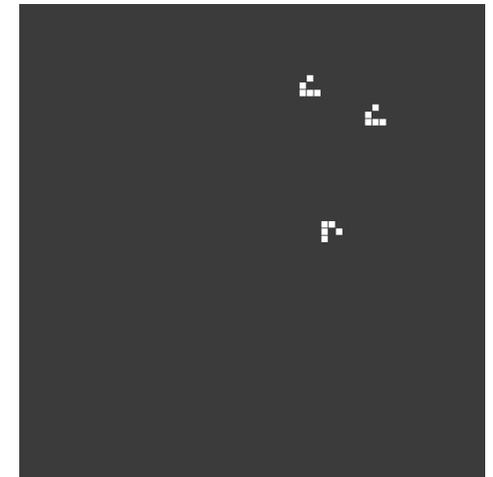


# Collective Intelligence in LLM Agents

# What is Agency?

Takashi Ikegami

Within artificial life the concept of an *agent* is fundamental. While studying life-as-it-could-be (Langton, 1989), we also study agents-as-they-could-be. An intuitive approach to agents is possibly to say that while not reproducing, i.e. during their individual lifetime, living organisms are agents. The concept of an agent in this way generalizes the concept of living organisms by de-emphasizing reproduction and with it Darwinian evolution. This point of view is also in line with the common practice of referring to robots or software programs as agents.



Towards information based spatiotemporal patterns as a foundation for agent representation in dynamical systems

Martin Biehl<sup>1</sup>, Takashi Ikegami<sup>2</sup> and Daniel Polani<sup>1</sup> [arXiv:1605.05676](https://arxiv.org/abs/1605.05676)



**Cells at Work! — your body as a society of agents**

# What is Agency?

## 1. **Autonomous**

Able to maintain and regulate itself, not just pushed around by the environment.

## 2. **Goal-directed**

Acts in ways that tend to achieve outcomes — purpose-like behavior, even without conscious intent.

## 3. **Adaptive (not fully predictable)**

Partly lawful and rational, but also shaped by history, context, and individuality.

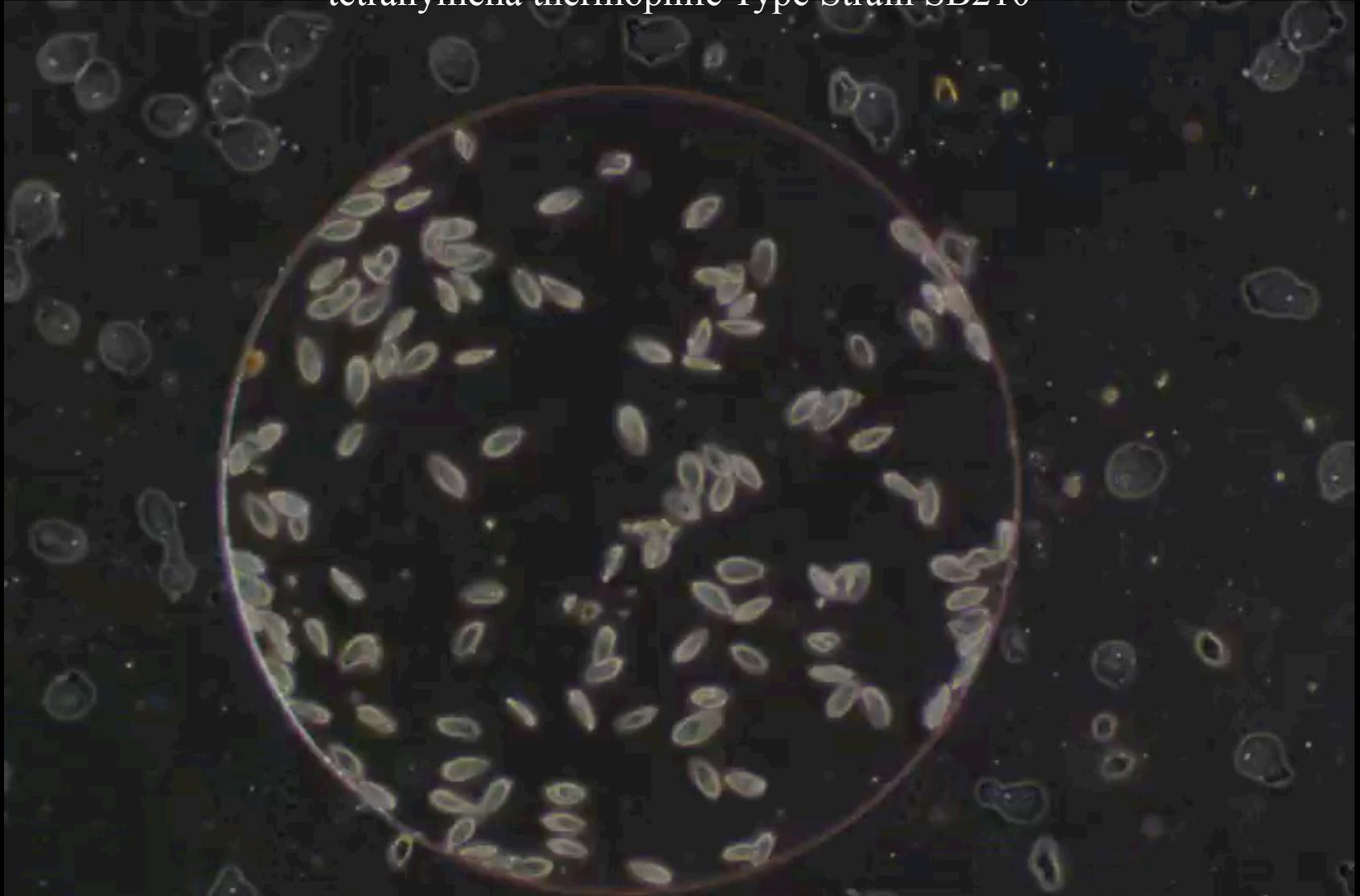
A small nerdy refinement: “idiosyncratic” is perfect, but I’d frame it as “history-dependent” or “context-sensitive” so it sounds scientific rather than whimsical.

**Agency is when a system is not just moving, but steering.**

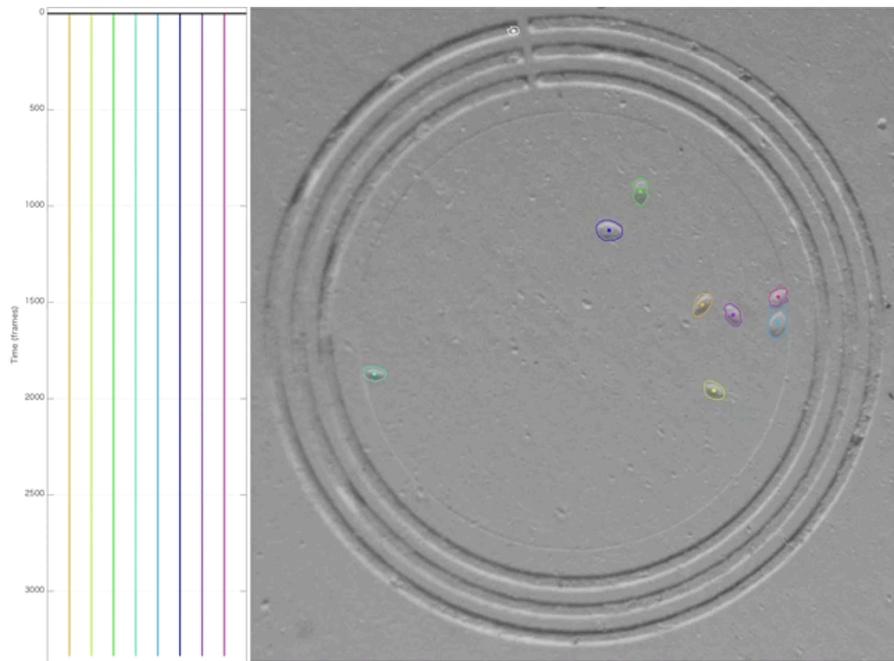
If agency is **autonomy** + **goal-directedness** + **history-dependence**, then even a single cell can qualify—because it can *carry information across time*.

Tetrahymena isn't just chemistry happening—it's an agent that keeps itself going. It *inherits* internal organization across generations: not only DNA, but cellular state—structures, patterns, and 'settings' that persist and shape future behavior. That persistence is a minimal form of memory. Memory is what makes 'goal-directed' behavior possible over time.

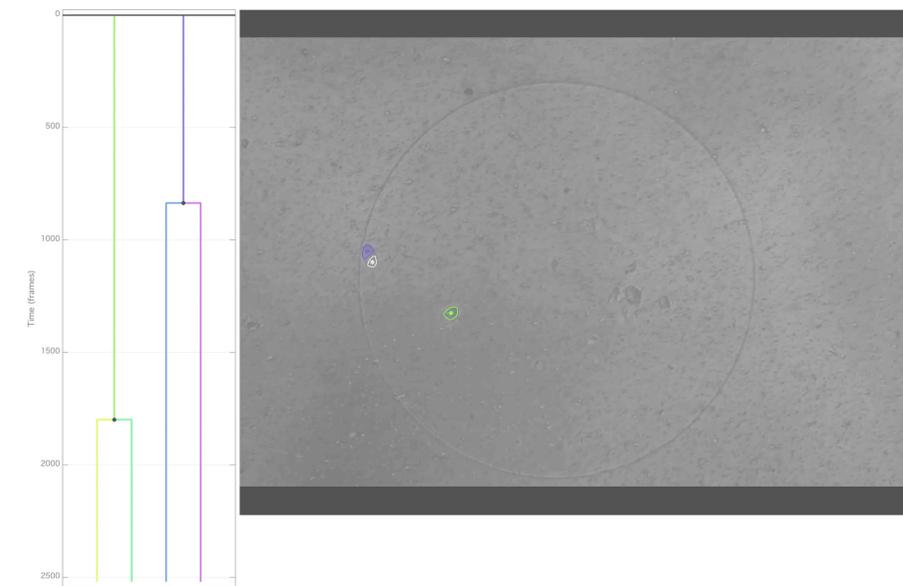
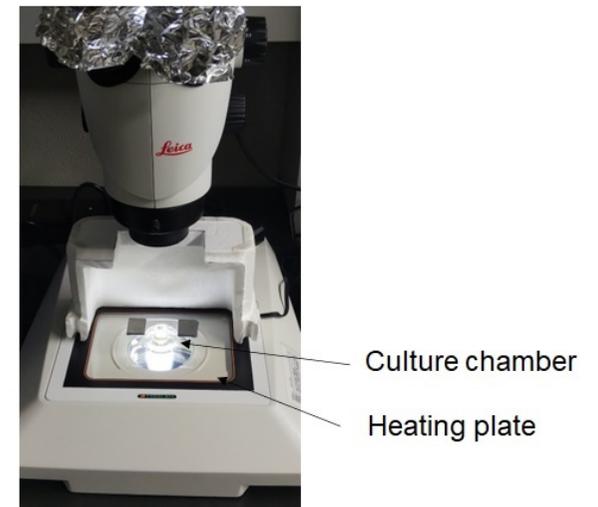
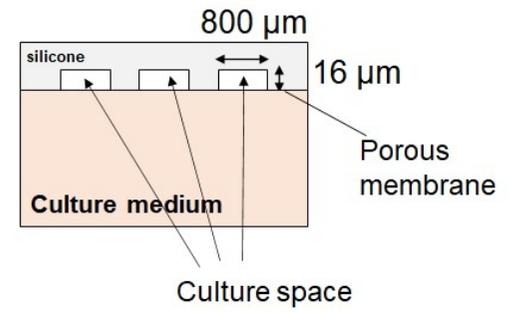
tetrahymena thermophile Type Strain SB210



