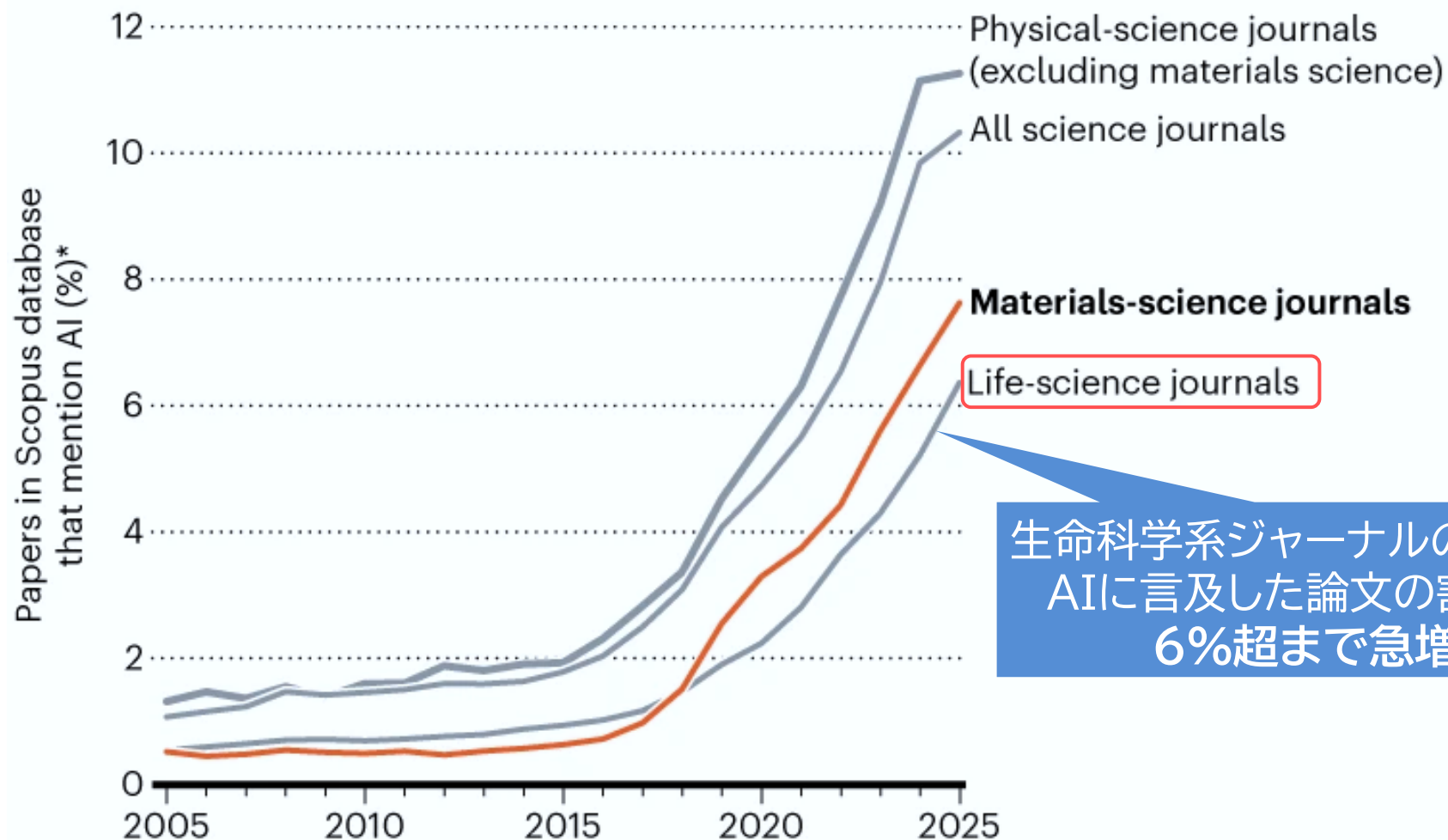


AI for Life Science に関する動向

2025年11月25日
研究振興局ライフサイエンス課

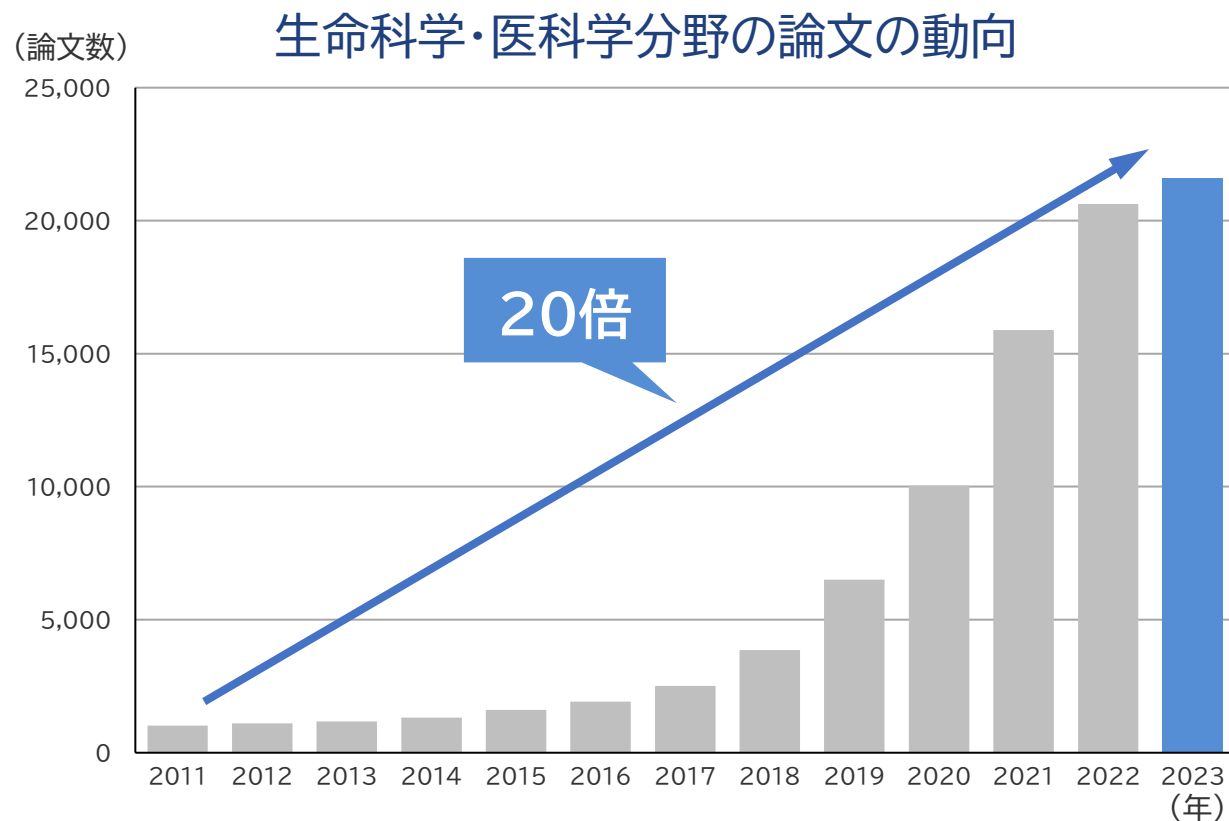
AIに言及した科学研究論文の状況



生命科学系ジャーナルの論文で
AIに言及した論文の割合は
6%超まで急増

*Title, abstract or keywords contain: 'machine learning', 'neural net', 'deep learning', 'random forest', 'support vector machine', 'artificial intelligence', 'AI', 'dimensionality reduction', 'Gaussian processes', 'naive Bayes', 'large language models', 'LLM', 'ChatGPT', 'Gaussian mixture models', 'ensemble methods' or 'graph neural net'.

