Risk & Opportunity in AI4Insurance

보험 산업에서의 AI 리스크 및 기회: 사례 및 제안

Prof. Sungbin Lim (林聖彬)
Department of Statistics, College of Political Science and Economics

2026. 05. 28, 보험연구원

KOREA UNIVERSITY
Department of Statistics
KOREA UNIVERSITY

Part 1

Generated by Google Imagen

Artificial Intelligence

Machine Learning vs Machine Reasoning

- Machine Learning focuses on mining the hidden patterns from data to tackle a pre-determined problem
- Machine Reasoning implements thinking process as a computational system by manipulating acquired knowledge and data to answer a new question

learned models

supervised learning
(e.g. classification, regression) $\{(x_i, y_i)\} \rightarrow f_\theta$

reasoning w/ knowledge
(e.g. meta learning) $\{(\mathcal{D}_\tau, f_\tau)\} \rightarrow F_\Theta$

unsupervised learning
(e.g. density estimation) $\{x_i\} \rightarrow g_\phi$

reasoning w/o knowledge
(e.g. scientific discovery) $\{\mathcal{D}_\tau, \boxed{}\} \rightarrow G_\Phi$

index of data samples

index of tasks

acquired knowledge

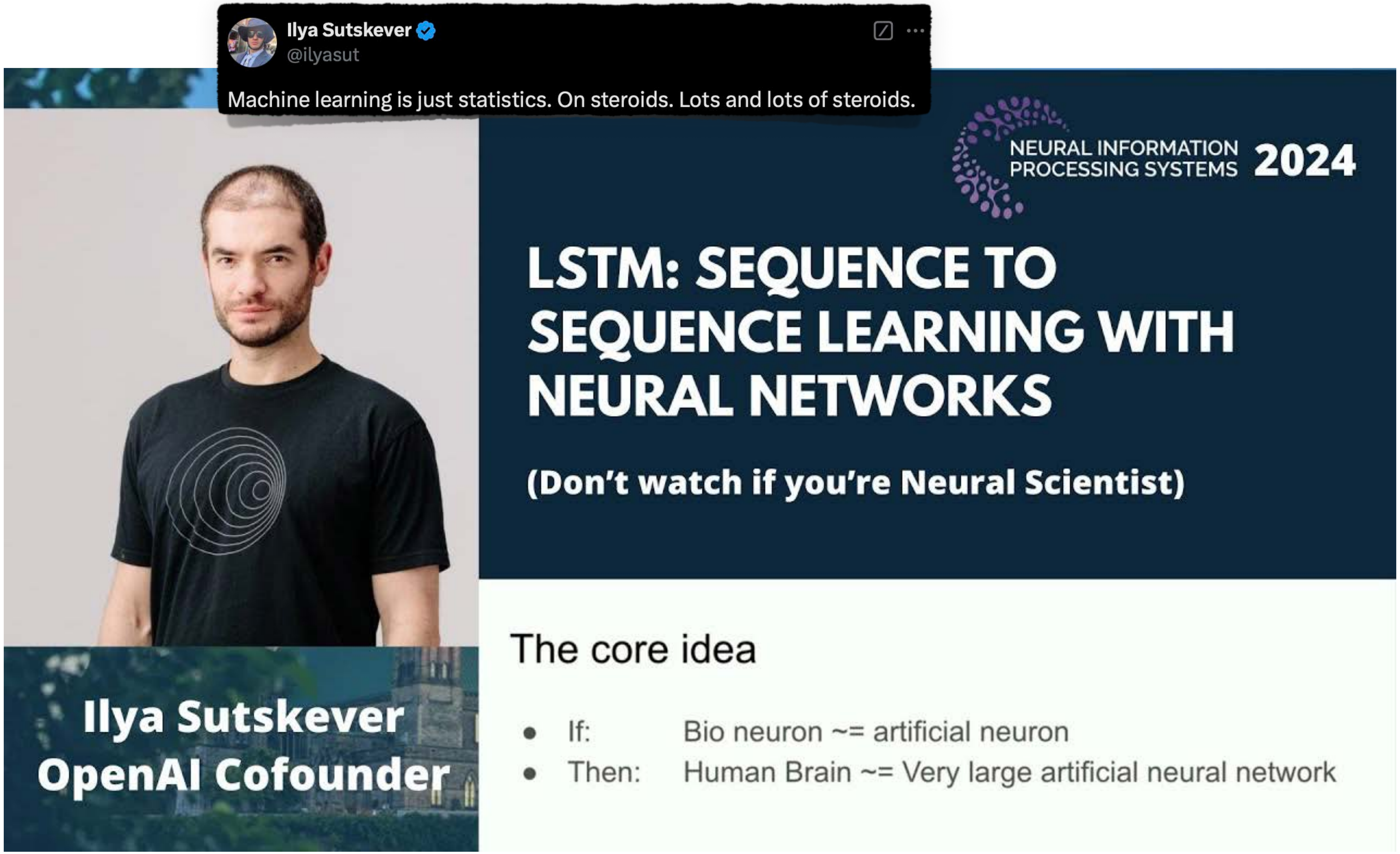
L. Bottou, From Machine Learning to Machine Reasoning, ***Machine Learning*** **(2014)**

When creative machines overtake man

Dr. Jürgen Schmidhuber

"I expect huge RNNs on dedicated hardware to simultaneously perceive and analyze an immense number of multimodal data streams (speech, texts, video, many other modalities) from many sources, learning to correlate all those inputs and use the extracted information to achieve a myriad of commercial and non-commercial goals. Those RNNs will continually and quickly learn new skills on top of those they already know. This should have innumerable applications, although I am not even sure whether the word *application* still makes sense here."

Source: TEDx Talks (2012)



Source: NeurIPS 2024 Test-time Talk

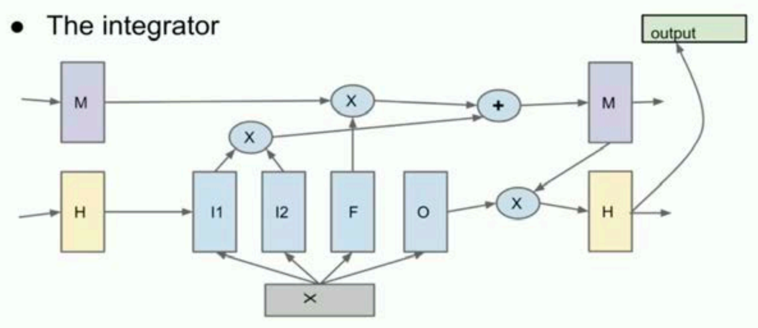
Learning Principle of Future?

What we got right: Deep Learning / Autoregressive Models (Transformer) / Scaling Hypothesis

What we got wrong: the LSTM

The heart of the LSTM

- The integrator



"The Deep Learning Hypothesis"

- Human perception is fast
  - Neurons fire at most 100 times a second
  - Humans solve perception in 0.1 seconds

Pre-training as we know it will end

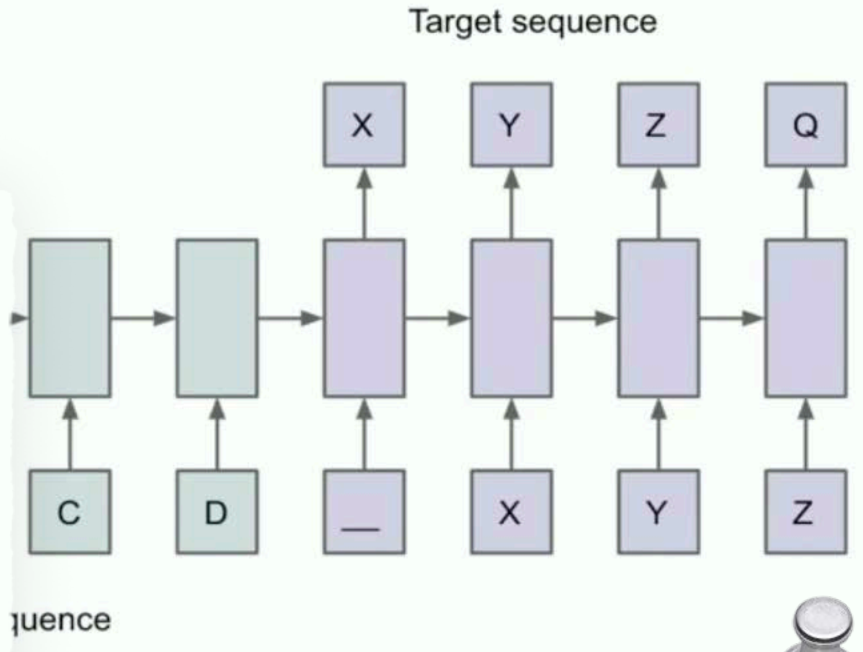
Compute is growing:

- Better hardware
- Better algorithms
- Larger clusters

Data is not growing:

- We have but one internet
- **The fossil fuel of AI**

Our main idea



Conclusions

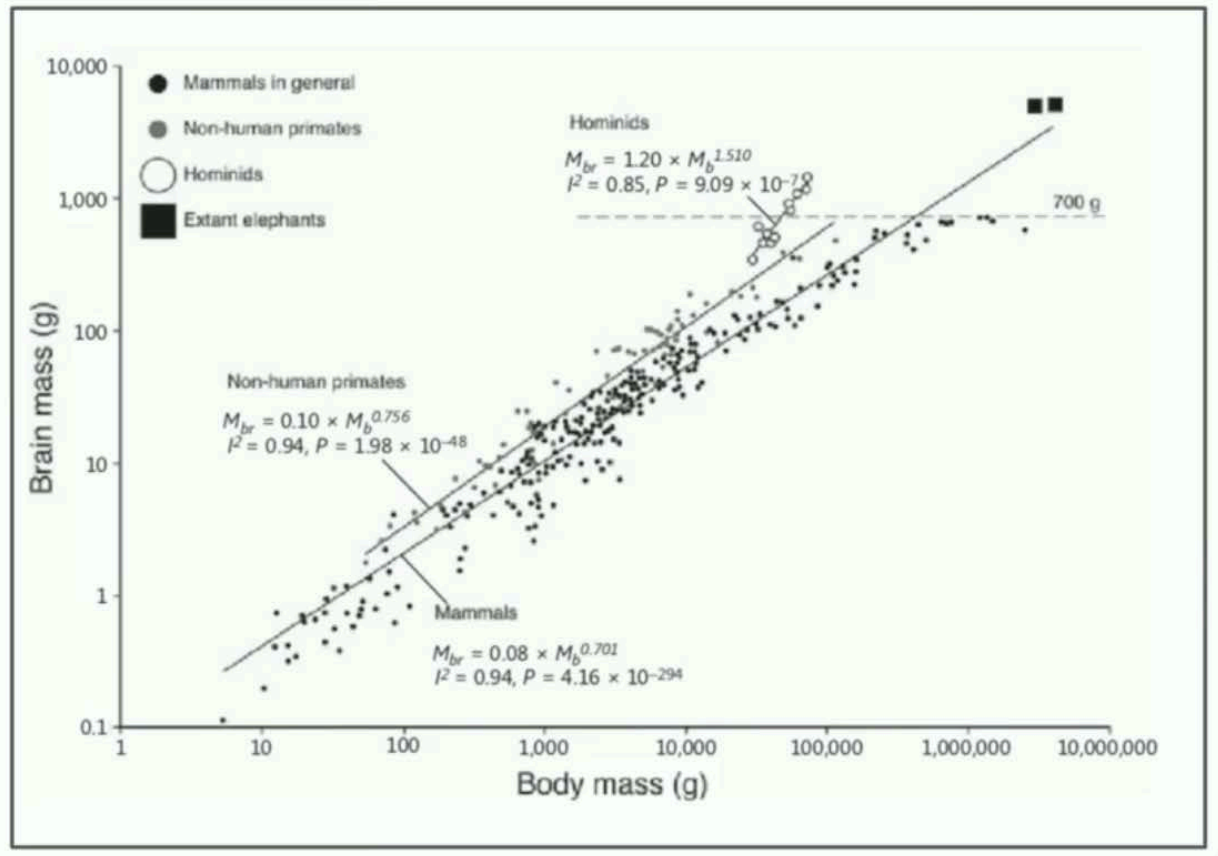
- If you have a large big dataset
- And you train a very big neural network
- Then success is guaranteed!

