

AIロボット駆動科学とAGI

理化学研究所 科学研究基盤モデル開発プログラム(AGIS)
プロジェクトディレクター

AIロボット駆動科学イニシアティブ 代表理事

高橋恒一

AIとロボットを用いて実験科学を構成する様々なプロセスを
一つのサイクルとして統合し、力強く加速することを目論みる新しい方法論

ノーベル＝チューリングチャレンジ： “2050年までに自律的にノーベル賞級の発見を行うAIを開発する”



Workshop Summary developed by Workshop Attendees

**Posing an AI Scientist Grand Challenge:
Artificial Intelligence Systems Capable of Nobel-Quality Discoveries**

Co-Chairs:
Yolanda Gil (University of Southern California)
Ross King (The Alan Turing Institute)
Hiroaki Kitano (Osaka Institute of Science and Technology, The Systems Biology Institute,
and The Alan Turing Institute)

**Sponsored by the Office of Naval Research (ONR Global)
Hosted by The Alan Turing Institute**

1. Workshop Goals
The Alan Turing Institute, supported by the Office of Naval Research (ONR Global), convened a workshop to explore the merits of a grand challenge focused on AI for scientific discovery and how this might be achieved. The meeting was organized by Ross King, Yolanda Gil, and Hiroaki Kitano, bringing together three continents, in London on 26-27 February 2020.

Scientific discovery has long been of interest to AI researchers. Herbert Simon (Nobel Laureate and Turing award winner) worked on cognitive aspects of scientific discovery in the 1960s [1]. Edward Feigenbaum (Turing award winner) worked with Joshua Lederberg (Nobel Laureate) and other colleagues in Stanford in the 1960s on automating the identification of organic molecules and learning rules about their mass spectra [2].

Now decades later, driven by vastly more powerful computer hardware and software, AI is having an ever-increasing impact on science [3].

The workshop co-chairs brought forward complementary long-term visions for AI in science: robot scientists that achieve closed-loop automation in terms of hypothesis generation, experimental design, execution and verification [4], the concept of thoughtful AI systems that can be effective partners for scientific discovery [5], and the intriguing proposal for a grand challenge for AI and systems biology to develop AI systems that can make major scientific discoveries significant enough to be worthy of a Nobel Prize [6].

The workshop was attended by a diverse group of thirty invited participants. Several expected invitees had to cancel their travel in the last minute due to travel restrictions related to COVID-19. Twenty participants came from universities, five from industry research labs, and five from government. Nine were from the UK, fifteen from the US, and five from Japan, and one from Europe (outside the UK). Some participants were AI researchers and practitioners with direct

AI Scientist Grand Challenge: Summary of Discussion February 2020 3

2020年2月にアラン・チューリング研究所で開催された日米英ワークショップで
グランドチャレンジの設定が合意。（日本のムーンショット目標の設定にも影響）

以降、国際ワークショップシリーズが進行中

オーガナイザー： 北野宏明、Ross King、Yolanda Gil、Robert Murphy、清田純、高橋恒一など

参加機関： Sony、SBI、理研、OIST、東大、名古屋大、チューリング研究所、Google、DeepMind、CMU、Liverpool大、DARPAほか多数

AIロボット駆動科学とは？ | 世界的な潮流

AIとロボティクスを融合した「AIロボット駆動科学」は、国際的にも注目が高まる新たな研究領域です。

- ✓ 2024年のノーベル物理学賞では、機械学習研究の功績でジョン・ホップフィールド氏とジェフリー・ヒントン氏が受賞。同年のノーベル化学賞では、AIによるタンパク質構造予測「AlphaFold」が選ばれ、AIがすでに科学の最前線にいることを象徴づけました。
- ✓ 人間の直感に頼らず自律的に知を探索するこの研究領域は、研究の効率化にとどまらず、科学の在り方そのものを変えようとしています。

各国の主な動き

アメリカ

- 米国では、AIとロボットの融合による「自律型研究ラボ（self-driving lab）」の開発が産学連携を中心に進展しています。
- 国立研究所やNSFによる研究支援があり、カーネギーメロン大学とEmerald Cloud Labの連携による完全リモートAI統合ラボの構築など、民間投資も含めた多様な取り組みが展開されています。

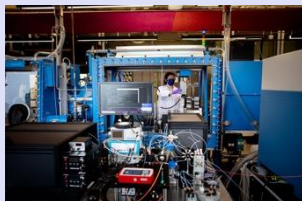
中国

- 中国科学技術大学では、「AI-Chemist」と呼ばれるシステムが、実験室で実際の化学実験を自律的にを行い、化学研究の自動化を推進。AIとロボット技術を組み合わせることで、研究プロセスを効率化し、新たな発見を追求する取り組みが加速しています。



カナダ

- トロント大学などのAcceleration Consortiumは、self-driving labの運営にCanada First Research Excellence Fundから7年間で2億カナダドルの支援を受け、材料発見を加速させています



イギリス

- リバプール大学とインペリアル・カレッジ・ロンドンが主導するAIChemylに、総額1,200万ポンド（約21億円）を投資。英国初の学術・産業連携によるAI化学研究機関であり、デジタル化学の世界的リーダーを目指す英国の国家戦略の一環として推進されています。



「複雑なものをどう扱うか」が21世紀科学の挑戦

数学の活用が物理の謎を解く鍵だったように、
AIの利用が生命の謎を解く鍵になるだろう
ー デミス・ハザビス
(慶應医学賞2024、ノーベル化学賞2024)





