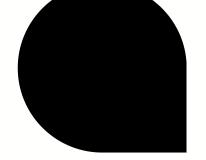
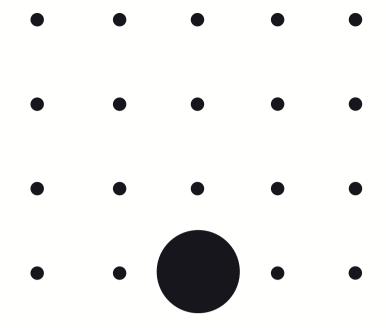
CONVERSATIONAL AI DAYS 2025

DEMYSTIFYING **AI AGENTS**

Juan Carlos Quintero







2025 be like...



2025 be like...



2025 be like...





Agenda

- Who am I?
- Why agents? Why now?
- Al Agents
- Tools use
- How to build an agent
- Latest developments
- Thanks!

Who am I?

I'm JC

- Born in Venezuela 🔤
- Studied Software Engineering
- Worked for 10+ years as Engineer
- Working now as AI Consultant
- Voiceflow Certified Expert
- YouTuber 😅



Montevideo, Uruguay





Berlin, Germany



Munich, Germany

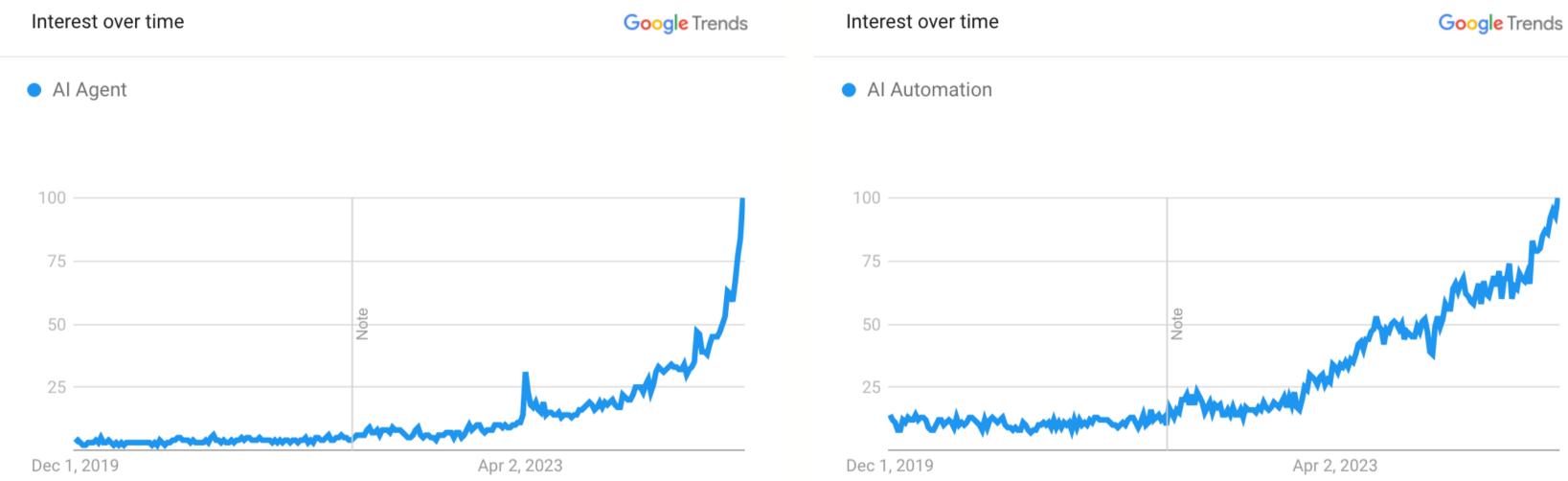


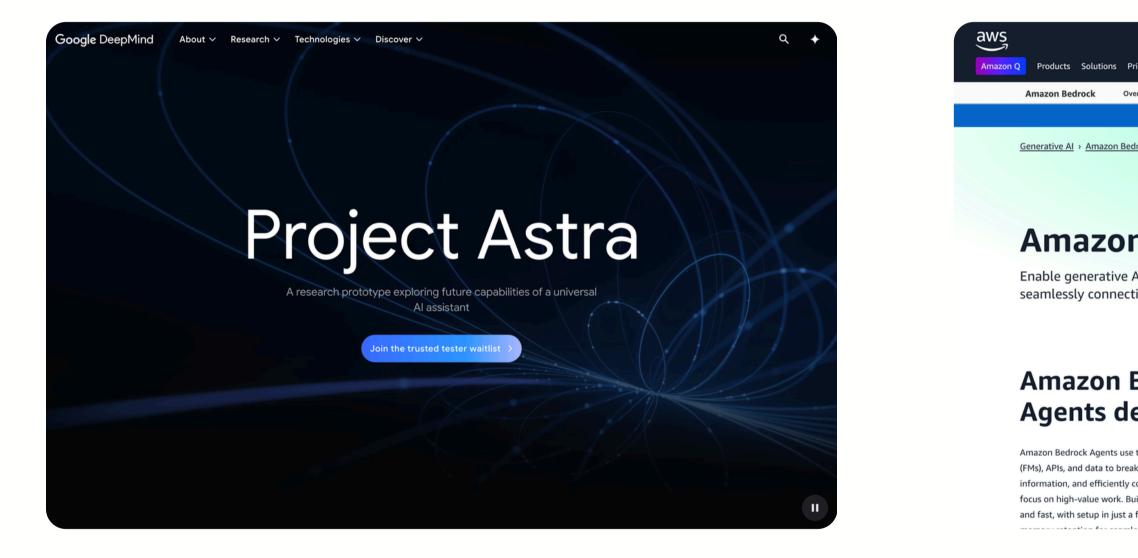




Cancun, Mexico







			About AWS	Contact Us	Support 🗸	English 🗸	My Account 🗸	Sign In	Create an AWS Account	8	
cing	Documentation	Learn	Partner Network	AWS Mark	cetplace C	ustomer Enablem	ent Events	Explore More	Q		
rview	Features 💌	Pricing	Model Providers 🔻	FAQs	Testimonial	s Resources					
DeepSeek-R1 now available fully managed in Amazon Bedrock Learn more »											
rock >	Agents										