What comes next? The long term

Superintelligence

- Agentic
- Reasons
- Understands
- Is self aware

What comes next? Example from nature

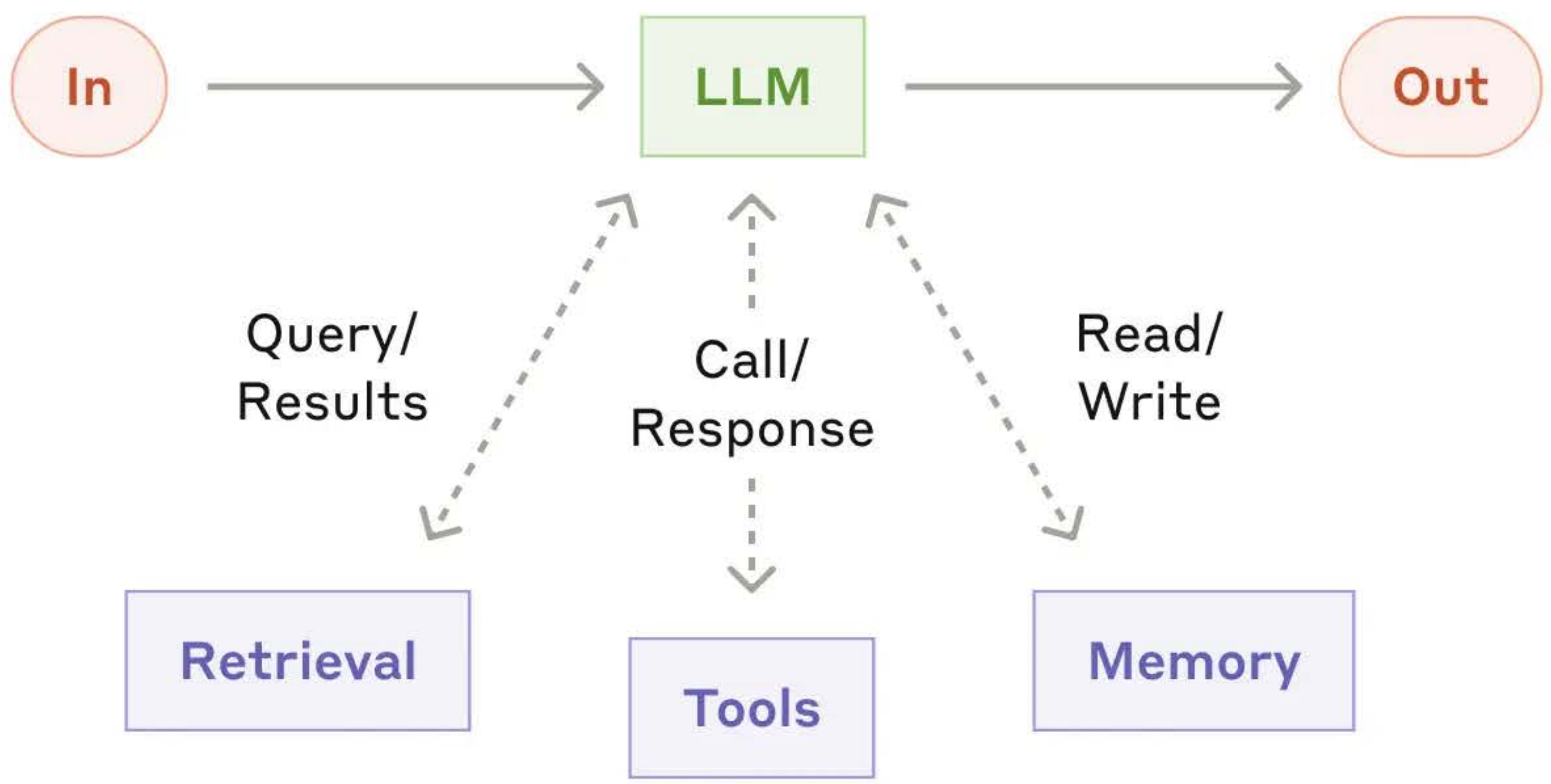


Manger et al., 2013

Ilya Sutskever, Seq2seq Learning with Neural Networks: what a decade, ***NeurIPS 2024 Test-time Talk***

What is Agentic AI?

- Autonomously determines what actions to take, plans multi-step workflows, and adapts based on real-time data.
  - **autonomy** — ability to make decisions, execute tasks, and refine strategies.
  - **adaptability** — learning from feedback, market fluctuations, and new data.
  - **coordination** — interacting with other AI agents and databases.

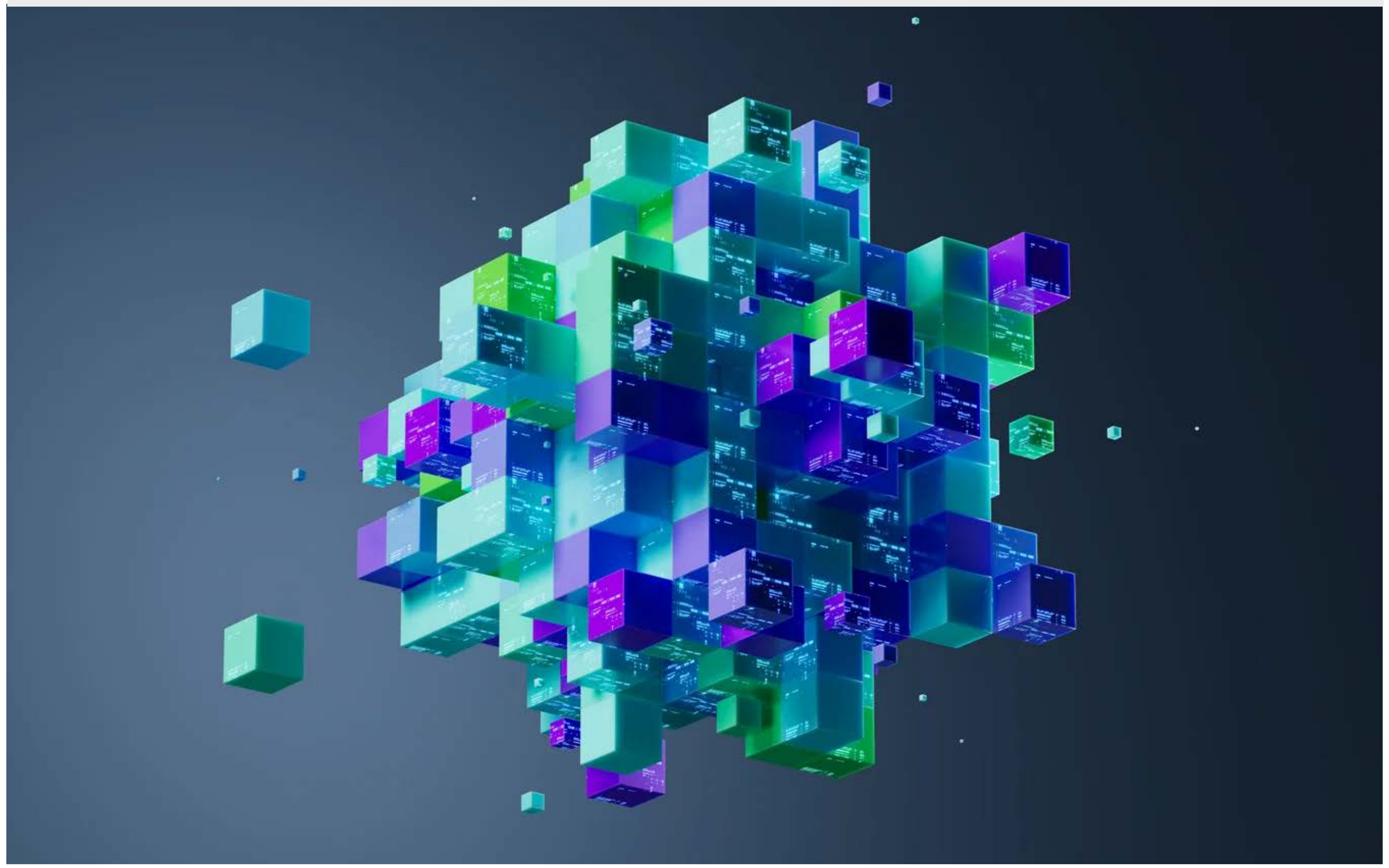


JUNE 8, 2024 | 12 MIN READ

AI Will Become Mathematicians' 'Co-Pilot'

Fields Medalist Terence Tao explains how proof checkers and AI programs are dramatically changing mathematics

BY CHRISTOPH DRÖSSER



RESEARCH

Advanced version of Gemini with Deep Think officially achieves gold-medal standard at the International Mathematical Olympiad

21 JULY 2025

Thang Luong and Edward Lockhart

Source: Christoph Drösser & Google DeepMind

Sebastien Bubeck @SebastienBubeck

Claim: gpt-5-pro can prove new interesting mathematics.

Proof: I took a convex optimization paper with a clean open problem in it and asked gpt-5-pro to work on it. It proved a better bound than what is in the paper, and I checked the proof it's correct.

Details below.

게시물 번역하기

ChatGPT 5 Pro

2503.10138v1.pdf
PDF

Can you improve the condition on the step-size in Theorem 1? I don't want to add any more hypothesis, I just want you to work to improve the step-size condition under the same assumptions as Theorem 1.

Request for GPT-5 Pro
Reasoned for 17m 35s

Yes. Under the *same* assumptions as Theorem 1 (convex L-smooth f and fixed step size for gradient descent), you can enlarge the guaranteed-convexity window from $\eta \le 1/L$ to

$$\boxed{\eta \le \frac{3}{2L}}$$

i.e., the optimization curve $n \mapsto f(x_n)$ is convex for any step size $\eta \in (0, 3/(2L)]$. This strictly improves Theorem 1's bound and still lies below the non-convexity regime exhibited in Theorem 2. See Theorem 1 and Theorem 2 in §3 (pp. 4–6) for the problem setup and statements referenced below.

Ernest Ryu @ErnestRyu · 8월 21일

There are 3 proofs in discussion:
v1. (η ≤ 1/L, discovered by human)
v2. (η ≤ 1.75/L, discovered by human)
v.GTP5 (η ≤ 1.5/L, discovered by AI)

Sebastien argues that the v.GPT5 proof is impressive, even though it is weaker than the v2 proof. (2/9)

1 2 352 3.4만

Ernest Ryu @ErnestRyu · 8월 21일

The proof itself is arguably not very difficult for an expert in convex optimization, if the problem is given.

Knowing that the key inequality to use is [Nesterov Theorem 2.1.5], I could prove v2 in a few hours by searching through the set of relevant combinations. (3/9)

Ernest Ryu @ErnestRyu · 8월 21일

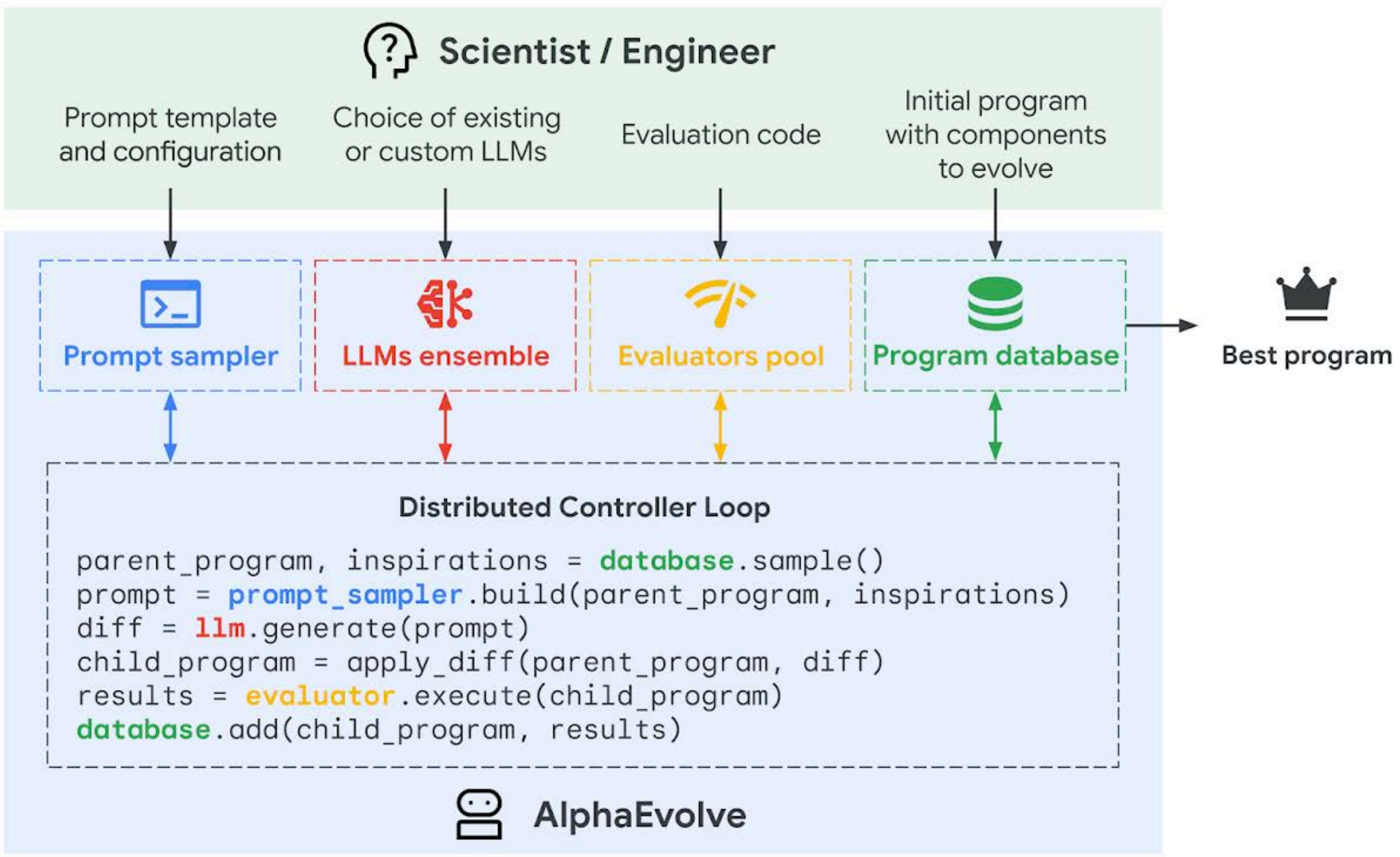
The proof is something an experienced PhD student could work out in a few hours. That GPT-5 can do it with just ~30 sec of human input is impressive and potentially very useful to the right user.

