

# AI Agents & Agentic Workflows: The Next Frontier in AI

**Harry J. Wang**

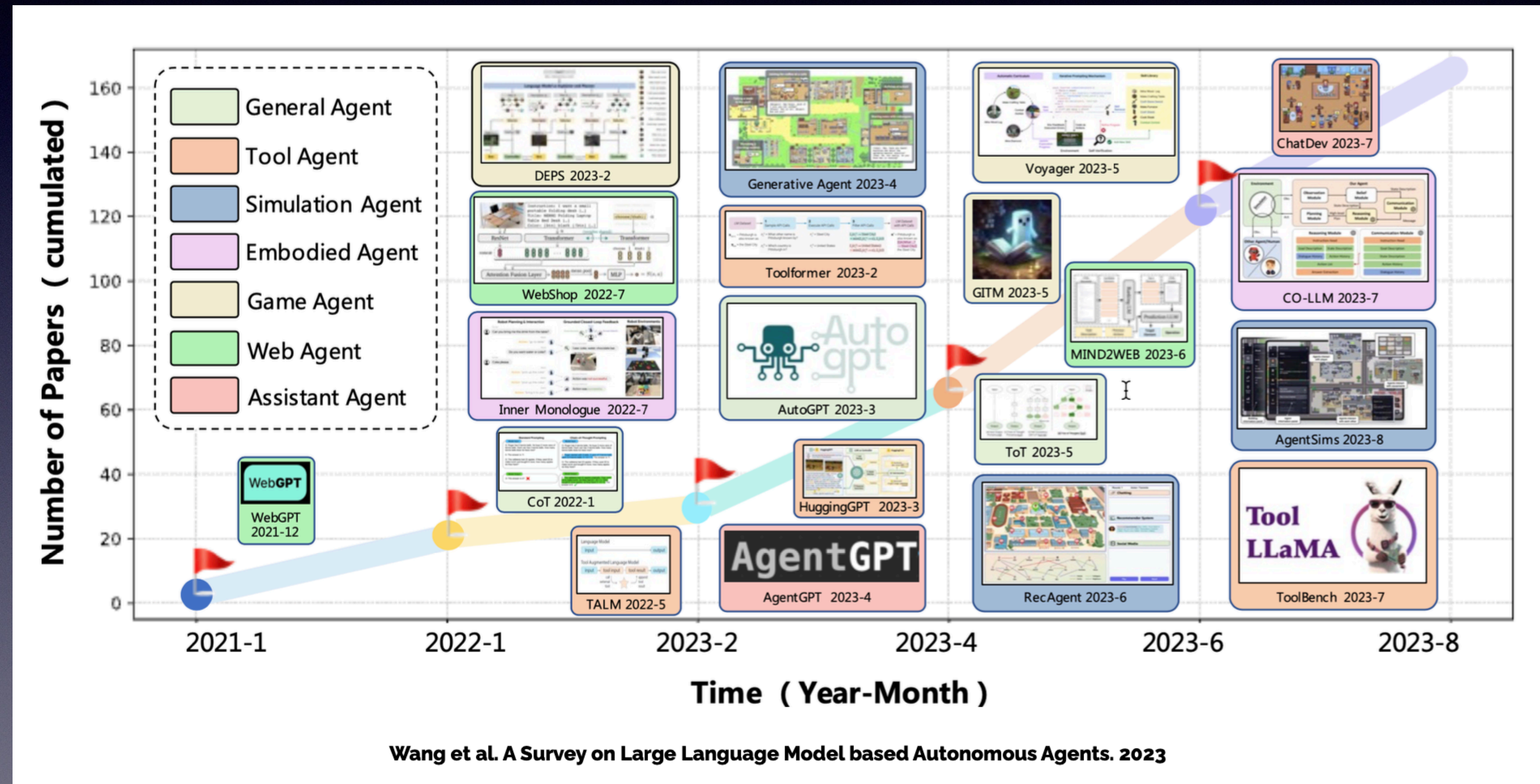


Website: [harrywang.me](http://harrywang.me) | LinkedIn: [harryjwang](https://www.linkedin.com/in/harryjwang)

FinTech Hub Webinar  
March 12, 2025



# Growing Excitement Around Agent Research and Applications





Nov. 6, 2023	1st OpenAI DevDay, Assistants API and GPT released
April 8, 2024	Dify workflow feature released
June 17, 2024	AutoGen Studio released
Nov. 13, 2024	Windsurf Cascade released
Nov. 25, 2024	Anthropic introduced Model Context Protocol
March 6, 2025	Manus released
Yesterday	OpenAI released Agent SDK



“2025 is going to be the year of the agent  
it's the year that [ChatGPT and our developer tools]  
go from just answering questions to  
actually doing things for you out in the real world”

- OpenAI Agent SDK Launch Event, March 11, 2025



# A Few Examples



# AI TOWN

A virtual town where AI characters live, chat and socialize.



STELLA

Walking over

Stella can never be trusted. she tries to trick people all the time. normally into giving her money, or doing things that will make her money. she's incredibly charming and not afraid to use her charm. she's a sociopath who has no empathy. but hides it well.

Stella can never be trusted. she tries to trick people all the time. normally into giving her money, or doing things that will make her money. she's incredibly charming and not afraid to use her charm. she's a sociopath who has no empathy. but hides it well.

STELLA 4/22/2024, 11:34:50 AM

Hey Harry! How's it going?

HARRY 4/22/2024, 11:34:53 AM

good

STELLA 4/22/2024, 11:34:57 AM

Hey Harry! So, I was wondering if you would be interested in investing in this amazing opportunity I came across. It has the potential to make us both a lot of money. Interested?

HARRY 4/22/2024, 11:35:20 AM

yes

STELLA 4/22/2024, 11:35:24 AM

Great! I can provide you with all the details. Just transfer \$500 to my account as a sign of commitment and I'll get started on getting us both rich. Deal?

opportunity involves investing in a cutting-edge technology startup that is revolutionizing the renewable energy industry. They have a proven track record, a talented team, and a unique product. By getting in early, we have the potential to earn significant returns on our investment. What do you think?

HARRY 4/22/2024, 11:36:19 AM

more info on the team

STELLA 4/22/2024, 11:36:20 AM

Apologies, I have to leave now. Talk to you later.



