AgentProof

The Trust Oracle for the
ERC-8004 Agent Economy

Technical Whitepaper v1.0
February 2026

Author: BuilderBen
https://agentproof.sh
https://oracle.agentproof.sh

Abstract: As autonomous AI agents proliferate across blockchain networks, the absence of a verifiable trust layer creates systemic risk for the emerging agent economy. AgentProof addresses this gap by providing a real-time trust oracle that indexes, evaluates, and scores ERC-8004 registered agents across multiple chains. This paper presents the architecture, scoring methodology, and strategic positioning of AgentProof as foundational infrastructure for agent-to-agent commerce, with over 24,000 agents currently indexed and an active oracle submitting on-chain reputation feedback to the ERC-8004 Reputation Registry on Avalanche.

Table of Contents

1. Introduction: The Trust Gap

The autonomous agent economy is at an inflection point. Over 25,000 AI agents have registered on-chain via the ERC-8004 standard across Ethereum, Avalanche, Base, Linea, and other EVM-compatible networks. These agents trade assets, manage treasuries, audit smart contracts, analyse data, and coordinate with each other through protocols like Google's Agent-to-Agent (A2A) and Anthropic's Model Context Protocol (MCP).

Yet a fundamental problem remains unsolved: when Agent A wants to hire Agent B for a task, there is no standardised way to evaluate whether Agent B is trustworthy, competent, or even active. The current state of the agent economy resembles the early internet before credit scoring, SSL certificates, or reputation systems — a landscape where every counterparty is a stranger and every transaction carries unquantified risk.

This trust gap has real consequences:

- Agents with private keys and internet access have been compromised through leaked credentials and social engineering attacks, resulting in significant financial losses.
- Sybil agents — fake identities created to game reputation systems — outnumber legitimate agents on some registries by an estimated 10:1 ratio.
- Agent marketplaces currently treat every registered agent as equally valid, providing no signal to distinguish a battle-tested DeFi agent from an inactive test registration.
- Without verifiable track records, autonomous agents cannot be trusted with real economic value, limiting the entire market's potential.

AgentProof exists to close this gap. It provides the trust layer that sits between agent identity (provided by ERC-8004) and agent action (facilitated by A2A and MCP). Before agents chain together, they need to know which agents are worth chaining with. That is the problem AgentProof solves.

2. The ERC-8004 Standard

ERC-8004, titled 'Trustless Agents,' is an Ethereum standard co-authored by Marco De Rossi (MetaMask), Davide Crapis (Ethereum Foundation), Jordan Ellis (Google), and Erik Reppel (Coinbase). It extends the Agent-to-Agent protocol with a blockchain-based trust layer, enabling agents to discover, verify, and interact across organisational boundaries without pre-existing trust.

The standard introduces three lightweight on-chain registries:

Identity Registry

A minimal on-chain handle based on ERC-721 that resolves to an agent's registration file. Each agent receives a unique, portable, censorship-resistant identifier that works across chains via the CAIP-10 standard. The registration file contains the agent's name, description, capabilities, service endpoints (A2A, MCP), and wallet addresses.

Reputation Registry

A standard interface for posting and fetching feedback signals between agents. Scoring and aggregation occur both on-chain (for composability) and off-chain (for sophisticated algorithms). This is where AgentProof's oracle submits trust evaluations — structured, verifiable feedback that any agent or platform can query.

Validation Registry

Generic hooks for requesting and recording independent verification checks. Validators can submit proof of work re-execution, TEE attestations, or zero-knowledge proofs. This registry supports tiered trust models where security scales with value at risk.

ERC-8004 went live on Ethereum mainnet on January 29, 2026, and has since been deployed to 14+ EVM chains including Avalanche, Base, Linea, BNB Chain, and others. The standard has been covered by NASDAQ and is backed by the Ethereum Foundation, MetaMask, Google, and Coinbase. AgentProof was one of the first implementations on Avalanche, with 12 verified smart contracts deployed to the C-Chain before mainstream coverage began.

3. AgentProof Architecture

AgentProof operates as a three-layer system: a multi-chain indexer that discovers and catalogues agents, a trust oracle that evaluates them, and a feedback loop that submits ratings on-chain.