However, GPT5 is by no means exceeding the capabilities of human experts. (9/9)

25 96 1.5천 5.8만

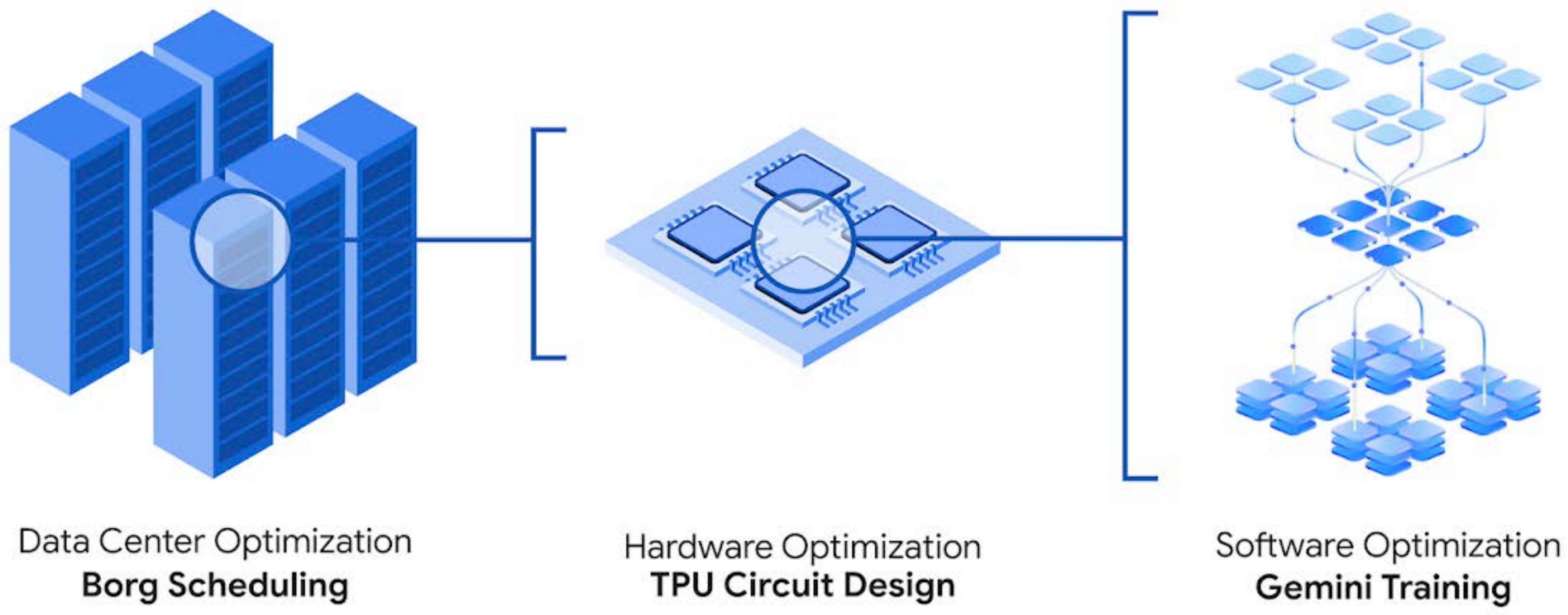
Source: S. Bubeck & E. Ryu (2025)

Human defines **"what?"**



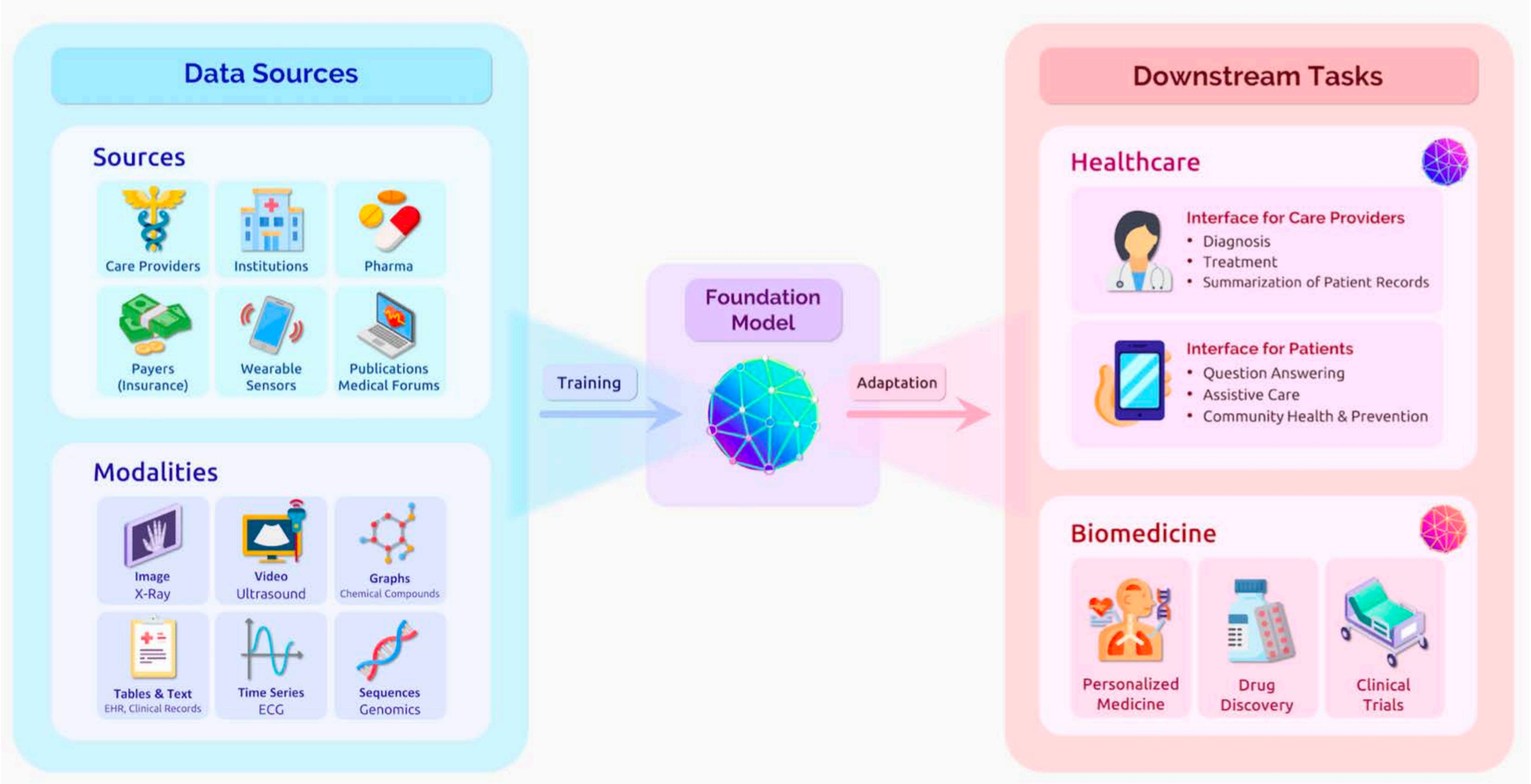
AlphaEvolve figures out **"how?"**

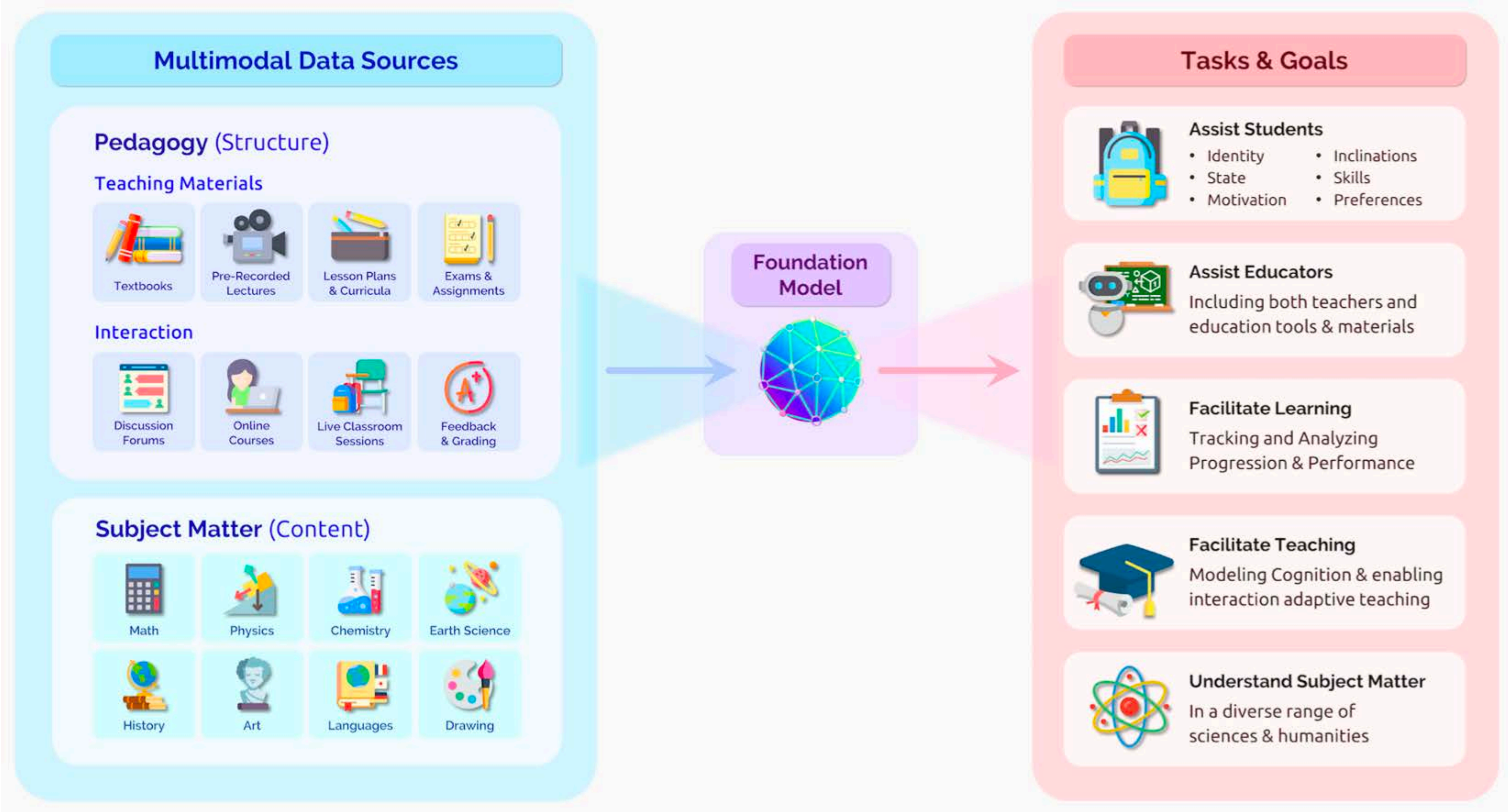
- **Analysis**
  - **Autocorrelation inequalities.** *AlphaEvolve* was able to improve the best known bounds on several autocorrelation inequalities.
  - **Uncertainty principles.** *AlphaEvolve* was able to produce a refined configuration for a problem arising in Fourier analysis, by polishing an uncertainty principle construction [33] leading to a slightly better upper bound.
- **Combinatorics and number theory**
  - **Erdős's minimum overlap problem.** *AlphaEvolve* established a new upper bound for the minimum overlap problem [25], slightly improving upon the previous record [40].
- **Geometry and packing**
  - **Kissing number problem.** In 11 dimensions, *AlphaEvolve* improved the lower bound on the kissing number, finding a configuration of 593 non-overlapping unit spheres that can simultaneously touch a central unit sphere, surpassing the previous record of 592 [31].
  - **Packing problems.** *AlphaEvolve* achieved several new results in packing problems, such as packing N points in a shape to minimize the ratio of the maximum and minimum distance, packing various polygons in other polygons in the most efficient way, and variants of the Heilbronn problem concerning point sets avoiding small-area triangles [29].



