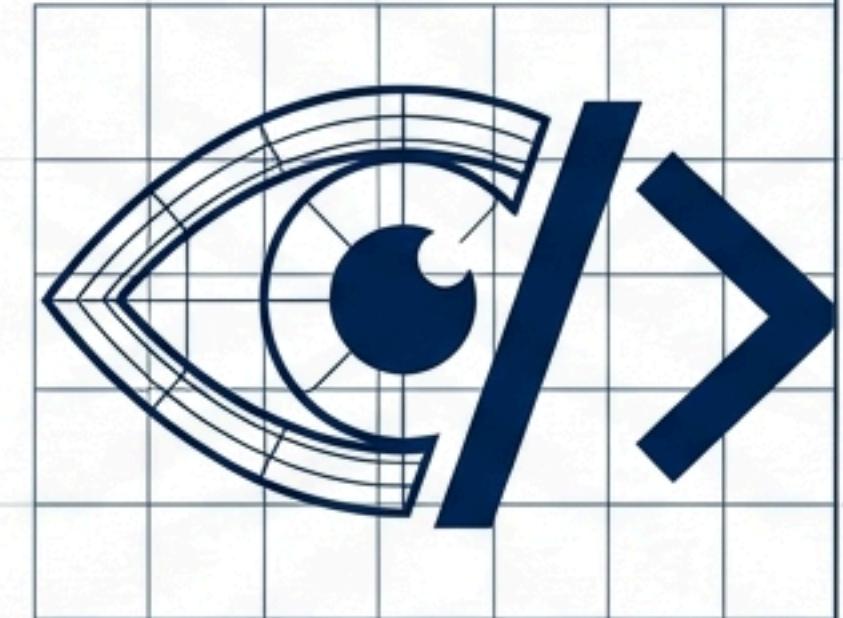


# THE MULTIMODAL AGENTIC ASSISTANT

A Technical Deep Dive into the Nexus Members Assistant



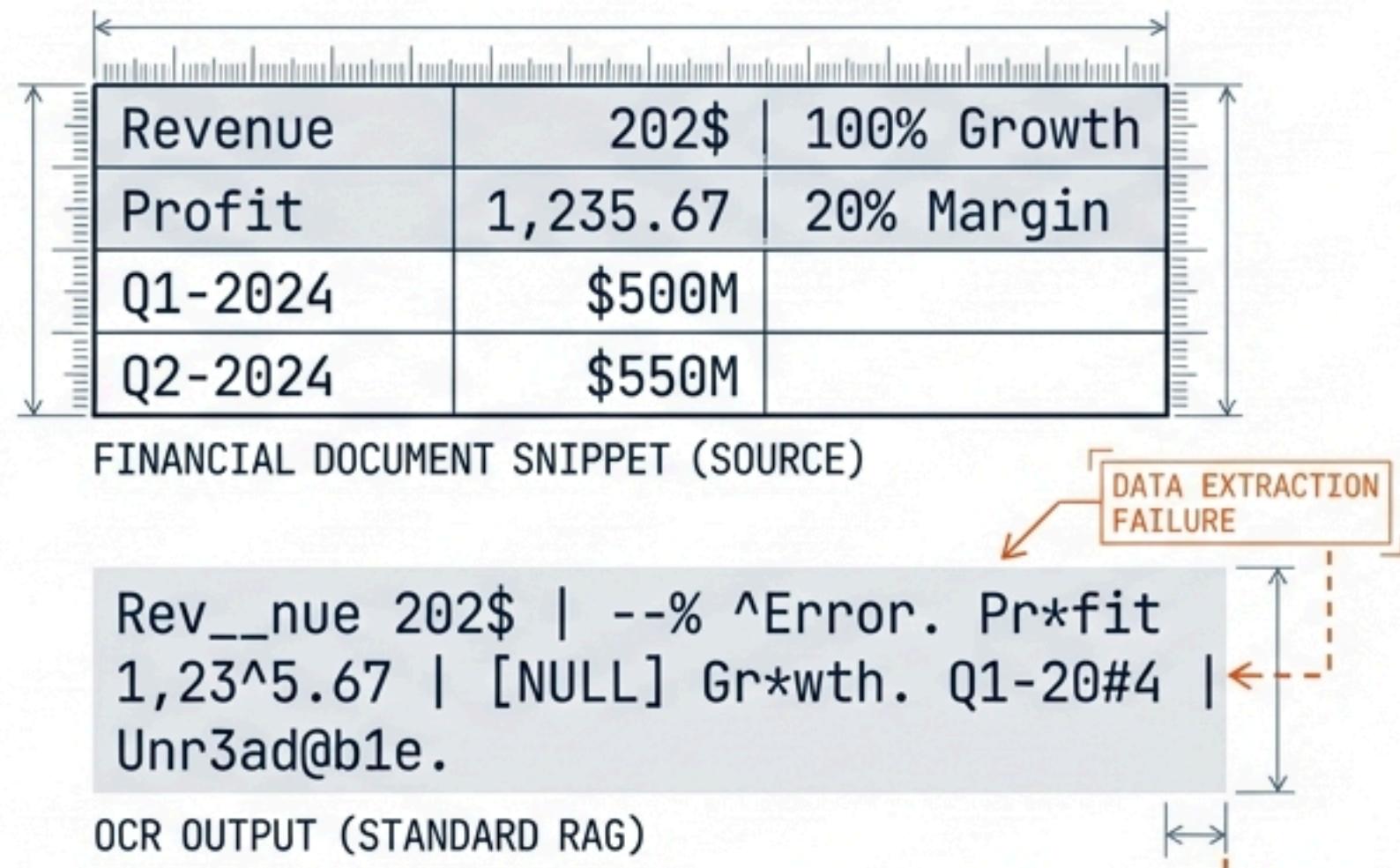
PROJECT: Round 1 Submission | HCLTech Hackathon

