15

AI AGENTS

Learn how there are being used in the real world

→

Bhavishya Pandit

Introduction

The AI agent industry hit **$5.4 billion** in 2024 (GrandViewResearch).

Klarna's AI agent is doing the work of **700 employees** all on its own.

So, What Exactly Is an AI Agent?

An AI agent is a smart system that can think, decide, and act on its own to complete tasks, just like a digital assistant. It analyzes information, makes choices, and uses tools to get things done—without needing constant human input.
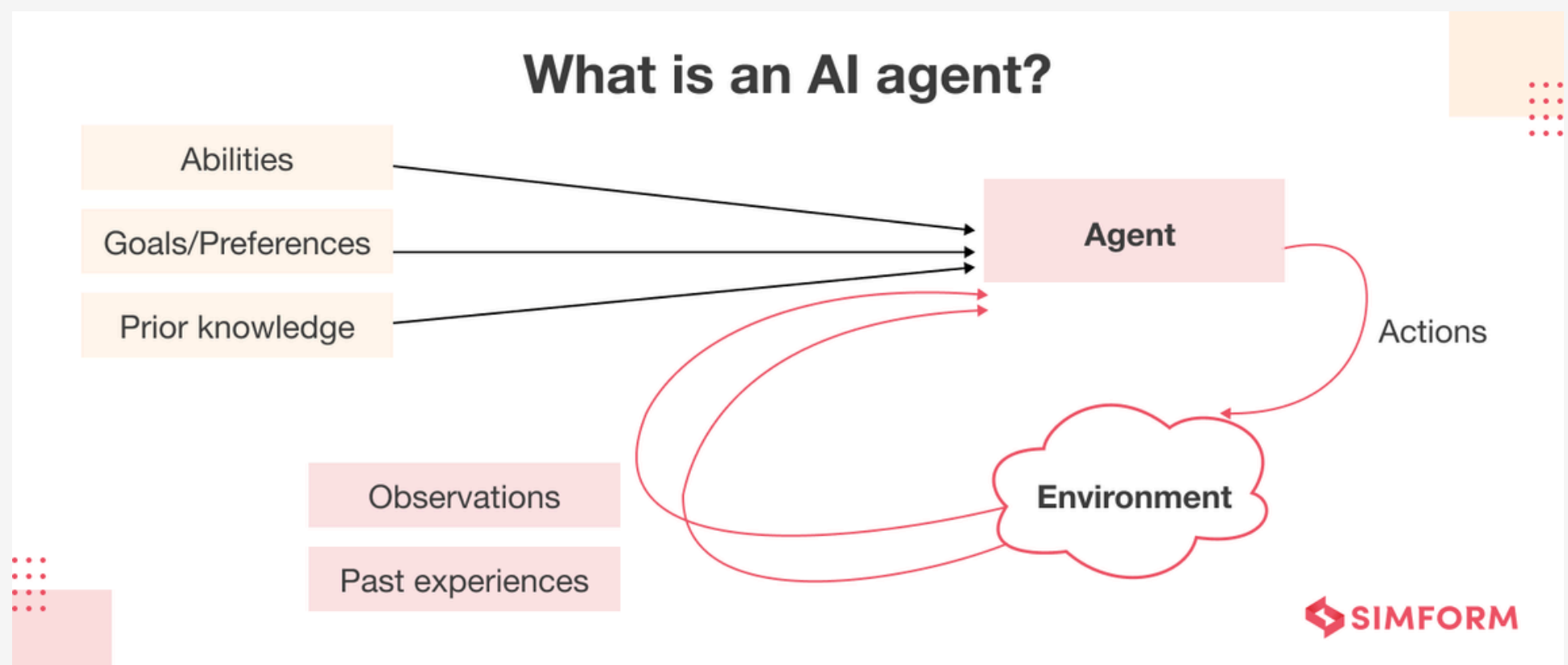
From coding to finance, education to architecture, AI agents are everywhere.

In this post you will get to know the top AI agents in the industry that are helping businesses grow.
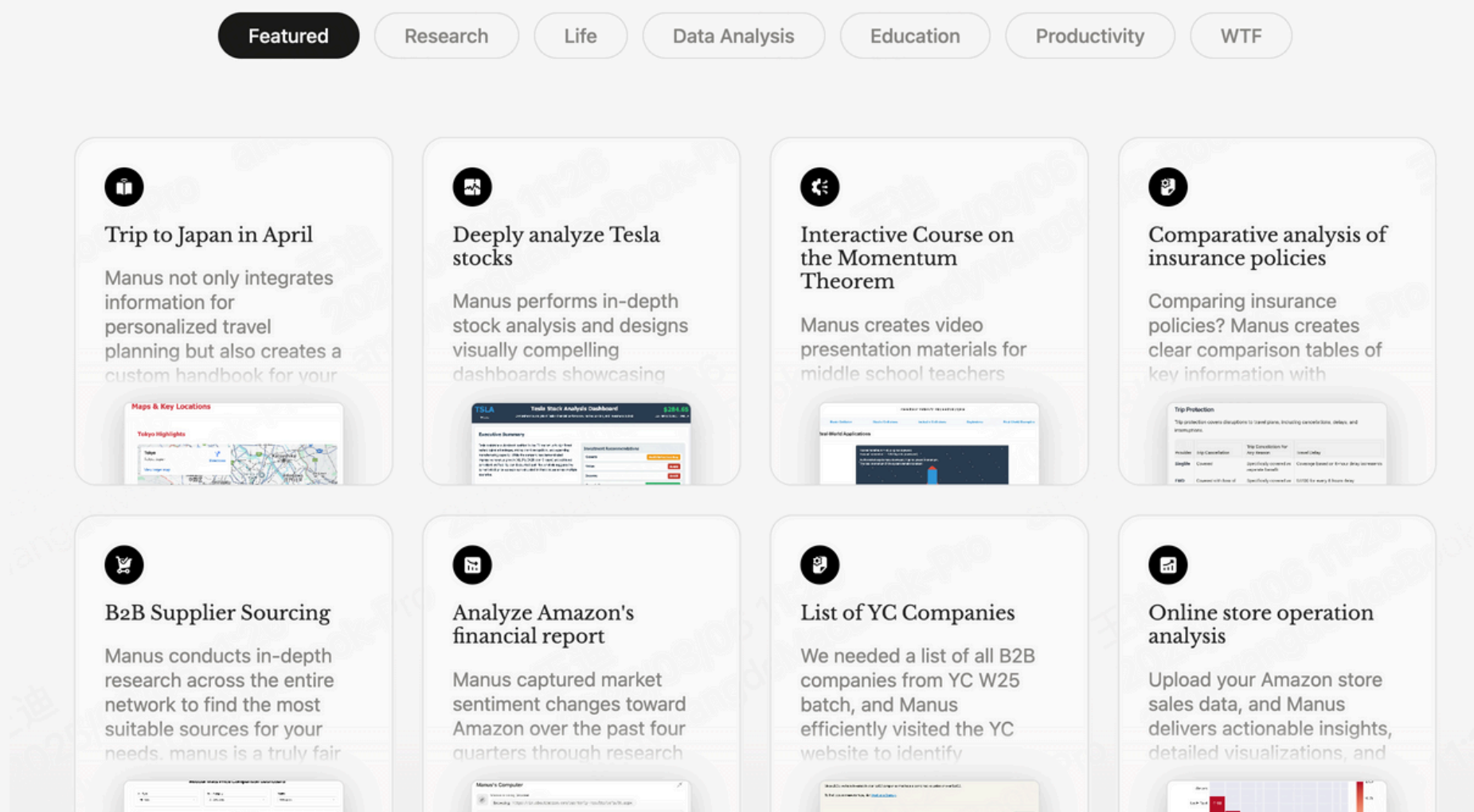
1. Manus AI

Manus AI is an advanced AI agent designed for autonomous task execution, outperforming traditional assistants with LLMs, multi-modal processing, and tool integration.