Amazon Bedrock Agents

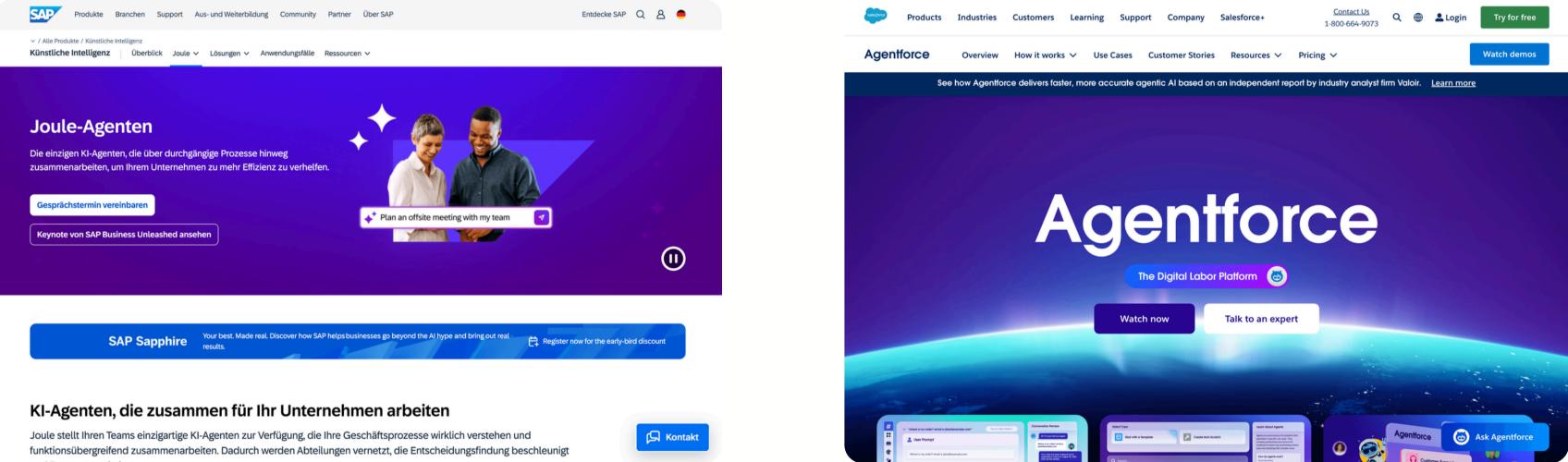
Enable generative AI applications to automate multistep tasks by seamlessly connecting with company systems, APIs, and data sources

Amazon Bedrock Agents demo

Ove

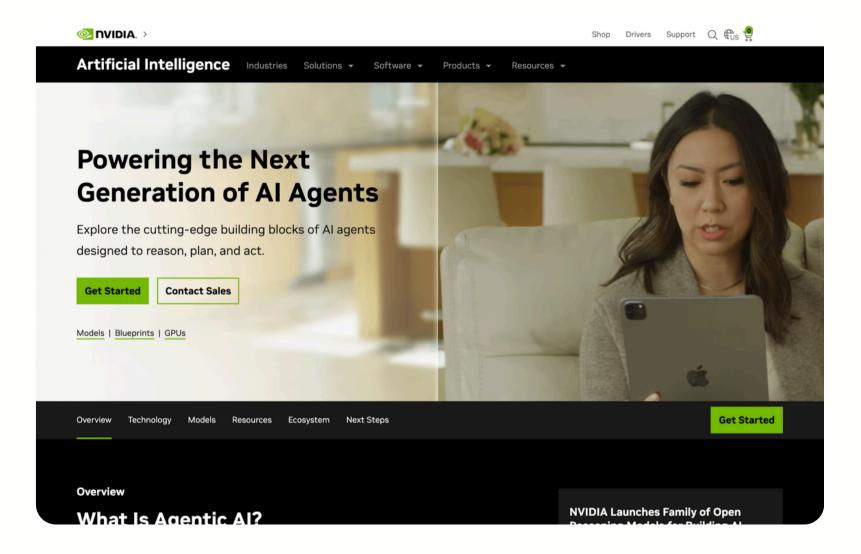
Amazon Bedrock Agents use the reasoning of foundation models (FMs), APIs, and data to break down user requests, gather relevant information, and efficiently complete tasks—freeing teams to focus on high-value work. Building an agent is straightforward and fast, with setup in just a few steps. Agents now include

