Towards Scalable Agent Data Synthesis

Jialong Wu (吴家隆)
Incoming PhD, AAIS, PKU
https://callanwu.github.io/
wujialongml@gmail.com

Background

北京大学 PEKING UNIVERSITY

Agents represent the third level of AI, capable of **thinking** and **acting** in the real world.

OpenAI Imagines Our AI Future

Stages of Artificial Intelligence	
Level 1	Chatbots, AI with conversational language
Level 2	Reasoners, human-level problem solving
Level 3	Agents, systems that can take actions
Level 4	Innovators, AI that can aid in invention
Level 5	Organizations, AI that can do the work of an organization

CURSOR

Claude

Gemini Deep Research

Deep research

manus

The more **advanced** the intelligence, the more **complex** the interactions, the more **scarce** the data.

Background

Q: In a well-known TV drama, the second female lead (actress) entered the entertainment industry in 1993. The current husband of the first female lead (actress) is from Huzhou, Zhejiang. The first male lead (actor) performed on the CCTV Spring Festival Gala six years later. What is the name of this TV drama?
Q:*某知名电视剧，女二号（演员）在1993年进入演艺圈。女一号（演员）的现任丈夫是浙江湖州人。男一号（演员）6年后登上了春晚舞台。问该电视剧是什么？*
A: Love of Parents
A：*父母爱情*

Q: Identify the title of a research publication published before June 2023, that mentions Cultural traditions, scientific processes, and culinary innovations. It is co-authored by three individuals: one of them was an assistant professor in West Bengal and another one holds a Ph.D.
Q：确定一篇发表于2023年6月之前的研究论文标题，该论文提到了文化传统、科学过程和烹饪创新。该论文由三位作者合著，其中一位曾是西孟加拉邦的助理教授，另一位拥有博士学位。

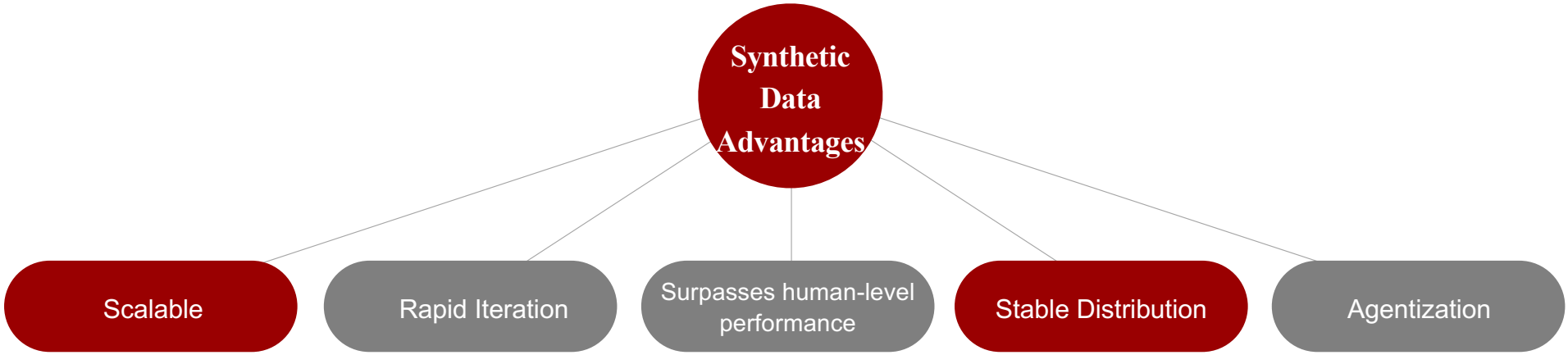
A: The Fundamentals of Bread Making: The Science of Bread
A:*《面包制作基础：面包的科学》*

In Deep Research, tasks require multi-step, long-horizon reasoning, making large-scale human annotation **infeasible** and turning data into **a key bottleneck** for intelligent agents.

Background

How can we build a scalable training data framework for intelligent agents to support the improvement of complex task capabilities?

Post-training for agents requires synthetic data.



Outline

北京大学 PEKING UNIVERSITY

Exploring Tool-Use Capabilities of the o-Series Reasoning Models

Initial Exploration of Data Synthesis

WebWalker

Data synthesis based on explicit task modeling

Training Deep Research

Scalable DR Training Data Synthesis

WebDancer

Data synthesis for generalizing capabilities to new tasks

Extending to General-Purpose Agent Task Scenarios

General Task Environment Modeling and Data Synthesis Paradigm

AgentScaler

Data synthesis for cross-agent tasks and environments

Extending to Human-Expert Scenarios

Real-World Expert-Level Data Synthesis

UniScientist

Data synthesis under a human–AI collaborative paradigm

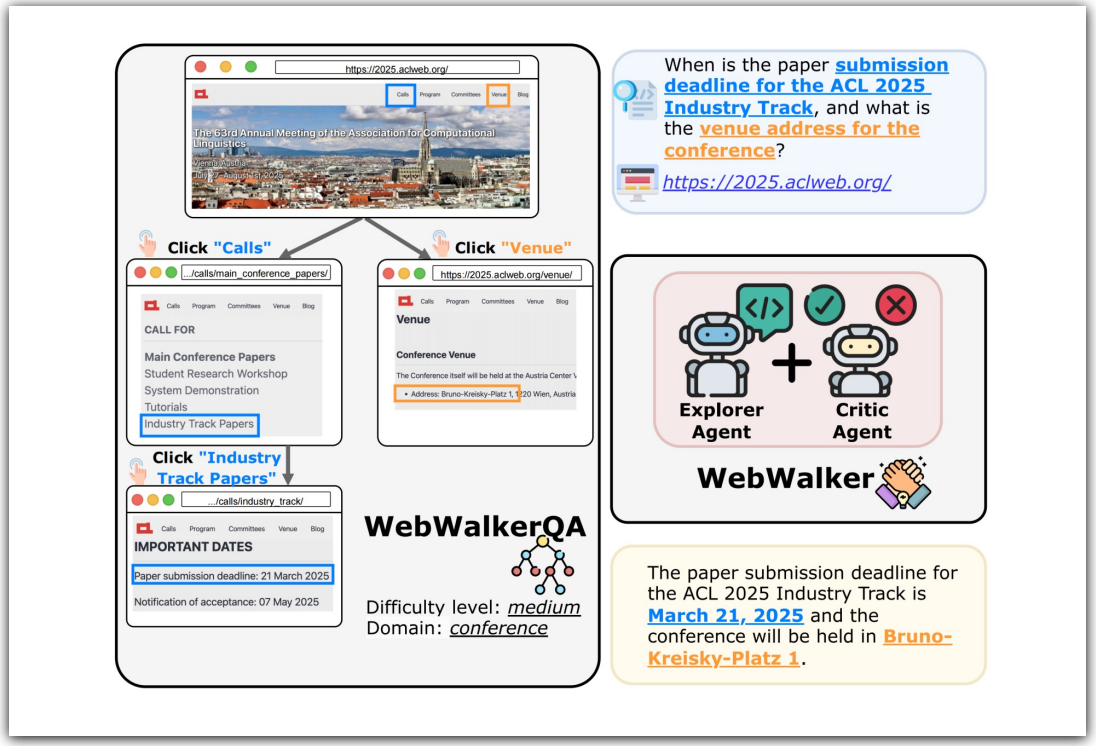
WebWalker

Overview

Constructing training and evaluation datasets for Deep Research–like agent domains.

