

Authorization for autonomous software has one final boundary, and it is the host, at the act.

SentoClaw builds that layer. Per-act authorization, enforced where the act commits, deployed as software beside the identity and policy systems you already run.

An agent holding a valid identity and a valid credential is steered by a prompt injection into reading a second repository its granted authorization does not name. The credential has not changed, so every control above the act permits what comes next. The kernel sees the act reach outside the granted region, and the read is denied before it commits.

ACT IS LEGIBLE

OBSERVATION UNFORGEABLE



Both properties are required, and they move in opposite directions. Legibility is strongest at the application and dissolves as the stack descends. Unforgeability is weakest there and strengthens. They overlap in one band, and that overlap is where the boundary falls.

WHAT PER-ACT AUTHORIZATION CLOSES

- **Possession stops being sufficient.** A credential without the matching observed origin cannot be redeemed. Theft gets you the artifact and none of the authority.
- **The name stops standing in for the act.** A tool name sits above the act and can be substituted. The act itself is what the kernel sees, and the act is what the grant is written against.
- **Authority stops carrying forward.** Short-lived permissions still rely on presentation. Authority here is observed at each act and does not propagate to the next one.
- **Delegation stops widening as it travels.** Scope is fixed when the grant is issued, and origin is established by what the kernel observes at each hop. A poisoned agent cannot credit itself as the origin or widen what it was given.
- **The tool does not have to cooperate.** The governed process is unmodified and holds no credential of its own. Nothing about it has to be trusted, because nothing it says enters the decision.
- **The decision stops requiring a round trip.** The observation and the verdict both happen on the host, with no network call in the path of the act.

The kernel has always been able to decide at the act. What it has never had is a grant to decide against.

THE BET THAT BROKE

A credential encodes, at issuance, the set of actions its holder may legitimately take. That prediction was nearly free while software was deterministic, because the action set was fixed at compile time. Every control above the act inherits that property: policy engines reason about the credential, gateways route on it, continuous evaluation re-checks it.

Autonomy defeats prediction. Agents now run for two weeks against more than a hundred tools, chaining sequences nobody enumerated. One autonomous intrusion was reconstructed at roughly 17,600 actions over two and a half days, with human review as the backstop under every reactive control in the stack.

The response has been to shorten credential lifetime, and the failure lives on the scope axis. A short-lived credential for a broad scope is still a broad grant, and as an agent's action space widens, the scope it needs converges on whatever it might need. Time was never the axis that was failing.

WHY THE BOUNDARY IS WHERE IT IS

Authorization needs two things at once. The act must be legible in terms of policy about acts can speak. And the observation must be unforgeable by the actor, meaning the party being authorized is not the source of the evidence used to authorize it.

These move in opposite directions. High up, every observation of an act is a claim made by the thing being judged. Low down, observation becomes incorruptible and the act dissolves. They overlap in one band, the kernel's view of the act. Everything above it is a prediction about a moment only it can see.

WHAT IDENTITY CANNOT EXPRESS

Non-human identity, NHI, takes as its object a principal and a permission. That is the right object for the question it answers, and it is a statement made before the act, about a holder. What it cannot express is a relationship between an authorization and one particular act, taken by one particular process, at one particular moment. A principal is not an act, and issuance is not a moment.

The observed binding supplies the missing term. Identity and policy decide who may and what may, and that decision authors the grant. The binding ties that grant to the act as the kernel measured it, at the instant the act commits. Without the grant there is nothing to decide against. Without the binding the grant is a permission again, and permissions are held rather than spent.

WHAT IT DOES NOT DO

It contains authority rather than influence: a response can still inform a plan the agent executes inside its legitimate scope. It binds on what was observed rather than parsing what was presented, so no payload inspection and no proxy appear in the architecture. And it requires control of the host, which a tenant inside a managed runtime lacks, though the operator of that runtime does not.