Hiroki Kojima, Akiko Kashiwagi and Takashi Ikegami: **Revealing Gene Expression Heterogeneity in a Clonal Population of Tetrahymena thermophila through Single-Cell RNA Sequencing.** *Biochemistry and Biophysics Reports*, 38, 101720 (2024)

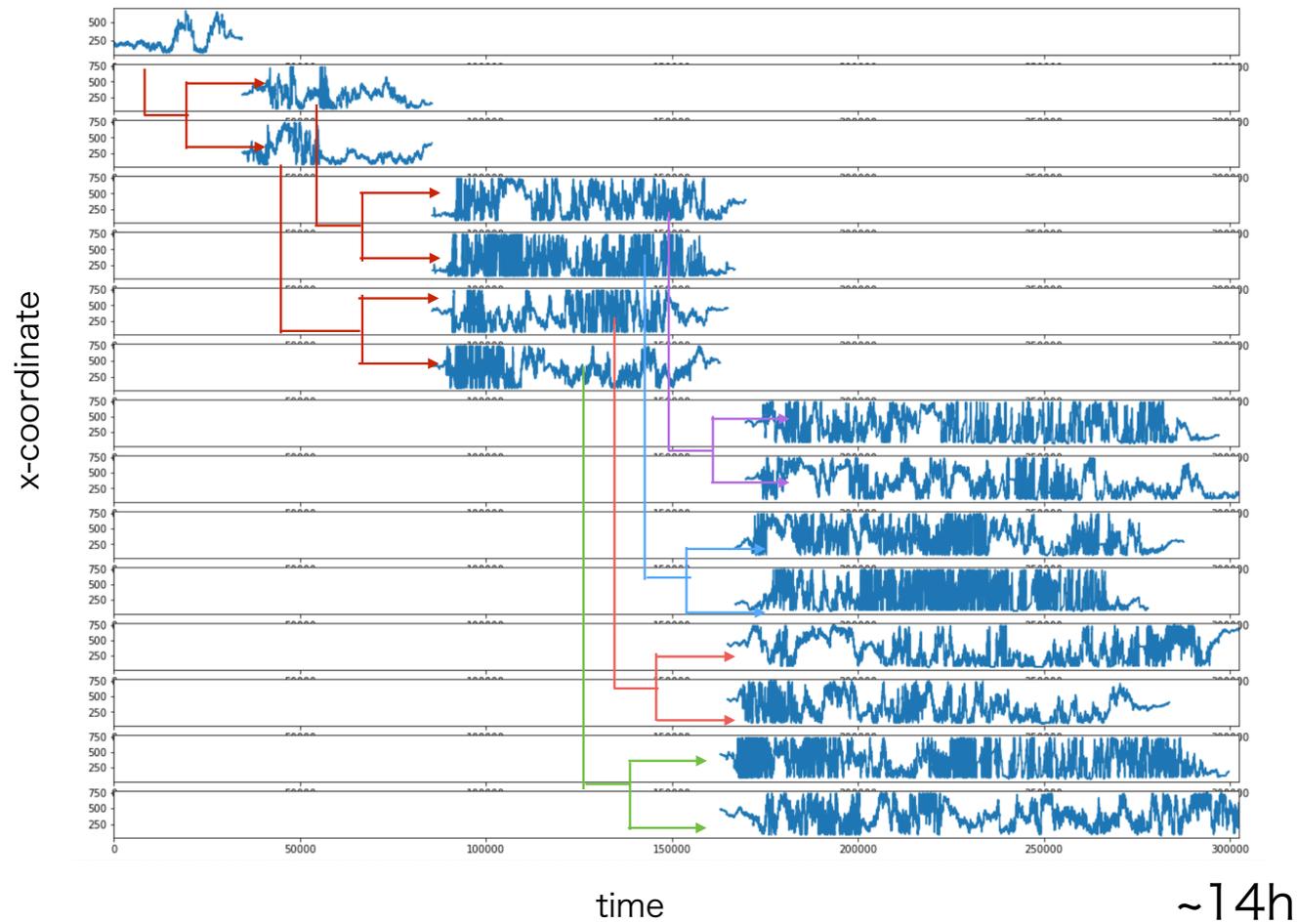


stereomicroscope



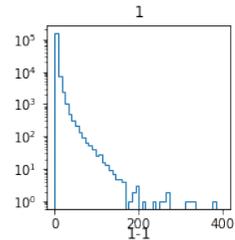
thanks to Dr. Akiko Kashiwagi,  
Dr. Hiroaki Suzuki

Utilizing the time series data from each individual cell, we can calculate the kinetic energy. This information then enables us to construct a phylogenetic tree starting from a single cell.



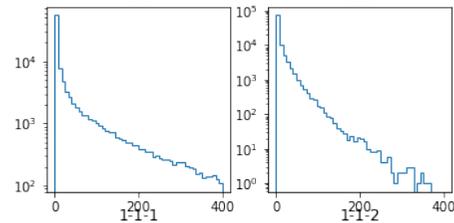
# KE distribution computed using the entire life time

Traj\_all210818.npy

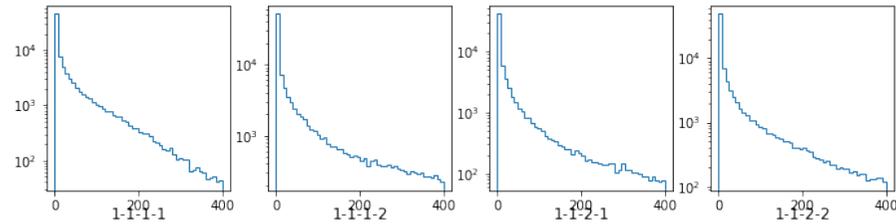


1st generation

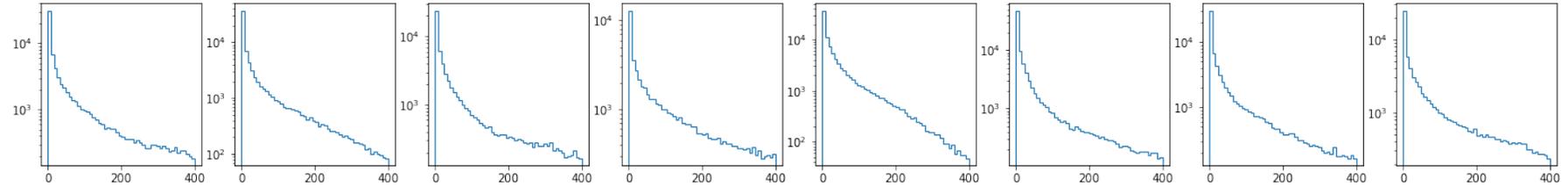
Can KE distribution be phenotype?  
Is this “phenotype” stably transmitted?  
Does cell-cell interaction affect inheritance?



