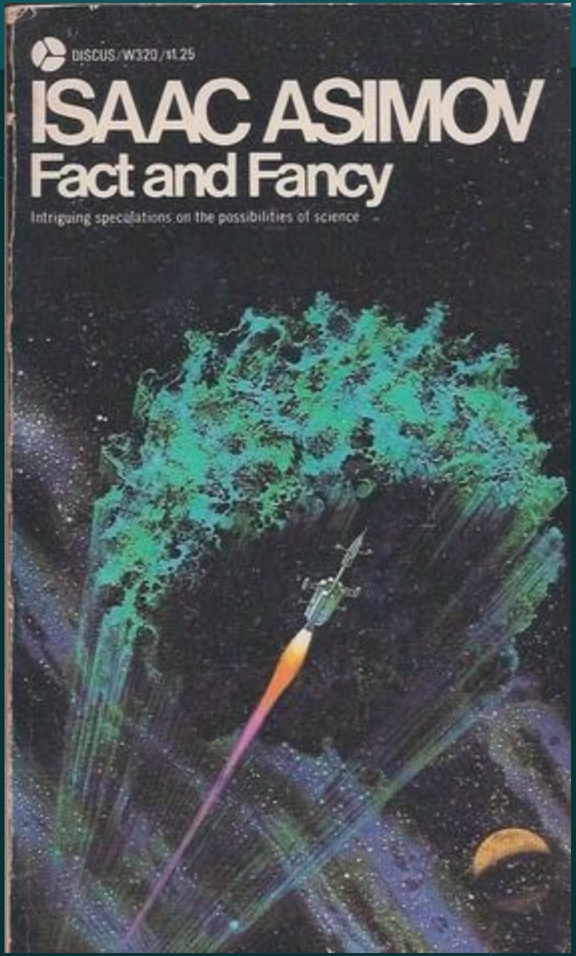


AutoDiscovery

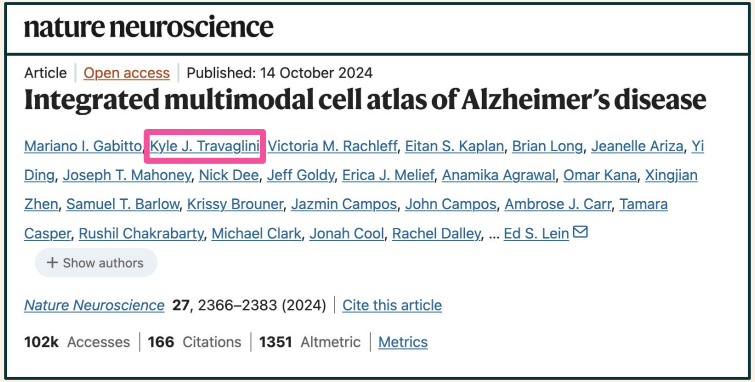
"Intriguing speculations on the possibilities of science"

Bodhi, Dhruv

Ai2

Kyle Travaglini
Scientist, Allen Institute

Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD)



nature neuroscience

Article | Open access | Published: 14 October 2024

Integrated multimodal cell atlas of Alzheimer's disease

Mariano I. Gabitto, Kyle J. Travaglini, Victoria M. Rachleff, Eitan S. Kaplan, Brian Long, Jeanelle Ariza, Yi Ding, Joseph T. Mahoney, Nick Dee, Jeff Goldy, Erica J. Melief, Anamika Agrawal, Omar Kana, Xingjian Zhen, Samuel T. Barlow, Krissy Brouner, Jazmin Campos, John Campos, Ambrose J. Carr, Tamara Casper, Rushil Chakrabarty, Michael Clark, Jonah Cool, Rachel Dalley, ... Ed S. Lein

+ Show authors

Nature Neuroscience **27**, 2366–2383 (2024) | Cite this article

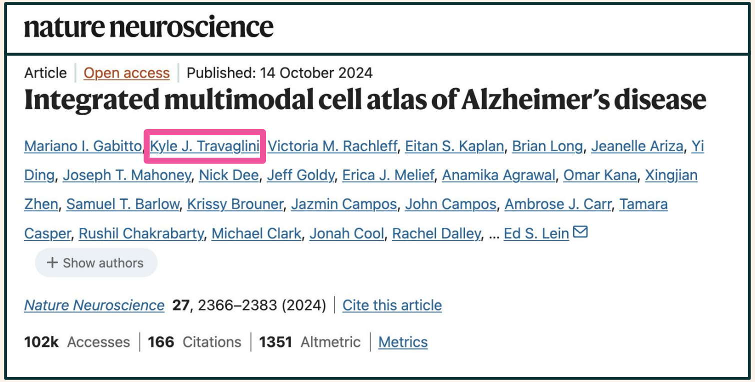
102k Accesses | **166** Citations | **1351** Altmetric | Metrics

Quantitative Neuropathology Data

Measurements of Abeta, pTau, pTDP43, a-synuclein, Neun+ cells, IBA1+ cells, and GFAP+ cells from quantitative analysis of stained neuropathology images from Middle Temporal Gyrus (MTG).

Kyle Travaglini
Scientist, Allen Institute

Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD)


nature neuroscience

Article | Open access | Published: 14 October 2024

Integrated multimodal cell atlas of Alzheimer's disease

Mariano I. Gabitto, Kyle J. Travaglini, Victoria M. Rachleff, Eitan S. Kaplan, Brian Long, Jeanelle Ariza, Yi Ding, Joseph T. Mahoney, Nick Dee, Jeff Goldy, Erica J. Melief, Anamika Agrawal, Omar Kana, Xingjian Zhen, Samuel T. Barlow, Krissy Brouner, Jazmin Campos, John Campos, Ambrose J. Carr, Tamara Casper, Rushil Chakrabarty, Michael Clark, Jonah Cool, Rachel Dalley, ... Ed S. Lein

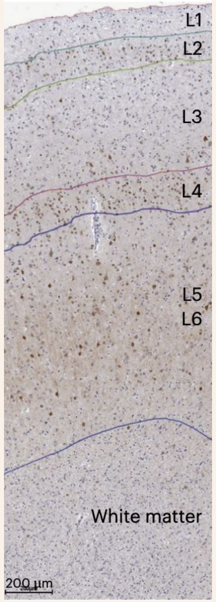
+ Show authors

Nature Neuroscience **27**, 2366–2383 (2024) | Cite this article

102k Accesses | **166** Citations | **1351** Altmetric | Metrics

Quantitative Neuropathology Data

Measurements of Abeta, pTau, pTDP43, a-synuclein, Neun+ cells, IBA1+ cells, and GFAP+ cells from quantitative analysis of stained neuropathology images from Middle Temporal Gyrus (MTG).



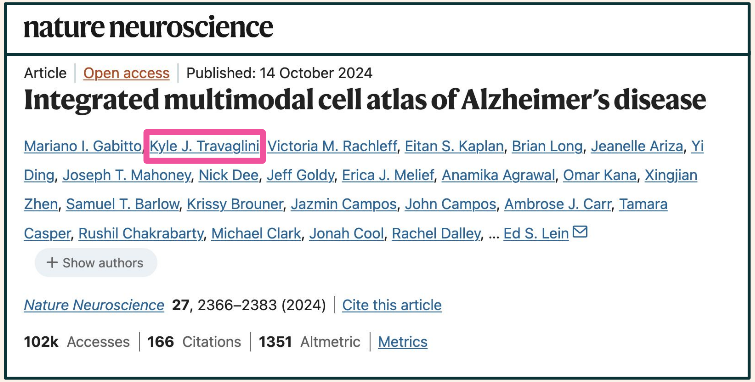
KyleT found:

AT8 + cells increased exponentially with disease progression and that **deep layers (5–6) accumulated more pathology** at advanced stages.

Ai2

Kyle Travaglini
Scientist, Allen Institute

Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD)


nature neuroscience

Article | Open access | Published: 14 October 2024

Integrated multimodal cell atlas of Alzheimer's disease

Mariano I. Gabitto, Kyle J. Travaglini, Victoria M. Rachleff, Eitan S. Kaplan, Brian Long, Jeanelle Ariza, Yi Ding, Joseph T. Mahoney, Nick Dee, Jeff Goldy, Erica J. Melief, Anamika Agrawal, Omar Kana, Xingjian Zhen, Samuel T. Barlow, Krissy Brouner, Jazmin Campos, John Campos, Ambrose J. Carr, Tamara Casper, Rushil Chakrabarty, Michael Clark, Jonah Cool, Rachel Dalley, ... Ed S. Lein

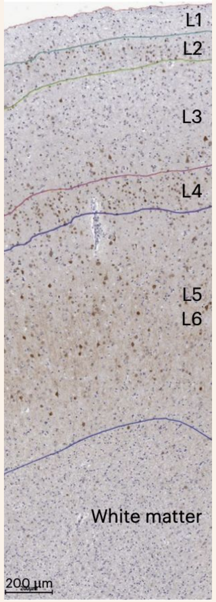
+ Show authors

Nature Neuroscience **27**, 2366–2383 (2024) | Cite this article

102k Accesses | **166** Citations | **1351** Altmetric | Metrics

Quantitative Neuropathology Data

Measurements of Abeta, pTau, pTDP43, a-synuclein, Neun+ cells, IBA1+ cells, and GFAP+ cells from quantitative analysis of stained neuropathology images from Middle Temporal Gyrus (MTG).



