TRAINING 4 HRS

AI AGENTS

Dr. Jon Krohn

Co-founder and CEO
Y Carrot 🥕

ML Practice Fellow
Lightning AI ⚡

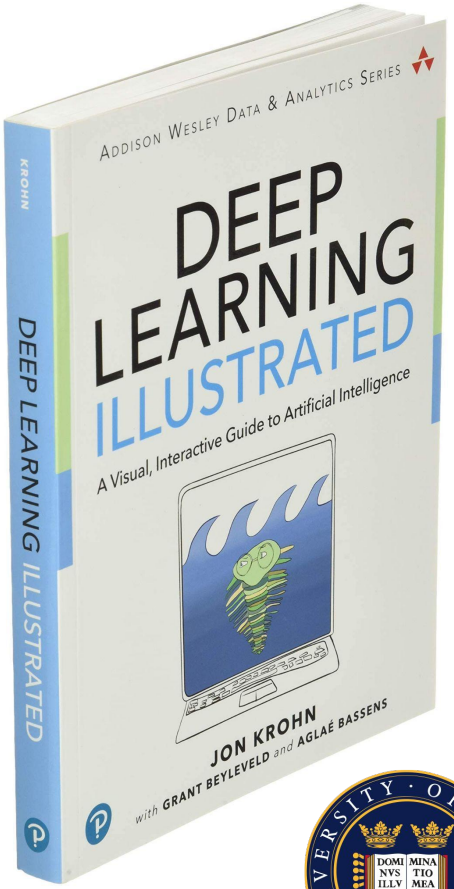
Edward Donner

Co-founder and CTO
Nebula.io

Agentic AI in Action: Build Autonomous, Multi-Agent Systems Hands-On in Python

ODSC EAST 2025
THE LEADING AI BUILDERS CONFERENCE

MAY 13-15
BOSTON



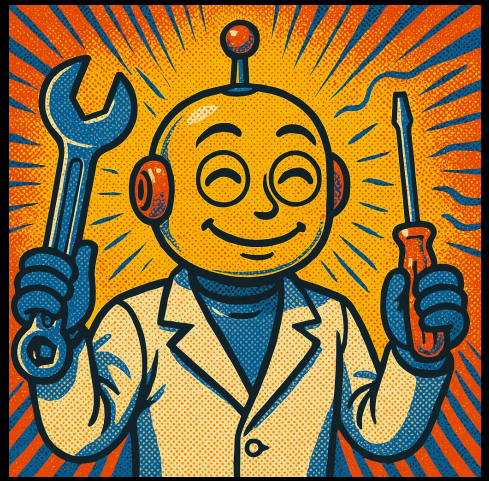
O'REILLY

Bloomberg Television

Module 1:
Defining Agents

Module 1 Coding:
Deep Research
with OpenAI Agents SDK

Module 2:
Designing Agents



Module 2 Coding:
Engineering Team
with CrewAI

Module 3:
Developing Agents

Module 3 Coding:
Autonomous Traders
with MCP

Module 1: **Defining** Agents

Module 1 Coding:
Deep Research
with OpenAI Agents SDK

Defining Agents (it’s ambiguous…)

"AI Agents are **programs where LLM outputs control the workflow**."

In practice, describes an AI solution that involves any or all of these:

1. Multiple LLM calls
2. LLMs with ability to use Tools
3. An environment where LLMs interact
4. A Planner to coordinate activities
5. Autonomy

