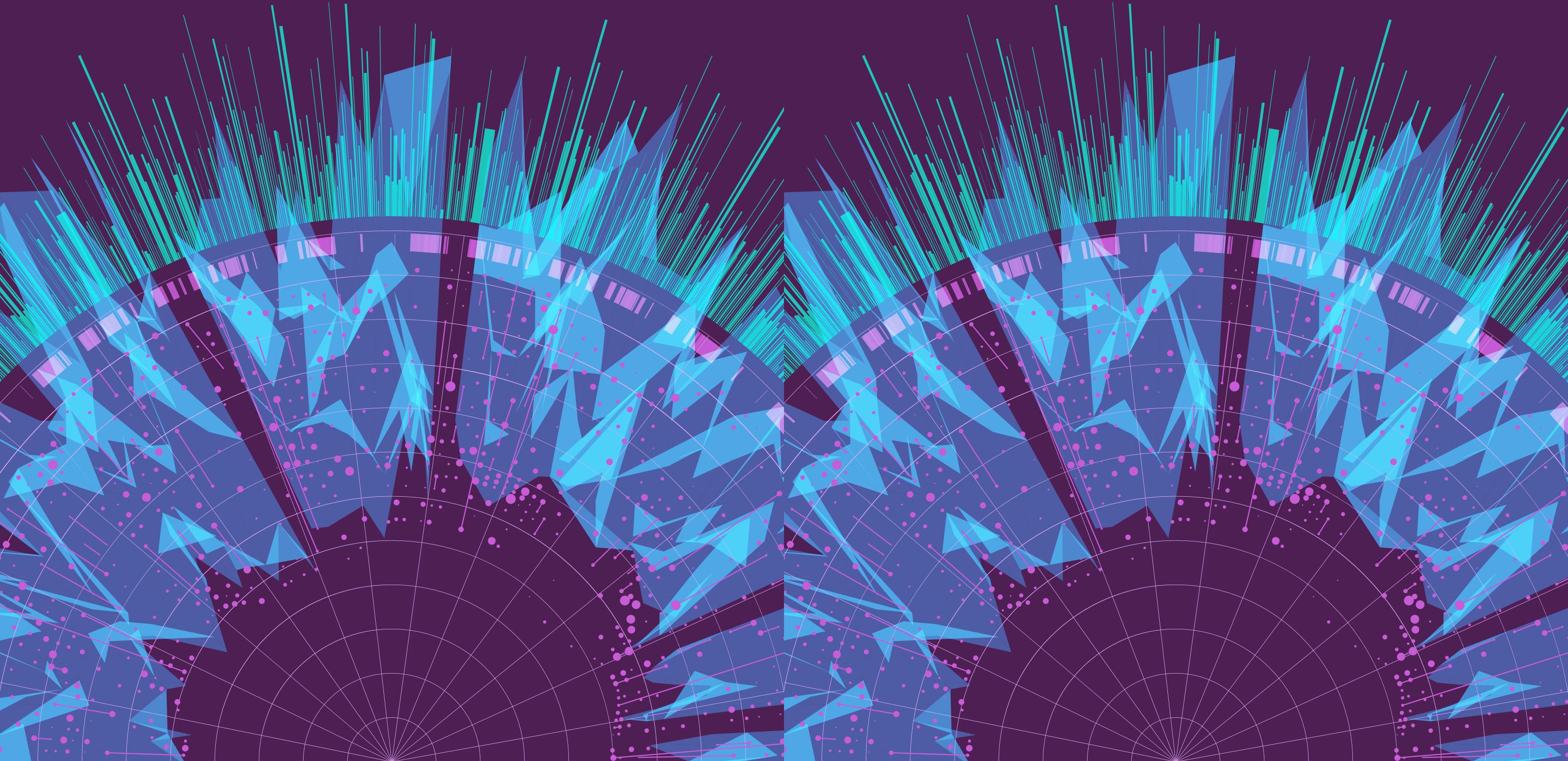


CHAPTER 5: Science and Medicine

本文档由 funstory.ai 的开源 PDF 翻译库 BabelDOC v0.3.17 (<http://yadt.io>) 翻译，本仓库正在积极的建设当中，欢迎 star 和关注。

第 5 章：
Science and
Medicine



Preview

Overview	298
Chapter Highlights	299
5.1 Notable Scientific Milestones	300
AlphaDev	300
FlexiCubes	301
Synbot	303
GraphCast	304
GNoME	305
Flood Forecasting	306
5.2 AI in Medicine	307
Notable Medical Systems	307
SynthSR	307
Coupled Plasmonic Infrared Sensors	309
EVEscape	310
AlphaMissence	312
Human Pangenome Reference	313
Clinical Knowledge	314
MedQA	314
Highlighted Research: GPT-4 Medprompt	315
Highlighted Research: MediTron-70B	317
Diagnosis	318
Highlighted Research: CoDoC	318
Highlighted Research: CT Panda	319
Other Diagnostic Uses	320
FDA-Approved AI-Related Medical Devices	321
Administration and Care	323
Highlighted Research: MedAlign	323

[ACCESS THE PUBLIC DATA](#)

レビュー

概要 298 章のハイライト 299

5.1 Notable Scientific Milestones	300
AlphaDev	300
FlexiCubes	301
Synbot	303
GraphCast	304
GNoME	305
Flood Forecasting	306
5.2 AI in Medicine	307
Notable Medical Systems	307
SynthSR	307
Coupled Plasmonic Infrared Sensors	309
EVEscape	310
AlphaMissence	312
Human Pangenome Reference	313
Clinical Knowledge	314
MedQA	314
Highlighted Research: GPT-4 Medprompt	315
Highlighted Research: MediTron-70B	317
Diagnosis	318
Highlighted Research: CoDoC	318
Highlighted Research: CT Panda	319
Other Diagnostic Uses	320
FDA-Approved AI-Related Medical Devices	321
Administration and Care	323
Highlighted Research: MedAlign	323

[ACCESS THE PUBLIC DATA](#)

Overview

This year's AI Index introduces a new chapter on AI in science and medicine in recognition of AI's growing role in scientific and medical discovery. It explores 2023's standout AI-facilitated scientific achievements, including advanced weather forecasting systems like GraphCast and improved material discovery algorithms like GNoME. The chapter also examines medical AI system performance, important 2023 AI-driven medical innovations like SynthSR and ImmunoSEIRA, and trends in the approval of FDA AI-related medical devices.

概要

今年のAIインデックスは、科学と医療におけるAIの役割の高まりを認識し、AIによる科学的および医療的発見に関する新しい章を紹介します。2023年の際立ったAIを活用した科学的成果、例えればGraphCastのような高度な天気予報システムやGNoMEのような改良された材料発見アルゴリズムを探ります。この章では、医療AIシステムのパフォーマンス、2023年の重要なAI駆動の医療革新であるSynthSRやImmunoSEIRA、FDAのAI関連医療機器の承認に関するトレンドも検討します。

Chapter Highlights

1. Scientific progress accelerates even further, thanks to AI. In 2022, AI began to advance scientific discovery. 2023, however, saw the launch of even more significant science-related AI applications—from AlphaDev, which makes algorithmic sorting more efficient, to GNoME, which facilitates the process of materials discovery.

2. AI helps medicine take significant strides forward. In 2023, several significant medical systems were launched, including EVEscape, which enhances pandemic prediction, and AlphaM issence, which assists in AI-driven mutation classification. AI is increasingly being utilized to propel medical advancements.

3. Highly knowledgeable medical AI has arrived. Over the past few years, AI systems have shown remarkable improvement on the MedQA benchmark, a key test for assessing AI's clinical knowledge. The standout model of 2023, GPT-4 Medprompt, reached an accuracy rate of 90.2%, marking a 22.6 percentage point increase from the highest score in 2022. Since the benchmark's introduction in 2019, AI performance on MedQA has nearly tripled.

4. The FDA approves more and more AI-related medical devices. In 2022, the FDA approved 139 AI-related medical devices, a 12.1% increase from 2021. Since 2012, the number of FDA-approved AI-related medical devices has increased by more than 45-fold. AI is increasingly being used for real-world medical purposes.

章のハイライト

1. 科学の進歩はAIのおかげでさらに加速しています。 2022年には、AIが科学的発見を進め始めました。しかし2023年には、アルファデブのような、アルゴリズムのソートをより効率的にする科学関連のAIアプリケーションがさらに多く登場しました。GNoMEは、材料発見のプロセスを促進します。

2. AIは医学の大きな進歩を助けています。 2023年には、パンデミック予測を強化するEVEscapeや、AI駆動の変異分類を支援するAlphaM issenceなど、いくつかの重要な医療システムが導入されました。AIは、医療の進歩を推進するためにますます利用されています。

3. 高度な知識を持つ医療AIが登場しました。 過去数年間で、AIシステムは、AIの臨床知識を評価するための重要なテストであるMedQAベンチマークで驚異的な改善を示しました。2023年の注目モデルであるGPT-4 Medpromptは、90.2%の精度を達成し、2022年の最高スコアから22.6ポイントの増加を記録しました。2019年にベンチマークが導入されて以来、MedQAにおけるAIのパフォーマンスはほぼ3倍になっています。

4. FDAはますます多くのAI関連医療機器を承認しています。 2022年には、FDAは139のAI関連医療機器を承認し、2021年から12.1%の増加を見せました。2012年以降、FDAに承認されたAI関連医療機器の数は45倍以上に増加しています。AIは、実際の医療目的でますます使用されています。

This section highlights significant AI-related scientific breakthroughs of 2023 as chosen by the AI Index Steering Committee.

5.1 Notable Scientific Milestones

AlphaDev

AlphaDev discovers faster sorting algorithms

AlphaDev is a new AI reinforcement learning system that has improved on decades of work by scientists and engineers in the field of computational algorithmic enhancement. AlphaDev developed algorithms with fewer instructions than existing human benchmarks for

fundamental sorting algorithms on short sequences such as Sort 3, Sort 4, and Sort 5 (Figure 5.1.1). Some of the new algorithms discovered by AlphaDev have been incorporated into the LLVM standard C++ sort library. This marks the first update to this part of the library in over 10 years and is the first addition designed using reinforcement learning.

AlphaDev vs. human benchmarks when optimizing for algorithm length

Source: Mankowitz et al., 2023 | Chart: 2024 AI Index report

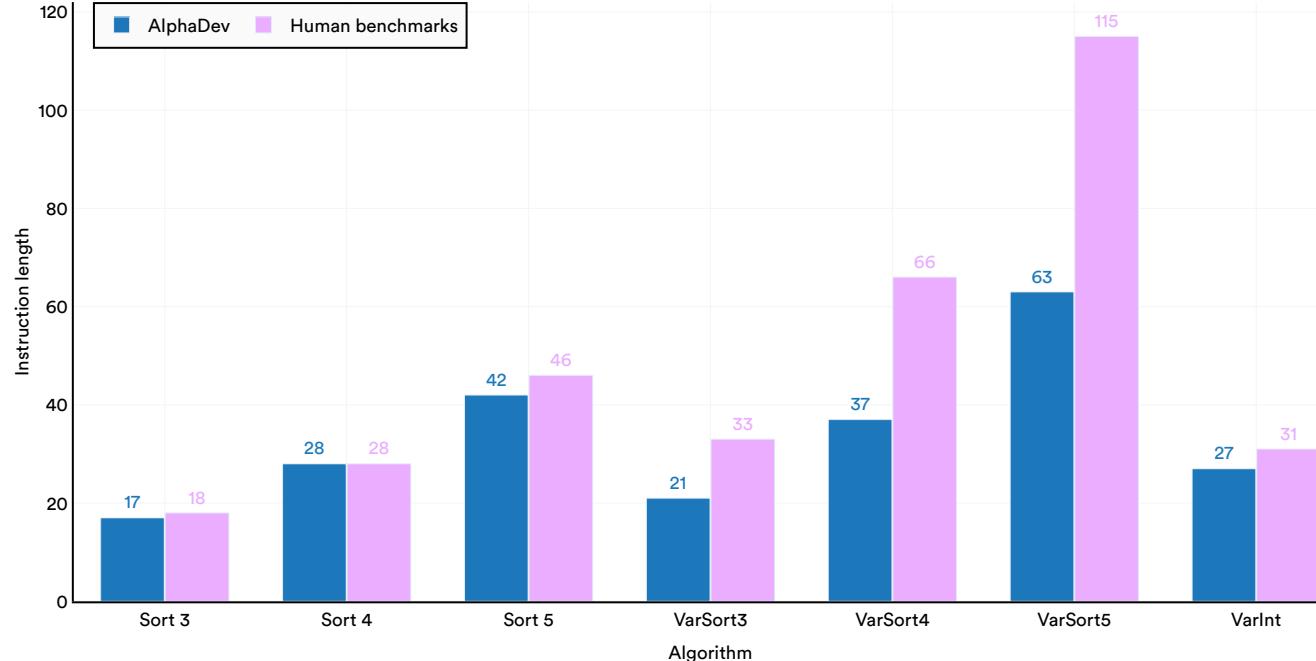


Figure 5.1.1

This section highlights significant AI-related scientific breakthroughs of 2023 as chosen by the AI Index Steering Committee.

5.1 注目すべき科学的マイルストーン

アルファデブ

AlphaDevはより高速なソートアルゴリズムを発見します。

アルファデブは、計算アルゴリズムの強化の分野における科学者やエンジニアの数十年の研究を改善した新しいAI強化学習システムです。アルファデブは、既存の人間のベンチマークよりも少ない命令でアルゴリズムを開発しました。

Sort 3、Sort 4、Sort 5などの短いシーケンスにおける基本的なソートアルゴリズムです（図 5.1.1）。アルファデブによって発見された新しいアルゴリズムのいくつかは、LLVM 標準 C++ sort ライブラリに組み込まれました。これは、ライブラリのこの部分への最初の更新であり、10年以上ぶりのもので、強化学習を使用して設計された最初の追加です。

アルファデブと人間のベンチマークのアルゴリズム長の最適化に関する比較

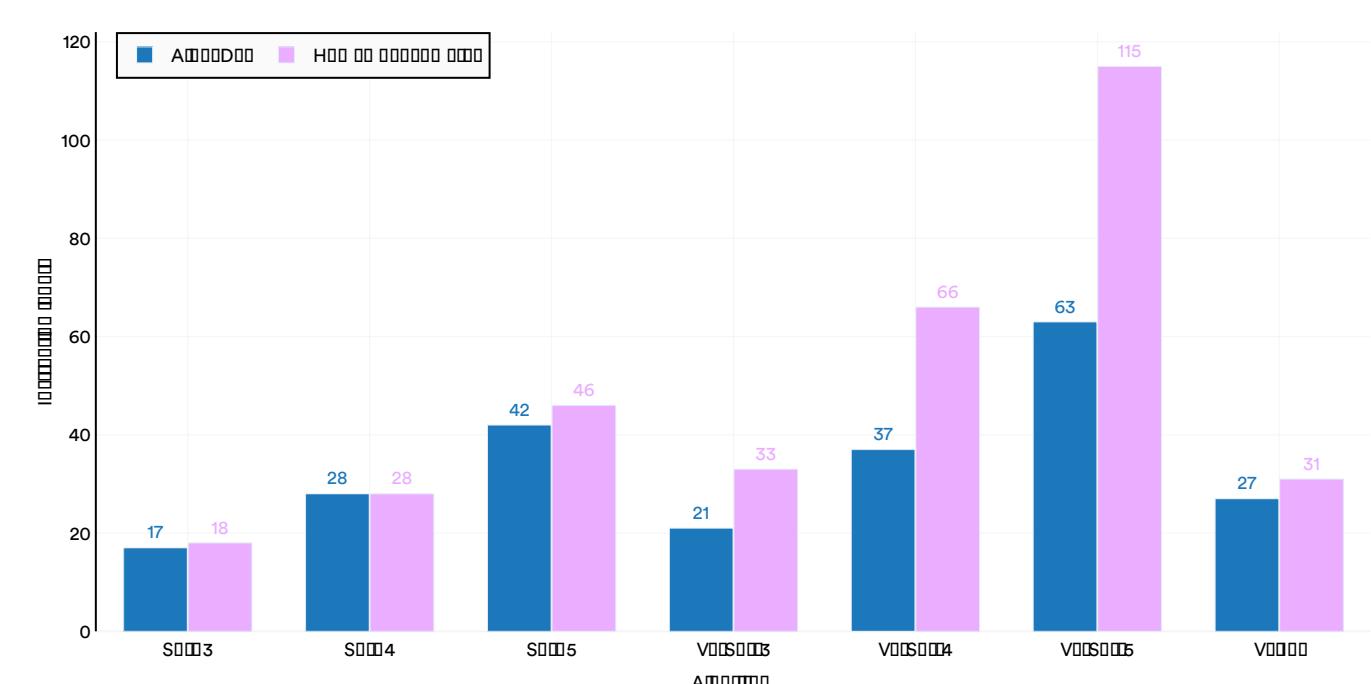


Figure 5.1.1

FlexiCubes

3D mesh optimization with FlexiCubes

3D mesh generation, crucial in computer graphics, involves creating a mesh of vertices, edges, and faces to define 3D objects. It is key to video games, animation, medical imaging, and scientific visualization. Traditional isosurface extraction algorithms often struggle with limited resolution, structural rigidity, and numerical instabilities, which subsequently impacts

quality. FlexiCubes addresses some of these limitations by employing AI for gradient-based optimization and adaptable parameters (Figure 5.1.2). This method allows for precise, localized mesh adjustments. Compared to other leading methods that utilize differentiable isosurfacing for mesh reconstruction, FlexiCubes achieves mesh extractions that align much more closely with the underlying ground truth (Figure 5.1.3).

Sample FlexiCubes surface reconstructions

Source: Nvidia, 2023

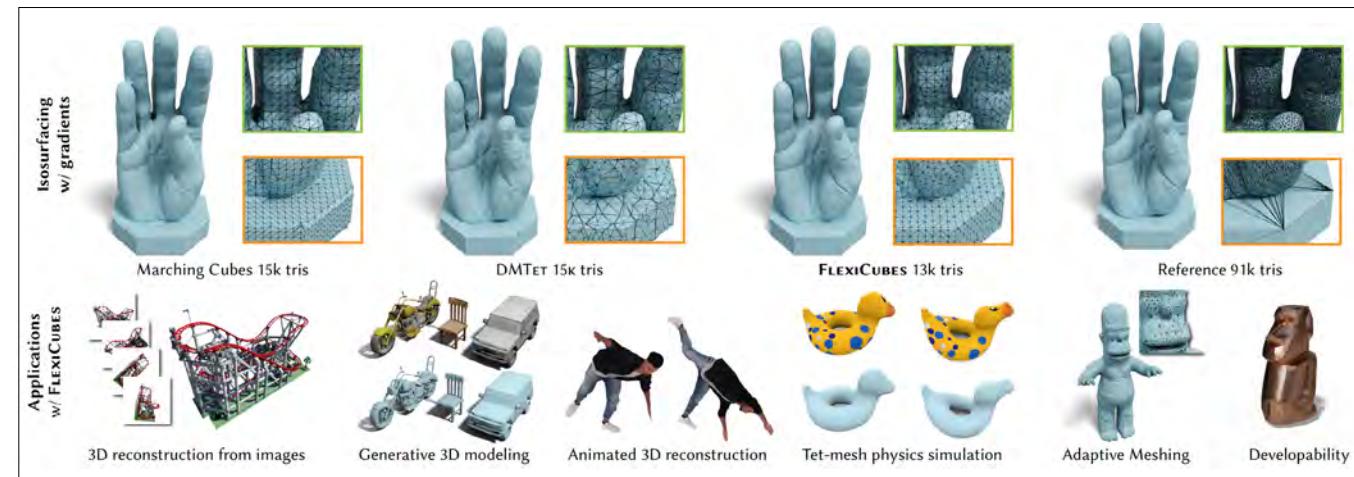


Figure 5.1.2

フレキシキューブ

フレキシキューブを用いた3Dメッシュ最適化

3Dメッシュ生成はコンピュータグラフィックスにおいて重要であり、3Dオブジェクトを定義するための頂点、エッジ、面のメッシュを作成することを含みます。これはビデオゲーム、アニメーション、医療画像、科学的可視化において重要です。従来の等値面抽出アルゴリズムは、限られた解像度、構造的剛性、数値的不安定性に苦しむことが多く、これがその後の影響を及ぼします。

FlexiCubesの表面再構築 サンプル

出典: Nvidia, 2023

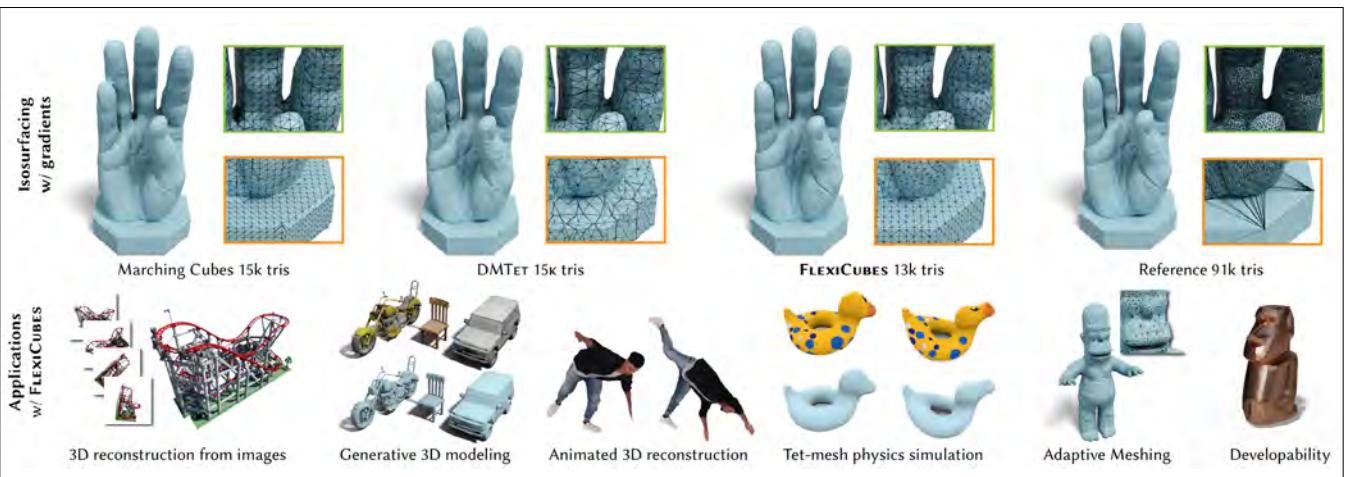


Figure 5.1.2

Select quantitative results on 3D mesh reconstruction

Source: Shen et al., 2023 | Chart: 2024 AI Index report

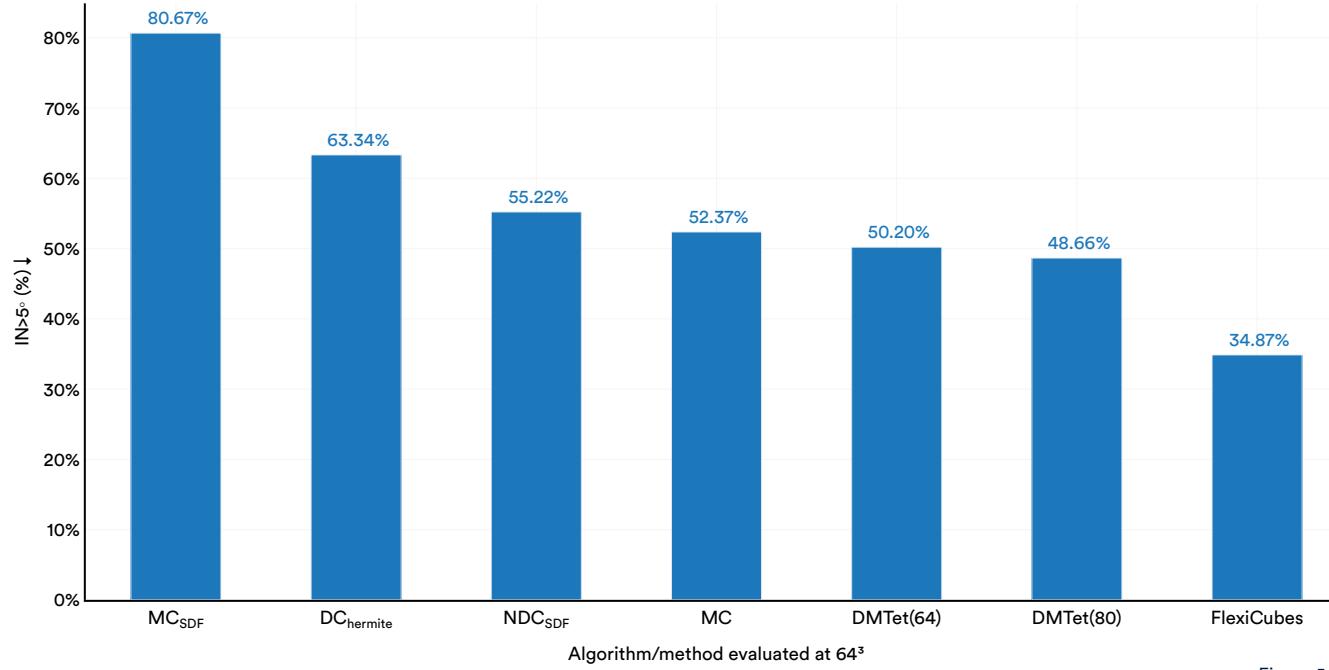


Figure 5.1.3

3Dメッシュ再構築に関する定量的結果を選択

出典: Shen et al., 2023 | チャート: 2024 AIインデックスレポート

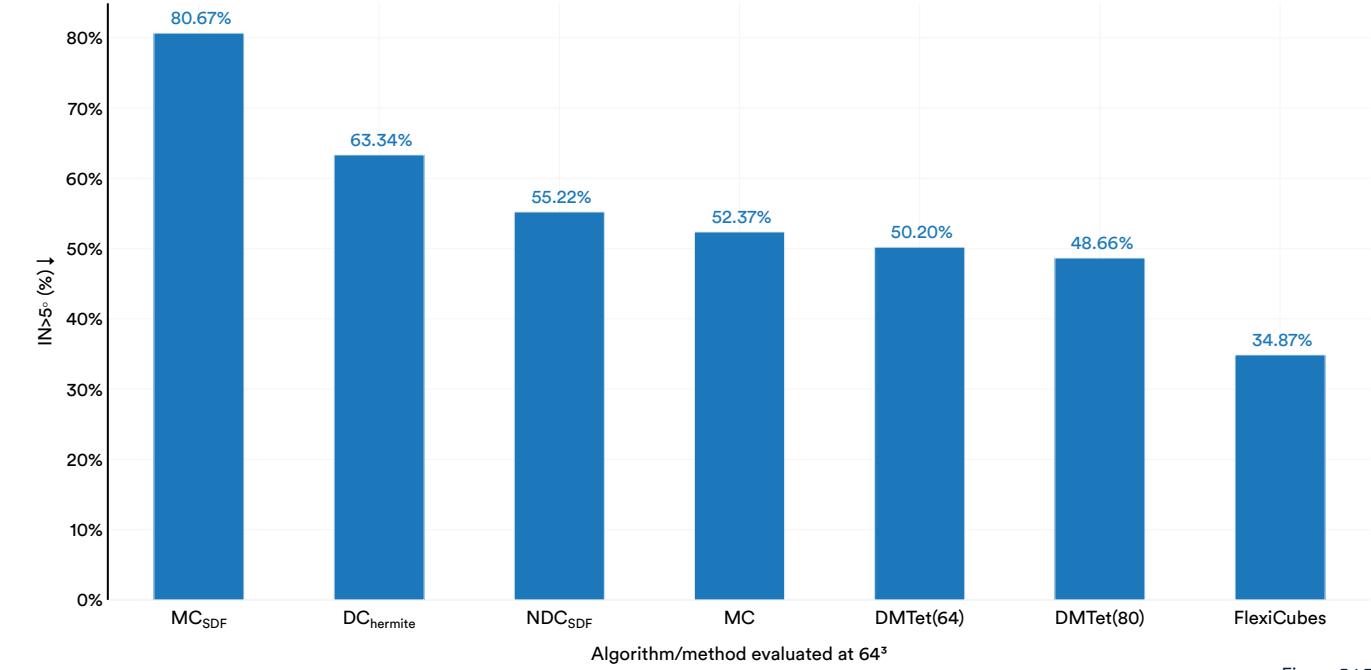


Figure 5.1.3

Synbot

AI-driven robotic chemist for synthesizing organic molecules

Synbot employs a multilayered system, comprising an AI software layer for chemical synthesis planning, a robot software layer for translating commands, and a physical robot layer for conducting experiments. The closed-loop feedback mechanism between the AI and the robotic system enables Synbot to develop synthetic recipes with yields equal to or exceeding established references

(Figure 5.1.4). In an experiment aimed at synthesizing M1 [4-(2,3-dimethoxyphenyl)-1H-pyrrolo[2,3-b]pyridine], Synbot developed multiple synthetic formulas that achieved conversion yields surpassing

Synbot design

Source: Ha et al., 2023

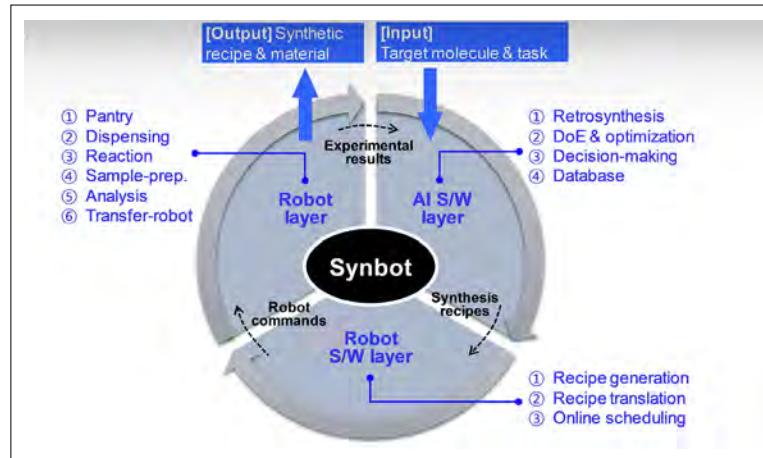


Figure 5.1.4

the mid-80% reference range and completed the synthesis in significantly less time (Figure 5.1.5). Synbot's automation of organic synthesis highlights AI's potential in fields such as pharmaceuticals and materials science.