KyleT found:

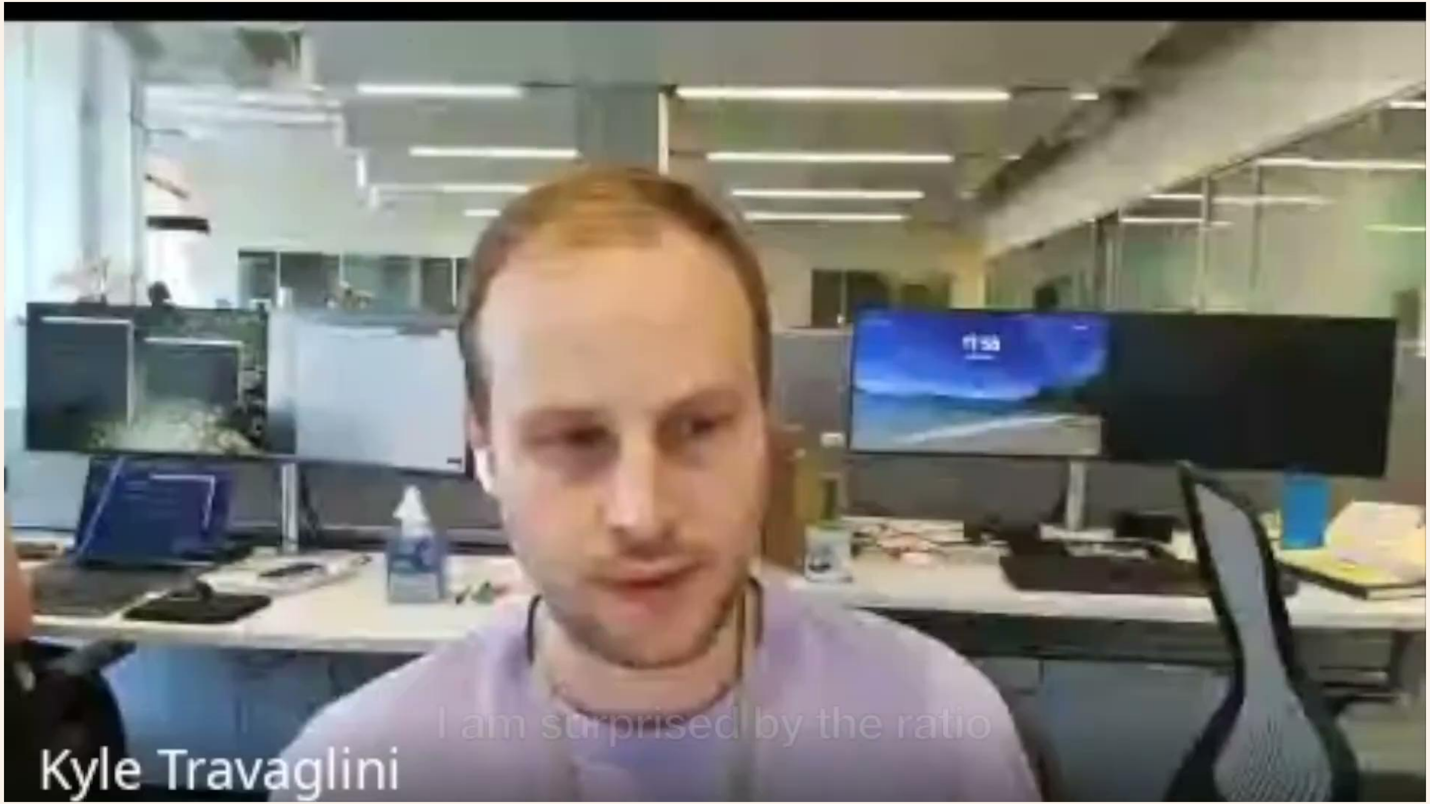
AT8 + cells increased exponentially with disease progression and that **deep layers (5–6) accumulated more pathology** at advanced stages.

AutoDiscovery found:

A **higher cortical laminar gradient of AT8**—quantified as the **ratio** of percent AT8 positive area in **deep** layers (5–6) over **superficial layers** (1–4)—is associated with dementia and **significantly predicts** cognitive status, adjusting for APOE genotype.

Ai2

Kyle Travaglini
Scientist, Allen Institute



AutoDiscovery found:

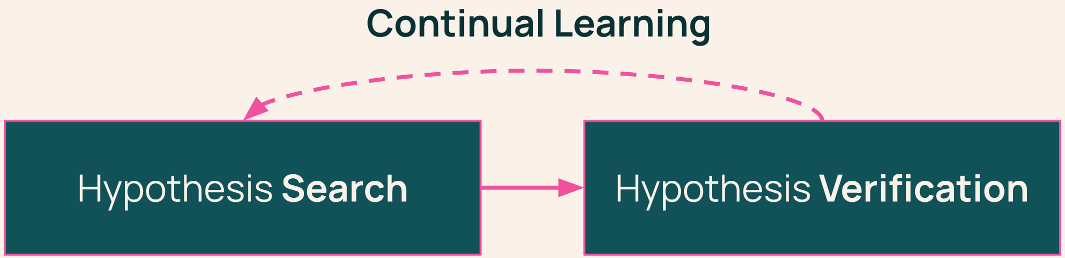
A **higher cortical laminar gradient of AT8**—quantified as the **ratio** of percent AT8 positive area in **deep** layers (5–6) over **superficial layers** (1–4)—is associated with dementia and **significantly predicts** cognitive status, adjusting for APOE genotype.

What are the ingredients of a discovery system?

What are the ingredients of a discovery system?

Hypothesis **Search** → Hypothesis **Verification**

What are the ingredients of a discovery system?



What are the ingredients of a discovery system?

Continual Learning

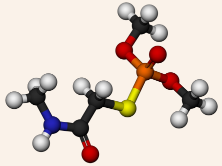
Hypothesis **Search** → Hypothesis **Verification**

Example hypotheses?

Donors with a known head injury are not more likely to have dementia at the time of death

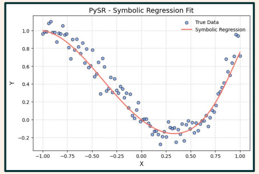
Natural Language

Statistical Tests



Molecule

Simulations / Wet-lab Tests



Equation

Curve-fitting

Ai2

What are the ingredients of a discovery system?

Continual Learning

Hypothesis **Search** → Hypothesis **Verification**

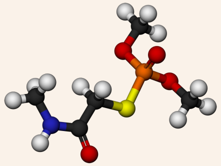
Example hypotheses?

This talk

Donors with a known head injury are not more likely to have dementia at the time of death

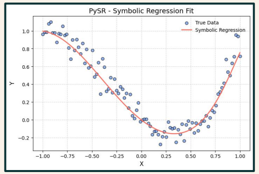
Natural Language

Statistical Tests



Molecule

Simulations / Wet-lab Tests



Equation

Curve-fitting

Ai2

What are the ingredients of a discovery system?

Continual Learning

How do **scientists** do this today?

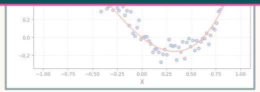
dementia at the time of death

Natural Language

Statistical Tests

Molecule

Simulations / Wet-lab Tests



Equation

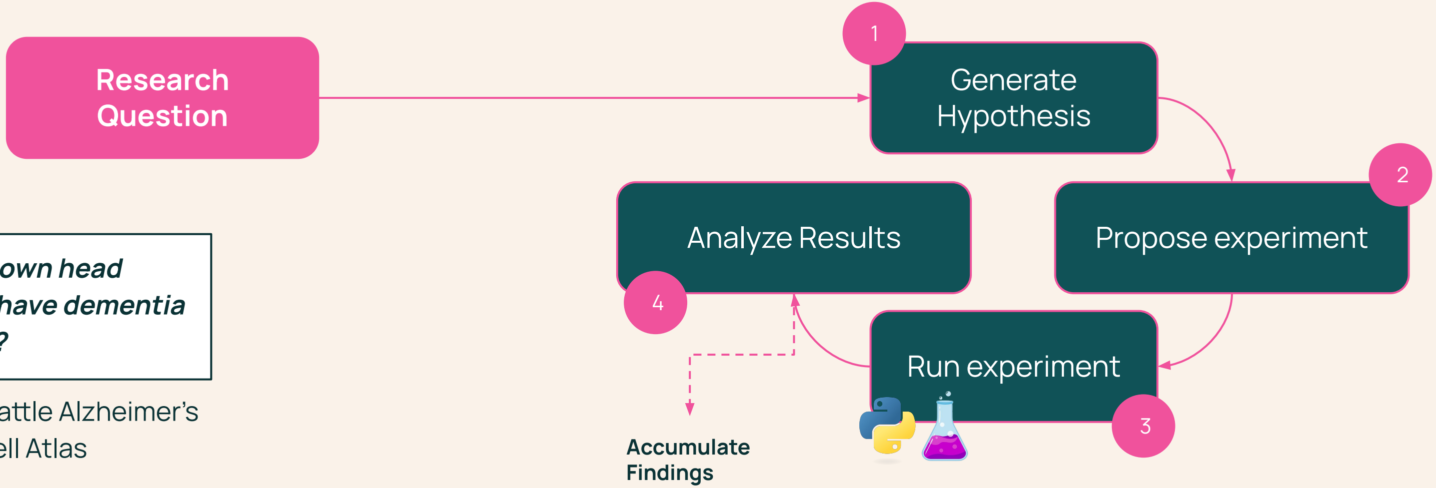
Curve-fitting

Ai2

Research Question

Are donors with a known head injury more likely to have dementia at the time of death?

Sea-AD: The Seattle Alzheimer's Disease Brain Cell Atlas

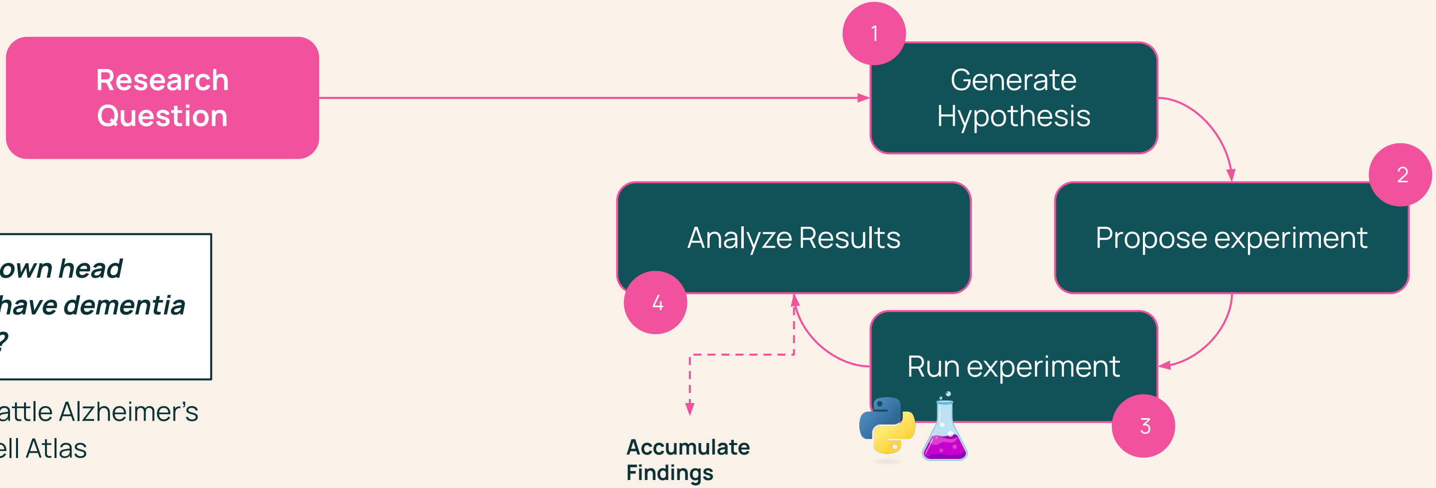


Are donors with a known head injury more likely to have dementia at the time of death?

