

Primera notificación de una brecha de datos personales causada por un ataque ejecutado mediante un agente de IA

PRIMARY

Publisher	Agencia Española de Protección de Datos
Published	2026-09-14
Retrieved	2026-09-18
Cite key	agenciaespa2026primera
URL	https://www.aepd.es/prensa-y-comunicacion/blog/primera-notiviacion-brecha-datos-personales-causada-por-ataque-ejecutado-mediante-agente-ia

Claims drawn from this source. CLAIM-2026-09-18-01, CLAIM-2026-09-18-02, CLAIM-2026-09-18-03, CLAIM-2026-09-18-04, CLAIM-2026-09-18-05, CLAIM-2026-09-18-09

AEPD blog, 14 September 2026 — "Primera notificación de una brecha de datos personales causada por un ataque ejecutado mediante un agente de IA", signed Francisco Pérez Bes.

Key passages (original Spanish, working translation follows each):

"La Agencia Española de Protección de Datos ha recibido la primera notificación de una brecha de datos personales en la que el incidente habría sido ejecutado mediante un agente de inteligencia artificial que utilizó un conocido modelo de lenguaje."

The AEPD has received the first notification of a personal data breach in which the incident was reportedly executed by means of an AI agent that used a well-known language model.

"El agente atacante inició una búsqueda de vulnerabilidades en archivos genéricos, y realizó un login correcto. Una vez accedió al sistema, comenzó a buscar, de forma autónoma, vulnerabilidades en la aplicación, lo que, una vez conseguido, le permitió modificar datos personales y acceder a facturas."

The attacking agent began a search for vulnerabilities in generic files, and performed a successful login. Once it had accessed the system it began to search, autonomously, for vulnerabilities in the application, which once achieved allowed it to modify personal data and access invoices.

"Asimismo, la utilización de un concreto modelo de IA tampoco implica que el modelo o la infraestructura de su proveedor hayan sido comprometidos ni que la herramienta haya sido diseñada para desarrollar actividades maliciosas. Ahora bien, lo relevante desde la perspectiva de la protección de datos es que un tercero habría utilizado un"

agente de IA como instrumento para encadenar exitosamente distintas fases del ataque."

Likewise, the use of a particular AI model does not imply that the model or its provider's infrastructure was compromised, nor that the tool was designed to carry out malicious activities. What is relevant from the data protection perspective is that a third party used an AI agent as an instrument to successfully chain together the different phases of the attack.

"Esta primera notificación no permite afirmar una tendencia estadística, aunque sí constituye una señal significativa de que los ataques apoyados en inteligencia artificial han dejado de ser un riesgo teórico."

This first notification does not permit the assertion of a statistical trend, although it does constitute a significant signal that AI-supported attacks have ceased to be a theoretical risk.

The four observations ("¿Por qué es relevante esta primera notificación?")

1. Risk analyses must expressly incorporate attacks assisted or executed by AI: "no basta con incluir una referencia genérica a malware, phishing o acceso no autorizado".
2. Response times must be revised: "Los procedimientos diseñados para ataques ejecutados manualmente pueden resultar insuficientes cuando un agente analiza simultáneamente múltiples activos, prueba distintas vías de acceso y adapta su comportamiento con rapidez."
3. Digital identities and credentials increase in importance: "Un agente que obtiene una cuenta, una clave API o un token con permisos excesivos puede operar a la velocidad de una máquina y acceder a diferentes servicios antes de que la organización detecte un comportamiento anómalo."
4. Security of processing cannot depend on manual intervention alone.

The post cites CCN-CERT BP/36, *Buenas prácticas frente al modelo de IA ofensiva*, published by the Centro Criptológico Nacional.

The post names neither the affected organisation, nor the agent, nor the model.