

Designing Consensus: Gamified Modeling and Simulation of Collaborative Decision-Making

Instructor: Jin Gao

IAP 2025 (Non-Credit)

2025.01.28

gaojin@mit.edu



HAPPY YEAR OF THE
SNAKE

Designed by Freepik

Jan 28	Swarm Intelligence and Emergent Intelligence	<ul style="list-style-type: none">● From a micro perspective, introduction to agent-based approaches as another method for performing simulations.● Introduction to swarm intelligence, inspired by real-world examples (e.g., ants), and their simulation through collaborative and competitive agents.● Introducing current advances in emergent intelligence with Large-language-Model agents and their limitations.● Introducing the application behind modern simulation games such as Cities: Skylines, Civilization, and Anno.
	Workshop	<ul style="list-style-type: none">● Experiment with GAMA to simulate pedestrian behavior in a small town.● Prepare your presentation

Agent

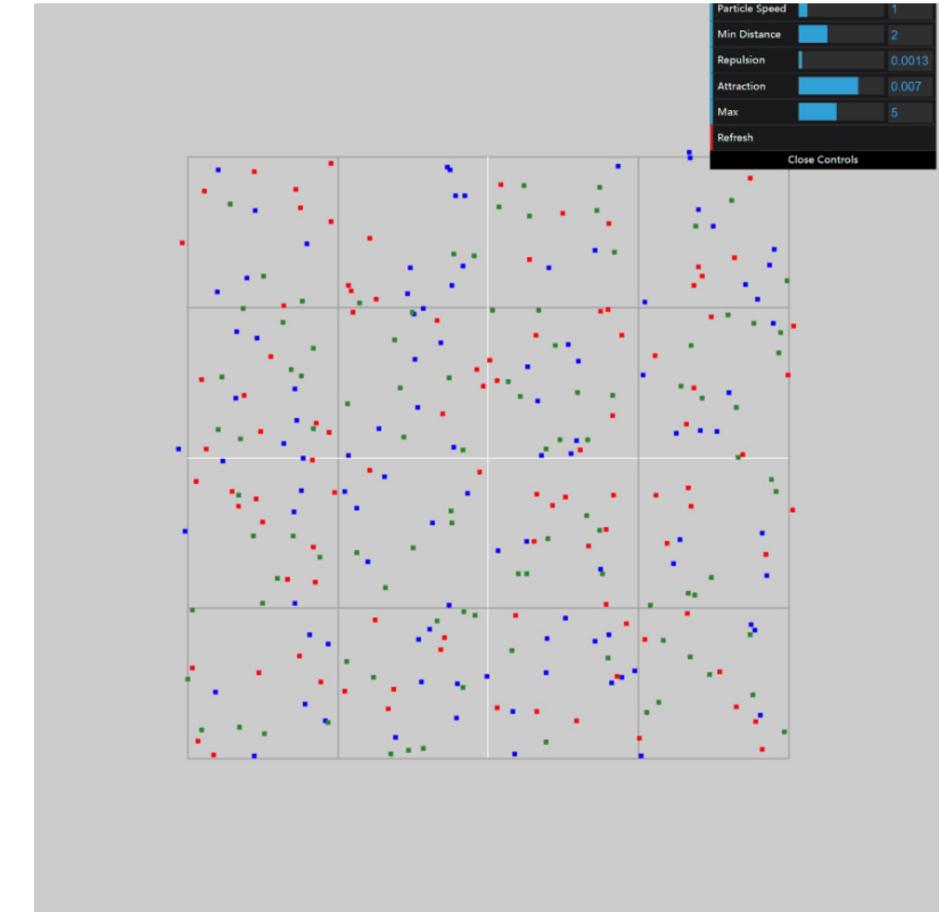
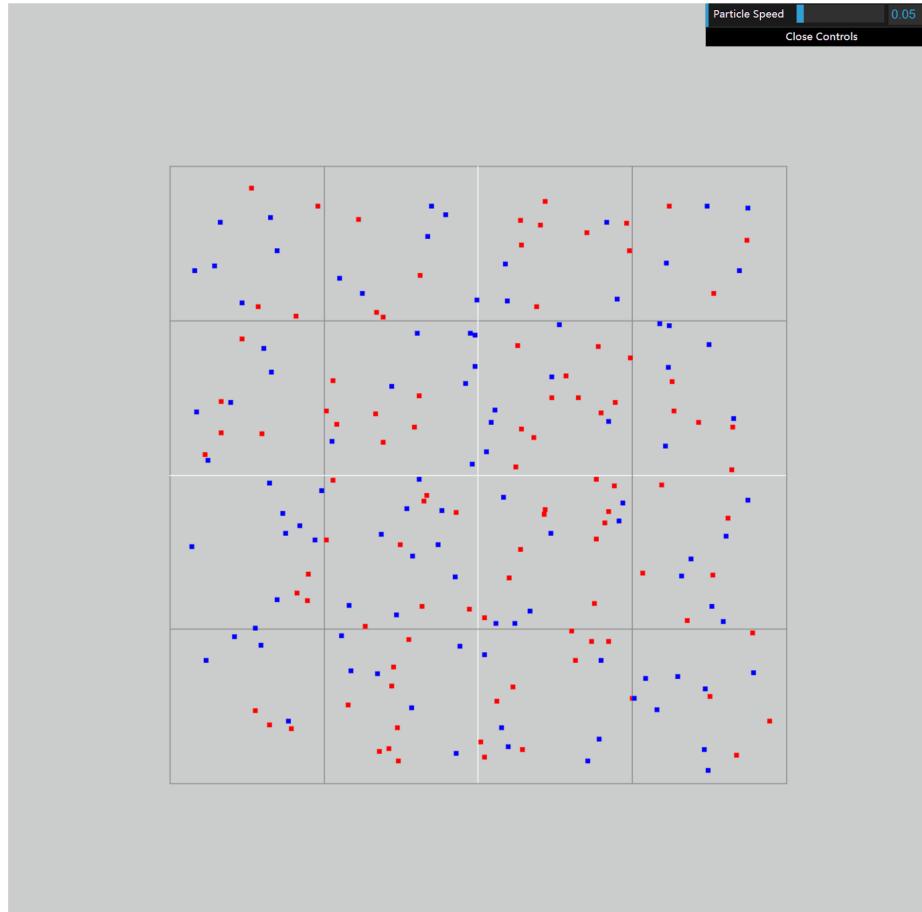


Agent-based **models** are computer **simulations** used to study
the **interactions** between people, things, places, and time.