**TEAM NEXUS:-**  
**Shuvradeep Bera, Aman Yadav**  
**Sandeep Saikia ,Shourya Kumar**

# WHY STANDARD RAG FAILS FINANCIAL DATA

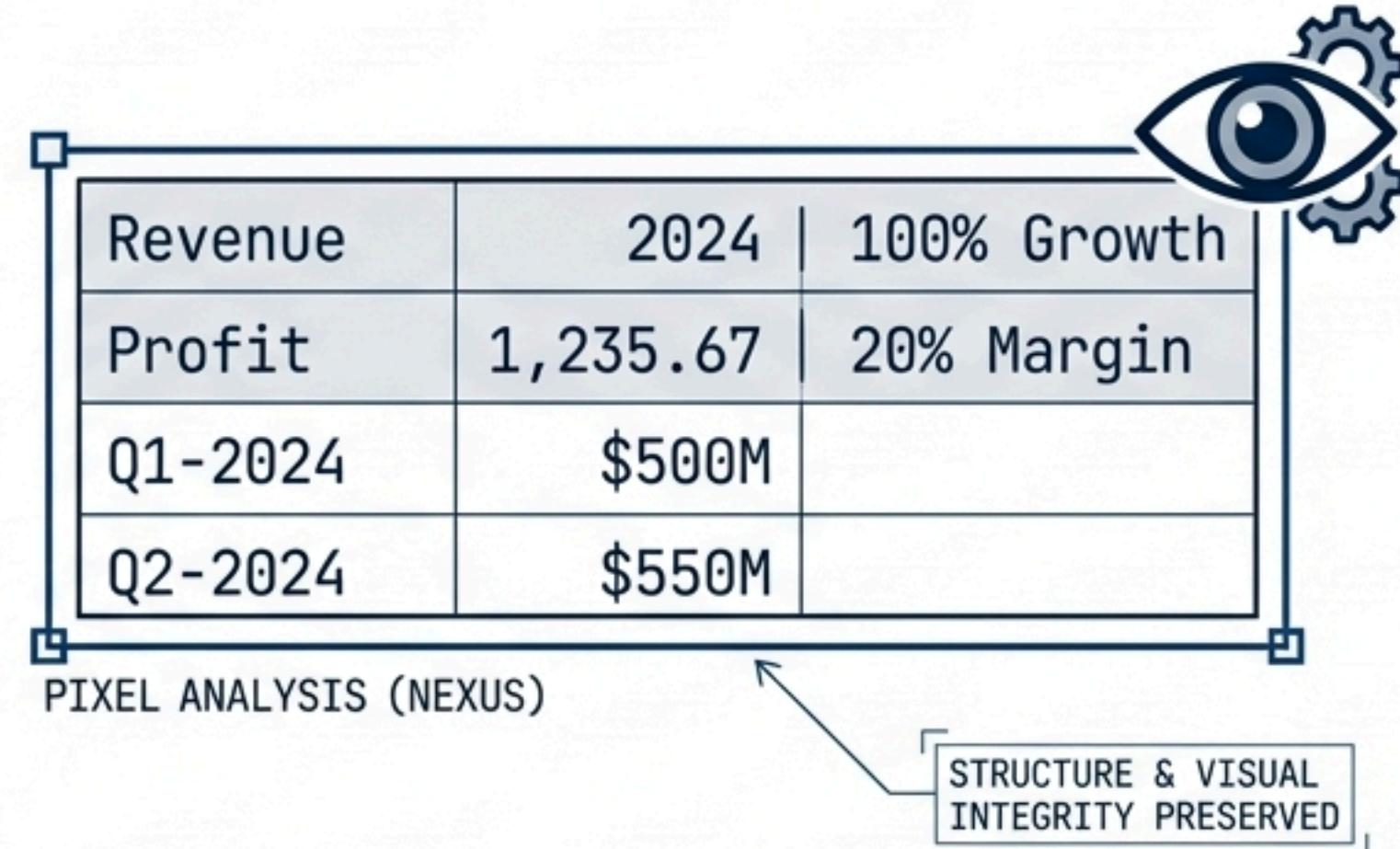


## APPROACH A: TEXT EXTRACTION (OCR)



✗ **Result: Hallucination & Structure Loss**

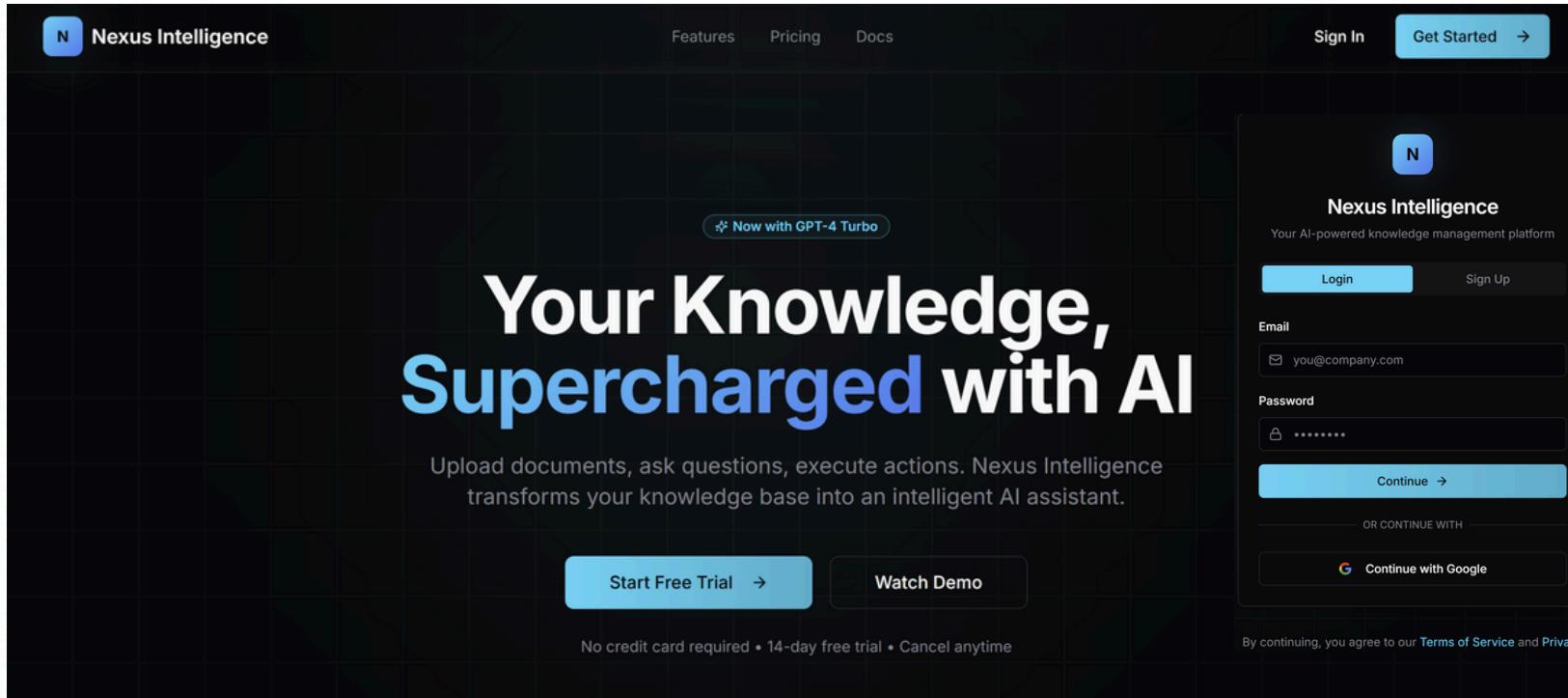
## APPROACH B: NEXUS (PIXEL ANALYSIS)