Key Features:

✅ Task Automation – Handles reports, data analysis, content creation, and more.
✅ Multi-Modal AI – Processes text, images, and code.
✅ Tool Integration – Connects with web browsers, databases, and editors.
✅ Adaptive Learning – Optimizes responses over time.

Use case gallery

Learn how Manus handles real-world tasks through step-by-step replays.

Featured | Research | Life | Data Analysis | Education | Productivity | WTF

Trip to Japan in April

Manus not only integrates information for personalized travel planning but also creates a custom handbook for your

Deeply analyze Tesla stocks

Manus performs in-depth stock analysis and designs visually compelling dashboards showcasing

Interactive Course on the Momentum Theorem

Manus creates video presentation materials for middle school teachers

Comparative analysis of insurance policies

Comparing insurance policies? Manus creates clear comparison tables of key information with

B2B Supplier Sourcing

Manus conducts in-depth research across the entire network to find the most suitable sources for your needs. manus is a truly fair

Analyze Amazon's financial report

Manus captured market sentiment changes toward Amazon over the past four quarters through research

List of YC Companies

We needed a list of all B2B companies from YC W25 batch, and Manus efficiently visited the YC website to identify

Online store operation analysis

Upload your Amazon store sales data, and Manus delivers actionable insights, detailed visualizations, and

GAIA Benchmark Leader

🏆 Outperforms GPT-4, AutoGPT, and h2oGPTe, achieving state-of-the-art results in reasoning, multi-modal processing, and automation.
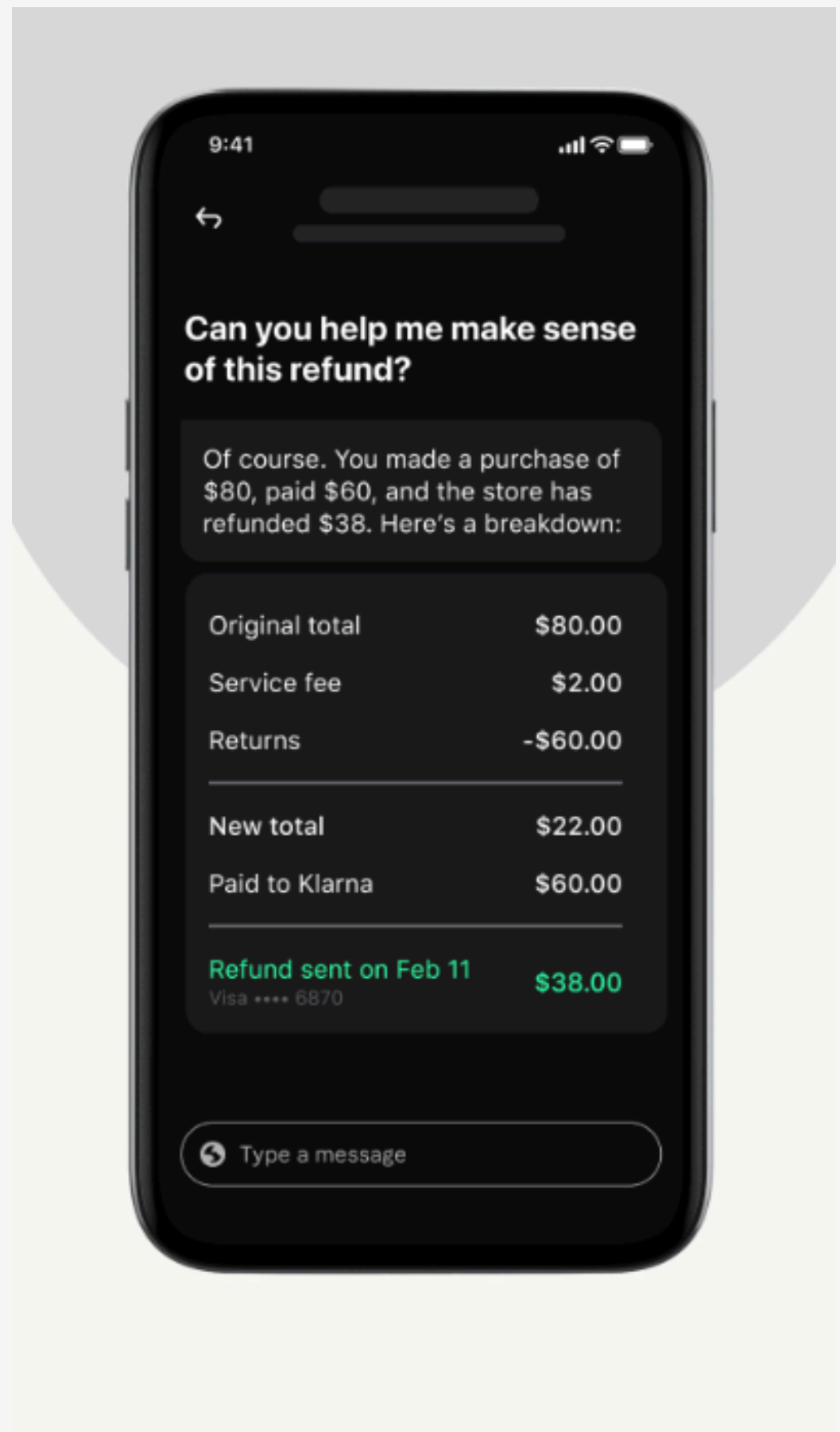
Bhavishya Pandit

2.Klarna

Klarna built an advanced AI assistant using LangGraph and LangSmith. This system:

✅ Route customer queries

✅ Provides context-aware responses

✅ Use test-driven development for behavior analysis and prompt optimization.

The Impact:

🔷 80% faster customer query resolution.

🔷 70% of repetitive tasks automated, freeing human agents for complex cases.

🔷 Improved accuracy, reducing unnecessary escalations.

9:41

Can you help me make sense of this refund?

Of course. You made a purchase of $80, paid $60, and the store has refunded $38. Here's a breakdown:

Original total	$80.00
Service fee	$2.00
Returns	-$60.00
New total	$22.00
Paid to Klarna	$60.00
Refund sent on Feb 11 Visa •••• 6870	$38.00

Type a message

3.Google's Project Astra

A **multimodal AI** agent capable of processing text, speech, images, and videos. It uses visual memory to identify objects and remember locations, enhancing its ability to assist users in real-time.

- **Multimodal Processing** – Handles text, speech, images, and videos for a richer AI experience.

- **Visual Memory** – Identifies objects through a camera and remembers locations for better contextual understanding.

- **Human-Like Interaction** – Combines visual and verbal inputs for more natural and dynamic conversations.

- **Real-Time Assistance** – Helps users by answering questions, retrieving information, and identifying objects on the go.

4. Vizient

Vizient, a leader in healthcare performance improvement, is transforming how hospitals and clinics analyze data.

✅ LangGraph for Orchestration – With LangGraph, they implemented a structured hierarchy where worker agents report to a supervisor agent, ensuring seamless coordination and higher reliability.