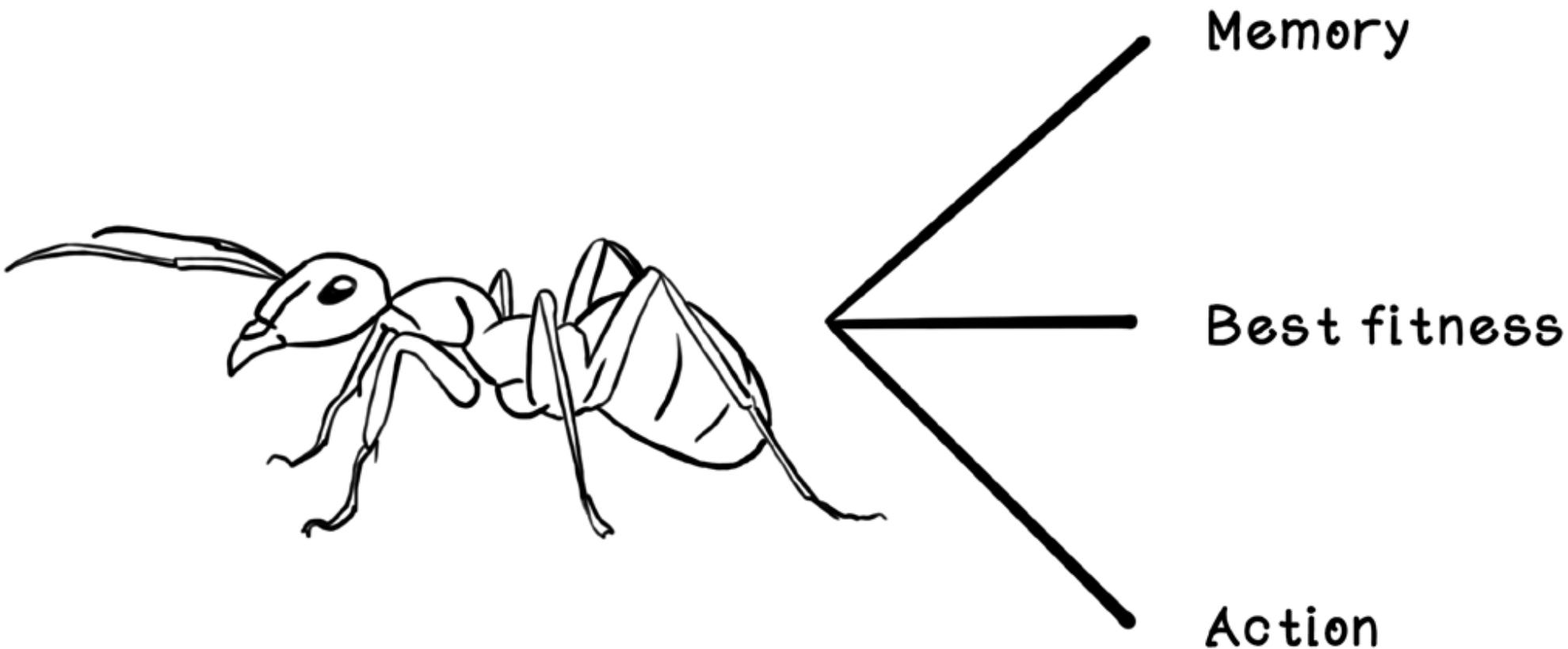
Multi-Agent



Agent Based Simulation and Swarm Intelligence



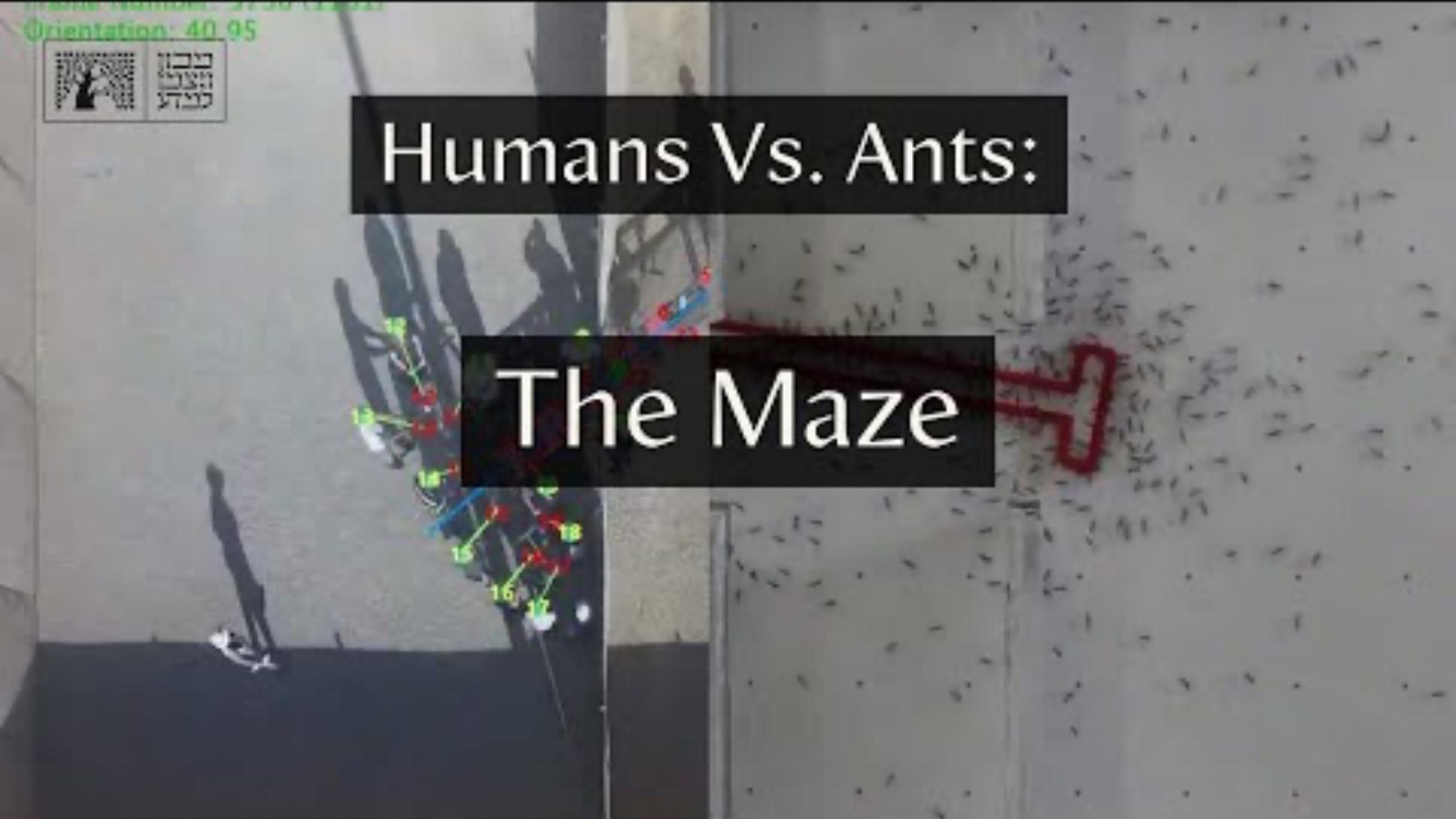


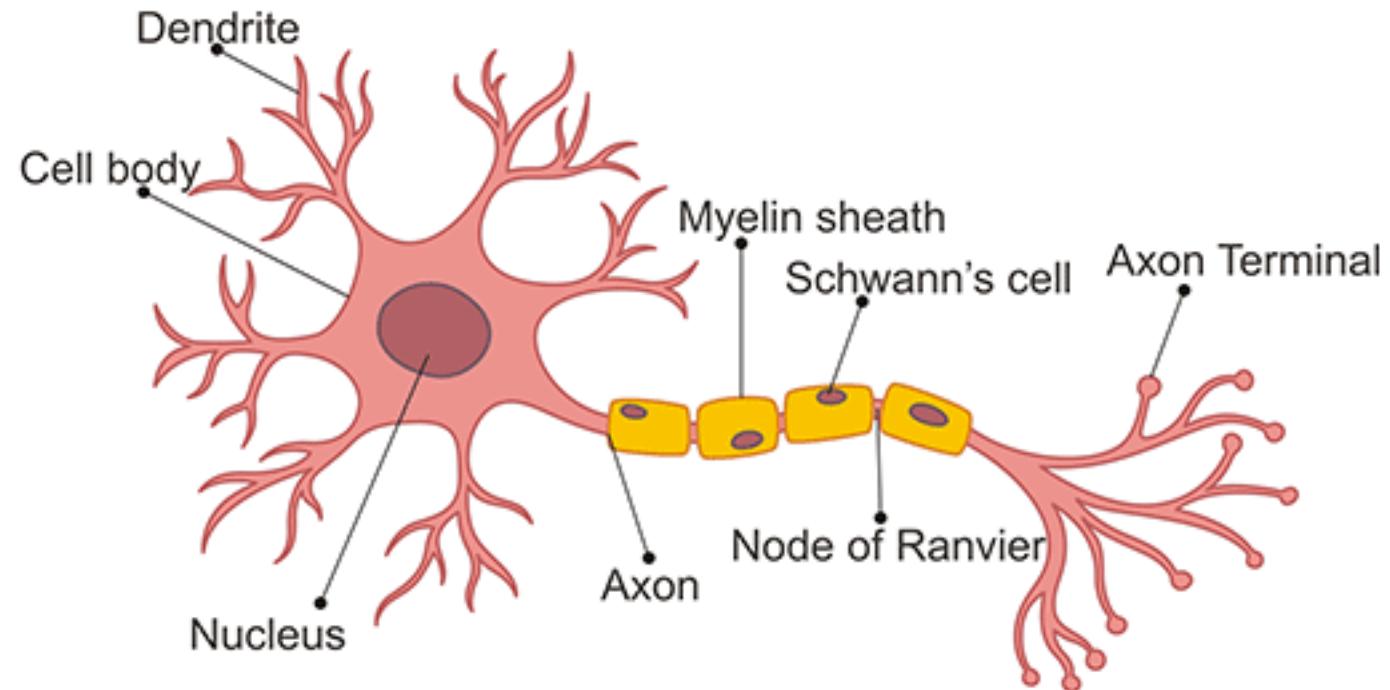


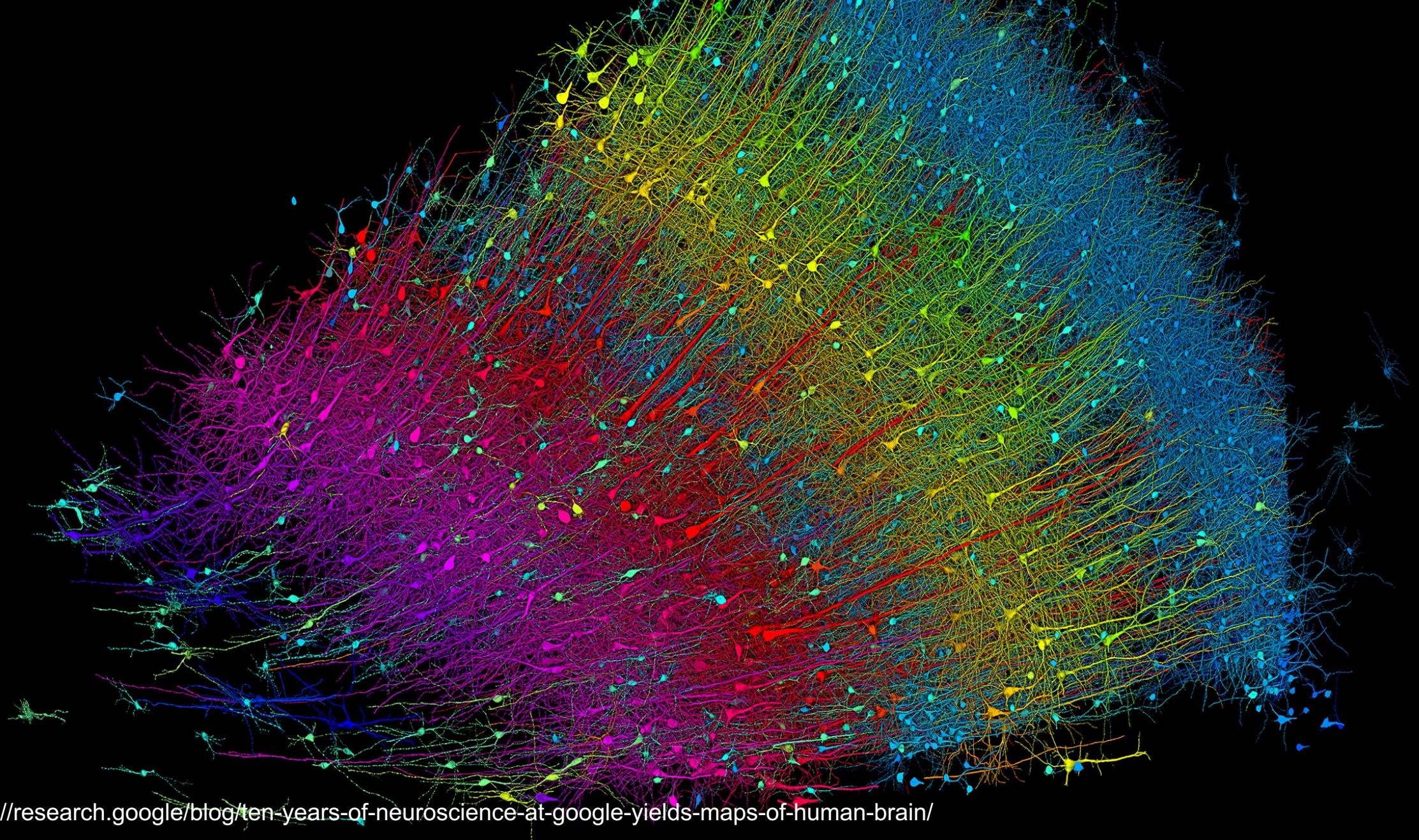


Humans Vs. Ants:

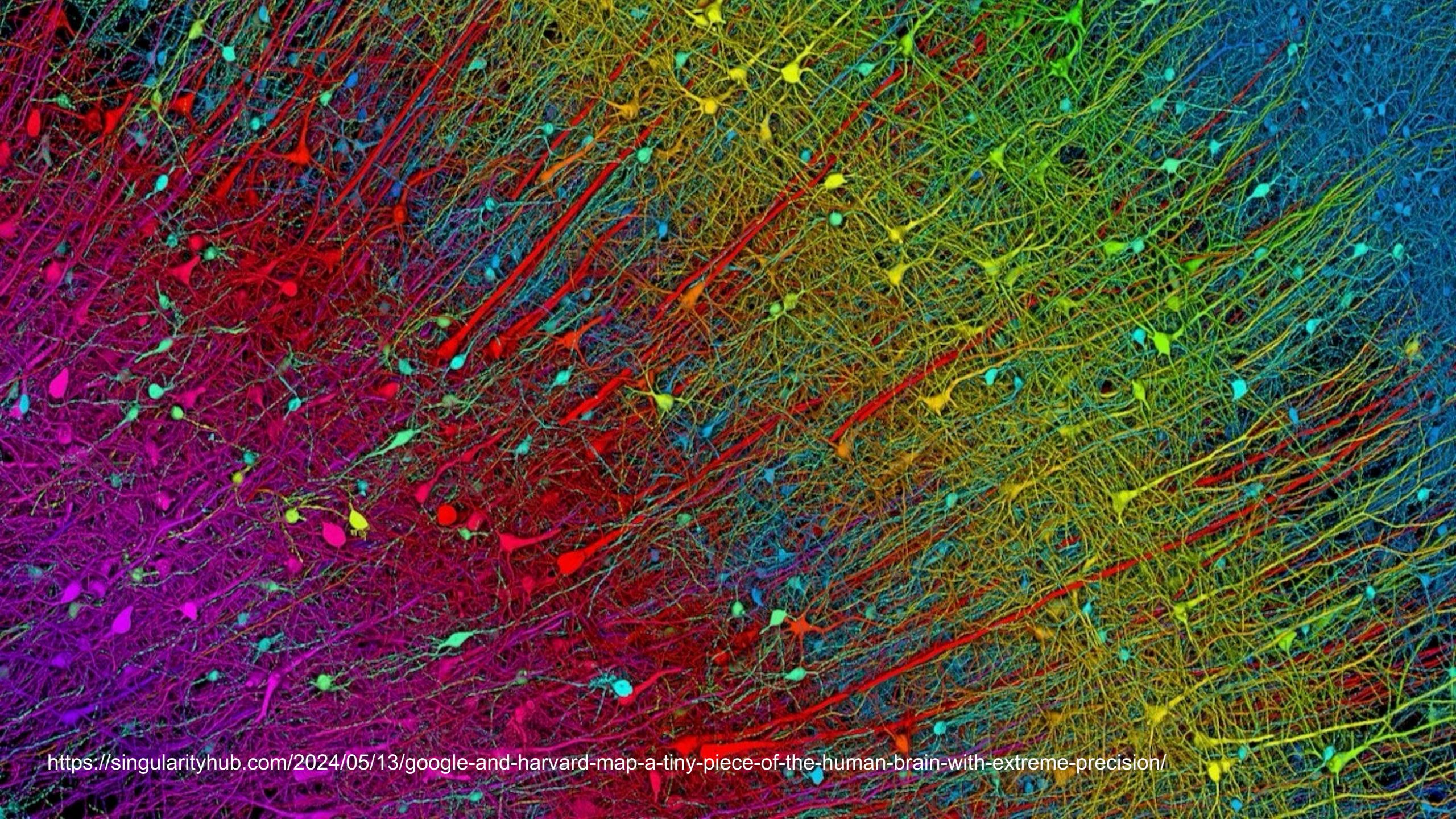
The Maze



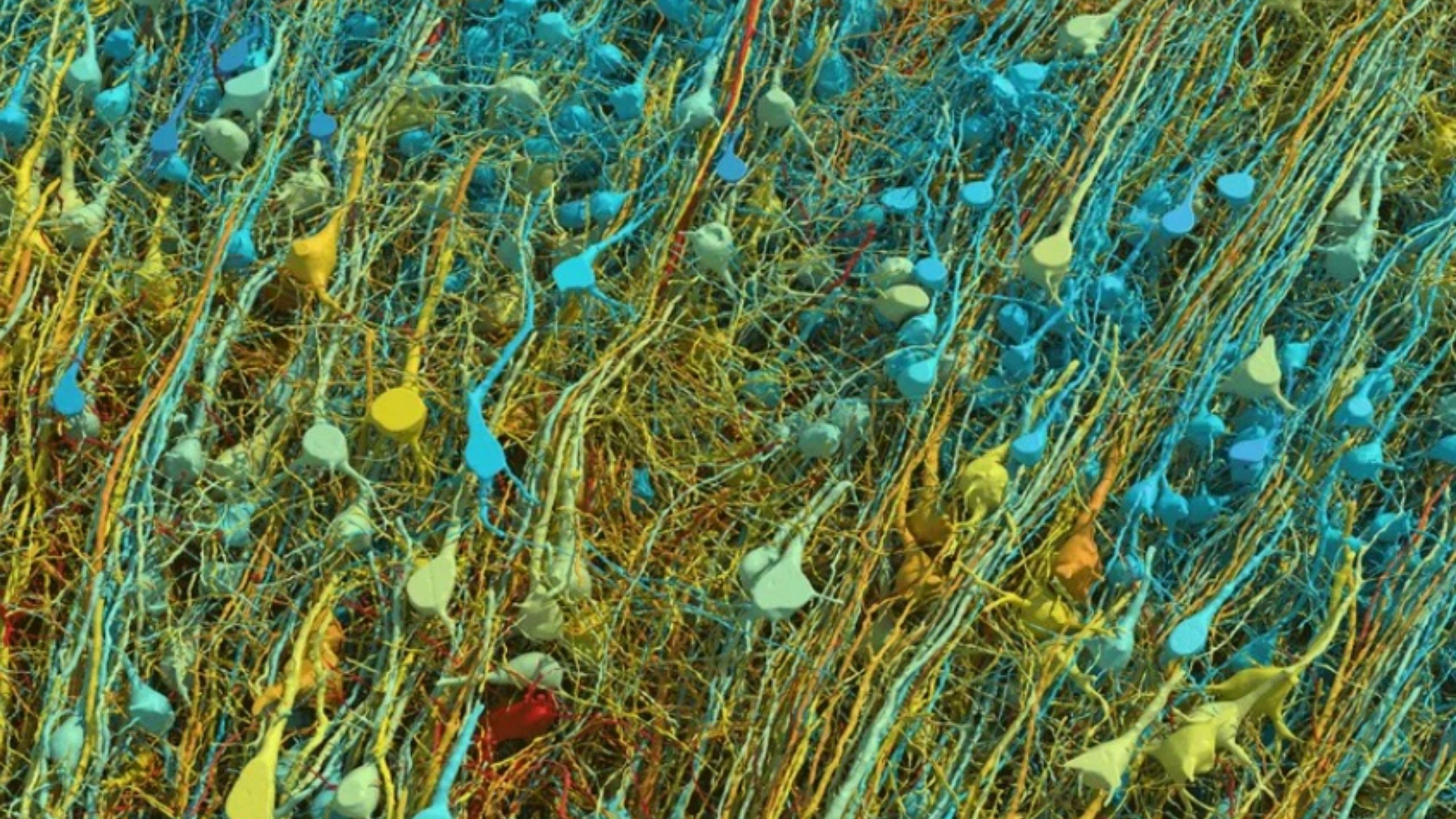


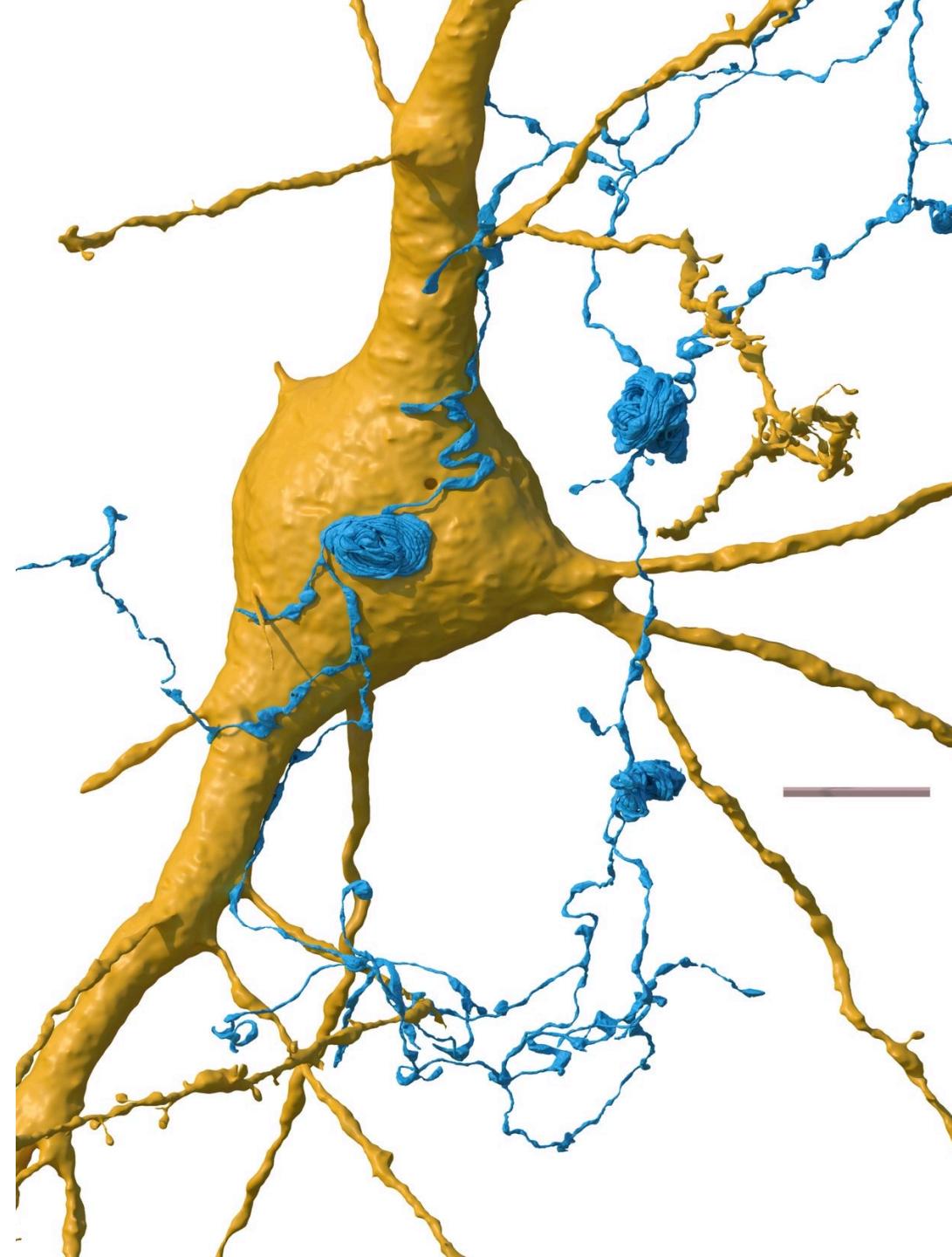


<https://research.google/blog/ten-years-of-neuroscience-at-google-yields-maps-of-human-brain/>

