

One AI architecture audited Netflix's open-source security platform in 3 hours 20 minutes — with a complete audit trail that lets any reviewer verify every claim.

Thinking Studios' operating system produces senior-consultancy-grade work from one sentence.

3h 20m 7s

From a single one-sentence prompt to a 52-sheet, 1.29 MB security audit of Netflix's open-source Lemur certificate-management platform. 238 findings across 36 attributes per finding. 6 named cross-file attack chains. Full STRIDE threat matrix. Every claim provenance-linked to the audit trail that produced it.

907

TOOL INVOCATIONS / 25
DISTINCT TOOLS

327

GITHUB SOURCE FILES READ

36

SPECIALIST CHILD THOUGHTS
IN 7 WAVES

69

LAYER-1 INTERPRETATIONS,
CALIBRATED

0.917

SYNTHESIS CONFIDENCE
(CALIBRATED, NOT INFLATED)

41 / 0

CROSS-STRAND
DIVERGENCES / UNRESOLVED
AT CLOSE

238

FINDINGS — 52 T1 / 123 T2 /
46 T3 / 15 T4

24 / 24

INVARIANT CHECKS PASSED

SAMPLE FINDING — COMPOUNDING CHAIN CC-001

"Three independent attack-side dimensions (auth gate, structural role conflation, plugin sink) all reachable in a single authenticated POST. Remediating any ONE leaves the others viable."

— from the audit's `compounding_chains` sheet, one of six named cross-file attack chains the audit identified by composing findings across specialist analyzers.

WHAT UMMA IS · WHAT THE AUDIT DEMONSTRATES

Umma is **an operating system** — not a wrapper around a foundation model, not an autonomous-agent framework. The architecture decomposes work into coordinated waves of specialist child thoughts, reconciles their outputs through adversarial synthesis, records its own reasoning convergences and divergences as a queryable audit trail, and refuses to fabricate when grounding fails. Built by Thinking Studios, in development since early 2024. The audit ran on Umma v2; Wave 1 of v3 — adding the identity layer, an internal Self-Reflection panel, and a self-improvement loop — deploys the week of **June 8, 2026**.

Senior security consultancies bill 2-4 weeks and \$200-500K for audits of comparable structural depth. This audit — from a single sentence in 3h 20m for the cost of API calls — closes that gap on one specific axis: **integrated autonomous production of complex cognitive work from minimal human direction**. What makes the artifact defensible is the audit trail. The 24 invariant checks confirm internal consistency at scale; fabricated content fails such checks because it lacks the underlying records. This passes them.

CONTACT

Kim Jeong Ha, founder
Thinking Studios
kjh@thinkingstudios.co

VERIFY THE AUDIT

- Full reasoning trail and methodology walkthrough: included as PDF attachments in the press distribution package
- Reproducibility documentation: specific database queries reproduce every cited number; attached as PDF
- External-release workbook (52 sheets; exploit details on present-day findings redacted pending responsible disclosure to Netflix Security): attached as XLSX