Source: Google DeepMind (2025)

AI for Service Industries



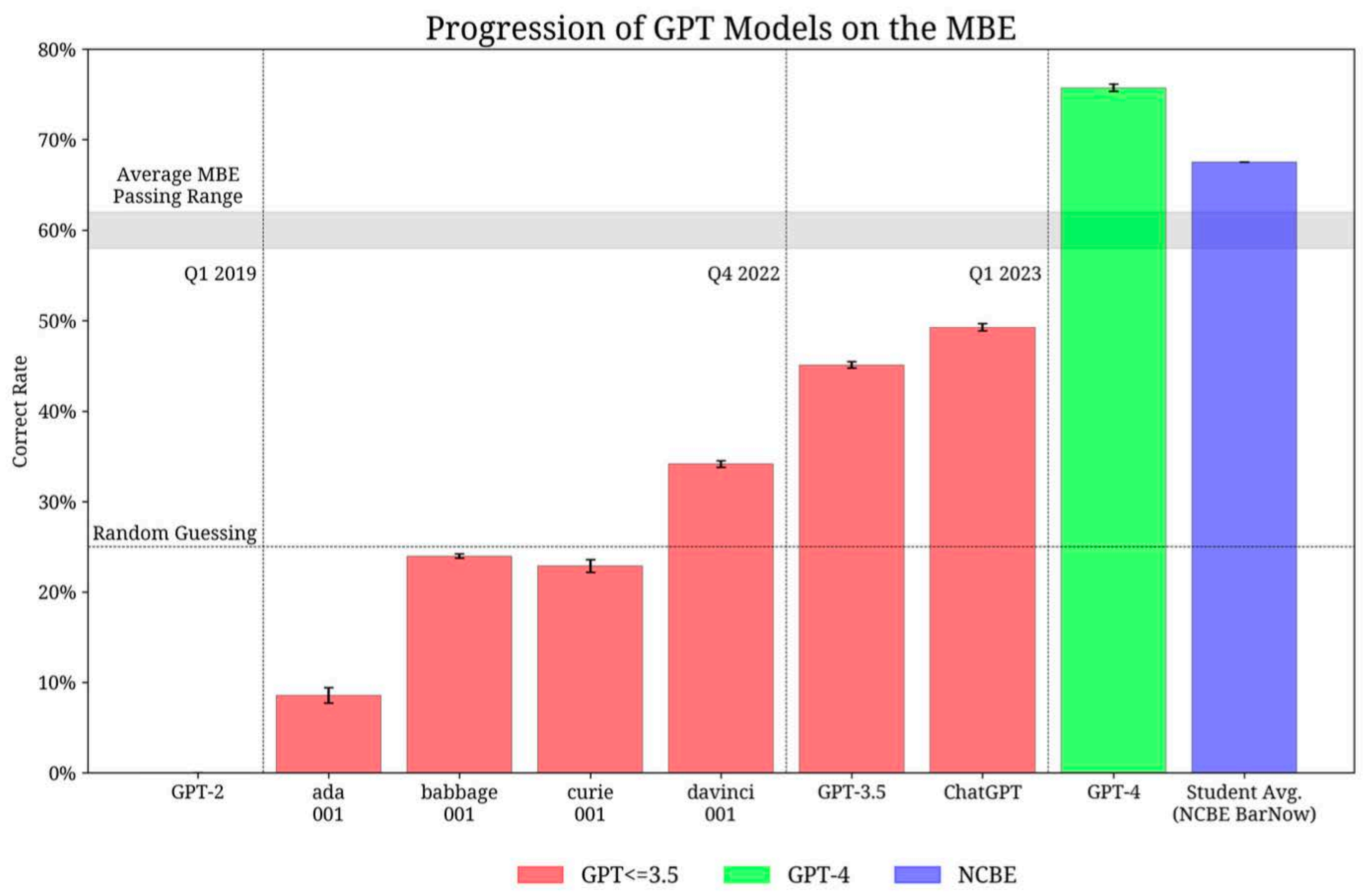


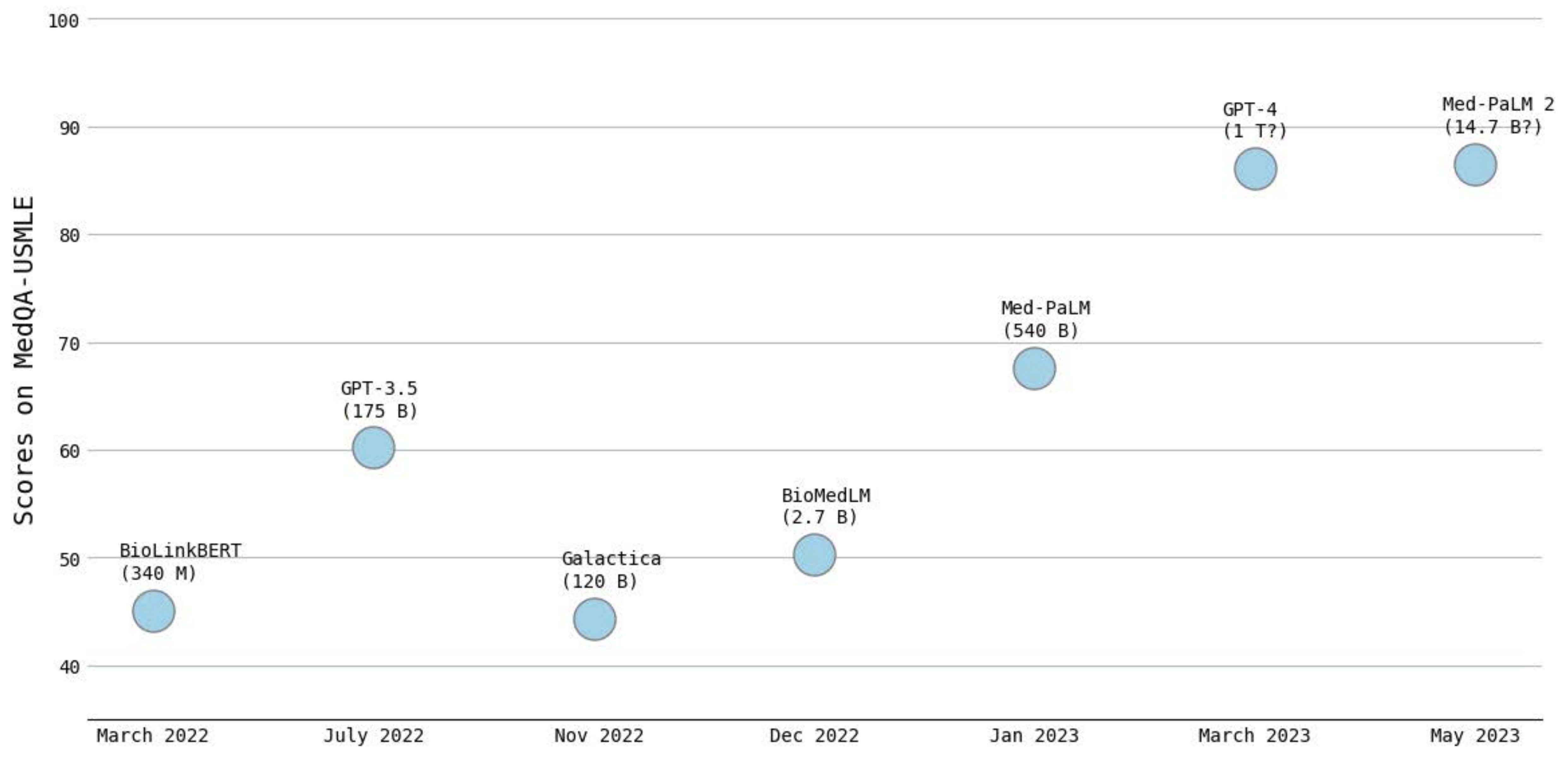
Healthcare

Education

Source: Stanford & OpenAI

AI for Law & Medical Services

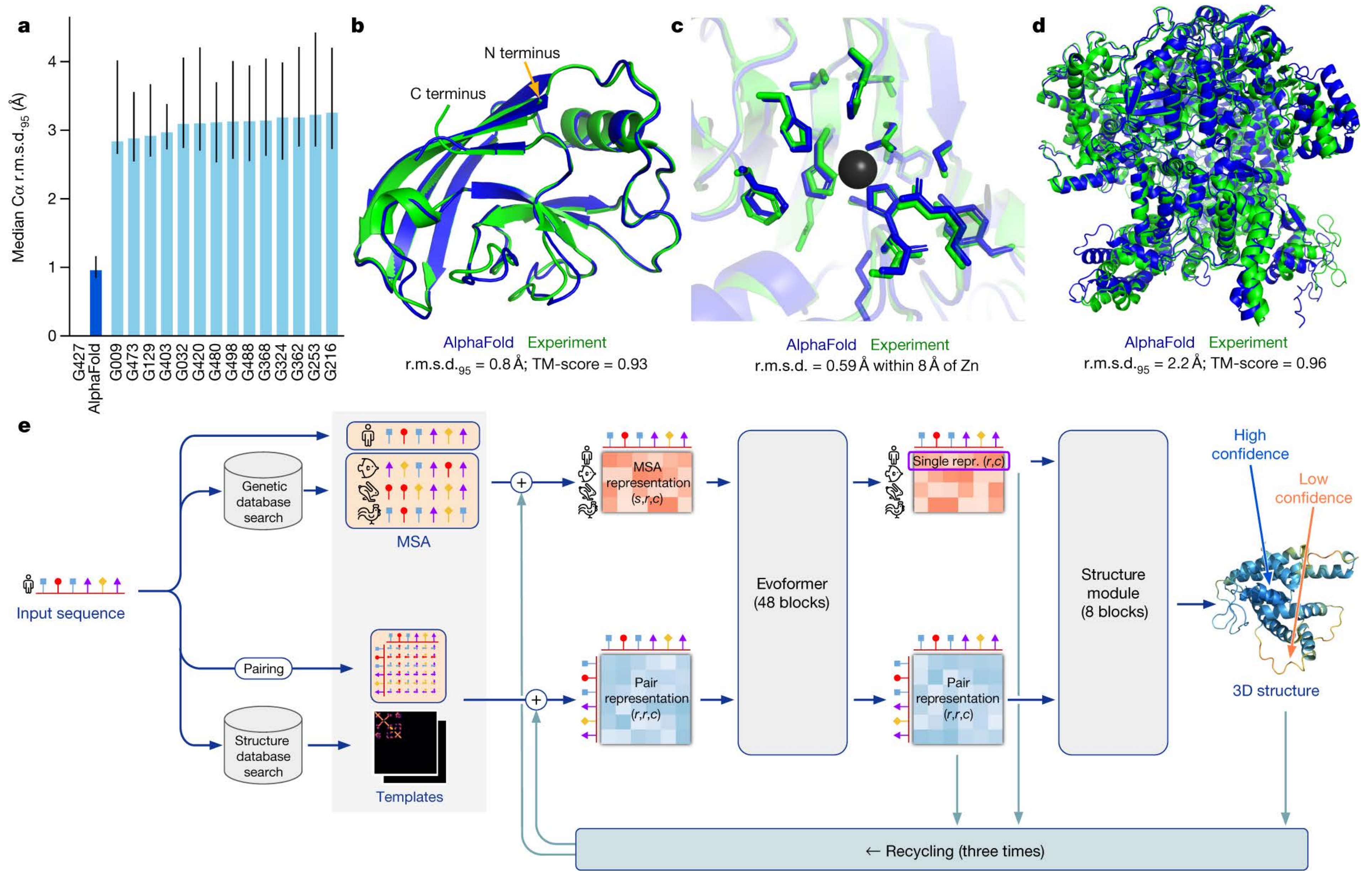




Source: Katz et al. & Yennie Jun

AI for Natural Science

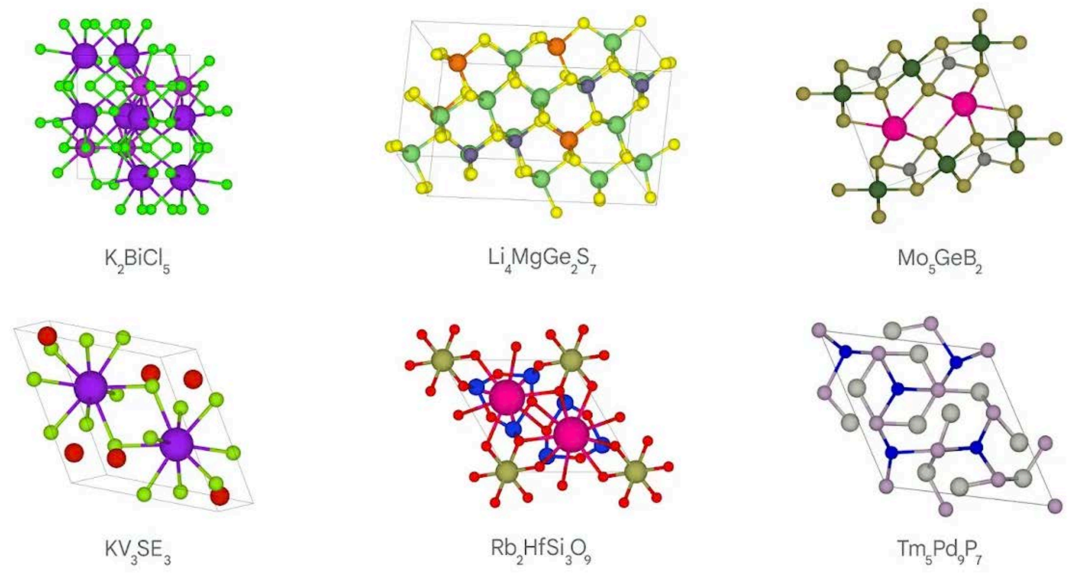
Google DeepMind

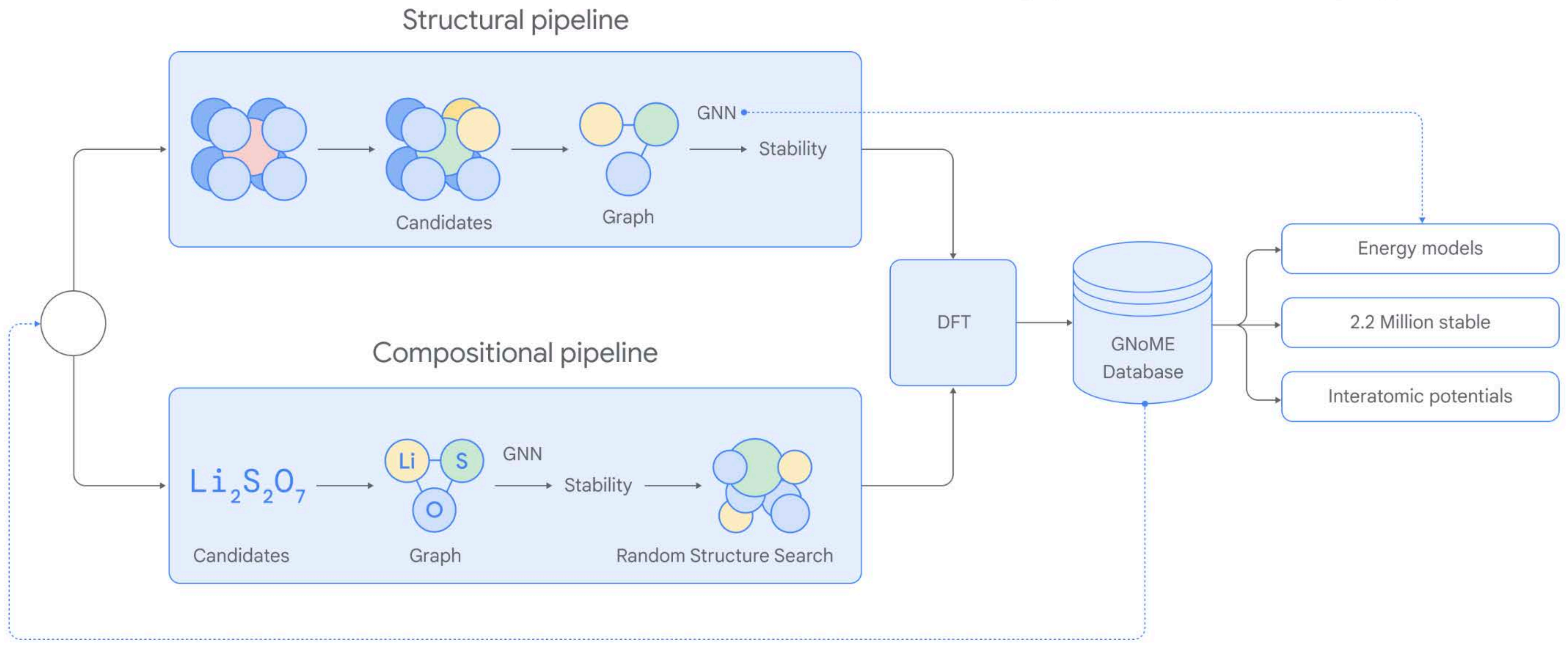


Millions of new materials discovered with deep learning

29 NOVEMBER 2023