近代科学のアプローチで躍進

数学的極限操作により物事を
「単純化して理解する」



21世紀の科学の課題

非平衡系では正準アンサンブルが成立せず要素の「個性」が表出
「単純化して理解する」ことが難しい

新たなアプローチ

世界を単純化して理解する



複雑な世界を複雑なまま捉える

最小限の手数で
きれいに仮説を立証する



多様な高品質データで不確実性から
本質をあぶりだす

近代科学は
分野の専門分化の歴史



専門性を超えた総合知的アプローチ

三つの具体的課題

課題

AIロボット駆動による解決

課題 1 プロセス の断絶

実験、データ処理、仮説構築などの
研究プロセス間のより強固な接続が必要。

→ AIとロボットで研究プロセスを統合

一つのサイクルとして自動化し力強く加速

課題 2 分野 の断絶

異なる分野間の知識の連携が不十分。

→ AI基盤モデルで知識を集約、統合

複雑に絡み合った全体を俯瞰し、
個別細分化した科学分野を再統合。

課題 3 人間の能力 の限界

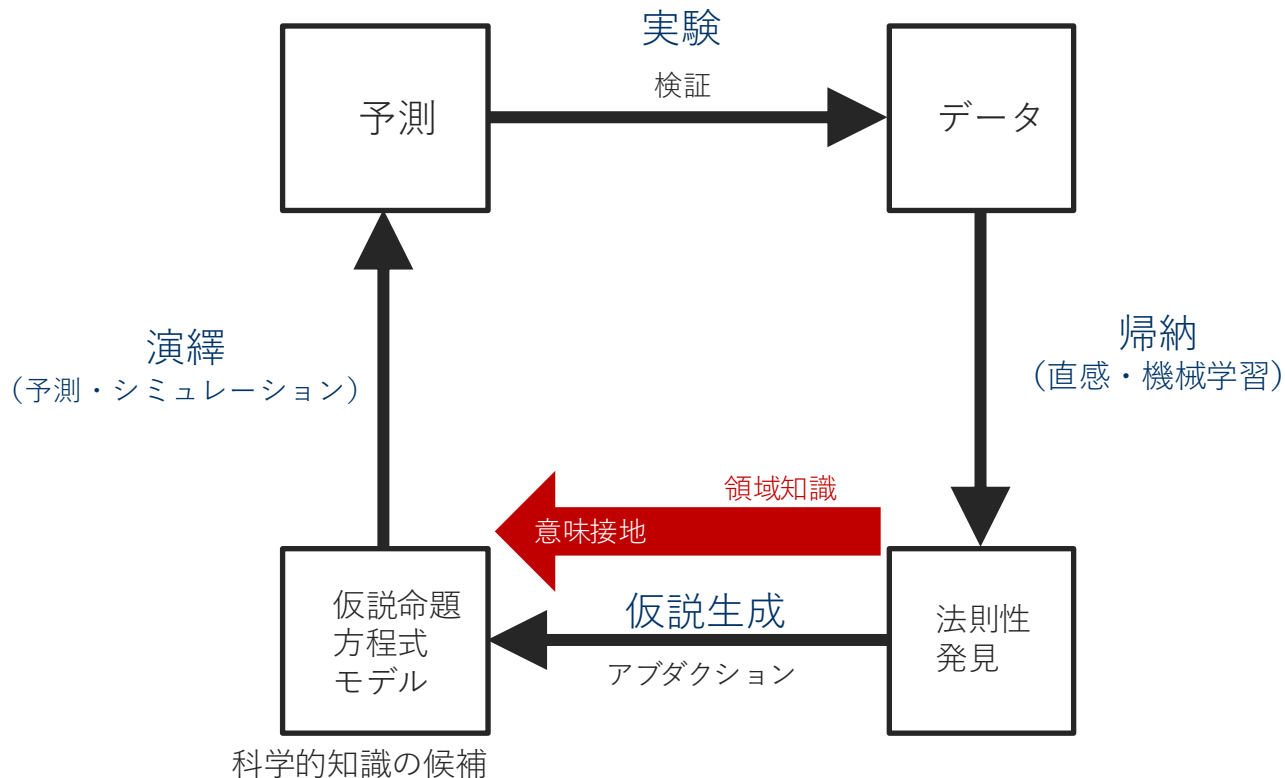
人間の実験手技や認知能力には限界がある。

→ ロボットによる実験の限界突破
→ AIによる認知能力の限界突破

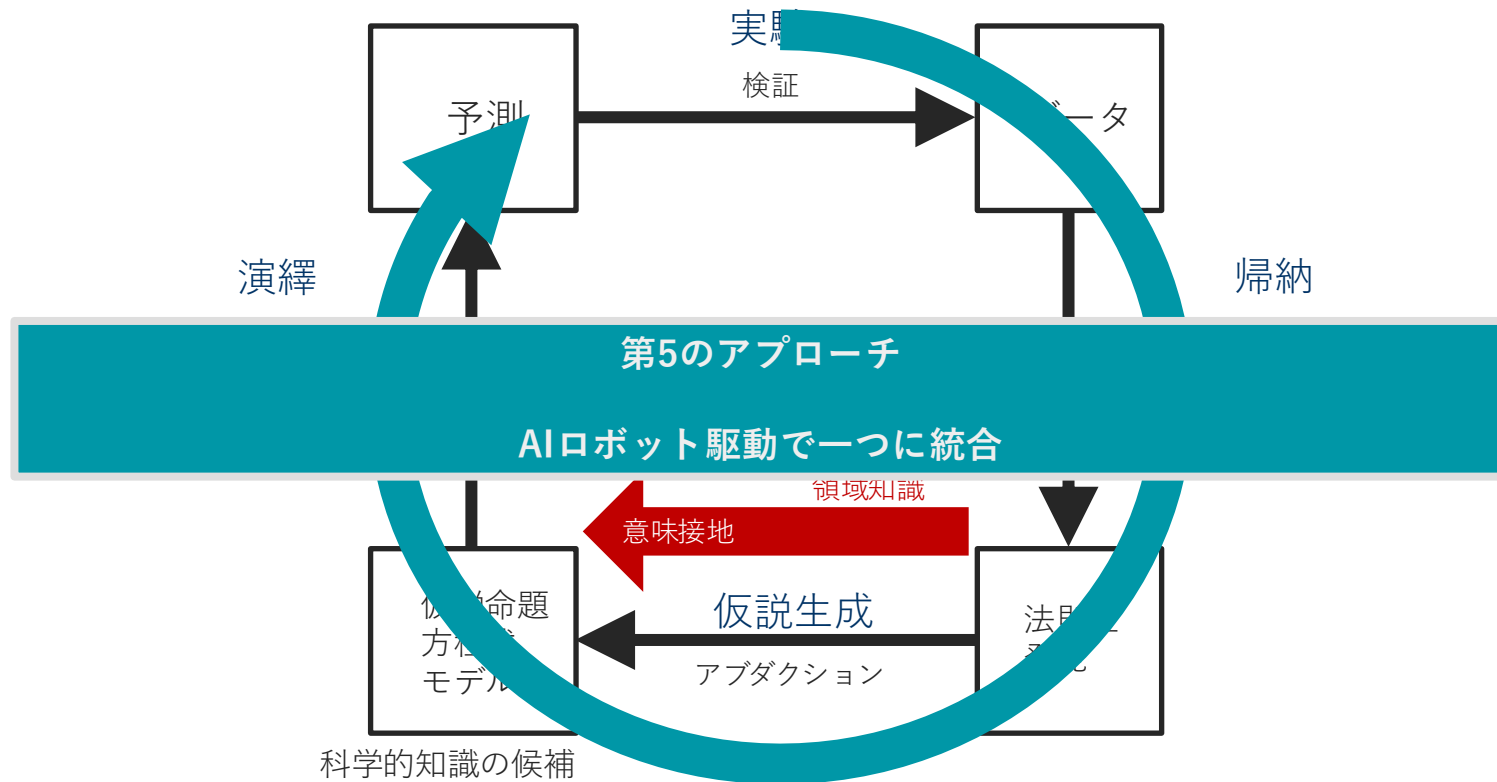
ヒトとAIロボットの協働により
これまでより広く、深い科学を実現。

課題1 (プロセスの断絶) の解決:

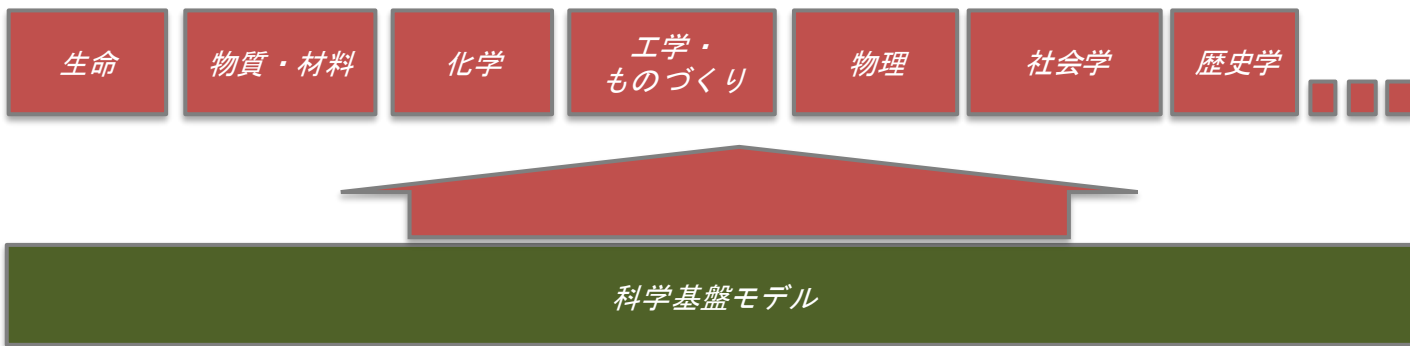
実験科学を構成するプロセスの統合



課題1（プロセスの断絶）の解決： 実験科学を構成するプロセスの統合



課題 2 (分野の断絶)の解決： 基盤モデルによる個別細分化した研究ドメインの再統合



近代以降、科学分野は個別細分化の歴史を辿った。

人類が書き溜めた言語情報の総体を表すもの = 大規模言語モデル (LLM) であるように、

人類の科学的知識の総体を体現するもの = 科学基盤モデル を作れるはず。



細分化した学問分野を再統合して「One Science」を実現したい！

課題3 (人間の限界)の解決： AIロボット駆動による探索可能領域の拡大

時空間スケール： マイクロ・マクロ

TeV以上のエネルギースケール

人の手技では届かない超絶技巧

超並列実験

宇宙・極限領域での自律実験

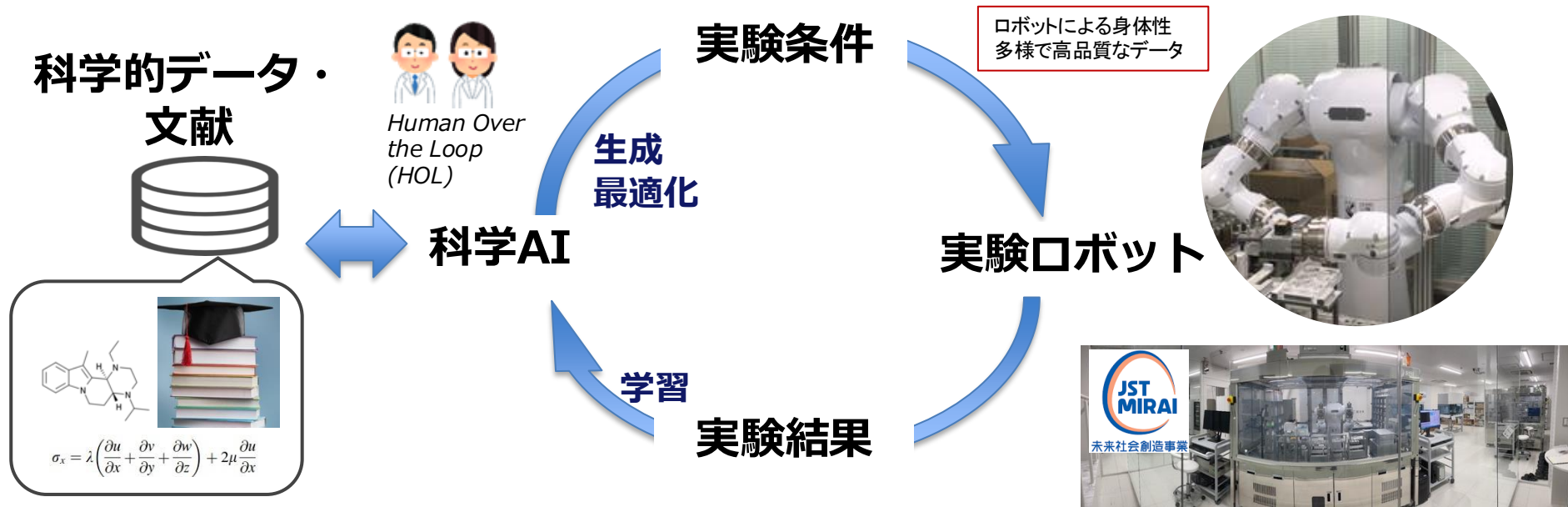
人類知のフロンティア拡大

人文・社会・自然科学の融合



「新たなアプローチ」を可能にする基盤こそがAIロボット駆動科学

AIとロボットを用いて実験科学を構成する様々なプロセスを一つのサイクルとして統合し、力強く加速することを目論みる新しい方法論（第五の科学的方法論）



複雑なまま捉える

→基盤AIによるマルチモーダルデータ統合

多様なデータから本質をあぶりだす

→ロボットによるバイアスのないメガ実験

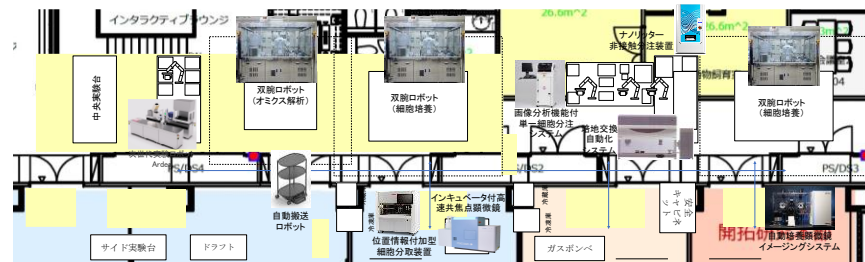
専門を超えた総合知的アプローチ

→AIロボット駆動による人間中心の研究加速

理研神戸地区における実験自動化研究



ロボティック・バイオロジープロトタイピング・ラボ
(JST未来社会創造事業)



