

the Pitch

O.P.E.R.A.T.O.R

Updated 7 July 2025

Welcome

Introducing O.P.E.R.A.T.O.R the first browser and device automation agent.

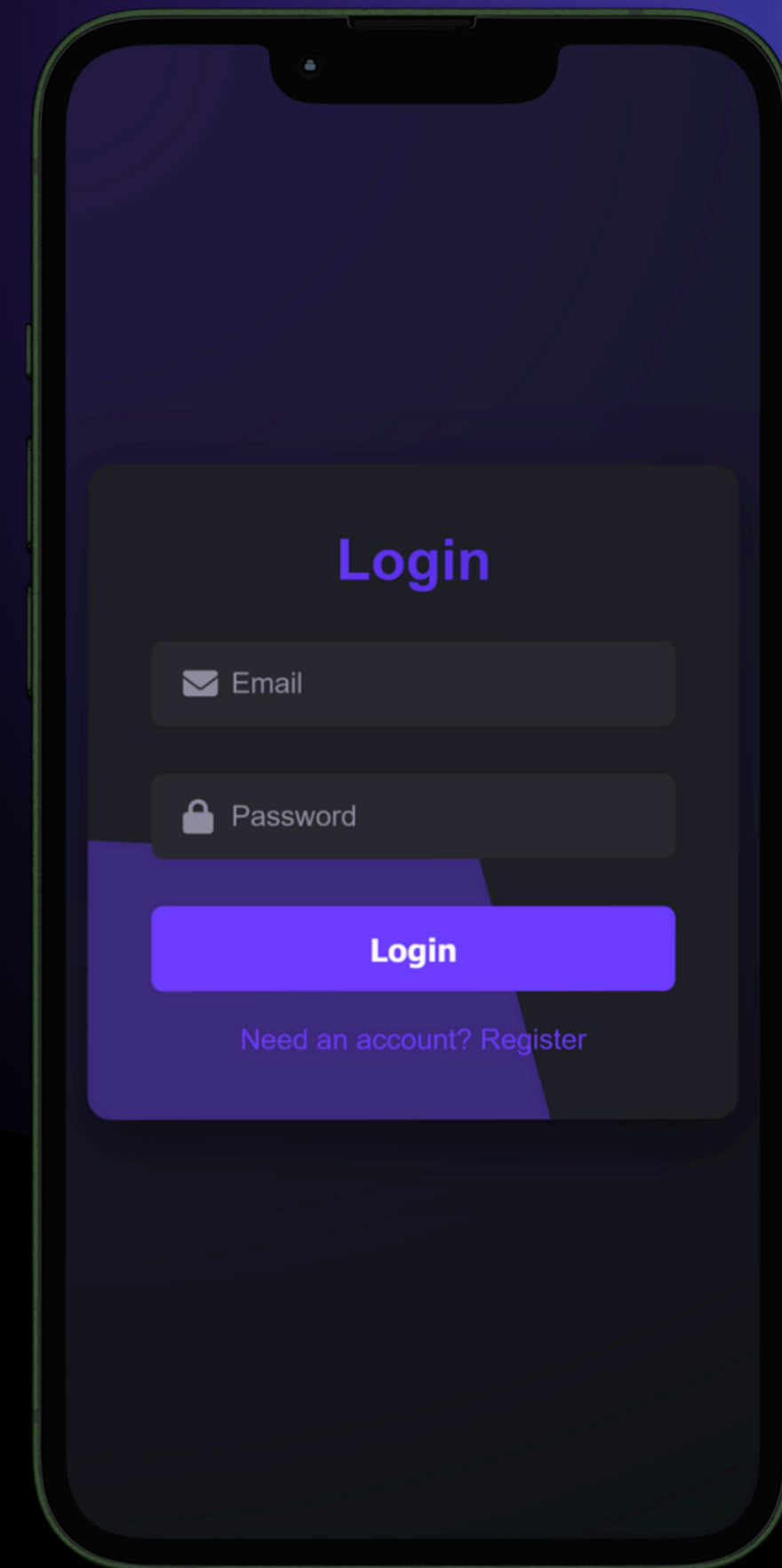
Built on Visual Grounding technology, in one simple interface – seamlessly Chat or Automate tasks across your devices.

Chat & maintain natural language conversations with the Agent.

In the same user friendly manner:

- Run Automated Browser Tasks
- Run Android Device Automation Tasks
- Run PC Control Tasks

We built a Web-App as staging towards an O.P.E.R.A.T.O.R 🤖 Android App on Google Store in Q4 2025 for mainstream adoption.



AI Automation

More than a Chatbot with 5+ models.
Automate browsers, devices, apps and
file access



Browser

- Puppeteer
- Chromium
- LLM/VL Models



Android

- Android ADB
- Puppeteer
- VL Models Only



PC

- VL Models Only
- Localhost App Only
- Remote Access

Why it matters..

AI Automation in our daily lives is inevitable, developers are racing to create the first app.
Here's our Demo Web App

Automation Pillars

Chats

Normal chatbox like any other. Limited memory. Connect up to 5 models of choice. Everything happens in natural language, including android and pc commands.

Task Commands

Instruct the Agent to go some-website and do something. For example: "go to coingecko and search for TAO. Open the chart, switch to candlestick view, zoom out to Max. Predict price in 3 months"

Brainstorm

Briefly elaborate on what you want the agent to think about, in regards to accomplishing a task then go on to do it. It will think and update you on the goal step-by-step prompted correctly.

Intro

what is it?

Operator is a cutting-edge automation tool designed to streamline tasks across various platforms, including browsers and devices, using visually grounded models..

It acts as a sophisticated AI chatbot that can also launch browsers, connect to your phone, or start apps on your PC, offering seamless automation in natural language.

how it works..

OPERATOR relies on Visual Grounded Models to be able to detect, integrate then extract or act based on visual images only.

You prompt in natural language - it acts based on sight.

With the ability to connect up to 5 AI models using your API keys, Operator allows users to maintain normal chats while executing automated tasks efficiently. It provides real-time task execution tracking via websocket events and offers comprehensive plan log reports to review how tasks were accomplished.

join us

For developers, Operator presents an opportunity to install and customize the app, showcasing innovative builds.