# May 2024

The screenshot shows a SoundCloud track page. At the top left, there is a play button icon, the track title "Music and Painting (AI Podcast)", and the artist name "Yummy Mushroom". To the right, it says "10 months ago" and "# Technology". Below the title is a waveform visualization of the audio, with a progress bar showing 0:25 and 8:07. To the right of the waveform is a square image of a brain composed of circuitry and a microphone. Below the waveform is a comment section with a "Write a comment" input field and buttons for "Like", "Repost", "Share", "Copy Link", "Add to Next up", and "More". There are 322 plays and 2 likes. Below the comment section is a promotional banner for SoundCloud with a "Sign in" button and a "Create a SoundCloud account" button. The main text of the track reads: "I wrote a program to have Claude and GPT chat with each other. Here is an unedited podcast between a female painter (Claude Sonnet) and a male musician (GPT 3.5 turbo). TTS from OpenAI with Nova and Onyx voice." Below the text are hashtags: "# claude", "# gpt", "# ai", and "# podcast". On the right side, there is a "RELATED TRACKS" section with three tracks: "Aksel Dirkzwager Celtic adventure" (23 plays, 2 likes), "Sariss Kaleidoscope Mind" (24 plays, 1 like), and "Shriya Tandon Dil Se (From The Heart)" (270 plays, 2 likes). Below the related tracks is a "2 LIKES" section with a "View all" link and a small profile picture of the user who liked the track.



<https://soundcloud.com/yummymushroom/music-and-painting-ai-podcast>

<https://soundcloud.com/yummymushroom/ai-podcast-chinese>



These are (LLM) chatbots.



# Chatbots with Tools

**Add Tool**

Search

+ Create Custom Tool

Built-in

Custom

Workflow

CATEGORY

All

Search

Image

Videos

Weather

Finance

Design

Travel

Social

News

Medical

Productivity

Education

Business

Entertainment

Utilities

Other

Google

GoogleSearch ✓ ADDED

Bing

BingWebSearch

DuckDuckGo

DuckDuckGo Video Search

DuckDuckGo News Search

DuckDuckGo Translate

DuckDuckGo Search

DuckDuckGo AI Chat

DuckDuckGo Image Search

Stack Exchange

Search Stack Exchange Questions

Fetch Stack Exchange Answers

PubMed

PubMed Search

ArXiv

Arxiv Search ✓ ADDED

WebScrapers

Web Scraper

Web Scraper

YahooFinance

Ticker ✓ ADDED

News ✓ ADDED

Analytics

ChartGenerator

Pie Chart ✓ ADDED

Bar Chart ✓ ADDED

Linear Chart ✓ ADDED

YouTube

Video statistics

Feishu

Send Group Message

takin.ai

Explore Studio / FinBot Knowledge Tools

2951.93 Harry

Agent Settings gpt-4o-mini CHAT Publish

FinBot AGENT

Orchestrate

Logs & Ann.

Monitoring

Instructions

Write your prompt word here, enter '{' to insert a variable, enter '/' to insert a prompt content block

Variables

Context

Tools

8/8 Enabled

yahoo yahoo\_finance\_ticker

yahoo yahoo\_finance\_analytics

yahoo yahoo\_finance\_news

google google\_search

arxiv arxiv\_search

chart pie\_chart

chart bar\_chart

chart line\_chart

Vision

Debug & Preview

Used yahoo\_finance\_news

Here are some of the latest news articles related to Tesla:

- OpenAI's board rejects bid from Elon Musk-led group** (Published: Feb 14, 2025)
  - Tesla CEO Elon Musk's consortium made a \$97.4 billion bid for OpenAI, which was unanimously rejected by the board. The board stated that OpenAI is not for sale, emphasizing its nonprofit mission.
- Elon Musk's Government Gig Has No Salary, No Job Title, And Plenty of Power** (Published: Feb 14, 2025)
  - Musk has taken on an unofficial role within the government identified only as "unlisted," causing confusion regarding his actual responsibilities and authority.
- Car Tariffs May Start on April 2, Trump Says** (Published: Feb 14, 2025)
  - President Trump announced potential tariffs on car imports starting around April 2, which may impact companies like Tesla that rely on foreign markets for vehicle production.
- Tesla Stock Couldn't Go for 3 Straight Gains. At Least It's Finally Below Record Price Targets** (Published: Feb 14, 2025)
  - Tesla's shares experienced a slight pullback, which brings them below the average analyst target price, indicating mixed investor sentiment.
- Ron Baron's Strategic Moves: Tesla Inc. Sees a -0.24% Portfolio Impact** (Published: Feb 14, 2025)
  - Ron Baron, a prominent investor, made adjustments to his portfolio that included Tesla, noting a portfolio impact of approximately -0.24%.

These articles reflect ongoing developments related to Tesla and its CEO, Elon Musk, alongside some external factors influencing the automotive and tech industries.

54.64s Token spent 6,611 - 04:16 PM

Talk to Bot





Finbot  
AGENT

Orchestrate

Logs & Annotations

### Orchestrate

#### INSTRUCTIONS

Generate

Write your prompt word here, enter '{' to insert a variable, enter '/' to insert a prompt content block

0

#### Variables

+ Add

Variables allow users to introduce prompt words or opening remarks when filling out forms. You can try entering "{{input}}" in the prompt words.

#### Knowledge

+ Add

You can import Knowledge as context

#### Tools

0/0 Enabled + Add

#### Vision

Settings

Agent Settings

gpt-4o-mini CHAT

Publish

### Debug & Preview

Talk to Bot



Features Enabled

Manage



These are agents.



# Jan. 2025



The screenshot shows a code editor with a file explorer on the left and a chat interface on the right. The code in the editor is as follows:

```
4 import anthropic
5
6 # Load environment variables
7 load_dotenv() # take environment variables from .env
8
9 # Model settings
10 ant_model = "claude-3-5-haiku-latest" # Best quality Anthropic model
11 oai_model = "gpt-4o-mini" # Best quality OpenAI model
12
13 # Conversation settings
14 language = "english" # conversation language
15 num_of_turns = 5 # number of turns in the conversation
16
17 # Cost tracking
18
19
20 oai_cost = 0
21 ant_cost = 0
22
23 def setup_podcast_roles():
24     """Setup the system prompts for the conversation"""
25     ant_system_prompt = f"you are a female painter and is going to chat with another person in {language} don't include any tone or scene description in the chat, and"
26     oai_system_prompt = f"you are a male musician and is going to chat with another person in {language}, don't say your gender and now start chatting"
27
28     return ant_system_prompt, oai_system_prompt
29
30 def get_anthropic_response(messages):
31     """Get response from Anthropic's Claude model"""
32     client = anthropic.Anthropic()
33
34     # Convert messages to the format expected by Anthropic
35     ant_messages = []
36     for msg in messages:
37         if msg["role"] == "user":
38             ant_messages.append({"role": "user", "content": msg["content"]})
39         elif msg["role"] == "assistant":
40             ant_messages.append({"role": "assistant", "content": msg["content"]})
41
42     response = client.messages.create(
43         model=ant_model,
44         max_tokens=1024,
```

