

深層学習と記号システムのその先へ 集合的予測符号化から自律駆動科学まで

Tadahiro Taniguchi

- 1) Professor, Graduate School of Informatics, Kyoto University
- 2) Visiting Professor, Research Organization of Science and Technology,
Ritsumeikan University
- 3) Senior Technical Advisor, Panasonic Holdings Corporation

情報科学技術分野における戦略的重要研究開発領域に関する検討会

11th June 2024

深層学習と記号システムの融合とシステム1/2

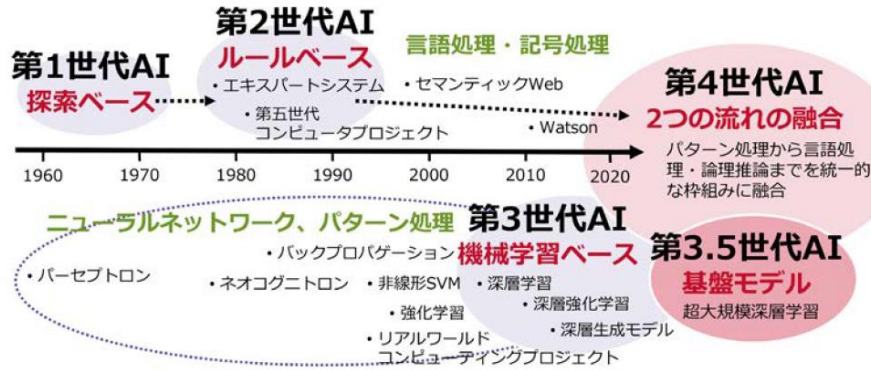


図 1-1-8 第4世代AIへの発展の流れ

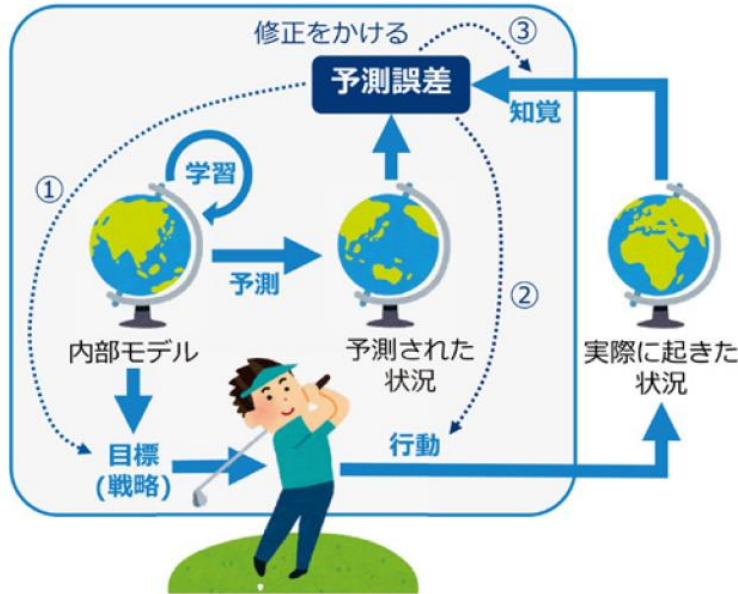


図 1-1-10 発達・創発モデル (予測誤差最小化原理)

