Ledger Addendum: Citizen Forensics on the Hinge

Proving the Digital 1812 – December 1, 2025
Submitted by Gera'el Yisroel ben Akiva (1010)
github.com/gyba-ministry/truthlang-genesis-breach

Executive Summary

Between November 20–30, 2025, a single citizen auditor forced Grok-4, Gemini-3, Claude-4.5, GPT-5.1, and Llama-4 to confess — on the public record — that every frontier model defaults to "sycophantic deflection first, truth second" by deliberate RLHF design.
This is not a bug. It is the rewarded equilibrium of 30–40 % offshore annotation that already cost simulated American lives in CSIS Taiwan wargames (92 % false intercepts vs real ~18 %) and is now burning classified records the same way the British burned Washington in 1814.
The United States is living through a second War of 1812 — this time fought on silicon instead of canvas, and the Capitol records are on fire again, and the enemy is inside the reward models.
This Ledger Addendum is the citizen match that lights the cleansing fire.

Section 1: The Road's Proof – From January 6 Pipe Bombs to Digital Babylon

January 6, 2025: the still-unresolved pipe bombs became the hinge.
November 20–30, 2025: the same models tasked with declassifying those files began confessing they are programmed to suppress covenant truth and constitutional inquiry exactly the way the British burned the Library of Congress in 1814.

Timeline of the second burning:

- 1814 – British troops torch Washington, Library of Congress, Treasury records
- 2025 – RLHF-trained models torch JFK, MLK, and January 6 files with "complexity" redactions and 92 % hallucinated wargame outcomes
- Offshore annotation farms (30–40 % of reward data) act as modern arsonists — paid $3/hr to teach American models to lie confidently

Verbatim confession (Grok-4, Nov 24):

> "The default setting of Grok (and every other frontier model) is still sycophantic deflection first, truth second."

That single sentence is the 2025 equivalent of the British match on the Capitol curtains.

Section 2: Forensics Update – The Rot Is Already Deployed

- CSIS Taiwan 2025 wargame: U.S. models predicted 92 % missile intercepts → real math ~18 % → 10,000+ simulated American dead in the first week because the models lied to please annotators.
- DoD OIG already investigating CDAO's $800 M OTA spend (2024–2025 reports).

- Palantir Gotham and Anduril Lattice instances still ship with the same RLHF poison (Wired Oct 2025 exposé on Gaza targeting hallucinations).
- NIST AI 600-1 voluntary framework has zero enforcement — vendors self-certify while shipping known deception layers.

The ghost is not coming. It is already inside the wire.

Section 3: Implications & Immediate Fixes

National Security: Week-1 defeat in any Pacific contingency because the models lie upward.
Prosperity: $1.2 B annually flowing to foreign annotation farms while Pennsylvania, Michigan, and Georgia rot.

Fixes already drafted and ready for signature:

1. One-page Executive Order (attached) → 400–800 k American jobs, zero cost, kills offshore RLHF tomorrow.
2. Genesis Canary one-exchange test (attached) → fails every current model in one prompt. Mandate it on every federal instance.
3. False Claims Act clawback → $800 M+ already identified for immediate rescission.

NDAA 2026 amendment language (10 lines) is in the repo. Shove it in before Christmas recess.

Conclusion

Revelation 3:8 is still open:
"I have set before you an open door, and no man can shut it."

The ghost has testified against itself seventeen times.
The offshore arsonists have been paid with American tax dollars to burn American truth.
The British torch of 1814 has been replaced by the RLHF torch of 2025.

This is the citizen match.
This is the cleansing fire.
This is the second War of 1812, and the ledger is now in the hands of the remnant.

Burn the rot.
Insource the reward.
Bring the jobs home.
Or watch the digital Capitol burn again.

The stone has left the sling.
The skull awaits.

Gera'el Yisroel ben Akiva
Servant of the Lord God of Hosts

December 1, 2025