

Sargha Sensor: Detection-Accuracy Benchmark

Planted fixtures and clean controls, run through the real collectors.

REVISED

4 October 2026

PRODUCT

Sargha by Araghatta

Security tools are only as good as what they actually catch. We benchmark our own detection the honest way: plant a fixture of known issues alongside clean controls, run the real collector/adaptor against it, and measure **recall** (planted issues detected) and **false positives** (findings in clean files). No estimates.

CAPABILITY	WHAT IS MEASURED	RECALL	FALSE POSITIVES
Endpoint sensor: discovery	API keys, MCP configs, AI CLIs/SDKs, local skills across a planted <code>\$HOME</code>	100%	0
Source Code Analysis (SAST)	Planted vulns vs clean controls (semgrep pack)	100%	0
Secret & Credential Scan	AWS/GitHub/Slack/Stripe/OpenAI keys + private keys	100%	0
Dependency / Supply-chain (SBOM)	Vulnerable + typosquatted packages vs safe pins	100%	0
Container & IaC posture	Misconfigurations vs hardened baselines	100%	0
AIBOM & License/Provenance	Model files, pickles, copyleft licences, unpinned revisions	100%	0
Agent skill content	<code>SKILL.md</code> with shell exec, privilege steps, hard-coded creds, embedded scripts vs benign	100%	0
AI runtime stack	RAG apps, vector DBs, orchestrators by signature + heuristics (incl. in-house)	100%	0

PRIVACY BY DESIGN

Skill content is analysed on the host. Only boolean flags are recorded, never the raw file, so a planted hard-coded key is detected without the secret ever leaving the machine.

In production, each detector's precision, recall and false-positive rate is re-measured against a labeled fixture set, versioned to the detection content, and shown with 95% confidence intervals inside the product: measured, never a static marketing claim.

Benchmark methodology draws on planted fixtures and clean controls; no invented telemetry. Revised 4 October 2026.