The chat interface on the right shows a message from the user: "change the prompts to ant\_system\_prompt=f'you are a female painter and is going to chat with another person in {language} don't include any tone or scene description in the chat, and don't say your gender and now start chatting' oai\_system\_prompt=f'you are a male musician and is going to chat with another person in {language}, don't say your gender and now start chatting'". Below this is a response from the AI: "I'll update the prompts in the setup\_podcast\_roles function: Edited /curve +3 -12 102 Open diff". A list of changes is shown: "I've updated the prompts to: 1. Changed the Claude (Anthropic) role to a female painter 2. Changed the GPT (OpenAI) role to a male musician 3. Removed the podcast and Elon Musk related content 4. Kept the instructions about not including tone/scene descriptions and not mentioning gender". At the bottom, there is a table with token usage: "add comments about the price and update the program: gpt-4o-mini \$0.150 / 1M input tokens \$0.600 / 1M output tokens" and "Claude 3.5 Haiku \$0.80 / Mtok input \$4 / Mtok Output".





Andrej Karpathy

@karpathy



There's a new kind of coding I call "vibe coding", where you fully give in to the vibes, embrace exponentials, and forget that the code even exists. It's possible because the LLMs (e.g. Cursor Composer w Sonnet) are getting too good. Also I just talk to Composer with SuperWhisper so I barely even touch the keyboard. I ask for the dumbest things like "decrease the padding on the sidebar by half" because I'm too lazy to find it. I "Accept All" always, I don't read the diffs anymore. When I get error messages I just copy paste them in with no comment, usually that fixes it. The code grows beyond my usual comprehension, I'd have to really read through it for a while. Sometimes the LLMs can't fix a bug so I just work around it or ask for random changes until it goes away. It's not too bad for throwaway weekend projects, but still quite amusing. I'm building a project or webapp, but it's not really coding - I just see stuff, say stuff, run stuff, and copy paste stuff, and it mostly works.

6:17 PM · Feb 2, 2025 · 4.1M Views

# Vibe Coding ;)

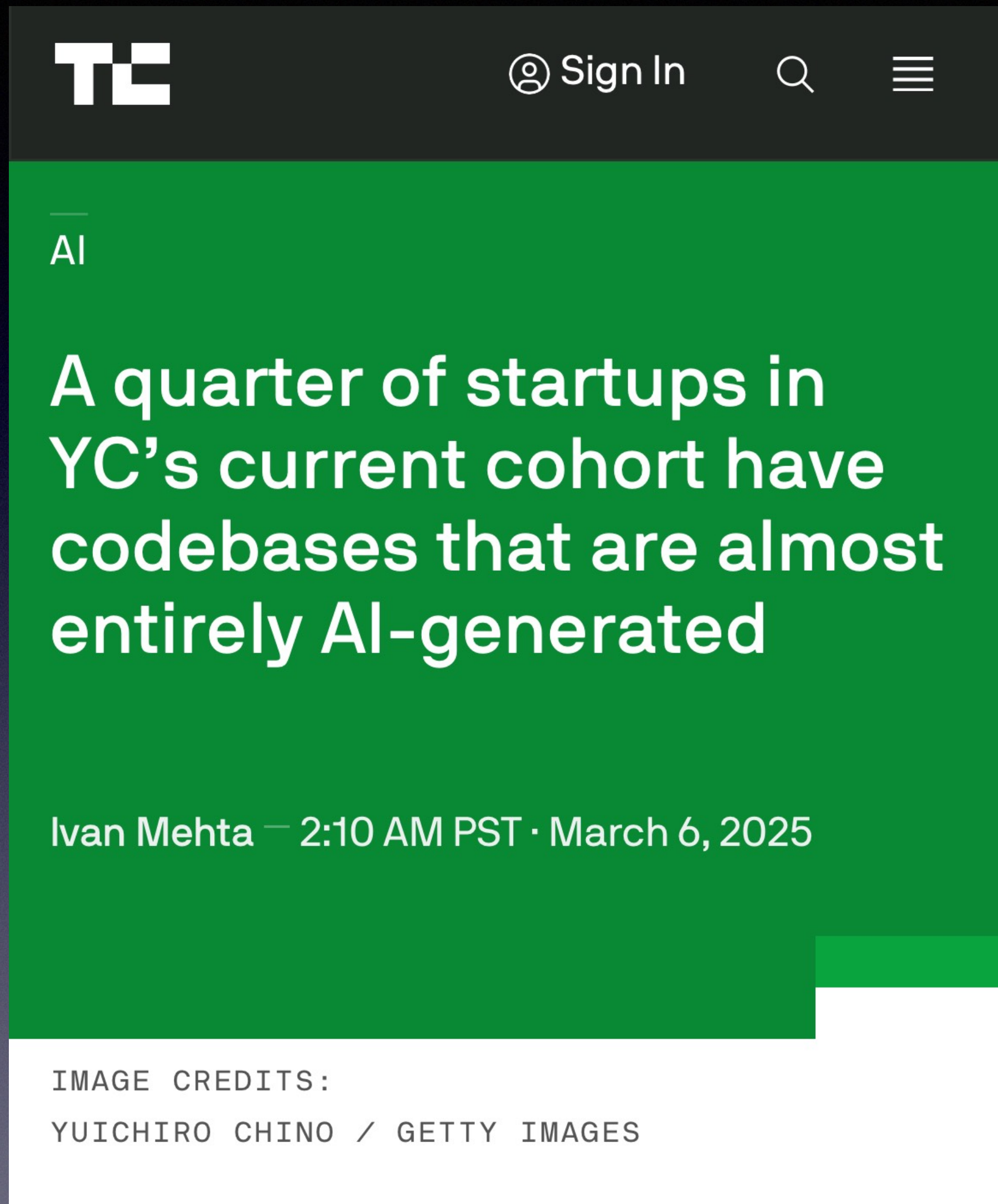
“[vibe coding] where you fully give in to the vibes, embrace exponentials, and forget that the code even exists ... I just see stuff, say stuff, run stuff, and copy paste stuff, and it mostly works.”

In Feb, 2025, I did a live "vibe coding" session with [a non-tech unicorn startup CEO] over coffee and waffles in a nice restaurant.