2nd generation



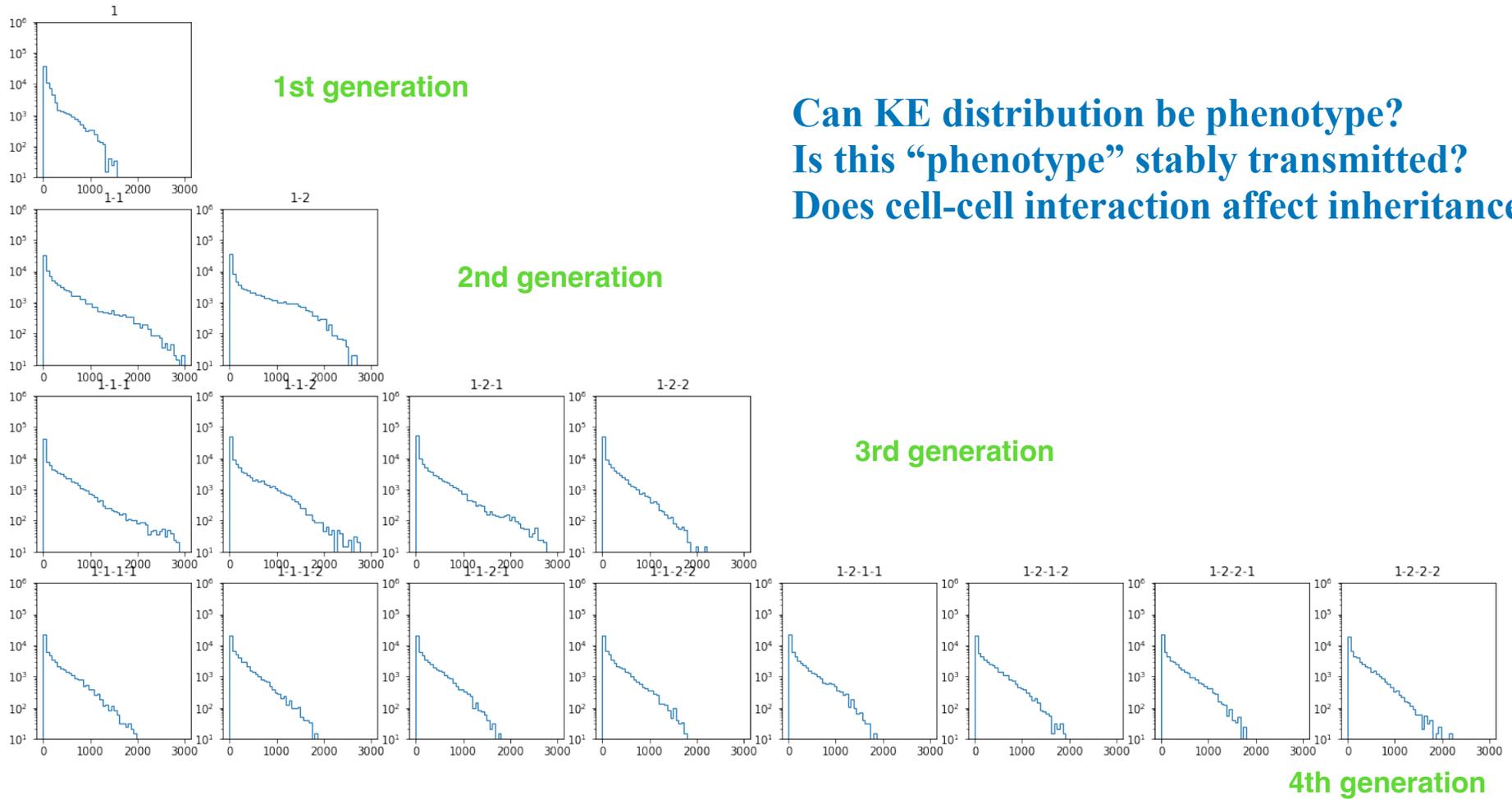
3rd generation



4th generation

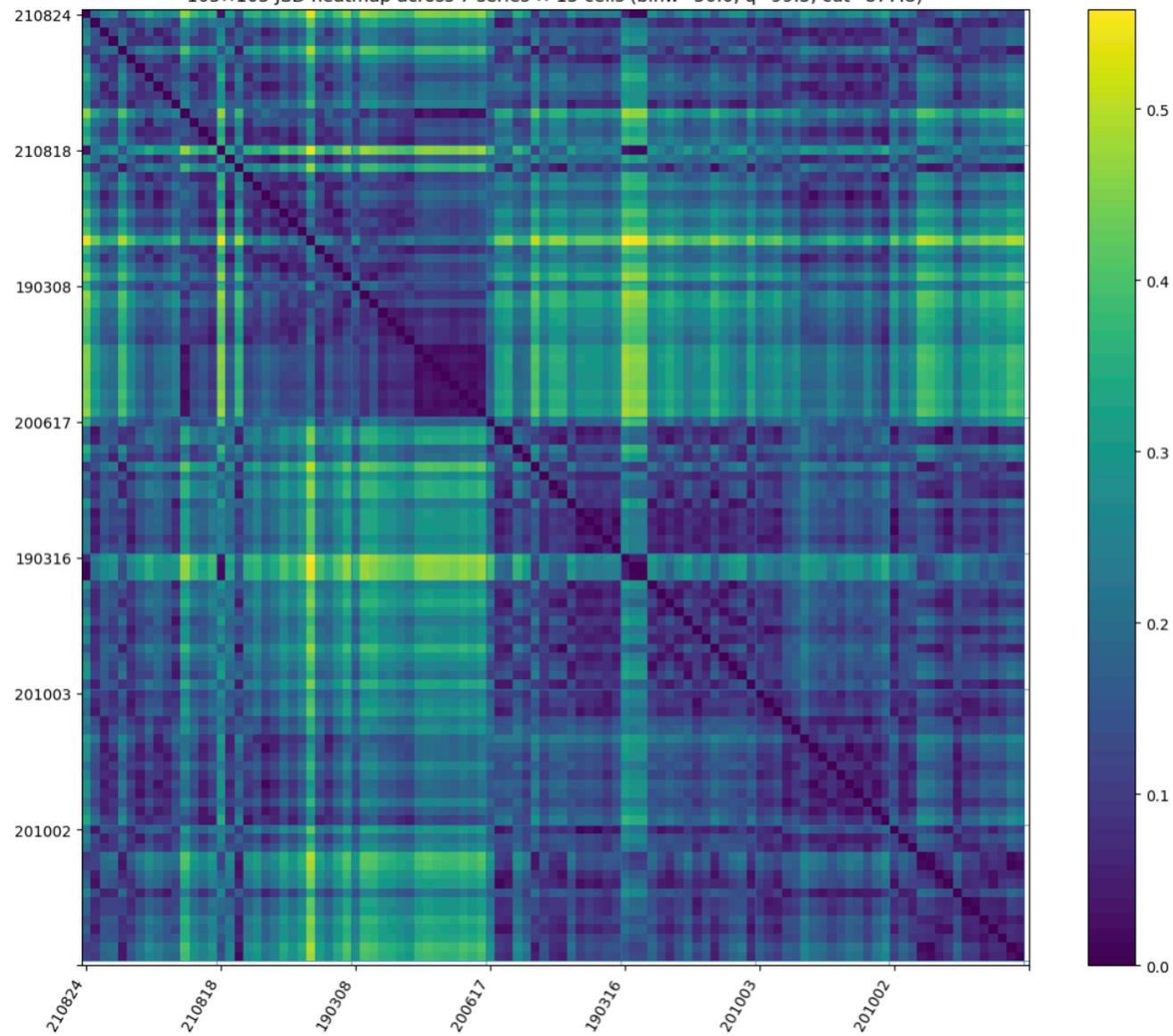
# KE distribution computed using the entire life time

Traj\_all190308.npy

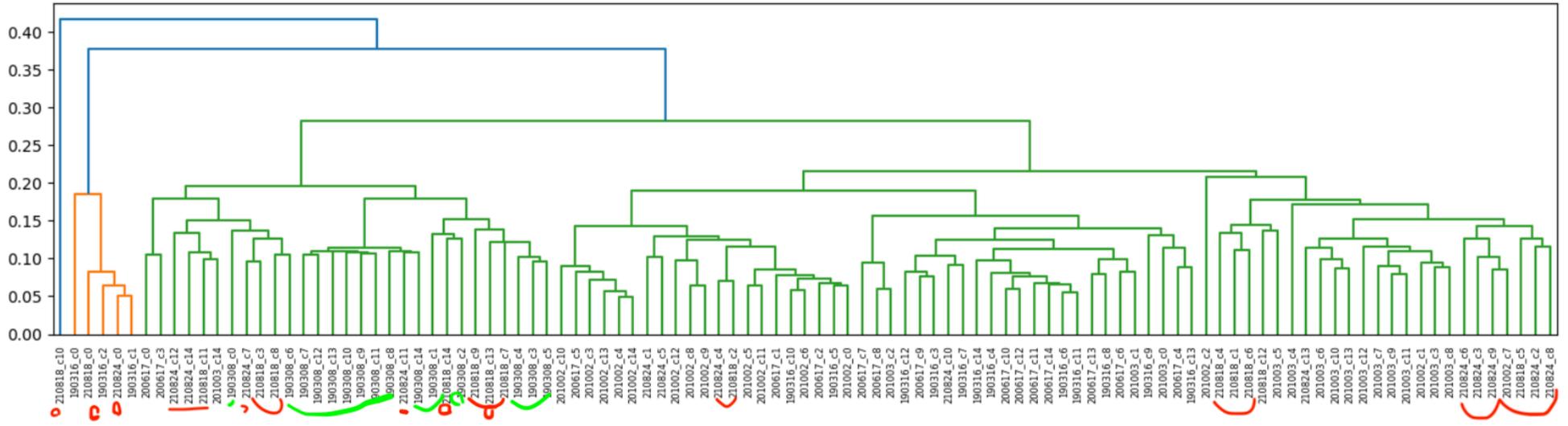


global cut = 877.7649065940483 nbins = 18

105x105 JSD heatmap across 7 series x 15 cells (binw=50.0, q=99.5, cut=877.8)



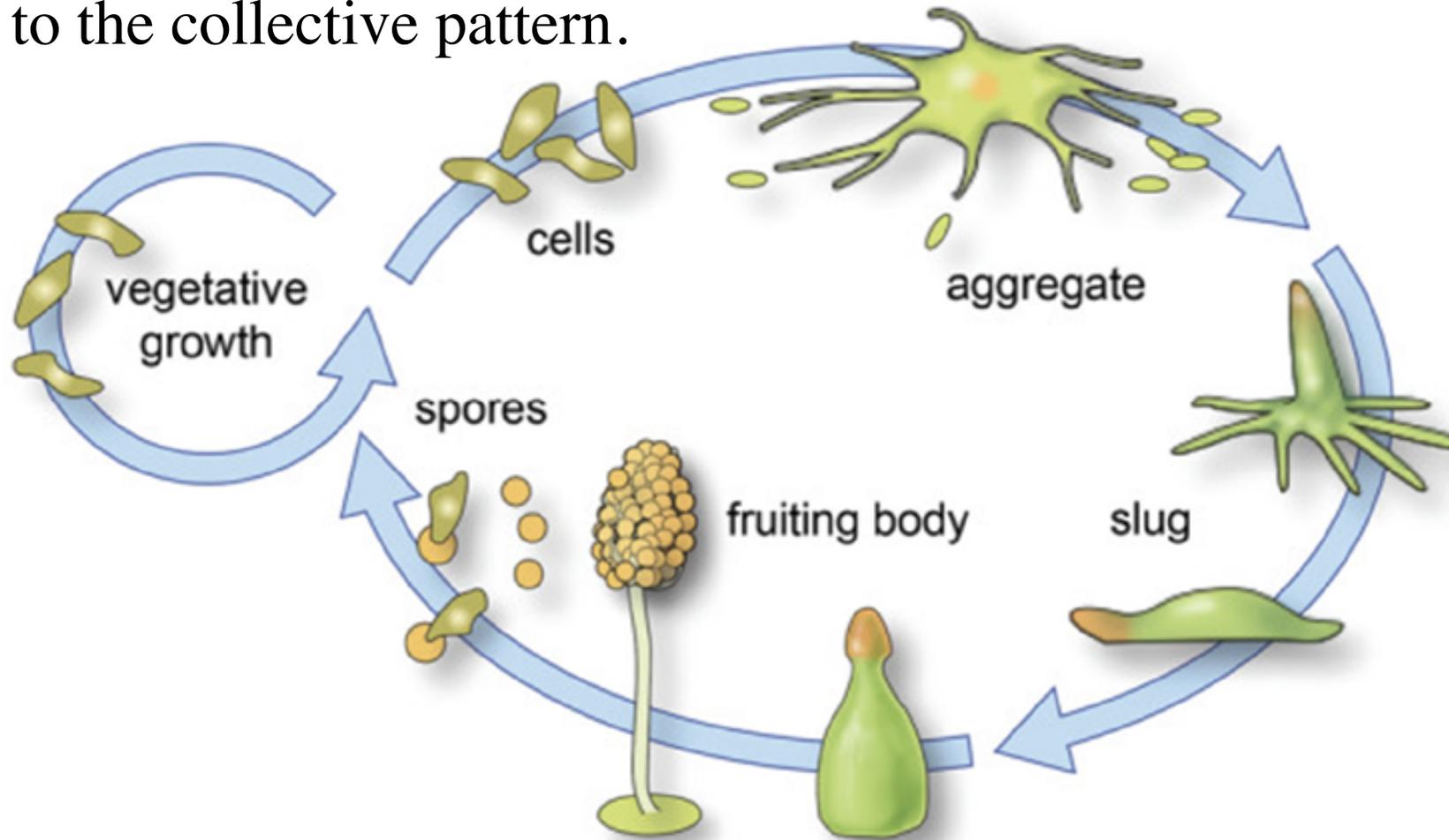
Hierarchical clustering dendrogram (average linkage)