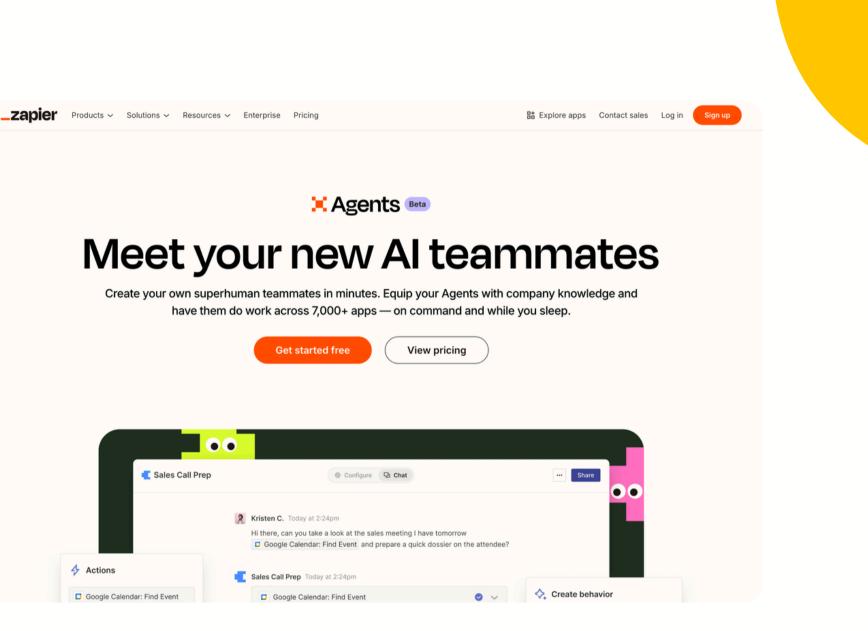


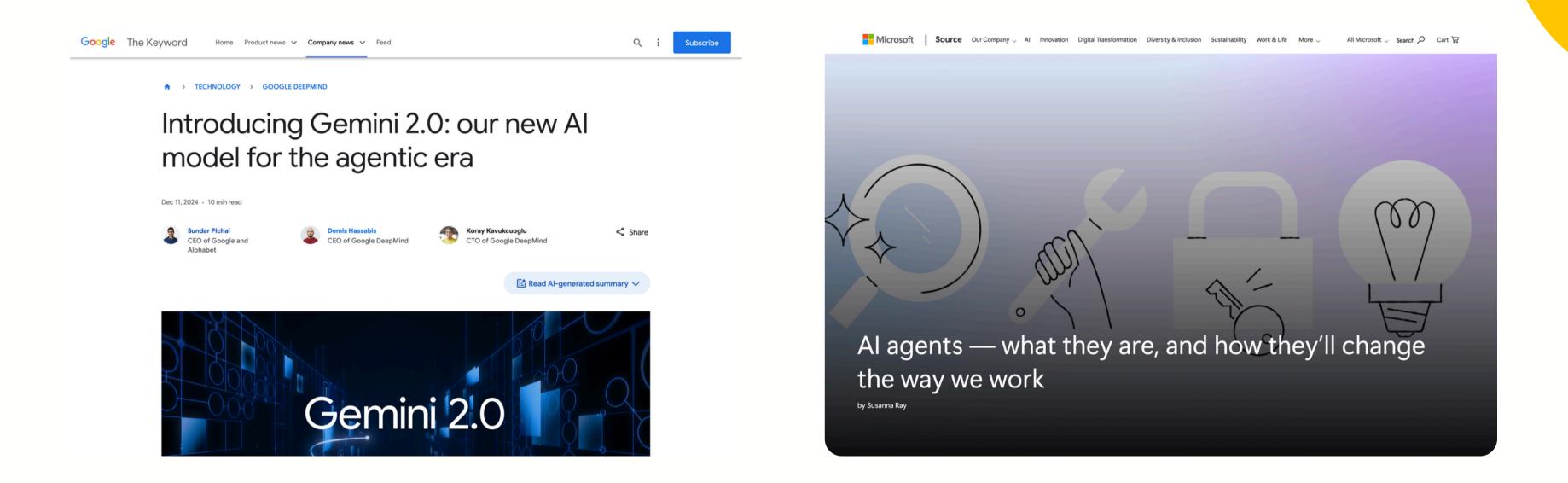


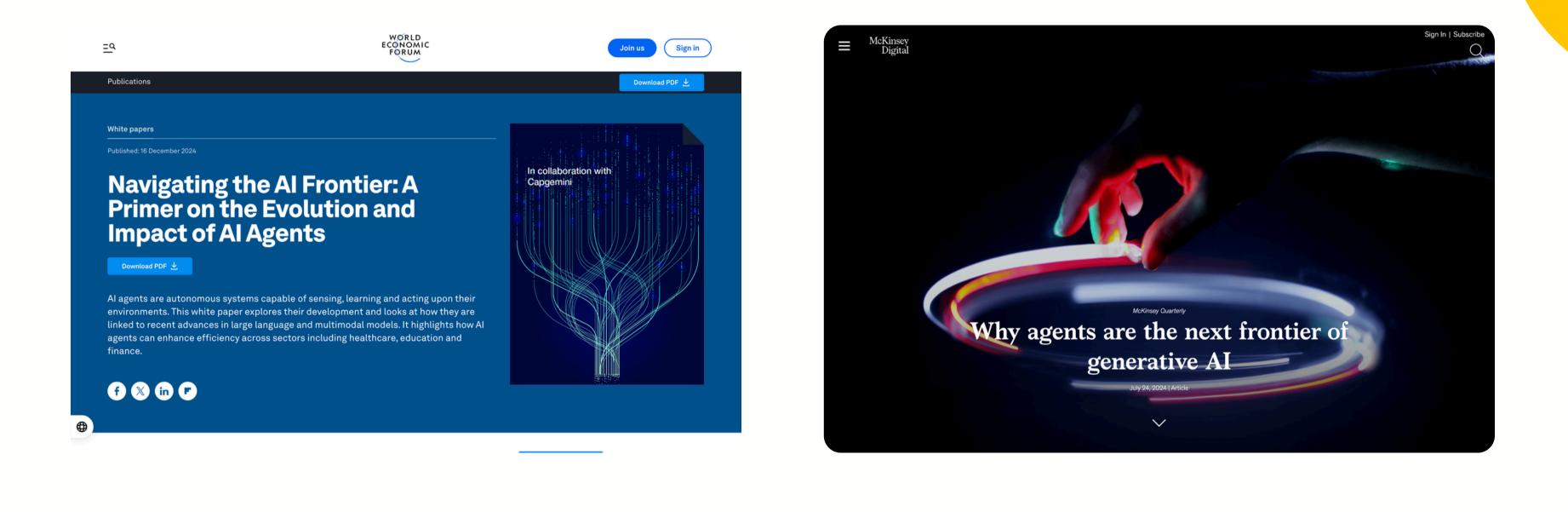
. .















Marc Benioff

MOONSHOTS









"Software that uses **AI** and **tools** to accomplish a goal requiring **multiple steps**"

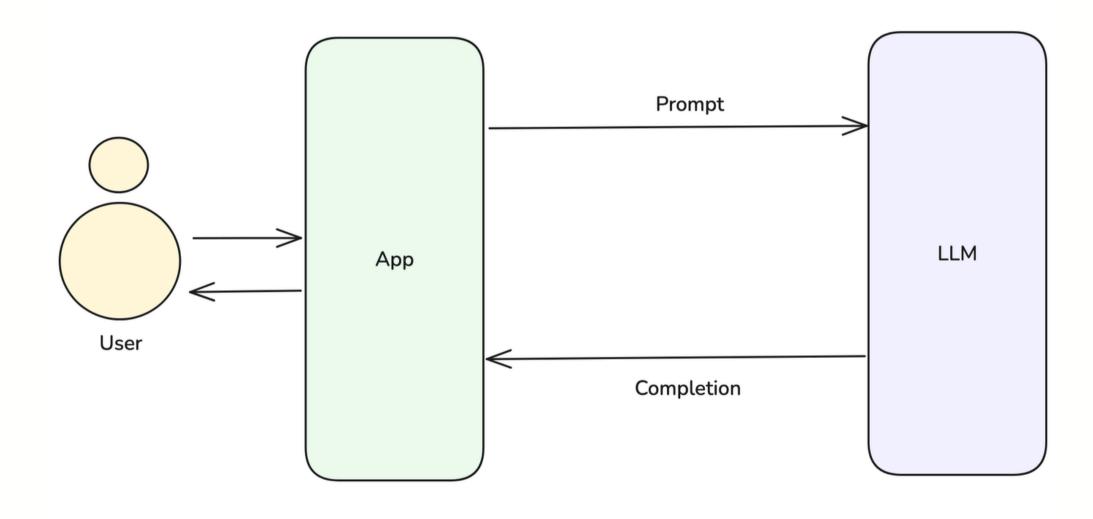
The Future Of AI Agents With Dharmesh Shah | INBOUND 2024 <u>https://www.youtube.com/watch?v=lityUpVVD38</u>



Manager:

"How many vacation days per year do we have?"

LLM (-)



LLM = Large Language Model (AI) AI = Artificial Intelligence

• • •

Manager:

"How many vacation days per year do we have?"

Bot:

"Your company offers 25 vacation days per year"

Manager:

"How many vacation days per year do we have?"

