

# 次世代計算基盤について： AI for Scienceの観点から

泰地 真弘人 (Makoto Taiji)    [taiji@riken.jp](mailto:taiji@riken.jp)

理化学研究所

生命機能科学研究センター 副センター長  
理事長補佐

## ■ 第4次産業革命 = 知的労働の自動化

▷ 研究活動はその中核で、**価値の源泉**

## ■ 大規模言語モデルの汎用性と 科学研究における有用性

▷ LLMは色々な概念を理解しているように  
ふるまい、適切なタスク選択が可能に

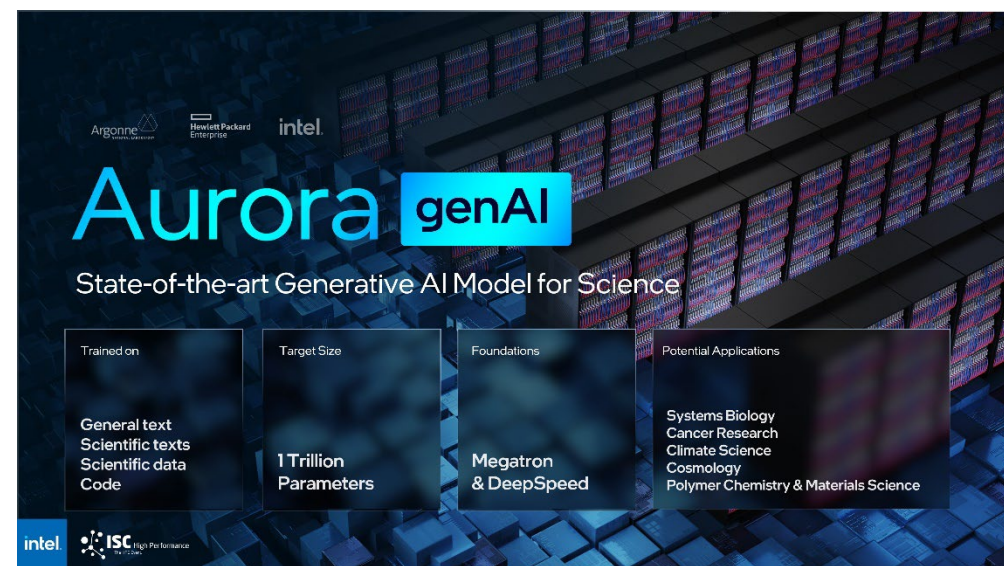
▷ 会話だけでなく、適切な道具や手法を  
能動的に用いることが可能

- 数値的な分析手法・シミュレーション手法などの選択
- 実験の自動化

## ■ 科学の各ドメイン固有の基盤モデルの重要性

▷ 既にタンパク質モデル、ゲノムモデルなどで成果

▷ 大規模データから基盤モデルを作成する流れが加速中



Aurora genAI:  
1兆パラメタの科学向け生成モデル  
(計画中)  
Argonne + Intel

## ■ 大規模言語モデル

(Large Language Model)

人間のこれまでの言語活動  
からの学習



**人間の知性を大きく超えることは  
難しい**

かなりかしこい人間止まり



## ■ マルチモーダル基盤モデル

(Multimodal Foundation Model)

自然言語

**+ 数値情報**

**+ 高速計算・情報処理**



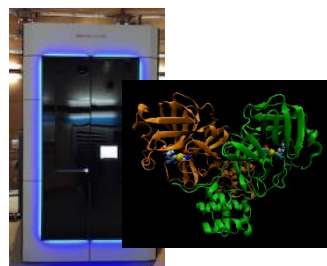
人間が本来苦手な数値情報も加えることで  
**超知性のような存在に進化可能**



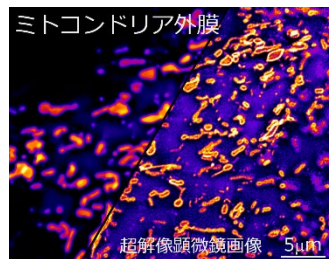
- 生命・医科学分野では、特に多くの種類のデータを扱い、統合していくことが求められており、**マルチモーダル基盤モデルのテストベッドに最適**