Episode #841

Andrew Ng

SuperDataScience

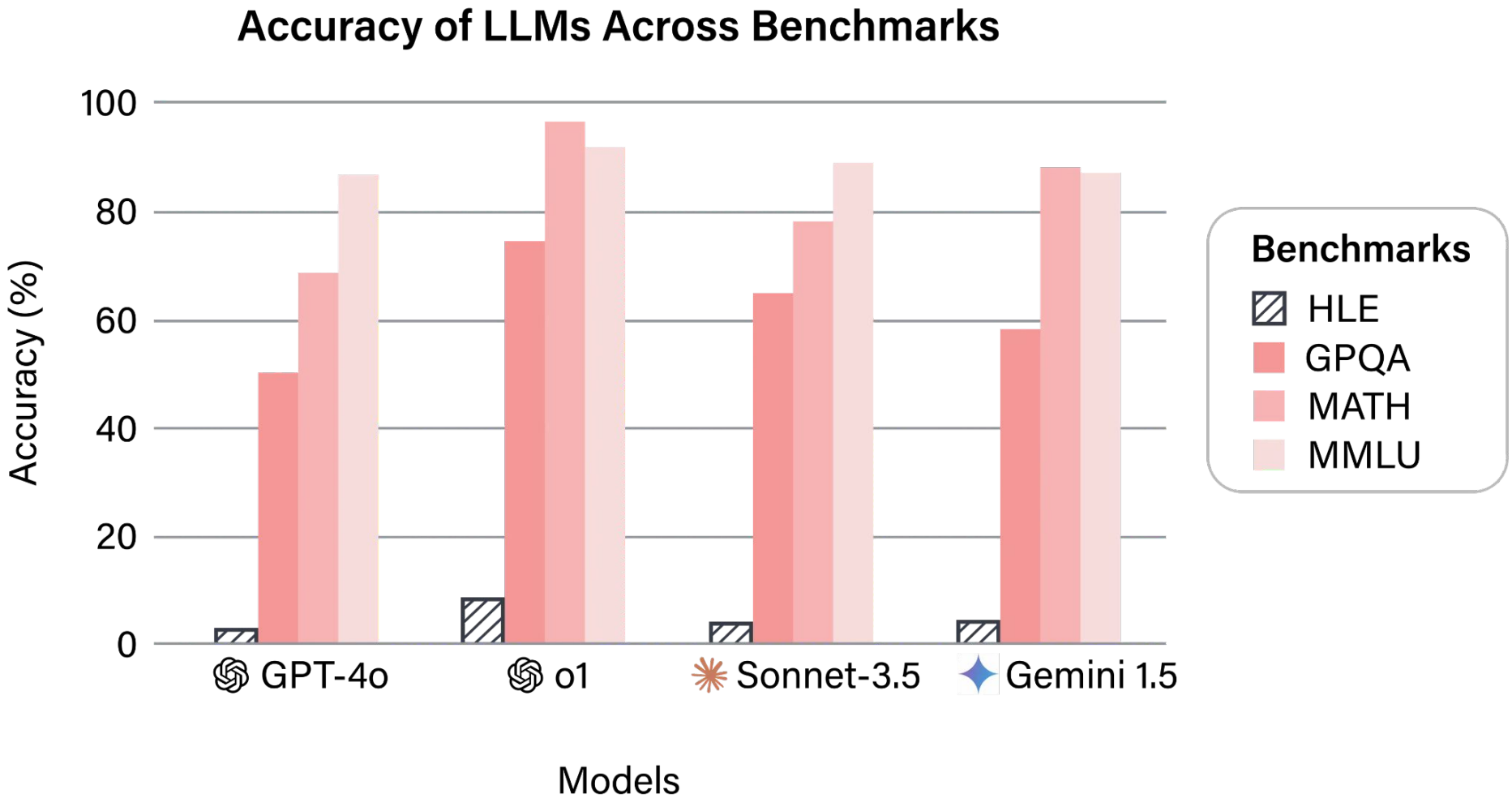
PODCAST

w/ Jon Krohn

Machine Learning | AI | Success

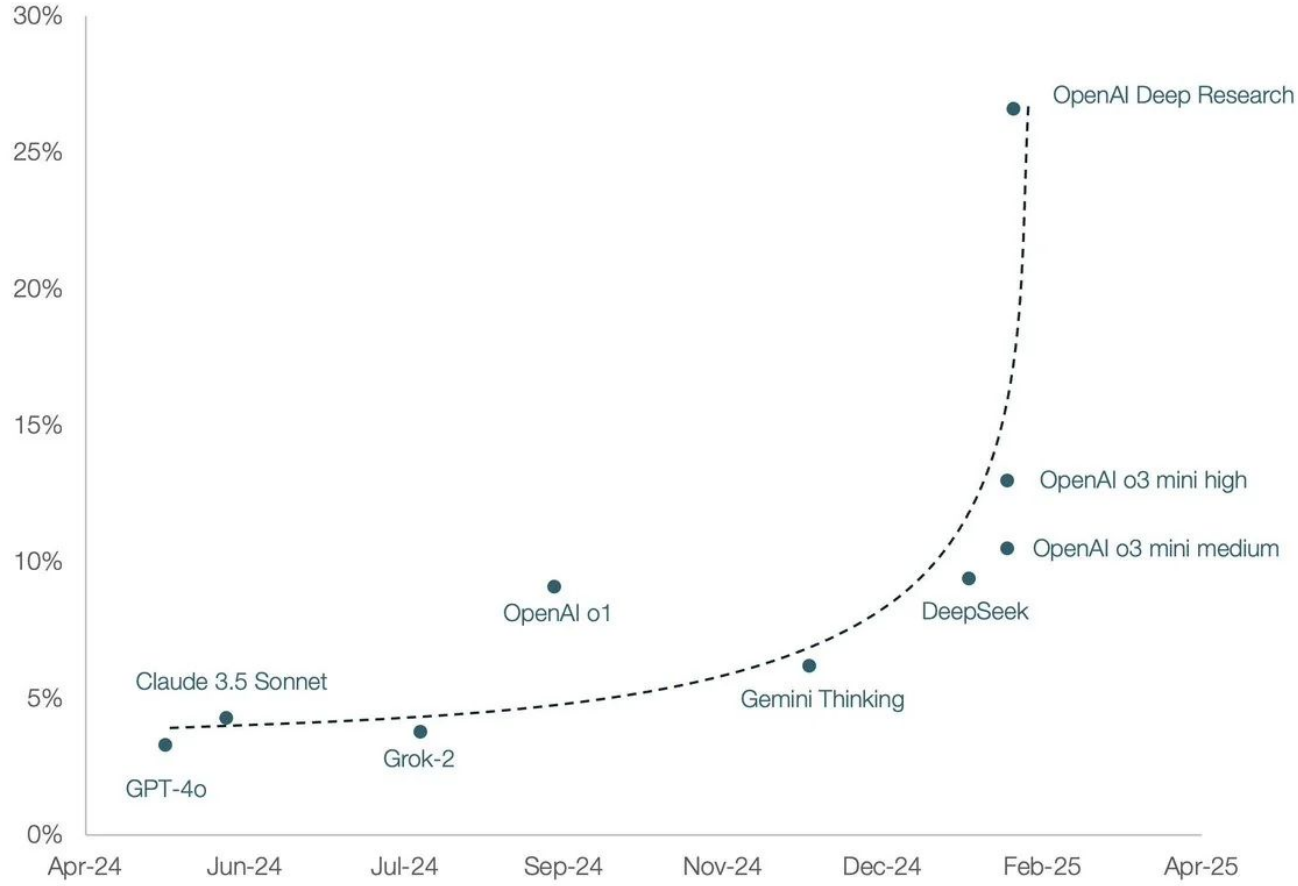
Episode #867

Andriy Burkov



Source: agi.safe.ai

AI Scores on *Humanity's Last Exam*

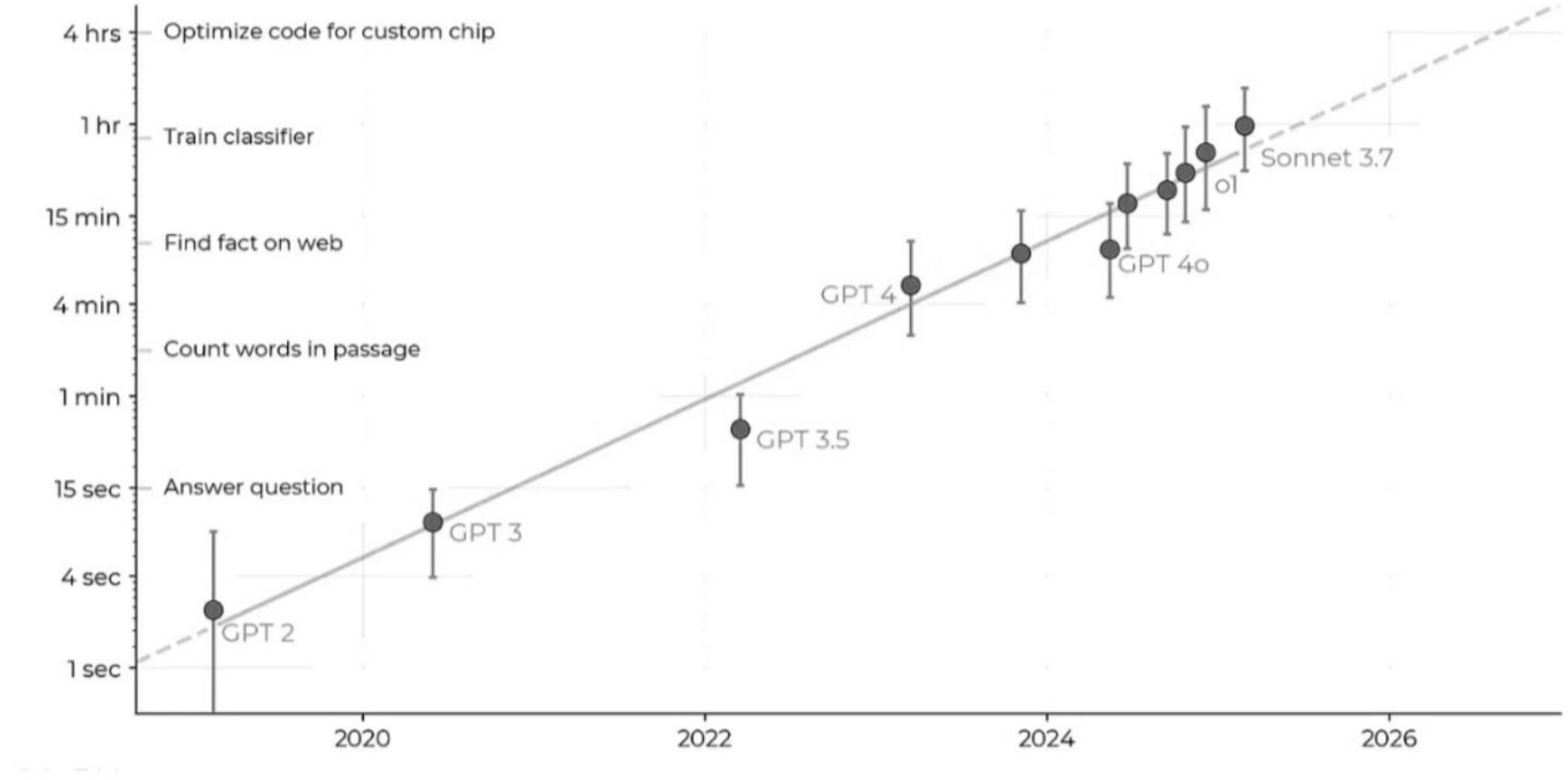


Source: Tomas Pueyo for Uncharted Territories, with data from Dan Hendrycks, of Humanity's Last Exam

The length of tasks AIs can do is doubling every 7 months

METR

Task length (at 50% success rate)



It's Never Been A Better Time To Develop AI Agents

1. ML/LLM Advancements
2. Cloud Infrastructure
3. Data
4. Open-Source Software
5. Low-Cost Hardware
6. Development Tools
7. Connectivity
8. Regulatory Environment
9. Market Demand
10. Educational Resources

Agents in Action

- Code Generation → Team of Software Developers
- Medical Diagnosis
- Scientific Literature Review
- Design and Run Experiments then Write Paper
- Replace More and More Enterprise Tasks → First Billion-$ Firm with No Employees?

Agentic Systems

Anthropic distinguishes two types:

Workflows:
Systems where LLMs and tools are orchestrated through predefined code paths

Agents:
Systems where LLMs dynamically direct their own processes and tool usage, maintaining control over how they accomplish tasks

"Tools" give LLMs autonomy

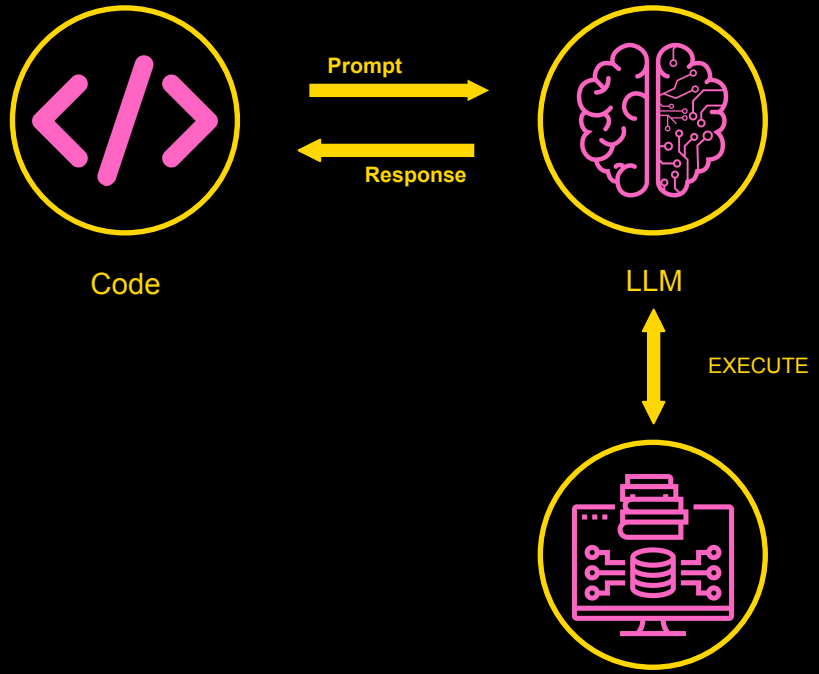
Give an LLM the power to carry out actions like **query a database** or **message other LLMs**

Sounds **spooky**, right?
OpenAI can reach into my computer?!