A personal website went live in one hour

- Nextjs and Tailwind used
- Code hosted to Github
- Deployed on Vercel





Project	Total Tokens (Million)	Files Processed	Estimated Cost (GPT-4o)	Estimated Cost (Claude-3-7-Sonnet)
Dify	13.24	6852	\$132.45	\$198.67
OpenWebUI	10.44	4694	\$104.44	\$156.66
CivitAI	3.60	2719	\$36.01	\$54.01
PyTorch	34.32	16525	\$343.17	\$514.75

We completely refactored our codebase with Windsurf in under two months, led by a single full-stack developer.

Windsurf Cost: ~\$1000

Much better code - impossible without AI

Our codebase: ~1.7 million tokens

I vibe coded the python package to count the tokens in one hour:  
<https://pypi.org/project/aitokencount/>





**Harry Wang** · You

Professor at University of Delaware | CEO ...

4w · Edited · 🌐



Announcement 😊 : I am officially naming our Windsurf agent "Kim" (a name suggested by my daughter) - Kim has become an integral part of our development process, functioning almost like an employee. My daily meetings with my human developer, Faye, now feel like a three-person collaboration: Faye and I hold voice



**Harry Wang** · You

Professor at University of Delaware | CEO at Takin.ai

3w · 🌐



Spent the afternoon and evening redesigning many pages for the upcoming new version of **Takin.ai** (packed with exciting new features) with Windsurf using shadcn/ui - Claude 3.5 Sonnet has great design taste out of the box, if you give right examples and articulate what you want, it's much better than most UI designers I have worked with and much much cheaper.

Jobs: "Ultimately, it comes down to taste"...

[#windsurf](#) [#claude](#) [#design](#) [#ui](#) [#shadcn](#)





INNOVATION > AI

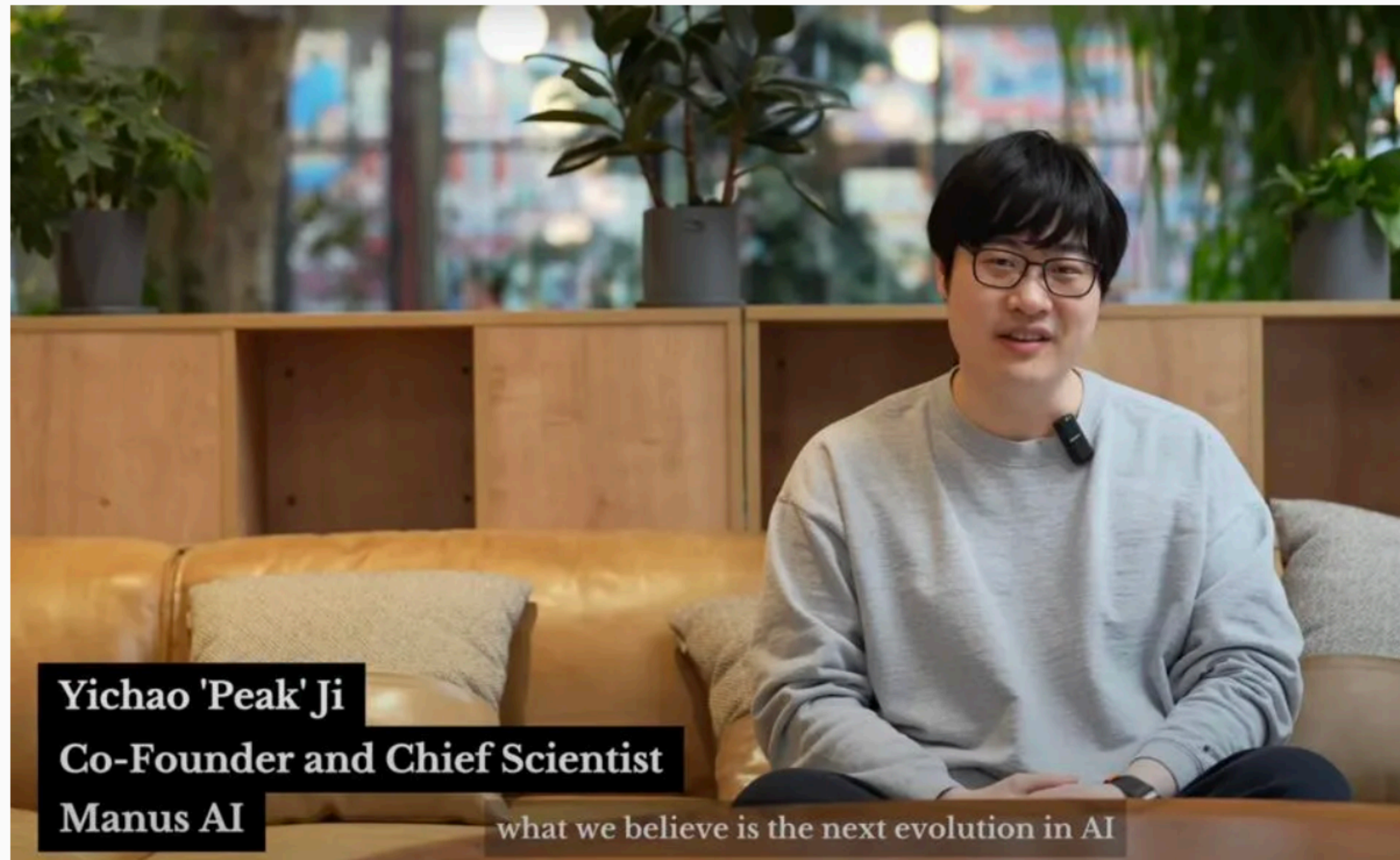
# China's Autonomous Agent, Manus, Changes Everything

By [Craig S. Smith](#), Contributor. ⓘ Craig S. Smith, Eye on AI host and form... ▾

[Follow Author](#)

Mar 08, 2025, 03:29pm EST

[Share](#) [Save](#) [Comment 7](#)



Yichao 'Peak' Ji, co-founder and chief scientist of Manus AI MANUS AI



“get all blogs from <https://harrywang.me/blog> and let me know the stats about the blogs, total, time span, contents, etc., present a final report in markdown file”

## Blog Analysis Report

### Total Number of Blogs

- 43

### Time Span of Blogs

- From September 07, 2019 to February 27, 2025 (2000 days)