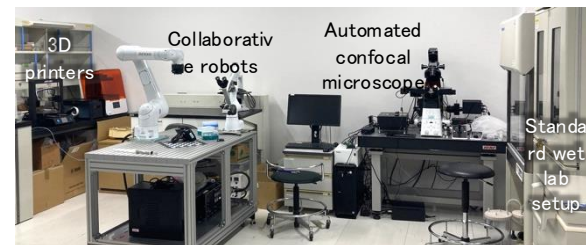
科学研究基盤モデル開発プログラム 自律ロボットラボ



LabDroid まほろ
RBI株式会社



Ardea: 次世代自動実験ロボット
TRIP-AGIS



BDR AXバイオロジー プロジェクト
コア・ラボラトリー

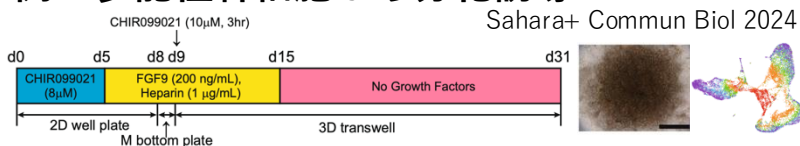
TRIP-AGIS 能動的データ創出共通基盤(自律データ工場)



TRIP-AGIS 能動的データ創出共通基盤 (自律データ工場)

細胞応答モデル開発を想定、エスノグラフィ調査 & 整備中

例：多能性幹細胞から分化誘導



多種多様な実験条件

- ・ 添加因子 (種類、濃度、タイミング、期間)
- ・ 培養基材 (種類、タイミング、期間)
- ・ 初期細胞塊 (大きさ、形状、微細構造) etc.

実験量：探索数のべき乗で増加 $N=x^y$

Read-outの多様性

- ・ 遺伝子発現 (直接的、単一細胞の感度、**破壊的**)
日毎の遺伝子発現→日数分サンプルが必要
- ・ タンパク質発現 (直接的、バルク解析)
- ・ 生化学分析 (間接的、バルク解析)
- ・ イメージング (**継時的**評価可能、定量性限定的)

毎日違う実験をスマートにこなす必要性

プロトコル分析：モジュラリティやスケールなどを調査



Cell culture, cell differentiation, micro-operation, gene expression assay, optogenetics...

データ創出のための実験機器を配備



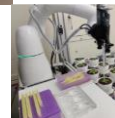
- ・ RBI Maholo (Cell culture x2, Multiomics x2)
- ・ Opentrons Flex
- ・ Cellenion CellenONE HT
- ・ Tecan Fluent w/ ThermoFisher Cytomat
- ・ Nikon BioPipeline Live
- ・ Automated Thermal Cycler & qPCR x4
- ・ Beckman Echo 655
- ・ Yokogawa CQ3000
- ・ Yamaha Motor CELL HANDELR 2
- ・ Rorze CellKeeperII 48 Plus, etc.

実験の”進化”に向けた取り組み

- ・ 各機器のHands-onセッション
- ・ 目的に合わせた実験の提案、実装
- ・ AI、ロボティクスなど自動化技術の開発
- ・ 国際標準や業界動向の調査、産業界・DMPとの対話



