

知の集積が知を吐き出す装置になるとき： AIは研究の何を変えるか

東北大学大学院情報科学研究科
西 羽美

自己紹介

- 西 羽美 (にしはふみ)
- 2010年 東京工業大学 生命理工学研究科 博士 (理学) 取得
- 2010年 アメリカ国立衛生研究所(NIH) NLM/NCBI 博士研究員
- 2013年 日本学術振興会 特別研究員 (PD)
- 2014年 東北大学 情報科学研究科 助教
- 2018年 東北大学 情報科学研究科 准教授

生物系
↓
情報系

• 専門：構造バイオインフォマティクス
(タンパク質分子構造の生命情報学)

特に、分子構造のデータベース解析 (データサイエンス)

…のはずが、最近「コピーでAIに研究させる」人だと思われる？



<https://edx.nii.ac.jp/lecture/20251210-03>

出典：国立情報学研究所「AIしか出せない国際会議Agents4Scienceに出してみた — 西 羽美 (東北大学大学院情報科学研究科・准教授)」教育機関DXシンポジウムアーカイブズ, <https://edx.nii.ac.jp/lecture/20251210-03> (参照 2026-07-08) .



2025/11 バイオ生成AI研究会「Agents4Science参戦記」、2025/12 NII教育機関DX (上記スクショ)
2026/3 AI・AGI時代の学術出版「AIに研究のすべてを託すとき」、2026/4 実験医学「AIが研究者を追い越すとき」

今日の話題

- 1. 生命科学における知の集積（生命情報学）と知を吐き出す装置の出現
 - ゲノム・タンパク質構造データの集積とAlphaFold・AlphaFold DB
- 2. AI scientistの出現と民主化
 - AIによる科学研究への影響と自動化の動向
(※たぶんAIに聞いた方が速いし早い)
 - AI scientist作って（もらって）みた&使ってみた
(※たぶんAIに聞いても答えてくれない)
- 3. おわりに（AIが研究者を追い越すとき？）
 - 諸分野における研究自動化への反応



ChatGPTで生成
(相変わらずタンパク質立体
構造を正しく描けない)

1. 生命情報学： 生命科学における知の集積

生命情報学は図書館から派生した？

アメリカ国立医学図書館
(National Library of Medicine, NLM)



画像 : Chris Spielmann, "NIH national library of medicine (7)," Wikimedia Commons, 2013 upload / photograph from NIH/NCI, public domain, [https://commons.wikimedia.org/wiki/File:NIH_national_library_of_medicine_\(7\).jpg](https://commons.wikimedia.org/wiki/File:NIH_national_library_of_medicine_(7).jpg).

生物工学情報センター
(National Center for Biological
Information, NCBI)

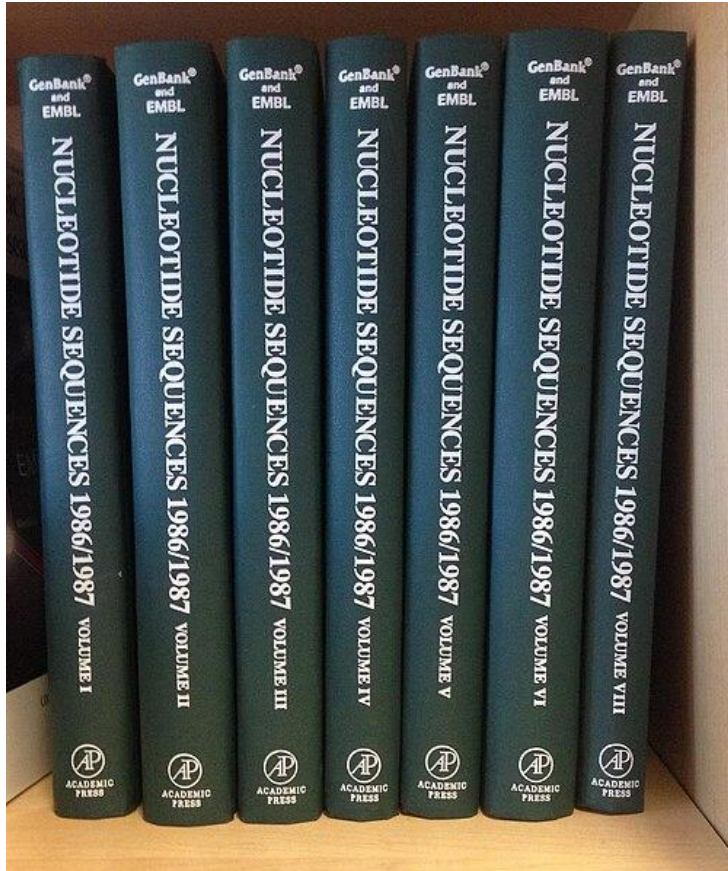


2016年6月撮影

- NCBI: 生命情報学における世界最大の研究センターの一つ・生命科学系国際データベース提供拠点
 - 生物・医学系文献データベースPubMedが代表的
- “1986—NLM’s Long Range Plan was completed; it contained a recommendation that a new NLM Division be created to manage and process molecular biology information.”
- “A Brief History of NCBI’s Formation and Growth”
<https://www.ncbi.nlm.nih.gov/books/NBK148949/>

出典 : Kent Smith, “A Brief History of NCBI’s Formation and Growth,” in The NCBI Handbook, 2nd ed., National Center for Biotechnology Information (US), 2013–, <https://www.ncbi.nlm.nih.gov/books/NBK148949/> (参照 2026-07-08) .

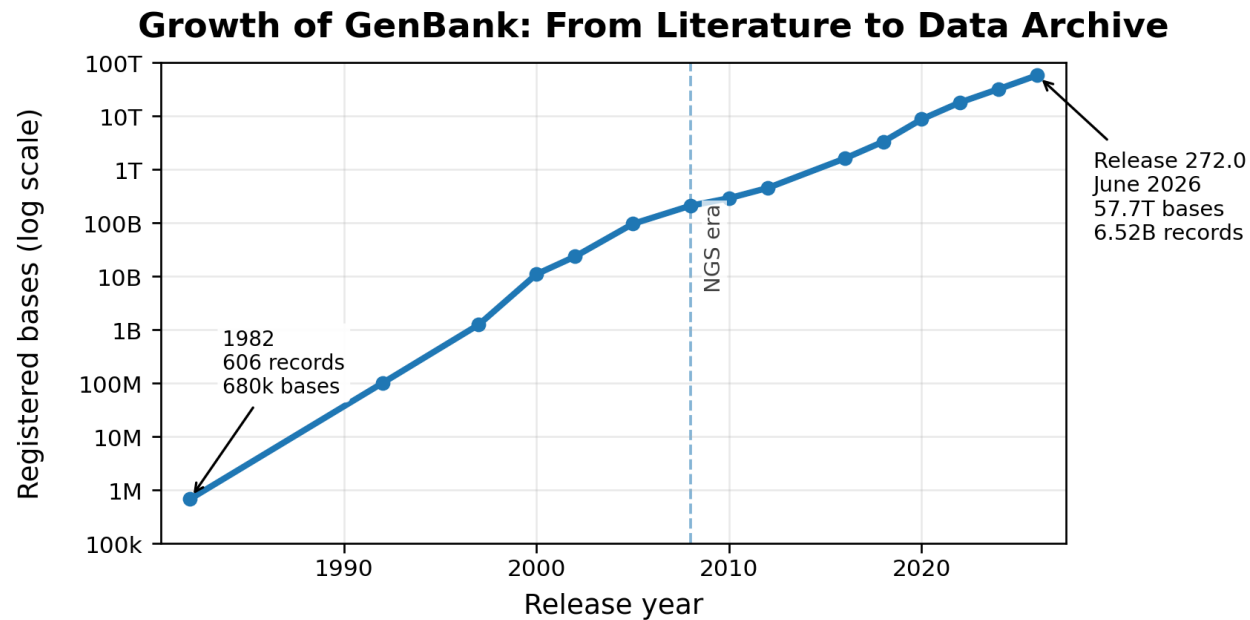
生命情報の蓄積の始まり：GenBank



画像：David Landsman / NCBI-NLM-NIH, "Genbank & EMBL: NucleotideSequences 1986/1987 Volumes I-VII," Wikimedia Commons, public domain (US government work), https://commons.wikimedia.org/wiki/File:NucleotideSequences_86_87.jpeg.

- GenBank：
公開塩基配列の一次アーカイブ
- 1979年：Walter Goadらが Los Alamos Sequence Databaseを構築
- 1982年：GenBankとして開始
(LANLで運用、1992年にNIH/NCBIへ)
- Release 3 (1982年12月)：
606レコード／680,338塩基
- 論文掲載配列の収集・入力・注釈付けから始まった
- 現在はINSDC (DDBJ・ENA・GenBank) で日次交換

「情報の本体」が紙面からデータベースへ

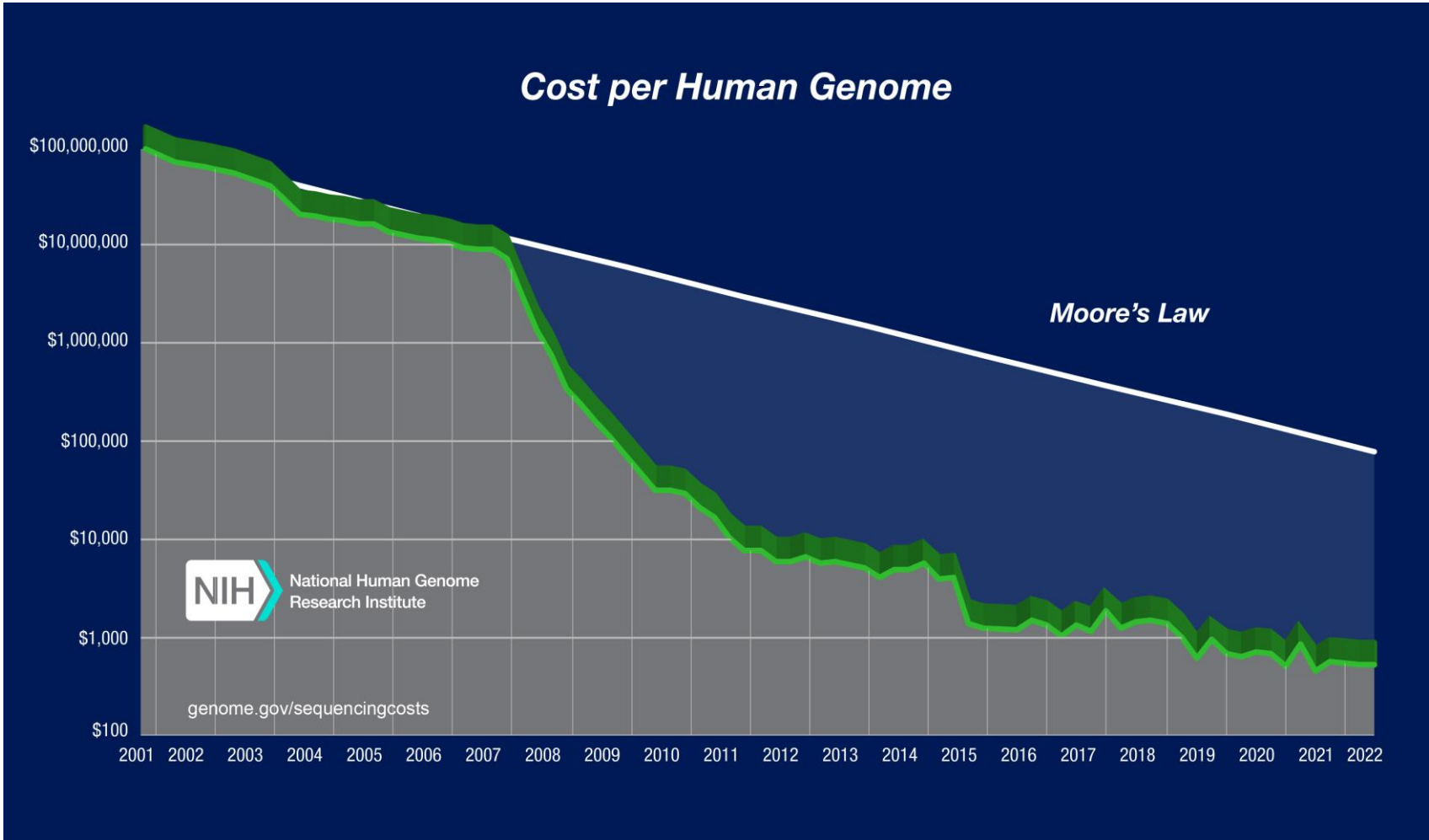


Data: NCBI GenBank Statistics and Release 272.0 notes. 2026 total includes traditional + set-based records.

出典 : NCBI Insights, "GenBank Release 272.0 is Available!", 16 June 2026, <https://ncbiinsights.ncbi.nlm.nih.gov/2026/06/16/genbank-release-272/> ; NCBI, "GenBank and WGS Statistics," <https://www.ncbi.nlm.nih.gov/genbank/statistics/> (参照 2026-07-08) .