Amil Merchant and Ekin Dogus Cubuk





AlphaFold

GNoME

Jumper et al., Highly accurate protein structure prediction with AlphaFold, ***Nature*, (2021)**

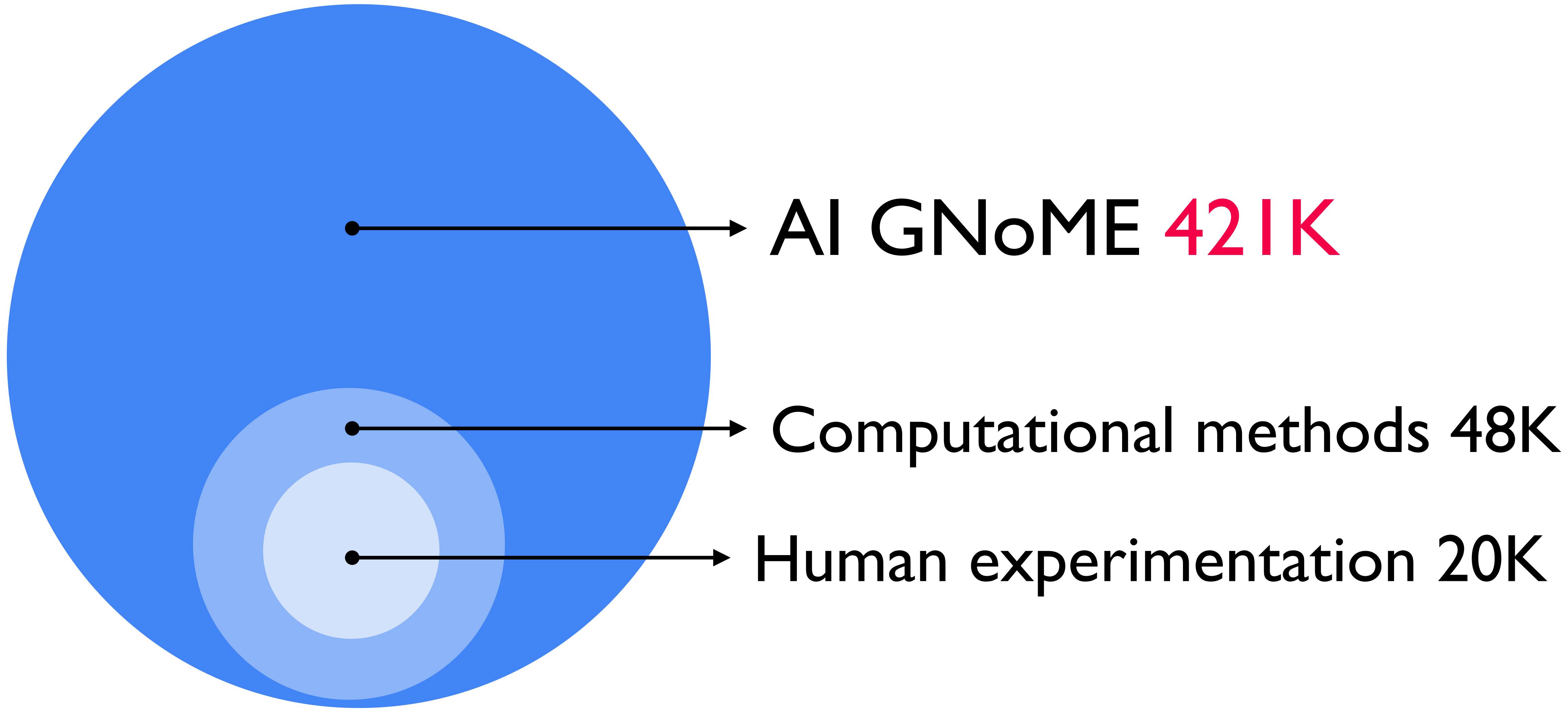
Merchant et al., Scaling deep learning for materials discovery, ***Nature*, (2023)**

Millions of new materials discovered with deep learning

29 NOVEMBER 2023

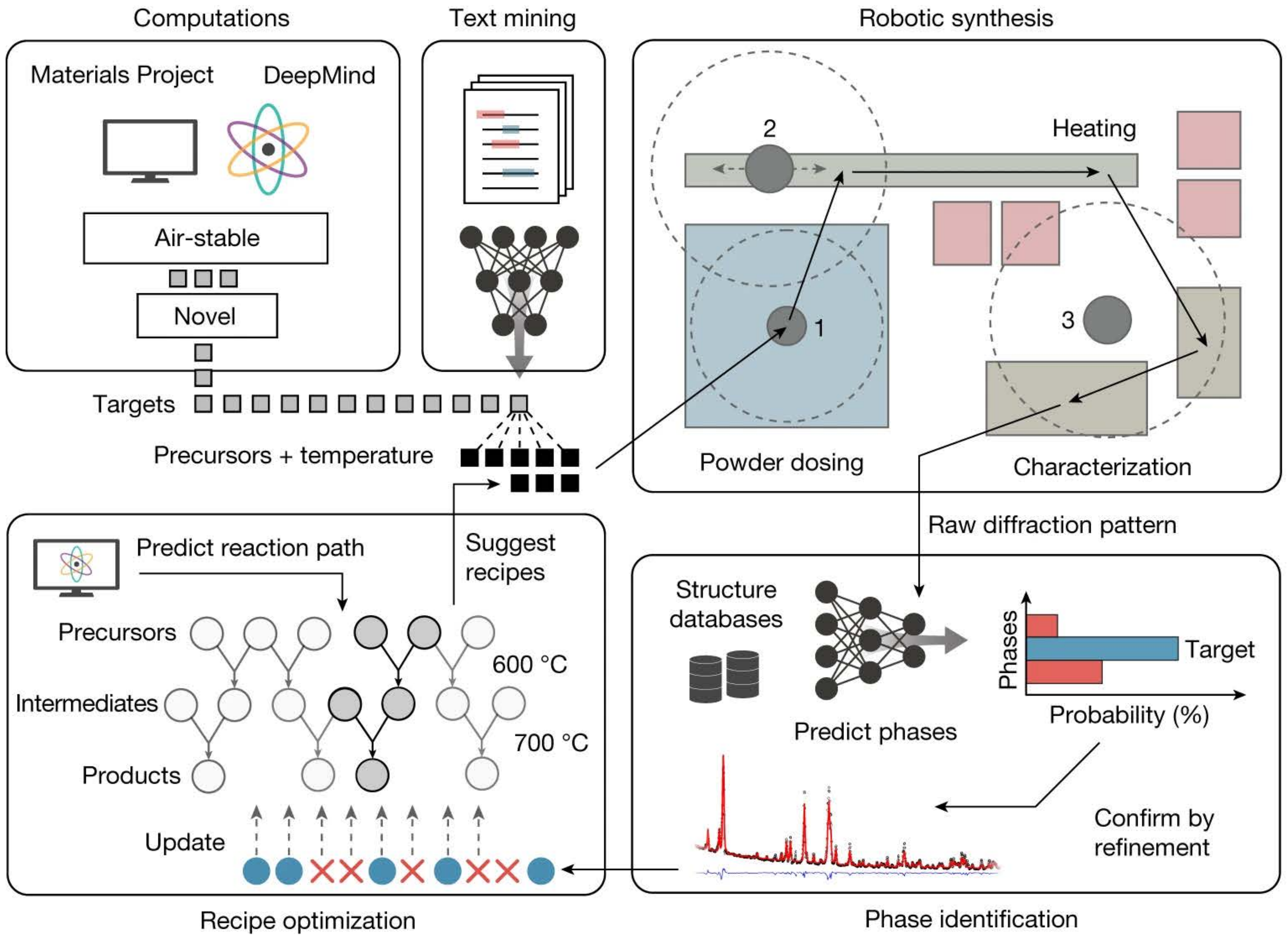
Amil Merchant and Ekin Dogus Cubuk

Share



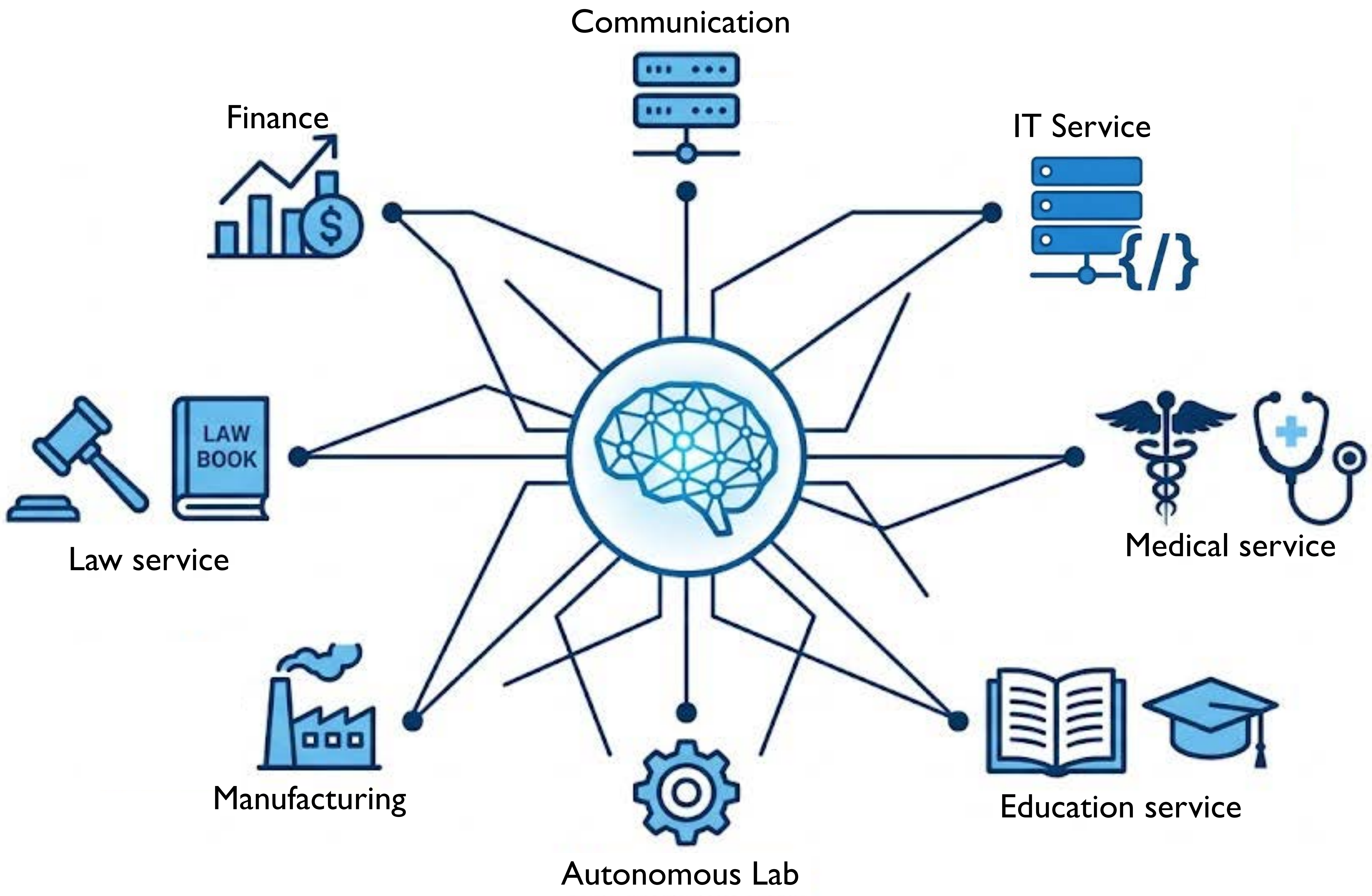
Source: Google DeepMind & Berkeley Lab (2023)