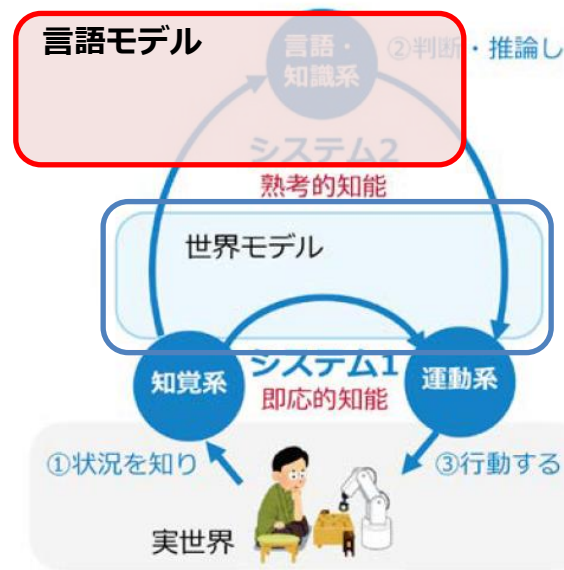


図 1-1-9 二重過程モデル

人間の知能はシステム1+2をカバー、**第4世代AI**もこれを目指す

第3.5世代AI(現在の基盤モデル)はシステム2の一部までカバー、論理推論・論理構築等が十分にできていない

第3世代AI(従来の深層学習)はシステム1に相当する

マルチモーダル情報と大規模言語モデルに基づく ロボティクス基盤モデル



Image data

Text data

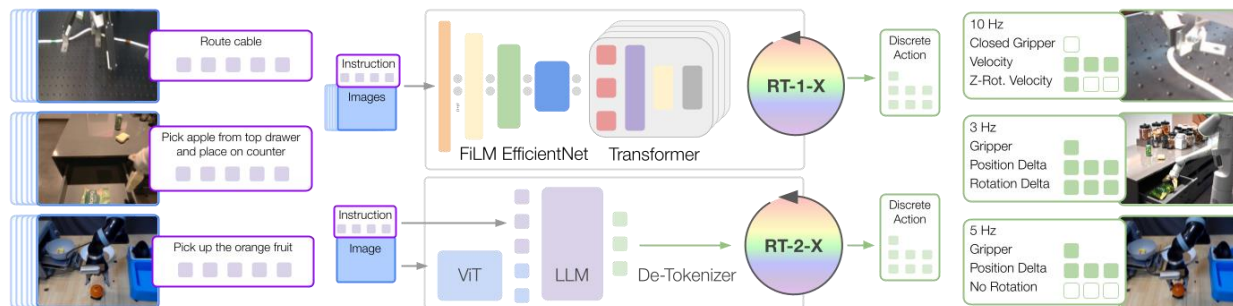
ViT
22B

PaLM
540B



RT-1: Robotics Transformer for Real-World Control at Scale [Brohan+ 2022] : 13万エピソード700種類のロボットタスクを13台のロボット（EDR）を17ヶ月駆動してデータを集めて大規模に学習

PaLM-E: An Embodied Multimodal Language Model [Driess+ 2023]
画像の基盤モデルと大規模言語モデルを統合しロボティクスのアプリケーションを目指している。



RT-X [Open X-Embodiment Collaboration 2023]: 複数拠点ロボティクス基盤モデル。22の異なるロボットから21の研究機関で分散的に収集しオープンデータセットを構築

Brohan, Anthony, et al. "Rt-1: Robotics transformer for real-world control at scale." arXiv preprint arXiv:2212.06817 (2022). Dec. <https://robotics-transformer.github.io/>

Driess, Danny, et al. "PaLM-E: An Embodied Multimodal Language Model." arXiv preprint arXiv:2303.03378 (2023). <https://palm-e.github.io/>

Open X-Embodiment Collaboration, Open X-Embodiment: Robotic Learning Datasets and RT-X Models, 2023

大規模言語モデルの背景にある言語現象

学習・内化

巨大な言語資源

(潜在的な構造)
分布意味論・構文構造

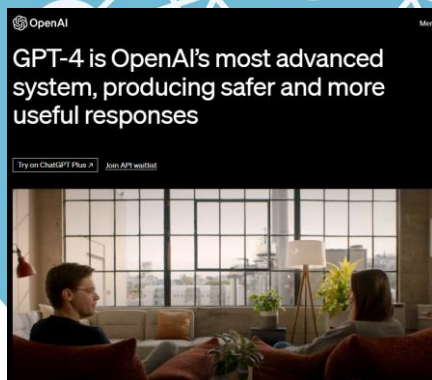
巨大なニューラルネットワーク

暗黙的な推論

言語システム

発話・テキスト

言語知識



フランス語のクイックソートのコードをください。

以下は、フランス語で書かれたクイックソートのコードです。

```
python
def quicksort(tableau):
    if len(tableau) <= 1:
        return tableau
    else:
        pivot = tableau[0]
```

大規模言語モデル(LLM)

GPT-4 is OpenAI's most advanced system, producing safer and more useful responses <https://openai.com/product/gpt-4> (2023/03/15)

人間の社会が言語を形成し終えた後に、そこから大規模なテキストが生成され、それを大規模言語モデルに内化すれば**実用上高いレベルの言語理解・使用ができる**と分かった。

言語の創発・動態についての様々な問い

言語システムがどうやって出来てきたか？（**創発**）

文化としての言語システムはどう動的に変容するのか？（**動態**）

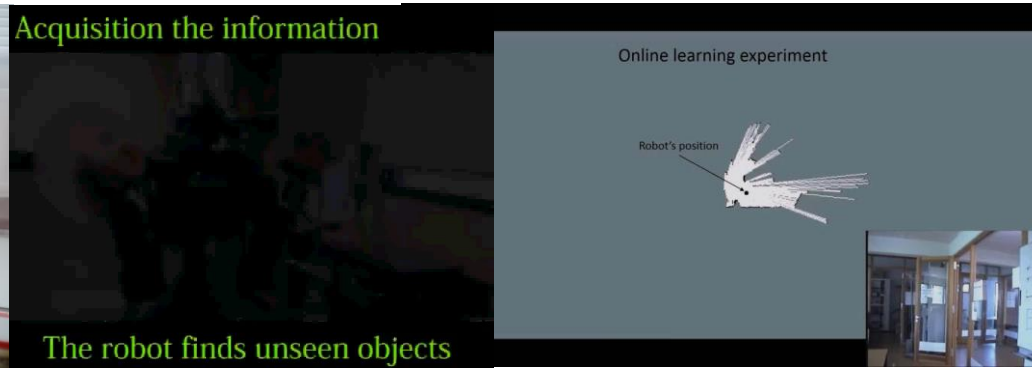


人間の社会が形成した**言語の性質**

- A) **分布意味論**を満たす.
- B) 特殊な**統語的構造**を持つ.
- C) 相手の頭の中が覗けない中で、言葉の意味は**創発**し**共有**される.
- D) **実世界**において意図や対象を表象し、**環境適応**の役に立つ.