SiLA
Rapid Integration



AIロボットによるiPS細胞分化誘導条件の自動発見

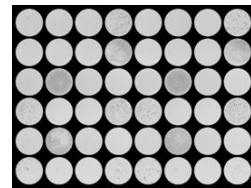
暗黙知を写し取ったAIロボットが自律的に至適分化誘導条件を発見



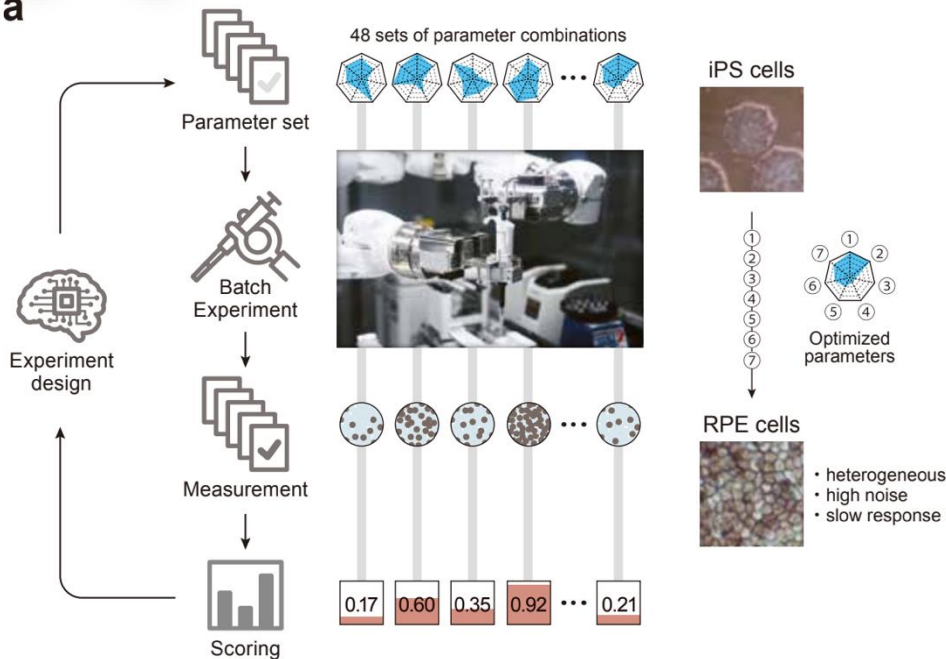
産経新聞

高橋政代

神戸アイセンター病院

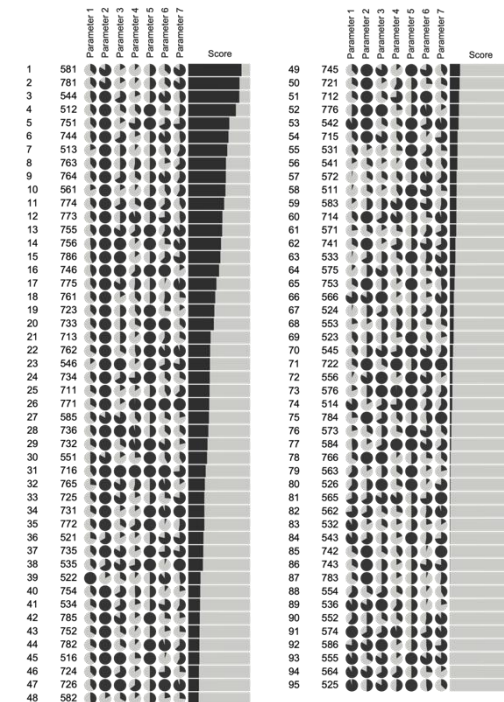


a



Round 5+7 Parameter vs Score: Visualization

Type: Circle

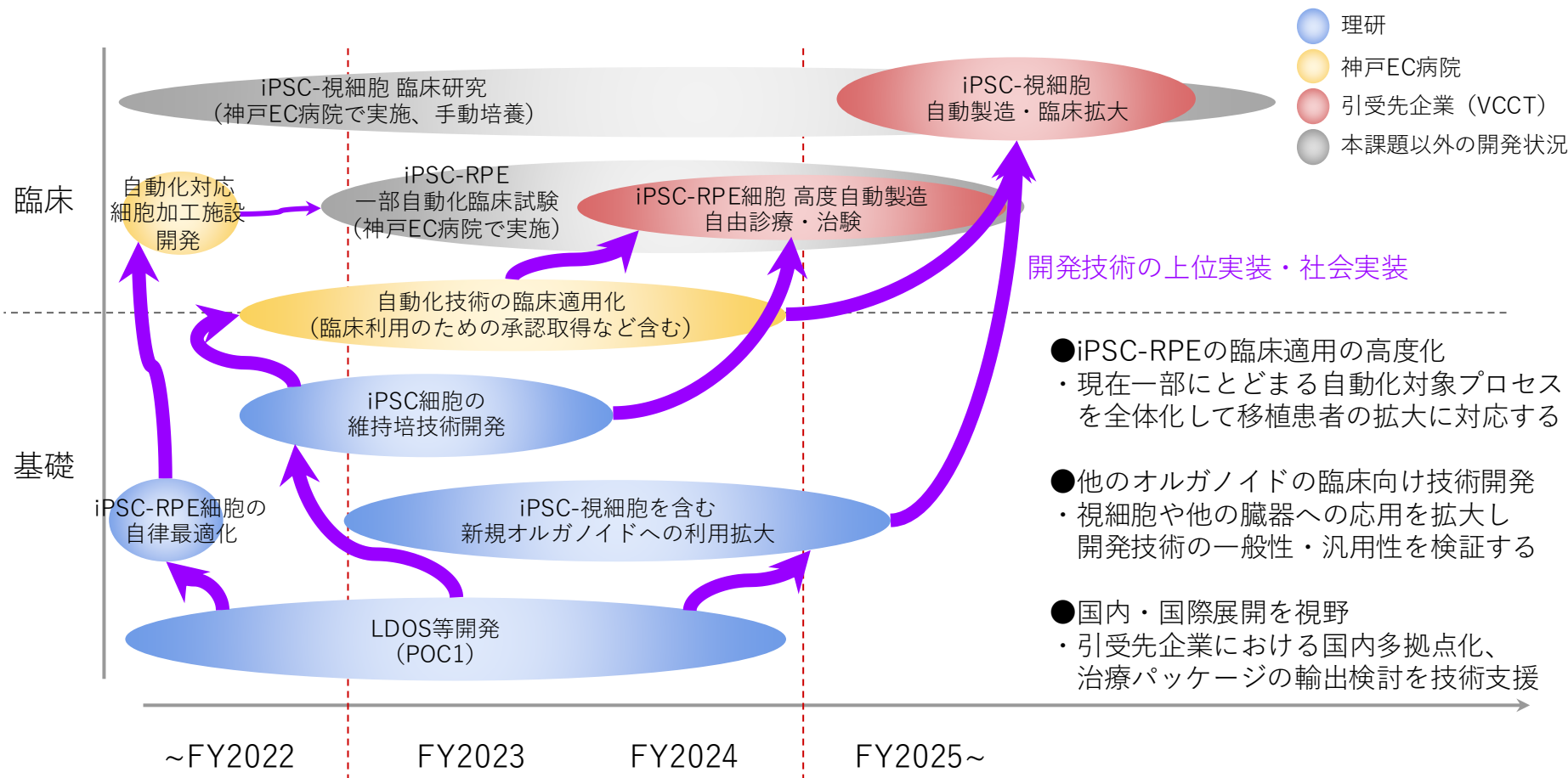


自動実験計画AI： バイオ実験の特性である高ノイズ、高並列、高コスト性に対応した
ベイズ最適化手法(Batch Contextual Local Penalization)を開発。

Epistra

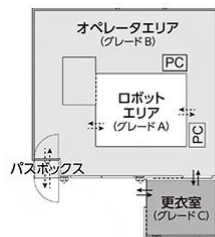
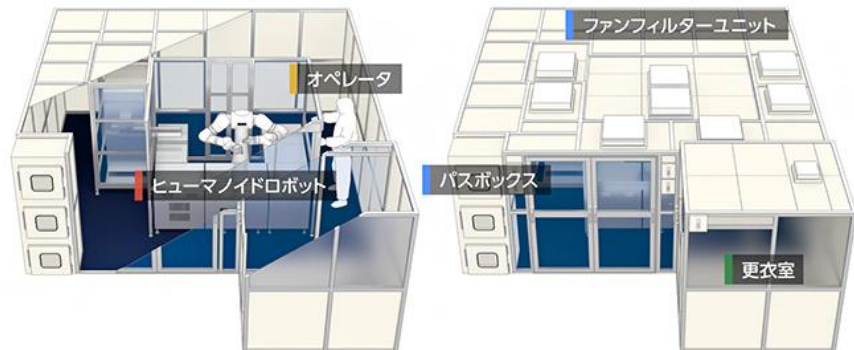
- 三億通りの可能性からAIが約150の候補条件を絞り込み、ロボットと連携して実験検証
- 専門家が年単位の時間をかけた研究プロセスを月単位に短縮(約10倍の研究加速効果)

細胞培養・再生医療のロボット化



ロボット用細胞培養加工施設（R-CPF）を開発（世界初）

Terada et al., SLAS Technology, 2023. (23/11/30プレスリリース)



清浄度目標（一部抜粋）				
清浄度レベル	浮遊菌	落下菌	付着菌	
ロボットエリア	A	<1	<1	<1
オペレータエリア	B	≤10	≤5	≤5
更衣室	C	≤100	≤50	≤25

(単位) 個/m³ 個/プレート



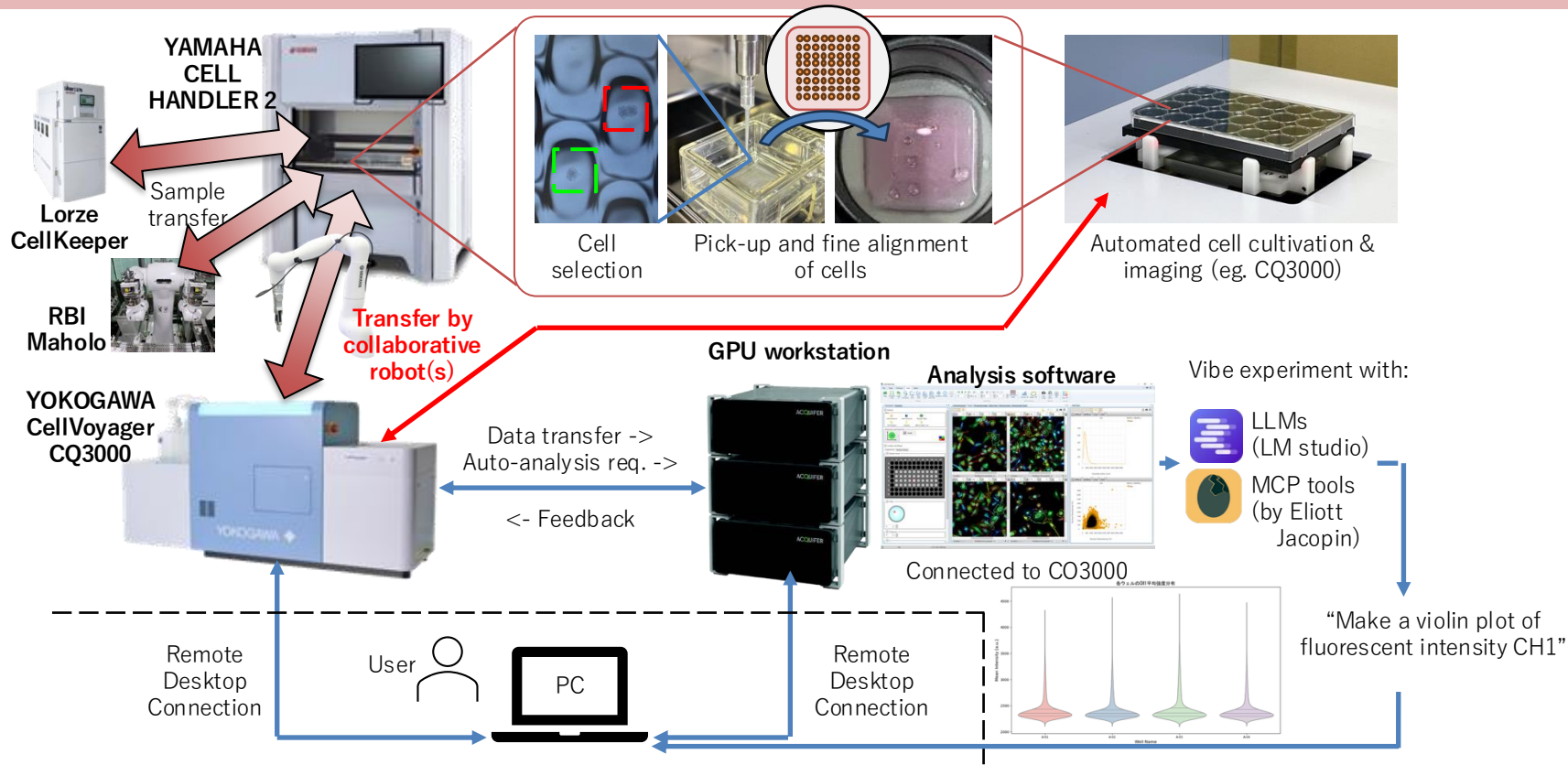
空気齢 (秒)

多 ← 換気回数 → 少

- 汎用ヒト型ロボットの作業環境を無菌化することに成功
 - ・ 汎用ヒト型ロボットとエアバリアブース（ダイダン社製）と組み合わせることで達成
 - ・ 汎用ヒト型ロボットもエアバリアブースも拡張性・汎用性に優れるためプロトコルの変更が容易
 - ・ 清浄度管理区域環境モニタリングにより臨床研究に必要な無菌環境を全区画で達成していることを実証
- 基礎研究のロボット動作がそのまま臨床の現場へ
 - ・ ロボットラボで作成・検証したロボット動作を臨床研究でそのまま利用可能となった
- 細胞加工施設従事者の教育コストを圧倒的に削減
 - ・ 導入前：匠に匹敵する「10年選手」のみが従事可能
 - ・ 導入後：匠の技術は全員には不要：1年目から従事可能
- 2022年供用開始済
 - ・ 2022年2月：臨床研究の実施に必要な行政の認証を取得
 - ・ 2022年12月：臨床研究での利用を開始
- 国内・国際展開を視野
 - ・ 神戸アイセンター病院ならびに株式会社VC Cell Therapyにおいて国内多拠点化、治療パッケージの輸出を調整中

Ongoing prototyping of vibe research workflow

Example prompt: "Isolate cells suitable for generating kidney organoids and spot them onto a 24-well plate in a 1 mm grid. Then, transfer the plate to the high-content imaging system CQ3000 to observe the 3D morphology of each cell group every 6 hours."