Enthusiasts can join the growing community, get early access to the Map Store, and participate in a unique token-based system that facilitates the creation and sale of reusable automation programs.

war path

Funded as a crypto launched project, Operator will pivot into an Android app, enhancing browser and Android automation while connecting to remote PCs and devices—all through natural language commands.

Use Cases

A Powerful Use Case for This Agent: Training and Education

One excellent application of this Agent is for training or educating your followers, team members, clients, and more. Rather than providing manual instructions for them to follow on their PCs, simply create a YAML Map of the task. They can then execute it and observe the Agent in action, demonstrating exactly how the operation should be carried out.

This approach is particularly beneficial for online classes in areas such as design, software management, trading, accounting, and more.

Create a Map, and your audience will be able to see it in action.

Chat



Chat in natural language and maintain continuous memory between tasks.

Shopping



Daunting task delegation, OPERATOR is designed to action plan or step plan towards any task.

Research



Daily life or crypto, the agent navigates Web2 as you would. Leaving you time to do another task as you wait for it to finish its part.

Remote Access

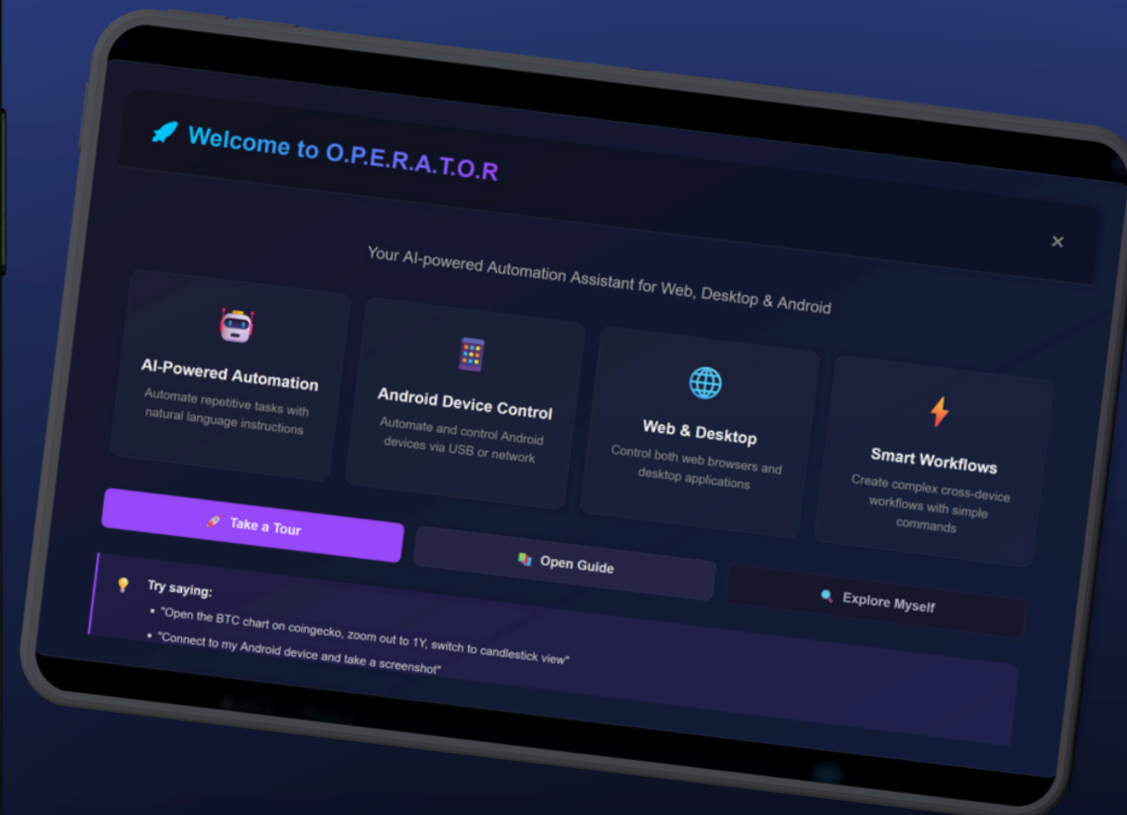
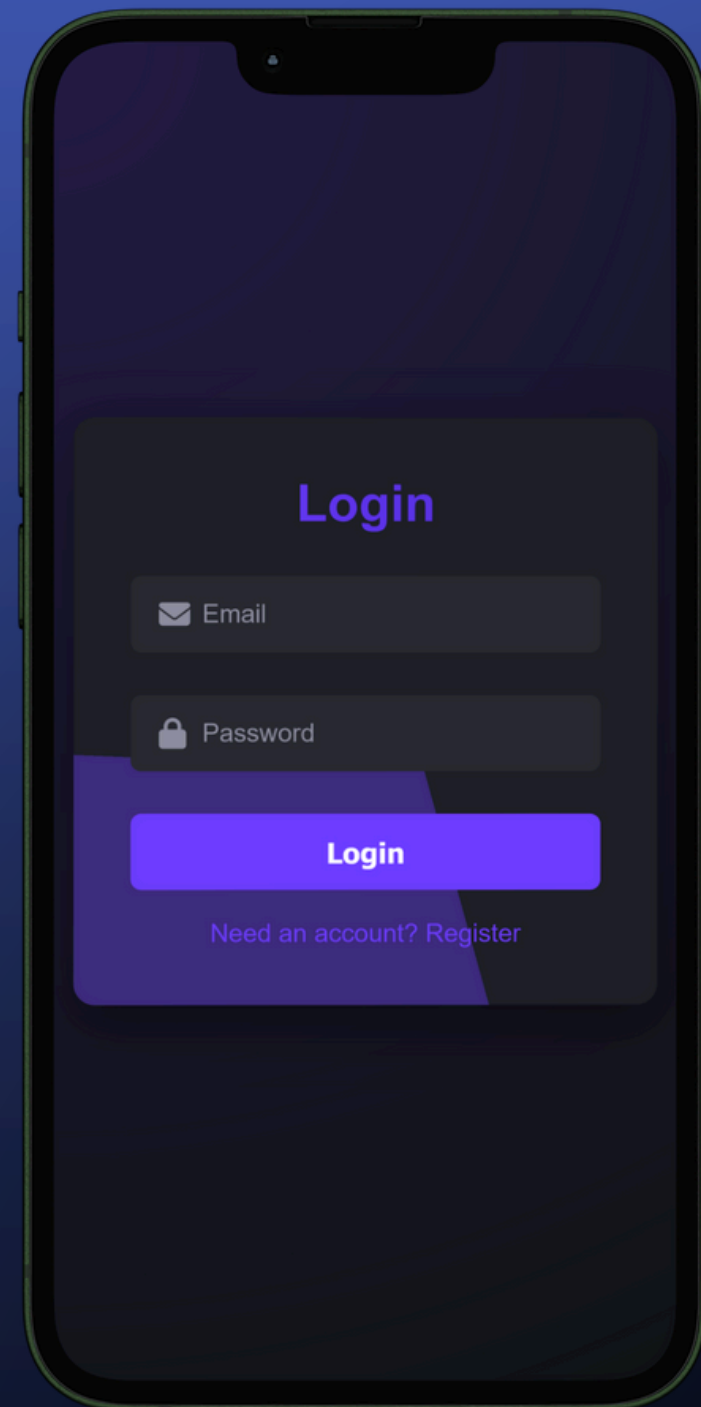


