnerabilidades fueron descubiertas en el controlador de automatización industrial ControlEdge Virtual UOC, que puede implementarse como una máquina virtual basada en Linux, eliminando la necesidad de tener un controlador físico.

Tras ser notificada por Claroty, Honeywell lanzó parches y publicó un aviso para advertir a sus clientes sobre estas fallas de seguridad.

4. **RECOMENDACIONES:**
   - Actualizar sus dispositivos Honeywell a la última versión estable recomendada.

5. REFERENCIAS:

- https://www.securityweek.com/critical-vulnerability-in-honeywell-virtual-controller-allows-remote-code-execution/
- https://nvd.nist.gov/vuln/detail/CVE-2023-5390

Sistemas de Agua en EE.UU. Vulnerables: Autoridades Exigen Plan de Ciberseguridad Urgente

Tipo de Ataque: Vulnerabilidades

Medio de Propagación: Internet

1. **PRODUCTOS AFECTADOS:**
   - Sistemas de Controles Industriales (ICS)
2. **RESUMEN:**

La Agencia de Protección Ambiental de Estados Unidos (EPA) emitió una alerta de cumplimiento para exigir medidas que protejan los sistemas de agua potable contra las ciberamenazas, después de que las inspecciones revelaran que más del 70% de los sistemas no cumplen plenamente con la Ley de Agua Potable Segura.

3. **DETALLE:**

Inspecciones realizadas por la Agencia de Protección Ambiental (EPA) desde septiembre de 2023 sacaron a la luz que una abrumadora mayoría de los sistemas de distribución de agua no cumplen cabalmente con la normativa federal de Agua Potable Inocua. Se detectaron vulnerabilidades cibernéticas críticas en algunos sistemas, incluida la utilización de contraseñas predefinidas y mecanismos de autenticación fácilmente violables.

En respuesta, el lunes la EPA emitió una advertencia de cumplimiento delineando los pasos que los operadores deben dar para blindar sus activos, como reducir la exposición a internet, realizar evaluaciones periódicas de riesgos, reemplazar claves predefinidas, inventariar activos de TI y OT, desarrollar y poner a prueba planes de respuesta ante incidentes, respaldar datos, mitigar fallas y brindar capacitación de concientización al personal.

La agencia anticipa un incremento de las inspecciones planificadas y advierte que, si corresponde, tomará acciones de cumplimiento civil y penal, incluso ante situaciones que representen un peligro inminente y sustancial.

Tras una serie de ataques cibernéticos disruptivos contra el sector hídrico estadounidense, las autoridades han actuado para robustecer los protocolos de los sistemas críticos y responder a los incidentes, publicando guías de ciberseguridad y sancionando a grupos ciberdelictivos vinculados a estados nacionales que estarían detrás de los ataques.

Algunos de los casos recientes incluyeron ataques de ransomware, hackers iranís que apuntaron a sistemas de control industrial (ICS) y hackers informáticos rusos que provocaron un desborde de agua en una pequeña localidad de Texas.

Expertos en ciberseguridad alertan que esta situación conducirá a más compromisos por parte de grupos delictivos radicados en China, Rusia e Irán. Recomiendan estrategias que incluyen escanear y mapear todos los dispositivos IoT, categorizarlos por niveles de riesgo, aislarlos en segmentos de red dedicados y restringir de forma extrema su acceso para administración y actualizaciones únicamente desde sitios e IPs aprobadas. También sugieren proteger esos equipos con dispositivos de seguridad endurecidos con

capacidad de conexión celular. Para empresas con recursos limitados, la recomendación es externalizar su programa de ciberseguridad y contratar servicios de seguridad gestionados.

4. RECOMENDACIONES:

- Se hace hincapié en la importancia de implementar medidas de seguridad adecuadas para proteger los sistemas de automatización industrial.

5. REFERENCIAS:

- https://www.securityweek.com/epa-issues-alert-after-finding-critical-vulnerabilities-in-drinking-water-systems/

Amenaza en Ascenso: Hacktivistas de Rusia Atacan Infraestructura Crítica

Tipo de Ataque: Vulnerabilidades

Medio de Propagación: Internet

1. **PRODUCTOS AFECTADOS:**
   - Interfaces hombre-máquina (HMI)
2. **RESUMEN:**

Agencias gubernamentales de EE.UU., Canadá y Reino Unido están alertando sobre una oleada de ataques contra sistemas de control industrial (ICS) y tecnología operativa (OT) por parte de presuntos hacktivistas prorrusos, instando a las organizaciones de infraestructura crítica a reforzar sus defensas cibernéticas.

3. **DETALLE:**

Un documento informativo elaborado por la agencia de ciberseguridad CISA y sus aliados revela que grupos hacktivistas han intentado comprometer sistemas ICS y OT en Norteamérica y Europa, particularmente en sectores como agua y aguas residuales, represas, energía, y alimentos y agricultura.

Los piratas informáticos han apuntado principalmente a interfaces hombre-máquina (HMI) expuestas a internet, aprovechando típicamente contraseñas predeterminadas y software VNC desactualizado.

Las agencias han estado rastreando este tipo de ataques desde 2022, pero la nueva alerta fue motivada por incidentes recientes en los que hacktivistas prorrusos se adjudicaron el crédito.

Según el aviso, estos atacantes manipularon HMI, haciendo que bombas de agua y equipos de aireación excedieran sus parámetros operativos normales, alterando configuraciones, desactivando mecanismos de alerta y cambiando contraseñas administrativas para bloquear el acceso a los operadores.

Si bien la mayoría de las víctimas lograron revertir a controles manuales rápidamente, algunas experimentaron derrames menores en tanques de almacenamiento.

Aunque no es raro que los hacktivistas exageren sus afirmaciones, las agencias advirtieron que si bien la actividad observada hasta ahora creó solo "efectos molestos", los piratas informáticos "son capaces de técnicas que representan amenazas físicas contra entornos OT inseguros y mal configurados".

Esta suposición se ve reforzada por un reciente reporte de Google Cloud Mandiant, que sugiere que al menos algunos de estos "hacktivistas" parecen ser personas vinculadas a una sofisticada unidad de piratería del gobierno ruso, específicamente Sandworm (APT44), conocida por ataques ICS altamente disruptivos.

El documento de CISA incluye recomendaciones para defensores de redes, fabricantes de dispositivos OT y organizaciones que han sido objetivo de estos ataques, instándolos a reforzar sus medidas de ciberseguridad.

4. **RECOMENDACIONES:**

- Se hace hincapié en la importancia de implementar medidas de seguridad adecuadas para proteger los sistemas de automatización industrial.

5. REFERENCIAS:

- https://www.securityweek.com/russian-hackers-target-industrial-systems-in-north-america-europe/