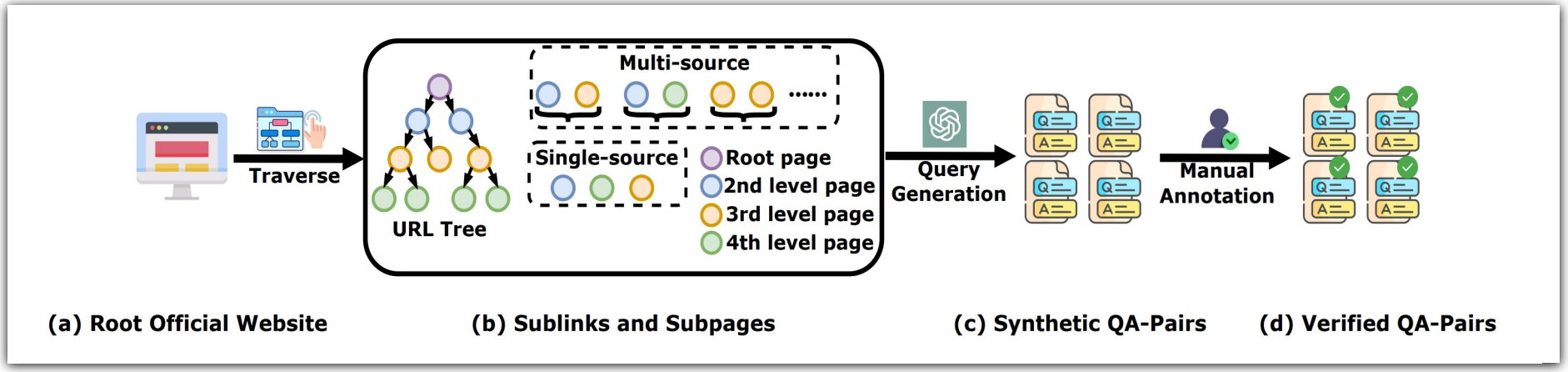
Motivation

Investigating the information gathering and autonomous navigation capabilities of large language models in complex web environments.



WebWalker

Data Generation Pipeline



Insights

- For tasks with explicit formulations, human annotation strategies can be effectively transferred to large language models, enabling scalable labeling and data synthesis.
- The data synthesis process can leverage tool ecosystems similar to those used by agents during problem solving.

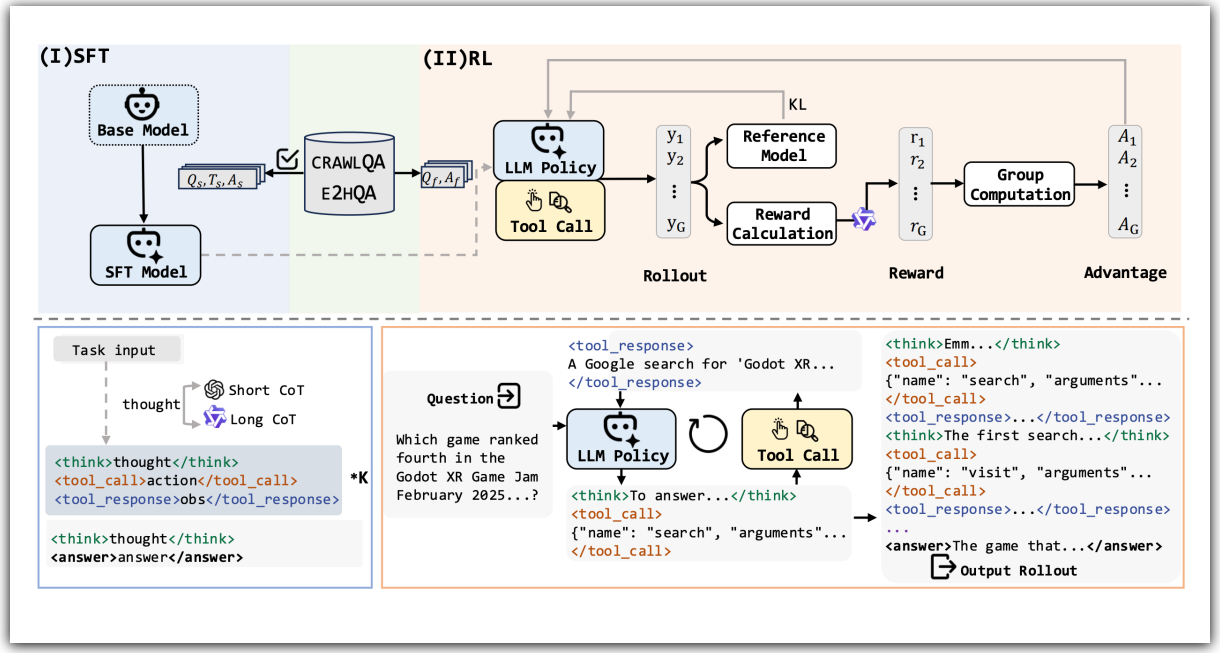
北京大学 PEKING UNIVERSITY

Overview

Training Deep Research agents with synthetic data.