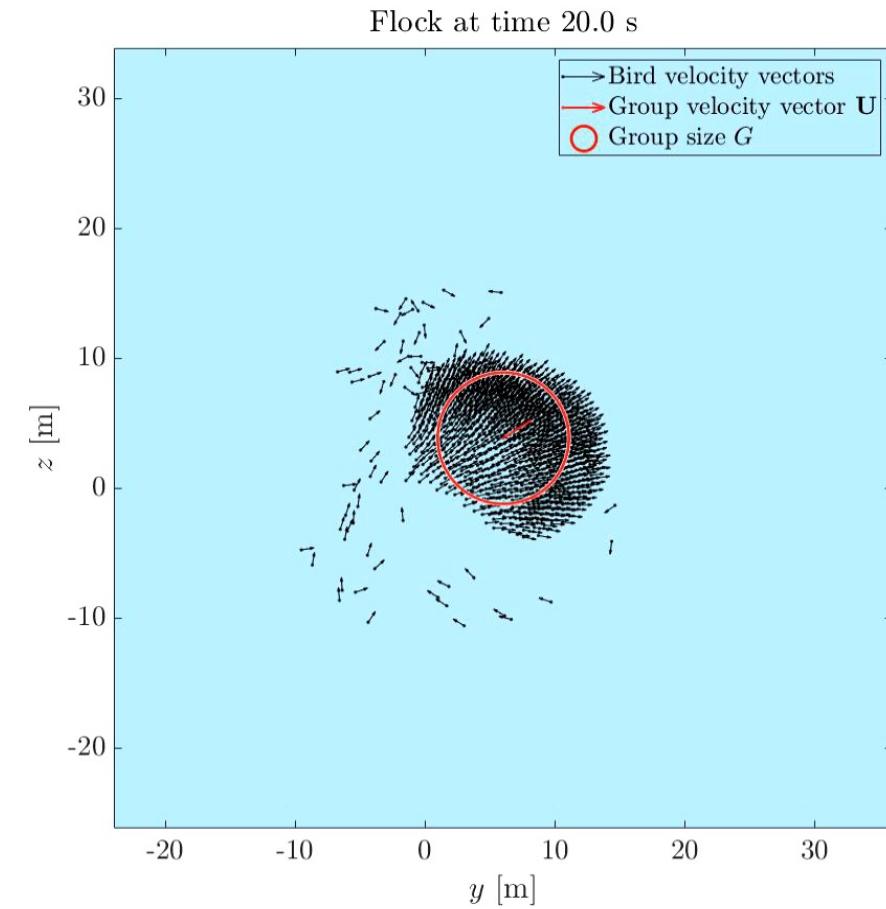


<https://singularityhub.com/2024/05/13/google-and-harvard-map-a-tiny-piece-of-the-human-brain-with-extreme-precision/>



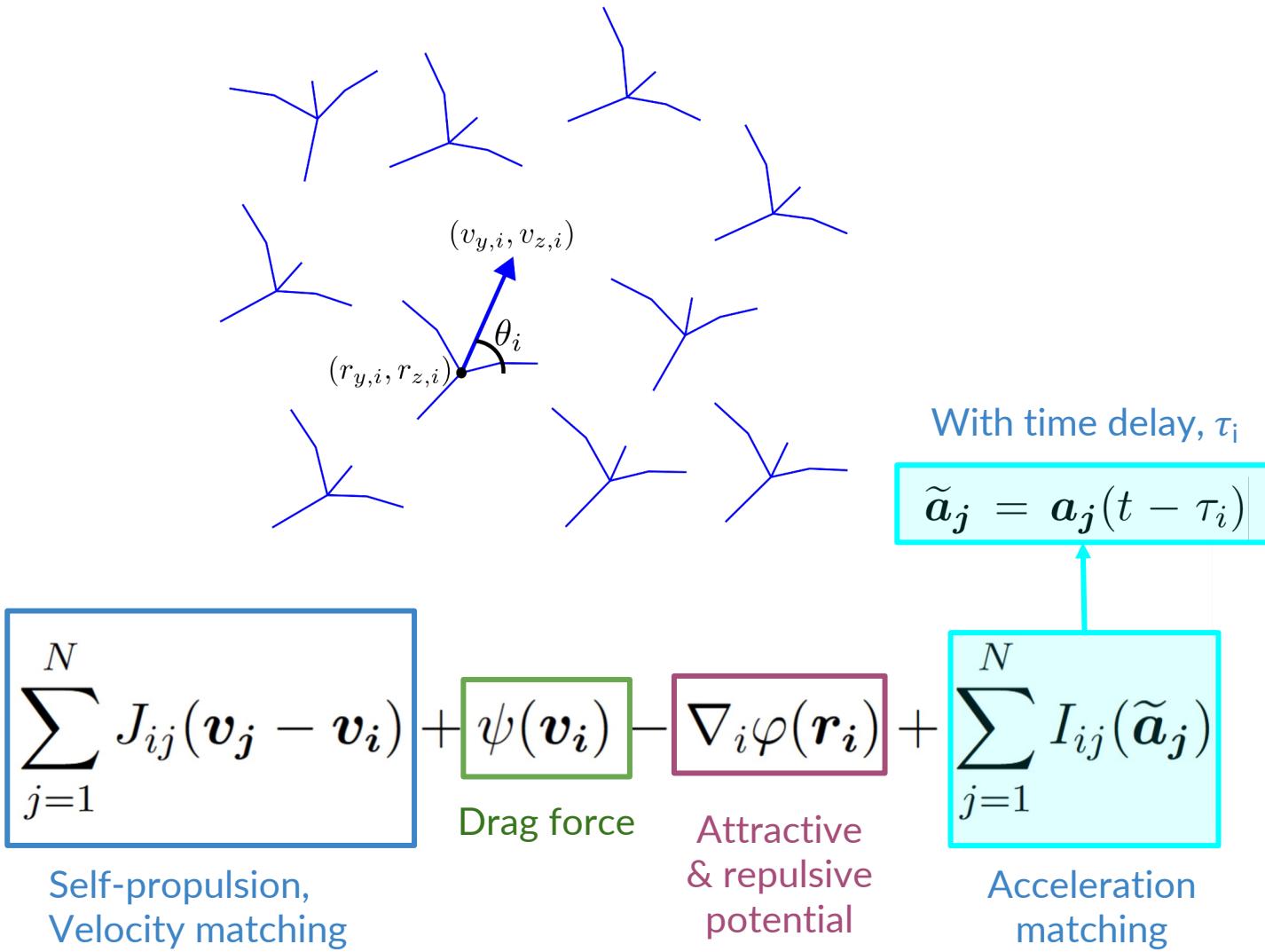


Swarm Intelligence



Swarm Intelligence

$$\boxed{x} = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_N \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_N \end{bmatrix} = \begin{bmatrix} r_{y1} \\ \vdots \\ r_{yN} \\ r_{z1} \\ \vdots \\ r_{zN} \\ v_{y1} \\ \vdots \\ v_{yN} \\ v_{z1} \\ \vdots \\ v_{zN} \end{bmatrix} \quad \left. \begin{array}{l} \text{2D Positions} \\ \text{2D Velocities} \end{array} \right\}$$



Model Formulation

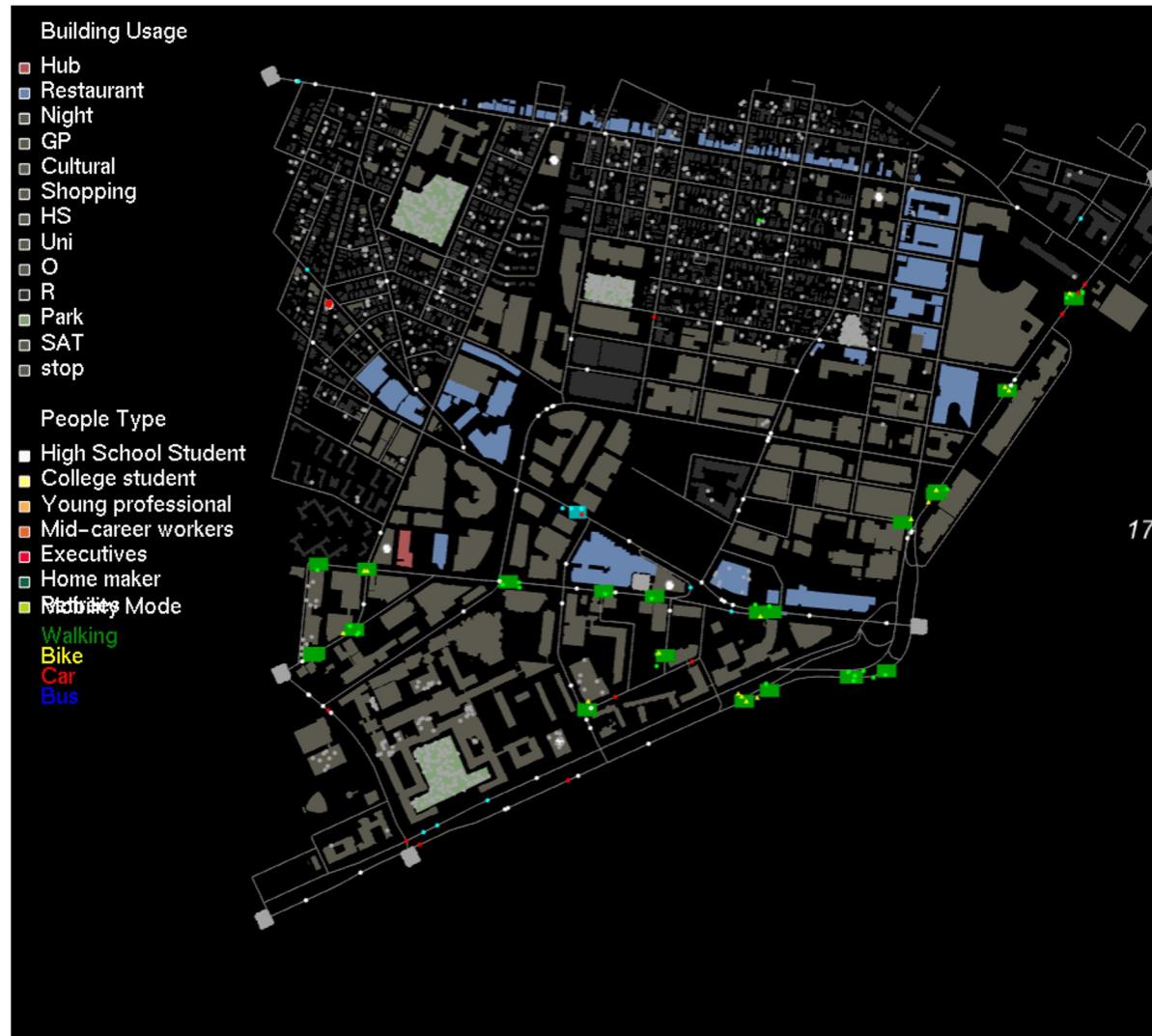
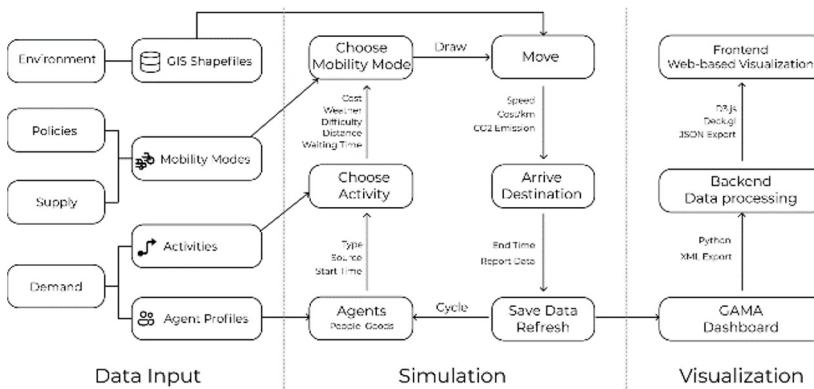
Multi-purpose Automatic Vehicles Simulation

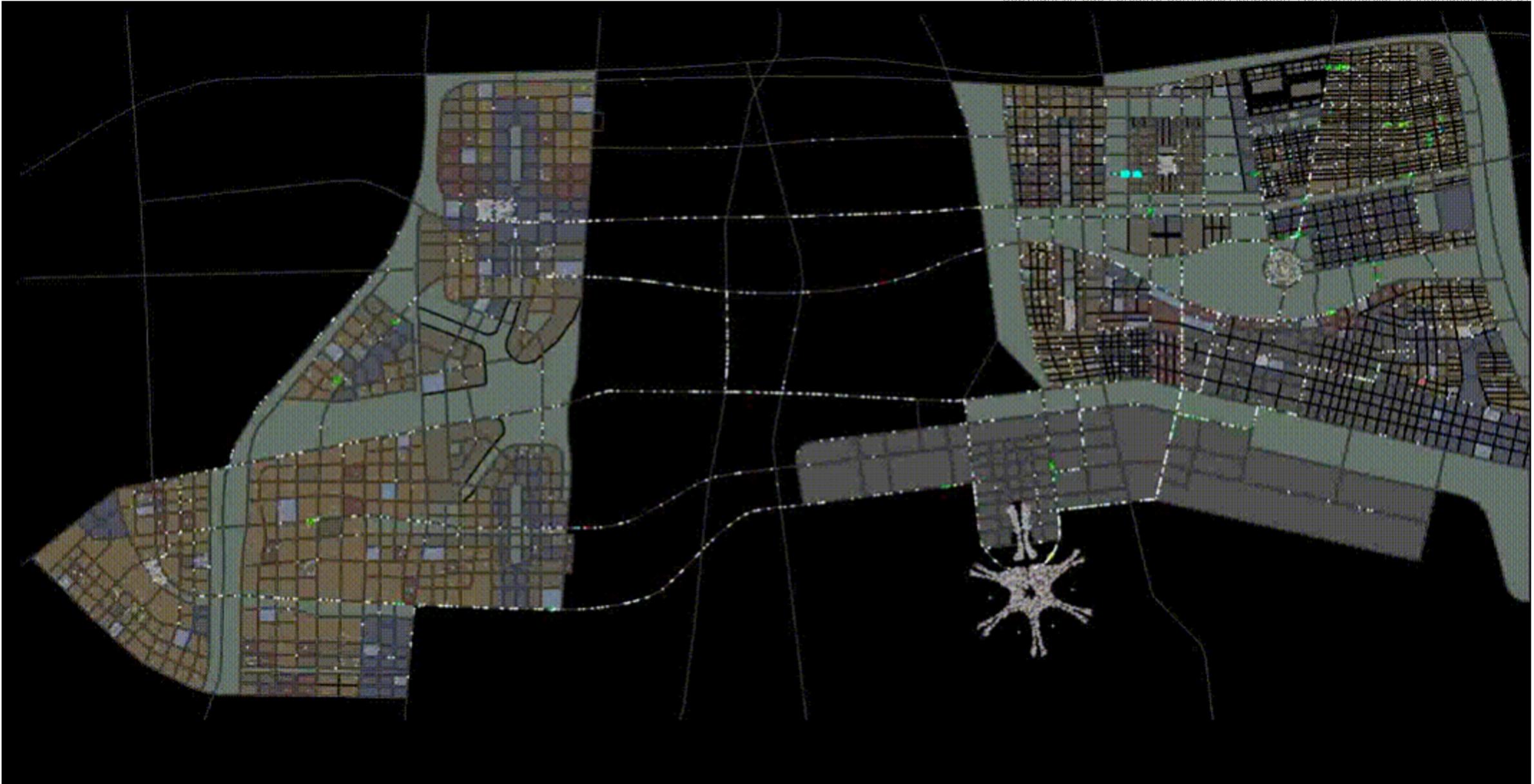
Description:

Experimenting how an automatic vehicle that can transport both passengers and goods benefits mobility efficiency.

Finding:

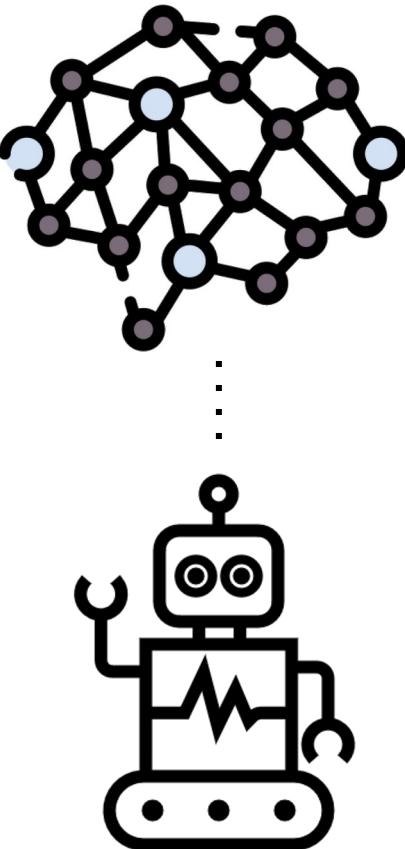
For Kendall Square, it is good to have a combination of large number of small-scale multi purpose vehicles and a few big ones.