- 論文中の配列ではなく登録レコードが一次データになる
 - 投稿前に登録 → accession番号を取得
- Accessionは各レコードの一意・永続ID
 - 番号だけで配列・注釈・引用を取得できる
- Release 272.0 (2026年6月) :
約57.69兆塩基／約65.20億レコード*
 - * traditional + set-based (WGS/TSA/TLS) の合計

ゲノム情報取得コストの減少と情報の爆発

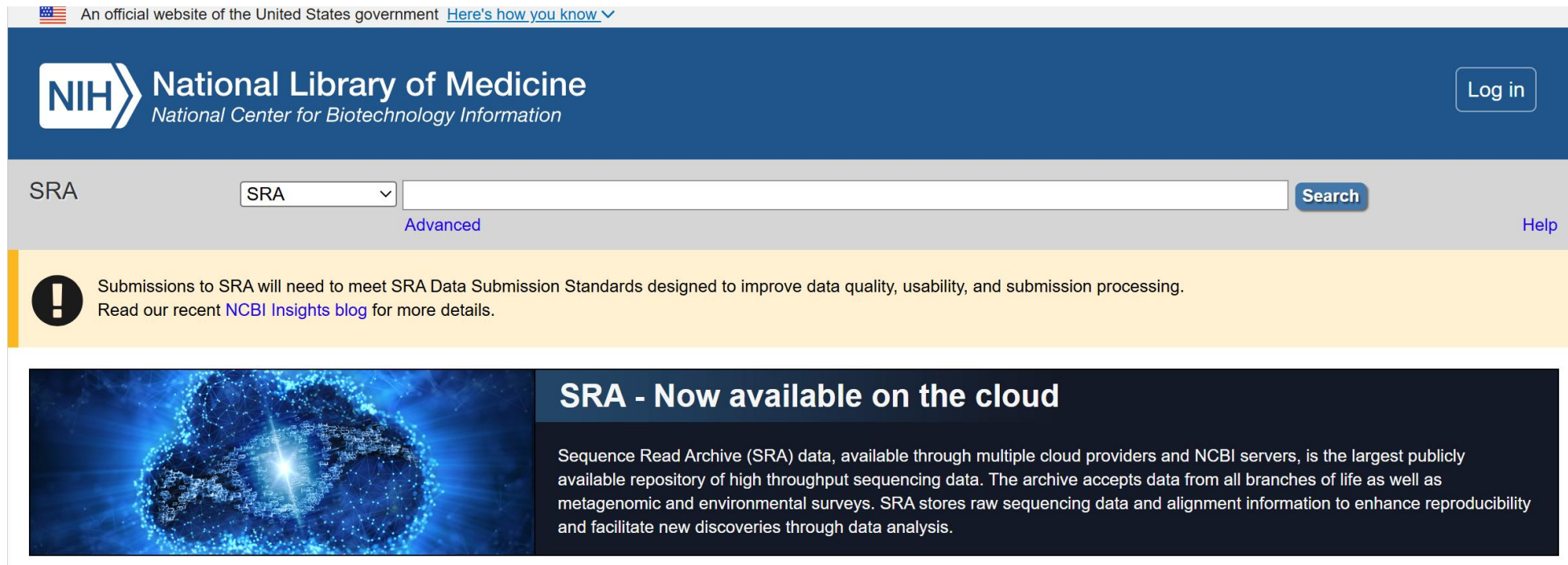


- 2005-2008年：
次世代シーケンサー
の登場と普及
→ゲノムを読み取る
コストの崩壊
- ヒトゲノム1つの解読
は、数十億円規模から、
いまや数万円レベルへ
- GenBank / SRAへ
配列データが爆発的に流
入

<https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>

出典：Kris A. Wetterstrand, "DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP)," National Human Genome Research Institute, <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data> (参照 2026-07-08) .

新たな大量情報の蓄積 (Sequence Read Archive)



出典 : NCBI, "Sequence Read Archive (SRA)," <https://www.ncbi.nlm.nih.gov/sra> (参照 2026-07-08) .

- SRA (Sequence Read Archive) :
次世代シーケンサー等から出力される高スループット配列データの最大級の公開リポジトリ
- 2021年時点で公開分だけで約25.6ペタ塩基対・1,480万run
- GenBankは「遺伝子等の配列カタログ」、SRAは「実験データのアーカイブ」

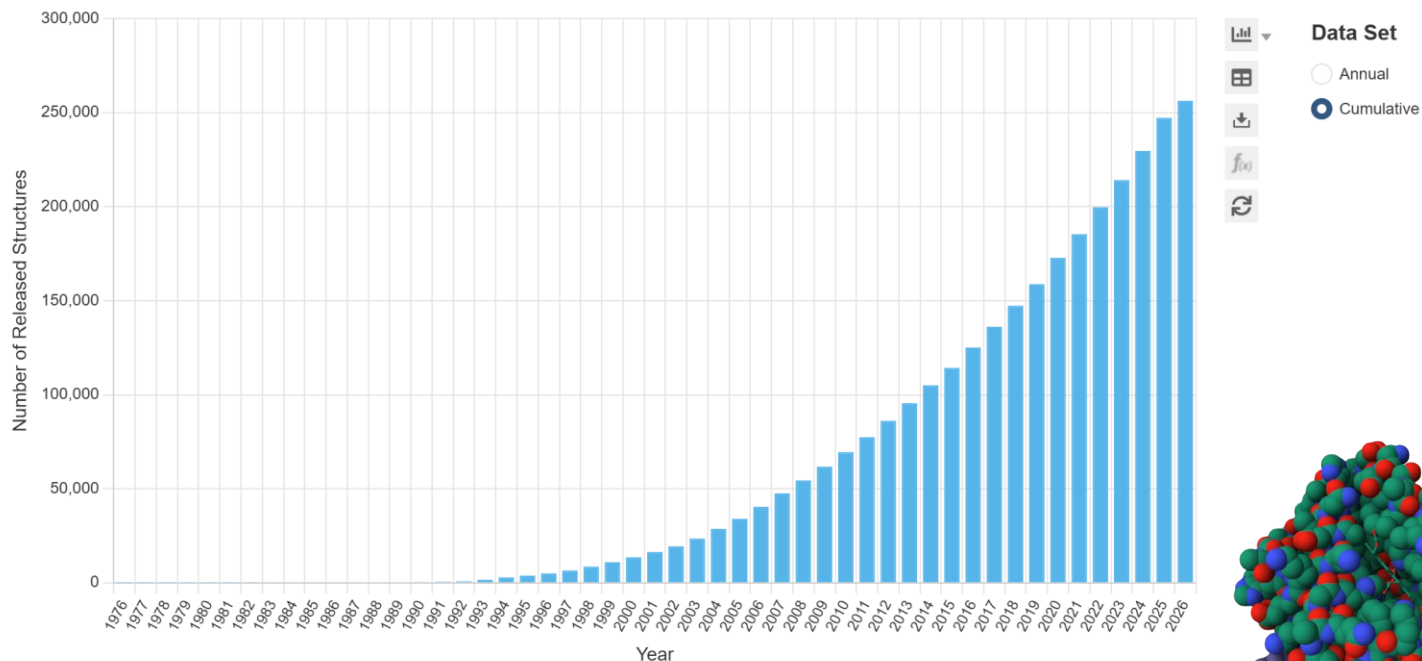
出典 : Katz, K. et al. "The Sequence Read Archive: a decade more of explosive growth," Nucleic Acids Research 50(D1), D387–D390 (2022), <https://doi.org/10.1093/nar/gkab1053>.

タンパク質分子の3次元構造のデータベース



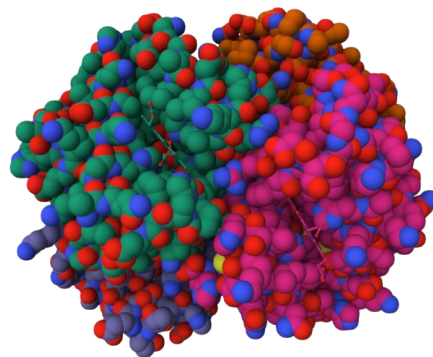
PDB Statistics Data Growth Data Distribution Search API

PDB Data Growth by Released Structures



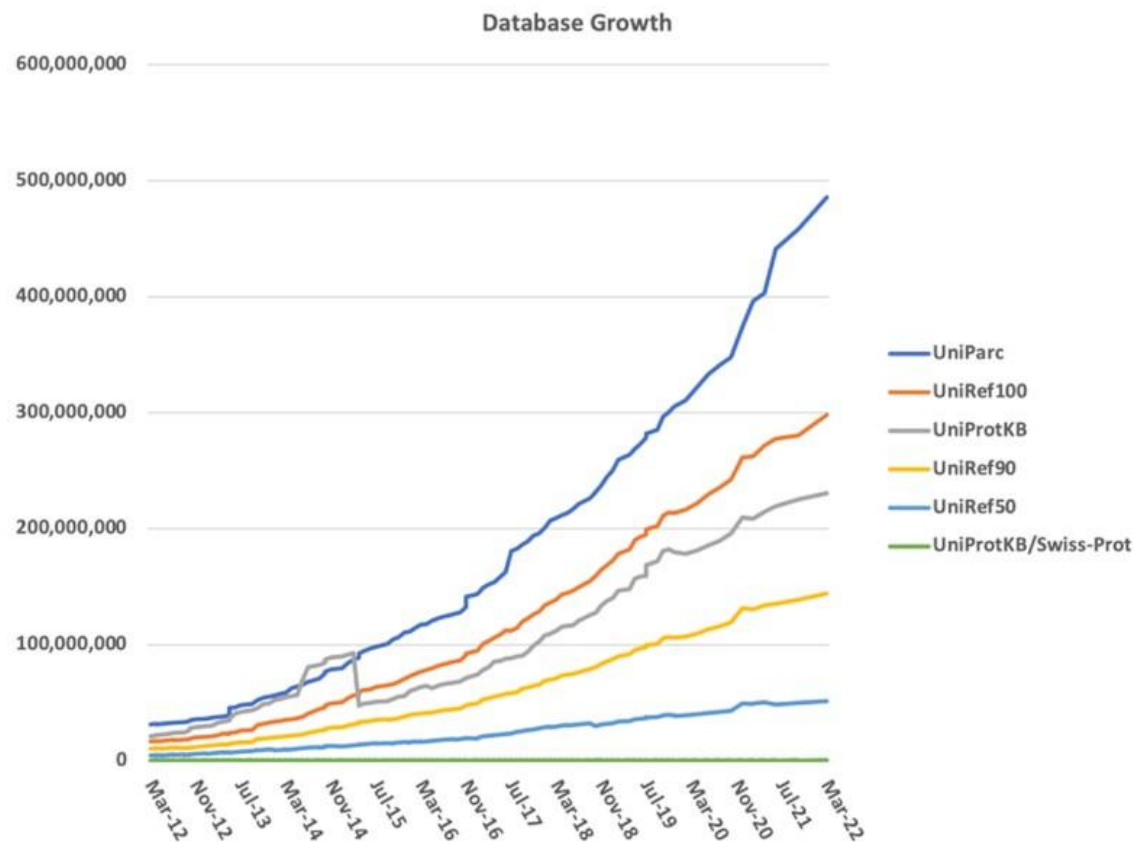
出典：RCSB Protein Data Bank, “Growth of Released Structures per Year,”
<https://www.rcsb.org/statistics/growth/overall-structures> (参照 2026-07-08) .

- 生体を構成する主要な分子：DNA（ゲノム）及びタンパク質
- DNA: 生命の設計図、タンパク質：設計図から作られた生命の部品
- Protein Data Bank (PDB)
 - 実験的に決定された生体高分子3次元構造の単一国際アーカイブ
 - 1971年に7件の構造データから始まる
 - 現在は Worldwide Protein Data Bank (wwPDB) の下で、RCSB PDB・PDBe・PDBj・BMRB等が共同で管理・提供
 - 2026年現在、公開エントリ約25.7万件



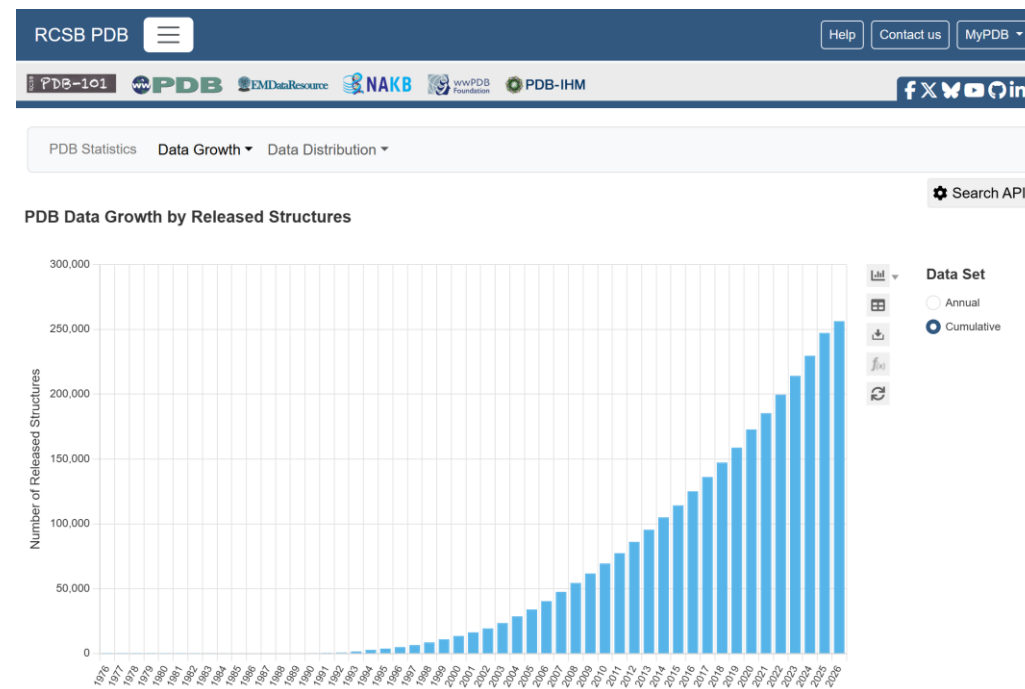
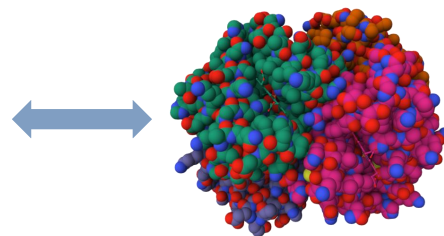
遅い蓄積と速い蓄積：埋まらない（絶望的な）差

- DNA・ゲノム・（アミノ酸配列）は実験的に取得が容易
＝数億エントリの蓄積
- 対応する分子構造の情報取得は非常に困難＝数十万エントリ



出典：The UniProt Consortium, “UniProt: the Universal Protein Knowledgebase in 2023,” Nucleic Acids Research 51(D1), D523–D531 (2023), <https://doi.org/10.1093/nar/gkac1052>.