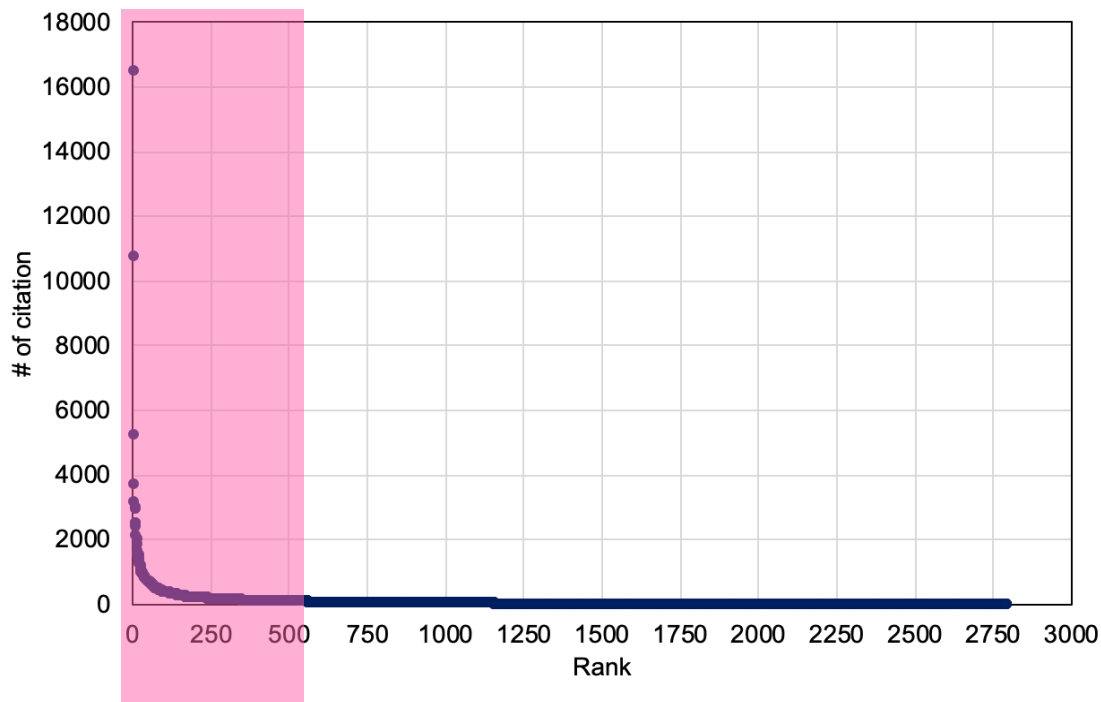
Integration ongoing in partnership with Yamaha, YOKOGAWA, etc.

生命科学実験の分散開発に対応した次世代型実験ロボットシステムの開発



次世代型実験ロボットシステム初号機（IIB N303に既設。開発コードネーム：Ardea）

生命科学実験プロトコルのロングテール性



Nature Protocol誌
(2006年創刊)

全論文数 2,792

全引用数 277,877

上位 20% の論文は
引用数の 75% を占める

ショートヘッド

ロングテール

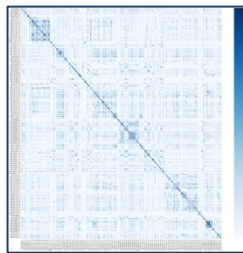
⇒ 分散開発の必要性・有効性

実験記述言語LabCode

オブジェクト・フロー・プログラミング(OFP)パラダイム:
データフロープログラミング(DFP)の実世界拡張

特徴

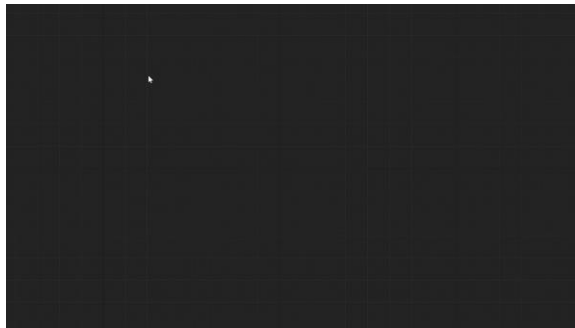
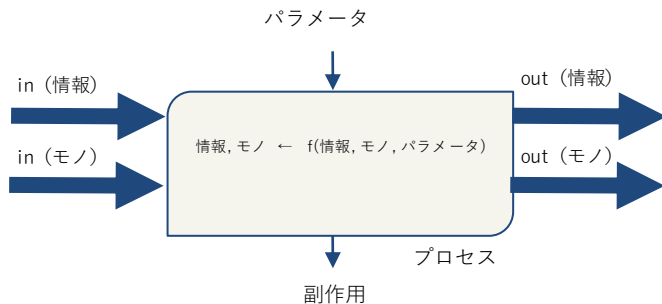
- ・実験におけるモノや情報の流れを直感的に表現する 高級言語
- ・フロー記述とロボット動作の詳細を分離
- ・概念的にはデータフローの上位互換。
既存のワークフローエンジンと連携し、ウェット実験と
データ処理、シミュレーション実行等をシームレスに記述可能
- ・DFPは数学的には関数型プログラミングと等価。
OFPにも同様の対応関係。様々な数学的テクニックが適用可能
- ・並行性判定が容易なため並列スケジューリングに適する
- ・次世代ロボットArdeaがネイティブ実行環境を提供



実験プロトコルのモジュール性解析

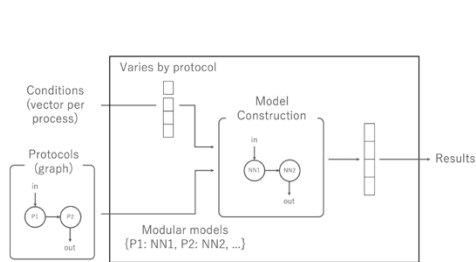
言語設計が前提とする仮説が
成立することを示唆する結果。

OFP: オブジェクトフロー・プログラミング

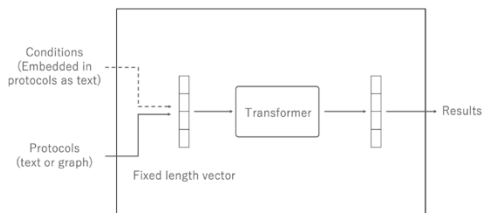


LabCode処理系自体を「知能化」し、未知の実験の結果を予測

- OFPプロセス毎にNNを構成して学習し、toy-problemで未知のプロトコルでも実行結果を予測出来ることを確認

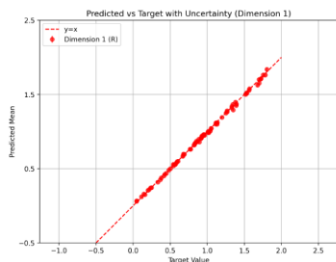
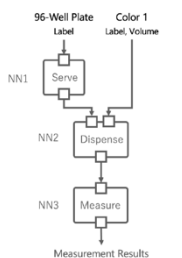


(b) Proposed method



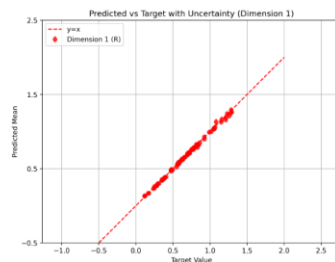
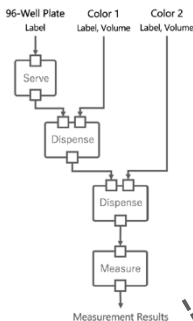
(c) Transformer-based method

Protocol 1: 1-color measure



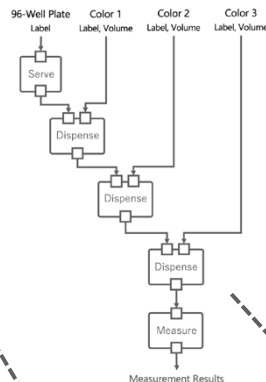
(a) 1-color mixing (学習済)

Protocol 2: 2-color mix

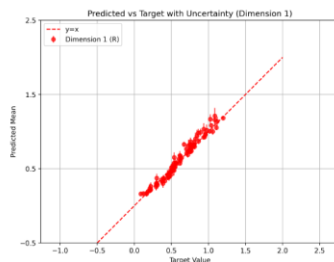


(b) 2-color mixing (学習済)

Protocol 3: 3-color mix



Tecan社自動分注ロボット
Fluent



(c) 3-color mixing (未学習)