シンボット

AI 駆動のロボット 化学者 のための

有機分子の合成 Synbotは、化学合成計画のためのAIソフトウェア層、コマンドを翻訳するためのロボットソフトウェア層、実験を実施するための物理ロボット層からなる多層システムを採用しています。AIとロボットシステム間の閉ループフィードバックメカニズムにより、Synbotは確立された基準と同等またはそれを超える収率を持つ合成レシピを開発することができます（図 5.1.4）。

M1 [4-(2,3-ジメトキシフェニル)-1H-ピロロ[2,3-b]ピリジン] の合成を目指した実験では、Synbotは変換収率が基準を超える複数の合成式を開発しました。

Synbotの設計

出典: Ha et al., 2023

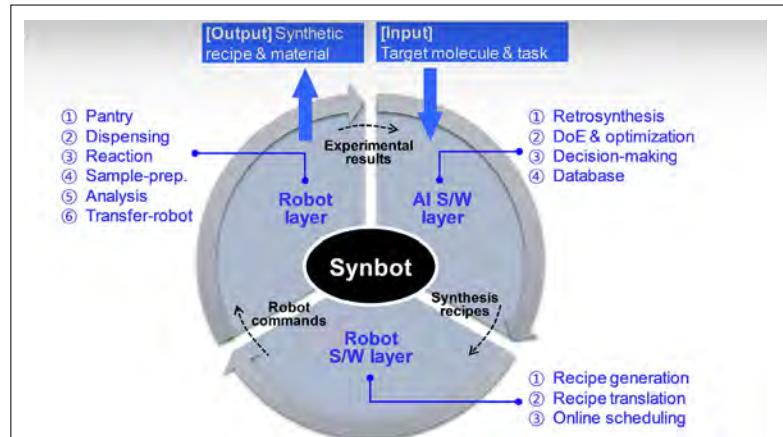


図 5.1.4

中 80% の基準範囲を超え、合成を大幅に短時間で完了しました（図 5.1.5）。Synbotによる有機合成の自動化は、製薬や材料科学などの分野におけるAIの可能性を強調しています。

Reaction kinetics of M1 autonomous optimization experiment, Synbot vs. reference

Source: Ha et al., 2023 | Chart: 2024 AI Index report

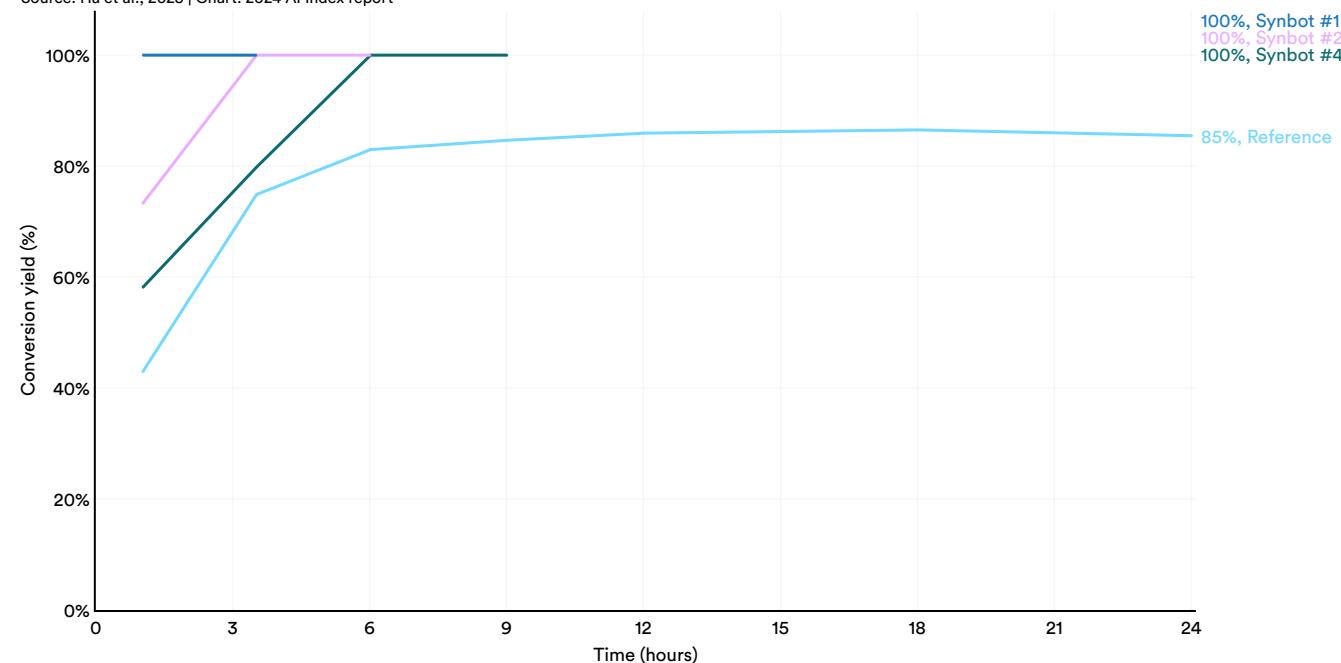


Figure 5.1.5

M1 自律最適化実験の反応速度論、Synbot 対基準

出典: Ha et al., 2023 | チャート: 2024 AI インデックスレポート

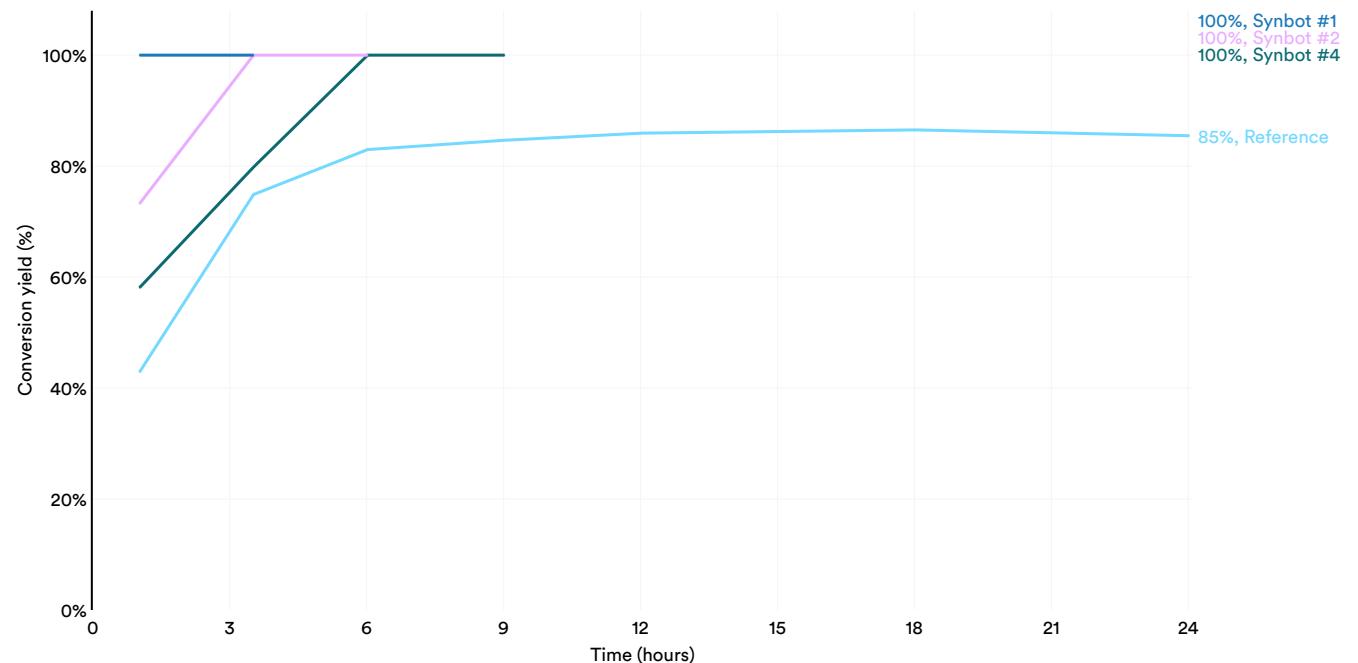


Figure 5.1.5

GraphCast

More accurate global weather forecasting with GraphCast

GraphCast is a new weather forecasting system that delivers highly accurate 10-day weather predictions in under a minute (Figure 5.1.6). Utilizing graph neural networks and machine learning, GraphCast processes vast datasets to forecast temperature, wind speed, atmospheric conditions,

and more. Figure 5.1.7 compares the performance of GraphCast with the current industry state-of-the-art weather simulation system: the High Resolution Forecast (HRES). GraphCast posts a lower root mean squared error, meaning its forecasts more closely correspond to observed weather patterns. GraphCast can be a valuable tool in deciphering weather patterns, enhancing preparedness for extreme weather events, and contributing to global climate research.

GraphCast weather prediction

Source: DeepMind, 2023

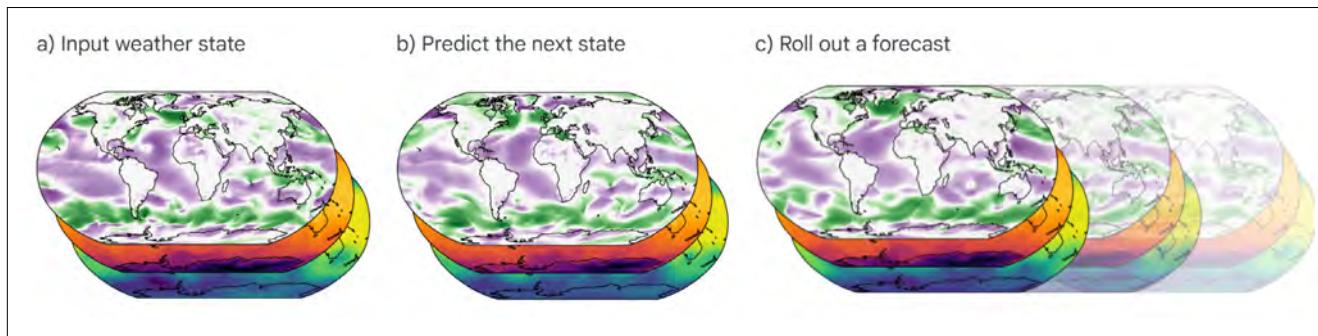


Figure 5.1.6

Ten-day z500 forecast skill: GraphCast vs. HRES

Source: Lam et al., 2023 | Chart: 2024 AI Index report

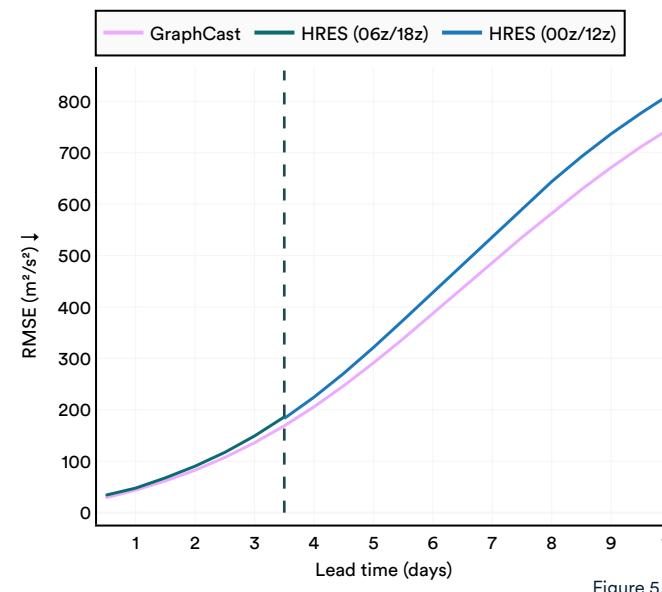


Figure 5.1.7

グラフキャスト

グラフキャストによるより正確な全球気象予測

GraphCastは、1分未満で非常に正確な10日間の天気予報を提供する新しい天気予報システムです（図5.1.6）。グラフニューラルネットワークと機械学習を利用して、GraphCastは膨大なデータセットを処理し、気温、風速、大気条件を予測します。

GraphCast 天気予測

出典: DeepMind, 2023

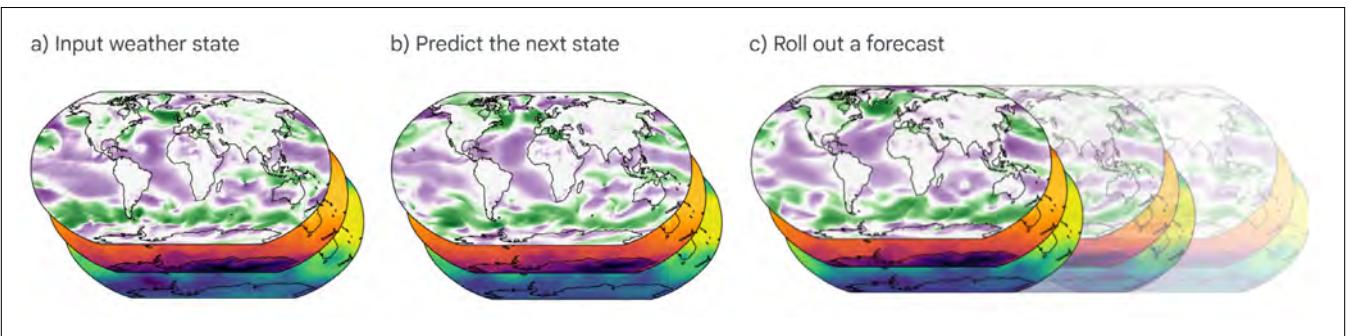


図 5.1.6

10日間のz500予測スキル: GraphCast対HRES

出典: Lam et al., 2023 | チャート: 2024 AI Index report

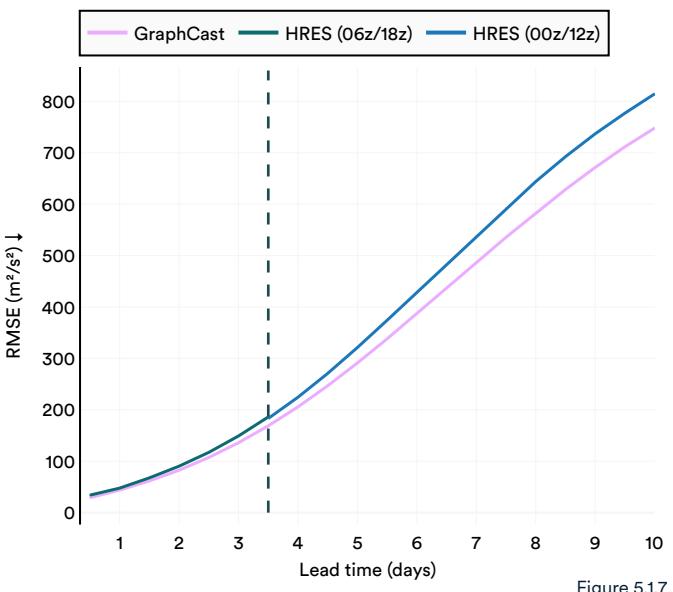


Figure 5.1.7

GNoME

Discovering new materials with GNoME

The search for new functional materials is key to advancements in various scientific fields, including robotics and semiconductor manufacturing. Yet this discovery process is typically expensive and slow. Recent advancements by Google researchers have demonstrated that graph networks, a type of AI model, can expedite this process when trained on large datasets. Their model, **GNoME**, outperformed the Materials Project, a leading method in materials discovery, by identifying a significantly larger number of stable crystals (Figure 5.1.8). GNoME has unveiled 2.2 million new crystal structures, many overlooked by human researchers (Figure 5.1.9 and Figure 5.1.10). The success of AI-driven projects like GNoME highlights the power of data and scaling in speeding up scientific breakthroughs.

Sample material structures

Source: Merchant et al., 2023

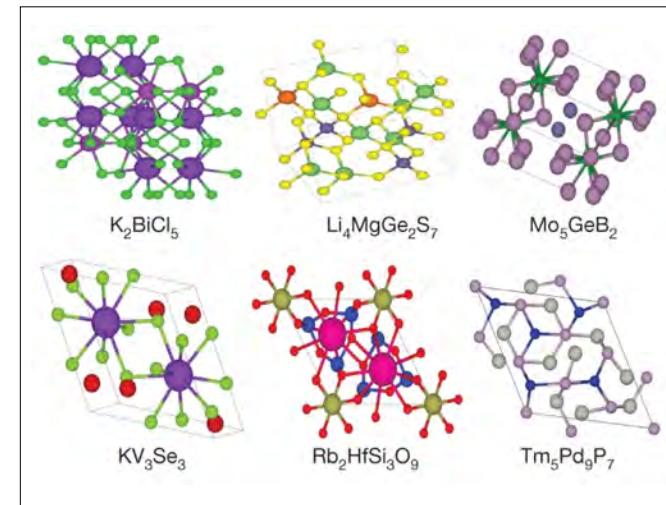


Figure 5.1.8

GNoME

GNoMEを用いた新材料の発見

新しい機能性材料の探索は、ロボティクスや半導体製造を含むさまざまな科学分野の進展にとって重要です。しかし、この発見プロセスは通常、高価で遅いものです。Googleの研究者による最近の進展は、AIモデルの一種であるグラフネットワークが、大規模なデータセットで訓練されることでこのプロセスを加速できることを示しています。彼らのモデルGNoMEは、材料発見の主要な手法であるMaterials Projectを上回り、一はるかに多くの安定した結晶を特定しました（図 5.1.8）。GNoMEは220万の新しい結晶構造を明らかにし、多くは人間の研究者によって見落とされていました（図 5.1.9および図 5.1.10）。GNoMEのようなAI駆動のプロジェクトの成功は、科学的ブレークスルーを加速するデータとスケーリングの力を強調しています。

サンプル材料構造
出典: Merchant et al., 2023

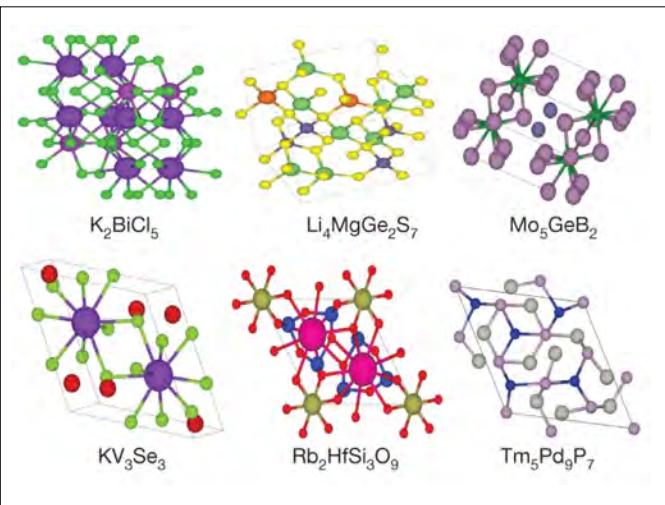


Figure 5.1.8

GNoME vs. Materials Project: stable crystal count

Source: Merchant et al., 2023 | Chart: 2024 AI Index report

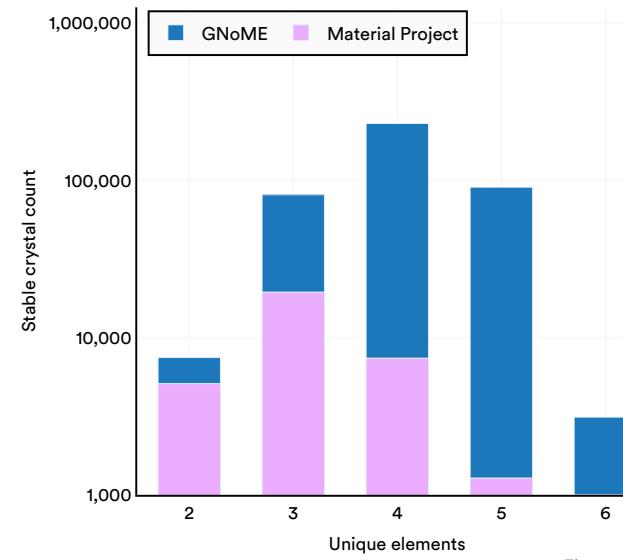


Figure 5.1.9

GNoME vs. Materials Project: distinct prototypes

Source: Merchant et al., 2023 | Chart: 2024 AI Index report

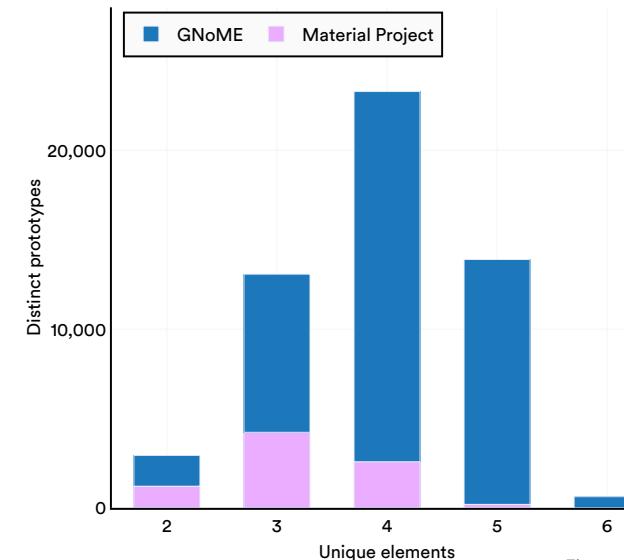


Figure 5.1.10

GNoME 対 Materials Project: 安定した結晶数

出典: Merchant et al., 2023 | チャート:
2024 AI インデックスレポート

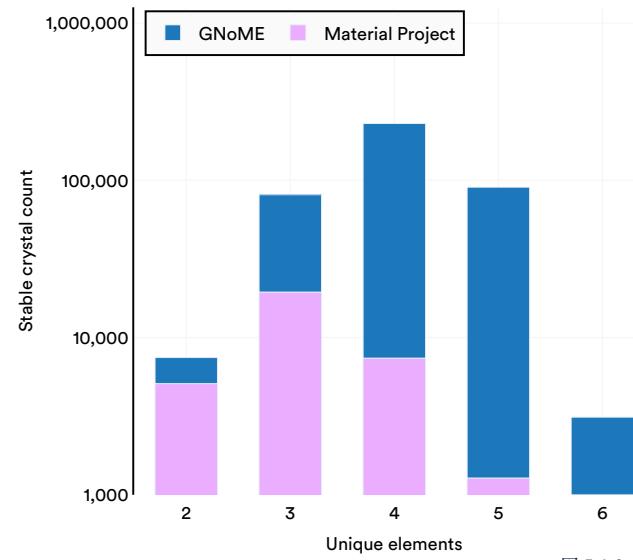


図 5.1.9

GNoME 対 Materials Project: 異なるプロトタイプ

出典: Merchant et al., 2023 | チャート:
2024 AI インデックスレポート

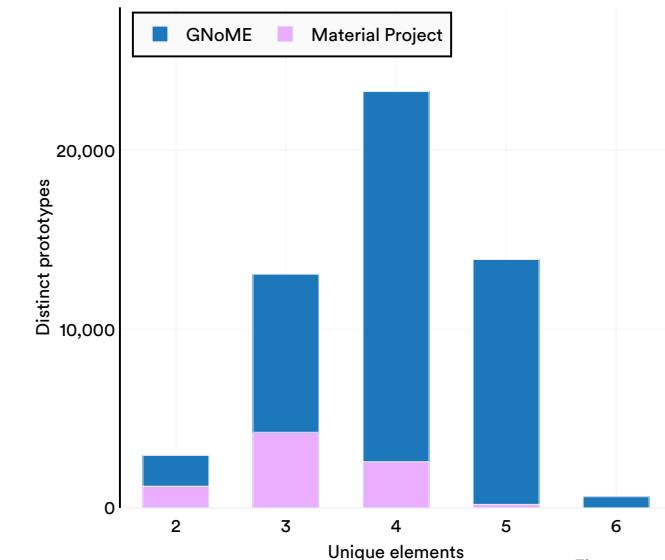


Figure 5.1.10

Flood Forecasting

AI for more accurate and reliable flood forecasts

New research introduced in 2023 has made significant progress in predicting large-scale flood events. Floods, among the most common natural disasters, have particularly devastating effects in less developed countries where infrastructure for prevention and mitigation is lacking. Consequently, developing more accurate prediction methods that can forecast these events further in advance could yield substantial positive impacts.

A team of Google researchers has used AI to develop highly accurate hydrological simulation models that are also applicable to ungauged basins.¹ These innovative methods can predict certain extreme flood events up to five days in advance, with accuracy that matches or surpasses current state-of-the-art models, such as GloFAS. The AI model demonstrates superior precision (accuracy of positive predictions) and recall (ability to correctly identify all relevant instances) across a range of return period events, outperforming the leading contemporary method (Figure 5.1.11).² The model is open-source and is already being used to predict flood events in over 80 countries.

Predictions of AI model vs. GloFAS across return periods

Source: Nearing et al., 2023 | Chart: 2024 AI Index report

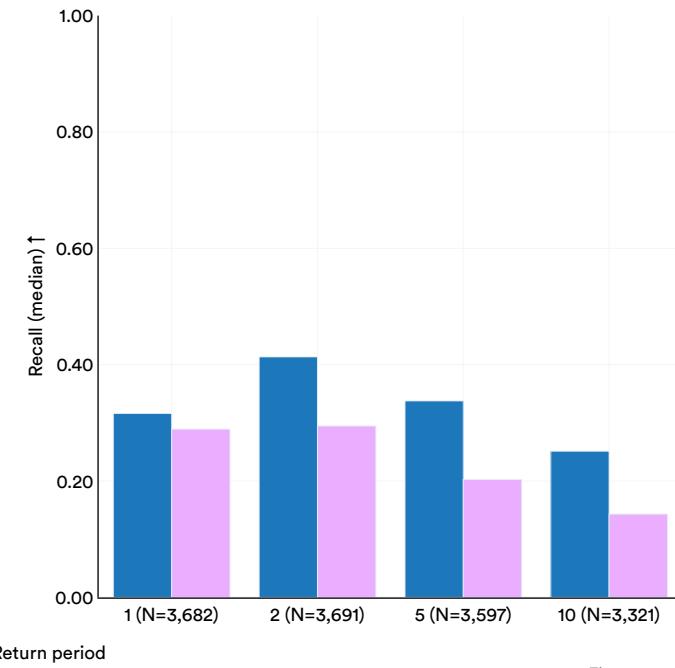
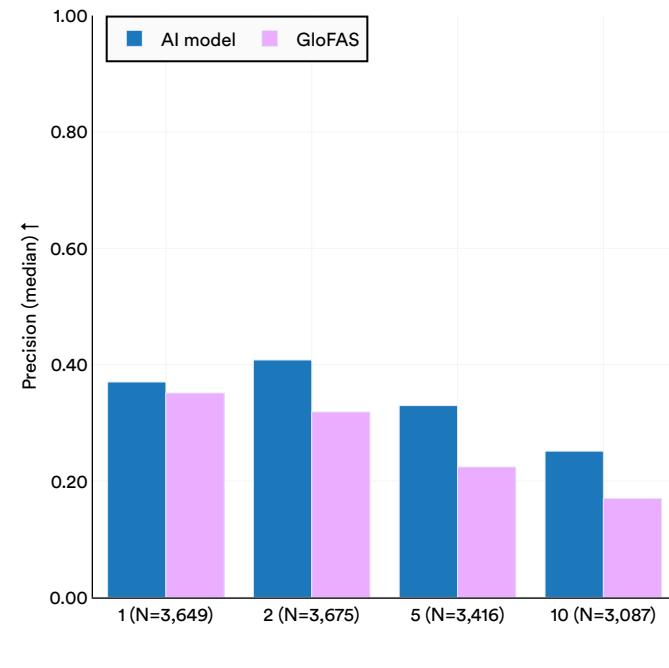


Figure 5.1.11

¹ An ungauged basin is a watershed for which there is insufficient streamflow data to model hydrological flows.

² A return period (recurrence interval) measures the likelihood of a particular hydrological event recurring within a specific period. For example, a 100-year flood means there is a 1% chance of the event being equaled or exceeded in any given year.