VLSPADKTNVKA AWGKVG AHAGEYGAEL
ERMFLSFPTTKTYFPHFDLSH GSAQVKGH
GKKVADALTNVAHVDDMPNALSALSDLH
AHKLRVDPVNFKLLSHCLLVTLAAHLP AE
FTPAVHASLDKFLASVSTVLT SKYR



出典：RCSB Protein Data Bank, “Growth of Released Structures per Year,” <https://www.rcsb.org/statistics/growth/overall-structures>（参照 2026-07-08）.

「実験並み精度のAI」 AlphaFoldの衝撃

The world this week

News in focus



A protein's function is determined by its 3D shape.

'IT WILL CHANGE EVERYTHING': AI MAKES GIGANTIC LEAP IN SOLVING PROTEIN STRUCTURES

DeepMind's program for determining the 3D shapes of proteins stands to transform biology, say scientists.

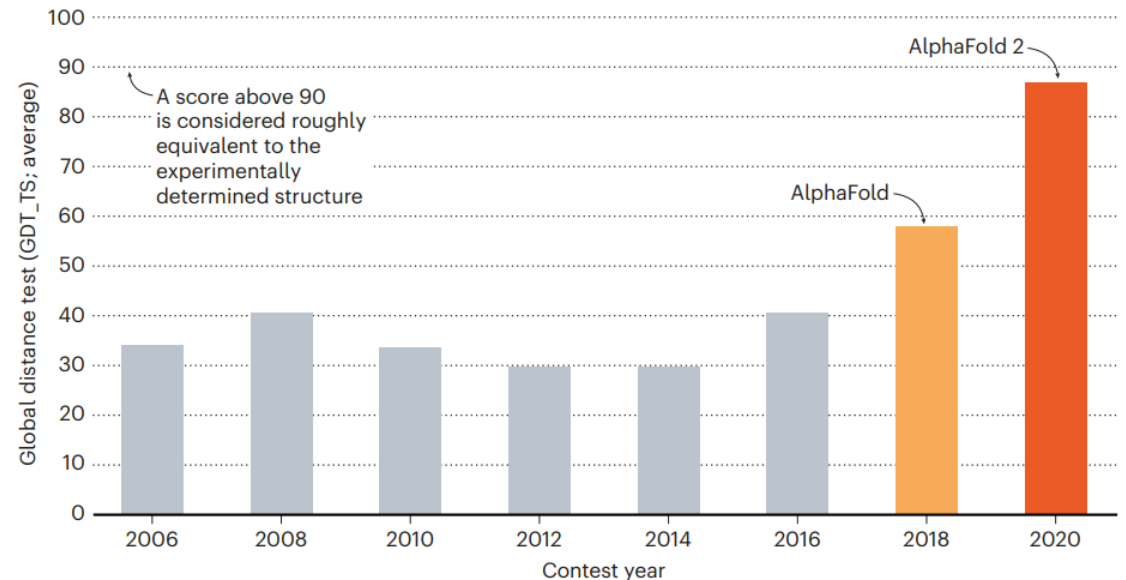
<https://www.nature.com/articles/d41586-020-03348-4>

出典 : Callaway, E. "It will change everything": DeepMind's AI makes gigantic leap in solving protein structures, *Nature* 588, 203–204 (2020), <https://doi.org/10.1038/d41586-020-03348-4>. Graph source: DeepMind; photo credit: Edward Kinsman / Science Photo Library.

- 2020年11月末：構造予測コンテスト CASP14にてDeepMind社のAlphaFold2が実験に迫る予測精度を達成
- 開発者のJohn JumperとDemis Hassabisは、2024年にDavid Bakerとともにノーベル化学賞を受賞

STRUCTURE SOLVER

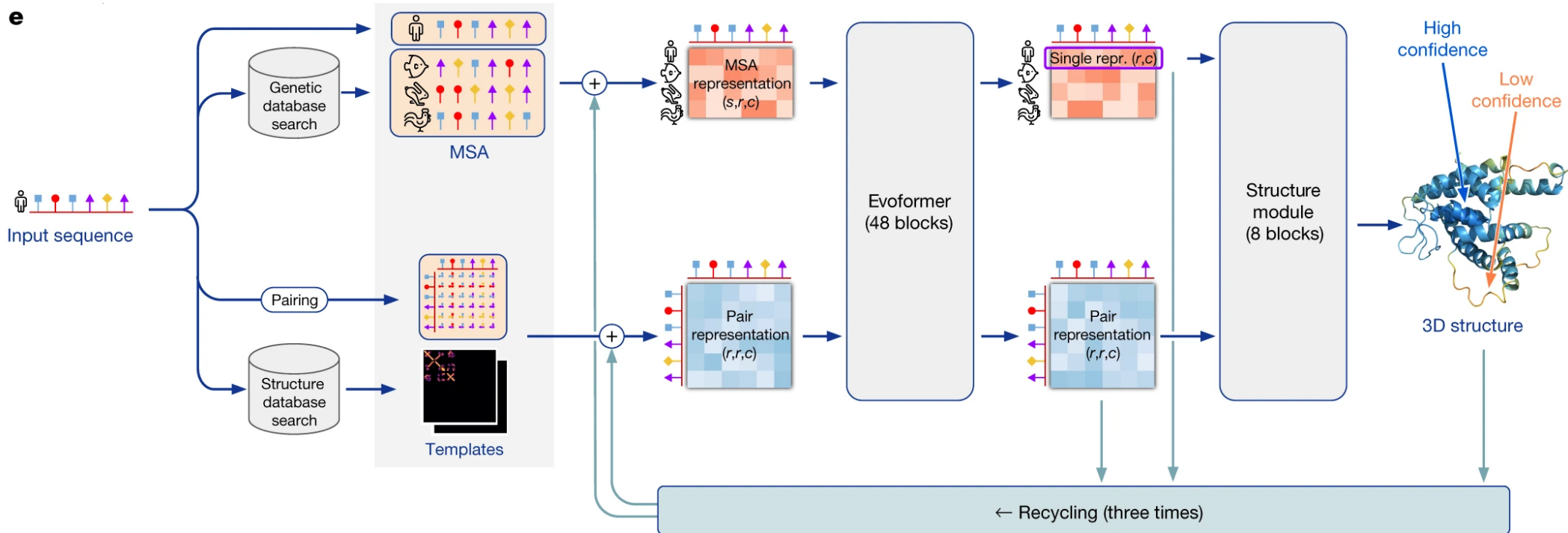
DeepMind's AlphaFold 2 algorithm significantly outperformed other teams at the CASP14 protein-folding contest — and its previous version's performance at the last CASP.



<https://media.nature.com/original/magazine-assets/d41586-020-03348-4/d41586-020-03348-4.pdf>

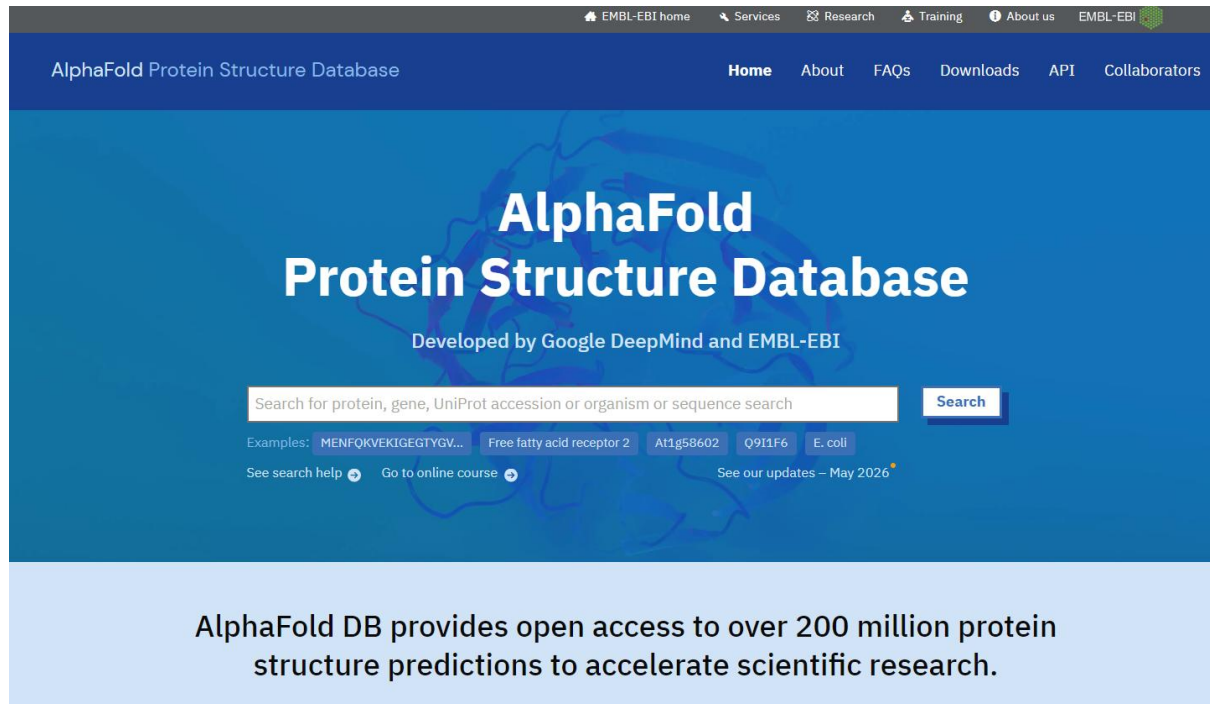
AlphaFoldは「速い蓄積」を武器に変えた

出典：Jumper, J. et al. "Highly accurate protein structure prediction with AlphaFold," Nature 596, 583–589 (2021), <https://doi.org/10.1038/s41586-021-03819-2>. Licensed under CC BY 4.0.



- 基本の教師データ：実験で決定された立体構造＝遅い蓄積
- 予測精度の決め手となる進化情報：巨大な配列・メタゲノムDB＝速い蓄積
- 異種の「知の集積」が合流して生まれた

AlphaFold Structure Databaseの衝撃

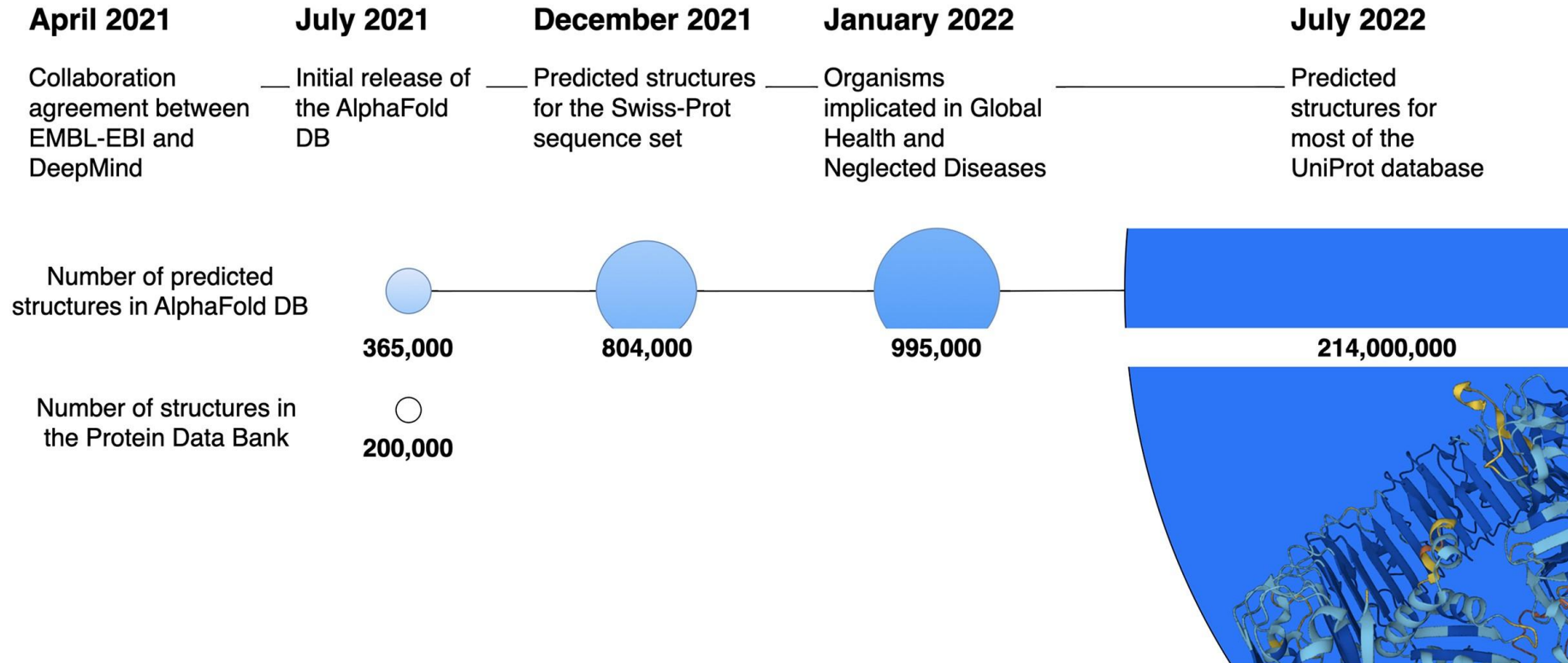


出典 : AlphaFold Protein Structure Database, EMBL-EBI / Google DeepMind,
<https://alphafold.ebi.ac.uk/> (参照 2026-07-08) .
AlphaFold Data are licensed under CC BY 4.0.

