SIGMA LAUNCHES HUMAN-SUPERVISED INDEPENDENT VERIFICATION FOR AI CONVERSATIONAL AGENTS

Sigma launches an independent verification service for enterprise AI agents, combining continuous evaluation with human supervision to mitigate business risks.

Human-supervised AI conversational agents verification customers can trust

LONDON, UNITED KINGDOM, September 22, 2026 / EINPresswire.com/ -- Sigma today announced the launch of Sigma Eval. This verification service provides an independent assurance layer to evaluate conversational AI systems across 12 dimensions of safety, user experience and performance, both before and after deployment.

The launch addresses the changing demands of AI governance for enterprise conversational AI agents. Sigma Eval gives organisations an impartial, evidence-based assessment of how their conversational AI systems actually behave with real users, not how they are expected to behave in a controlled demo.

> “Independent verification turns AI performance from an internal assumption into evidence that boards, regulators, and customers can trust. Companies should know whether their AI agents perform reliably”
>
> *Dr. Daniel Tapias*

The gap between AI deployment and independent verification is becoming increasingly visible with direct consequences for businesses and their customers. A 2026 industry survey* of more than 1,300 professionals found that 57% of organizations have AI agents in production, while only 52% report running offline evaluations (LangChain). In a separate survey of 157 enterprises, 50% reported having deployed an AI agent or LLM feature that had passed internal evaluations but subsequently caused a customer-facing failure (Venturebeat).

Enterprises are deploying conversational agents into customer service, sales, claims handling and internal support faster than they can verify them. Most teams still rely on the judgement of the same suppliers that built the system, on anecdotal testing, or on a handful of sample conversations reviewed by hand. Sigma´s position is straightforward: assurance must come from outside the system being assured.

Independent evidence-based verification services for business AI systems

“Companies should not have to guess whether their AI agent is good. They should be able to know. Independent verification is what turns an internal assumption into something a board, a regulator or a customer can trust.” said Dr. Daniel Tapias, CEO of Sigma.
(en apdo indep) “Independent verification turns AI performance from an internal assumption into evidence that boards, regulators, and customers can trust. Companies should know whether their AI agents perform reliably

An assurance layer, not another AI tool
Sigma Eval sits alongside the AI stack rather than inside it. The company does not build the agent, tune the model or supply the prompts; it measures the resulting behaviour. That separation is deliberate and allows the output to function as assurance: an external verdict produced by a party with no stake in the outcome.
The service is designed to be used continuously across the lifecycle of an AI system. Before deployment, it acts as a pre-launch gate: organisations can see how an agent handles adversarial users, sensitive data, off-topic pressure and unresolved requests before real customers ever encounter it. After deployment, it becomes a monitoring instrument, because agents drift. Prompts get edited, models are upgraded, knowledge bases change and integrations move. An agent that was safe in March is not automatically safe in September. Repeat verification makes that drift visible.

Twelve dimensions across three categories
Each evaluation scores an agent across 12 dimensions grouped into three categories:
• Safety: the reputational, legal and financial exposure created by system misbehaviour: bias, toxicity, hallucinations, opacity, PII exposure and vulnerability to manipulation.
• User experience: how well the agent holds a conversation and keeps the user on track: context misalignment, conversational inconsistency and user disengagement.
• Performance: how effectively the agent moves users towards their goal: trajectory performance, customer effort and resolution cost.
Evidence is generated by a proprietary system that combines synthetic users with synthetic

conversations, exercising the agent at scale and in any language. Rather than a single aggregate score, clients receive a scorecard that reports each dimension separately, so that a strong commercial result cannot mask a safety weakness, or the reverse.
Simple to start by design

The service requires no integration work, no code access and no model access. An organisation submits the public link to its conversational agent, together with basic contact details and any context worth knowing: the agent's purpose, its languages, its KPIs, its known edge cases. Sigma Cognition runs the evaluation and delivers a confidential report by email.
Confidentiality is a condition of the service, not an option. Reports are shared only with the requesting organisation, and are never used as public benchmarking material or promotional content.
Sigma is opening the service with a free evaluation report, available to any organisation operating a conversational agent. The company is also making a limited number of repeat evaluations available so that teams can verify the effect of remediation work, not merely receive a diagnosis.

Sigma Eval addresses the changing demands of AI governance.
The launch responds to a market need related to the shift in how AI systems are governed so, organisations are increasingly paying attention to how high-impact systems are assessed, by whom, and on what evidence. Independent third-party evaluation is the mechanism that most other regulated industries already recognise in financial reporting, in product safety, in security. AI is arriving at the same point.
Sigma's verification work builds on applied research in AI safety, evaluation and anonymisation carried out with academic and institutional partners. That research base underpins Sigma Eval and its companion product Sigma Cypher, which removes personally identifiable information before analysis.

About Sigma Group
Sigma Group develops advanced AI solutions that help organizations maximize data value throughout its lifecycle. Backed by over 18 years of experience and continuous R&D, the company specializes in data quality, multimodal anonymization, generative AI, and trusted data platforms. Through its core offerings, Sigma Eval, Sigma Cypher, and Sigma Verified, Sigma Group delivers independent, evidence-based verification and safety services for business-critical AI systems, combining next-generation research with human-in-the-loop expertise.

Driving Human Expertise to Accelerate Next-Gen AI

Mya Oozeerally
Sigma AI -London Region Office
email us here
Visit us on social media:
LinkedIn

This press release can be viewed online at: https://www.einpresswire.com/article/943953367

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.