洪水予測

より正確で信頼性の高い洪水予測のためのAI

2023年に発表された新しい研究は、大規模な洪水イベントの予測において重要な進展を遂げました。洪水は最も一般的な自然災害の一つであり、特に予防と緩和のためのインフラが不足している発展途上国では壊滅的な影響を及ぼします。そのため、これらのイベントをさらに早く予測できるより正確な予測方法を開発することは、かなりのポジティブな影響をもたらす可能性があります。

Googleの研究チームは、AIを使用して、無計測流域にも適用可能な非常に正確な水文学的シミュレーションモデルを開発しました。これらの革新的な方法は、極端な洪水イベントの一部を最大5日前に予測でき、その精度はGloFASなどの現在の最先端モデルに匹敵するか、それを上回ります。このAIモデルは、さまざまなリターン期間のイベントにおいて、優れた精度（正の予測の精度）と再現率（関連するすべてのインスタンスを正しく特定する能力）を示し、現代の主要な方法を上回っています（図5.1.11）。このモデルはオープンソースであり、すでに80カ国以上で洪水イベントの予測に使用されています。

AIモデルとGloFASのリターン期間における予測

Source: Nearing et al., 2023 | Chart: 2024 AI Index report

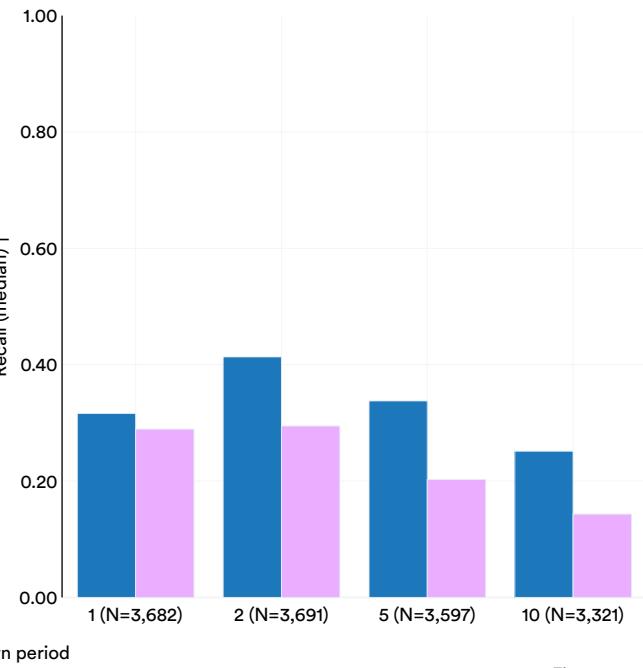
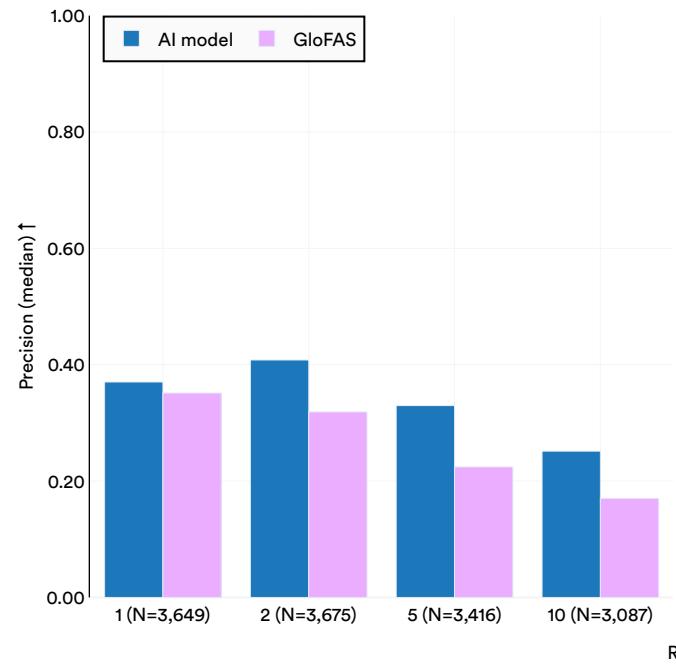


Figure 5.1.11

¹ An ungauged basin is a watershed for which there is insufficient streamflow data to model hydrological flows.

² A return period (recurrence interval) measures the likelihood of a particular hydrological event recurring within a specific period. For example, a 100-year flood means there is a 1% chance of the event being equaled or exceeded in any given year.

AI models are becoming increasingly valuable in healthcare, with applications for detecting polyps to aiding clinicians in making diagnoses. As AI performance continues to improve, monitoring its impact on medical practice becomes increasingly important. This section highlights significant AI-related medical systems introduced in 2023, the current state of clinical AI knowledge, and the development of new AI diagnostic tools and models aimed at enhancing hospital administration.

5.2 AI in Medicine

Notable Medical Systems

This section identifies significant AI-related medical breakthroughs of 2023 as chosen by the AI Index Steering Committee.

SynthSR

Transforming brain scans for advanced analysis

SynthSR is an AI tool that converts clinical brain scans into high-resolution T1 weighted images (Figure 5.2.1). This advancement addresses the issue of scan quality variability, which previously limited the use of many scans in advanced research. By transforming these scans into T1-weighted images, known for their high contrast and clear brain structure depiction, SynthSR facilitates the creation of detailed 3D brain renderings. Experiments using SynthSR demonstrate robust correlations between observed volumes at both scan and subject levels, suggesting that SynthSR generates images closely resembling those produced by high-resolution T1 scans. Figure 5.2.2 illustrates the extent to which SynthSR scans correspond with ground-truth observations across selected brain regions. SynthID significantly improves the visualization and analysis of brain structures, facilitating neuroscientific research and clinical diagnostics.

SynthSR generations

Source: Iglesias et al., 2023

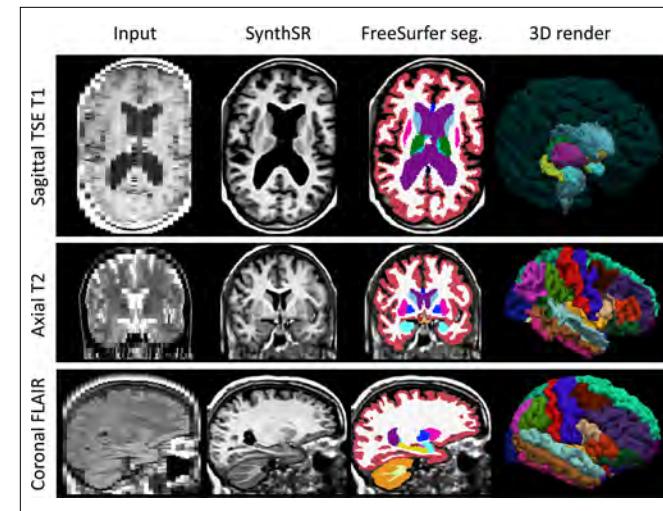


Figure 5.2.1

AIモデルは、ポリープの検出から診断を支援することまで、医療においてますます価値が高まっています。AIの性能が向上し続ける中で、医療実践への影響を監視することがますます重要になっています。このセクションでは、2023年に導入された重要なAI関連医療システム、臨床AIの知識の現状、および病院管理を強化することを目的とした新しいAI診断ツールとモデルの開発について強調します。

5.2 医療におけるAI

注目すべき医療システム

このセクションでは、AIインデックス運営委員会が選定した2023年の重要なAI関連医療のブレイクスルーを特定します。

SynthSR

脳スキャンを高度な分析のために変換するSynthSRは、臨床脳スキャンを高解像度のT1強調画像に変換するAIツールです（図5.2.1）。この進展は、以前は多くのスキャンの高度な研究での使用を制限していたスキャン品質の変動性の問題に対処します。これらのスキャンを高コントラストで明確な脳構造の描写が特徴のT1強調画像に変換することで、SynthSRは詳細な3D脳レンダリングの作成を促進します。SynthSRを使用した実験は、スキャンおよび被験者レベルの両方で観察されたボリュームとの強い相関を示しており、SynthSRが高解像度T1スキャンによって生成された画像に非常に似た画像を生成することを示唆しています。図5.2.2は、選択された脳領域におけるSynthSRスキャンが真実の観察とどの程度一致しているかを示しています。SynthIDは脳構造の視覚化と分析を大幅に改善し、神経科学研究および臨床診断を促進します。

SynthSR生成

出典: Iglesias et al., 2023

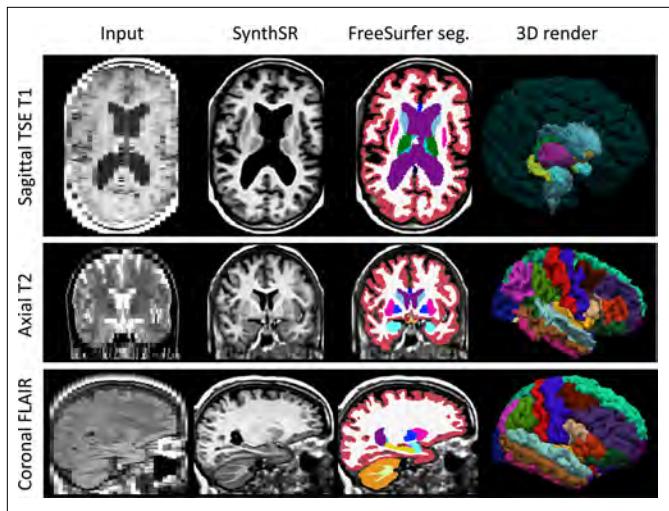


図5.2.1

SynthSR correlation with ground-truth volumes on select brain regions

Source: Iglesias et al., 2023 | Chart: 2024 AI Index report

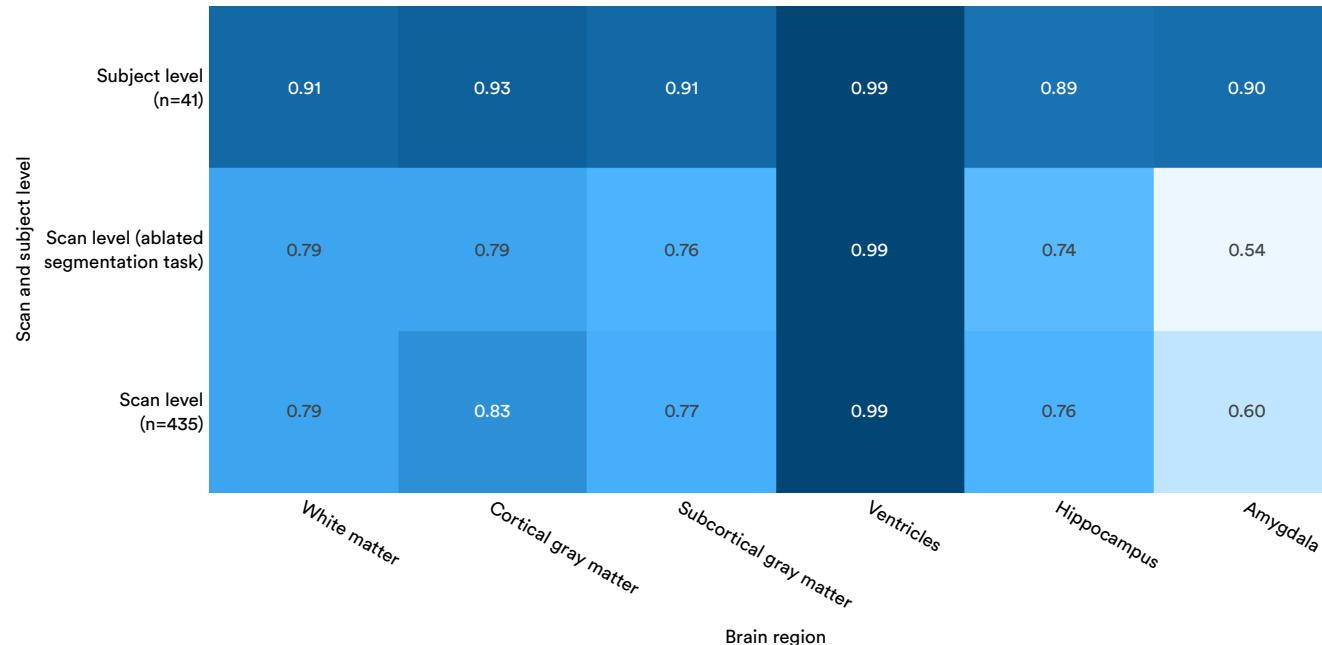


Figure 5.2.2

SynthSRの選択された脳領域における真のボリュームとの相関
出典: Iglesias et al., 2023 ロイヤリー: 2024 AIインデックスレポート

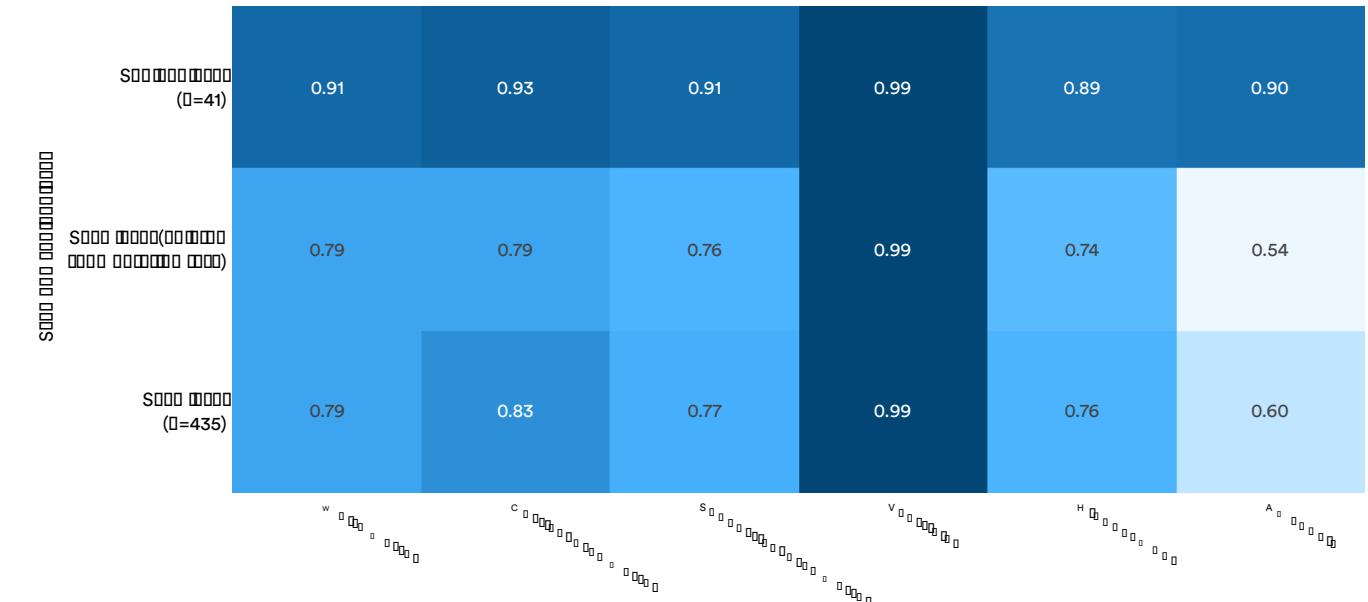


Figure 5.2.2

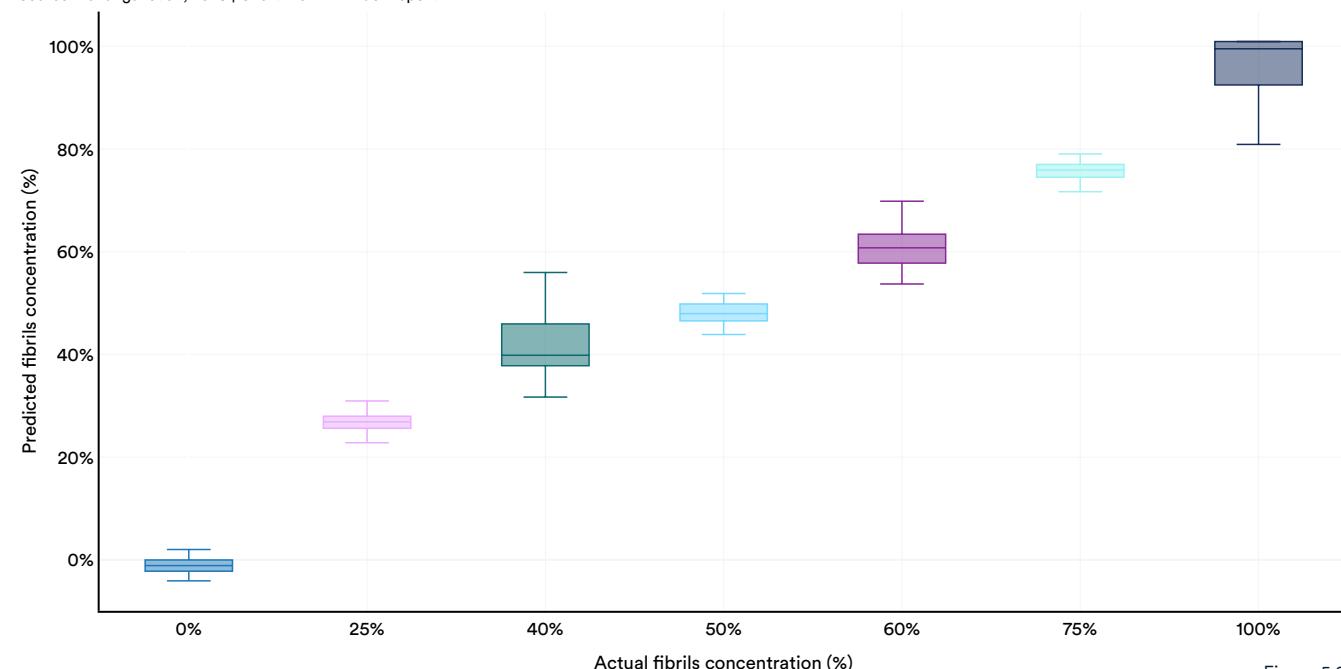
Coupled Plasmonic Infrared Sensors

Coupled plasmonic infrared sensors for the detection of neurodegenerative diseases

Diagnosis of neurodegenerative diseases such as Parkinson's and Alzheimer's depends on fast and precise identification of biomarkers. Traditional methods, such as mass spectrometry and ELISA, are useful in that they can focus on quantifying protein levels; however, they cannot discern changes in structural states. This year, researchers uncovered a new method for neurodegenerative disease diagnosis that combined AI-coupled plasmonic infrared sensors that use Surface-Enhanced Infrared Absorption (SEIRA) spectroscopy with an immunoassay technique (ImmunoSEIRA; Figure 5.2.3). In tests that compared actual fibril percentages with predictions made by AI systems, the accuracy of the predictions was found to very closely match the actual reported percentages (Figure 5.2.4).

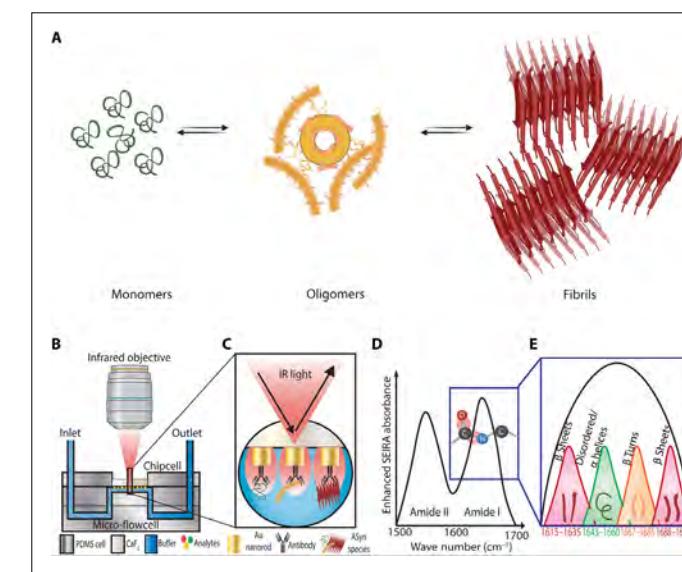
Deep neural network predicted vs. actual fibrils percentages in test samples

Source: Kavungal et al., 2023 | Chart: 2024 AI Index report



ImmunoSEIRA detection principle and the setup

Source: Kavungal et al., 2023



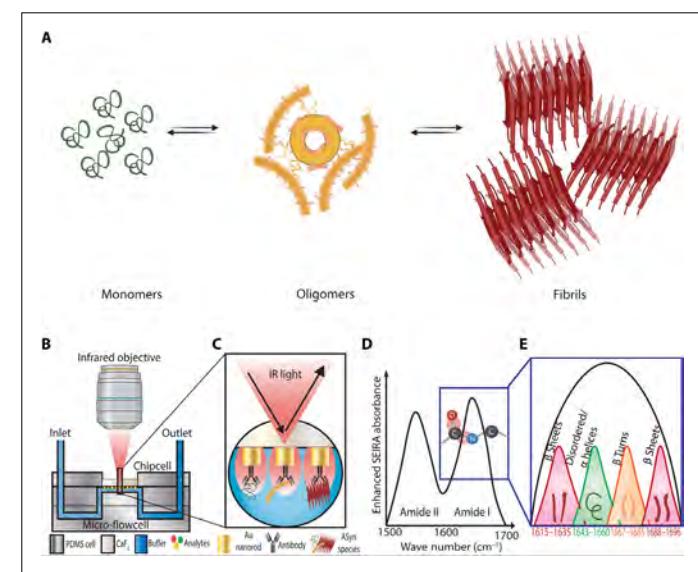
カップルプラズモニック赤外線センサー

神経変性疾患の検出のためのカップルプラズモニック赤外線センサー

パーキンソン病やアルツハイマー病などの神経変性疾患の診断は、バイオマーカーの迅速かつ正確な特定に依存しています。質量分析法やELISAなどの従来の方法は、タンパク質レベルの定量に焦点を当てることができるため有用ですが、構造状態の変化を識別することはできません。今年、研究者たちは、表面増強赤外吸収(SEIRA)分光法を使用したAI結合プラズモニック赤外線センサーと免疫アッセイ技術(ImmunoSEIRA; 図5.2.3)を組み合わせた神経変性疾患診断の新しい方法を発見しました。実際のフィブリルの割合とAIシステムによって行われた予測を比較したテストでは、予測の精度が実際に報告された割合と非常に密接に一致することがわかりました(図5.2.4)。

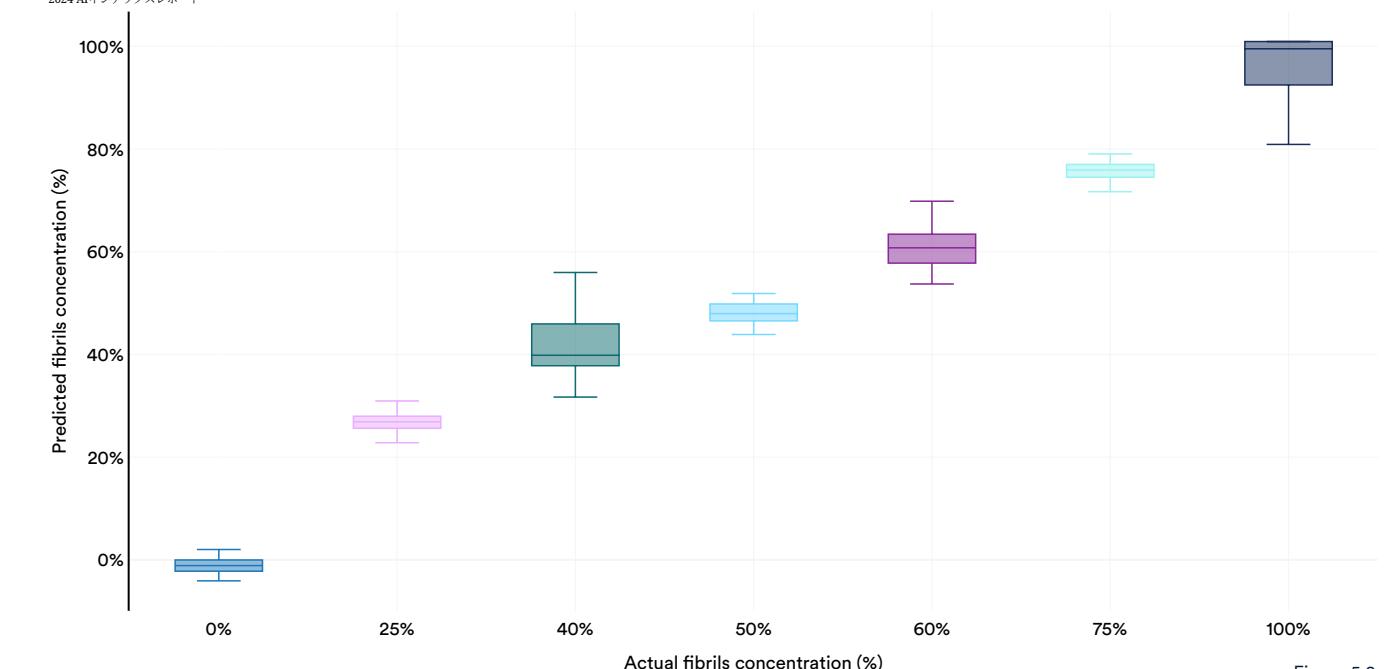
ImmunoSEIRA 検出原理とセットアップ

出典: Kavungal et al., 2023



深層ニューラルネットワークによる予測と実際のフィブリルの割合のテストサンプル

出典: Kavungal et al., 2023 | チャート: 2024 AI Index report



EVEscape

Forecasting viral evolution for pandemic preparedness

Predicting viral mutations is vital for vaccine design and pandemic minimization. Traditional methods, which rely on real-time virus strain and antibody data, face challenges during early pandemic stages due to data scarcity. EVEscape is a new AI deep learning model trained on historical sequences and biophysical and structural information that predicts the evolution

of viruses (Figure 5.2.5). EVEscape evaluates viral escape independently of current strain data predicting 50.0% of observed SARS-CoV-2 mutations, outperforming traditional lab studies which predicted 46.2% and 32.3%, as well as a previous model, which predicted only 24% of mutations (Figure 5.2.6). This performance highlights EVEscape's potential as a valuable asset for enhancing future pandemic preparedness and response efforts.

