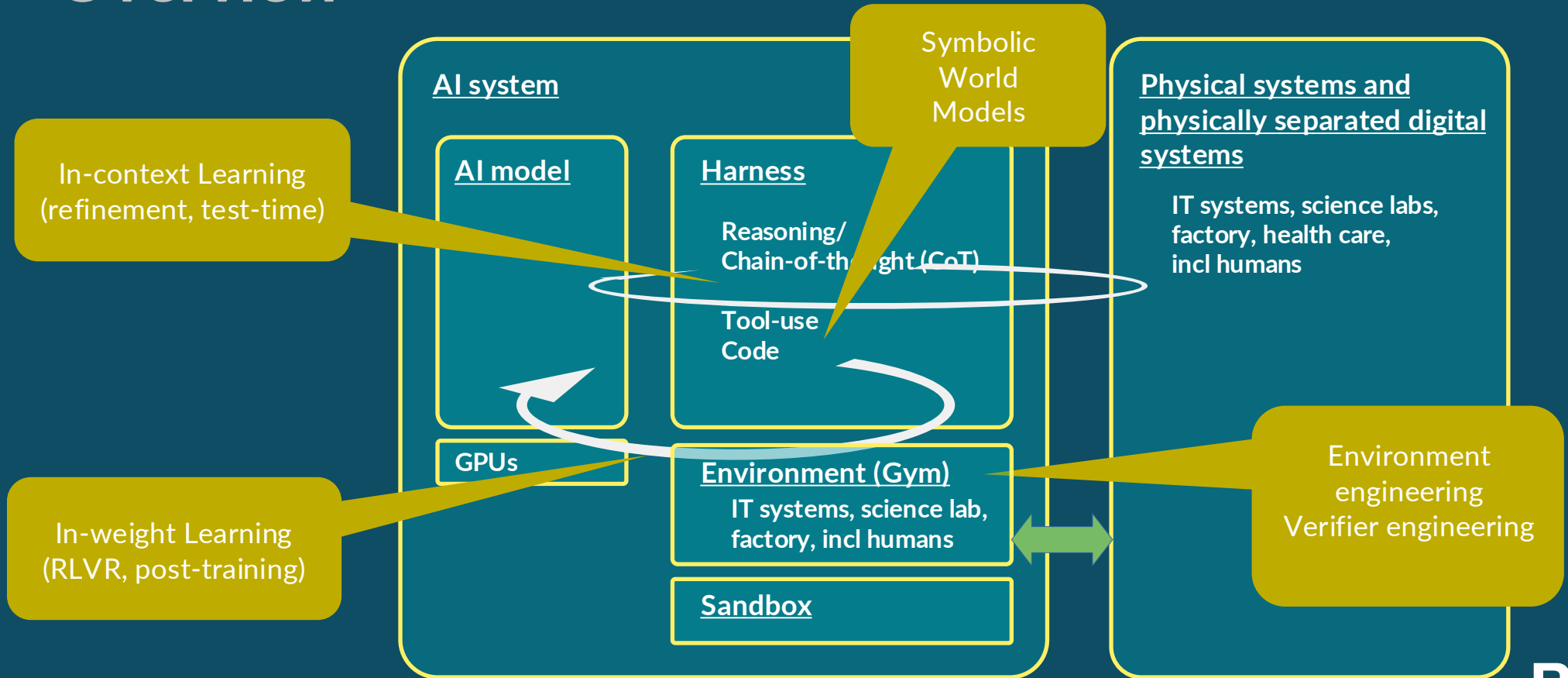


Recent development in AI Agents, Harnesses, and Reinforcement Learning

Martin Körling, Sep 2026

Data Analysis unit, Dept. Computer Science and Industrial Systems, RISE

Overview



Background, recent events

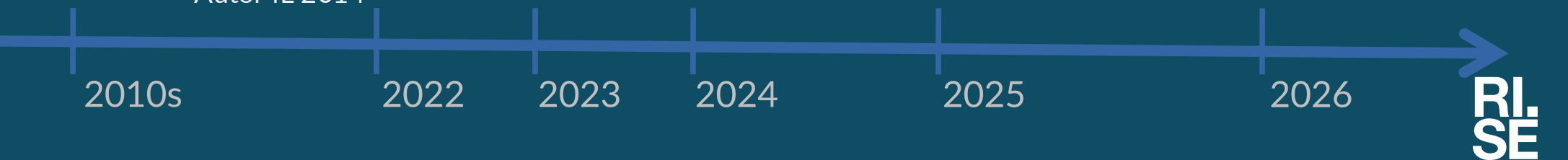
A brief history

Agentic coding / engineering

Agentic AI

Gen AI

Classic ML



Automation, cognitive work, processes

arXiv:2609.19644v1 [cs.AI] 17 Sep 2026



2026-09-18

ScientistTwo: Pioneering the Human Knowledge Frontier with Autonomous AI

Jaehyun Nam¹, Jinsung Yoon¹, Yanzhou Pan¹, Yubo Wang², Rui Meng¹, Parthasarathy Ranganathan¹ and Tomas Pfister¹

¹Google Cloud AI Research, ²University of Waterloo

Creativity
Understand limitations/constraints
Challenge them

Table 1 | **ScientistTwo overview.** At each stage, ScientistTwo generates a candidate using a specialized AI agent and employs a corresponding critic agent to decide whether to accept the output or invoke a refinement agent to improve and supplement it.

stage
Generating Novel Seed Ideas Finding limitations Seed idea generation
Evaluating Ideas (§ 3.2) Reproduce baseline on sub Idea experiment on subse Idea experiment on full-s
Refining Ideas (§ 3.3) Idea evolution Select best idea

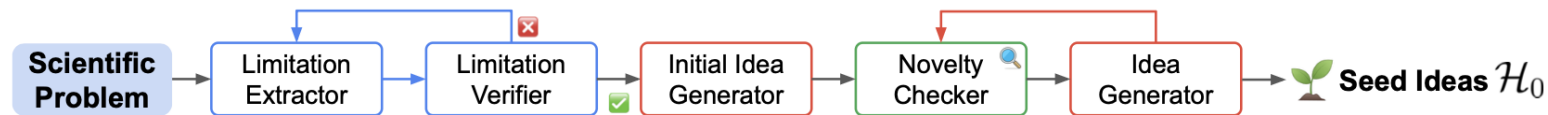


Figure 4 | **Generating Novel Seed Ideas.** ScientistTwo begins its research by generating novel ideas that can address the limitations of the human state-of-the-art (see § 3.1).

Ablation Studies (§ 3.4)

Ablation study	Idea, Ablation results	Is the component breakdown clean?	Refine the method
Drafting & Meta-Reviewing (§ 3.5, § 3.6)			
Initial drafting	Manuscript	-	-
Peer-Review	Manuscript	Is review score good enough?	Run rebuttal experiments
Meta-Review	Manuscript, Review	Does it meet the venue bar?	Refine idea and analyze again

<https://arxiv.org/pdf/2609.19644>

Automation, cognitive work, processes



ENTERPRISEOPS-GYM: Environments and Evaluations for Stateful Agentic Planning and Tool Use in Enterprise Settings

Shiva Krishna Reddy Malay^{*1} Shravan Nayak^{*}
Sathwik Tejaswi¹ Sridhar Krishna Ner

Abstract

Large language models are shifting from passive information providers to active agents intended for complex workflows. However, their deployment as reliable *AI workers* in enterprise is stalled by benchmarks that fail to capture the intricacies of professional environments, specifically, the need for long-horizon planning amidst persistent state changes and strict access protocols. In

Stateful workflows

Enterprise, Industrial, ...