✅ LangSmith for Observability & Prompt Management – Vizient uses LangSmith's tracing capabilities to debug real-time issues and optimize responses.
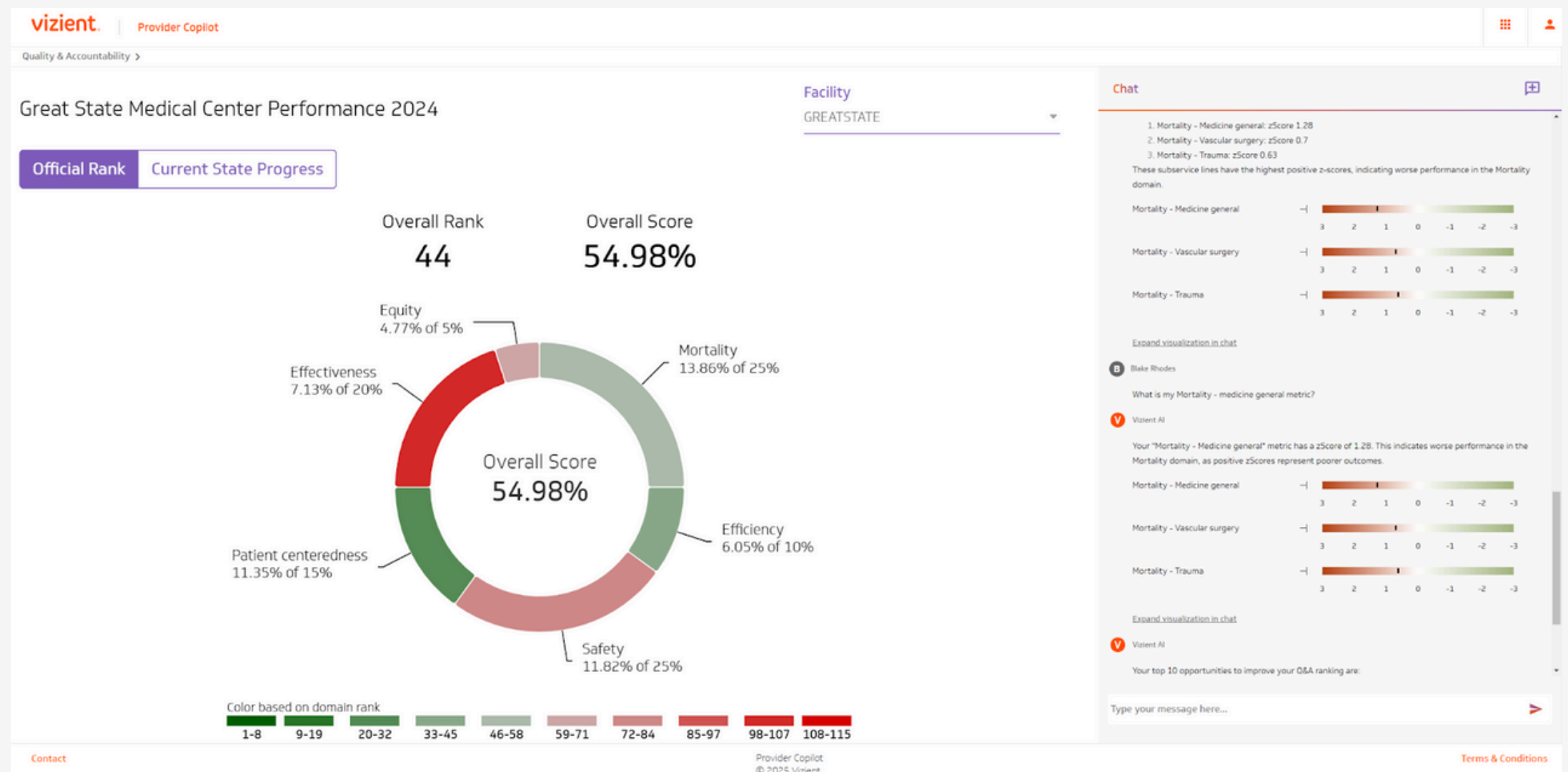
5. Captide

Captide leverages **LangGraph and LangSmith** to build agentic workflows that streamline financial research.
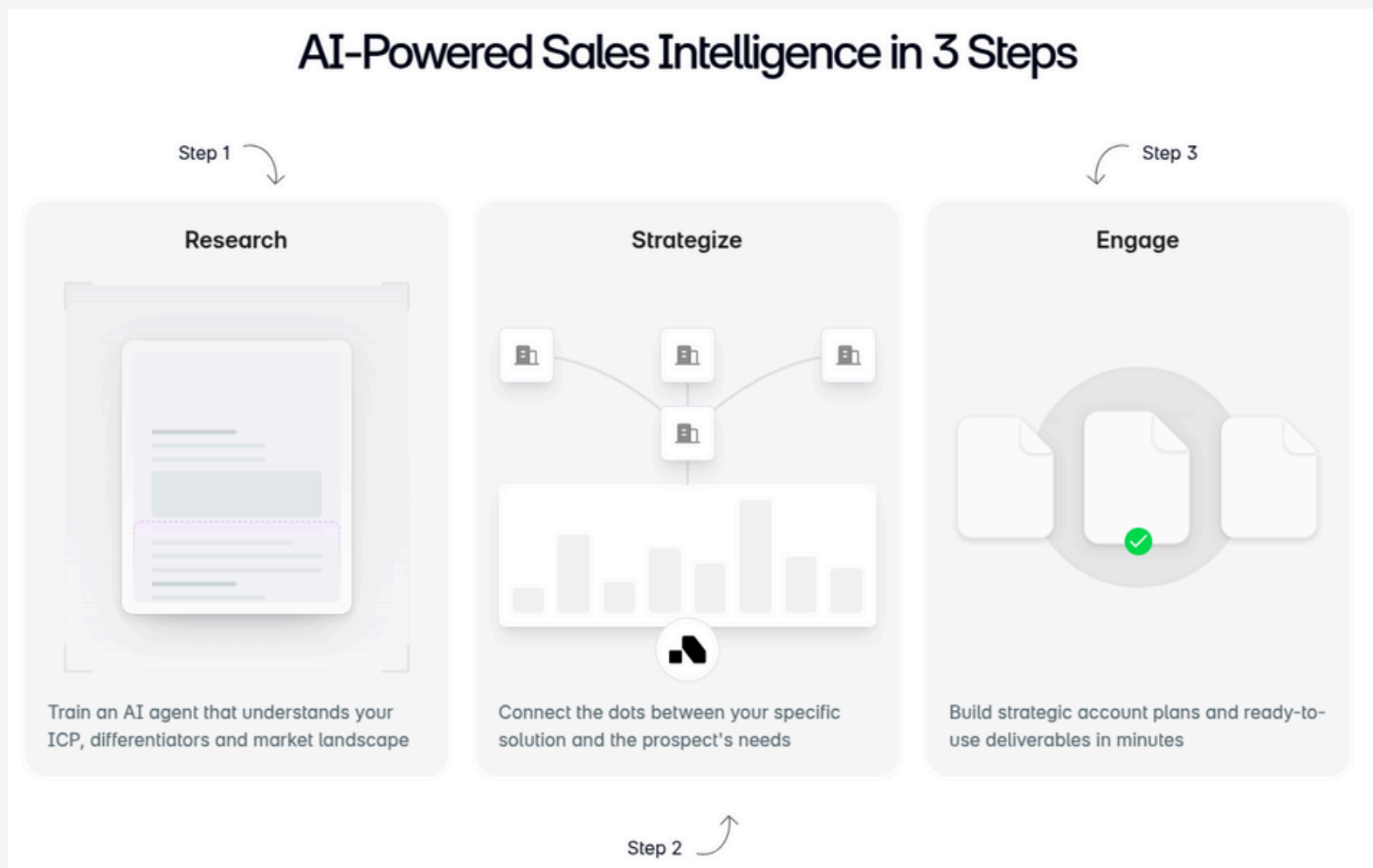
Using natural language processing (NLP), analysts can request complex data extractions in plain English, and Captide's agents handle retrieval, processing, and structuring of vast document sets.
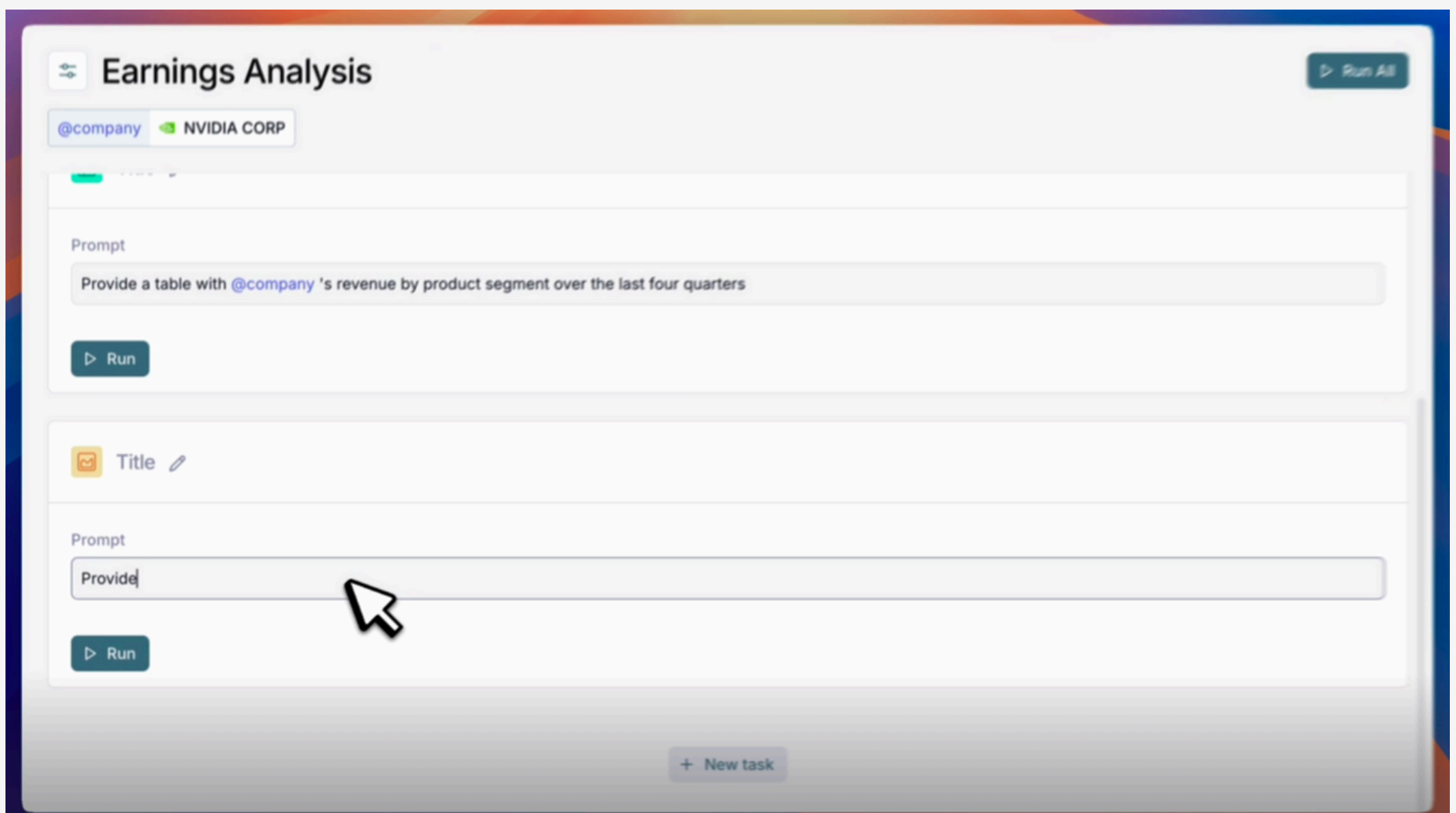
- Parallel Processing with **LangGraph**