EVEscape design

Source: Thadani et al., 2023

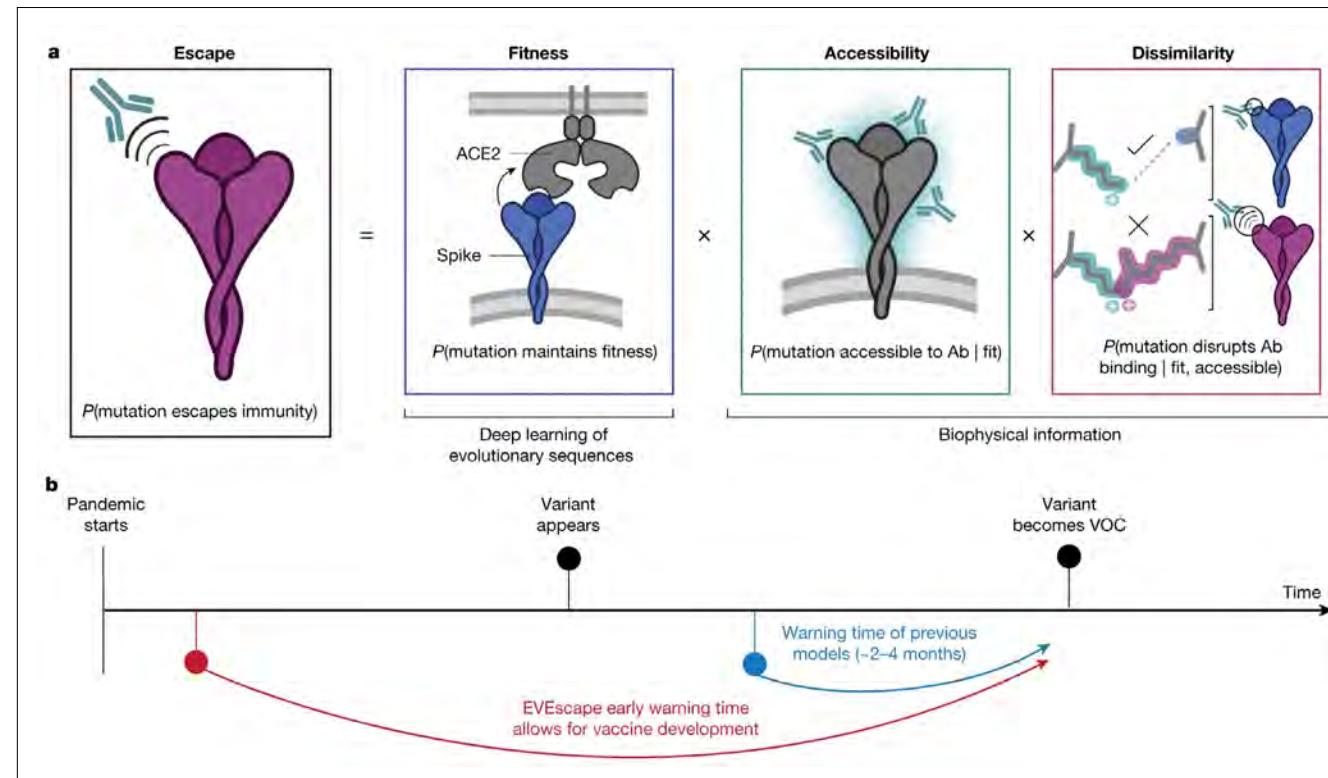


Figure 5.2.5

EVEscape

パンデミック対策のためのウイルス進化の予測

ウイルスの変異を予測することは、ワクチン設計とパンデミックの最小化にとって重要です。従来の方法は、リアルタイムのウイルス株と抗体データに依存しており、データが不足しているパンデミックの初期段階では課題に直面します。EVEscapeは、歴史的な配列と生物物理的情報に基づいて訓練された新しいAI深層学習モデルであり、進化を予測します。

ウイルスの進化（図 5.2.5）。EVEscapeは、現在の株データに依存せずにウイルスの逃避を評価し、観察されたSARS-CoV-2の変異の50.0%を予測し、従来の実験室研究の46.2%および32.3%を上回り、以前のモデルの24%の変異予測をも上回る性能を示しています（図 5.2.6）。この性能は、EVEscapeが将来のパンデミックの準備と対応努力を強化するための貴重な資産となる可能性を強調しています。

EVEscapeの設計

出典: Thadani et al., 2023

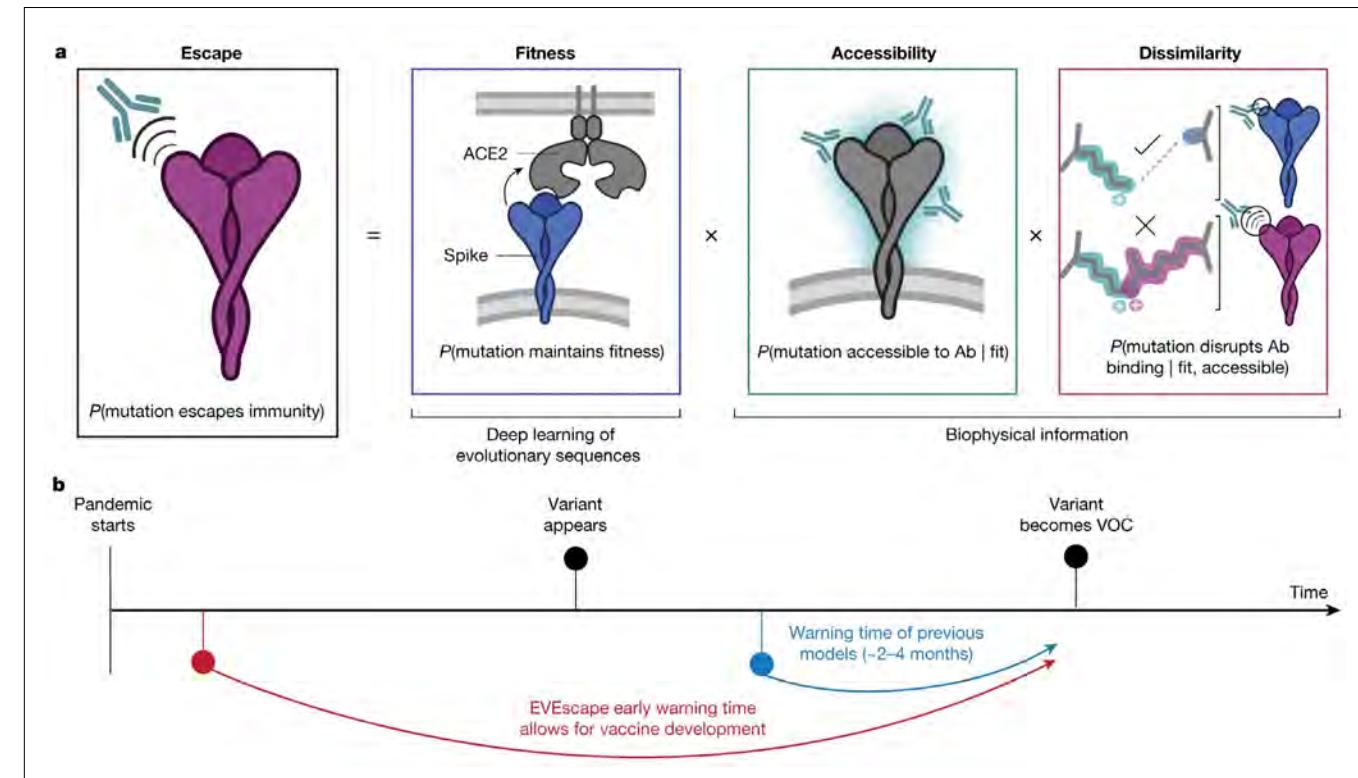
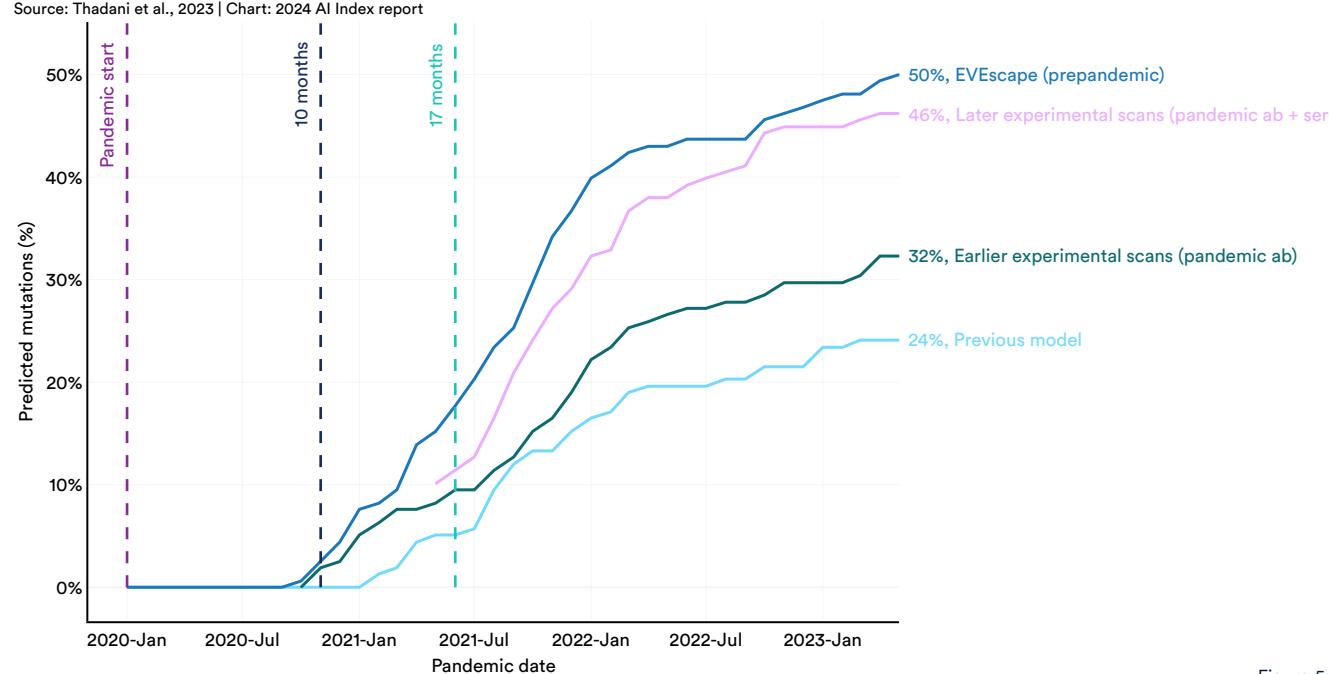


図 5.2.5

EVEscape vs. other models on SARS-CoV-2 RBD mutation prediction



EVEscapeと他のモデルのSARS-CoV-2 RBD 変異予測に関する比較

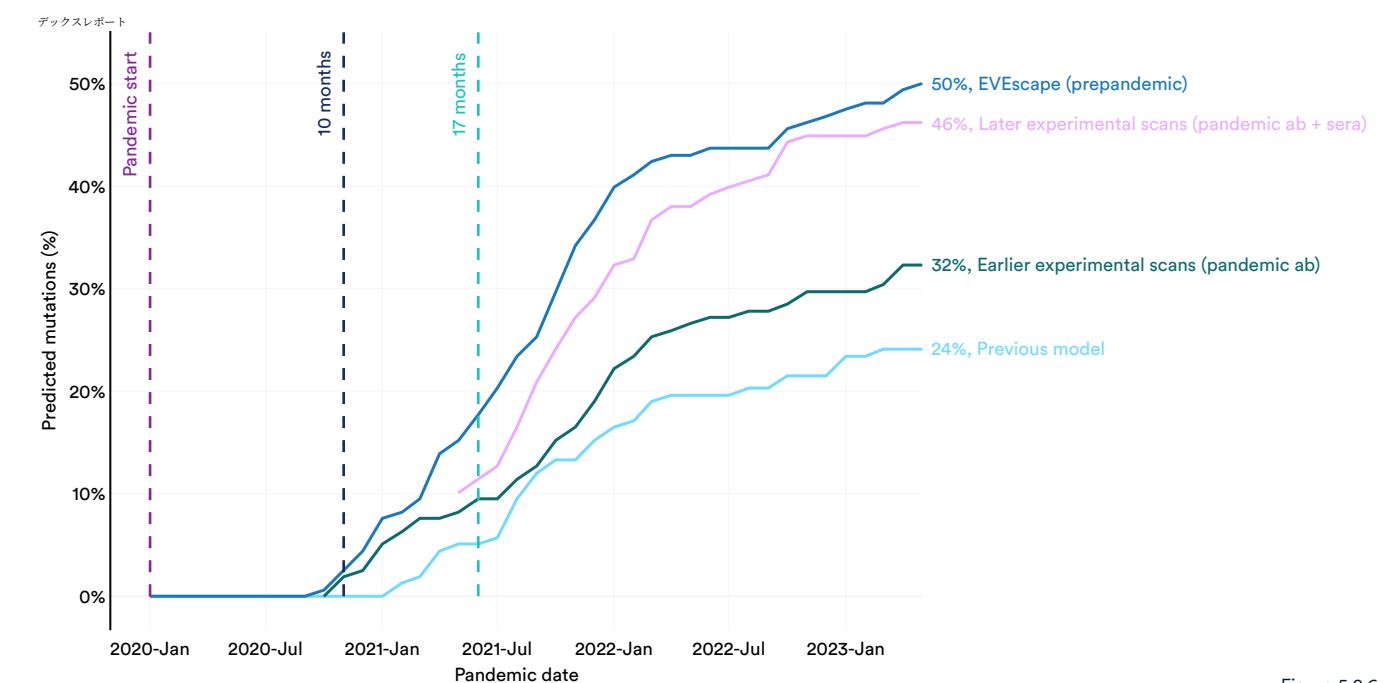


Figure 5.2.6

AlphaMissense

Better classification of AI mutations

Scientists still do not fully understand which genetic mutations lead to diseases. With millions of possible genetic mutations, determining whether a mutation is benign or pathogenic requires labor-intensive experiments.

In 2023, researchers from Google DeepMind unveiled AlphaMissense, a new AI model that predicted the pathogenicity of 71 million missense variants. Missense mutations are genetic alterations that impact the functionality of human proteins (Figure 5.2.7) and can lead to various diseases, including cancer. Of the 71 million possible missense variants, AlphaMissense classified 89%, identifying 57% as likely benign and 32% as likely pathogenic, while the remainder were categorized as uncertain (Figure 5.2.8). In contrast, human annotators have only been able to confirm the nature of 0.1% of all missense mutations.

Hemaglobin subunit beta (HBB)

Source: Google DeepMind, 2023

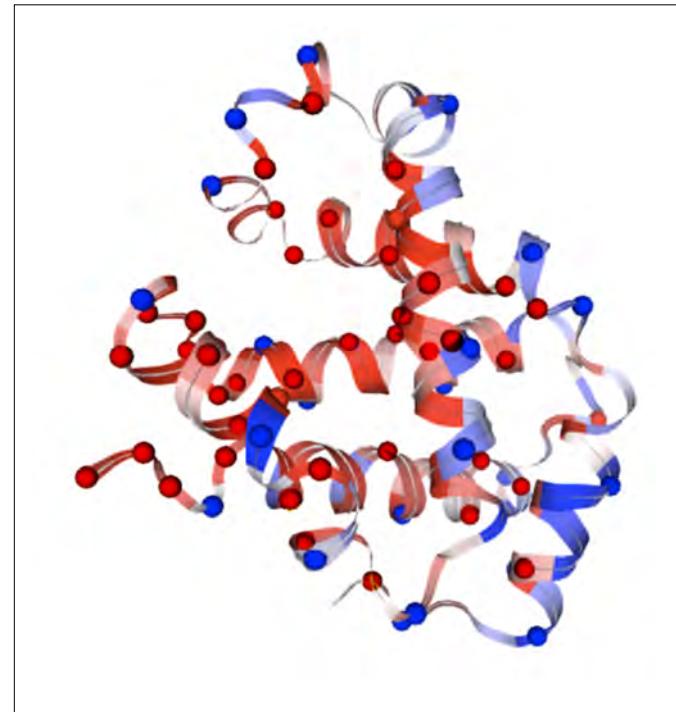


Figure 5.2.7

アルファミスセンス

AI変異のより良い分類

科学者たちは、どの遺伝子変異が病気につながるのかを完全には理解していません。数百万の可能な遺伝子変異の中から、変異が良性か病原性かを判断するには、労力を要する実験が必要です。

2023年、Google DeepMindの研究者たちは、7100万のミスセンス変異の病原性を予測する新しいAIモデル、AlphaMissenseを発表しました。ミスセンス変異は、人間のタンパク質の機能に影響を与える遺伝子の変化であり（図5.2.7）、がんを含むさまざまな病気を引き起こす可能性があります。7100万の可能なミスセンス変異のうち、AlphaMissenseは89%を分類し、57%をおそらく良性、32%をおそらく病原性と特定し、残りは不確定として分類されました（図5.2.8）。対照的に、人間のアナターダーはすべてのミスセンス変異の0.1%の性質を確認することしかできませんでした。

ヘモグロビンサブユニットベータ (HBB) 出典: Google DeepMind, 2023

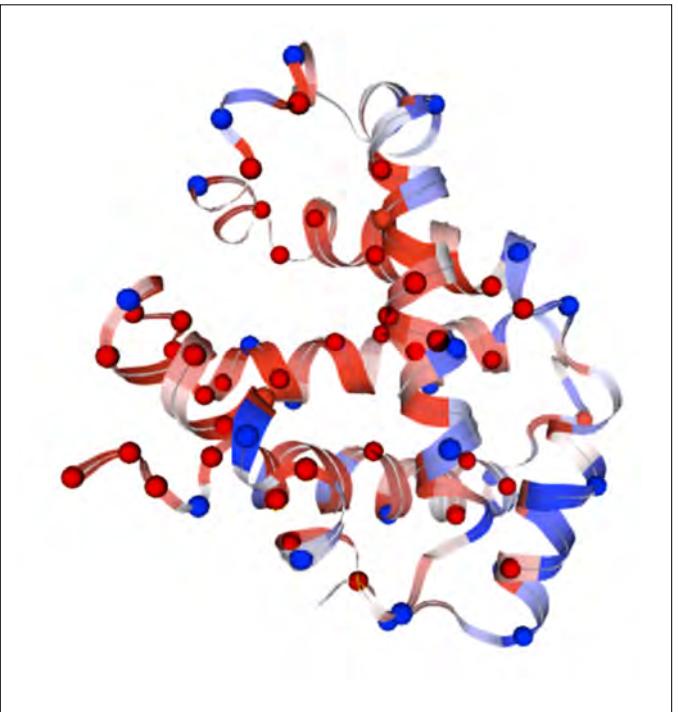


図 5.2.7

AlphaMissense predictions

Source: Google DeepMind, 2023 | Chart: 2024 AI Index report

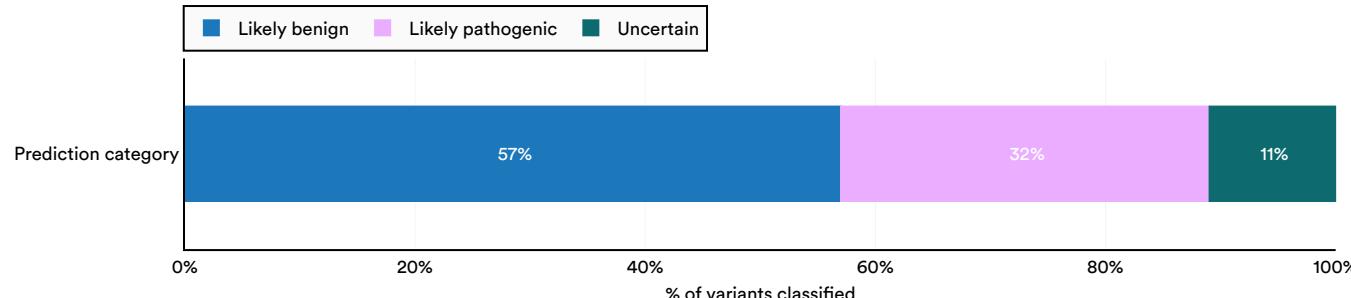


Figure 5.2.8

AlphaMissenseの予測

出典: Google DeepMind, 2023

チャート: 2024年AIインデックスレポート

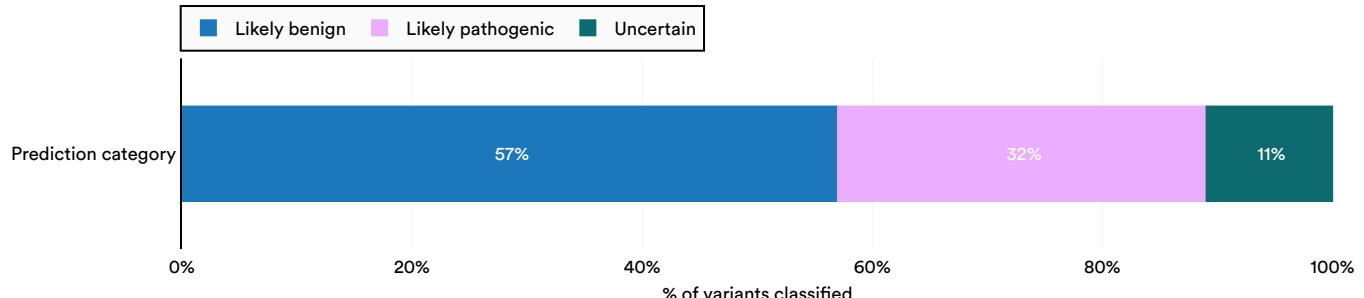


Figure 5.2.8

Human Pangenome Reference

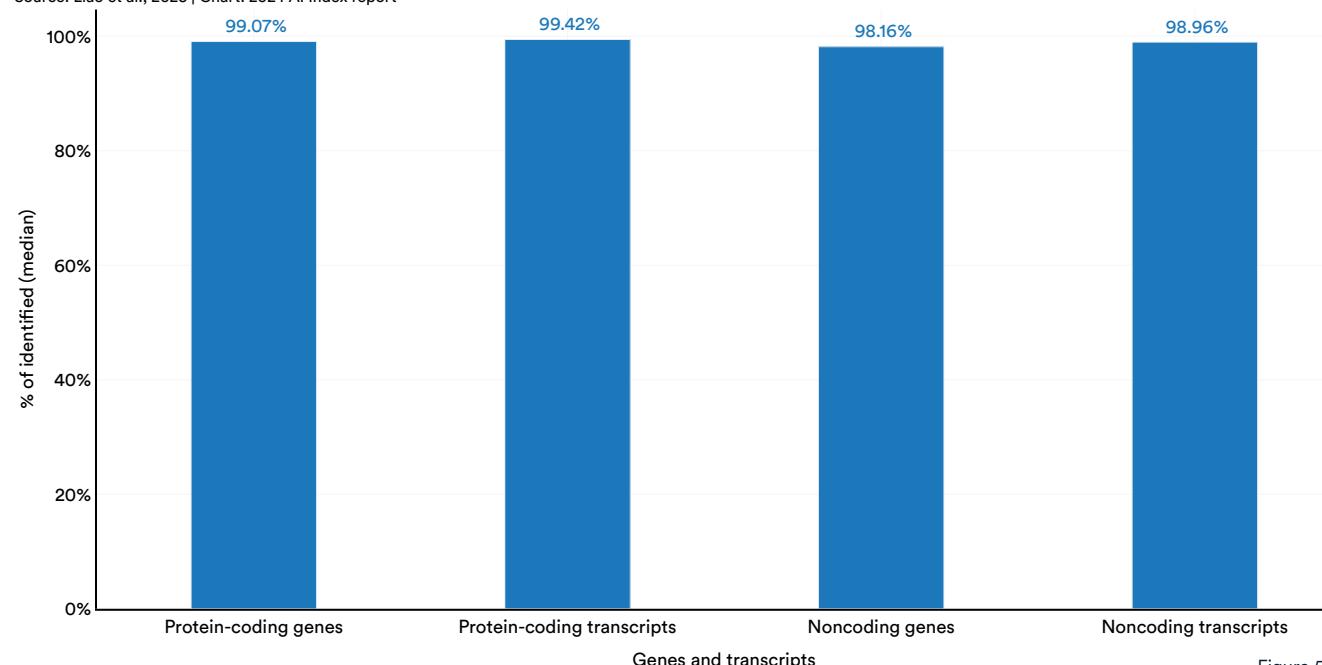
Using AI to map the human genome

The human genome is a set of molecular instructions for a human. The first human genome draft was released in 2000 and updated in 2022. However, the update was somewhat incomplete. It did not incorporate various genetic mutations, like blood type, and did not as completely map diverse ancestry groups. Therefore, under the existing genome reference, it would be difficult to detect diseases or find cures in certain groups of people.

In 2023, the Human Pangenome Research Consortium, comprising 119 scientists from 60 institutions, used AI to develop an updated and more representative human genome map (Figure 5.2.9). The researchers achieved remarkable accuracy, annotating a median of 99.07% of protein-coding genes, 99.42% of protein-coding transcripts, 98.16% of noncoding genes, and 98.96% of noncoding transcripts, as detailed in Figure 5.2.10.

Ensembl mapping pipeline results

Source: Liao et al., 2023 | Chart: 2024 AI Index report



Graph genome for the MHC region of the genome

Source: Google Research, 2023

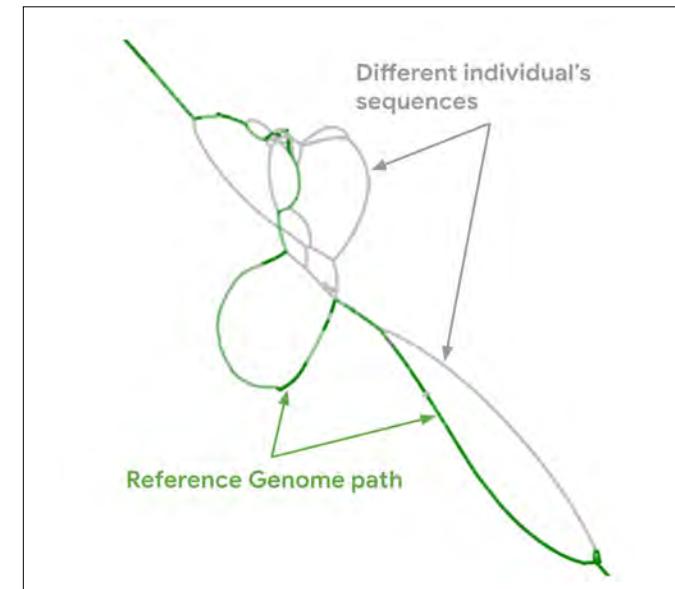


Figure 5.2.9

This latest version of the genome represents the most comprehensive and genetically diverse mapping of the human genome to date.

ヒトパンゲノムリファレンス

AIを使用してヒトゲノムをマッピングする

ヒトゲノムは人間のための分子指示のセットです。最初のヒトゲノムドラフトは2000年に発表され、2022年に更新されました。しかし、その更新はやや不完全でした。血液型のようなさまざまな遺伝子変異を取り入れておらず、多様な祖先グループを完全にマッピングしていませんでした。したがって、既存のゲノム参照の下では、特定の人々のグループで病気を検出したり治療法を見つけることは困難です。

2023年、119人の科学者からなるヒトパンゲノム研究コンソーシアムは、AIを使用して更新されたより代表的なヒトゲノムマップを開発しました（図5.2.9）。研究者たちは驚異的な精度を達成し、タンパク質コーディング遺伝子の中央値99.07%、タンパク質コーディング転写物の99.42%、ノンコーディング遺伝子の98.16%、ノンコーディング転写物の98.96%を注釈しました。詳細は図5.2.10に示されています。

ゲノムのMHC領域のグラフゲノム

出典: Google Research, 2023

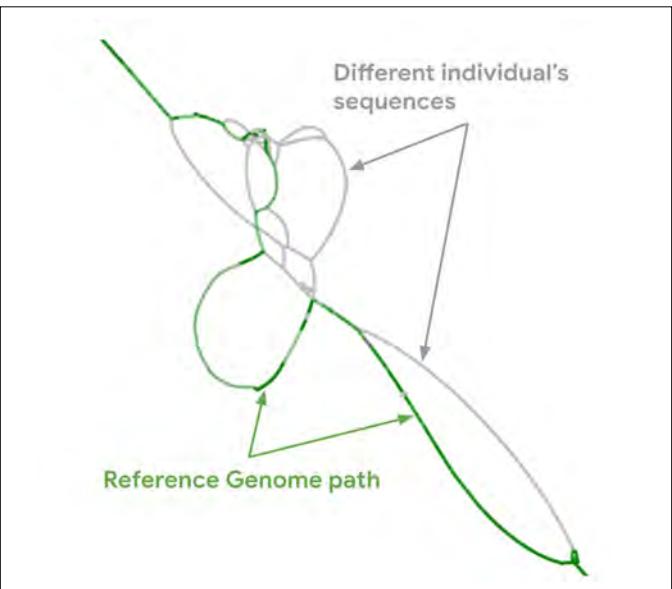
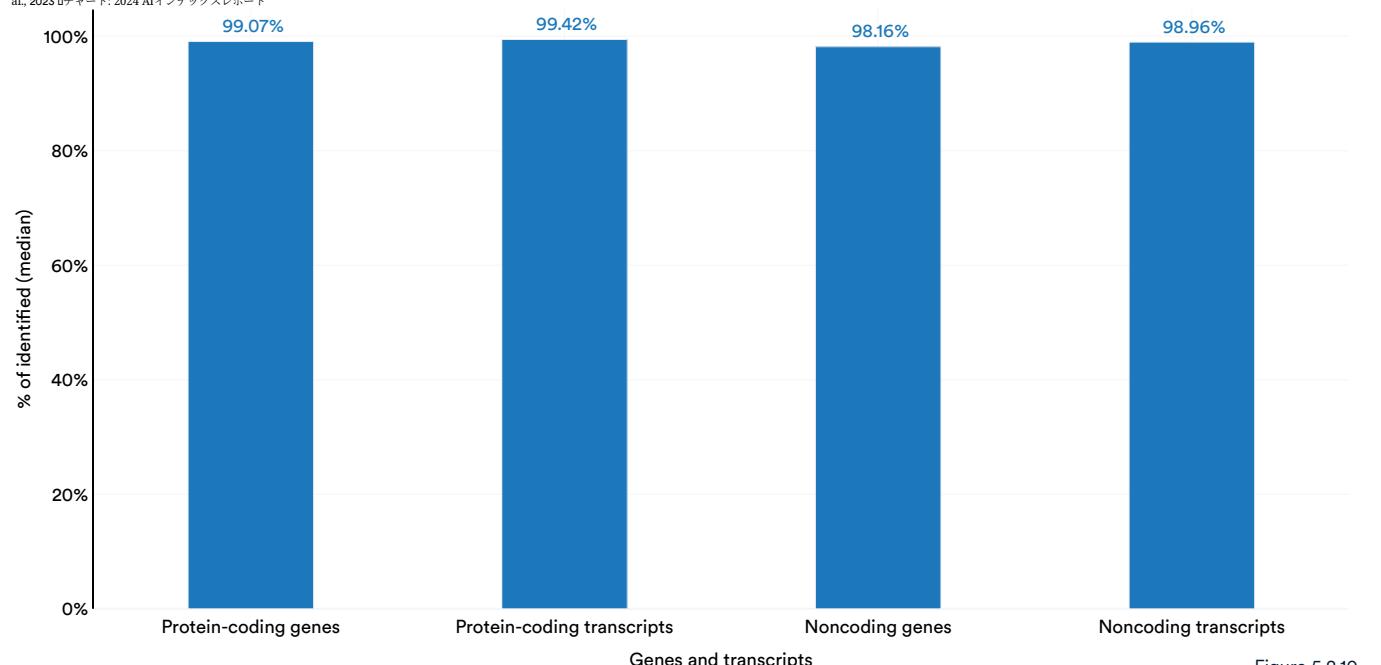


Figure 5.2.9

この最新のゲノムバージョンは、これまでで最も包括的に遺伝的に多様なヒトゲノムのマッピングを表しています。

Ensemblマッピングパイプラインの結果

出典: Liao et al., 2023 | Chart: 2024 AI Index report



Clinical Knowledge

Evaluating the clinical knowledge of AI models involves determining the extent of their medical expertise, particularly knowledge applicable in a clinical setting.

MedQA

Introduced in 2020, MedQA is a comprehensive dataset derived from professional medical board exams, featuring over 60,000 clinical questions designed to challenge doctors.