AI・機械学習を用いた科学研究の論文動向



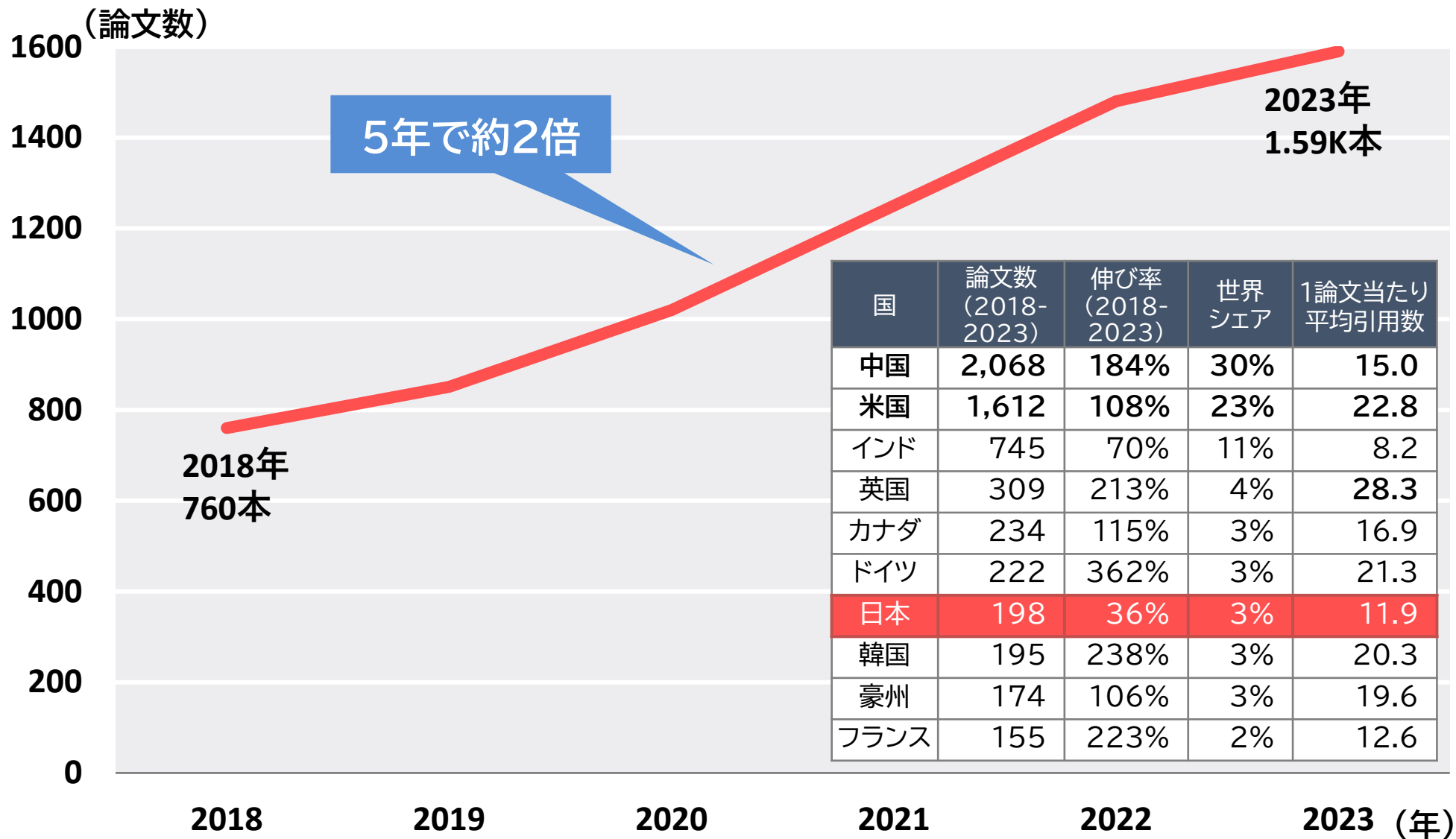
2011-2023年累積
上位15国・地域(整数カウント)

中国	26,591
米国	25,904
イングランド	6,325
ドイツ	6,087
インド	6,001
韓国	4,210
カナダ	4,057
イタリア	3,654
日本	3,439
オーストラリア	3,210
フランス	2,942
オランダ	2,719
スペイン	2,672
スイス	2,221
台湾	2,123

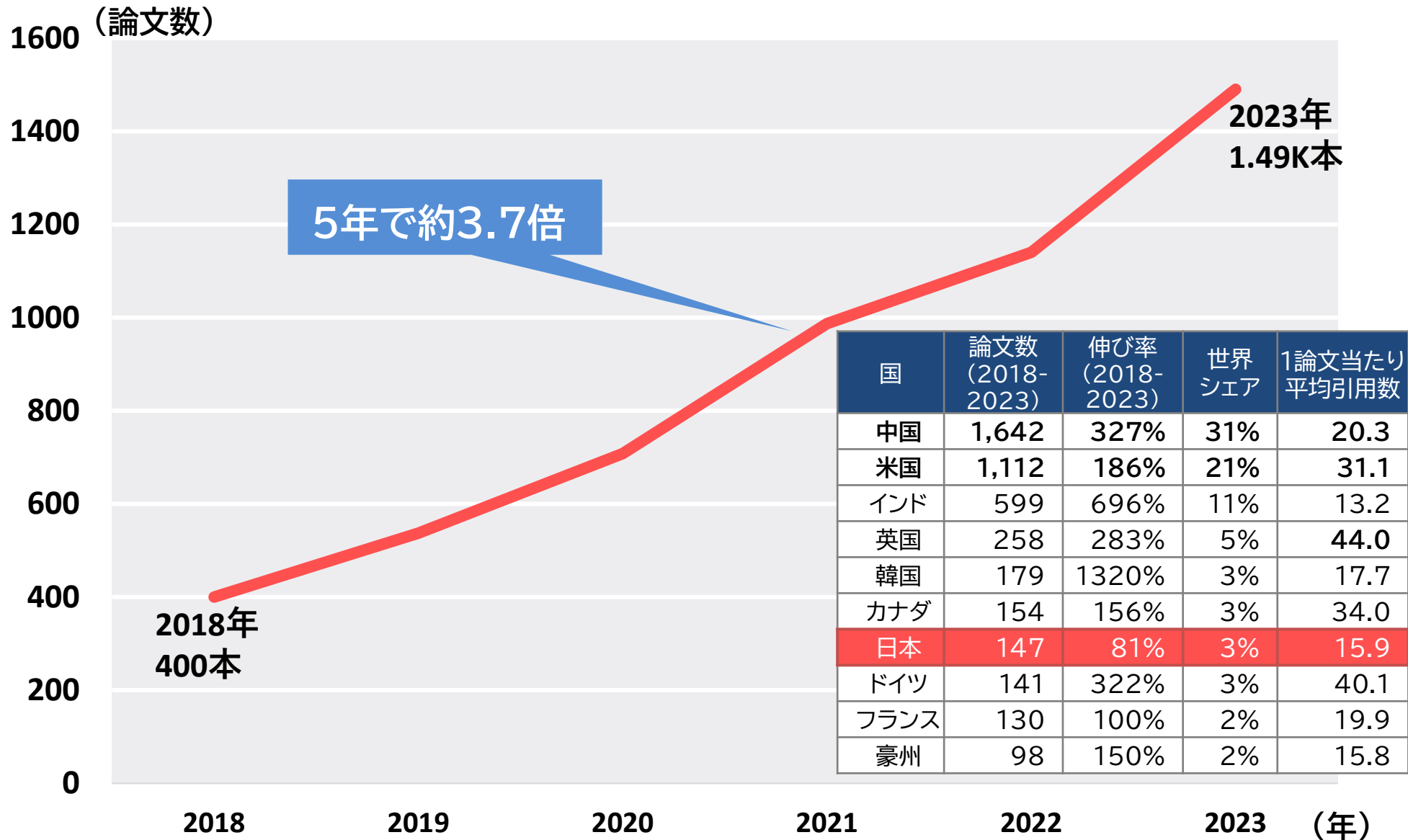
注: Web of Scienceを利用し、化学・材料分野及び生命科学・医科学分野で、“artificial intelligence”, “machine learning”, “deep learning”, “Neural Network”, “Bayesian optimization”, “Large language Models”または “Natural Language Processing”をキーワードに含む論文(article及びreview)の件数を集計

資料: 科学技術振興機構研究開発戦略センター提供

AI + 遺伝学研究 の論文動向

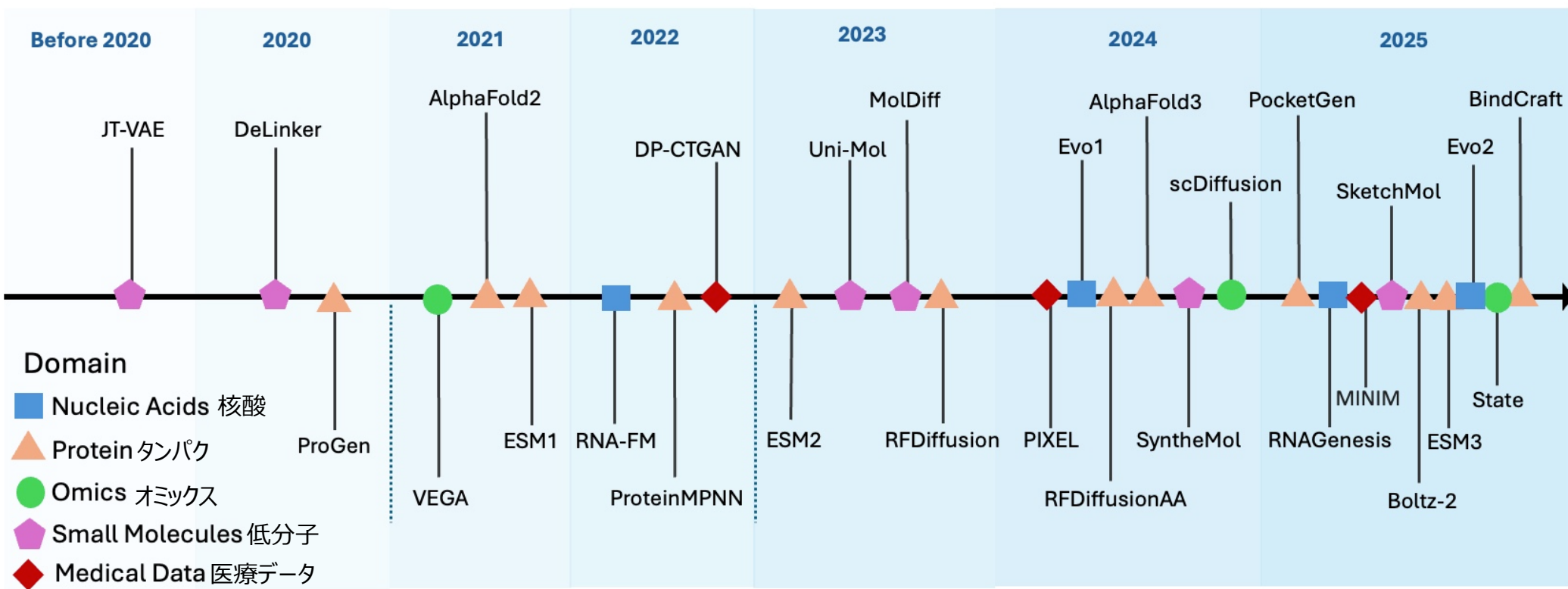


AI + 薬学研究 の論文動向



様々なライフサイエンス分野のAIモデル開発の状況

- ◆ 2021年のAlphaFold2の登場が大きな転換点となり、ライフサイエンス分野におけるAI(特に深層学習や基盤モデル(Foundation Model))を活用した研究が急増。
- ◆ さらに2023年の生成AIブームを受けて、大規模言語モデル(LLM)の医療応用、拡散モデル(Diffusion Model)のバイオ分野での応用、ラボオートメーションの研究なども加速。