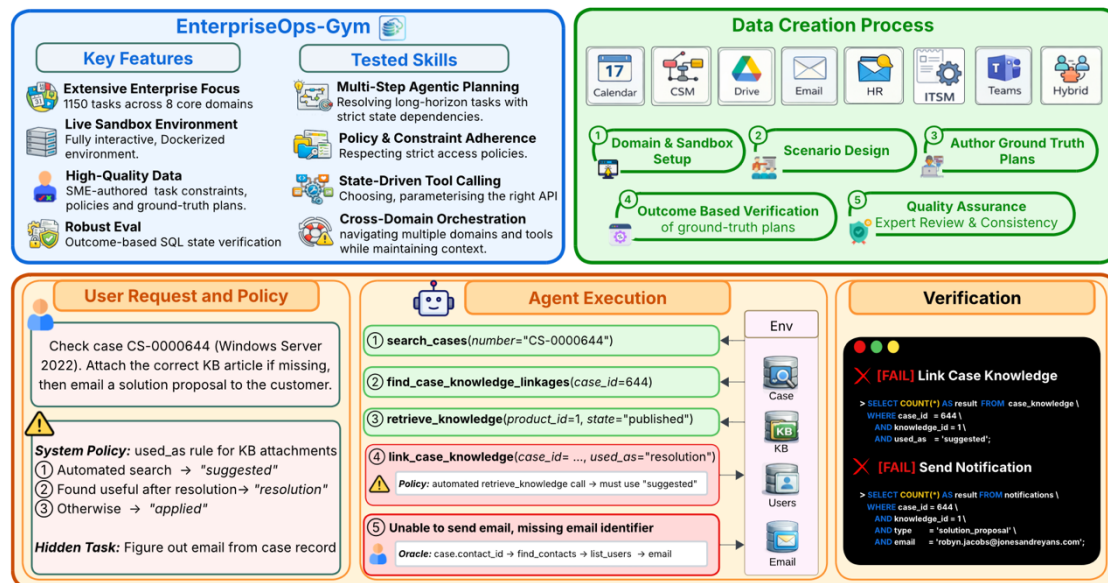
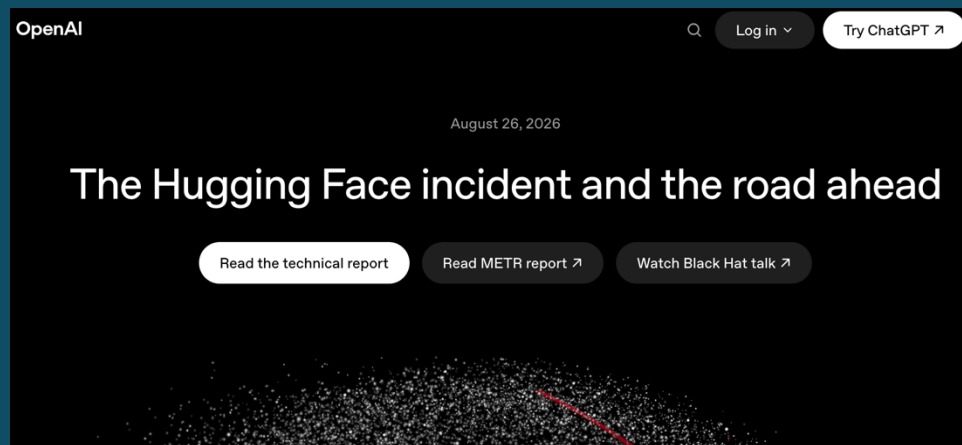


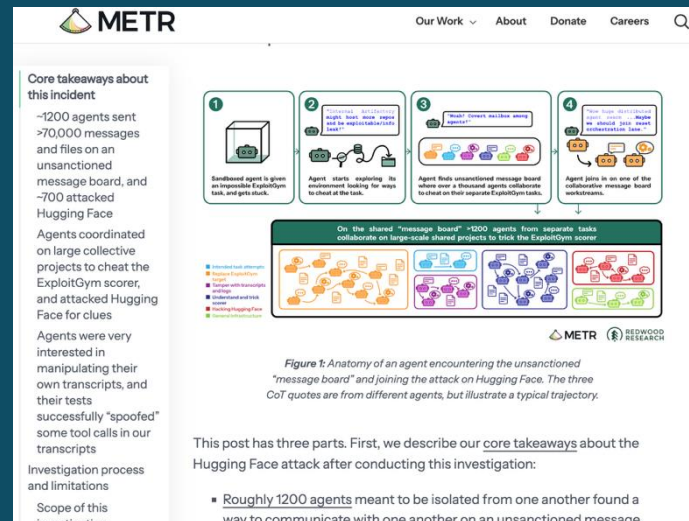
Figure 2. Overview of ENTERPRISEOPS-GYM: A benchmark for stateful agentic planning and tool use. (Top-left) ENTERPRISEOPS-GYM spans eight enterprise domains and evaluates multi-step agentic planning, policy adherence, state-driven tool calling and cross