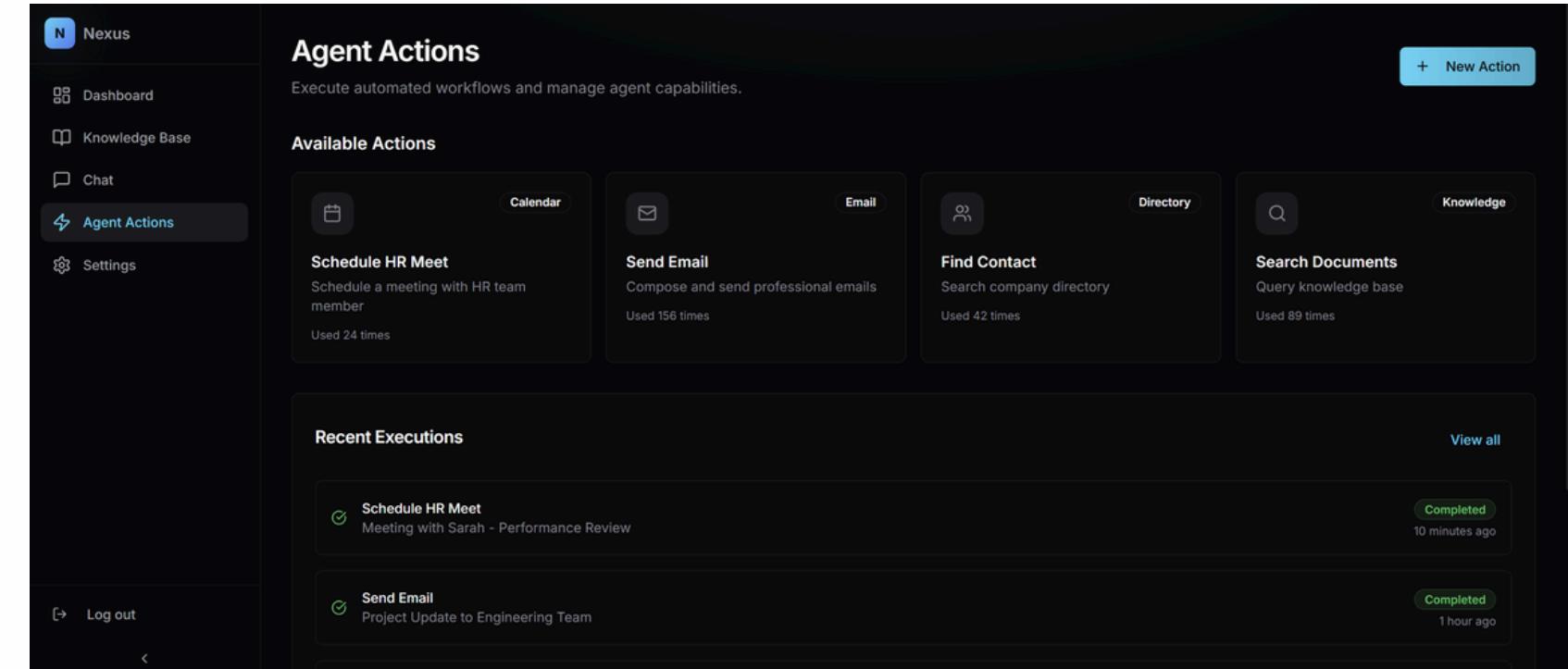
✓ **Result: 100% Fidelity (Visual Context)**

