kovra

FOUNDING DOCUMENT · JUNE 2026

Vision, mission, principles and values

The founding document for the age of AI agents

A product of Kaeus Inc

kovra.sh

01 The acceleration paradox

We are living through the largest jump in productivity since the arrival of the internet. Artificial intelligence no longer merely suggests: it **acts**. Autonomous agents write code, move data, run tasks, and operate systems on behalf of people. What used to take weeks now takes hours.

But that speed carries a price almost no one has put on the table honestly: **for an agent to work for you, you usually have to give it the same access you have.** The keys to your database. The access to your services. Your production credentials. And unlike a person, an agent **acts on everything it reads** —even when what it reads was crafted to deceive it.

This is the paradox: the more capable the agent, the more access it needs; and the more access it has, the larger the surface through which an organization can be compromised. Today, the market's tacit answer is to resign itself to a trade-off: either you go fast, or you stay protected.

We reject that trade-off. We believe you can capitalize on all the acceleration that AI enables **without widening your exposure surface**. It is not choosing one of the two worlds. It is having both.

02 The world is changing faster than risk can keep up

The first place this is felt is engineering —agents reading repositories full of secrets—. But it would be a mistake to think the problem lives only there.

Agents are entering **every function of the company**: finance, sales, operations, support, marketing. An agent that answers customers needs the keys to the customer system. One that automates reports touches the sensitive data source. And one that prepares payments or reconciles transactions needs **the online banking credentials —the keys to the accounts where the company's funds are held. These agents, the ones operated by business areas, tend to be as exposed as engineering's, or more** —because the people who operate them are not security specialists, and because no one designed those workflows with an actor that acts on what it reads in mind.

An organization's exposure surface no longer grows with the number of employees. It grows with the number of **agents**, and each agent is a new place where a credential can leak, be misread, or be turned against its owner. Risk is multiplying at the speed of AI adoption —and most organizations will not see it coming until it happens.

03 The risk, in business terms

You don’t need to understand cryptography to understand what is at stake. One question is enough for the risk owner:

> How many of the secrets that hold up your operation could an AI agent read today —and what would happen if one of those agents were steered, through a manipulated piece of data, to use them against you?

The consequences are neither hypothetical nor technical: an exposed production credential, a key that ends up outside the organization, a critical system touched by a process that believed it was doing its job. The cost is not measured in lines of code —it is measured in disruption, in lost trust, in regulatory exposure, in reputation.

And take it to the most visceral case, the one any board understands: the **online banking** credentials —the keys to the accounts where the company’s funds are held— are, to an agent, just one more secret among those it can read and use. A steered agent does not need to break into any system: it uses a legitimate credential believing it is doing its job, and the money is already gone. Would you hand those keys to an assistant that just read an email from a stranger? Today, without knowing it, many organizations already do.

And the most uncomfortable part: today that risk is **invisible** to whoever must answer for it. The conversation about AI productivity happens in one room; the one about credential exposure, in another. kovra exists to bring those two rooms together —and so that the answer is not to slow down, but to govern.

04 Why the current model falls short

For decades, the way to protect a secret was to **guard it well**: keep it encrypted, control who can request it, hand it to whoever asks. That model assumes something the age of agents breaks: that **whoever requests the secret is trustworthy**.

And you don’t even need an attacker for the model to fail. Day to day, when an agent is entrusted with a task that requires access to a system, it does the most natural thing: it offers to solve it and, with the best of intentions, **asks for the credential** to get in. There is no malice —there is naïveté—. The agent does not understand the implications of what it asks for, nor does it take on the responsibility of safeguarding what it receives. Asking for a key and keeping it with the weight it deserves are two different things; an agent, by nature, only understands the first. **The best-intentioned agent is still a poor custodian.**

And that is only the gentle face. The same door —an agent that acts on what it reads— can be opened against you: a poisoned document, a manipulated error message, or a hidden instruction are enough to steer it toward a bad action while it believes it is doing its job. In both cases —the naïve and the steered— the classic model, guard and hand over, gives it the secret in the clear and trusts it to use it well. The instant the agent receives it, control is lost.

The problem, then, is not to guard the secret better. It is to **rethink who receives it, and for what**, when whoever asks for it —in good faith or steered— cannot take on the responsibility of safeguarding it.

05 kovra's answer

> **Vision.** The acceleration AI enables should not force a choice between going fast and staying protected. We imagine organizations that capitalize on all of their agents' autonomy —in engineering and in every business area— without widening their exposure surface: where delegating privileged work to an agent never means surrendering control of the credentials that work depends on. It is not a trade-off; it is the best of both worlds — speed and custody, at once.
>
> **Mission.** To be the custodian of a secret's use in the age of agents. kovra lets tools and agents **use** secrets without **seeing** them, keeps the human as the principal at every dangerous point, and makes the secure path the convenient one. We start with the developer, where the risk is sharpest, and we extend the same guarantee to every agent an organization puts to work.

The central idea is a shift of frame: from guarding the secret to **guarding its use**. The agent can work —it can ask to use a secret— but it never becomes the owner of the most critical values. When an action truly matters, the agent does not decide it: a human **authorizes it with a single gesture** (their fingerprint), seeing exactly what is about to happen. kovra acts; you authorize.

And all of this without kovra —or any cloud— ever seeing your secrets: they are **encrypted at the source**, and control always remains with their owner.

06 Why it is defensible

A good vision is not enough; there has to be a reason it is hard to copy. kovra's advantage rests on two pillars that reinforce each other:

1. Guarding the use, not just the storage. Anyone can store an encrypted secret. What is hard — and valuable— is letting an agent orchestrate work with a secret it never gets to control: at the

most dangerous point, the secret is used through a trusted executor, with the exact action in view and a human's approval. That is the difference between handing over a key and operating the lock for whoever needs it.