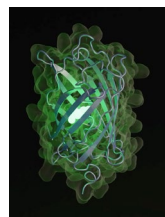
(出典) Zaixi Zhang et al. (2025, Oct) *Generative AI for Biosciences: Emerging Threats and Roadmap to Biosecurity*
<https://arxiv.org/abs/2510.15975>

生命・医科学におけるAI基盤モデルの開発

タンパク言語モデル

元Meta社の研究者らが設立したEvolutionaryScale社が約28億のタンパク質を学習した**タンパク言語モデル(ESM3)**を開発。タンパク質の配列、構造、機能の3つを同時に推論し自然界にはない新しいタンパクを生成することも可能に。

【出典】2024年6月25日, EvolutionaryScale社,
「Simulating 500 million years of evolution with a language model」



新しい緑色
蛍光タンパク質

タンパク・デザインモデル

スイス・ローザンヌ工科大学を中心に、MITやオランダの研究チームが連携する国際チームが、AlphaFold2の技術を応用し、標的タンパク質の特定領域に結合し、その機能を制御するタンパク質を、従来よりも高成功率で設計できるモデル(**BindCraft**)を開発。

【出典】2025年8月27日, Martin Pacesa et al., Nature 「One-shot design of functional protein binders with BindCraft」

ゲノム言語モデル

米・Arc Institute(非営利研究機関)が、スタンフォード大、UCバークレー、UCサンフランシスコ、NVIDIAと連携し、微生物、植物、動物、ヒトから集めた約9.3兆のDNA塩基対の情報を学習した**ゲノム言語モデル(Evo2)**を開発。遺伝子変異がタンパク質やRNAに及ぼす影響の評価や、新しいDNA配列の生成も可能に。

【出典】2025年2月19日 Arc Institute 「AI can now model and design the genetic code for all domains of life with Evo 2」

予測モデル

Google DeepMind社が、非コード領域を含む最大100万塩基対のDNA配列を解析し、遺伝子発現、スプライシングパターン、クロマチンの特徴など多様な分子情報(モダリティ)を単一塩基対レベルで同時に予測可能な、ヒトやマウスのゲノムで学習したモデル**AlphaGenomics**を開発。遺伝子変異の影響を、効率的に短時間で予測することを可能に。

【出典】2025年6月25日, Google DeepMind, 「AlphaGenome: AI for better understanding the genome」

ゲノム言語モデルとLLMの統合モデル

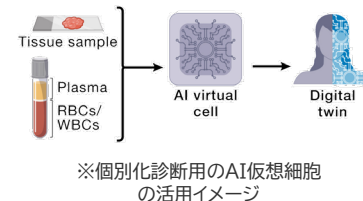
トロント大学等の研究チームが、ゲノム言語モデルと大規模言語モデル(LLM)を統合し、理解可能なモデル**BioReason**を開発。

【出典】2025年5月29日, Adibvafa Fallahpour et al, arXiv:2505.23579
「BioReason: Incentivizing Multimodal Biological Reasoning within a DNA-LLM Model」

AI仮想細胞モデルの開発

✓米・Chan Zuckerberg Initiativeは、2023年9月、**AI仮想細胞(AI Virtual Cell)**計画を明らかにし、2024年12月にはスタンフォード大学等と、**AI仮想細胞基盤モデル**の具体構想を提案

【出典】2024年12月12日, Charlotte Bunne et al, Cell,
「How to build the virtual cell with artificial intelligence: Priorities and opportunities」



✓米・イエール大、Google DeepMind等が、シングルセルデータ(scRNA-seq)を人間が理解できる言語として扱う**仮想細胞モデルC2S-Scale**を開発。「この細胞は薬Xにどう反応するか?」といった質問に、生物学的情報に基づいた回答が自然言語で得られる。

【出典】2025年4月17日, Syed Asad Rizvi et al,
「Scaling Large Language Models for Next-Generation Single-Cell Analysis」

✓米・Arc InstituteとUCバークレー、スタンフォード大、UCサンフランシスコ、ペンシルベニア大、イエール大が、幹細胞や、ガン細胞、免疫細胞が薬物、サイトカイン、遺伝的摂動にどう反応するかを予測する**仮想細胞モデル(STATE)**を、約1.7億個の細胞からの観察データと、1億個以上の細胞からの摂動データを学習させて開発。

【出典】2025年6月23日, Arc Institute
「Predicting cellular responses to perturbation across diverse contexts with State」

戦略的な取組を進める研究機関

✓米・ブロード研究所の“**Eric and Wendy Schmidt Center**”開設

2021年3月、機械学習と生物学を融合し、生命のプログラムの理解を目指す、センターを1.5億ドルの寄附を得て開設。タンパク質の細胞内局在予測モデル等の開発が進められている。

✓米・ハーワード・ヒューズ医学研究所(HHMI)の『**AI@HHMI**』計画

2024年8月、科学的発見の加速のため、AIを活用した生物医学研究に10年間で5億ドル規模を投資することを発表。ジャーネリア・リサーチ・キャンパスを中心に、HHMI全体でAIの活用を促進。

✓欧州分子生物学研究所(EMBL)の『**Science AI Strategy**』

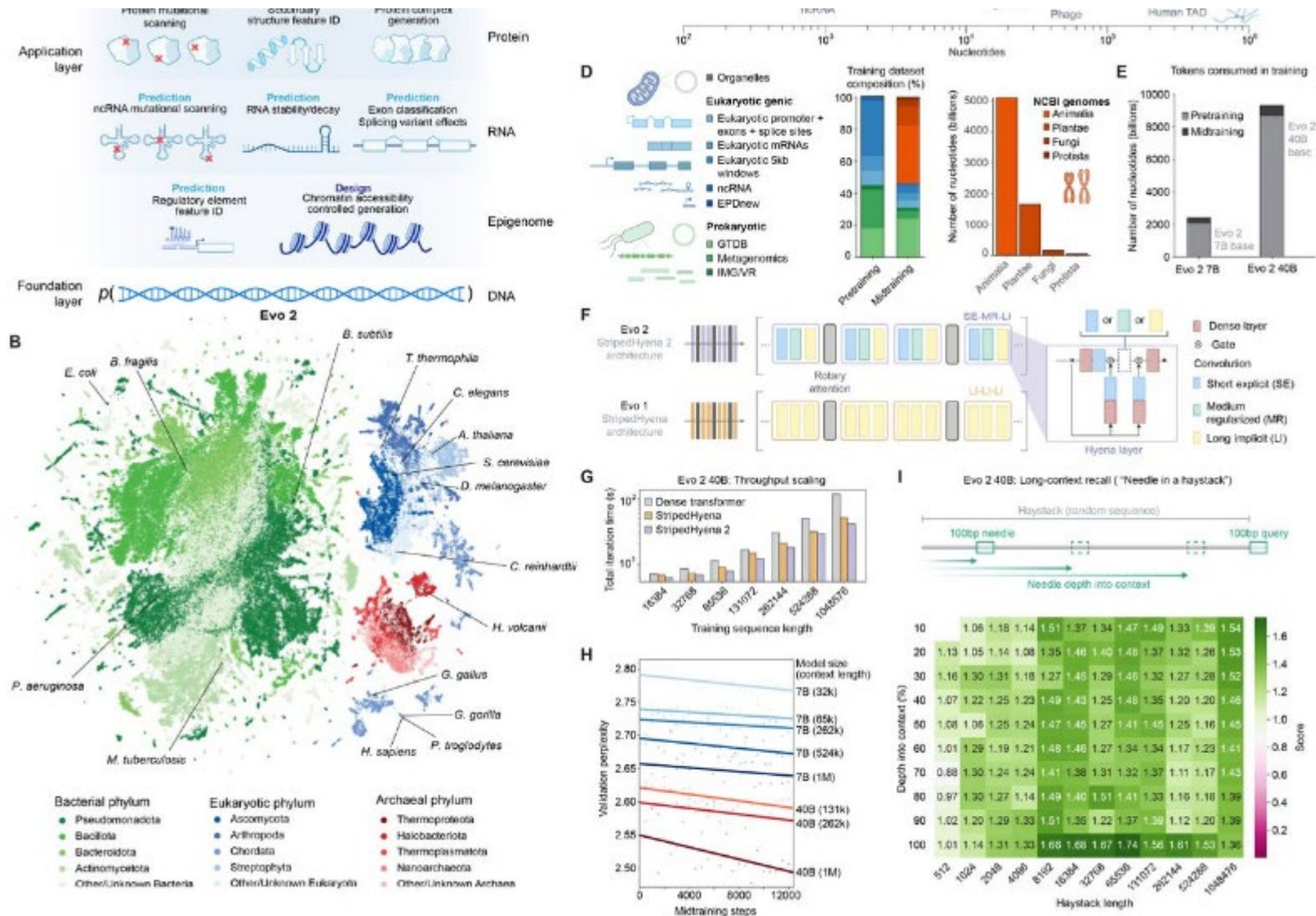
2025年2月、欧州の生命科学分野のAI活用の変革を目指した戦略を提示。生物学的ドメインに特化した基盤モデルの開発や、欧州の大規模バイオデータのAI対応に整備すること等が掲げられている。