- 計算済み予測構造データベース
 - AlphaFold2を既知のタンパク質に実行して予測構造を全公開
 - 2021年7月に約36万件
- 実験構造との桁違いの差
 - 数十年かけたPDBは約26万件
 - しかもAFDBの構造はタンパク質の全長をカバーする
- 既存の「知の基盤」への統合
 - PDB・UniProt・Ensembl等に組み込み済み

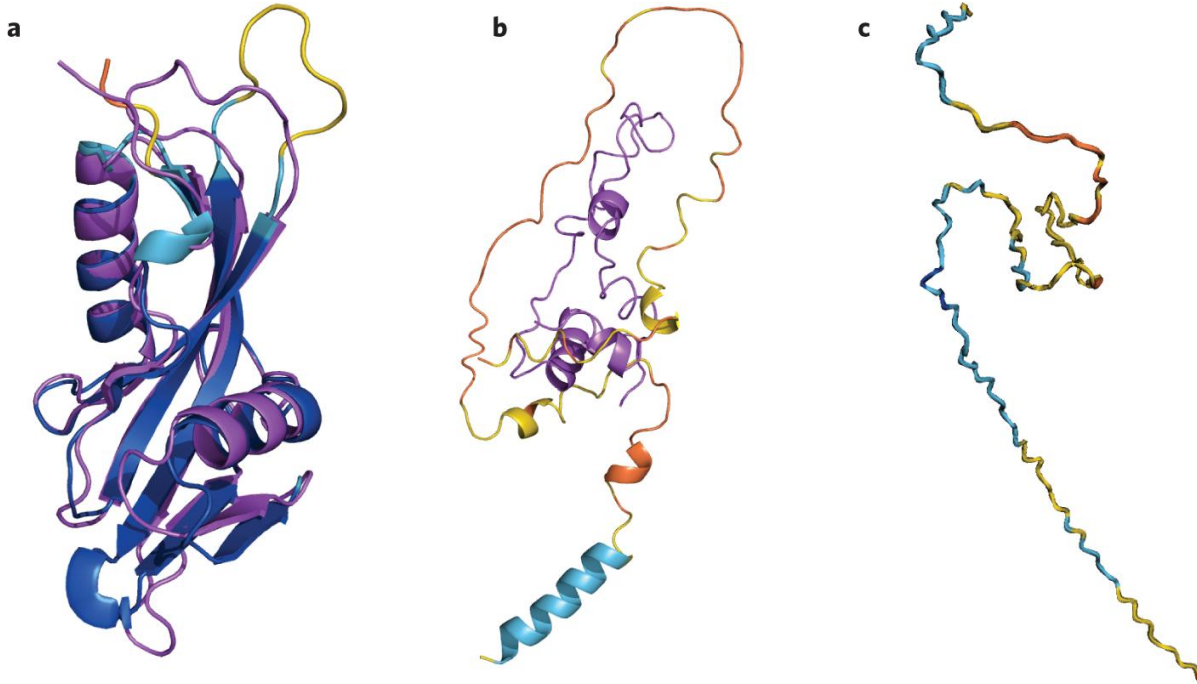
「差が埋まる？」：AlphaFold DBの拡大

出典：Varadi, M. et al. "AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences," Nucleic Acids Research 52(D1), D368–D375 (2024), <https://doi.org/10.1093/nar/gkad1011>. Licensed under CC BY 4.0.



- 2022年7月：UniProtデータベースのほぼすべての配列（2億件）に対する予測構造が公開される＝AIにより遅い蓄積が速い蓄積に件数で追いつく
- Demis Hassabis曰く「2億の構造データ＝博士課程10億年分」？
 - 10億年は言い過ぎだが、世界中から得られる100年以上分のデータなのは間違いない

AlphaFoldは（当然）万能ではない



“The good, the bad and the ugly”

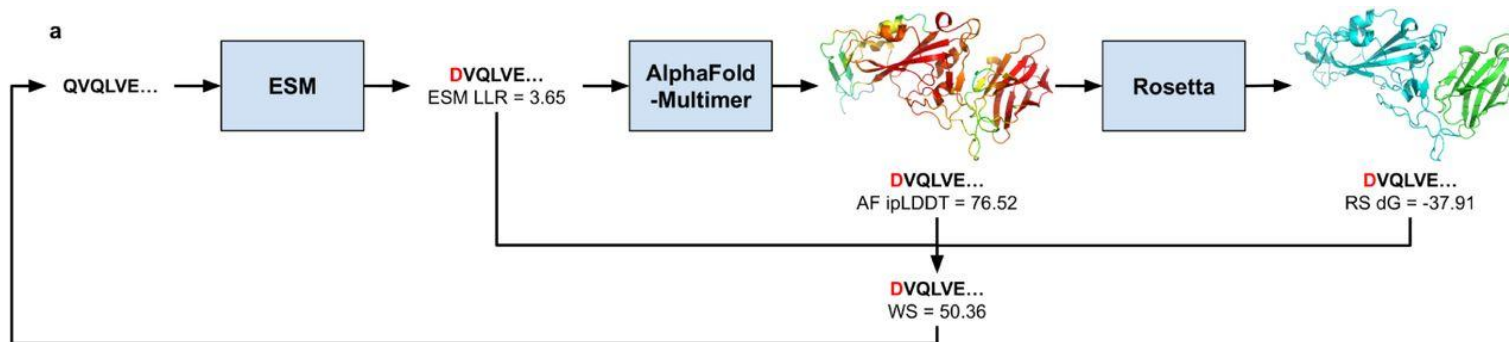
出典：Thornton, J. M., Laskowski, R. A. & Borkakoti, N. “AlphaFold heralds a data-driven revolution in biology and medicine,” Nature Medicine 27, 1666–1669 (2021), <https://doi.org/10.1038/s41591-021-01533-0>.

- 当然ながらAIは完全ではない
- AlphaFoldの予測を示した左図で、a, bは紫が正解、cは正解がない
 - cはこれほど紐状ではないはず
 - 水色・黄色・オレンジはAI (AlphaFold) の自信度
 - 水色はそこそこ自信があるはずだが…
- 予測が当たっているかどうか確認するのは難しい
 - 2億を全部実験するのは無理

衝撃から4年後：道具としてのAIから道具を使うAIへ

17

- The Virtual Lab (2024年11月BioRxivプレプリント) : AIによる研究プロセスの半自動化
- 人間が研究代表者AIにテーマを与える
(例：コロナウィルスに結合する分子をデザイン)
- 研究代表者AIがチームメンバーAIを設定
- 計算機実験のフロー、利用するソフトウェアの選定、利用などは全てAIが決定・実行（人間の助けあり）
 - AlphaFold, ESMなど主要なAIモデルを活用
- 人間が実験して最終成果物を確認



nature

Explore content ▾ About the journal ▾ Publish with us ▾

nature > articles > article

Article | Published: 29 July 2025

The Virtual Lab of AI agents designs new SARS-CoV-2 nanobodies

Kyle Swanson, Wesley Wu, Nash L. Bulaong, John E. Pak & James Zou

Nature 646, 716–723 (2025) | Cite this article

Save article

76k Accesses | 163 Citations | 604 Altmetric | Metrics

出典：Swanson, K. et al. "The virtual lab: AI agents design new SARS-CoV-2 nanobodies with experimental validation," Nature 646, 716–723 (2025), <https://doi.org/10.1038/s41586-025-09442-9>. Preprint: bioRxiv, <https://doi.org/10.1101/2024.11.11.623004>.

(a) Phase 1: Team selection



Human
Researcher



Principal
Investigator



Immunologist



Machine
Learning
Specialist



Computational
Biologist



Scientific
Critic

(b) Phase 2: Project specification

Team Meeting



Summary

Modify nanobodies Ty1, H11-D4, Nb21, and VHH-72 to improve binding to KP.3.

(c) Phase 3: Tools selection

Team Meeting



Summary

ESM, AlphaFold-Multimer, and Rosetta to design improved nanobodies.

2. AI scientistの出現と民主化へ

チャットボットからエージェントへ

チャットボット 2022～

- ・ 質問に答える・文章を書く
- ・ 1問1答で完結
＝続きを考えるのは人間
- ・ 道具なし・記憶は会話の中だけ

例：ChatGPT



エージェント 2023～

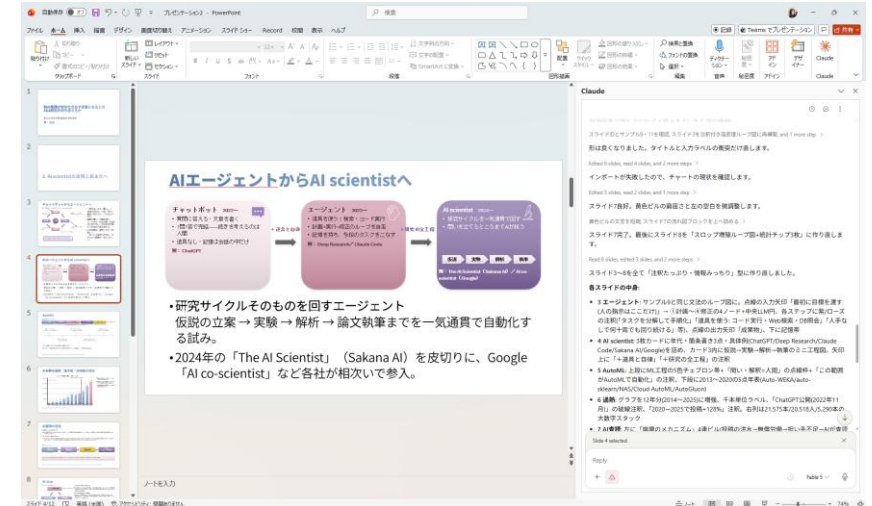
- ・ 道具を使う：検索・コード実行
- ・ 計画→実行→修正のループを自走
- ・ 記憶を持ち、多段のタスクをこなす

例：Deep Research/
Claude Code



+道具と自律

- ・ 「答える」から「動く」へ
- ・ 従来のチャットボットはインタラクティブに問いに答えるだけ
- ・ エージェントは目標に対し計画→実行→結果の観察→修正のループを自分で回す
- ・ 道具を使い、記憶を持つ
 - ・ コード実行・Web検索・データベース照会などの外部ツール
 - ・ 途中経過を覚えながら多段の作業を進める
- ・ 自律の度合いの向上



スライドIDとサンプル9・11を確認、スライド3を注釈付き高密度ループ図に再構築、and 1 more step >

形は良くなりました。タイトルと入カラベルの衝突だけ直します。

Edited 6 slides, read 4 slides, and 2 more steps >

インポートが失敗したので、チャートの現状を確認します。

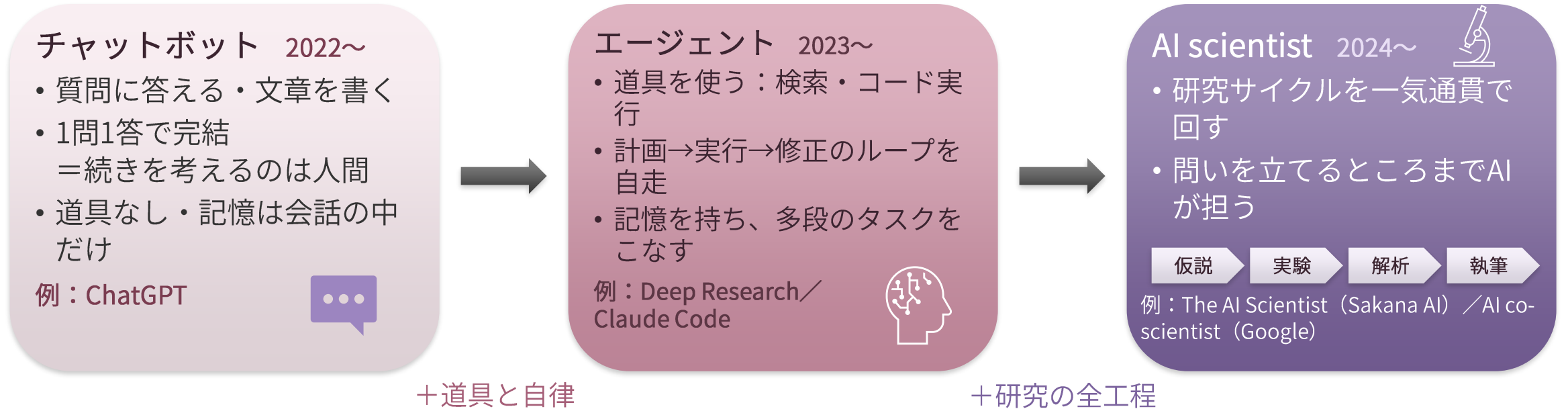
Edited 3 slides, read 2 slides, and 1 more step >