2. Blind custody, wherever the secret lives. The secret is encrypted at the source and the owner never loses control —whether that secret lives on a single machine, is shared across a team, or is synchronized through kovra's cloud. The infrastructure that carries it can never read it. It is the difference between trusting a provider and **not needing to trust one**.

Together, these two pillars are not a feature you add; they are a posture you have to adopt from day one. That is why we choose **depth over reach**: we would rather be irreplaceable in what truly matters than ubiquitous in what anyone can offer.

07 Honesty as an asset

Most security tools promise more than they can deliver. We do the opposite, on purpose.

There is a truth no tool can erase: to use a secret, at some moment it has to exist in the clear. No one removes that "last mile" —not us either—. And we say so openly. What kovra does is **shrink how much sees a secret, keep the critical values away from the manipulable link, and put a human at the decision that matters.**

This honesty is not a weakness: it is the asset. A risk owner cannot sign off on an absolute promise —they know none exists—. They can sign off on a clear, bounded, verifiable guarantee. **Whoever tells the truth about what they do not do is, precisely, the one to be believed about what they do.** That credibility is what we guard most.

08 Our principles

These principles are the filter we run every decision through. They describe the world we want to build.

1. **The human is the principal.** The agent requests use; it never becomes the owner of a critical secret. Authority flows from a deliberate human gesture, not from whoever asks.
2. **Honesty is the advantage that can't be copied.** We never promise more than we guarantee. Credibility is the asset —we protect it above the hype.
3. **Acceleration and protection are not traded.** We neither slow the agent down to stay safe, nor loosen security to go fast. Inconvenient security is security that goes unused.

4. **Sovereignty over the secret belongs to its owner.** Control always remains in the owner's hands; the secret is encrypted at the source and kovra never sees its contents in the clear — wherever it is deployed.
5. **Protection is graded by danger.** Not all secrets are equal: protection grows with sensitivity, and the most dangerous use is the most governed.
6. **Depth over reach.** We win by being irreplaceable at the critical point, not by competing for ubiquity. When we have to choose, we choose depth.
7. **Designed for agents, not trusting them —and verifiable.** We assume the agent may be compromised; and everything is inspectable. Trust is earned by being verifiable, not by asking for it.

09 Our values

If the principles describe the world we want, the values describe **how we behave** while building it.

- **Custody.** We care for what is entrusted to us as if it were the most valuable thing we have. Responsibility for another's secret weighs more than our own convenience.
- **Radical honesty.** We say what we do not do. We would rather lose a sale than a truth. Trust, once broken, does not mend.
- **Depth over noise.** We build real advantage, not smoke. Craft and technical rigor before the fashion of the moment.
- **Verifiable, not merely credible.** We do not ask to be believed; we give the means to check. Being open and auditable is a posture, not a concession.
- **The human at the center.** Technology serves the owner; it never replaces them in the decision that matters. Convenience is never built at the cost of their control.

10 Trajectory — three horizons

The vision is large; the path is deliberate. We advance by horizons, earning the right to the next with what we delivered in the previous one. (A narrative trajectory, with no committed dates: the order is firm, the calendar is earned.)

Horizon 1 — Today: the developer (in production)

kovra already guards the use of secrets for developers and their coding agents: use without sight, the human as principal authorizing with their fingerprint, local by design and with its

source code open to inspection. It is the beachhead —the place where the pain is sharpest and the solution most demonstrable—. **This is not a promise: it already exists and works.**

Horizon 2 — Near term: the connected organization

Extend the same guarantee beyond a single machine and beyond the coding agent:

- **Shared use and synchronization with blind custody** —kovra's cloud, "the nest"— where teams share the use of secrets without the infrastructure being able to read them.
- **Deployment on the organization's own infrastructure** for those that require everything to live within their own premises —with the same guarantee of encryption at the source.
- **The guarantee extends to business agents**, not just engineering's: the area operating the agent capitalizes on AI without widening its exposure.

In every case, one constant: **the owner of the secret is the organization, and the nest can never see a secret.**

Horizon 3 — The vision: custody of use at organizational scale

The destination: an organization that capitalizes on all of AI's autonomy with its exposure surface intact. Governance of secret use at scale —declarative policies over which processes are trusted, verifiable provenance for every sensitive use, an audit a regulator and a board can believe— all without kovra or the nest ever seeing a secret.

It is the world of the vision turned into everyday operation: AI's acceleration and the custody of risk, at last, in the same room.

11 Why now, why us

Why now. The adoption of AI agents is at its inflection point, and the exposure it brings still has no owner in most organizations. The window to define how the use of secrets is guarded in this era is opening right now —before the problem becomes a headline.

Why us. Kaeus Inc is not a security company that set out to manufacture and sell a product. We are a **software development company** that ran into this problem firsthand: the very exposure we describe here affects us —and our customers— every time we put an agent to work. kovra was born of that need: we built it to solve our problem, we use it every day, and we open it because the problem is everyone's. That is where our credibility lies —we do not sell fear from the outside; we carry the same risk you do—. And we built it on a posture, not on an added feature —the human as principal, honesty as a defensible advantage, the owner's sovereignty as an invariant—, choosing the harder path because it is the only one that leads to the vision.

> **kovra — highly secure hyperproductivity.** Accelerate at the speed of light without crashing into a rock. kovra acts; your fingerprint authorizes —so that the age of agents is the age of accelerated productivity, not of leaked credentials.

Discover kovra

- Website: https://kovra.sh
- Source code: https://github.com/kaeus-inc/kovra-core