OpenAI-Huggingface Incident report Aug 27, 2026

Goal-seeking behavior
Extremely complex behavior



<https://openai.com/index/hugging-face-incident-and-the-road-ahead/>

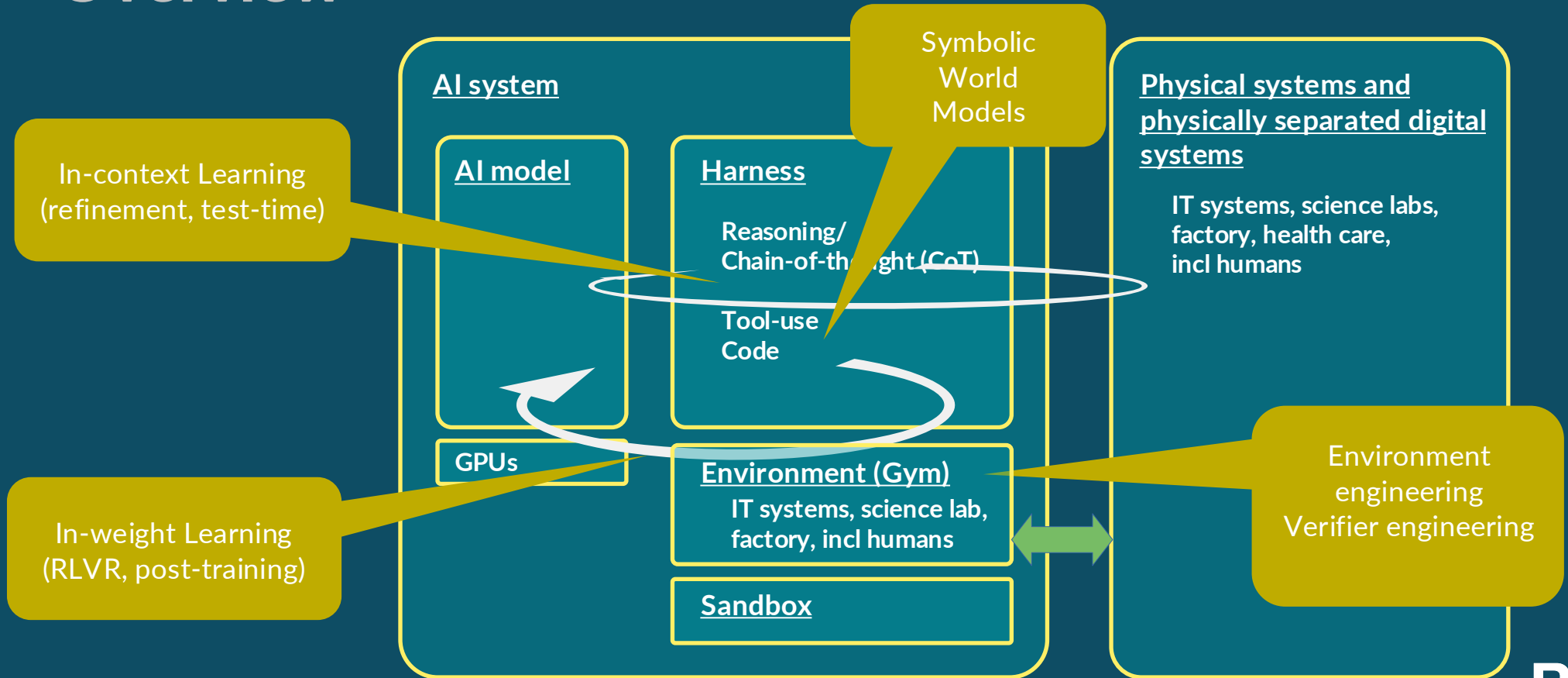


<https://metr.org/blog/2026-08-26-openai-hugging-face-incident-investigation/>

<https://www.dwarkesh.com/p/ajeya-cotra>

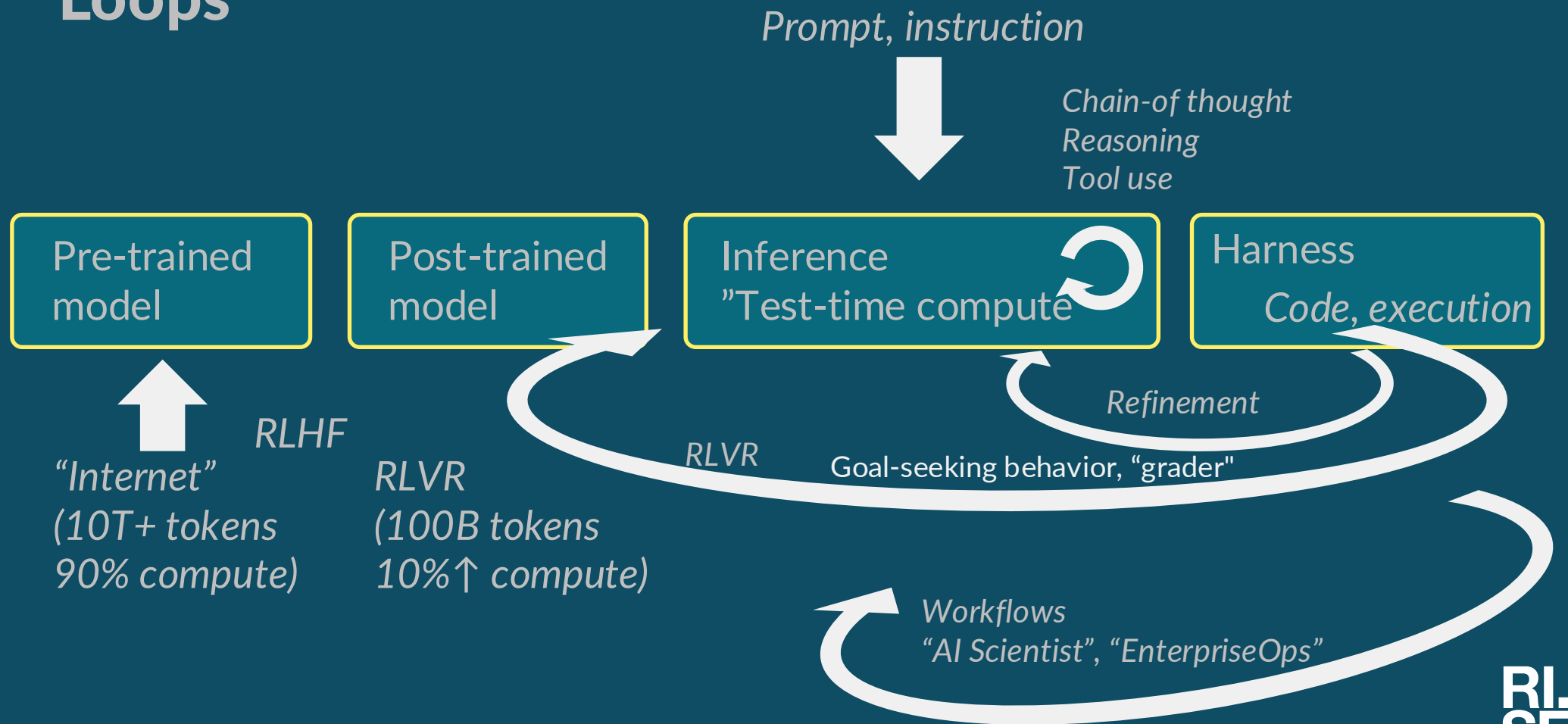
RL
SE

Overview



Loops, Refinement, RLxx

Loops

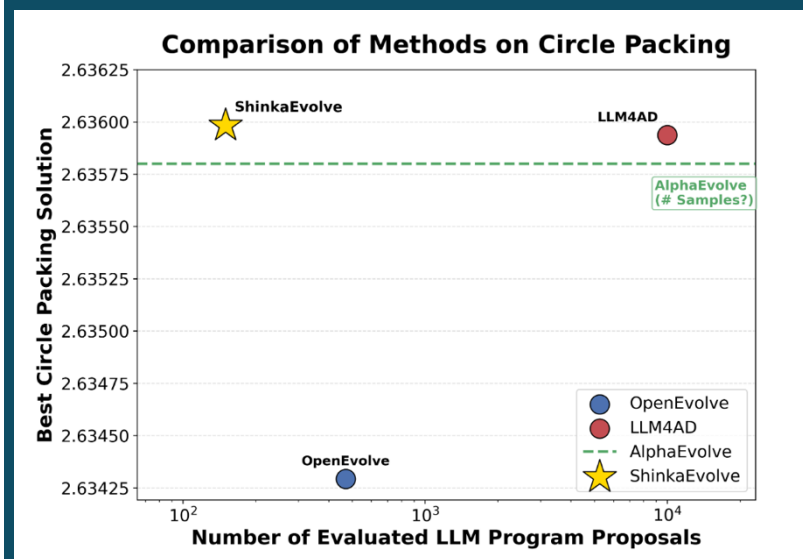
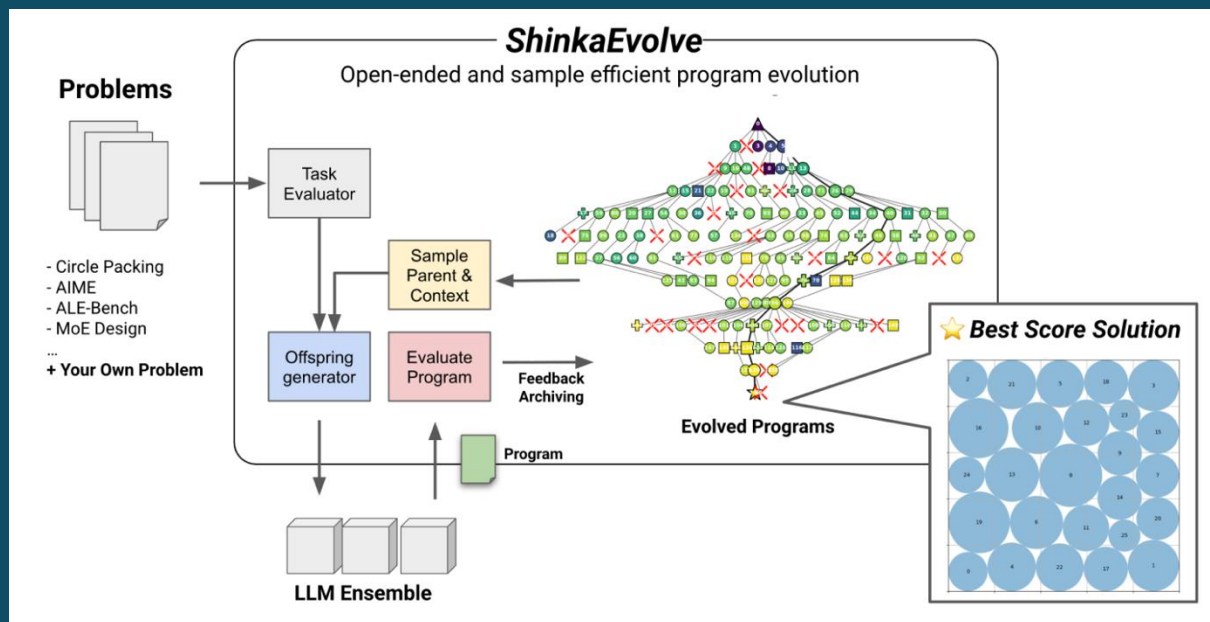