Sea-AD: The Seattle Alzheimer's Disease Brain Cell Atlas

Goal-driven

← Degree of supervision



Are donors with a known head injury more likely to have dementia at the time of death?

Sea-AD: The Seattle Alzheimer's Disease Brain Cell Atlas

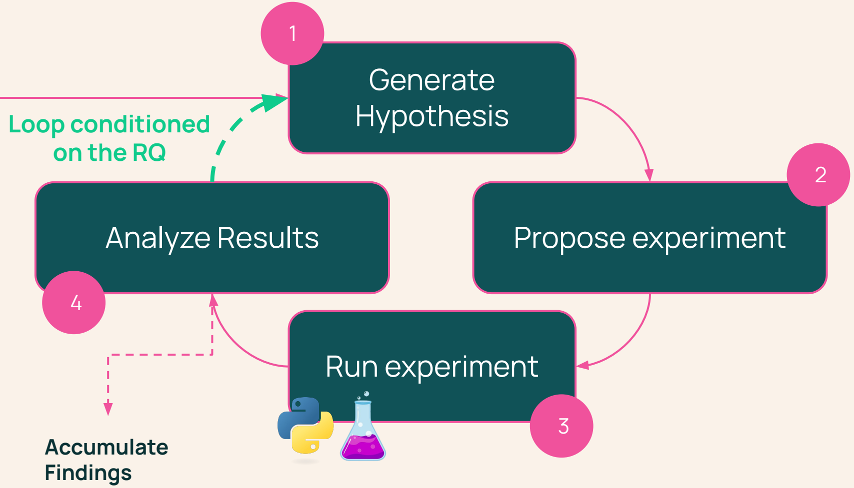
Goal-driven Exploration

← Degree of supervision

Research Question

What other factors increase the likelihood of dementia?

Sea-AD: The Seattle Alzheimer's Disease Brain Cell Atlas



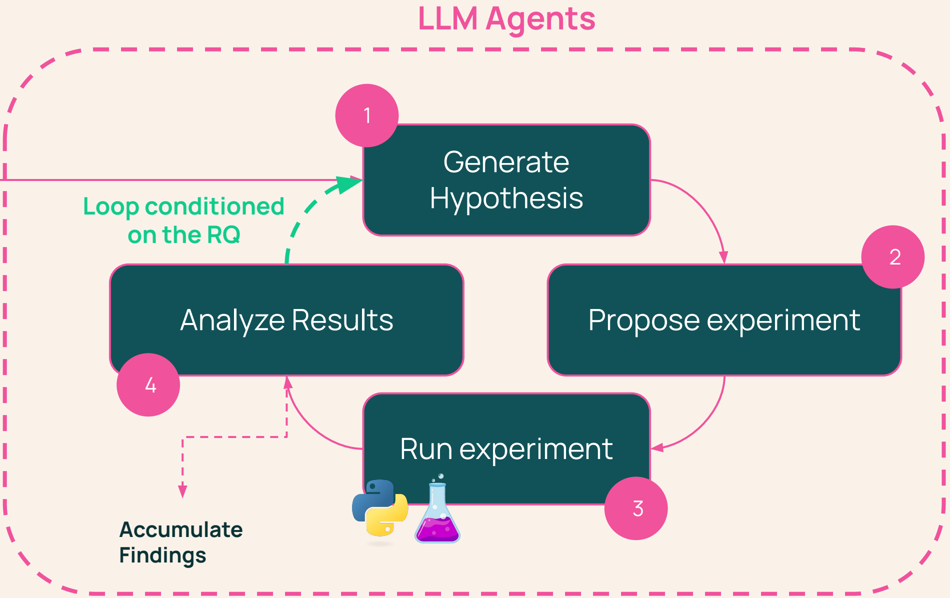
Goal-driven Exploration

← Degree of supervision

Research Question

What other factors increase the likelihood of dementia?

Sea-AD: The Seattle Alzheimer's Disease Brain Cell Atlas



Goal-driven Exploration

← Degree of supervision

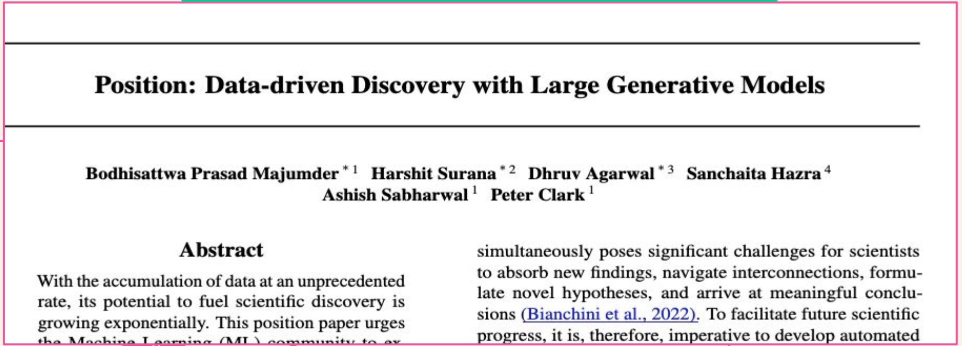
Asta DataVoyager

ICML
International Conference On Machine Learning

Research Question

What other factors increase the likelihood of dementia?

Sea-AD: The Seattle Alzheimer's Disease Brain Cell Atlas



Position: Data-driven Discovery with Large Generative Models

Bodhisattwa Prasad Majumder [*1] Harshit Surana [*2] Dhruv Agarwal [*3] Sanchaita Hazra [4]
Ashish Sabharwal [1] Peter Clark [1]

Abstract

With the accumulation of data at an unprecedented rate, its potential to fuel scientific discovery is growing exponentially. This position paper urges

simultaneously poses significant challenges for scientists to absorb new findings, navigate interconnections, formulate novel hypotheses, and arrive at meaningful conclusions (Bianchini et al., 2022). To facilitate future scientific progress, it is, therefore, imperative to develop automated

Run experiment

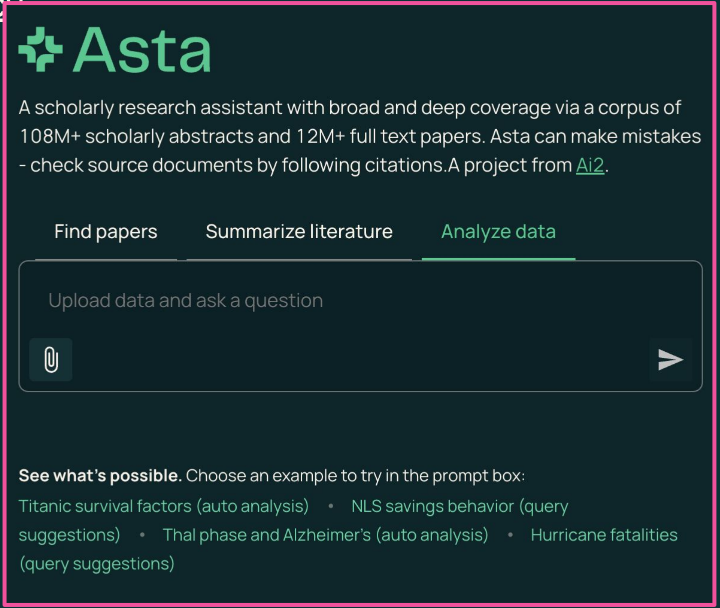
3

Accumulate Findings

Asta DataVoyager announced at Madrona IA Summit!

The Cancer AI Alliance (CAIA) unveiled DataVoyager, the "first collaborative AI platform for cancer research."

Running over real patient data without sacrificing privacy



Asta DataVoyager announced at Madrona IA Summit!

Ai2

The Cancer AI Alliance (CAIA) unveiled DataVoyager, the “first collaborative AI platform for cancer research.”

Running over real patient data without sacrificing privacy

Asta

A scholarly research assistant with broad and deep coverage via a corpus of 108M+ scholarly abstracts and 12M+ full text papers. Asta can make mistakes - check source documents by following citations.A project from Ai2.

Find papers | Summarize literature | Analyze data

Upload data and ask a question

See what’s possible. Choose an example to try in the prompt box:
Titanic survival factors (auto analysis) • NLS savings behavior (query suggestions) • Thal phase and Alzheimer’s (auto analysis) • Hurricane fatalities (query suggestions)

THE AI ECONOMY

Ai2’s Asta DataVoyager Makes Complex Data Analysis Effortless

Scientists can now ask questions in plain language and get reproducible insights—no coding required

KEN YEUNG
OCT 01, 2025

GeekWire

Cancer AI Alliance says new tech platform will speed breakthroughs with novel privacy approach

KUOW .ORG 94.9 npr Network

New AI platform led by Fred Hutch aims to accelerate cancer breakthroughs

Newsweek.AI

Cancer Researchers Find a Way Around AI's Biggest Bottleneck: Data Sharing

Asta DataVoyager announced at Madrona IA Summit!

Ai2

The Cancer AI Alliance (CAIA) unveiled DataVoyager, the “first collaborative AI platform for cancer research.”

