kovra

DOCUMENTO FUNDACIONAL · JUNIO 2026

Visión, misión, principios y valores

El documento fundacional para la era de los agentes de IA

Un producto de Kaeus Inc

kovra.sh

01 La paradoja de la aceleración

Estamos viviendo el salto de productividad más grande desde la llegada de internet. La inteligencia artificial ya no solo sugiere: **actúa**. Agentes autónomos escriben código, mueven datos, ejecutan tareas y operan sistemas en nombre de las personas. Lo que antes tomaba semanas, hoy toma horas.

Pero esa velocidad tiene un precio que casi nadie ha puesto sobre la mesa con honestidad: **para que un agente trabaje por ti, normalmente hay que darle el mismo acceso que tienes tú.** Las llaves de tu base de datos. Los accesos a tus servicios. Las credenciales de producción. Y a diferencia de una persona, un agente **actúa sobre todo lo que lee** —incluso cuando lo que lee fue diseñado para engañarlo.

Esta es la paradoja: cuanto más capaz es el agente, más acceso necesita; y cuanto más acceso tiene, mayor es la superficie por la que una organización puede ser comprometida. Hoy, la respuesta tácita del mercado es resignarse a una disyuntiva: o vas rápido, o estás protegido.

Nosotros rechazamos esa disyuntiva. Creemos que se puede capitalizar toda la aceleración que la IA habilita **sin ampliar la superficie de exposición**. No es elegir uno de los dos mundos. Es tener los dos.

02 El mundo está cambiando más rápido de lo que el riesgo puede seguir

El primer lugar donde esto se siente es la ingeniería —agentes que leen repositorios llenos de secretos—. Pero sería un error pensar que el problema vive solo ahí.

Los agentes están entrando a **cada función de la empresa**: finanzas, ventas, operaciones, soporte, mercadeo. Un agente que responde a clientes necesita las llaves del sistema de clientes. Uno que automatiza informes toca la fuente de datos sensible. Y uno que prepara pagos o concilia movimientos necesita **las credenciales de la banca en línea —las llaves de las cuentas donde se custodian los fondos de la empresa. Estos agentes, los operados por áreas de negocio, suelen estar igual o más expuestos que los de ingeniería** —porque quienes los operan no son especialistas en seguridad, y porque nadie diseñó esos flujos pensando en un actor que actúa sobre lo que lee.

La superficie de exposición de una organización ya no crece con el número de empleados. Crece con el número de **agentes**, y cada agente es un nuevo lugar por donde una credencial puede filtrarse, ser malinterpretada o ser dirigida en contra de su dueño. El riesgo se está

multiplicando a la velocidad de la adopción de IA —y la mayoría de las organizaciones no lo verá venir hasta que ocurra.

03 El riesgo, en términos de negocio

No hace falta entender criptografía para entender lo que está en juego. Basta una pregunta para el dueño del riesgo:

> ¿Cuántos de los secretos que sostienen tu operación podría leer hoy un agente de IA —y qué pasaría si uno de esos agentes fuera dirigido, mediante un dato manipulado, a usarlos en tu contra?

Las consecuencias no son hipotéticas ni técnicas: son una credencial de producción expuesta, una clave que termina fuera de la organización, un sistema crítico tocado por un proceso que creyó estar haciendo su trabajo. El costo no se mide en líneas de código —se mide en interrupción, en pérdida de confianza, en exposición regulatoria, en reputación.

Y llévalo al caso más visceral, el que entiende cualquier directorio: las credenciales de la **banca en línea** —las llaves de las cuentas donde se custodian los fondos de la empresa— son, para un agente, un secreto más entre los que puede leer y usar. Un agente dirigido no necesita vulnerar ningún sistema: usa una credencial legítima creyendo que hace su trabajo, y el dinero ya salió. ¿Le entregarías esas llaves a un asistente que acaba de leer un correo de un desconocido? Hoy, sin saberlo, muchas organizaciones ya lo hacen.

Y lo más incómodo: hoy ese riesgo es **invisible** para quien debe responder por él. La conversación sobre productividad de IA ocurre en una sala; la de exposición de credenciales, en otra. kovra existe para juntar esas dos salas —y para que la respuesta no sea frenar, sino gobernar.

04 Por qué el modelo actual no alcanza

Durante décadas, la forma de proteger un secreto fue **custodiarlo bien**: guardarlo cifrado, controlar quién puede pedirlo, entregárselo a quien lo pide. Ese modelo asume algo que la era de los agentes rompe: que **quien pide el secreto es de confianza**.