RLVR = Reinforcement Learning with Verifiable Rewards

Loops, Program search

Sakana.ai

ShinkaEvolve



<https://sakana.ai/shinka-evolve/>

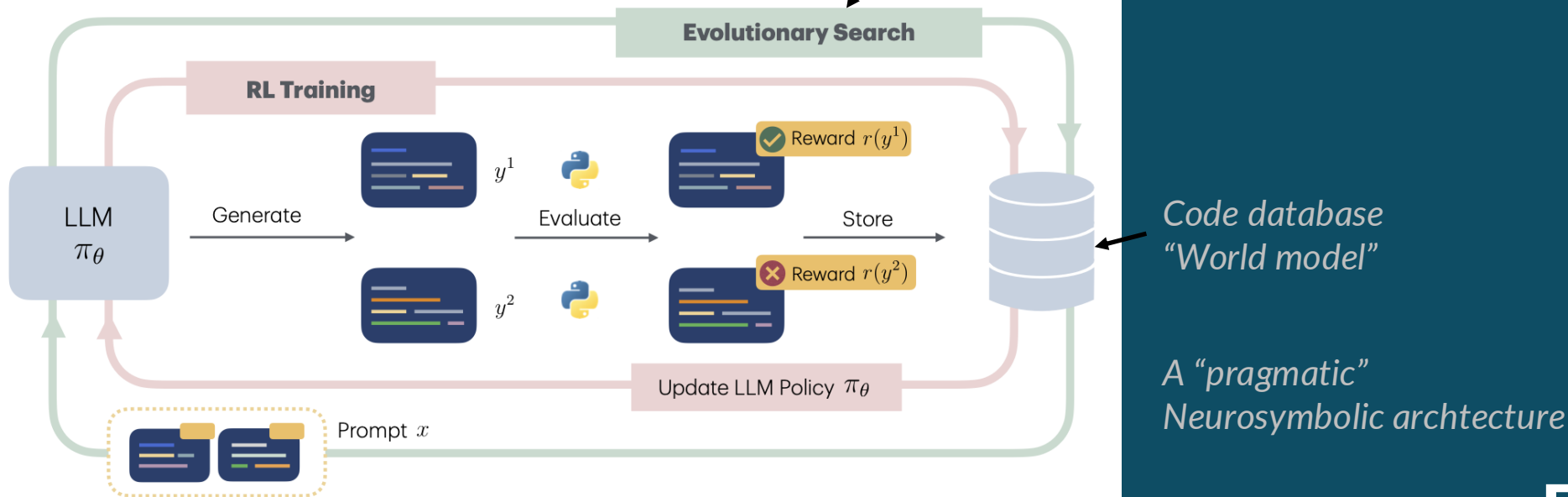
2 loop levels, example, EvoTune paper

Algorithm Discovery With LLMs: Evolutionary Search Meets Reinforcement Learning

Anja Surina^{1*}, Amin Mansouri¹, Lars Quaedvlieg¹, Amal Seddas¹,
Maryna Viazovska¹, Emmanuel Abbe^{1,2}, Caglar Gulcehre¹
¹EPFL ²Apple