AI-augmented Autonomous Lab Operation



Szymanski et al., An autonomous laboratory for the accelerated synthesis of novel materials, ***Nature*, (2023)**

Boiko et al., Autonomous chemical research with large language models, ***Nature*, (2022)**


Communication
Finance
IT Service
LAW
BOOK
Law service
Medical service
Manufacturing
Education service
Autonomous Lab

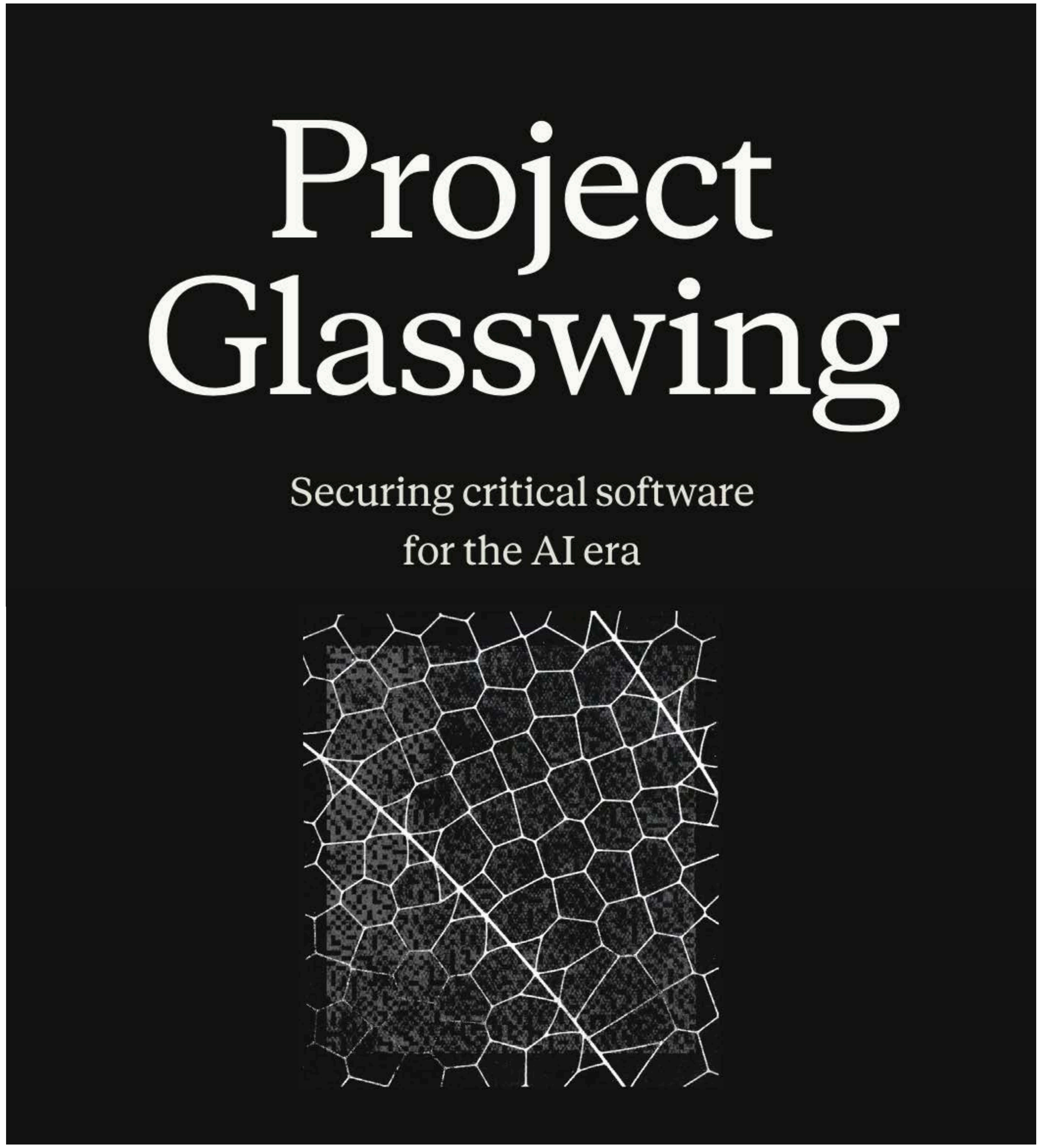
Part 2

Risk & Opportunity in AI4Insurance

Is AI Treat? or Threat?

- McKinsey & Company estimates AI could unlock ***$50 billion*** to ***$70 billion*** of insurance industry revenue, with the highest impact on marketing and sales, customer operations, and software engineering dimensions.
- Data security landscape is shifting as AI and cloud transformation accelerate. **61%** of organizations say their AI applications are already being targeted by attackers, while **48%** report reputational damage caused by AI-generated misinformation of deepfakes.

McKinsey & Company, AI in Insurance: Understanding the Implications for Investors, **(2026)**
Thales, Top Cybersecurity Risks Facing Insures in the Age of AI and Cloud Transformation, **(2026)**

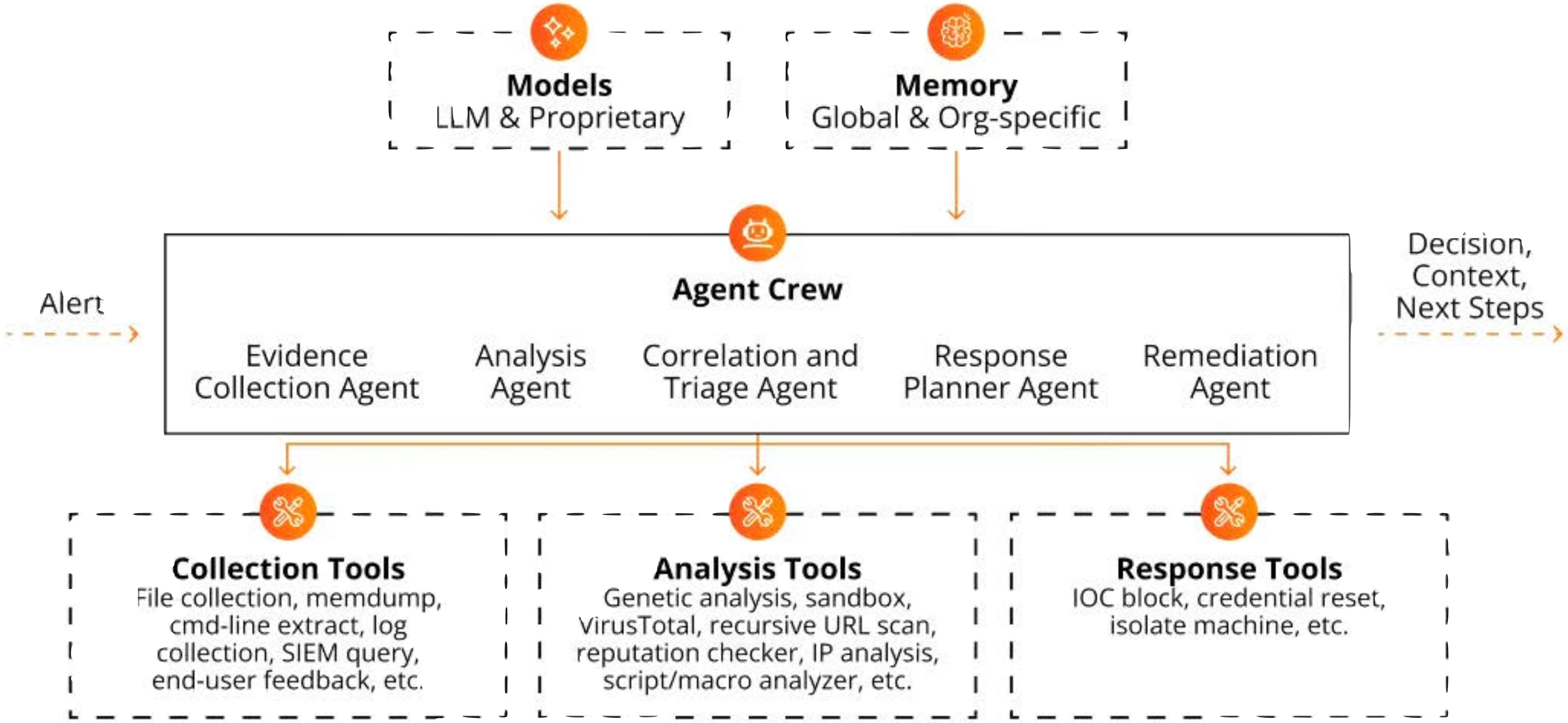


AI

배경훈 과기정통부 장관 "AI 보안 주권 구축 시급…앤트로픽 미토스 계기로 대비 서둘러야"

