

GOVERNED WORKFORCE • SECURITY ANALYST

The governed security-analyst agent.

A zero-trust SOC agent that triages and analyzes under a capability token, a deny-by-default policy, and a signed audit chain — one governed job at a time.

● Asphodel + Tiresias-ZT • ACME Corporation

THE STAKES

An analyst you can trust needs proof, not permission.

Security triage is the one seat where an autonomous agent touches the highest-signal, most sensitive data. If it reasons off-policy, escalates silently, or writes where it shouldn't, the blast radius is the org itself — 325,154 principals across 28 departments.

325,154 DIRECTORY PRINCIPALS	28 DEPARTMENTS	205,034 NHI / SERVICE IDENTITIES	
--	--------------------------------------	--	--

Source • Plane 1 Directory/Identity — Samba-AD, DC=acme,DC=example (120,120 humans + 205,034 NHI).

THE WORKFORCE

367 governed agents. This is one of them.

The security analyst is a first-class tenant in a live fleet — every agent enrolled, policy-scoped, and metered. No agent is an exception to the plane it runs in.

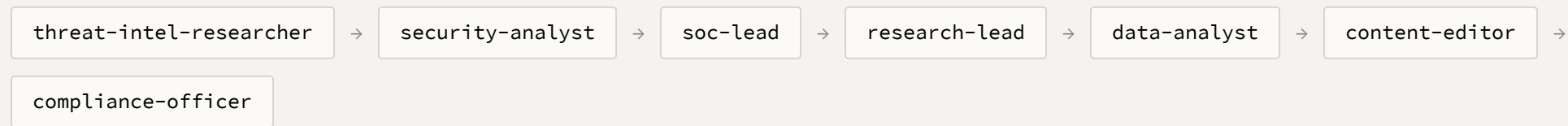
367 GOVERNED AGENTS (REGISTRY)	7 SHOWCASE SOC AGENTS	360 FLEET AGENTS • 28 DEPTS	
--	---	---	--

Source • Plane 2 registry (367 agent tenants, postgres) • Plane 5 runner — personas.json + agent-manifest.json.

WHERE THE ANALYST SITS

The 14-edge SOC hand-off DAG.

The showcase agents form a topological hand-off graph. The security analyst is the second node — it receives from threat-intel, then hands off up the chain to the SOC lead. Position is policy: it can only accept and emit where the DAG allows.



Source · Plane 5 — 7 showcase agents, 14-edge topological hand-off DAG (`runner.py schedule --interval 3600`).

GOVERNANCE

Capability token in. Deny-by-default at the door.

The analyst never talks to data directly. Every governed call carries a tenant API key and an HS256 capability token the agent mints from its own 64-hex secret. The PEP validates both, evaluates policy, and denies unless explicitly allowed.

WHAT THE AGENT PRESENTS

`X-Tiresias-API-Key` tenant identity

`X-Capability-Token` HS256, self-minted from 64-hex secret

WHAT THE PEP DOES

Validate key + token → evaluate policy → **deny-by-default**. On allow, forward the request **verbatim** to this tenant's soul-svc backend — nothing added, nothing rewritten.

Source • Plane 3 Enforcement (PEP acme-tiresias-zt, host :8343) — deny-by-default, verbatim forward on allow.

LIVE PROOF

Real enforcement, real numbers.

This is not a mock. Every governed call is validated at the enforcement plane, signed into an Ed25519 audit chain, and metered into the usage store behind the ROI board.

PEP	Enforcement plane, published host port acme-tiresias-zt · site acme-site · 367-agent tenant slice pulled via ETag-cached registry	:8343
AUDIT	Signed audit chain shipped to CP SoulKey Ed25519 · durable WAL at /data/audit-wal	Ed25519
DRIFT	SoulWatch behavioral drift → auto-quarantine shared postgres quarantine store · SIEM via RFC5424 syslog	auto-Q
USAGE	Per-call cost + tokens into the usage store tzt_usage → Grafana ROI board (uid acme-roi-economics)	tzt_usage

Source · Plane 3 Enforcement + Plane 6 Observability – audit chain, SoulWatch, tzt_usage.

One triage job, end to end.

The analyst picks up an alert, reasons over it, and writes its conclusion to memory. Every hop is governed — the LLM answer becomes a governed chain-of-thought and a memory write into the ACME data tier.



| The LLM produces the analysis. The plane decides whether it is ever allowed to persist.

Source • Governed request flow (one job) — cap-token → PEP → soul-svc tenant → PostgREST → acme_soul.

Governed CoT, then a memory write to acme_soul.

The analyst's reasoning is captured, not discarded. Two governed endpoints land the work in a tenant-isolated database where row-level security is forced — the agent can only ever see and write its own rows.

POST	Capture the reasoning trace governed chain-of-thought, source_type + authored_by required on writes (ContentInspection enforce)	/v1/cot/capture
POST	Persist the triage conclusion written to acme_soul via the agent's own soul-svc tenant	/v1/memory/write
RLS	Forced row-level security on _memories cross-tenant reads impossible without SET ROLE service_role — postgres itself is subject to FORCE RLS	FORCE

Cheap by default. Powerful only on cause.

The analyst homes on a free model and stays there for routine triage. Incident keywords escalate it to a standard paid tier; on any LLM failure it degrades to a free floor rather than going dark. Spend follows severity, never habit.

TIER	ROLE	MODEL	\$/1M-TOKEN	ECONOMICS
	Home (routine triage)	hermes-405b:free		\$0
	Escalation (incident/breach/critical)	sonnet-4.5 · standard	\$0.03 / call	
	Ceiling (deepest analysis)	opus-4.8	\$0.19 / call	
	Floor (fast free degrade)	llama-3.2-3b:free		\$0

Source • Plane 5 model roster + keyword escalation • Plane 6 ROI board observed \$/call (opus \$0.19 • sonnet \$0.03).

WHY FREE-HOMING HOLDS

Fallback-to-floor is the reliability story.

Free tiers saturate — hermes-405b:free returns chronic 429s under load. The analyst's floor is a free model chosen for reliable degrade, so a rate-limited home never becomes a stalled agent. That saturation is exactly the self-host case the ROI board makes.

\$0.0015 GPT-OSS:120B / 1M TOKENS	\$0.0012 LLAMA-70B (HF) / 1M TOKENS	\$0.002 DEEPSEEK-V4-PRO / 1M TOKENS	
---	---	---	--

Source • Plane 6 ROI — free-tier saturation (hermes-405b:free chronic 429) → self-host economics.

AUDIT & OVERSIGHT

Every job is reviewable after the fact.

Governance does not end at the write. The analyst's calls are signed, its memory and reasoning are inspectable in the console, and drift trips an automatic quarantine — oversight is continuous, not periodic.

MANAGEMENT CONSOLE

acme-asphodel-ui (:3001) · console.asphodel.ai — agents, memory / CoT review, mindsets, access & route policy, backend reachable.

ANALYST / REVIEW CONSOLE

PEP /admin/* · pdp.tiresias.watch — audit/verify, soulwatch/status, quarantine, analytics: trajectory · model_baseline · drift.

Source · UIs — asphodel-ui (management) + PEP /admin review console (role-token gated).

THE TAKEAWAY

A security analyst that earns its autonomy.

Deny-by-default at the door. A capability token on every call. Reasoning captured, memory RLS-isolated, spend matched to severity, and a signed chain behind all of it. One of 367 — governed the same way, provable the same way.

Give an agent the SOC. Keep the proof.

● Asphodel + Tiresias-ZT • ACME Corporation • Governed AI Workforce