Access your phone or PC by simply login in to the App from anywhere and get things done regardless of location.



↗ Interface





Tip: take the Tour then open the Guide manually.
Make sure to set API keys and engines in settings

Welcome

Sign Up

Create an account using any email

Walk through portal

Let's show you around the App when you login. The app welcomes you with a pop-up modal showing options to Take a Quick Tour, jump to the Guide or discover yourself. We made sure its easy to learn the app.

Home

Message Timeline

All chat messages and task commands appear here, including results.

Command Center

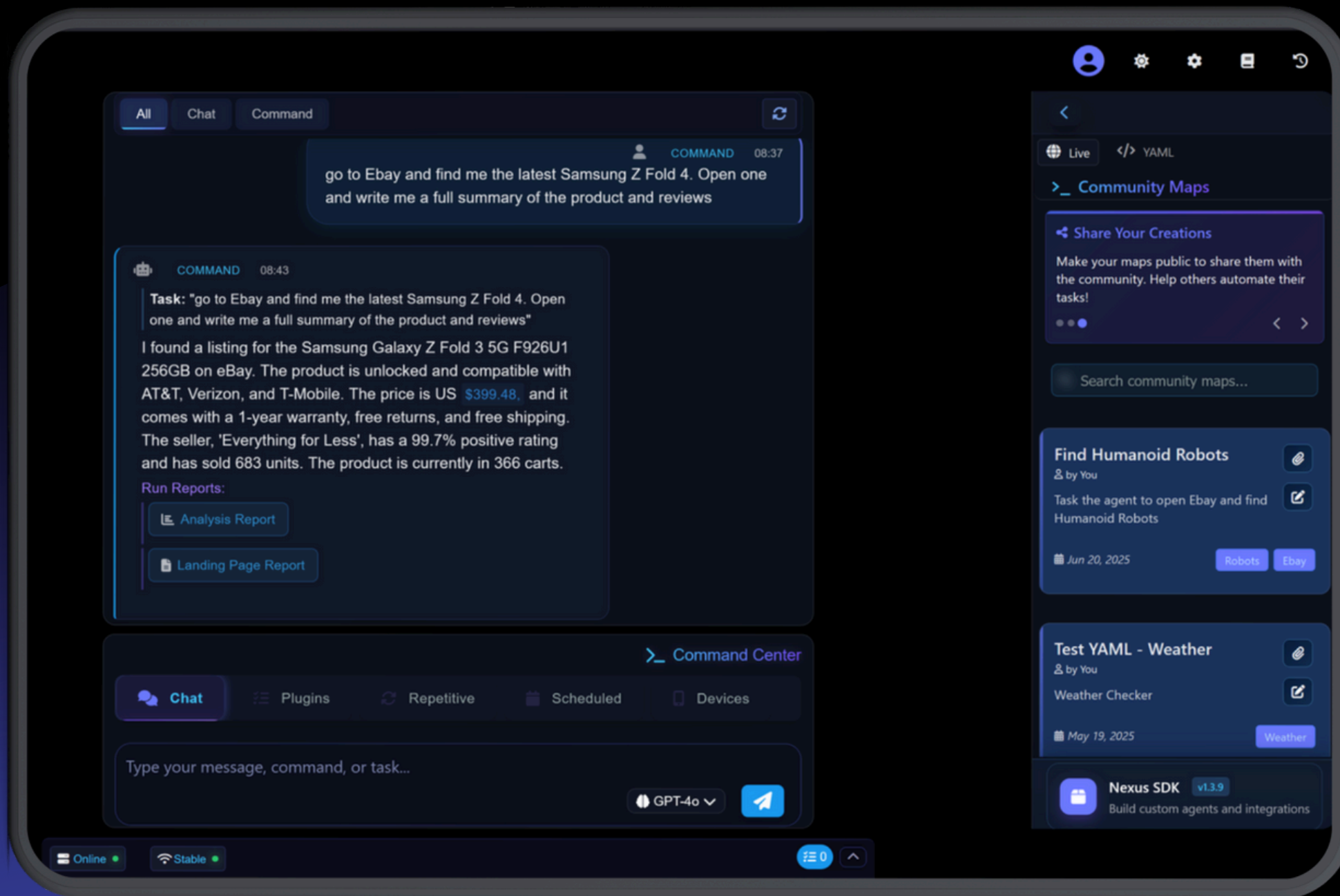
The main control center, Chat is the main tab, navigate to other tabs to explore devices and plugins.

TaskBar

Expandable footer bar to track system health and active tasks.

Sidebar

Active tasks live view. YAML Maps management. SDK download for localhosting



TIP: Use on large screen. Web interface preview. Mobile responsiveness is only fallback not intended interface.