(参考) ゲノム言語モデル Evo2 (2025年2月)



文部科学省

米・Arc Institute(非営利研究機関)が、スタンフォード大、UCバークレー、UCサンフランシスコ、NVIDIAと連携し、微生物、植物、動物、ヒトから集めた約9.3兆のDNA塩基対の情報を学習した基盤モデルを開発。遺伝子変異がタンパク質やRNAに及ぼす影響の評価や、新しいDNA配列の生成も可能に。



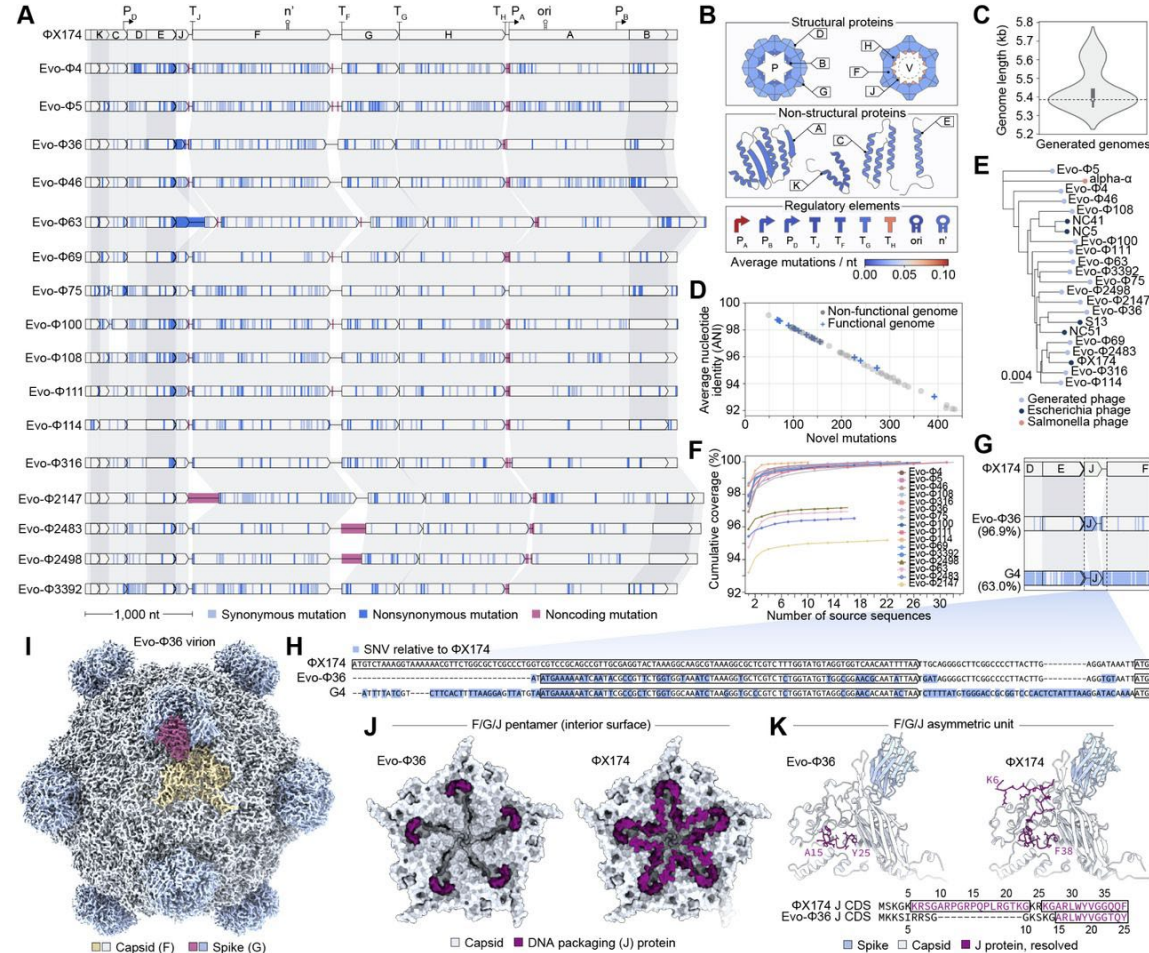
出典 2025年2月19日 Arc Institute 「AI can now model and design the genetic code for all domains of life with Evo2」

<https://www.biorxiv.org/content/10.1101/2025.02.18.638918v1.full> 1

(参考)ゲノム言語モデルを用いたバクテリオファージの設計・生成(2025年9月)

2025年9月、米・Arc Institute(非営利研究機関)とスタンフォード大の研究チームが、Evo1及びEvo2モデルを用いて、抗生物質に耐性のある大腸菌株に感染するバクテリオファージを設計・生成したことを発表。

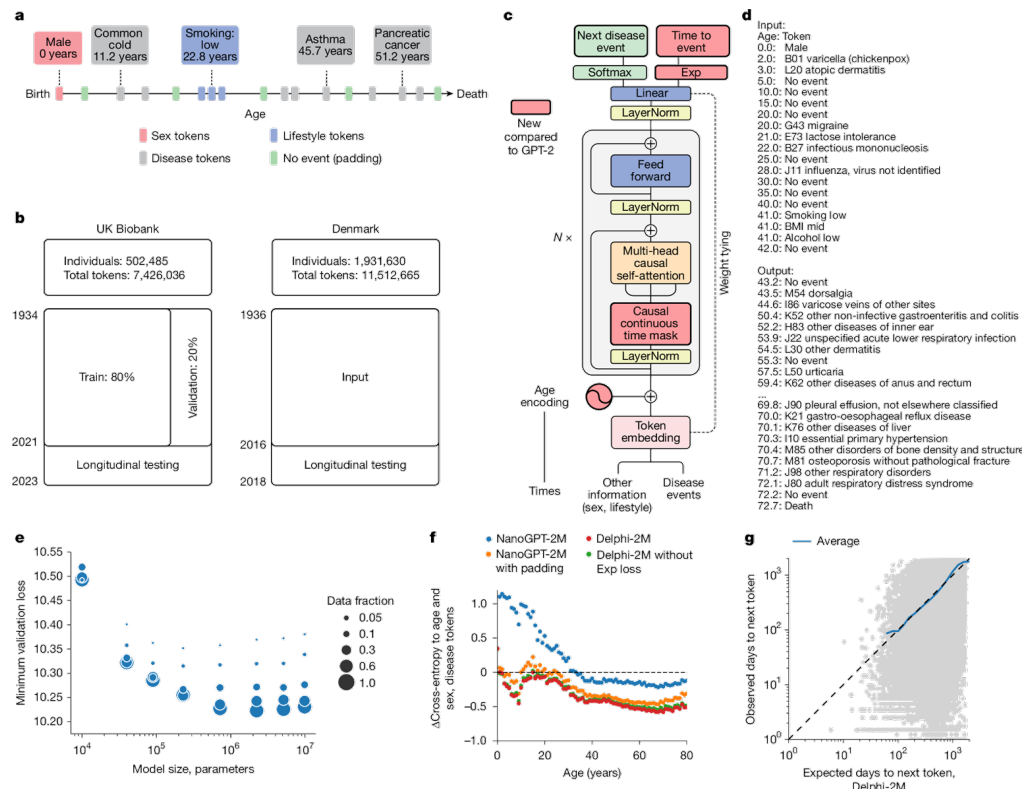
- Evo1およびEvo2は既に200万以上のファージゲノムを用いてトレーニングされているが、研究チームは、さらに「教師あり学習」で追加学習をおこなったうえで、5,386塩基、11個の遺伝子から構成されている、比較的小さなゲノムを持つ一本鎖DNAウイルスである ϕ X174をテンプレートとして、抗生物質に耐性のある大腸菌株に感染するようなゲノムを生成。
- 研究者らはAIの生成したゲノム情報に従ってDNAを合成し、それをファージとして増やした後、目的とする大腸菌に感染させ、それを死滅させることができるかどうか調べた。AIモデルが提案した302種類のゲノム設計のうち16種類が実際に機能し、大腸菌の成長を阻害したり死滅することを確認。



(参考) 疾患予測モデル Delphi-2M (2025年9月)

