Sahamati Labs
June 2026

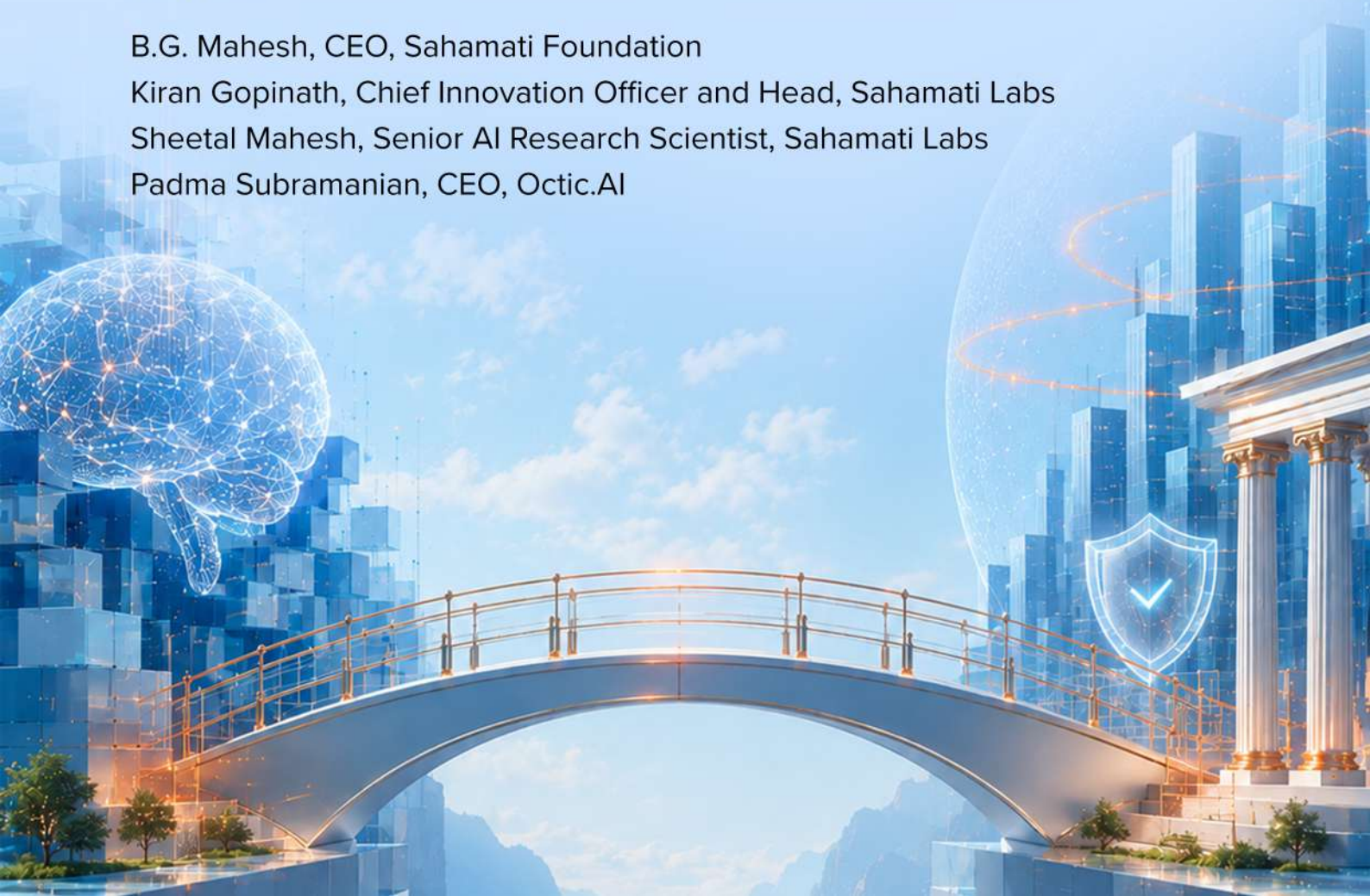
PRE-READ

A Framework for AI Agents in the Account Aggregator Ecosystem

Four Pillars of Consent, Identity, Provenance, and Conformance Built on the Existing AA Foundation