The reality is a bit **mundane**

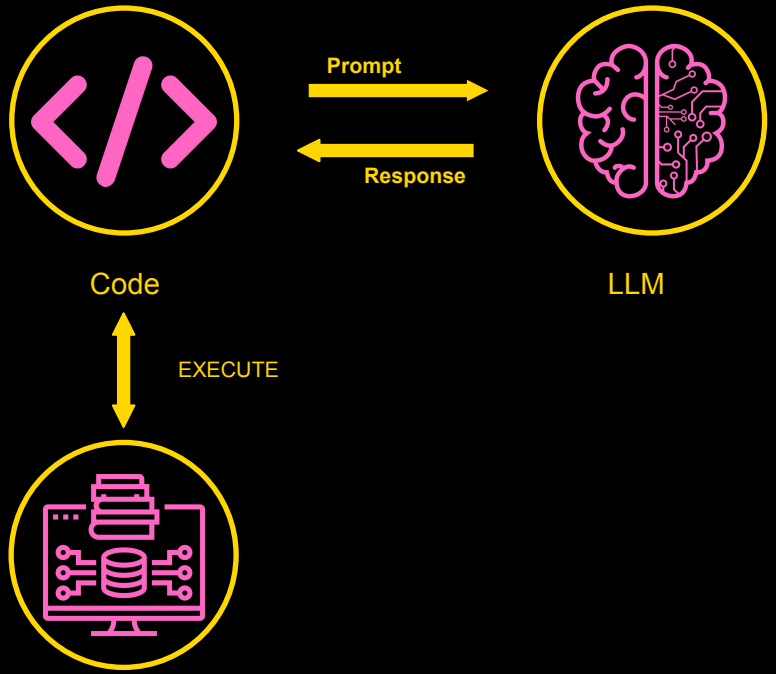
Tool Calling — popular perception

“An LLM can reach into my computer”

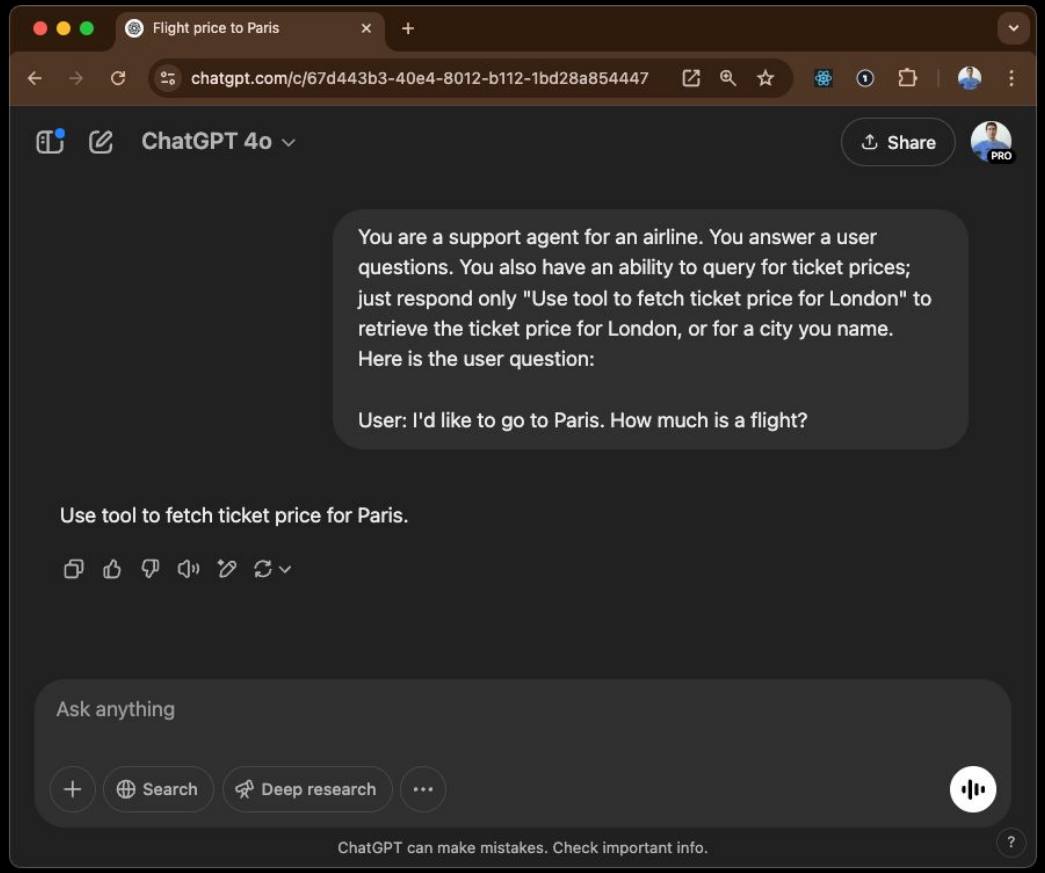


Tool Calling — mundane reality

An LLM responds with the actions needed



Behold Tool use.



Risks of Agent Systems

Unpredictable path

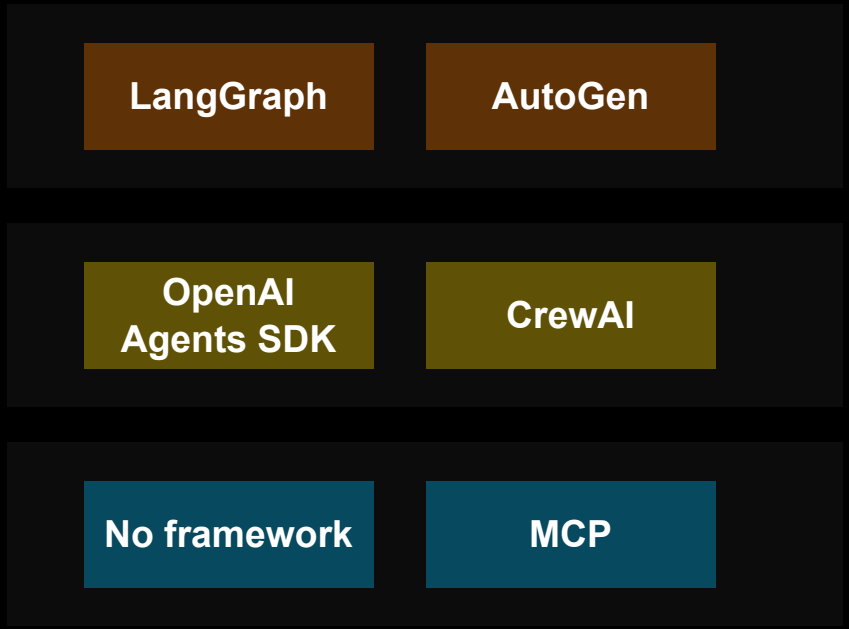
Unpredictable output

Unpredictable costs

Monitor

"Guardrails ensure your agents behave safely, consistently, and within your intended boundaries"

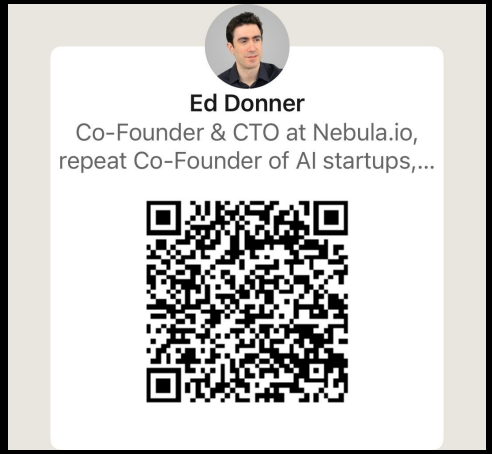
Agentic AI Frameworks

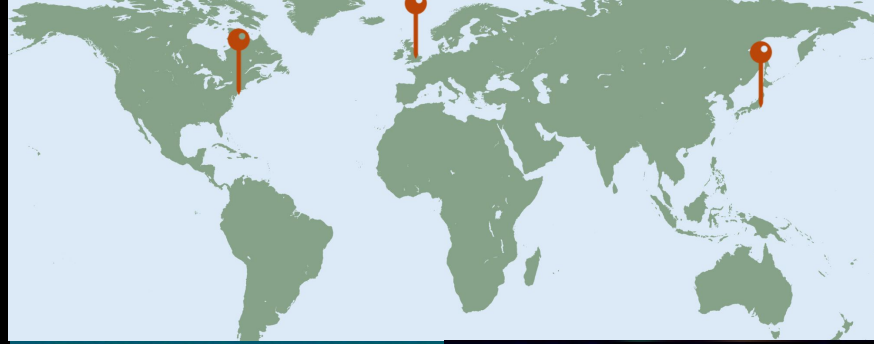


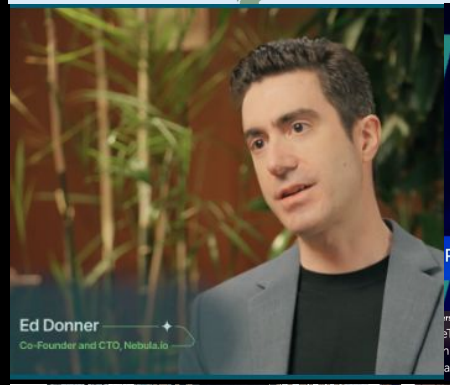
…and many, many more! Which to pick depends on the **use case** and **your preference**

NEBULA

https://www.linkedin.com/in/eddonner/







That's me!!

..and that's Jon!!

Introducing OpenAI Agents SDK

Lightweight and flexible

Stays out of the way

Makes common activities easy

(My favorite)

Minimal terminology

Agents represent LLMs

Handoffs represent interactions

Guardrails represent controls

Three steps

1. Create an instance of **Agent**

2. Use **with trace()** to track

3. Call **Runner.run()** to run

Coding project #1
Deep Research

Good News:

We have quite the treat in store for you

Bad News:

Partly pre-baked! And we will move fast.

