

生成系AI 科学研究への波及

泰地 真弘人 (Makoto Taiji) 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

■ 基盤モデル(Foundation Model, FM)

= 汎用的に利用可能で、高度な推論を内包する大規模モデル

GPTシリーズなどの、大規模言語モデル (Large Language Model, LLM)が代表的な成功例

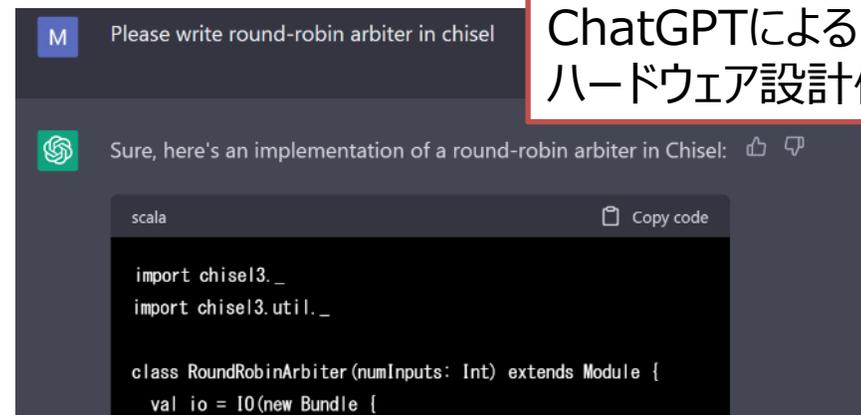
■ 生成AI (Generative AI)

= 指示に従って画像・文章・化学構造などを生成するモデル

大規模言語モデルに基づく文章生成・潜在拡散アルゴリズムによる画像生成

などで大きく発展し、**内部にロジックらしきものを内包。**

並の人間が作るものと遜色ないレベルに到達。



Stable diffusionによる画像生成

Prompt: A girl scientist is doing a biological experiment in a dark laboratory
Negative: mask
<https://stablediffusionweb.com>



大規模言語モデルに創発した知性

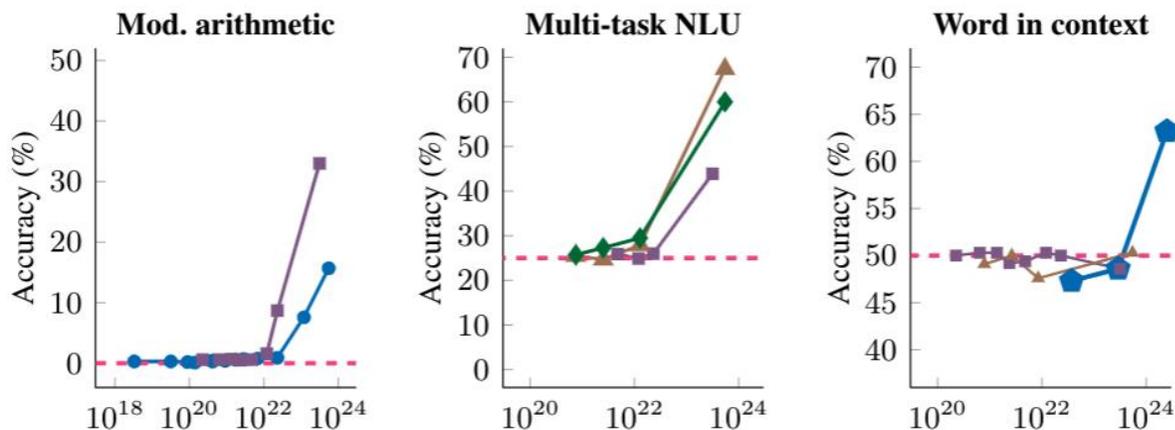
- スケーリング則の発見によって、計算量に応じた最適なモデル規模・データ量が決められるようになった

→ 計算量を指標に性能が決まる

- **1YFLOPs** (10^{24} = 1 紓演算, 富岳半精度ピーク性能の約6日分)

規模の大規模言語モデルで、急に能力が創発する現象が観測

—●— LaMDA —■— GPT-3 —◆— Gopher —▲— Chinchilla —◆— PaLM - - - Random



Model scale (training FLOPs)