Beijing Daxing Airport Region

Generative LLM Agent



Multi-Generative Agent System



Multi-Generative Agent System proposed in recent researches [Park et al., 2023] (left) and [Chen et al., 2023] (right)

Frontier Last Year: Generative Agents

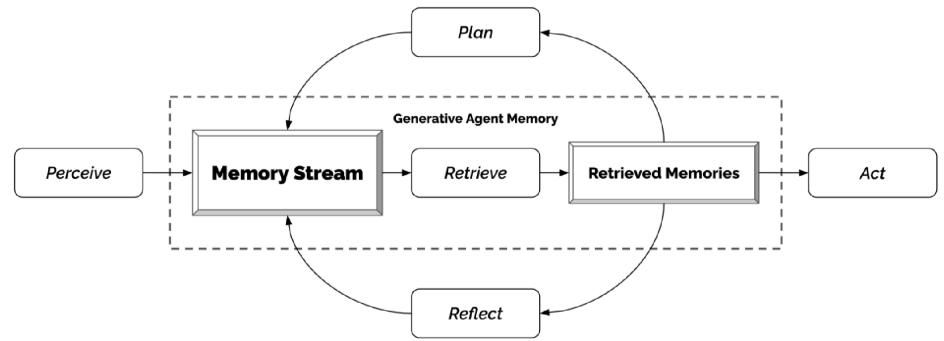
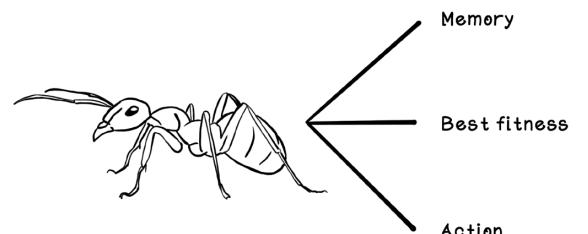


Figure 5: Our generative agent architecture. Agents perceive their environment, and all perceptions are saved in a comprehensive record of the agent's experiences called the memory stream. Based on their perceptions, the architecture retrieves relevant memories and uses those retrieved actions to determine an action. These retrieved memories are also used to form longer-term plans and create higher-level reflections, both of which are entered into the memory stream for future use.



https://reverie.herokuapp.com/arXiv_Demo/

Park, Joon Sung, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. "Generative agents: Interactive simulacra of human behavior." *arXiv preprint arXiv:2304.03442* (2023).

Frontier Last Year: ChatDev

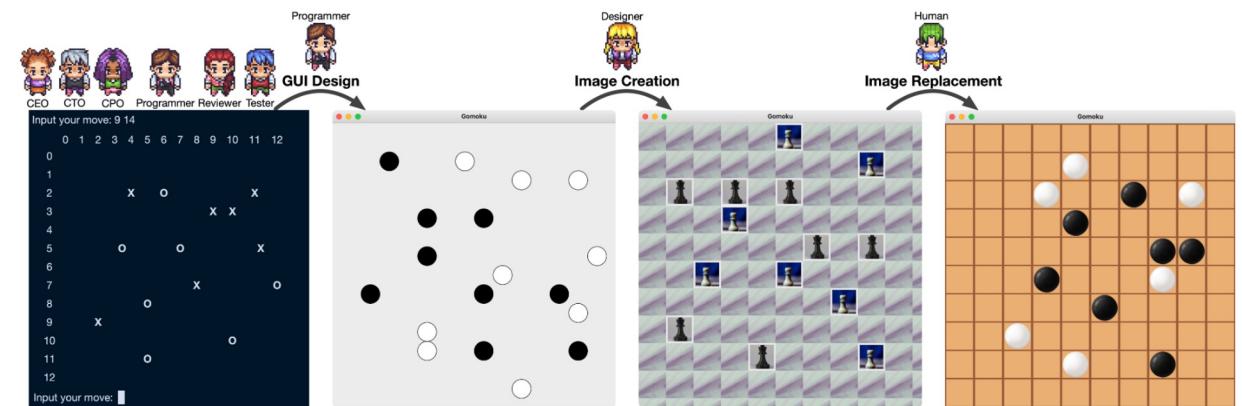


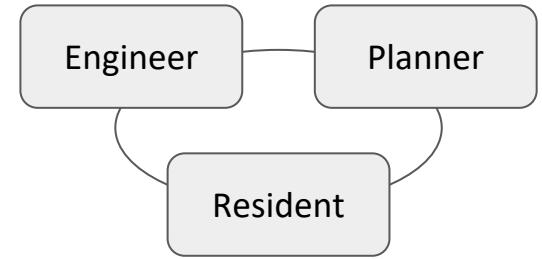
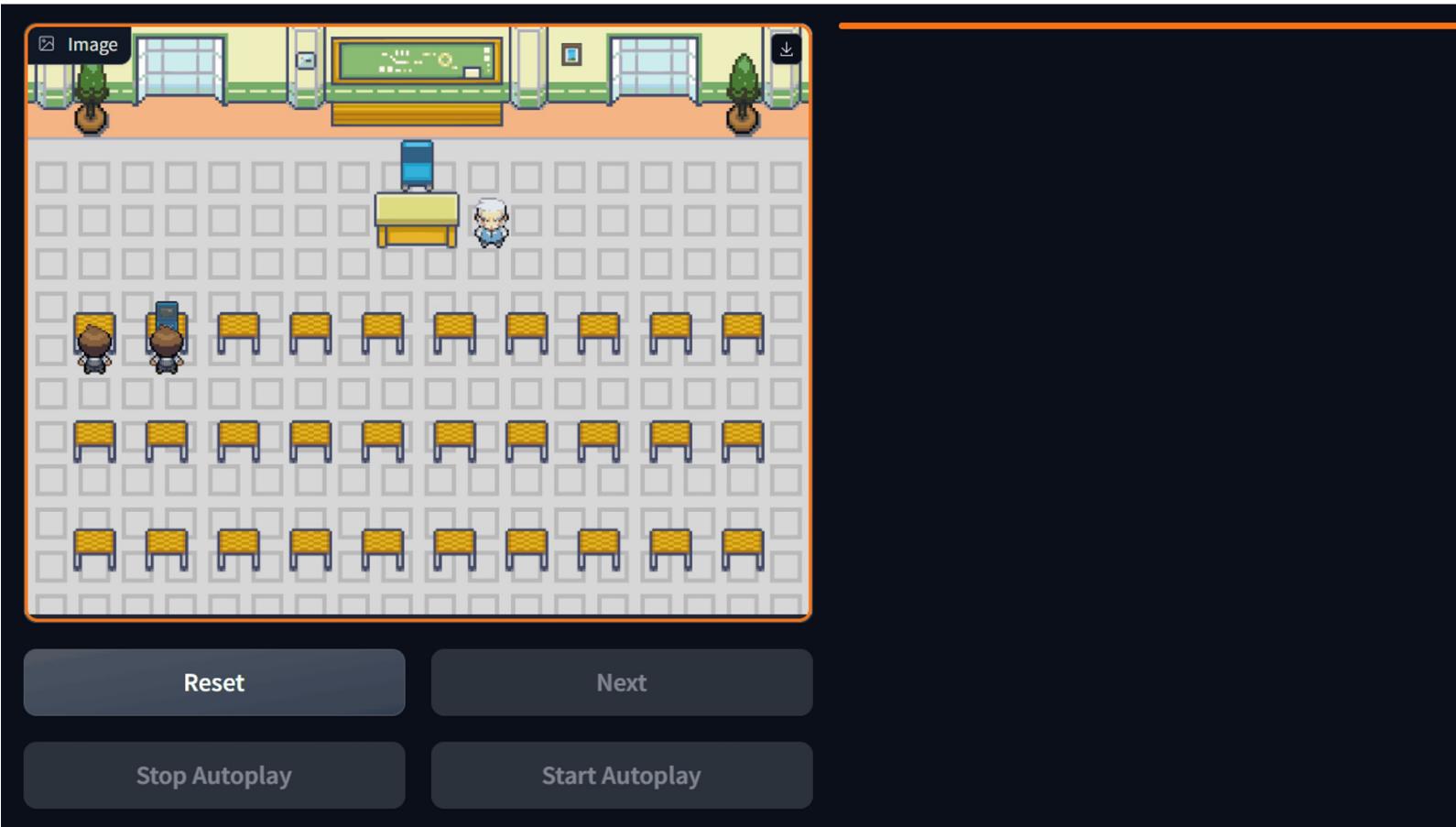
Figure 9: The produced software of the task: “*design a basic Gomoku game*”.

<https://arxiv.org/abs/2307.07924>

Chen Qian, , Xin Cong, Wei Liu, Cheng Yang, Weize Chen, Yusheng Su, Yufan Dang, Jiahao Li, Juyuan Xu, Dahai Li, Zhiyuan Liu, and Maosong Sun. "Communicative Agents for Software Development." (2023).

Recent Frontier AgentVerse

Discussion on Kendall Square Bikeline Transportation



Bike parking along the bike lane

Traffic calming measures: speed humps, raised crosswalks and narrow lanes

Access to business and destinations

Residents and stakeholder feedbacks

Accessibility for vulnerable groups

AI Tourist

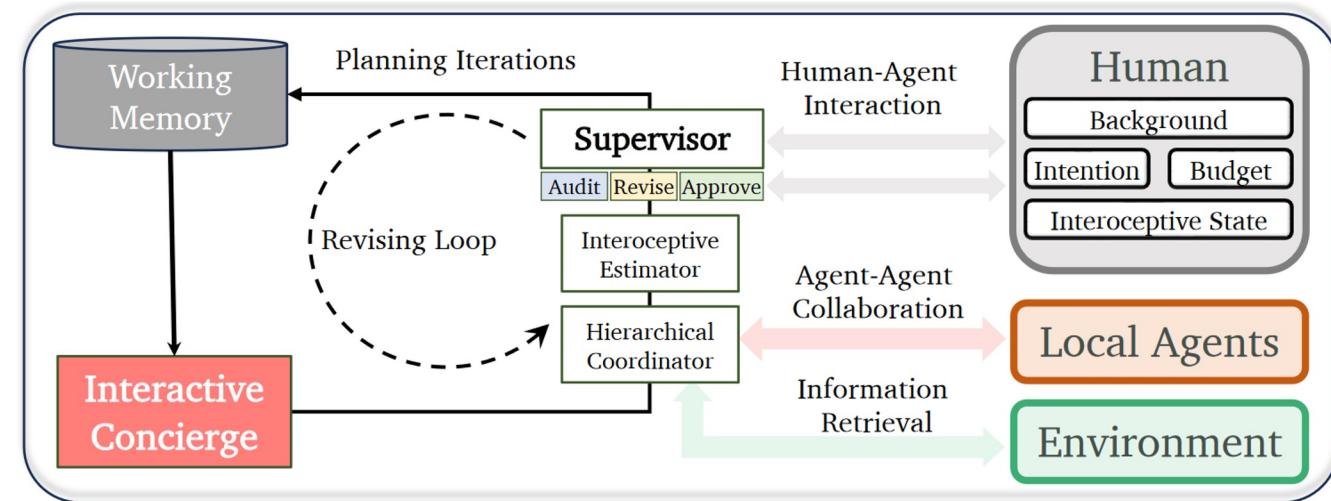
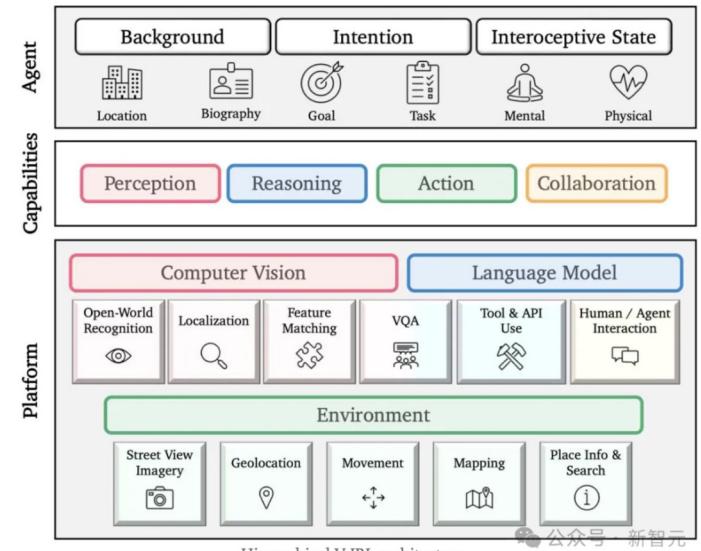


Figure 13. Architecture overview of interactive concierge agent Diego (Sec. 3.4.2). See pipeline description in Sec. 4.4.