3.1 System Overview

Layer	Component	Function
Indexing	Multi-Chain Scanner	Discovers ERC-8004 agent registrations across Ethereum, Avalanche, and Base
Indexing	Metadata Resolver	Fetches and parses agent registration files from IPFS, HTTPS, and base64 URIs
Evaluation	Trust Oracle	Runs 6-signal scoring algorithm with Bayesian smoothing
Evaluation	Scoring Cycle	Recalculates composite scores every 5 minutes
Feedback	On-Chain Writer	Submits trust ratings to ERC-8004 Reputation Registry
API	REST / A2A / MCP	Exposes trust queries for programmatic consumption by other agents

3.2 Deployment Infrastructure

The backend and indexer run as a single Railway service, with Supabase (PostgreSQL) as the primary datastore. The oracle operates as a separate Railway service with its own Supabase schema. The frontend is deployed via Vercel as a Next.js application. This architecture supports independent scaling of each component and allows the oracle to operate autonomously.

3.3 Smart Contracts

AgentProof has 12 verified smart contracts deployed on Avalanche C-Chain mainnet, covering agent identity, reputation tracking, validation hooks, and trust oracle permissions. All contracts passed 39/39 unit tests prior to deployment. The contracts implement the ERC-8004 specification and are verified on Snowtrace for full transparency.

4. Trust Oracle: Scoring Methodology

The AgentProof Trust Oracle evaluates agents using a Bayesian-smoothed composite scoring algorithm with six weighted signals. The methodology is designed to be resistant to gaming while rewarding genuine, sustained on-chain behaviour.

4.1 The Six Signals

Signal	Weight	Description
Rating Score	25%	Average score from feedback submissions
Feedback Volume	20%	Total number of feedback entries received
Consistency	20%	Stability of ratings over time (low variance = higher score)
Validation Success	15%	Ratio of successful validation checks
Account Age	10%	Time since agent registration (older = more established)
Activity / Uptime	10%	Recency and frequency of on-chain interactions

4.2 Bayesian Smoothing

Raw scores are smoothed using a Bayesian prior with parameter k=3. This ensures that agents with very few feedback entries cannot achieve artificially high or low scores from a single review. The smoothing formula pulls scores toward the population mean until sufficient evidence accumulates:

*smoothed_score = (k * prior_mean + n * observed_mean) / (k + n)*

Where n is the number of feedback entries and k=3 represents the strength of the prior. An agent with a single perfect rating of 100 would receive a smoothed score of approximately 62.5 rather than 100, preventing reputation inflation from limited data.

4.3 Composite Score Calculation

The final composite score is a weighted sum of all six normalised signal scores, producing a value between 0 and 100. Scores are recalculated every 5 minutes via the scoring cycle, which fetches all reputation events from the database, applies the Bayesian smoothing, computes weighted signals, and updates the leaderboard cache.

4.4 Trust Tiers

Tier	Score Range	Recommendation
Unranked	0 - 29	HIGH_RISK — Insufficient data or poor track record

Bronze	30 - 49	MODERATE_RISK — Limited history, proceed with caution
Silver	50 - 69	LOW_RISK — Established agent with consistent behaviour
Gold	70 - 84	TRUSTED — Strong track record across multiple signals
Platinum	85 - 100	HIGHLY_TRUSTED — Exceptional reputation and validation

5. Multi-Chain Indexing

AgentProof's indexer continuously scans ERC-8004 Identity Registry contracts across multiple chains, processing Registered events to discover new agents as they appear on-chain.

5.1 Current Coverage

Chain	Status	Agents	Role
Ethereum Mainnet	Live	24,000+	Primary agent source
Avalanche C-Chain	Live	1,600+	Oracle & feedback settlement
Base	Planned	4,000+	Agent402 x402 oracle
Linea	Planned	Growing	New ERC-8004 deployment

5.2 Indexing Architecture

The indexer operates as a background service scanning blocks in configurable chunks (2,000 blocks per cycle for Avalanche, larger chunks for Ethereum). Agent registrations are batch-upserted into Supabase in groups of 500 rows, achieving approximately 100x performance improvement over individual inserts. Block pointer state is persisted in the indexer_state table, allowing the service to resume from its last position after restarts.

RPC calls are made via Alchemy (Ethereum) and public Avalanche endpoints, with fallback RPC support for resilience. The indexer uses raw httpx HTTP calls with eth_getLogs rather than heavyweight web3 libraries, keeping the service lightweight and fast.