遺伝子配列・発現



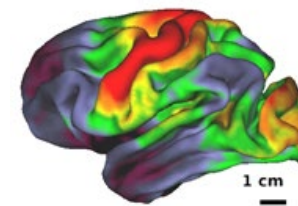
タンパク構造



イメージング



プロテオーム  
リポーム



脳・神経活動



その他表現型

- 大規模言語モデル（とそれに内包される人間的知性）をベースとして活用し、生命・医科学分野の文献データ・実験データ・シミュレーションデータを統合した基盤モデルを作成
- 生命・医科学基盤モデルの応用・産業展開
  - ▷ 医療の高度化、創薬、生命・医科学研究の加速
  - ▷ マルチモーダル基盤モデル作成技術は、材料科学・工業・農業等に広く応用可能

- Evolutionary Scale Model-2, MetaAI
- タンパク質のアミノ酸配列を学習させたモデル
  - ▷ Data: UniRef 50(タンパク質のアミノ酸配列のDB)の配列- $\beta+\alpha$  約60M
  - ▷ Transformerで自己教師あり学習
  - ▷ 最大15Bパラメタ
- ESM-2 + Protein Data Bankの立体構造(おそらく $10^4\sim 10^5$ )で立体構造予測
  - ▷ タンパク質内のアミノ酸のコンタクトを予測。単語間の関連に相当する。
  - ▷ AlphaFold2と同等の性能を高速に実現
  - ▷ 最大モデルで最高の性能

<https://doi.org/10.1101/2022.07.20.500902>

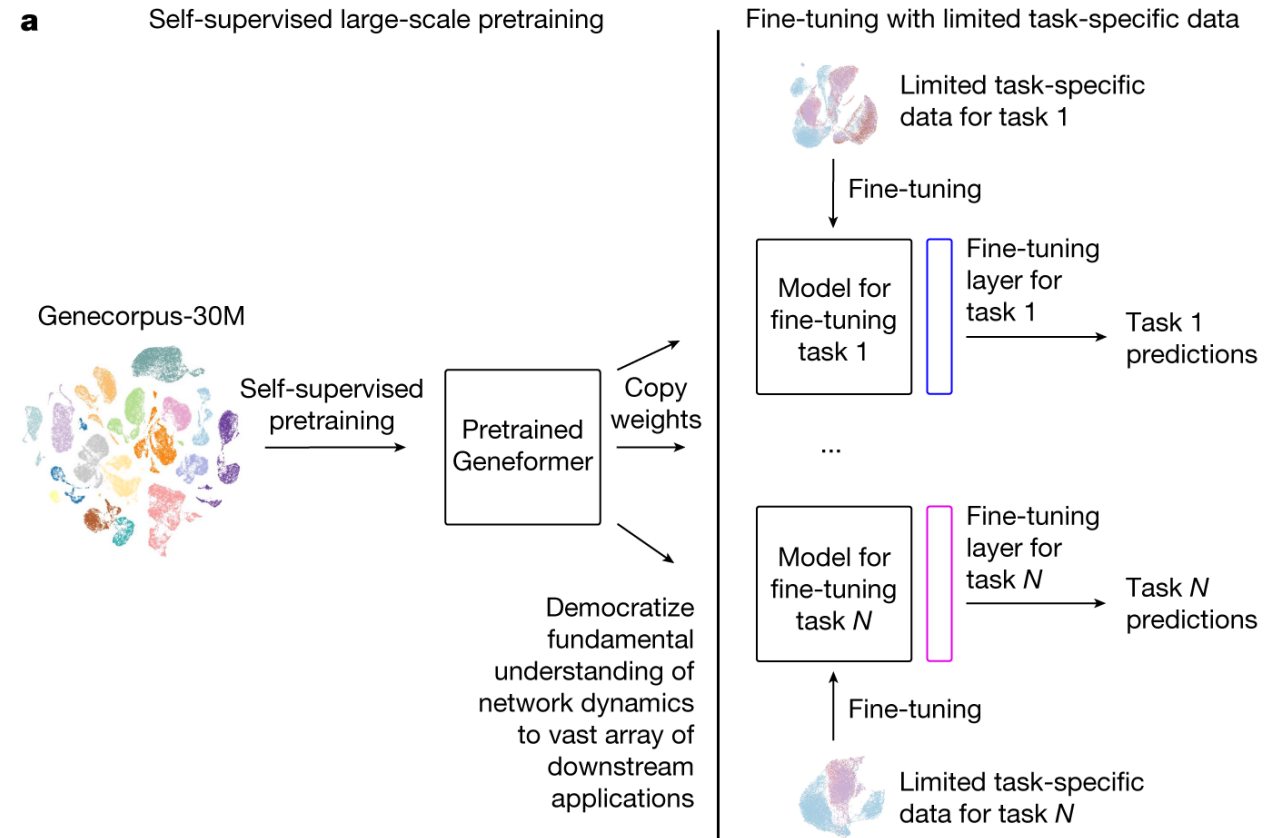
<https://github.com/facebookresearch/esm>