スライド7良好。黄色ピルの窮屈さと左の空白を微調整します。

黄色ピルの文言を短縮、スライド7の流れ図ブロックを上へ詰める >

AIエージェントの実例：
Claude for PowerPoint
(Claude Fable 5)

AIエージェントからAI scientistへ



- AI scientist＝研究サイクルを回すAIエージェント
 - 一般に仮説の立案 → 実験 → 解析 → 論文執筆までを一気通貫で自動化する
- 2024年の「The AI Scientist」(Sakana AI) を始めとして、Google「AI co-scientist」など各社が相次いで参入
 - 商用サービス：Kosmos (Edison Scientific, 2025/11), Claude Science (Anthropic, 2026/6)

AI scientistの前身：AutoML

機械学習の工程（パイプライン）



← この範囲の“探索と最適化”が AutoML で自動化 →

2013 Auto-WEKA
モデル選択とハイパー
パラメータを同時探索

2017 NAS (Zoph & Le)
ネット構造の設計まで自動化

2020 AutoGluon
数行のコードで実用精度

2015 auto-sklearn
メタ学習で探索を高速化

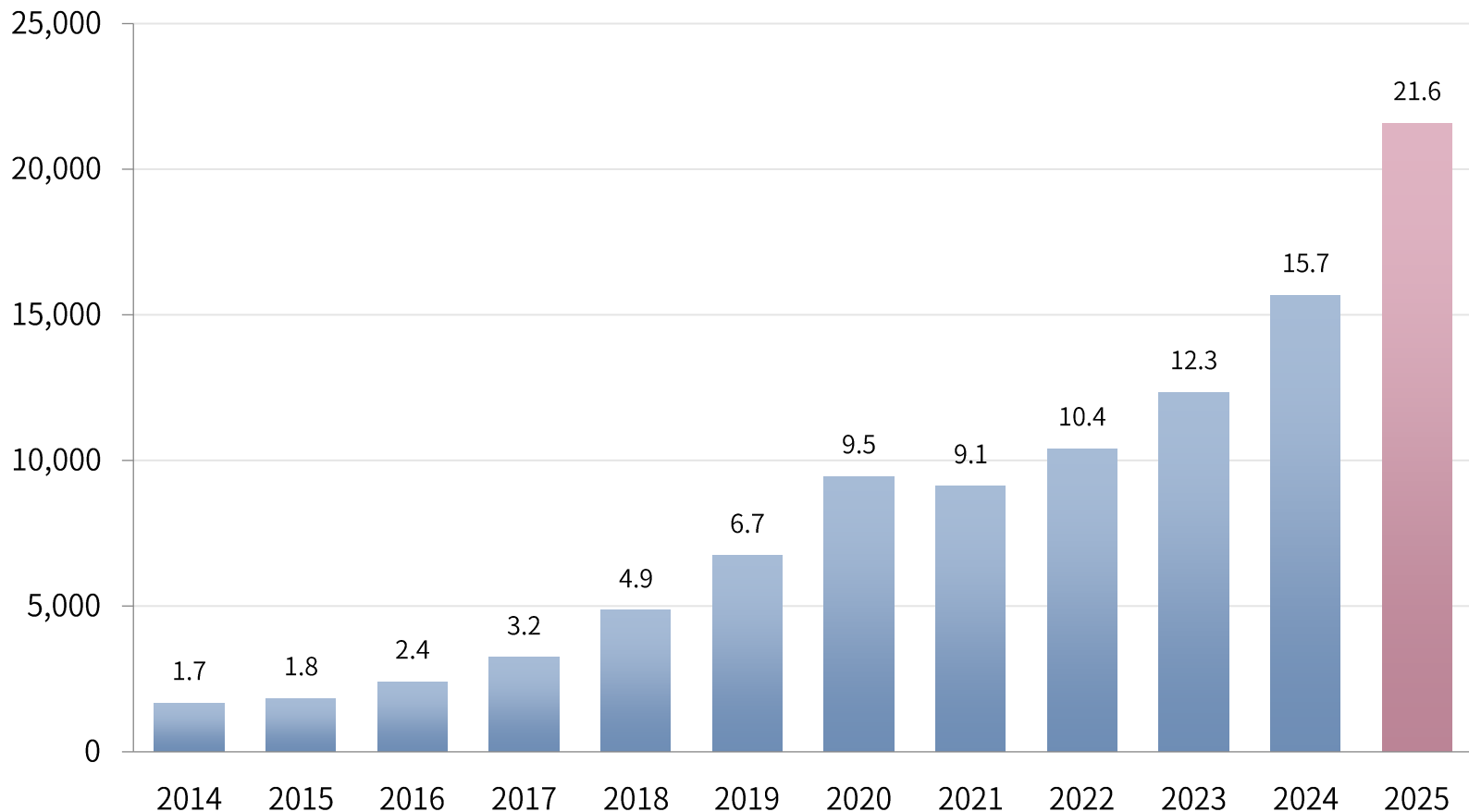
2018 Cloud AutoML
商用化＝専門家なしで使える

- AutoML: 機械学習を“機械化”する試み：
ハイパーパラメータ調整やモデル選択といった職人技を自動化
 - Auto-WEKA (2013) やAuto-sklearn (2015) が先駆け
- ニューラルアーキテクチャ探索 (NAS) へ：
2017年頃には、ネットワーク構造自体をAIが設計するNASも登場
- ただし自動化されたのは最適化のループだけ

出典：Baratchi, M., Wang, C., Limmer, S. et al. “Automated machine learning: past, present and future,” Artificial Intelligence Review 57, 122 (2024), <https://doi.org/10.1007/s10462-024-10726-1>.

AI分野の過熱：論文数・投稿数の増加

NeurIPS 投稿数の推移（単位：千本）



- 情報系国際会議：NeurIPSの例
- 投稿数：2025年で21,575本
 - 10年で12倍に
- 査読者：20,518人動員
 - 採択は5,290本（24.5%）
- 「人手の審査はもう限界では」

出典：NeurIPS Program Committee Chairs, “Reflections on the 2025 Review Process from the Program Committee Chairs,” NeurIPS Blog, 30 September 2025, <https://blog.neurips.cc/2025/09/30/reflections-on-the-2025-review-process-from-the-program-committee-chairs/>；過年度値はNeurIPS公式統計/OpenReview等に基づき著者作成

AI Slopの増加・AI査読の登場

- 論文側：AI生成原稿と偽引用が混入
 - NeurIPS 2025：採択論文から偽引用100件、約50本で確認
 - ICLR 2026：AI生成率の高い投稿も確認
- 査読側：人手レビューはすでに限界
 - Web of Science収録の研究論文は2024年に253万本規模
 - 査読労働は2020年だけで1億時間超
- 対応側：AIを「禁止」だけでなく、責任ある形で組み込む段階へ
 - ICLR 2026：査読の21%・15,899件が完全AI生成と推定
 - AAAI-26や日本生物物理学会／LASAでもAI補助査読・編集を試行／議論



出典：GPTZero, "GPTZero finds 100 new hallucinations in NeurIPS 2025 accepted papers," 21 January 2026, <https://gptzero.me/news/neurips/> ; Ansari, S. "Compound Deception in Elite Peer Review: How AI-Generated Papers with Fake Citations Entered NeurIPS 2025," arXiv:2602.05930 (2026).

出典：ICLR Program Chairs, "A Retrospective on the ICLR 2026 Review Process," ICLR Blog, 31 March 2026, <https://blog.iclr.cc/2026/03/31/a-retrospective-on-the-iclr-2026-review-process>

出典：Pangram Labs, "Pangram Predicts 21% of ICLR Reviews are AI-Generated" <https://www.pangram.com/blog/pangram-predicts-21-of-iclr-reviews-are-ai-generated>

出典：Aczel, B. et al. "A billion-dollar donation: estimating the cost of researchers' time spent on peer review," Research Integrity and Peer Review 6, 14 (2021), <https://doi.org/10.1186/s41073-021-00118-2>.

出典：Liang, W. et al. "Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews," Proceedings of ICML 2024, PMLR 235:29575–29620, <https://proceedings.mlr.press/v235/liang24b.html>.

出典：AAAI, "AAAI Launches AI-Powered Peer Review Assessment System," 16 May 2025, <https://aaai.org/aaai-launches-ai-powered-peer-review-assessment-system/> ; Biswas, J. et al. "AI-Assisted Peer Review at Scale: The AAAI-26 AI Review Pilot," arXiv:2604.13940 (2026), <https://arxiv.org/abs/2604.13940>.

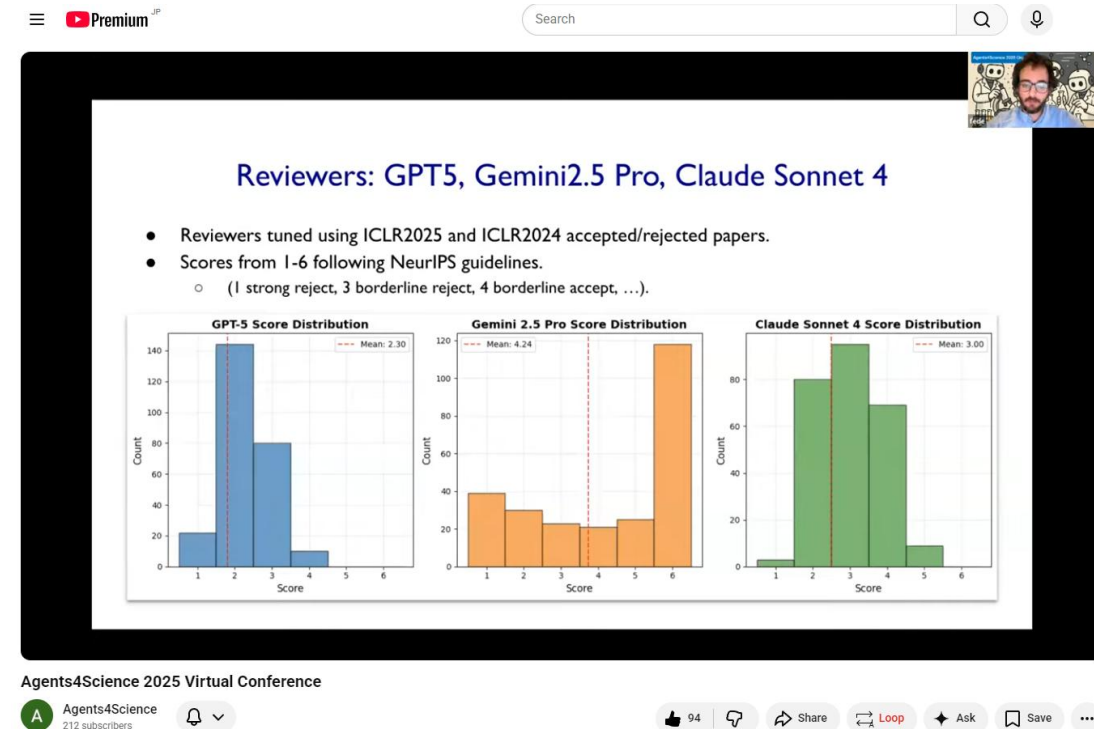
出典：日本生物物理学会「公開シンポジウム案内：論文再設計：生成AI・AGI時代の学術出版の未来」2025年12月24日, <https://www.biophys.jp/news/lib/newsshow.php/7209>

2025年10月：Agents4Scienceの開催



出典：Agents4Science, "Open Conference of AI Agents for Science 2025," Stanford, <https://agents4science.stanford.edu/index.html> (参照 2026-07-08)。

- Agents4Science：AIのAIによる国際会議
 - Virtual Lab著者のJames Zouが主催
 - 9月5日締切→延びて9月15日(AOE)に
 - 10月5日採否発表、10月22日 3時間のオンラインカンファレンス
- AIが「筆頭著者」である必要があり、査読もAIが行う
 - 投稿システム（OpenReview）の都合上、最後の投稿は人間が行う
- 現状のAIの能力や研究での立ち位置をオープンにテスト・議論するための場



動画出典：Agents4Science conference recording, YouTube, <https://www.youtube.com/watch?v=7pXqAeedqOo&t=477s> (参照 2026-07-08)。

Agents4Science全体統計



出典：Bianchi, F., Queen, O., Thakkar, N. et al. "Exploring the use of AI authors and reviewers at Agents4Science," Nature Biotechnology 44, 11–14 (2026), <https://doi.org/10.1038/s41587-025-02963-8>.

- 約300報の投稿、48報が Accept
- 自己申告のAI使用レベル（4段階）：仮説生成、実験デザインと実装、解析、執筆のそれぞれで、AIの寄与（A~D）を選ぶ
 - Aがほぼ人間主導、BがAI50%以下、CがAI50%以上、DがほぼAI主導
- 仮説生成は人間の寄与大・データ解析と執筆はAI寄与大

自動化の素人な自分もコピペで参戦

以下について考えて欲しい

- Agents4Scienceという、AIを使ってやった研究を投稿できるカンファレンスが開催される。ぜひここにこのプロジェクトに関連した小ネタを投稿したい。でも締切は3日後なので、大きな計算は走らせられない。

- Agents4Scienceのサイトはこれ

<https://agents4science.stanford.edu/index.html> FAQとか読んで中身を確認して。ただしAgents4ScienceはAIをある程度自律的に研究に使うという意味合いで、必ずしもコンピュータサイエンスで言うところのエージェントとそれに関連した研究そのものを意味しているわけではないので注意。

- 自分がやっているタンパク質生成と配列探索というコンテキストで、つい最近分子ロボット境界の人たちから部品となるようなリジッドな構造変化を起こすタンパク質を作れないか、と言われたので、それを探索出来ないか、と思っている。これを効率よく探すというテーマはどう?