多段の推論や、状況に応じたタスクの選択など、**人間らしい知性が創発**している模様

J. Wei and Y. Tay, “Characterizing Emergent Phenomena in Large Language Models”.

<https://ai.googleblog.com/2022/11/characterizing-emergent-phenomena-in.html>

- 大規模言語モデルに創発した知性は、今後の社会全般・研究活動に大きなインパクトを有する。

これを活用しつつ、発展させる基盤は国内に必須

- ▷ 計算基盤 : GPGPUスパコン・低電力実行環境・プロセッサ・ソフトウェア
- ▷ モデル基盤 : 大規模言語モデル・その他モデル開発
- ▷ 応用基盤 : 各ドメインへのチューニング
- ▷ 理論基盤 : スケーリング則、創発がなぜ起こるかの研究
- ▷ 実用化基盤 : サービス化など

理研は計算基盤・応用基盤・理論基盤で強い基盤をもつ

今後強化が必要なポイント :

- ▷ 計算基盤の一層の強化
- ▷ 応用に向けたAI研究と各ドメイン研究の融合促進 (TRIP推進)
- ▷ 基盤モデルの理論研究

■ 今後は言語に留まらない**マルチモーダル化**が重要

▷ 既にGPT-4は画像入力を受けつけ、**視覚を有している**

▷ 様々なデータを直接入力
= Virtual Brain-Machine Interface
を開発し、モデルにデータを入力

▷ **まだ手つかずで、発展の余地が大きい**

- 画像・音声の統合処理、ロボティクス、医療などで検討が進んでいる
- 今後モデル圧縮・省電力化が進めば、自動運転や生活支援などへ応用も可能

画像の面白さの意味を言語化できている

Example of GPT-4 visual input:

User What is funny about this image? Describe it panel by panel.



Source: <https://www.reddit.com/r/hmmn/comments/ubab6v/hmmn/>

GPT-4 The image shows a package for a "Lightning Cable" adapter with three panels.

Panel 1: A smartphone with a VGA connector (a large, blue, 15-pin connector typically used for computer monitors) plugged into its charging port.

Panel 2: The package for the "Lightning Cable" adapter with a picture of a VGA connector on it.

Panel 3: A close-up of the VGA connector with a small Lightning connector (used for charging iPhones and other Apple devices) at the end.

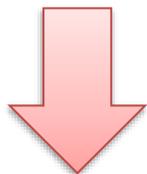
The humor in this image comes from the absurdity of plugging a large, outdated VGA connector into a small, modern smartphone charging port.

Table 3. Example prompt demonstrating GPT-4's visual input capability. The prompt consists of a question about an image with multiple panels which GPT-4 is able to answer.

■ 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



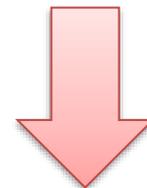
■ マルチモーダル基盤モデル

(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・医科学分野でのマルチモーダル基盤モデルの可能性

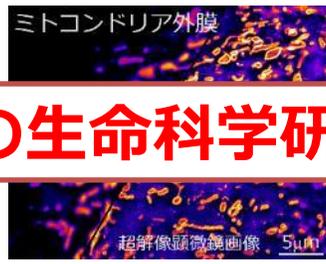
- 生命・医科学分野では、特に多くの種類のデータを扱い、統合していくことが求められており、**マルチモーダル基盤モデルのテストベッドに最適**



遺伝子配列・発現



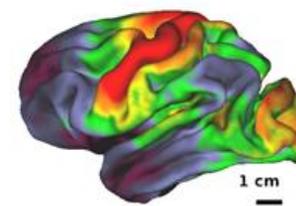
タンパク構造



イメージング



プロテオーム
リピドーム



脳・神経活動



その他表現型

理研の生命科学研究基盤を活用

- 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
 - ▷ 防衛・サイバーセキュリティへのデュアルユース
 - ▷ 基盤モデル作成・利用に必須の技術としての、計算機・半導体産業の振興
 - ▷ AI活用による新しい働き方・新しい社会の模索の促進