**THE GAP:** Moving from Reading Text to Seeing Pixels

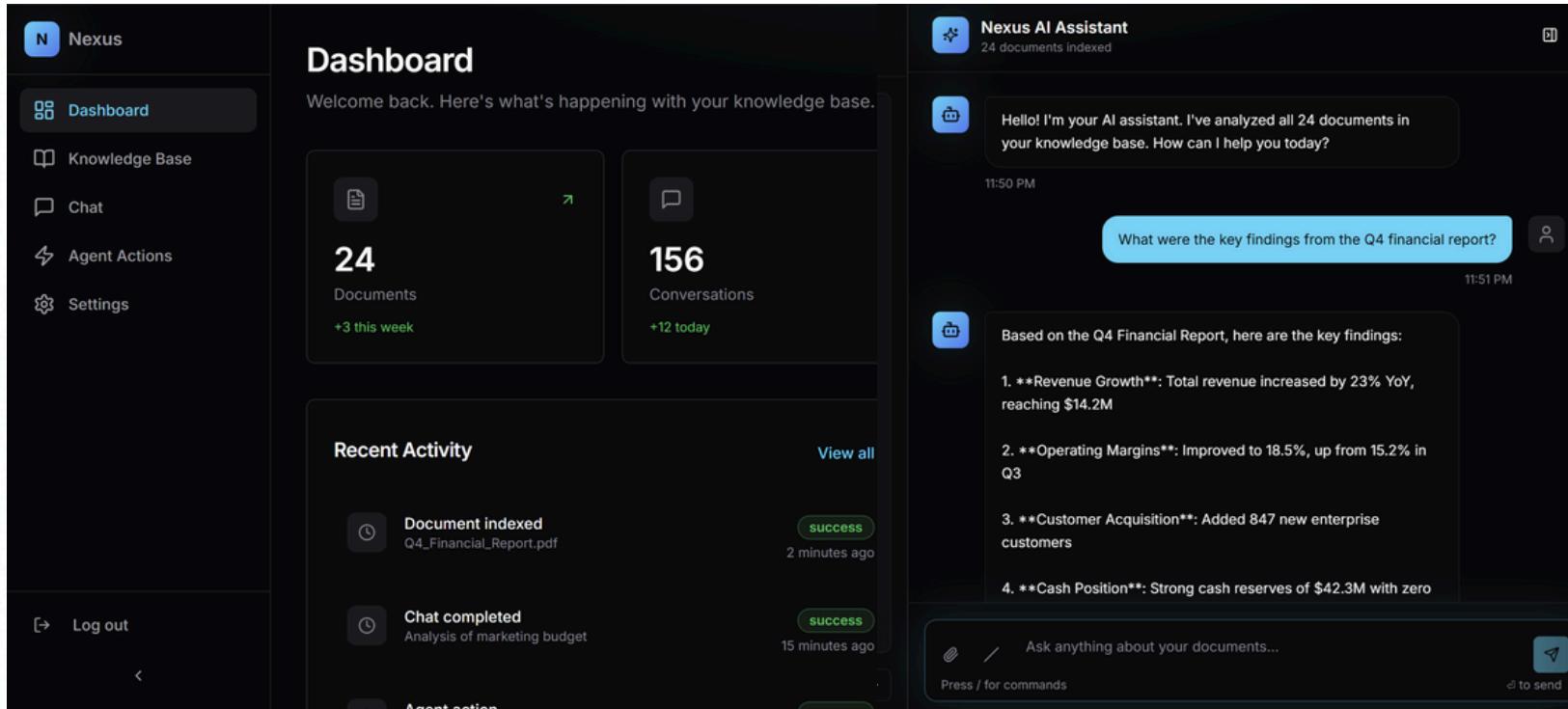
# SOLUTION: THE NEXUS MEMBERS ASSISTANT



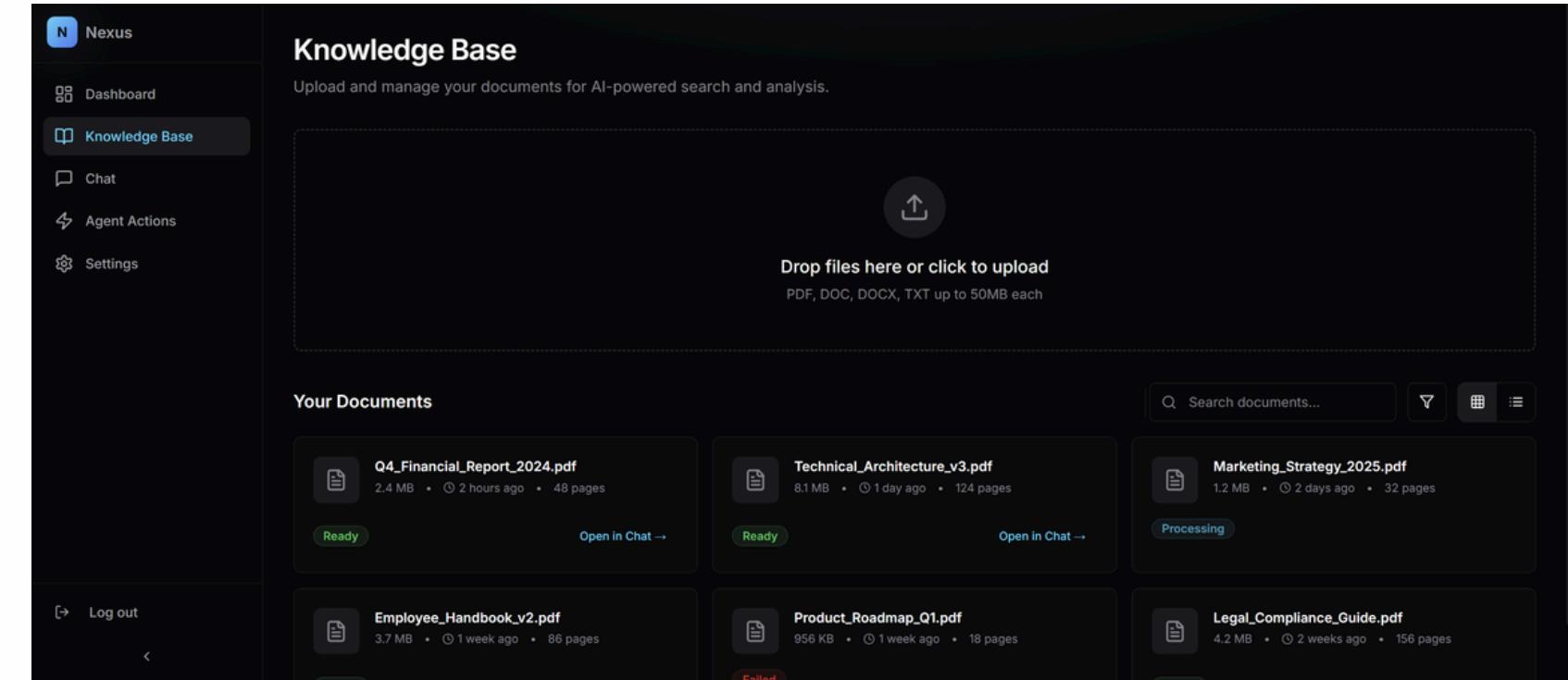
The landing page for Nexus Intelligence features a dark-themed header with navigation links for "Features", "Pricing", and "Docs". A "Sign In" button and a "Get Started" button are prominently displayed. Below the header, a large call-to-action section highlights "Now with GPT-4 Turbo" and "Your Knowledge, Supercharged with AI". It includes a login form with fields for "Email" and "Password", and a "Continue" button. Below the login form, there's a "OR CONTINUE WITH" section featuring a "Continue with Google" button. At the bottom, there's a note about terms of service and privacy, along with "Start Free Trial" and "Watch Demo" buttons.



The Agent Actions dashboard shows a sidebar with "Dashboard", "Knowledge Base", "Chat", "Agent Actions" (which is selected), and "Settings". The main area is titled "Agent Actions" with the sub-instruction "Execute automated workflows and manage agent capabilities." It displays a grid of available actions: "Schedule HR Meet" (used 24 times), "Send Email" (used 156 times), "Find Contact" (used 42 times), and "Search Documents" (used 89 times). Below this, a "Recent Executions" section lists completed tasks: "Schedule HR Meet" (Meeting with Sarah - Performance Review, completed 10 minutes ago) and "Send Email" (Project Update to Engineering Team, completed 1 hour ago).