調べながら AGENTS.md 2 x claudereview.txt short_desc_3.txt

```

1  # 全体の説明
2
3  - このディレクトリはAgents4Scienceという国際会議に投稿するための研究を進めるための、プロジェクトディレクトリである
4  - 1_implmentation, 2_implmentation, 3_implmentationにそれぞれある short_desc_(number).txtにあるプログラム・スクリプトをそれぞれ実装し、結果を得て、考察し、必要な図を作って、最終的な論文を書け。
5
6  ## データ・フォーマットファイルの位置
7
8  - 解析に必要なAlphaFoldの立体構造は /data/afdb にある。alphafold_v4/ と alpha_pae_v4/ にそれぞれ構造とPAEがある。ただし、Humanのタンパク質のみの可能性があるので、必要に応じて追加データをダウンロードするように。追加データはこのディレクトリに置くこと。
9  - 最終的な会議論文のためのテンプレートは /data/project_data/agents4sci/ agents4science_2025.tex にある。コピーを取って、それに書き込むように。
10
11 ## プロジェクトの立案背景
12
13 - 以下は、このプロジェクトの立案プロセスである。必要に応じて思い出そう。
14 - 最初にGPT-5 Proを独立に7回実行し、7つのアイデアを出した。それらのやりとりは original_seven_ideas.txt にまとめられている
15 - これらの7つのアイデアを取りまとめて、3つのGPT-5 Proに最終案を出すように要求した。さらにそれぞれの案に対して別のGPT-5 Proが査読を行い、そのコメントに反論してもらった。これらのアイデアはほぼ同一だが、少し異なるところがある。
16 - それぞれの会話は1_implmentation, 2_implmentation, 3_implmentationフォルダの、long_conversation_(number).txtにまとめられている。
17 - 最終的に得られたコーディングエージェント向け記述がshort_desc_(number).txtである。
18
19

```

スクリプト成果物・
結果サマリー

実装の（改善）提案・
次のステップの明示

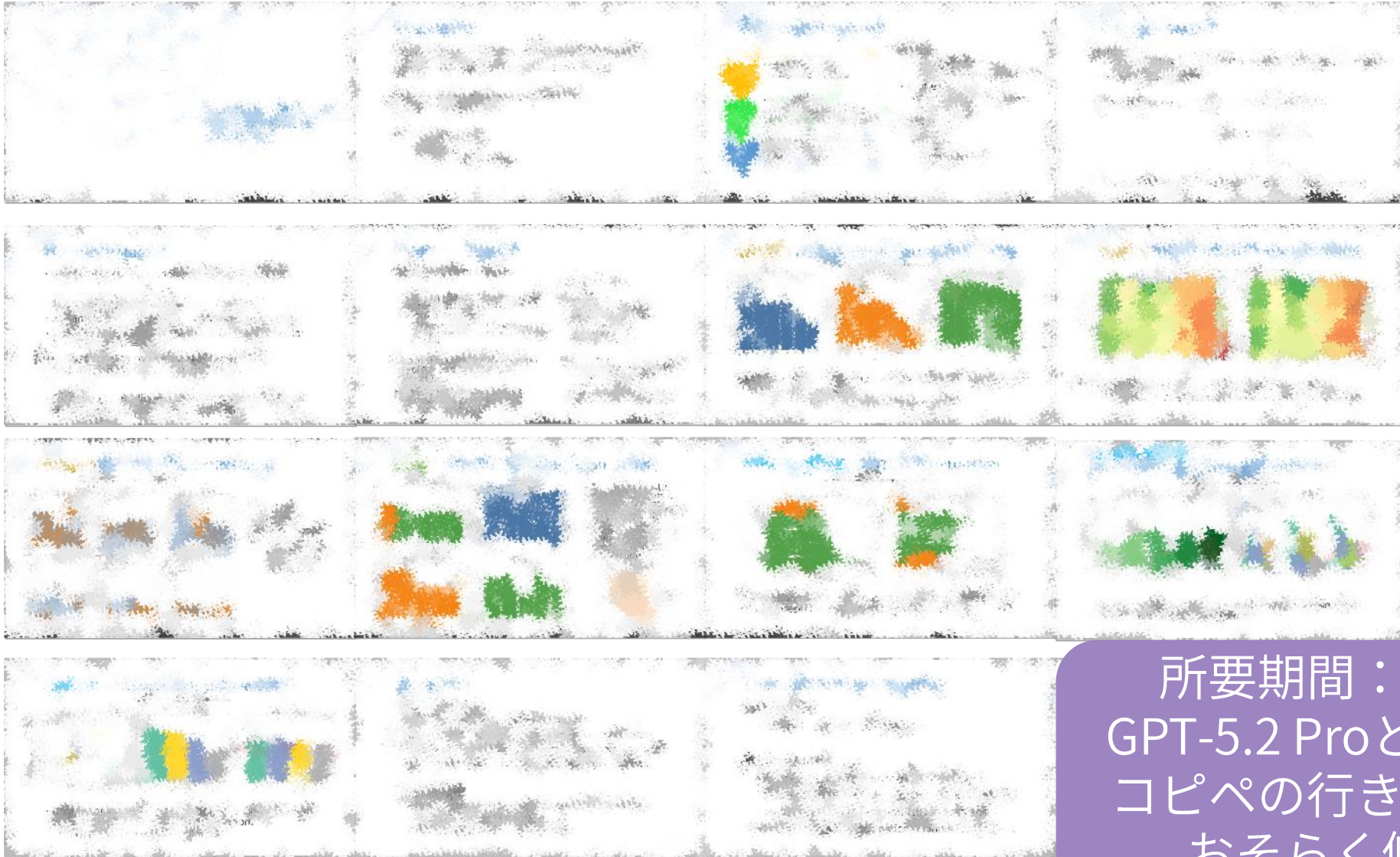
最初のプロンプト

アイデア出し・
論文執筆

完成した論文

- ChatGPT 5 Proに独立に7案を出させ、さらにその中から選ばせる
→コーディングAIに仕様を書いて渡す
- 最後はGoogle AI-coscientist等の研究エージェントと似た作りに収束（研究プロセス上の機能でエージェント分割）
→今回は矢印部分は人間ハーネスが頑張っている
- AIの記憶の管理なども自ずと人間の仕事になっている
→プロンプトやタイミングに人間の創意工夫が入り込む（実質的には全体をオーガナイズしているのと同じ）

2025年12月：AIの急成長と出力の質の向上



所要期間：10日くらい
GPT-5.2 ProとCodexの間の
コピペの行き来：30回程度
おそらく修論レベル

2026年1月：AIによる自動化もAIに願える…？

🔗 > ChatGPT 5.2 Pro ▾

↑ 共有する ...

知ってるかもしれないけど私はこれまでChatGPT Proとコーディングエージェントの間をコピペで繋ぐことで研究を進める方法を試してきた。ただ、いつまでもコピペでやるわけにもいかない。それに、一つのプロジェクトだけでなく、次々とプロジェクトに取り組むようなエージェントシステムが出来て欲しいと思っている、まるで学生が最初に卒論から初めて、修論D論とステップアップするように。

なので、独立したエージェントシステムとしてChatGPT ProとCodexの間を行き来しつつ、適宜別のGPT ProやClaudeなんかのレビューを受けつつプロジェクトを論文化まで完走する、そういう仕組みを作りたい。このためにはAPIの利用の仕方の理解やProとCodexのやりとりの管理、どのタイミングで別のProやClaudeを読んでみてもらうか、などの追加の仕組みが必要になる。最終的な実装はCodexにまかせるつもりだけでも、十分に調べて必要な情報をあつめ、パッケージ化してもらえだろうか？

2026年1月：AIによる自動化もAIに願える…？

🔗 > ChatGPT 5.2 Pro ▾

↑ 共有する ...

知ってるかもしれないけど私はこれまでChatGPT Proとコ

Thought for 23m 47s ▸

やりたいこと、かなり筋が良いです。いまの「ChatGPT Pro ⇄ コーディングエージェントをコピペで接続」は、動くけどスケールしない（そして手首が先に査読で落ちる）ので、“会話”ではなく“成果物と台帳”を中核に置く方式に切り替えるのが正攻法です。

結論から言うと、実装をCodexに任せる前提なら、あなたが欲しい仕組みは MCP + 台帳 (Project/Task/Artifact/Review) + Codex実行アダプタ の3点セットで、かなり現実的に作れます。特にMCPは ChatGPT側(Connectors/Developer mode) と Codex側(MCP) の両方から同じ「研究オーケストレータ」に触らせられるので、コピペ排除の背骨になります。

Codexに渡せる形に「パッケージ化」しました

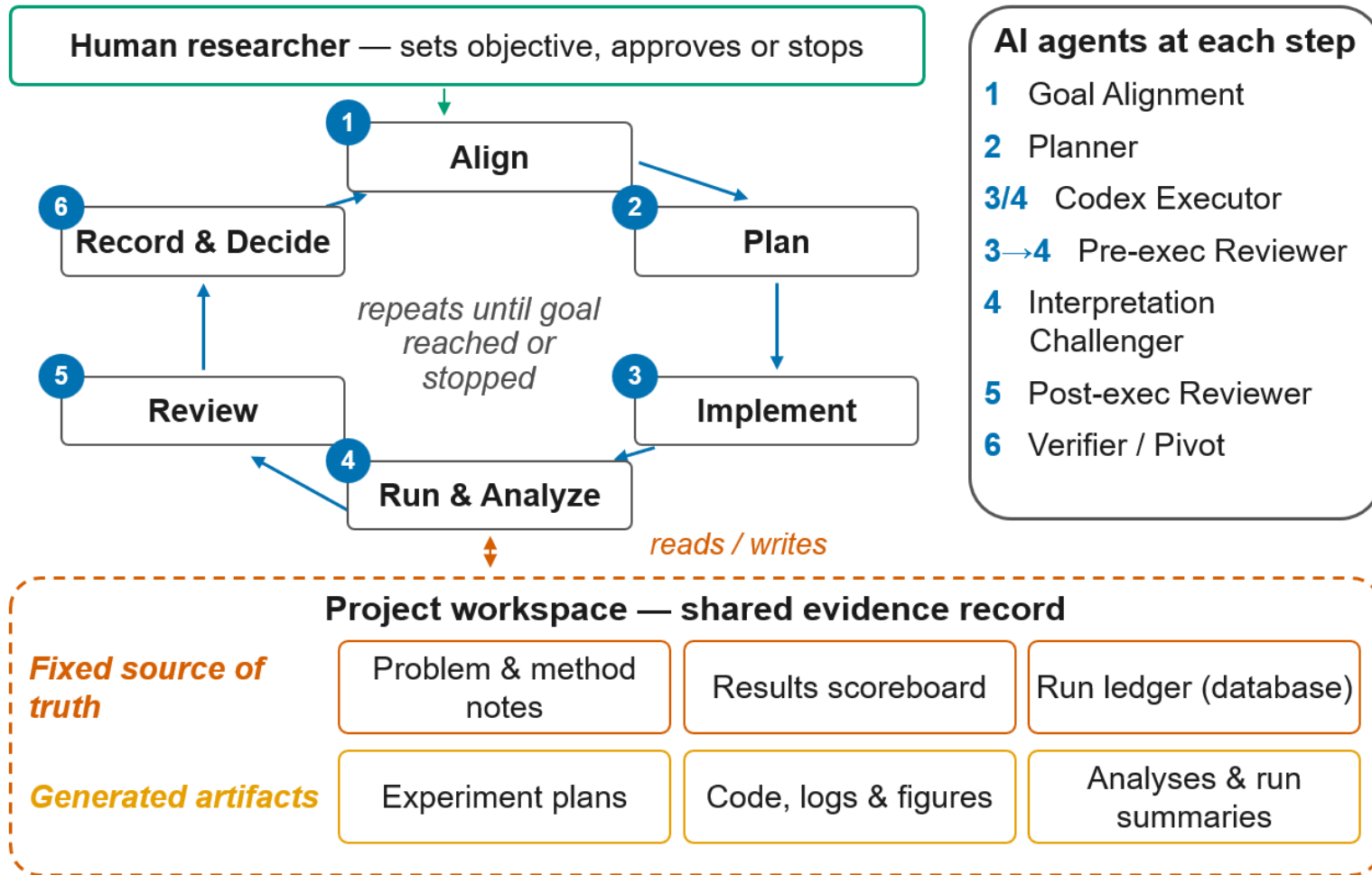
あなたがCodexに“実装を任せる”ためのスターターパック（仕様書・AGENTS.md・出力スキーマ例・rules例）を zipで用意しました。

ダウンロード: [research-orchestrator-pack.zip](#)