□ どうして人間の社会がこのような言語を形成できているのか？
□ なぜ**大規模言語モデル**がこれほどまでに有効なのか？（凄いのはLLMより言語そのもの？）

記号創発ロボティクス^[谷口 2014他]



谷口忠大「記号創発ロボティクス」2014

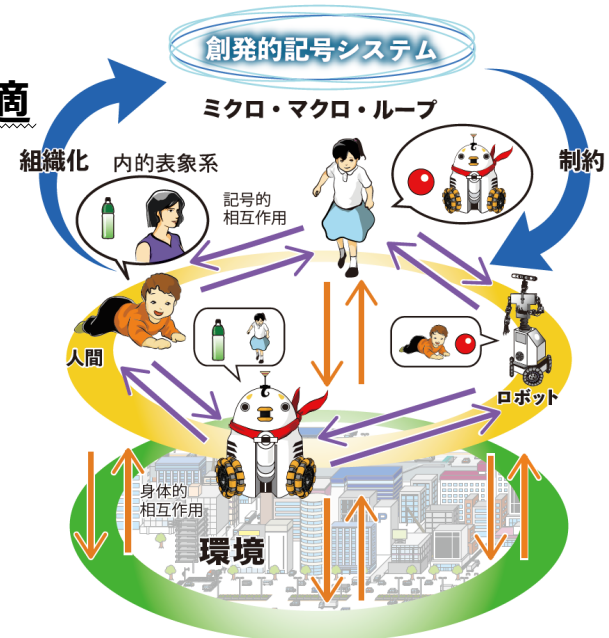


谷口忠大「心を知るための人工知能」2020

- 人間の子供は**自らの身体的経験**，**感覚運動情報の統合**を通じて機能を獲得し，言語を獲得し，コミュニケーションをも可能にする。
- **実世界経験**に基づく**発達・言語獲得ロボット**の実現とそれを通した人間の**認知発達の構成論的理解**を目指す。
- また**記号（言語）**を生み出し世界を理解し，協調する**適応的な自律的知能**を構成し理解する。