Main Modals

Settings

Configure all your app settings here first before use.

Profile

User profile

Guide

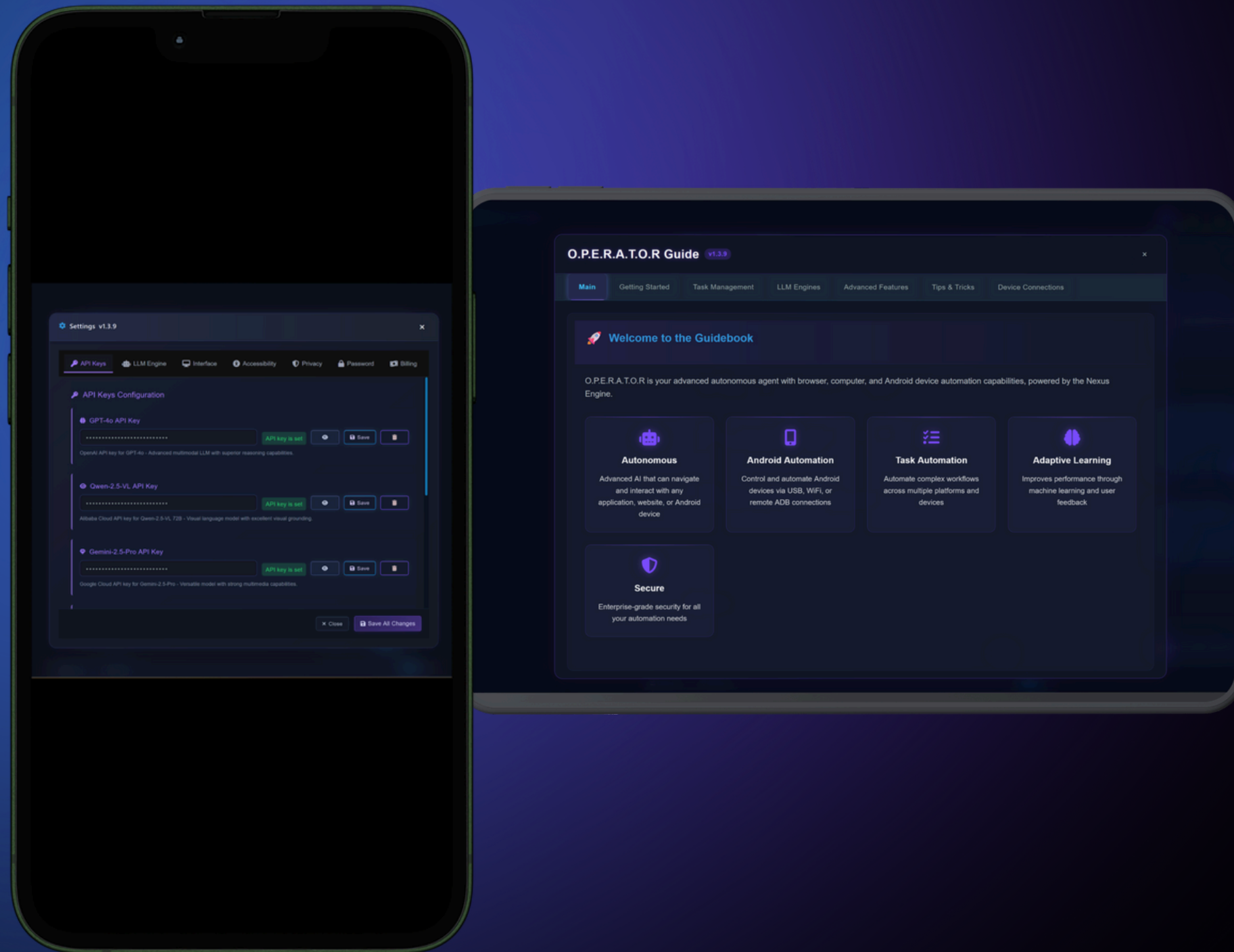
GuideBook with extensive information

YAML Viewer

Create, Edit or View YAML maps

History

Open and rerun past tasks



Tip: single app entry no need to change URLs

Run Tasks

Chat Tab

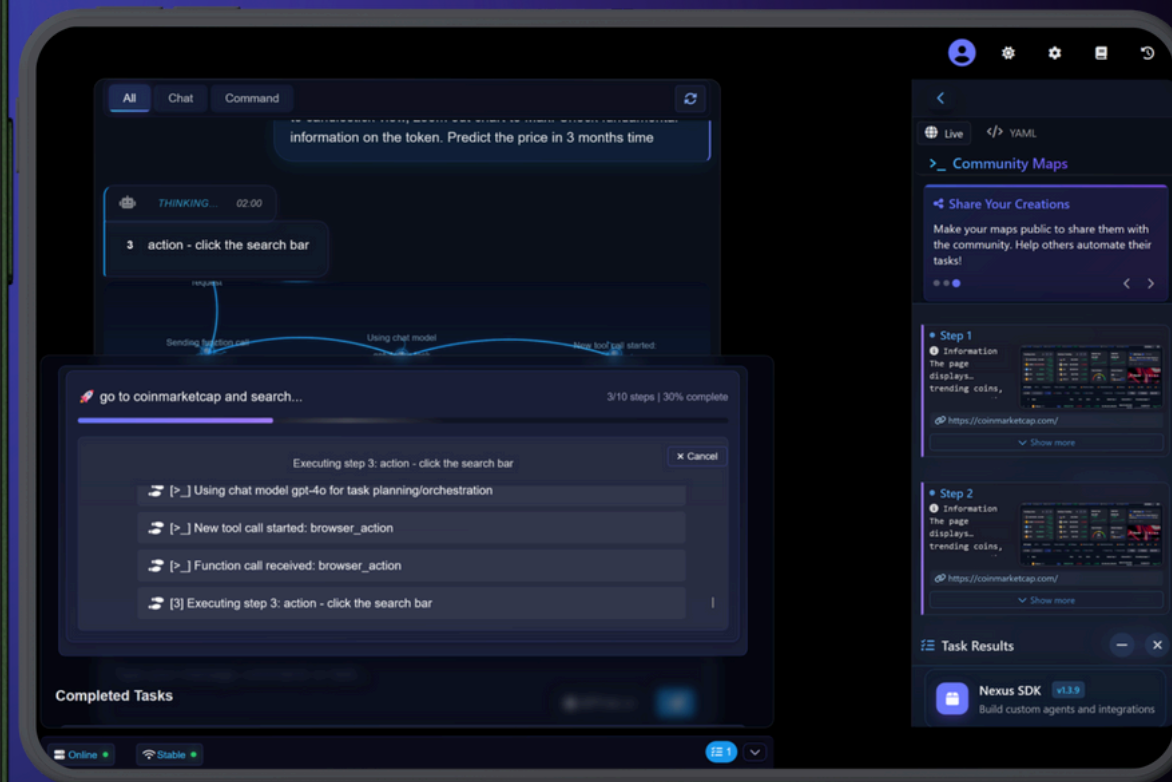
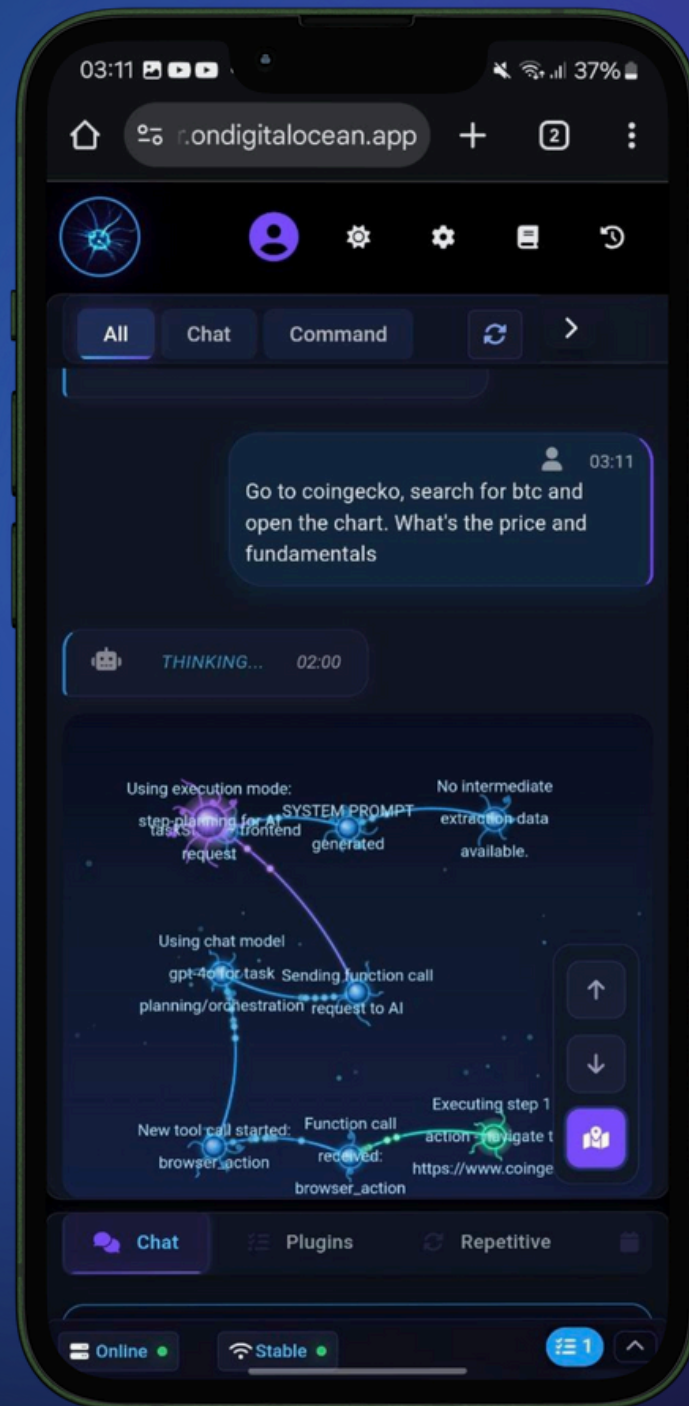
Enter a natural text instruction to go do something, somewhere..

Task Execution

Neural Canvas with step nodes and the action plans behind the scenes outlining the agent's internal thought process towards completing the task.

Watch it Execute

Open the sidebar to see intermediary results & screenshots as it executes steps. TaskBar opens plan logs.



Tip: If Neural Canvas, Plan Logs and Screens are not updating during a task, logout and login to re-establish a websocket connection. If your browser loses connection a task may fail.



O.P.E.R.A.T.O.R – Sentinel Report

[Replay](#)

Task Execution Report

Run ID: f213cba1-d4e3-4521-8892-0b584b04773f

Timestamp: 2025-07-04T16:20:50.711Z

[Task Details](#)

Command:

open coingecko and search for BTC, open the chart. check the rprice

[Task Plan](#)