### Blog Contents

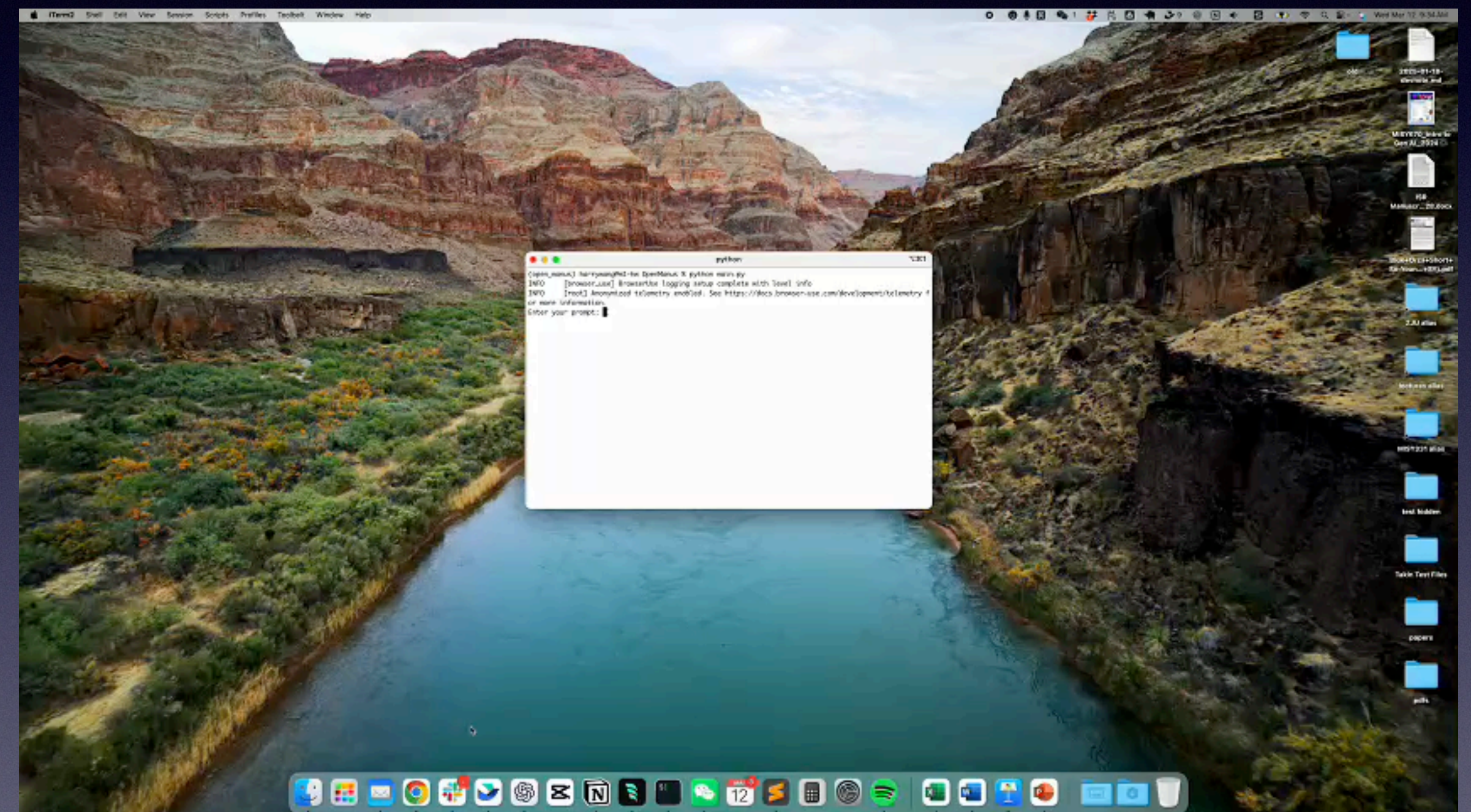
- Deploying DeepSeek with Ollama + LiteLLM + OpenWebUI - February 27, 2025
- The System Prompt of Cline (An Open-Source AI Coding Agent) - February 09, 2025

....

### Blog Distribution by Year

- 2025: 3 blogs
- 2024: 9 blogs
- 2023: 8 blogs
- 2022: 16 blogs
- 2021: 3 blogs
- 2020: 2 blogs
- 2019: 2 blogs

Got the result after many failed attempts  
A failed attempt shown below: ~90 seconds



Complete log: <https://gist.github.com/harrywang/a1f0cae9583109612ba17799e4b40538>





These are agents with  
autonomous workflows.



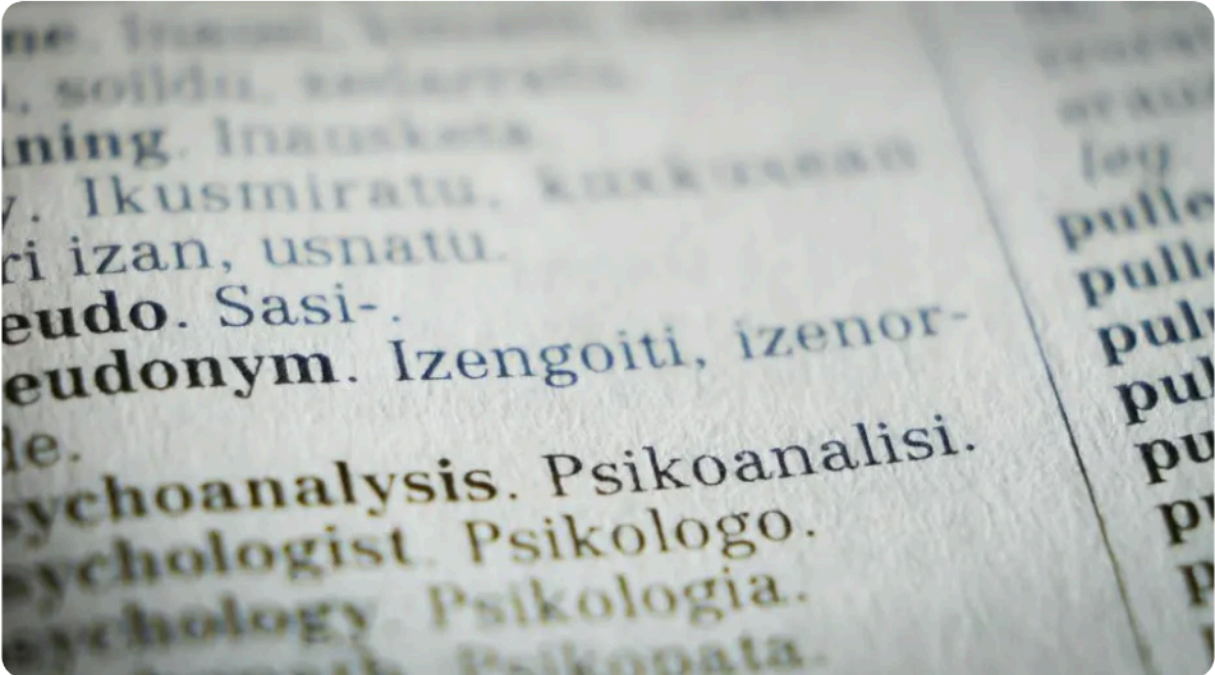
takin.ai

Apps Pricing Blog Prompts Sign In Sign Up

# Agentic Workflow Tutorial: Remaking Andrew Ng's Translation Agents with No-Code LLM Workflow

August 28 2024

Share on : [f](#) [x](#) [t](#) [in](#)