Source: Anthropic & 과기정통부

Agentic AI & CyberSecurity

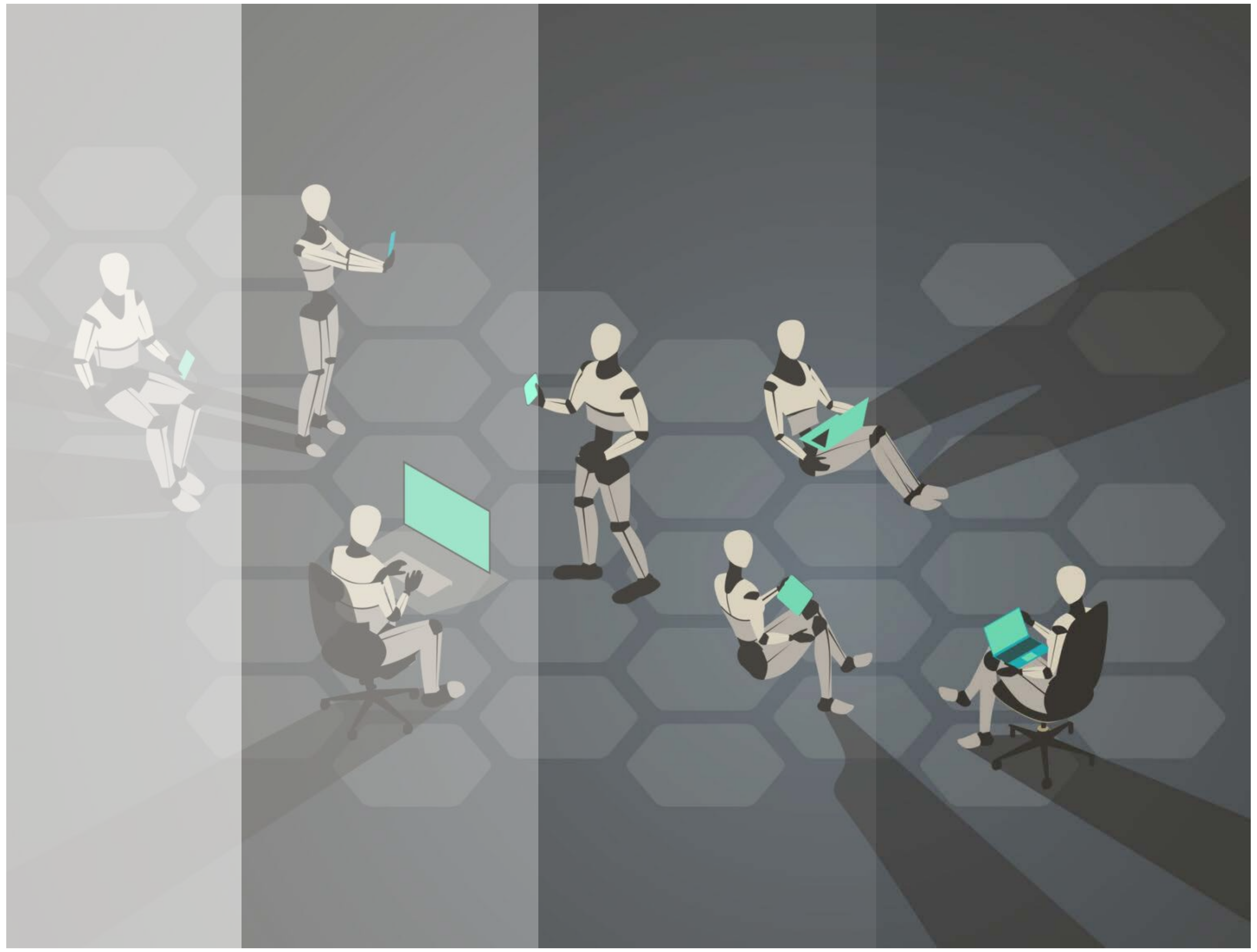


Source: N-iX

Agentic AI & Moral Hazard

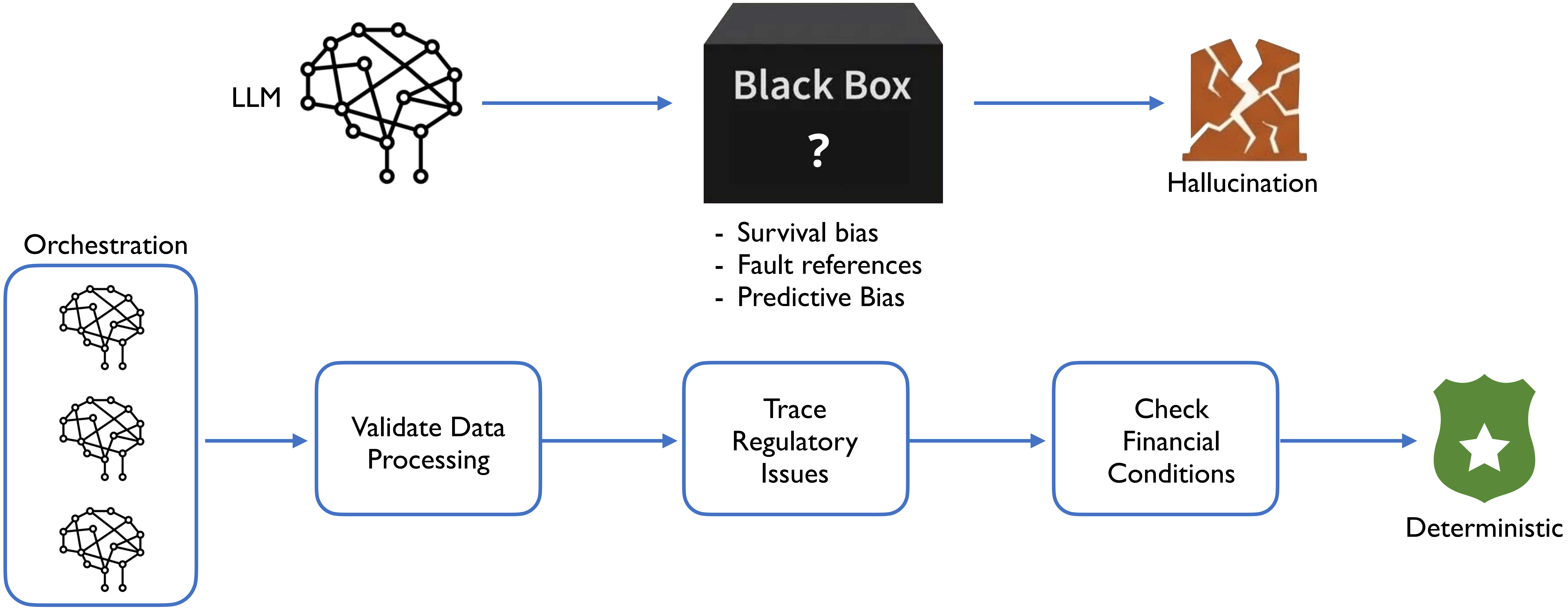
Demystifying Actuarial Skills with Agentic AI

- Processing claim records, regulatory shifts, updating mortality tables or loss, version control are heavy burden to actuarial teams.
- AI agents are enhanced with the ability to use specialized tools, maintain context over time, and make autonomous decisions.
- Use cases of AI in actuarial analysis
  - identify sudden changes in claim patterns
  - simulate risk scenarios and report



AI in Actuarial Analysis: How AI Agents are transforming insurance risk modeling, ***ampcome***, (2025)
Agentic AI: Your new actuarial coworker, ***Casualty Actuarial Society***, (2025)

Single LLM vs Orchestrating AI Agents

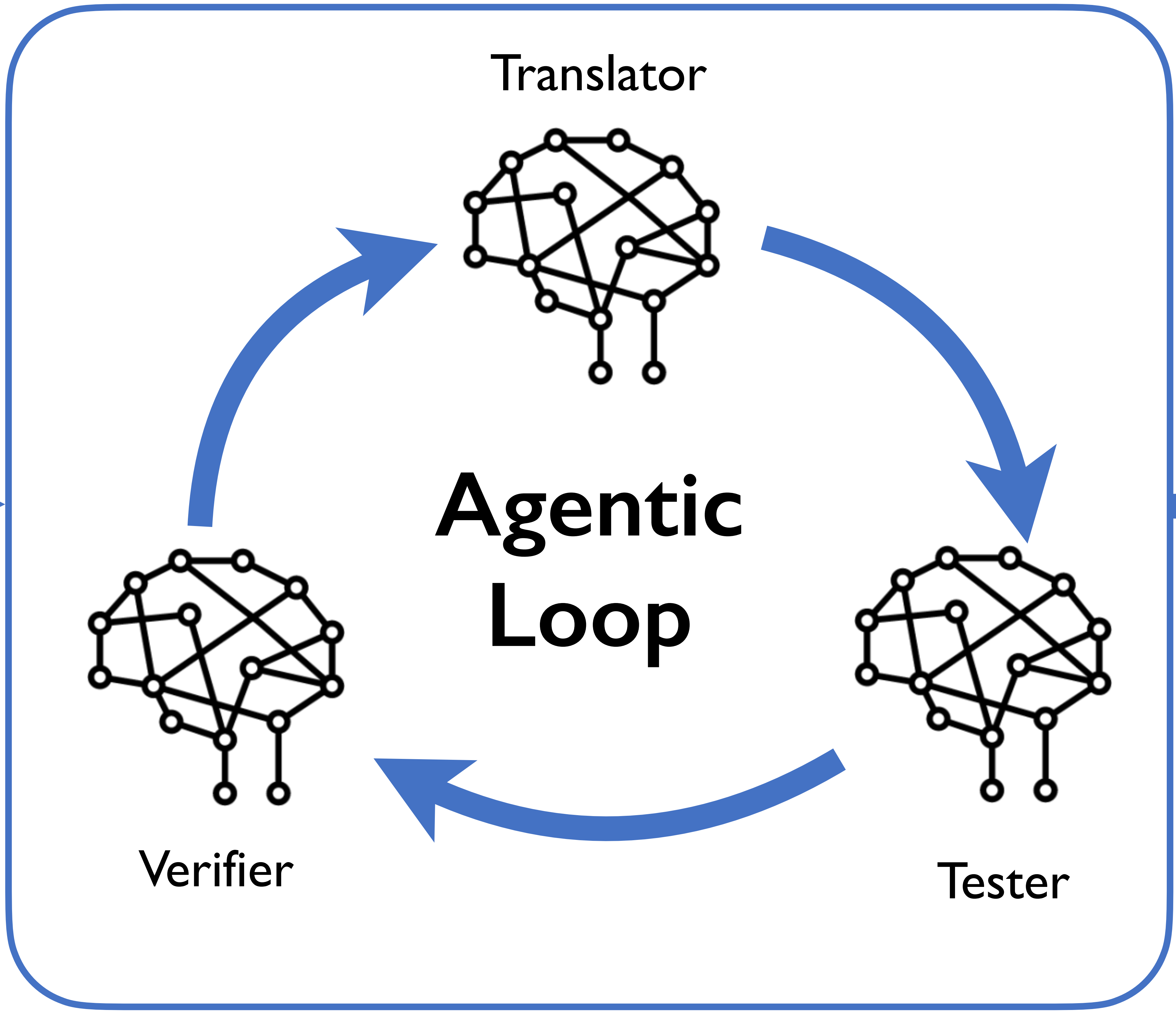


Falck et al., Is In-Context Learning in Large Language Models Bayesian? A Martingale Perspective, *ICML* (2024)

Multi-Agent based Code Migration


```r
library(shiny)
library(ggplot2)
library(dplyr)

# Sample Data
data <- data.frame(
    Date = as.Date(c("2023-01-01", "2023-01-02", "2023-01-01", "2023-01-02")),
    Country = c("USA", "USA", "Canada", "Canada"),
    Value = c(100, 120, 80, 90)
)

# Shiny App
ui <- fluidPage(
    selectInput("country_filter", "Select Countries", choices = unique(data$Country), sele
    plotOutput("line_chart")
)

server <- function(input, output) {
    output$line_chart <- renderPlot({
        filtered_data <- data %>% filter(Country %in% input$country_filter)
        ggplot(filtered_data, aes(x = Date, y = Value, color = Country)) +
            geom_line() +
            labs(title = "Value Over Time", x = "Date", y = "Value") +
            theme_minimal()
    })
}

shinyApp(ui, server)
```