Theodoris, C.V., Xiao, L., Chopra, A. *et al.* **Transfer learning enables predictions in network biology.** *Nature* **618**, 616–624 (2023). <https://doi.org/10.1038/s41586-023-06139-9>

## ■ Transcriptomeの基盤モデル

- ▷ 30Mのsingle-cell transcriptome データ
- ▷ 複数の下流のタスクに適応可能
  - Gene dosage sensitivity predictions
  - Chromatin dynamics predictions
  - Network dynamics predictions
- ▷ 様々な応用
  - In silico gene network analysis
  - In silico treatment analysis

## ■ 類似のものとしてscGPTもある



Cui, H. *et al.* **scGPT: Towards Building a Foundation Model for Single-Cell 2 Multi-omics Using Generative AI.** <https://doi.org/10.1101/2023.04.30.538439>

シミュレーションに加え、**深層学習は非線形の複雑システムを扱うための一般的な方法**として、科学の全域で重要になりつつある

**HPCのほぼ全領域を、AI + HPCで再定義していく必要がある**

- 1. AI作成・利用のためのHPC: スケールへの対応**
  - ▷ 大規模並列学習・演繹
  - ▷ ハードウェア : Accelerators / Quantum NN
- 2. シミュレーション結果のAI化・サロゲートモデル**
- 3. AIを利用した大規模シミュレーション**
  - ▷ AIによるシミュレーション～解析の自動化
- 4. AIを進化させる枠組**
  - ▷ AI間の対話・競争、社会科学・工学への応用