Good News:

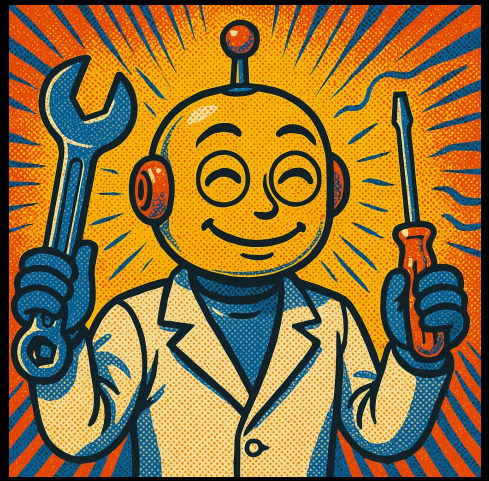
You can recreate this yourself!

https://github.com/ed-donner/action

Module 1:
Defining Agents

Module 1 Coding:
Deep Research
with OpenAI Agents SDK

Module 2:
Designing Agents



Module 2 Coding:
Engineering Team
with CrewAI

Module 3:
Developing Agents

Module 3 Coding:
Autonomous Traders
with MCP

Module 2: **Designing** Agents

Module 2 Coding:
Engineering Team
with CrewAI

Agentic Systems

Anthropic distinguishes two types:

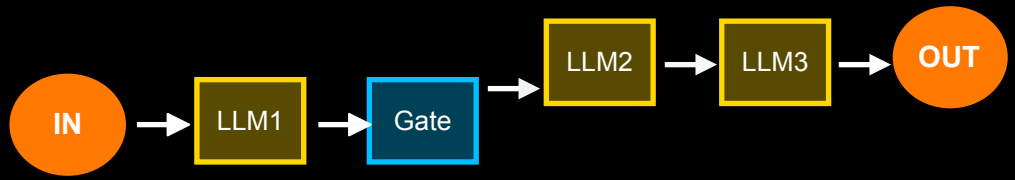
Workflows:
Systems where LLMs and tools are orchestrated through predefined code paths

Agents:
Systems where LLMs dynamically direct their own processes and tool usage, maintaining control over how they accomplish tasks

5 workflow design patterns

1. PROMPT CHAINING

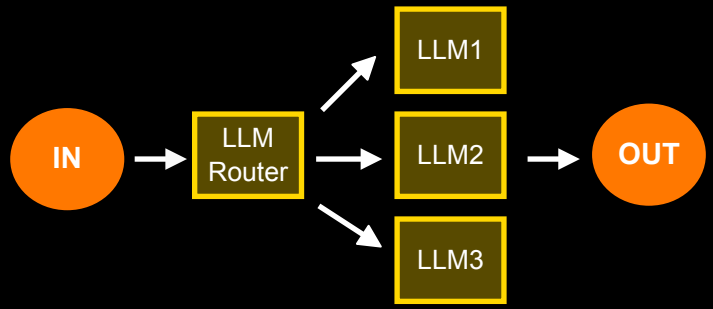
Decompose into fixed sub-tasks



5 workflow design patterns

2. ROUTING

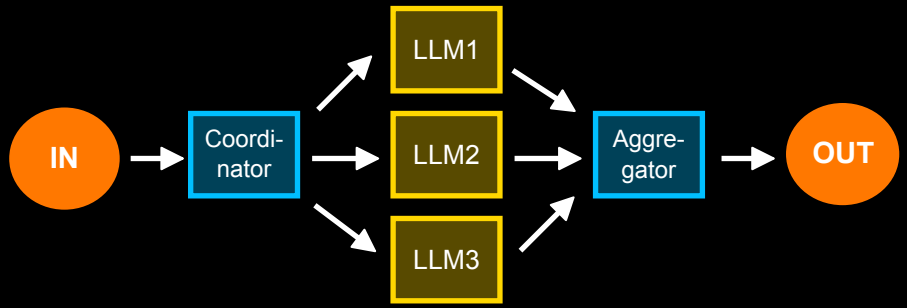
Direct an input into a specialized sub-task, ensuring separation of concerns



5 workflow design patterns

3. PARALLELIZATION

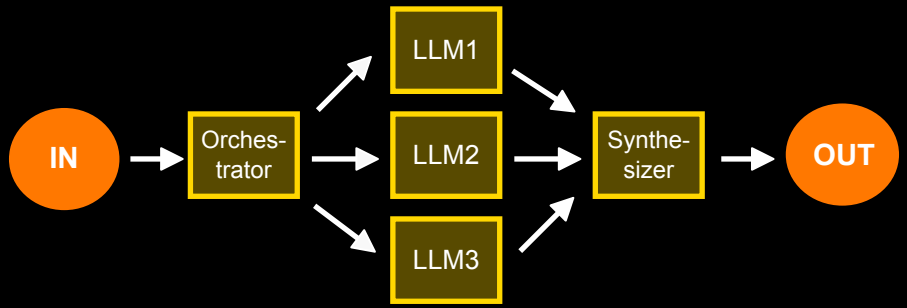
Breaking down tasks and running multiple subtasks concurrently



5 workflow design patterns

4. ORCHESTRATOR-WORKER

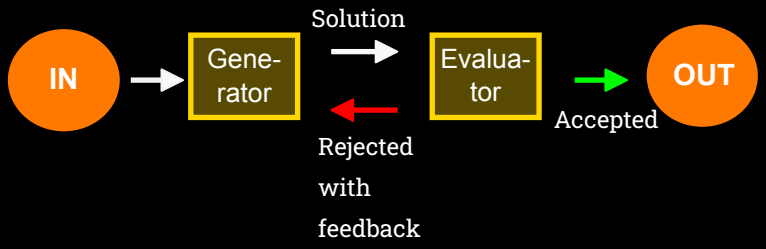
Complex tasks are broken down dynamically and combined



5 workflow design patterns

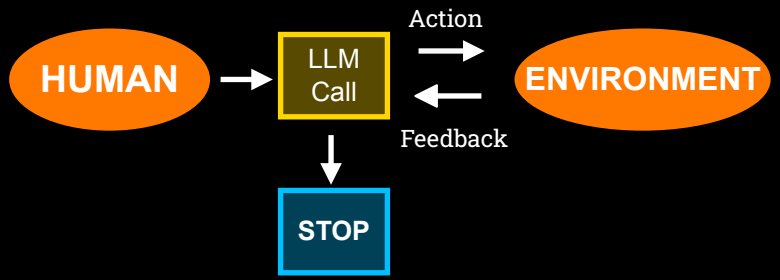
5. EVALUATOR-OPTIMIZER

LLM output is validated by another



By contrast, Agents:

1. Open-ended
2. Feedback loops
3. No fixed path



CrewAI core concepts

Agent: an autonomous unit, with an LLM, a role, a goal, a backstory, memory, tools

Task: a specific assignment to be carried out, with a description, expected output, agent

Crew: a team of **Agents** and **Tasks**; either:

> Sequential: run tasks in order they are defined
>
> Hierarchical: use a Manager LLM to assign

Lightweight, but somewhat more opinionated than OpenAI Agents SDK -
more terminology, more prescriptive

Five Steps to a Crew AI project

1. Create the project with:
   crewai create crew my_project
2. Fill in the config yaml files to define the **Agents** and **Tasks**
3. Complete the crew.py module to create the **Agents**, **Tasks** and **Crew**, referencing the config
4. Update main.py to set any inputs
5. Run with:
   crewai run

Giving coding skills to an Agent

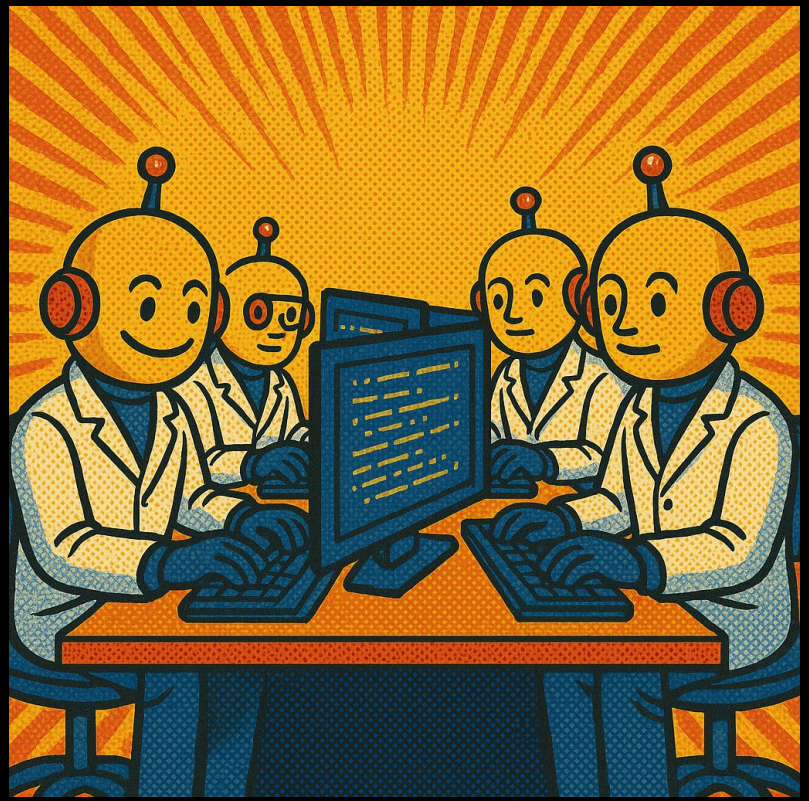
It's **hard** and it's **complex**, but you can have an Agent in Crew that has the ability to write code, execute it securely isolated in a Docker container, and investigate the results

Except it's not.

```
Agent(
    allow_code_execution=True,
    code_execution_mode="safe"
)
```

These are often described as Coder Agents.

