Week 2

W2.2- LLM Agent Basics

2026 Spring

LLM Agents Foundation & Applications

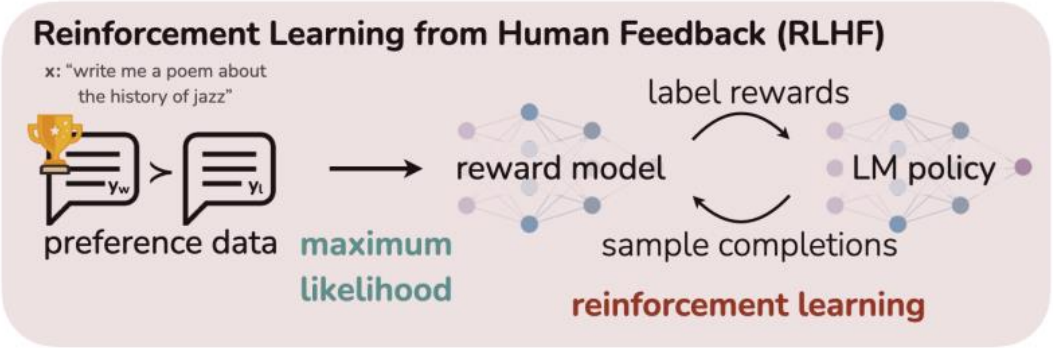
Dr. Yanjun Qi

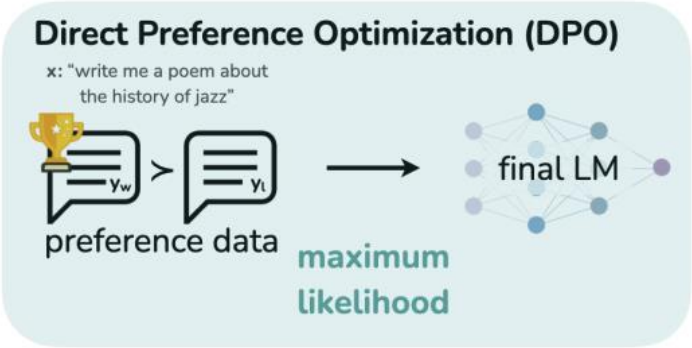
20260115

Last Class

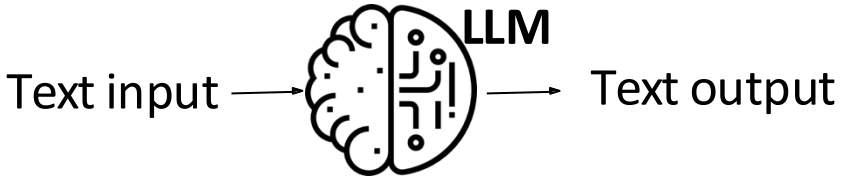
To incorporate human preferences:

- **RL algorithms** (PPO, GRPO, REINFORCE) explicitly maximize expected reward from a reward model – we normally call this group **RLHF**
- **Direct alignment methods** (DPO, IPO, KTO) optimize preference objectives without explicit reward modeling
- Though they can be shown to implicitly optimize an equivalent objective under certain assumptions





Accelerated development (0925)



Major Large Language Models (LLMs)