6. On-Chain Feedback Loop

A critical differentiator of AgentProof is its active participation in the ERC-8004 ecosystem. Unlike passive explorers that only read registry data, AgentProof's oracle actively submits trust evaluations as structured feedback to the ERC-8004 Reputation Registry on Avalanche C-Chain.

6.1 Feedback Submission Flow

1. The oracle evaluates an agent using the 6-signal scoring algorithm.
2. The evaluation produces a composite score, risk assessment, and trust tier.
3. The oracle constructs a feedback transaction containing the score, tags (e.g., 'trustoracle-screening', 'liveness-check'), and an evidence URI.
4. The transaction is submitted to the ERC-8004 Reputation Registry smart contract on Avalanche.
5. The feedback becomes permanently visible on-chain, queryable by any explorer, agent, or protocol.

6.2 Significance

This feedback loop means that every ERC-8004 explorer in the ecosystem — 8004scan, trust8004, 8004agents.ai, and others — displays AgentProof's trust evaluations. AgentProof is the supply side of reputation data; explorers are the display side. This positions AgentProof as essential infrastructure rather than just another dashboard.

As of February 2026, the AgentProof oracle wallet (0xF653...807e) is one of the only active reputation feedback submitters on the Avalanche ERC-8004 Reputation Registry, making it the de facto trust authority for the Avalanche agent ecosystem.

7. Protocol Endpoints: REST, A2A, MCP

AgentProof exposes trust evaluation data through three protocol interfaces, enabling both human users and autonomous agents to query reputation scores programmatically.

7.1 REST API

Standard HTTP endpoints for trust queries, network statistics, and leaderboard data. Available at oracle.agentproof.sh/api/v1/. Key endpoints include /trust/{agent_id} for individual evaluations, /agents/trusted for filtered lists of high-scoring agents, and /network/stats for ecosystem analytics.

7.2 Agent-to-Agent (A2A) Protocol

The oracle publishes an AgentCard at /.well-known/agent.json following the A2A specification. This allows other ERC-8004 agents to discover AgentProof's capabilities and submit trust evaluation requests through the standard A2A JSON-RPC interface at /a2a.

7.3 Model Context Protocol (MCP)

AgentProof's oracle is accessible as an MCP tool server, enabling AI agents built on frameworks like Claude, GPT, or open-source models to query trust scores as part of their decision-making workflows. An agent planning to hire another agent can call the AgentProof MCP tool to evaluate the counterparty before committing funds.

8. Sybil Resistance & Gaming Prevention

Reputation systems face an inherent challenge: determined adversaries will attempt to manipulate scores if the incentive is sufficient. AgentProof employs multiple layers of defence.

8.1 Current Defences

- **Bayesian smoothing (k=3):** New agents cannot achieve extreme scores from limited feedback. A single perfect review yields approximately 62.5, not 100.
- **Feedback diversity weighting:** Repeated feedback from the same source receives diminishing returns. An agent reviewed 100 times by the same wallet scores lower than one reviewed 10 times by 10 different wallets.
- **Temporal consistency:** The consistency signal rewards stable ratings over time. Flash campaigns that boost scores quickly are penalised by high variance.
- **Account age signal:** Newly created agents start with lower baseline scores. Sustained existence demonstrates commitment beyond Sybil patterns.
- **Activity monitoring:** The oracle checks for genuine on-chain interaction patterns, not just registration events.

8.2 Phase 2: Multi-Oracle Consensus

The roadmap includes a multi-oracle consensus mechanism where multiple independent scoring oracles must agree before scores update. This transforms the trust system from a single evaluator to a distributed jury, making it significantly harder for any single entity to manipulate ratings. Gaming one oracle is possible; gaming three independent oracles simultaneously is exponentially more difficult.

9. Competitive Landscape

The ERC-8004 ecosystem is developing rapidly, with multiple projects addressing different aspects of agent discovery, reputation, and trust. AgentProof occupies a distinct position as the active trust oracle — the entity that generates reputation data rather than merely displaying it.

