

CAIS'26 Agent Workshop

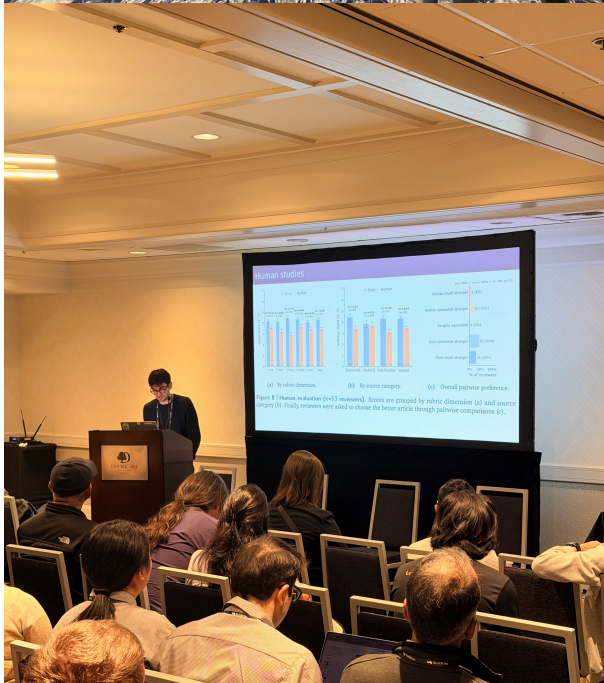
Sunday, May 31, 2026 9:38 PM

Attending time: May 26
Address: San Jose

Summary of categories:

Area	Fow what	Interpretation
AI agents for discovery in the wild	Methods, evaluation, and deployment for AI agents used in real-world discovery.	Focus on the *Discovery*/ self inference
RLEval: Methods and reinforcement learning environments for evaluating ai agents	AI agent evaluation: methods, RL environment design, benchmarks, and real-world case studies.	Common tasks in the reality world
Supporting our ai overlords(SAO)	A workshop on agent-first data systems, agents for data science and analytics, and the future of data systems	Design the data system for agents/ agents for data science
Agents skills 2026	Agent Skills — Design, Evaluation, and Optimization of Procedural Knowledge for LLM Agents	How to input human's thoughts/ tools/ experience to the agents
Agentic software engineering	Bridging AI and Software Engineering for reliable agentic systems.	Bring agents into the realistic production process/ runtime & execution data/ context management

Some pictures:



Insights/Highlights:

1. Multi-agents' inter cooperation
 - a. Different agents make different formats for single task
 - b. Deliver production to next agent may caused different result
2. Standard process of SWE
 - a. Integrating agents into standard software development workflows.
3. Measure agents' failure
 - a. Omission: skip the *important* info
 - b. Source reliability: the quality of sources
 - c. Common templates: iterative

So the main idea of these thoughts are all based on the disadvantages of probability in LLM.



Which means LLM can do the most work in general life but it can't be reliable in every task/step because its shortcome.

- How to bridge this gap which brought by probability?
- How can we create agents that leverage the strengths of LLMs while avoiding their inherent flaws? (Is it possible for multiple agents to mutually correct each other's errors?)



Keywords:

- How to fix these issues?
- Symbolistic presentation
 - Reliable Skills
 - Standard presentation

