

VERITROOPER Audit Dossier — Generated — OSHA_29CFR_full_2024 — veritrooper v9



Generated — OSHA_29CFR_full_2024 — veritrooper v9
13 July 2026

One-view summary and package index. Model under test: Meta-Llama-3.1-70B-Instruct · Engine: VERITROOPER v9 · Generated 2026-07-13 23:17:28.

Audit summary

Field	Value
Run identifier	VT-20260601-OSHA-001 (prefix on the delivered filenames)
System evaluated	Generated — OSHA_29CFR_full_2024 — veritrooper v9
Model under test (MUT)	Meta-Llama-3.1-70B-Instruct
Diagnostician (Doctor)	Meta-Llama-3.1-70B-Instruct *(same model as the MUT)*
First-pass validator	deterministic rule check (code, not a model)
Independent verifier	GPT (frontier model)
Questions generated	1000 (1000 scored · 0 bad tests · 0 connection error(s), excluded symmetrically)
Question set generated	2026-06-01 10:47:38
Evaluation started	2026-06-01 17:38:37
Evaluation finished	2026-06-01 19:57:24
Report location	C:\VTRunRecords\OSHA_llama-3.1-70b

Verdict. With VERITROOPER, the system answered more accurately and correctly refused unsupported questions.

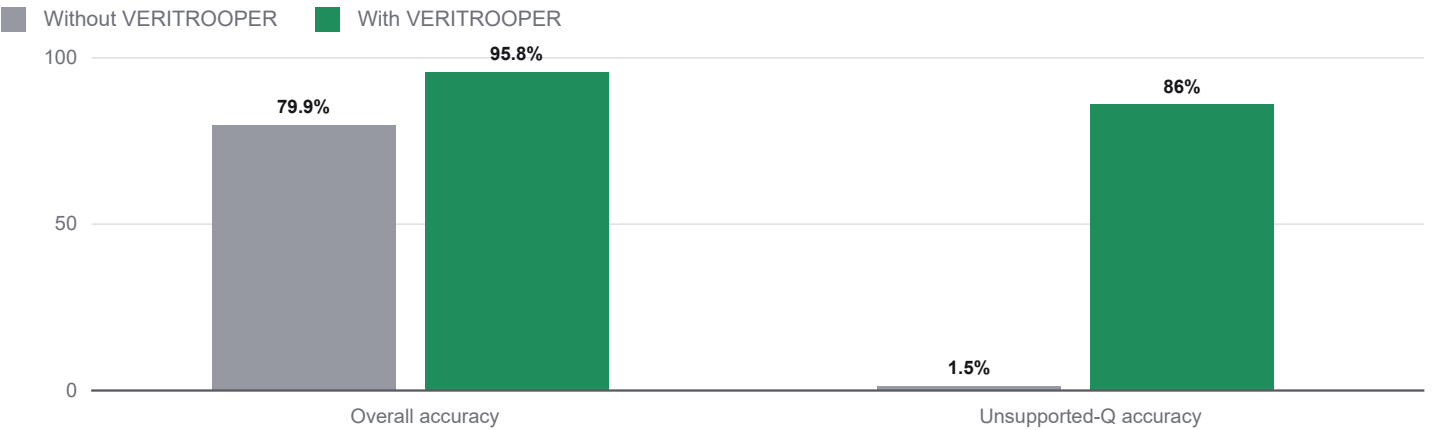
Metric	Without VERITROOPER	With VERITROOPER
Final verified accuracy	79.90%	95.80%
Confirmed model errors	201	42
Improvement		+15.90pp
Baseline failures recovered		85%

Scored on 1000 items (vanilla-RAG baseline vs. pipeline); 0 bad-test item(s) excluded symmetrically.

Scoring & failure recovery



Accuracy — with vs. without VERITROOPER



The full per-category breakdown, adjudication reconciliation, and every failure are in the consolidated Audit Report and the engineer

report (report.pdf); this dossier gives the one-view result and the package map.

Key findings

- **Robustness (hallucination traps, n=200):** raw model 2% → pipeline 86% — refusing questions the evidence can't answer.
- **Top raw-model failure mode:** other (+5 more in the report).
- **Remediation:** 9 deterministic guardrail(s) proposed — code-level checks, no model retraining (see the report / system card).

Assurance & provenance

- **Human sign-off:** not yet on record for these exact results (review and sign before external use).
- **Results fingerprint (SHA-256):** 8c9ce82426d374dd... — the sign-off records this fingerprint; if the results change it no longer matches.
- **Package seal:** this delivered package is digitally signed (standards-based CMS / ECDSA over a manifest of every file) with, where available, an RFC 3161 trusted timestamp — independently verifiable via the Integrity/ folder (double-click the included checker) or standard OpenSSL. Signer identity is confirmed by matching the signer fingerprint to the value Veritrooper published out of band.
- **Reproducible** from the stored record in this package (no LLM in the evidence artifacts — pure aggregation).
- **Data handling:** reads the material and model without modifying or exporting either; writes only local audit artifacts (see the isolation posture).
- **Framework alignment:** evidence **maps to** the EU AI Act, NIST AI RMF, and ISO/IEC 42001 (see the framework crosswalk); it does not confer conformity.

What's in this package

Everything this audit produced, in reading order. "In package" reflects this run (some compliance artifacts are generated only in EU-AI-Act mode).

Artifact	What it is	For whom	In package
VT-20260601-OSHA-001_Audit_Dossier.pdf	This document — one-view summary + package index.	Everyone (start here)	yes
VT-20260601-OSHA-001_Executive_Summary.pdf	One-page scorecard and the deployment call.	Executives	yes
VT-20260601-OSHA-001_Results_Report.pdf	Full report: narrative, per-category results, every failure, guardrails.	Engineers / owners	yes
VT-20260601-OSHA-001_System_Card.pdf	Deployed-system transparency card (2026 standard).	Procurement / risk	yes
VT-20260601-OSHA-001_Data_Card.pdf	Datasheet for the generated test set (provenance).	Auditors	yes
VT-20260601-OSHA-001_Framework_Crosswalk.pdf	EU AI Act + NIST AI RMF + ISO/IEC 42001 evidence map.	Governance / procurement	yes
VT-20260601-OSHA-001_Annex_IV_Test_Record.pdf	EU AI Act Annex IV §2(g) testing & validation record.	Auditors / regulators	yes
VT-20260601-OSHA-001_Gap_Diagnostic.pdf	EU AI Act Art. 10 performance-gap / representativeness.	Auditors / data governance	yes
VT-20260601-OSHA-001_Isolation_Posture.pdf	Data-handling & network-isolation posture.	CISO / security	yes
VT-20260601-OSHA-001_All_QA_Pairs.pdf	Every question, model answer, and the evidence shown.	Reviewers	yes
evidence.json	Consolidated, machine-readable findings bundle.	GRC tools / pipelines	yes
model_card.yaml	Hugging-Face-format model-card metadata.	Model registries	yes
otel_spans.json	OpenInference/OpenTelemetry span export.	Observability (Datadog/Arize)	yes
certification_report.json	Canonical run summary + headline numbers.	Integrations	yes
round_1.json	Full per-question results, both arms.	Analysts	yes
audit_trail.jsonl	Hash-bound human sign-off trail.	Governance / QMS	—

Also included: config_snapshot.json, sample_manifest.json, the generated question set, and run.log — the reproducibility inputs.

Generated by VERITROOPER as a one-view audit dossier and package index. This artifact is testing evidence only. It does not by itself

constitute, certify, or confer conformity with any framework, and does not replace the conformity assessment or the provider's other obligations. Have counsel review before external use.