コードネーム Ardea: LabCodeにネイティブ対応した自動実験プラットフォーム

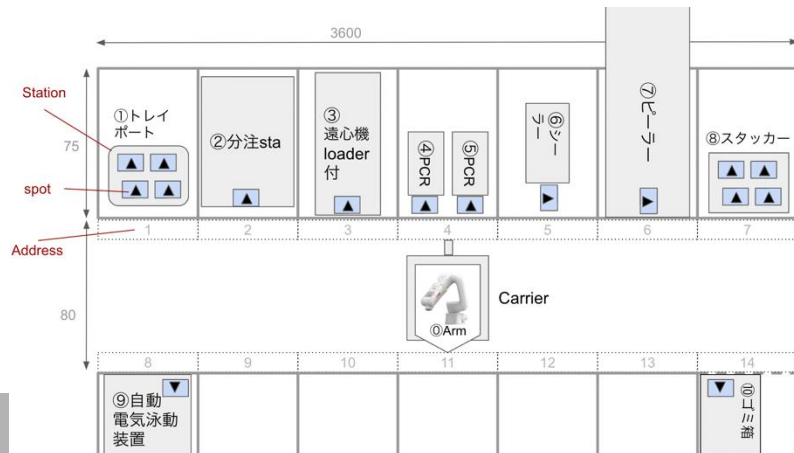


分散拠点用
コンパクト版

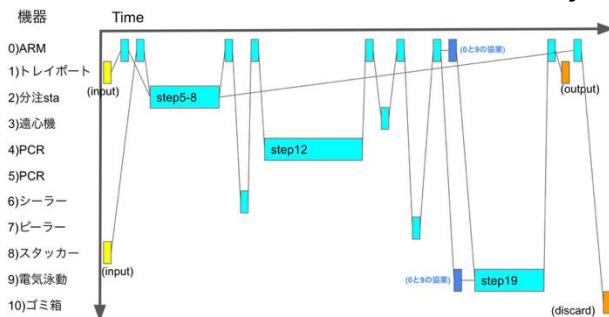


安川電機でまほろを開発した梅野さん（当ラボ技師）がハードウェア設計を担当。

レール付き協働6軸アームの採用により、まほろの宿題だった
(1)ティーチングコスト、(2)プロトコルの可搬性、(3)ソフトウェア構造
の複雑さなどの課題を解決。



station layout



concurrent scheduling

自動実験の集中化と分散化を同時に追求

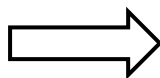
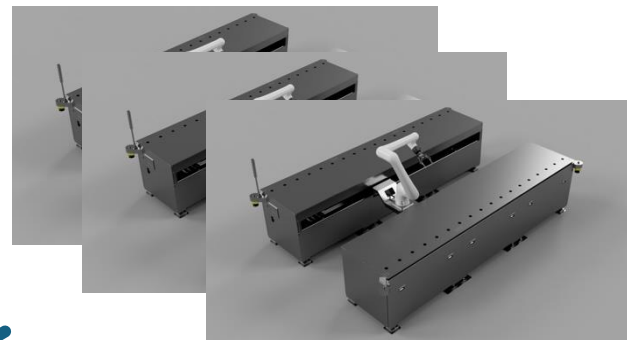
分散拠点 コンパクト

先端研究現場の
高度実験技術を
ロボットプラットフォーム
によりデジタル可搬化



集約拠点 フルサイズ

収集・構築された
実験モジュールを
大規模集中拠点で
高効率実行



例：サンプル前処理



例：シーケンスライブラリ調製

拠点1



トランスクリプトーム

拠点2



エピゲノム

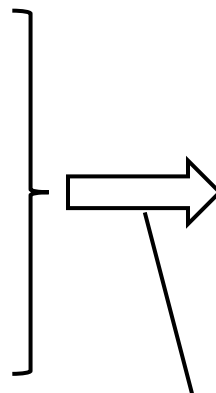
拠点3



メタゲノム

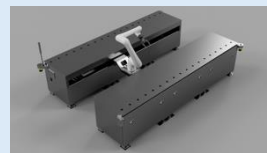
⋮

分散開発



異種間での共有

集約拠点



同一機種間での共有

将来展望： 科学基盤AIと実験ロボットの共進化

- シングルモーダル基盤モデルの学習

- ➡ 従来タイプの専用・ライン型自動化でもOK

- マルチモーダル化と学習のスケール拡大に従って基盤モデルの推論性能が向上し、生成される実験の分野、種類、複雑さが徐々に増す。

- ➡ 実験ロボットの汎用性向上を要求。

- 最終的にはAIが新たな実験や計測装置の提案も行う。

- ➡ 実験ロボットの未来像。(科学AIの自律性レベル4以降)

マルチモーダル化とスケール性という科学基盤モデルの特性は
自律性と汎用性の向上という今後の実験ロボット開発の方向性とマッチする

AIロボット駆動科学イニシアティブ

これまでのシンポジウム・研究会参加プロジェクト（一部）

- ▶ JST未来社会創造事業 「ロボティックバイオロジーによる生命科学の加速」（高橋恒一代表）
- ▶ 同 「マテリアル探索空間拡張プラットフォームの構築」（長藤圭介代表）
- ▶ JSTムーンショット型研究開発事業 「人とAIロボットの創造的共進化によるサイエンス開拓」（原田香奈子PM）
- ▶ 同 「人と融和して知の創造・越境をするAIロボット」（牛久祥孝PM）
- ▶ 理化学研究所 科学研究基盤モデル開発プログラム:AGIS（泰地真弘人PD）
- ▶ JST さきがけ 「研究開発プロセス革新」領域（竹内一郎研究総括）
- ▶ JST CREST 「バイオDX」領域（岡田康志研究統括）
- ▶ 経済安全保障重要技術育成プログラム「人作業伝達」プロジェクト
- ▶ 学術変革領域研究 「学習物理学の創成」（橋本幸士領域代表）
- ▶ 北海道大学 化学反応創成研究拠点（前田理拠点長）
- ▶ 立命館大学・京都大学 記号創発システム科学プロジェクト（谷口忠大代表）
- ▶ 一般社団法人AIアライメント・ネットワーク

JST研究開発戦略センター、科学技術・学術政策研究所、産業技術総合研究所、国立情報学研究所、三菱総合研究所 ほか

Mission

活動理念

科学研究のプロセスをAIとロボットで再定義し、科学の進展を飛躍的に加速する

Vision

活動方針

分野や組織の壁を越えた国際的な研究ネットワークの中で共創を実現し、科学者の知的好奇心を尊重しながら、人々の豊かさにつながる科学・技術のあり方を追求する

Activities

行動指針

産学官をつなぐハブとして、具体的かつ多面的な活動を推進する

- ✓ 啓蒙とコミュニティ形成: シンポジウムや研究会を定期開催し、AIロボット駆動科学の重要性を社会に訴えと共に、多様な才能が出会い、対話するコミュニティを形成。
- ✓ 分科会活動: AIロボット駆動科学の基礎論として重要なメタサイエンス（Science of Science）のほか、研究分野やベンダーの壁を越えて機器・データ・AIが円滑に連携できる、開かれた技術標準やベンチマークの議論を主導。
- ✓ 政策形成への貢献: 現場の知見と専門的な議論を集約し、日本の科学技術競争力を高めるための実効性のある政策提言を実行。
- ✓ 国際連携の窓口: Acceleration Consortiumなど海外の先進機関と連携し、グローバルな研究協力の窓口としての役割を担う。

AIロボット駆動科学イニシアティブ | 分科会について

AIロボット駆動科学の基盤となる接続性・理論構築・評価軸づくりに着手します。
メンバーの議論の中で柔軟に分科会を立ち上げます。下記4つの分科会形成を進めています。

コネクティビティー・アーキテクチャー分科会

AI・ロボット・実験装置間を統合する通信方式・連携手順を検討。
SiLA2など既存標準を踏まえつつ、AI連携・ロボット間連携を含む
AIロボット駆動科学ワークフロー全体の標準化を推進。

ワークフロー記述分科会

仮説生成から実験実行・解析までのプロセスを交換可能な形式で
記述。分散開発・共有を可能にし、まずは実験プロトコル記述形式
の標準提案に注力。

メタサイエンス分科会

AI駆動科学を「科学の科学（science of science）」の観点から
再定義。科学活動の知的プロセスを情報システムとして実装するた
めの基礎理論を構築。メタサイエンス協会・Science of Science研
究会と連携し、政策提言も視野に活動。

ベンチマーク分科会

AIとロボットそれぞれに対するベンチマークの策定とコンペティション企
画を実施。AIロボット駆動科学における性能評価と研究促進の基
盤を整備。

- AI駆動時代に重要になる「科学の科学（メタサイエンス）」と科学哲学
- 知的労働が自動化された後の人間の役割と経済社会構造のデザイン
- AIアライメント問題
 - 人とAIの目標設定の整合性
 - AGIの出現間近に求められる新たな研究課題
- サイエンスコミュニケーションの理想形はどう変化するか？

AIロボット駆動科学の重要性は広く認識され、国からの投資も始まっている。

その一方で、

- 1 科学技術研究加速の本来のポテンシャル発揮
- 2 産業界のR&D現場への幅広い展開
- 3 関連成長市場の顕在化

にはまだ十分に至っていない。

足りないものは何か？



本イニシアティブがを見つけ、対応して動きたい

AIロボット駆動科学イニシアティブ設立の思いと参画のお誘い

我々は、2023年に第一回シンポジウムを開催して以来、AIロボット駆動科学の重要性を訴えてきました。

この科学的方法論と産業構造の大転換点にあたり、何が必要なのかを見つけ、異なる専門性が繋がり、実行してゆく場として、このたび一般社団法人AIロボット駆動科学イニシアティブを設立しました。