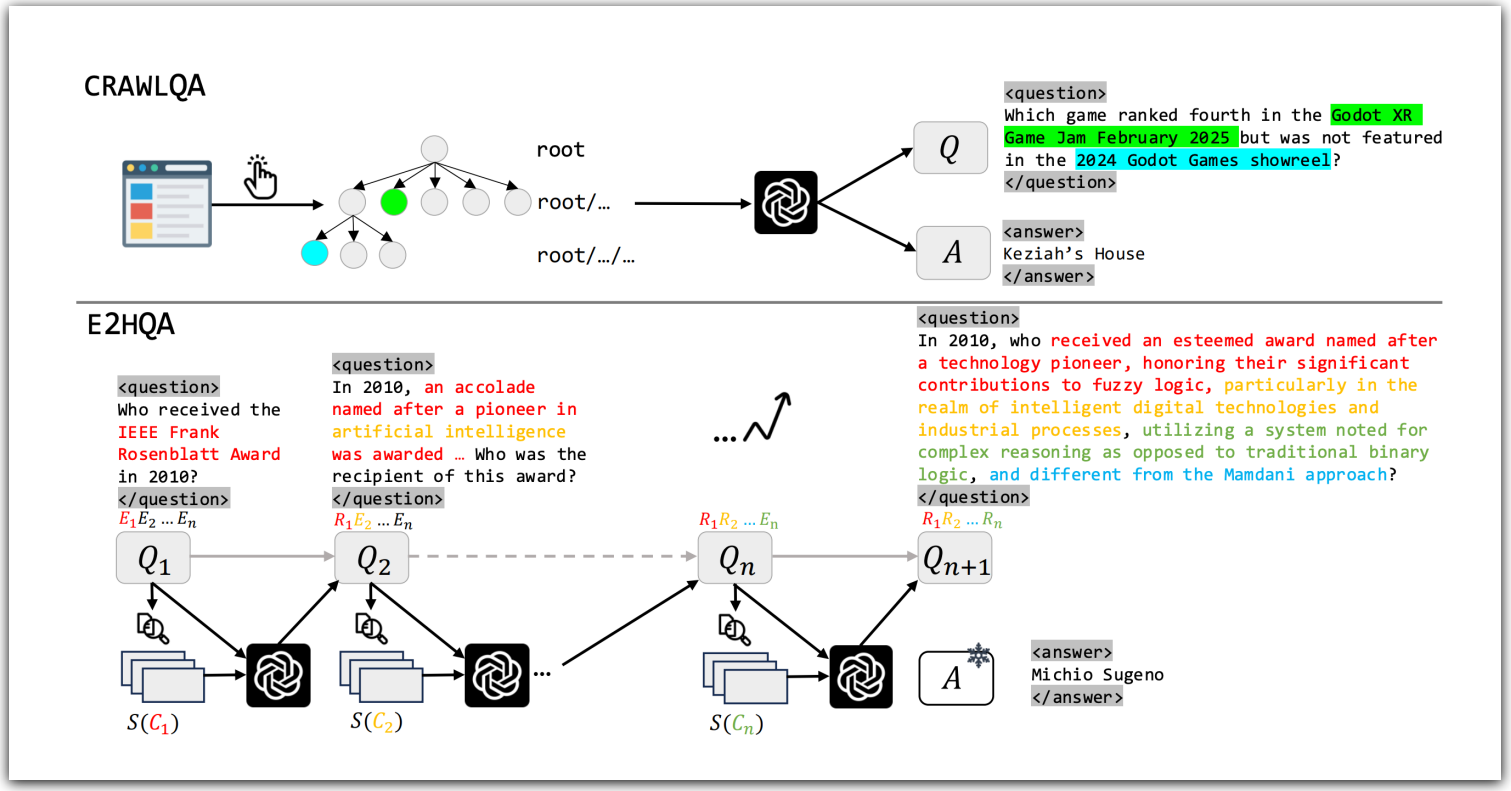
Motivation

Advancing model performance in Deep Research tasks.



WebDancer

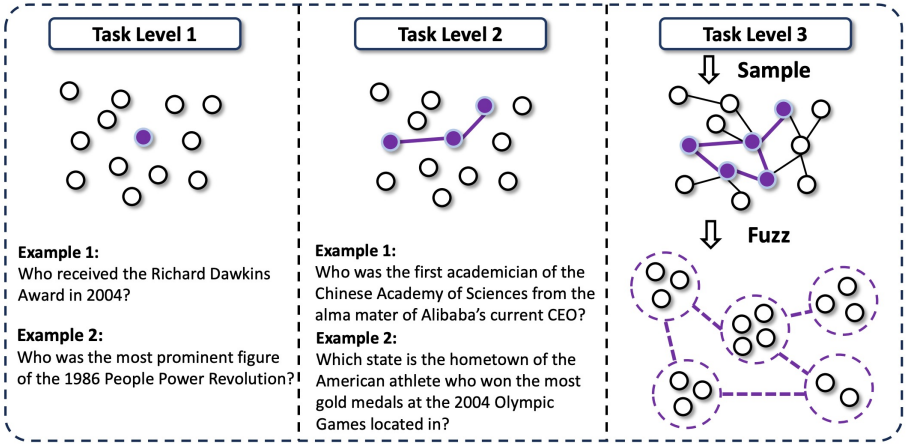
Data Synthesis Pipeline

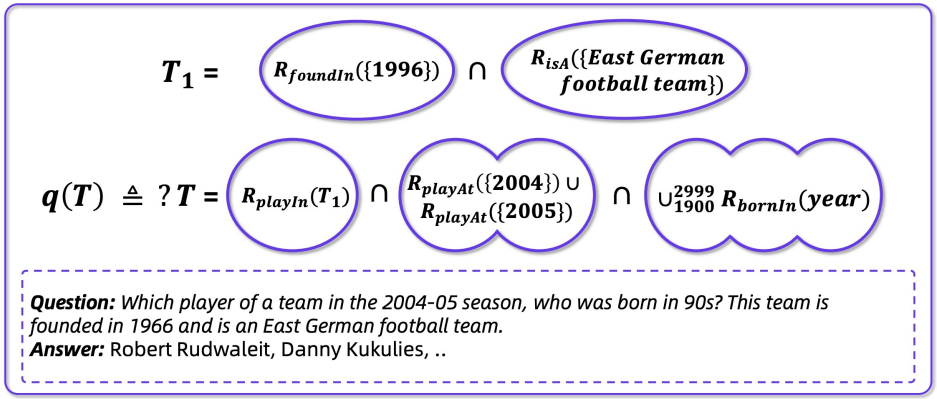


北京大学 PEKING UNIVERSITY

Insights

- New tasks require scalable formulization paradigms.
- Leveraging agents for autonomous data synthesis can surpass human annotation in data quality.
- Post-training can be conducted entirely with synthetic data.
- Obtaining high-quality data hinges on effective selection from large-scale synthetic data.



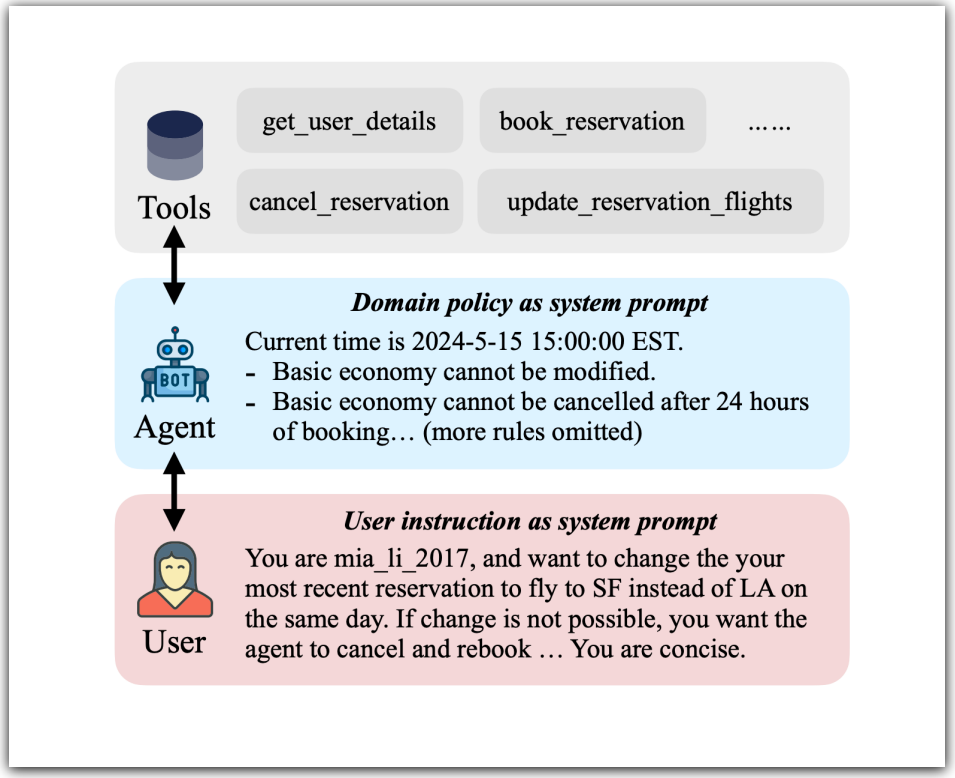


AgentScaler

Overall Modeling general agent tool use as database read–write operations.