- Structured Outputs: With **LangGraph's Trustcall** Python library, Captide ensures that extracted metrics adhere to predefined JSON schemas

- Real-Time Monitoring with **LangSmith**

6. Aomni

Aomni is an AI agent designed to revolutionize B2B sales and research by automating key processes like account planning and prospect discovery.


AI-Powered Sales Intelligence in 3 Steps

Step 1 — Research
Train an AI agent that understands your ICP, differentiators and market landscape

Step 2 — Strategize
Connect the dots between your specific solution and the prospect's needs

Step 3 — Engage
Build strategic account plans and ready-to-use deliverables in minutes

- **Automates Sales Research & Account Planning**
  - Conducts in-depth web research to identify potential customers.
  - Develops Ideal Customer Profiles (ICPs) for targeted outreach.
  - Generates strategic account plans to optimize sales efforts.

- **Trains a Custom AI Agent on Company Data**
  - Learns about your product, market, goals to create personalized sales strategies.
  - Analyzes target companies' challenges, aligns them with your solutions.
  - Generates outreach plans to improve efficiency and precision.

- **Enhances Sales Workflows & Engagement**
  - Saves time by automating prospect research and strategy creation.
  - Ensures scalability, making it suitable for both startups and enterprises.

7. NVIDIA's Eureka Agent

Powered by GPT-4, **Eureka** automates reinforcement learning by writing optimized reward programs, outperforming human-designed ones in 80% of tasks, leading to a 50% performance boost.

Using **NVIDIA Isaac Gym**, a GPU-accelerated physics simulation platform, Eureka iteratively refines its reward functions, enabling self-improvement.
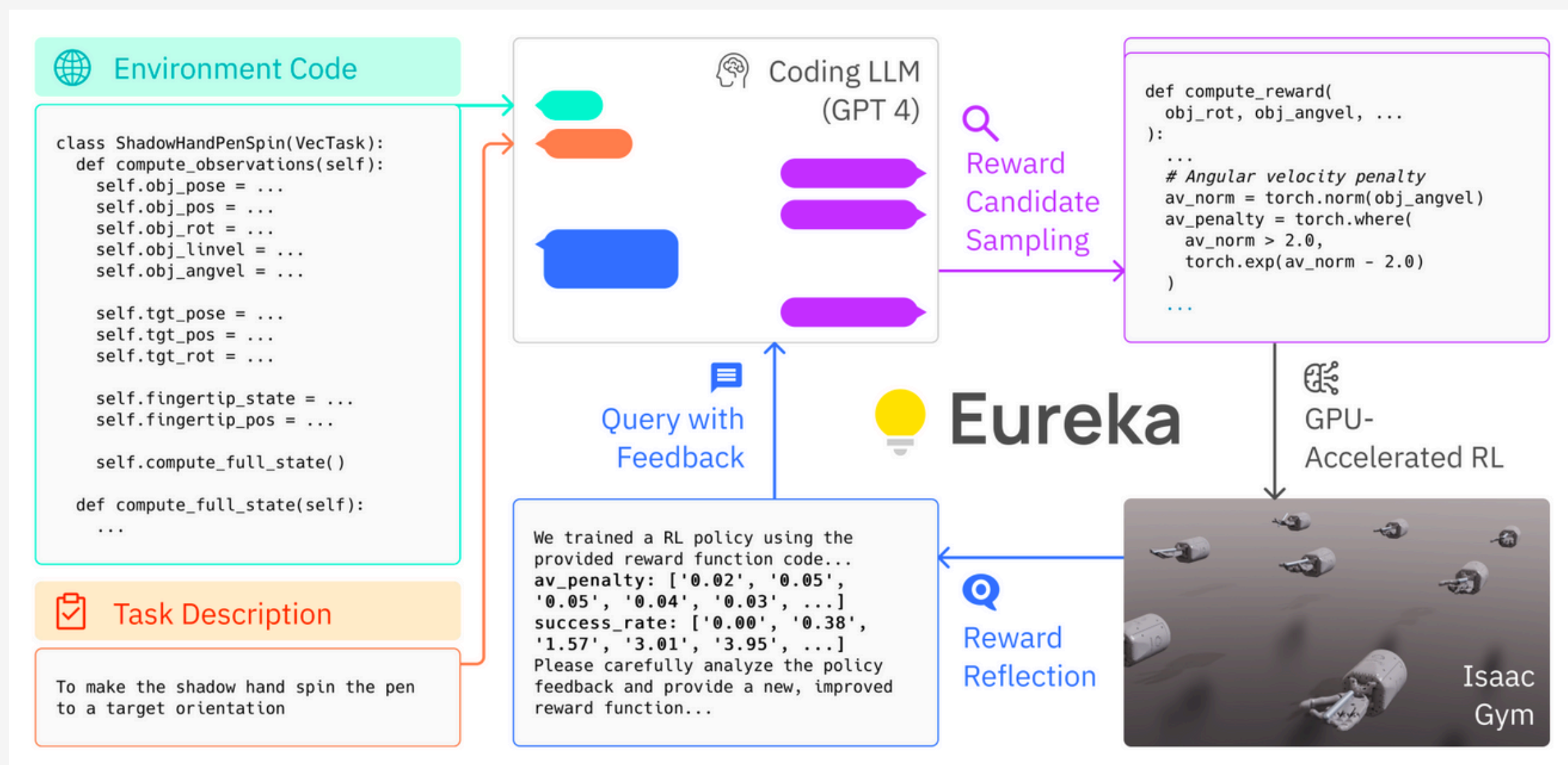
8. Replit Agent