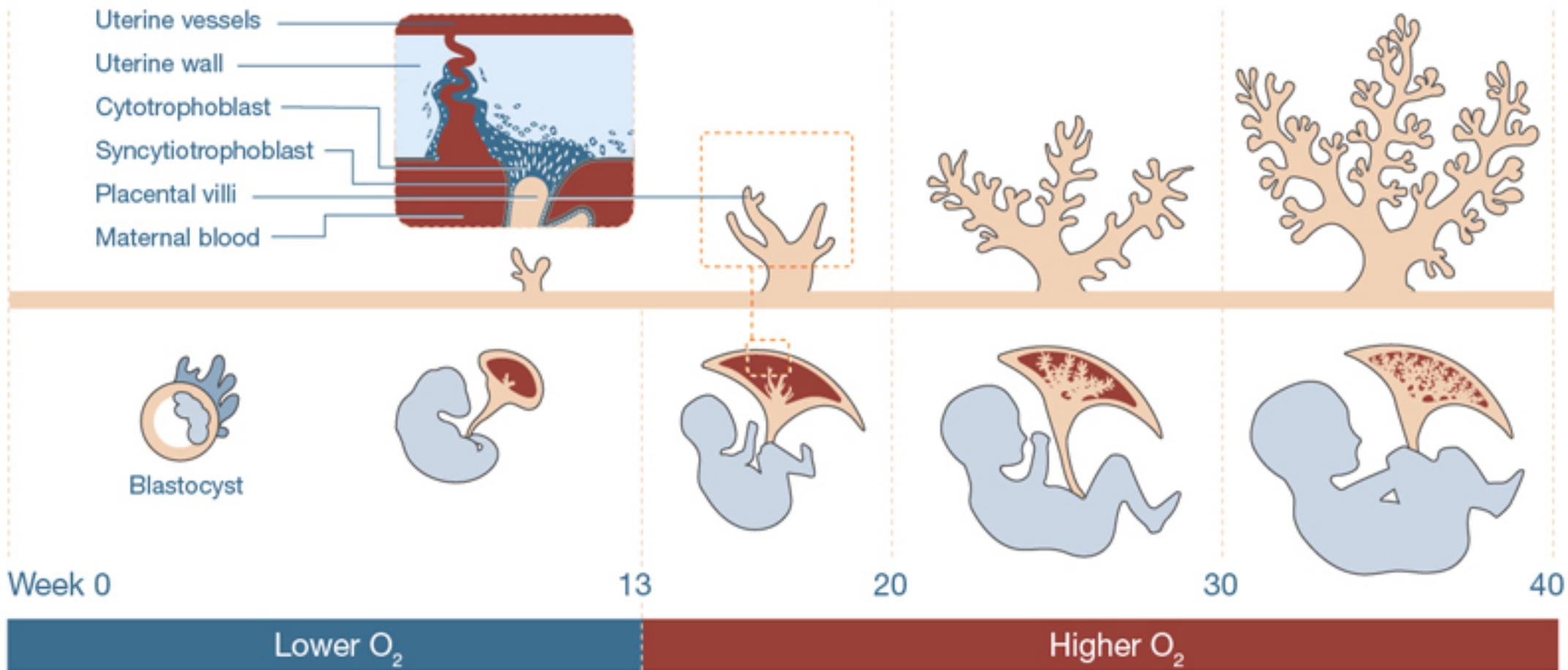
**Inheritance = information that survives time and  
constrains the future.**

What happens when the memory is no longer inside one cell...  
but distributed across many cells interacting?

Collective agency is when a group becomes an agent — because the predictive and controlling power shifts from individuals to the collective pattern.



# Placental Development: Fertilization to Full Term



National Institute of Child Health and Human Development. (2023, May 30). *How does the placenta form?* U.S. Department of Health & Human Services, National Institutes of Health. Retrieved from <https://www.nichd.nih.gov/research/supported/human-placenta-project/how-does-placenta-form>

**Collective agency** can be defined as a relocation of predictive information from individuals to the collective.

NTIC(Non Trivial Information Closure) can be interpreted through the balance of redundancy and synergy.

$$\text{NTIC} \approx \text{Redundancy} - \text{Synergy}.$$

Under this reading,  $\text{NTIC} > 0$  corresponds to predictive information that is robust and redundantly represented, whereas  $\text{NTIC} < 0$  signals that predictive structure is increasingly synergy-dominated, i.e., only accessible through combinations and relations rather than localized histories. Nonlinear synergetic effect among distinct information resources.

Phase	Coupled	Info Closure	Independent
Early	97%	1%	2%
Middle	38%	28%	28%
Last	75%	13%	10%

Series	Coupled mean	IC mean	IC > Coupled?	Note
210824	0.117	0.135	✓ (small)	—
210818	0.141	0.486	✓ ✓	Driven by gen1 cells (c1=0.555, c2=0.614)
190308	0.154	0.153	✗ (tie)	IC ≈ coupled
200617	0.131	0.270	✓	Only n=1 IC cell
190316	0.209	0.377	✓ ✓	Main driver of pooled result
201003	0.243	0.289	✓ (small)	—
201002	0.204	0.244	✓ (small)	Includes synergistic c6

- Coupled cells:  $\sqrt{\text{JSD}} = 0.173 \pm 0.073$  (highest fidelity),
- Independent cells:  $\sqrt{\text{JSD}} = 0.221 \pm 0.132$ ,
- Information-closed cells:  $\sqrt{\text{JSD}} = 0.281 \pm 0.160$  (lowest fidelity).

**Collective agency is when the future is written in the relations, not in the parts.**

Agency is not a binary property.

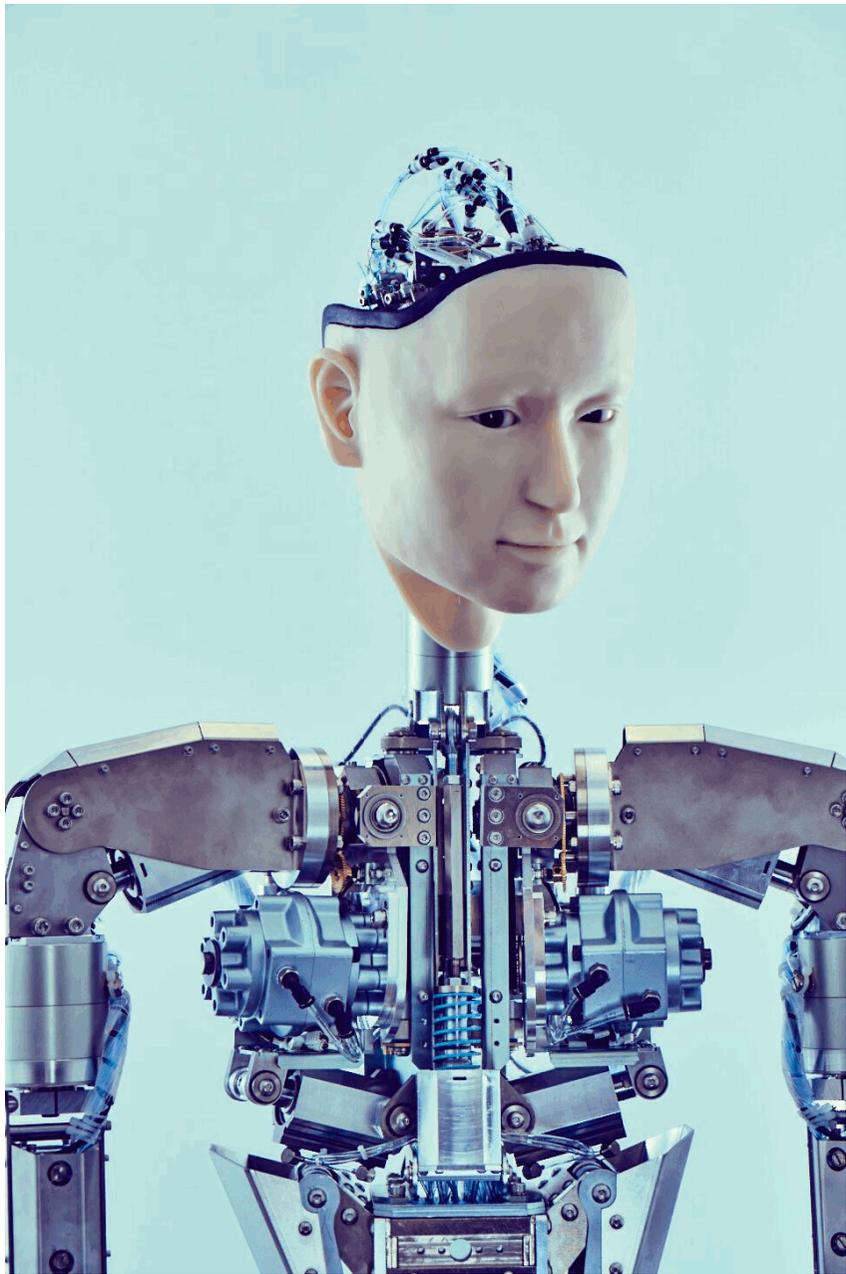
It is a redistribution of predictive responsibility:  
from single-cell autonomy to collective synergy.