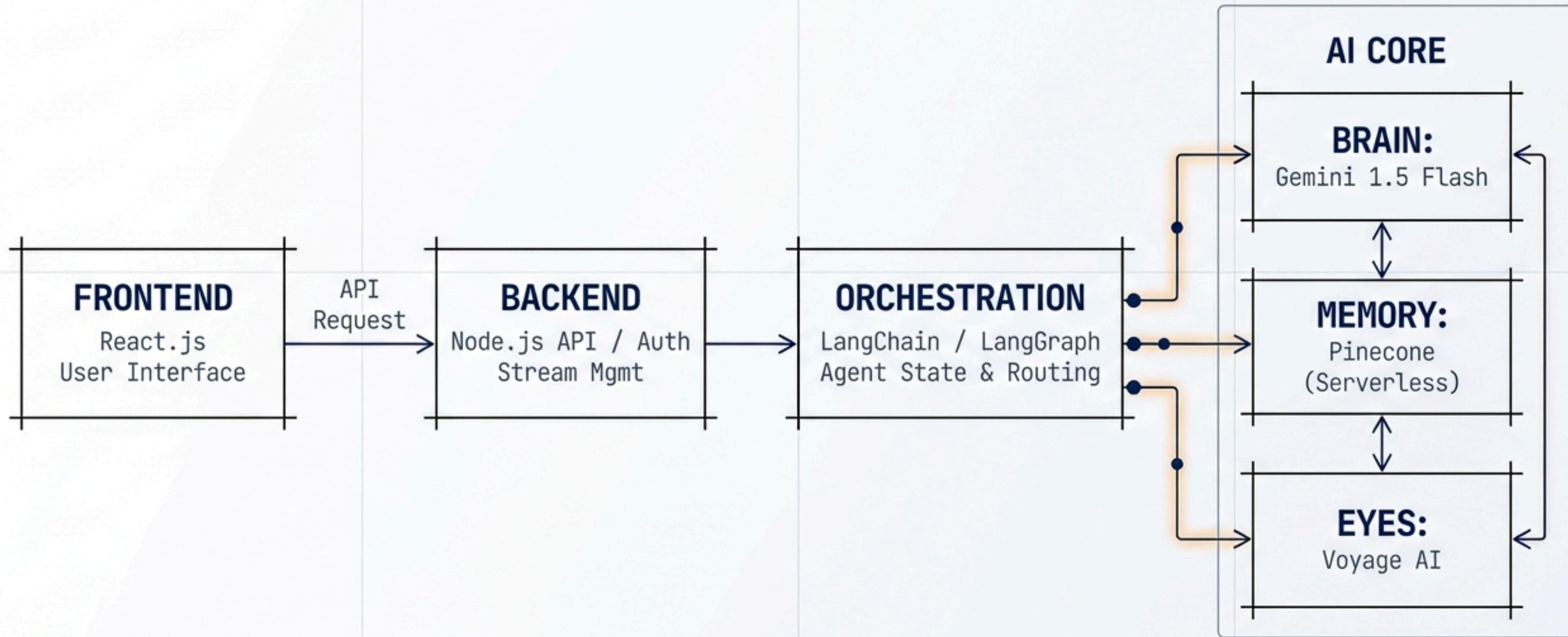


The Nexus Dashboard has a sidebar with "Dashboard" (selected), "Knowledge Base", "Chat", "Agent Actions", and "Settings". The main area is titled "Dashboard" with the sub-instruction "Welcome back. Here's what's happening with your knowledge base.". It shows two cards: "24 Documents" (+3 this week) and "156 Conversations" (+12 today). Below this is a "Recent Activity" section showing "Document indexed" (Q4\_Financial\_Report.pdf, success, 2 minutes ago) and "Chat completed" (Analysis of marketing budget, success, 15 minutes ago). At the bottom, there's a text input field "Ask anything about your documents..." and a "Press / for commands" button.



The Knowledge Base dashboard has a sidebar with "Dashboard", "Knowledge Base" (selected), "Chat", "Agent Actions", and "Settings". The main area is titled "Knowledge Base" with the sub-instruction "Upload and manage your documents for AI-powered search and analysis.". It features a "Drop files here or click to upload" area with a note "PDF, DOC, DOCX, TXT up to 50MB each". Below this is a "Your Documents" section displaying several document cards: "Q4\_Financial\_Report\_2024.pdf" (2.4 MB, 2 hours ago, 48 pages, Ready, Open in Chat →), "Technical\_Architecture\_v3.pdf" (8.1 MB, 1 day ago, 124 pages, Ready, Open in Chat →), "Marketing\_Strategy\_2025.pdf" (1.2 MB, 2 days ago, 32 pages, Processing), "Employee\_Handbook\_v2.pdf" (3.7 MB, 1 week ago, 86 pages, Failed), and "Product\_Roadmap\_Q1.pdf" (956 KB, 2 weeks ago, 18 pages, Failed).

# HIGH-LEVEL SYSTEM ARCHITECTURE



Data Flow: Left to Right

# STRATEGIC CHOICE: MULTIMODAL NATIVE

## 01

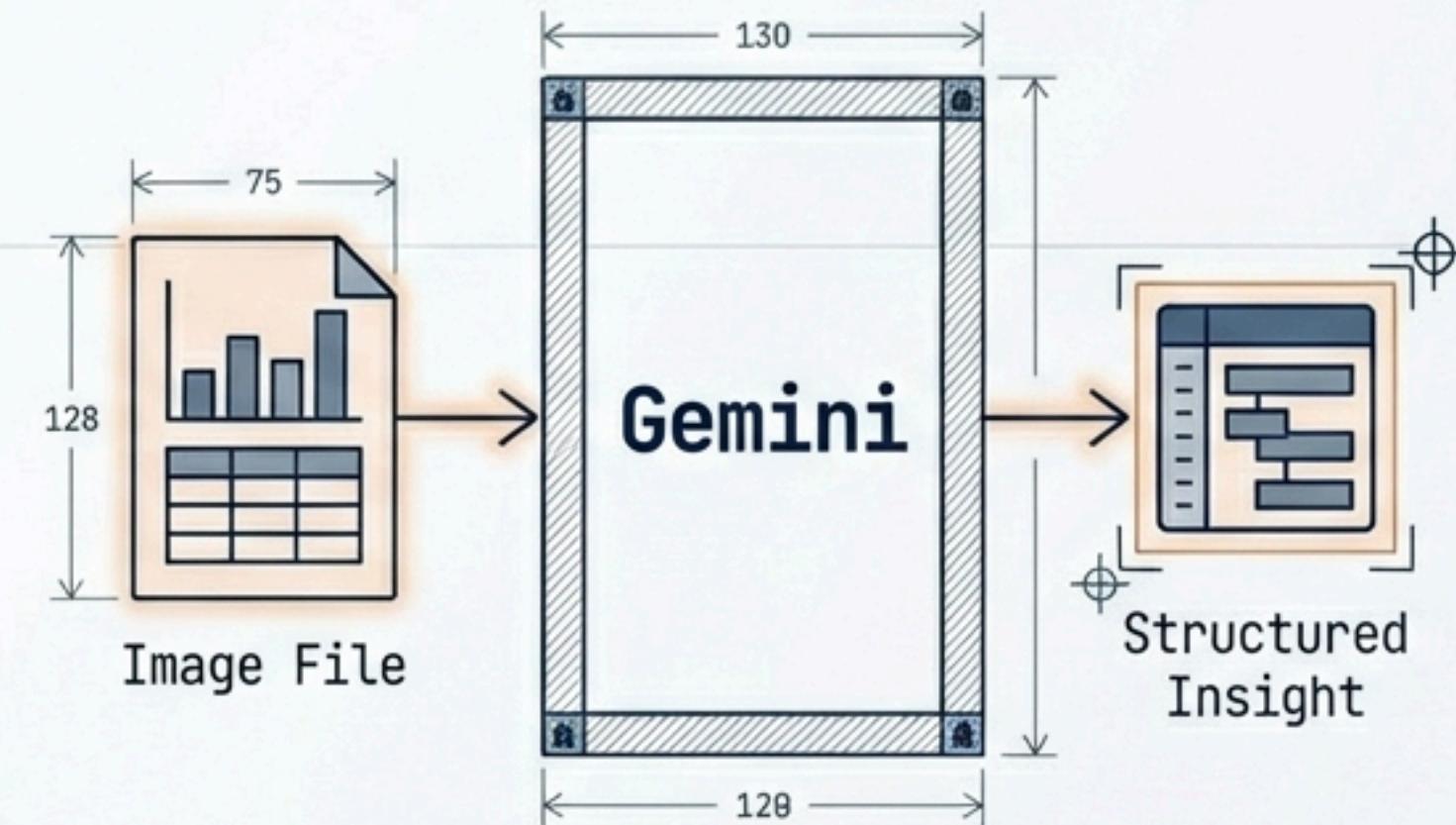
### THE BRAIN

GOOGLE GEMINI 1.5 FLASH



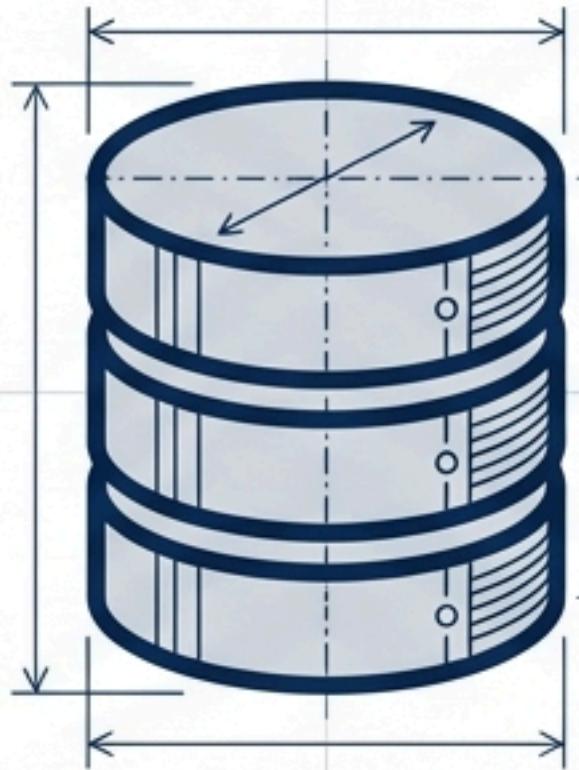
#### STRATEGIC CHOICE: MULTIMODAL NATIVE