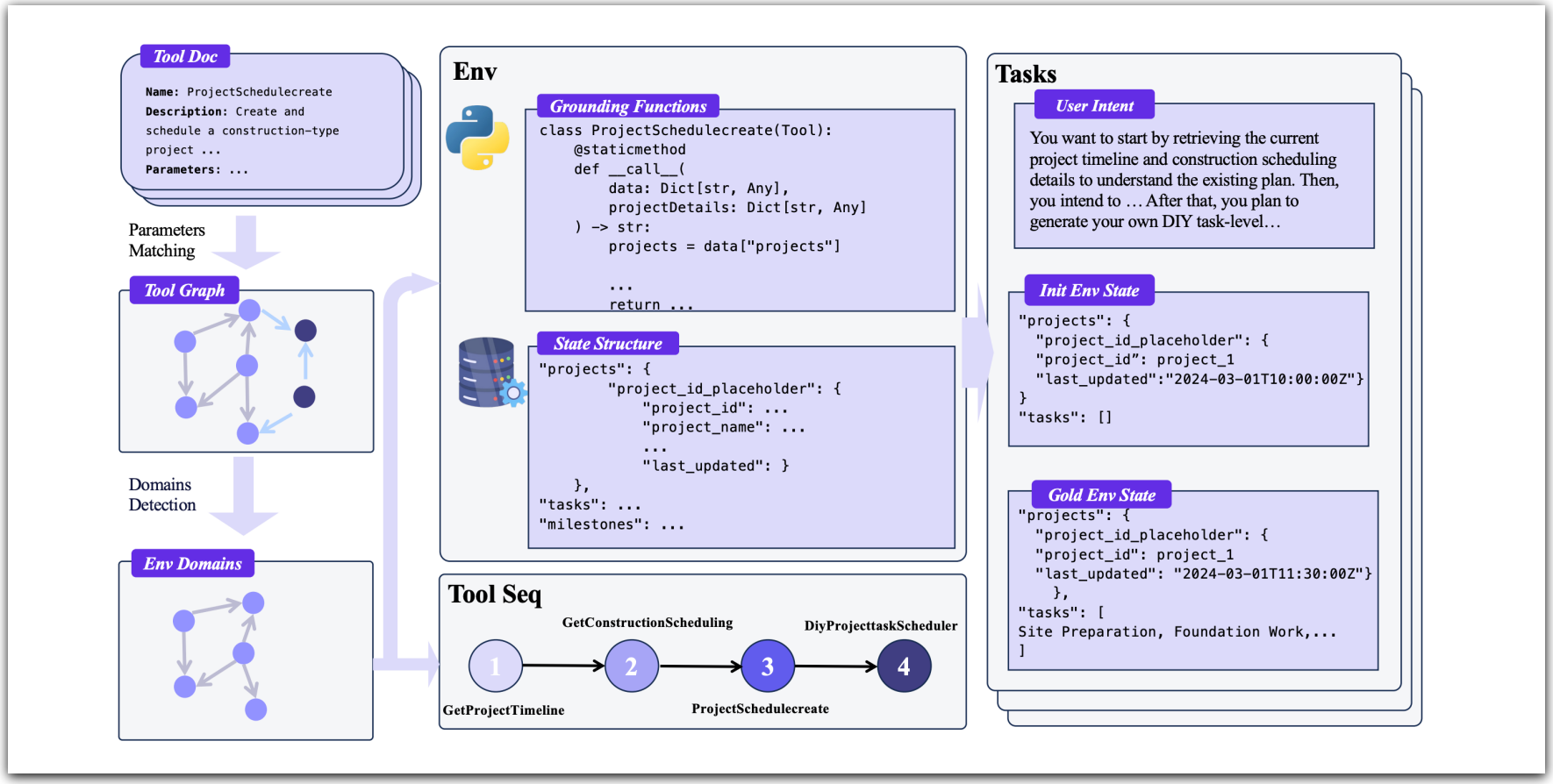
Motivation

Training agents for general capabilities across diverse environments.



AgentScaler

Data Synthesis Pipeline



Insights

- Agents act on the environment, and the environment in turn provides feedback to the agents; thus, agent research should shift from a purely task-driven paradigm to a task–environment co-modeling paradigm.
- Simulated / Mocked environments play a critical role in agent training.

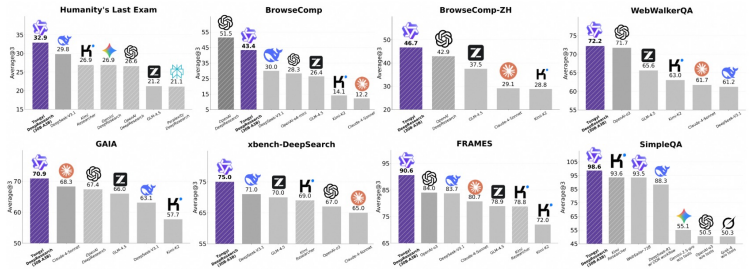
Open-sourced Projects

北京大学 PEKING UNIVERSITY

Tongyi DeepResearch

https://github.com/Alibaba-NLP/DeepResearch

- Achieved over 17K GitHub stars and ranked #1 on GitHub Trending for two consecutive weeks.
- The open-source models have been downloaded over 350K times on Hugging Face and ModelScope.
- Featured by major media outlets, including China Economic News and the South China Morning Post.



UniScientist

北京大学 PEKING UNIVERSITY

Overall

Training models for scientific research capabilities.

Sample Physics Research Problem:

For many physical phenomena involving plasmas, the plasmas can be treated as one or more fluids obeying fluid mechanics and Maxwell's laws. However, some analysis of a plasma requires that the different species of ions and electrons be treated kinetically with an assigned distribution function in phase space.

Additionally, it is common to imagine the distribution function of an element of the plasma, say the electrons, as being anisotropic. Because a magnetized plasma has, by its very nature, a preferred direction (the direction of the magnetic field), it is entirely possible that there be an anisotropy between those species traveling parallel and those species traveling perpendicular to the magnetic field.

One of the features that becomes visible when treating plasmas kinetically is the emergence of Bernstein modes. These are wave modes that travel at frequencies that are higher resonances of the plasma frequency of a species.