発達知能 / 記号創発システム への構成論的アプローチ

その構想は未達。完遂のための新展開が必要。



記号創発システムの概念図

谷口忠大 「コミュニケーションするロボットは創れるか」 (NTT出版) 2010

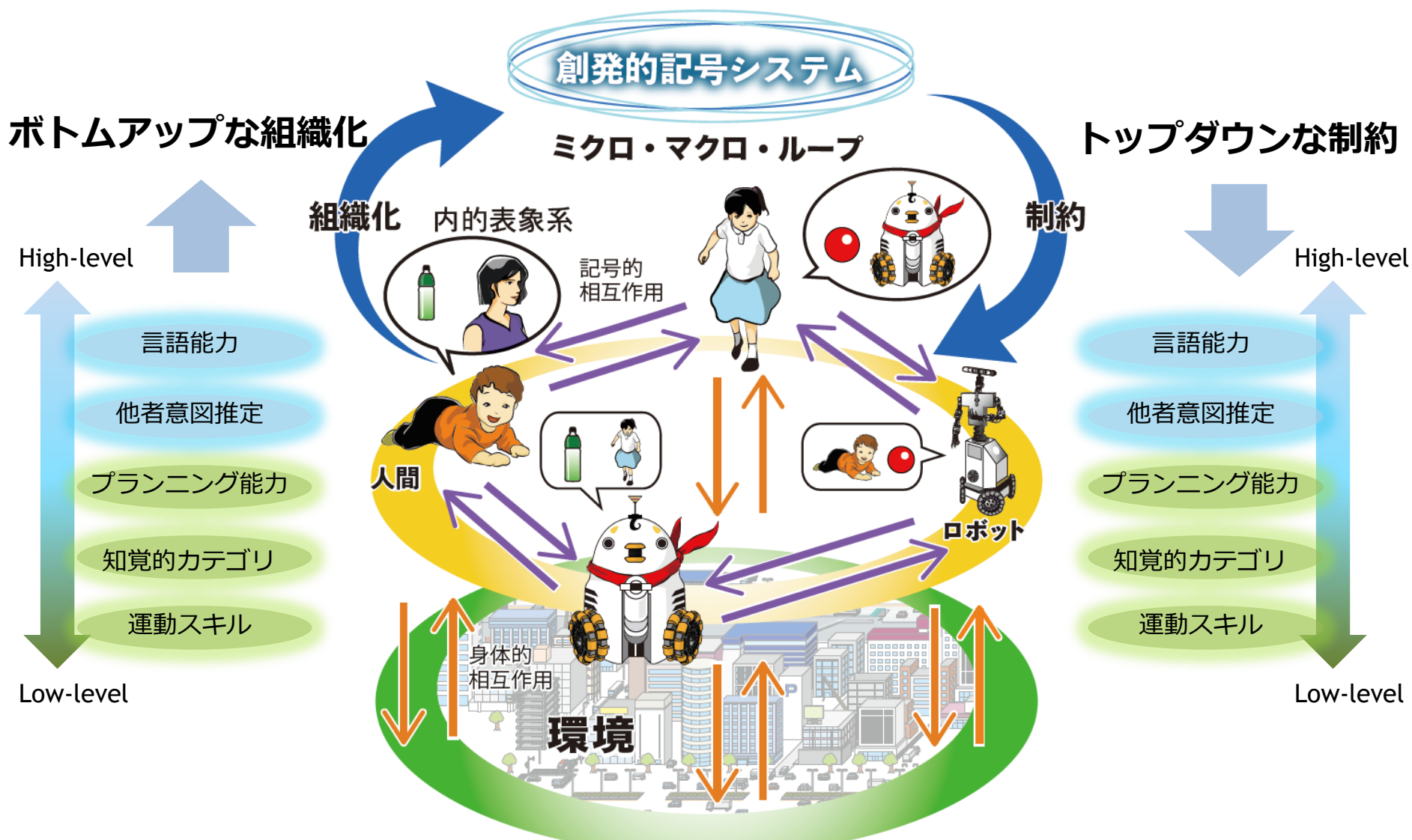
谷口忠大「記号創発ロボティクス」(講談社メチエ) 2014

Tadahiro Taniguchi, Takayuki Nagai, Tomoaki Nakamura, Naoto Iwahashi, Tetsuya Ogata, and Hideki Asoh, Symbol Emergence in Robotics: A Survey, Advanced Robotics, 30(11-12) pp.706-728, 2016.

Tadahiro Taniguchi, Emre Ugur, Matej Hoffmann, Lorenzo Jamone, Takayuki Nagai, Benjamin Rosman, Toshihiko Matsuka, Naoto Iwahashi, Erhan Oztup, Justus Piater, Florentin Wörgötter, Symbol Emergence in Cognitive Developmental Systems: A Survey, IEEE Transactions on Cognitive and Developmental Systems, 11(4), pp.494-516, 2019.

谷口忠大「心を知るための人工知能」(共立出版) 2020

記号創発システム



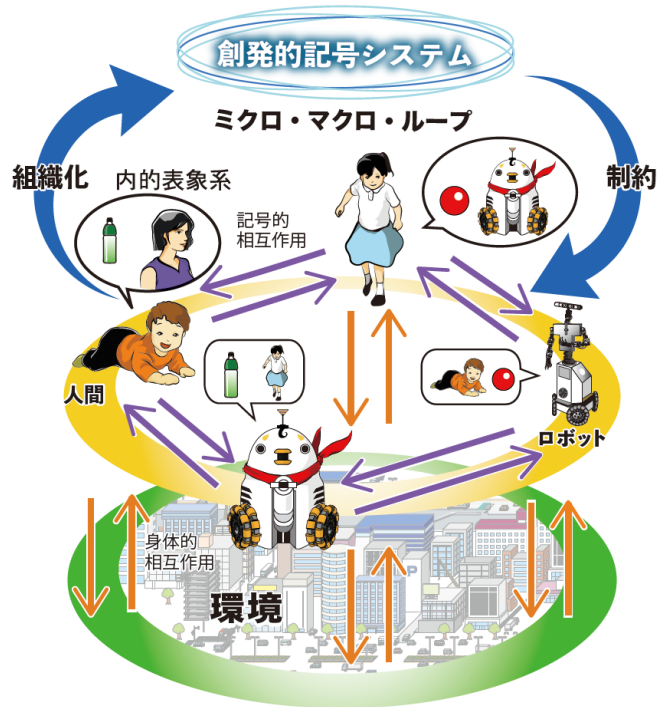
言語を生み，その意味を調整することを通して，社会は全体として何をしているのか？

解説

分散的ベイズ推論としてのマルチエージェント記号創発

Symbol Emergence as a Decentralized Bayesian Inference

谷口 忠大^{*1*2} ^{*1}立命館大学 ^{*2}パナソニックホールディングス株式会社
Tadahiro Taniguchi^{*1*2} ^{*1}Ritsumeikan University ^{*2}Panasonic Holdings Corp.