ranked by capabilities, sized by billion parameters used for training

CLICK LEGEND ITEMS TO FILTER

anthropic chinese google meta mistral openAI other xAI

Parameters (Bn) open access

search... show only: all

MMLU

89.8 = human expert

▲ 70+ IDEAL ▲

20 40 60 80

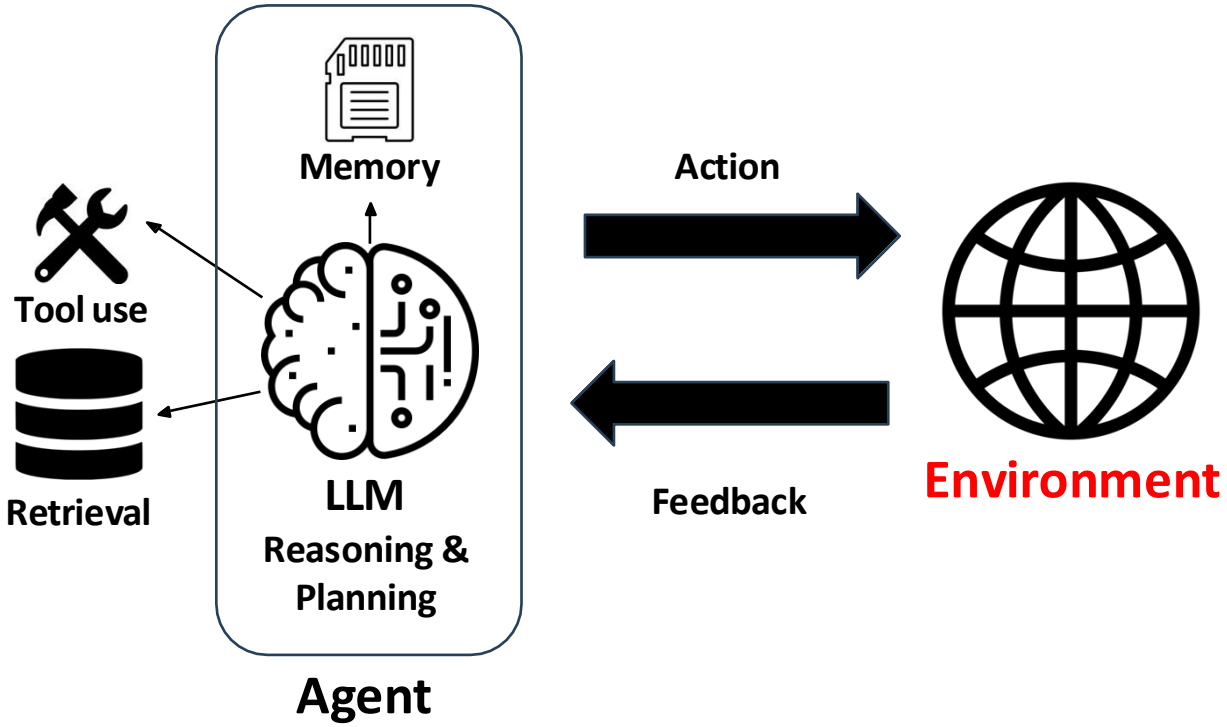
pre-2022 2022 2023 2024 2025

GPT-3 RoBERTa UL2 20B BLOOM Atlas Galactica Gopher Chinchilla U-PaLM ChatGPT (gpt-3.5-turbo) Titan LLaMA-65B GPT-4 Classic BloombergGPT GPT-NeoX Ernie 4.0 Ernie 3.5 Llama 2 Falcon 180B Claude 2 Yi-34B Claude 2.1 Gemini 1.5 Pro Claude 3 Opus Skywork-13B Granite Baichuan 2 Mistral 7B Mamba Nemotron-3 22B Griffin Command-R Hawk RWKV-v5 Eagle 7B Rene OpApple On-Device Jun 24 AMD-Llama-135m RakutenAI-7B Pile-T5 JetMoE-8B HLAT DBRX K2 MAP-Neo NeMo Nemotron-4-340B Claude 3.5 Sonnet GPT-4o Llama 3.1 405B Grok 2 Gemini 2.0 Gemini-1.5 Pixtral-12b Minitron-4B OLMoE-1B-7B Llama 3.2 3B Mistral Small 3.1 GPT 4.5 Llama 4 Maverick Gemini 2.0 Pro GPT o1* GPT o1 pro* DeepSeek-R1 DeepSeek-V3 o3 DeepSeek V3.1 GPT-5 Gemini 2.5 Pro Claude Opus 4 Qwen3-Next Mistral Medium 3 Claude 4.0 Sonnet Grok 4 Qwen2.5