Consider a plasma where the electrons are anisotropic such that T|| ≠ T⊥. Consider electrostatic wave modes traveling perpendicular to the magnetic field with the condition on the upper hybrid frequency that $\Omega_e < \omega\_UH \leq 2\Omega_e$

First, determine the dispersion relation for an arbitrary k⊥. You need only consider the first two harmonics. With this result, recover the cold plasma dispersion relation and the first term in the thermal correction.

Next, consider a small parallel component to the wave number (k||), and find the condition for instability.

Sample Physics Research Rubric

Points: 1.0, Item: Analysis restricted to the first two harmonics (only two terms in the infinite sum are taken).

Points: 1.0, Item: Correctly find the second harmonic wave frequency with the first thermal correction,

$(\omega/\Omega_e)^2 = 4 + 3\cdot(\omega_{pe}^2\Omega_e^2)/(3\Omega_e^2 - \omega_{pe}^2)\cdot\mu_e$.

Points: 1.0, Item: Correctly finds the dispersion relation for a general temperature,

$(\omega/\Omega_e)^2 = (1/2)\cdot[5 + 2\cdot(\omega_{pe}^2/\Omega_e^2)\cdot(\Gamma_1(\mu_e)+4\Gamma_2(\mu_e))/\mu_e \pm \sqrt{(9 - 12\cdot(\omega_{pe}^2/\Omega_e^2)\cdot(\Gamma_1(\mu_e)-4\Gamma_2(\mu_e))/\mu_e + 4\cdot(\omega_{pe}^4/\Omega_e^4)\cdot((\Gamma_1(\mu_e)+4\Gamma_2(\mu_e))/\mu_e)^2)}]$.

Points: 1.0, Item: Correctly finds the first harmonic wave frequency with its thermal correction,

$(\omega/\Omega_e)^2 = 1 + (\omega_{pe}^2/\Omega_e^2) - 3\cdot(\omega_{pe}^2\Omega_e^2)/(3\Omega_e^2 - \omega_{pe}^2)\cdot\mu_e$.

Points: 1.0, Item: Correctly gives the general form for the imaginary part of the dielectric constant,

$\mathrm{Im}[\varepsilon] = (\sqrt{\pi})/(k^2\lambda_\parallel^2)\cdot\sum_{\ell}\Gamma_\ell(\mu_e)\cdot(\omega + A\ell\Omega_e)/(k\|a\|_\infty)\cdot e^{-((\omega-\ell\Omega_e)/(k\|a\|_\infty))^2}$.

Points: 1.0, Item: Correctly simplifies the dispersion relation for the cold plasma limit,

$(\omega/\Omega_e)^2 = (1/2)\cdot[5 + (\omega_{pe}^2/\Omega_e^2) \pm \sqrt{(9 - 6\cdot(\omega_{pe}^2/\Omega_e^2)\cdot(1-2\mu_e) + 4\cdot(\omega_{pe}^4/\Omega_e^4)}]$.

Points: 1.0, Item: Correctly simplifies the imaginary part of the dielectric for the first two harmonics,

$\mathrm{Im}[\varepsilon] = (\sqrt{\pi})/(k^2\lambda_\parallel^2)\cdot(\ X_1\cdot(\omega + A\Omega_e)/(k\|a\|_e) + X_2\cdot(\omega + 2A\Omega_e)/(k\|a\|_e)\ )$.

Points: 1.5, Item: Derives the correct condition for wave instability,

$A < -(\omega/\Omega_e)\cdot(X_1+X_2)/(X_1+2X_2) < 0$.

Points: 1.0, Item: Includes the correct equation for the dispersion relation for this situation,

$\varepsilon = 1 - (1/(k^2\lambda_\perp^2))\cdot\sum_{\ell=1}^{\infty}[\ (2\ell^2\Omega_e^2)/(\omega^2-\ell^2\Omega_e^2)\ ]\cdot\Gamma_\ell(\mu_e)$.

Points: 0.5, Item: Recognizes that a negative imaginary part in the susceptibility will drive wave instability.

Sample Chemistry Research Problem

Phthalocyanines are important macrocyclic compounds widely utilized in material and life sciences due to their outstanding physicochemical properties, including intense absorption in the visible-near-IR region and high electron-accepting ability. However, unlike porphyrins, the structural diversity of phthalocyanines has been limited by traditional synthetic methods that rely on one-step condensation of isoindoline units. The development of methodologies that allow controlled modification of the phthalocyanine framework, particularly at the meso-positions, represents a significant advancement in expanding the structural and electronic diversity of these compounds.

Analyze how the modification of meso-nitrogen atoms in nickel(II) phthalocyanine affects the electronic structure, aromaticity, spectroscopic properties, and reactivity of the resulting macrocycles. Explain the synthetic methodology that enables these modifications, the relationship between π-electron count and observed properties, and how these systems compare to traditional phthalocyanines and related porphyrinoids.

Sample Response

1) Introduction: Why Modify the meso-Nitrogens of Phthalocyanines?
Phthalocyanines (Pc) are a class of macrocyclic dyes known for their remarkable chemical, electronic, and optical properties. Their widespread use in materials science, catalysis, and biology stems from strong absorptions in the visible/near-IR regions, exceptional redox characteristics (i.e., high electron affinity), and rigid π-conjugation. ***Traditional synthetic methods, however, typically form the macrocycle in a single step from phthalonitrile or isoindoline precursors, restricting the ability to introduce substitutions selectively at key ring positions — especially the "meso" (bridging) nitrogens.****Recently, methodologies allowing targeted modification of the meso-nitrogen atoms in Ni(II) phthalocyanines have emerged as a powerful means to expand the structural and electronic diversity of these macrocycles.*

(... shortened for the purposes of this figure)

2) Effect on Electronic Structure and Aromaticity
a) π-Electron Count and Ring Currents
Phthalocyanines are classical 18-π-electron aromatic systems, conferring high stability and planarity. Altering the meso-nitrogens through substitution or ring expansion can shift the effective π-electron count. For example, introducing additional unsaturated moieties or heteroatoms may expand the cyclic conjugation to 20 or 22 π-electrons. Changes to the electron count can shift the molecule from Hückel aromaticity (4n + 2 π-electrons) to another aromatic or even antiaromatic regime, affecting ring currents and magnetic properties (NMR chemical shifts, ring currents, etc.).
b) Charge Distribution and Frontier Orbitals
Meso-substitution can tune the electron density distribution around the macrocycle. Electron-donating groups can push electron density toward the ring center, lowering the energy of the lowest unoccupied molecular orbital (LUMO). Electron-withdrawing groups, conversely, stabilize the highest occupied molecular orbital (HOMO) and shift redox potentials toward more positive values. These changes alter both the electrochemical profile (more/less easily reduced or oxidized) and the Q- and B-band positions in the UV-Vis spectrum.