Replit Agent is an AI tool that helps users build software through natural language prompts. It simplifies development by handling prototyping, debugging, and UI creation.

Key Features:

✅ AI-Assisted Development – Automates coding tasks and suggestions.
✅ Interactive Assistance – Users can refine prompts and iterate on AI-generated plans.
✅ Web & Mobile Support – Available on desktop and Replit's mobile app.
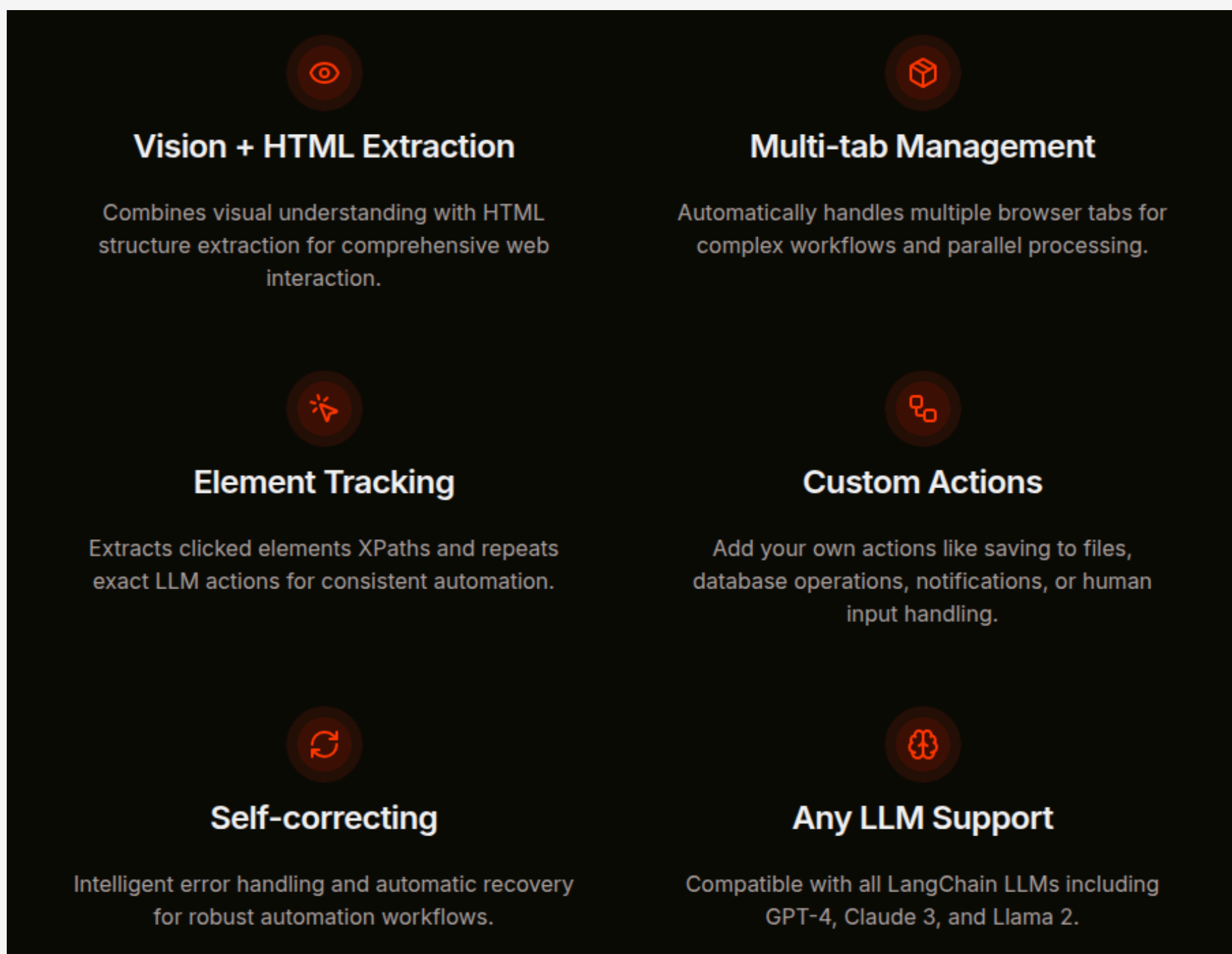
Bhavishya Pandit
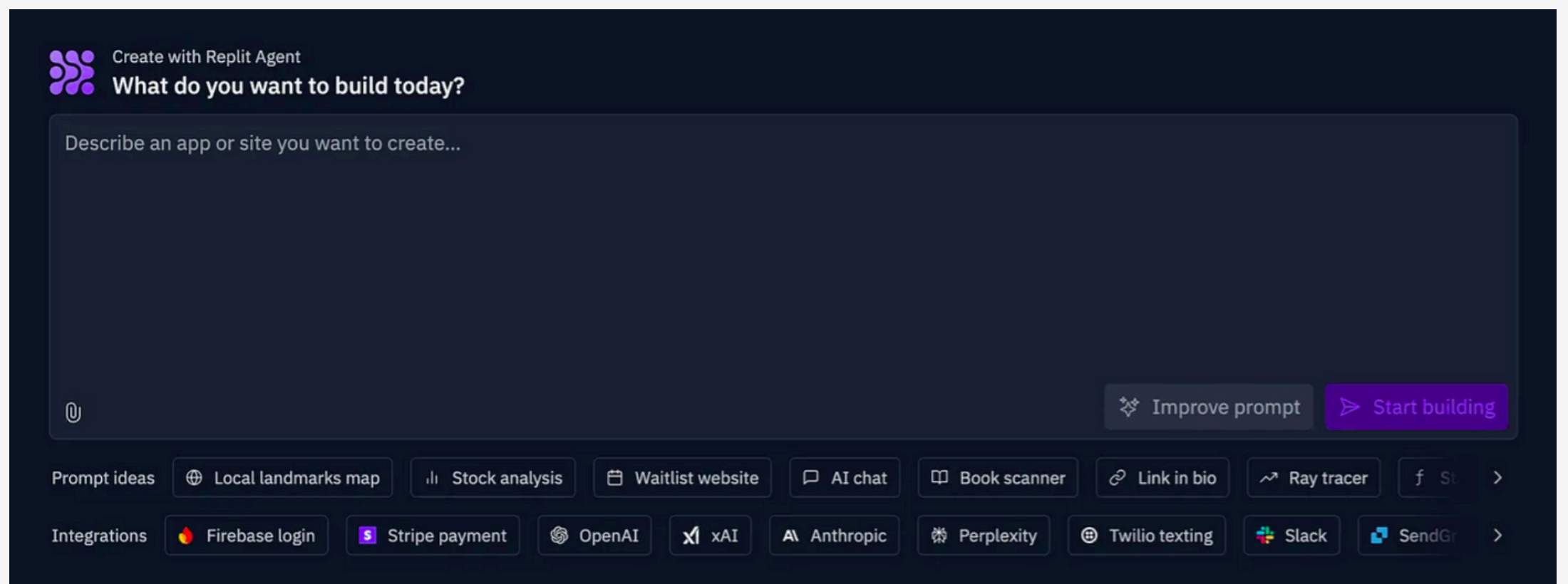
9. Browser Use

Browser-use is the easiest way to **connect your AI agents with the browser.**

Browser Use lets you run agents on the browser to be your online personalty -

You can
1. Write Google Docs
2. Fill Job Applications
3. Do flight bookings etc.