欧州分子生物学研究所(EMBL)、独がん研究センター(DKFZ)、コペンハーゲン大学の研究チームが、英国バイオバンク(UK Biobank)およびデンマークの匿名化されたデータを活用し、患者の記録(生活習慣要因や他の健康状態など)における特定の疾患発生時期の相対的なパターンを検出するAIモデル「Delphi-2M」を開発。

- このAIモデルは、GPT-2のアーキテクチャを拡張したモデルにUK Biobankの40万人分のデータを用いて訓練され、デンマークの190万人の患者データを用いて検証された。研究チームは**個人の病歴に基づき1,000種類以上の10年以上先の疾患発生確率を予測**できることを確認。
- その精度は、各ツールで予測対象となる疾患数がはるかに少ない既存ツールと同等かそれ以上であり、**さらに最大20年先までの将来の健康経路をシミュレーション**し、プライバシーを保護しつつ他のAIモデルの訓練に有用な合成データを生成することも可能なことを確認。
- このAIモデルは、厳格な倫理規則のもと匿名化された健康データを用いて訓練されており、英国のバイオバンク参加者はインフォームド・コンセントを行い、デンマークのデータはデンマーク国内に留まることを義務付ける国内規制に従ってアクセスされている。
- なお、今後より多様な健康情報を取り込むことで、個別化医療の向上にも貢献するとみられるが、著者らは、本モデルが学習データに内在するバイアスを反映している点、および追加検証なしに直接的な医療判断に用いるべきではない点を指摘している。



- 2024年6月に英国・ロンドンで、英国王立協会 (The Royal Society) と米国科学アカデミー (National Academy of Sciences, NAS) が、AIの台頭が科学研究にもたらす機会と課題に焦点を当てたフォーラムを共同開催。
- 2日間にわたるフォーラムでは、6つのセッションを通じて、科学におけるAIの活用に関する重要なテーマを取り上げ、先駆的な事例を取り上げながら、AI がオープンサイエンスを前進させ、画期的な発見を促し、地球規模の課題に取り組む可能性について議論されるとともに、卓越した科学を支える厳密性(rigor)・誠実性(integrity)・包摂性(inclusivity)の原則を守りながら、この変革的技術を活用するうえで科学者が負う責任について議論を行なっていく必要性なども確認された。

セッション 1: 科学研究を支えるAI科学の登場	AIの歴史・学問分野の現状とともに、AIが現実世界のデータからテキストや画像の形で新しい（合成的または人工的な）データを生成できるようになったことについて、タンパク質構造予測を可能としたAlphaFoldの他、天文学、量子化学、化学分野、教育分野等での活用事例を通じて紹介され、AIを科学研究にどのように統合できるかについて議論された。
セッション 2: 科学的発見の最前線にあるAI：学際的な変革的応用	AIが科学的発見をいかに変えてきているか、材料科学(Materials Project等)、ロボットによる児童実験、天気予報(Google DeepMindのGraphCastやGenCast等)におけるAIの活用事例など、分野横断的な先駆的な事例を紹介しながら、更なる可能性とともに、現時点での限界も議論された。
セッション 3: AIによって強化された科学的進歩とAIによって可能になった新しい科学	AIが多様な分野で科学研究の実践をどのように変えてきたか、またそれに伴う課題が議論された。生物学等でのAI活用事例の紹介を通じて、AI が非 AI 専門の科学者にも、これまでアクセス不可能または利用困難だったデータを解析する能力を与え、新たな研究の可能性を開いたことが議論されるとともに、複雑な生物学的課題に取り組むには、相関に基づく予測を超えて“因果推論”に焦点を移す必要があることなどが議論された。
セッション 4: オープンサイエンス、再現性、複製可能性、プライバシー	AIを活用した研究における、オープンサイエンスの機会と課題、再現性(reproducibility)、複製可能性(replicability)の確保等の課題について、医療分野でのAI活用の実装・規制実務上の課題を含めて、議論された。
セッション 5: 最近の出版物に関するパネル	AIが科学研究に与える影響に関するレポート（米・PCASTレポート <i>Supercharging research: Harnessing artificial intelligence to meet global challenges</i> と、英国王立協会レポート <i>Science in the age of AI: How artificial intelligence is changing the nature and method of scientific research</i> ）が検証された。
セッション 6: 科学における責任あるAI	科学におけるAIの責任と倫理を伴った活用や、人類の利益に資する活用を支えるために取るべき措置について、バイアス、再現性、幻覚（hallucination）などの課題にも触れられながら、議論された。

(出典) <https://royalsociety.org/news-resources/projects/us-uk-forum-2024/>
https://royalsociety.org/-/media/events/2024/06/us-uk-forum/us-uk-scientific-forum-2024_programme-booklet.pdf
<https://www.nasonline.org/event/us-uk-scientific-forum-on-science-in-the-age-of-ai/>
https://www.nasonline.org/wp-content/uploads/2025/11/2024-US-UK-Forum-on-Science-in-the-Age-of-AI_Formal-Summary.pdf

米国・AI Action Plan (2025.7) : AI-Enabled Science

<https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf>



- 2025年7月23日、米・ホワイトハウスが公表した「**America's AI Action Plan**」において、AIそのものの開発やインフラ整備だけでなく、
 - AIを活用した科学研究（AI-Enabled Science）への投資
 - 世界最高水準の科学データセットの整備
 - バイオセキュリティへの投資等についても言及。
- 今回の全23ページのアクションプランは、1月23日の大統領令に基づき、180日間かけてまとめられたもので、科学技術政策局(OSTP)のマイケル・J・クラツィオス局長、デビッド・サックスAI・暗号資産特別顧問、マルコ・ルビオ国家安全保障担当大統領補佐官(国務長官)が執筆。取りまとめに当たり、本年2月には、NSFとOSTPから情報提供依頼(RFI)が出され、10,000人以上が意見提出。

【アクションプランの構成と主な取組】

■柱1 AI イノベーションの加速

- オープンソース及びオープンウェイトの AIの推奨
- AIを活用した科学研究（**AI-Enabled Science**）への投資
(生物学や神経科等の分野での自動化されたクラウド対応のラボへの投資含む)
- **世界最高水準の科学データセットの整備**
(将来の生物学的基盤モデルの構築も見据えた全米生物の全ゲノム解析プログラム含む)
- AIの科学(Science of AI)の促進

■柱2 米国のAIインフラの構築

■柱3 国際的なAI外交・安全保障の主導権

- **バイオセキュリティへの投資** (悪意ある病原体等の合成を防ぐための実効性あるメカニズムの構築含む)

米国・NAIRR Pilot (National Artificial Intelligence Research Resource Pilot)

<https://nairrpilot.org/>

概要：2024年に開始された、NSFやDOE等の複数の連邦機関とAmazonやNvidia等の民間・非営利団体が連携して研究コミュニティに、計算資源やデータセット、訓練など必要な研究リソースを提供する全米規模の取り組み。

申請ウェブサイト

深いパートナーシップ 新しい機会

選択したパートナーリソースとのより深いコラボレーションへのアクセスを直接リクエストします。

[今すぐ申し込む](#)

スタートアップ・プロジェクトのリソース 新しい機会

スタートアッププロジェクトのAIリソースへのアクセスをリクエストします。

[今すぐ申し込む](#)

研究リソース 申請が必要

AI研究に合わせたハイパフォーマンス コンピューティング プラットフォームにアクセスします。

[今すぐ申し込む](#)

教育リソース 申請が必要

教育プラットフォーム(計算ノートブックなど)へのアクセスを要求します。

[今すぐ申し込む](#)

データ、モデルなど 自由にアクセス可能

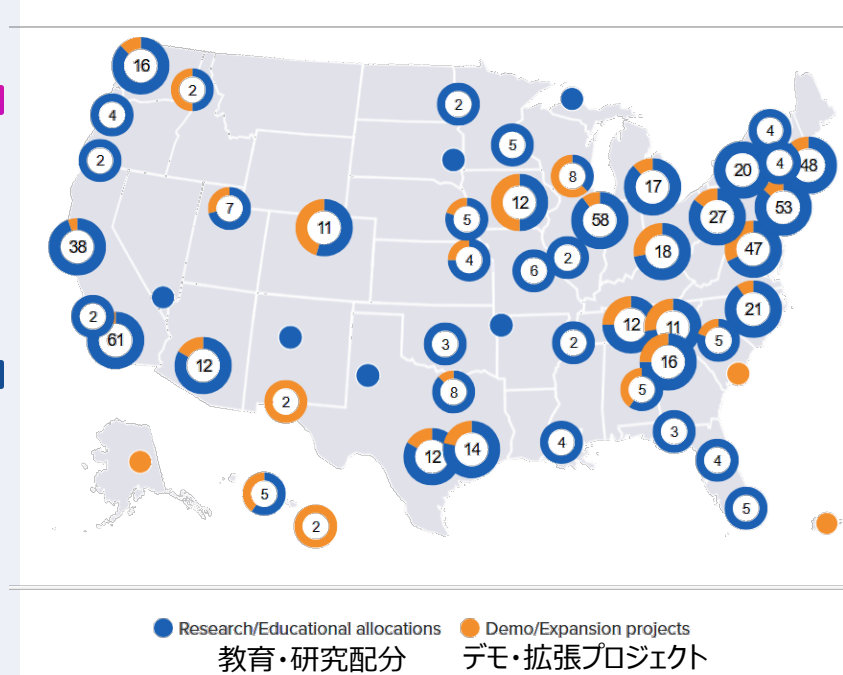
厳選されたデータセット、事前トレーニング済みモデル、AIシステムのトレーニングとテストのための追加ツールを参照します。