生命科学における基盤モデルの例

■ 自然言語系：BERT/GPTをベースに生命科学向けに追加学習したモデル

- ▷ BioBERT/PubMedBERT
BERTにPubMed/PMCデータを追加学習
- ▷ BioGPT
GPT-2にPubMedデータを追加学習
- ▷ MedGPT
GPT-2に電子カルテ情報を追加学習
- ▷ GatorTron
電子カルテをスクラッチから学習(9B)

■ タンパク質

- ▷ ESM-2 (Meta, Protein Language Model)
- ▷ RF Diffusion (Baker Lab., Protein design)：生成モデル

■ ゲノム

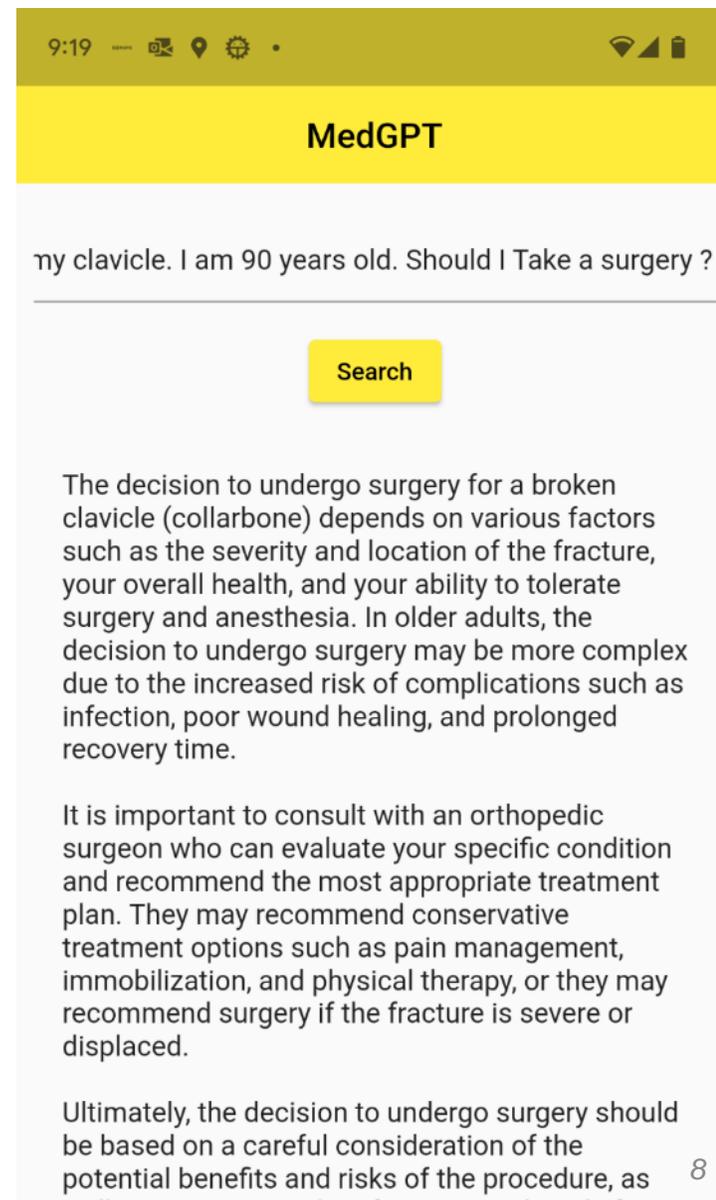
- ▷ GenSLM
- ▷ scGPT

■ 化合物生成:化合物ライブラリを学習させた生成モデル

- ▷ ChemBERT
- ▷ MUDiff

■ イメージや遺伝子発現データの統合は今後の課題

MedGPT Android app



9:19 — 9:19 9:19 9:19

MedGPT

my clavicle. I am 90 years old. Should I Take a surgery ?

Search

The decision to undergo surgery for a broken clavicle (collarbone) depends on various factors such as the severity and location of the fracture, your overall health, and your ability to tolerate surgery and anesthesia. In older adults, the decision to undergo surgery may be more complex due to the increased risk of complications such as infection, poor wound healing, and prolonged recovery time.