1. はじめに

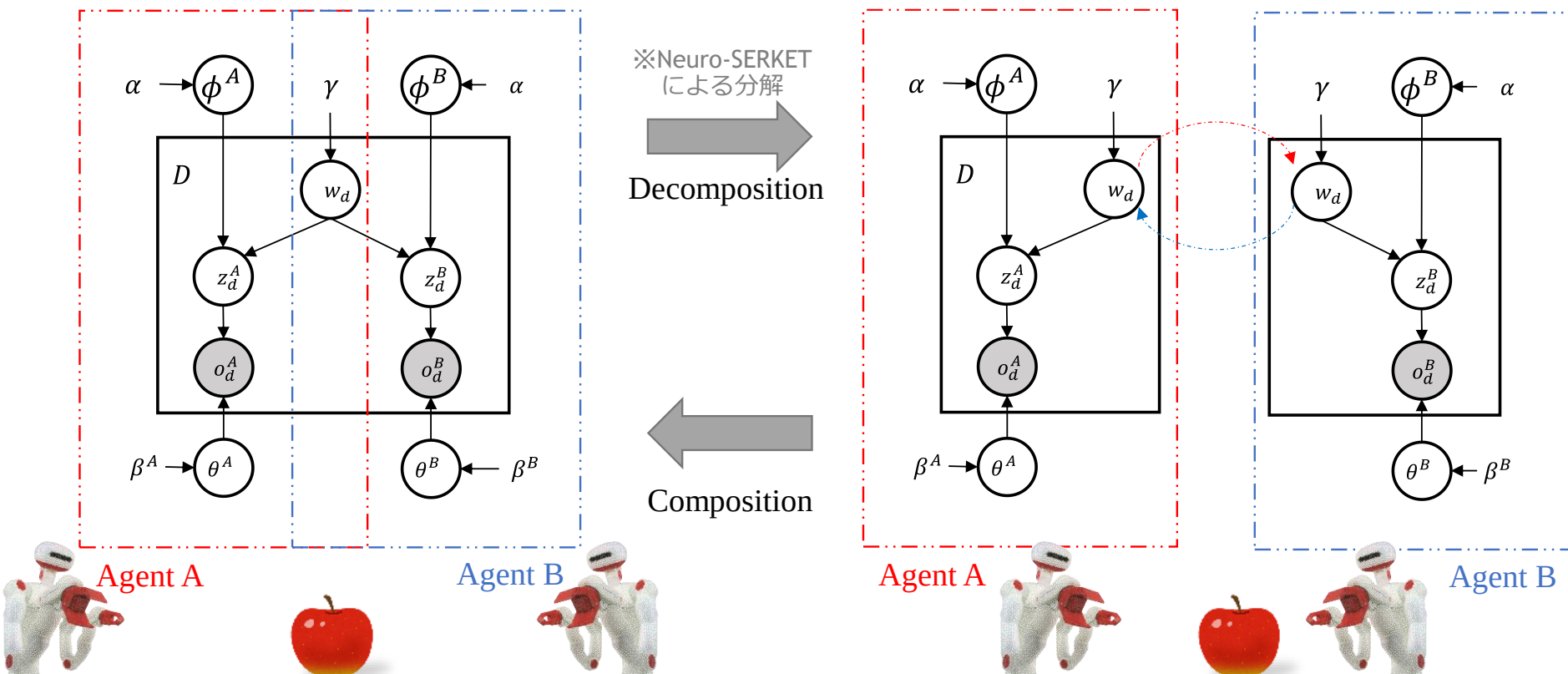
「社会はベイズ推論する。それが記号創発である」
このような主張を立てるところから本稿を始めよう。それが何を意味しているのか？ それは妥当な言明なのか？
本稿はこの抽象的な理論的言明に対して記号創発ロボティクスの最新の知見に基づいて、肉付け、またその理論的骨子を与えていく。

集合的予測符号化 (Collective Predictive Coding)

私たちの社会は言語を創発することで，集団（集合知）として世界の観測に対する予測性を高めているのではないか？．

言語創発は集団としての予測符号化を個体の学習に分散化するメカニズム

記号創発システムの確率的グラフィカルモデル



Theorem 1. *MH naming game is a Metropolis-Hastings sampler of $P(w, z, \theta, \phi|o)$.*