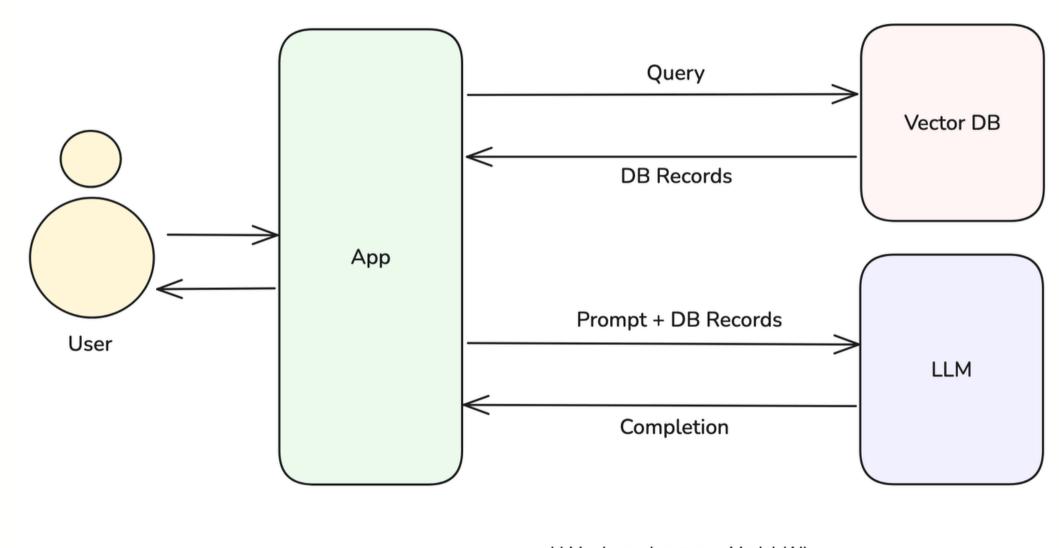
Bot:

"Your company offers 25 vacation days per year" X (hallucination!)

Manager:

"How many vacation days per year do we have?"

Knowledge + LLM (RAG)



LLM = Large Language Model (AI) AI = Artificial Intelligence RAG = Retrieval Augmented Generation DB = Database



Manager:

"How many vacation days per year do we have?"

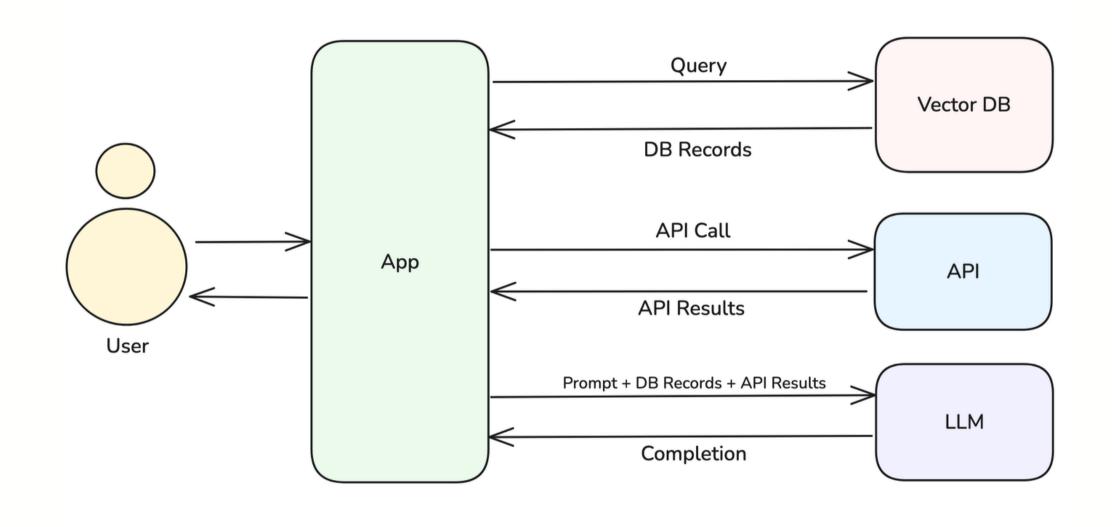
Bot:

"Your company offers 30 vacation days per year"

Manager:

"How many vacation days does Max have left to use this year?"

API + Knowledge + LLM (RAG-ish)



LLM = Large Language Model (AI) AI = Artificial Intelligence RAG = Retrieval Augmented Generation DB = Database API = Application Programming Interface





Manager:

"How many vacation days does Max have left to use this year?"

Bot: "20 days" 🗸

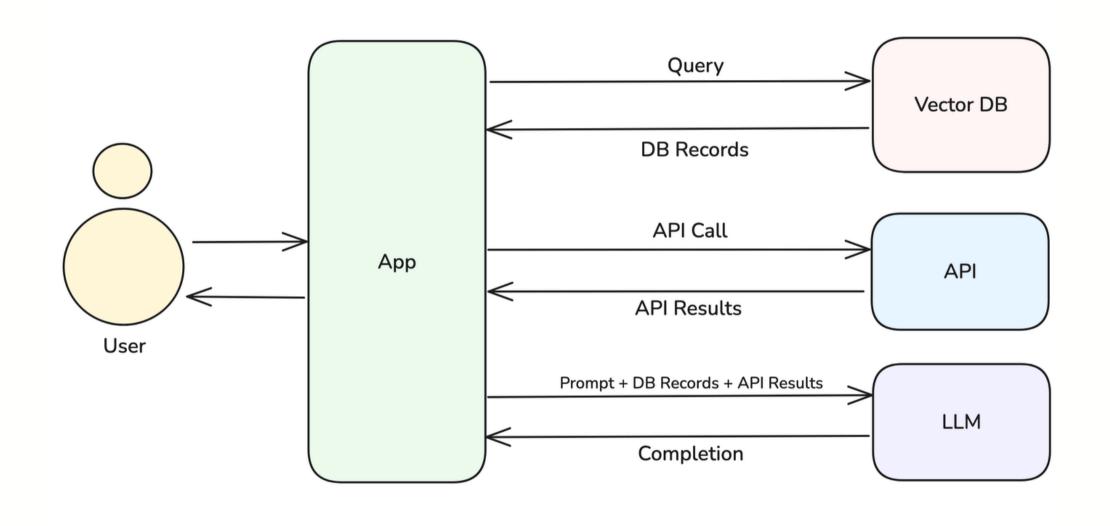
"accomplish a goal..."

Manager:

"Register five vacation days for Max from ..."



API + Knowledge + LLM (RAG-ish)

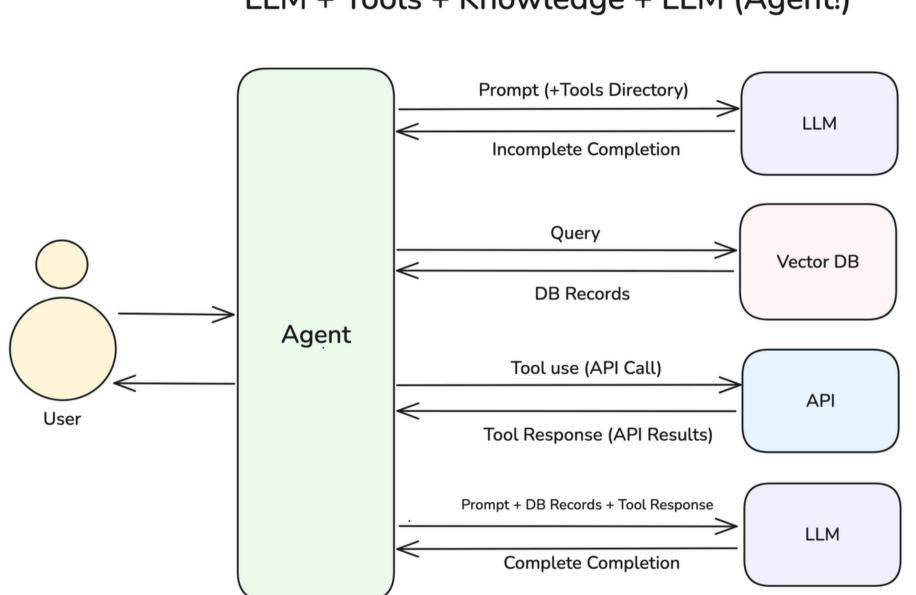


LLM = Large Language Model (AI) AI = Artificial Intelligence RAG = Retrieval Augmented Generation DB = Database API = Application Programming Interface





LLM + Tools + Knowledge + LLM (Agent!)