**理研R-CCSを中心にAI for Science roadmapの作成中**

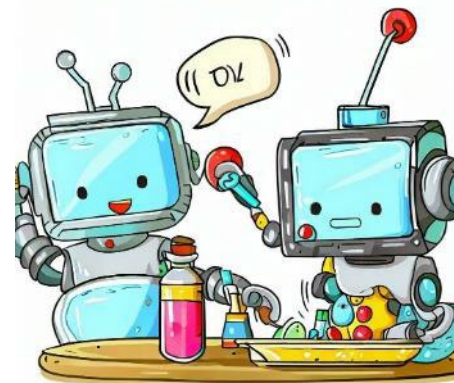
## 実験計画作成

LLMを利用した  
仮説検証のための  
実験計画の自動生成



**Plan**

## 大規模 シミュレーション



**Do**

**Check**



## 自動解析

LLM/ドメイン  
基盤モデルを用い、  
データの仮説に沿った  
解釈、レポート作成

**Action**



## 仮説生成

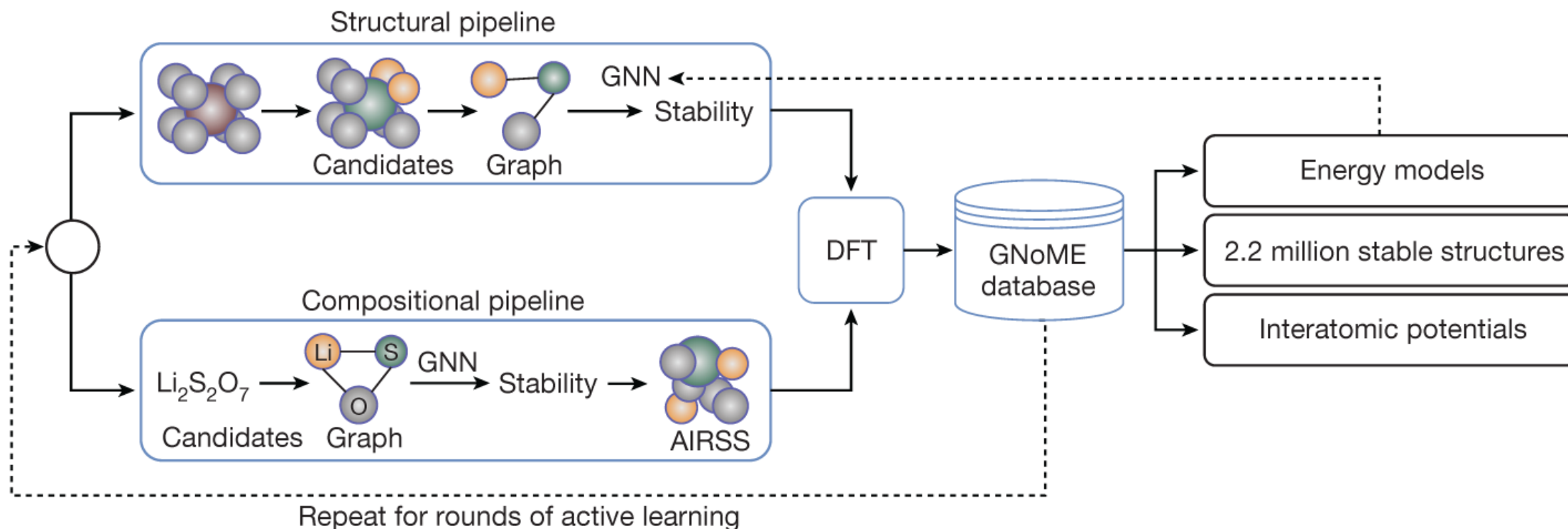
LLM/ドメイン  
基盤モデルを用いた  
仮説生成・  
仮説修正

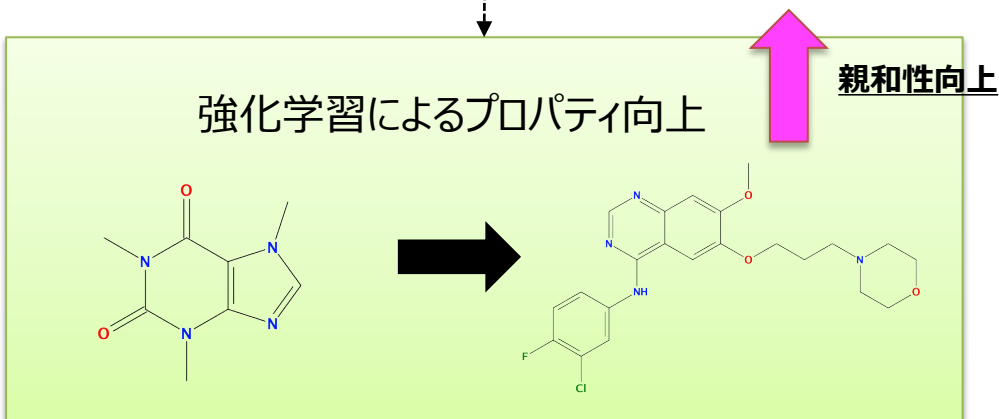
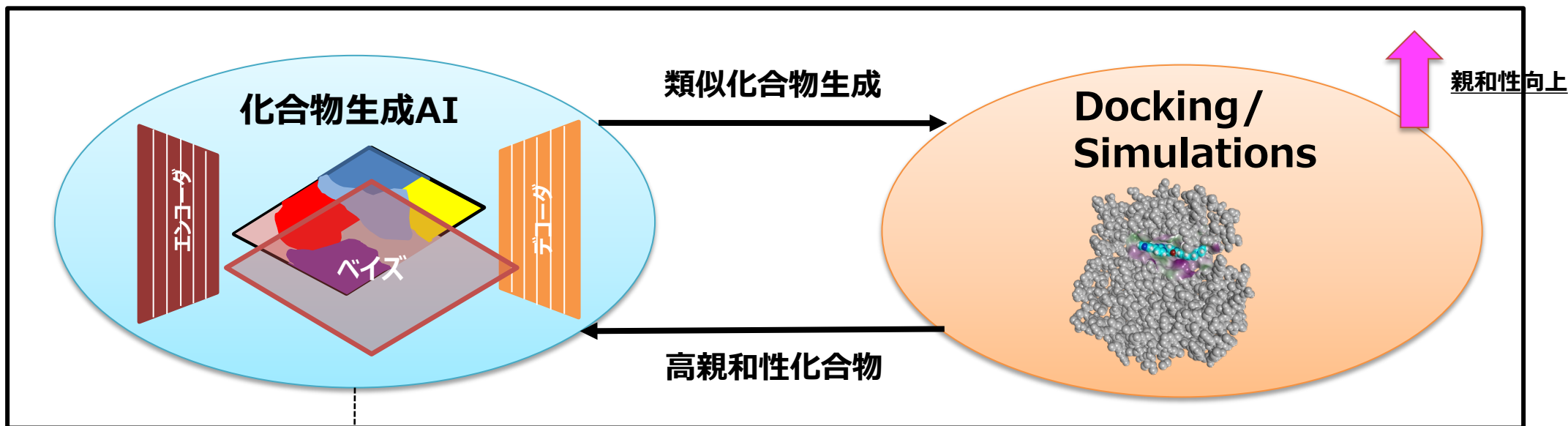
回路設計・プログラム作成なども  
同様のサイクルで自動化が可能