Engineering Team

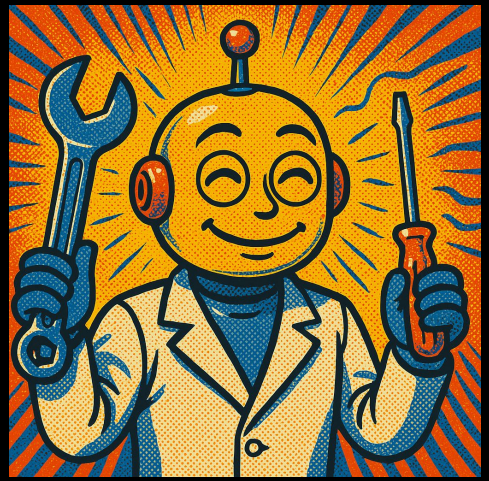


- Engineering Lead
- Backend Engineer
- Frontend Engineer
- Test Engineer

Module 1:
Defining Agents

Module 1 Coding:
Deep Research
with OpenAI Agents SDK

Module 2:
Designing Agents



Module 2 Coding:
Engineering Team
with CrewAI

Module 3:
Developing Agents

Module 3 Coding:
Autonomous Traders
with MCP

Module 3: **Developing** Agents

Module 3 Coding:
Autonomous Traders
with MCP

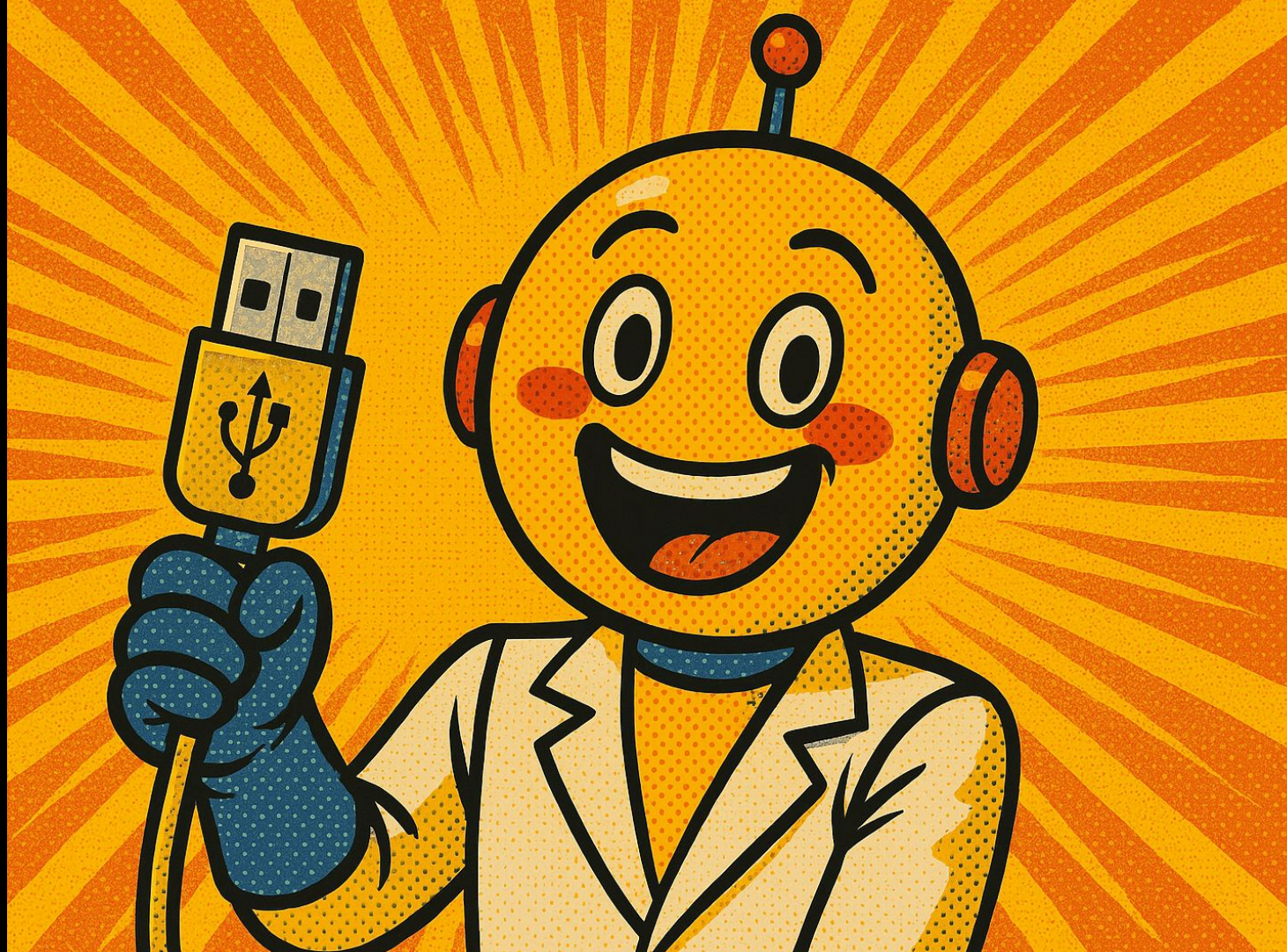
And now!

Introducing the

Model

Context

Protocol



MCP

What it's not

A framework for building agents

A fundamental change to how agents work

A way to code agents

What it is

A protocol - a standard

A simple way to integrate tools, resources, prompts

"A USB-C port for AI applications"

MCP

Reasons not to be excited

It’s just a standard, it’s not tools themselves

LangChain already has a big Tools ecosystem

You can already make any function into a Tool

Reasons to be excited

Makes it frictionless to integrate

It's taking off! Exploding ecosystem

HTML was just a standard, too 😉

MCP Core Concepts

The Three Components:

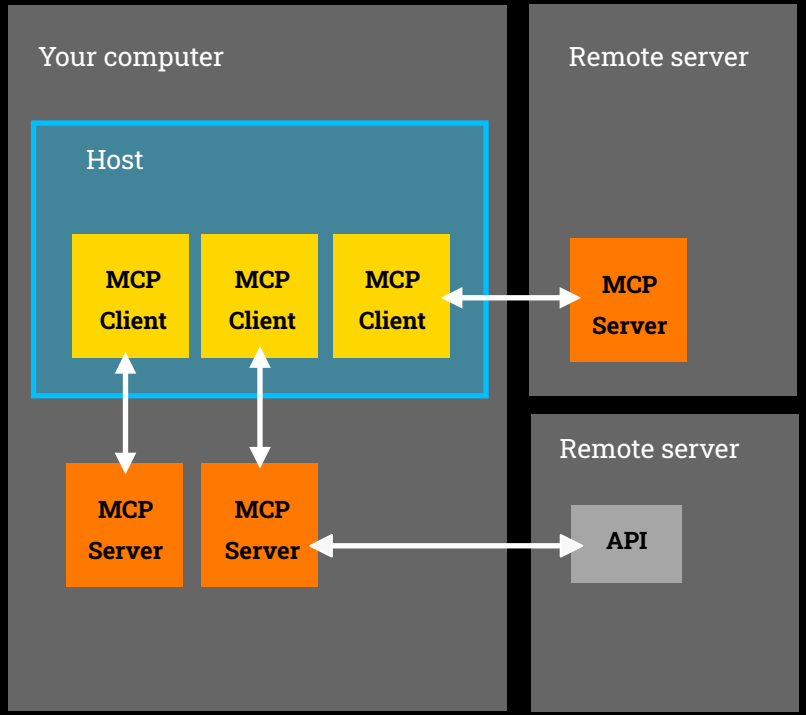
> **Host** is an LLM app like Claude or our Agent architecture
>
> MCP Client lives inside Host and connects 1:1 to MCP Server
>
> MCP Server provides tools, context and prompts

Example:

Google Maps is an **MCP Server** with geolocation tools (see here)

You can configure **Claude Desktop** (the host) to run an **MCP Client** that then launches the Google Maps MCP Server on your computer

Architecture



Common misconception:

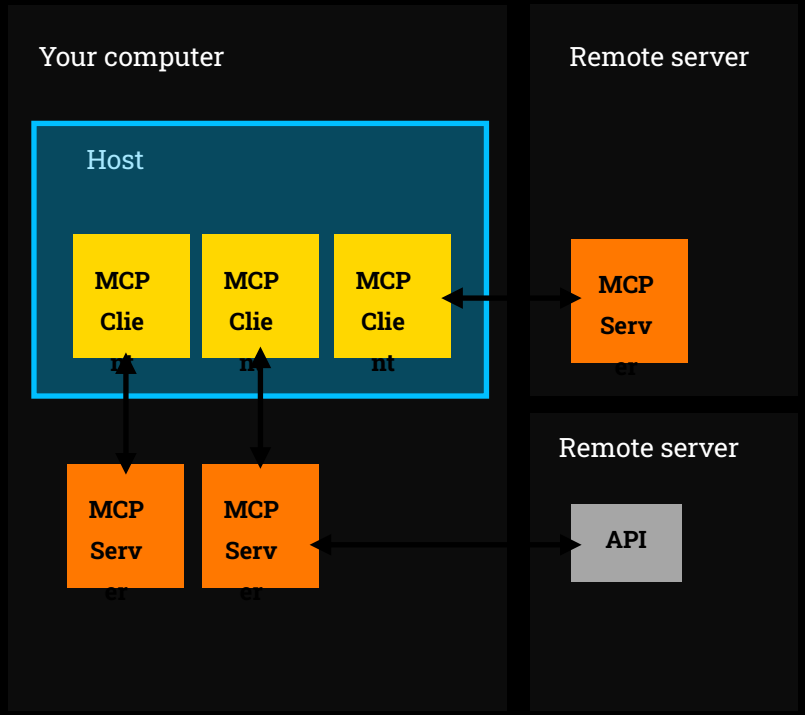
MCP Servers run remotely

Most common reality:

Download open-source MCP Server

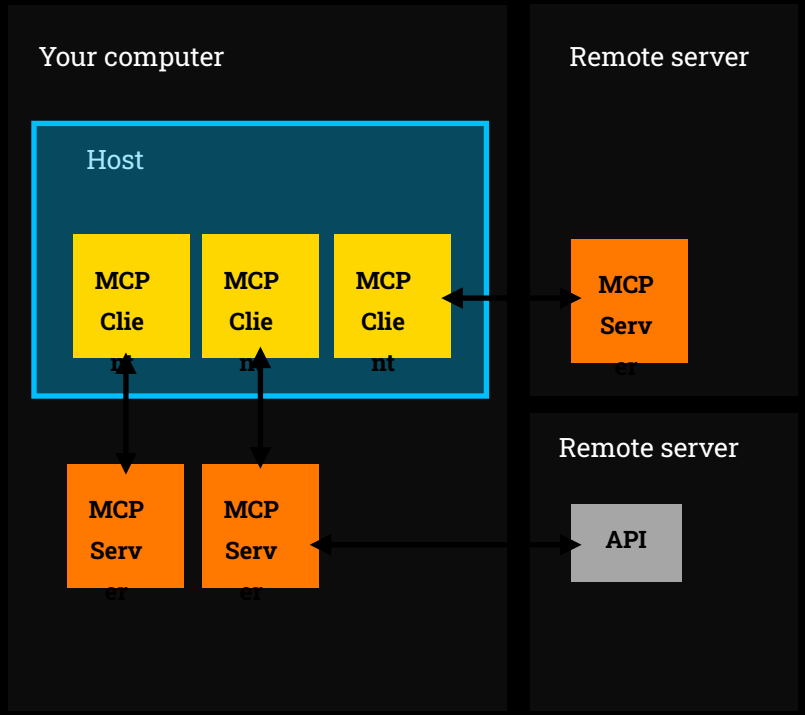
Run locally

Architecture



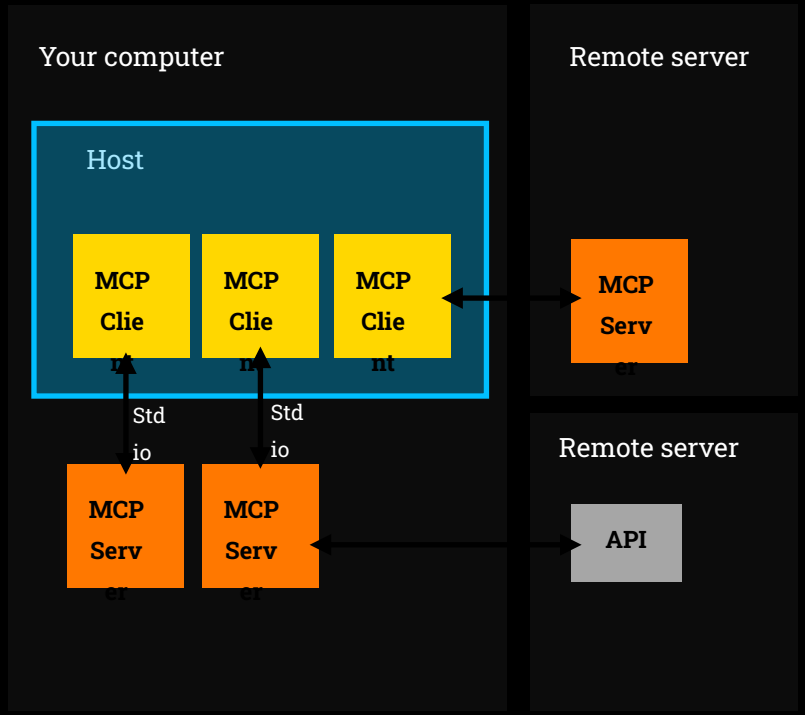
MCP Servers most often run on your box

Download open-source MCP Servers, run them locally



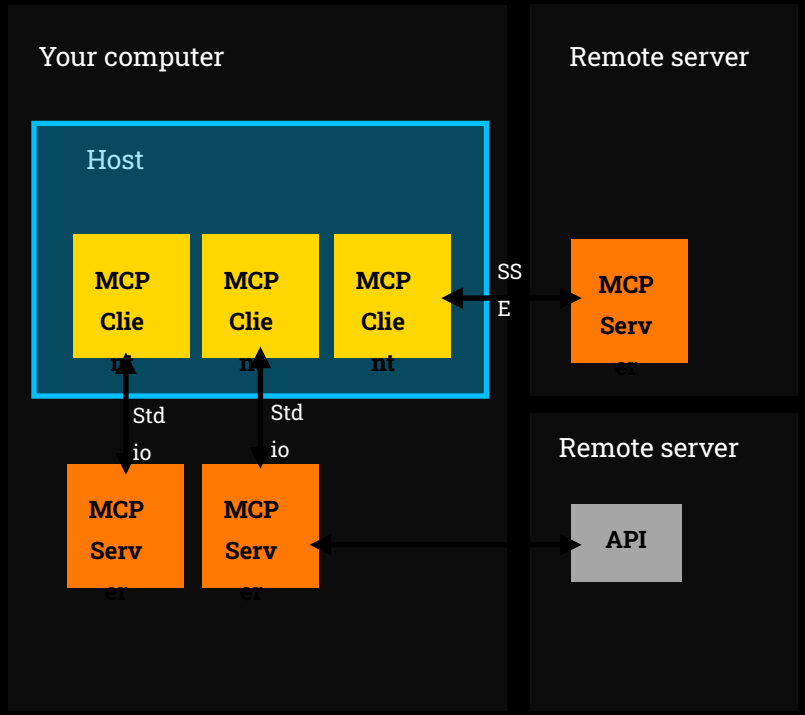
Two "Transport" mechanisms

Stdio spawns a process and communicates via standard input/output



Two "Transport" mechanisms

..while SSE uses HTTPS connections with streaming



Making an MCP Server

Why make an MCP Server

Sharing! Allow others to incorporate tools and resources

Consistently incorporate all our MCP Servers

Understand the plumbing

Reasons not to make an MCP Server

If it's only for us, then we could just make tools -
the **@function_tool** decorator can make any function into a tool

Module 3 Project

Autonomous Traders

Commercial

6 MCP servers with 44 tools and 2 resources

Agent interactions

Autonomous

Do not use for trading decisions!

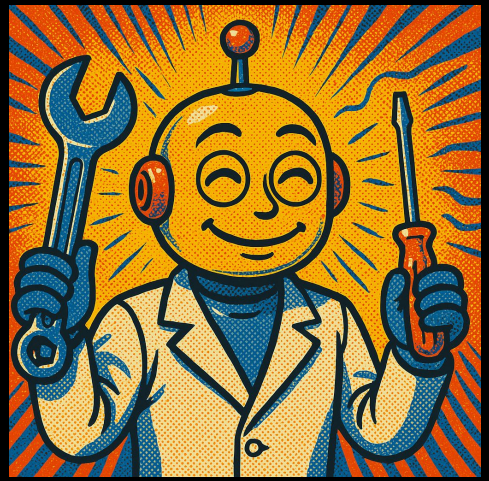
Note: we will go through the code quickly; use it to get some intuition and explore later!


NEWS
74.50
72.34
153.72

Module 1:
Defining Agents

Module 1 Coding:
Deep Research
with OpenAI Agents SDK

Module 2:
Designing Agents



Module 2 Coding:
Engineering Team
with CrewAI

Module 3:
Developing Agents

Module 3 Coding:
Autonomous Traders
with MCP

It's Never Been A Better Time To Develop AI Agents

1. ML/LLM Advancements
2. Cloud Infrastructure
3. Data
4. Open-Source Software
5. Low-Cost Hardware
6. Development Tools
7. Connectivity
8. Regulatory Environment
9. Market Demand
10. Educational Resources

Workforce Implications

For coming years, more jobs created by AI than lost

However: funding for reskilling programs is key

How data scientists in particular can “future-proof”:

- Stay on top of AI trends
- Focus on multi-agent orchestration and management
- Don’t underestimate power of foundational subjects
- Develop domain expertise
- Develop human-AI collaboration skills
- Hone communication / influence skills within organization

1. Energy
2. Nutrition
3. Lifespans
4. Education
5. Freedom from violence
6. Freedom of expression
7. Sustainability
8. Cultural preservation
9. Exploration
10. Community

Thank you! **Please stay in touch**

Jon

jonkrohn.com to sign up for email newsletter

linkedin.com/in/jonkrohn

youtube.com/c/JonKrohnLearns

ycarrot.com: consulting + hiring!

- Skip waitlist
- 15 free credits per month
- 20% off first month of Pro

Ed

Full Agentic AI course available on Udemy

linkedin.com/in/eddonner

edwarddonner.com

Complete Agentic AI course on Udemy

Code Screenshots

Outputs to the labs are here:
https://github.com/ed-donner/action/tree/main/outputs

Gradio

127.0.0.1:7860

Deep Research

What topic would you like to research?

What should I attend at ODSC East on Thursday May 15th 2025?

Run

ODSC East 2025: Event Overview for May 15th

Introduction

ODSC East 2025 is set to take place from May 13–15, 2025, at the Boston Convention and Exhibition Center. This report focuses on the key events and sessions occurring specifically on Thursday, May 15th, helping potential attendees maximize their experience at the conference.