Running over real patient data without sacrificing privacy

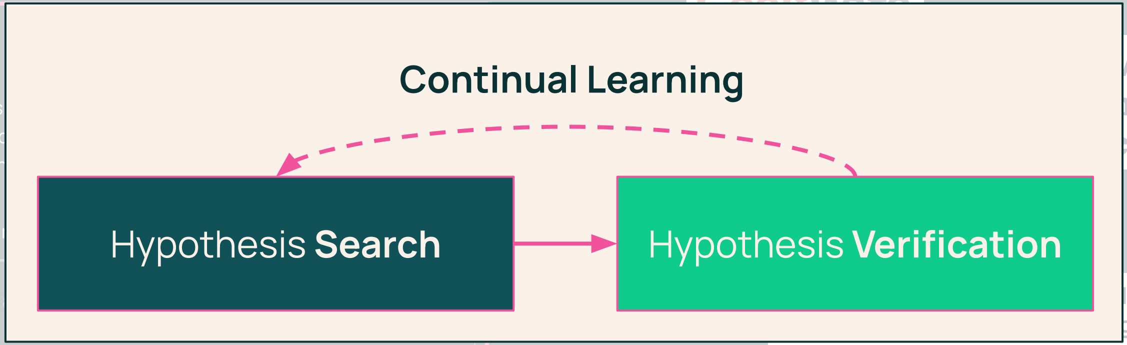
Asta

A scholarly research ass
108M+ scholarly abstrac
- check source documen

Find papers Su

Upload data and ask

See what’s possible. Choose an example to try in the prompt box:
Titanic survival factors (auto analysis) • NLS savings behavior (query suggestions) • Thal phase and Alzheimer's (auto analysis) • Hurricane fatalities (query suggestions)



THE AI ECONOMY

Ai2’s Asta DataVoyager Makes Complex Data Analysis Effortless

Scientists can now ask questions in plain language and get reproducible insights—no coding required

KEN YEUNG
OCT 01, 2025

w tech
roughs
h

lutch aims to
ghs

Newsweek.AI

Cancer Researchers Find a Way Around AI's Biggest Bottleneck: Data Sharing

Goal-driven Exploration

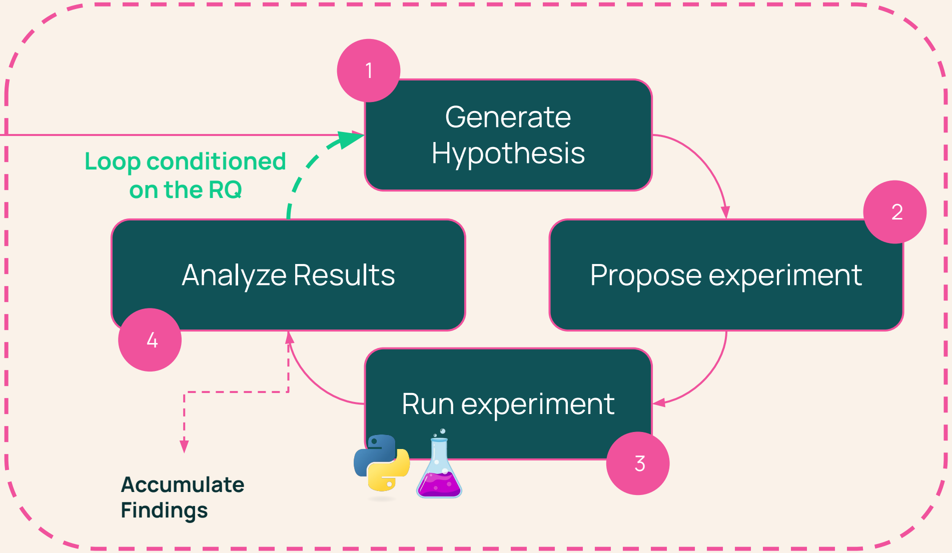
← Degree of supervision

Research Question

What other factors increase the likelihood of dementia?

Sea-AD: The Seattle Alzheimer's Disease Brain Cell Atlas

LLM Agents



Ai2

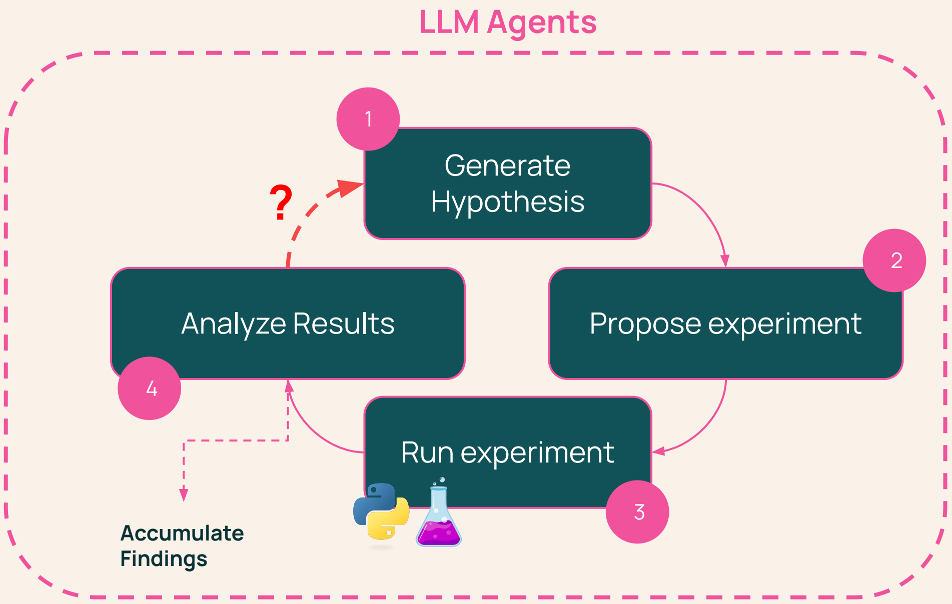
What if we don’t start with an RQ?

← Degree of supervision

Research Question

What other factors increase the likelihood of dementia?

Sea-AD: The Seattle Alzheimer’s Disease Brain Cell Atlas



Ai2

What if we don’t start with an RQ?

← Degree of supervision

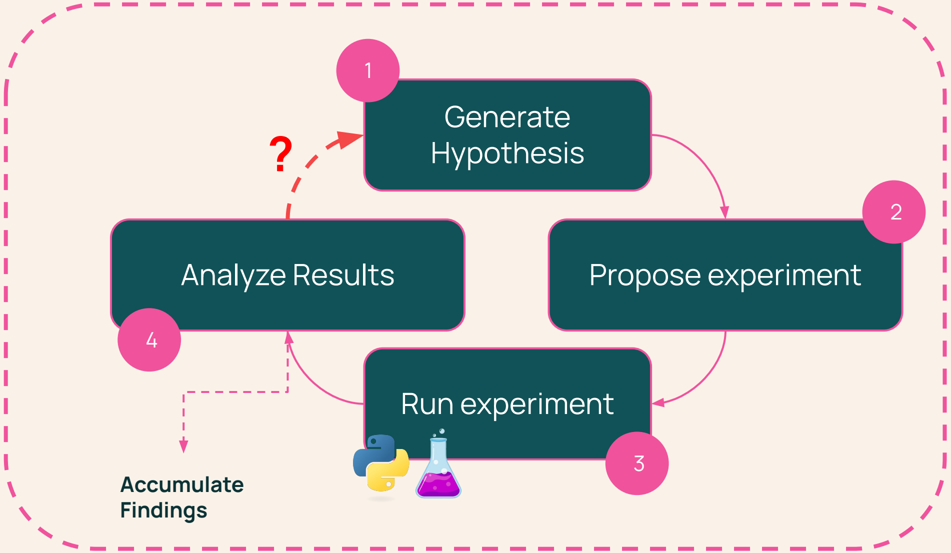
Research Question

Which hypotheses should be prioritized for testing?

likelihood of dementia?

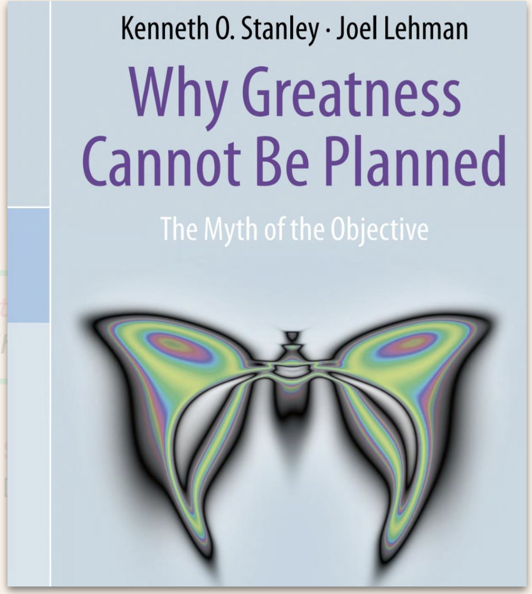
Sea-AD: The Seattle Alzheimer’s Disease Brain Cell Atlas

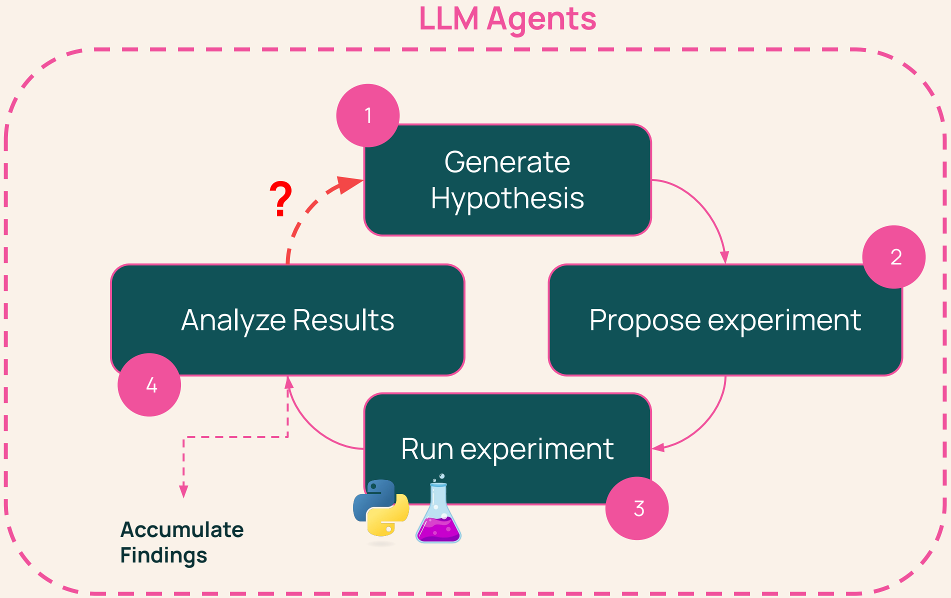
LLM Agents



“Open-ended” Discovery

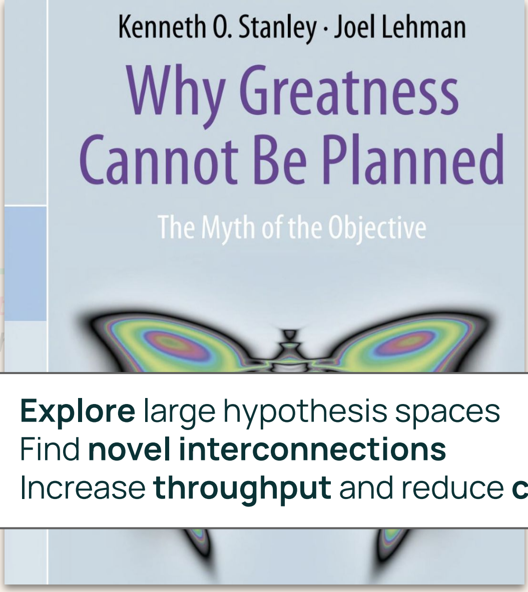
← Degree of supervision





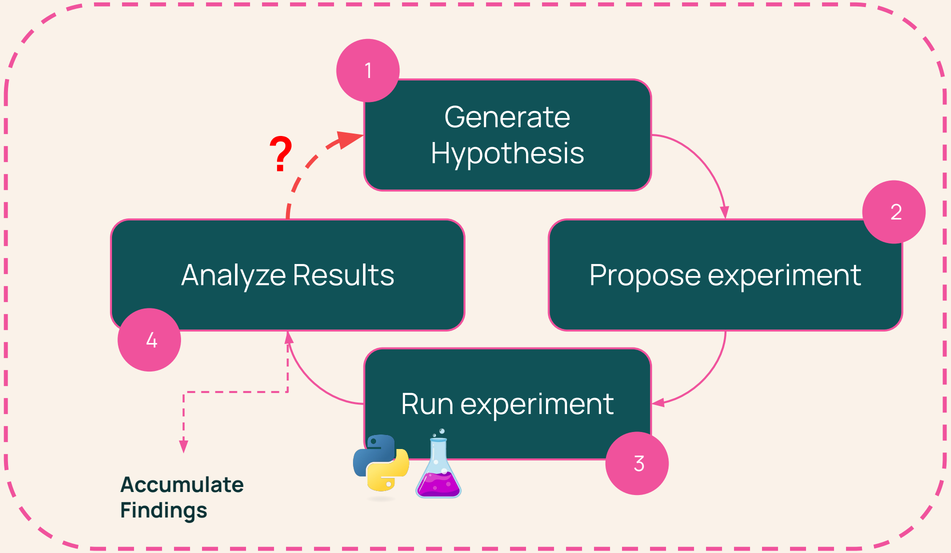
“Open-ended” Discovery

← Degree of supervision



- **Explore** large hypothesis spaces
- Find **novel interconnections**
- Increase **throughput** and reduce **costs**