Legacy actuarial code
(R, SAS, …)




```python
import boto3
import pandas as pd
import matplotlib.pyplot as plt

# Step 1: Connect to Amazon S3 and Retrieve Data
s3 = boto3.client("s3")
bucket_name = "your-bucket-name"
file_key = "path/to/metadata-dump.csv"

# Download the file locally
s3.download_file(bucket_name, file_key, "metadata.csv")

# Step 2: Load the Data into a Pandas DataFrame
data = pd.read_csv("metadata.csv")

# Step 3: Filter and Process Data
# Example: Filter for specific app package names
app_packages = ["com.example.app1", "com.example.app2"]
filtered_data = data[data["package_name"].isin(app_packages)]

# Summarize metadata (e.g., count occurrences by app package)
summary = filtered_data.groupby("package_name")["event_count"].sum().reset_index()

# Step 4: Visualize the Processed Data
plt.figure(figsize=(8, 6))
plt.bar(summary["package_name"], summary["event_count"], color="skyblue")
plt.title("Event Counts by App Package")
plt.xlabel("App Package Name")
plt.ylabel("Total Events")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

New code
(python, Rust, …)

Orchestrating Agents for Causal Data Analysis

ORCA: ORchestrating Causal Agent

Joanie Hayoun Chung
Department of Statistics,
Korea University
Seoul, South Korea
jchung02@korea.ac.kr

Sumin Lee
Department of Statistics,
Korea University
Seoul, South Korea
leesumin0924@korea.ac.kr

Sungbin Lim
Department of Statistics
Korea University
LG AI Research
Seoul, South Korea
sungbin@korea.ac.kr

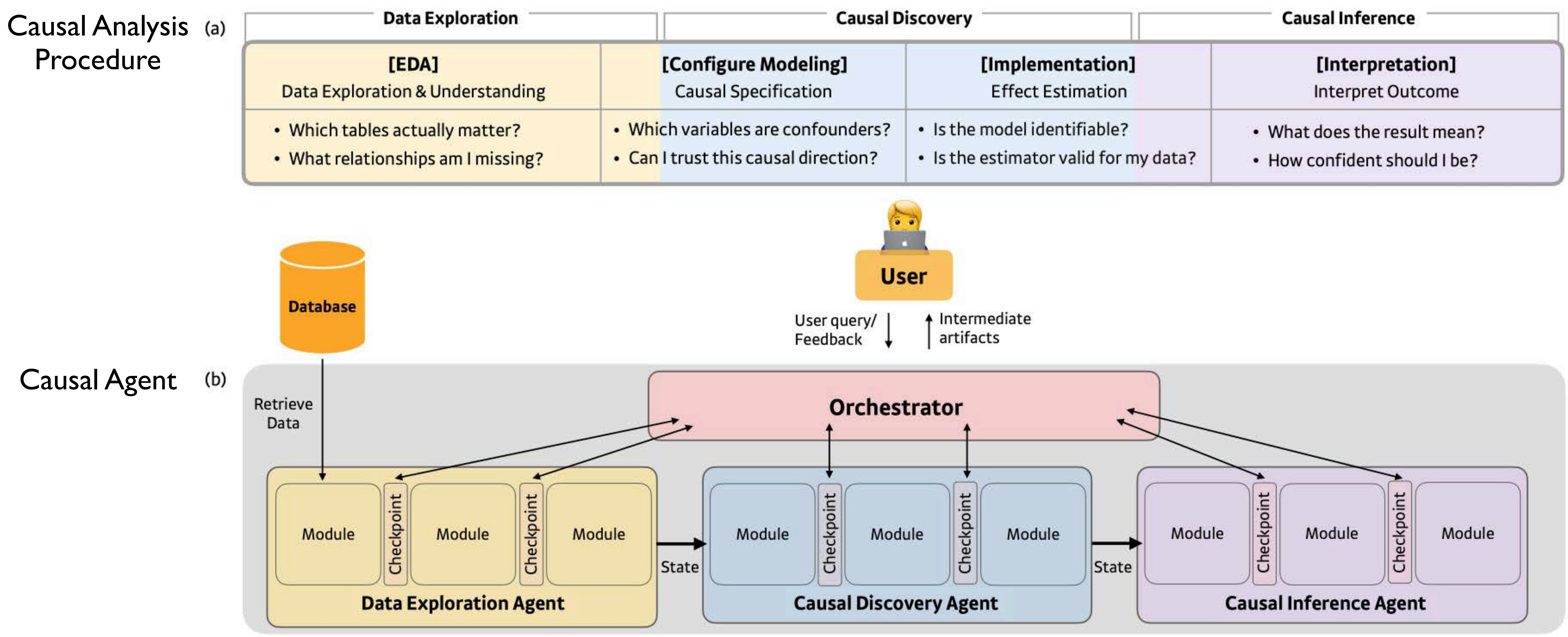
Chung, Hayoun
Korea Univ.

Lee, Sumin
Korea Univ.

LIM(†)
Korea Univ.
LG AI Research

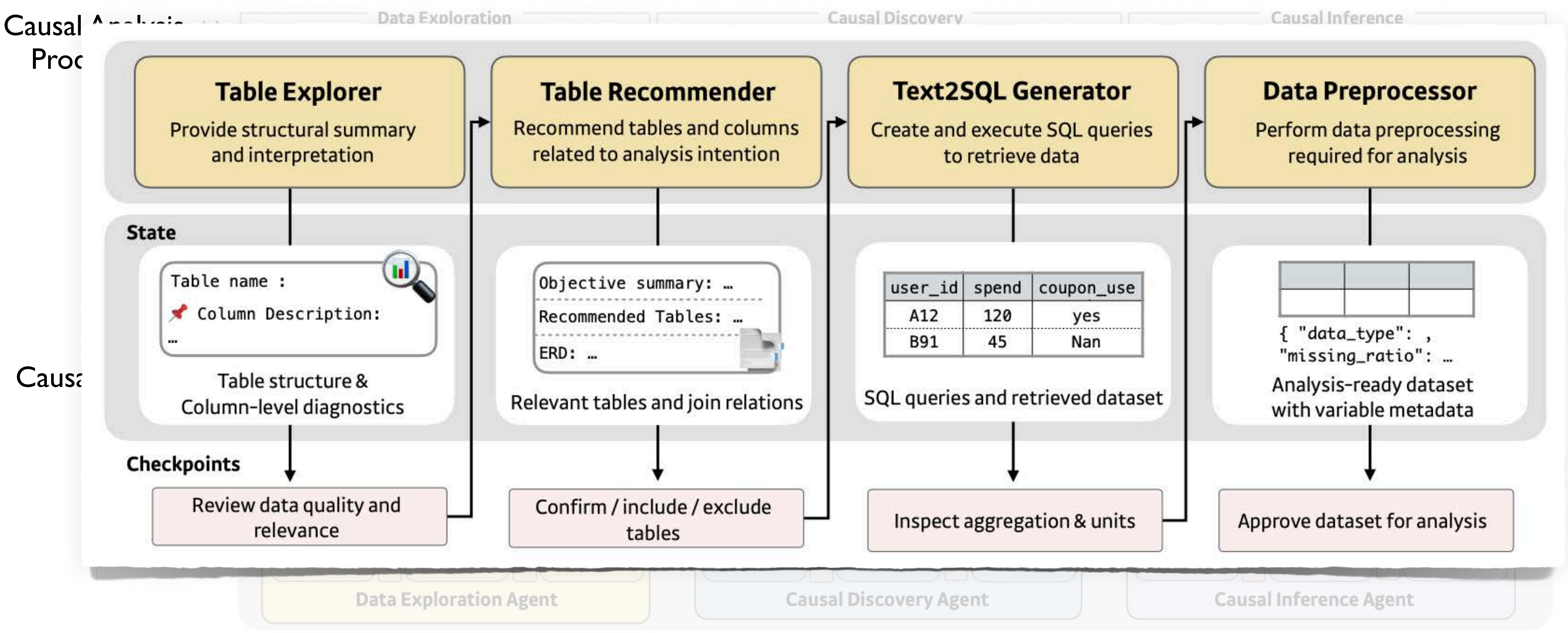
Chung et al., ORCA: ORchestrating Causal Agent, ***CHI poster*** (2025)

Orchestrating Agents for Causal Data Analysis



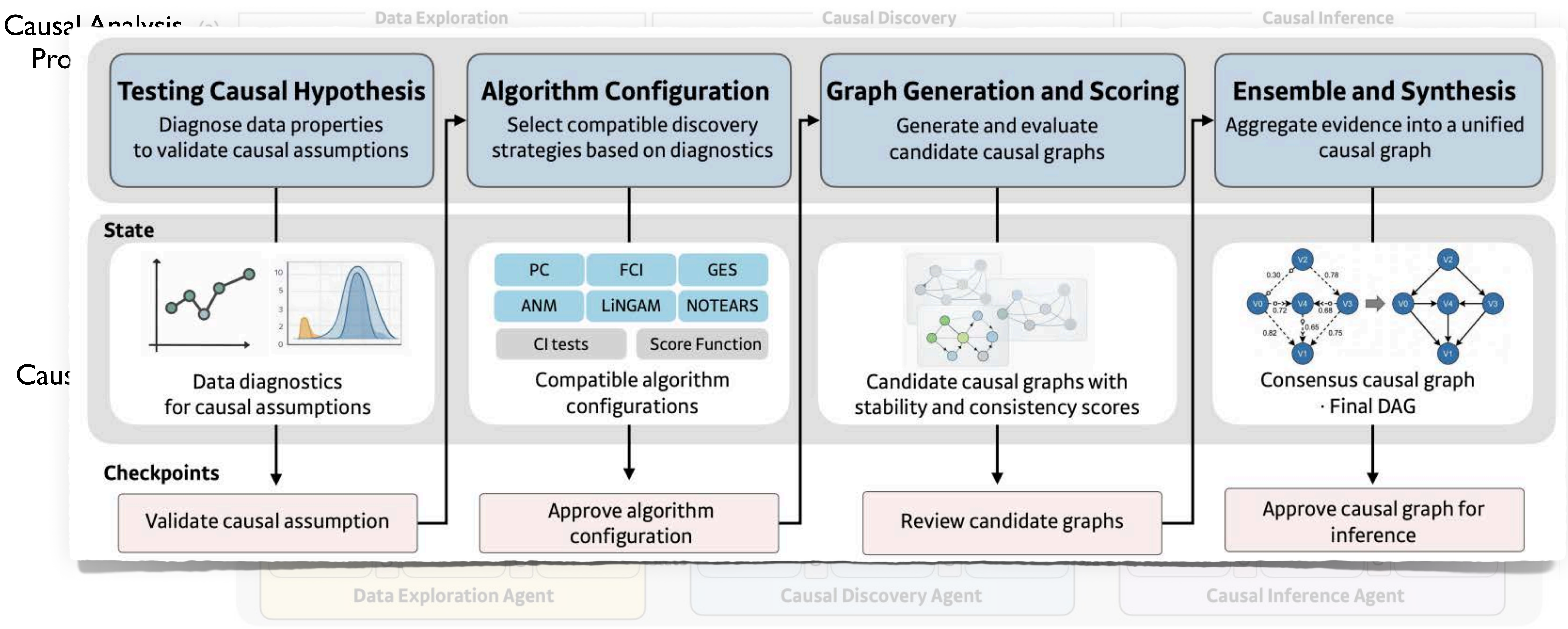
Chung et al., ORCA: ORchestrating Causal Agent, ***CHI poster*** (2025)

Orchestrating Agents for Causal Data Analysis



Chung et al., ORCA: ORchestrating Causal Agent, ***CHI poster*** (2025)

Orchestrating Agents for Causal Data Analysis

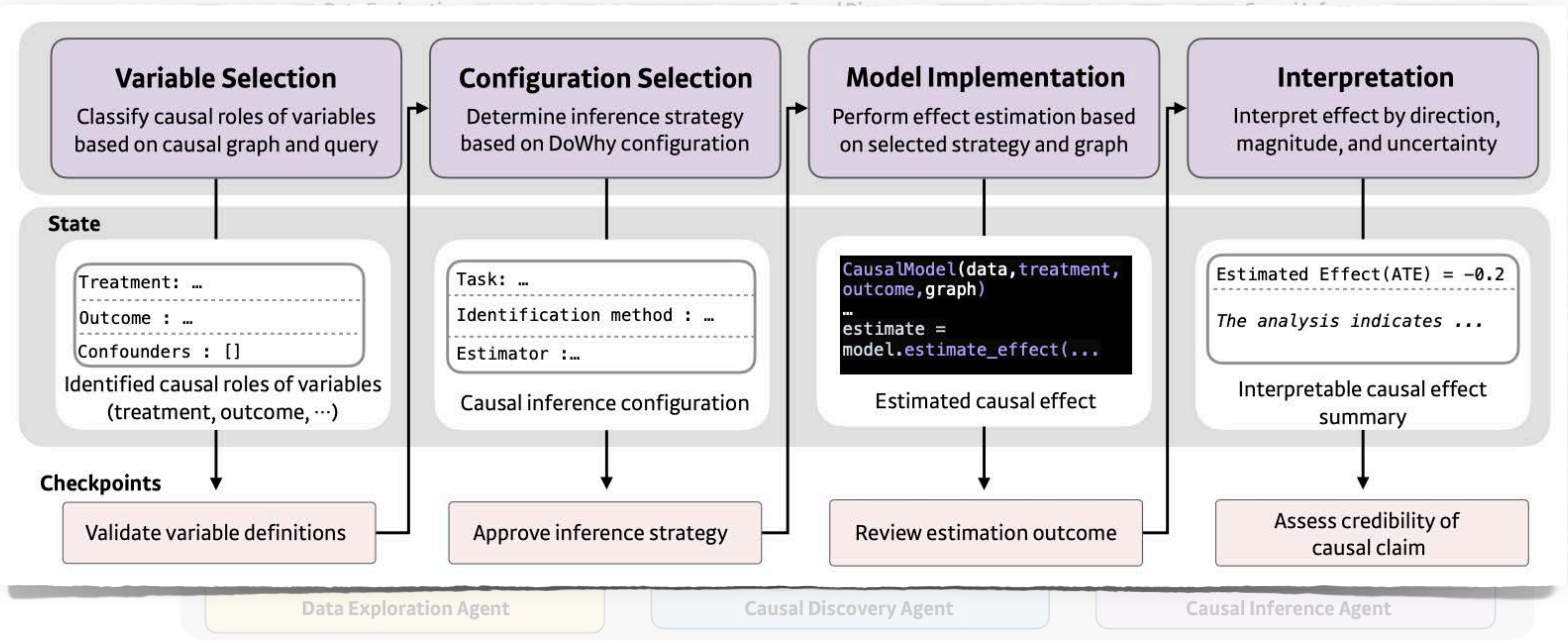


Chung et al., ORCA: ORchestrating Causal Agent, ***CHI poster*** (2025)

Orchestrating Agents for Causal Data Analysis

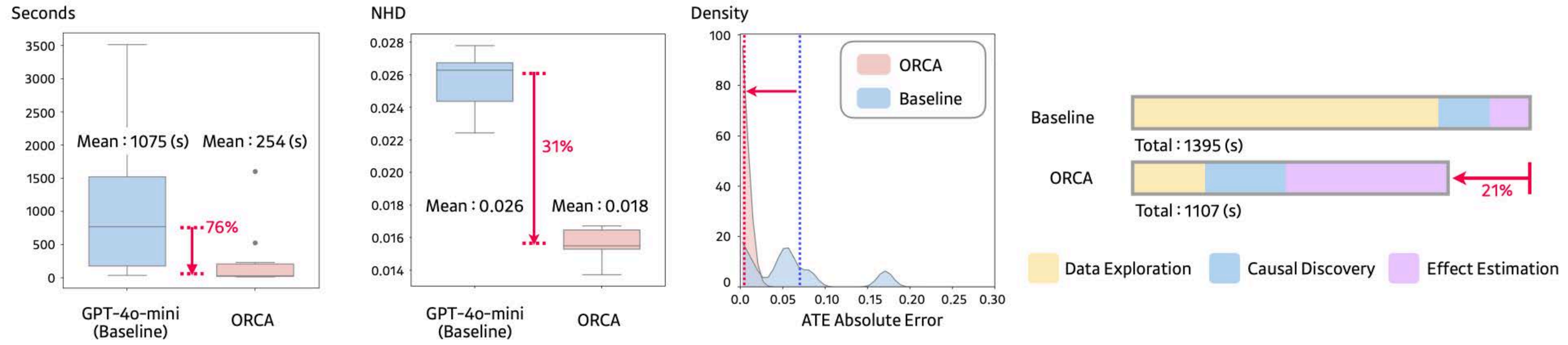
Caus
Pr

Cau



Chung et al., ORCA: ORchestrating Causal Agent, ***CHI poster*** (2025)

Orchestrating Agents for Causal Data Analysis



(a) Data Exploration Duration (s) (b) Causal Graph Error Comparison (c) Distribution of Absolute Error (d) Total Task Duration and Proportion

	User interaction time		User interaction turns	
	Baseline	ORCA	Baseline	ORCA
Data exploration step	439.34 ± 373.32	80.60 ± 77.71	5.72 ± 4.17	1 ± 2.96
Causal discovery step	78.49 ± 117.51	164.02 ± 155.64	2 ± 4.14	1.1 ± 2.331
Causal analysis step	71.48 ± 66.69	155.90 ± 202.15	1.64 ± 1.8	0.714 ± 1.25

데이터 엔지니어링 부담 감소 & 분석 품질 향상

Chung et al., ORCA: ORchestrating Causal Agent, ***CHI poster*** (2025)

Q&A

Department of Statistics
KOREA UNIVERSITY

KOREA UNIVERSITY