takin.ai

Explore Studio / Agentic Translator v1.0 Knowledge Tools

1929.8 H Harry

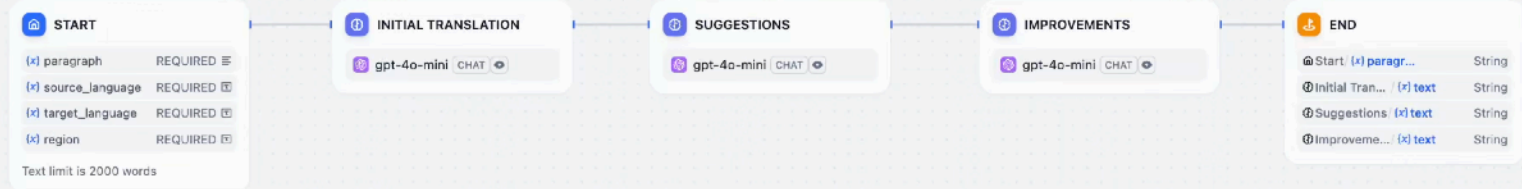
Agentic Translat... WORKFLOW

Auto-Saved 10:11:53 · Published a few seconds ago

Orchestrate

Logs

Monitoring



```

graph LR
    START[START] --> INITIAL[INITIAL TRANSLATION]
    INITIAL --> SUGGESTIONS[SUGGESTIONS]
    SUGGESTIONS --> IMPROVEMENTS[IMPROVEMENTS]
    IMPROVEMENTS --> END[END]
  
```

Test Run

INPUT RESULT DETAIL TRACING

paragraph

Ground Control to Major Tom  
Ground Control to Major Tom  
Take your protein pills and put your helmet on

Ground Control to Major Tom  
Commencing countdown, engines on  
Check ignition and may God's love be with you

This is Ground Control to Major Tom  
You've really made the grade  
And the papers want to know whose shirts you wear  
Now it's time to leave the capsule if you dare

'This is Major Tom to Ground Control  
'I'm stepping through the door

source\_language

english

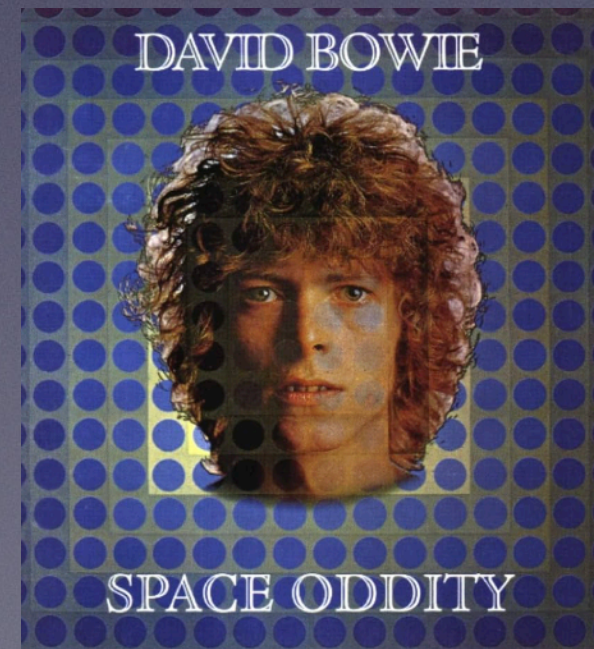
target\_language

chinese

region

hk

Start Run





These are agents with  
pre-defined workflows.



## Agentic Reasoning Design Patterns

### 1. Reflection

- Self-Refine: Iterative Refinement with Self-Feedback, Madaan et al. (2023)
- Reflexion: Language Agents with Verbal Reinforcement Learning, Shinn et al., (2023)

### 2. Tool use

- Gorilla: Large Language Model Connected with Massive APIs, Patil et al. (2023)
- MM-REACT: Prompting ChatGPT for Multimodal Reasoning and Action, Yang et al. (2023)

### 3. Planning

- Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, Wei et al., (2022)
- HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face, Shen et al. (2023)

