

# 人と融和して知の創造・越境をする AI ロボット

## 1. プログラムにおける位置づけ

本研究プロジェクトでは、研究者と AI が融和して、2050 年にノーベル賞級の研究成果を生み出せる世界を目指します。研究や開発に従事する人材の生産性を、論文のみならず、特許も含めた創造的活動まで含めて加速する基盤技術を開発します。目標3の3つのターゲットでは、「原理・解法を発見する AI ロボット(自律実証)」の中に位置づけられます。

現在でも世界中で実験科学を加速する AI やロボットの研究開発が進んでいます。実験科学では、実験をしてその解析を踏まえて資料を作成しつつ、新たな発見の気づきを得て仮説を立ててまた実験し…というループを繰り返します。仮説が先行する場合も実験が先行する場合もあり、ビジネスにおける PDCA サイクル (Plan-Do-Check-Act cycle) や OODA ループ (Observe-Orient-Decide-Act) ループと似ています。現在の実験科学のための AI やロボットは、これらの各ステップのいずれかに注目して行われています(図 1)。

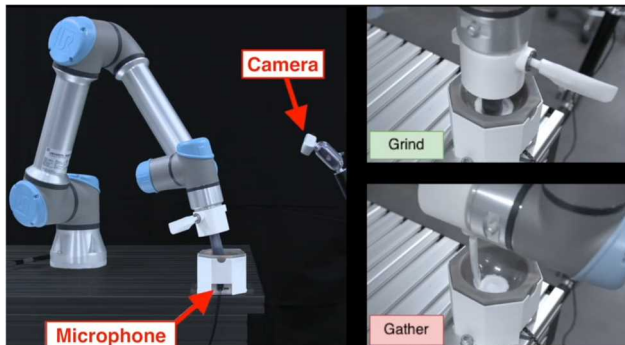


図 1 材料科学の研究開発のための粉混ぜロボット。目や耳といった複数の感覚情報を活用しており、実験の準備だけでなく混ぜて反応を促すような用途も期待されます。

本研究プロジェクトは、ループ全体を担う AI ロボットを開発することで、自らループを通して学び、人間の研究者と共進化できるような AI サイエンティストを実現します。

## 2. 研究開発の概要及び挑戦的な課題

現在、ChatGPT に代表される大規模言語モデルが注目を集めています。このバックボーンになっている GPT (Generative Pre-trained Transformer) も含めた「基盤モデル」と呼ばれるものが、その多様な使い方で注目されています。では、なぜ多様な使い方ができるようになったのでしょうか。それは、あらかじめ大規模なテキストデータの単語の並び(言語モデル)を学習させた上で少しだけ入出力の例を教えると、多様なタスクを十分な精度で推論できることが分かったためです。

本研究プロジェクトで目指す AI サイエンティストは、なるべく少数の実験科学のデータや、人間の研究者からの教示をもとに賢く学習できる必要があります。また、先ほど述べたようなループの多種多様なプロセスをこなす必要があります。



図 2 科学用基盤モデルの概念図。汎用の基盤モデルを科学用途にさらに発展させて、AI サイエンティストの諸機能を効率的に学習させることを目指しています。

そのためには、論文や特許、書籍などの文献や実験データ、そして人間の研究者からの教示などの多種多様なデータから学習できる科学用の基盤モデルが必要です(図 2)。

またもちろん、テキストだけでなく図表や実験データなどの多種多様なデータ(マルチモーダル)を入出力できる必要がありますし、AI やロボットを通じて実世界の実験や人間の研究者とのインタラクションを行う必要があります。そのためにも、AI サイエンティストは人間の研究者が理解して納得できるような情報提示を行う必要があります、説明性のある AI をさらに発展させた、共理解可能な AI (Co-understandable AI) が必要です。

## 3. 今後の展開

本研究プロジェクトでは、2025 年までにまず人間の研究者が行う研究を文献や実験作業を通じて学び、文献のレビューや追実験が行える AI サイエンティストを実現します。

その後 2030 年までに、AI サイエンティスト自身が実験結果や文献から仮説形成(アブダクション)に至って新たな研究の主張を人間の研究者と企画し、実験計画と実行、そして結果の解析を担い、資料としてまとめたり論文文化したりして再び研究者に提示するといった、新たな実験科学の知見を切り拓く機能を実現していく予定です。