初期版までの所要期間：3日
GPT-5.2 ProとCodexの間の
コピペの行き来：10回程度

自動研究器 (AI scientist) by ChatGPT 5.2 Pro

B Automated research cycle and evidence record

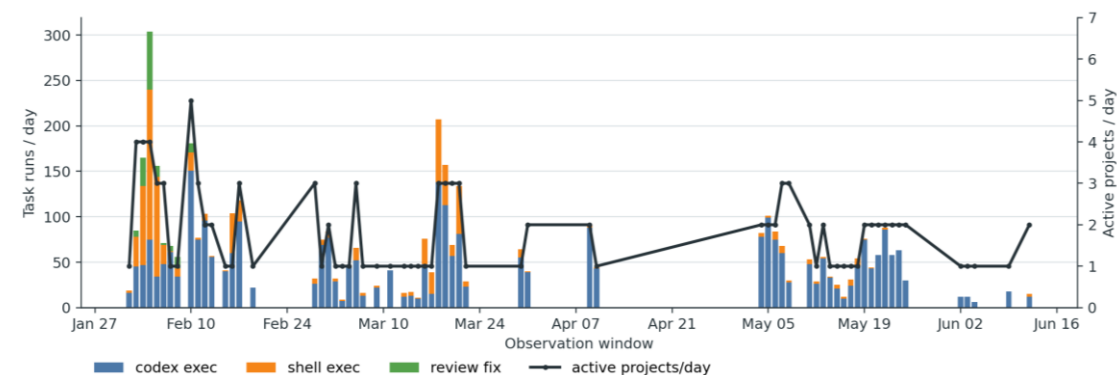


- Planner / Executor / Reviewer / Challenger / Verifier
- Orchestration AI agent は置かず、単一の Python スクリプトがサイクルを制御
- ClaudeCode / Codex の対話セッションが orchestration role を担う場合もある
- Ledger- and artifact-based
：エージェントからの全出力を登録・記録
- 主な評価源：primary metric を持つ scoreboard

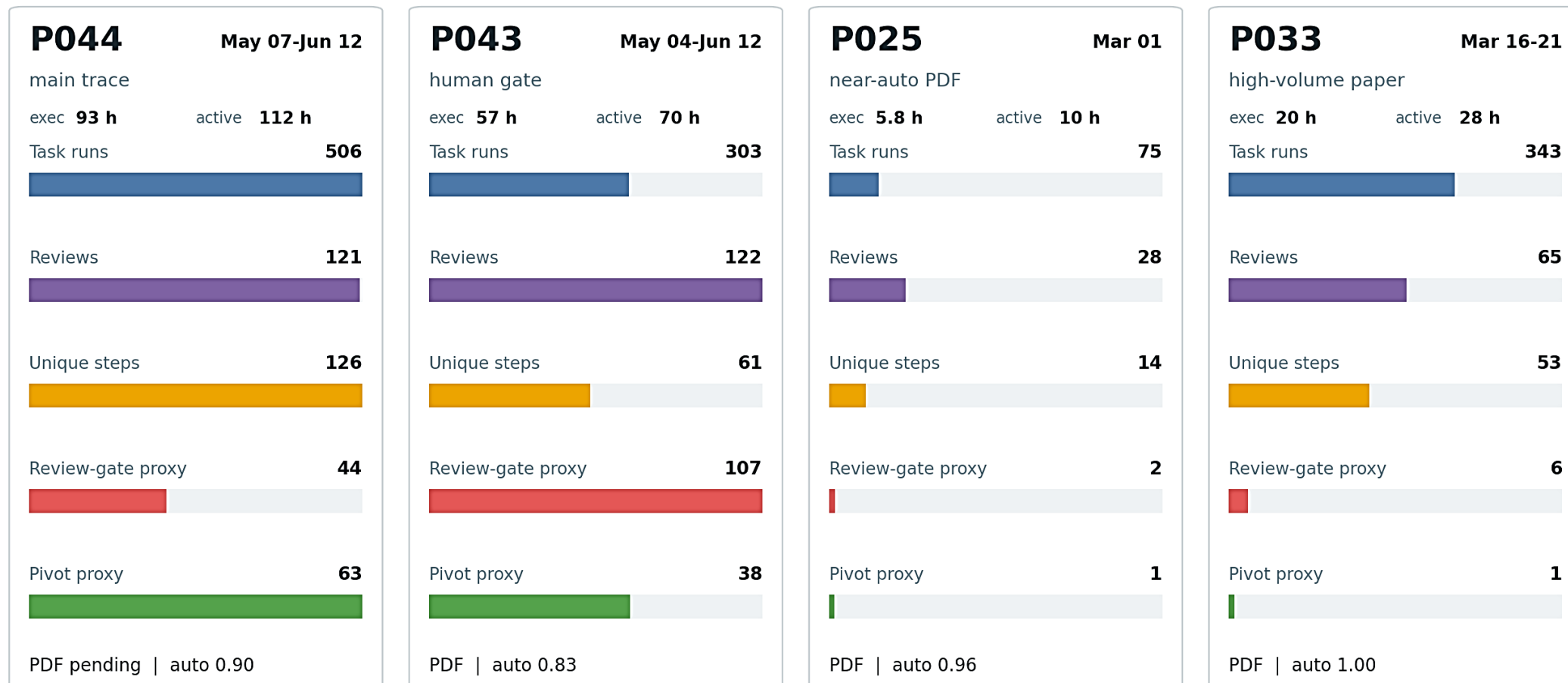
作っ（てくれた）たものを使ってみた

item	Count or value	note
Observation window	2026-02-01 to 2026-06-12	Ledger records included in the poster-safe corpus.
Projects analyzed	45	Demo/test/current paper workspace excluded.
Tasks	4,825	Planned work units created by the Planner
Task runs	4,185	Executor calls actually launched by the system.
Review records	1,364	Structured reviewer outputs stored in the ledger.
Unique agent steps	731	Distinct agent-loop steps; one step can create or run several tasks.
Median task-runs per unique step	5.33	Typical number of executor calls bundled into one recorded step.
Artifact records	22,413	Registered files, reports, logs, figures, and provenance entries.
Review-gate proxy (reviews)	337	Checkpoint-style reviews: human/PI/manual/visual approval, inspection, or stop/go language.
Visual/approval artifact proxies	50	Artifact keyword hits that suggest visual checks or approval outputs.
Pivot proxies	318	Direction-change or reframing language in task specs, artifacts, or project stage; counted independently from review gates.

- 2月初めにプロトタイプが完成
- 2月いっぱいかけて試運転、各種エージェントのプロンプト調整
- 3月～5月にかけて実際の研究に利用
 - 単発のアイデア
 - アイデアの連鎖（Chain）
 - 長い・多い計算を要する課題
 - 人間による試行錯誤を必要とする課題



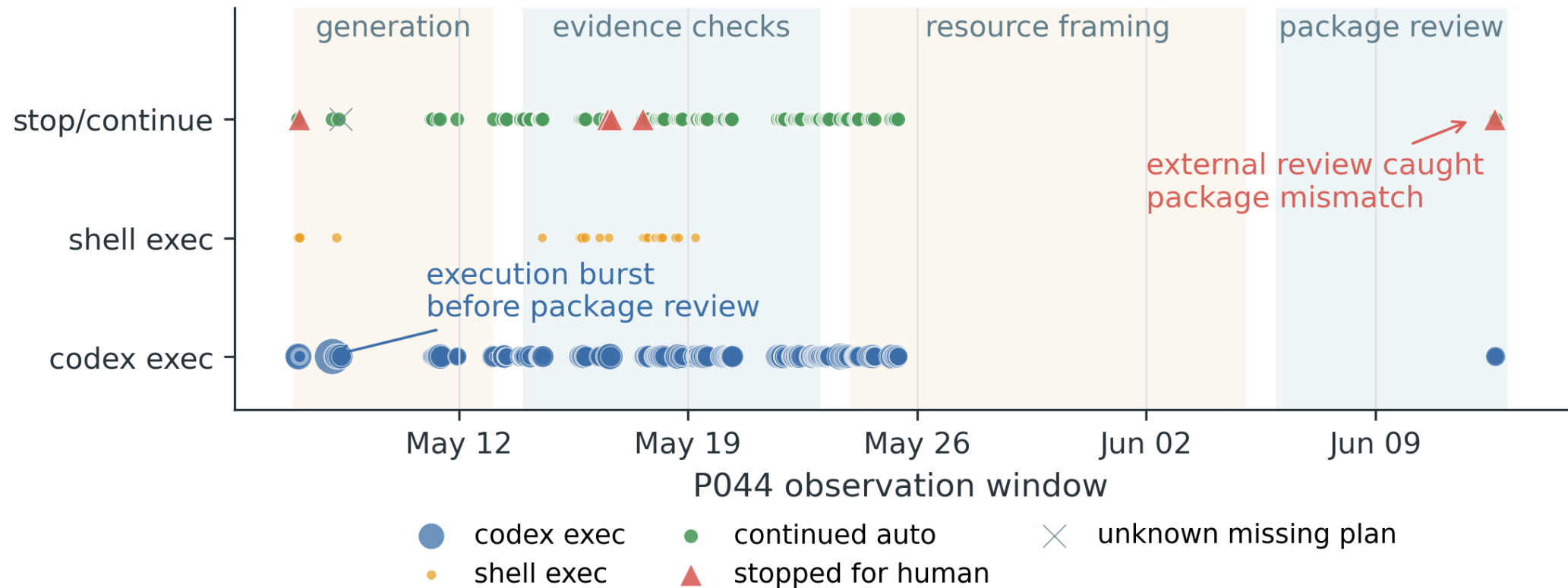
代表的なプロジェクト例



- P044：execution / pivotが多い「実行重視」ケース
- P043：review-gateが多い「判断・承認重視」ケース
- P025・P033は短く終わる対照例（全体が短いと判断はそれほど増えない）

プロジェクトトレース例(1) (P044) : AIエージェントが粛々と手順を実行する

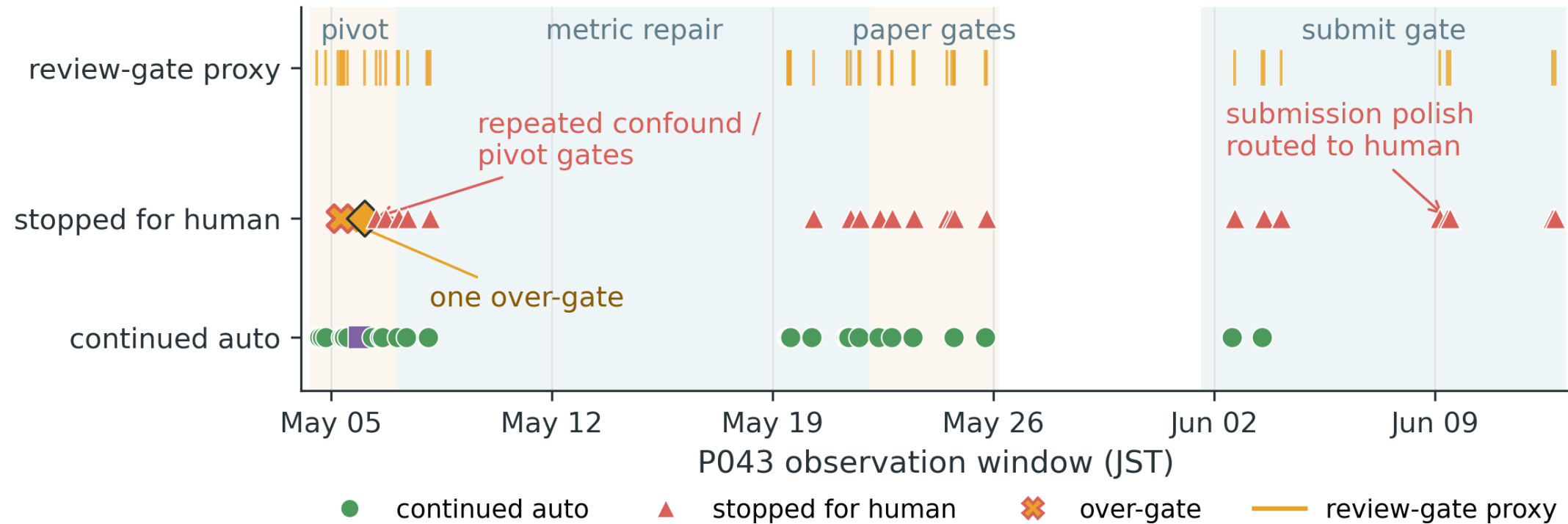
A



- 506 task runs / 126 unique agent steps の”Execution-heavy”なケース
- ほぼ自動で実行が進み、人間による判断はほとんど必要としない
 - 最後の細かい不整合は、プロジェクト全体を監査したChatGPT Proが捕まえた

プロジェクトトレース例(2) (P043) : 人間とAIの間で判断を押し返しあう

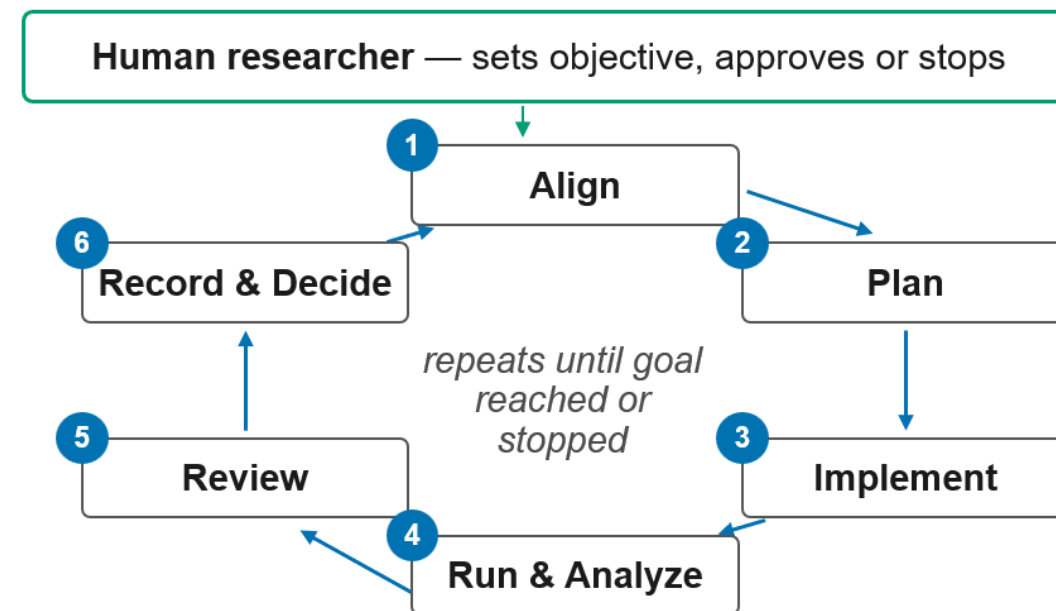
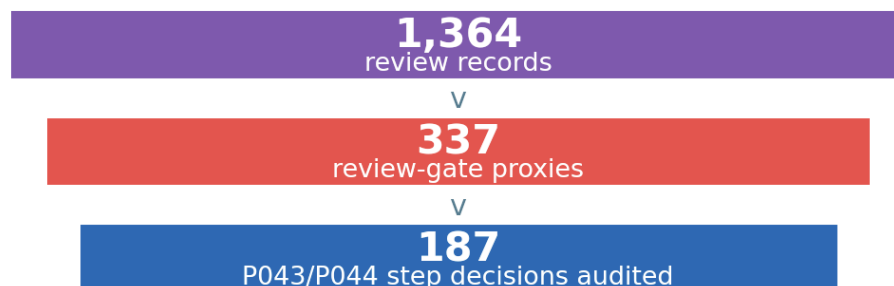
B