- Merchant, A., Batzner, S., Schoenholz, S.S. *et al.* Scaling deep learning for materials discovery. *Nature*(2023). <https://doi.org/10.1038/s41586-023-06735-9>
- 安定な結晶構造 + エネルギー評価
- いわゆる生成AIとはちょっと違うが、  
既存構造を改変 → GNNで物性・構造予測 → DFTで評価

a





Sample and dockなど

機械学習による提案・  
シミュレーションによる検証を  
組み合わせた自動探索

Remi Lam et al. , Learning skillful medium-range global weather forecasting. Science **0**, eadi2336 DOI:10.1126/science.adi2336

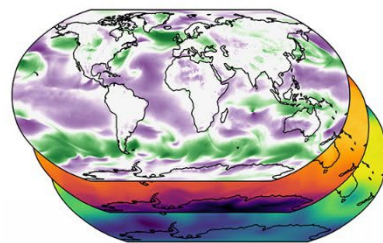
- 気象予報モデル
- ヨーロッパ中期予報センター(ECMWF)の高解像度モデル(HRES)を上回る精度の10日間予測が可能

- ▷ GNN 36.7M parameters

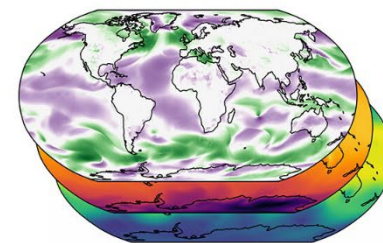
- ▷ TPU v4 single node (275TFLOPS) 1時間で10日分の予測が可能

- Simulationのサロゲートではないが、AIモデルでの置き換えの可能性を示すもの

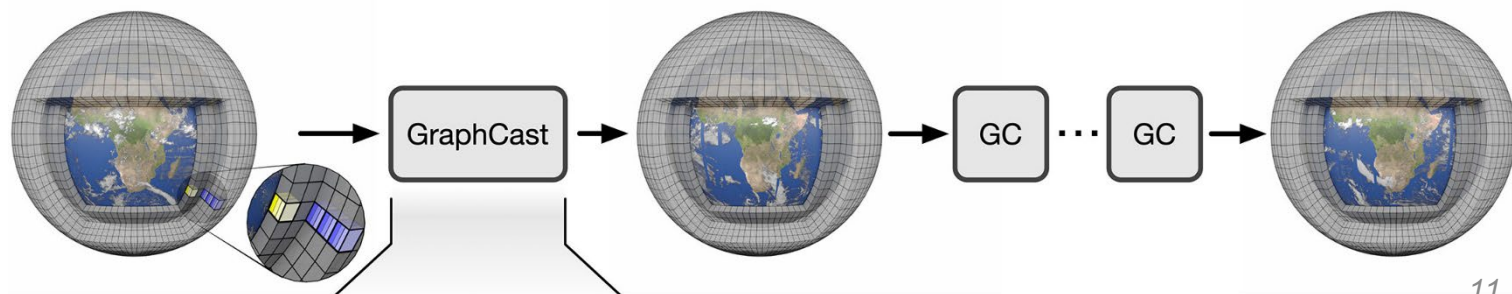
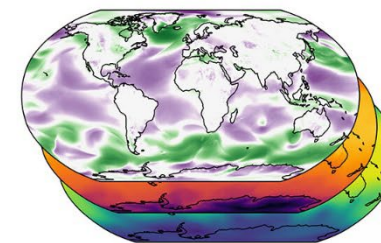
a) Input weather state