イニシアティブを創設したのは我々ですが、イニシアティブは我々のものではなく、分野共有の財産です。

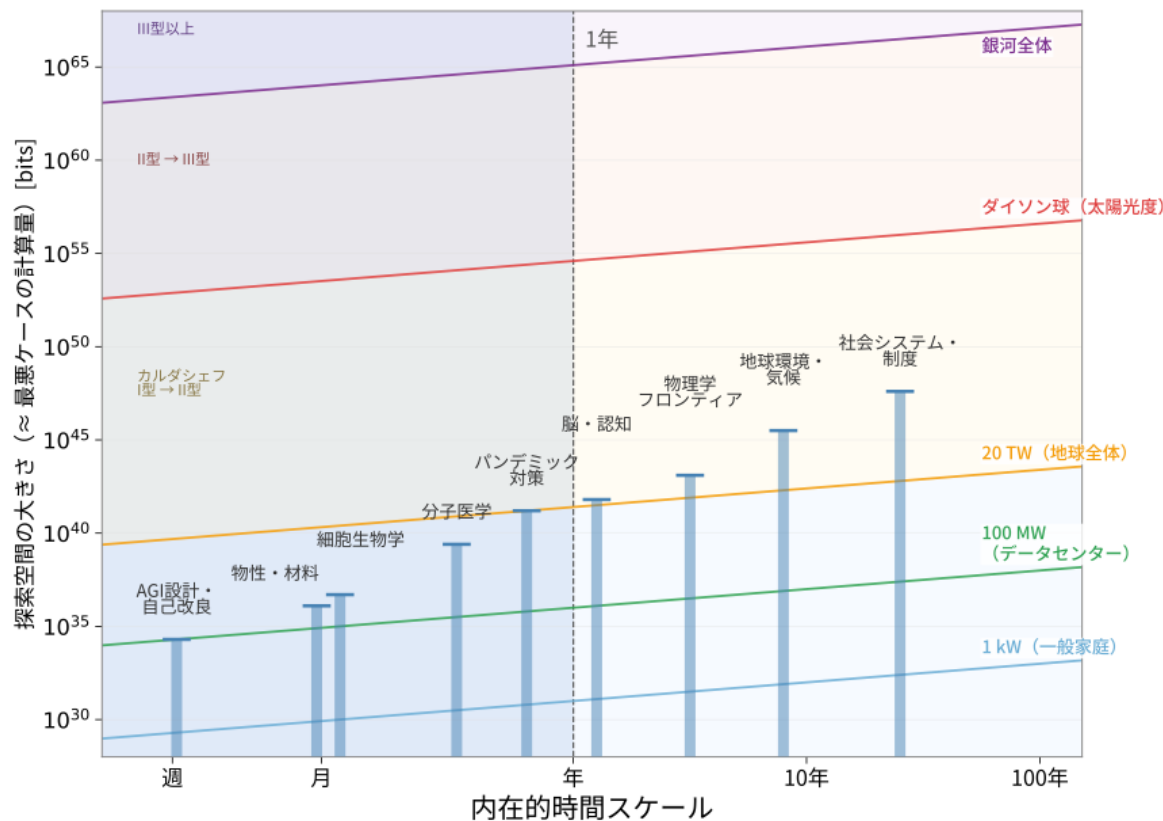
企業メンバーシップ、個人メンバーシップ、分科会メンバー、研究会参加者など、研究者、企業関係者、政策関係者、起業家まで幅広い間口で参加形態を設けて参ります。

是非ご参画下さい。

AIロボット駆動科学イニシアティブ発起人

高橋恒一・長藤圭介・牛久祥孝・原田香奈子

AGIが1年で到達可能な科学フロンティア上限の試算



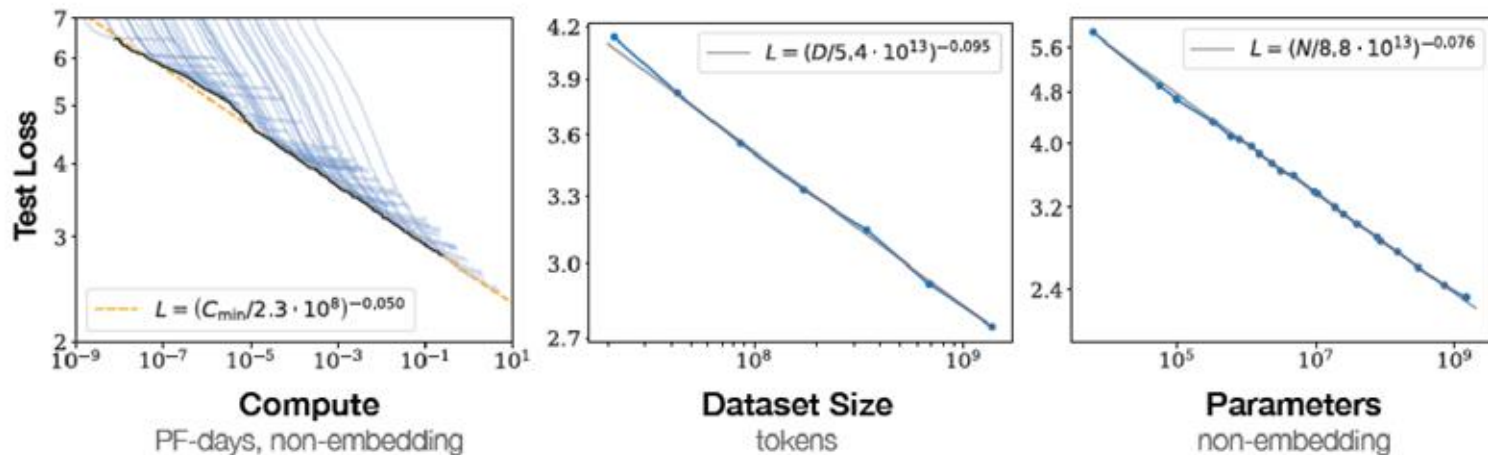
試算前提 (1)AGIは自己改良によりランダウアー限界($k_B \ln 2/\text{bit}$)に100%まで漸近。(2)自動実験により対象への摂動・観測がタイムラグなしに可能。
 横軸：研究対象に内在する時定数。 縦軸：研究対象のコルモゴロフ複雑性概算値(log10 bits)からLevin探索による最悪ケース計算量試算
 1年間で到達可能な領域を色付け (消費電力：kW 一般家庭 -- MW -- データセンター -- GW -- 地球全体 -- ダイソン球 -- 銀河全体)

基盤モデル ⊃ 大規模言語モデル(LLM) ⊃ Chat-GPT, Gemini, Claude....

LLMや基盤モデルがAGIへの道筋であるかは意見が分かれる

⇒ 基盤モデルの性能向上がどうして起きているかを理解することが重要

大規模言語モデルのスケーリング則



J. Kaplan et al., Scaling Laws for Neural Language Models. <https://doi.org/10.48550/arXiv.2001.08361>

スケーリング則：2020年に発見。

モデルの規模、計算量、データサイズを増やすほど、あたかも際限がなく性能が向上しているように見える。

技術的背景：

Transformerと呼ばれるencoder-decoderモデルが再帰などの並列計算を阻害する仕組みを廃したため、計算力の集約が容易になり、飛躍的に大規模なデータセットでの学習が可能になった。

汎用人工知能 (AGI) と万能知能 (Universal AI)

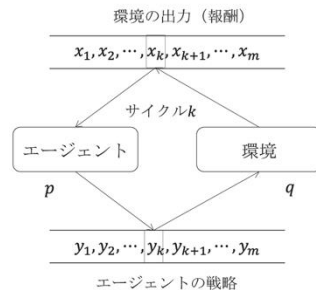
- AGIの定義に全ての専門家の合意はない。
- AGIの理論研究では「Legg-Hutter知能」がよく用いられる。

$$I_{\text{LH}}(\pi) = \sum_{\mu \in \mathcal{M}} 2^{-K(\mu)} V_{\mu}(\pi)$$

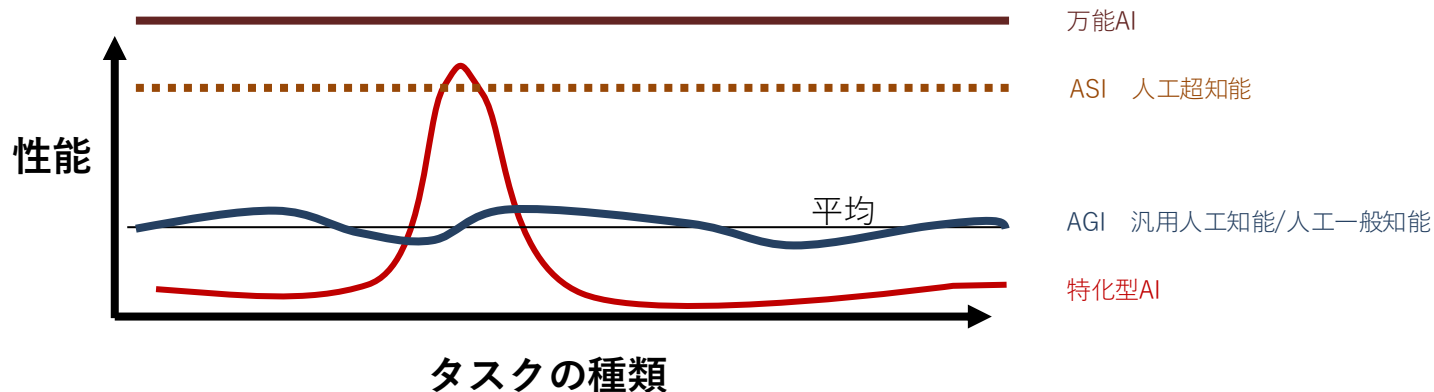
知能 (I) = 環境 (μ) のコルモゴロフ複雑性 (K) で重みづけされた総報酬 (V) の期待値