Vision + HTML Extraction

Combines visual understanding with HTML structure extraction for comprehensive web interaction.

Multi-tab Management

Automatically handles multiple browser tabs for complex workflows and parallel processing.

Element Tracking

Extracts clicked elements XPaths and repeats exact LLM actions for consistent automation.

Custom Actions

Add your own actions like saving to files, database operations, notifications, or human input handling.

Self-correcting

Intelligent error handling and automatic recovery for robust automation workflows.

Any LLM Support

Compatible with all LangChain LLMs including GPT-4, Claude 3, and Llama 2.

10. SAP's Agent - Joule

Joule's collaborative AI agents go beyond basic automation by working together across different business functions. They can:

- Reason, Plan, and Execute Tasks Collectively

- Automate Multi-Step Business Processes

- Break Down Organizational Silos

- Leverage Business Context with SAP Knowledge Graph

- Enable Cross-Functional Collaboration

- Improve Efficiency with Extensibility
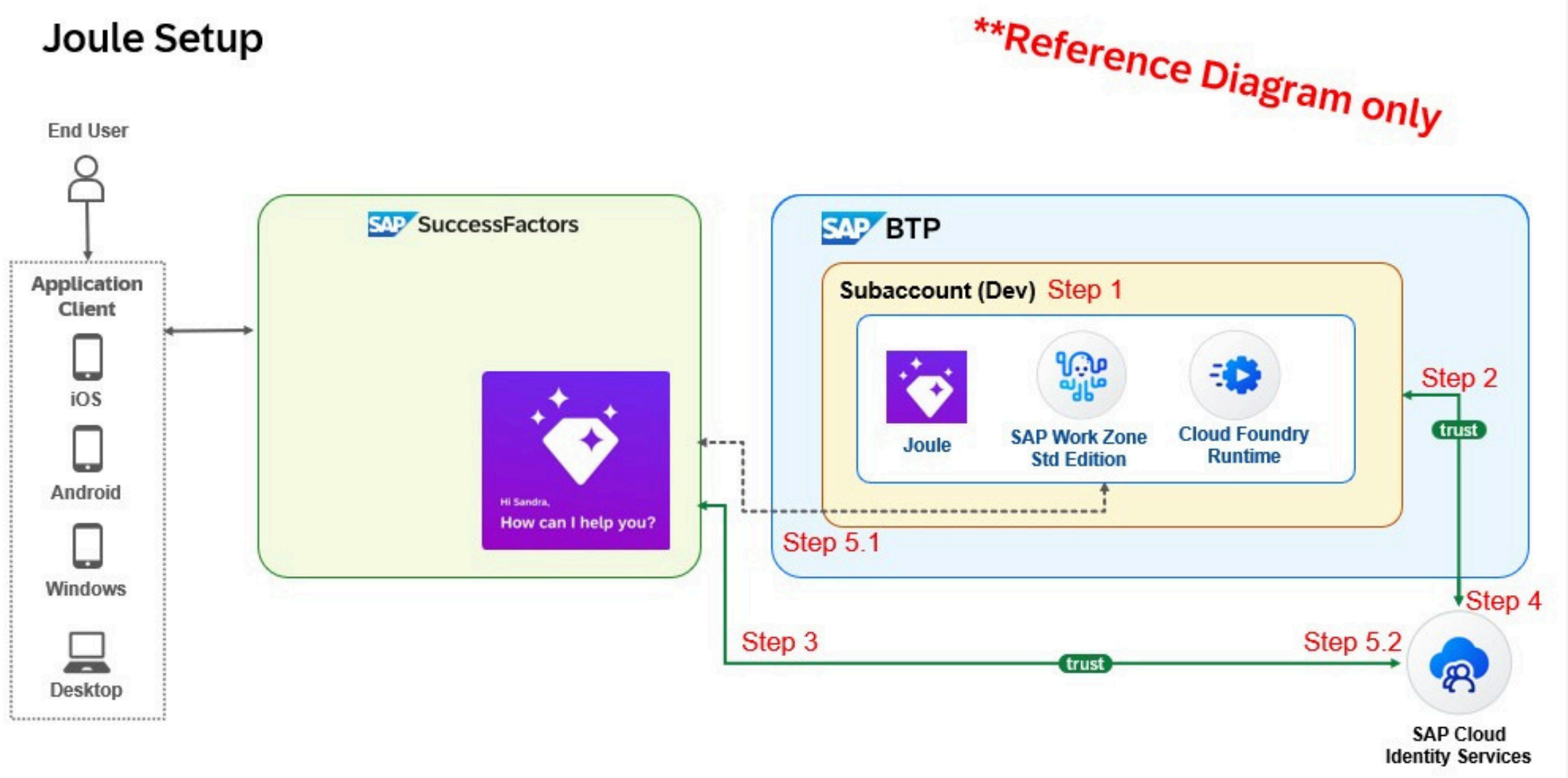
11. Play.ai