b) Predict the next state



c) Roll out a forecast



## AIで作る・AIの作る社会の研究

### ■ LLM等の基盤モデルを用いたマルチエージェントモデル

### ■ 人間社会のモデルとしての意味

- ▷ 人間行動・社会行動・SNS等
- ▷ ロボット・車等の異なる身体性での行動
- ▷ 経済行動・ファイナンス

### ■ AIを進化させるプラットフォームとしての意味

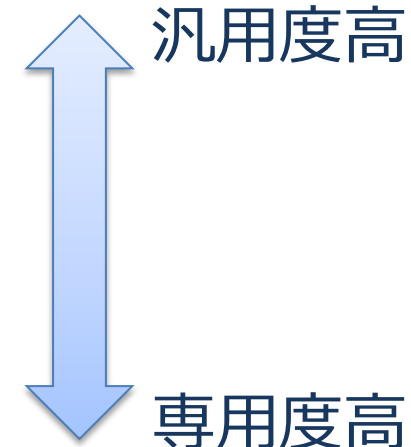
- ▷ AI自身の競争による進化



<https://doi.org/10.48550/arXiv.2304.03442>

## ■ ワークロード別の最適ハードウェア特性

- ▷ シミュレーション 広メモリ帯域
- ▷ AI inference 広メモリ帯域・低精度
- ▷ AI training テンソル演算・その他専用化
- ▷ 古典分子動力学 強スケーリングアクセラレータ
- ▷ 量子系・一部最適化 量子計算、量子古典ハイブリッド



## ■ ワークロードに合わせた「ベストミックス」の実現

- ▷ ヘテロジニアスなアーキテクチャ要素の取り込み
  - 均一なシステムの中にアーキテクチャ要素を組み込むアプローチ  
シミュレーション+AI演繹は比較的容易、+AI学習ぐらいまでが可能
  - 専用計算システム
- ▷ ヘテロジニアス性が高まるほど、Flagshipでのクラウド的統合が有効。使い分け、利用支援の強化を行いやすい。

- まだアルゴリズム面でも変化の可能性があり、Transformerを加速するというだけでは済まない
- Red oceanであるが、**AI向けの加速は今後の全ての計算基盤の軸になっていく**と考えられる。

**今後も国内でプロセッサ開発を続けるのなら、必須の要素**

## ■ 例

- ▷ Spiking NN・RWKV等向けの検討
- ▷ TransformerのBayes化（信頼度付きの推論）
- ▷ アダプティブな混合精度・スパースネスへの対応
- ▷ 低電力化・低コスト化（推論エンジンでのFlashメモリ利用など）

- 実際に関係を行うような、ハードウェアを含めたアーキテクチャ研究は、一旦アカデミアの研究としてはほぼ絶滅。
- 理由
  - ▷ **汎用**プロセッサのアーキテクチャは枯れたものになってきた
  - ▷ 産業的な重要性から、開発が民間の仕事になった
  - ▷ 開発費の高騰
- 一方で、Mooreの法則終了後多様なアーキテクチャの重要性は高まる。
  - ▷ Hardware-algorithm co-designの推進
  - ▷ AI向けプロセッサ・MD向けプロセッサ
- ハード研究の再興は、HPC業界では米国・欧州共に必要性を感じつつも、苦勞している。
  - ▷ SC2022でも、Co-design推進のために必要との議論に。
  - ▷ 欧州は、European Processor Initiativeがあり、HPC分野ではBarcelona Supercomputing Centerで開発しようとしている。

## ■ 一方、プロセッサ開発は昔に比べるとだいぶ楽に

- ▷ RISC-Vのフリー実装、Armの研究用コア他
- ▷ 論理合成機能の高度化：HDL設計からトランジスタまでの自動化  
同規模の設計リソースで、複雑度は10倍以上、  
回路規模は1,000倍以上のものが設計可能に
- ▷ FPGAの大規模化

## ■ Rapidusの立ち上げに合わせ、国内のプロセッサ研究も再興したい。

- ▷ **HPC・AI向けアーキテクチャの探求**
- ▷ ユーザと開発者の連携が重要：Hardware-Algorithm Co-design



## ■ MDGRAPE-5: 次世代MD専用計算機

▷ 並列データフロー型計算機: **強スケーリング性能の強化**

▷ **中粒度**のデータフロー処理

– 結合カー個、セル間の非結合力計算などを単位として処理

– これらの処理間をイベント（粒子のNW経由での到着、前の計算処理完了など）をベースで駆動。データフロー基盤はなるべく汎用的に設計。

▷ Bond-engine : 結合力計算の高速化

▷ 粗視化分子動力学への対応

▷ パラメトライズド設計 : 最適化の最適化を容易に

▷ 富岳Next FSで一部実施。富士通研究所と共同研究

## ■ Movitan: 深層ベイズ学習向け専用計算機

▷ RISC-V GeminiのBayes-by-Backprop向け拡張