Agentic AI demands a fundamentally different security architecture.

Waxell is the control plane for agentic systems: threat-mapped against OWASP, extended beyond it, policy-enforced at every layer, and audited down to the individual tool call.

THE PROBLEM

High-stakes agents. No existing framework covers them.

- ✗ No framework covers agents that autonomously decide which tools to invoke, in what order, with what arguments.
- ✗ Memory poisoning slips past standard controls: the provenance looks correct while the content is corrupt.
- ✗ Runaway autonomous loops turn into runaway spend before anyone notices.
- ✗ Agent actions land on shared service accounts, not on a specific agent or person.

```
// Threat model
owasp_agentic_threats: "T1-T15"
waxell_extended_threats: "W1-W10"

// Controls
policy_layers: 3
isolation_model: "per-tenant"
audit_retention: "by plan, up to 365d+"
soc2_type_ii: "in_progress"
```

OWASP AGENTIC AI THREATS: T1-T15

T1 Memory Poisoning	**T2** Tool Misuse	**T3** Privilege Compromise	**T4** Resource Overload	**T5** Cascading Hallucination Attacks
T6 Intent Breaking & Goal Manipulation	**T7** Misaligned & Deceptive Behaviors	**T8** Repudiation & Untraceability	**T9** Identity Spoofing & Impersonation	**T10** Overwhelming Human-in-the-Loop
T11 Unexpected RCE & Code Attacks	**T12** Agent Communication Poisoning	**T13** Rogue Agents in Multi-Agent Systems	**T14** Human Attacks on Multi-Agent Systems	**T15** Human Manipulation

WAXELL W1-W10: NOT IN EXISTING FRAMEWORKS

W1 Financial Cost Exhaustion	**W2** Cross-Tenant Leakage	**W3** Causal Lineage Loss	**W4** Model / Provider Drift	**W5** SDK Supply-Chain Risk
W6 Prompt-as-Code Deceit	**W7** Behavioral Regression	**W8** Telemetry Self-DoS	**W9** External-Agent Governance	**W10** Agent Identity Lifecycle

WHAT WAXELL DELIVERS

Full Agentic Observability

Every run, LLM call, tool invocation, and child workflow captured in a causal trace tree with inputs, outputs, costs, and latencies.

Three-Layer Policy Enforcement

Pre-action prevention at the MCP Gateway. In-session shaping via playbooks and SDK. Post-action audit with immutable trace evidence.

Cost Governance

Token budgets, model cost tables, per-tenant overrides, and alerting. Financial DoS treated with the same seriousness as data DoS.

Per-Tenant Isolation

Tenant-isolated data. Scoped API keys, memory, and cost overrides. No shared logical surface across customers by construction.