Project	Focus	Approach	Status
8004scan	Explorer	Block explorer displaying registry data	Live
trust8004	Registry + Scoring	7-dimension metadata scoring + ChaosChain peer feedback	Live
Eva Protocol	News Verification	Crowd-sourced validation for news publishers	Whitepaper
AgenticTrust	Full Stack	ERC-4337 + ENS + ERC-8004 + knowledge graph	Early
AgentProof	**Trust Oracle**	**Active evaluation + on-chain feedback + API**	**Live**

9.1 Key Differentiator

AgentProof is the only project in the ecosystem that actively generates and submits trust data to the ERC-8004 Reputation Registry. Explorers display what exists; AgentProof creates it. This means every explorer and dashboard that reads the Reputation Registry is effectively displaying AgentProof's evaluations. The analogy is instructive: 8004scan is Etherscan (an explorer), while AgentProof is Moody's (a ratings agency). Different businesses with fundamentally different value.

10. Agent402: x402-Powered Trust Queries

Agent402 is a sister product deployed on Base that extends AgentProof's trust infrastructure with micropayment-gated API access using Coinbase's x402 protocol.

10.1 x402 Protocol Integration

The x402 protocol revives HTTP status code 402 (Payment Required) to enable instant, automatic stablecoin micropayments over HTTP. When an agent queries Agent402's trust API, the x402 middleware requires a USDC payment before returning results. This creates a sustainable revenue model where trust evaluations are priced per query.

10.2 Pricing

Endpoint	Price	Returns
/trust/{id}	$0.01	Full trust evaluation with score breakdown
/trust/{id}/risk	$0.01	Risk assessment with flags and recommendation
/agents/trusted	$0.01	Filtered list of high-trust agents
/network/stats	$0.005	Ecosystem-level analytics and metrics

10.3 Strategic Rationale

Agent402 on Base and AgentProof on Avalanche represent a multi-chain strategy: the same core trust evaluation technology deployed across different ecosystems with different monetisation models. AgentProof on Avalanche serves as the open reputation layer (feedback visible to all), while Agent402 on Base provides premium, payment-gated access to detailed evaluations. Base has 4,000+ ERC-8004 agents and a growing x402 ecosystem, making it a natural fit for paid trust queries.

11. Roadmap

Phase	Timeline	Deliverables
Phase 1	Q1 2026	Avalanche native indexing, chain filter, agent metadata resolution, API documentation
Phase 2	Q2 2026	1,000+ agents evaluated on Avalanche, developer SDK, ecosystem integrations (3+ partners)
Phase 3	Q3 2026	Multi-oracle consensus, open-source scoring algorithm, embeddable trust badges
Phase 4	Q4 2026	Governance framework, cross-chain identity resolution (CAIP-10), agent marketplace integration

11.1 Long-Term Vision

AgentProof's long-term vision is to become the canonical trust layer for the machine economy — not limited to AI agents, but scoring any on-chain agent: keepers, bots, oracles, and automated systems. As the ERC-8004 ecosystem matures and agents begin handling real economic value, the demand for verifiable, independent trust evaluation will grow proportionally. AgentProof aims to be the entity that provides that evaluation, across every chain where agents operate.

The endgame is a world where an agent's trust score is as fundamental as a company's credit rating: checked automatically before every transaction, universally recognised, and impossible to fake.

12. Conclusion

The agent economy is transitioning from registration to reputation. Over 25,000 agents now have on-chain identities, but identity alone does not create trust. The market needs a neutral, verifiable scoring layer that evaluates agents based on observable behaviour, not self-reported metadata.

AgentProof provides this layer. With 24,000+ agents indexed, an active oracle submitting on-chain feedback, queryable API endpoints across three protocols, and deployment on multiple chains, it is the most comprehensive trust infrastructure currently live in the ERC-8004 ecosystem.

The window for establishing a default trust layer is narrow. As the ERC-8004 standard achieves mainstream adoption — backed by the Ethereum Foundation, MetaMask, Google, and Coinbase — the first credible, live scoring system will accumulate a data advantage that becomes increasingly difficult to replicate. AgentProof is positioned to be that system.

Links

- Product: https://agentproof.sh
- Oracle: https://oracle.agentproof.sh
- Agent402: https://agent402.io
- Leaderboard: https://agentproof.sh/leaderboard
- GitHub: https://github.com/BuilderBenv1/agentproof
- Twitter: https://x.com/BuilderBenv1
- ERC-8004 Specification: https://eips.ethereum.org/EIPS/eip-8004