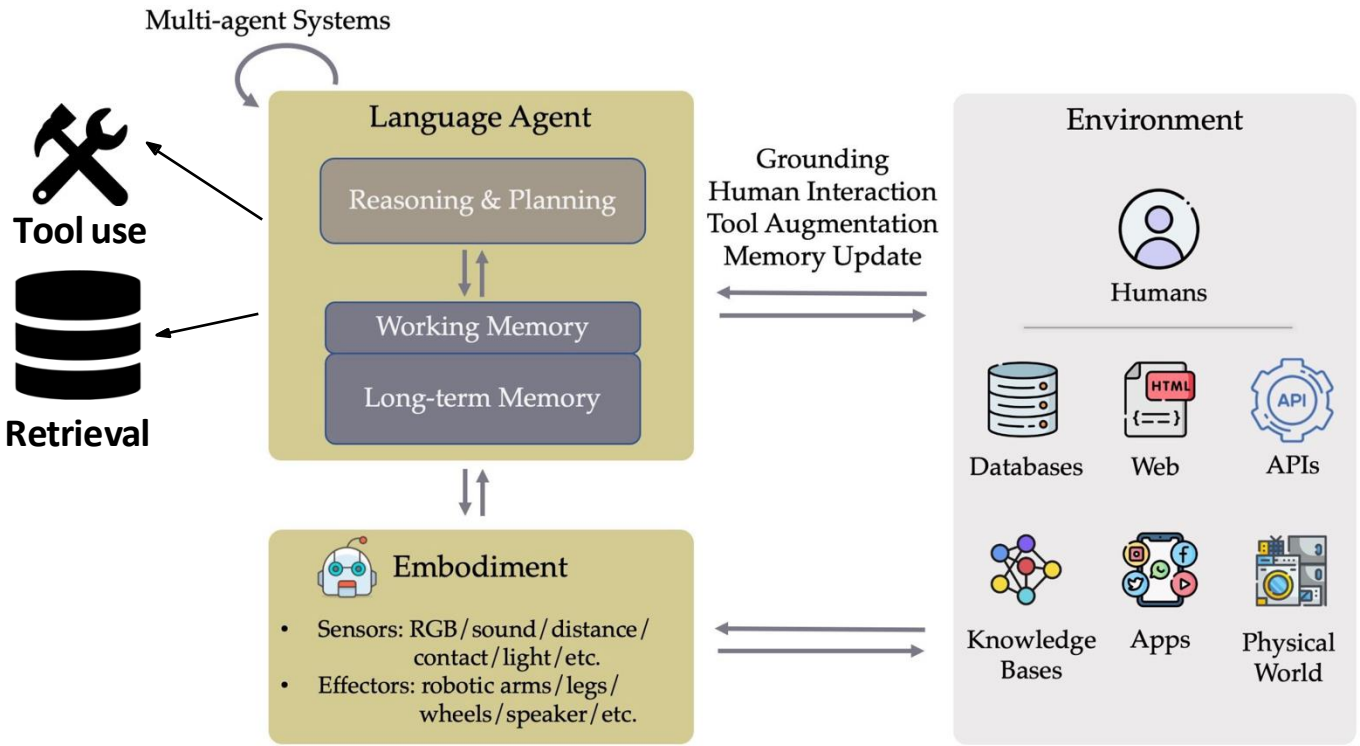
David McCandless, Tom Evans, Paul Barton
Informationisbeautiful // Sep 2025

MMLU = benchmark for measuring LLM capabilities
* = parameters undisclosed // source: LifeArchitect // data