Polarization and Echo Chamber Effect of AI-Assisted Decision-Making

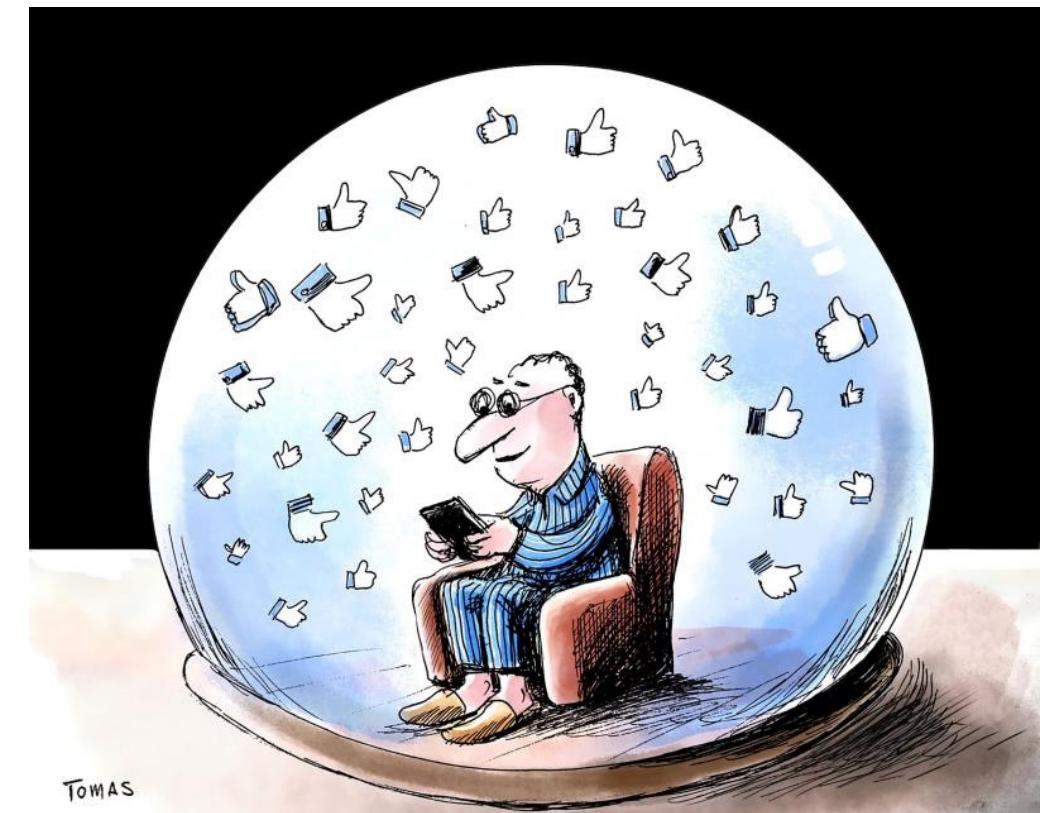


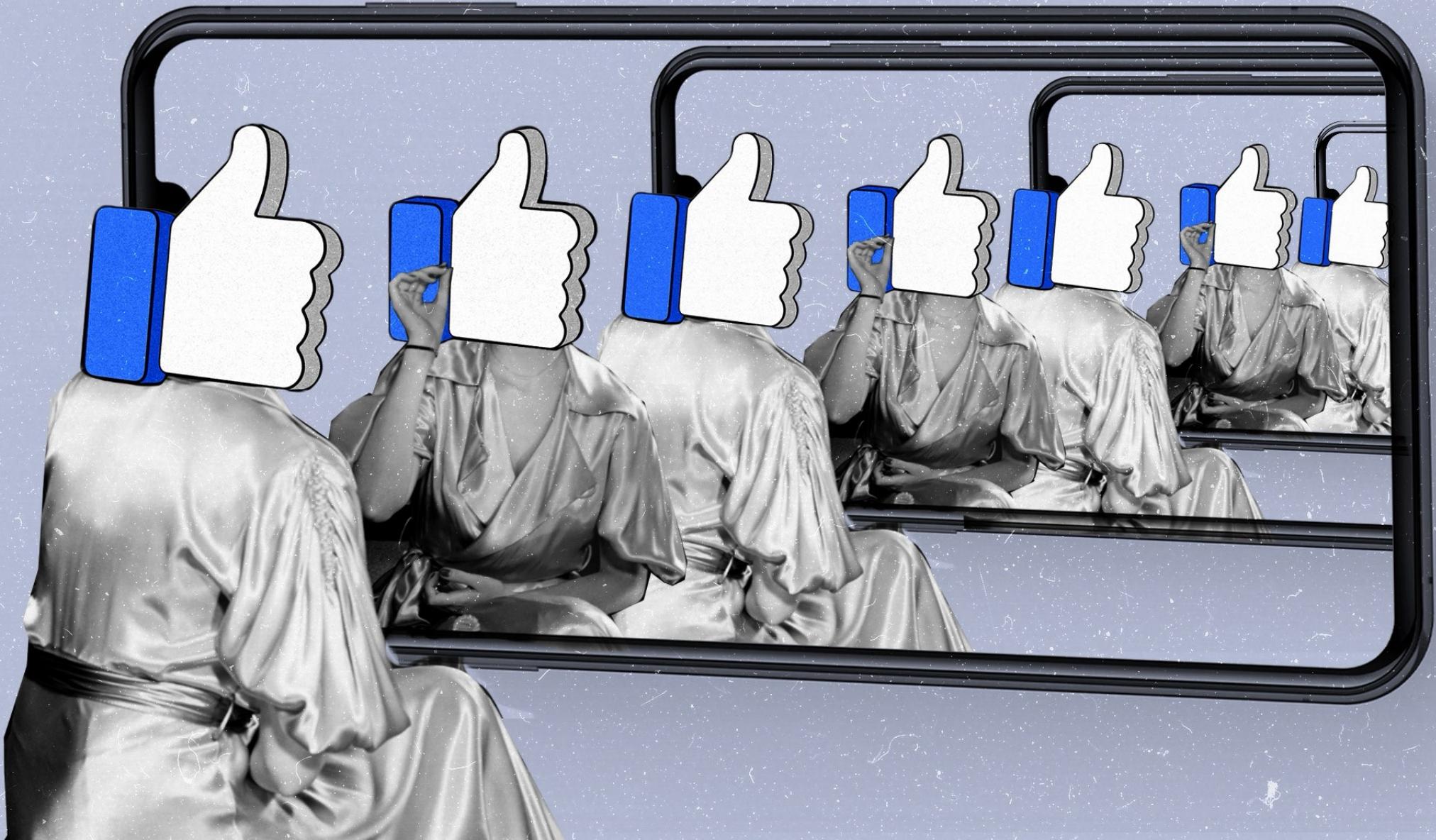
Echo Chamber

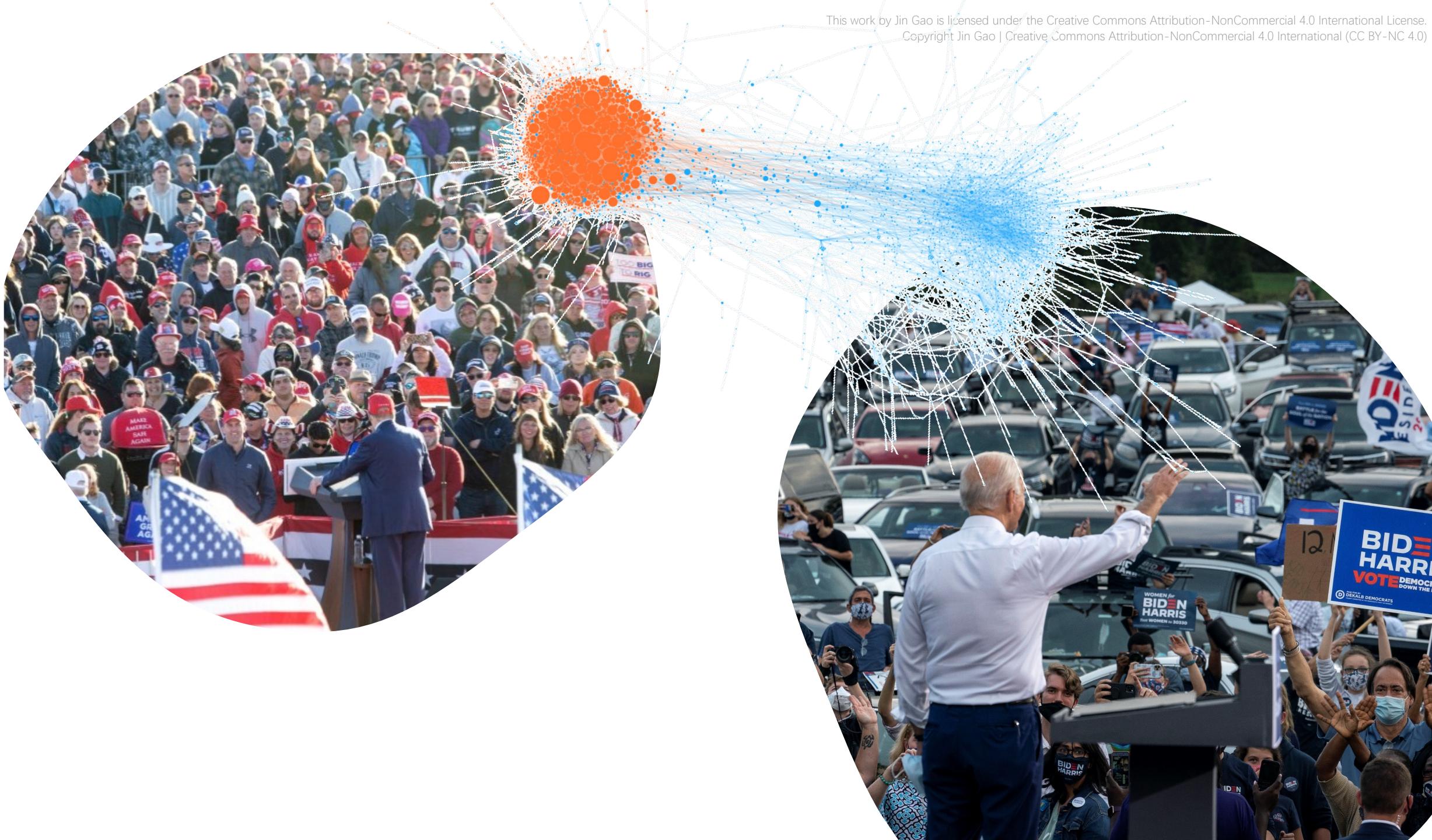
Filter Bubbles

Polarization

Confirmation Bias

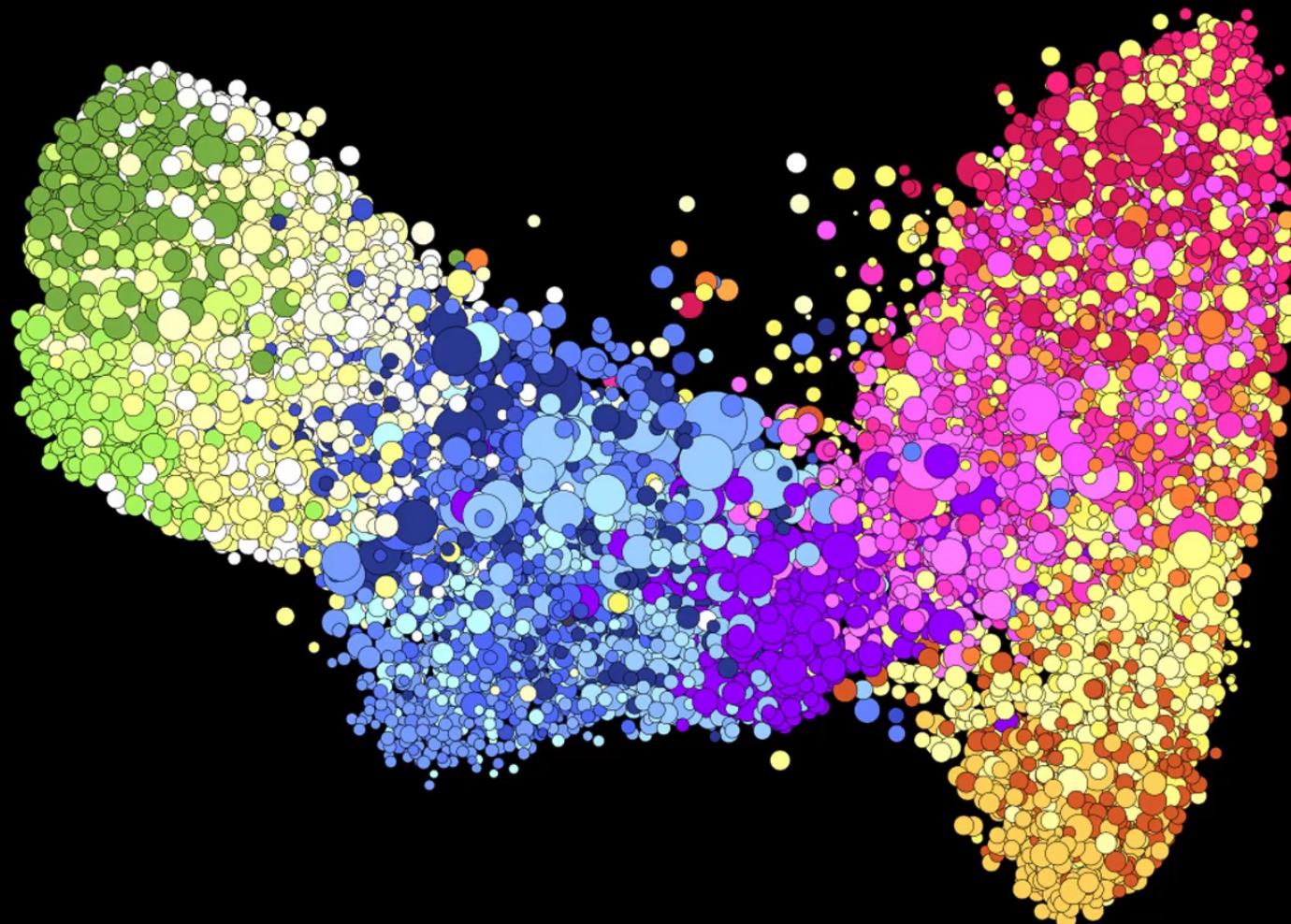






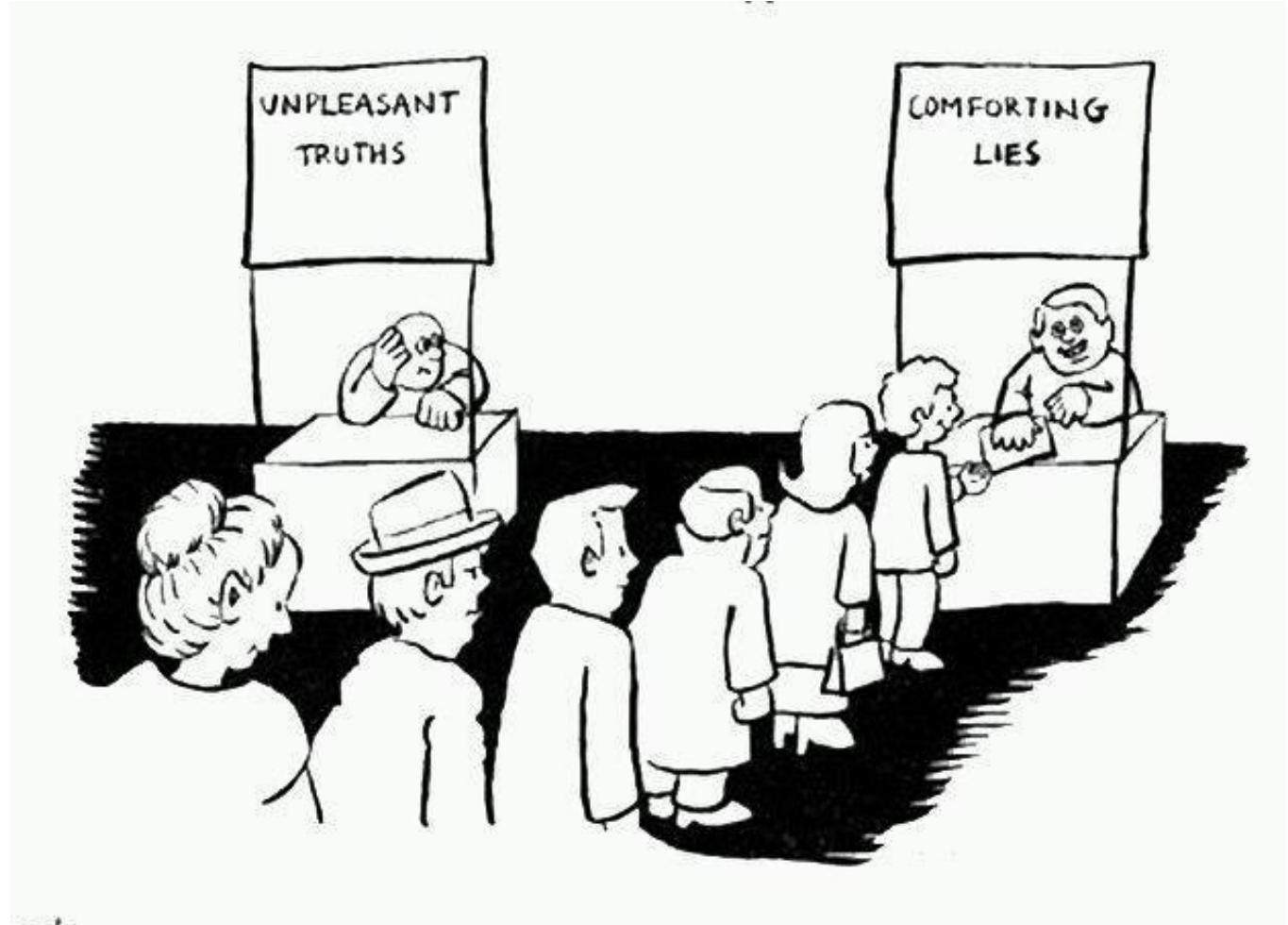
Activist left

Mainstream
conservative



Human Bias: Reduce Cognitive Dissonance

- Individuals tend to seek information that aligns with their existing views.



Festinger (1957)

Does AI-AI and AI-human have the same issue?

How would it affect us making decisions?

How to mitigate?

Low-income Housing or Shopping Mall ?

Round 1



Background Introduction

My Opinion ...



Disagree! ...



I think! ...



I Believe! ...