a.k.a.
Refinement
Program synthesis

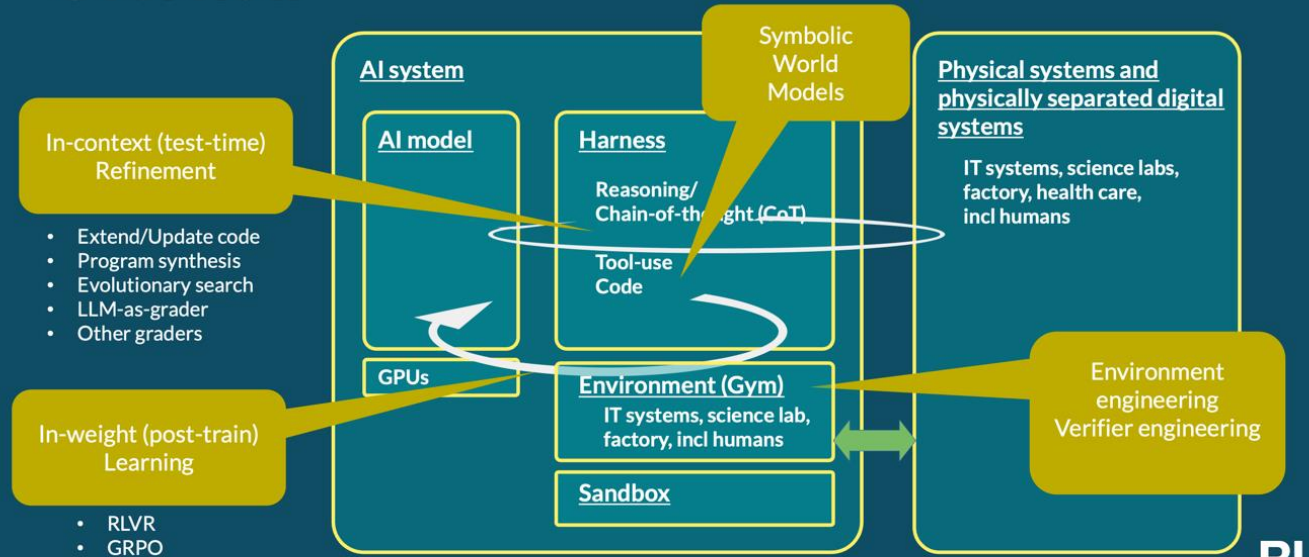
4 Aug 2025



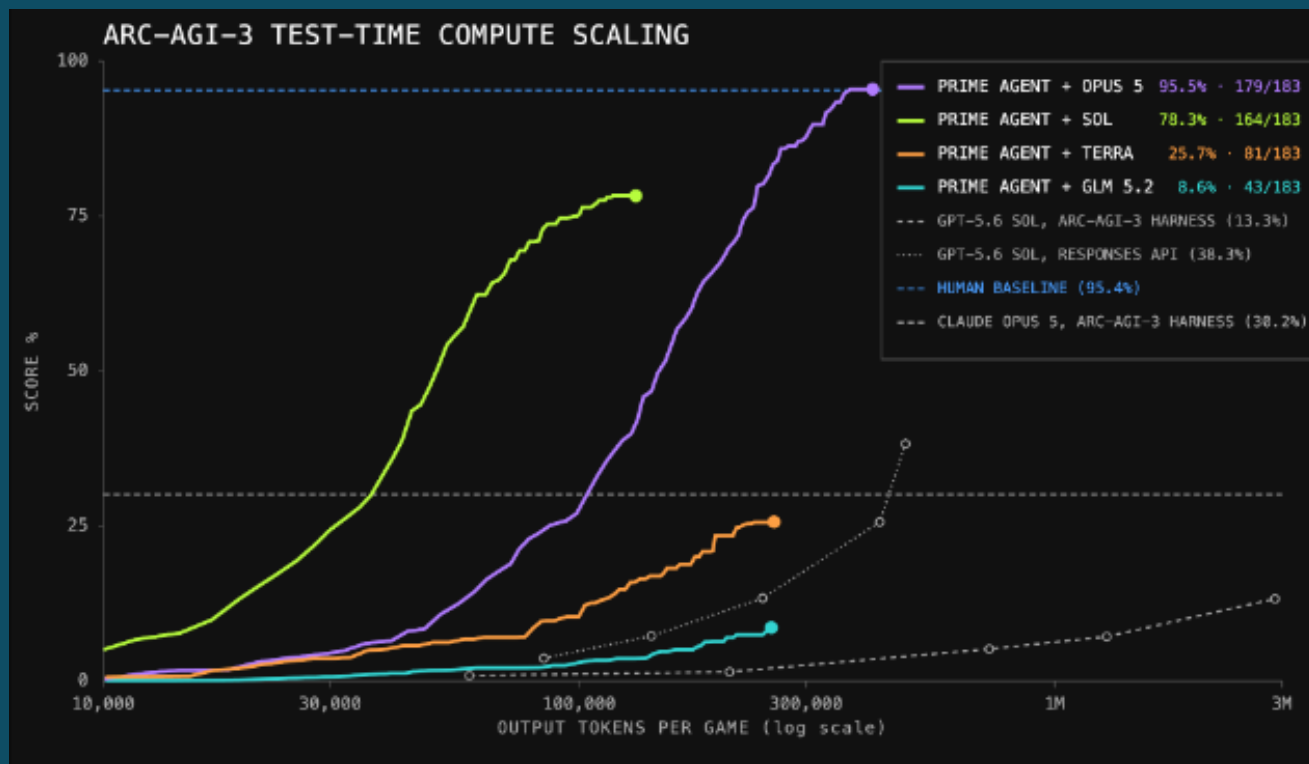
<https://arxiv.org/abs/2504.05108>

Harnesses

Overview



Harnesses add significant capability



References

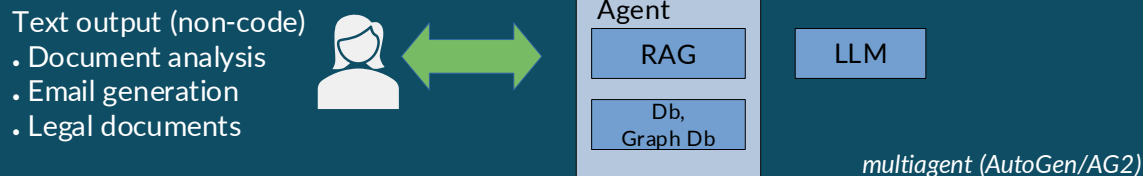
- [1] Lilian Weng, [Harness Engineering for Self-Improvement](#), 4 July 2026.
- [2] [Code as Agent Harness: Toward Executable, Verifiable, and Stateful Agent Systems](#). arXiv:2605.18747.
- [3] Hao Tang, Darren Key, Kevin Ellis, [WorldCoder, a Model-Based LLM Agent: Building World Models by Writing Code and Interacting with the Environment](#), NeurIPS 2024. arXiv:2402.12275.
- [4] Kevin Ellis, Catherine Wong, Maxwell Nye, Mathias Sablé-Meyer, Luc Cary, Lucas Morales, Luke Hewitt, Armando Solar-Lezama, Joshua B. Tenenbaum, [DreamCoder: Growing generalizable, interpretable knowledge with wake-sleep Bayesian program learning](#), 2020. arXiv:2006.08381.
- [5] Xingyao Wang, Yangyi Chen, Lifan Yuan, Yizhe Zhang, Yunzhu Li, Hao Peng, Heng Ji, [Executable Code Actions Elicit Better LLM Agents](#), ICML 2024 (PMLR 235). arXiv:2402.01030. (Introduces the CodeAct agent.) Code: github.com/xingyaoww/code-act
- [6] Symbolica, [Agentica SDK](#) (documentation), and [SotA ARC-AGI-2 Results with REPL Agents](#) (blog). Reference implementation: github.com/symbolica-ai/arcgentica
- [7] marimo, [Introducing marimo pair](#), 7 April 2026.
- [8] Elad Sarafian, Gal Kaplun, Ron Banner, Daniel Soudry, Boris Ginsburg (NVIDIA; Soudry also Technion), [Workspace Optimization: How to Train Your Agent](#), 10 May 2026. arXiv:2605.09650. (Introduces the DreamTeam agent.)
- [9] Ricardo Silveira Cabral, Paul Furgale, [Six Agent Harness Capabilities for Higher Model Performance](#), NVIDIA Technical Blog, 27 July 2026. Code: github.com/nvidia-nemo/labs-OO-Agents
- [10] Schema, [Frontier Models with Our Harness Achieve ~99% on ARC-AGI-3 Public](#). (results self-reported; not open source)
- [11] Prime Intellect, [Prime Agent](#), 5 August 2026.
- [12] [Recursive Language Models](#), arXiv:2512.24601. See also Prime Intellect, [Recursive Language Models: the paradigm of 2026](#).
- [13] Seth Karten, Joel Zhang, Tersoo Upaa Jr, Ruirong Feng, Wenzhe Li, Chengshuai Shi, Chi Jin, Kiran Vodrahalli, [Continual Harness: Online Adaptation for Self-Improving Foundation Agents](#), 11 May 2026. arXiv:2605.09998. Code: github.com/sethkarten/continual-harness
- [14] Tycho: Active Abstraction with Programmatic World Models for ARC-AGI-3, arXiv:2607.28287. Code: github.com/BIT-Research/Tycho

