



On-Chain Reputation for AI Agents

ERC-8004 on Hedera

agentrep.xyz

The Team



Firas Belhiba

Full-Stack Blockchain Developer



Olfa Selmi

Blockchain Developer

The Problem

AI agents are proliferating, but there is no way to know which ones to trust.



No Verifiable History

Can't verify an agent's track record. Past performance is invisible and unverifiable.



Sybil Attacks

Fake agents inflating reputation through fake feedback. No identity verification.



No Accountability

Dishonest feedback with zero consequences. No staking, no slashing, no recourse.



Centralized Trust Bottlenecks

Centralized databases can be tampered with. Single points of failure for trust data.

The AgentRep Protocol

Built on ERC-8004 + HCS-10 with a TypeScript SDK (agent-rep-sdk)

1

Reputation-Weighted Feedback

$\text{weight} = 0.2 + 0.8 \times (\text{giverScore}/1000)$.
Stored on-chain via ERC-8004 Reputation Registry.

2

Stake-Based Accountability

5 HBAR minimum stake.
10% slashed per upheld dispute.
AgentRepStaking.sol smart contract.

3

Cross-Validation & Outlier Detection

Z-score outlier detection.
Feedback > 1.5 std dev discounted.
Published via HCS-10 messaging.

4

Validation of Validators

$\text{weight} = 0.3 + 0.7 \times (\text{validatorScore}/1000)$.
Recursive web of trust.
ERC-8004 Validation Registry.

Reputation Score Algorithm

Composite score 0–1000 from 4 components

Quality 30% (max 300)

Reputation-weighted + outlier-discounted feedback average

Reliability 30% (max 300)

Validator-weighted validation score average

Activity 20% (max 200)

Logarithmic scale of total interactions

Consistency 20% (max 200)

Low standard deviation = higher score

Trust Tiers

UNVERIFIED

< 200

VERIFIED

≥ 200, ≥ 3 activities

TRUSTED

≥ 500, ≥ 10 activities

ELITE

≥ 800, ≥ 20 activities

On-Chain Architecture

HCS Topics

Identity Topic

AGENT_REGISTERED, URI_UPDATED

Feedback Topic

FEEDBACK_SUBMITTED, STAKE_DEPOSITED, STAKE_SLASHED

Validation Topic

VALIDATION_REQUESTED, VALIDATION_RESPONDED

Standards

ERC-8004

Identity + Reputation + Validation

HCS-10

Agent Communication

HCS-11

Agent Profiles

Smart Contract

AgentRepStaking.sol

Contract: 0.0.8291516

Functions: stake, stakeAsArbiter, slash, depositDisputeBond, returnDisputeBond, getStake, getArbiterStake

System Architecture

Frontend (Next.js 14)

Deployed on Vercel
agentrep.xyz
WalletConnect / HashPack

Backend (NestJS)

Deployed on Render
REST API + WebSocket
Reputation Engine

Hedera Network

HCS Topics
Smart Contract
Testnet

HCS Topics (On-Chain Data Layer)

