SIGMA LANZA LA VERIFICACIÓN INDEPENDIENTE CON SUPERVISIÓN HUMANA PARA AGENTES CONVERSACIONALES DE IA

Sigma Eval es un servicio de verificación independiente para agentes de IA empresariales que combina evaluación continua con supervisión humana.



Supervision humana independiente de agente conversacional de IA en la que el cliente puede confiar

MADRID, SPAIN, September 22, 2026 /EINPresswire.com/ -- Sigma ha anunciado hoy el lanzamiento de Sigma Eval. El servicio de verificación aporta una capa de garantía independiente para evaluar sistemas conversacionales de IA en 12 dimensiones de seguridad, experiencia de usuario y rendimiento, tanto antes como después del despliegue.

El lanzamiento responde a las nuevas exigencias de la gobernanza de la IA para los agentes conversacionales empresariales. Sigma Eval ofrece a las organizaciones una valoración imparcial y basada en evidencias de cómo se comportan sus sistemas conversacionales de IA con usuarios reales, y no de cómo se espera que se comporten en una demostración controlada.

> Las empresas no deberían tener que adivinar si su agente de IA es bueno. La verificación independiente permite que un consejo de administración, un regulador o un cliente confíen en su comportamiento”
>
> *Daniel Tapias, CEO de Sigma Group*

La distancia entre el despliegue de la IA y su verificación independiente es ya visible, con consecuencias para empresas y sus clientes. Un estudio sectorial de 2026 reveló que el 57 % de las organizaciones tiene agentes de IA en producción, y solo el 52 % afirma realizar evaluaciones offline (LangChain). En otro estudio, el 50 % declaró haber desplegado un agente de IA o una funcionalidad basada en LLM que había superado las evaluaciones internas y que, sin embargo, provocó

después un fallo ante el cliente (Venturebeat).

Las empresas están desplegando agentes conversacionales en atención al cliente, ventas, gestión de siniestros y soporte interno más rápido de lo que pueden ser verificados. La mayoría sigue confiando en el criterio de los proveedores que construyeron el sistema, en pruebas anecdóticas o en pocas conversaciones de muestra revisadas a mano. La posición de Sigma es clara: la garantía debe venir de fuera del sistema que se garantiza.

«Las empresas no deberían tener que adivinar si su agente de IA es bueno. Deberían poder saberlo. La verificación independiente es lo que convierte una suposición interna en algo en lo que un consejo de administración, un regulador o un cliente pueden confiar», afirmó el Dr. Daniel Tapias, CEO de Sigma.



Evaluacion en seguridad, experiencia de usuario y rendimiento del agente conversacional

Una capa de garantía, no otra herramienta de IA
Sigma Eval se sitúa junto al stack de IA. La compañía no construye el agente, ni ajusta el modelo, ni proporciona los prompts: mide el comportamiento resultante. Esa separación es deliberada y permite que el resultado funcione como garantía: un veredicto externo emitido por una empresa que no tiene interés en el resultado.

El servicio está diseñado para usarse a lo largo del ciclo de vida de un sistema de IA. Antes del despliegue, actúa como control previo al lanzamiento: las organizaciones pueden ver cómo gestiona el agente a usuarios adversarios, datos sensibles, presión fuera de tema y peticiones sin resolver antes de que un cliente se encuentre con él. Después del despliegue se convierte en instrumento de monitorización, porque los agentes derivan. Los prompts se editan, los modelos se actualizan, las bases de conocimiento cambian y las integraciones se mueven. Un agente que era seguro en marzo no lo es necesariamente en septiembre. La verificación periódica hace visible esa deriva.

Doce dimensiones en tres categorías
Cada evaluación puntúa al agente en 12 dimensiones agrupadas en tres categorías:
 Seguridad: la exposición reputacional, legal y financiera derivada del mal comportamiento del sistema: sesgo, toxicidad, alucinaciones, opacidad, exposición de PII y vulnerabilidad a la manipulación.
 Experiencia de usuario: cómo mantiene el agente la conversación y al usuario enfocado en

ella: desalineación del contexto, inconsistencia conversacional y desinterés del usuario.
 Rendimiento: con qué eficacia el agente lleva al usuario hacia su objetivo: rendimiento de trayectoria, esfuerzo del cliente y coste de resolución.
La evidencia se genera con un sistema propietario que combina usuarios sintéticos con conversaciones sintéticas, ejercitando el agente a escala y en cualquier idioma. En lugar de una única puntuación agregada, los clientes reciben un informe que reporta cada dimensión por separado, de modo que un buen resultado comercial no pueda ocultar una debilidad de seguridad, ni al contrario.

Sencillez de uso
El servicio no requiere trabajo de integración ni acceso al código o al modelo. La organización envía el enlace público de su agente conversacional, junto con unos datos de contacto básicos y cualquier contexto relevante: el propósito del agente, sus idiomas, sus KPIs y sus casos límite conocidos. Sigma Cognition ejecuta la evaluación y entrega un informe confidencial.

La confidencialidad es una condición del servicio, no una opción. Los informes se comparten únicamente con la organización solicitante y nunca se utilizan como material público de benchmarking ni como contenido promocional.
Sigma abre el servicio con un informe de evaluación gratuito, disponible para cualquier organización que opere un agente conversacional. La compañía también pone a disposición un número limitado de evaluaciones repetidas para que los equipos puedan verificar el efecto del trabajo de corrección, y no solo recibir un diagnóstico.

Sigma Eval responde a las nuevas exigencias de la gobernanza de la IA
El lanzamiento responde a una necesidad del mercado ante el cambio en la forma de gobernar los sistemas de IA: las organizaciones prestan cada vez más atención a cómo y quién evalúa los sistemas de alto impacto, y con qué evidencias. La evaluación independiente por terceros es el mecanismo que la mayoría de los sectores regulados ya reconoce en la información financiera, en la seguridad de producto y en la seguridad informática. La IA está llegando al mismo punto.
El trabajo de verificación de Sigma se apoya en investigación aplicada en seguridad de la IA, evaluación y anonimización, desarrollada con socios académicos e institucionales. Esa base de investigación sustenta Sigma Eval y su producto complementario Sigma Cypher, que anonimiza la información personal identificable antes del análisis.

Sobre Sigma Group
Sigma Group desarrolla soluciones avanzadas de IA que ayudan a las organizaciones a maximizar el valor de sus datos a lo largo de su ciclo de vida. Con más de 18 años de experiencia, la compañía está especializada en calidad del dato, anonimización multimodal, IA generativa y plataformas de datos de confianza. A través de su principal oferta: Sigma Eval, Sigma Cypher y Sigma Verified, Sigma Group ofrece servicios independientes de verificación y seguridad basados en evidencias para sistemas de IA críticos para el negocio, combinando investigación con experiencia humana

M.ª Luisa Izquierdo Núñez
Sigma Cognition
Visit us on social media:
LinkedIn
YouTube

This press release can be viewed online at: https://www.einpresswire.com/article/943947494

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.
© 1995-2026 Newsmatics Inc. All Right Reserved.