

PROJECT 042 - WORKFLOW ARCHITECTURE

AI Agent + Tools + Human Approval System | Architecture & Future Roadmap

SYSTEM ARCHITECTURE OVERVIEW

The architecture separates AI reasoning from protected execution. The AI Operations Agent handles natural-language requests and read-oriented PostgreSQL tools, while consequential order-status changes are converted into approval requests and processed by a deterministic human-authorization workflow.

1. USER REQUEST Operational question or requested order change.	2. AI OPERATIONS AGENT Interprets intent, uses memory, and selects the appropriate tool.
3. READ / INTELLIGENCE TOOLS Order Lookup • Customer Lookup • Operations Intelligence	4. PROTECTED ACTION GATE Creates a Pending approval instead of modifying the order.
5. HUMAN DECISION Slack presents Approve / Reject controls to the reviewer.	6. CONTROLLED EXECUTION Revalidates, routes, executes authorized outcome, confirms, and logs.

WORKFLOW 01 - MAIN AI OPERATIONS AGENT

Component	Role
Receive Operations Request	Conversational entry point for operational questions and commands.
AI Operations Agent	Reasoning/orchestration layer that identifies intent and selects tools.
AI Reasoning Model	Interprets requests and generates grounded responses.
Conversation Memory	Maintains short conversational context.
Order Lookup	Reads order records from PostgreSQL.
Customer Lookup	Reads customer records from PostgreSQL.
Operations Intelligence	Supports broader operational queries and summaries.
Request Human Approval	Creates Pending authorization for protected order changes; no direct order write.

WORKFLOW 02 - HUMAN APPROVAL & ORDER UPDATE

Approval Request & Notification: Receive Approval Request -> Lookup Pending Approval -> Validate Approval Status -> Send Slack Approval Card.

Human Decision & Validation: Receive Human Decision -> Parse Approval Decision -> Retrieve Approval Request -> Validate Still Pending -> Route Approval Decision.

Controlled Execution & Audit: Approved requests update the order, send confirmation, and log the action. Rejected requests preserve the order, send rejection confirmation, and log the rejection.

KEY DESIGN PRINCIPLES

- Separation of reasoning and authority: AI requests; humans authorize protected writes.
- Validate before mutation: approval state is checked before database changes.
- Deterministic execution: explicit routing controls the approval and rejection paths.
- Auditability: consequential outcomes are recorded with request and execution context.

PROJECT 042 - FUTURE ROADMAP

From portfolio simulation to production-ready AI operations platform

PHASE 1 HARDEN CORE	Authenticated webhooks, role-based authorization, least-privilege database credentials, strict validation, idempotency, error handling, and separate dev/staging/production environments.
PHASE 2 APPROVAL GOVERNANCE	Approval expiry, decision reasons, escalation rules, SLAs, manager routing, and multi-level approval for higher-risk actions.
PHASE 3 EXPAND AI TOOLS	Inventory, shipment tracking, exception analysis, customer history, refund eligibility, fulfillment intelligence, and KPI tools while preserving read/write boundaries.
PHASE 4 OBSERVABILITY	Dashboards for approval volume, response time, tool usage, failures and overrides; plus retries, alerts, dead-letter handling, and health monitoring.
PHASE 5 AI GOVERNANCE	Fine-grained permissions, durable audit retention, model/tool version tracking, sensitive-data controls, evaluations, regression tests, and rollback procedures.

RECOMMENDED NEXT CAPABILITIES

Capability	Purpose
Approval policy engine	Make approval requirements configurable instead of relying only on fixed branches.
Role-based access control	Restrict who may approve specific action categories.
Idempotency / replay protection	Prevent duplicate mutations from repeated events or button clicks.
Approval expiry & escalation	Prevent stale Pending requests and improve response time.
Monitoring dashboard	Track requests, approvals, outcomes, failures, and business impact.
Automated evaluations	Test tool selection, guardrails, and grounded agent behavior.
Production audit store	Provide durable evidence for troubleshooting and governance.

TARGET PRODUCTION ARCHITECTURE

User / Operations Team -> AI Operations Interface -> AI Agent & Tool Router -> Read Tools OR Protected Action Gateway -> Approval Policy -> Slack / Approval UI -> Human Decision -> Validation & Authorization -> Transactional Database Update -> Confirmation -> Audit & Monitoring
--

LONG-TERM BUSINESS OUTCOME

The target is a governed AI operations layer: employees interact with business systems conversationally, AI handles low-risk retrieval and analysis, and higher-risk actions pass through explicit authorization controls. This increases automation while preserving human accountability and traceability.