LUMETRACE

AI Detection Capabilities

Capability overview for prospective clients - August 2026

Every organisation is now running AI it did not procure, and defending AI it did procure. Three questions follow, and a traditional security stack answers none of them: is the AI-driven traffic on my network legitimate or hostile, is my data leaking into third-party AI services, and is anyone attacking the AI systems I run myself?

Lumetrace answers all three from the network alone - passively, with no endpoint agents and no changes to your applications. Everything described here is running in production today.

SHADOW AI

Data leaving to AI services

Which staff and systems send company data to external AI tools.

AGENT ACTIVITY

AI agents on your network

Autonomous and scripted activity separated from human traffic.

IMPERSONATION

Fake AI crawlers

Bots claiming to be a known AI company when they are not.

CONTENT

Harvesting of your site

Collection to train a model, told apart from competitive monitoring.

YOUR AI ESTATE

Attacks on your own models

Theft, extraction, corpus poisoning and AI supply-chain exposure.

INDUSTRIAL

AI-driven OT reconnaissance

Machine-speed mapping of an OT estate, before anything is touched.

What Lumetrace Detects LIVE

- **Shadow AI and data leakage.** Which internal hosts are sending data to external AI services, which tools are in use, sanctioned or not, and when the volume leaving looks like a leak rather than a question. Most organisations are surprised by this inventory on day one.
- **AI agents and automated traffic.** Legitimate AI agents and business automation separated from malicious bots, scrapers and AI-driven attacks. Corroboration across multiple independent signals before anything is raised, because a false positive on automation is worse than silence.
- **Impostor AI crawlers.** Anyone can label their traffic as a well-known AI company's crawler, and many do to get past robots rules and rate limits. Lumetrace establishes whether the visitor is genuinely who it claims to be, so you can act on impersonation rather than on a name.
- **Content harvesting, classified.** High-volume retrieval of your site has two very different meanings: your content being collected to train or ground a model, and a competitor watching your prices. Lumetrace reports which of the two it observed, not a single request count.
- **Attacks on the AI you run yourself.** Your inference endpoints, models and retrieval corpus are assets that conventional monitoring does not recognise. Lumetrace discovers them, attributes who reaches them, and detects attempts to clone a model through its own interface, move model weights off the estate, tamper with the retrieval corpus, or extend your agents onto third-party tool servers outside your inventory.
- **AI-driven reconnaissance of OT.** For industrial operators, the cheapest place to stop an intrusion is before anything reaches a controller. Bulk interest in engineering artefacts and control-system naming is surfaced early - and attributed to machine agency only when that is independently evidenced, never inferred from speed alone.

Evidence an Analyst Can Act On LIVE

A finding nobody can corroborate is a finding nobody acts on. Every AI detection carries its own evidence, built only from what was actually observed on the wire: what was touched, why it was judged suspicious, and the same host's related activity on one timeline. Findings are mapped to MITRE ATLAS, the adversarial-ML counterpart to ATT&CK, with industrial findings additionally mapped to ATT&CK for ICS - so your analysts read them in a framework they already use.

- **Board-level rollup.** A single AI exposure score with a severity band and the breakdown behind it.
- **Your SIEM, not just ours.** Findings stream into Splunk, Microsoft Sentinel, QRadar and any syslog or webhook SOAR target, as structured fields rather than prose - so AI findings can be correlated and actioned in the console your team already lives in.

AI Detection Capabilities - continued

- **Executive and technical reporting.** Scheduled reports written for a board, with the supporting technical appendix and per-incident detail behind them.

Your incident data does not go to a third-party model
The summaries and reports are generated deterministically inside the platform. We do not send your telemetry to an external LLM to write them. Nothing leaves for a model we do not control, and no summary can invent a finding that did not occur - which matters both for data residency and for what reaches your board.

What a Passive Sensor Cannot See BOUNDARY

We would rather state the boundary than let you find it. Prompt injection and unsafe model output live inside the prompt and tool-call layer, where no network sensor can observe them. We can evidence that an autonomous agent is operating and what it reached; we cannot tell you it exceeded the remit it was given, because that remit is not on the network. Closing those would need a different class of telemetry - it is a product boundary we have drawn deliberately, not a gap we have quietly left out of this document.

How It Deploys LIVE

- **Passive.** A mirrored copy of traffic. We are never inline, so we cannot become an outage or a bottleneck in your network.
- **No agents, no code changes.** Nothing to install on endpoints, servers or applications.
- **Live traffic or captures.** Runs continuously on a mirror, or against packet captures you already hold - which is also the fastest way to evaluate us.
- **On-premise or hosted.** Available as a sealed on-premise appliance for operators who require that data never leaves their estate.
- **Per-tenant isolation.** Each client sees only their own data.

Where This Fits In Your Stack

Lumetrace is not a replacement for the controls you already run - it covers the ground none of them stand on. Endpoint tooling cannot see an unmanaged device, a contractor's laptop or a controller it will never be installed on. Cloud access brokers see sanctioned applications, not the AI service someone reached directly. A firewall sees the connection, not whether the thing on the other end of it was a person, an agent, or an attacker wearing an agent's name.

- **Detection, not interception.** We read a mirrored copy and report what happened. We do not sit in the path of your traffic, and we do not block - so adopting us cannot break a production process, which is the objection that stops most industrial deployments.
- **The unmanaged estate.** Anything with an IP address is in scope, including the equipment and third-party devices that will never accept an agent.
- **One view across IT, OT and AI.** The same sensor covers all three, so an intrusion that crosses from the corporate network into the industrial one is one story, not three consoles.

See It For Yourself

Every claim in this document is demonstrable live, on request - including the negative controls, where the same activity at human pace deliberately produces no AI finding. We will show you what we do not alert on as readily as what we do.

The fastest evaluation is a scoped pilot on your own environment. Within days you will see your live AI-service inventory, which hosts are sending data out to AI tools, any impersonated crawlers reaching your public sites, access to and attacks against your internal AI systems, any machine-speed reconnaissance of an OT estate, and your executive AI exposure score - with nothing to deploy on your endpoints.

TALK TO US
lumetrace.com - a scoped pilot, or a live demonstration against your own packet capture.