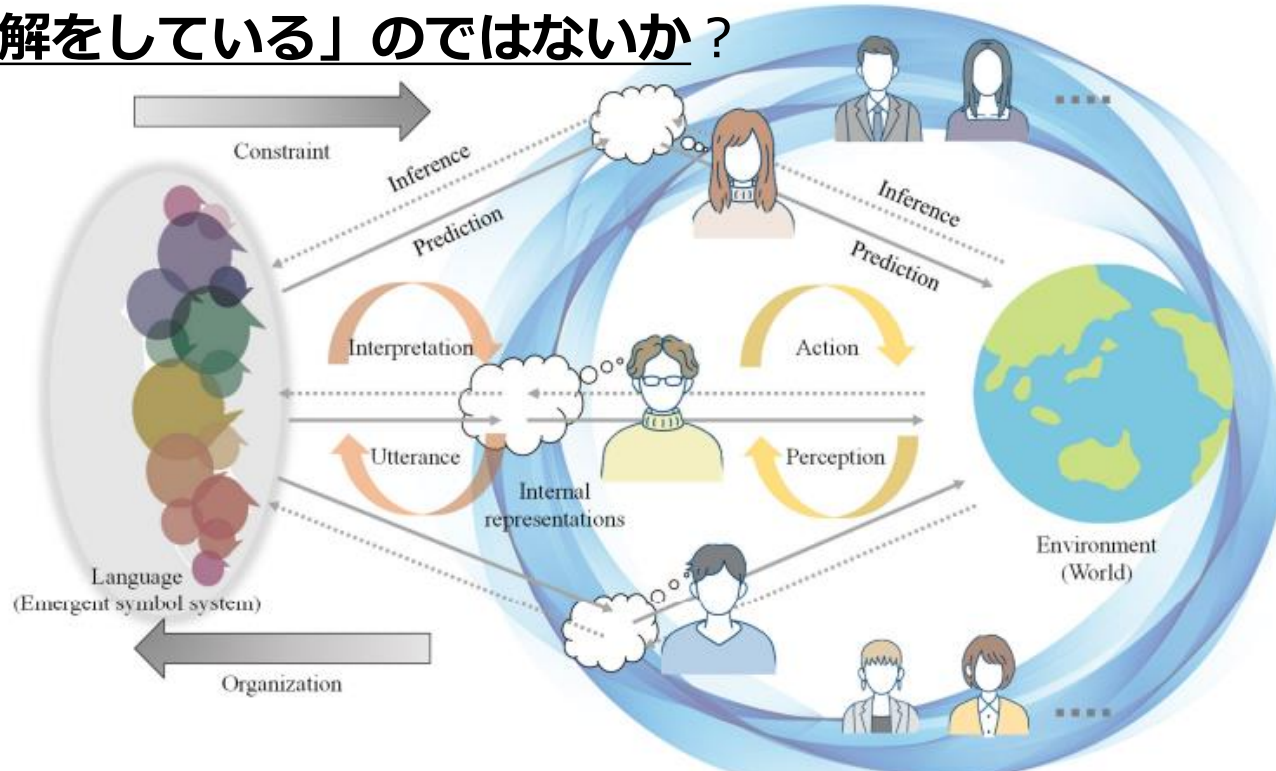
MH名付けゲームは分散型MCMCベイズ推論になる

私たちは言語を生み出すことで「脳をつなぐ」のと同等の認識統合を行える？

なぜ大規模言語モデルは世界を理解しているように見えるのか？

集合的予測符号化仮説 [Taniguchi '23]

- 言語そのものが**集合的な予測符号化**によって形成されるために、世界の情報が分布意味論の中にコーディングされている。
- 記号創発システムの**社会的な表現学習装置**として理解できる。
- 故に**大規模言語モデルは「まるで身体を持っているかのように世界を理解をしている」のではないか？**



Taniguchi, T. (2023, August 15). Collective Predictive Coding Hypothesis: Symbol Emergence as Decentralized Bayesian Inference. <https://doi.org/10.31234/osf.io/d2ty6>

集合的予測符号化に基づく言語と認知のダイナミクス ～記号創発ロボティクスの新展開に向けて～, 認知科学, 2023, accepted

人間の(社会的)知能の三重過程※

人間の知能は個に閉じない適応性に
その特殊性がある

※現状は谷口の造語です。
三重過程 = Tri-Process Theory??