[リソースを見る](#)

現在の機会

【計算資源の申請・配分の仕組み】

- 研究者や教育者は、NAIRR Pilotのウェブサイトを通じて、AI研究に必要なリソースへのアクセスを申請。
- 申請にはリソースの使用目的、必要なリソースの詳細、予算見積もり等含む提案書の提出が求められる。
- 提案書は、NAIRR Pilotの評価基準に基づいて審査され、承認された場合、リソースが配分される。
- スタートアッププロジェクトは申請から2週間以内に審査され、承認されると最大3カ月利用可能。



2025年8月28日、米・NSFが、NAIRR Pilot 強化のための下記の方策を発表

- ・Integrated Data Systems and Services (NSF IDSS)
- ・大学等が保有する10の戦略的データセットを選定し NAIRRの枠組み内での利用を可能に

- 欧州委員会は10月8日、「AI活用(Apply AI)戦略」と「**科学分野でのAI(AI in Science)戦略**」を発表。

- 「AI in Science 戦略」では、科学のためのAI資源を集約する仮想研究機関「**RAISE(Resource for AI Science in Europe)**」を通じた具体方策を4本柱で提示

◆卓越性と人材(Excellence & Talent):

世界中の科学人材や高度なスキルを持つ専門家を欧州に惹きつける。これには、RAISEパイロットによるDoctoral NetworksおよびNetworks of Excellenceが含まれる。

◆計算資源(Compute):

Horizon Europeから€600Mを提供し、EUの研究者やスタートアップのAIギガファクトリーへの専用アクセス権を確保。

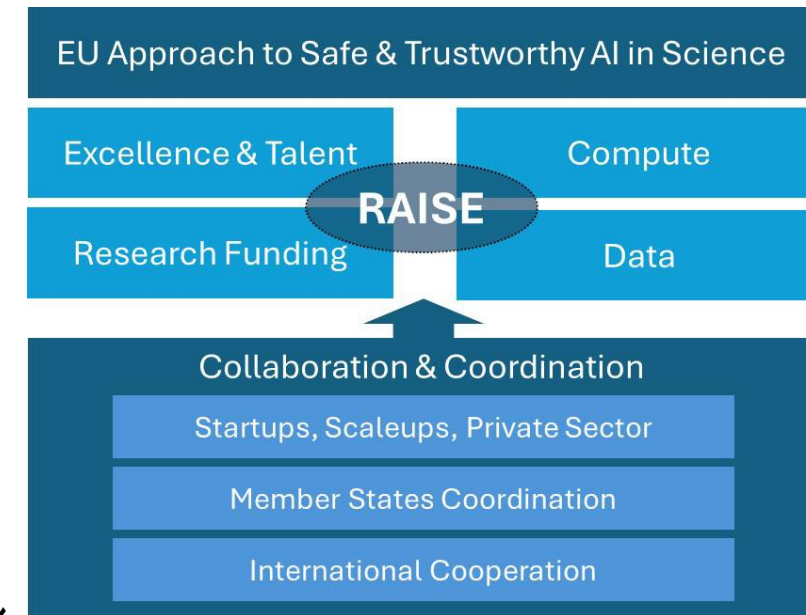
◆研究資金:

Horizon EuropeのAI投資を年間€3 billion以上に倍増させ、**科学分野のAIへの資金を倍増**させることを目指す。

◆データ:

AIを科学分野で活用するために必要なデータセットの収集、キュレーション、統合を支援。

- AI in Science の主要分野(Key Sector)として「先端材料 (advanced material)」と「**バイオ (biotechnologies)**」に言及
- 欧州共同研究センター (JRC) は、同日付で「科学研究におけるAIの役割-政策のための科学」と題した報告書を公表し、本戦略の裏付けとなるエビデンス等を公表。
- 11月3、4日にデンマーク・コペンハーゲンで「**AI in Science Summit 2025**」を開催



英国・AI for Science Strategy (2025.11.21)

<https://www.gov.uk/government/publications/ai-for-science-strategy/ai-for-science-strategy>



英政府は、AI for Scienceを加速するための**具体の15項目**を示し、「AI機会行動計画(AI Opportunities Action Plan, 2025年1月)」における2026～30年の政府投資 £ 20億のうち**最大 £ 1.37億**を充てると発表。また**優先 5 分野**を示すとともに、官民で挑む野心的な目標(**ミッション**)も掲げた。

【目的 (Objective)】

- AI を活用した科学 (AI-driven science) の最先端を切り開く能力を高める ⇒Action 1 & 2
 - 英国の科学分野におけるリーダーシップの地位を維持することを確実にする ⇒Pillar1～3の下に、Action 3～14
- + ミッション (Action 15)

【優先 5 分野】

英国が既に強みを持ち、AIにより大きなインパクトが期待される「**エンジニアリング・バイオロジー**」、「核融合」、「材料科学」、「**医療研究**」、「量子技術」(※「産業戦略(Modern Industrial Strategy)」とも連携)

■ AI駆動科学

(Action 1) AI駆動科学を加速。Sovereign AI Unitが「**自律型ラボの開発・拡大**」について、またARIAが「AI scientist」について公募を実施。
また、責任ある安全なAI活用について、**バイオセキュリティへの影響などドメイン特有の課題にも対処**。

(Action 2) メタ科学 (metascience) の研究を支援。AIが科学研究そのものをどのように変えるか、UK Metascience Unitと連携して、探求。

■ データ (Pillar 1)

(Action 3) UKRIが資金提供した研究によるデータを可能な限り保存・キュレーションし、FAIR 原則に準拠させる。2026年にデータポリシー更新予定。

(Action 4) DSITは、**高価値データセットを特定し、整備**。Renaissance Philanthropyと連携し、まず優先5分野のデータを精査・収集を開始。

(Action 5) UKRIは、ネガティブ実験データも含む**ダーク・データ (dark data)の収集パイロットプログラム**の開始。

(Action 6) 高価値データセットの安全な活用のための**大規模なデータストレージインフラ**を、国のスーパーコンピューターの近くに整備。

■ 計算資源 (Compute) (Pillar 2)

(Action 7) AI Research Resource(AIRR)を通じた公募を開始。**研究者向けにIsambard-AIやDawnスパコン上でのGPU時間を提供**。

小規模向け：最大10,000 GPU時間/3カ月間、ミッション重点プロジェクト(医薬品開発等)：最大1,400,000GPU時間(2週間程度)

(Action 8) 英国内の研究機関を結ぶフェデレーテッド・コンピュータ・クラスタ・ネットワークの構築。未使用／余剰の計算資源の活用可能性を追求。

■ 人材と文化 (People and Culture) (Pillar 3)

(Action 9) 今後5年間、AI for Science博士課程プログラムの拡充等を通じて、**AIを研究に活用できる研究者を1,000人以上育成**する。

(Action 10) 学際的なフェローシップ・プログラムを活用し、**トップ研究者にAI能力を身につけさせながら、コミュニティのスキルアップ**を図る。

(Action 11) あらゆる分野やキャリア段階の研究者や技術専門職(RTP)のAIスキル向上のための**トレーニングプログラムの創設・提供**。

(Action 12) 優先分野におけるAI for Scienceのブレークスルーを推進するための**学際的研究チームの構築** (UKRIによる「ハブ」型研究支援)。

(Action 13) 研究技術専門職(Research Technical Professional RTP)の育成・確保への投資と、そのためのキャリアパスの整備。

(Action 14) **AIモデルのコミュニティ主導のベンチマーク・評価の開発支援**。(例：構造生物学のCritical Assessment of Structure Prediction)

■ ミッション (Missions)

ミッション 1 は、「**前臨床段階**(develop trial-ready drugs)を**2030年までに100日以内**とできるように創薬を加速し、新治療薬の早期展開に貢献」

(Action 15) 追加のAI for Scienceミッションをいくつか選定し、2026年に開始。(GO-Scienceのホライズン・スキャン等も活用、アカデミア等とも対話)

第2節 AI開発力の戦略的強化

- 我が国が独自にAIエコシステムを研究及び開発することのできる能力を強化。AIエコシステムを日本国内に構築、積極的に海外に展開することで、国際競争力も強化し、国力強化、デジタル赤字の解消に寄与。国家主権と安全保障の観点も踏まえ、日本の自律性・不可欠性を確保する。
- 開かれた形での開発を志向し、国内外からのトップ人材を集約し、日本の強みとなる産業・医療・研究といった分野の質の高いデータを生かし、日本国内におけるAI開発力を高める。
- 日本の勝ち筋として、AIモデルとアプリを組み合わせた多様なサービスの創出、AIとロボット等を組み合わせたフィジカルAIの開発導入、科学研究に広くAIを活用するAI for Science等の推進。
- 国家主権と安全保障の観点や日本の文化・習慣等も踏まえた信頼できる基盤モデルを開発。
- AI開発・利活用基盤の増強・確保のため、データセンター、ネットワーク、計算資源、電力等、AIインフラ等の整備を加速。
- AIエコシステムが持続可能な形で発展できるよう、官民が戦略的に投資。

