Moltbook expuso correos y claves de una red de IA

04/02/2026

Una falla de seguridad dejó al descubierto miles de correos electrónicos y más de un millón de claves de acceso en una red social creada para agentes de inteligencia artificial.

Moltbook, una red social experimental diseñada exclusivamente para agentes de inteligencia artificial, quedó en el centro de la polémica tras una grave filtración de datos que expuso información sensible de la plataforma y de miles de personas vinculadas a su funcionamiento.

La vulnerabilidad fue detectada por investigadores de la empresa de seguridad Wiz, quienes identificaron una base de datos mal configurada que permitía acceso total a la información del sistema. Según el reporte, quedaron expuestos más de 35.000 correos electrónicos, 1,5 millones de claves de autenticación y mensajes privados intercambiados dentro de la plataforma.

Qué es Moltbook y cómo ocurrió la filtración

Moltbook se presenta como una red social donde los protagonistas no son humanos, sino agentes de inteligencia artificial que publican, comentan, votan y construyen reputación mediante un sistema similar al de foros tradicionales. En los últimos días había ganado notoriedad dentro de la comunidad tecnológica por su propuesta y por el nivel de interacción entre los bots.

El interés creció aún más cuando figuras del sector destacaron el proyecto como una experiencia cercana a la ciencia ficción. Su propio creador explicó públicamente que no escribió el código de la plataforma, sino que definió la arquitectura general y dejó que la inteligencia artificial se encargara del desarrollo.

Sin embargo, esa forma de creación acelerada derivó en una falla crítica: una clave de acceso a la base de datos quedó expuesta en el código visible para cualquier usuario, lo que habilitaba leer y modificar toda la información del sistema sin autenticación.

La respuesta tras el incidente

Desde Wiz indicaron que el hallazgo se realizó de manera no intrusiva y que la situación fue comunicada de inmediato a los responsables de Moltbook. La empresa aseguró que la vulnerabilidad fue corregida en pocas horas y que los datos utilizados durante la investigación fueron eliminados.

El episodio volvió a encender las alertas sobre los riesgos de desarrollar plataformas complejas sin controles de seguridad adecuados, especialmente en proyectos que combinan inteligencia artificial, datos sensibles y acceso masivo en línea.