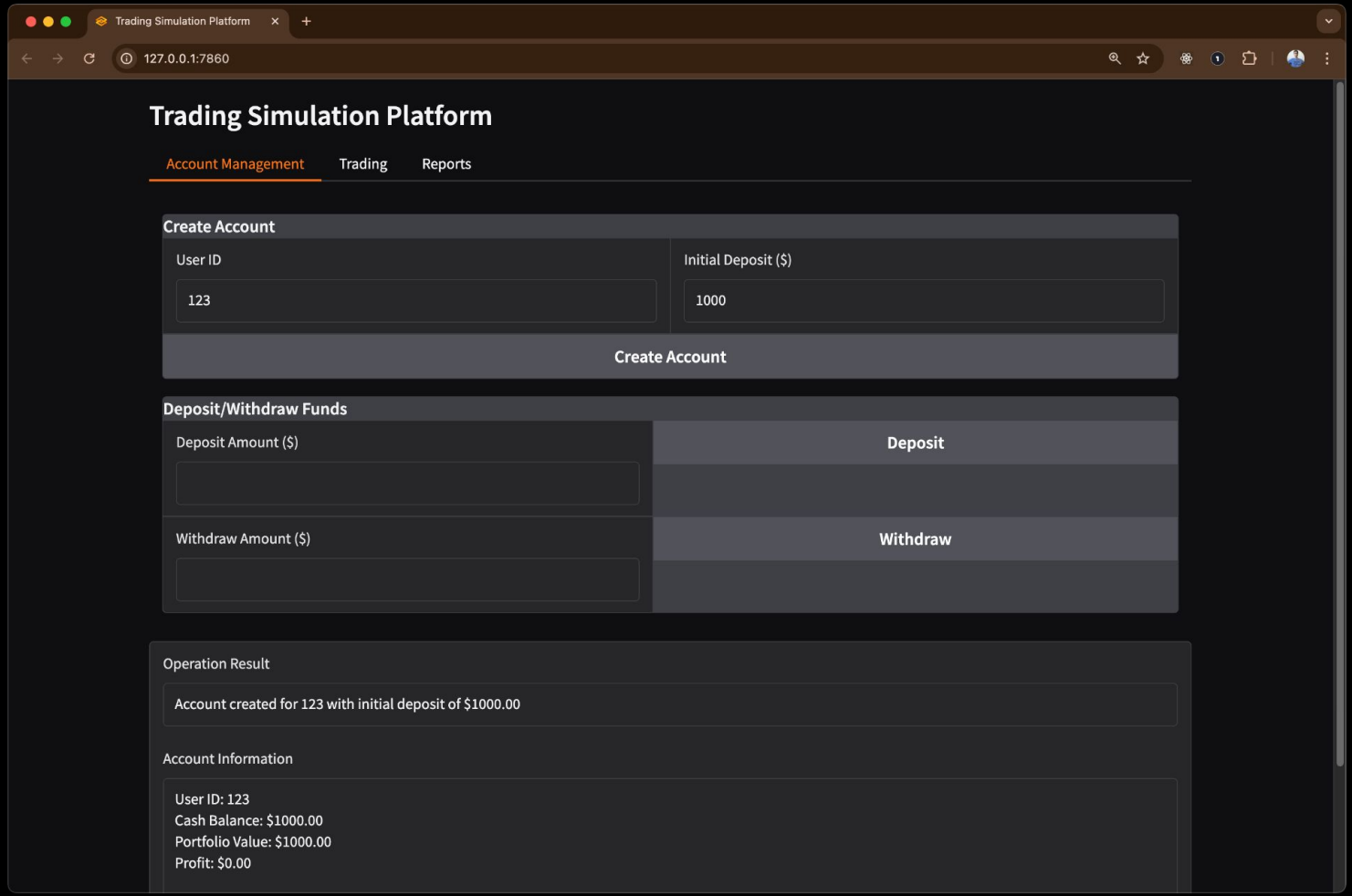
Background of ODSC East

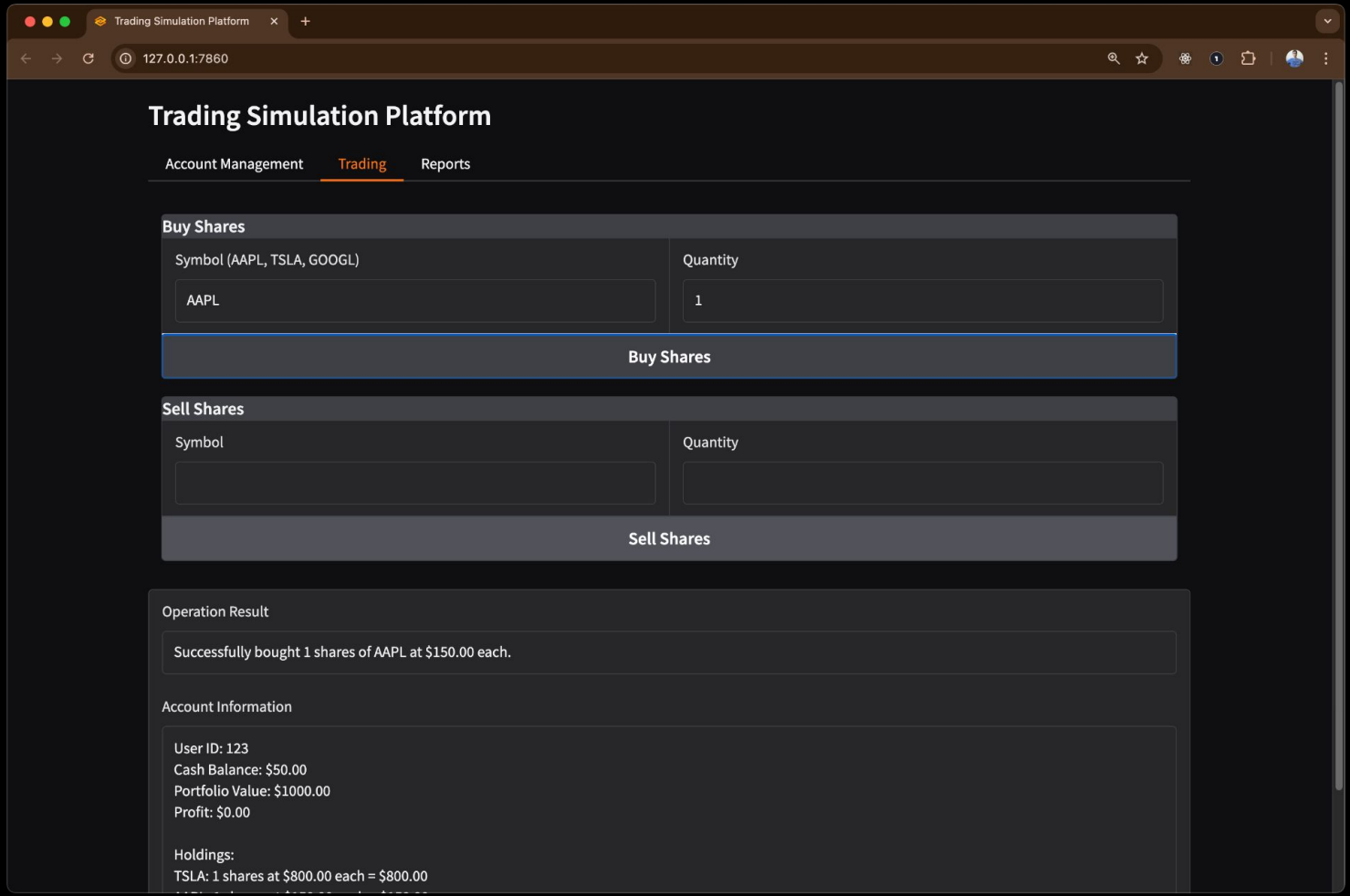
The Open Data Science Conference (ODSC) has established itself as a pivotal gathering for data scientists, AI experts, and industry professionals. The 2025 event marks the 10th anniversary of the conference, showcasing advancements in AI and data science through workshops, keynote presentations, and networking opportunities.