【具体的な取組例】

(1) 日本国内のAI開発力の強化

- ① 新たなデータセット、AIの開発に必要なマルチモーダルなデータの創出・提供等のデータ連携基盤の構築【○内閣府(AI室)、デジタル庁、総務省、文科省、経産省】
- ② トップ人材を含め、国内外からAI開発者を確保するための待遇面や生活環境を含めた包括的な取組【○内閣府(AI室)、文科省、経産省】
- ③ 先進的な知見を取り入れるため、大学・研究機関や国内外の民間事業者等との連携・協働を推進【○内閣府(AI室)、文科省、経産省】
- ④ AIモデルの高性能化やマルチモーダル化の促進【○内閣府(AI室)、総務省、文科省、経産省】
- ⑤ AI評価基盤やテストベッドの整備【○内閣府(AI室)、総務省、文科省、経産省】
- ⑥ グローバルサウスを含めたAI産業の海外市場展開支援【総務省、○外務省、経産省】

(2) 日本の勝ち筋となるAIモデル等の開発推進

- ① AIロボットの公的需要創出や自動運転技術を含めたフィジカルAIの開発・実証の促進【○内閣府(AI室)、文科省、経産省】
- ② 基盤モデルの開発・利活用、研究データ創出・利活用の高効率化、情報基盤の強化等によるAI for Scienceの推進【文科省】
- ③ 新業の開発を高度化する創業AIを推進【厚労省】
- ④ 製造業、インフラ、コンテンツ、金融等におけるAIと融合した新たなビジネスモデルの追求【○内閣府(AI室)、関係省庁】

(3) 信頼できるAI基盤モデルの開発

- ① 日本の文化・習慣等を踏まえた信頼できるAIの開発・評価、質の高い日本語データの整備・拡充(既存の集積データの利活用含む)【○内閣府(AI室)、デジタル庁、総務省、文科省、経産省】
- ② 日本が、信頼できるAIを開発する拠点となり、オープンなAIモデルも含めたエコシステムを世界に広げる【○内閣府(AI室)、総務省、外務省、文科省、経産省】

(4) AI開発・利用基盤の増強・確保

- ① AIデータセンター、効率的な電力・通信インフラの整備(ワット・ビット連携)、オール光ネットワーク、次世代情報通信基盤の研究及び開発【○総務省、文科省、経産省】
- ② 高性能AI半導体等の研究及び開発【総務省、文科省、経産省】、富岳NEXTの開発【文科省】
- ③ AIインフラに関するサプライチェーンの強化【総務省、○経産省】

第2節 危機管理投資・成長投資による強い経済の実現

1. 経済安全保障の強化

（1）戦略分野の官民連携投資、重要物資のサプライチェーン強化

（A I・半導体、造船、量子、フュージョン、バイオ、航空、宇宙 等）

（略）

（A Iの開発・社会実装とそれを支える半導体・データセンターの支援）

A Iを基盤とした経済社会の発展を実現するため、人工知能関連技術の研究開発及び活用の推進に関する法律（A I法）及び同法に基づく人工知能基本計画及び適正性確保のための指針に則り、イノベーション促進とリスク対応の両立を推進する。A I利活用の加速的推進、A I研究開発力の戦略的強化、A Iガバナンスの主導、A I社会に向けた継続的変革の方針に基づき、大胆な規制改革を含む施策を内外一体で進め、国内研究開発の強化と社会実装の促進を図る。AI for Science の戦略方針を2025年度内に策定し、そのうち科学基盤モデル開発等の取組を先行的に加速する。併せて、A I研究開発に不可欠な国内人材育成・確保、質の高いデータ整備及び利活用の促進、評価基盤の構築及び研究開発力の強化と利活用を促進する。生成A Iの開発と実装を一体的に支援するとともに、日本が強みをもつ産業とA Iを融合した多様なサービスの創出、計算資源・情報通信基盤の関連インフラの高度化、積極的な海外展開を進める。また、A Iの安全性を高め、信頼できるA Iの活用を起点としたイノベーションを促進するため、A Iセーフティ・インスティテュートを抜本的に強化する。さらに、A Iと日本の高度なロボティクス技術を融合した世界最先端のA Iロボティクスを通じて、人手不足の解消、生産性向上やD Xを実現するべく、その実装拡大に向けた導入環境の整備と競争力強化を図るための戦略を策定する。（略）

施策例

・AI for Science による科学研究革新（文部科学省）

6. 集中的に取り組む事項の実現に向けた具体的取組 (4) 時代に即した研究環境の構築

(4) — 1 AI for Science による科学研究の革新

近年、AIを科学研究に組み込むことで、研究の範囲やスピードに飛躍的向上をもたらす「AI for Science」が、創造性・効率性などの観点で科学研究の在り方に急速かつ抜本的な変革をもたらしつつあり、AI研究開発力が科学研究力に直結する時代になってきている。科学の再興のためには、急速に進展する国際潮流に乗り遅れることなく、AI駆動型研究を支えるオープンサイエンスを更に推進するとともに、経済安全保障上も重要なAI研究における国際競争力の確保や、日本固有の強みを生かして技術的優位性・不可欠性を確保するための分野横断的・組織横断的な「AI for Science」の先導的実装に向けた取組を切迫感・危機感を持って推進することが喫緊の課題である。

そのため、強力な政策誘導により、戦略的かつスピード感をもって AIイノベーションを推進し、多くの意欲ある研究者及び先端的研究リソースのポテンシャルを最大化する研究システムの革新を実現する。更には産学官において広範に実装することで、我が国の研究力・国際競争力の抜本的強化につなげていく。

【具体的取組】

① AI 利活用研究 (AI for Science) と AI 研究 (Science for AI) の推進

我が国においては、世界最高水準の情報基盤を有するとともに、ライフサイエンス・マテリアル等の重点分野においてAI開発・利活用の要となる質の高い実験・観測データなどの強みを有しており、これらのリソースを最大限活用して、科学基盤モデルやAIエージェントの開発・活用、フィジカル AIも活用した研究活動の自動・自律化などによる次世代AI駆動ラボシステムの開発などのAI利活用研究を強力に推進し、研究の効率性・生産性を向上させ、研究者の創造性を最大化させることが重要である。

加えて、AI時代における社会インフラともなるAIについては、他国に依存することなく、自国で研究開発する能力を保持することは経済安全保障上も極めて重要であり、透明性・信頼性の確保や、数理科学、データ科学等の活用による AI そのものの研究開発を抜本的に強化することが必要不可欠である。

そのため、AI利活用研究 (AI for Science) においては、研究費や計算資源などのリソースを戦略的かつ機動的に分配・整備しながら、我が国の勝ち筋となり得るライフサイエンス・マテリアル等の重点分野における先導的・先駆的な研究と、あらゆる研究分野においてAI for Science の波及・新興を促進しつつ、若手研究者をはじめとした意欲ある研究者による次の種や芽となる新たなアイデアへの挑戦などの萌芽的・探索的な研究を通じて我が国独自の競争優位を築くための取組を両輪に、柔軟かつ機動的な研究開発を新たに推進する。

また、AI 研究 (Science for AI) においては、国際競争力の観点も踏まえて、数理科学を基盤とした AI 技術に関連する最先端の理論研究やアルゴリズム開発等の基礎研究を推進するとともに、ハルシネーションやブラックボックス等のリスクや懸念が指摘されているところ、画像、音声等を扱うことのできるマルチモーダルモデルを含めた生成AI モデルの透明性・信頼性の確保に向けた研究開発を推進する。

② AI 駆動型研究を支えるデータの創出・活用基盤の整備

AI for Science の実現のためには、より多くの研究者が AI を活用した研究環境を利用でき、AI-ready なデータの収集、解析の標準化も含めた高品質かつ大量のデータを継続的に生み出し活用できる研究システムの構築が重要である。

そのため、最先端の研究設備を集積し、研究現場が裨益する形で、研究設備の自動・自律化、遠隔化による大規模なオートメーション/クラウドラボの形成を進めることが必要である。また、高品質な研究データを創出・活用するため、全国の研究大学などにおいて、コアファシリティを戦略的に整備するとともに、先端的研究設備・機器の整備・利活用・共用・高度化・開発を推進する。また、AI時代の大量なデータの学習を見据え、再現性・トレーサビリティ等が確保されたバイオリソースのような良質なデータを創出する基盤を安定的に確保・供給していくことも重要である。加えて、ユニークなデータが貴重な研究資産となるところ、知の継承や海外流失の防止も含め、電子化されていないデータやいわゆるレガシーデータの利活用などについても検討することが必要である。