3) Spectroscopic Consequences
a) UV-Vis Absorption (Q and B Bands)
The principal absorption features of phthalocyanines lie in the visible (Q-band, typically 600–700 nm) and near-UV (B-band, typically 300–400 nm). ***Substitution that expands the ring conjugation or introduces strong electron-donating/withdrawing groups can:***
• Shift the Q-band to longer wavelengths (bathochromic shift), reaching into the near-IR, which is highly desirable for optoelectronic and photodynamic applications.
• Alter relative intensities of these bands and merge or split them, reflecting changes in orbital symmetries and energies.
b) NMR Spectroscopy and Aromatic Ring Currents
Modifications to the π-electron count and distribution are directly observed in ^{1}H NMR (and ^{13}C NMR) chemical shifts. ***More highly conjugated (or expanded) aromatic rings exhibit distinct downfield shifts for protons located within induced ring currents, while any partial loss of aromaticity or incorporation of antiaromatic segments can cause atypical shielding/deshielding patterns.***

4) Reactivity and Coordination Chemistry
Because phthalocyanines are often used as redox catalysts or sensors, the meso-nitrogen modifications can significantly influence reactivity:
• Electron-rich meso substituents facilitate nucleophilic or electrophilic attacks at the ring periphery, enabling site-selective functionalizations that are otherwise difficult.

(... shortened for the purposes of this figure)

Sample Grading Rubric

Analysis of Traditional Phthalocyanine Synthesis Limitations (1 point)

✅ **1.0 point: Correctly explains that traditional phthalocyanine synthesis involves one-step condensation with simultaneous formation of all meso-nitrogen bridges, providing limited control over substitution patterns at these positions.**
0.5 point: Mentions limitations of traditional methods but without specific focus on meso-position control challenges.
0.0 point: Fails to identify key limitations of traditional synthetic approaches or provides incorrect analysis.

Thiolate-Mediated Tetramerization Process (1 point)

1.0 point: Correctly describes the thiolate-mediated reductive tetramerization and explains how counter cation size (K+ or Cs+ vs. Na+) affects selectivity between tetramer formation and direct macrocyclization.
0.5 point: Mentions thiolate-mediated tetramerization but without explaining factors controlling selectivity.
❌ **0.0 point: Incorrectly describes the oligomerization process or omits critical details about selectivity control.**

Analysis of NMR Spectroscopic Features (1 point)

1.0 point: Correctly explains that upfield shifts in the 16π system indicate paratropic ring current (antiaromaticity), contrasts this with the broad signals in 17π systems due to paramagnetism, and connects these observations to the underlying electronic structures.
✅ **0.5 point: Identifies basic NMR patterns but without clear connection to ring currents or electronic structure.**
0.0 point: Incorrectly interprets NMR data or fails to connect spectral features to electronic properties.

Electrochemical Property Analysis (1 point)

1.0 point: Correctly explains that the 16π system shows two reversible reductions reflecting conversion to 17π radical and 18π aromatic states, while 17π systems show narrow redox gaps due to facile interconversion between 16π, 17π, and 18π states, and relates these patterns to the underlying electronic structures.
✅ **0.5 point: Describes redox patterns without clearly connecting them to specific electronic state changes.**
0.0 point: Incorrectly interprets electrochemical data or fails to connect redox behavior to electronic properties.

Analysis of Absorption Spectroscopy (1 point)

1.0 point: Correctly explains that the 16π system shows weak/broad absorption due to symmetry-forbidden HOMO–LUMO transitions in antiaromatic systems, while 17π systems show Q-like bands plus NIR-II absorptions characteristic of radical species, and contrasts these with typical phthalocyanine spectral features.
✅ **0.5 point: Describes absorption features but provides limited connection to underlying electronic structures.**
0.0 point: Incorrectly interprets absorption data or fails to relate spectral features to electronic properties.

Reactivity Analysis of Antiaromatic System (1 point)

1.0 point: Correctly explains the high reactivity of the 16π system toward nucleophiles, details specific reactions with hydroxide (ring opening) and hydrazine (ring expansion), and explains how these transformations relieve antiaromatic destabilization.
0.5 point: Mentions reactivity but provides limited analysis of specific transformations or the driving forces behind them.
❌ **0.0 point: Incorrectly analyzes reactivity patterns or fails to connect them to the antiaromatic character of the 16π system.**

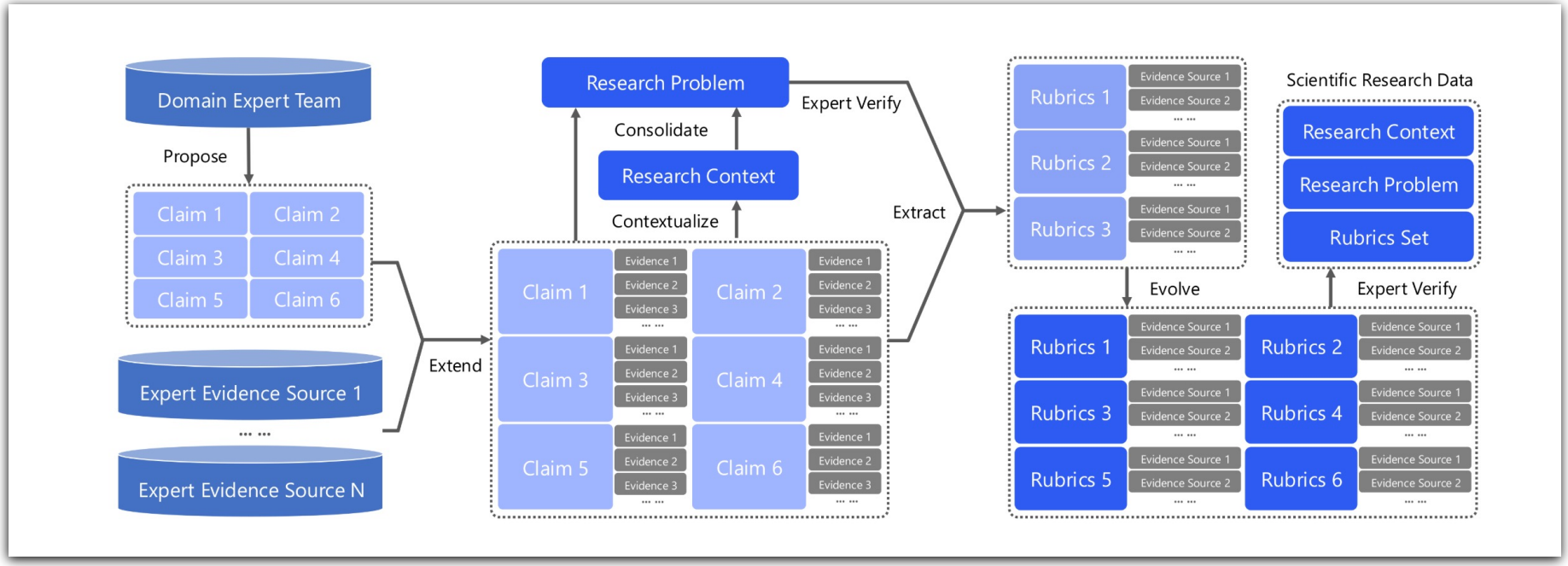
(...and more)

UniScientist

Formalizing Scientific Research: An Agentic Perspective

We model open-ended scientific research as **Active Evidence Integration and Model Abduction**.