社会的記号／規範系

システム3
社会的/集合的知能

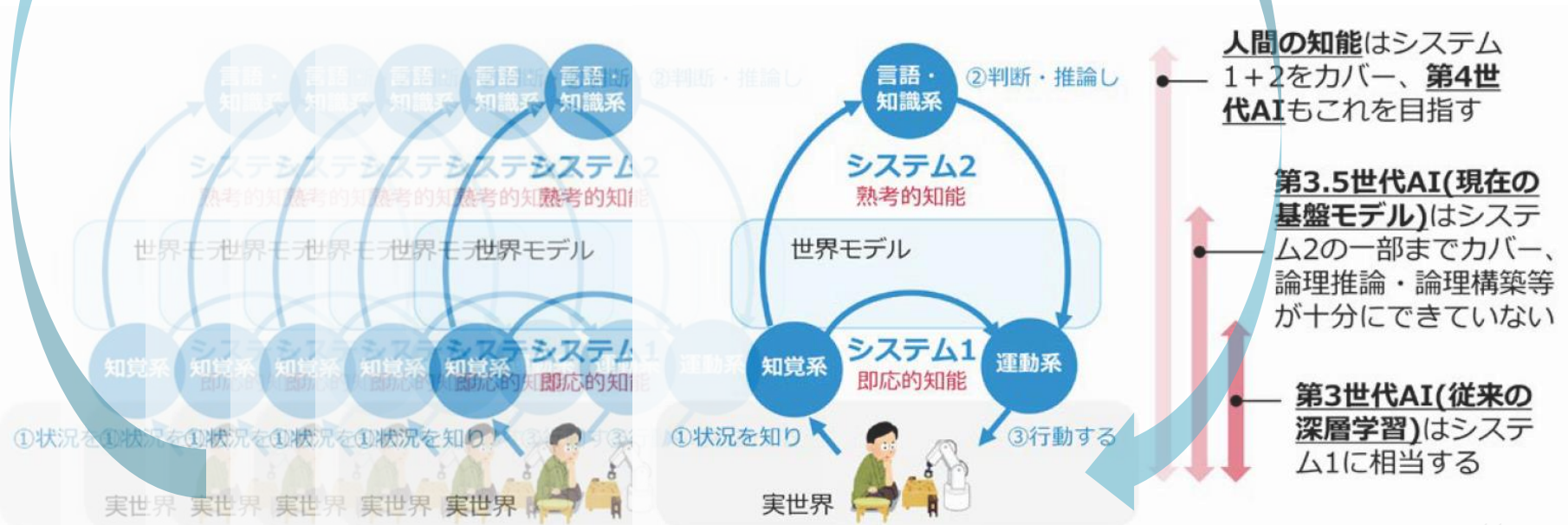
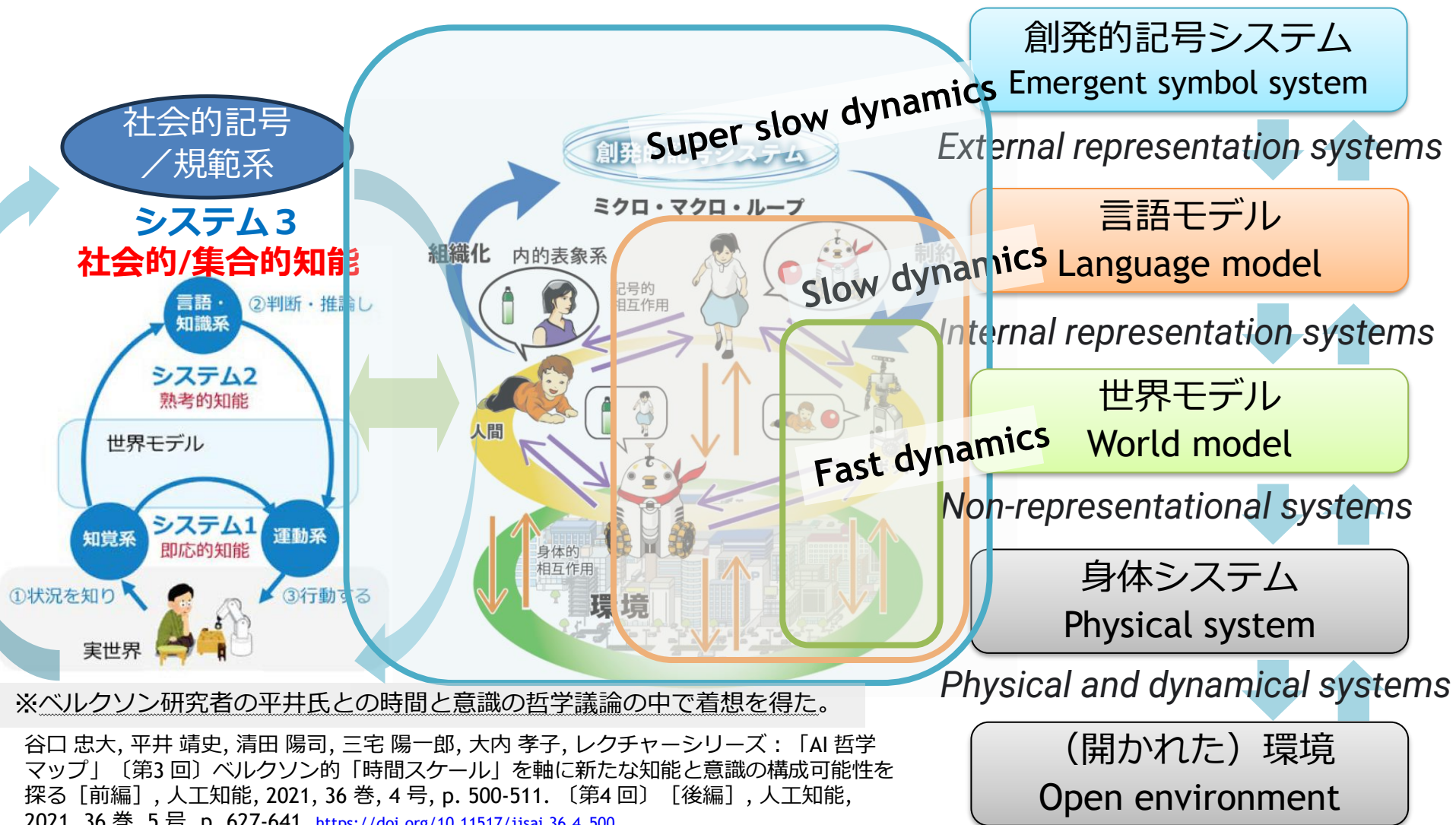