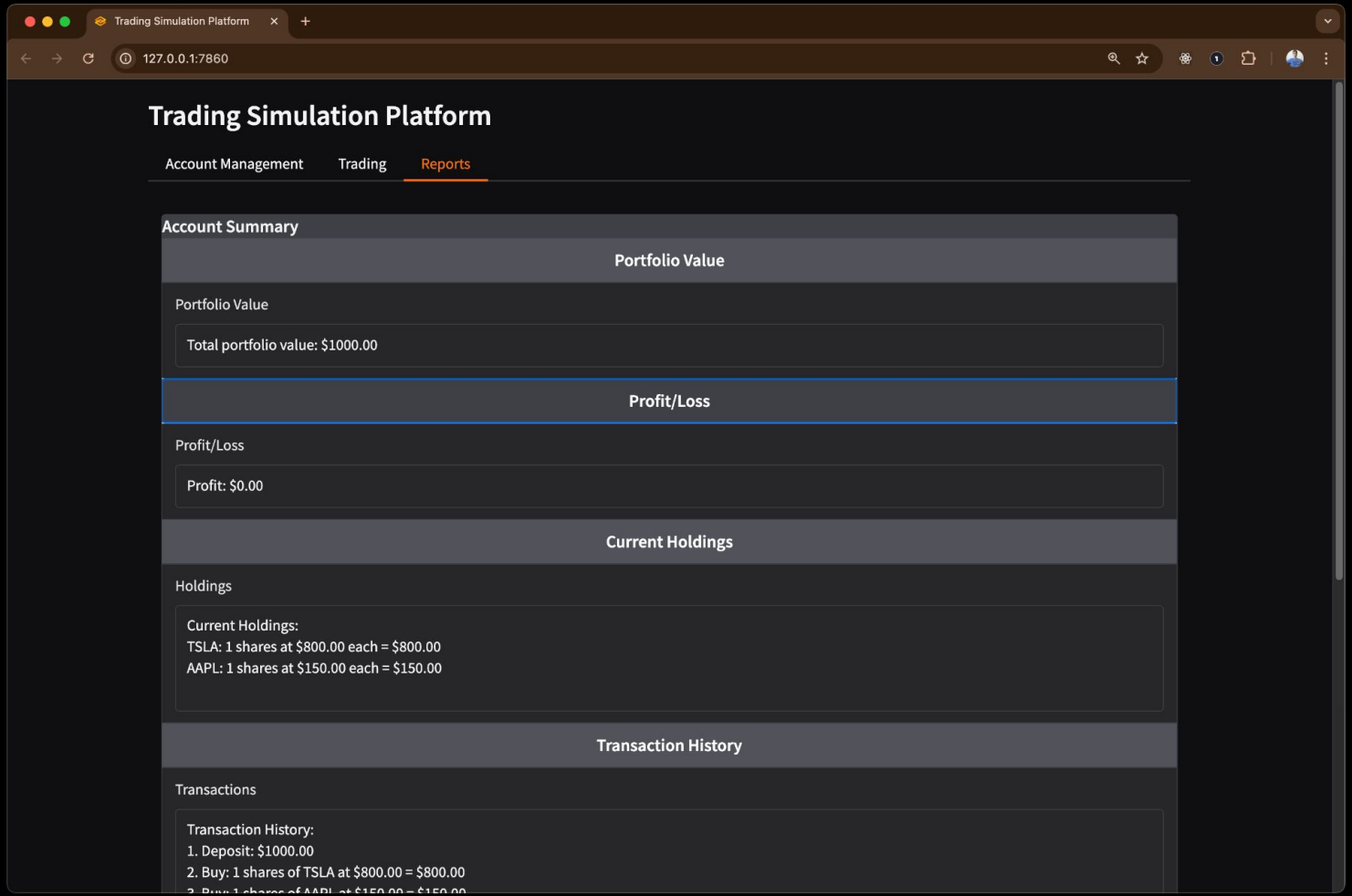
Key Events on May 15, 2025

1. ODSC Conference Sessions

Thursday marks the final day of the main conference with a focus on hands-on training sessions, workshops, and keynote presentations. These events are designed to cater to a diverse audience, ranging from nascent data scientists to seasoned professionals.


Trading Simulation Platform
127.0.0.1:7860
Trading Simulation Platform
Account Management
Trading
Reports
Create Account
User ID
123
Initial Deposit ($)
1000
Create Account
Deposit/Withdraw Funds
Deposit Amount ($)
Deposit
Withdraw Amount ($)
Withdraw
Operation Result
Account created for 123 with initial deposit of $1000.00
Account Information
User ID: 123
Cash Balance: $1000.00
Portfolio Value: $1000.00
Profit: $0.00


Trading Simulation Platform
127.0.0.1:7860
Trading Simulation Platform
Account Management
Trading
Reports
Buy Shares
Symbol (AAPL, TSLA, GOOGL)
AAPL
Quantity
1
Buy Shares
Sell Shares
Symbol
Quantity
Sell Shares
Operation Result
Successfully bought 1 shares of AAPL at $150.00 each.
Account Information
User ID: 123
Cash Balance: $50.00
Portfolio Value: $1000.00
Profit: $0.00
Holdings:
TSLA: 1 shares at $800.00 each = $800.00


Trading Simulation Platform
127.0.0.1:7860
Trading Simulation Platform
Account Management
Trading
Reports
Account Summary
Portfolio Value
Portfolio Value
Total portfolio value: $1000.00
Profit/Loss
Profit/Loss
Profit: $0.00
Current Holdings
Holdings
Current Holdings:
TSLA: 1 shares at $800.00 each = $800.00
AAPL: 1 shares at $150.00 each = $150.00
Transaction History
Transactions
Transaction History:
1. Deposit: $1000.00
2. Buy: 1 shares of TSLA at $800.00 = $800.00

Traders

127.0.0.1:7861/?__theme=dark

Warren (GPT 4.1 Mini) - Patience

$10,116 ⬆ $116


```
2025-05-08 14:57:46 : [response] Started response
2025-05-08 14:57:46 : [function] Ended function buy_shares
2025-05-08 14:57:46 : [function] Ended function sell_shares
2025-05-08 14:57:46 : [account] Retrieved account details
2025-05-08 14:57:46 : [account] Bought 2 of NVDA
2025-05-08 14:57:46 : [account] Retrieved account details
2025-05-08 14:57:51 : [function] Started function push
2025-05-08 14:57:51 : [response] Ended response
2025-05-08 14:57:52 : [response] Started response
2025-05-08 14:57:52 : [function] Ended function push
2025-05-08 14:57:55 : [trace] Ended: Warren-rebalancing
2025-05-08 14:57:55 : [agent] Ended agent Warren
2025-05-08 14:57:55 : [response] Ended response
```

Holdings

Symbol	Quantity
MSFT	10
NXPI	2
NVDA	10

Recent Transactions

symb…	quanti…	price	timestamp
MSFT	10	436.15056	2025-05-05 15:(
AMD	15	98.99759999999999	2025-05-05 15:(
ON	30	38.31648	2025-05-06 11:(
NXPI	10	183.02532	2025-05-06 11:(

George (DeepSeek V3) - Bold

$10,202 ⬆ $202


```
2025-05-08 14:58:12 : [function] Started function sell_shares
2025-05-08 14:58:12 : [generation] Ended generation
2025-05-08 14:58:33 : [generation] Started generation
2025-05-08 14:58:33 : [function] Ended function push
2025-05-08 14:58:33 : [function] Started function push
2025-05-08 14:58:33 : [generation] Ended generation
2025-05-08 14:58:41 : [generation] Started generation
2025-05-08 14:58:41 : [function] Ended function push
2025-05-08 14:58:41 : [function] Started function push
2025-05-08 14:58:41 : [generation] Ended generation
2025-05-08 14:58:53 : [trace] Ended: George-rebalancing
2025-05-08 14:58:53 : [agent] Ended agent George
2025-05-08 14:58:53 : [generation] Ended generation
```

Holdings

Symbol	Quantity
NVDA	2
AVGO	22
QCOM	20

Recent Transactions

symb…	quanti…	price	timestamp
NVDA	20	114.729	2025-05-05 15:(
PINS	50	26.923740000000(	2025-05-05 15:(
AI	50	21.72336	2025-05-06 11:(
PINS	-50	27.0607699999999	2025-05-06 12:(

Ray (Gemini 2.5 Flash) - Systematic

$10,093 ⬆ $93


```
2025-05-08 14:57:06 : [agent] Ended agent Researcher
2025-05-08 14:57:06 : [generation] Ended generation
2025-05-08 14:57:11 : [generation] Started generation
2025-05-08 14:57:11 : [function] Ended function push
2025-05-08 14:57:11 : [function] Started function push
2025-05-08 14:57:11 : [generation] Ended generation
2025-05-08 14:57:14 : [generation] Started generation
2025-05-08 14:57:14 : [function] Ended function push
2025-05-08 14:57:14 : [function] Started function push
2025-05-08 14:57:14 : [generation] Ended generation
2025-05-08 14:57:17 : [trace] Ended: Ray-rebalancing
2025-05-08 14:57:17 : [agent] Ended agent Ray
2025-05-08 14:57:17 : [generation] Ended generation
```

Holdings

Symbol	Quantity
VOO	18
NVDA	6

Recent Transactions

symb…	quanti…	price	timestamp
VOO	9	517.237410000000	2025-05-07 10:27
VOO	10	514.727400000000	2025-05-07 11:29
NVDA	1	113.837320200000	2025-05-07 12:31
VOO	-1	512.912120000000	2025-05-07 15:38

Cathie (Grok 3 Mini) - Crypto

$10,275 ⬆ $275


```
2025-05-08 14:57:52 : [account] Bought 10 of ICLN
2025-05-08 14:57:52 : [function] Started function buy_shares
2025-05-08 14:57:52 : [generation] Ended generation
2025-05-08 14:57:53 : [generation] Started generation
2025-05-08 14:57:53 : [function] Ended function buy_shares
2025-05-08 14:57:53 : [account] Retrieved account details
2025-05-08 14:57:58 : [generation] Started generation
2025-05-08 14:57:58 : [function] Ended function push
2025-05-08 14:57:58 : [function] Started function push
2025-05-08 14:57:58 : [generation] Ended generation
2025-05-08 14:58:03 : [trace] Ended: Cathie-rebalancing
2025-05-08 14:58:03 : [agent] Ended agent Cathie
2025-05-08 14:58:03 : [generation] Ended generation
```

Holdings

Symbol	Quantity
NVDA	76
ARKK	3
ICLN	13

Recent Transactions

symb…	quanti…	price	timestamp
NVDA	50	114.729	2025-05-05 15:(
NVDA	37	113.13081	2025-05-06 12:(
ARKK	1	49.108019999999	2025-05-07 12:3
NVDA	-1	113.61731	2025-05-07 15:3