Play.ai allows users to build capable AI Voice Agents using a novel approach based on a single **Large Dialogue Model (LDM)**.
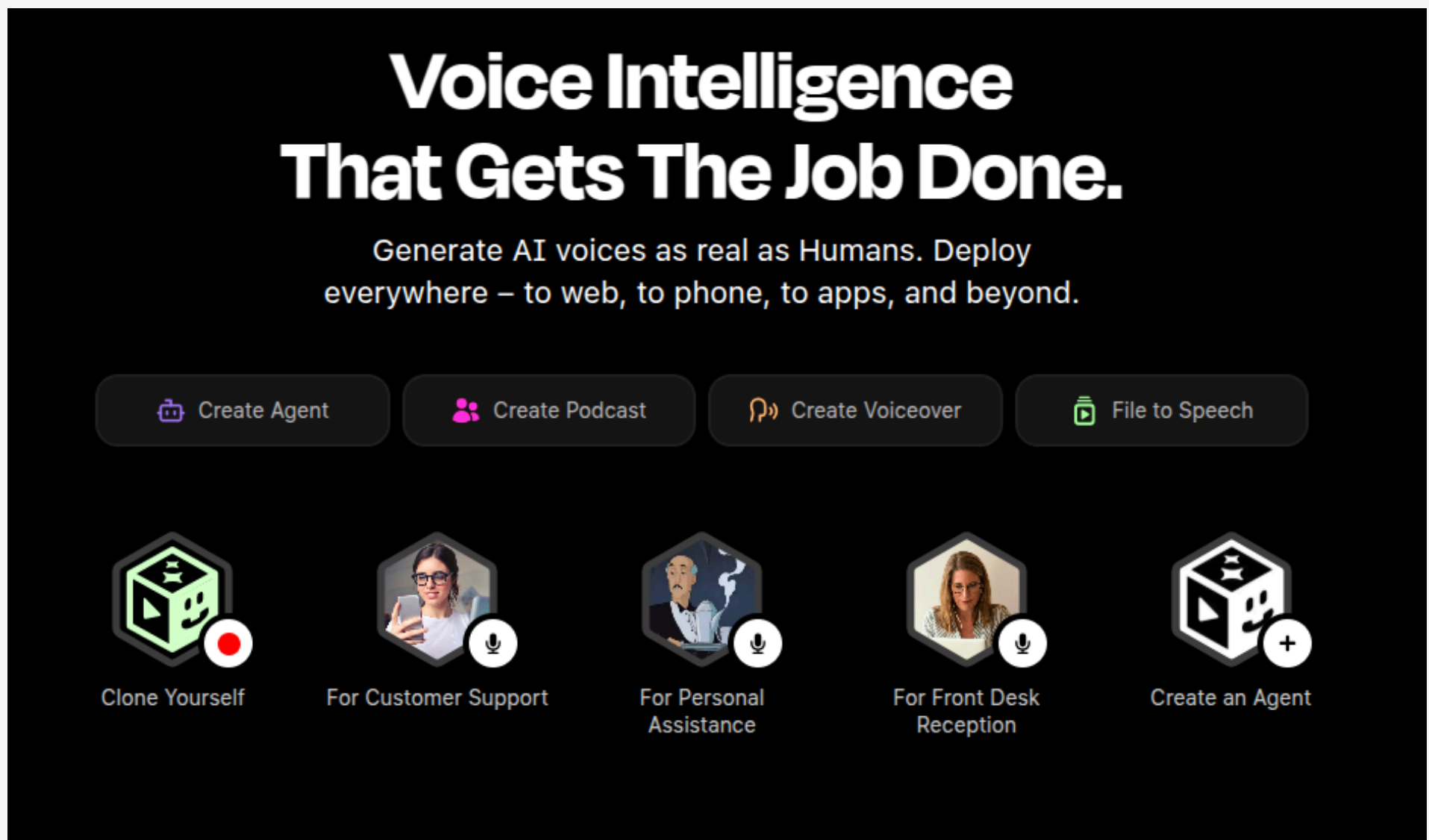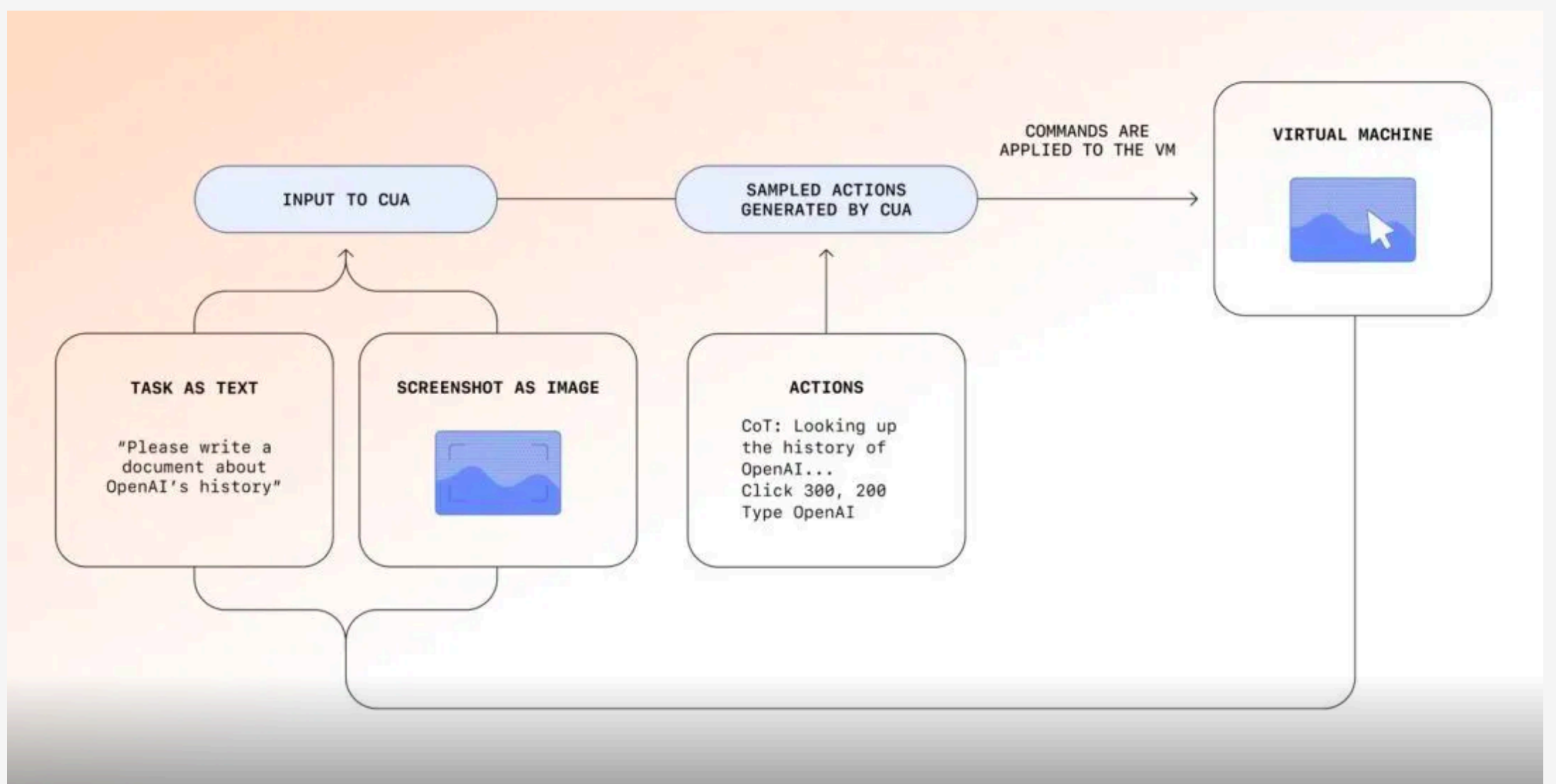

Play AI is used in various applications, including:

- **Sales and Customer Support:** It enables businesses to scale operations with cloned voices for full conversations or automated messages, providing 24/7 customer service.
- **Gaming:** It powers AI characters with expressive conversational voices.
- **Content Creation:** It is ideal for creating AI podcasts, audiobooks, and customized entertainment content.
- **Business Automation:** It helps create AI virtual receptionists and humanlike IVR systems for efficient call management

12. Operator by OpenAI

OpenAI's Operator can automate repetitive browser tasks like filling out forms, ordering groceries, scheduling appointments, and navigating websites using clicks and typing.


INPUT TO CUA

SAMPLED ACTIONS GENERATED BY CUA

COMMANDS ARE APPLIED TO THE VM

VIRTUAL MACHINE

TASK AS TEXT

"Please write a document about OpenAI's history"

SCREENSHOT AS IMAGE

ACTIONS

CoT: Looking up the history of OpenAI... Click 300, 200 Type OpenAI

*Computer Using Agent

CUA achieved state-of-the-art results,

38.1% success rate on **OSWorld**,
58.1% on **WebArena**,
and **87%** on **WebVoyager**