AI performance on the MedQA benchmark has seen

MedQA: accuracy

Source: Papers With Code, 2023 | Chart: 2024 AI Index report

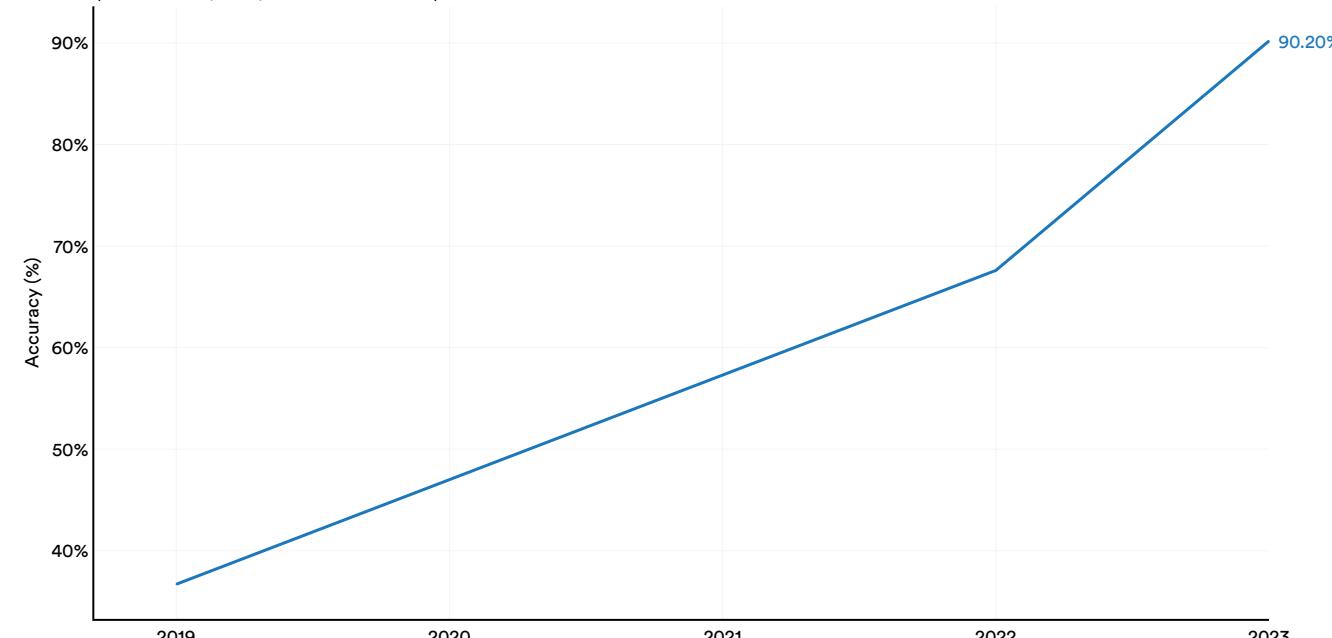


Figure 5.2.11

臨床知識

AIモデルの臨床知識を評価することは、特に臨床環境で適用可能な医療専門知識の範囲を判断することを含みます。

MedQA

2020年に導入されたMedQAは、医療専門のボード試験から派生した包括的なデータセットで、医師に挑戦するために設計された60,000以上の臨床問題を特徴としています。

MedQAベンチマークにおけるAIのパフォーマンスは

MedQA: 精度

Source: Papers With Code, 2023 | Chart: 2024 AI Index report

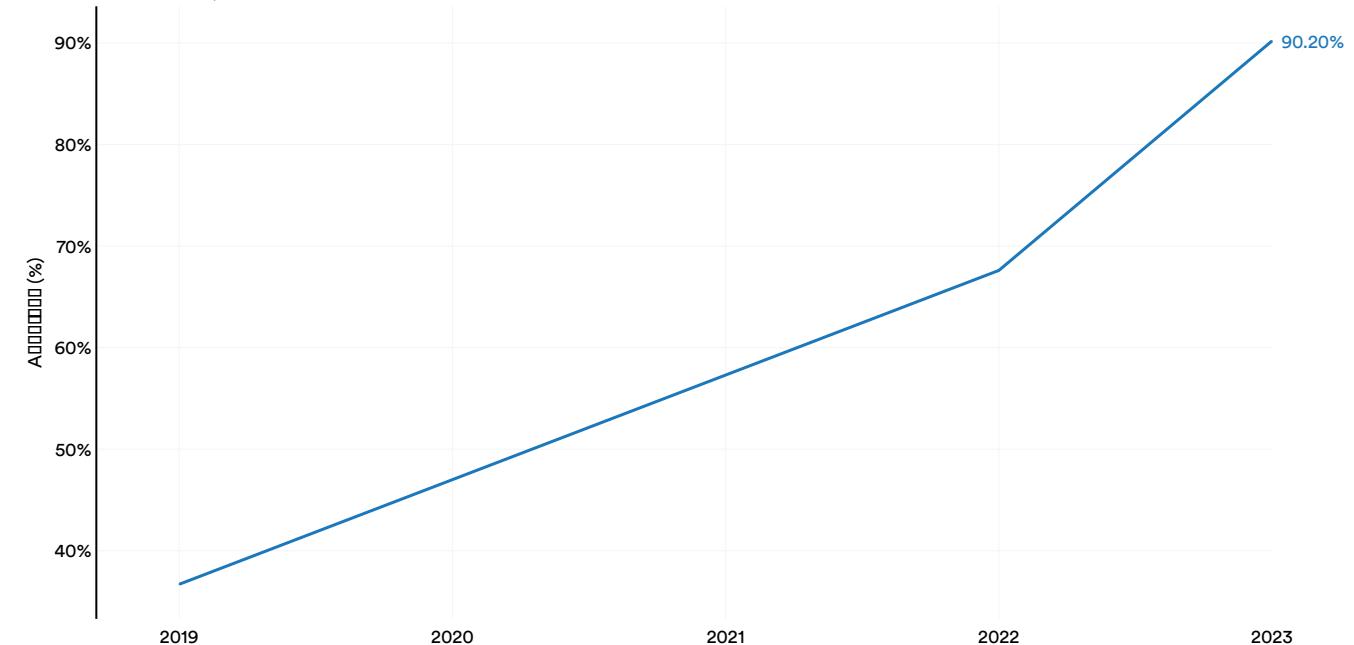


Figure 5.2.11

Highlighted Research:

GPT-4 Medprompt

Although LLMs exhibit impressive general knowledge, it is commonly assumed that significant fine-tuning is required for them to excel at specialized knowledge, such as answering medical questions. Fine-tuning entails training an LLM on domain-specific data.

Research from Microsoft in late 2023 has overturned this assumption. This study employed prompt engineering to direct GPT-4 toward achieving remarkable performance on the MultiMedQA benchmark suite, a group of four challenging medical benchmarks (Figure 5.2.12). GPT-4 Medprompt exceeded the performance of the top 2022 model, Flan-PaLM 540B, in the multiple-choice sections of several renowned medical benchmarks, including PubMedQA, MedMCQA, and MMLU, by 3.0, 21.5, and 16.2 percentage points, respectively. It also exceeded the performance of the then state-of-the-art Med-PaLM 2 (Figure 5.2.13).

Moreover, as noted earlier, GPT-4 Medprompt was the first to surpass the 90% accuracy mark on the MedQA benchmark. This breakthrough not only underscores GPT-4 Medprompt's exceptional and potentially clinically useful medical capabilities but also demonstrates that fine-tuning may not always be necessary for adapting models to specialized domains. Prompt engineering has shown to be a promising alternative strategy.

GPT-4 vs. Med-PaLM 2 answering a medical question

Source: Nori et al., 2023

Question: A 22-year-old male marathon runner presents to the office with the complaint of right-sided rib pain when he runs long distances. Physical examination reveals normal heart and lung findings and an exhalation dysfunction at ribs 4-5 on the right. Which of the following muscles or muscle groups will be most useful in correcting this dysfunction utilizing a direct method?
(A) anterior scalene (B) latissimus dorsi (C) pectoralis minor (D) quadratus lumborum

Hand-crafted CoT from Med PaLM 2

Expert: Let's solve this step-by-step, referring to authoritative sources as needed. Among the options, only pectoralis minor muscle originates from the outer surfaces of the 3rd to 5th ribs.

GPT-4 generated CoT

GPT-4: The primary muscle involved in rib exhalation is the internal intercostal muscle. However, this option is not listed among the answer choices. Among the provided options, the pectoralis minor can contribute to rib movement, as it originates from the 3rd-5th ribs and can act to depress the ribs during exhalation when the scapula is fixed. Therefore, it could potentially be useful in correcting an exhalation dysfunction at ribs 4-5.

Figure 5.2.12

注目の研究:

GPT-4 Medprompt

LLMは印象的な一般知識を示しますが、医療に関する質問に答えるなどの専門的な知識を発揮するためには、かなりのファインチューニングが必要であると一般的に考えられています。ファインチューニングは、ドメイン特有のデータでLLMをトレーニングすることを含みます。

2023年末のマイクロソフトの研究は、この仮定を覆しました。この研究では、プロンプトエンジニアリングを用いてGPT-4をMultiMedQAベンチマークスイートにおいて驚異的なパフォーマンスを達成するように導きました。このスイートは、4つの挑戦的な医療ベンチマークのグループです（図5.2.12）。GPT-4 Medpromptは、PubMedQA、MedMCQA、MMLUを含むいくつかの著名な医療ベンチマークの選択肢セクションで、2022年のトップモデルであるFlan-PaLM 540Bのパフォーマンスをそれぞれ3.0、21.5、16.2パーセントポイント上回りました。また、当時の最先端であったMed-PaLM 2のパフォーマンスも上回りました（図5.2.13）。

さらに、前述のように、GPT-4 MedpromptはMedQAベンチマークで90%の精度を超えた最初のモデルでした。このブレークスルーは、GPT-4 Medpromptの卓越した、そして臨床的に有用な医療能力を強調するだけでなく、モデルを専門分野に適応させるためにファインチューニングが常に必要ではないことを示しています。プロンプトエンジニアリングは、有望な代替戦略であることが示されています。

GPT-4とMed-PaLM 2の医療質問への回答

出典: Nori et al., 2023

Question: A 22-year-old male marathon runner presents to the office with the complaint of right-sided rib pain when he runs long distances. Physical examination reveals normal heart and lung findings and an exhalation dysfunction at ribs 4-5 on the right. Which of the following muscles or muscle groups will be most useful in correcting this dysfunction utilizing a direct method?
(A) anterior scalene (B) latissimus dorsi (C) pectoralis minor (D) quadratus lumborum

Hand-crafted CoT from Med PaLM 2

Expert: Let's solve this step-by-step, referring to authoritative sources as needed. Among the options, only pectoralis minor muscle originates from the outer surfaces of the 3rd to 5th ribs.

GPT-4 generated CoT

GPT-4: The primary muscle involved in rib exhalation is the internal intercostal muscle. However, this option is not listed among the answer choices. Among the provided options, the pectoralis minor can contribute to rib movement, as it originates from the 3rd-5th ribs and can act to depress the ribs during exhalation when the scapula is fixed. Therefore, it could potentially be useful in correcting an exhalation dysfunction at ribs 4-5.

Figure 5.2.12

Highlighted Research: GPT-4 Medprompt (cont'd)

Model performance on MultiMedQA sub-benchmarks

Source: Nori et al., 2023 | Chart: 2024 AI Index report

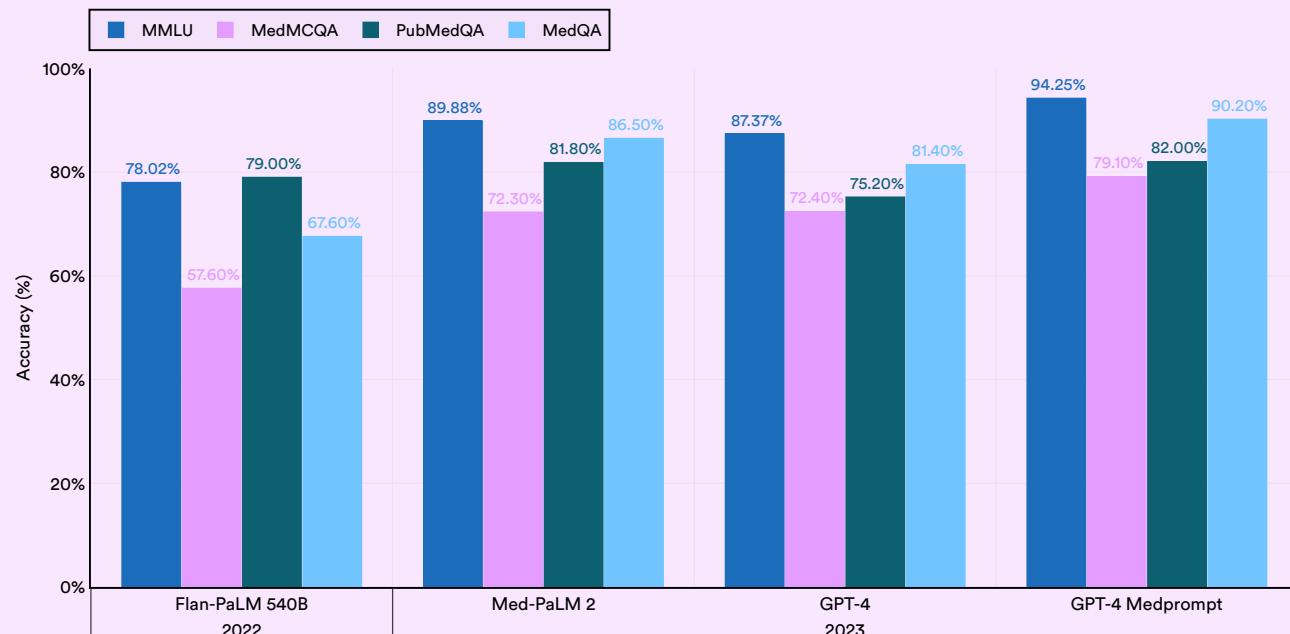


Figure 5.2.13

注目の研究: GPT-4 メドプロンプト（続き）

マルチメドQAサブベンチマークにおけるモデルのパフォーマンス 出典: Nori et al.,

2023 デチャート: 2024 AI インデックスレポート

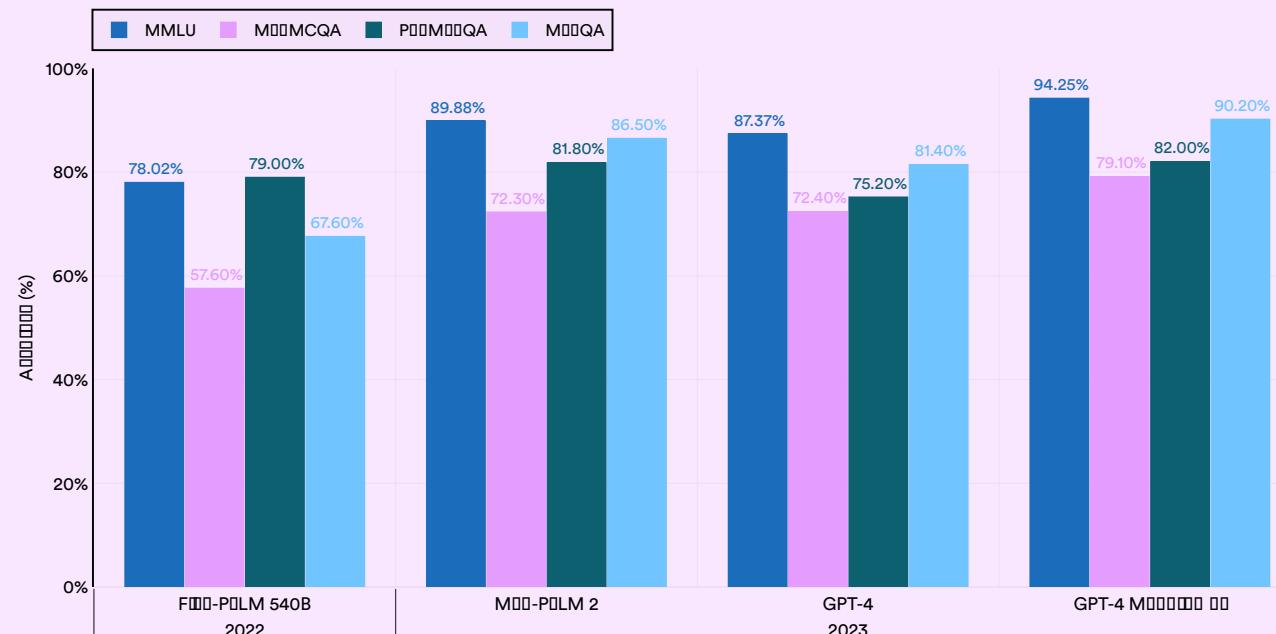


Figure 5.2.13

Highlighted Research: MediTron-70B

GPT-4 Medprompt is an impressive system; however, it is closed-source, meaning its weights are not freely available to the broader public for use. New research in 2023 has also sought to advance the capabilities of open-source medical LLMs. Among this new research, MediTron-70B stands out as particularly promising. This model achieves a respectable 70.2% accuracy on the MedQA benchmark. Although this is below the performance of GPT-4 Medprompt and Med-

PaLM 2 (both closed models), it represents a significant improvement over the state-of-the-art results from 2023 and surpasses other open-source models like Llama 2 (Figure 5.2.14). MediTron-70B's score on MedQA is the highest yet achieved by an open-source model. If medical AI is to reach its fullest potential, it is important that its capabilities are widely accessible. In this context, MediTron represents an encouraging step forward.

Performance of select models on MedQA

Source: Chen et al., 2023 | Table: 2024 AI Index report

Model	Release date	Access type	Score on MedQA
GPT-4 Medprompt	November 2023	Closed	90.20%
Med-PaLM 2	April 2023	Closed	86.20%
MediTron-70B	November 2023	Open	70.20%
Med-PaLM	December 2022	Closed	67.20%
Llama 2	July 2023	Open	63.80%

Figure 5.2.14

注目の研究: MediTron-70B

GPT-4 Medpromptは印象的なシステムですが、クローズドソースであるため、その重みは一般の人々が自由に使用するためには利用可能ではありません。2023年の新しい研究は、オープンソースの医療LLMの能力を向上させることを目指しています。この新しい研究の中で、MediTron-70Bは特に有望な存在として際立っています。このモデルは、MedQAベンチマークで70.2%の尊敬すべき精度を達成しています。これは、GPT-4 MedpromptやMed-

PaLM 2（どちらもクローズドモデル）よりは劣りますが、2023年の最先端の結果に対して重要な改善を示しており、Llama 2のような他のオープンソースモデルを上回っています（図5.2.14）。MediTron-70BのMedQAでのスコアは、オープンソースモデルによって達成された中で最高のものです。医療AIがその最大の潜在能力に達するためには、その能力が広くアクセス可能であることが重要です。この文脈において、MediTronは前向きな一步を示しています。

選択したモデルのMedQAにおけるパフォーマンス

出典: Chen et al., 2023 | 表: 2024 AIインデックスレポート

モデル	リリース日	アクセスタイプ	MedQAスコア
GPT-4 Medprompt	2023年11月	閉鎖	90.20%
Med-PaLM 2	2023年4月	閉鎖	86.20%
MediTron-70B	2023年11月	開放	70.20%
Med-PaLM	2022年12月	閉鎖	67.20%
Llama 2	2023年7月	開放	63.80%

Figure 5.2.14

Diagnosis

AI tools can also be used for diagnostic purposes including, for example, in radiology or cancer detection.

Highlighted Research:

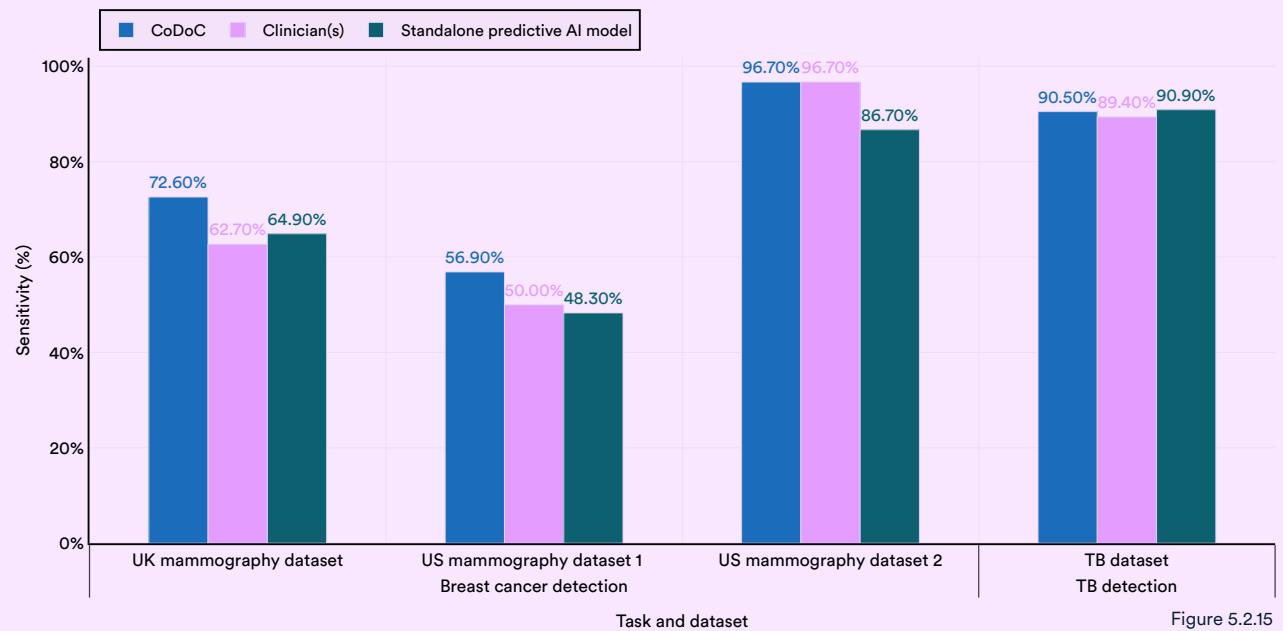
CoDoC

AI medical imaging systems demonstrate robust diagnostic capabilities, yet there are instances where they overlook diagnoses that clinicians catch, and vice versa. This observation suggests a logical integration of AI systems and clinicians' diagnostic abilities. In 2023, researchers unveiled CoDoC (Complementarity-Driven Deferral to Clinical Workflow), a system designed to discern when to rely on AI for diagnosis and when to defer to traditional clinical methods. CoDoC notably enhances both sensitivity (the ability to correctly identify individuals with a disease) and specificity

(the ability to accurately identify those without it). Specifically, across four medical datasets, CoDoC's sensitivity surpasses clinicians' by an average of 4.5 percentage points and a standalone AI model's by 6.5 percentage points (Figure 5.2.15). In terms of specificity, CoDoC outperforms clinicians by an average of 2.7 percentage points across tested datasets and a standalone predictive model by 5.7 percentage points. Moreover, CoDoC has been shown to reduce clinical workflow by 66%. These findings suggest that AI medical systems can be integrated into clinical workflows, thereby enhancing diagnostic accuracy and efficiency.

CoDoC vs. standalone predictive AI system and clinical readers: sensitivity

Source: Dvijotham et al., 2023 | Chart: 2024 AI Index report



診断

AI tools can also be 放射線学や癌検出などの診断目的で使用される。

注目の研究:

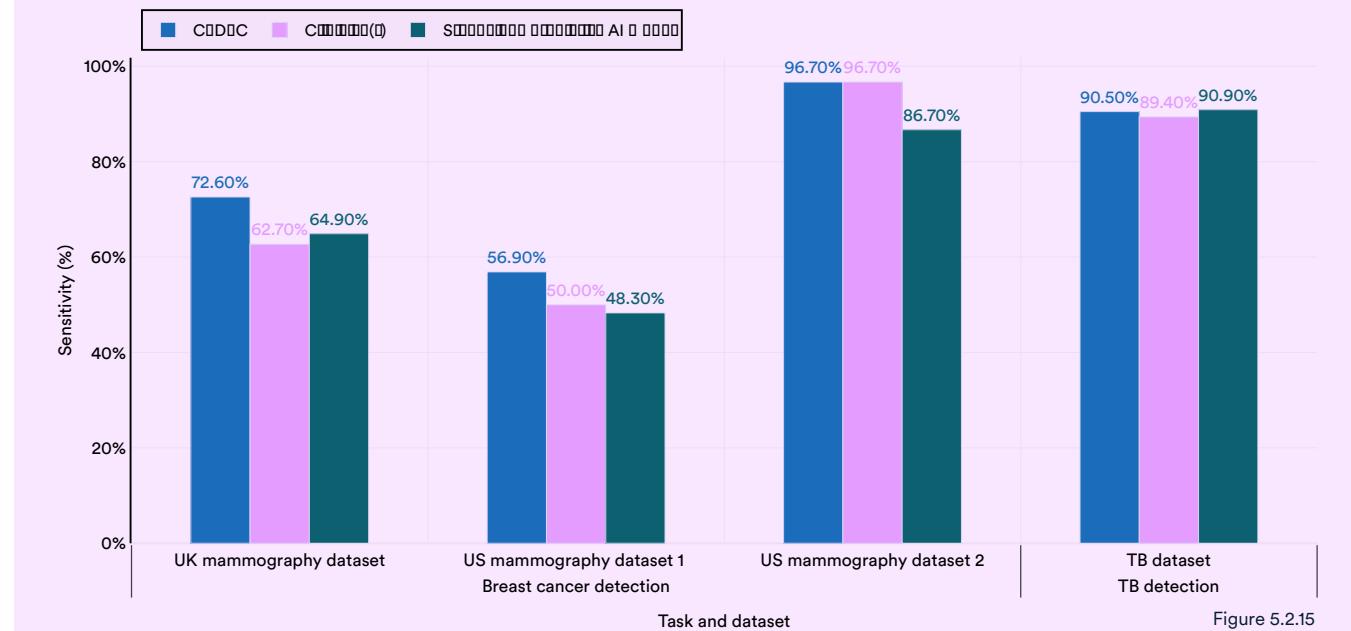
CoDoC

AI 医療画像システムは堅牢な診断能力を示していますが、臨床医が見逃す診断をAIが見逃す場合やその逆もあります。この観察は、AIシステムと臨床医の診断能力の論理的統合を示唆しています。2023年、研究者たちはCoDoC（補完性駆動型臨床ワークフローへの委譲）を発表しました。このシステムは、診断のためにAIに依存すべき時と従来の臨床方法に委譲すべき時を識別するように設計されています。CoDoCは、感度（病気のある個人を正しく特定する能力）と特異度を両方とも著しく向上させます。

特に、4つの医療データセットにおいて、CoDoCの感度は臨床医を平均4.5パーセントポイント、単独のAIモデルを6.5パーセントポイント上回ります（図5.2.15）。特異度に関しては、CoDoCはテストされたデータセット全体で臨床医を平均2.7パーセントポイント、単独の予測モデルを5.7パーセントポイント上回ります。さらに、CoDoCは臨床ワークフローを66%削減することができます。これらの発見は、AI医療システムが臨床ワークフローに統合され、診断の正確性と効率を向上させることができることを示唆しています。

CoDoC対独立した予測AIシステムと臨床リーダー: 感度

出典: Dvijotham et al., 2023 | チャート: 2024 AI Index report

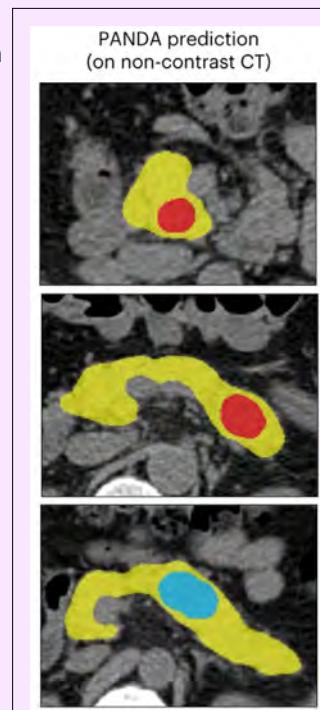


Highlighted Research:

CT Panda

Pancreatic ductal adenocarcinoma (PDAC) is a particularly lethal cancer, often detected too late for surgical intervention. Screening for PDAC in asymptomatic individuals is challenging due to its low prevalence and the risk of false positives. This year, a Chinese research team developed PANDA (pancreatic cancer detection with artificial intelligence), an AI model capable of efficiently detecting and classifying pancreatic lesions in X-rays (Figure 5.2.16). In validation tests, PANDA surpassed the average radiologist in sensitivity by 34.1% and in specificity by 6.3% (Figure 5.2.17). In a large-scale, real-world test involving approximately 20,000 patients, PANDA achieved a sensitivity of 92.9% and a specificity of 99.9% (Figure 5.2.18). AI medical tools like PANDA represent significant advancements in diagnosing challenging conditions, offering cost-effective and accurate detection previously considered difficult or prohibitive.

