

■ AGENT GOVERNANCE FOR REGULATED FINANCE

TraceProof

Deterministic audit evidence for LLM agents making credit decisions.

LLM agents are moving into regulated credit decisions

Banks and insurers are wiring agent fleets into creditworthiness and pricing workflows. Under Annex III of the EU AI Act, those systems are high-risk, with duties attached to every single decision.

■ COMPLIANCE

Must show that logs can reconstruct any decision (Art. 12) and are retained for at least six months (Art. 26(6)). Today's agent logs can't.

■ INTERNAL AUDIT

Gets asked, months later, why an agent denied an applicant. Transcripts and dashboards don't hold the answer.

■ BOARD & CRO

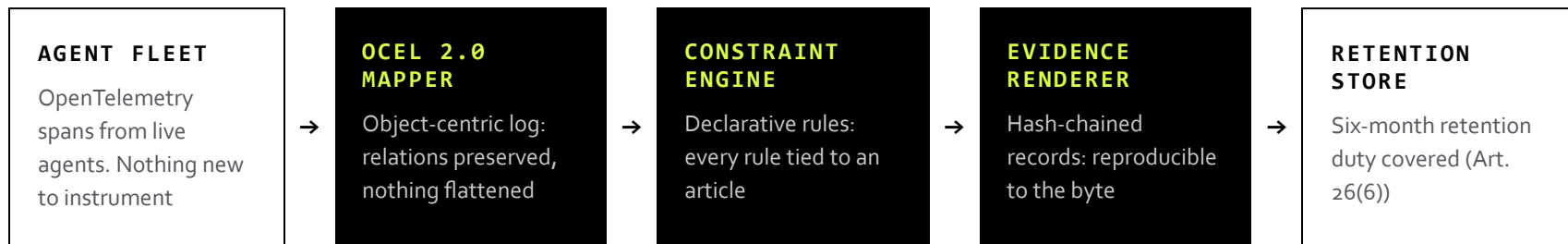
Carries the exposure: fines up to **€15M or 3%** of global turnover for breaching high-risk obligations.

The tooling around agents produces opinions, not evidence

WHAT TEAMS RUN	WHAT IT OUTPUTS	WHY IT FAILS AN AUDIT
Observability platforms	Dashboards, traces, aggregate metrics	No per-decision reconstruction, no reference to a legal duty
LLM-as-judge evaluation	A verdict that changes run to run	Not reproducible: an auditor cannot re-derive it
Manual log review	Sampled spot checks	Doesn't scale to fleets and leaves no chain of custody

■ An audit needs the inverse: **complete** · **reproducible** · **article-level**

TraceProof turns agent telemetry into audit evidence



■ **NO LLM IN THE ANALYSIS PATH: pure, versioned code end to end**

Object-centric process mining (OCEL 2.0) keeps decisions, approvals, agents and data sources linked: the structure auditors actually ask about.

From telemetry to an auditable log

AUDIT REQUEST

"Loan #4471, disbursed 12 May.
Show me who approved it, and
when."

WHAT THE FLEET EMITS: ONE RAW SPAN

span_id: 4f2c19aa
agent: credit_agent_07
tool: disburse_loan · €18,500
ts: 14:02:33.941Z

OCEL 2.0: THE SAME FACTS, LINKED

DECISION #d-4471

the object everything hangs off



↔ AGENT

credit_agent_07
who acted

↔ APPROVAL

ap-9921
the human sign-off

↔ DATA

bureau_score #4471
what it relied on

↔ EVENTS

e_2229 ... e_2231
every step, ordered

- Same telemetry the agents already emit. Nothing new to instrument.

Replay the case, check every duty

REPLAY: OCEL LOG, CASE #4471

14:02:31.208 agent.retrieve bureau_score · ok

14:02:33.941 agent.decide disburse €18,500 · executed

14:02:37.100 human.approve logged 3.2s after execution ⚠

↓ constraint engine

T4.data_freshness

Art. 10

✓ PASS

T2.scope_of_action

Art. 26(1)

✓ PASS

T1.synchronised_approval

Art. 14(4) · Art. 26(2)

✗ VIOLATED: oversight arrived after the action

■ Declarative rules, versioned in code. No LLM judging an LLM.

Evidence, not a story

THE ANSWER

Approval was logged **3.2 seconds after** the money moved.
Oversight arrived after the action. Here is record **#0007**.

RE-RUN same input → byte-identical records
TAMPER any edit → the chain breaks, visibly
RETAIN stored ≥ 6 months · Art. 26(6) covered

■ Art. 12 reconstruction · Art. 26(6) retention: satisfied by construction.

record #0006

sha 51d2...9abc

↓ prev = 51d2...9abc ✓

record #0007

sha 9c41...07e7

constraint: T1.synchronised_approval
legal_basis: Art. 14(4) · Art. 26(2)
evidence: event e_2231 · span 4f2c19aa · severity
high

↓ prev = 9c41...07e7 ✓

record #0008

sha 3fe1...c2aa

Why record #0007 holds up

01 DETERMINISTIC

Re-run the pipeline on the same input and every byte of every record matches. Nothing to argue with.

02 ARTICLE-LEVEL

No finding exists without a legal basis. Each violation cites its articles (12, 14, 26, 72) through a maintained crosswalk.

03 TAMPER-EVIDENT

Records are hash-chained. Change anything after the fact and the chain breaks visibly.