It is important to consult with an orthopedic surgeon who can evaluate your specific condition and recommend the most appropriate treatment plan. They may recommend conservative treatment options such as pain management, immobilization, and physical therapy, or they may recommend surgery if the fracture is severe or displaced.

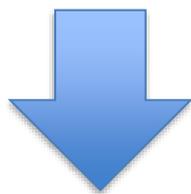
Ultimately, the decision to undergo surgery should be based on a careful consideration of the potential benefits and risks of the procedure, as

- 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- β + α 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>

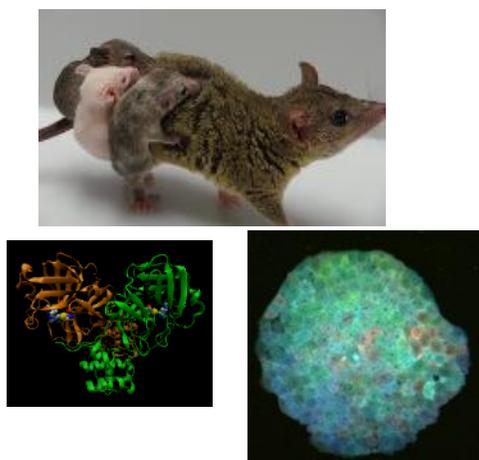
- データ量は60Mとまあまあの量だが、実際のタンパク質の多様性は 10^5 程度。
- それでも、モデル規模を大きくするほど予測性能が上がる。



- 生命科学分野での基盤モデルの可能性、現状データでも思ったよりあるのではないか

- まだマルチモーダルなものは端緒レベル
 - ▷ 画像関連で診断情報と合わせたものがある
 - ▷ 化合物生成で物性と合わせたモデル
- マルチモーダル基盤モデルは、生命科学の長年の課題
 - = データ統合を実現するためのキーテクノロジー
 - ▷ 真のトランスオミックス（オミックス情報の統合）の実現
 - ▷ **複雑な生命システムを、複雑なままモデル化して予測を可能に**

生命システム



←→
*Digital
Twin*



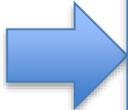
マルチモーダル
基盤モデル

Generated by DALL-E on Bing for
“multimodal foundation model for
life science”

1. 各データの基盤モデル化

- ▷ 遺伝子配列 : 遺伝子構造データとの融合
- ▷ タンパク質配列 : ESM-2 (MetaAI)などの成功例
- ▷ 画像情報 : 顕微鏡動画画像の生成モデル
- ▷ 文献情報 : 一般の大規模言語モデルに追加学習

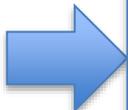
利用イメージ



ゲノム遺造予測
タンパク質立体構造予測
疾患モデル作成
モデル・パラメタ抽出
等

2. 大規模言語モデルと各モデルの融合

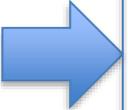
- ▷ マルチモーダル基盤モデル技術の開発
- ▷ 大規模言語・生命データ融合モデルの開発
- ▷ 計算システム・サービスの開発