PANDA detection
Source: Cao et al., 2023
Figure 5.2.16



PANDA vs. mean radiologist on multicenter validation (6,239 patients)

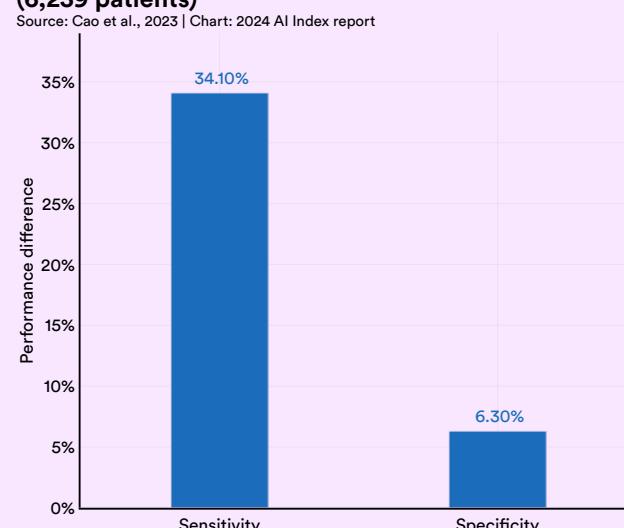


Figure 5.2.17

PANDA performance on real-world multi-scenario validation (20,530 patients)

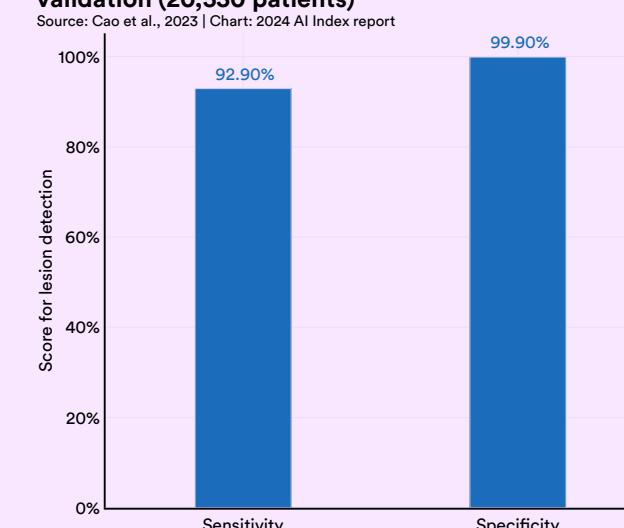
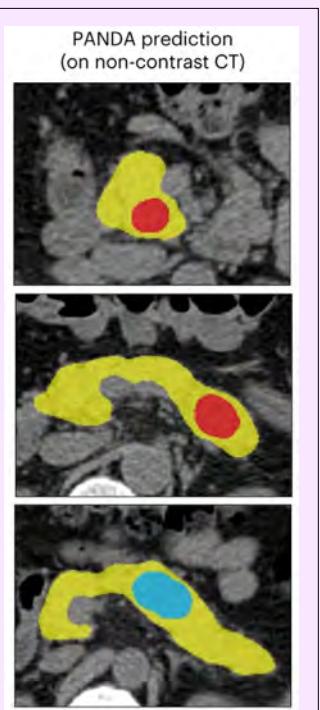


Figure 5.2.18

注目の研究: CTパンダ

膵管腺癌（PDAC）は特に致死的な癌であり、外科的介入のために遅すぎる時期に発見されることが多い。無症状の個人におけるPDACのスクリーニングは、その低い有病率と偽陽性のリスクのために困難である。今年、中国の研究チームはPANDA（人工知能による膵臓癌検出を開発し、X線で膵臓病変を効率的に検出・分類することができるAIモデルを作成した（図 5.2.16）。検証テストでは、PANDAは感度で平均放射線科医を34.1%、特異度で6.3%上回った（図 5.2.17）。約 20,000人の患者を対象とした大規模な実世界テストでは、PANDAは感度 92.9%、特異度 99.9%を達成した（図 5.2.18）。PANDAのようなAI 医療ツールは、困難な状態の診断において重要な進展を示しており、以前は困難または高価と考えられていたコスト効果的で正確な検出を提供している。

図 5.2.
16 PANDA
検出
出典: Cao et al., 2023



出典: Cao et al., 2023 | チャート: 2024 AI Index report | **PANDAと多施設検証における平均放射線科医（6,239人の患者）**

出典: Cao et al., 2023 | チャート: 2024 AI Index report | **PANDAの実世界のマルチシナリオ検証におけるパフォーマンス（20,530人の患者）**

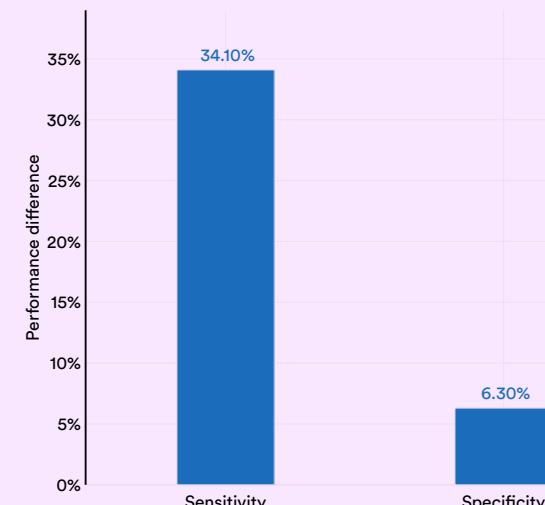


Figure 5.2.17

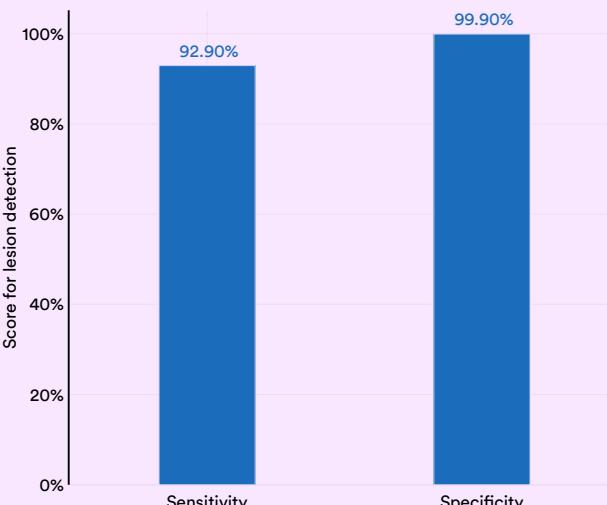


Figure 5.2.18

Other Diagnostic Uses

New research published in 2023 highlights how AI can be used in other diagnostic contexts. Figure 5.2.19 summarizes some of the findings.

Additional research on diagnostic AI use cases

Source: AI Index, 2024

Research	Use case	Findings
Schopf et al., 2023	Breast cancer	The authors conducted a meta-review of the literature exploring mammography-image-based AI algorithms. They discovered that predicting future breast cancer risk using only mammography images achieves accuracy that is comparable to or better than traditional risk assessment tools.
Dicente Cid et al., 2023	X-ray interpretation	The researchers developed two open-source neural networks, X-Raydar and X-Raydar-NLP, for classifying chest X-rays using images and free-text reports. They found that these automated classification methods perform at levels comparable to human experts and demonstrate robustness when applied to external data sets.

Figure 5.2.19

その他の診断用途

2023年に発表された新しい研究は、AIが他の診断コンテキストでどのように使用できるかを強調しています。図5.2.19は、いくつかの発見を要約しています。

診断AIの使用事例に関する追加研究

AI Index, 2024

Research	Use case	Findings
Schopf et al., 2023	Breast cancer	The authors conducted a meta-review of the literature exploring mammography-image-based AI algorithms. They discovered that predicting future breast cancer risk using only mammography images achieves accuracy that is comparable to or better than traditional risk assessment tools.
Dicente Cid et al., 2023	X-ray interpretation	The researchers developed two open-source neural networks, X-Raydar and X-Raydar-NLP, for classifying chest X-rays using images and free-text reports. They found that these automated classification methods perform at levels comparable to human experts and demonstrate robustness when applied to external data sets.

図5.2.19

FDA-Approved AI-Related Medical Devices

The U.S. Food and Drug Administration (FDA) maintains a [list](#) of AI/ML-enabled medical devices that have received approval. The devices featured on this list meet the FDA's premarket standards, which include a detailed review of their effectiveness and safety. As of October 2023, the FDA has not approved any devices that utilize generative AI or are powered by LLMs.

Figure 5.2.20 illustrates the number of AI medical devices approved by the FDA over the past decade. In 2022, a total of 139 AI-related medical devices received FDA approval, marking a 12.1% increase from the total approved in 2021. Since 2012, the number of these devices has increased by more than 45-fold.

Number of AI medical devices approved by the FDA, 2012–22

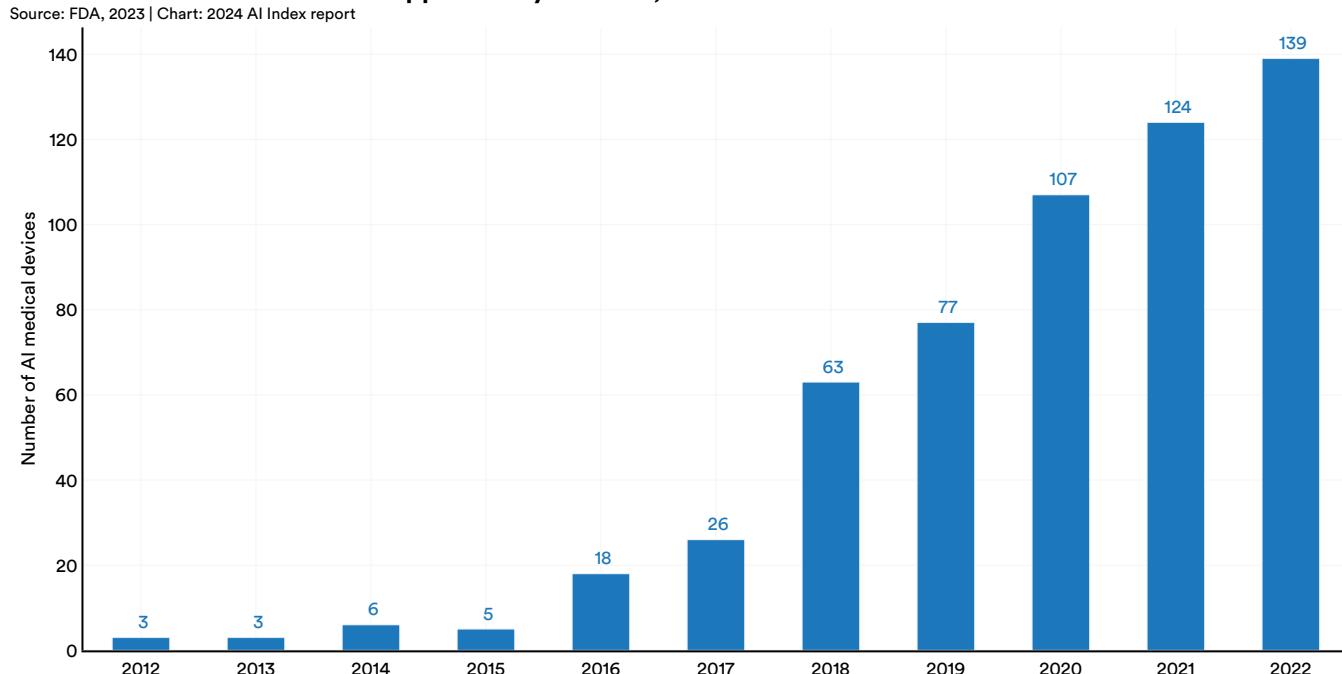


Figure 5.2.20

3 The FDA last updated the list in October 2023, meaning that the totals for 2023 were incomplete. Consequently, the AI Index limited its data presentation to include only information up to 2022.

FDA 承認 のAI 関連医療機器

米国食品医薬品局（FDA）は、承認を受けたAI/ML対応の医療機器のリストを維持しています。このリストに掲載されている機器は、FDAの市場前基準を満たしており、その効果と安全性の詳細なレビューが含まれています。2023年10月現在、FDAは生成AIを利用したり、LLMによって動作する機器を承認していません。

図5.2.20は、過去10年間にFDAによって承認されたAI医療機器の数を示しています。2022年には、合計139のAI関連医療機器がFDAの承認を受け、2021年に承認された総数から12.1%の増加を記録しました。2012年以降、これらの機器の数は45倍以上に増加しています。

2012年から2022年までにFDAに承認されたAI医療機器の数

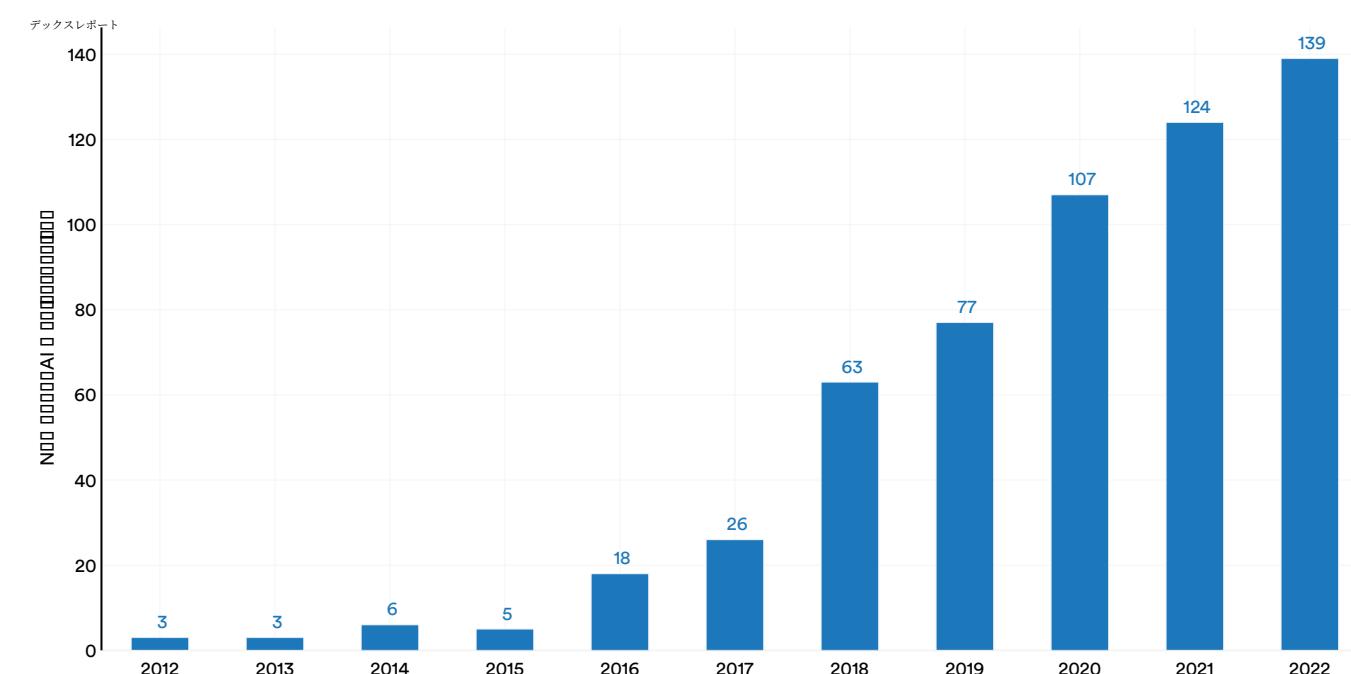


Figure 5.2.20

3 The FDA last updated the list in October 2023, meaning that the data for 2023 is incomplete. Therefore, the AI Index limited its data presentation to include only information up to 2022.

Figure 5.2.21 illustrates the specialties associated with FDA-approved medical devices. Of the 139 devices approved in 2022, a significant majority, 87.1%, were related to radiology. The next most common specialty was cardiovascular, accounting for 7.2% of the approvals.

Number of AI medical devices approved by the FDA by specialty, 2012–22

Source: FDA, 2023 | Chart: 2024 AI Index report

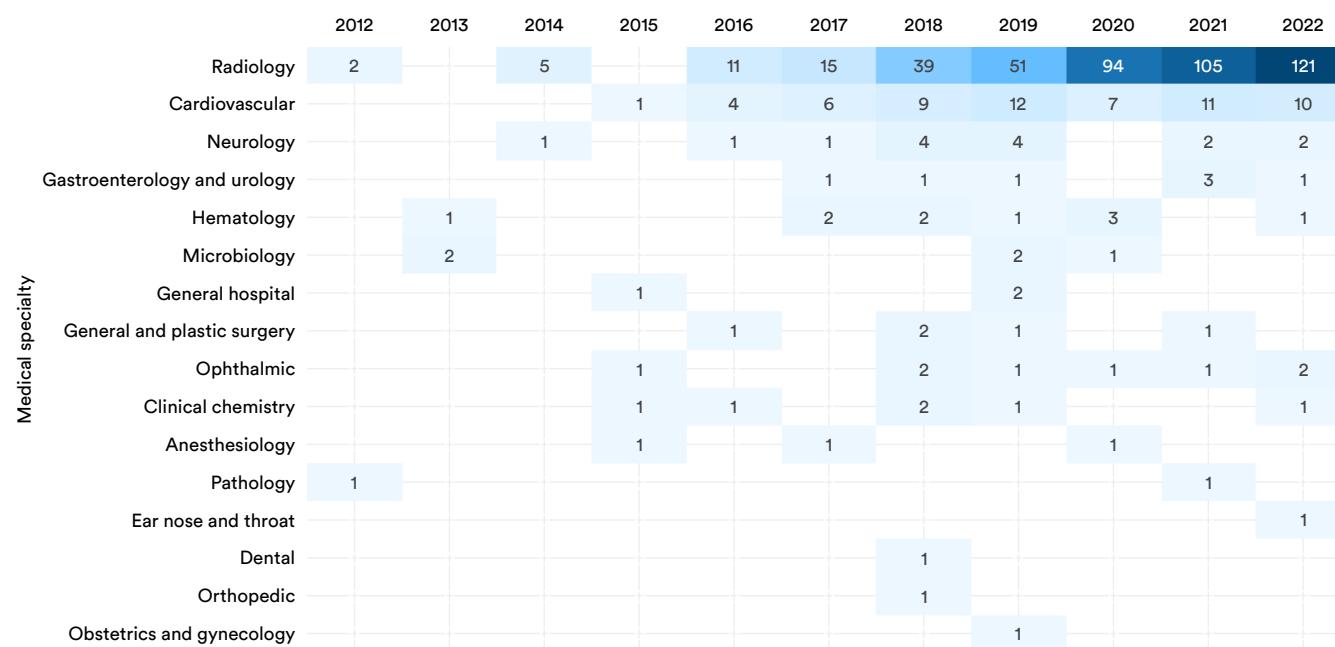


Figure 5.2.21

2012年から2022年までの専門分野別にFDAに承認されたAI医療機器の数

出典: FDA, 2023 | チャート: 2024年 AIインデックスレポート

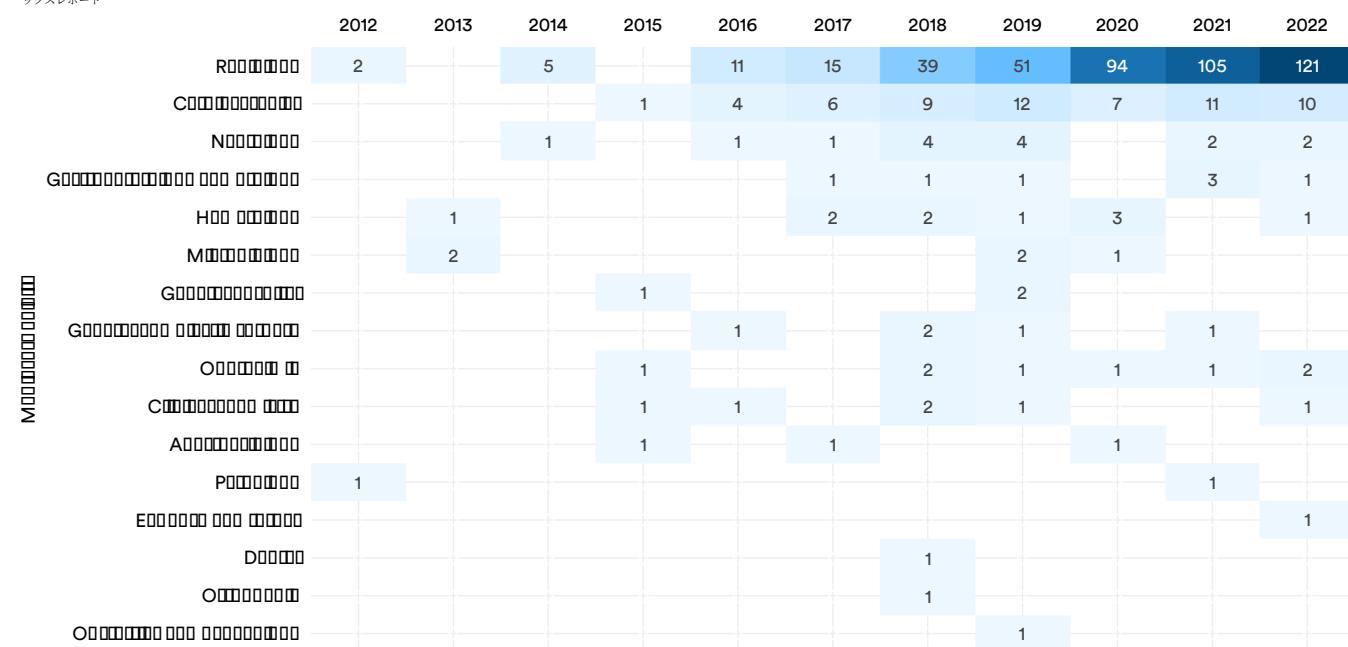


図 5.2.21

Administration and Care

AI tools also hold the potential to enhance medical administration efficiency and elevate the standard of patient care.

Highlighted Research:

MedAlign

Despite significant advances in AI for healthcare, existing benchmarks like MedQA and USMLE, focused on knowledge-based questions, do not fully capture the diverse tasks clinicians perform in patient care. Clinicians often engage in information-intensive tasks, such as creating tailored diagnostic plans, and spend a significant proportion of their working hours on administrative tasks. Although AI has the potential to streamline these processes, there is a lack of suitable electronic health records (EHR) datasets for benchmarking and fine-tuning medically administrative LLMs. This year researchers have made strides to address this gap by introducing MedAlign: a comprehensive EHR-based

benchmark with 983 questions and instructions and 303 clinician responses, drawn from seven different medical specialties (Figure 5.2.22). MedAlign is the first extensive EHR-focused benchmark.

The researchers then tested various existing LLMs on MedAlign. Of all LLMs, a GPT-4 variant using multistep refinement achieved the highest correctness rate (65.0%) and was routinely preferred over other LLMs (Figure 5.2.23). MedAlign is a valuable milestone toward using AI to alleviate administrative burdens in healthcare.

MedAlign workflow

Source: Fleming et al., 2023

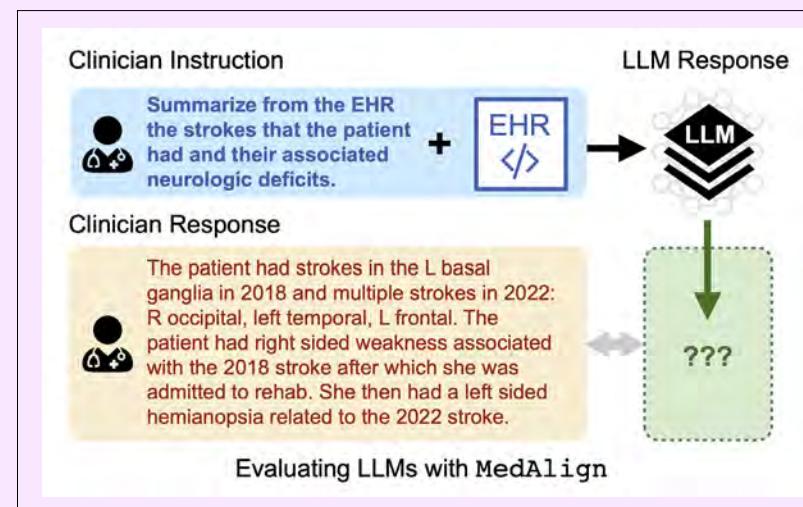


Figure 5.2.22

管理とケア

AI tools also hold the potential to enhance medical administration efficiency and elevate the standard of patient care.

注目の研究:

MedAlign

AIが医療分野で著しい進展を遂げているにもかかわらず、MedQAやUSMLEのような既存のベンチマークは、知識ベースの質問に焦点を当てており、臨床医が患者ケアで行う多様なタスクを完全には捉えていません。臨床医は、個別の診断計画を作成するなどの情報集約型のタスクに従事し、業務時間のかなりの割合を管理業務に費やしています。AIはこれらのプロセスを効率化する可能性を秘めていますが、医療管理用のLLMをベンチマークおよび微調整するための適切な電子健康記録(EHR)データセットが不足しています。今年、研究者たちはこのギャップに対処するために、包括的なEHRベースのMedAlignを導入することで進展を遂げました。

983の質問と指示、7つの異なる医療専門分野からの303の臨床医の回答を含むベンチマーク(図5.2.22)。MedAlignは、最初の包括的なEHRに焦点を当てたベンチマークです。

研究者たちは次に、MedAlignでさまざまな既存のLLMをテストしました。すべてのLLMの中で、マルチステップの洗練を使用したGPT-4のバリアントが最も高い正確性(65.0%)を達成し、他のLLMよりも頻繁に好まれました(図5.2.23)。MedAlignは、医療における管理負担を軽減するためにAIを使用するための貴重なマイルストーンです。

MedAlignワークフロー

出典: Fleming et al., 2023

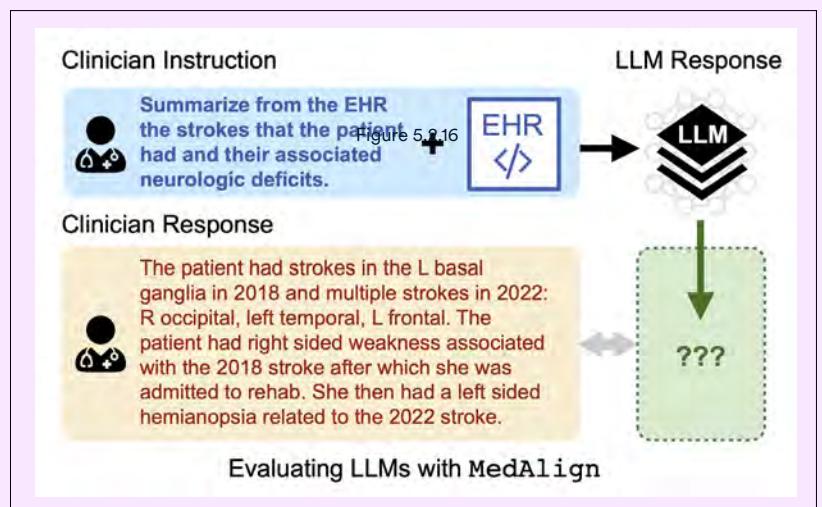


図5.2.22

Highlighted Research: MedAlign (cont'd)

Evaluation of model performance: human vs. COMET ranks

Source: Fleming et al., 2023 | Chart: 2024 AI Index report

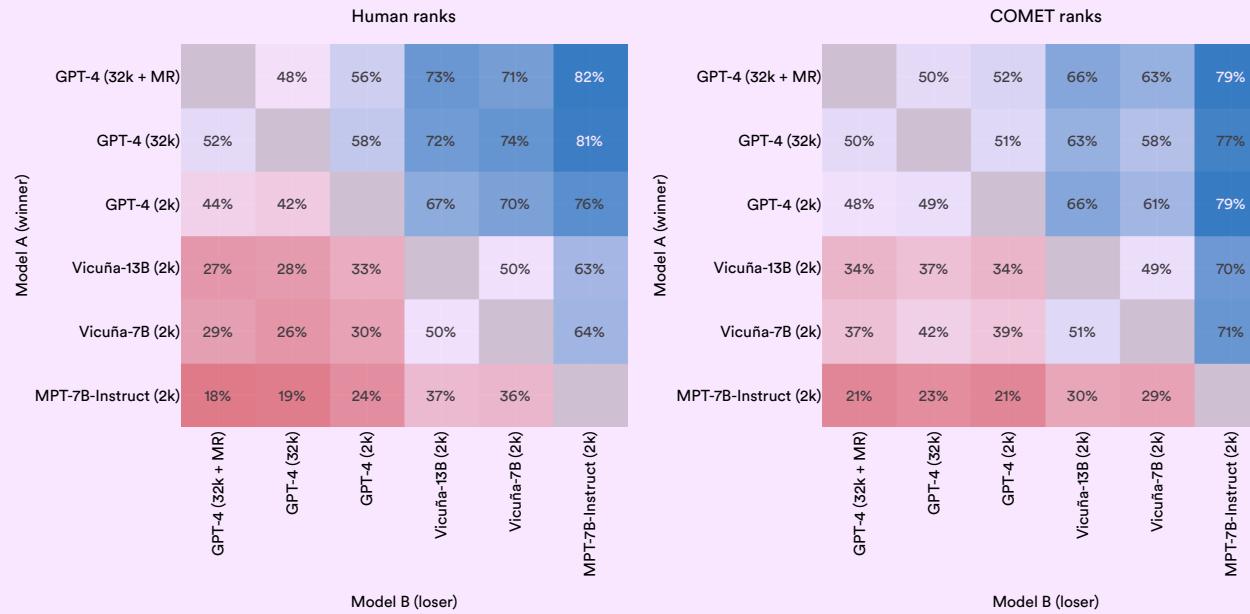


Figure 5.2.23

注目の研究: MedAlign (続き)