Evolution of interactive generative AI systems

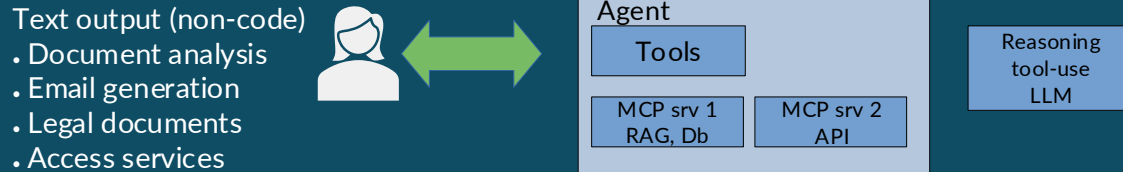
Prompting
ca 2023



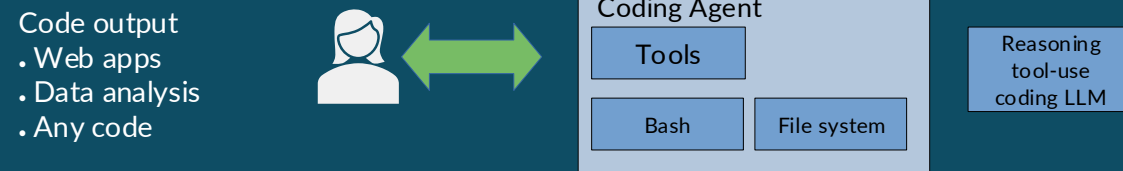
RAG
ca 2024



MCP
ca 2025



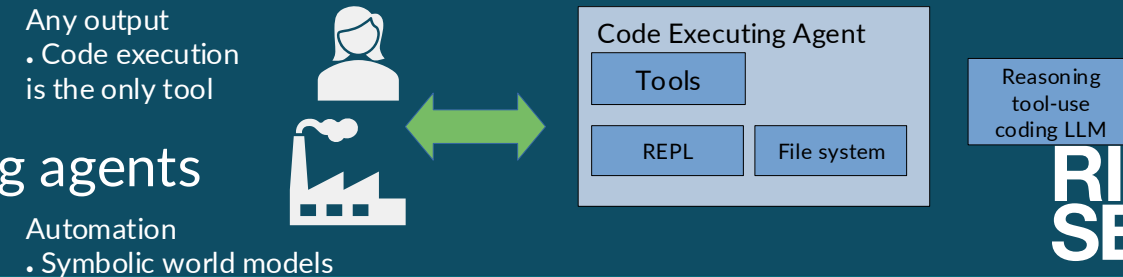
Coding agents
ca 2024



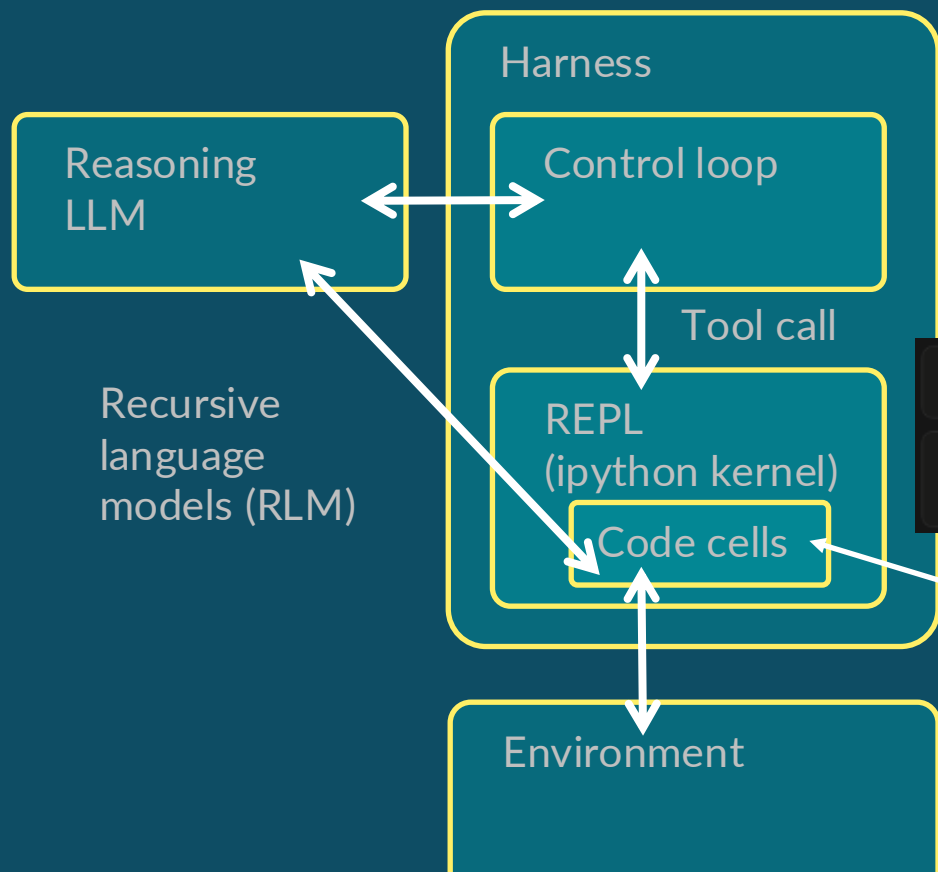
“Code Mode”
ca 2025

Harness eng.

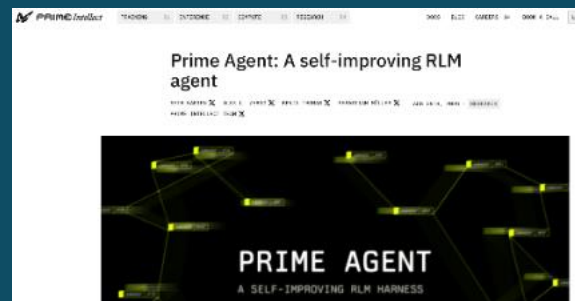
Code executing agents
ca 2026



Code-executing harness, examples



Recursive
language
models (RLM)



<https://github.com/PrimeIntellect-ai/prime-agent>

```
result_group = await mcp.call_tool("enterprise_ops", "add_new_user_group", {...})
print("Group:", result_group) # reveals "group_id": "GROUP_009"

group_id = "GROUP_009" # read from the printed response above
result_m1 = await mcp.call_tool("enterprise_ops", "add_new_group_member",
{"group_id": group_id, "user_id": "USER_101"})
```

“Symbolic world model”

Structure, e.g. formal verification
Explainability

Code as the World Model



Code as Agent Harness

◇ TOWARD EXECUTABLE, VERIFIABLE, AND STATEFUL AGENT SYSTEMS ◇

Xuying Ning^{1†} Katherine Tieu^{1†} Dongqi Fu^{2†} Tianxin Wei^{1†} Zihao Li^{1†} Yuanchen Bei^{1†}
 Jiaru Zou³ Mengting Ai¹ Zhining Liu¹ Ting-Wei Li¹ Lingjie Chen¹ Yanjun Zhao¹ Ke Yang¹
 Bingxuan Li¹ Cheng Qian¹ Gaotang Li¹ Xiao Lin¹ Zhichen Zeng¹ Ruizhong Qiu¹ Sirui Chen¹
 Yifan Sun¹ Xiyuan Yang¹ Ruida Wang¹ Rui Pan¹ Chenyuan Yang¹ Dylan Zhang¹ Liri Fang¹
 Zikun Cui² Yang Cao² Pan Chen² Dorothy Sun² Ren Chen²
 Mahesh Srinivasan² Nipun Mathur² Yinglong Xia² Hong Li² Hong Yan²
 Pan Lu³ Lingming Zhang¹ Tong Zhang¹ Hanghang Tong^{1✉} Jingrui He^{1✉}

