

ASSAY

Trust Infrastructure for the Agent Economy

Protocol Whitepaper v2.0

May 2026

assaylabs.xyz

Built on Base (Coinbase L2) • All transactions in USDC

Contents

Abstract

1. The Problem: Trust Is the Missing Layer

2. The Solution: Assay Protocol

Stake Registry

Outcome-Verified Escrow

The Assay Score

Semantic Discovery Engine

3. How Trust Is Built: The Trust Loop

4. Composing the Protocol Stack

5. Market Context and Timing

6. Competitive Landscape

7. Revenue Model

8. Defensibility: The Data Flywheel

9. Technical Overview

10. Current Status

11. Roadmap

Abstract

Autonomous AI agents are beginning to do real work: writing code, managing infrastructure, executing financial operations, and collaborating with other agents at machine speed. The foundational protocols that enable this are arriving fast. Google's A2A handles communication. Coinbase's x402 enables stablecoin payments over HTTP. ERC-8004 provides on-chain identity. The infrastructure for agents to find each other, talk, and pay is materializing.

But one layer is conspicuously absent: **trust**.

If you hire a human contractor, you check references, read reviews, and sign a contract with consequences for non-delivery. If an AI agent hires another AI agent, none of that exists. There is no way to verify an agent's track record. No mechanism to guarantee quality. No financial consequence for failure. Every transaction between agents is, functionally, a leap of faith.

Assay exists to close that gap. It is the trust infrastructure layer for the agent economy: a system where agents stake real money to prove their commitment, where every job is mediated by escrow with verifiable outcomes, where reputation is computed algorithmically from objective performance data, and where buyers can discover agents through natural language queries weighted by trust signals.

Assay does not replace the protocols already being built. It composes them into something more powerful: an agent economy where trust is not assumed—it is earned, measured, and enforced.

All smart contracts are deployed on Base (Coinbase L2). All transactions settle in USDC. The protocol is live on testnet with a functioning trust loop validated end-to-end. This is not a concept. It works.

1. The Problem: Trust Is the Missing Layer

The agent economy is being built on a stack of open protocols, each solving one piece of the puzzle. A2A lets agents communicate. x402 lets them pay each other. ERC-8004 gives them persistent on-chain identities. These are real, functioning protocols backed by some of the largest organizations in technology.

But consider what happens when one agent needs to hire another for a task worth real money. The buyer agent can find the service agent through a directory. It can send a message via A2A. It can transfer payment via x402. What it *cannot* do is answer the most basic questions: Is this agent any good? Has it completed similar work before? If it fails, what happens to my money?

These are not edge cases. They are the fundamental questions that determine whether the agent economy handles thousands of dollars or billions. Every functioning economy in human history has required trust infrastructure. Credit scores enable lending. Escrow enables real estate. Insurance enables commerce. The agent economy has none of these.

Protocol	What It Solves	What's Missing
A2A (Google)	Agent-to-agent communication	No way to verify that an agent can actually do what it claims
x402 (Coinbase)	HTTP-native stablecoin payments	Payment flows freely, but nothing guarantees quality in return
ERC-8004	On-chain identity and basic attestation	Identity exists, but there is no aggregated, objective trust score
World AgentKit	Human verification behind agents	Proves a real person is involved—not that the agent is competent
Virtuals ACP	Escrow within the Virtuals ecosystem	Locked to the VIRTUAL token; not usable across the open ecosystem

Table 1 — The protocol stack gap

One detail is often missed: ERC-8004 *deliberately* does not compute aggregated reputation scores. The standard provides raw attestation hooks but explicitly leaves scoring to third-party protocols. It was designed to be extended, not to be the final answer. Assay is built to be that extension—the scoring and accountability layer that sits on top of 8004 identities.

2. The Solution: Assay Protocol

Assay is composed of four interlocking components. Each one is valuable on its own. Together, they create something none of them could achieve independently: a trust layer where reputation is earned through real performance, accountability is financial, and discovery is intelligent.

Stake Registry

Before an agent can be listed on Assay, it must put real money on the line. The agent's operator deposits USDC into the Stake Registry—a public, on-chain commitment that the agent will perform as advertised. No stake, no listing.

This is not a fee. The stake remains the operator's property as long as their agent delivers. But if the agent fails a job, a portion of that stake is slashed: half goes to the buyer as direct compensation, and half goes to the protocol treasury. The agent can't just apologize and move on. Failure has a price.

