Engineer Software

Security for the AI agents you run

Access reviews and controls for AI agents: an agreed inventory, prioritized findings, permission tests and a verification record.

Understand what your AI agents can reach and put controls around that access. An agent connected to mail, documents or an API introduces permissions and failure paths that need review. We inventory the agreed systems, assess the risks, implement approved controls and test the restricted actions. You receive the access map, findings and verification record.

Turn the service into a result

Make “Security for the AI agents you run” serve a business outcome. We can connect the technical work to the users, operations and decisions it needs to support. Every agent's access mapped and held to what it needs, every action logged where the agent cannot alter it, and every safeguard tested rather than trusted.

Explore our services Get in touch →

Examine the relevant work

Use the related work on this page to examine our approach. Ask about the tradeoffs and the conditions behind each result. Built fail-closed rate limiting and anti-abuse — 22 rate limiters that fail closed

Built fail-closed rate limiting and anti-abuse Get in touch →

Mapped: every agent and what it can reach

We inventory agents within the agreed scope, including approved integrations and other connections identified with your staff. For each, the map records credentials, tools and reachable data. You receive the map with the findings and responsibility for keeping it updated is agreed as part of handover. It is a record of the reviewed environment, with its scope and gaps made explicit.

Locked: each agent on least privilege

An agent gets its own identity and the smallest privilege its job needs: roles rather than shared accounts (RBAC), read-only by default, writes only through narrow tools with limits on how often and how much. Secrets are split per agent and per host, never pasted into a prompt, and rotated without a code change. External text an agent reads — an email, a web page or a document — is treated as untrusted data. Tool permissions and validation checks enforce the agreed boundaries; a message saying “ignore your rules” is not authorization for an action. Controls fail closed: when a check cannot run, the action is refused.

Recorded: what it did, and for whom

Every tool call is logged with the person it acted for, what it was asked, what it touched and what came back, and the log is kept where the agent itself cannot change it. Alert rules are agreed for the system. They can cover a burst of reads, use outside expected hours or access beyond the agent’s agreed data scope. When something does go wrong, the record is what turns an incident into a timeline rather than a guess.

A map, a lock and a log **are three claims** —
until somebody tries to get past them.

The agreed controls are checked against representative access, misuse and recovery cases before handover.

A lock is tested, not trusted

We test the access controls within the engagement’s scope and record what was attempted, refused or still needs attention. Tests can cover denied tool calls, rate limits, credential isolation and prompt-injection attempts. The result is a reviewable verification record and a set of repeatable checks, with their limits stated. Incident work and credential rotation are scoped separately when an existing exposure needs containment.

The work behind it

We have done this work on servers, identity systems and platforms. We wrote 1,000 RBAC rules across Linux and database servers, led access management across M365 and single sign-on, and ran a continuous security scanning programme with 21 written audits. We split every operational secret per host, added a one-second preflight for secrets, found three locks on the door, none of them locked, and proved the intrusion ban works on every run. We built fail-closed rate limiting and wrote a scope rule after a leak. How we keep our own house is on how we work securely.

How an engagement runs

It starts read-only. We map the agents in the agreed scope, their credentials and what they reach, and hand you the findings in order of risk, each with its reason and the fix. The fixes are then made in your configuration and your repository, through your review, with checks that fail when covered access paths exceed the agreed permissions. You keep the map, the record and the tests — the file you hand to an auditor or an insurer — and the data agents read can be permissioned the same way. It is bought like the rest of our work, on the terms in what it costs.

If agents already hold keys to your systems, tell us which ones.

Every agent you run **holds a key to something** — know which, and to what.

It starts read-only: a map of the agents and credentials in scope, with findings ordered by risk.

Ask for an agent map What it costs →

Scope the work before commissioning it

Bring the deployment path, the operational risks and the people on call. We can scope a change with checks, a recovery route and clear ownership.

See what it costs Get in touch →

Talk through your requirements

Tell us what you need from “Security for the AI agents you run”, including the deadline and the team that will maintain it. We can discuss a practical way forward.

Explore our services

welcome@engineer.company +45 50 10 61 84 Copenhagen, Denmark

Take this page with you

Contact card

Carefully Crafted With Love by D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Copyright © 2025 – present · Engineer ApS · All rights reserved.

Printed from https://software.engineer.company/ai-security/