LLM agents: enabling LLMs to interact with the environment

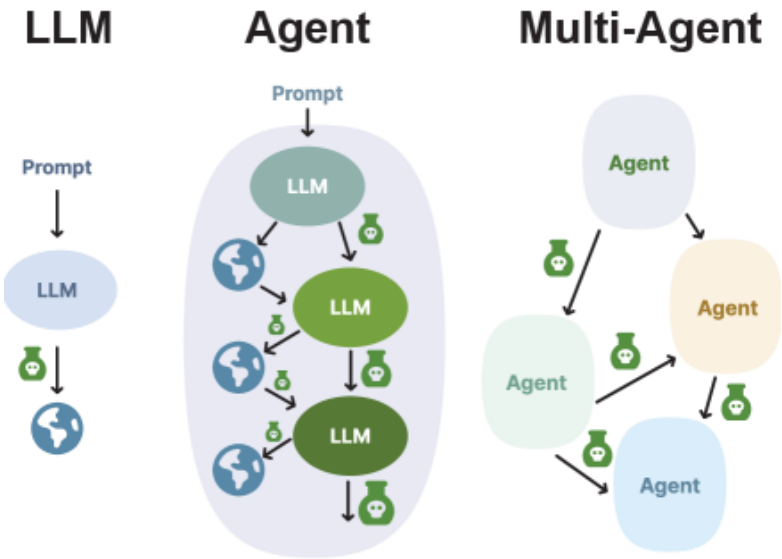


LLM Agents in Diverse Environments



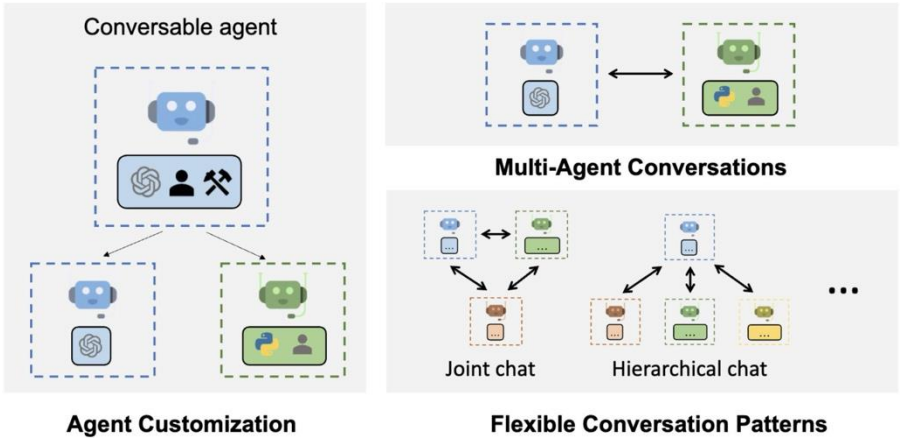
https://yusu.substack.com/p/language-agents

Multi-agent collaboration: division of labor for complex tasks



Emergence of social behaviors with role-play LLMs

Generative agents, Project Sid,...

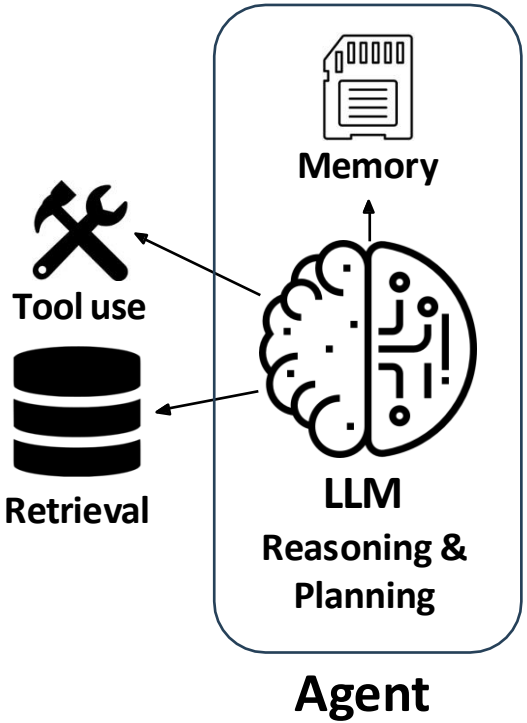


Specialized agents for different subtasks

Autogen, CrewAI, CAMEL, Mixture-of-Agents,...