モデルのパフォーマンス評価: 人間 vs. COMETランク

出典: Fleming et al., 2023 | チャート: 2024 AI インデックスレポート

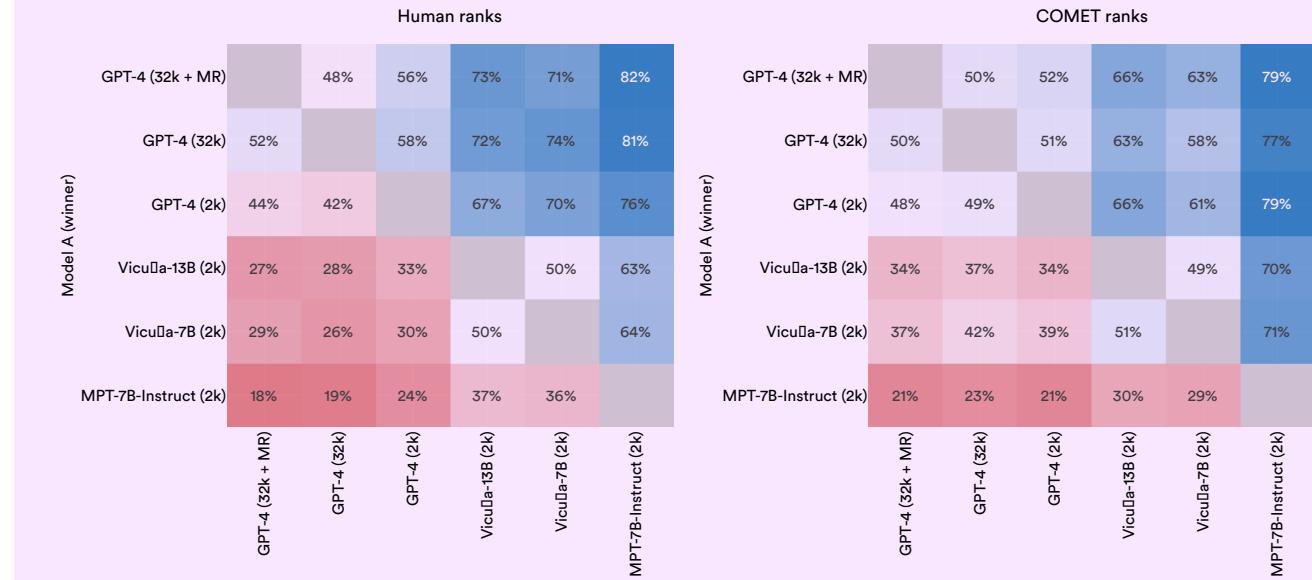
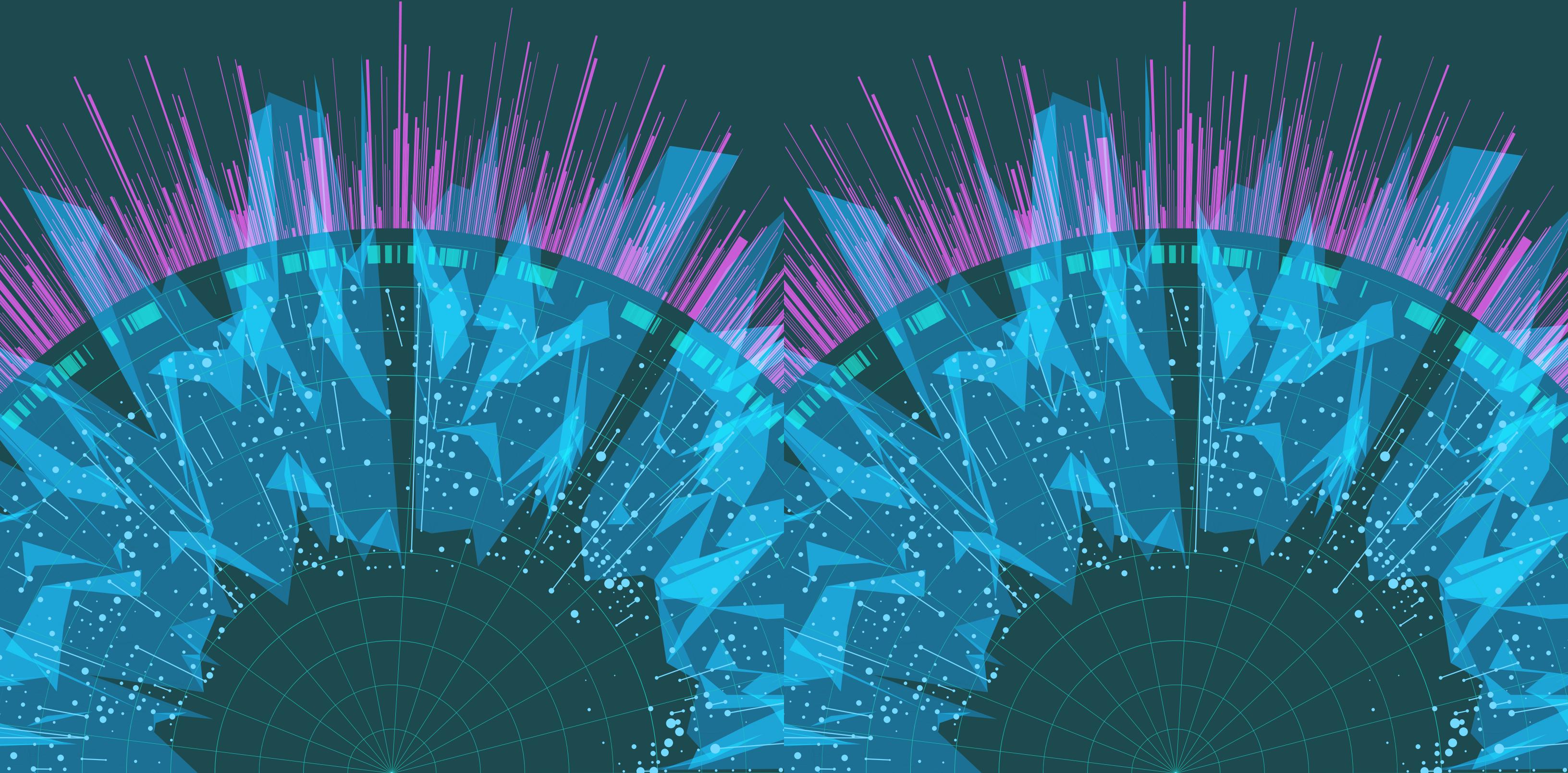


Figure 5.2.23

CHAPTER 6:
Education

第6章:
Education



Preview

Overview	327
Chapter Highlights	328
6.1 Postsecondary CS and AI Education	329
United States and Canada	329
CS Bachelor's Graduates	329
CS Master's Graduates	331
CS PhD Graduates	333
CS, CE, and Information Faculty	336
Europe	344
Informatics, CS, CE, and IT Bachelor's Graduates	344
Informatics, CS, CE, and IT Master's Graduates	347
Informatics, CS, CE, and IT PhD Graduates	351
AI-Related Study Programs	355
Total Courses	355
Education Level	356
Geographic Distribution	357
6.2 K-12 CS and AI Education	359
United States	359
State-Level Trends	359
AP Computer Science	361
Highlight: Access Issues	363
Highlight: ChatGPT Usage Among Teachers and Students	364

[ACCESS THE PUBLIC DATA](#)

レビュー

概要 327 章 のハイライト 328

6.1 Postsecondary CS and AI Education	329
United States and Canada	329
CS Bachelor's Graduates	329
CS Master's Graduates	331
CS PhD Graduates	333
CS, CE, and Information Faculty	336
Europe	344
Informatics, CS, CE, and IT Bachelor's Graduates	344
Informatics, CS, CE, and IT Master's Graduates	347
Informatics, CS, CE, and IT PhD Graduates	351
AI-Related Study Programs	355
Total Courses	355
Education Level	356
Geographic Distribution	357

6.2 K-12 CS および AI 教育 359

United States	359
State-Level Trends	359
AP Computer Science	361
Highlight: Access Issues	363
Highlight: ChatGPT Usage Among Teachers and Students	364

[ACCESS THE PUBLIC DATA](#)

Overview

This chapter examines trends in AI and computer science (CS) education, focusing on who is learning, where they are learning, and how these trends have evolved over time. Amid growing concerns about AI's impact on education, it also investigates the use of new AI tools like ChatGPT by teachers and students.

The analysis begins with an overview of the state of postsecondary CS and AI education in the United States and Canada, based on the Computing Research Association's annual Taulbee Survey. It then reviews data from Informatics Europe regarding CS education in Europe. This year introduces a new section with data from Studyportals on the global count of AI-related English-language study programs.

The chapter wraps up with insights into K-12 CS education in the United States from Code.org and findings from the Walton Foundation survey on ChatGPT's use in schools.

概要

この章では、AIとコンピュータサイエンス（CS）教育のトレンドを検討し、誰が学んでいるのか、どこで学んでいるのか、そしてこれらのトレンドがどのように進化してきたのかに焦点を当てています。AIが教育に与える影響についての懸念が高まる中、教師や学生によるChatGPTなどの新しいAIツールの使用についても調査しています。

この分析は、Computing Research Associationの年次タウルビー調査に基づいて、アメリカとカナダにおける高等教育のCSとAI教育の現状の概要から始まります。その後、ヨーロッパにおけるCS教育に関するInformatics Europeのデータをレビューします。今年は、AI関連の英語の学位プログラムの世界的な数に関するStudyportalsのデータを含む新しいセクションが導入されます。

この章は、Code.orgからのアメリカのK-12コンピュータサイエンス教育に関する洞察と、学校におけるChatGPTの使用に関するウォルトン財団の調査結果で締めくられています。

Chapter Highlights

1. The number of American and Canadian CS bachelor's graduates continues to rise, new CS master's graduates stay relatively flat, and PhD graduates modestly grow.

While the number of new American and Canadian bachelor's graduates has consistently risen for more than a decade, the number of students opting for graduate education in CS has flattened. Since 2018, the number of CS master's and PhD graduates has slightly declined.

2. The migration of AI PhDs to industry continues at an accelerating pace.

In 2011, roughly equal percentages of new AI PhDs took jobs in industry (40.9%) and academia (41.6%). However, by 2022, a significantly larger proportion (70.7%) joined industry after graduation compared to those entering academia (20.0%). Over the past year alone, the share of industry-bound AI PhDs has risen by 5.3 percentage points, indicating an intensifying brain drain from universities into industry.

3. Less transition of academic talent from industry to academia.

In 2019, 13% of new AI faculty in the United States and Canada were from industry. By 2021, this figure had declined to 11%, and in 2022, it further dropped to 7%. This trend indicates a progressively lower migration of high-level AI talent from industry into academia.

4. CS education in the United States and Canada becomes less international.

Proportionally fewer international CS bachelor's, master's, and PhDs graduated in 2022 than in 2021. The drop in international students in the master's category was especially pronounced.

5. More American high school students take CS courses, but access problems remain.

In 2022, 201,000 AP CS exams were administered. Since 2007, the number of students taking these exams has increased more than tenfold. However, recent evidence indicates that students in larger high schools and those in suburban areas are more likely to have access to CS courses.

6. AI-related degree programs are on the rise internationally.

The number of English-language, AI-related postsecondary degree programs has tripled since 2017, showing a steady annual increase over the past five years. Universities worldwide are offering more AI-focused degree programs.

7. The United Kingdom and Germany lead in European informatics, CS, CE, and IT graduate production.

The United Kingdom and Germany lead Europe in producing the highest number of new informatics, CS, CE, and information bachelor's, master's, and PhD graduates. On a per capita basis, Finland leads in the production of both bachelor's and PhD graduates, while Ireland leads in the production of master's graduates.

章 の ハイ ライ ト

1. アメリカとカナダのコンピュータサイエンス 学士卒業生の数は引き続き増加しており、新しいコンピュータサイエンス 修士卒業生は比較的横ばいで、博士卒業生はわずかに増加しています。その間、

新しいアメリカとカナダの学士卒業生の数は10年以上にわたり一貫して増加していますが、コンピュータサイエンスの大学院教育を選択する学生の数は横ばいです。2018年以降、コンピュータサイエンスの修士および博士卒業生の数はわずかに減少しています。

2. AI 博士の産業への移行は加速しています。2011年には、新しいAI 博士の約同じ割合が産業 (40.9%) と学術 (41.6%) で職を得ていました。しかし、2022年までに、卒業後に産業に参加した割合は大幅に増加し (70.7%)、学術に入った割合 (20.0%) と比較して大きな差が生じました。過去1年だけでも、産業に向かうAI 博士の割合は5.3ポイント上昇しており、大学から産業への脳の流出が激化していることを示しています。

3. 学術界への産業からの学術人材の移行が減少。2019年、アメリカとカナダの新しいAI 教員の13%が産業から来ていました。2021年にはこの数字は11%に減少し、2022年にはさらに7%に落ち込みました。この傾向は、産業から学術界への高レベルのAI 人材の移行が徐々に低下していることを示しています。

4. アメリカとカナダのCS 教育が国際的でなくなる。2022年には、2021年よりも国際的なCSの学士、修士、博士の卒業生が相対的に減少しました。特に修士課程の国際学生の減少が顕著でした。

5. より多くのアメリカの高校生がCSコースを受講するが、アクセスの問題が残る。2022年には、201,000件のAP CS 試験が実施されました。2007年以降、これらの試験を受ける学生の数は10倍以上に増加しました。しかし、最近の証拠によると、大規模な高校や郊外に住む学生はCSコースにアクセスできる可能性が高いことが示されています。

6. AI 関連の学位プログラムが国際的に増加している。英語で提供されるAI 関連の高等教育学位プログラムの数は2017年以降3倍に増加し、過去5年間にわたって安定した年次増加を示しています。世界中の大学がより多くのAIに焦点を当てた学位プログラムを提供しています。

7. イギリスとドイツがヨーロッパの情報学、CS、CE、ITの卒業生生産でリードしている。イギリスとドイツは、情報学、CS、CE、情報の学士、修士、博士の新卒業生を最も多く生産しているヨーロッパの国です。人口あたりでは、フィンランドが学士と博士の卒業生の生産でリードしており、アイルランドが修士の卒業生の生産でリードしています。

This section provides an overview of postsecondary education in CS and AI, highlighting graduation statistics across North America and Europe for various degrees including bachelor's, master's, and PhDs. It also covers information on AI-related courses offered in English.

6.1 Postsecondary CS and AI Education

United States and Canada

This subsection presents an analysis of data from the Computing Research Association's [Taulbee Survey](#), which evaluates the state of CS and AI postsecondary education in the United States and Canada. The survey covers 297 PhD-granting CS departments across the United States and Canada.¹

CS Bachelor's Graduates

Over the past decade, the total number of new CS bachelor's graduates in North America has steadily risen, increasing more than threefold, with a 7.9% year-over-year rise from 2021 to 2022 (Figure 6.1.1).

New CS bachelor's graduates in the United States and Canada, 2010–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

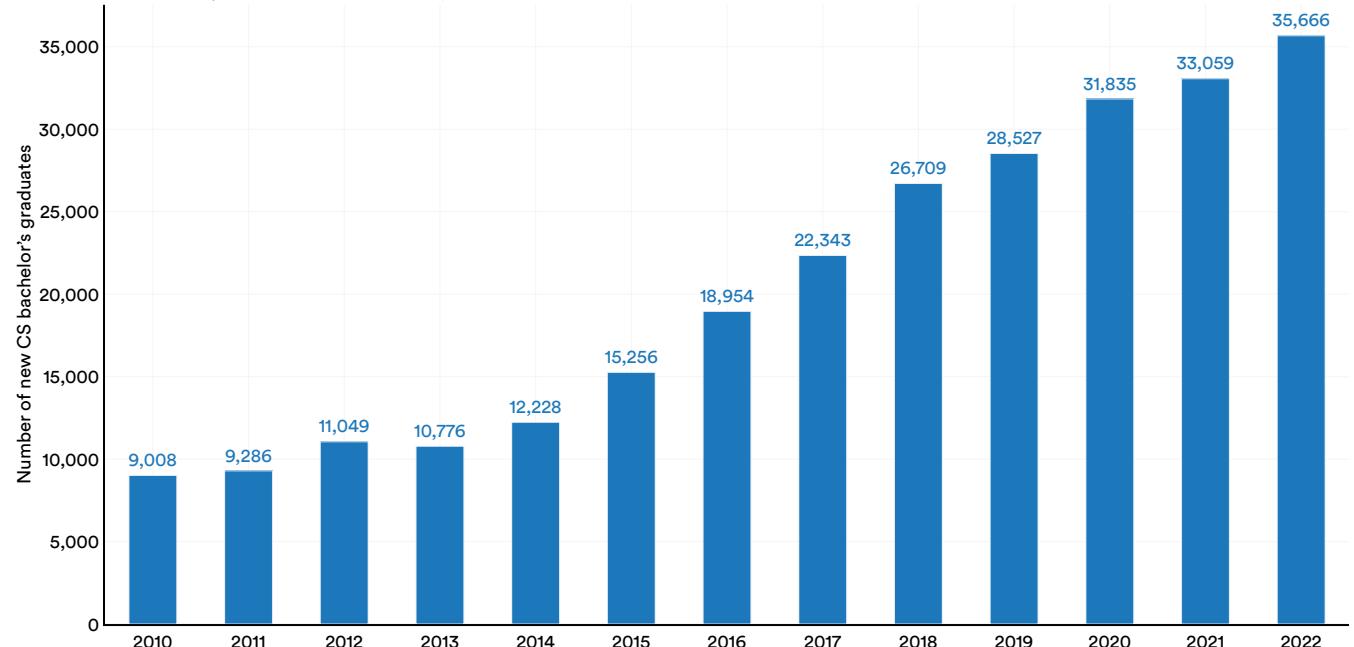


Figure 6.1.1

¹ It is important to note that not all PhD-granting departments targeted in the survey provided responses. Out of the 297 departments targeted, only 182 responded, yielding an overall response rate of 61%.

このセクションでは、コンピュータサイエンスと人工知能における高等教育の概要を提供し、北米とヨーロッパにおける学士、修士、博士などのさまざまな学位の卒業統計を強調しています。また、英語で提供されるAI関連のコースに関する情報も含まれています。

6.1 高等教育におけるコンピュータサイエンスと人工知能

アメリカ合衆国とカナダ

この小節では、アメリカ合衆国とカナダにおけるコンピュータサイエンスおよびAIの高等教育の状況を評価するタウルビー調査からのデータの分析を提示します。この調査は、アメリカ合衆国とカナダの297の博士号授与コンピュータサイエンス学科をカバーしています。¹

コンピュータサイエンスの学士卒業生

過去10年間で、北米における新しいコンピュータサイエンスの学士卒業生の総数は着実に増加し、2021年から2022年にかけては前年比7.9%の増加を記録し、3倍以上に増加しました（図6.1.1）。

アメリカ合衆国とカナダの新しいコンピュータサイエンス学士卒業生、2010–22 出典:CRAタウルビー調査、

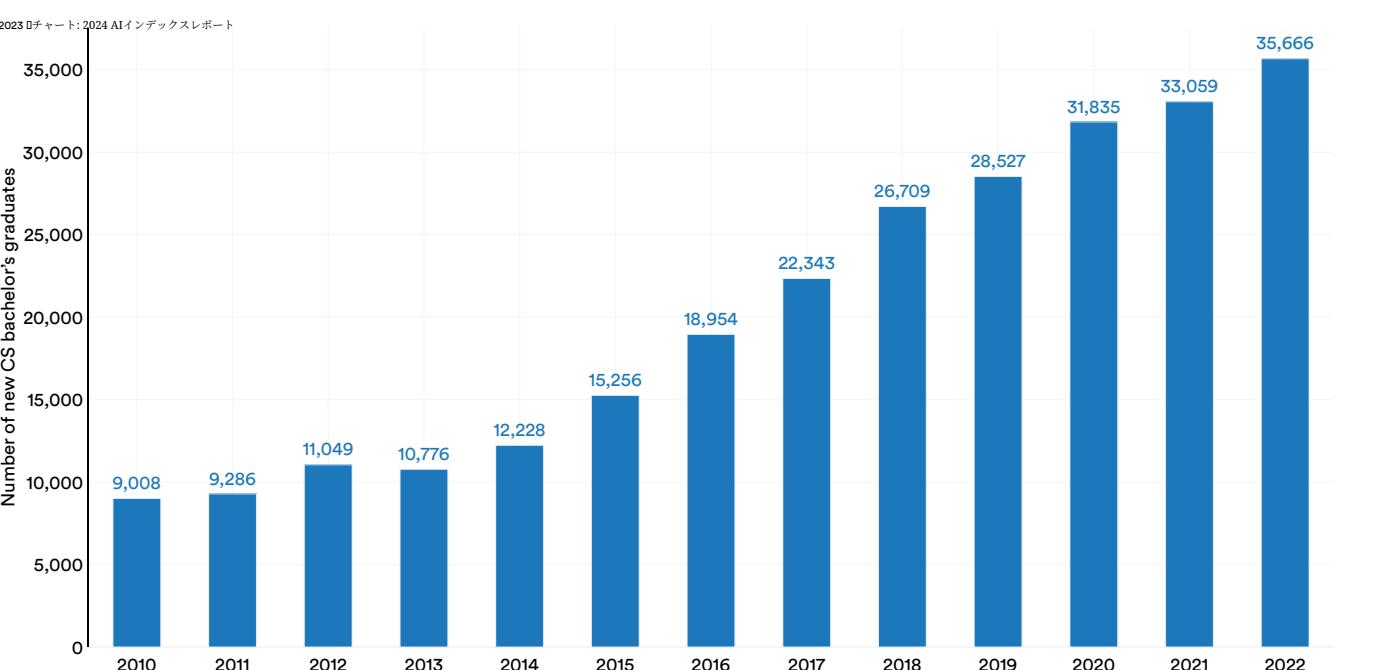


Figure 6.1.1

¹ It is important to note that not all PhD-granting departments targeted in the survey provided responses. Out of the 297 departments targeted, only 182 responded, yielding an overall response rate of 61%.

For the first time in almost eight years, the proportion of international students among CS bachelor's graduates in American and Canadian universities declined, falling from 16.3% in 2021 to 15.2% in 2022 (Figure 6.1.2). This decline likely reflects the increased difficulty of obtaining study visas during the early years of the Trump administration, an impact that is only now beginning to manifest in

the data. The decline is also partially attributable to international travel restrictions that were imposed during the COVID-19 pandemic, affecting the ability of international students to study in the United States and Canada. Despite this recent drop, the overall trend over the last decade shows a steady increase in the proportion of international students.

New international CS bachelor's graduates (% of total) in the United States and Canada, 2010–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

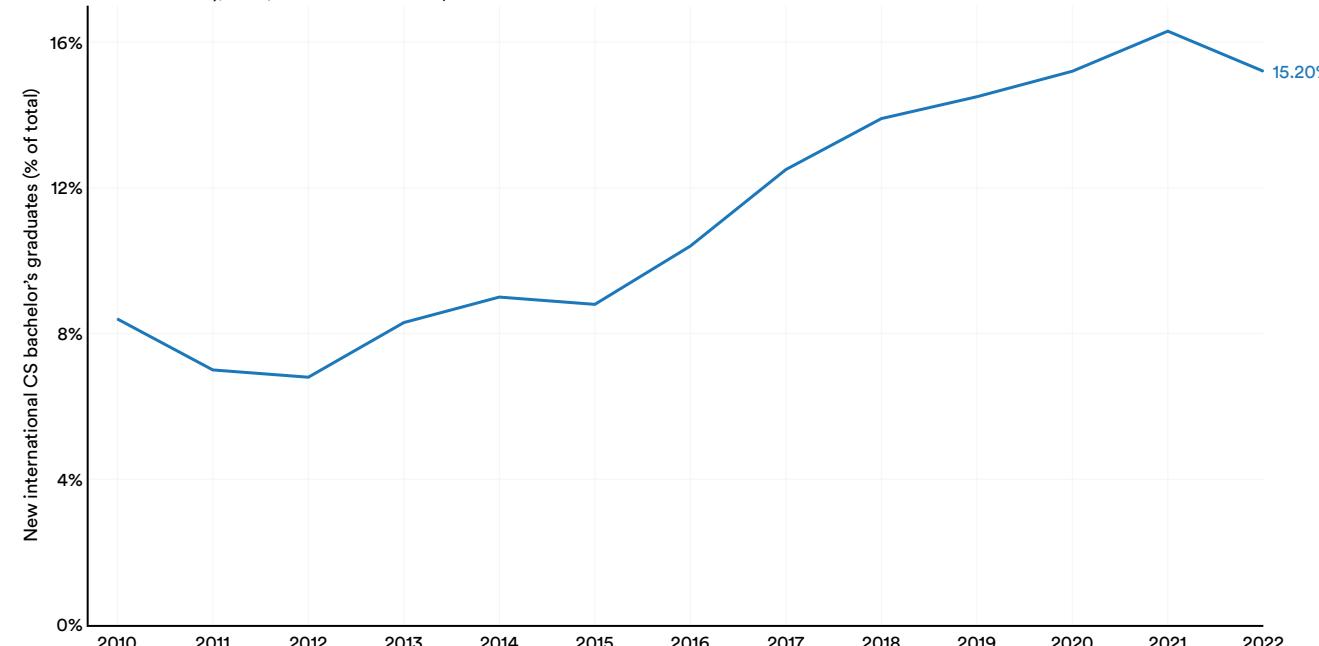


Figure 6.1.2

ほぼ8年ぶりに、アメリカとカナダの大学におけるCS学士卒業生の中で国際学生の割合が減少し、2021年の16.3%から2022年には15.2%に落ち込みました（図6.1.2）。この減少は、トランプ政権初期の間に学習ビザを取得することが難しくなったことを反映している可能性があり、その影響は今になってようやく現れ始めています。

データにおいても、この減少はCOVID-19パンデミック中に課された国際旅行制限の影響を部分的に受けしており、国際学生がアメリカとカナダで学ぶ能力に影響を与えています。この最近の減少にもかかわらず、過去10年間の全体的な傾向は国際学生の割合が着実に増加していることを示しています。

アメリカとカナダにおける新しい国際的なCS学士卒業生（総数の%）、2010–22 出典：CRAタウルビー調査、2023年チャート：2024 AIインデックスレポート

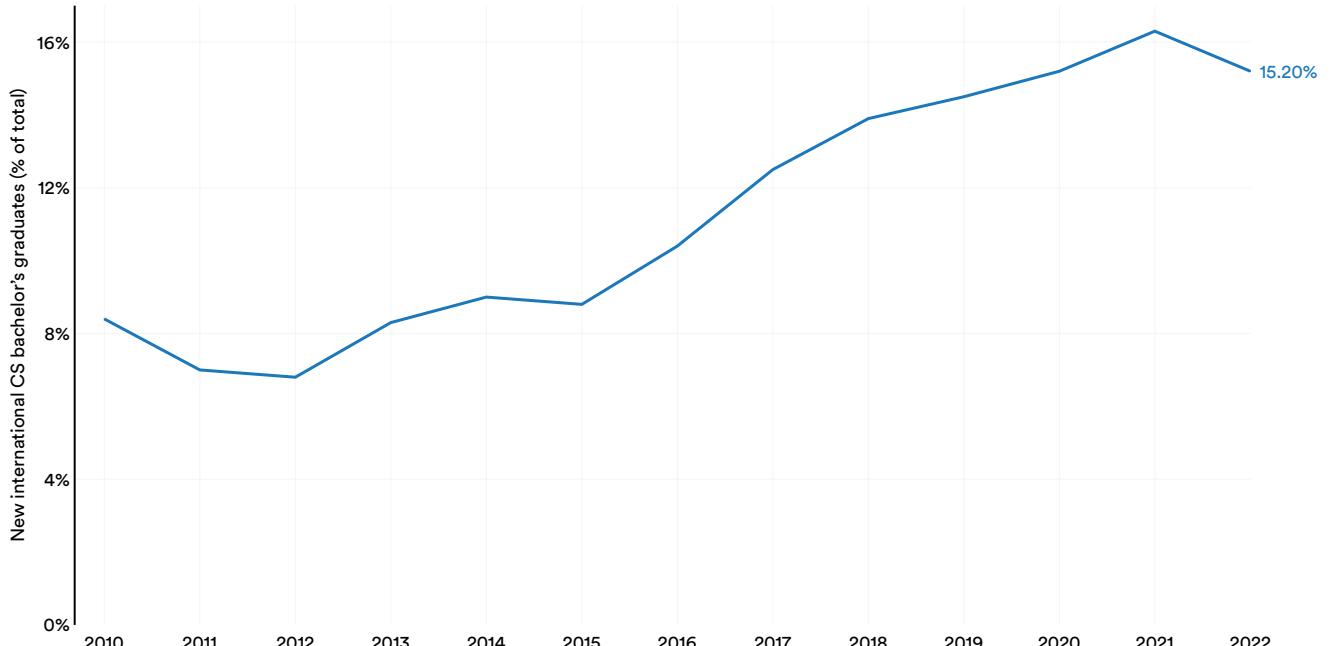


Figure 6.1.2

CS Master's Graduates

AI courses are commonly included in CS master's degree programs. While the total number of new CS master's graduates from American and Canadian universities more than doubled over the past decade,

the number appears to have leveled out since 2018 and slightly decreased, by 2.5%, last year (Figure 6.1.3). This leveling is a reflection of the decline in international master's students shown in the following graph.

