

PAQUETE SECTORIAL · IA Y AGENTES AUTÓNOMOS

Agentes de IA que actúan, acciones que se demuestran

Control, aprobación humana y evidencia verificable para agentes de inteligencia artificial, automatizaciones y flujos autónomos.

Sus agentes ya no solo recomiendan: **utilizan herramientas, modifican información, ejecutan código y activan flujos completos**, con frecuencia sin intervención humana continua. Una credencial válida ya no basta como control. *¿Quién limita lo que cada agente puede hacer, quién aprueba lo sensible y qué evidencia queda de cada acción?*

QUÉ APORTA VELANTIA A ESTE SEGMENTO

AUTONOMY — gobierna cada agente: **capacidades, alcances, condiciones, aprobación humana, revocación y trazabilidad**, también para robots y procesos automáticos.

GATE — decide cada acción crítica que un agente solicita: **permitir, condicionar, reducir, derivar a aprobación o bloquear**, según política y contexto.

SENTINEL — observa el comportamiento, **detecta anomalías** y atribuye cada actuación a su agente, identidad y herramienta.

EVIDENCE — conserva una **historia verificable** de cada decisión y resultado, lista para auditoría, cumplimiento e investigación.

QUÉ PUEDE HACER CON ELLO

- ▶ **Definir límites por agente y herramienta:** reglas, permisos y políticas de ejecución por agente, herramienta, dato y acción.
- ▶ **Exigir aprobación humana** antes de acciones sensibles, irreversibles o privilegiadas — sin frenar el resto de la automatización.
- ▶ **Supervisar en tiempo real** la actividad de agentes, flujos automatizados e integraciones, con su contexto.
- ▶ **Introducir agentes con menos riesgo:** el daño de un agente comprometido queda acotado a lo que su permiso autoriza, con registro de todo lo que hace.

El agente puede actuar — **dentro de límites explícitos y con una historia atribuible de lo que hace.**

DESPLIEGUE

Software, máquina virtual, nube privada o Trust Appliance, con gobierno unificado en **VELANTIA ONE**. Para control dedicado y local de agentes, el appliance **AUTONOMY GATEWAY 500** (en pre-reserva). Se integra con sus plataformas de IA, identidad y seguridad **sin reemplazar el stack existente**.

POR QUÉ AHORA

Cada mes, **más acciones críticas se ejecutan sin intervención humana continua**, y las identidades no humanas ya superan a las humanas. Reguladores y clientes empiezan a **exigir demostrar** qué control se aplicó a cada acción autónoma. Quien no gobierna al agente hoy, solo podrá reconstruir lo que hizo mañana.

Próximo paso: una reunión inicial y una demostración.

Flujo completo en entorno controlado: acción del agente → política → aprobación o bloqueo → evidencia. Disponible propuesta de piloto de seguridad de agentes de IA, acotada y con métricas.