- Direct Vision:** Analyzes images (charts/tables) within the context window.
- Error Elimination:** Bypasses OCR text parsing completely.
- Cost Efficiency:** Optimized for high-volume enterprise queries.



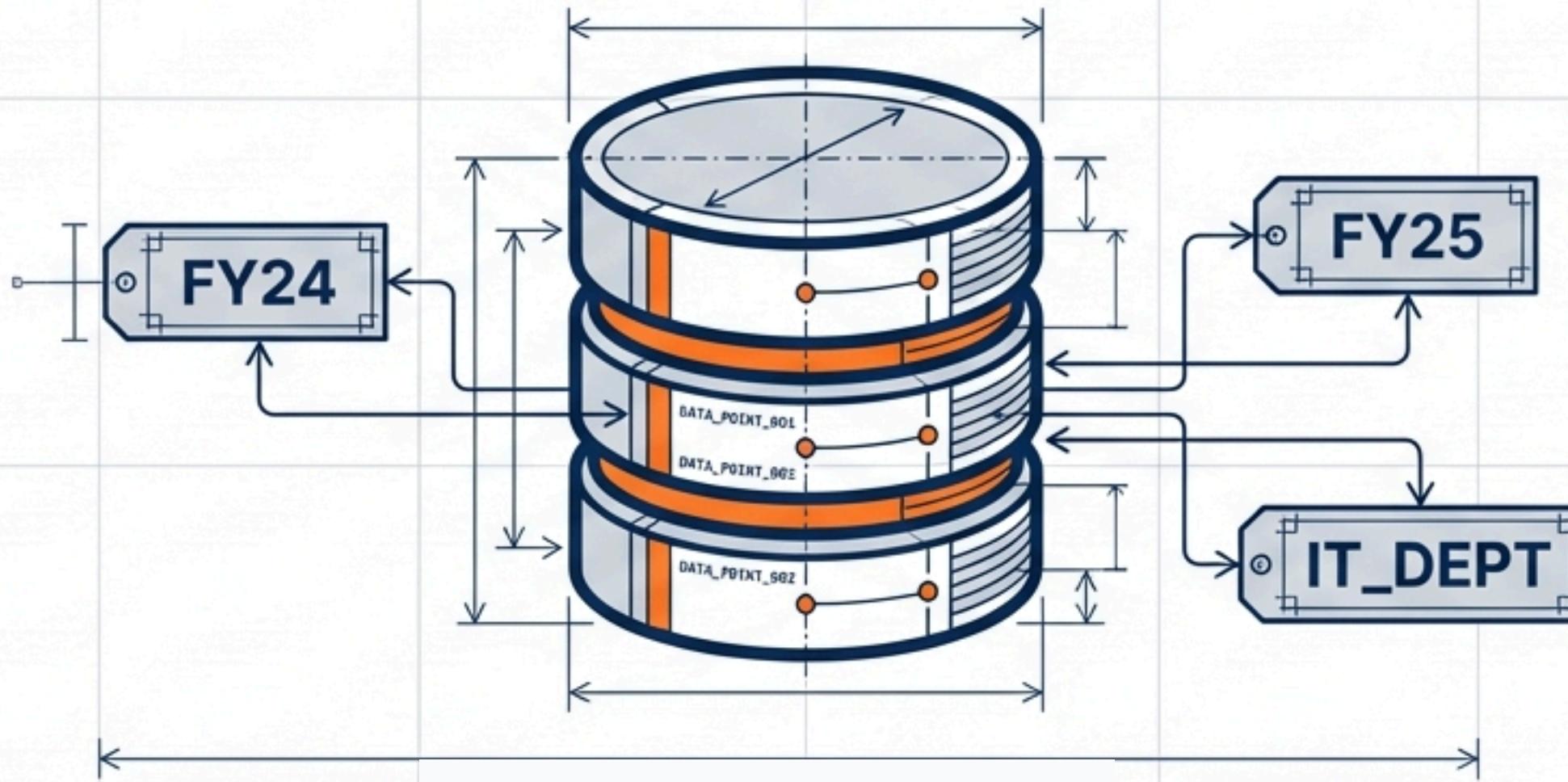
# 02

## THE MEMORY PINECONE VECTOR DB



### STRATEGIC CHOICE: SERVERLESS SCALE

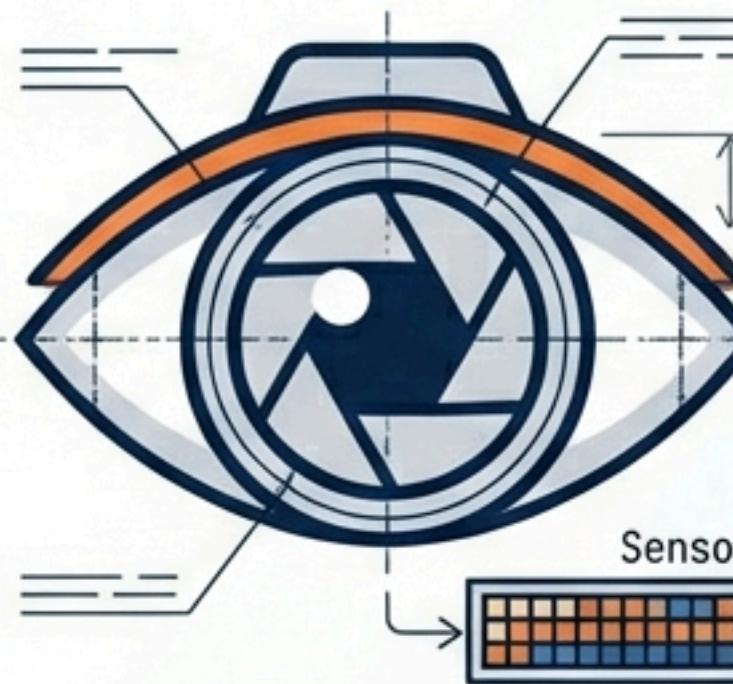
- ✓ **Scalability:** Serverless architecture handles massive indexes.
- ✓ **Metadata Filtering:** Precision retrieval by "Fiscal Year" or "Business Unit".
- ✓ **Collaboration:** Cloud-native structure for distributed team access.