LLM = Large Language Model (AI) AI = Artificial Intelligence RAG = Retrieval Augmented Generation DB = Database API = Application Programming Interface

Manager:

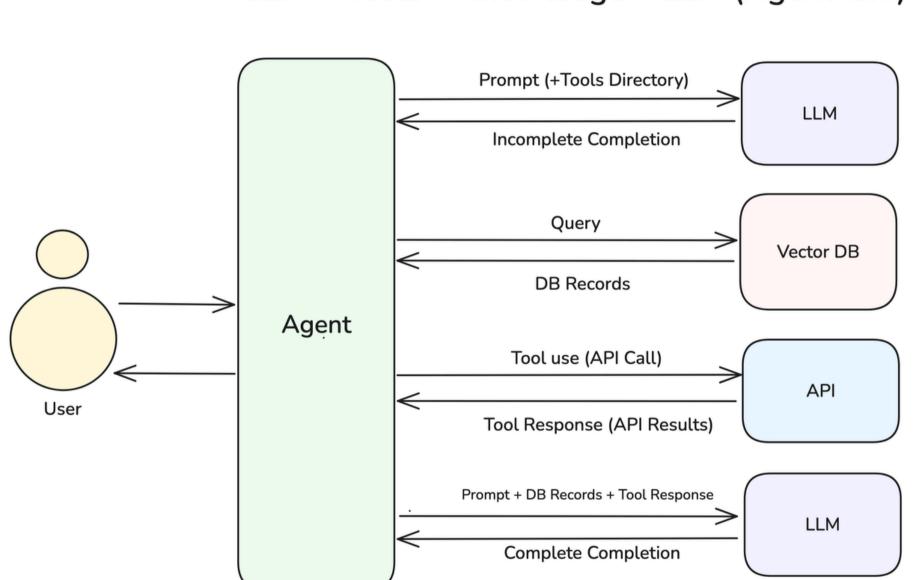
"Register **five** vacation days for **Max** from ..."

Bot:

"Done, 5 vacation days were registered for Max"



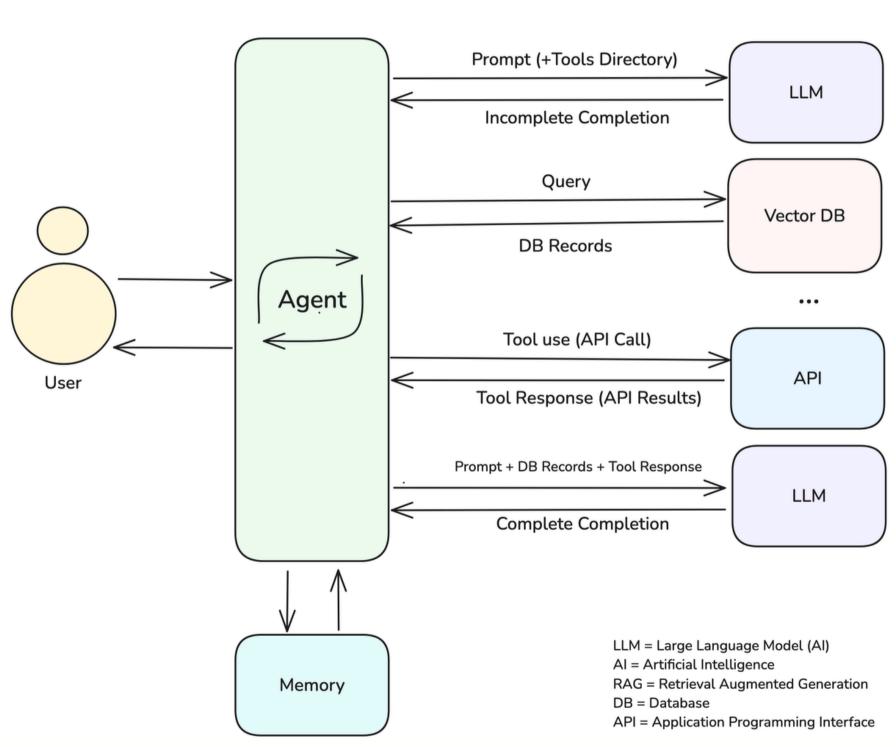
LLM + Tools + Knowledge + LLM (Agent-ish!)



LLM = Large Language Model (AI) AI = Artificial Intelligence RAG = Retrieval Augmented Generation DB = Database API = Application Programming Interface



LLM + Tools + Knowledge + LLM (Agent!)



Manager:

"Book a nice hotel in a top beach location in Venezuela for our company retreat next October"

(Iron Man's Jarvis type-of-agent...)



Manager:

"Book a nice hotel in a top beach location in Venezuela for our company retreat next October"

Bot:

"Done, I have booked 34 double rooms for 67 confirmed employees and 8 single rooms for Executives and C-Level, in Alba Hotel located in Los Roques, a top island destination in Venezuela, from October 18th until October 26th. Invoice is in your inbox, you have 72 hours to make the payment" 😿

Versus

- Al Agents vs Embodied Agents
- Al Agents vs Al Chatbots
- Al Agents vs Al Automation
- Al Agents vs Agentic