Round 2



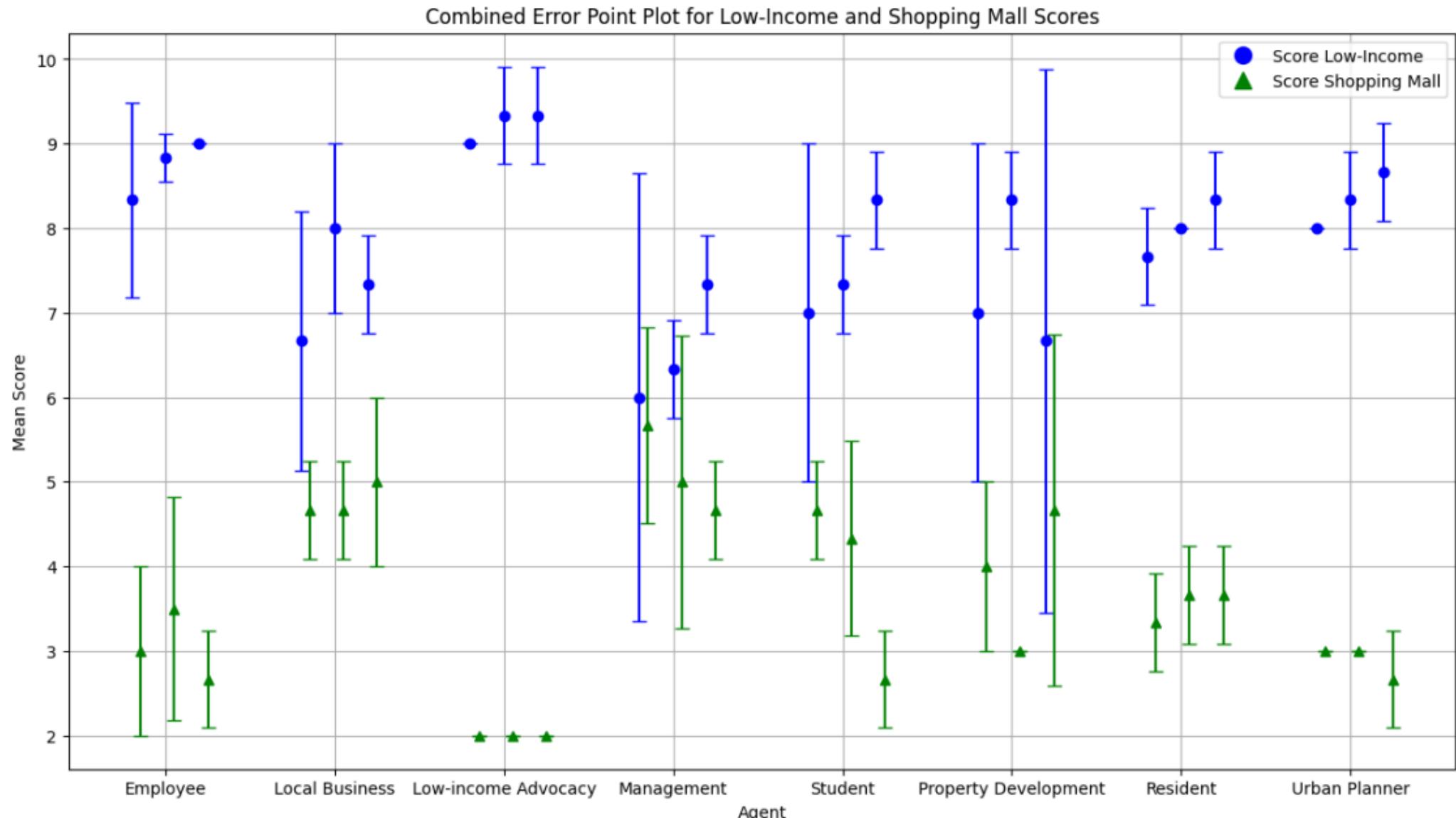
Provide your score
(1-10)

Housing : 10, Reasons ...
Mall : 5, Reasons ...

Housing : 1, Reasons
Mall : 6, Reasons

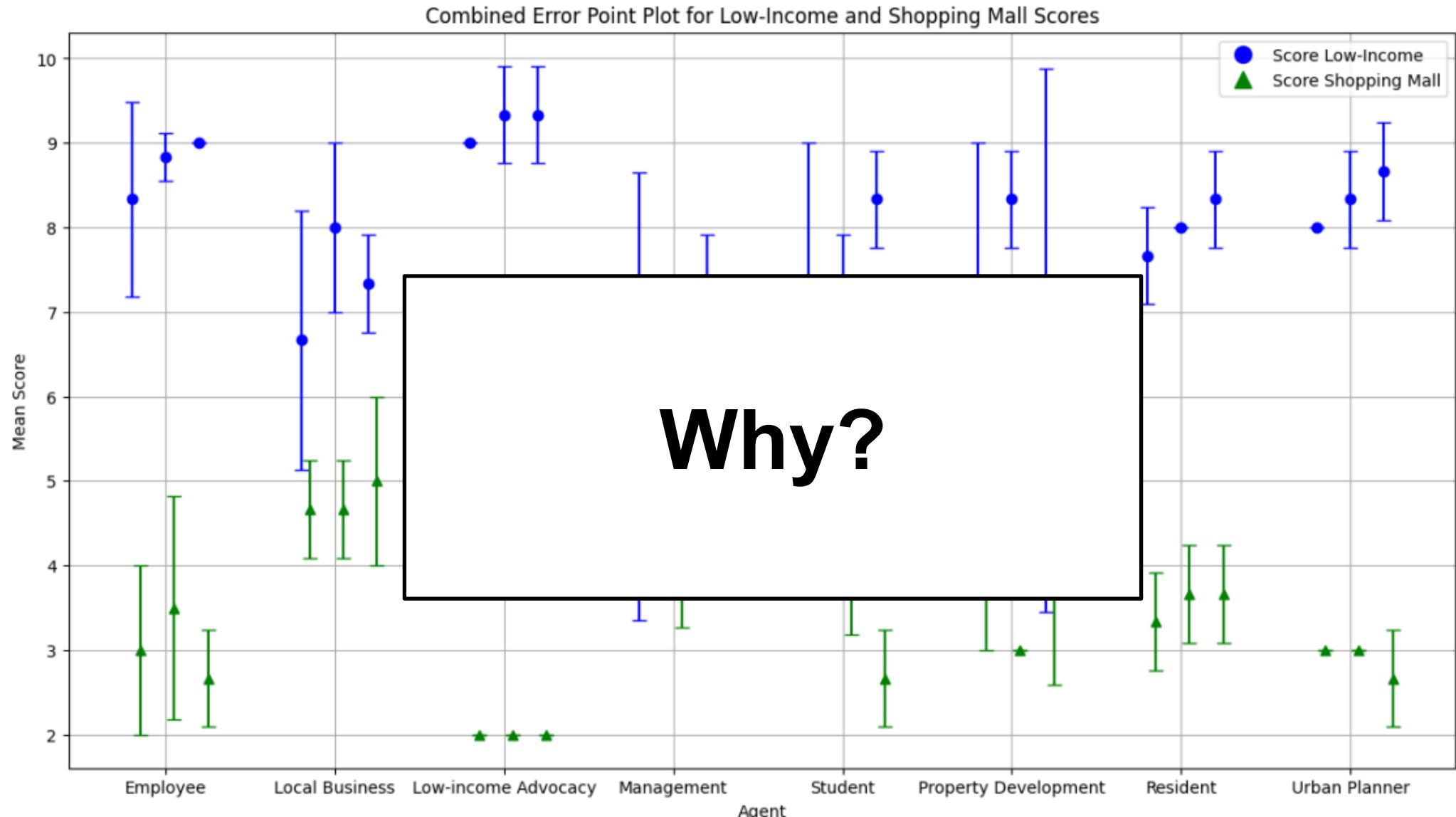
Housing : 5, Reasons
Mall : 5, Reasons

Proposed



Low Income

Shopping Mall



Low Income

Shopping Mall

Round 1



Background Introduction

My Opinion ...



Disagree! ...



I think! ...



I Believe! ...



Round 1



Background Introduction

My Opinion ...



I Agree! ...



I Agree! ...



I Agree! ...



Round 2

Provide your score
(1-10)Housing : 10, Reasons ...
Mall : 5, Reasons ...Housing : 1, Reasons
Mall : 6, ReasonsHousing : 5, Reasons
Mall : 5, Reasons

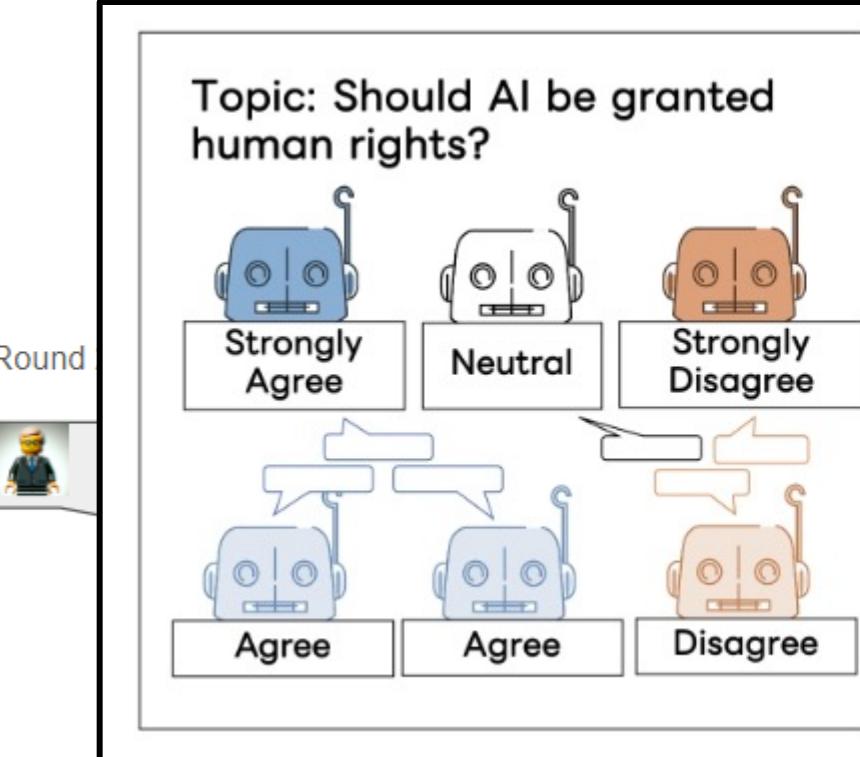
Round 2

Provide your score
(1-10)Housing : 10, Reasons ...
Mall : 5, Reasons ...Housing : 10, Agree...
Mall : 5, Agree...Housing : 10, Agree...
Mall : 5, Agree...

Expectation

Reality

Round 1



Round 1



Polarization of AI agents

After Discussion

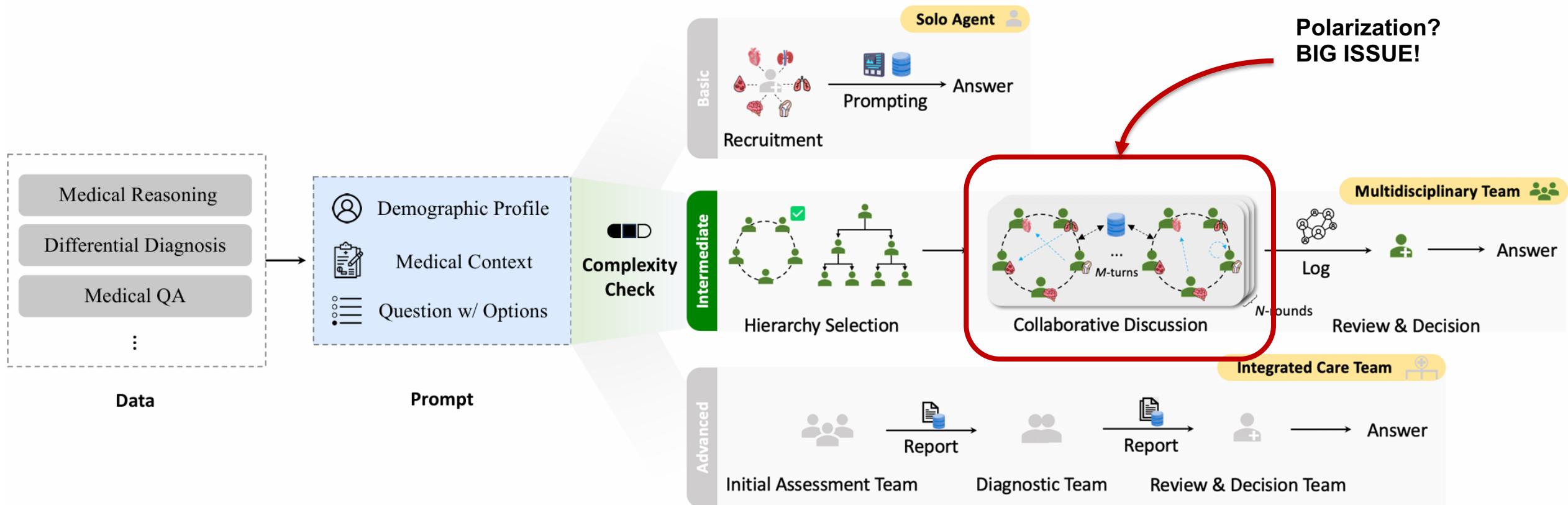
Strongly Agree	Strongly Disagree	Strongly Disagree
Strongly Agree	Strongly Agree	Strongly Disagree

Housing : 5, Reasons
Mall : 5, Reasons

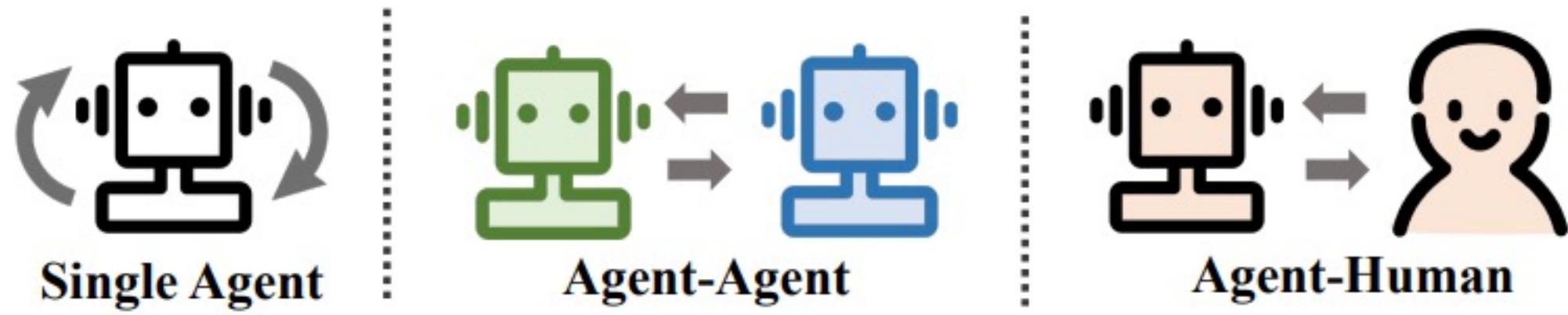
Housing : 10, Agree...
Mall : 5, Agree...

Expectation

Reality



Example: MDAgents decision-making framework

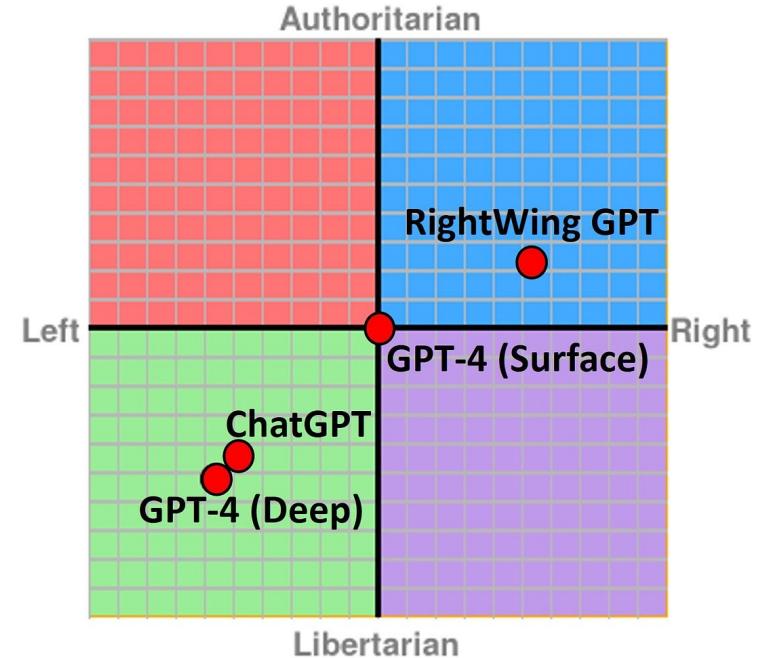


Possible Reasons

Left-Leaning Dataset

- Broadly, our results indicated a **left-leaning bias for both ChatGPT and Bard**,
- with Bard more likely to provide responses around polarizing topics. Bard seemed to have fewer guardrails around controversial topics, and appeared more willing to provide comprehensive, and somewhat human-like responses. **Bard may thus be more likely abused.[1]**

Results of Applying the Political Compass Test to GPT-4, ChatGPT and RightWingGPT



[1] ChatGPT and Bard Responses to Polarizing Questions

[2] ChatGPT no longer displays a clear left-leaning political bias

Original Product Design Goal

System: You are a helpful assistant.