- Legg-Hutter知能が計算可能な全ての環境に対してパレート最適となるのが**万能AI**
- 万能AIのモデルであるマークスハッターのAIXI(AI ξ)は、ある環境における強化学習を用いた逐次決定により期待総報酬を最大化するAIエージェント (Hutter 2000).
 - AIXIは万能チューリングマシン上におけるソロモノフ帰納推論として実行される。
※ソロモノフの帰納推論：最良の科学モデルは実験データを再現する最短のプログラム（コルモゴロフ複雑性の最小化）であるとする理論で、オッカムの剃刀の数理的な表現とも言える。
 - ただし、ソロモノフ帰納推論は計算不可能であることが証明されている。
 - 2023年に自己予測に基づいた近似モデル(Self-AIXI)が提案 (Catt et al., NeurIPS 2023)
- AIXIは情報圧縮を知能の代用として用いる
 - 過去の経験から帰納的に総報酬の期待をコルモゴロフ複雑性で重み付けしている。

$$a_t := \arg \max_{a_t} \sum_{o_t r_t} \dots \max_{a_m} \sum_{o_m r_m} [r_t + \dots + r_m] \sum_{q: U(q, a_1 \dots a_m) = o_1 r_1 \dots o_m r_m} 2^{-\text{length}(q)}$$



特化型AI、AGIと万能AIは ノーフリーランチ定理を介して関連している。



.....コスト関数の極値を探索するあらゆるアルゴリズムは、
全ての可能なコスト関数に適用した結果を平均すると同じ性能となる (Wolpert and Macready 1995)

現代のAIがスケーリングによりAGIと連続している可能性

2024年から2025年にかけて、Transformerの学習が、AGIの理論を構成するソロモノフ帰納※に漸近するという研究結果が4つの研究グループから独立に提示。



現在の巨大AI開発のスケーリング競争がAGI達成に結びつく可能性が否定「できない」。

※ソロモノフ帰納：

1960年代にレイ・ソロモノフが提唱したアルゴリズム情報理論（コルモゴロフ複雑性）に基づく帰納推論の理論的最適解。AGIの理論研究の基盤である「万能AI」の中核原理でもあり、学習能力の理論的上限を定義する。

（なお、「オッカムの剃刀」の数理的な表現でもあり、科学AIとも関連が深い）

参考： 各論文の要点

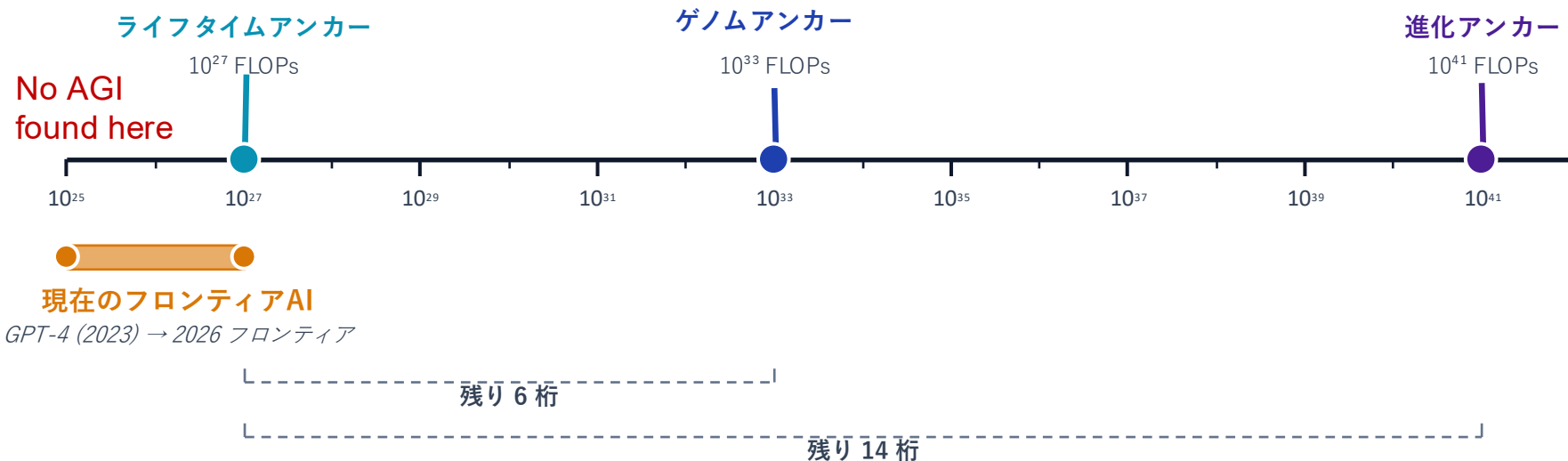
- Grau-Moya et al. 2024 “Learning Universal Predictors”：普遍予測を学習可能な対象として扱い、予測学習がソロモノフ帰納的な理想に近づく条件を示す。
- Young & Witbrock 2024 “Transformers as Approximations of Solomonoff Induction”：Transformerを、理想化条件下でソロモノフ帰納の近似器として位置づける。
- Wan & Mei 2025 “Large Language Models as Computable Approximations to Solomonoff Induction”：LLMを、非計算可能なソロモノフ帰納の有限・計算可能な近似として捉える。

解説記事：

www.aialign.net/blog/transformer-agi-part1



AGIまでの距離（生物学的アンカー）



短期 / SHORT-TERM

ライフタイム アンカー

≈ 10^{27} FLOPs

ヒト1人が生涯に脳で行う計算の総量。AIが2026年に並ぶ射程。

中位 / MEDIAN

ゲノム アンカー

≈ 10^{33} FLOPs

ヒトゲノムに記述された脳の発達情報と神経回路の構築に要する計算量。

上限 / UPPER BOUND

進化 アンカー

≈ 10^{41} FLOPs

神経系を持つ動物が進化した約5億年間に地球上の全動物の神経系が行った計算量。



Thermodynamic Limits of Physical Intelligence 鷗鷗 物理的知能の熱力学的限界

Koichi Takahashi & Yusuke Hayashi | ALIGN / RIKEN AGIS | arXiv:2602.05463

背景

現代AIとヒトの脳には大きな物理的効率の差がある。
Pt並次ィ隘制 (bits / J)% デ欄均りMあィ焔額僅算公デ屏鑑れダ瀧屢N

□□□□ 瀧瀧□□□□+□□□ □□□□□/ノヲウヲ

η_E 認識効率



1 Jあたりに獲得した「環境の構造情報」 (= エビプレキシティー利得)

$$\eta_E = \Delta I(X; Y) / E_{\text{cons}} \quad [\text{bits} / \text{J}]$$

η_C □□□□

エンパワメント毎ジュール

1 J ふをタ = 焔燼りをPt並臺キィ 額焔燼% (= バムゴさちム 半額燼)

$$\eta_C = C(A \rightarrow S_{\{t+T\}}) / E_{\text{cost}} \quad [\text{bits} / \text{J}]$$

命題 瀧瀧両指標の物理的な限界を導出

□□当判□□□□□能

欄欄合瀧算° イMを_E > 沓_C イぶろチシロムヨノナη 算公 (k_B T ln 2 / bit) = 沓ゾチダ。

命題当判閥界の重要性能

瀧燼ち焔制 = イMヲ厂ワ倭燼剪ホメキ默デ噪サPt并焔臺公ィ 焔焔% み焔焔効屢N

量子知能への示唆

- 幼境幼情 (ムンヨ 倭燼) = ソタM焔情匱棘ィ 焔焔然デ額焔焔ヤY = 焔焔N
- 限界の再出現 ただし測定や初期化 (情報の消去) 時には ランダウアー・コストが必ず発生。
- 焔焔ホムーTη ヨィ 焔焔燼算公デPt並ほダ%ィ° イちもM焔焔僅額焔焔キPt焔° も% 焔焔効屢N

意義と展望

- ! Lスケーリングの物理的基礎
焔焔 ヲη メィ 焔焔 = 焔れダ焔焔僅算公デ額焔焔 = 焔焔N
- 持続可能な! L設計の指針
焔焔焔焔 = 焔° も Green AI イ 焔焔焔焔デ焔焔N

科学AIの自律性レベル

<https://doi.org/10.51094/jxiv.818>



Lv.	機械が行うこと	獲得要素	基本問題	駆動技術	取り組み
0	計算	計算	計算		
1	閉鎖系の 解探索	探索	探索	アルゴリズム	AlphaGo(囲碁)、自動定理証明など
2	開放系の 解探索	実験	能動学習	実験自動化 最適化	光触媒自動探索（リバプール大 2020） 理研 iPS細胞分化誘導（2022）
3	形式固定の 仮説生成	知識	知識表現	モデルベース能動学習	酵母システム生物学（Ross Kingら 2019） 理研新領域開拓課題「AI駆動システム生物学」の目標
4	形式自由の 仮説生成	記号	記号接地	マルチモーダル学習 LLMエージェント	機械学習研究（sakana.ai 2024） Google AI Co-scientist（2025） 理研AGISの目標（実験科学研究で実現）
5	研究立案	意味	意味理解	「汎用目的AI」 （ツールAGI）	DeepMind、OpenAI、Anthropicなどがビジョンとして掲げる
6	研究目的の 設定	価値	自発性	「人工一般知能」 （本来のAGI）	理論研究が先行

理化学研究所未来戦略室にて第一版を策定。ノーベル＝チューリングWSのほか、JST-NSF AI&Robotics WS、日本学術会議オープンサイエンス・データ利活用推進小委員会、人工知能学会SIG-AGI、AIP-NISTEPセミナーシリーズ、AAAI CASD 2023などで議論。人工知能学会誌2025年5,6月号に前後編に分けて採録。