自動解析・統合解析
実験・治療計画
実験自動化との融合
創薬
医療・研究支援サービス
等

3. マルチモーダル生命・医科学基盤モデルの開発

- ▷ モデルの統合
- ▷ シミュレーション・演繹モデルとの融合
- ▷ 量子計算技術との融合



AIによる画期的発見
研究開発自動化
自律進化
全計算機上での自動研究
等

生命・医科学分野の広範なデータ＋自然言語データを基に、データ解析・実験/治療方針作成・創薬など、**複数の用途に汎用可能な基盤モデル**の構築

基盤モデル構築に必須の要素

1. 大規模データ

遺伝子配列・発現
イメージング・構造情報
表現形・疾患関連データ



2. エクサスケール計算資源

富岳・富岳Next
DL専用システム



3. 先進機械学習アルゴリズム・ソフトウェア・ミドルウェア

生命・医科学 基盤モデル



- ・日本語対応
- ・マルチモーダル：
自然言語以外
もデータ統合

理研の総力を結集

データ解析自動化
実験計画自動化

診断支援
治療計画策定

創薬加速

基盤モデル向け
低電力LSI開発

補足

■ 基盤モデル = Foundation Model

“On the Opportunities and Risks of Foundation Models”,

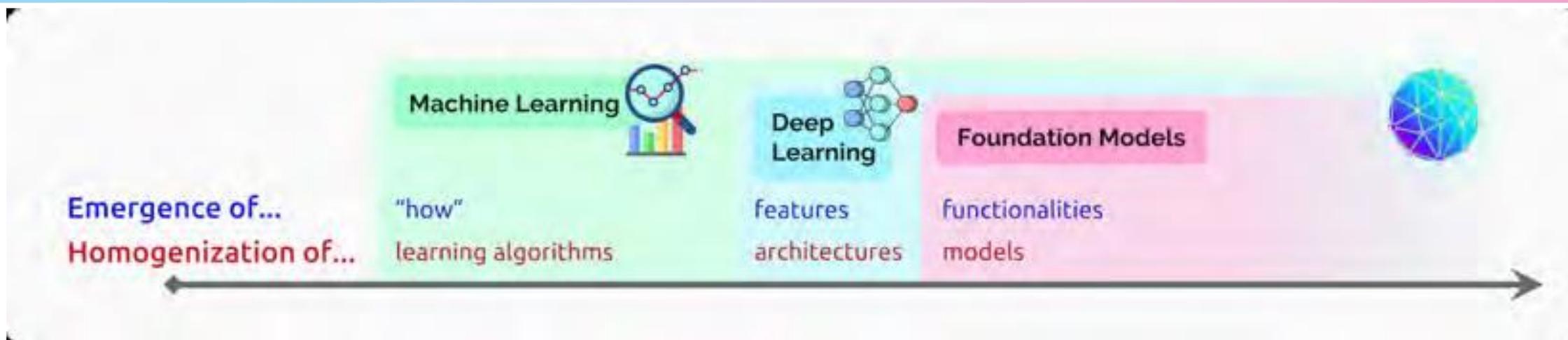
R. Bommasani et al. (2021), <https://arxiv.org/abs/2108.07258>

Abstractからの抜粋

幅広いデータで訓練され（多くはself-supervisionも大規模に適用し）、下流の幅広いタスクに適応できるモデル（BERT、DALL-E、GPT-3など）の台頭により、AIのパラダイムシフトが起きつつある。未だ不完全ではあるが、その極めて中心的な地位を強調するため、これらのモデルを基盤モデルと呼ぶことにする。…

訳語としては基盤モデルが用いられているが、「根本モデル」のほうがニュアンス的には正しいかも

基盤モデル(FM) ⊃ 大規模言語モデル(LLM)



■ 創発性

- ▷ 多段の推論や、複数の情報を統合した複雑な答を創出可能

■ 均一化 (Homogenization)

- ▷ 一つのモデルで、多くの課題に対応可能
- ▷ 人間のような汎用性に近づく

■ これまで

- ▷ データからDeep Learningを使って学習させたモデルを領域ごとに作る

■ これから

- ▷ これまでのDeep Learningの重要さは失われていないが

- ▷ 均一化：基盤モデルを使いこなす

- 追加学習

- Prompt engineering :

- 基盤モデルから欲しい出力を得るための入力を作る技術

- 技術が進むにつれ、一過性のものとは思う。ただ、基盤モデルの動作原理の理解は難しいので、常に付きまとう問題であろう。

- 人に指示を出すときの難しさみたいなものか。

- Transformer (2017, Google)
 - ▷ 文脈があるものを並列処理可能に
- BERT (2019, Google. *Bidirectional Encoder Representations from Transformers*)
 - ▷ Transformerベースの大規模モデル
 - ▷ 340Mパラメータ
- GPT-3 (2020, OpenAI. *Generative Pre-trained Transformer*)
 - ▷ 175Gパラメータ(最大のもの)
- GPT-4 (2023, OpenAI)
 - ▷ パラメータ数非公開
 - ▷ 画像も入力可能に
- その他多数、BLOOMのようなOpenな取り組みもある。

大規模言語モデルの創発性：ChatGPT

Q. As a professional novelist, please write a science fiction such that RIKEN plays an important role.

Sure, here's a short science fiction story featuring RIKEN:

Title: The RIKEN Experiment

It was the year 2050, and the world was on the brink of a major scientific breakthrough. The RIKEN organization in Japan had been conducting a top-secret experiment for the past ten years, and the results were finally about to be revealed.