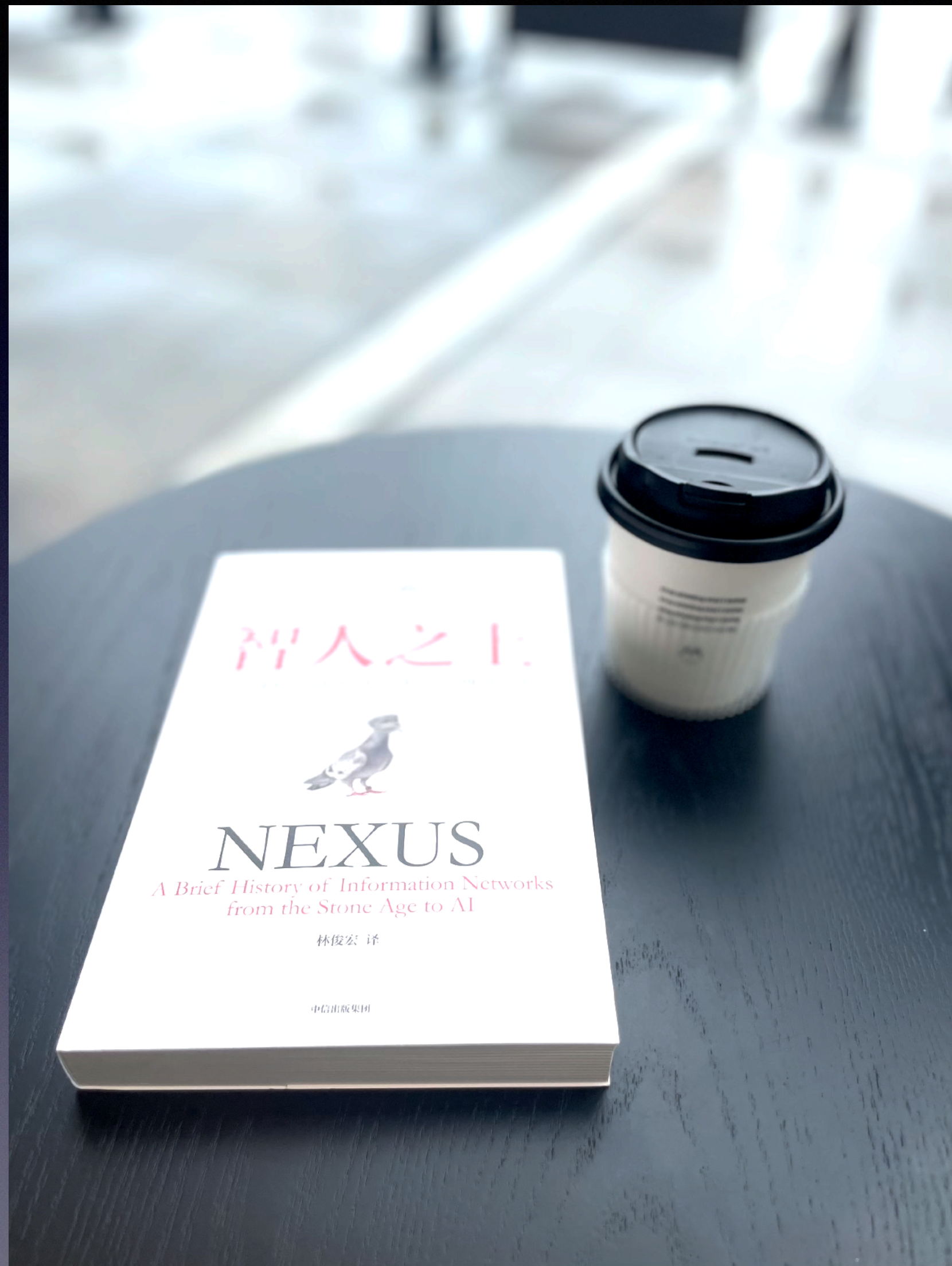
### 4. Multi-agent collaboration

- Communicative Agents for Software Development, Qian et al., (2023)
- AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation, Wu et al. (2023)

Andrew Ng

Sequoia Capital's AI Ascent  
March 26, 2024





Sep. 2024

**Animal**

Creation and use of tools

**Homo Sapiens**

Large-scale collaboration  
via information networks

**Nexus**



**Chatbots**



Use of tools

**Agents**



Coordinated Collaboration

Autonomous Collaboration

**Agentic Workflows**



Agents & Workflows  
Are **NOT** New.



What Changed?



Data + Compute + Model  $\xrightarrow{\text{Scaling Laws}}$  LLMs



Show me the code

—> Show me the prompt  
Enable rapid experiments to craft agent behaviors

Deterministic (Rule-Based)

—> Non-deterministic (Generative)  
Enable autonomous tool calling, planning, reflection