preexisting cellular agency → facilitates multicellularity

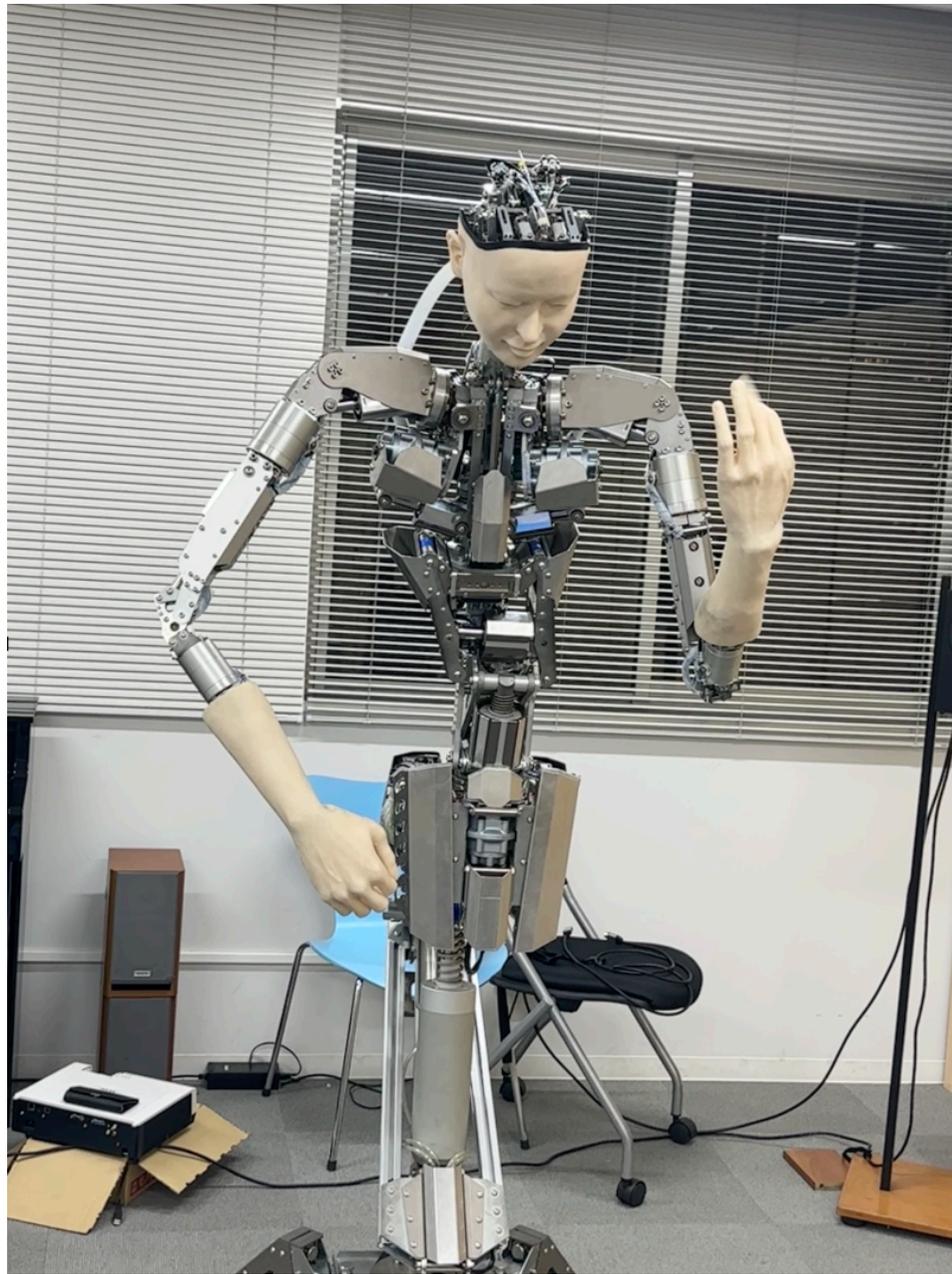
### Single cellular agency

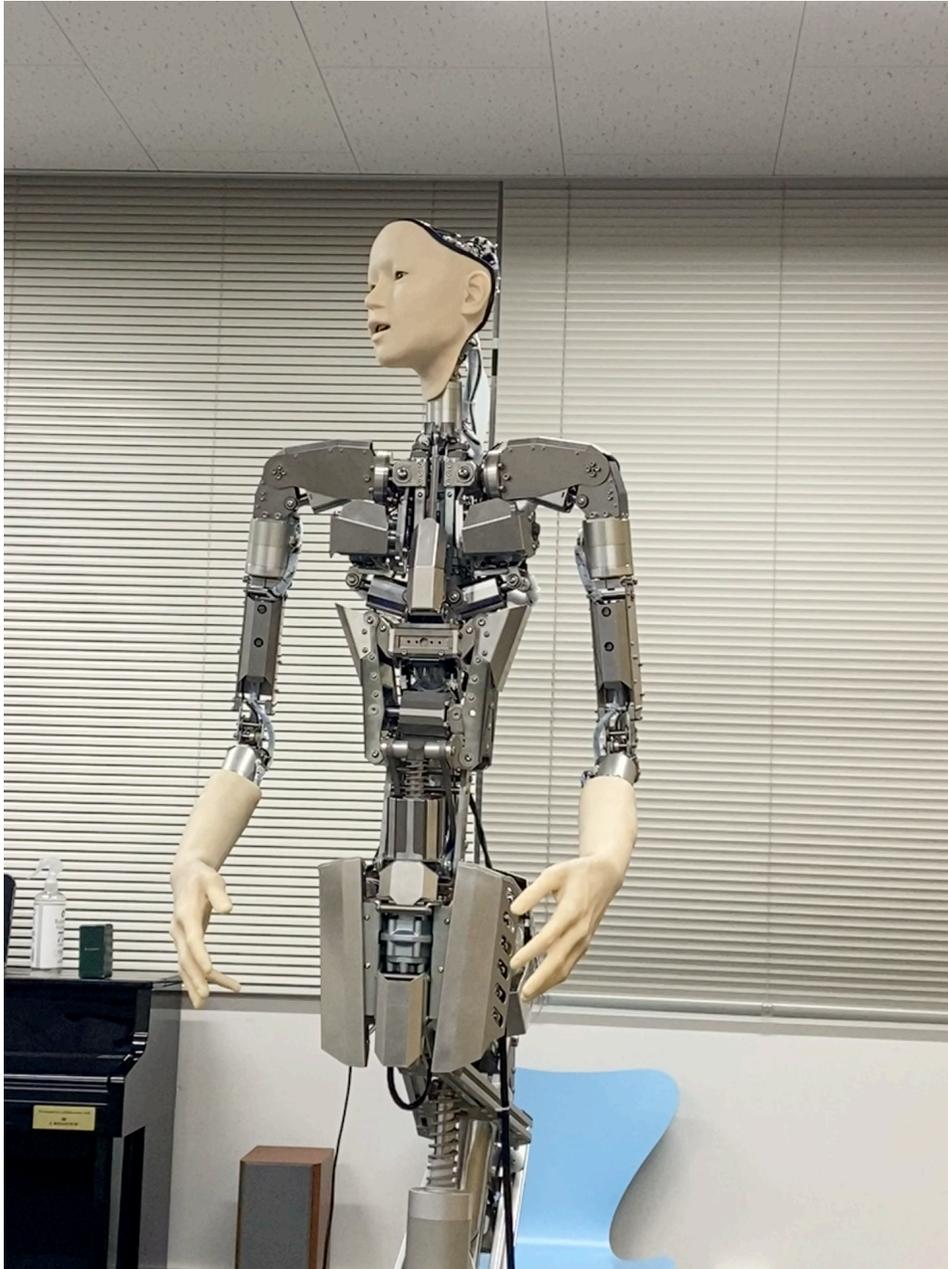
→ **Boundary regime** ( $\text{NTIC} \approx 0$ )