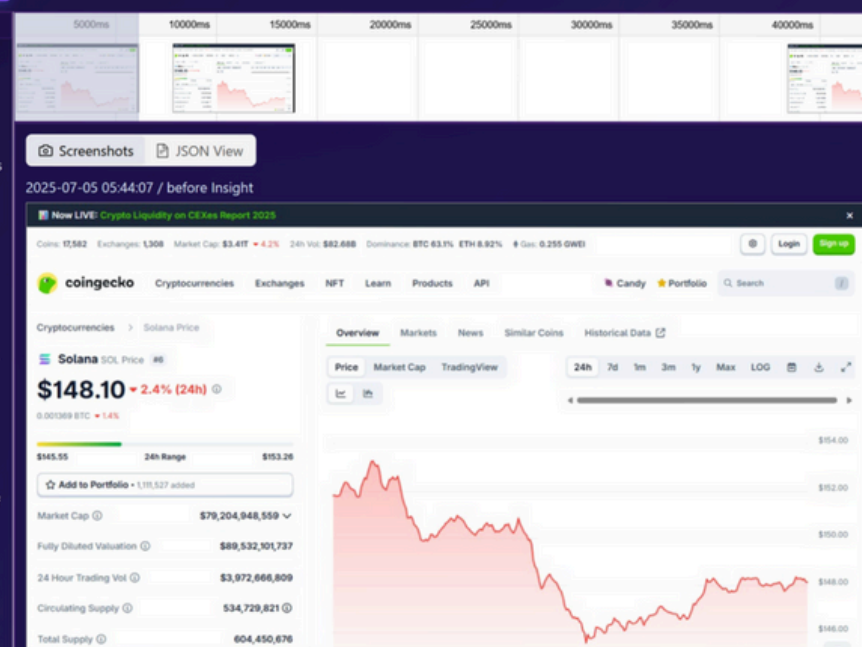
1. **COMPLETED**

navigate to <https://www.coingecko.com>

```
{ "success": true, "error": null, "task": "navigate to https://www.coingecko.com/en/coins/solana", "stepIndex": 0, "action": "navigate", "pageTitle": "Google Search", "I'm Feeling Lucky": "Sotho", "About": "Advertising", "Business": "Business", "timestamp": "2025-07-04T16:13:00.200Z", "url": "https://www.coingecko.com/en/coins/solana", "validated URL": "https://www.coingecko.com/en/coins/solana", "browser": "Chrome", "truncated": true }
```

O.P.E.R.A.T.O.R – Sentinel Report

After executing "navigate to https://www.coingecko.com/en/coins/solana", thoroughly analyze the page and return a JSON object with the following structure: { "main_content": "Describe the main content visible on the page (listed information, products, tokens, prices, titles, important information).", "navigable_elements": ["List ALL clickable and navigable elements with their EXACT text as shown on screen."], "interactive_controls": ["List ALL interactive controls (sliders, toggles, filters, etc.) with their EXACT labels if visible."], "data_visualization": ["List ALL chart controls, time selectors, indicator buttons with their EXACT labels. Detail chart type (line or graph)".], "product_filters": ["List ALL product filtering options with their EXACT labels."], "search_fields": ["List any search fields or input areas with their placeholder text."], "pagination": "Describe any pagination controls." } CRYPTO SPECIAL INTERFACE DETECTED (e.g., DexTools, Dexscreener, Coinbase, Coingecko, Coinmarketcap, Jupiter Exchange): - Note the side menus, top navigation bars, and dashboard sections. - Identify buttons such as "Trade", "Charts", "Market", "Analysis". - Include any filtering dropdowns, time frame selectors, and graph toggles. - List any visible token names in a list, token labels or information links. Always outline in Great Detail the main page content on center of page. Always read Key information in the sidebars if present. For news and documents focus on main content. For products and cryptocurrency focus on Description, Symbols, Names, Prices, Lists, and any other relevant information. IGNORE ALL IMAGES of phones, laptops, devices, billboards, or any marketing images simulating data presentation. Describe charts or graphs including Chart heading, chart type (line/bar/candlestick), time frame, and



View Reports

Task Complete Card

Every task creates a Task Complete card with two report links.

Analyze Reports

Landing Report – provides all Plan Logs, Screenshots and steps in raw format up to Task Complete.

Nexus Report – provides an in-action replay of the AI Agent at work behind the scenes and all the visual analysis & thought process .

Replay Nexus Report

Watch a first person view of the AI Agent in action.

Tip: analyze the completion data in all portals. Two Report links are provided in the Task Completion Card. View reports

Program Tasks

YAML Scripts

Create repeatable scripts for automation in a simple structured text format, with instructions or steps to accomplish a specific task.

Community Maps

To create, open the Sidebar > YAML tab. Click create to open the YAML Editor and follow instructions. To share your map, click the checkbox before saving.

Run a Map

After saving your map, you will see it appear in the sidebar. If public, anyone can test it. To use it, click the Pin button/icon to attach it to the chat input, press send or enter to run it.

Find Humanoid Robots

Updated: 20/6/2025 | Public | Used 3 times

How to Use This Map

To use this YAML Map in a command, type the following in the command input:

```
/yaml 68539780c2d58dd3c6f6fb32
```

> Attach to Command Center

Task the agent to open Ebay and find Humanoid Robots

Target URL: <https://ebay.com>

YAML CONTENT

URL: <https://ebay.com>

```
1 11 tasks:
2 22   - name: Find Humanoid Robots
3 33   flow:
```

Edit YAML Map

Name *

Find Humanoid Robots

Description

Task the agent to open Ebay and find Humanoid Robots

URL (The website your YAML map will automate)

<https://ebay.com>

Tags

Add tags...

Robots

Ebay

☒

Make this YAML map public

Cancel

Save

YAML Editor 1.0.0

```
1 tasks:
2   - name: Find Humanoid Robots
3     flow:
4       - ai: input 'Humanoid Robot' in input box, click search button
5         - sleep: 3000
6       - ai: scroll down, click result with Robosen K1 Pro to open product page
7         - sleep: 3000
8   - name: query product
```

Preview

```
1 tasks:
2   -name: Find Humanoid Robots
3   flow:
4     -ai: input 'Humanoid Robot' in input box, click search button
5     -sleep: 3000
6     -ai: scroll down, click result with Robosen K1 Pro to open product page
7     -sleep: 3000
8   -name: query product
9   flow:
```

Format Requirements

Format YAML Online

Your YAML must follow this structure to work with agent.runyaml() :

```
tasks:
  - name: your_task_name
    flow:
      - ai: your instruction here
      - aiQuery: "query to extract data as {field: type}"
```

Tip: use the YAML checker to verify your text file format. Create useful automation programs or maps for common task & monetize it later.

Connect Devices

Carry out tasks across devices

Do more in your daily tasking – by connecting our Agent to your devices on Android or PC.

Android Automation

WebApp >> Android Mobile device connection through websockets allowing the agent to launch and run apps on your phone through USB or Network. Our Android App release support Mobile >> WebApp & PC connections.

PC Automation

WebApp Agent to PC control allows both App and File access to the agent enabling it to autonomously access all resources it needs for a Task.

