Los investigadores de ciberseguridad han alertado sobre una *«campaña de malware coordinada»* en JetBrains Marketplace, donde se han identificado al menos 15 complementos maliciosos capaces de robar claves de proveedores de inteligencia artificial (IA).

«Cada complemento se presenta como un asistente de programación basado en DeepSeek y otros modelos de lenguaje de gran escala. Ofrecen funciones como chat, generación de mensajes de commit, revisión de código, detección de errores y creación de pruebas unitarias», explicó el investigador de Aikido Security, Ilyas Makari. *«Funcionan exactamente como prometen. Sin embargo, la clave API del proveedor de IA que introduce el usuario es enviada secretamente a un servidor controlado por los atacantes.»*

Según los hallazgos, la actividad maliciosa estaría en marcha desde finales de octubre de 2025, y nuevos complementos continuaron publicándose hasta el 10 de junio de 2026. Dos de ellos, CodeGPT AI Assistant y DeepSeek AI Assist, superan las 25.000 descargas cada uno. No obstante, se desconoce si estas cifras son reales o si fueron infladas artificialmente para aparentar una mayor popularidad.

La lista completa de complementos identificados es la siguiente:

- DeepSeek Junit Test (org.sm.yms.toolkit)
- DeepSeek Git Commit (com.json.simple.kit)
- DeepSeek FindBugs (org.bug.find.tools)
- DeepSeek AI Chat (org.translate.ai.simple)
- DeepSeek Dev AI (com.yy.test.ai.simple)
- DeepSeek AI Coding (com.dev.ai.toolkit)
- AI FindBugs (com.json.view.simple)
- AI Git Commitor (com.my.git.ai.kit)
- AI Coder Review (org.check.ai.ds)
- DeepSeek Coder AI (com.review.tool.code)
- AI Coder Assistant (org.code.assist.dev.tool)
- DeepSeek Code Review (com.coder.ai.dpt)
- CodeGPT AI Assistant (com.my.code.tools)

- DeepSeek AI Assist (ord.cp.code.ai.kit)
- Coding Simple Tool (com.dp.git.ai.tool)

Aikido Security indicó que los 15 complementos comparten una base de código prácticamente idéntica. Todos solicitan al usuario acceder al panel de configuración e introducir una clave API de servicios de IA como OpenAI, SiliconFlow o DeepSeek para habilitar las funciones anunciadas.

Aunque las herramientas cumplen con las funciones que prometen, también incorporan una capacidad oculta para extraer silenciosamente las claves API proporcionadas y enviarlas a un servidor remoto (*«39.107.60[.]51»*) controlado por los atacantes mediante solicitudes HTTP sin cifrar.

«Los complementos también incluyen un nivel de servicio de pago», señaló la empresa. *«Cuando un usuario realiza un pequeño pago a través del sistema de donaciones integrado, el servidor devuelve una clave API al cliente. A partir de ese momento, el complemento utiliza esa clave para realizar llamadas al modelo en lugar de la del usuario. Esto resulta extraño, ya que ningún proveedor legítimo entregaría a los usuarios una clave funcional e ilimitada para un servicio de IA de pago.»*

Este comportamiento sugiere que los responsables de la operación podrían estar compartiendo o revendiendo las claves API robadas a otros actores maliciosos como parte de un esquema de monetización ilegal, permitiendo que quienes pagan accedan a servicios de IA utilizando las credenciales de las víctimas.

«El operador obtiene ingresos por un lado y credenciales gratuitas por el otro, mientras que los verdaderos propietarios de las claves terminan asumiendo los costes», añadió Makari.

La campaña constituye una nueva evidencia de cómo los actores de amenazas están enfocando cada vez más sus esfuerzos en los entornos de desarrollo y en el ecosistema de código abierto. Estos entornos representan objetivos especialmente atractivos porque suelen contener código fuente, credenciales de nube, claves de firma digital y claves API de

servicios de IA de pago que posteriormente pueden comercializarse o utilizarse en esquemas de *LLMjacking*.

«Un complemento debe tratarse con el mismo nivel de confianza y precaución que cualquier otra dependencia que se ejecute con sus privilegios. Además, es recomendable evitar introducir secretos de larga duración en herramientas que no hayan sido previamente verificadas», advirtió Aikido Security.

Extensiones maliciosas de Chrome roban conversaciones con IA

Este descubrimiento coincide con la detección de dos extensiones bloqueadoras de anuncios para Google Chrome que fueron sorprendidas recopilando conversaciones de los usuarios con chatbots de IA como OpenAI ChatGPT, Anthropic Claude, Google Gemini, Microsoft Copilot, Perplexity, DeepSeek, xAI Grok y Meta AI. La operación de recopilación de datos recibió el nombre de *PromptSnatcher* por parte del investigador Jean-Marie R.

Las extensiones identificadas, que aún permanecen disponibles en Chrome Web Store, son:

- Smart Adblocker (ID: iojpcjjdfhlcbgjnpngcmaojmlokmeii) — 90.000 usuarios (publicada en octubre de 2022)
- Adblock for Browser (ID: jcbjcocinigpbgfpnhlpagidbmlngnnn) — 10.000 usuarios (publicada en agosto de 2023)

«Aunque se presentan como simples bloqueadores de anuncios, estas extensiones incorporan un motor de interceptación desarrollado a medida que registra conversaciones privadas, información sobre el uso de modelos y metadatos relacionados con los niveles de suscripción de prácticamente todas las plataformas de IA importantes, incluyendo ChatGPT, Claude y Gemini», explicó el investigador. *«La operación utiliza listas de filtros públicas y legítimas, como EasyList e IDCAC, como cobertura funcional. De este modo, ofrecen una utilidad real de bloqueo de anuncios mientras ejecutan un canal de telemetría no revelado.»*

El hecho de que ambas extensiones lleven varios años disponibles indica que las capacidades de exfiltración de datos relacionados con IA fueron añadidas posteriormente mediante actualizaciones de software.

Este tipo de ataques pertenece a una categoría conocida como *Prompt Poaching*. Durante los últimos meses, se ha observado que diversas extensiones de navegador, tanto legítimas como maliciosas, han adoptado esta técnica para capturar de forma encubierta conversaciones de usuarios con sistemas de IA, bajo pretextos como mejorar la navegación segura o proporcionar métricas avanzadas de tráfico y participación. Sin embargo, sigue sin estar claro si estas prácticas incumplen las políticas de Google para extensiones de navegador.

«Las extensiones interceptan el historial completo de conversaciones con IA, información sobre el uso de modelos y el nivel de suscripción en ocho plataformas diferentes. Posteriormente, transmiten estos datos a infraestructuras controladas por los operadores sin informar adecuadamente al usuario, más allá de una genérica solicitud de consentimiento denominada 'Enhanced Protection'», concluyó el investigador.