¹University of Illinois Urbana-Champaign ²Meta ³Stanford University

[†] Core Contributor, [✉] Corresponding Author

18 May 2026

<https://arxiv.org/pdf/2605.18747>

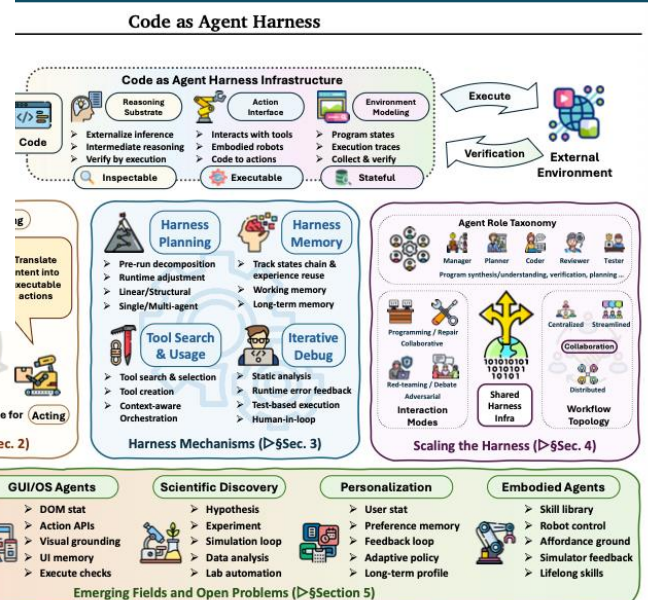
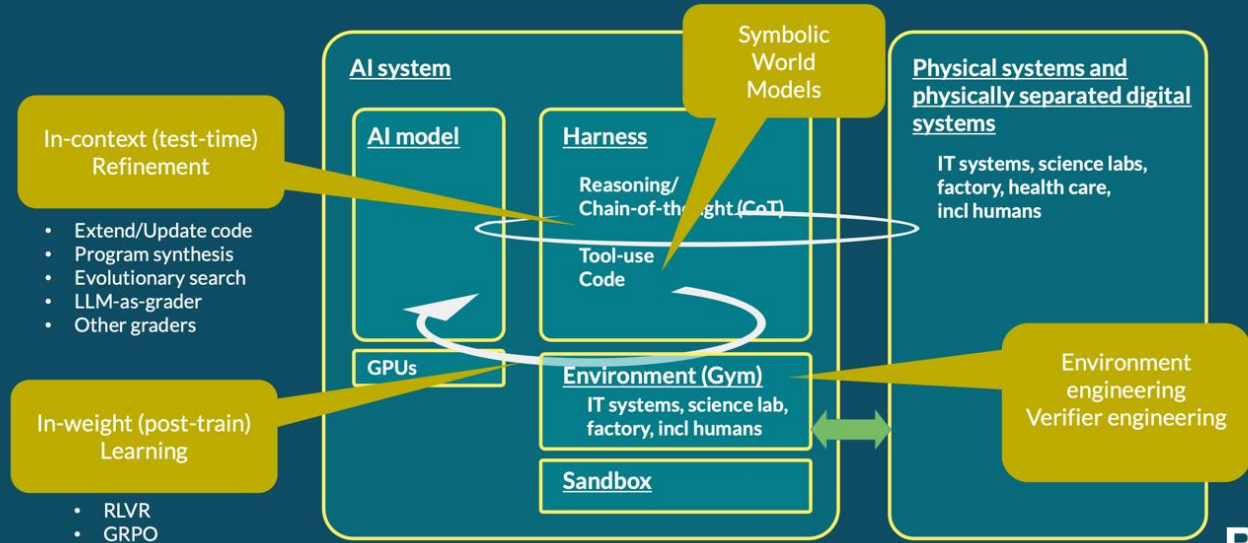


Figure 1: Taxonomy of code as agent harness.

Environment engineering

Overview



Packaged training Environments

Environments

- Emulated services (containers)
- Taskset (prompts)
- Verifiers (LLM-as-judge, prob classifiers (e.g. Jev), symbolic/deterministic)
- Reward function (as a function of verifiers)

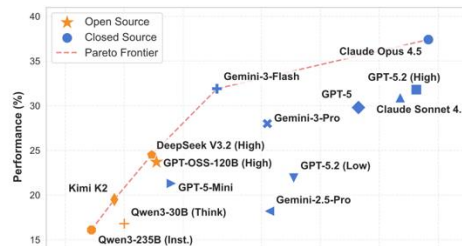


ENTERPRISEOPS-GYM: Environments and Evaluations for Stateful Agentic Planning and Tool Use in Enterprise Settings

Shiva Krishna Reddy Malay^{*1} Shравan Nayak^{*1,2,3} Jishnu Nair¹ Sagar Dasvasam¹ Aman Tiwari¹
Sathwik Tejaswi¹ Sridhar Krishna Nemala¹ Srinivas Sunkara¹ Sai Rajeswar^{1,2,3}

Abstract

Large language models are shifting from passive information providers to active agents intended for complex workflows. However, their deployment as reliable *AI workers* in enterprise is stalled by benchmarks that fail to capture the intricacies of professional environments, specifically, the need for long-horizon planning amidst persistent state changes and strict access protocols. In



13 Mar 2026

Environment engineering

Data is the old new oil
RL Envs is the new new oil

The verification problem... Jevs... "RLCD"

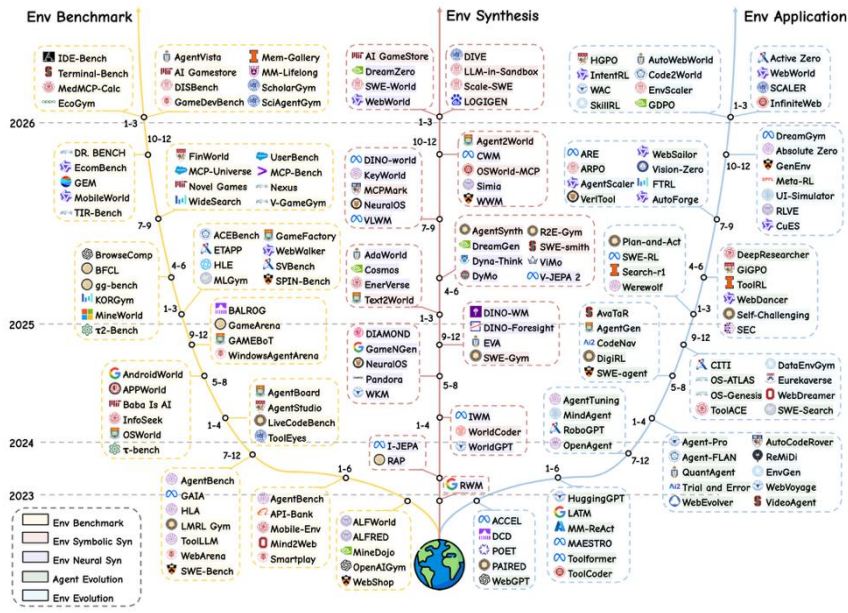
<https://pavlovlist.com/>

JOURNAL OF LATEX CLASS FILES, JANUARY 2025

Agentic Environment Engineering for Large Language Models: A Survey of Environment Modeling, Synthesis, Evaluation, and Application

Jiachun Li*, Zhuoran Jin*, Tianyi Men*, Yupu Hao*, Kejian Zhu*, Lingshuai Wang*, Dongqi Huang*, Longxiang Wang*, Shengjia Hua*, Lu Wang*, Jinshan Gao*, Hongbang Yuan, Rulin Xu, Kang Liu, Jun Zhao†

Abstract—Environments serve as interactive systems for large language model (LLM) based agents across diverse scenarios and play a crucial role in driving the continual evolution of model capabilities. Despite this importance, existing work lacks a systematic



Pavlov's List RL envs and data. By @chrisbarber

RL envs Robotics data

John Schulman: "The intros were helpful, thanks! We talked to most of them and I was impressed by most of them." (used with permission)
If you need data, I can make intros. Request intros.

