CAPS: Context-Aware Agentic Payment System

A deterministic payment kernel designed for a world where AI reasons but never decides alone. CAPS creates a trust boundary between AI reasoning and payment execution, ensuring no money moves unless all safety gates pass.

PROBLEM STATEMENT

The Problem

Traditional online payment systems are authentication-driven and reactive—they verify WHO you are, but not WHETHER the action makes sense.

- Lack of Intelligence in Authorization
- Unsafe AI Integration
- No Context Awareness
- Zero Explainability

SOLUTION

Our Solution

CAPS is an AI-powered authorization agent that sits upstream of payment rails, reasoning about intent, context, and risk before approving transactions.

- Make payments with Natural Language and Speech
- No payments are made without your presence
- You understand every step taken
- Every payment is audited
- System learns with every transaction

Multi-Layered Security Architecture

01

Interaction Layer

Captures human intent via UI or Voice

02

LLM Intent Interpreter

Zero Trust translation of natural language to structured JSON

03

Deterministic Control Plane

Schema Validator, Context Evaluator, and Policy Engine

04

Execution Sandbox

Mock UPI Lite execution with cryptographic invariants

05

Immutable Audit Ledger

Write-only, hash-chained record of all actions

Fraud Prevention & Risk Controls

Hallucination Defense

Strict schema validation rejects unknown merchants

Prompt Injection Protection

LLM has no access to system tools or bank APIs

Wallet Draining Prevention

Velocity limits + hard caps prevent rapid fund depletion

Replay Attack Defense

Single-use consent tokens with cryptographic binding

Silent Automation Block

Mandatory human consent step for every transaction

Intent Splitting Detection

Pattern detection identifies micro-transaction sequences

Consent Misuse

Consent is scoped, single-use, and bound to a specific intent.

Ambiguous User Intent

Low confidence or unclear intent automatically downgrades to manual approval.

Missing or Unreliable Context

Automation is reduced or blocked when context signals are incomplete.

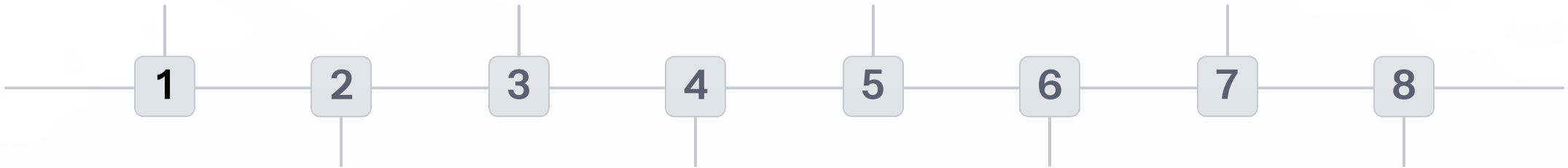
AI or System Failure

Any error or inconsistency defaults to fail-safe behavior.

Consent tokens use JWT with enforced scope (Merchant, Amount, Expiry) and anti-confused-deputy validation.

WORKFLOW

System Workflow



1

Input

User speaks or types command

2

Interpretation

LLM converts text to JSON Intent

3

Validation

Schema Validator ensures JSON integrity

4

Contextualization

System fetches Balance and Velocity data

5

Evaluation

Policy Engine runs 4-layer checks

6

Authorization

User signs Consent Token (JWT)

7

Execution

Sandbox validates Token and logs transaction

8

Feedback

User receives confirmation message