Dr. Masaru Ito, the head of the RIKEN team, stood before the world's top scientists and researchers, ready to unveil his team's discovery. He took a deep breath and began his presentation.

"We have successfully discovered a new method of harnessing the power of the human mind," Dr. Ito announced. "Using advanced neuroimaging techniques, we have found a way to tap into the brain's electrical signals and convert them into usable energy."

The room erupted in applause, and Dr. Ito smiled proudly. This was the moment he had been waiting for his entire life.

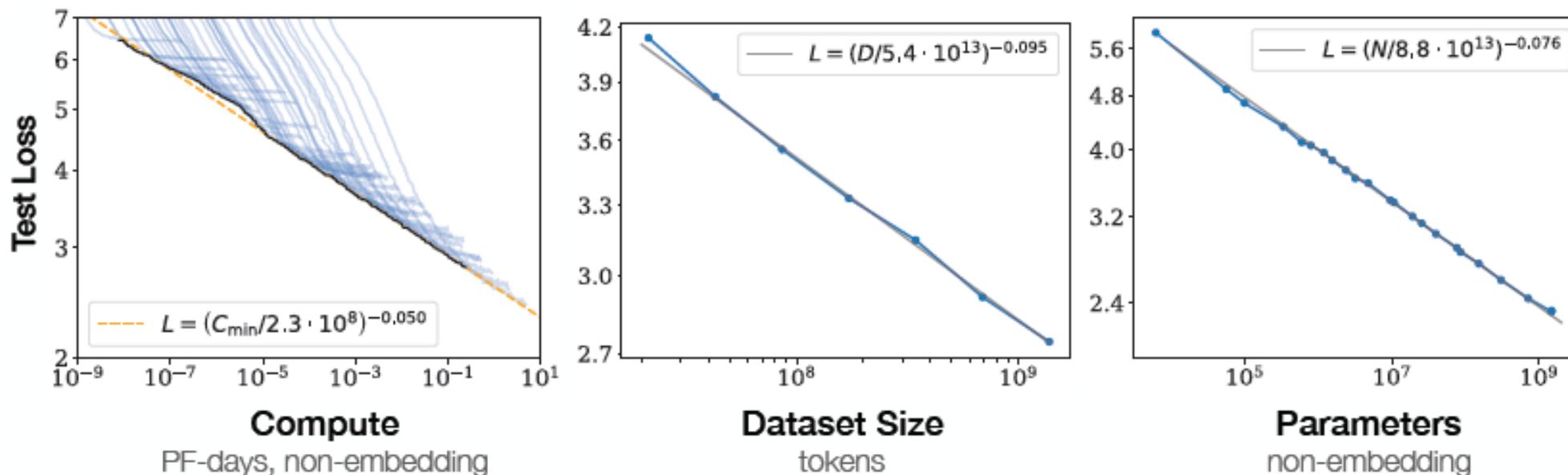
As news of the RIKEN experiment spread throughout the world, governments and corporations began vying for access to the technology. The potential applications were limitless: powering homes and cities, revolutionizing transportation, and even creating new forms of entertainment.

But as the world celebrated the RIKEN experiment, a group of rogue scientists began to express concern. They worried that the new technology could have unintended consequences, and that the power of the human mind should not be tampered with.

Their fears were realized when a group of hackers infiltrated the RIKEN system and began manipulating the brain signals. Suddenly, people all over the world began experiencing uncontrollable thoughts and emotions. Some even reported feeling physical pain and discomfort.

Dr. Ito and his team raced to find a solution, but it seemed like they were too late. The RIKEN experiment...