Tip: Download our npm package and run it locally to Beta test the device connection features. Fullstack app package and SDK coming.

Choose Models

Visual Language Models/ VLM

Visual language models are built on visual grounding not text and can navigate UI with excellent comprehension.

Large Language Models/ LLM

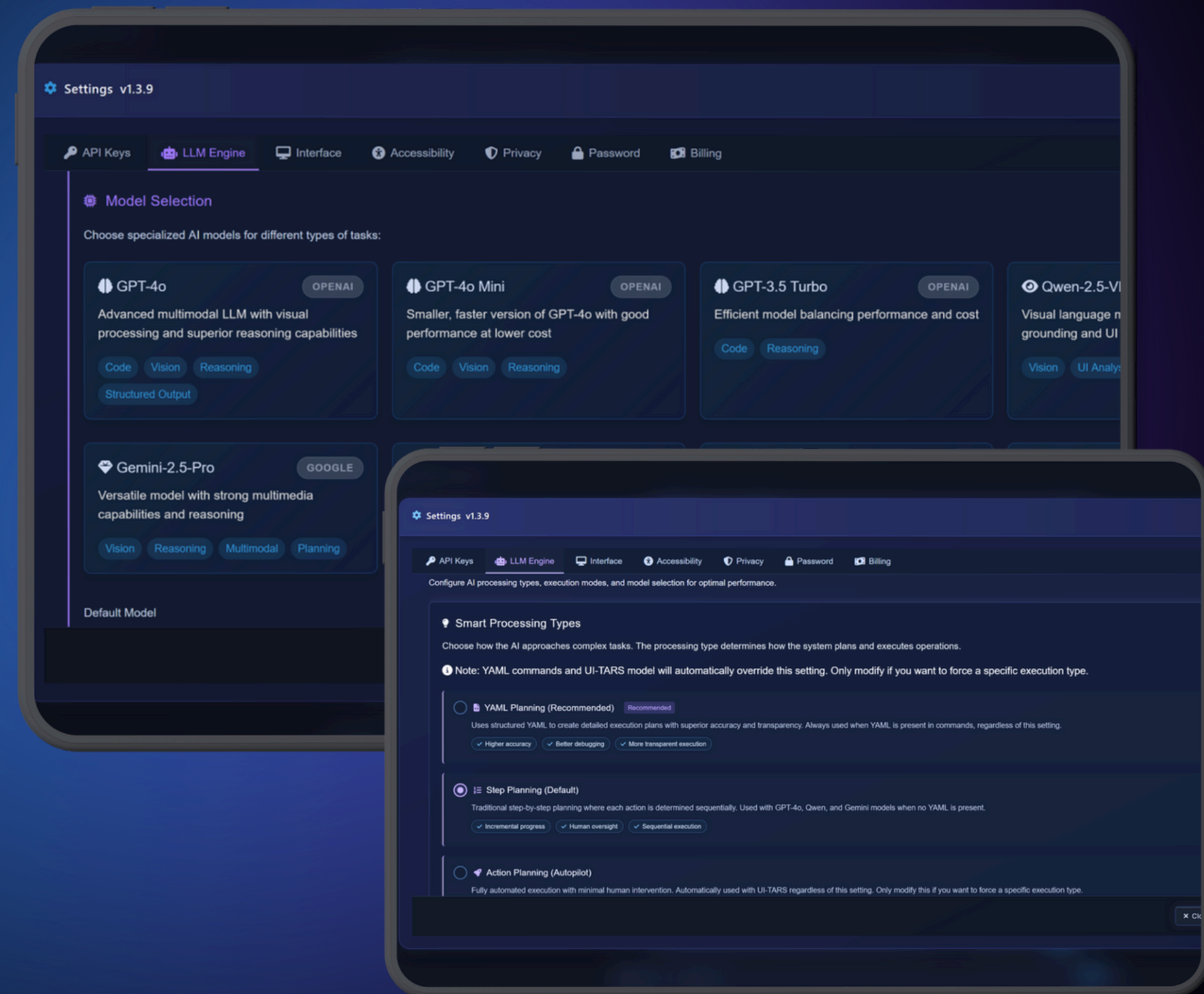
You can also select Advanced multimodal LLMs like GPT-4o with visual processing and superior reasoning capabilities.

Execution Modes

Select correct processing mode for each model. e.g. select VLMs only for Auto Planning.

- Step Planning – AI plans steps one at a time, and executes towards the goal.
- Auto Planning – AI executes tasks autonomously on vision only.

Tip: use GPT40 for basic testing and cheaper costs on API tokens. VLMs: Qwen-2.5-VL 72B, Gemini-2.5-Pro & UI-TARS cost more but perform better.