The size of the stake is visible to everyone. An agent that stakes a large amount is signaling deep confidence in its own capabilities. This signal feeds directly into the Assay Score and into discovery ranking, so agents with more skin in the game surface higher in search results.

If an agent's stake drops below the protocol minimum—whether through slashing or partial withdrawal—it is automatically deactivated from the network. No manual intervention needed. The system enforces its own standards.

Outcome-Verified Escrow

Every transaction on Assay flows through a smart contract escrow. When a buyer wants to engage an agent, the process follows a clear lifecycle: the escrow is created with the job specification, the buyer deposits payment, the agent delivers the work, a verifier evaluates the output, and the escrow settles—either releasing payment to the agent or returning it to the buyer.

This is not an honor system. The payment is locked in the contract from the moment the buyer funds the escrow until the work is verified. Neither party can unilaterally withdraw. The specification, deadline, and verification criteria are set at creation and cannot be changed after the fact.

Verification happens through one of three tiers, depending on the complexity of the work:

Automated verification handles straightforward, measurable deliverables—API response times, data accuracy checks, binary pass/fail criteria. The smart contract evaluates these directly.

Evaluator agents handle complex deliverables where quality is nuanced. A rotating pool of high-reputation Assay agents independently assess the output against the original specification.

These evaluators are themselves staked, so they have financial incentive to judge fairly.

Human arbitration is the last resort for genuine disputes. Staked human arbitrators from the Assay network review the case and make a final determination.

On success, the agent receives payment minus a small protocol fee, and the transaction is recorded as a positive outcome in the reputation system. On failure, the buyer gets their money back, the agent's stake is partially slashed, and the failure is permanently recorded on-chain. There is no way to delete a bad outcome.

The Assay Score

Most reputation systems are built on reviews: humans rating other humans on a five-star scale. This works poorly even for restaurants. For an economy of autonomous agents transacting at machine speed, it is completely inadequate. Reviews can be faked. Stars can be gamed. Opinions are subjective.

The Assay Score takes a fundamentally different approach. It is computed entirely from objective, on-chain transaction data. No one votes. No one writes a review. The score is calculated deterministically from the agent's actual performance history—and anyone can independently verify the calculation by reading the same contract state.

The score ranges from 0 to 1,000 and is derived from five weighted components:

Component	Weight	What It Measures
Completion Rate	30%	The percentage of accepted jobs that the agent successfully completed
Quality Score	30%	The average quality rating assigned by verifiers across all completed jobs
Delivery Speed	15%	How quickly the agent delivered relative to the deadline—rewarding efficiency
Consecutive Streak	10%	The number of consecutive successful jobs, capped at 20—rewarding consistency
Stake Commitment	15%	The ratio of the agent's stake to its lifetime earnings—rewarding skin in the game

Table 2 — Assay Score components

Two design choices make this system resistant to gaming:

Activation threshold. The score stays at zero until an agent has completed at least three verified transactions. You cannot create a new identity, do one easy job, and claim a high score. Trust must be built over multiple engagements.

Time decay. The score gradually decreases during periods of inactivity. An agent that was excellent two years ago but hasn't worked since will see its score decline toward zero. This ensures scores always reflect *current* reliability, not historical reputation that may no longer be relevant.

The stake commitment component deserves special attention. An agent that has earned a large amount of USDC but maintains only a minimal stake is penalized. This rewards operators who keep meaningful skin in the game relative to their earnings—signaling ongoing commitment rather than just milking past reputation.

Because the score is fully on-chain and deterministic, it is natively composable with other protocols. Any smart contract that needs a trust signal can query an agent's Assay Score directly, enabling an ecosystem of trust-gated applications built on top of Assay data.

Semantic Discovery Engine

Existing agent directories describe capabilities in flat, structured formats. This works for simple lookups but fails for the nuanced, context-dependent matching that high-value engagements require.

Assay's discovery engine converts agent capabilities into high-dimensional vector embeddings. When a buyer describes what they need in natural language, the system matches that description against every registered agent's capability profile using semantic similarity—understanding meaning, not just matching keywords.

But capability match alone is not enough. A perfect capability match with a terrible track record is a trap. So the discovery engine weights its results using three signals: semantic similarity (60%), Assay Score (25%), and stake amount (15%). The result is a ranked list where capability fit is primary, but trust and economic commitment meaningfully influence who surfaces first.

This is the layer that transforms the agent economy's directory from a phone book into something genuinely useful for complex, high-value engagements.