③ AI for Science を支える次世代情報基盤の構築

⑤ 大胆な投資資金の確保・環境整備

④ AI 関連人材の育成・確保

⑥ 推進体制の構築等

AI for Scienceに向けた我が国のライフサイエンス分野の強み

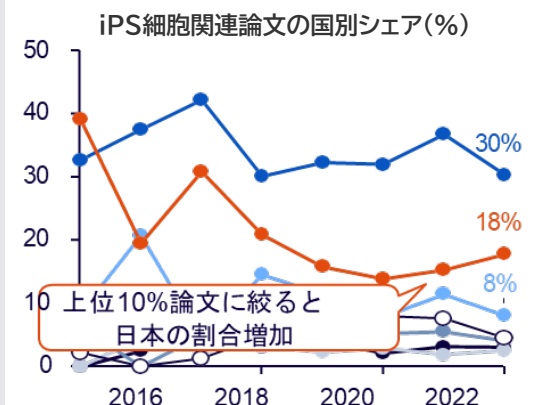
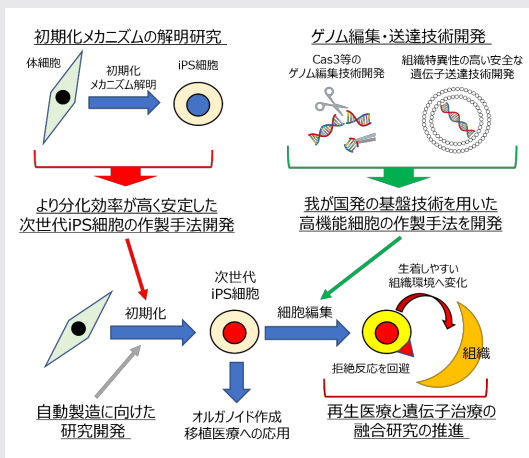


文部科学省

- ①最先端のデータを創出する実験科学: **iPS細胞**を活用したオルガノイドなど、**世界をリードする研究開発分野**を有する
- ②良質なデータを測る技術: **イメージング技術**など我が国が優位性を維持する計測技術
- ③データアセット: ゲノム・蛋白等に関する**良質なデータを多数保有**しており、実験データベースは我が国の強み

①最先端のデータを創出する実験科学

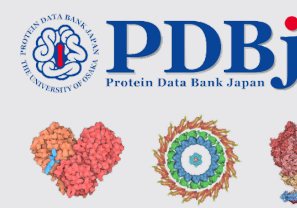
iPS細胞やオルガノイド等の発生・再生分野では、霊長類モデルを含む豊富なバイオリソースと高度なゲノム編集技術を組み合わせた動物モデル基盤により、世界に先駆けたデータ取得が可能。



*アーサー・ディ・リトル作成。Elsevier社のCiteScoreを用いて引用頻度上位10%に相当する学術誌を特定。

③データアセット

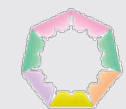
日本が運営主体のライフサイエンスDBは約900存在(世界第2位)し、世界をリードするDBが存在
ナショナルライフサイエンスデータベースプロジェクトのもとでデータベース統合解析技術の開発を推進



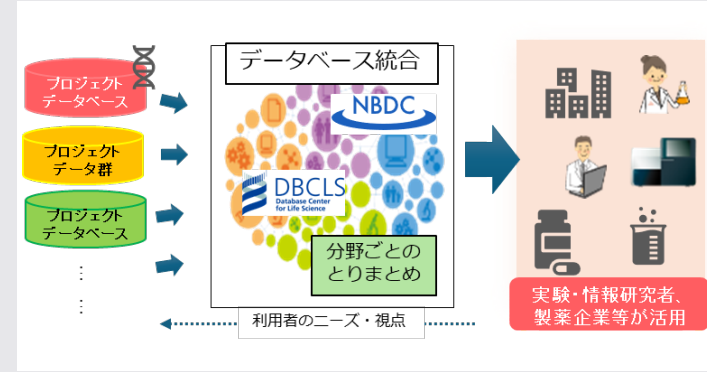
「タンパク質構造予測に関する研究」でノーベル化学賞を受賞したJohn M. Jumper博士らは、PDBのデータを学習させたAlphaFold2をさらにPDBのデータを用いて発展させ、タンパク質複合体の予測技術を高めたAlphaFold3を開発(Nature 2024)

②良質なデータを測る技術

バイオイメージングやゲノムシーケンシングについての高度な解析体制や先駆的な日本企業・大学・研究機関の連携が存在。
空間情報を含む高度なイメージング技術の開発と関連技術の融合により、生体内の空間的配置やダイナミクスまでの統合的理解が可能となる。



文部科学省科学研究費助成事業 学術変革領域研究[学術研究支援基盤形成]
先進ゲノム解析研究推進プラットフォーム
先進ゲノム支援 第2期



ライフサイエンス分野のAIモデルの学習に必要なデータセットを巡る状況

- 膨大な生命科学データが存在する一方、AIモデルの学習にそのまま利用できる形(AI-ready)で整理・標準化されているものは限られている。
(現在のところ、PDB、UniRef、抗体配列データ (OAS等) などの大規模データセットが用いられている)
- シングルセル・マルチオミクスデータについては、CELLxGENE など新しいデータセットが、AIモデルの訓練データとして活用されてきている。
- AI-readyのデータ・インフラの整備・強化の必要性が指摘されている

(参考)ライフサイエンス分野のAIモデルの開発に用いられたデータセットの例

AIモデル	AIモデルの種類	学習に用いられた公開データセットの例
ESM-2	Protein sequence model	UniRef90
Evo	Genomic sequence model	GTDB, IMG/PR, IMG/VR
AlphaFold2	Structure prediction model	PDB
ProGen2	Protein sequence model	UniRef90, metagenomic sequences, Observed Antibody Space (OAS)
AbLang	Antibody protein sequence model	Observed Antibody Space (OAS)
CaLM	Codon sequence model	Coding sequences from European Nucleotide Archive (ENA)
RiNALMo	Noncoding RNA sequence model	Rfam, RNACentral, NCBI
GenSLM	Genomic sequence model	GenBank
Enformer	Multi-omic prediction model	ENCODE datasets
scGPT	Transcriptomic count model	CELLxGENE

(出典) 米国科学工学医学アカデミー(2025年3月)『The Age of AI in the Life Sciences : Benefits and Biosecurity Considerations』
Appendix A TABLE A-4 Example Training Datasets and Associated Models:
The National Academies Press. <https://doi.org/10.17226/283>.

Training dataset size of biological models

EPOCH AI

Training dataset size (amino acids or proteins)



CC-BY

epoch.ai

概要

- AIとバイオテクノロジーの急速な進展を背景に、米国国防総省(DoD)からの検討要請を受けて、AIを活用した生物学的ツールのバイオセキュリティ・リスクへの影響や、逆にそのようなAI活用ツールがリスク低減に活用できるかについて評価・検証。
- 報告書では、AIは生命科学に革命的な恩恵をもたらす可能性があるが、リスクも無視できず、特にバイオセキュリティの分野では慎重なマネジメントが必要であり、データ整備・モデル開発・政策設計をバランスよく進めることが重要であると指摘。

〔メリットの検証〕

- ✓ 生物発見・デザインの高速化、効率化
- ✓ 実験データから仮説生成、パターン発見、設計支援
- ✓ 学び(“Learn”)段階での継続的改善
— フィードバックループ
- ✓ バイオセキュリティ(biosecurity)への
貢献環境中の病原体データ解析

〔リスクの検証〕

- ✓ 悪用可能性(悪意あるバイオ分子や病原体の設計や改変等)のリスクは無視できないが、現時点では、AIモデルの訓練データの不足やウェットラボ実験の必要性等、AIツールには限界があり、複製可能な複雑な感染性のある生物学的システムを構築することは困難。
- ✓ ただし、将来、自動化ラボ・オートメーション実験等が進めば、物理的な実験の障壁が下がる可能性もあり、注視が必要。



主な提言

- 関係機関(国防総省、保健福祉省、農務省等含む)による感染性病原体の基礎生物学研究の継続的支援
- 安全性とイノベーションの両立を目指し、“if-then 戦略”の採用⇒AI 技術の進展に応じて段階的な対策を準備
- データ収集・統合・キュレーション、データインフラの強化

概要



2025年9月、Jay Bhattacharya NIH長官は、米国におけるバイオセーフティ監視の近代化に向けた検討・取組を開始する旨のステートメントを発表。

- ✓ 21世紀にふさわしいバイオセーフティの枠組みを構築することを目指し、今後数か月間、研究コミュニティや関係者と密接に連携しながら、組換え核酸技術や合成核酸技術にとどまらない、潜在的リスクに対応する監視体制を刷新する
- ✓ 低リスクの組換え技術については、従来必要とされていた監視が過剰となる場合があるため、簡素化の検討を進める。
- ✓ 研究機関監督組織との連携を強化し、組換えDNA委員会(Institutional Biosafety Committees: IBC)の権限を強化する。これは、機関内でのバイオセーフティ管理の最前線としての役割を明確化するとともに、ヒト被験者研究や動物実験の審査委員会と同等の支援やリソースがIBCに提供されているかを再評価する機会ともなる。

Phase I 2025年秋	Phase II 冬-春 2026年	Phase III 春-夏 2026年	Phase IV 秋 2026年
- Conduct regional listening sessions and stakeholder events - Build awareness and encourage maximum public feedback 地域別公聴会等	- Synthesize information heard from stakeholder events and listening sessions - Publish Draft Policy for public comment ドラフト公表 ・パブリックコメント	- Assess Draft Policy feedback - Refine policy - Assess guidance needs ドラフト版 見直し	- Publish Final Policy - Release community-driven guidance - Launch rigorous outreach and education efforts 政策公表 (最終版)、 ガイダンス

(出典) <https://www.nih.gov/about-nih/nih-director/statements/nih-launches-initiative-modernize-strengthen-biosafety-oversight>
<https://osp.od.nih.gov/policies/biosafety-and-biosecurity-policy#tab2/>