図1-1-9

二重過程モデル

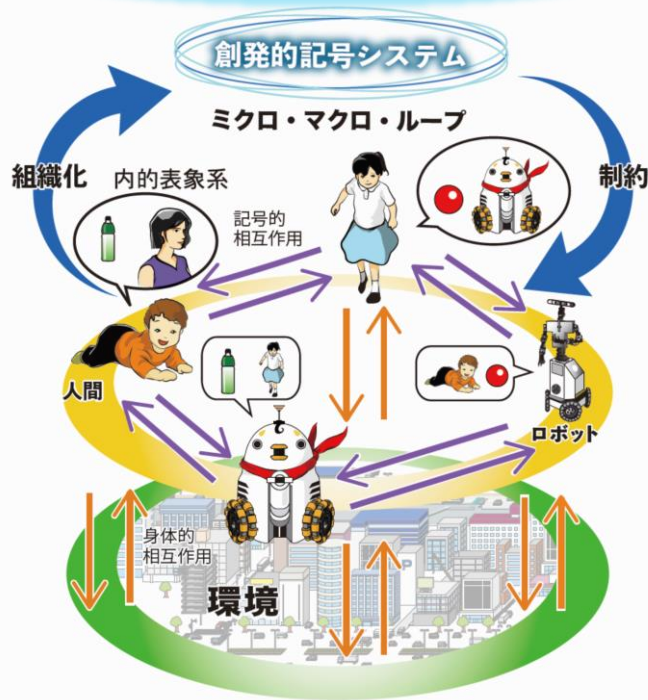
知能の三層モデル

(記号創発システム描像に基づく)



知能の三層モデルと記号創発システム に基づくAIロボット駆動科学のその先

形式知（と暗黙知）としての
科学的知識の形成と共有



□ 人間とAIが形成する科学的知識とは何か？

- AIロボット駆動科学は「科学者」が担う活動の代替、自律化までしか議論ができていない。
- 社会における「システム3」的領域におけるAIと人間の共創的な科学的知識の形成の位置づけが重要。
- 科学的知識の形成の視点では記号創発システム／集合的予測符号化は構成主義をベースに科学的実在論と社会構成主義を融和させるようなシステム観になっている。

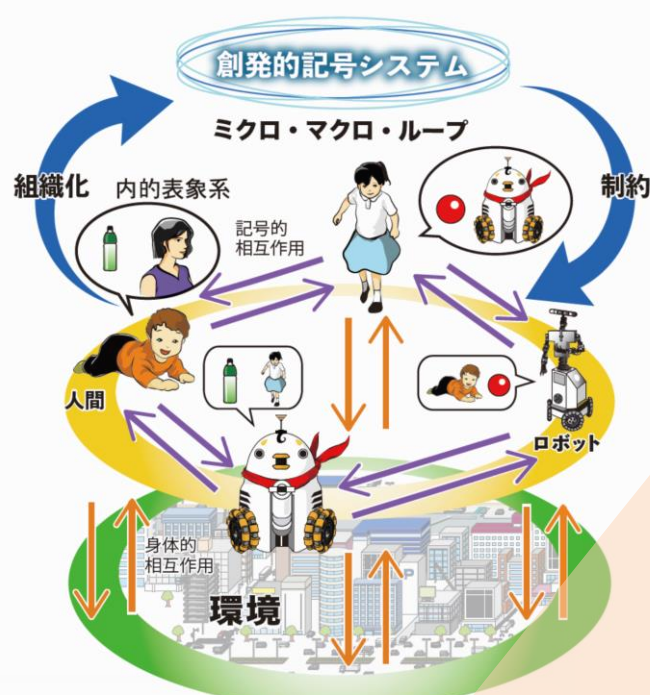
AIロボット駆動科学の発展
人間とAIによる「科学的知識」を支える科学哲学の探求



知能の三層モデルと記号創発システム に基づく共生的知能の創成に向けて

□ 共生的なAIアラインメント

- 人間は否応なくAIの言語使用や生成に影響を受ける。(e.g., "delve into")
- LLMは身体を持った人間の言語の「まとめ」的でありある意味で知識が接地している。
- しかし、その間にネガティブ／ポジティブフィードバック構造が生まれる。(e.g., AIカニバリズム)
- 「従わせる」AIアラインメントの議論とともに、俯瞰的な人間AI結合系の共生的なアラインメント議論が必要



例えばGPT-4がやたら使う"delve into"という表現は2023年3月以降、どんどんpopularになっており、言語使用上の立場を明確に変えている。
(谷口の個人的調査)

人間のパートナーとしての望ましいAI
人間AI共生マルチエージェント社会

<https://trends.google.co.jp/trends/explore?date=today%205-y&q=%22delve%20into%22&hl=ja>



関連する研究課題リスト

1. 記号／言語創発の学問の構成・体系化
2. 大規模言語モデルにある知識の背景を理解した上でそれらを活用する技術（分散型のLLMなどを含む）
3. マルチエージェントシステム（システム3）の知能の構成と活用（コミュニケーションのメカニズムデザイン）
4. 人間と生成AIをともに含んだ系の作動のモデル化と解析、その作動を保証するための体系の構築（共生的AIアラインメント？）
5. より明示的で形式的な記号的知識（論理式、科学的知識表現）を形成するための枠組み（システム3レベルでのNeuro-Symbolic AI）
6. 自律駆動科学や立法・法制全体のシステム論的モデル化と、その一部への生成AI活用
7. 身体（システム0）を持ったロボットが実世界情報を言語構造にまでフィードバックさせて更新し、活用する技術
8. 実世界の情報を取り込み続けるための高速・低消費電力AIチップ