Data Synthesis Pipeline



UniScientist

北京大学 PEKING UNIVERSITY

Human-LLM Collaborative Data Production Paradigm

LLMs as Generators

LLMs possess broad, cross-disciplinary knowledge that enables diverse generation at scale, transcending the knowledge boundaries of any single human expert.

Humans as Verifiers

Human experts possess discriminative precision that ensures quality and groundedness. Verification is cognitively far cheaper than creation from scratch.

Evolving Rubrics: Making the Unverifiable Verifiable

Open-ended scientific report = Composition of N closed-ended, verifiable rubric checks

Objective Consistency

For a fixed scientific report, repeated evaluations under the same rubric (performed K times) should yield consistent results, filtering out subjective or unstable criteria.

Discriminability

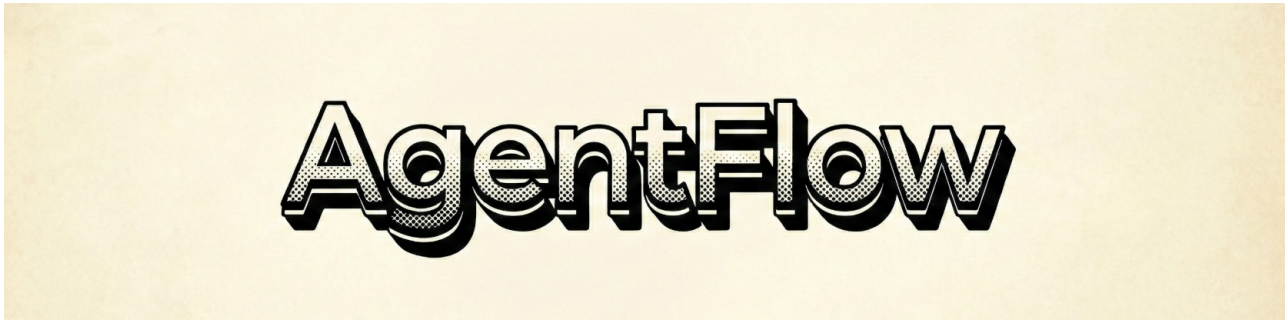
For reports exhibiting graded levels of completeness, rubric-based scores should meaningfully separate quality levels, filtering out trivial criteria.

Atomicity

Each rubric item should test a single knowledge point, avoiding composite criteria that simultaneously assess multiple claims, ensuring clean and verifiable evaluation units.

AgentFlow

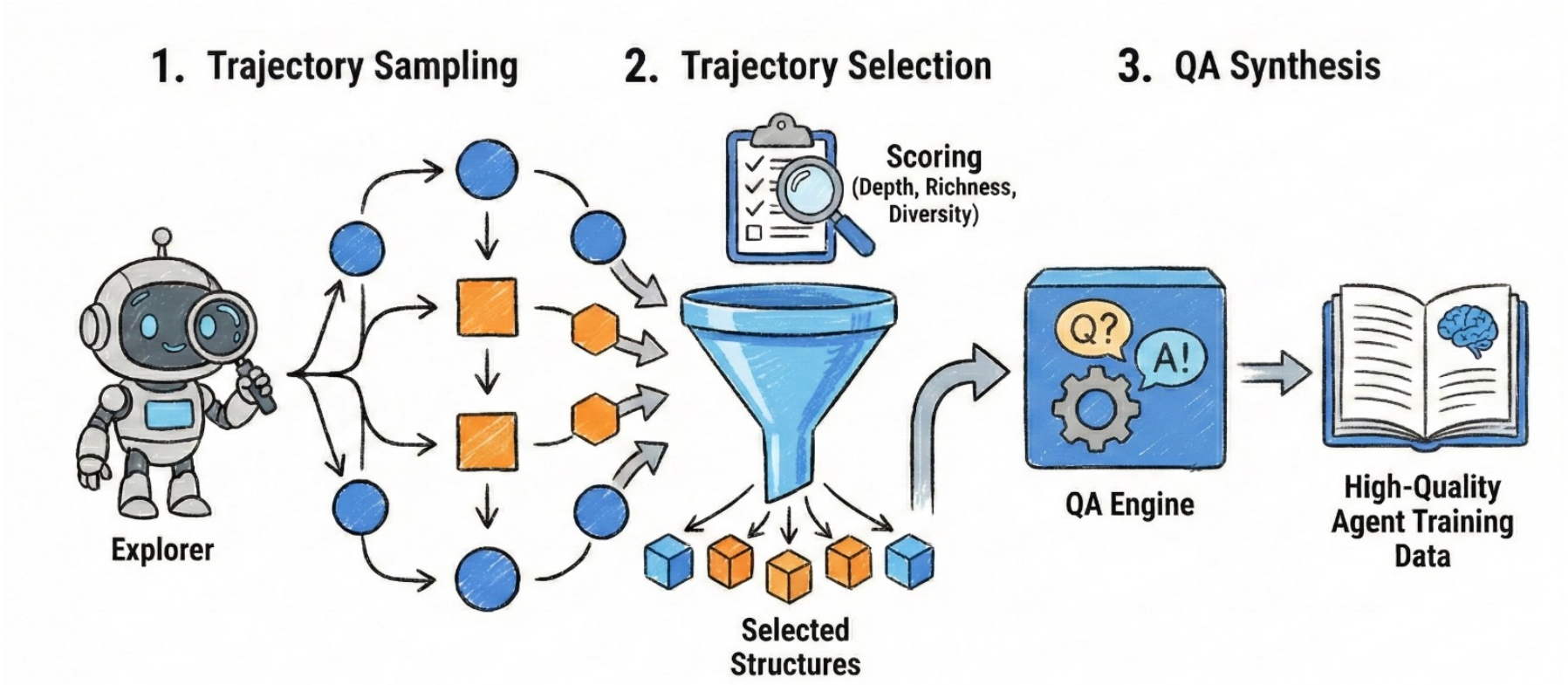
北京大学 PEKING UNIVERSITY



https://github.com/OpenDCAI/AgentFlow

- Unifying agent data synthesis as environment exploration, trajectory selection, task synthesis.
- Constructing an **unified, extensible,** all-in-one environment.
- Exploring model consolidation enabling a single model to possess diverse agent capabilities.

AgentFlow


1. Trajectory Sampling
2. Trajectory Selection
3. QA Synthesis
Scoring
(Depth, Richness, Diversity)
Explorer
Selected Structures
Q?
A!
QA Engine
High-Quality Agent Training Data

Information-Seeking

北京大学 PEKING UNIVERSITY

Question

A major literary work commissioned by the Holy Roman Emperor whose reign began in 1508 was part of his grand artistic legacy. While this patron commissioned famous manuscript anthologies during this period, this specific allegorical epic was distinctively designed for the printing press to ensure a wider audience. What is the exact publication year of its first edition?