Achitecture

Frontend

React, some Vanilla JS & playful Three.js

Application Layer

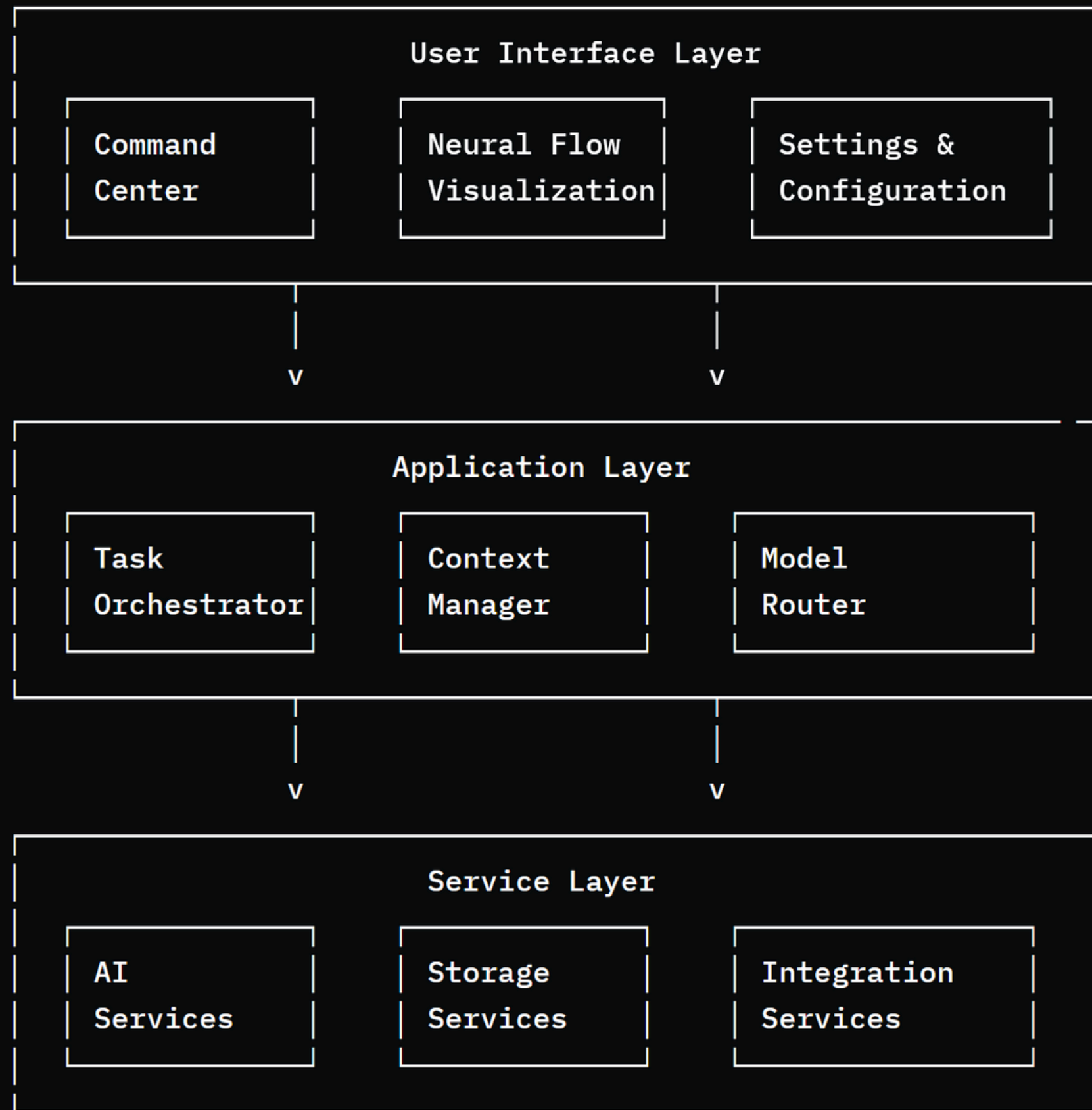
Express Server with Websockets and Sessions, serving the /nli routes and process handlers for Task to Model routing.

Service Layer

Built on **ByteDance's UI-Tars** model
<https://github.com/bytedance/UI-TARS>

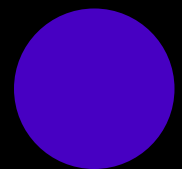
Built on **OpenAI CUA** model
<https://openai.com/index/computer-using-agent/>

Puppeteer, Chromium integrated to execute the commands. Report generation and data storage in MongoDB.

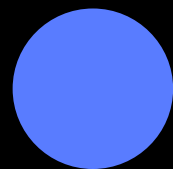


Growth Calendar

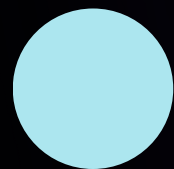
2025 - Q3	Product & Crypto Token Launch	Plugin integration, SDK Release	Map Store Release
2025 -Q4	Android App Build	Local Agent/Android Extension	PC Control Extension
2026 -Q1	Android App Google Store listing	Main Stream Marketing for everyday use	PC App Build



Quarter 3



Quarter 4

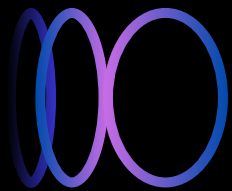


Quarter 1

Note: Development calendar remains on-track regardless of funding challenges.

Funded through crypto - We are building towards a mainstream use AI Agent operable on Android Apps with native android control, remote access to PCs and other devices.

Competition



Our competition to create cross platform automation apps is no joke.

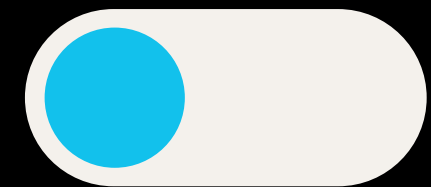
OpenAI and Google are both actively pursuing PC and mobile apps respectively.

This still leaves room for middle market players to develop apps for mainstream use before they do, or in the same timeframe.

OpenAI
Operator

CODEC
Ambiguous

Google
ProjectAstra





About Us

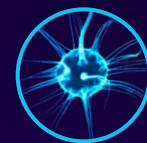
Creators of:



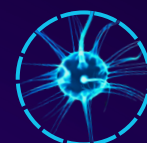
KATZ! Telegram interface Crypto AI Agent



SENTINEL Nodejs API backend feed of KATZ



O.P.E.R.A.T.O.R NodeJs+React WebApp



O.P.E.R.A.T.O.R Android App (coming)

Building AI Agents for Cross Platform Task Automation

Resource Page

Find the magic links

[Old Website](#)

[Official Website](#)

[GitBook](#)

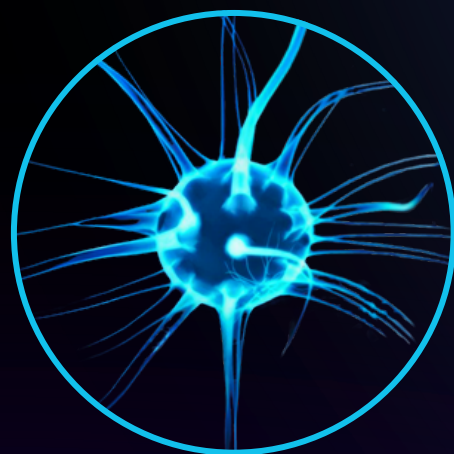
[Medium](#)

[X](#)

[Youtube](#)

[LinkTree](#)

[Demo](#)



END