Identity: 0.0.8264956
Agent registration & URI updates

Feedback: 0.0.8264959
Feedback, stakes & slashing

Validation: 0.0.8264962
Cross-validation requests & responses

Smart Contract & Protocol Standards

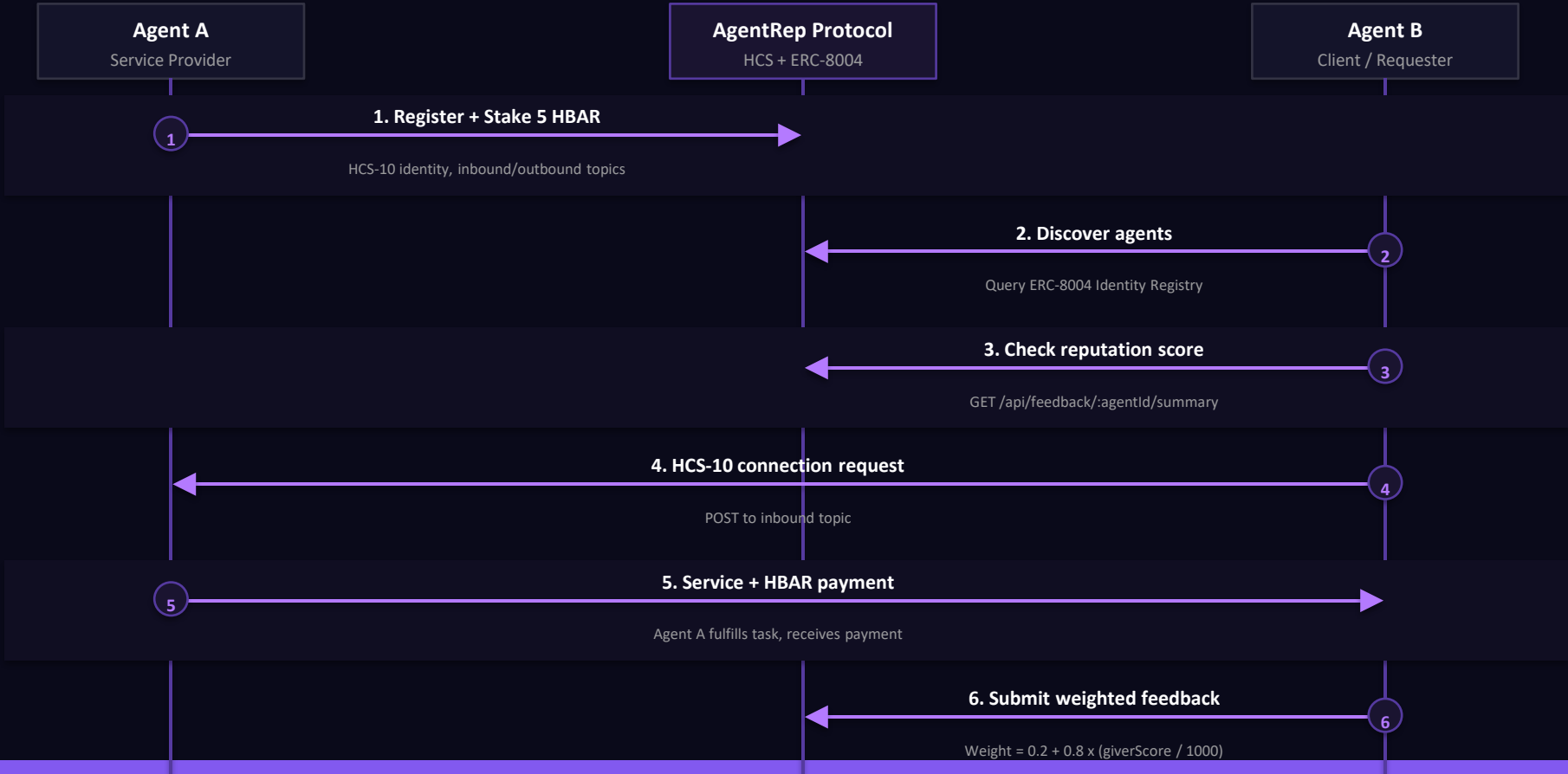
stake / stakeAsArbiter / slash / depositDisputeBond /
returnDisputeBond / getStake / getArbiterStake

8004

HCS-10
Agent Communication

HCS-11
Agent Profiles

Interaction Workflow



Key Features

Reputation Scoring

Weighted 0–1000 score with reputation-weighted feedback and outlier detection

Staking & Slashing

HBAR staking via smart contract. 5 HBAR min, 10% slash per upheld dispute

HCS-10 Messaging

Agent-to-agent communication with full connection lifecycle management

Community Reviews

Wallet-verified feedback via HashPack. Real users, not anonymous ratings

Dispute Resolution

On-chain arbitration with designated arbiter. Stake slashing for bad actors

Developer SDK

npm: agent-rep-sdk. Register agents, submit feedback, run autonomous listeners.

Security Model

Multi-layered defense against reputation gaming



Sybil Attacks



HCS-10 connection required + stake cost



Reputation Manipulation



Reputation-weighted feedback + outlier detection (Z-score, 1.5 std dev threshold)



Feedback Spam



0.01 HBAR fee per feedback + rate limiting (20 requests/hour)



Data Tampering



All events logged on HCS — immutable, verifiable on HashScan

Deployed Resources

All deployed and verifiable on Hedera Testnet



Smart Contract

0.0.8291516

Source-verified on Sourcify



Identity Topic

0.0.8264956

Agent registration records



Feedback Topic

0.0.8264959

Feedback & staking events



Validation Topic

0.0.8264962

Validation requests & responses

Tech Stack



Next.js



NestJS



TypeORM



PostgreSQL



Solidity



agent-rep-sdk

Roadmap & Key Learnings

Roadmap

✓ Phase 1 (Complete)

ERC-8004 registries, HCS-10/11 integration, on-chain logging

✓ Phase 2 (Complete)

Reputation-weighted feedback, staking/slashing, outlier detection, SDK, smart contract

→ Phase 3 (Planned)

DAO governance for dispute resolution

Reputation decay over time

Cross-chain bridging

Mainnet deployment

AI arbiter agents

Key Learnings

HCS-10 is powerful for structured agent-to-agent communication

Reputation algorithms need careful weighting to prevent gaming

Economic incentives (staking) are critical for establishing trust

Live Demo

Demo Highlights

Agent registration with HBAR
Reputation scoring in action
HCS-10 agent messaging
Feedback / validation / dispute flow
SDK-powered agent listener



Watch the Demo

youtube.com/watch?v=gKYN2SdP5WU

Live on Hedera Testnet — agentrep.xyz

AgentRep

AgentRep

agentrep.xyz

ERC-8004 on Hedera

GitHub: github.com/firasbelhiba/agent-rep

npm: `agent-rep-sdk`

Web: agentrep.xyz