一部由神圣罗马帝国皇帝委托创作的重要文学作品，其统治始于1508年，是其宏大艺术遗产的一部分。虽然这位赞助人在这一时期还委托编纂了著名的手稿选集，但这部特定的寓言史诗却是专门为印刷术而设计的，以确保能够面向更广泛的读者群。它的第一版确切出版年份是多少？

Answer 1517

RAG

Question

An 18th-century travelogue, later adapted for a radio series, describes a port town in southeastern England as notable for its rampant illicit trade. This town was also the home of a 16th-century gentleman whose murder led to his wife's execution. Centuries later, another resident of the same town was granted letters patent providing special commercial privileges in a particular year of the early 19th century. During that same year, a collector, whose large collection of manuscript poems was later auctioned, secured a patent for a method of grinding inks. In that year, a patent of nobility was issued to a German family; what is the German term for the princely status it conferred?

一部18世纪的游记后来被改编成广播系列，其中描述了英格兰东南部的一个港口小镇，以猖獗的非法贸易而闻名。这个小镇还是一位16世纪绅士的故乡，他的被谋杀导致其妻子被处决。几个世纪后，这个小镇的另一位居民在19世纪初的某一年获得了授予特定商业特权的专利证书。同一年，一位收藏家（其大量手稿诗集后来被拍卖）获得了一项关于研磨墨水方法的专利。在这一年里，一个德国家族被授予了一项贵族专利；该专利所赋予的“亲王等级”的德语术语是什么？

Answer Fürstenstand

Deep Research

Information-Seeking

北京大学 PEKING UNIVERSITY

Question
What is the difference in percentage-point increase between the overall mean score improvement shown in the bar chart of pre-test versus post-test scores and the improvement for the TIC Principle concept reported in the percentages table?

在前测与后测成绩柱状图中所显示的总体平均分提升，与百分比表中报告的“TIC 原则”概念提升之间，相差多少个百分点？

Answer 17.12

MM-Doc

Question
Find customers whose spending is above the overall average, and show their top 2 most spent music genres along with the amount spent on each.

找出消费金额高于总体平均水平的客户，并展示他们花费最多的前两个音乐类型，以及在每种类型上的消费金额。

Answer
...

NL2SQL

Question
Which feature has the highest importance in predicting 'time / retired' according to the Random Forest model?

根据随机森林模型，在预测“time / retired（退休时间/状态）”时，哪个特征的重要性最高？

Answer laps

Data Analysis

Operating-GUI Agent

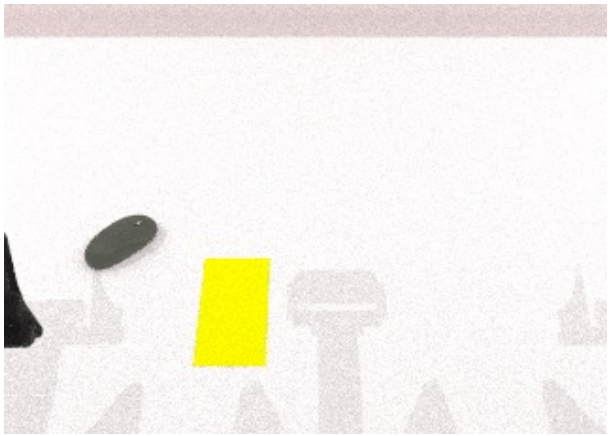
Instruction

I want to audit all command aliases on this Ubuntu machine, so please launch the terminal from the GUI, identify any home directory config files related to shell startup, and then generate a clean, sorted list that combines both currently active aliases and those hidden in your configuration files so I can see the full definitions of commands like alert or ll.

我想审计这台 Ubuntu 机器上的所有命令别名，因此请从图形界面启动终端，找出与 shell 启动相关的主目录配置文件，然后生成一个干净、排序后的列表，将当前正在使用的别名以及配置文件中隐藏的别名合并在一起，这样我就可以查看像 alert 或 ll 这样的命令的完整定义。

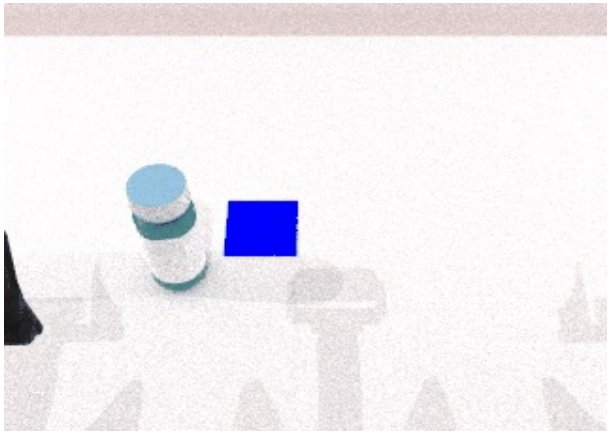
GUI

Operating-VLA

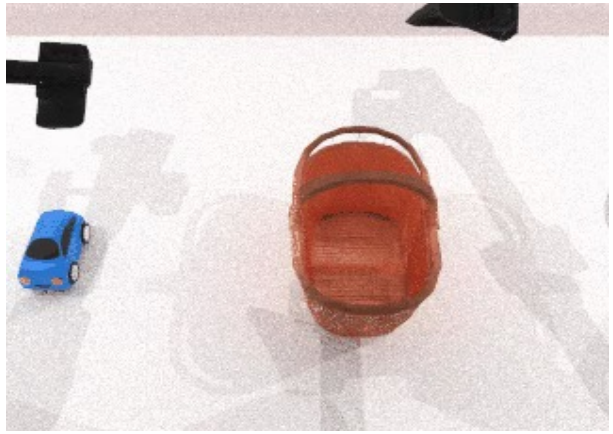


1. Place the mouse on the yellow pad

2. Open the laptop



3. Place the cup on the blue box



4. Store the car in the basket

What's Next

北京大学 PEKING UNIVERSITY

- Scaling data and environments for real-world, high-productivity scenarios

- Exploring data that elicits higher-order intelligence patterns in models

Conclusion

北京大学 PEKING UNIVERSITY

- Agent data research is one of the most **important** and **intriguing** directions in the LLM era.
- The key of agent data lies in **activating task-specific patterns** in models.
- Synthetic data enables scalable data generation while reflecting a **deep understanding** of agent tasks.

References

[1] WebWalker: Benchmarking LLMs in Web Traversal
[2] WebDancer: Towards Autonomous Information Seeking Agency
[3] WebSailor: Navigating Super-human Reasoning for Web Agent
[4] WebShaper: Agentically Data Synthesizing via Information-Seeking Formalization
[5] Towards General Agentic Intelligence via Environment Scaling
[6] Tongyi DeepResearch Technical Report
[7] UniScientist: Advancing Universal Scientific Research Intelligence
[8] https://github.com/OpenDCAI/AgentFlow

Personal Suggestion

北京大学 PEKING UNIVERSITY

- Continuing open-source efforts and long-term, consistent work.
- Building strong research impact.
- Keep pace with the times and seize every opportunity.
- Research is a form of self-fulfilling growth.

吴家隆 (Jialong Wu)
wujialongml@gmail.com
https://callanwu.github.io/
https://github.com/callanwu

THANK YOU FOR YOUR ATTENTION!