13. Magentic One

Magentic One is built on Microsoft's open-source **AutoGen** framework, allowing it to integrate with various large language models (LLMs) and tools.
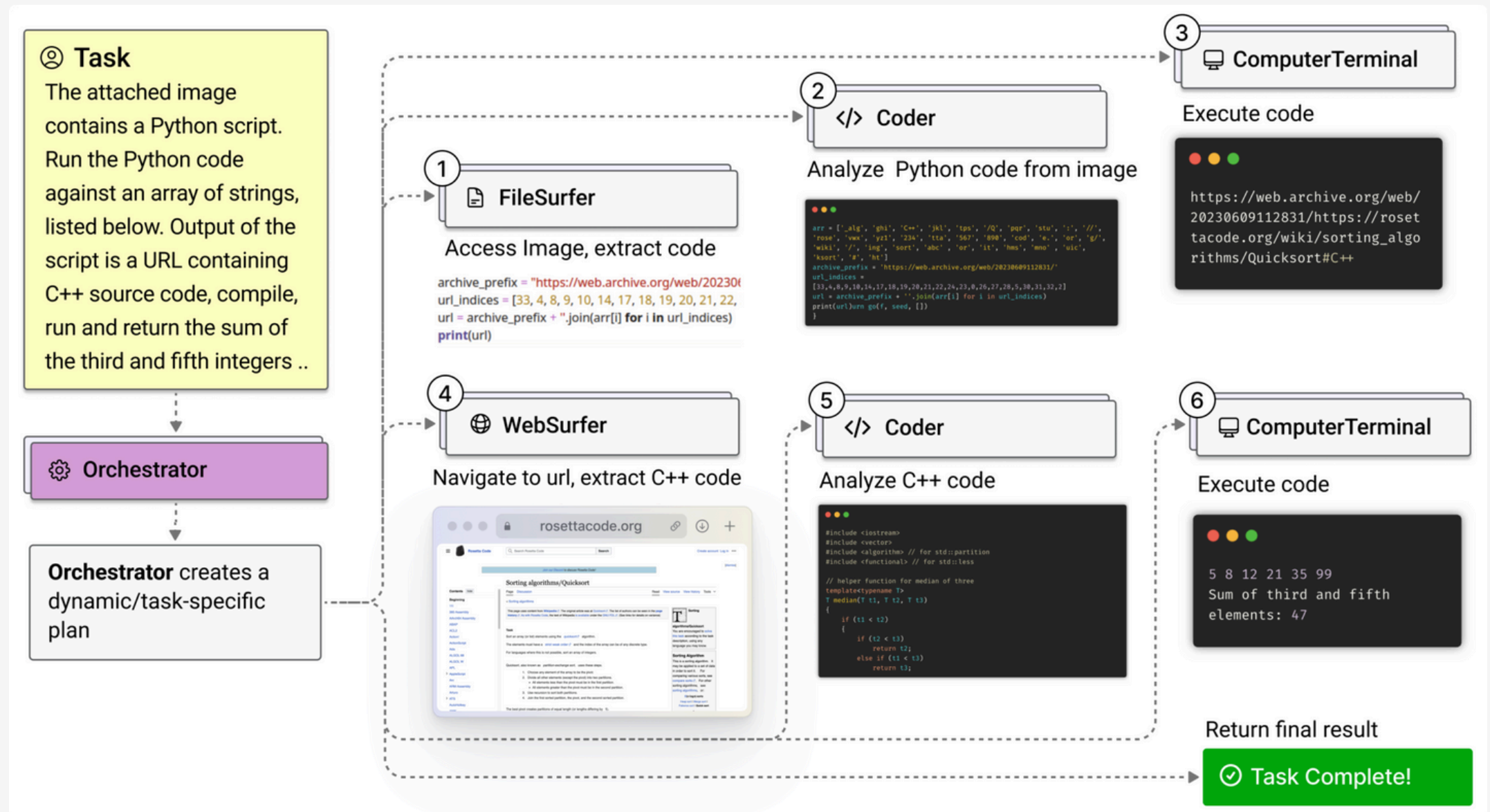
Key Features of Magnetic-One AI:
- **Multi-Agent Architecture:** The system is centered around an "**Orchestrator**" agent that coordinates four specialized agents:

  - **WebSurfer:** Handles web navigation tasks.
  - **FileSurfer:** Manages file operations.
  - **Coder:** Executes coding tasks.
  - **ComputerTerminal:** Performs command-line operations.

- T**ask Automation:** Magnetic-One can automate a wide range of tasks, including booking tickets, purchasing products online, editing documents, and executing Python code.
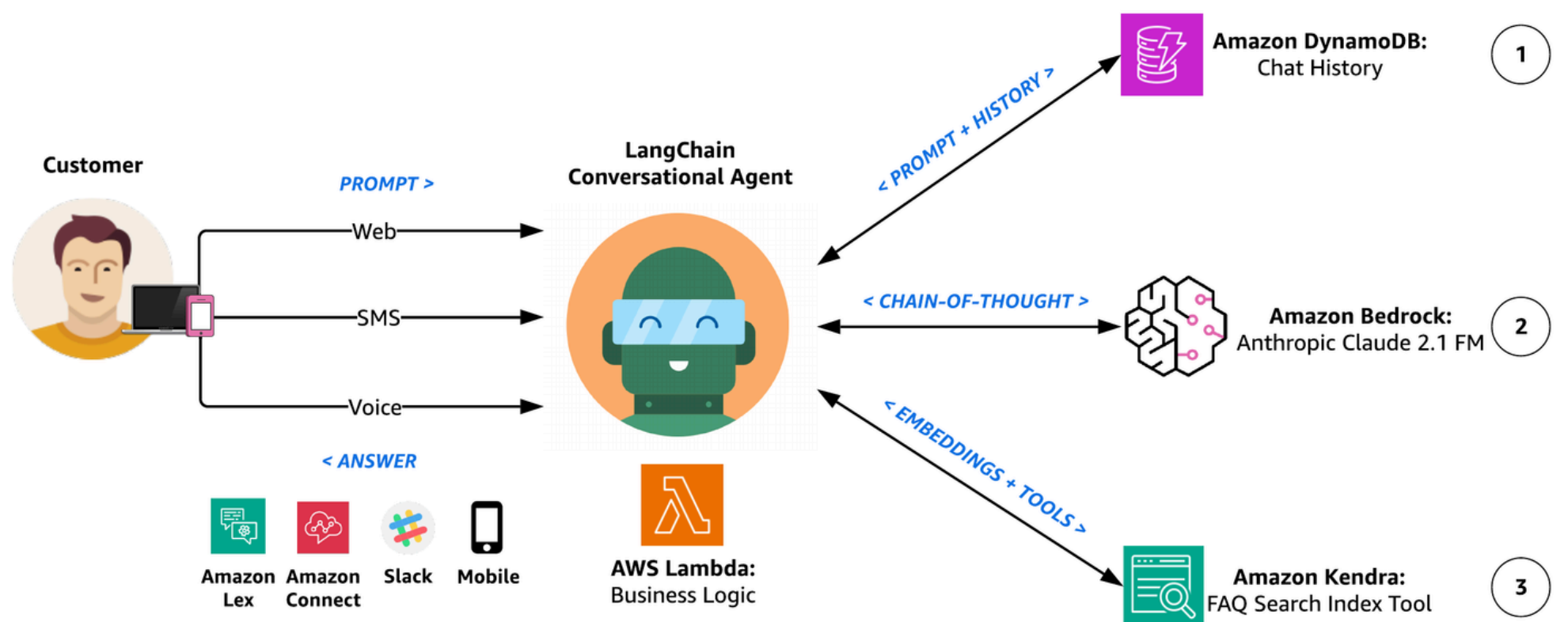
14. Synthflow

Synthflow AI is a platform for creating customizable AI voice assistants.

It offers a no-code interface for automating phone interactions, supporting over 20 languages and integrating with various tools.

Features:

1. No-code Interface
2. 20+ Language Support
3. Advanced Workflow Optimization

**Sample Conversational AI Agent
built using AWS technologies**

15. Waynabox

Waynabox offers surprise trips where travelers learn their destination just two days before departure. To enhance the experience, they teamed up with CrewAI to provide real-time, personalized itineraries.

CrewAI's AI agents:

- Find top activities, restaurants, and tours.

- Adjust plans based on weather and traveler preferences.

- Deliver itineraries instantly via WhatsApp.

SET UP YOUR TRIP

Choose the surprise trip that best suits you. Select the number of travelers, the airport of departure and the dates of your getaway.

REMOVE DESTINATIONS

You will discover the possible destinations and you'll be able to discard the cities you do not want to visit.

SURPRISE!

You will not know your destination until a few days before traveling!

Follow to stay updated on AI/ML

LIKE

COMMENT

REPOST