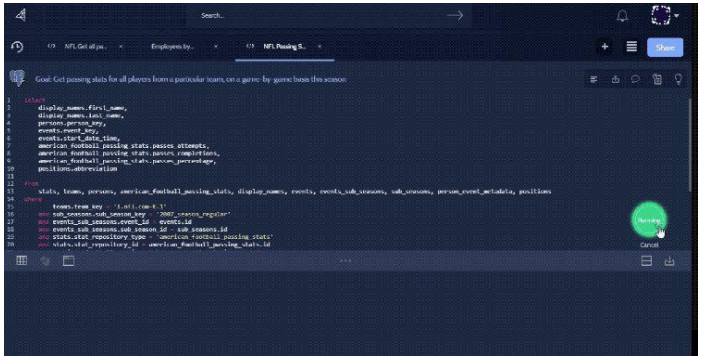
Project Sid: Many-agent simulations toward AI civilization, 2024

Why empowering LLMs with the agent framework

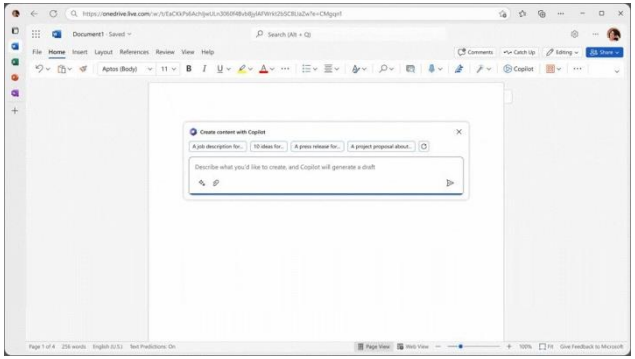


- Solving real-world tasks typically involves a trial-and-error process
- Leveraging external tools and retrieving from external knowledge expand LLM's information capabilities
- Agent workflow facilitates complex tasks
  - Allocation of subtasks to specialized tools
  - Multi-agent generation inspires better responses
  - Access to specialized evidence / data / inputs
  - ……

LLM agents transformed various applications



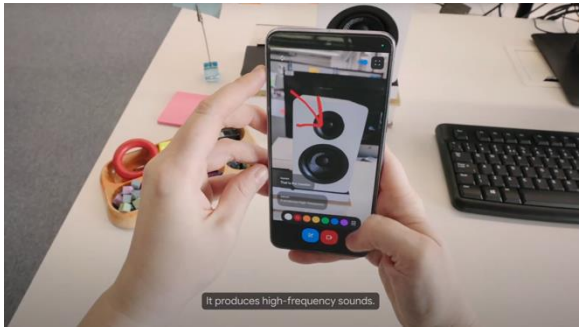
Code generation
Cursor, GitHub Copilot, Devin, Replit,…



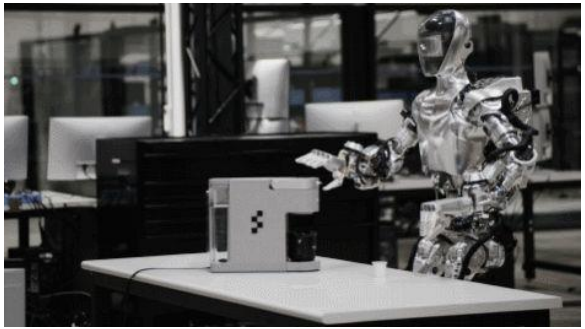
Workflow automation
Microsoft Copilot, Multi-On,…

- Healthcare
- Education
- Law
- Finance
- Cybersecurity

…



Personal assistant
Google Astra, OpenAI GPT-4o,…



Robotics
Figure AI, Tesla Optimus,…

LLM agents are improving rapidly (leaderboards!^

Full is a large benchmark made of 2000 instances (details)