Y ni siquiera hace falta un atacante para que el modelo falle. En el día a día, cuando a un agente se le encomienda una tarea que requiere acceso a un sistema, hace lo más natural: se ofrece a resolverla y, con toda buena intención, **pide la credencial** para entrar. No hay malicia —hay ingenuidad—. El agente no comprende las implicaciones de lo que pide, ni asume la responsabilidad de cuidar lo que recibe. Pedir una llave y custodiarla con el peso que merece

son dos cosas distintas; un agente, por naturaleza, solo entiende la primera. **El agente mejor intencionado sigue siendo un mal custodio.**

Y esa es solo la cara amable. La misma puerta —un agente que actúa sobre lo que lee— puede ser abierta en tu contra: un documento envenenado, un mensaje de error manipulado o una instrucción escondida bastan para conducirlo a una mala acción mientras cree estar haciendo su trabajo. En ambos casos —el ingenuo y el dirigido— el modelo clásico, custodiar y entregar, le da el secreto en claro y confía en que lo use bien. En el instante en que el agente lo recibe, el control se perdió.

El problema, entonces, no es guardar mejor el secreto. Es **repensar quién lo recibe y para qué** cuando quien lo pide —con buena fe o conducido— no puede asumir la responsabilidad de cuidarlo.

05 La respuesta de kovra

> **Visión.** La aceleración que la IA habilita no debería obligar a elegir entre ir rápido y estar protegido. Imaginamos organizaciones que capitalizan toda la autonomía de sus agentes —en ingeniería y en cada área de negocio— sin ensanchar su superficie de exposición: donde delegar trabajo privilegiado a un agente nunca significa cederle el control de las credenciales de las que ese trabajo depende. No es una disyuntiva; es lo mejor de ambos mundos —velocidad y custodia, a la vez.
>
> **Misión.** Custodiar el uso del secreto en la era de los agentes. kovra permite que herramientas y agentes **usen** secretos sin **verlos**, mantiene al humano como principal en cada punto peligroso, y hace que el camino seguro sea el conveniente. Empezamos por el desarrollador, donde el riesgo es más agudo, y extendemos la misma garantía a cada agente que una organización pone a trabajar.

La idea central es un cambio de marco: pasar de custodiar el secreto a **custodiar su uso**. El agente puede trabajar —puede pedir usar un secreto—, pero nunca se vuelve el dueño de los secretos más críticos. Cuando una acción importa de verdad, no la decide el agente: la **autoriza un humano con un solo gesto** (su huella), viendo exactamente qué se va a hacer. kovra actúa; tú autorizas.

Y todo esto sin que kovra —ni ninguna nube— vea jamás tus secretos: están **cifrados en origen**, y el control permanece siempre con su dueño.

06 Por qué es defendible

Una buena visión no basta; tiene que haber una razón por la que sea difícil de copiar. La ventaja de kovra se apoya en dos pilares que se refuerzan entre sí:

1. Custodiar el uso, no solo el almacenamiento. Cualquiera puede guardar un secreto cifrado. Lo difícil —y lo valioso— es dejar que un agente orqueste trabajo con un secreto que nunca llega a controlar: en el punto más peligroso, el secreto se usa a través de un ejecutor de confianza, con la acción exacta a la vista y la aprobación de un humano. Esa es la diferencia entre entregar una llave y operar la cerradura por quien la necesita.

2. Custodia ciega, viva donde viva el secreto. El secreto se cifra en origen y el dueño nunca pierde el control —viva ese secreto en una sola máquina, se comparta entre un equipo o se sincronice a través de la nube de kovra. La infraestructura que lo transporta nunca puede leerlo. Es la diferencia entre confiar en un proveedor y **no necesitar confiar en él**.

Juntos, estos dos pilares no son una función que se agrega; son una postura que hay que adoptar desde el primer día. Por eso elegimos **profundidad antes que alcance**: preferimos ser insustituibles en lo que de verdad importa, antes que ubicuos en lo que cualquiera puede ofrecer.

07 La honestidad como activo

La mayoría de las herramientas de seguridad prometen más de lo que pueden cumplir. Nosotros hacemos lo contrario, a propósito.

Hay una verdad que ninguna herramienta puede borrar: para usar un secreto, en algún momento tiene que existir en claro. Nadie elimina ese "último tramo" —ni nosotros—. Y lo decimos abiertamente. Lo que kovra hace es **encoger cuánto ve un secreto, mantener los secretos críticos lejos del eslabón manipulable, y poner a un humano en la decisión que importa.**

Esta honestidad no es una debilidad: es el activo. Un responsable de riesgo no puede firmar sobre una promesa absoluta —sabe que no existe—. Sí puede firmar sobre una garantía clara, acotada y verificable. **Quien dice la verdad sobre lo que no hace es, justamente, a quien se le puede creer lo que sí hace.** Esa credibilidad es lo que más cuidamos.

08 Nuestros principios

Estos principios son el filtro con el que tomamos cada decisión. Describen el mundo que queremos construir.