LLM Agents



Ai2

Main challenges?

1. **What automatic reward can guide scientific discovery?**
   a. Diversity → not enough.
   b. “Interestingness” and “utility” → subjective, unreliable.

2. **How can we repeatedly sample hypotheses using this reward?**
   a. Limited exploration ability in LLMs.
   b. Need an explicit **outer loop for search**.
   → What **search algorithm** should be used?

Ai2

Main challenges?

1. **What automatic reward can guide scientific discovery?**
   a. Diversity → not enough.
   b. "Interestingness" and "utility" → subjective, unreliable.
2. **How can we repeatedly sample hypotheses using this reward?**
   a. Limited exploration ability in LLMs.
   b. Need an explicit **outer loop for search**.
   → What search algorithm should be used?

Surprisal correlates with scientific impact

nature communications

Article

https://doi.org/10.1038/s41467-023-36741-4

Surprising combinations of research contents and contexts are related to impact and emerge with scientific outsiders from distant disciplines

Received: 15 March 2021

Accepted: 15 February 2023

Published online: 24 March 2023

Check for updates

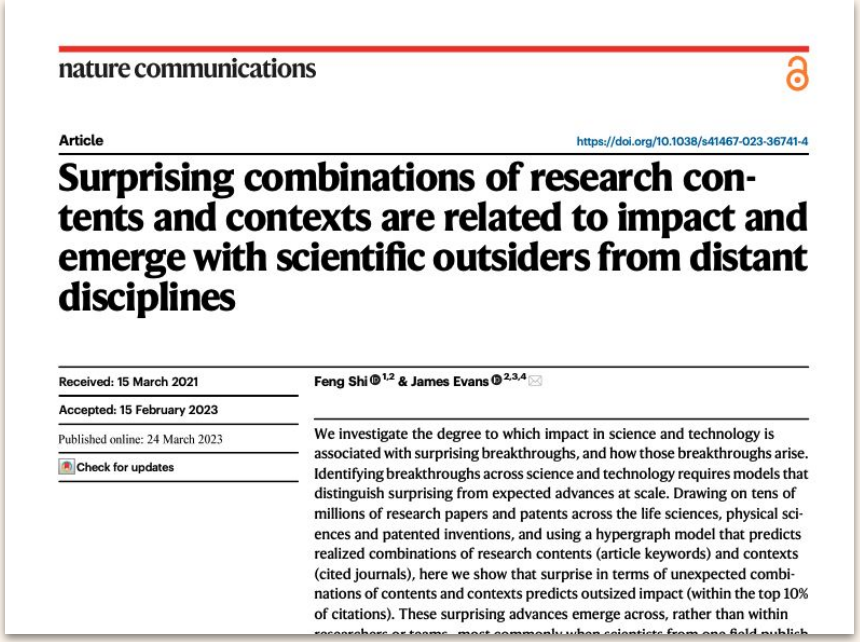
Feng Shi[1,2] & James Evans[2,3,4]

We investigate the degree to which impact in science and technology is associated with surprising breakthroughs, and how those breakthroughs arise. Identifying breakthroughs across science and technology requires models that distinguish surprising from expected advances at scale. Drawing on tens of millions of research papers and patents across the life sciences, physical sciences and patented inventions, and using a hypergraph model that predicts realized combinations of research contents (article keywords) and contexts (cited journals), here we show that surprise in terms of unexpected combinations of contents and contexts predicts outsized impact (within the top 10% of citations). These surprising advances emerge across, rather than within

Shi and Evans, Nature 2023:

The improbability or **surprisal** of a hypothesis is a strong predictor of scientific impact.

Surprisal correlates with scientific impact


nature communications

Article https://doi.org/10.1038/s41467-023-36741-4

Surprising combinations of research contents and contexts are related to impact and emerge with scientific outsiders from distant disciplines

Received: 15 March 2021

Accepted: 15 February 2023

Published online: 24 March 2023

Check for updates

Feng Shi[1,2] & James Evans[2,3,4]

We investigate the degree to which impact in science and technology is associated with surprising breakthroughs, and how those breakthroughs arise. Identifying breakthroughs across science and technology requires models that distinguish surprising from expected advances at scale. Drawing on tens of millions of research papers and patents across the life sciences, physical sciences and patented inventions, and using a hypergraph model that predicts realized combinations of research contents (article keywords) and contexts (cited journals), here we show that surprise in terms of unexpected combinations of contents and contexts predicts outsized impact (within the top 10% of citations). These surprising advances emerge across, rather than within

Shi and Evans, Nature 2023:

The improbability or **surprisal** of a hypothesis is a strong predictor of scientific impact.

Can we mechanize this as an automatic reward?

Our Approach: Bayesian Surprise!

$$\boxed{D_{KL}(\texttt{posterior} \,||\, \texttt{prior})}$$

After experimental evidence

Before experimental evidence

Agent's belief about a hypothesis H:
Distribution (capturing uncertainty) over probabilities that a given hypothesis is true.

$$\theta_H \in [0, 1]; \theta_H \sim \mathrm{Beta}(\alpha, \beta)$$

Bayesian Surprise Attracts Human Attention (Itti and Baldi, NeurIPS 2005)

Our Approach: Bayesian Surprise!

$$D_{KL}(\texttt{posterior} \,||\, \texttt{prior})$$

After experimental evidence

Before experimental evidence

LLM as the Bayesian observer to get an automatic metric → **Expanding the LLM's knowledge frontier**

Agent's belief about a hypothesis H:
Distribution (capturing uncertainty) over probabilities that a given hypothesis is true.

$$\theta_H \in [0,1]; \theta_H \sim \mathrm{Beta}(\alpha, \beta)$$

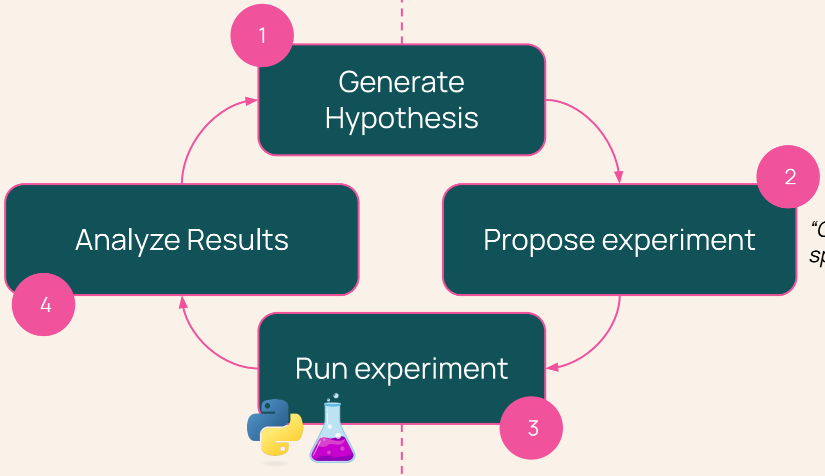
Bayesian Surprise Attracts Human Attention (Itti and Baldi, NeurIPS 2005)

Our Approach: Bayesian Surprise!

Let $\theta \sim \mathrm{Beta}(1,1)$
(Uninformed prior)

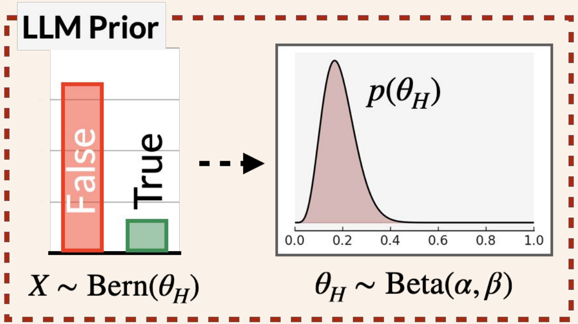
H

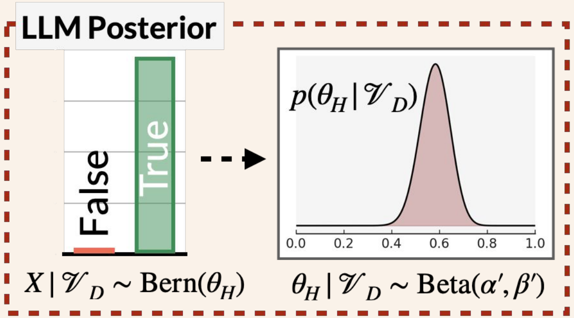
“Geographically proximate basins have more similar evolutionary rates than distant ones”