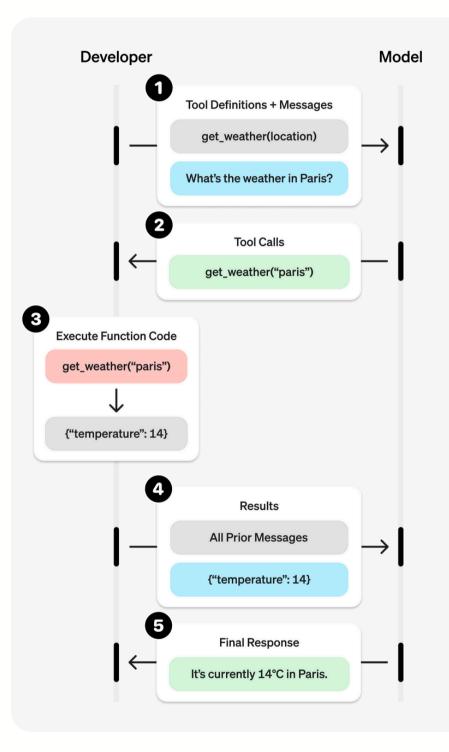
Tools use a.k.a. Function Calling

Anthropic

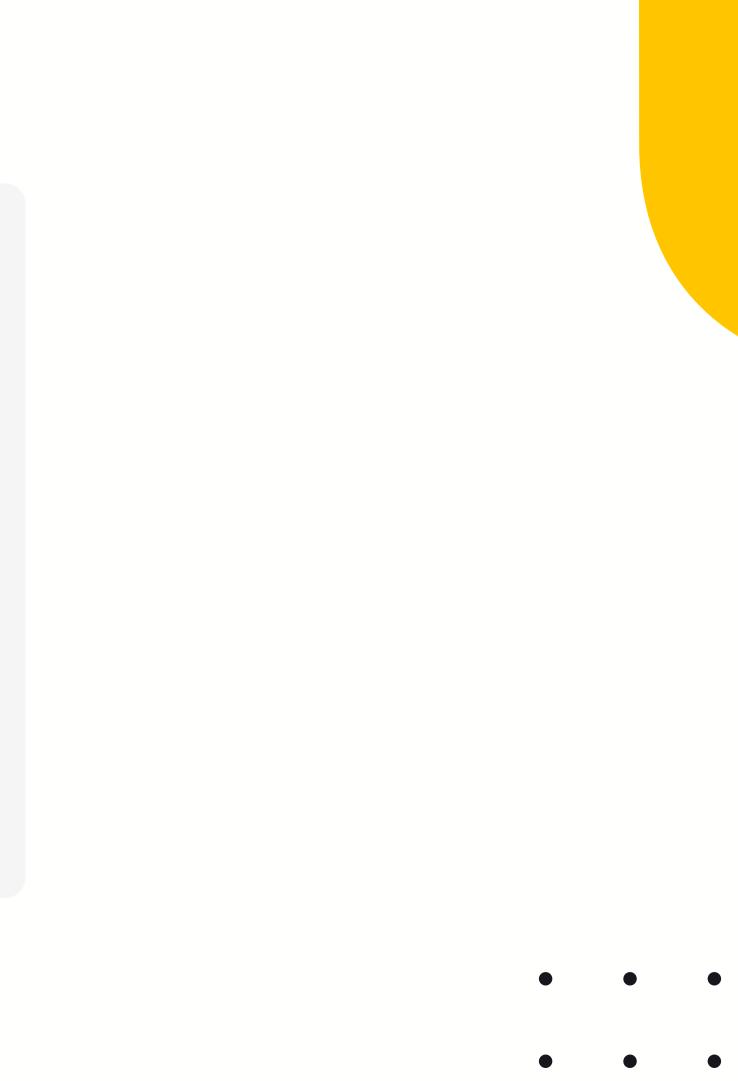
"Claude is capable of interacting with external client-side tools and functions, allowing you to equip Claude with your own custom tools to perform a wider variety of tasks."

OpenAl

"Function calling provides a powerful and flexible way for OpenAI models to interface with your code or external services"



https://platform.openai.com/docs/guides/function-calling?api-mode=responses



Tools use Prompt (+ Tools Directory)

```
"model": "claude-3-7-sonnet-20250219",
"max_tokens": 1024,
"tools": [{
    "name": "get_weather",
    "description": "Get the current weather in a given location",
    "input_schema": {
        "type": "object",
        "properties": {
            "location": {
                "type": "string",
                "description": "The city and state, e.g. Berlin, CA"
            "unit": {
                "type": "string",
                "enum": ["celsius", "fahrenheit"],
                "description": "The unit of temperature, either \"celsius\" or \"fahrenheit\""
        },
        "required": ["location"]
}],
"messages": [{"role": "user", "content": "What is the weather like in Berlin?"}]
```

Incomplete Completion (Use tool "X")

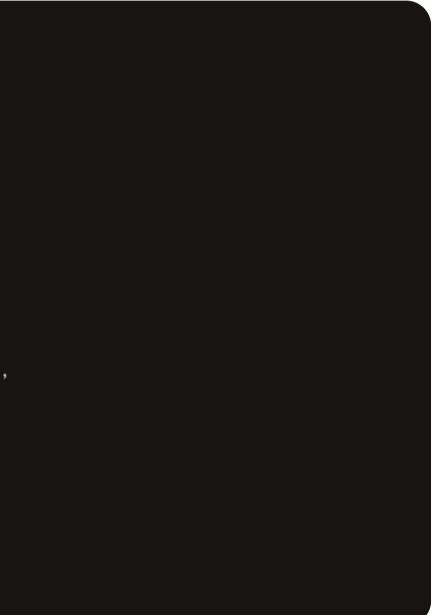
```
"id": "msg_01Aq9w938a90dw8q",
  "model": "claude-3-7-sonnet-20250219",
  "stop_reason": "tool_use",
  "role": "assistant",
  "content": [
      "type": "text",
      "text": "<thinking>I need to call the get_weather function, and the user wants Berlin, which
is likely Berlin, Germany.</thinking>"
    },
      "type": "tool_use",
      "id": "toolu_01A09q90qw90lq917835lq9",
      "name": "get_weather",
      "input": {"location": "Berlin, Germany", "unit": "celsius"}
```

Tool execution

if (type === "tool_use" && name === "get_weather") { const response = fetch_weather(input) // API Call get_llm_completion(prompt, response) // Reprompt