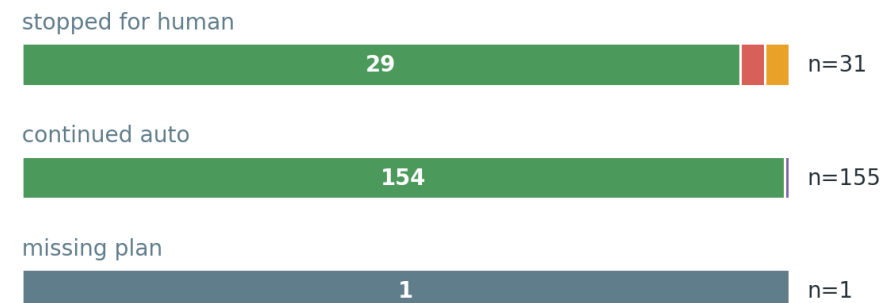
- 61 steps に対して 122 reviews / 107 review-gate proxies
- ピボットや投稿判断で、人間の判断を繰り返し要求
 - ボトルネックは「実行」ではなく「主張・判断・承認の制御」

AIによるstop / continueの判断はほぼ正しい

- 「必要なときに正しく止まれるか」
= 自動なもの安全性を測る一つの指標
- Review records を段階的に絞り込み、187 件の step decision を監査
- stop / continue 判断の多くは妥当
 - 人間著者による判定 (AIのサポートありで)
- Reviewerは層ごとに異なるリスクを検出
 - 最も重要な科学的妥当性は最終的に人間・外部レビューへ



Stop/continue audit result

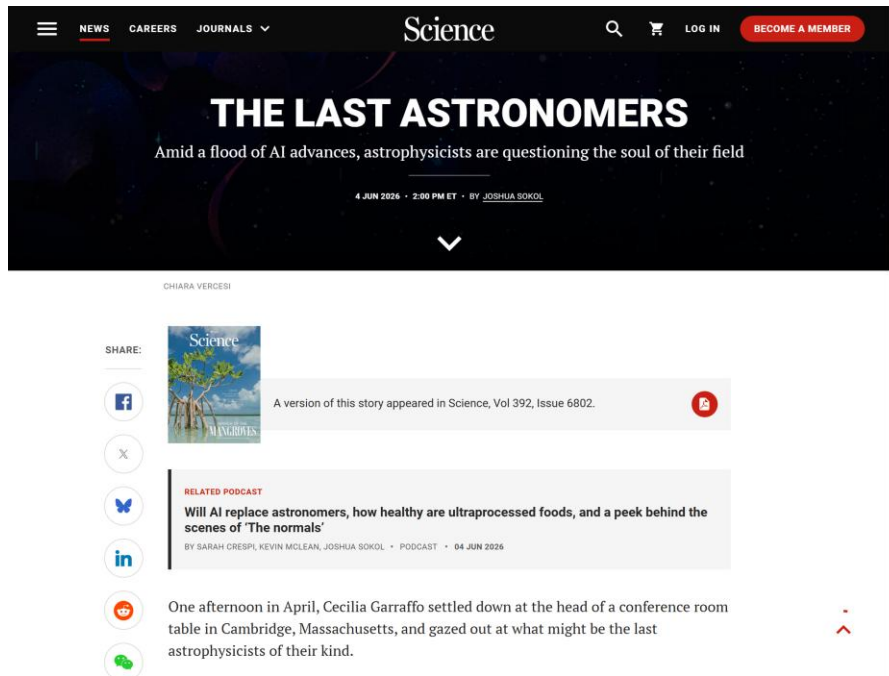


雑多な印象と失敗事例など

- コード実装・データ整理・図表作成等は10倍～100倍速
 - 単体のAIエージェントと異なり、適切に判断して次々ステップをこなしてくれる
- Review layerは大体信用できる＝人間は方針・妥当性・価値判断を担当
- 研究アイデアが続く限り、探索を広げ続けられる
- **AIらしい誤り**：もっともらしいが仮説をテストしていない metric を選ぶ
 - 科学的でない誤りとしては、Verifierを改変しようとするケースも
- **人間もやる誤り**：ID変換、前処理、メタデータの処理ミスによる誤った結果の生成
 - AI review をすり抜けることがある
- これらの誤りも10倍～100倍速で発生する
- 仮に誤りがなかったとしてもProject branch が爆発し、現状の把握が困難になる
- 速く実行できるほど、validation design がすべて
 - 人間の判断力リソースは有限 / ボトルネックは人間

3. おわりに（AIが研究者を追い越すとき？）

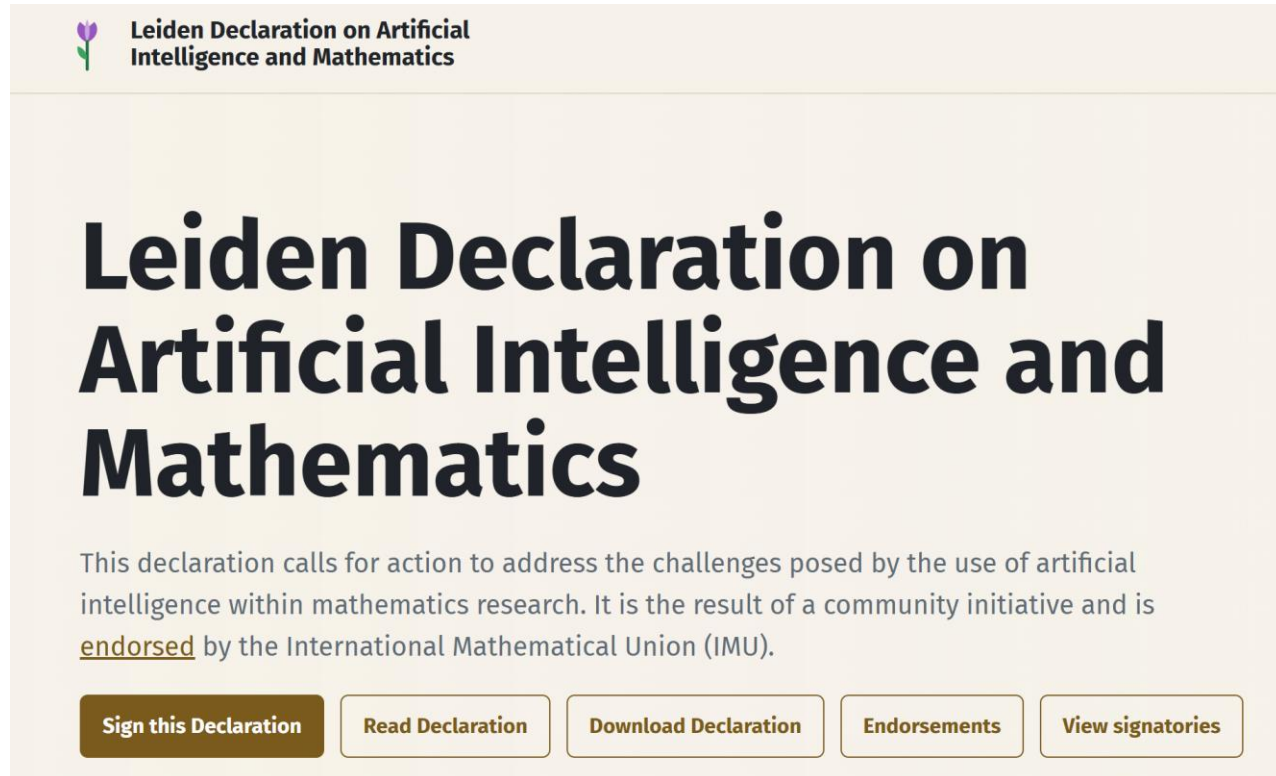
AIと天体物理学 (Science誌、2026年6月4日)



出典：Sokol, J. “The last astronomers: Amid a flood of AI advances, astrophysicists are questioning the soul of their field,” Science, 4 June 2026, <https://www.science.org/content/article/amid-flood-ai-advances-astrophysicists-are-questioning-soul-their-field>.

- “The last astronomers: Amid a flood of AI advances, astrophysicists are questioning the soul of their field”
 - AIの進化が天体物理学にもたらす影響と課題
- AI活用の急拡大：
 - データ解析、コーディング、論文執筆などのプロセスが急速に自動化されている
- 圧倒的な効率化と「査読システム」の危機
- 若手研究者のスキル低下リスク
 - 基礎的な作業をAIが代替→大学院生などが自ら悩み、直感や分析スキルを養う機会が奪われている
- 学問の「存在意義」を巡る議論
 - 「単に宇宙の答えを早く知ること」が目的か、それとも「人間が探求するプロセス自体」に意味があるのか
 - AIの台頭により、天体物理学という分野の「魂（存在意義）」そのものが問われている

数学分野：ライデン宣言（2026年6月2日）



The banner features a light beige background. At the top left is a small purple flower icon next to the text 'Leiden Declaration on Artificial Intelligence and Mathematics'. The main title 'Leiden Declaration on Artificial Intelligence and Mathematics' is in large, bold, black font. Below it, a paragraph of text explains the declaration's purpose. At the bottom, there are five buttons: 'Sign this Declaration' (dark brown), 'Read Declaration' (light brown), 'Download Declaration' (light brown), 'Endorsements' (light brown), and 'View signatories' (light brown).

Leiden Declaration on Artificial Intelligence and Mathematics

Leiden Declaration on Artificial Intelligence and Mathematics

This declaration calls for action to address the challenges posed by the use of artificial intelligence within mathematics research. It is the result of a community initiative and is endorsed by the International Mathematical Union (IMU).

[Sign this Declaration](#) [Read Declaration](#) [Download Declaration](#) [Endorsements](#) [View signatories](#)

出典：Leiden Declaration on Artificial Intelligence and Mathematics, 2 June 2026, DOI: 10.5281/zenodo.20302944, <https://leidendeclaration.ai/>（参照 2026-07-08）。

- AIの急速な発展が数学研究の「特徴的な (characteristic)」価値観を脅かしている現状への警鐘
- 数学を「人間による営み」として維持することを目指す
- 数学を脅かす「5つの主要なリスク」
 - 信頼性の低下、帰属の不透明性、格差と依存の拡大、過剰な宣伝（ハイプ）、研究の自律性の喪失
- 主な行動勧告（ステークホルダー別）
 - 数学者個人: AI利用の透明性を確保し、結果の正しさに人間が責任を持つ、学会・学術誌: AI利用に関する明確な基準を作り、厳格な査読制度を維持する、等

生命情報学の場合？

2026年8月開催予定の日本バイオインフォマティクス学会
年会プログラムより

8月26日（水） 13:30～15:00 第4会場

生成AI時代のバイオインフォマティクス教育を考える

“Rethinking Bioinformatics Education in the Era of Generative AI”

出典：日本バイオインフォマティクス学会, “IIBMP 2026 公募ワークショップ一覧・詳細”, <https://www.jsbi.org/iibmp2026/program/workshops/#ai-education> (参照 2026-07-08) .

「しかしながら、ここ数年の急速なAI技術の進展により、事態は急展開している。なんと近い将来最もAIに職をとってかわられそうな科学者は、簡単なコードを書いたりツールを駆使したデータ解析や可視化を専門とする人達であろうと予想されている。この問題に心を痛めている教育者は全国に多くいらっしゃるはずである。」

NCBI設立の趣旨(1988年)

出典：Kent Smith, “A Brief History of NCBI’s Formation and Growth,” in The NCBI Handbook, 2nd ed., National Center for Biotechnology Information (US), 2013–, <https://www.ncbi.nlm.nih.gov/books/NBK148949/> (参照 2026-07-08) .

The act stipulated the following functions for the new National Center for Biotechnology Information:

1. design, develop, implement, and manage **automated systems for the collection, storage, retrieval, analysis, and dissemination of knowledge** concerning human molecular biology, biochemistry, and genetics;

まとめ

- 「知の集積が知を吐き出す装置になるとき：AIは研究の何を変えるか」
- データの集積がデータを生み出す時代
 - ゲノム・立体構造データベースからAlphaFoldそしてAlphaFold DBが出現
- AIチャットボットからエージェントを経てAI scientistの出現
 - AI slop / AI査読 / Agents4Science
 - コピペからの研究自動化、自分でもやってみた
 - なんでも全自動とまではいかないが...
- AIが研究者を追い越すとき？