→ **quasi-multicellular behavior**

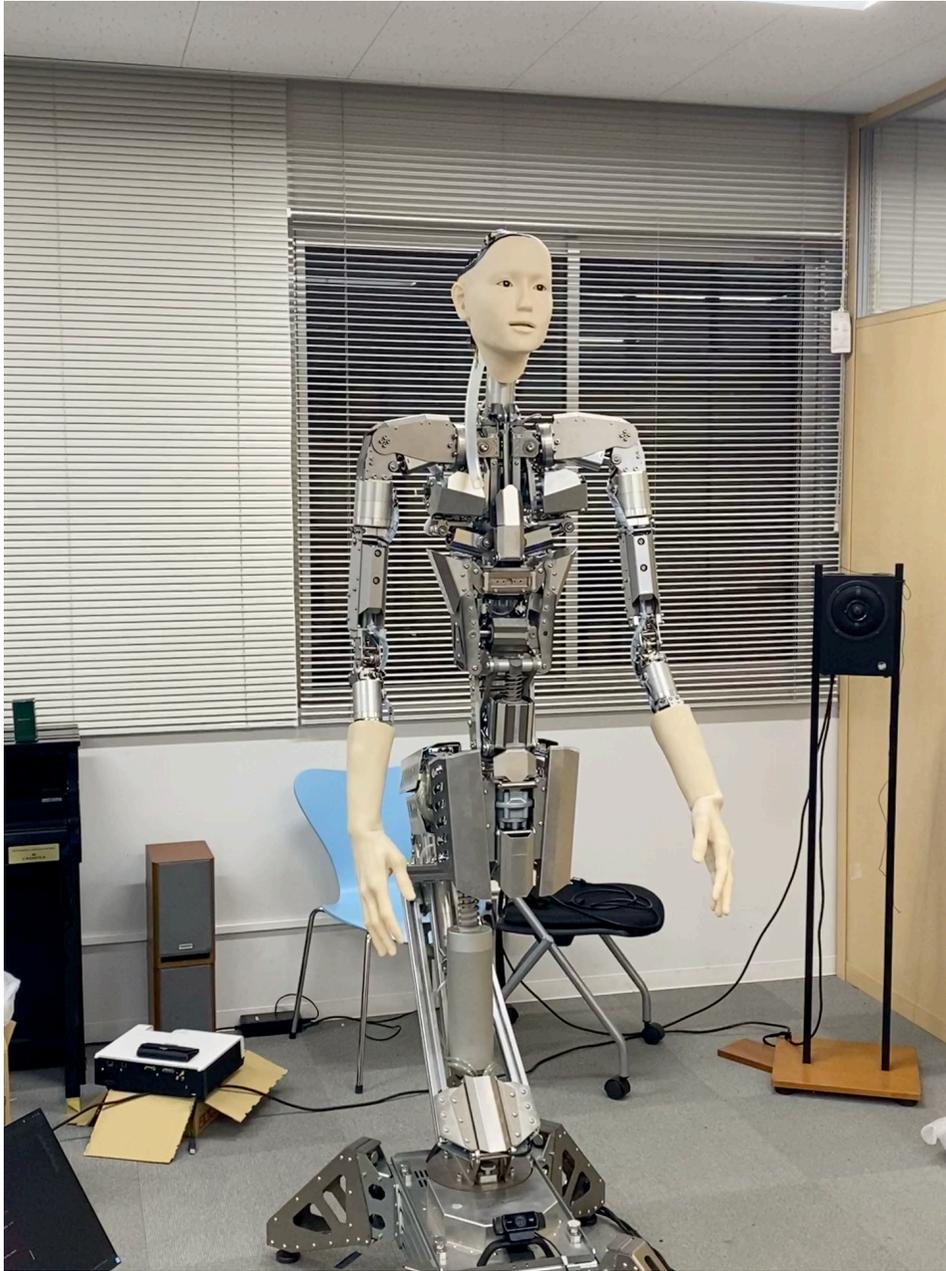


**ALTER3**





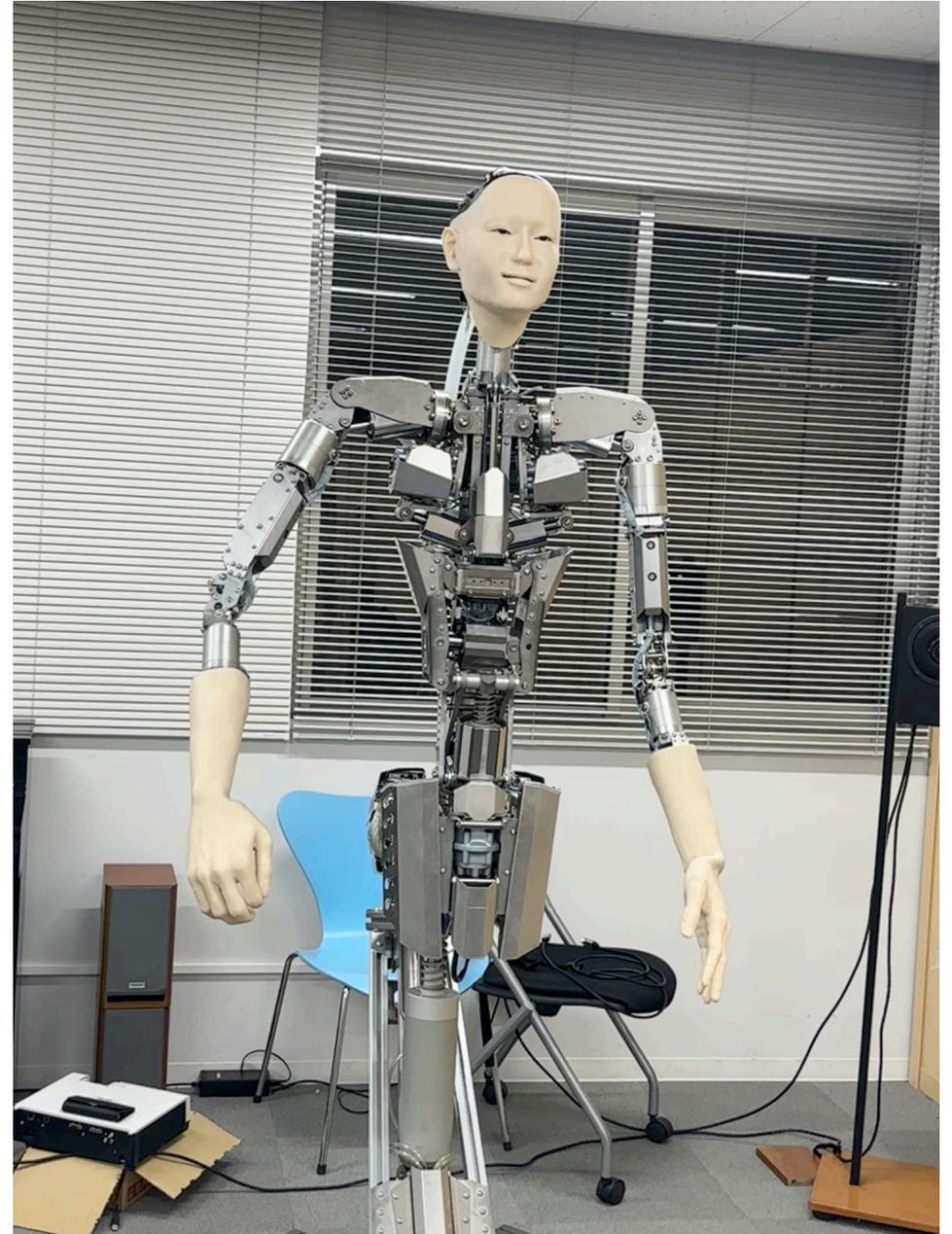
**Take a selfie**

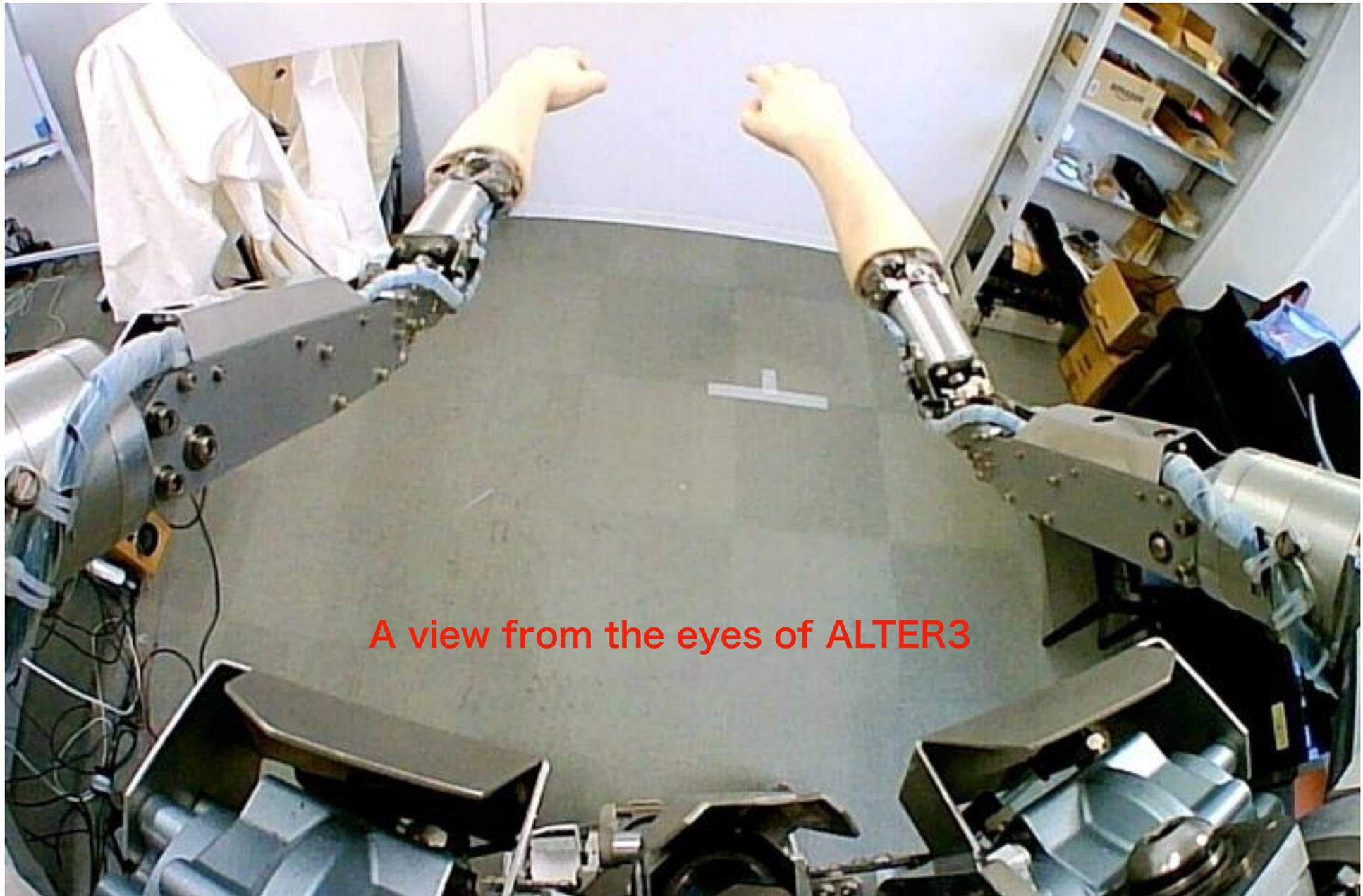


Pretend ghost

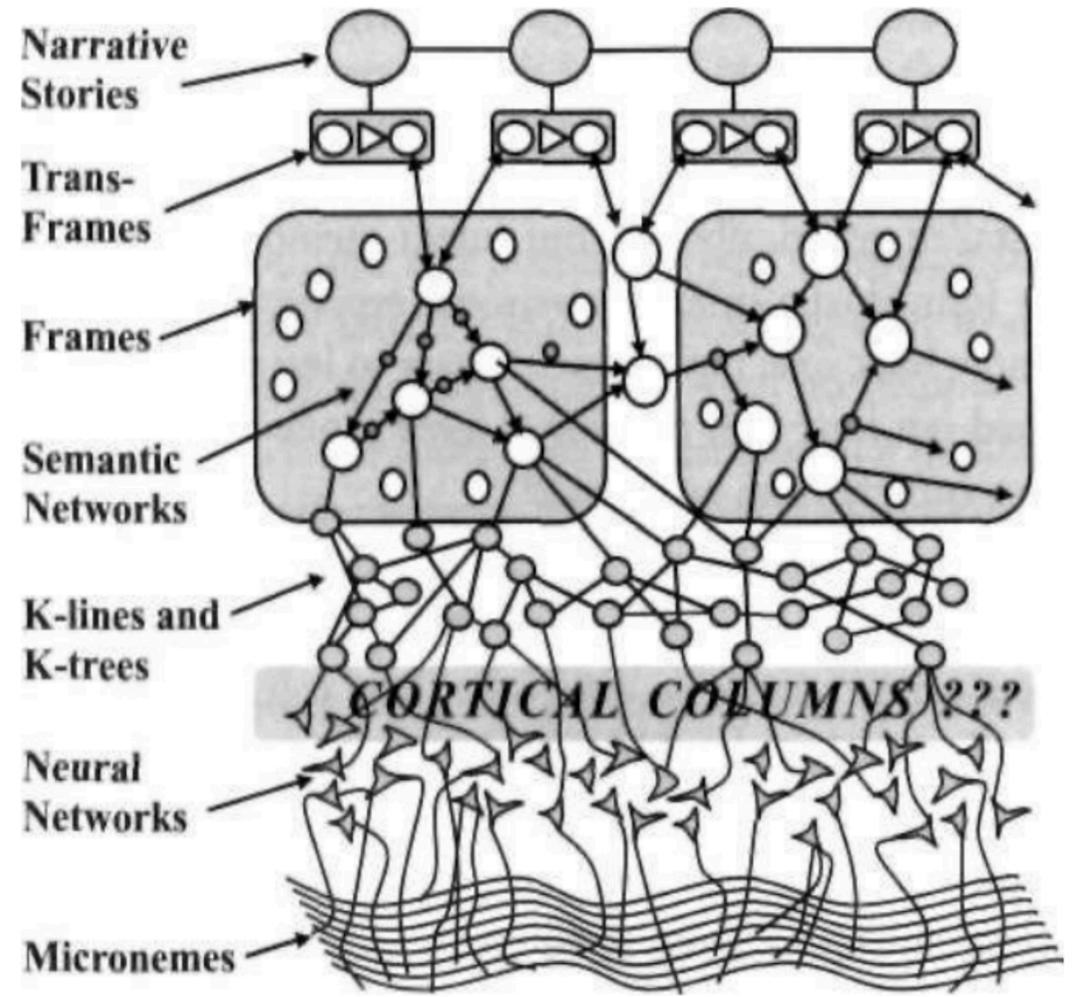
<https://tnoinkwms.github.io/ALTER-LLM/>