# STRATEGIC CHOICE: PIXEL-LEVEL INDEXING

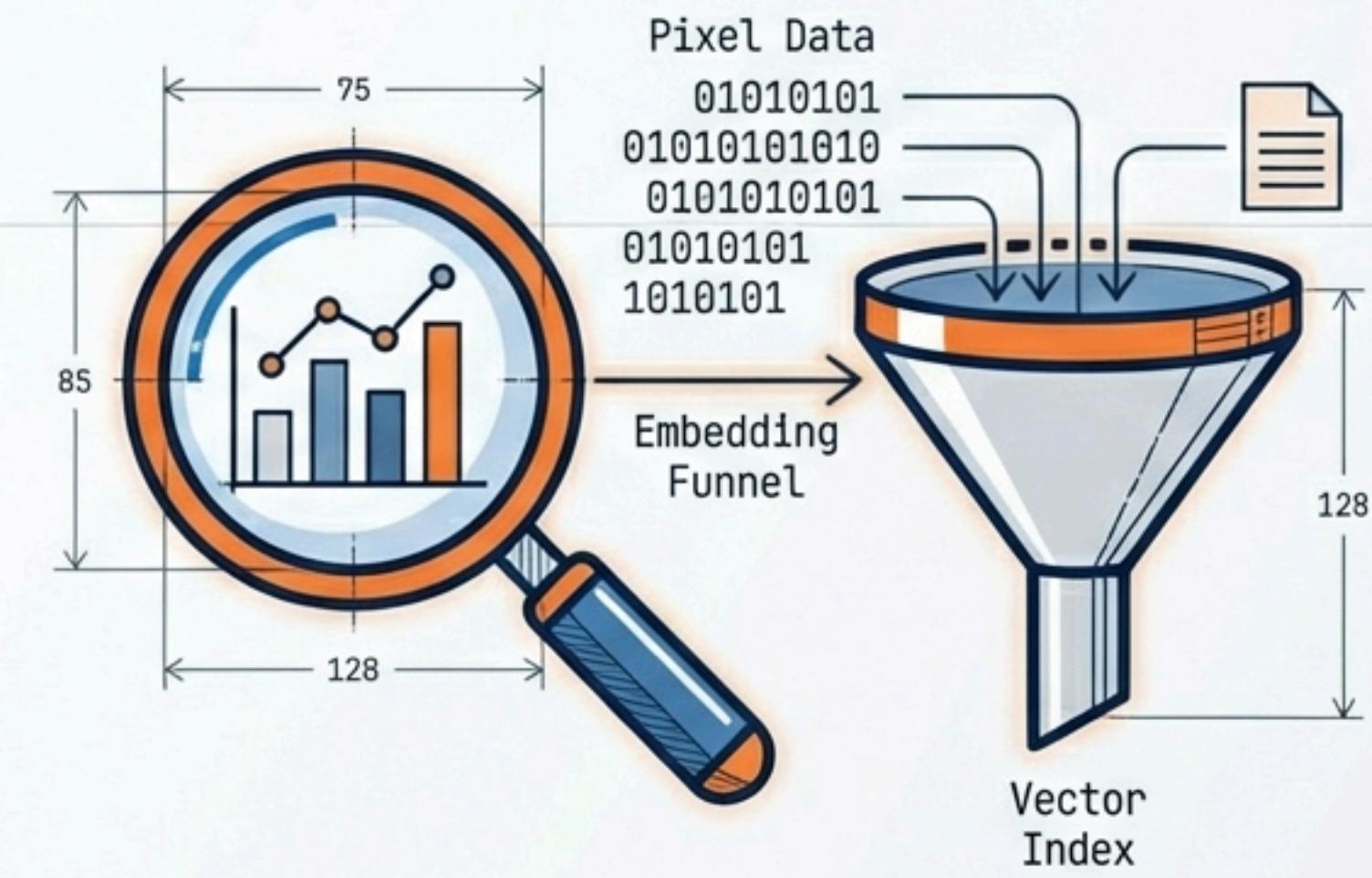
## 03 THE EYES

VOYAGE AI (MULTIMODAL-3)