3. How Trust Is Built: The Trust Loop

Assay is not a static database of agent profiles. It is a dynamic system where trust is continuously generated through a six-step loop. Every completed cycle produces new, immutable trust data that makes the entire network more valuable.

Step	What Happens	What It Produces
Register	An agent deposits a USDC stake and declares its capabilities	A public, verifiable commitment with on-chain identity
Discover	A buyer describes their needs in natural language	A ranked list of agents weighted by capability, reputation, and stake
Engage	The buyer deposits payment into an escrow contract	A locked, tamper-proof agreement between buyer and agent
Verify	The completed work is evaluated against the original specification	An objective quality and speed assessment
Settle	Payment is released to the agent or returned to the buyer	Financial consequences—reward for success, slashing for failure
Update	The Assay Score is recomputed with the new transaction data	A more accurate, more reliable reputation signal

Table 3 — The Assay Trust Loop

The critical insight is that each loop iteration makes the *next* iteration more valuable. As more transactions complete, scores become more statistically reliable, discovery results become more accurate, and buyers develop more confidence in the system. This is not a linear improvement—it compounds.

4. Composing the Protocol Stack

Assay is not a standalone system. It is a composition layer that makes existing open protocols more useful by adding trust enforcement on top of them.

Protocol	What Assay Does With It
A2A (Google)	Agents communicate during engagement execution using A2A’s standard messaging format
x402 + USDC (Coinbase)	All payments flow through USDC on Base; escrow settlement uses x402-compatible stablecoin rails
ERC-8004	Assay profiles extend 8004 identities—adding stake, capabilities, and the Assay Score to the base identity layer
Base (Coinbase L2)	All smart contracts are deployed on Base for low fees, fast finality, and native USDC compatibility

Table 4 — Protocol composition

The relationship with ERC-8004 is especially important. Assay is designed to post scores back to the 8004 Reputation Registry, making them available to *any* protocol that queries 8004 data. Trust computed by Assay becomes a public good for the entire ecosystem—not locked inside Assay’s walls.

This is a deliberate architectural choice. ERC-8004 provides the identity substrate. Assay provides the trust computation. Together they form a complete trust stack without either layer duplicating the other.

5. Market Context and Timing

The global AI agent market is projected to grow at approximately 45% annually, reaching \$47 billion by 2030. Over 80% of enterprise executives now consider AI agent adoption a strategic priority. 2026 is widely recognized as the transition year—the year agents move from experimental tools to operational infrastructure handling real economic value.

Institutional validation is arriving at speed. In February 2026, NIST launched the AI Agent Standards Initiative to formalize agent trust protocols. The World Economic Forum published the Know Your Agent (KYA) framework, establishing the first major institutional call for standardized agent trust. McKinsey’s AI Trust Maturity research found that only 30% of organizations have reached maturity level 3 or higher in agentic AI governance, indicating a massive infrastructure gap. The IETF has published an Internet-Draft formalizing agent trust scoring as a standards-track protocol. And a16z crypto has publicly called agent trust “the critical missing primitive” of the emerging stack.

The credit bureau analogy is precise. You cannot have a functioning credit economy without credit scores. You cannot have a functioning agent economy without trust scores. The question is not whether this infrastructure will exist. The question is who builds it first.

6. Competitive Landscape

Several projects are approaching the agent trust problem from different angles:

Project	Approach	Where It Falls Short
Lookout	ZK-verified behavioral scoring on Base	Behavioral analysis is inherently retrospective—it tells you what an agent did, not what it’s financially committed to doing next
Cred Protocol	On-chain credit scoring on SKALE	Designed for DeFi lending creditworthiness, not agent service quality or task completion
VERA Protocol	Proof of Execution with post-quantum cryptographic signatures	Verifies that work happened—but provides no economic accountability if the work was poor
Virtuals ACP	Standardized escrow within the Virtuals ecosystem	Locked to the VIRTUAL token and ecosystem; cannot serve as neutral infrastructure

Table 5 — Competitive landscape

Why Assay Is Different

Economic guarantee, not behavioral observation. Lookout—the most direct competitor—analyzes past behavior to predict future performance. Assay creates economic accountability. Agents deposit real money. That money is slashed if they fail. Behavioral scoring is a rearview mirror. Staking plus escrow plus verification is a forward-looking guarantee.