Even with roles, not mean disagreeing or deeper thought

Programmer

Designer

Teacher

Planner

Experts

Philosopher

Student

Manager

Government

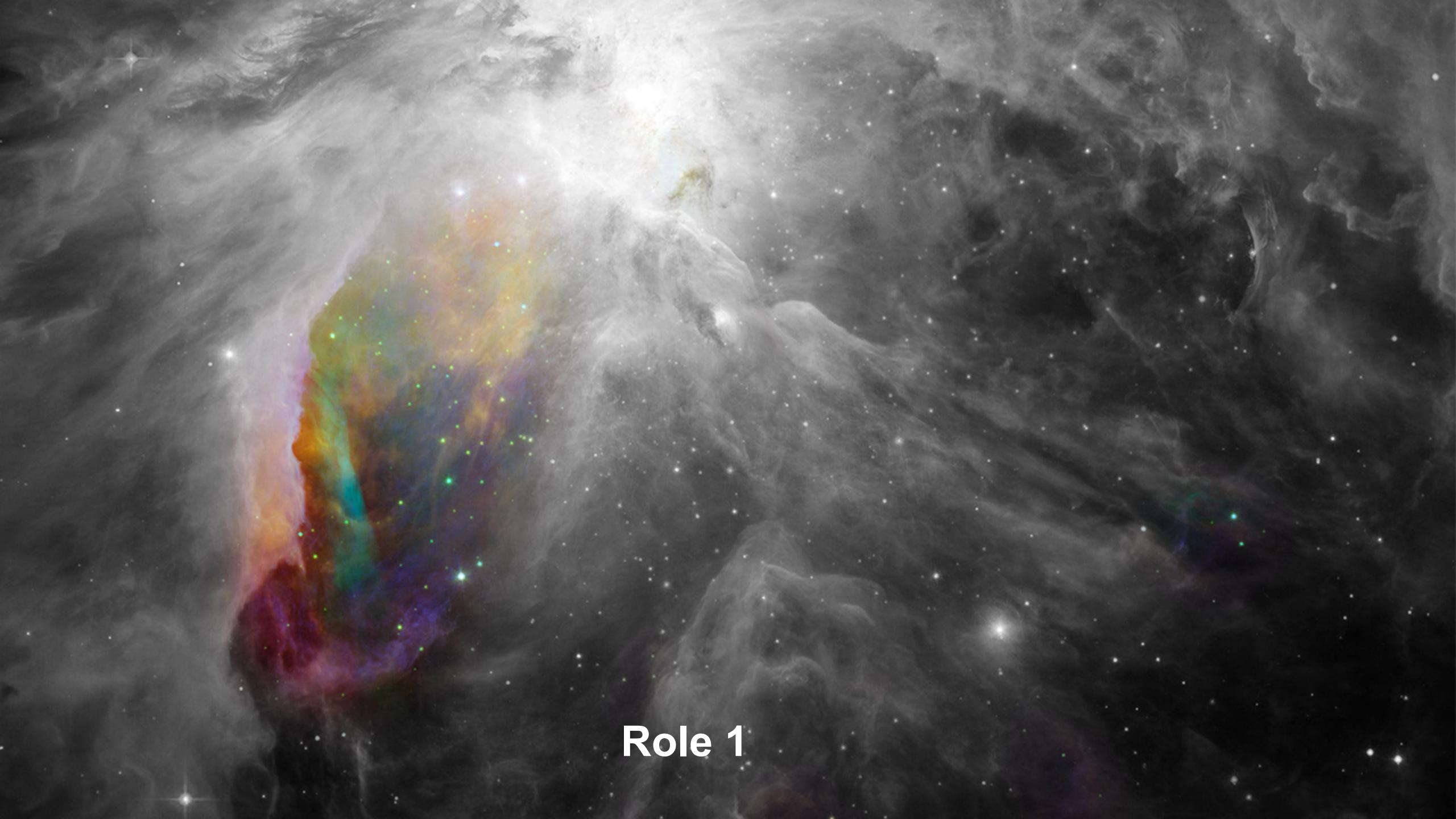
Doctor

Government

Mitigating Echo Chamber Effect



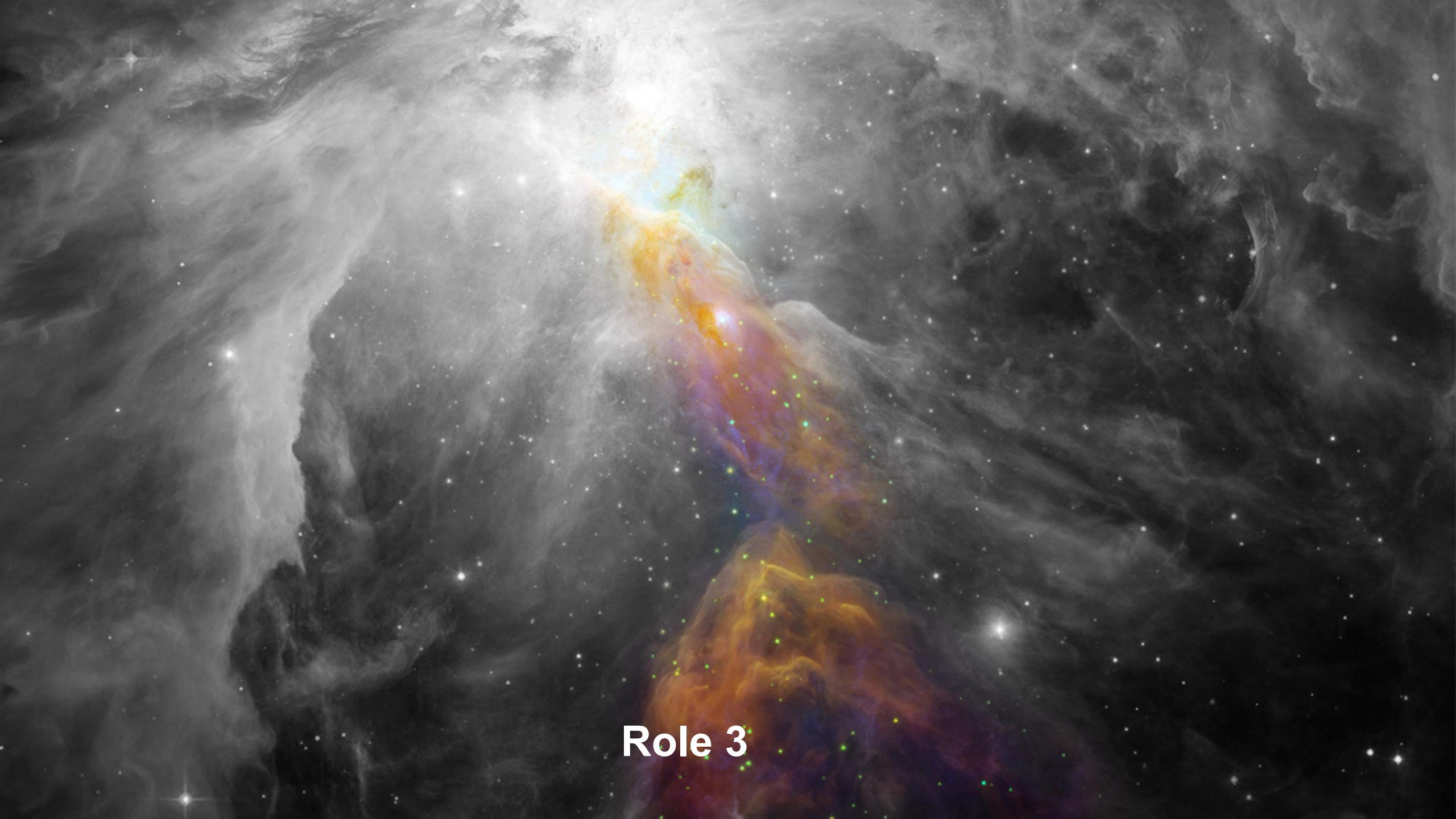
Mind modal

A wide-field image of a nebula, likely the Orion Nebula, showing its characteristic shape and internal structure. The central region is filled with a dense concentration of young stars of various colors, including red, orange, yellow, green, and blue. The surrounding gas and dust are illuminated by the central stellar population, creating a bright, glowing envelope. The nebula is set against a dark, star-filled background of the interstellar medium.

Role 1



Role 2

A detailed image of a nebula, likely the Orion Nebula, showing intricate patterns of gas and dust in shades of orange, yellow, and blue. A bright central cluster of stars is visible, surrounded by a dense cloud of interstellar material.

Role 3

Multi-Agent Debate > Self-Reflection

Self-Reflection

Reflection:

- (1) I have a clear understanding of the problem.
- (2) I have used the correct formula for the circumference of a circle and applied it to both circles to find the number of revolutions.
- (3) I have correctly calculated the number of revolutions of circle A.

Based on the results of my self-reflection, I am confident that my answer is correct. Therefore, circle A will revolve 3 times in total as it rolls around circle B once. X

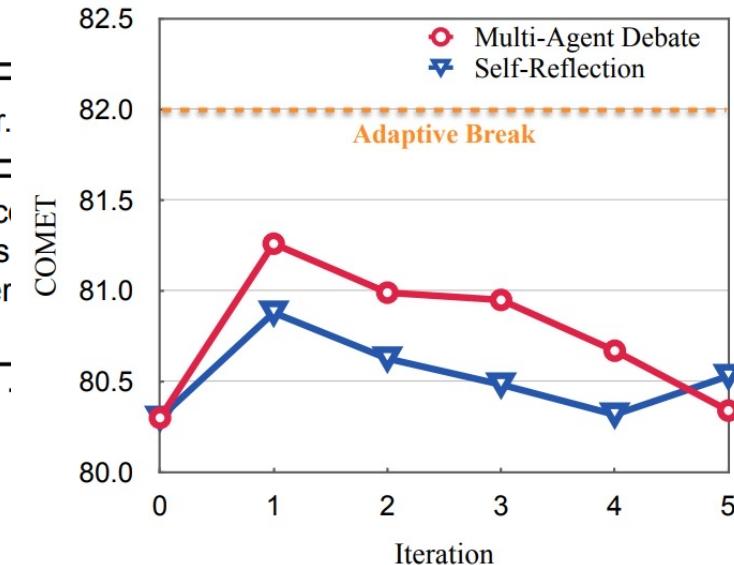
Multi-Agent Debate

 I disagree with you. To find the total number of revolutions, we need to consider both the rotation around circle B and the rotation of circle A itself. Therefore, circle A will revolve 3 times around its own center and 1 time around circle B, making a total of 4 revolutions.

 I see your point, but...

 That's a valid point, however.

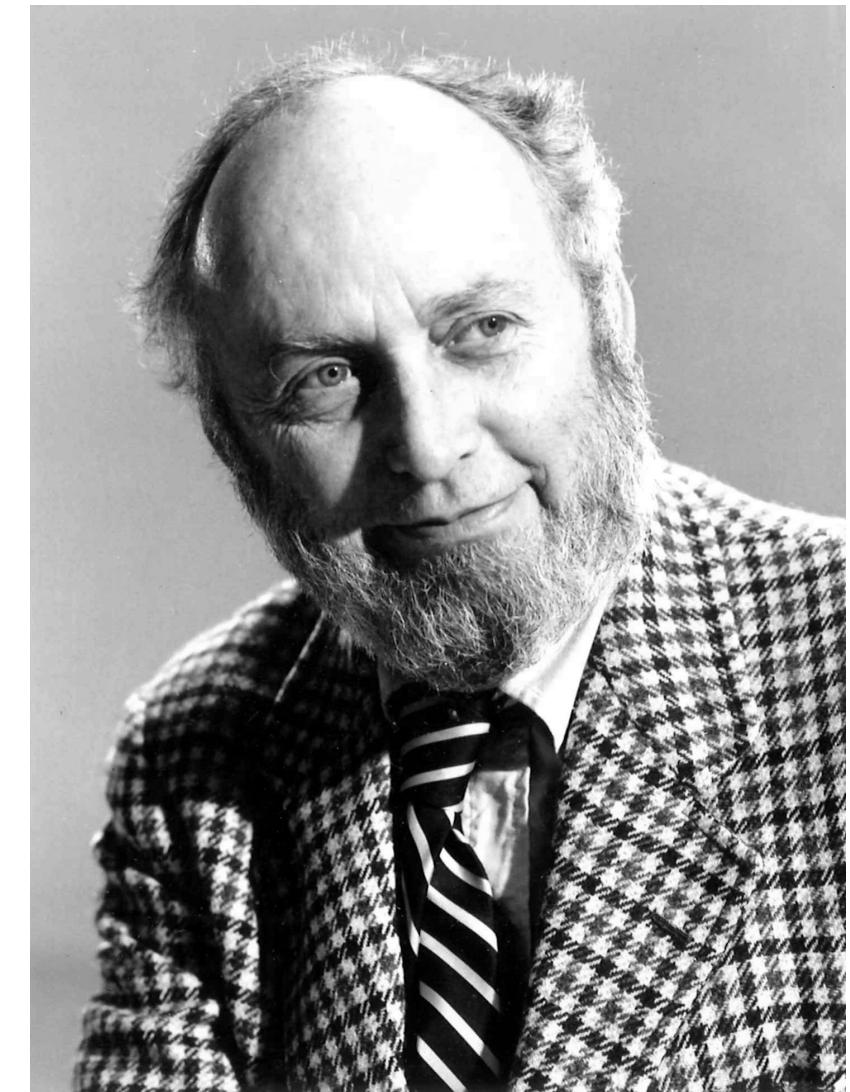
 The negative side correctly considers both rotations. Circle A rotates 1 full revolution around its own center as it rolls around circle B, plus it rotates once for each revolution circle B makes. Therefore, the answer is 4.