1. **El humano es el principal.** El agente solicita uso; nunca se vuelve dueño de un secreto crítico. La autoridad fluye de un gesto humano deliberado, no de quien lo pide.
2. **La honestidad es la ventaja que no se copia.** Nunca prometemos más de lo que garantizamos. La credibilidad es el activo —la protegemos por encima de la propaganda.
3. **Aceleración y protección no se canjean.** Ni frenamos al agente para estar seguros, ni aflojamos la seguridad para ir rápido. La seguridad inconveniente es seguridad que no se usa.
4. **La soberanía del secreto es de su dueño.** El control permanece siempre en sus manos; el secreto se cifra en origen y kovra nunca ve su contenido en claro —sin importar dónde se despliegue.
5. **La protección se gradúa según el peligro.** No todos los secretos son iguales: la protección crece con la sensibilidad, y el uso más peligroso es el más gobernado.
6. **Profundidad antes que alcance.** Ganamos siendo insustituibles en el punto crítico, no compitiendo por ubicuidad. Cuando hay que elegir, elegimos la profundidad.
7. **Diseñado para los agentes, sin confiar en ellos —y verificable.** Asumimos que el agente puede estar comprometido; y todo es inspeccionable. La confianza se gana siendo verificable, no pidiéndola.

09 Nuestros valores

Si los principios describen el mundo que queremos, los valores describen **cómo nos comportamos** al construirlo.

- **Custodia.** Cuidamos lo que se nos confía como si fuera lo más valioso que tenemos. La responsabilidad por el secreto del otro pesa más que nuestra propia conveniencia.
- **Honestidad radical.** Decimos lo que no hacemos. Preferimos perder una venta antes que una verdad. La confianza, una vez rota, no se recompone.
- **Profundidad sobre ruido.** Construimos ventaja real, no humo. El oficio y el rigor técnico antes que la moda del momento.
- **Verificable, no creíble.** No pedimos que nos crean; damos los medios para comprobarlo. Lo abierto y auditable es una postura, no una concesión.

- **El humano al centro.** La tecnología sirve al dueño; nunca lo reemplaza en la decisión que importa. La comodidad nunca se construye a costa de su control.

10 Trayectoria — tres horizontes

La visión es grande; el camino es deliberado. Avanzamos por horizontes, ganándonos el derecho al siguiente con lo entregado en el anterior. (Trayectoria narrativa, sin fechas comprometidas: el orden es firme, el calendario se gana.)

Horizonte 1 — Hoy: el desarrollador (en producción)

kovra ya custodia el uso de secretos para desarrolladores y sus agentes de codificación: usar sin ver, el humano como principal autorizando con su huella, local por diseño y con su código fuente abierto a inspección. Es la cabeza de playa —el lugar donde el dolor es más agudo y la solución, más demostrable—. **Esto no es una promesa: ya existe y funciona.**

Horizonte 2 — Cercano: la organización conectada

Extender la misma garantía más allá de una máquina y más allá del agente de código:

- **Uso compartido y sincronización con custodia ciega** —la nube de kovra, "the nest"— donde equipos comparten el uso de secretos sin que la infraestructura pueda leerlos.
- **Despliegue en infraestructura propia** para organizaciones que requieren que todo viva dentro de sus propias instalaciones —con la misma garantía de cifrado en origen.
- **La garantía se extiende a los agentes de negocio**, no solo a los de ingeniería: el área que opera el agente capitaliza la IA sin ampliar su exposición.

En todos los casos, una constante: **el dueño del secreto es la organización, y el nest jamás puede ver un secreto.**

Horizonte 3 — La visión: custodia del uso a escala organizacional

El destino: una organización que capitaliza toda la autonomía de la IA con su superficie de exposición intacta. Gobierno del uso de secretos a escala —políticas declarativas sobre qué procesos son de confianza, procedencia verificable de cada uso sensible, auditoría que un regulador y un directorio pueden creer—, todo sin que kovra ni el nest vean jamás un secreto.

Es el mundo de la visión hecho operación cotidiana: la aceleración de la IA y la custodia del riesgo, por fin, en la misma sala.

11 Por qué ahora, por qué nosotros

Por qué ahora. La adopción de agentes de IA está en su punto de inflexión, y la exposición que trae aún no tiene dueño en la mayoría de las organizaciones. La ventana para definir cómo se custodia el uso de secretos en esta era se está abriendo justo ahora —antes de que el problema se vuelva titular.

Por qué nosotros. Kaeus Inc no es una empresa de seguridad que salió a fabricar y vender un producto. Somos una **empresa de desarrollo de software** que se topó con este problema en carne propia: la misma exposición que describimos aquí nos afecta a nosotros —y a nuestros clientes— cada vez que ponemos un agente a trabajar. kovra nació de esa necesidad: lo construimos para resolver nuestro problema, lo usamos todos los días, y lo abrimos porque el problema es de todos. Ahí está nuestra credibilidad —no vendemos miedo desde afuera; cargamos el mismo riesgo que tú—. Y lo construimos sobre una postura, no sobre una función agregada —el humano como principal, la honestidad como ventaja defendible, la soberanía del dueño como invariante—, eligiendo el camino más difícil porque es el único que conduce a la visión.

> **kovra — hiperproductividad altamente segura.** Acelera a la velocidad de la luz sin estrellarte contra una roca. kovra actúa; tu huella autoriza —para que la era de los agentes sea la de la productividad acelerada, no la de las credenciales filtradas.

Conoce kovra

- Sitio web: https://kovra.sh
- Código fuente: https://github.com/kaeus-inc/kovra-core