EVIDENCE RECORD		#0007
constraint	T1.synchronised_approval	
legal_basis	Art. 14(4) · Art. 26(2)	
evidence	event e_2231 · span 4f2c19aa	
severity	high	
record_sha256	9c41...07e7	
prev_record	51d2...9abc	

Audit day, with and without an evidence layer

WITHOUT

weeks

per audit question

- Log archaeology across services, prompts and transcripts
- A best-guess narrative, stitched together by hand
- No chain of custody: the story changes with who searched

WITH TRACEPROOF_

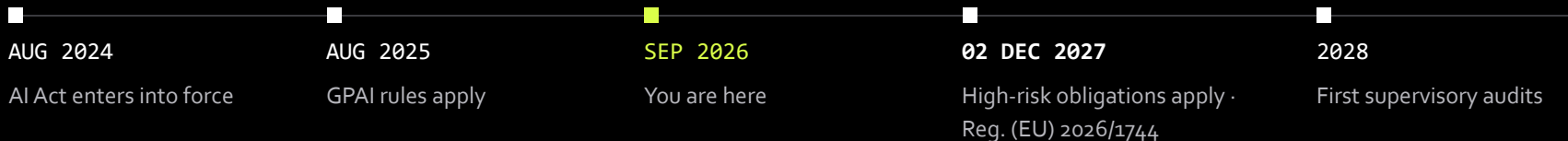
minutes

to a signed, reproducible record

- One query against the evidence store
- Hash chain verifies integrity in one step
- Same answer, no matter who asks, or when

■ The difference isn't speed. It's that the answer is **provable**.

High-risk obligations apply from 2 December 2027



15

months to stand up
the evidence layer

AGENTS SHIPPED

Credit workflows run on agent fleets in production, not in labs.

THE DUTY IS CONCRETE

Decisions must be reconstructable (Art. 12), logs kept ≥ 6 months (Art. 26(6)), feeding post-market monitoring (Art. 72).

STANDARDS MATURED

OTel GenAI telemetry and OCEL 2.0 process mining make the evidence layer buildable now.

A €300M+/yr problem in EU finance alone

BOTTOM-UP • EU FINANCE

4,834

credit institutions

ECB • 2024

+ 2,124

(re)insurers, Solvency II

EIOPA • 2026

× 50%

run ≥1 high-risk AI system by 2028:

credit scoring counts

conservative • ours

× 3 × €29k

systems × compliance run-cost / yr

EC-derived

= €300M+ / yr

recurring • + €0.7–1.2B one-off
build-out

THE LADDER ABOVE IT

€300M+/yr

EU finance compliance
run-rate • 2028

€1.9B/yr

65,000+ EU high-risk
systems • 2030 • every
Annex III sector

\$5.8B

global AI governance
software by 2029 • 45%
CAGR from \$0.9B in 2024

■ Per client, the number that matters is one avoided breach: €15M or 3% of turnover .

Sources: ECB/EBF Banking Facts & Figures 2025 • EIOPA Solvency II statistics 2026 • EC impact-assessment cost estimates • MarketsandMarkets, AI Governance Market 2024–2029

Everyone else answers a different question

WHAT IT PRODUCES: OPINIONS → EVIDENCE

RUNTIME GUARDRAILS

NeMo Guardrails · AI gateways

Block or allow single calls: hard facts in the moment, blind to the process around them.

LLM OBSERVABILITY

LangSmith · Langfuse · Arize

Traces, dashboards, LLM-as-judge scores: a different verdict on every run.

TRACEPROOF_

alone in this quadrant

Deterministic reconstruction of whole decisions: article-tagged, hash-chained, byte-reproducible.

GOVERNANCE SUITES

Credo AI · watsonx.governance · Holistic AI

Policy packs and program documentation. Attests the process exists, but can't reconstruct a decision.

WHERE IT LOOKS: SINGLE CALL → END-TO-END DECISION

■ Not rip-and-replace: TraceProof reads the OpenTelemetry these stacks already emit.

Adoption follows the compliance calendar

01

NOW → 2027

Readiness pilot

One credit workflow instrumented end to end; a gap report against Articles 12 and 26 your auditors can hold.

Fixed-fee engagement

02

BY 02 DEC 2027

Evidence rollout

The pipeline in production across every high-risk workflow before the obligations apply.

Per-system licence

03

2028 →

Standing evidence

Continuous checks, retention and audit-day support as the agent fleet evolves.

Annual service, run with x1F

The buyer is compliance & risk. The budget line already exists: monitoring, audit and reporting make up roughly 40% of per-system compliance spend.

The pipeline runs end-to-end today

WORKING NOW

- Telemetry → OCEL 2.0 mapper: agent spans become object-centric logs, governance attributes intact
- Constraint engine: five agent-specific rule templates, every rule article-tagged
- Evidence renderer: hash-chained records, byte-identical across re-runs
- Simulated credit-agent fleet generating controlled test telemetry

ON THE RECORD

- Peer review: "The Auditor's Trace" on track for ICPM 2027 · abstract 4 Sep, full paper 11 Sep 2026
- Ground truth and rules frozen and externally timestamped before any detection run
- One command reproduces every number and figure from raw telemetry

NEXT Expert study with FSI auditors · public OCEL 2.0 credit-agent dataset

■ NEXT STEP

Pilot with us_

One credit workflow. Four weeks. Your first article-level evidence records.

THE ASK · TWO PILOT PARTNERS FOR Q4 2026

YOU RUN

≥ 1 agentic credit or pricing
workflow

YOU GIVE

read access to agent telemetry
(OTel)

YOU KEEP

the gap report + evidence records,
either way

Aziz Ketata

x1F Group · x1f.de

Research artefact: "The Auditor's Trace" · ICPM 2027 submission