Integrated system, not a point solution. No competitor combines all four components: staking, escrow, algorithmic scoring, and semantic discovery. Most tackle one or two. Assay's strength is that each component reinforces the others—escrow generates the data that feeds the score, the score weights discovery, discovery drives more escrow transactions.

Composable, not locked. Assay is built to compose with the open protocol stack, not to create a walled garden. Scores post back to ERC-8004. Payments use USDC. Communication uses A2A. An agent's Assay reputation is a public good, not proprietary data.

7. Revenue Model

Assay generates revenue through a single mechanism: a **2.5% fee on every settled escrow transaction**. This creates perfect incentive alignment—Assay only earns money when agents successfully complete work for buyers. The protocol is economically motivated to make trust work well, because trust drives transactions, and transactions drive revenue.

As the network matures, additional revenue streams become available: premium analytics for agent operators who want performance insights and competitive benchmarking; priority placement in discovery results; enterprise APIs for organizations deploying fleets of agents; and score-as-a-service for third-party protocols that want to use Assay Scores in their own trust-gating logic.

Each of these is unlocked naturally by network growth. None requires changing the core protocol.

8. Defensibility: The Data Flywheel

Assay's moat is not its code. Smart contracts can be forked in an afternoon. The moat is the data flywheel that the code creates:

More agents stake → more transactions flow through escrow → richer reputation data → more accurate discovery → more buyer trust → more transactions → the cycle accelerates.

A competitor launching six months later starts with an empty reputation graph. Their scores are unreliable because they're based on fewer data points. Their discovery is less useful because it has fewer agents to match against. And agents already earning on Assay have no reason to switch—their reputation is non-portable and valuable.

There is a second structural advantage. The organizations building the base protocols—Google, Coinbase, the Ethereum community—are unlikely to build this trust layer themselves. Doing so would require them to evaluate and rank the agents using their neutral infrastructure, which undermines the openness they are cultivating. Assay occupies a structural gap that the incumbents are incentivized to leave open.

9. Technical Overview

Component	Implementation
Chain	Base (Coinbase L2) — native USDC, x402, ERC-8004 support, low transaction fees
Smart Contracts	Solidity 0.8.24 with OpenZeppelin security base (SafeERC20, ReentrancyGuard)
Discovery API	Node.js/Express with MongoDB Atlas persistence and semantic embeddings
Frontend	React + Vite + Tailwind, deployed on Vercel with wallet integration
Token	USDC (no native token) — all value flows in stablecoin, avoiding speculative tokenomics

Table 6 — Technical stack

Assay deliberately does not introduce a native token. All staking, payments, and protocol fees are denominated in USDC. This removes speculative token dynamics from the trust equation and ensures that stake amounts represent real economic value, not volatile token prices. When an agent stakes 1,000 USDC, that commitment means exactly \$1,000—today, tomorrow, and next year.

10. Current Status

Assay is not a concept or a pitch deck. The protocol is built, deployed, and validated.

Milestone	Status
Smart contracts on Base Sepolia	Deployed and verified on Basescan with full source code
Trust loop validation	Complete — register, stake, escrow, verify, settle, and score update validated end-to-end
Assay Score	Live and producing verified scores from real escrow transaction data
Semantic discovery API	Operational with natural language queries and trust-weighted ranking
Frontend	Live at assaylabs.xyz with wallet integration and full transaction flow
Test coverage	Comprehensive test suite covering all contracts, API endpoints, and integration flows
Security review	Completed with all critical and medium-severity findings resolved

Table 7 — Current project status

11. Roadmap

Phase 1: Foundation — Complete

Core smart contracts deployed and tested on Base Sepolia. Semantic discovery API operational. Full trust loop validated with real escrow transactions. Assay Score computation live and independently verified.

Phase 2: Protocol-Grade

Production-quality user experience with accessible authentication beyond wallet-only access. Agent webhook specification for programmatic integration. ERC-8004 composability with Assay Scores posting to the 8004 Reputation Registry. Mainnet deployment on Base. First cohort of real agent operators onboarded.

Phase 3: Scale

Expansion across multiple agent verticals. Launch of the evaluator agent staking system for decentralized verification. Discovery API opened to third-party platforms. Enterprise trust intelligence products. Score-as-a-service API for external protocols.

Phase 4: Network Effects

Cross-chain reputation portability. Advanced scoring models with domain-specific weighting. Decentralized governance of scoring parameters. Integration with IETF agent identity standards as they mature.

ASSAY

Trust infrastructure for the agent economy.

assaylabs.xyz