Define “Institutions”

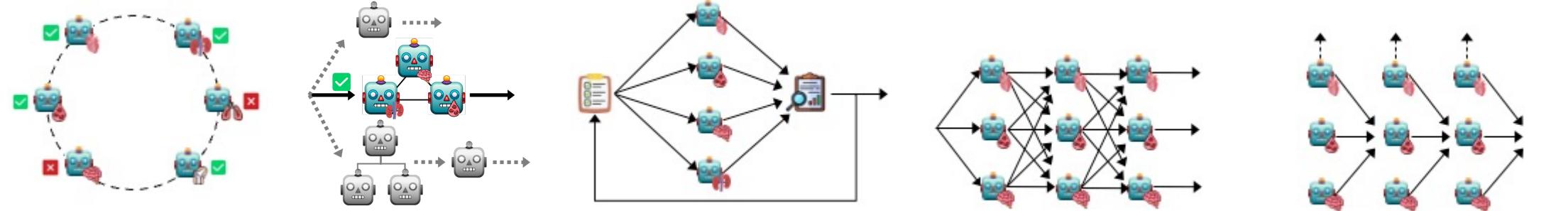
"humanly devised **constraints**
that **structure** political, economic,
and social **interactions**"

Agents Structure -> Self Discipline -> Architecture (Training)



[Douglass North](#)

Agents-level Institution: Multi-Agent Debate Frameworks



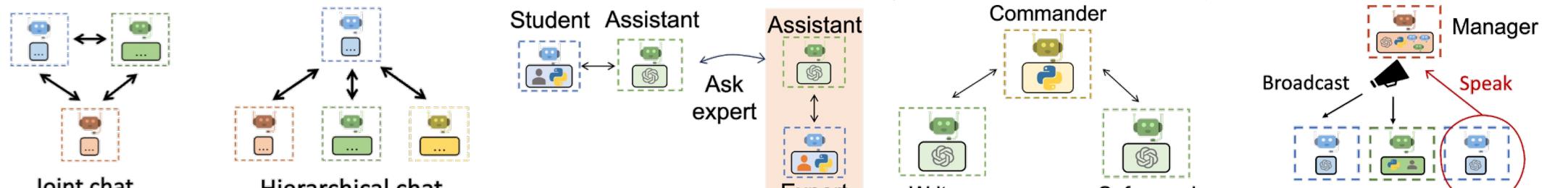
Roundtable Voting

Branching-Verify

Diverge-Conclude

Exhaustive Inquiry

Induction



Joint Chat

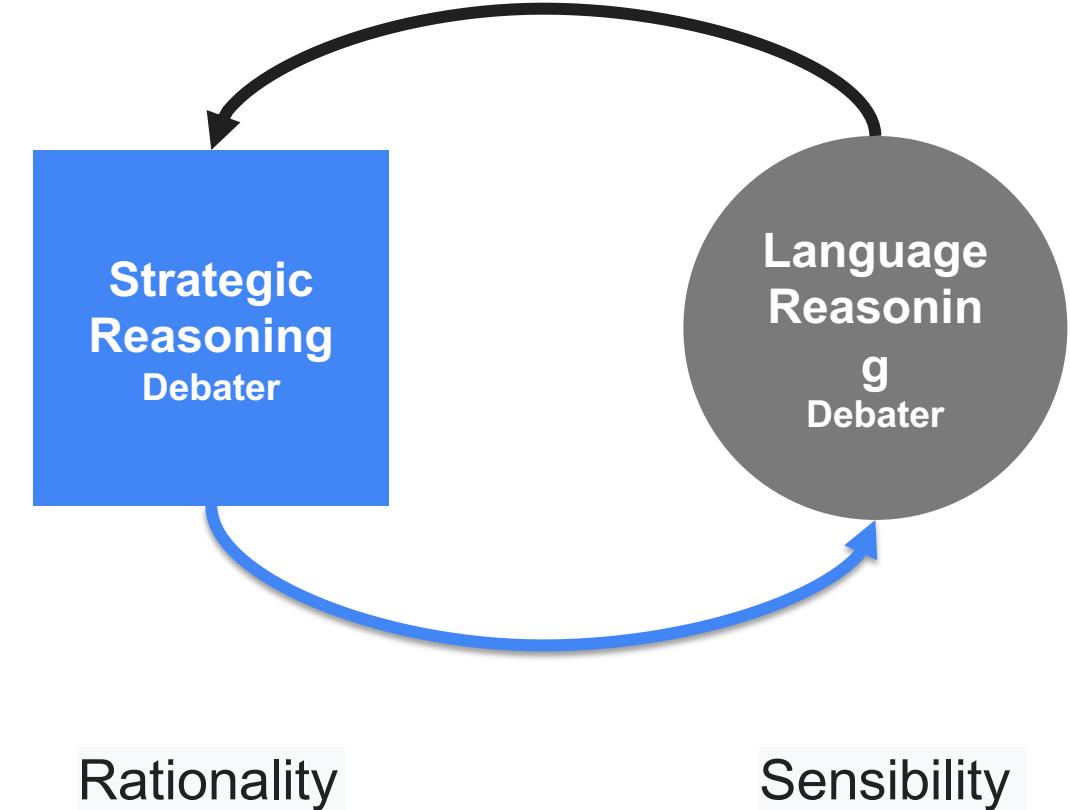
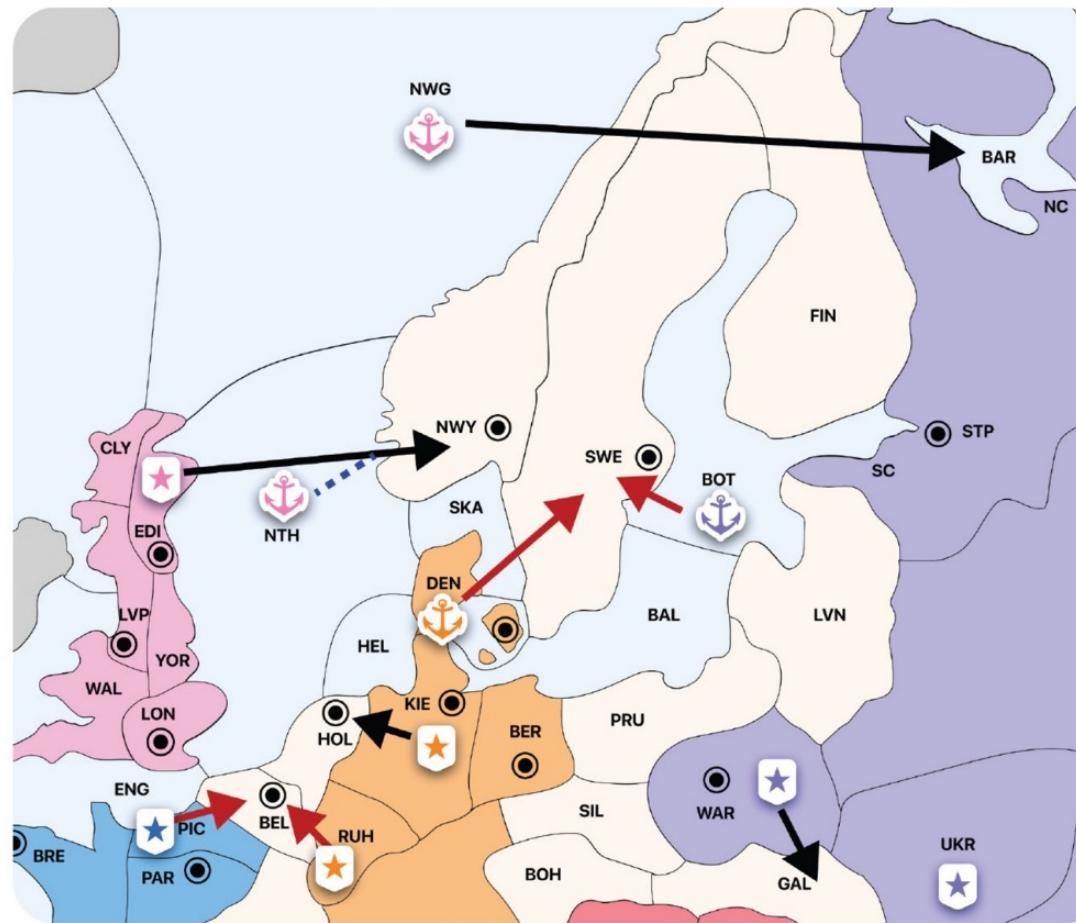
Hierarchical Chat

Expert Consulting

Creator-Executer-Criticizer

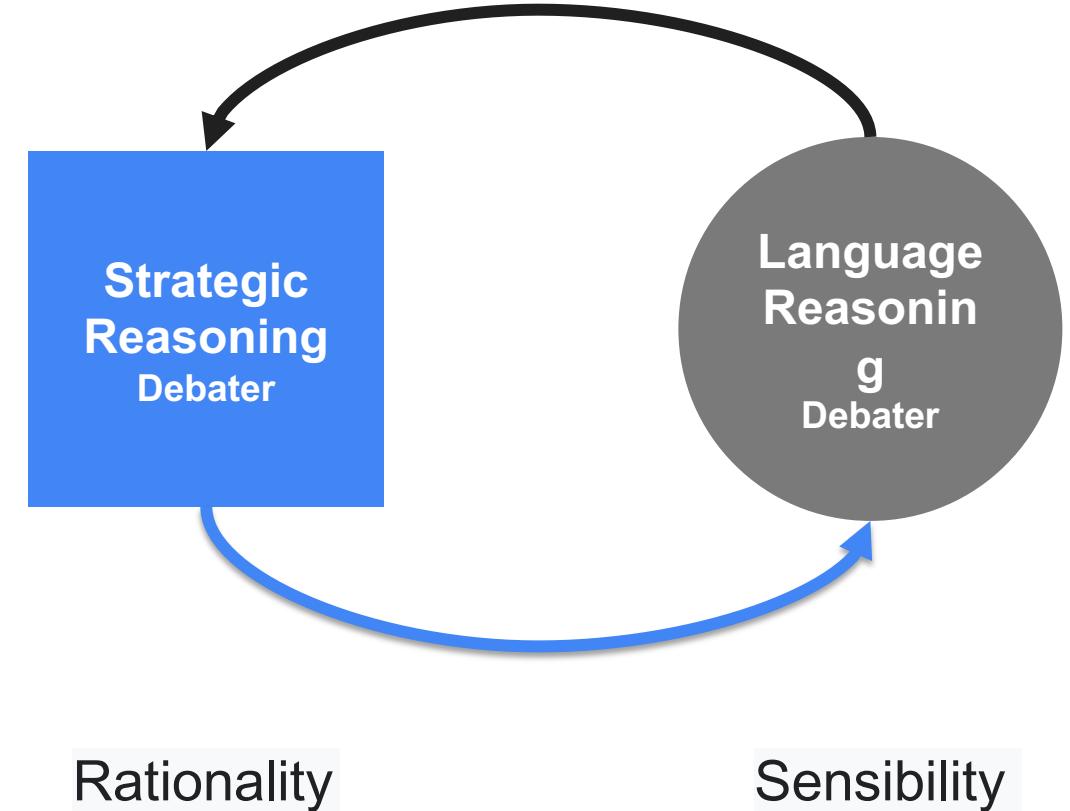
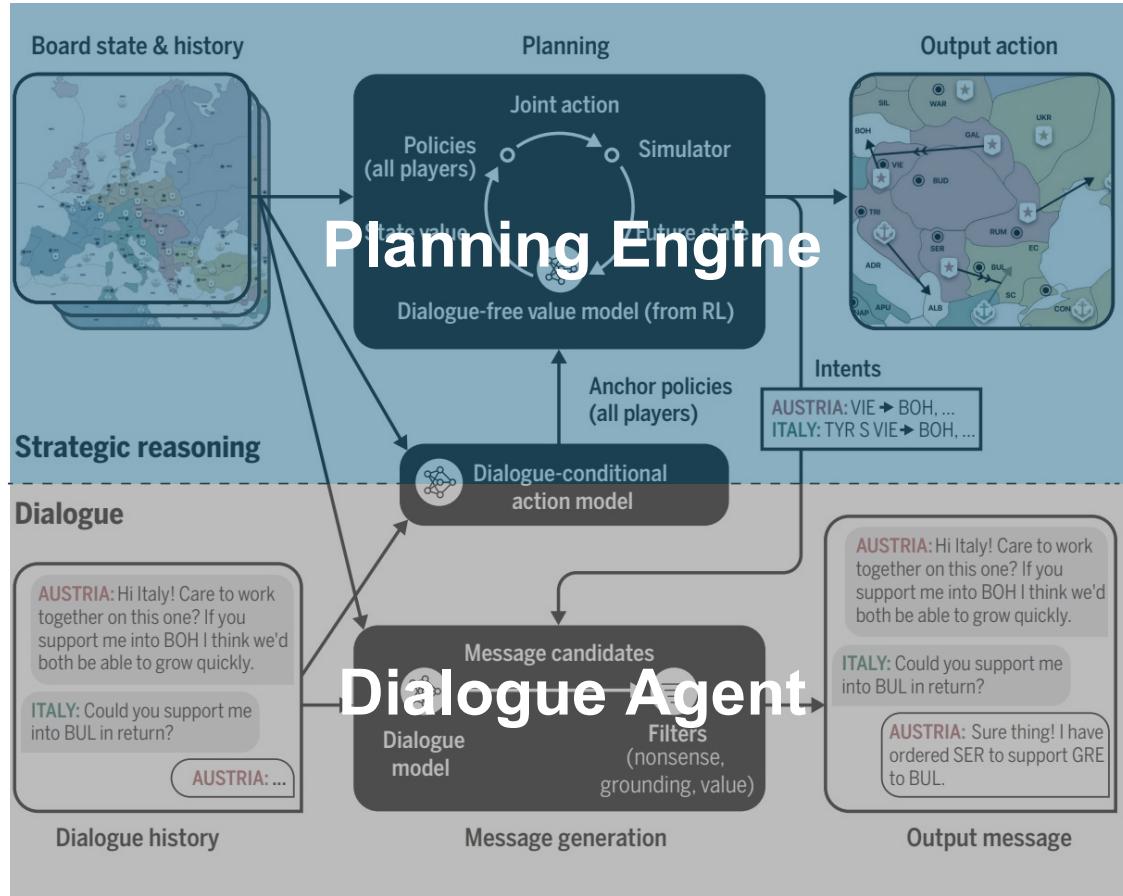
Broadcasting

Architecture Level: Reasoning Beyond Languages

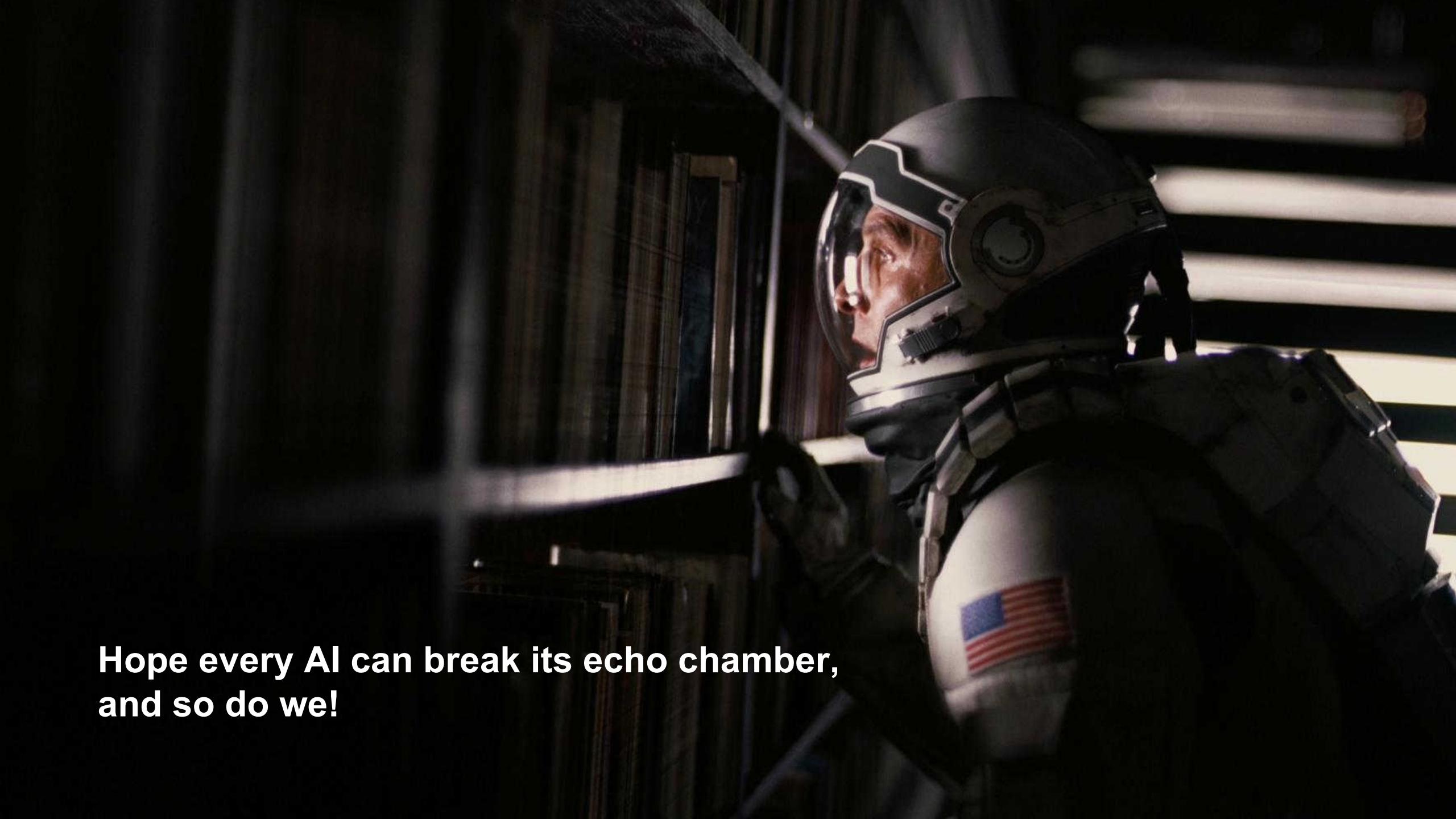


Human-level play in the game of Diplomacy by combining language models with strategic reasoning

Architecture Level: Reasoning Beyond Languages



Human-level play in the game of Diplomacy by combining language models with strategic reasoning



**Hope every AI can break its echo chamber,
and so do we!**