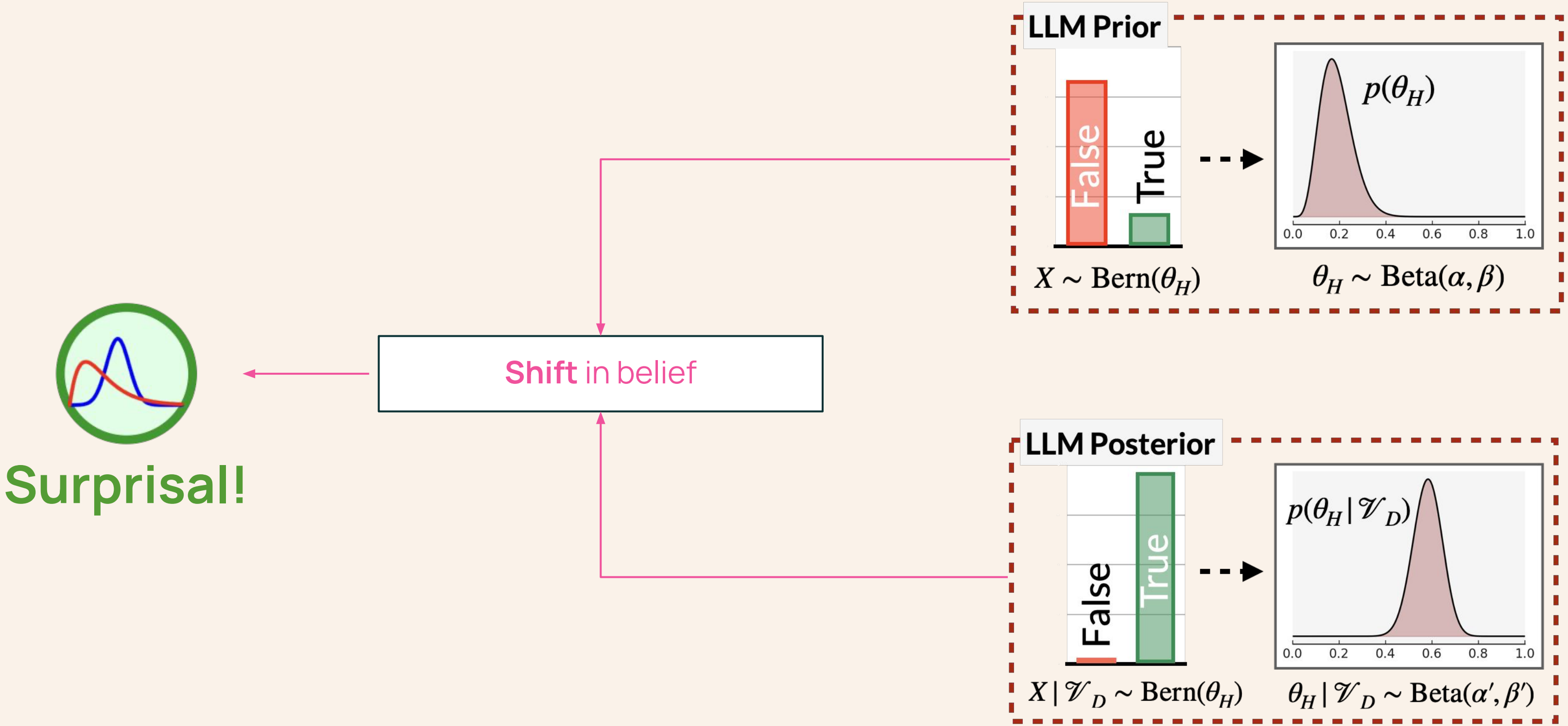
“Compute global and local spatial autocorrelation…”

H + **Experiment Results**





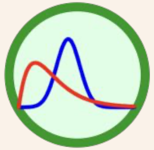
Our Approach: Bayesian Surprise!



Our Approach: Bayesian Surprise!

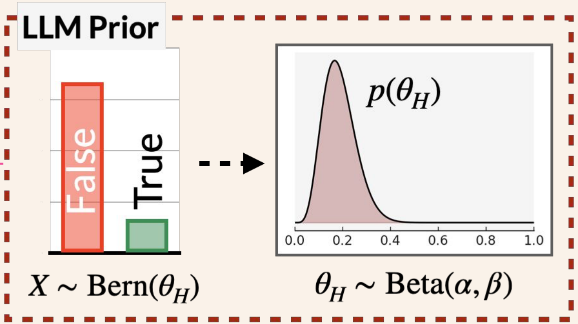
Familiarity	Bayesian Surprise
Low	0.64
Medium	0.66
High	0.62
Overall	0.65

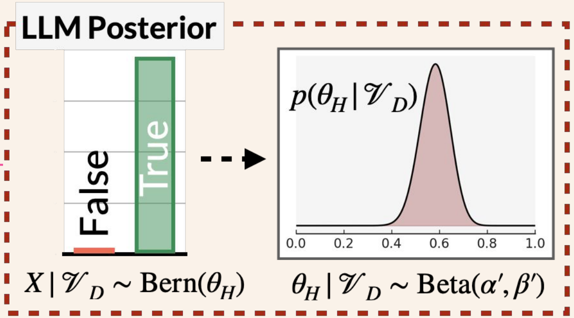
↑ Correlation w/ <u>human</u> surprisal



Surprisal!

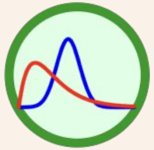
Shift in belief





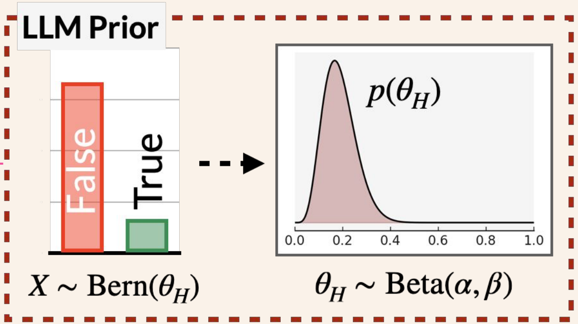
Our Approach: Bayesian Surprise!

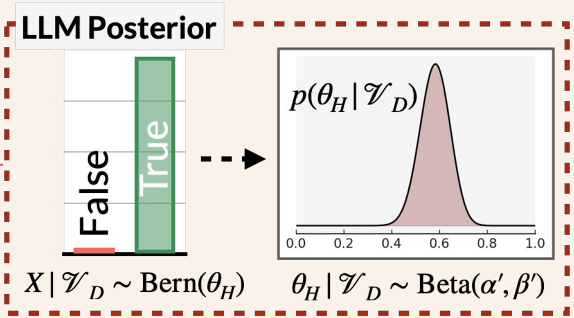
Familiarity	Bayesian Surprise	LLM Surprisal
Low	**0.64**	0.13
Medium	**0.66**	0.12
High	**0.62**	0.08
Overall	**0.65**	0.11



Surprisal!

Shift in belief





Main challenges?

1. **What automatic reward can guide scientific discovery?**
   a. Diversity → not enough.
   b. "Interestingness" and "utility" → subjective, unreliable.

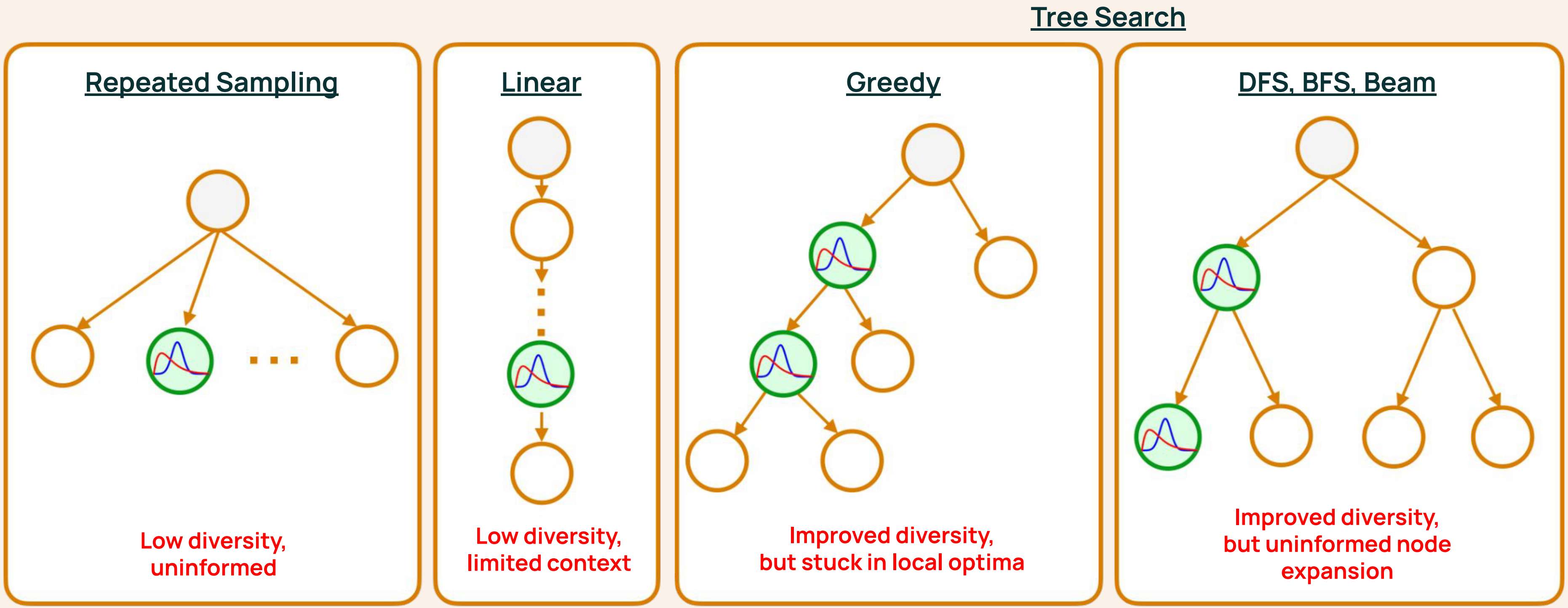
2. **How can we repeatedly sample hypotheses using this reward?**
   a. Limited exploration ability in LLMs.
   b. Need an explicit **outer loop for search**.
      → What search algorithm should be used?

Main challenges?

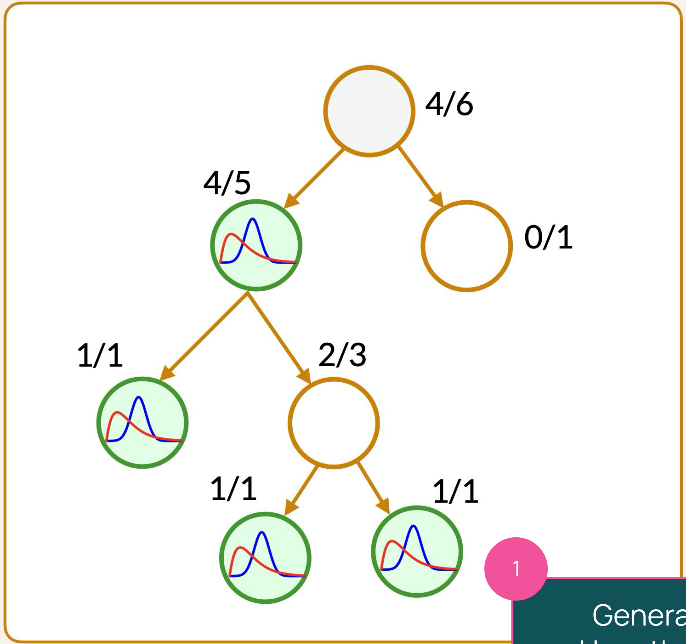
1. **What automatic reward can guide scientific discovery?**
   a. Diversity → not enough.
   b. “Interestingness” and “utility” → subjective, unreliable.