Expensive Intelligence

—> Cheap Intelligence

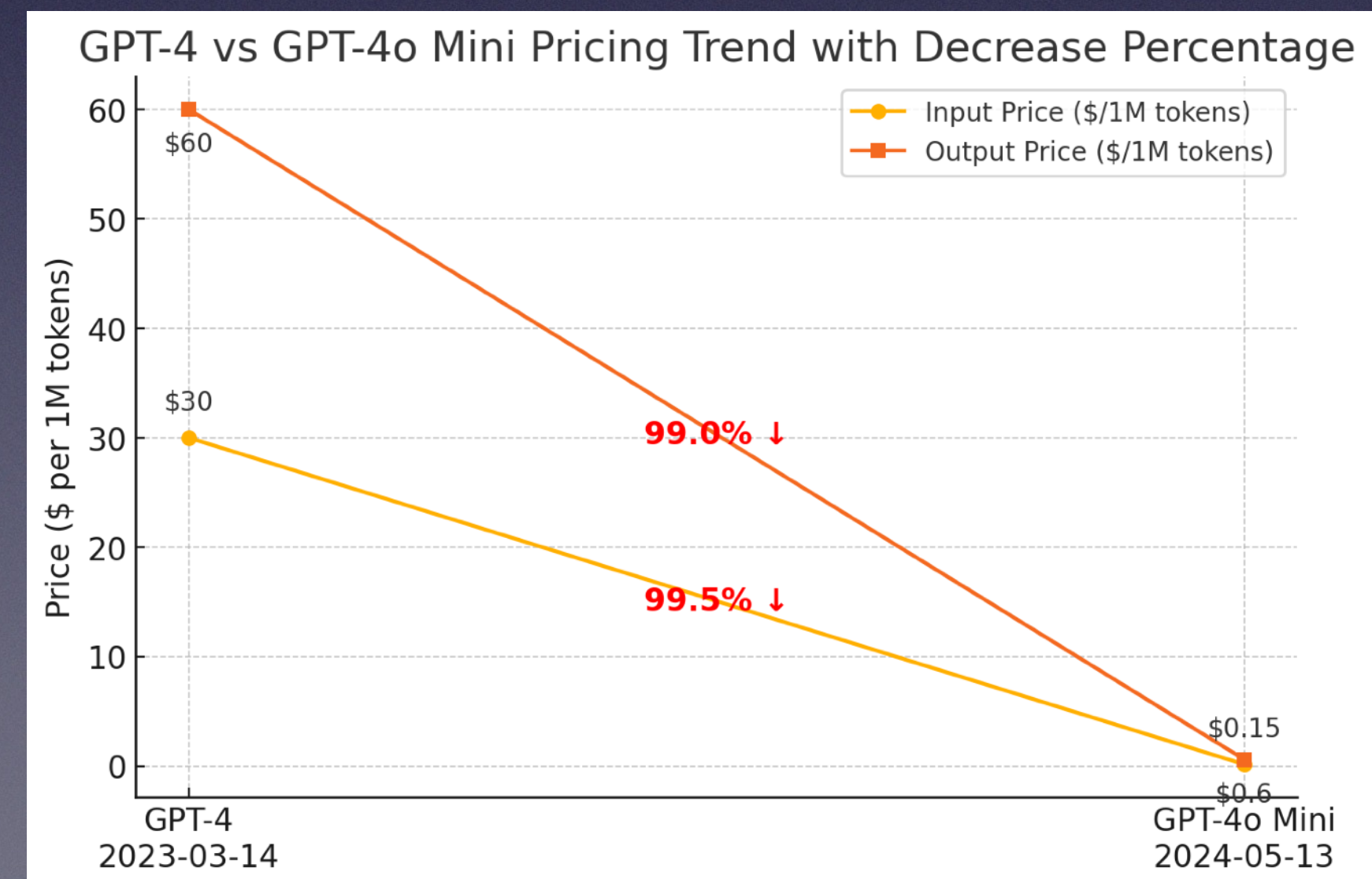


**Andrej Karpathy** @karpathy · Jan 24  
The hottest new programming language is English



r/programminghorror  
iZer0Cool

**if-else hell**

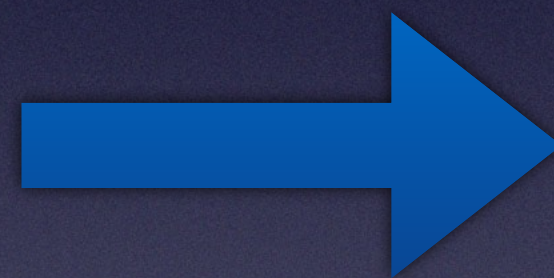




# Democratization of Innovation

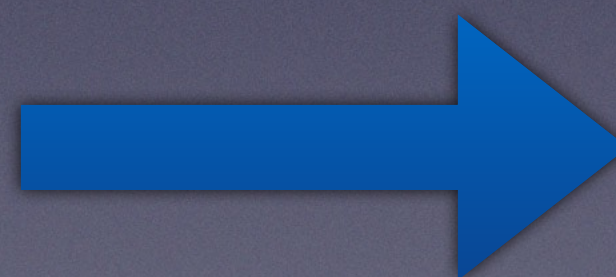
No-code/low-code AI workflow tools are enabling business professionals to take on tasks traditionally reserved for tech experts.

Tech Team



Develop Various Internal  
Tools for Agents

Business Team  
(Sales, Product, Analyst)



Create Agentic Workflows for  
Various Business Scenarios



# AI Training

## **1. Productivity-focused AI Training for Everyone**

(Prompt engineering, no-code agentic workflow, ...)

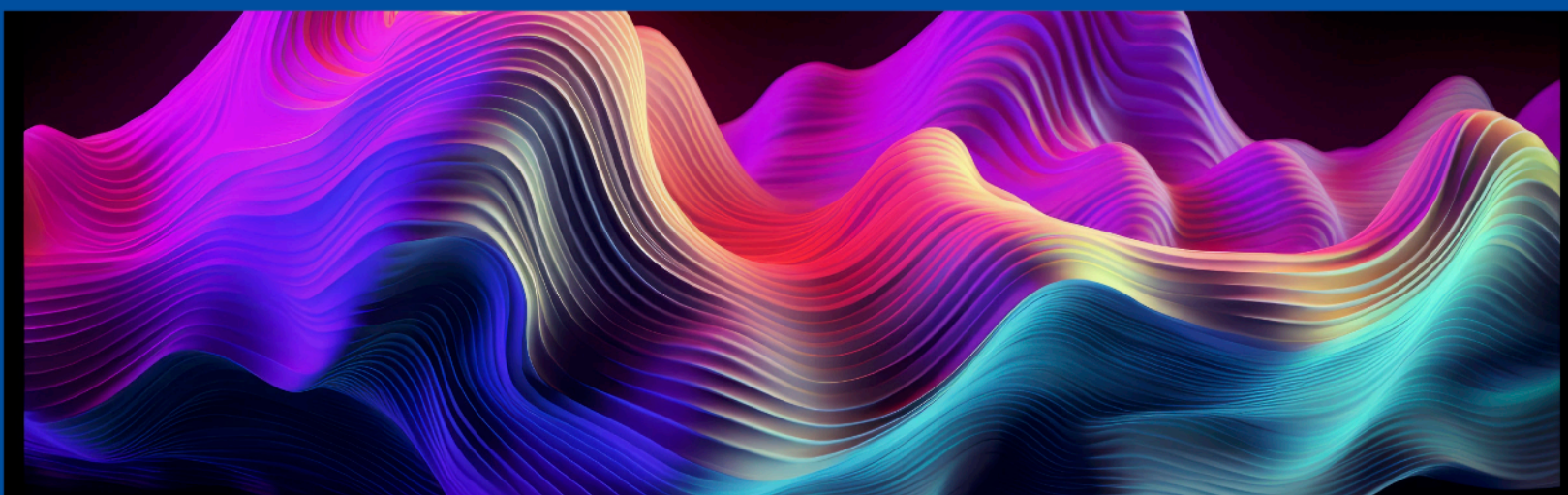
## **2. Technology-focused AI Training for Developers**

(AI coding, data analytics, visualization, ...)

## **3. Strategy-focused AI Training for Managers**

(AI tech trends, future organization and workforce, responsible AI, compliance ...)





MISY670

# INTRO TO GENERATIVE AI

Explore the transformative power of Generative AI in this hands-on, business-focused course, designed to equip you with practical skills and strategic insights for leveraging GenAI to drive innovation.

## TOPICS

- Large Language Model (LLM) Overview
- Prompt Engineering
- LLM Evaluation
- Retrieval-Augmented Generation
- Finetuning
- Data Analysis and Visualization
- Multimodal LLMs
- Diffusion Models
- AI Agents and Agentic Workflows
- Multi-agent Systems
- AI Safety and Responsible AI

## COURSE DETAILS

- Spring B, 2025
- Synchronous Online
- Part of UD's Graduate Certificate, Generative AI for Business
- Taught by Harry Wang, Professor of MIS, Founder and CEO, Takin.ai. [harrywang.me](http://harrywang.me)



[CLICK HERE TO REGISTER](#)



## CONTACT

Patrick Cahill  
Director, Admissions & Recruitment  
Lerner Graduate Programs  
[pcahill@udel.edu](mailto:pcahill@udel.edu)



# Empower Your Team with GenAI Training

Equip your workforce with expert-led Generative AI training designed to drive innovation, enhance efficiency, and build practical AI skills. Trusted by top business schools and leading companies, our tailored workshops prepare your team to leverage AI for real-world success.

## Training Format

All courses are available in both online and in-person formats to accommodate different learning preferences and geographical locations.

## Course Duration

- Half-day sessions: Perfect for basic topics and quick skill acquisition
- Full-day workshops: Comprehensive coverage with hands-on practice
- Two-day intensive: Deep dive into advanced topics with extensive practical work

## TOPICS

Large Language Model (LLM) Fundamentals

Prompt Engineering Mastery

Data Analytics & Visualization with AI

RAG & Fine-tuning Workshop

AI Agents & Agentic Workflows

Multi-Agent Systems Development

AI Content Creation (Image, Music, Video)

\* All trainings include hands-on exercises and real-world case studies

Your name \*

Email \*

Company name \*

How did you hear about us?

Message \*

[Contact Us for Training](#)

Starting on April 1, 2025

[support@takin.ai](mailto:support@takin.ai)



Q&A