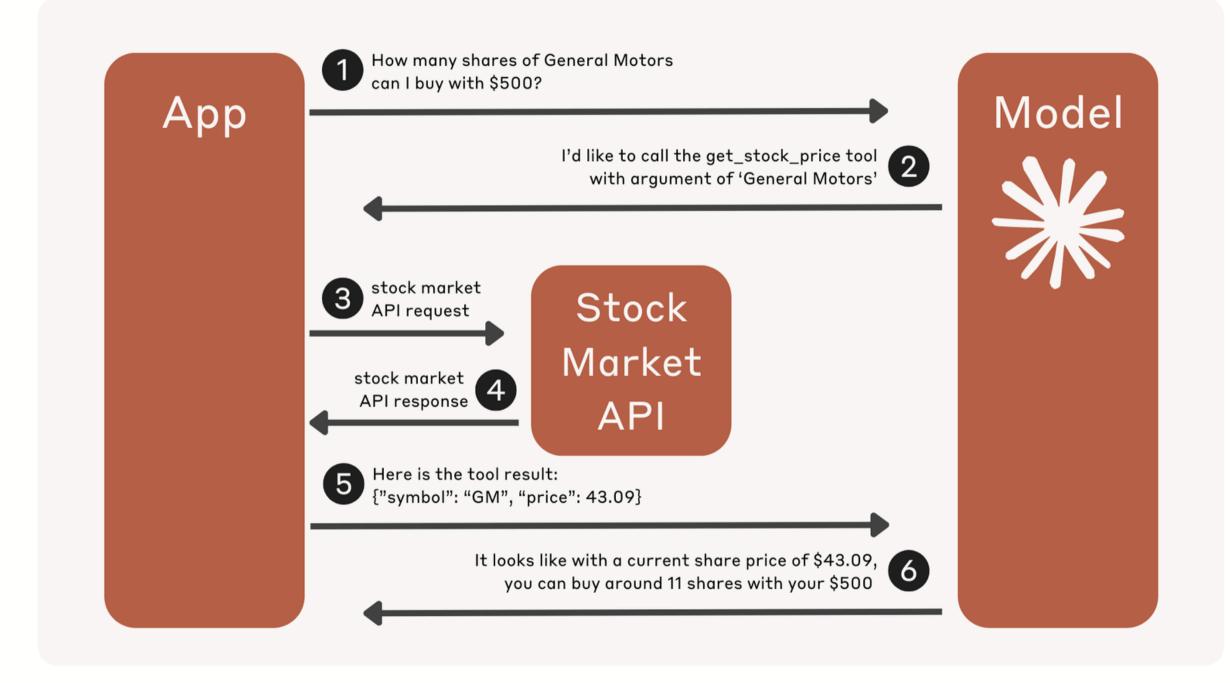
Reprompt (With tool result)

```
"messages": [
        "role": "user",
        "content": [
                "type": "tool_result",
                "tool_use_id": "toolu_01A09q90qw90lq917835lq9",
                "content": "15 degrees"
```



Complete Completion (Using tool response)

```
"id": "msg_01Aq9w938a90dw8q",
  "model": "claude-3-7-sonnet-20250219",
  "stop_reason": "stop_sequence",
  "role": "assistant",
  "content": [
     "type": "text",
      "text": "The current weather in Berlin is 15 degrees Celsius (59 degrees Fahrenheit). It's a
cool day in the city!"
```



https://docs.anthropic.com/en/docs/build-with-claude/tool-use/overview https://github.com/anthropics/courses/blob/master/tool_use/01_tool_use_overview.ipynb

How to build an agent

High-Code / Yes-Code

- OpenAl
- Anthropic
- Mistral
- Cohere
- Llama
- •



How to build an agent

Agent Frameworks

- CrewAl
- AutoGen
- LangChain
- AI SDK by Vercel
- Agents SDK by OpenAI (new)

• ...



How to build an agent

Low-Code / No-Code

- Flowise
- Replit
- n8n
- LangFlow
-



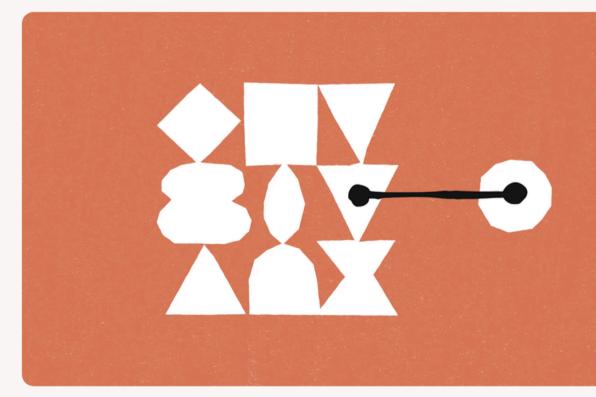
A\

Claude API Research Commitments Learn News

(Announcements)

Introducing the Model Context Protocol

Nov 25, 2024 • 3 min read



Try Claude



OpenAl 🗆	
	February 2, 2025 Release
Research	Introducing deep res
Safety	0 1
ChatGPT	An agent that uses reasoning to synthesize large amounts of online complete multi-step research tasks for you. Available to Pro users today
Sora	
API Platform	Try on ChatGPT 7
For Business	
Stories	
Company	
News	E ChatGPT 40 ~
	Compile a research report on how the retail industry has changed in the past 3 years. Use builtets and tables where necessary for clarity.
	 Could you specify which aspects of the retail industry you're most interested in? For example: E-commerce vs. brick-and-mortar trends Consumer behavior shifts Supply chain challenges Emerging technologies (AI, automation, etc.) Impact of inflation and economic changes Sustainability and ethical retailing Also, do you need a global perspective, or should the focus be on a specific region? If I I I I I I I I I I I I I I I I I I

Q Log in

search

information and Plus and Team next.



Ask ChatGPT

All, global per

A\

Claude API Research Commitments Learn News



Claude 3.7 Sonnet is now available on all <u>Claude</u> plans—including Free, Pro, Team, and Enterprise—as well as the Anthropic API, Amazon Bedrock, and Google Cloud's Vertex AI. Extended thinking mode is available on all surfaces except the free Claude tier.

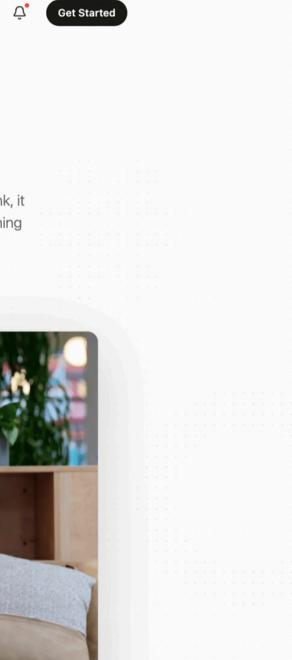
In both standard and extended thinking modes, Claude 3.7 Sonnet has the same price as its predecessors: \$3 per million input tokens and \$15 per million output tokens—which includes thinking tokens.



ട്ര്യ manus 🌐 English 🗸 Use cases Benchmarks Leave it to Manus Manus is a general AI agent that bridges minds and actions: it doesn't just think, it delivers results. Manus excels at various tasks in work and life, getting everything done while you rest. Introducing Manus

Intro





OpenAl	

Research

Safety

ChatGPT

API Platform

For Business

Stories

Company

News

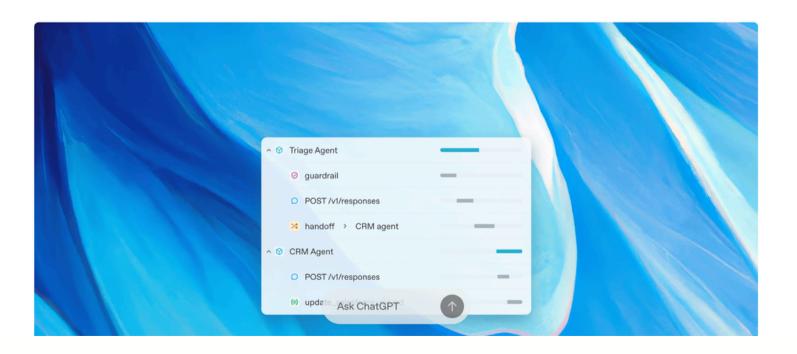
Sora

March 11, 2025 Product

New tools for building agents

We're evolving our platform to help developers and enterprises build useful and reliable agents.