2. **How can we repeatedly sample hypotheses using this reward?**
   a. Limited exploration ability in LLMs.
   b. Need an explicit **outer loop for search**.
      → What **search algorithm** should be used?

Search algorithms for discovery



Our Approach:
MCTS with progressive widening



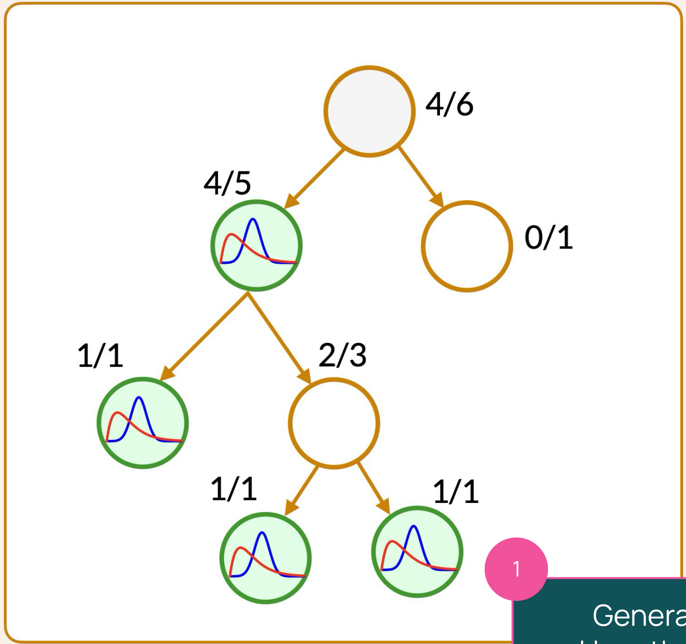
1

Generate Hypothesis

MCTS reward: **Bayesian Surprise**

Claim: Nodes in the search tree with high surprisal counts are likely to elicit surprisal on further expansion.

Our Approach:
MCTS with progressive widening



1

Generate Hypothesis

<u>MCTS reward</u>: **Bayesian Surprise**

Claim: Nodes in the search tree with high surprisal counts are likely to elicit surprisal on further expansion.

*In practice, we use a **UCT-based selection policy** to balance exploration and exploitation.*

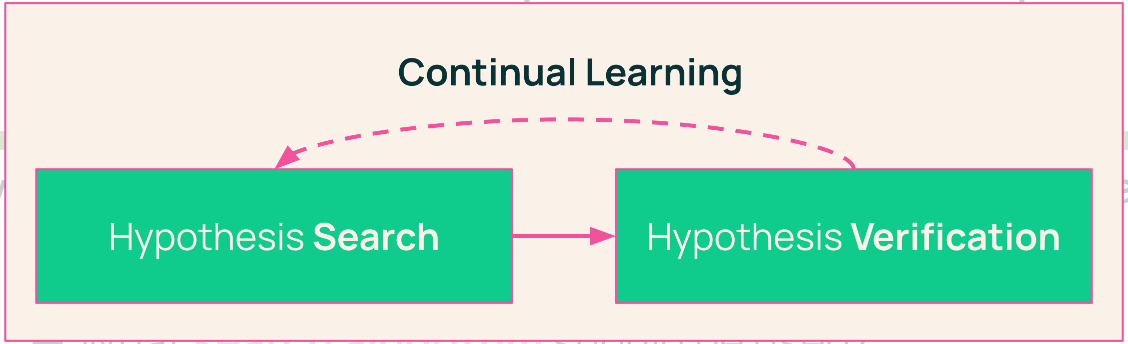
Main challenges?

1. **What automatic reward can guide scientific discovery?**
   a. Diversity → not enough.
   b. “Interestingness” and “utility” → subjective, unreliable.
2. **How can we repeatedly sample hypotheses using this reward?**
   a. Limited exploration ability in LLMs.
   b. Need an explicit **outer loop for search**.
      → What **search algorithm** should be used?

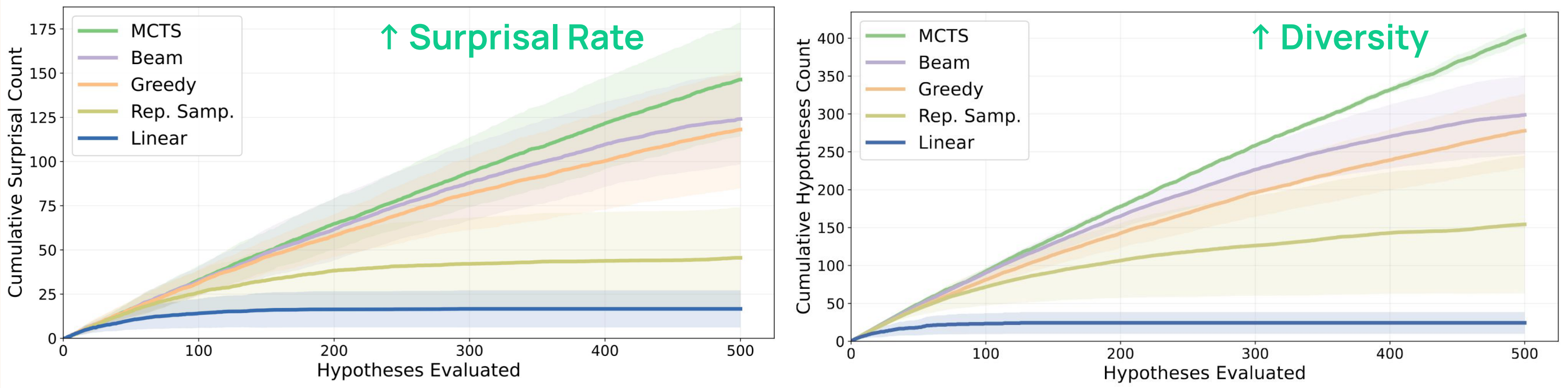
Main challenges?

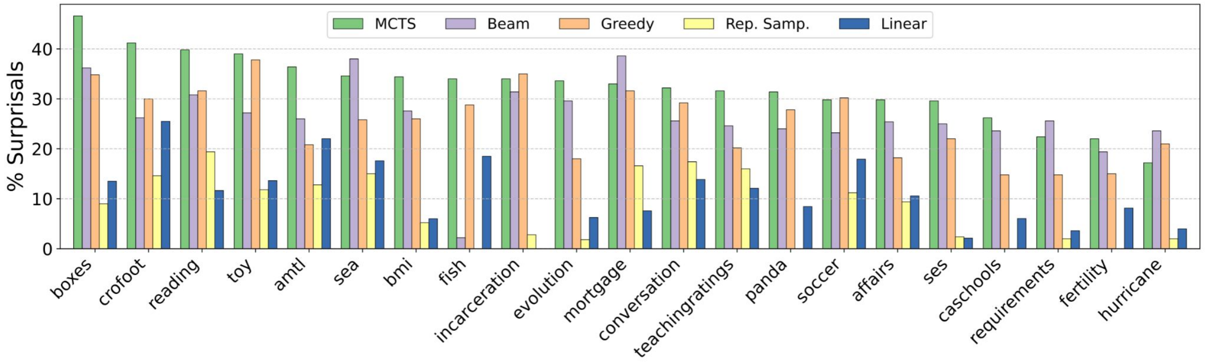
1. What automatic reward can guide scientific discovery?
   a.
   b.

2. How ... ard?
   a.
   b.
   → What search algorithm should be used?



Results from 21 research domains





Robust Across Domains

Ai2

AutoDiscovery via LLM Surprisal

NEURAL INFORMATION PROCESSING SYSTEMS

Check out our paper for more experiments and analyses!

Repo: github.com/allenai/autods

AUTODISCOVERY: Open-ended Scientific Discovery via Bayesian Surprise

Dhruv Agarwal[*α] Bodhisattwa Prasad Majumder[*β]

Reece Adamson[*α] Megha Chakravorty[*α] Satvika Reddy Gavireddy[*α]
Aditya Parashar[γ] Harshit Surana[β] Bhavana Dalvi Mishra[β]

Andrew McCallum[α] Ashish Sabharwal[β] Peter Clark[β]

[α]University of Massachusetts Amherst [β]Allen Institute for AI [γ]Capital One
[*]equal contributions

dagarwal@cs.umass.edu, bodhisattwam@allenai.org
https://github.com/allenai/autods

AutoDiscovery via LLM Surprise

NEURAL INFORMATION PROCESSING SYSTEMS

Check out our paper for more experiments and analyses!

Repo:
github.com/allenai/autods

AUTODISCOVERY:
Open-ended Scientific Discovery via Bayesian Surprise

Dhruv Agarwal[*α] **Bodhisattwa Prasad Majumder**[*β]

Reece Adamson[*α] **Megha Chakravorty**[*α] **Satvika Reddy Gavireddy**[*α]
Aditya Parashar[γ] **Harshit Surana**[β] **Bhavana Dalvi Mishra**[β]

Andrew McCallum[α] **Ashish Sabharwal**[β] **Peter Clark**[β]

[α]University of Massachusetts Amherst [β]Allen Institute for AI [γ]Capital One
[*]equal contributions

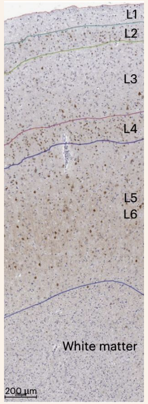
dagarwal@cs.umass.edu, bodhisattwam@allenai.org
https://github.com/allenai/autods

Next:
Impact so far...

Kyle Travaglini
Scientist, Allen Institute

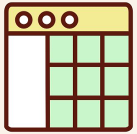
Each QND record encodes:

- What (biomarker) × Where (grey matter / cortical layers) × How (area %, count, density, size, co-expression, etc.).
- APOE genotype of the donor, a genetic risk factor for AD.
- Cognitive Status
- Thal, Braak
- CERAD score



for illustration purposes

Quantitative Neuropathology Data



AutoDiscovery
500 hypotheses, ~20-30 hours

- Select top-10 surprising hypotheses acc. to the degree of surprisal, dir. of belief change
- Evaluate surprisingness with Kyle, along with implementation details and analysis
- Follow-up (e.g., he wanted to only focus on Alzheimer's, and not dementia)

Kyle Travaglini
Scientist, Allen Institute

QND:

I am surprised by the ratio. It is really relevant with the major hypotheses we presented.