スケーリング則



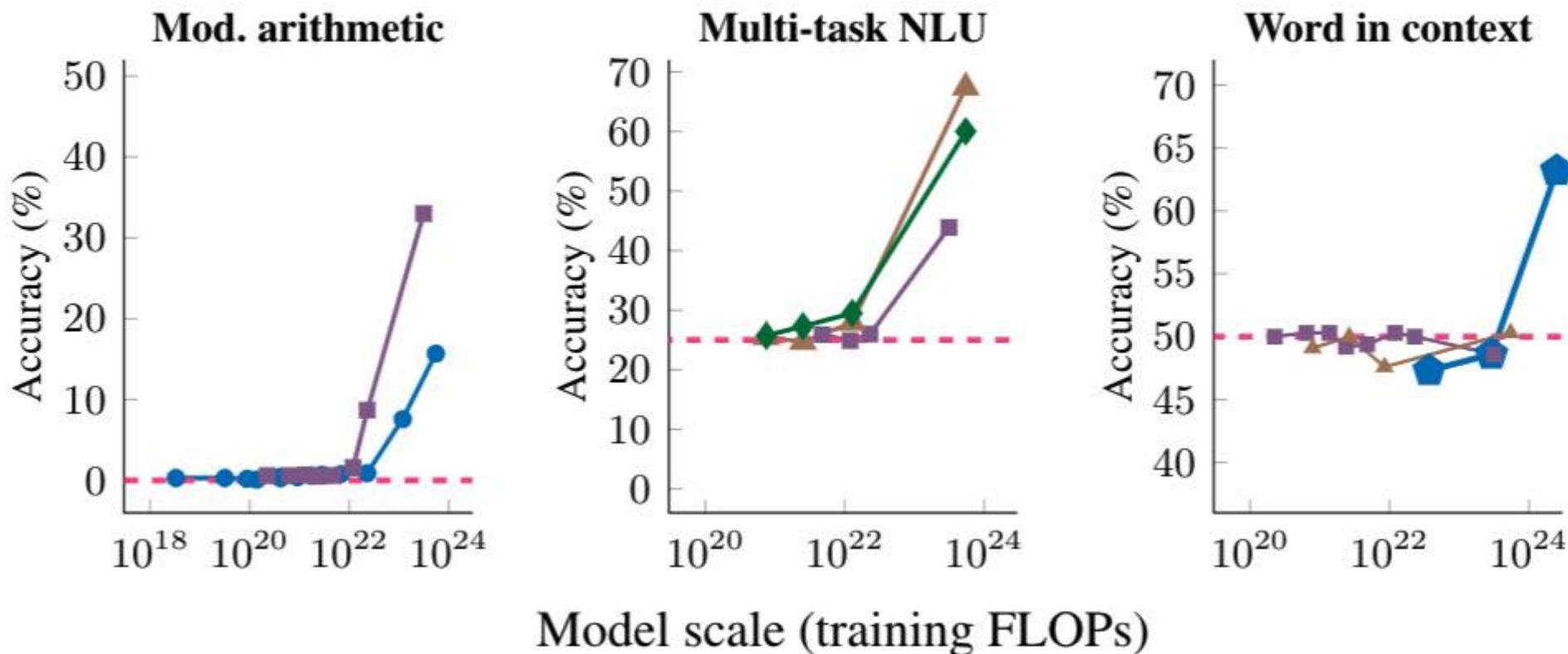
J. Kaplan et al., *Scaling Laws for Neural Language Models*. <https://doi.org/10.48550/arXiv.2001.08361>

- 言語モデルは、パラメタ数・計算量・データサイズを増やすほど精度が上がる
- 最適モデルサイズ \sim 計算量^{0.7}
- GPT-3の最大規模では、2 EFLOPS・day（効率10%として、富岳単精度の20日分）でも収束しきっていない

T. Henighan et al., *Scaling Laws for Autoregressive Generative Modeling*. <https://doi.org/10.48550/arXiv.2010.14701>

1YFLOPs(12EFLOPS・day)での創発性

● LaMDA ■ GPT-3 ◆ Gopher ▲ Chinchilla ◆ PaLM - - - Random



J. Wei and Y. Tay, “Characterizing Emergent Phenomena in Large Language Models”.
<https://ai.googleblog.com/2022/11/characterizing-emergent-phenomena-in.html>