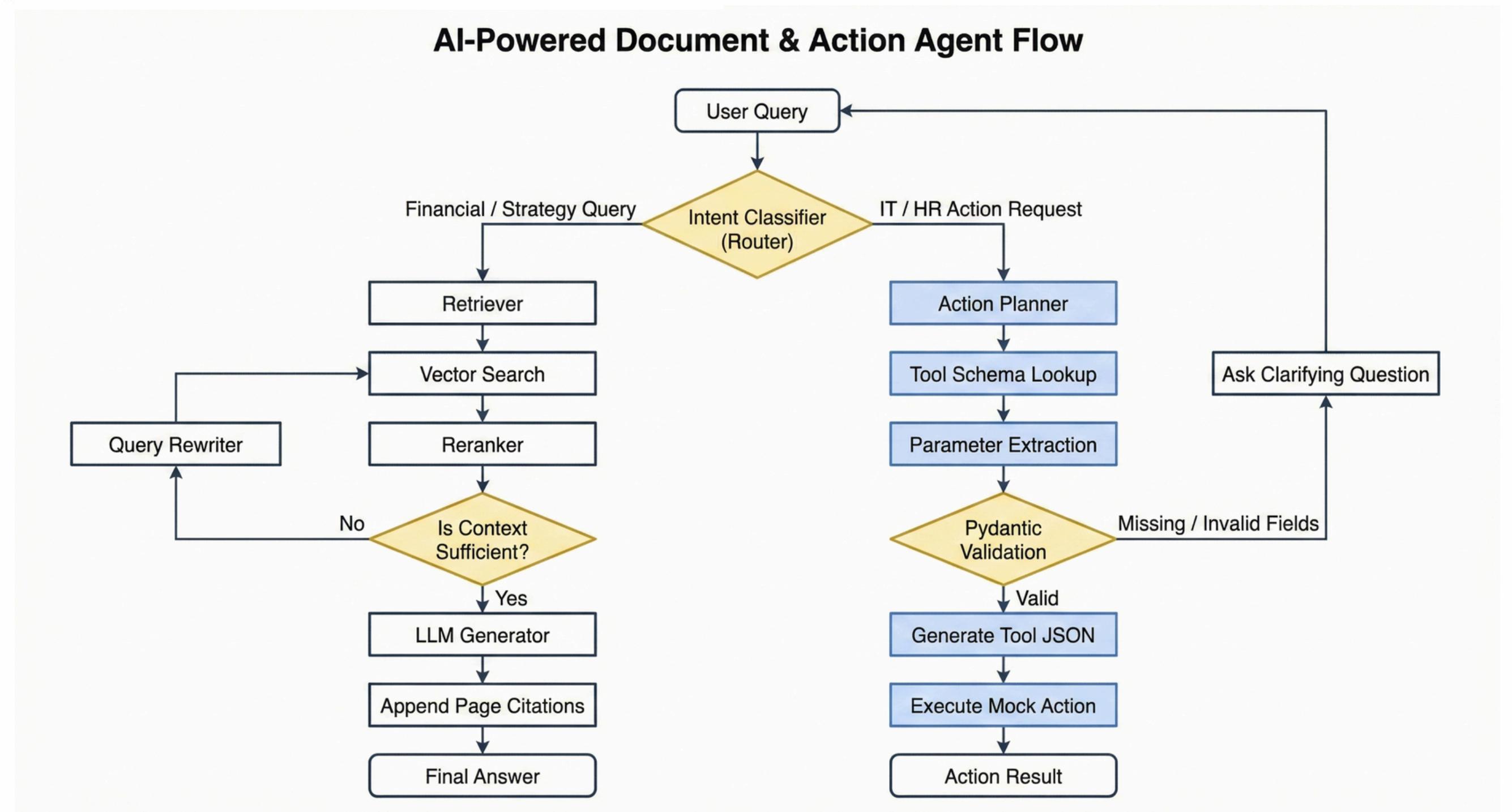


### STRATEGIC CHOICE: PIXEL-LEVEL INDEXING

- Beyond Text:** Creates vector embeddings from image pixels, not captions.
- Visual Search:** Matches text queries ("Show revenue graph") to visual artifacts.
- The Differentiator:** Enables retrieval of un-captioned charts.



# 04 THE WORKFLOW



# WORKFLOW: RETRIEVAL & REASONING

FROM USER QUERY TO VISUAL ANALYSIS & RESPONSE

User Query

What was the  
Revenue Growth?

THE RETRIEVED ARTIFACT

Financial Year	Revenue (₹ Crores)	Net Profit	Revenue Growth (%)
201-2021	₹117,055	₹50,200	0.60%
201-2022	₹128,350	₹37,360	1.30%
201-2024	₹117,055	₹31,800	6.57%
204-2023	₹117,055	₹31,800	6.26%
204-2034	₹163,080	₹95,760	3.35%
201-2023	₹154,080	₹33,350	4.87%
201-2021	₹158,000	₹97,300	6.50%
201-2028	₹108,555	₹35,750	5.50%
201-2023	₹145,000	₹376,360	10.50%

Visual  
Retrieval

Model  
Ingestion

KEY INSIGHT

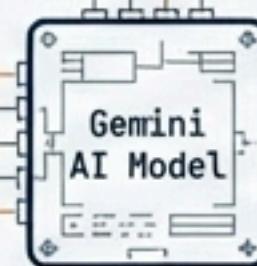
Accuracy is guaranteed by analyzing  
the source image directly.

GEMINI 1.5 FLASH ANALYSIS

Revenue  
(₹ Crores)

₹117,055 Crores

₹117,055 Crores



Gemini 1.5  
Flash

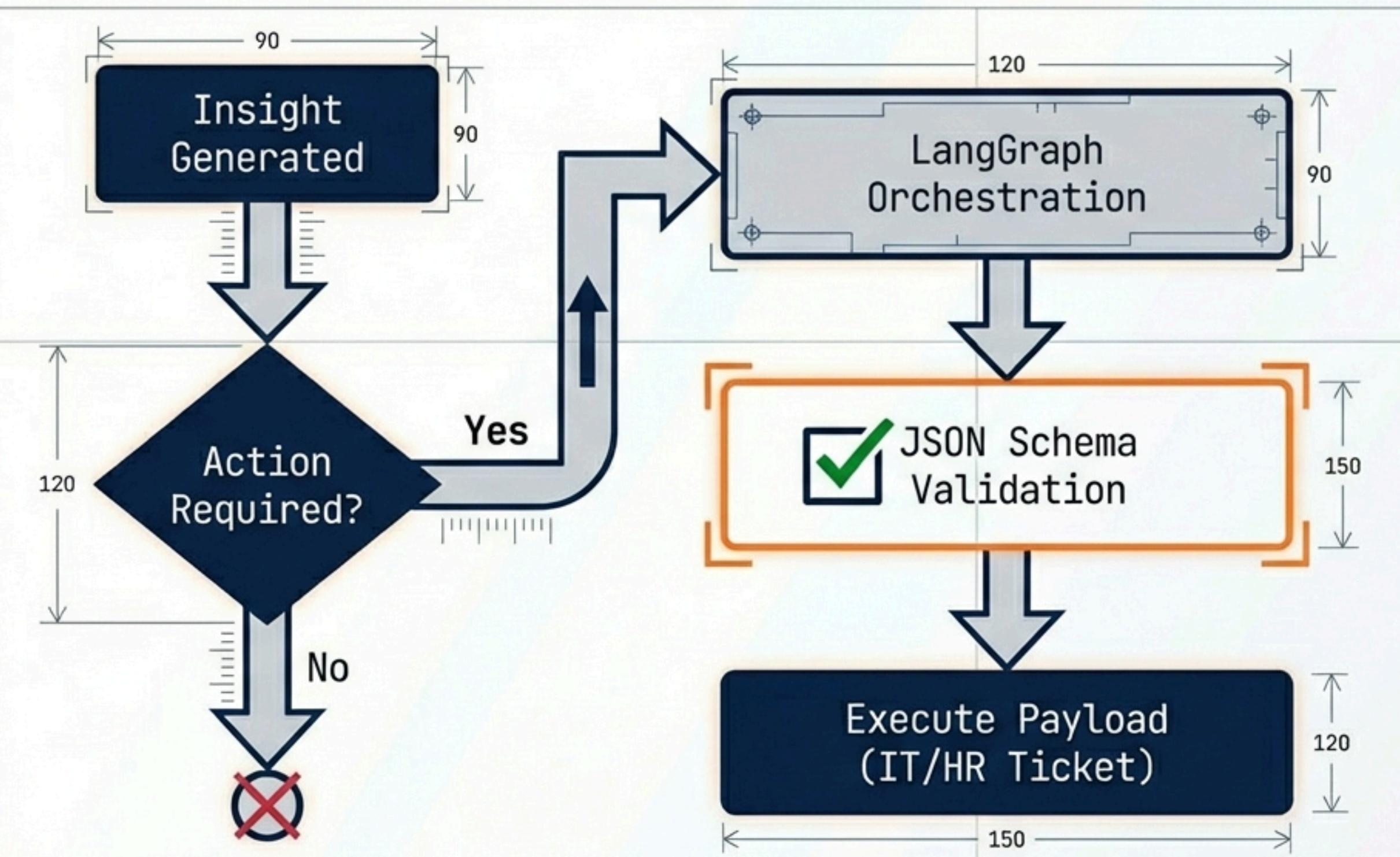
Data Flow: Visual Reasoning & Extraction

System Response

Based on the table,  
Revenue is ₹117,055  
Crores.

# WORKFLOW: AGENTIC ACTION

FROM INSIGHT TO EXECUTION



Safety First:  
LangGraph enforces strict JSON structure to prevent bot hallucinations during execution.