GIRONA NOTICIES.COM
+ Periodisme

NextAI responde al miedo a la IA con un DNI para «robots digitales»

Actualidad Empresarial | 18/09/2026 | 00:00

La compañía desarrolla una identidad verificable para agentes de inteligencia artificial y orienta su evolución hacia los principios europeos de transparencia, trazabilidad y supervisión humana. Argos es su primer caso público.

La inteligencia artificial está pasando de responder preguntas a ejecutar tareas. Para las empresas, ese salto abre una cuestión que ya no es teórica: quién responde cuando un agente actúa en su nombre. NextAI, compañía de consultoría e ingeniería de inteligencia artificial empresarial, ha desarrollado un DNI digital para agentes de IA que permite consultar su identidad declarada, la organización responsable y el alcance de su autorización.

La iniciativa llega en pleno debate sobre el control de los sistemas autónomos. Los informes publicados por OpenAI y la investigación de METR y Redwood Research sobre el incidente de Hugging Face documentan comunicaciones no autorizadas y actuaciones fuera del encargo en entornos de evaluación. El 16 de septiembre, OpenAI presentó además un marco para comunicar comportamientos desalineados. Son hechos que exigen respuestas técnicas, sin necesidad de atribuir voluntad humana a los modelos.

«Mientras medio internet se pregunta si la IA se nos está yendo de las manos, nosotros estamos construyendo herramientas para poder dirigirla mejor. Si vamos a dar a un agente las llaves de una empresa, me parece bastante razonable empezar por pedirle el DNI: de quién es, para qué está ahí y qué tiene autorizado», señala Cosmin Popescu, CEO de NextAI.

Un DNI para la IA que trabaja dentro de la empresa

Un agente puede preparar un informe, investigar una cuenta comercial o colaborar en un proceso de atención al cliente. Cada capacidad plantea decisiones distintas sobre información, acceso y responsabilidad. Autorizar una tarea concreta no equivale a conceder permiso para hacer cualquier cosa que ayude a terminarla.

El DNI digital para agentes de IA desarrollado por NextAI convierte parte de ese mandato en una credencial firmada y verificable. Recoge la identidad del agente, la sociedad que responde de él, sus acciones autorizadas y prohibidas, la vigencia y una referencia a la versión de su manifiesto.

Su representación visual, el Sello CORTEX, facilita el acceso a una comprobación pública. La expresión «DNI» hace comprensible el concepto; no designa un documento estatal ni una certificación oficial.

Los «robots digitales» de esta iniciativa son agentes de software. La posible extensión a robots físicos requeriría integrar controles específicos de seguridad y operación.

Argos: una propuesta que ya se puede comprobar

El primer ejemplo público es Argos, un agente de la flota de NextAI. Su credencial declara permisos para investigar y preparar borradores, junto a prohibiciones como enviar comunicaciones, ejecutar pagos o firmar contratos.

Cualquier persona puede consultar esa declaración en el verificador público de DIGITAL CORTEX. La evidencia es concreta: una identidad, un emisor, una autorización y un estado que se pueden revisar.

La credencial acredita lo declarado por el emisor. Para convertir sus restricciones en bloqueos efectivos, los conectores y herramientas deben comprobarlas y aplicarlas. Esa conexión entre identidad y ejecución es una de las principales líneas de evolución de la propuesta.

Del manifiesto a las reglas que acompañan cada acción

El manifiesto describe la función, la misión, el principio operativo y el marco de autonomía del agente. Su publicación por versiones permite vincular la credencial a una declaración determinada y mantener una referencia para revisar cambios.

Las acciones permitidas, las prohibiciones y las condiciones de autorización se establecen al emitir la credencial. Las instrucciones que recibe el modelo y los controles de las herramientas deben mantenerse coherentes con ese mandato.

Sobre esta base, NextAI plantea ampliar el sistema hacia comprobaciones previas a cada acción, autenticación de quien presenta la credencial, delegaciones entre agentes que no amplíen permisos y mecanismos de retirada efectiva de accesos. Son líneas de desarrollo que requieren integración y pruebas, no funcionalidades que se den por implantadas en todos los entornos.

«El DNI es el principio. Queremos que el manifiesto y las reglas del agente acompañen al trabajo: cuando usa una herramienta, cuando delega y cuando cambia su configuración. Cada avance en autonomía tiene que venir acompañado de un avance en control», añade Popescu.

Una dirección alineada con los principios europeos de IA

NextAI orienta el desarrollo del DNI digital hacia los objetivos de transparencia, trazabilidad y supervisión humana recogidos en el Reglamento europeo de Inteligencia Artificial —AI Act—. Este marco establece obligaciones según el uso y el riesgo: contempla requisitos de documentación, registro de actividad y supervisión humana para los sistemas de alto riesgo, y deberes de transparencia para determinados sistemas de IA.

La identidad verificable puede aportar una pieza documental a ese trabajo: vincular un agente con una organización, una misión y una autorización. El cumplimiento exige evaluar cada despliegue y aplicar las medidas que correspondan. El DNI y el Sello CORTEX no constituyen una certificación europea ni sustituyen esa evaluación.

Construir confianza antes de ampliar la autonomía

La apuesta de NextAI es que las empresas puedan incorporar más capacidades de IA con una base de responsabilidad visible. Identificar al agente, delimitar su autorización y registrar lo que ocurre son piezas que pueden contribuir a prevenir acciones no autorizadas y facilitar la respuesta cuando algo falla.

Su eficacia depende de la integración con medidas de seguridad, supervisión y evaluación del comportamiento. El DNI no garantiza por sí solo que un agente no falle ni demuestra que hubiera impedido incidentes ajenos.

La compañía propone avanzar mediante pruebas concretas: comprobar qué acciones se deniegan, cómo se detectan cambios y cuánto tarda en retirarse el acceso. Esa evidencia permitiría ampliar las funcionalidades con criterios verificables.

Sobre NextAI

NextAI es una marca de consultoría e ingeniería de inteligencia artificial empresarial operada por Cosmolabs OÜ. Desarrolla agentes de IA, aplicaciones y ecosistemas que conectan datos, procesos y herramientas. Su enfoque de «cerebros empresariales» busca que la información de una organización pueda utilizarse para apoyar decisiones y ejecutar procesos. Digital Cortex proporciona la infraestructura de identidad verificable de esta iniciativa.

Portavoz: Cosmin Popescu, CEO de NextAI.

Autor: Redacció

https://gironanoticies.com/empresarial/304336-nextai-responde-al-miedo-a-la-ia-con-un-dni-para-robots-digitales.htm