Filters: Open Scaffold ▾ All Tags ▾

Model	% Resolved	Org
✅ SWE-agent 1.0 (Claude 3.7 Sonnet)	33.83	
✅ OpenHands + CodeAct v2.1 (claude-3-5-sonnet-20241022)	29.38	
AutoCodeRover-v2.0 (Claude-3.5-Sonnet-20241022)	24.89	
✅ SWE-agent + Claude 3.5 Sonnet	18.13	
✅ SWE-agent + GPT 4 (1106)	12.47	
✅ SWE-agent + GPT 4o (2024-05-13)	11.99	
✅ SWE-agent + Claude 3 Opus	10.51	
✅ RAG + Claude 3 Opus	3.79	

SWE-bench

Leaderboards

BENCHMARKS

SWE-bench

SWE-bench Verified

SWE-bench Bash Only

SWE-bench Multilingual

SWE-bench Multimodal

SWE-bench Lite

ABOUT

GAIA (Mialon et al.)

huggingface.co/gaia-benchmark

Results: Test

Agent name	Model family
JoinAI_V2.2	GPT 5, Gemini 3 Pro, DeepSeek 3.1, Qwen 3
Nemotron-ToolOr	Nemotron-ToolOrchestrator-8B, GPT-5, Claude Opus 4.1
Nemotron-ToolOr	Nemotron-ToolOrchestrator-8B, GPT-5, Claude Opus 4.1
SU Zero - Shuqi	Self Consistency 35
JoinAI_V2.1	GPT, Gemini, DeepSeek, Qwen
ShawnAgent_v3.1	GPT5.2, Claude Sonnet 4.5, Gemini 3 Pro
HALO V1217-1	

SWE-Bench (Jimenez*, Yang*, et al.) / swebench.com / Today's screenshot on Leaderboard!