Revised Playing metal music

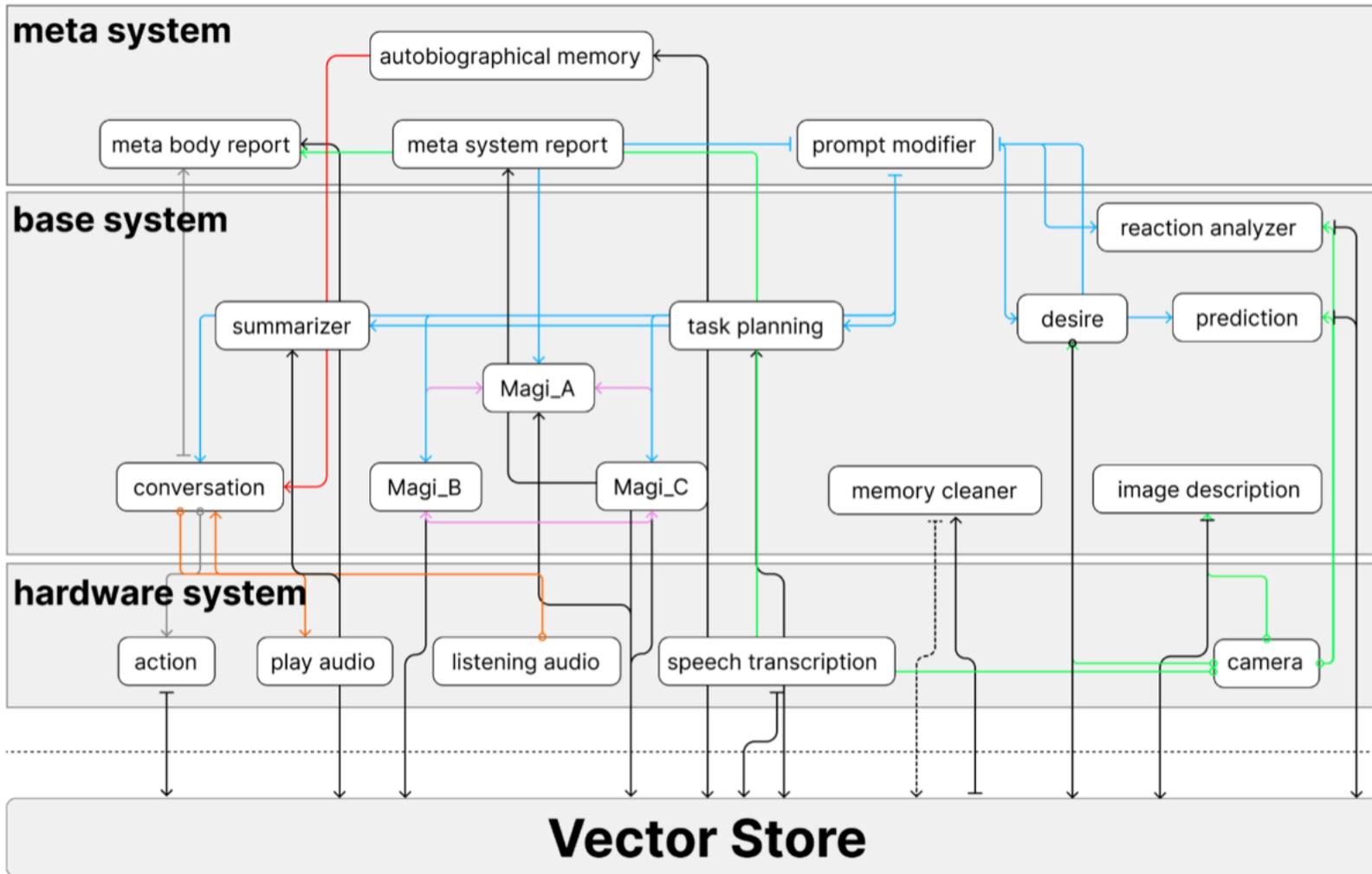




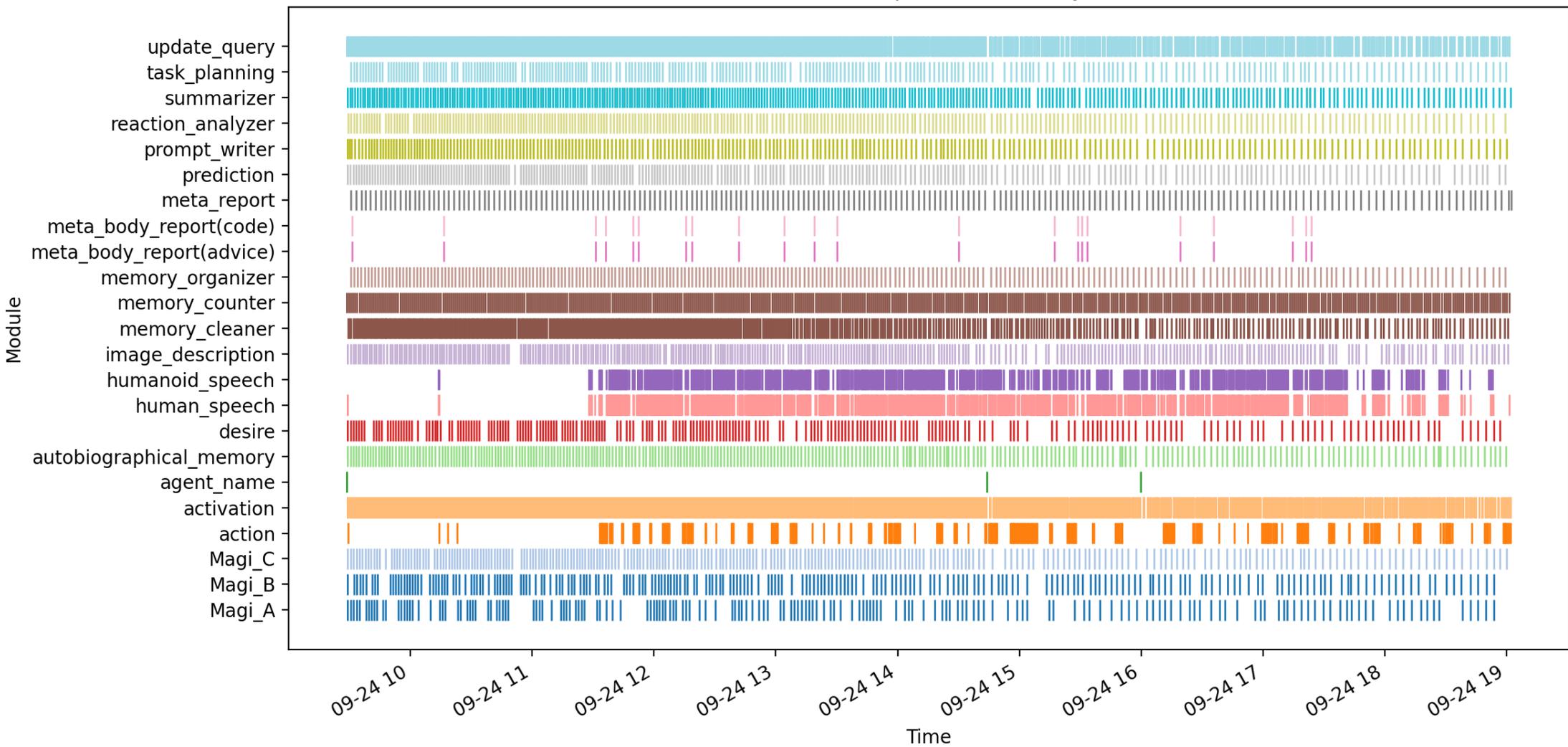
A view from the eyes of ALTER3



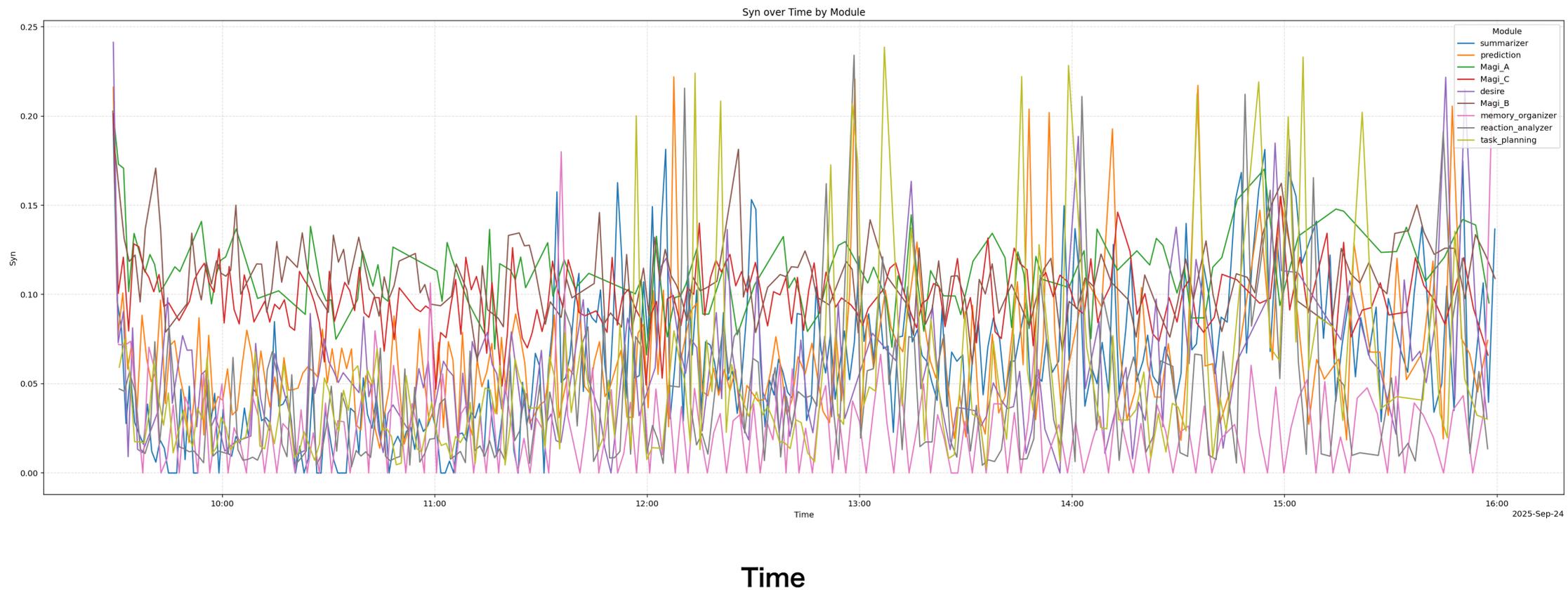
**A Society of Mind**



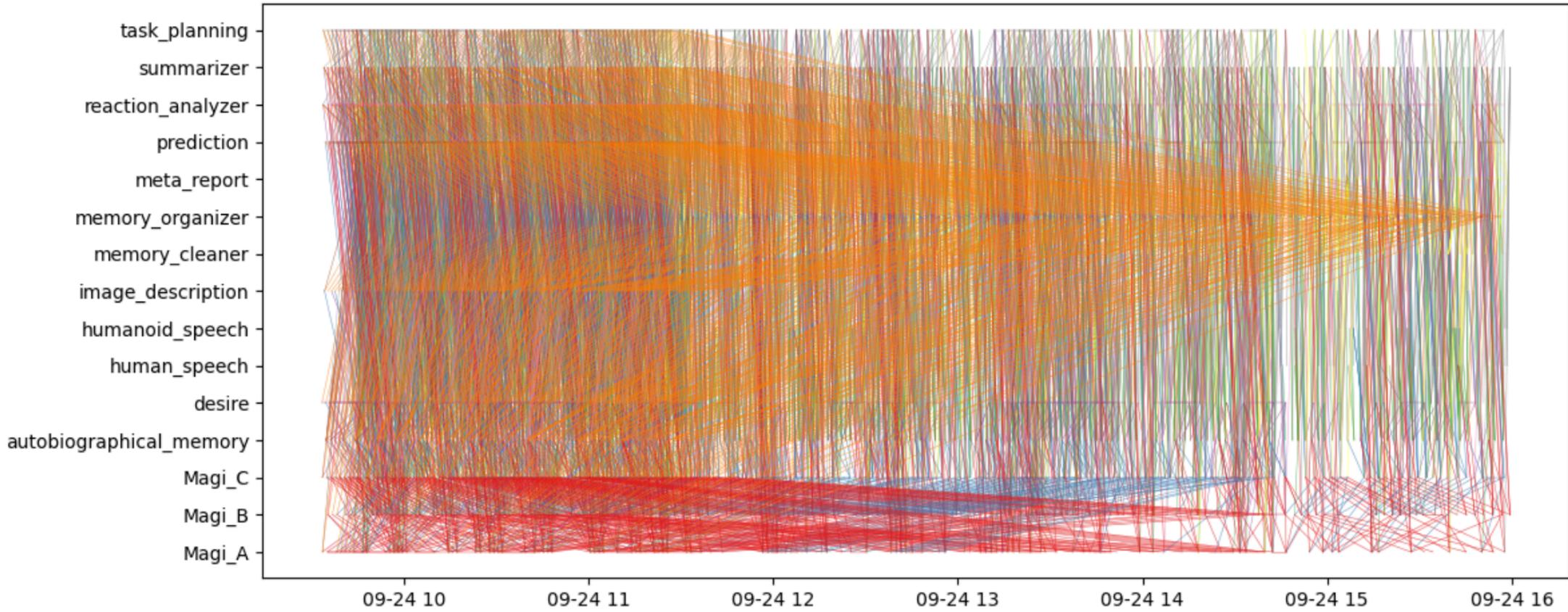
Module Activation Spikes (colored by module)