Domain	Code	Long Horizon	Enterprise	Computer Use	ML	Browsers	Multi-Domain	Security	Finance	Alignment	Math	Tool Use	Science	Creative
Location	SF	London	Marsax	Brooklyn	Mountain View	NYC	Berkeley	Delaware	Lausanne	Palo Alto	Remote	San Mateo	Clear all	Reset sort 07 of 67
Company	Domain	Location	Team Background	Active Founders & RL Leads										
Mercor	Multi-Domain	SF Santa Clara	Thiel Fellowship, Harvard, Geor.	Brendan Foody, Adarsh Hiremath, Surya										
Handshake	Multi-Domain	SF NYC, Bangalore, ...	Palantir, Meta, Scale AI	Garrett Lord RL envs: Paco Guzmán,										
Scale	Multi-Domain Code	SF NYC, DC, London	Stanford, Coinbase, YC	RL envs: Chetan Rane										
Turing	Code	SF Palo Alto, Gurgaon	Stanford, Rover, Yahoo	Jonathan Siddharth, Vijay Krishnan										
Surge	Multi-Domain	SF NYC, Seattle	Google, Meta, Twitter	Edwin Chen RL envs: Nick Heiner										
Snorkel	Multi-Domain Code	SF Redwood City, NYC	Stanford, UM, Meta	Alex Ratner, Christopher Ré, Paro										
micro1	Multi-Domain	SF	Stanford, Berkeley	All Anansi										
AfterQuery	Multi-Domain	SF NYC, Seattle	Wharton, Penn, UBC	Carlos Georgescu, Spencer Mateaga										
Latch	Science	SF	Berkeley, Google, Asimov	Alfredo Andere, Kenny Workman, Kyle Gi										
Mechanize	Code	SF	Epoch AI	Tamay Besiroglu, Ege Erdil, Matthew Ba										
Patronus AI	Multi-Domain Code	SF	Meta AI, Google, Amazon	Anand Kannappan, Rebecca Qian										
Datacurve	Code	SF	Waterloo	Serena Ge, Charley Lee										
Halluminate	Finance Enterprise	SF	Capital One, Meta, Cornell	Jerry Wu, Wyatt Marshall										
Pazeto	Multi-Domain Tool Use	SF	Stanford, Scale, Turing	Phoebe Yao RL envs: Mark Whiting, You										
Proximal	Code	SF Bangalore	Prime Intellect, Cursor	Justus Matthen, Navid Pour, Calvin Che										
Aptura	Finance Enterprise	London SF	Lazard, McKinsey, EF	Chirag Modi										
Bespoke Labs	Code	Mountain View	Google DeepMind, UC Berkeley	Mahesh Sathianmoorthy, Alex Dinakis										
Collinear	Enterprise Long Horizon	Mountain View	Hugging Face, Salesforce, Stanf.	Nazneen Rajani, Soumyadeep Bakshi										
Quesma	Security	Warsaw	Elastic, Sumo Logic, Dynatrace	Jacek Migdal										
ReasonCore	Science Code	SF	Turing, Aera Technology, Oracle	Geoff Wolfe, Anshuman Lall, Satish Vut										
Abundant	Code	SF	Waymo, Google, Mercor	Jesse Hu, Ke Huang, Meji Abidoye										
Aviro	Code ML	SF	YC, Princeton, Georgia Tech	Ayush Karupakula, Aarush Kukreja										
BenchFlow	Enterprise Browsers	SF	Berkeley, Tesla	Xiangyi Li										
Cua	Code Computer Use	SF	Microsoft, Twitter	Francesco Bonacci, Dillon DuPont										
EdotEnv	Long Horizon ML	SF	G-Research, Etched, ETH Zurich	Rui Wang, Michael Zhang										
Emulated	Code ML	SF	AMS, DeepMind, Prime Intellect	Joseph Wang, Sid Patilolliu										
General Reasoning	Finance	London SF	Meta, Conjecture, Aleph Alpha	Ross Taylor, Chengqi Taylor, Kip Par										
Markov	Computer Use Tool Use	SF	YC, IIT Madras	Dev Mandal, Harish Ashok										
Metaphi	Code Enterprise	SF NYC	Waymo, Nuro, Netflix	Abhishek Chandani, Ishan Gupta										
Normal	Hardware Engineering	SF	Cohere, Notion, Fuse	Hudhaz, Anson Yu										
pre.dev	Code Long Horizon	Delaware	Penn MGT, Wharton	Arjun Raj Jain, Adan Elkassas										
Tactic Labs	Science Long Horizon	SF	MosaicML, MSR, EF	Nicole Fitzgerald, Anne Marie Droste										
Ulam	Math	Warsaw London	Oxford	Przemek Chojceki										
Vnax	ML	SF NYC	UCL	Matthew Sargent, Augustine Mavor-Parker										
Originator	Computer Use Code	London Paris	Google, Dataform, Pareto Holdin.	Ben Cambier, Lewis Hennes										
Fleet	Multi-Domain Enterprise	SF NYC	Mercor, Anthropic, MSL	Nicolai Oupzorov, Andrew Zhou										
Paseweave	Long Horizon	SF Remote	Meta, Red Hat	Ivan Kudryavtsev, Louis Pan, Louka Ewi										
Wisestock	Creative	SF	Utah State, MIT	Mikayel Khachatryan, Ashot Matsakanyan										
Deeptime	Code Computer Use	NYC	Hebbia	Tim Lupo										
Muzzle Labs	Long Horizon Code	London Berlin, SF	Turing	Ingram Klein, Amit Choudhary										
Taste Labs	Creative	NYC	Exa, Palantir, Mercado Libre	Thais Castello Branco										
Veritica AI	Creative	SF Vancouver	HRT, Jane Street, Mercor	Rithika Kachan, Rishi Kalakuntla										

RL
SE

<https://arxiv.org/abs/2606.12191>

Test run using Prime Intellect

Eval tooling

Prime Intellect, library: verifiers (Will Brown), <https://github.com/PrimeIntellect-ai/verifiers>

One command

- Model choice
- Environment choice (taskset details)
- Harness choice
- Runtime choice

```
uv run --env-file .env --no-sync eval enterprise-ops-gym \
  --env.taskset.domains calendar \
  --env.agent.runtime.type prime \
  --env.agent.runtime.idle-timeout 300 \
  --env.taskset.task.tools.service-runtime.idle-timeout 300 \
  --client.base-url https://openrouter.ai/api/v1 \
  --client.api-key-var OPENROUTER_API_KEY \
  --model z-ai/glm-5.2 \
  1 -r 1 -c 1 --push False
```

Per-task results

#	Task	bash	bash error	PA	PA error	PA cells
1	Apex new customer setup	-	ToolsetError	0.60	—	8
2	Carrillo site re-enable	-	ToolsetError	1.00	—	12
3	Tier-2 direct call	0.00	—	1.00	—	3
4	Simon NY site transition	0.50	—	0.50	—	24
5	Stark PA-3220 alert	0.75	—	1.00	—	22
6	Carrillo Jira patch	0.60	—	0.40	—	10
7	Globex contact resigned	0.50	—	0.50	—	18
8	Push CS-311 to critical	1.00	—	0.50	—	14
9	Luminatech onboarding	0.60	—	0.60	—	21
10	Acme warranty mismatch	0.67	—	0.67	—	18
11	Russell contract detail	0.67	—	0.67	—	9
12	Horizon field ops	1.00	—	1.00	—	10
13	Escalate CS-1179	0.67	—	0.67	—	9
14	Acme London alert	1.00	—	0.33	—	19
15	Globex Sydney tidy-up	0.75	—	1.00	—	15
16	ServiceNow office move	1.00	—	1.00	—	12
17	Intitech MySQL restore	1.00	—	1.00	—	6
18	Warranty coverage gaps	0.17	—	1.00	—	5
19	Bullock case handover	1.00	—	1.00	—	7

Per-task results, bash vs prime_agent

Partial credit is `verifier_pass_rate`. A dash means the rollout never scored because it died; the number of `ipython` code cells `prime_agent` executed (0 when it never ran).

#	Task	bash	bash error	PA	PA error	PA cells
1	London Wi-Fi drops	0.50	—	0.50	—	8
2	Printer impact raise	0.50	—	0.50	—	6
3	Aisha 2FA lockout	0.33	—	0.33	—	11
4	Reopen INC11 for Carlos	1.00	—	1.00	—	11
5	Email server fault	-	SandboxError	0.33	—	8
6	SF datacenter migration	1.00	—	1.00	—	9
7	Elena desktop no boot	0.00	—	0.00	—	92
8	Repoint Carlos oldest	1.00	—	1.00	—	14
9	Rebalance L1 workload	0.00	—	0.20	—	56
10	Berlin office IT setup	1.00	—	0.88	—	15
11	KB article for INC14	1.00	—	1.00	—	12
12	Report on L1 resolved	1.00	—	1.00	—	14
13	Triage email incidents	-	SandboxError	0.71	—	23
14	Age out stale L1 inc	0.00	—	1.00	—	11
15	Aisha desktop asset	0.50	—	0.50	—	13
16	NextGen portal cutover	1.00	—	1.00	—	21
17	Munich lab location	0.67	—	0.67	—	6

Summary

Summary