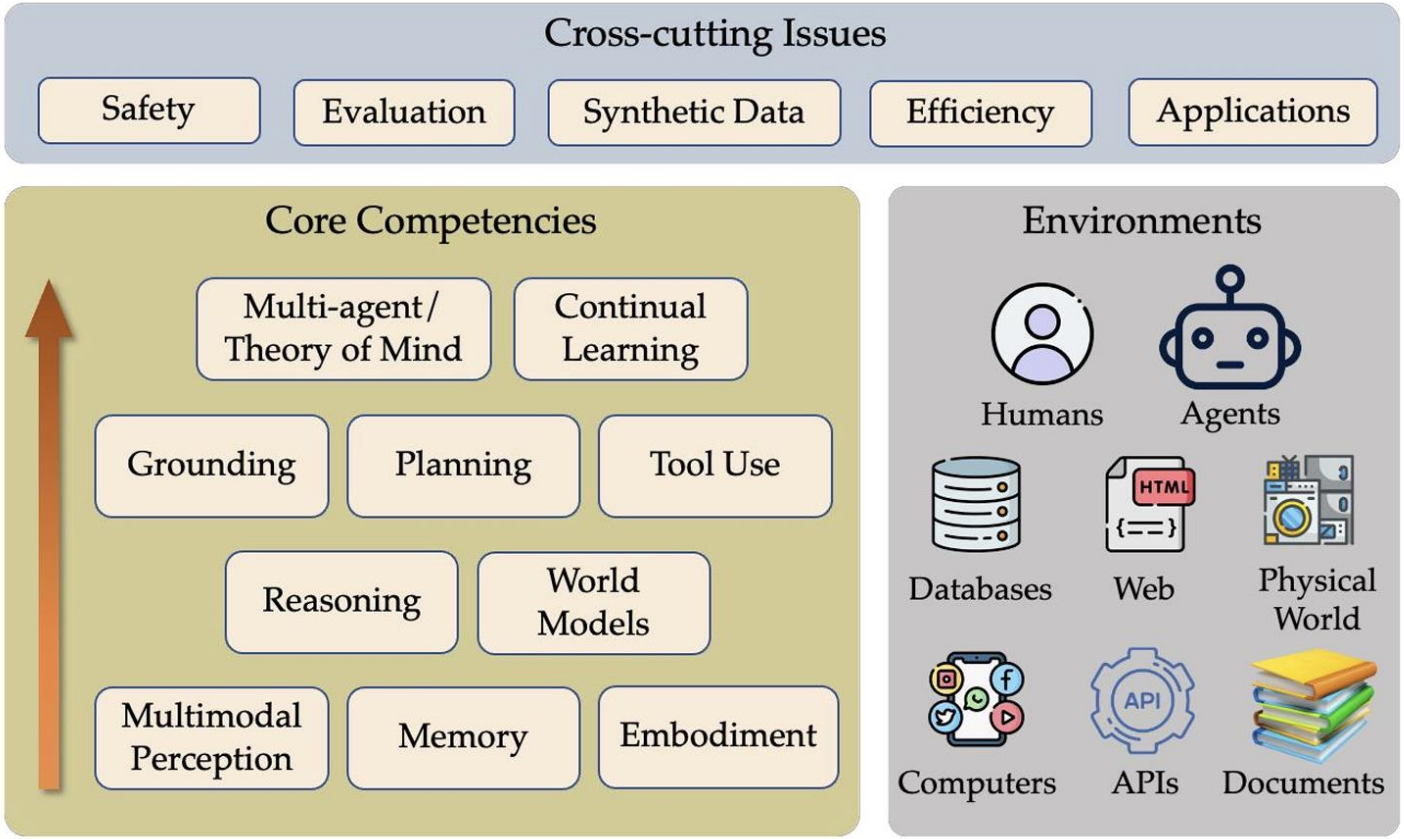
F2 | OAgent

	A	B	C	D	E	F	G	H
1	a	Open?	Model Size (billion)	Model	Success Rate (%)	Result Source	Work	Traj
2	01/2026	✗		OAgent	71.6	OAgent	OAgent	Link
3	12/2025	✓	GPT-5	ColorBrowserAgent	71.2	ColorBrowserAgent	ColorBrowserAgent	Link
4	10/2025	✓	-	Claude Code + GBOX MCP	68	GBOX AI	GBOX AI	Link
5	09/2025	✗	-	DeepSky Agent	66.9	Self-reported	DeepSky Agent	Link
6	10/2025	✗		Narada AI	64.2	Self-reported	Narada AI	Link
7	02/2025	✓	-	IBM CUGA	61.7	IBM CUGA	IBM CUGA	html+ json
8	01/2025	✗	-	OpenAI Operator	58.1	OpenAI CUA	OpenAI CUA	Link

WebArena (Zhou et al.)

webarena.dev

Topics We will cover in this course:



Topics covered in this course. ➔ Potential Projects:

- Applications
  - Software development
  - Workflow automation
  - Multimodal applications
  - Industrial applications like Healthcare, Legal, Fin, …
- Model core capabilities
  - Reasoning
  - Planning
  - Multimodal understanding
- LLM agent frameworks
  - Workflow
  - Tool use
  - Retrieval-augmented generation
  - Multi-agent systems
- Safety and ethics

- Applications
  - Build LLM agents applications in specialized / novel domains
- Core Fundamentals
  - Enhance core agent capabilities (memory, planning, tool use, alignments, efficiency, …)
  - Enhance decentralized multi-agent systems
- Benchmarks / Build Novel Frameworks
  - Create and improve benchmarks for Evaluating LLM agents
  - Reimplement or Build novel frameworks for agent workflow
- Safety and ethics
  - Reveal safety concerns in deployment (misuse, privacy, etc.)
  - Defense safety converns

Challenges for LLM agent deployment in the wild

- Reasoning and planning
  - LLM agents tend to make mistakes when performing complex tasks end-to-end
- Embodiment and learning from environment feedback
  - LLM agents are not yet efficient at recovering from mistakes for long-horizon tasks
  - Continuous learning, self-improvement
  - Multimodal understanding, grounding and world models
- Multi-agent learning, theory of mind
- Safety and privacy
  - LLMs are susceptible to adversarial attacks, can emit harmful messages and leak private data
- Human-agent interaction, ethics
  - How to effectively control the LLM agent behavior, and design the interaction mode between humans and LLM agents