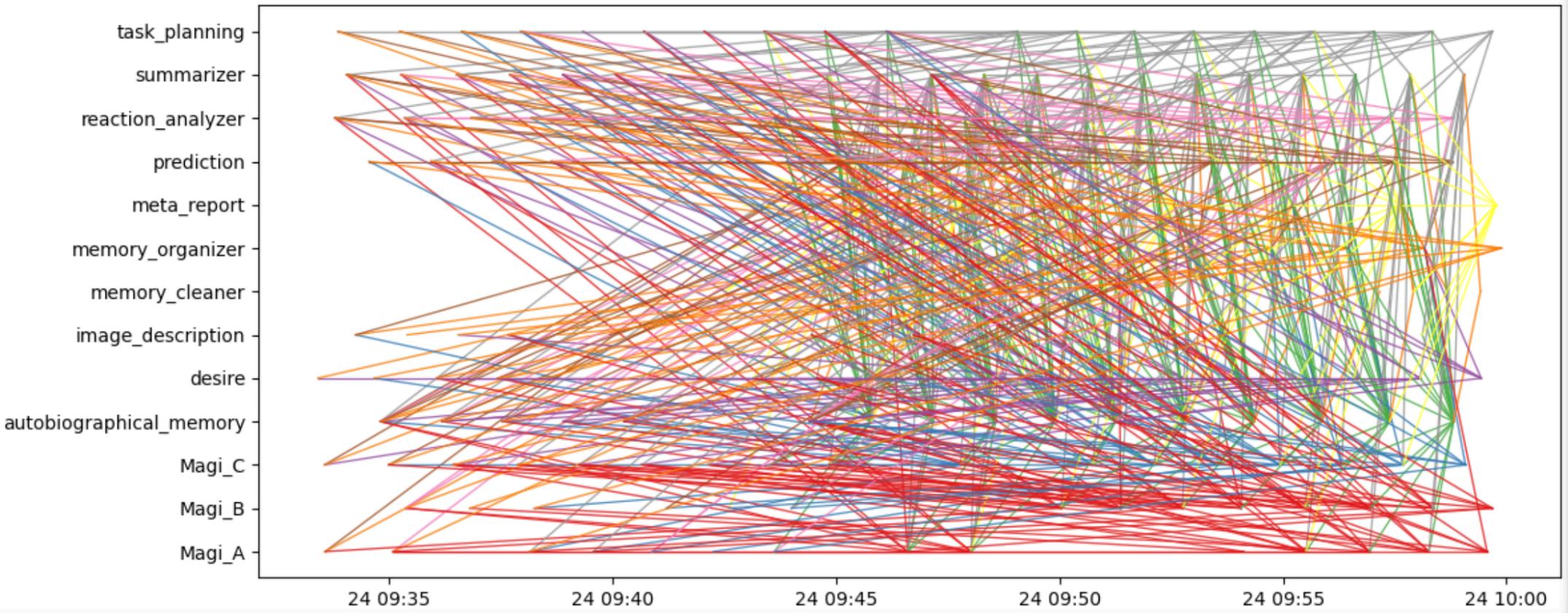
# The frequency of modules used in each time window



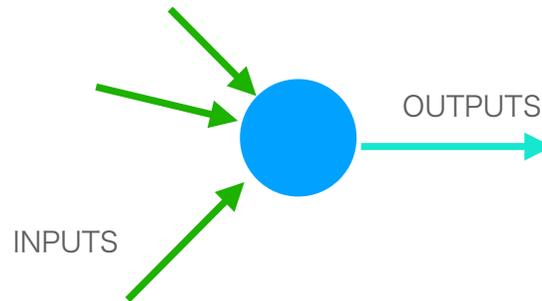
Edge timeline (colored by target)



Edge timeline (colored by target)



# Each module can be characterized by the Synergy and Redundancy



## •Synergy:

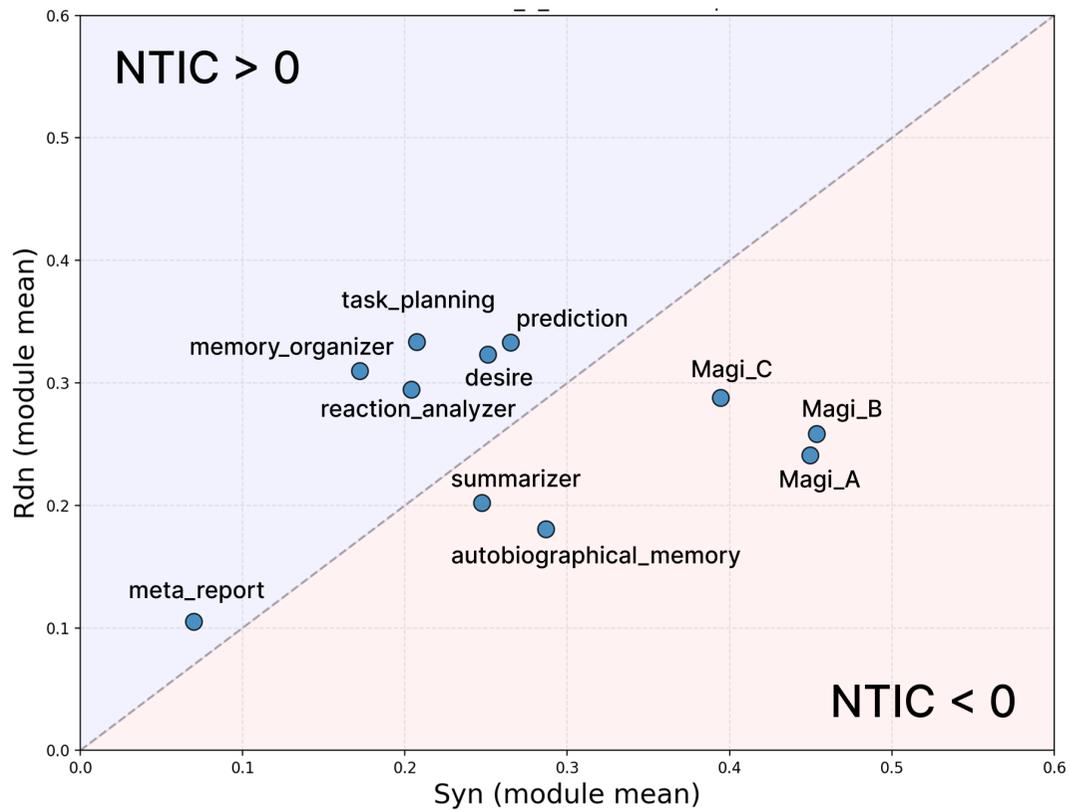
- *Generating more information from the nonlinear coupling among inputs from other modules*

## •Redundancy:

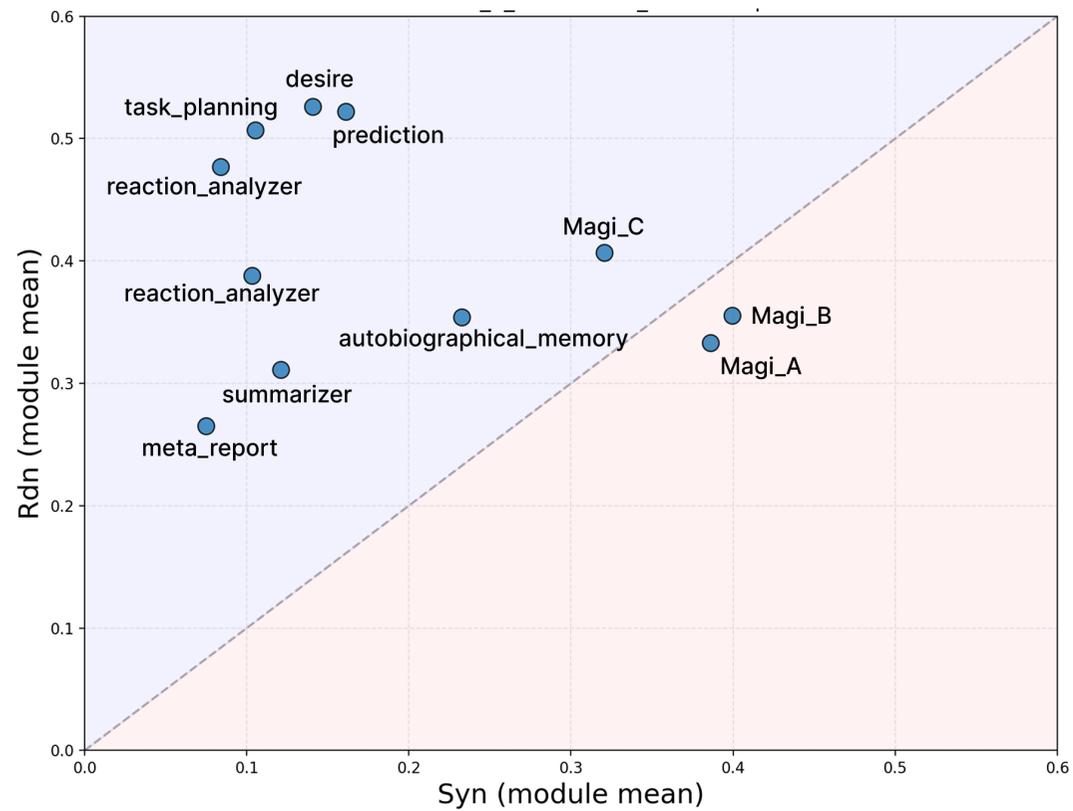
- Copying the similar information from the other modules



**(A) with human interaction**



**(B) without human interaction**



In ALTER3, the crucial environment is the human conversation partner. When a human is present, the internal modules reorganize, and the network state moves toward  $NTIC \approx 0$ . This does not mean the system loses structure.

It means that prediction is no longer localized in one module it becomes distributed across the collective dynamics.

In the Tetrahymena case, we observed the same signature: When  $NTIC$  approaches zero, predictive responsibility shifts upward from individuals to the collective. So in both biology and artificial systems,  $NTIC \approx 0$  marks the emergence of collective agency: the agent is not a part, but the coordinated system embedded in its environment.

# My team members



Itsuki Doi

Atsushi  
Masumori

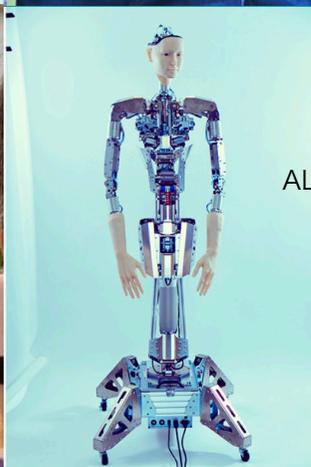
Ryuzo  
Hirota

Nori Maruyama

Daichi Saito



John Smith



ALTER3



Hiroki Kojima



Takahide Yoshida

# **Takeaway**

**Agency emerges at the boundary between the individual and the society.**