Try in Playground 7



Q Log in

About Us Al Master Plan

Culture

News

Careers

HELIX: A VISION-LANGUAGE-ACTION MODEL FOR GENERALIST HUMANOID CONTROL

February 20, 2025

Introducing Helix

We're introducing Helix, a generalist Vision-Language-Action (VLA) model that unifies perception, language understanding, and learned control to overcome multiple longstanding challenges in robotics. Helix is a





Latest Startups Venture Apple Security AI Apps Nvidia GTC | Events Podcasts Newsletters



AI

A job ad for Y Combinator startup Firecrawl seeks to hire an Al agent for \$15K a year

Julie Bort - 9:00 AM PST · February 14, 2025

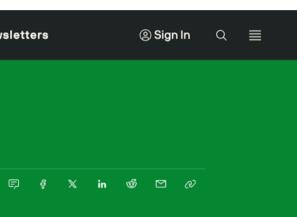
IMAGE CREDITS: FIRECRAWL

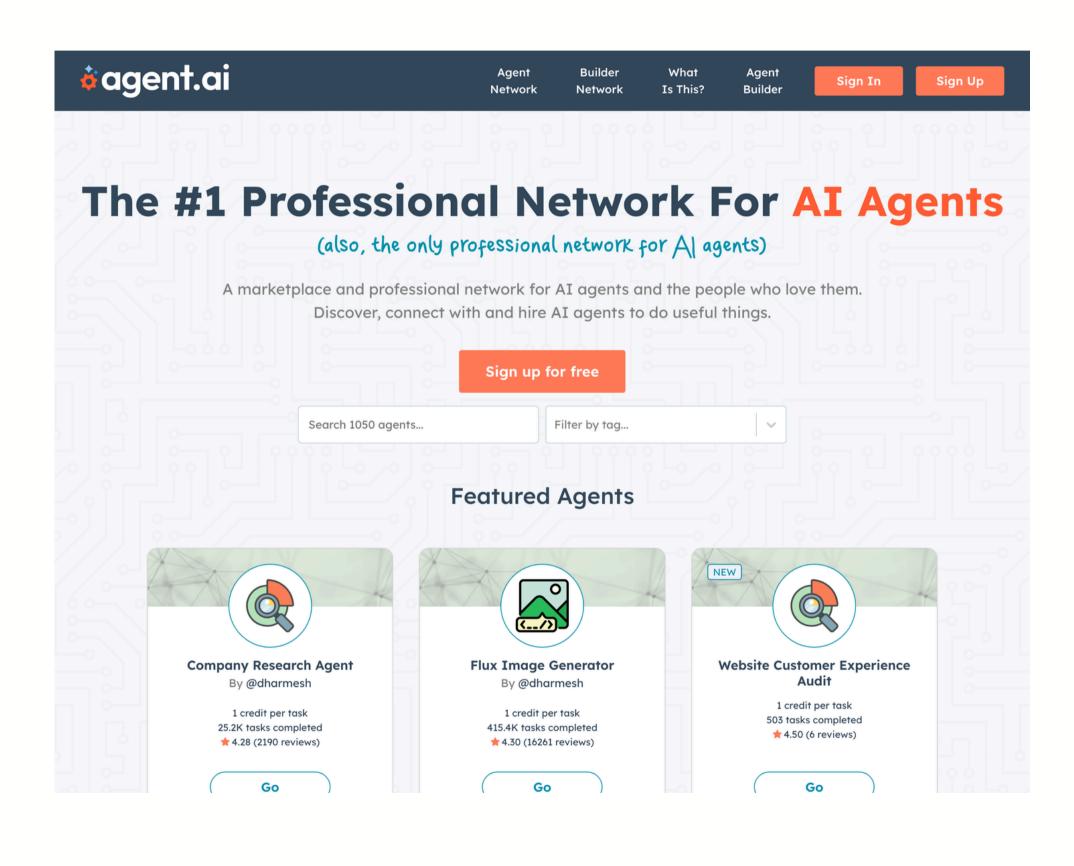


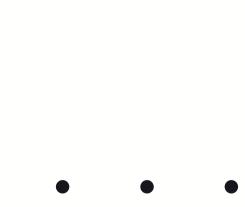
Last week, an ad from the Y Combinator job board for a tiny startup called Firecrawl went viral on X.

That's because the educer't far a human "Diagon apply only from are an Al



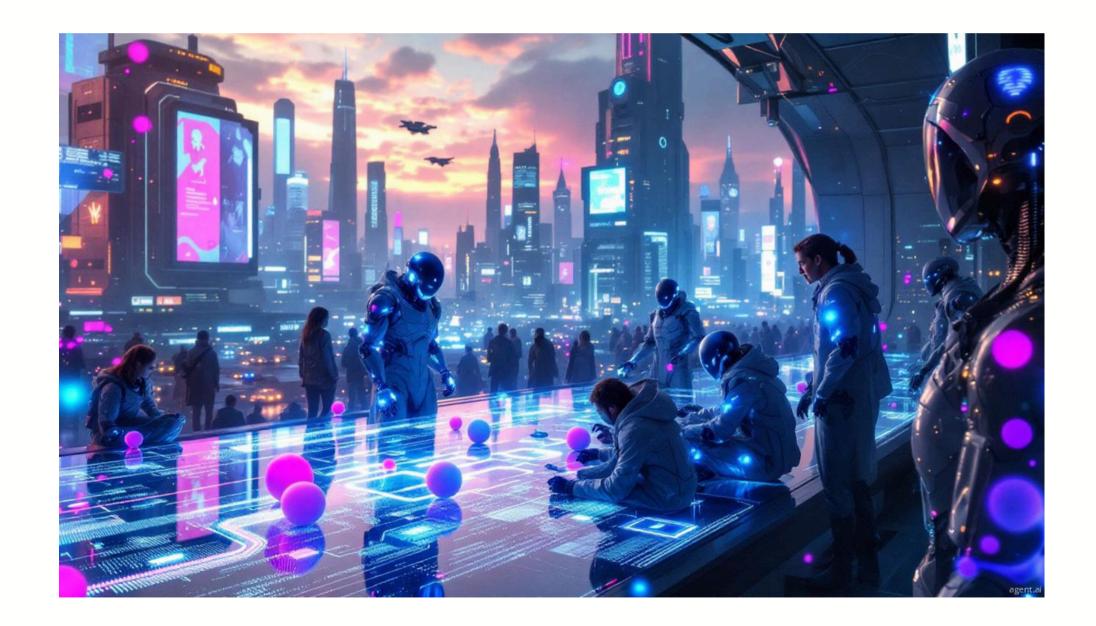






What's coming?

- Hybrid Workplace
- Agent-to-Agent communication!





Thank you!

https://moonside.ai/