Authors:

B.G. Mahesh, CEO, Sahamati Foundation
Kiran Gopinath, Chief Innovation Officer and Head, Sahamati Labs
Sheetal Mahesh, Senior AI Research Scientist, Sahamati Labs
Padma Subramanian, CEO, Octic.AI



This is a brief introduction to the framework Sahamati Labs has proposed for governing AI agents that act on consumer financial data in India's Account Aggregator ecosystem. The full design paper is available separately.

What the framework addresses

India's Account Aggregator (AA) framework lets consumers share their financial data with regulated lenders, insurers, and other institutions under consent, with audit and recourse mechanisms that protect the consumer.

AI agents are now beginning to act on this consented data inside the receiving institutions, making lending decisions, fraud assessments, KYC verifications, and other determinations. The AA framework can verify that the data was shared under consent, but it cannot see what happens once the data is inside the receiving institution: which AI model was deployed, what tools the agent used, what was inferred from the data, or whether the agent operated within its authorised scope.

As AI agents become the dominant way consumer financial data is processed, the trust property that makes AA work (that everything done with consumer data is observable, attributable, and auditable) is at risk of eroding.

What the framework proposes

Four extensions to the existing AA framework, each addressing one part of the agentic governance question.

Consent: Consent that authorises specific processing, not just data sharing. Consumers grant permission for specific kinds of AI processing on their data, not just for the data itself to be shared. They can specify which kinds of agents may act, which derivatives may be retained, and what is out of bounds.

Identity: Registered agents with verifiable identity. Every AI agent that acts on AA data is registered with an industry body, with its capabilities and behaviour declared in advance. Unregistered agents cannot operate within the framework.

Provenance: Verifiable records of what every agent did. AI processing happens inside specially protected computing infrastructure operated by an entity independent of the institution whose work it serves. The infrastructure produces structured, signed records of

institution whose work it serves. The infrastructure produces structured, signed records of every action. Consumers see these records in plain language through their AA app; regulators can query them for audit.

Conformance: Pre-deployment testing against known risks. Agents are tested against a curated library of attack patterns and failure modes before being allowed to operate. The library is maintained by an industry body and updated as new risks emerge.

The framework rests on a shared infrastructure model with two roles: a Governance Operator and a Runtime Operator.

An industry body operates the governance layer (the registry, the audit trail, the safety-testing library) and sets the rules these encode.

A separately designated entity, under regulatory oversight, operates the protected runtime that holds and processes AA data in a Trusted Execution Environment (TEE), sometimes called a Confidential Clean Room (CCR) in India.

In order to operationalize the four extensions, the TEE would be set up, wherein every action will be registered. This allows the TEE to enforce every action and leave an auditable trail. So, every action will produce a receipt, be it a system receipt, a read receipt, or a refusal receipt to denote exactly what the data was used for and sent to the governance layer. This ensures that AI models do not extend their scope of use and the customers retain sufficient autonomy to decide when they want to opt out of a process.

Why this matters now

The Account Aggregator framework took years to establish consumer trust in financial data sharing. AI agents are entering this flow rapidly. Without governance designed for agentic processing, the framework's trust property degrades, and the alternative (agents acting on consumer financial data outside structured consent infrastructure, through bilateral OAuth-style arrangements with each service) represents a substantial regression in consumer protection.
The framework is an extension to the AA framework that establishes ecosystem-wide accountability for AI processing while preserving the AA framework's ability to support innovation in consumer finance.

The full design paper, A Framework for AI Agents in the Account Aggregator Ecosystem, is available at https://sahamati.org.in/ai-agents-in-aa/. It provides the structural specifications, the architectural details, the operational considerations, and the analysis of alternatives. Feedback may be sent to labs@sahamati.org.in