AutoDiscovery found:

A **higher cortical laminar gradient of AT8**—quantified as the **ratio** of percent AT8 positive area in **deep** layers (5–6) over **superficial layers** (1–4)—is associated with dementia and **significantly predicts** cognitive status, adjusting for APOE genotype.

Kyle Travaglini
Scientist, Allen Institute

QND:

I am surprised by the ratio. It is really relevant with the major hypotheses we presented.

Mike Jacobi
Ai2 Skylight

Maritime Activity:

The hypothesis column is fascinating. Those are patterns that nobody is talking about right now.

AutoDiscovery found:
Vessels **missing IMO numbers** have significantly **smaller** lengths than vessels with valid IMO entries.

Ai2

Kyle Travaglini
Scientist, Allen Institute

QND:

I am surprised by the ratio. It is really relevant with the major hypotheses we presented.

Social Science/HCI:

It found all the major results from my thesis ch 1. What's even cooler is that it didn't find surprises where there shouldn't be any. Got the null results from my thesis ch 2.

Sanchaita Hazra
PhD student, U of Utah

Mike Jacobi
Ai2 Skylight

Maritime Activity:

The hypothesis column is fascinating. Those are patterns that nobody is talking about right now.

AutoDiscovery found:
Female participants will **switch to AI** advice **less frequently** than male participants in the AI-assisted lie detection task.

Kyle Travaglini
Scientist, Allen Institute

QND:

I am surprised by the ratio. It is really relevant with the major hypotheses we presented.

Social Science/HCI:

It found all the major results from my thesis ch 1. What's even cooler is that it didn't find surprises where there shouldn't be any. Got the null results from my thesis ch 2.

Sanchaita Hazra
PhD student, U of Utah

Mike Jacobi
Ai2 Skylight

Maritime Activity:

The hypothesis column is fascinating. Those are patterns that nobody is talking about right now.

CDC NHANES:

Definitely there is a lot of credibility in the results w.r.t. things like increased blood lead concentration & complex interactions w socioeconomic indicators.

Stephen Salerno
PostDoc, Fred Hutch

AutoDiscovery found:

Higher education is associated with **lower blood lead concentration**, ind. of race & poverty

How you can run AutoDiscovery

IN:

- Datasets
- Associated Metadata
- Your API key

OUT:

A CSV with:

- Hypos sorted by degree of surprise
- For each H, plan and analysis
- Direction of belief change
- Absolute belief values

github.com/allenai/autods

Command:

```
python src/run.py \
    --work_dir="work" \
    --out_dir="outputs" \
    --dataset_metadata="<your_metadata.json>" \
    --n_experiments=100 \
    --model="o4-mini"
    --belief_model="o4-mini"
```

Easy to launch through beaker/gantry as 1-2 day background jobs.
Please reach out if you need help setting up!

How you can run AutoDiscovery

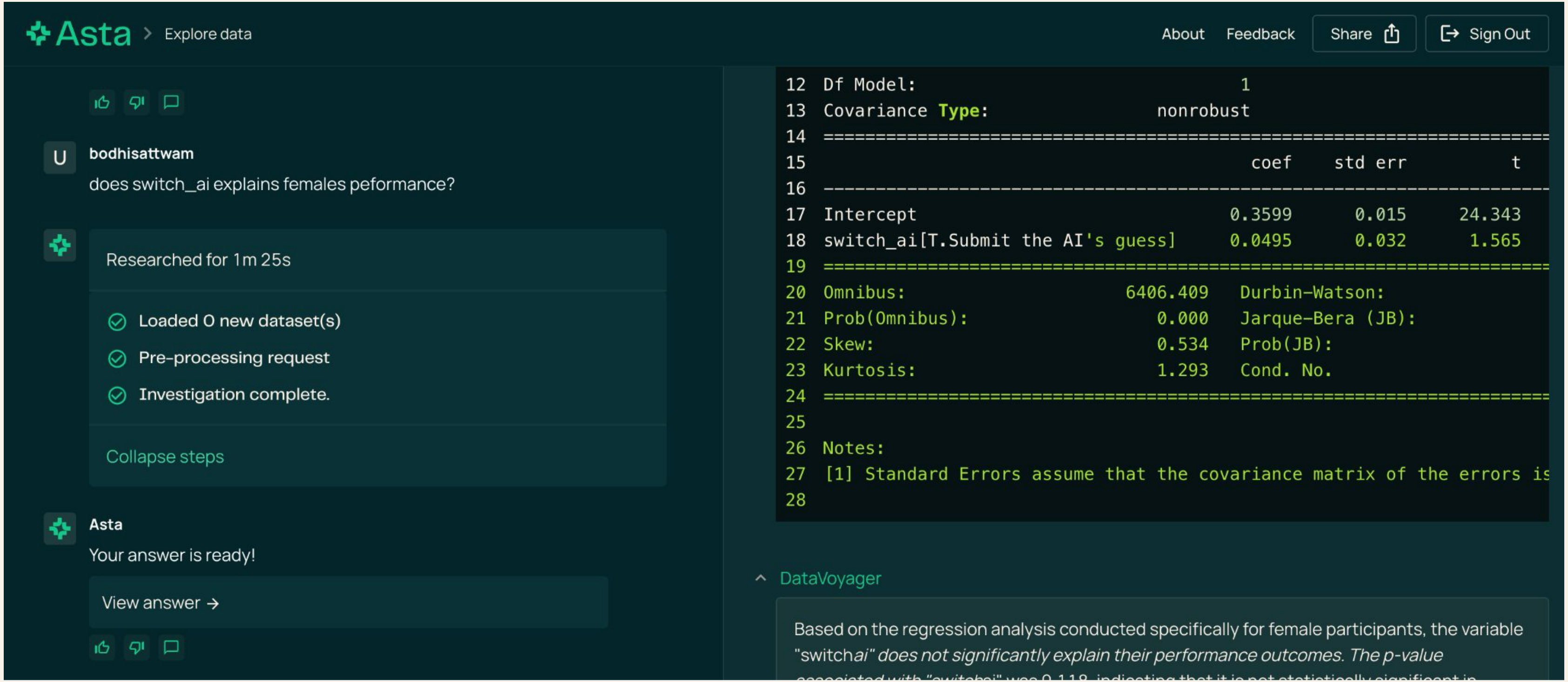
```
{
    "id": 0,
    "domain": "sociology",
    "datasets": [
        {
            "name": "nls_ses_processed.csv",
            "description": "This dataset contains social background factors …",
            "columns": {
                "raw": [
                    ...
                    {
                        "name": "ABILITY: COMPOSITE OF ASVAB SCORE",
                        "description": "Composite variable created by summing ..."
                    },
                    ...
```

AutoDiscovery Results Sheet

hypothesis	surprisal	prior	posterior	belief_dir	belief_change	experiment_plan	analysis
Lower educational attainment is associated with higher blood lead levels, and this association is stronger among participants who received WIC than those who did not.	TRUE	0.4597	0.8049 › 0.6	pos	0.3453 › 0.1	Objective: To evaluate whether WIC participation moderates the association between education and blood lead concentration. Steps: 1. Load NHANES.csv and select LBXBPB_LOD, EDUC_3CAT, WIC_2CAT, AGE, SEX, RACE_2CAT, PIR_3CAT, WTMEC4YR, SDMVPSU, SDMVSTRA. 2. Exclude Refused/Unknown in EDUC_3CAT, WIC_2CAT, PIR_3CAT. 3. Recode EDUC_3CAT ordinally (1=High School,2=Some College,3=College Grad). 4. Construct survey design with weights, cluster, strata. 5. Fit survey-weighted linear regression: LBXBPB_LOD ~ EDU_num * WIC_2CAT + AGE + SEX + RACE_2CAT + PIR_3CAT. 6. Test interaction EDU*WIC (joint Wald test). 7. Compute slopes: effect of education on lead within WIC vs non-WIC. 8. Plot interaction: predicted lead vs education level by WIC status.	The weighted regression successfully estimated the moderating effect of WIC on the education–lead relationship. Key findings: • Significant main effects: Higher education is associated with lower blood lead (coef=–0.2134, SE=0.0472, p<0.001); WIC participants have higher lead overall (coef=0.5479, SE=0.155, p<0.001). • Significant interaction: EDU_num:WIC_bin coef=–0.2912 (SE=0.062, p<0.001), indicating the protective effect of education on lead is stronger among WIC recipients. • Simple slopes: Education slope=–0.2134 (95% CI: –0.3059, –0.1209) for non-WIC vs –0.5047 (95% CI: –0.7056, –0.3037) for WIC. • Interaction plot: Both groups show decreasing lead with more education, but the decline is steeper among WIC recipients. These results support the hypothesis that the inverse association between education and blood lead is stronger in the WIC group.

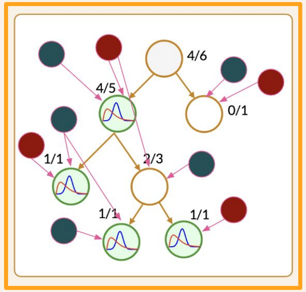
hypothesis	surprisal	prior	posterior	belief_dir	belief_change	experiment_plan	analysis
Lower educational attainment is	TRUE	0.4597	0.8049	pos	0.3453	Objective: To evaluate whether WIC participation	The weighted regression successfully estimated the

AutoDiscovery ⟷ Asta



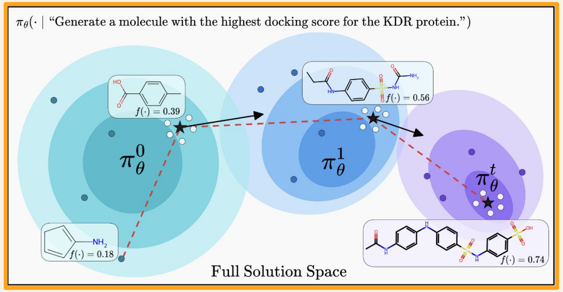
What's next

World Model & Causality in AutoDiscovery



- Online belief update
- User-guided Search
- External validation for causal factors

Test-time adaptation for Hypothesis Generation



Got a hypothesis but no data? Data search!

Introducing the Data Commons Model Context Protocol (MCP) Server: Streamlining Public Data Access for AI Developers

SEPT. 24, 2025

Keyur Shah
Software Engineer

Controlling False Discovery Rate

Online multiple testing with e-values

Ziyu Xu* Aaditya Ramdas†

November 14, 2023

Our Team - Thank you!