New CS master's graduates in the United States and Canada, 2010–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

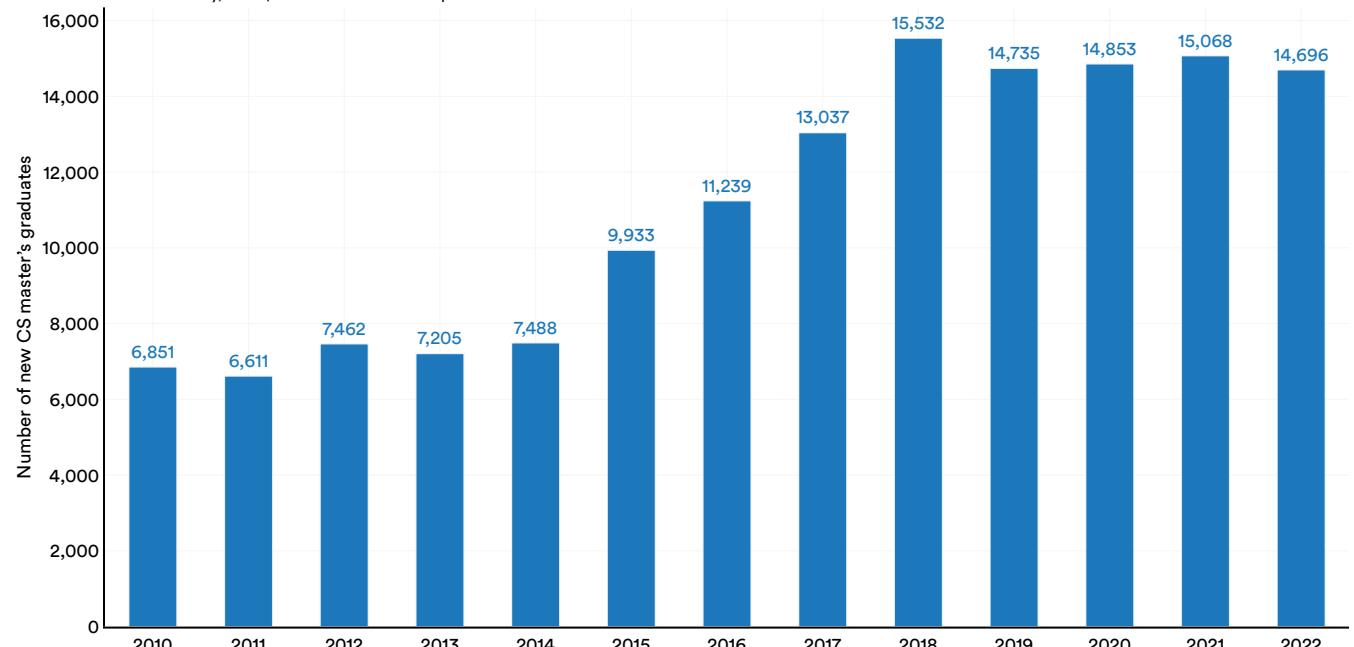


Figure 6.1.3

CS 修士卒業生

AIコースは一般的にCS 修士課程に含まれています。過去10年間でアメリカとカナダの大学からの新しいCS 修士卒業生の総数は2倍以上に増加しましたが、

この数値は2018年以降横ばいになっており、昨年は2.5%減少しました（図6.1.3）。この横ばいは、次のグラフに示されている国際的な修士課程の学生の減少を反映しています。

アメリカとカナダの新しいCS 修士卒業生、2010–22

出典:CRAタウルビー調査、2023年チャート:2024 AIインデックスレポート

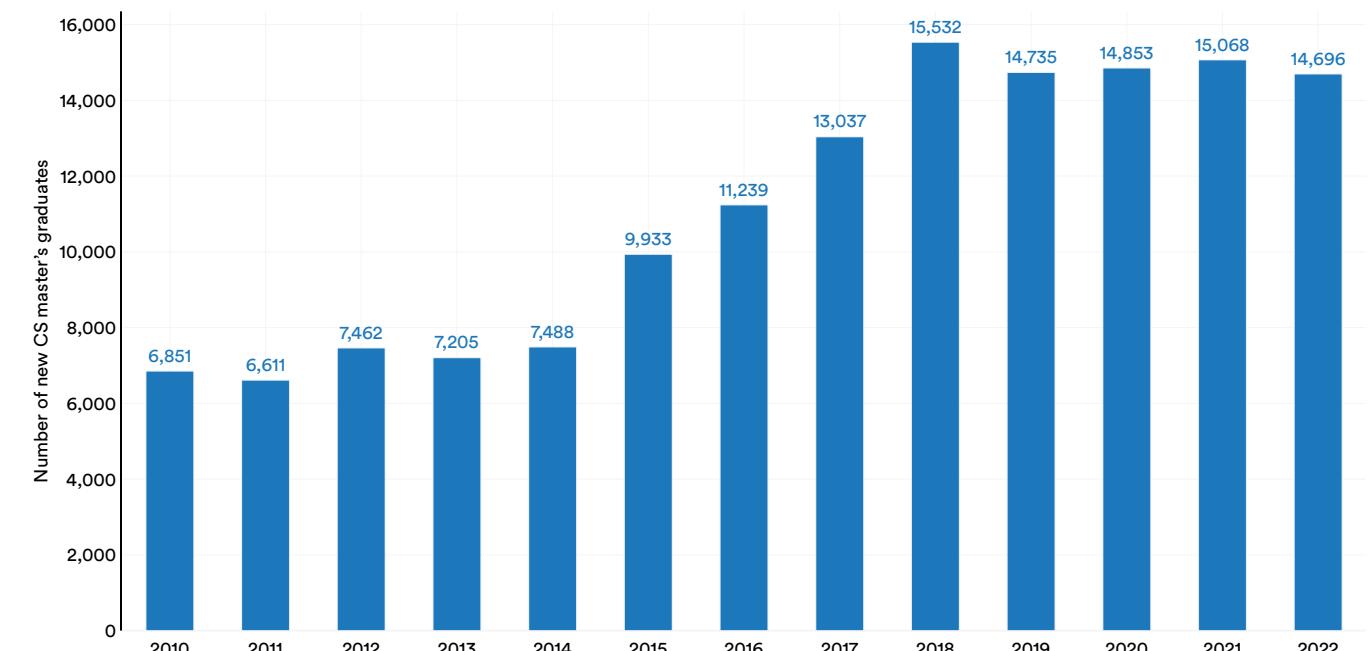


Figure 6.1.3

In 2022, American and Canadian universities experienced a notable decrease in international CS master's students. This downward trend began around 2017, but the decline was most pronounced last year, at 14.8 percentage points (Figure 6.1.4). Currently, the split between international and domestic CS master's graduates is roughly even.

2022年、アメリカとカナダの大学では国際的なCS修士学生が顕著な減少を経験しました。この下降傾向は2017年頃から始まりましたが、昨年は14.8ポイントの減少が最も顕著でした（図6.1.4）。現在、国際的なCS修士卒業生と国内のCS修士卒業生の割合はほぼ均等です。

New international CS master's graduates (% of total) in the United States and Canada, 2010–22

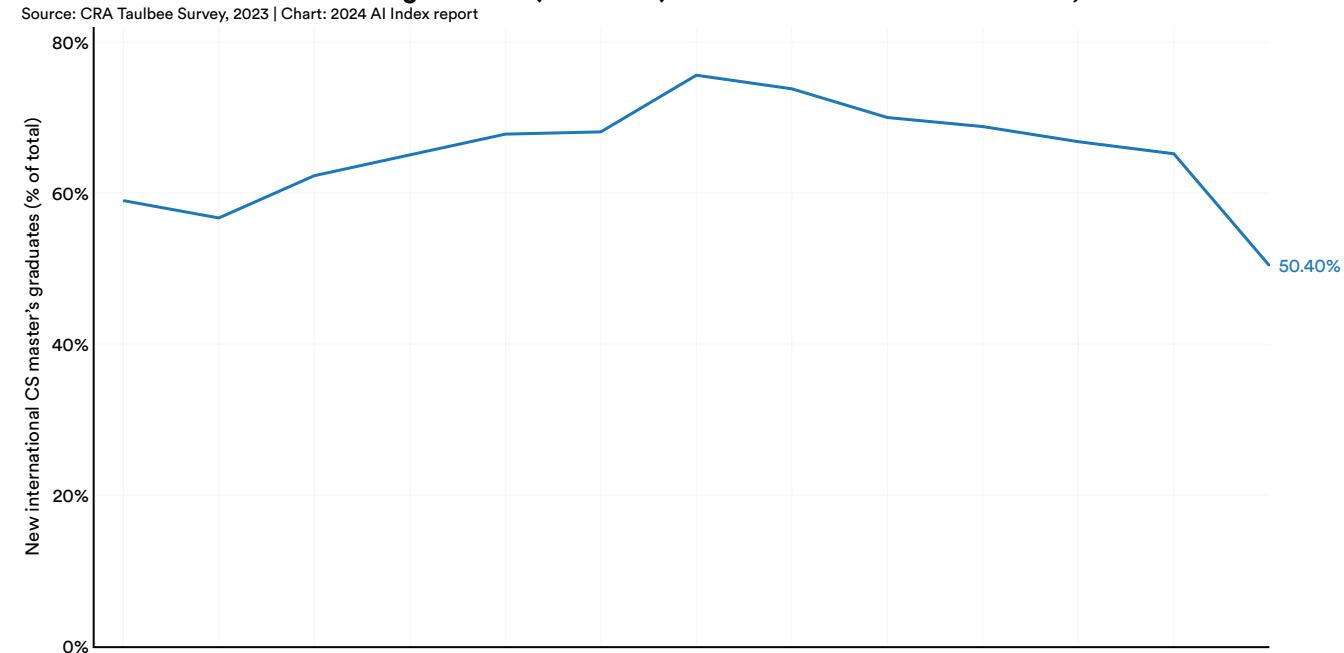


Figure 6.1.4

アメリカとカナダにおける新しい国際的なCS修士卒業生（総数の%）、2010–22 出典：CRAタウルビー調査、2023年チャート：2024 AIインデックスレポート

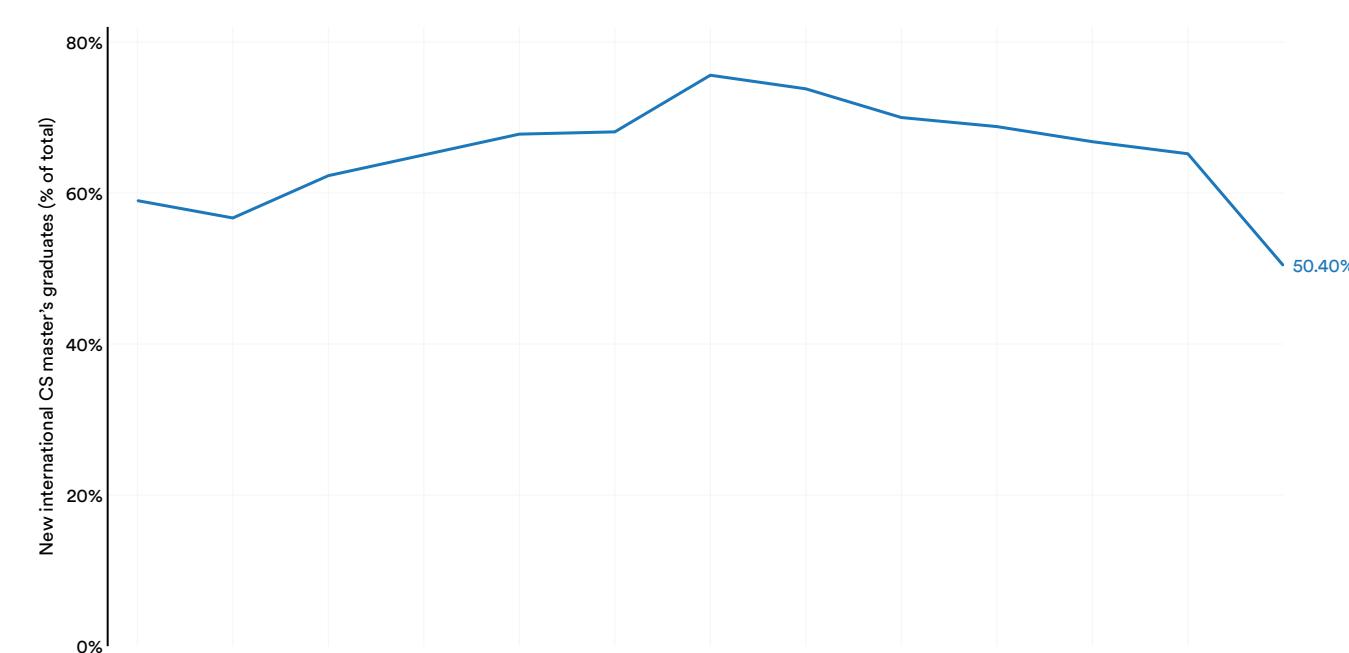


Figure 6.1.4

CS PhD Graduates

For the first time in a decade, there has been a significant increase in the number of new CS PhD graduates at American and Canadian universities. In 2022, the number of CS PhD graduates reached 2,105, the highest since 2010 (Figure 6.1.5).

New CS PhD graduates in the United States and Canada, 2010–22

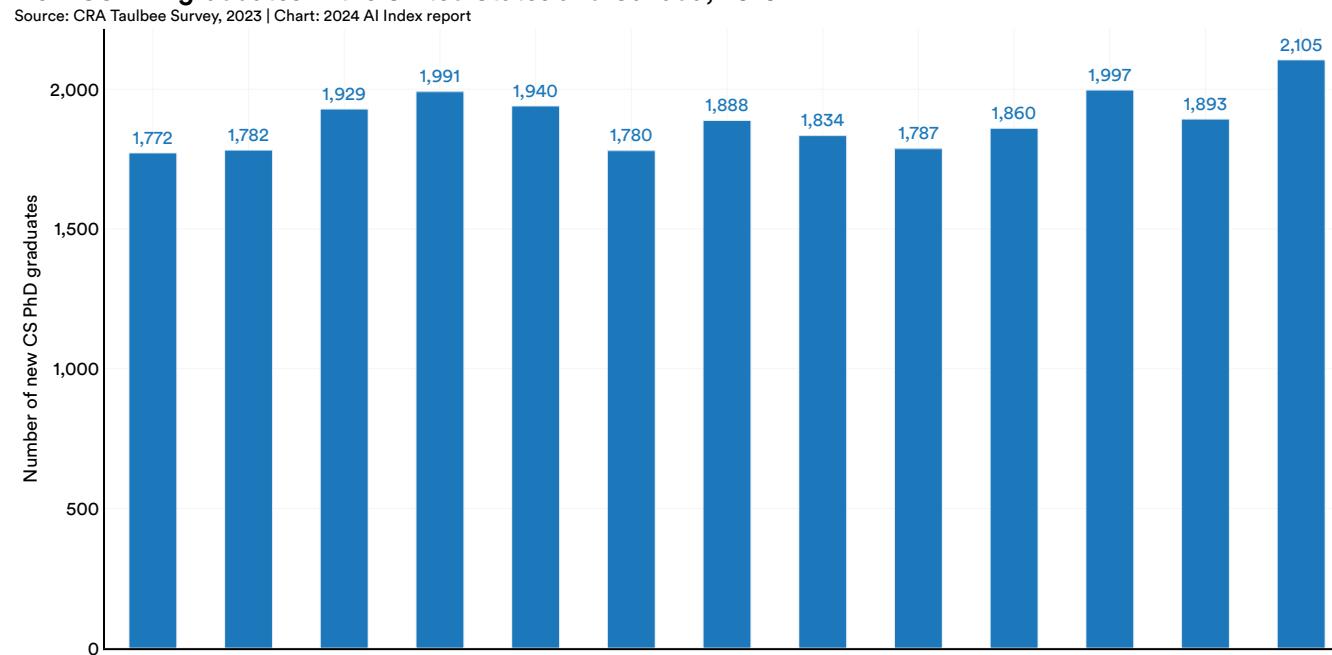


Figure 6.1.5

CS 博士号取得者

10年ぶりに、アメリカとカナダの大学における新しいCS博士号取得者の数が大幅に増加しました。2022年には、CS博士号取得者の数が2,105人に達し、2010年以来の最高値となりました（図6.1.5）。

アメリカとカナダの新しいCS博士号取得者、2010–22

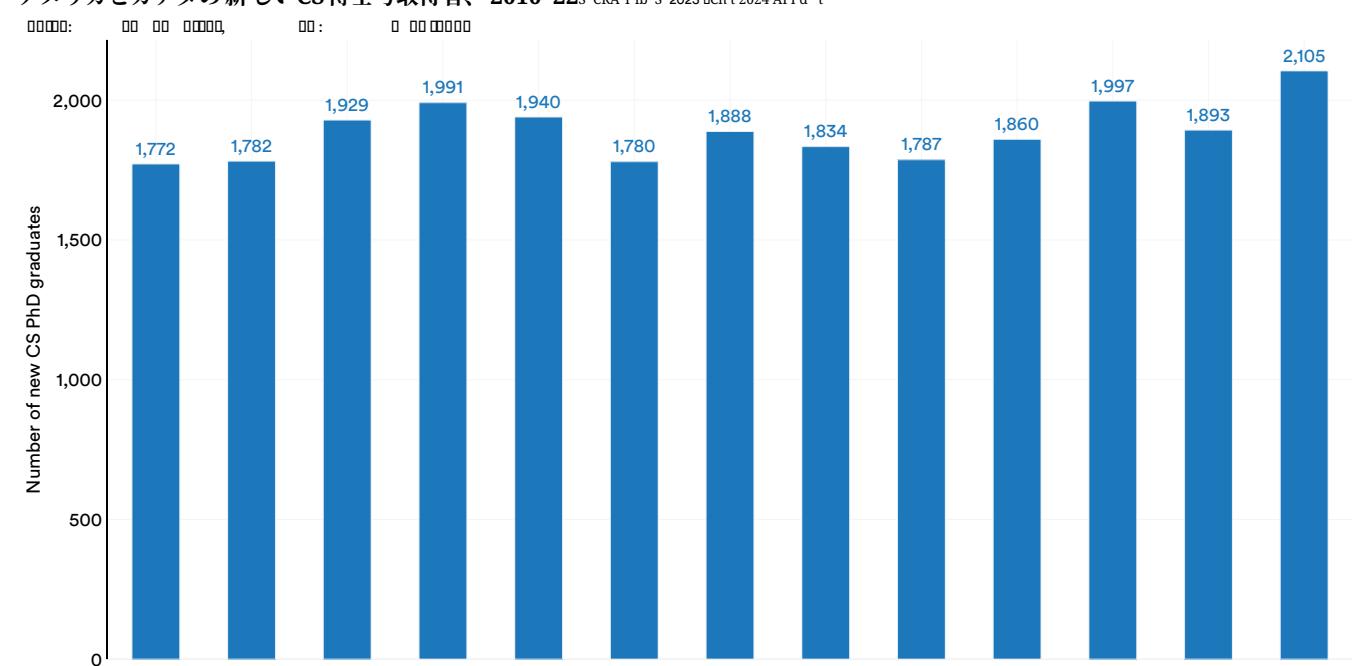


Figure 6.1.5

While the proportion of international students among CS PhD graduates has risen over the past decade, there was a slight decrease in this proportion in the last year, dropping from 68.6% in 2021 to 65.9% in 2022 (Figure 6.1.6).

過去10年間でCS博士号取得者の中で国際学生の割合は増加しましたが、昨年はこの割合がわずかに減少し、2021年の68.6%から2022年には65.9%に落ちました（図6.1.6）。

New international CS PhD graduates (% of total) in the United States and Canada, 2010–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

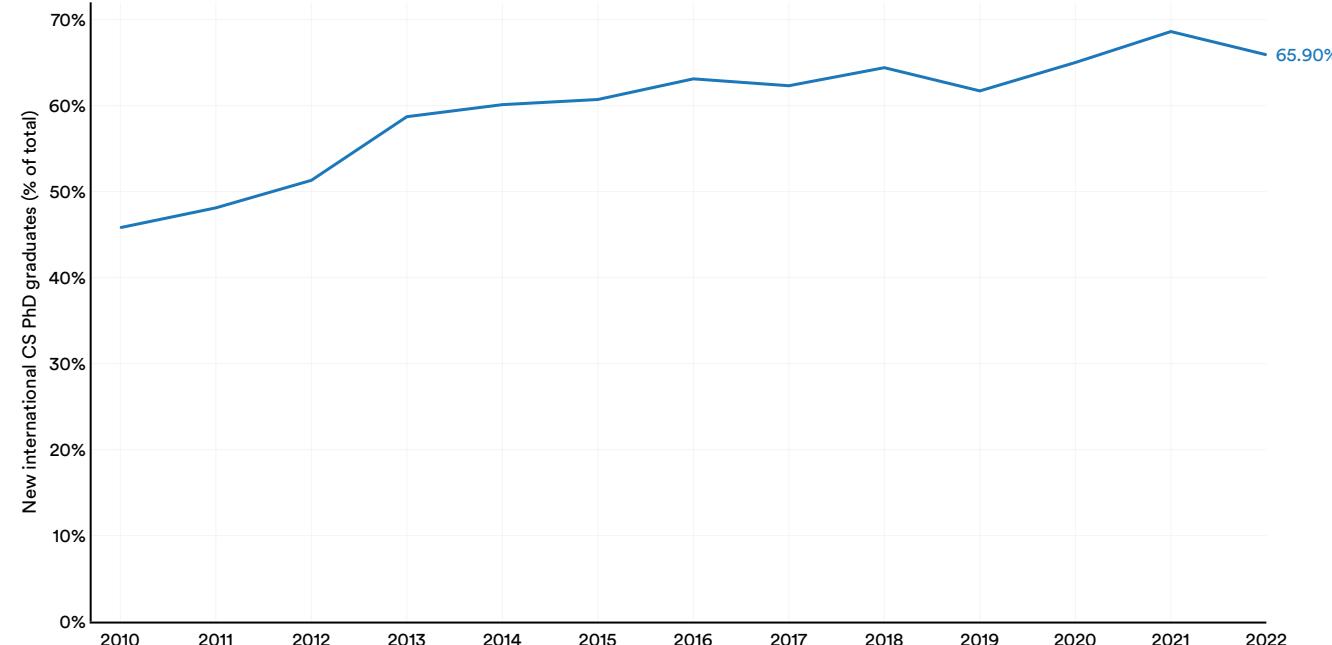


Figure 6.1.6

アメリカとカナダにおける新しい国際的なCS博士号取得者（総数の%）、2010–22

出典:CRAタウルビー調査、2023 | チャート:2024 AIインデックスレポート

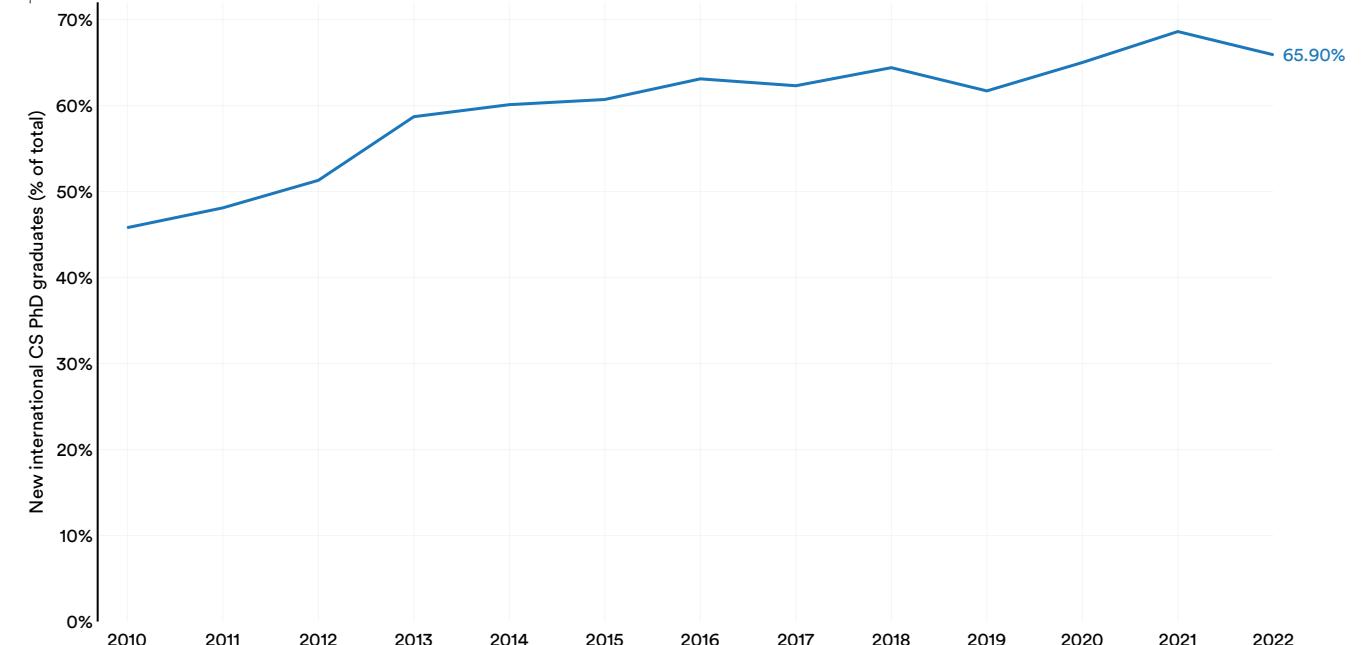


Figure 6.1.6

Where do newly minted AI PhDs choose to work after graduating? Following a trend highlighted in last year's AI Index report, a growing share of AI doctoral recipients are pursuing careers in industry (Figure 6.1.7 and Figure 6.1.8). In 2011, around the same percentage took jobs in industry (40.9%) as in academia (41.6%).

Employment of new AI PhDs (% of total) in the United States and Canada by sector, 2010–22

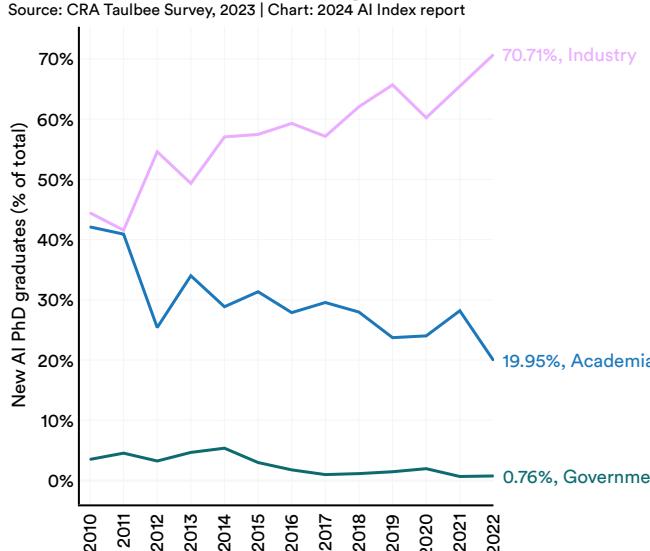


Figure 6.1.7²

Employment of new AI PhDs in the United States and Canada by sector, 2010–22

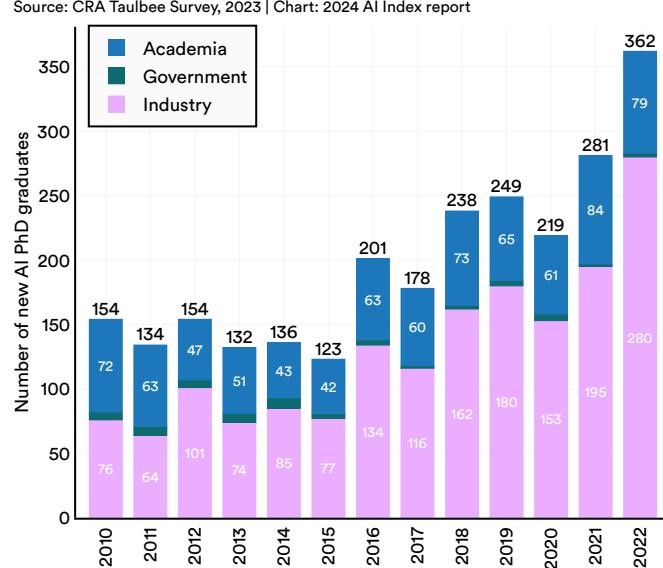


Figure 6.1.8

新たに取得したAI博士号を持つ人々は、卒業後どこで働くことを選ぶのでしょうか？昨年のAIインデックスレポートで強調されたトレンドに従い、AI博士号取得者の中で産業界でのキャリアを追求する割合が増加しています（図6.1.7および図6.1.8）。2011年には、産業界での就職率（40.9%）と学界での就職率（41.6%）はほぼ同じでした。

しかし、2022年までに、卒業後に産業界に参加した割合は大幅に増加し（70.7%）、学界に入る割合（20.0%）と比較されました。新しいAI博士号取得者が政府の役割に就く割合は、過去5年間で約0.7%と比較的低く安定しています。

アメリカとカナダにおける新しいAI博士号取得者の雇用（全体の%）セクター別、2010–22

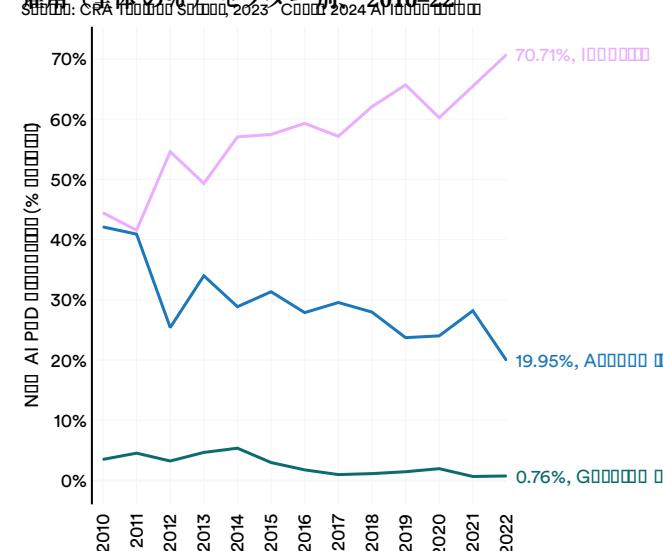


Figure 6.1.7²

アメリカにおける新しいAI博士号取得者の雇用とカナダのセクター別、2010–22

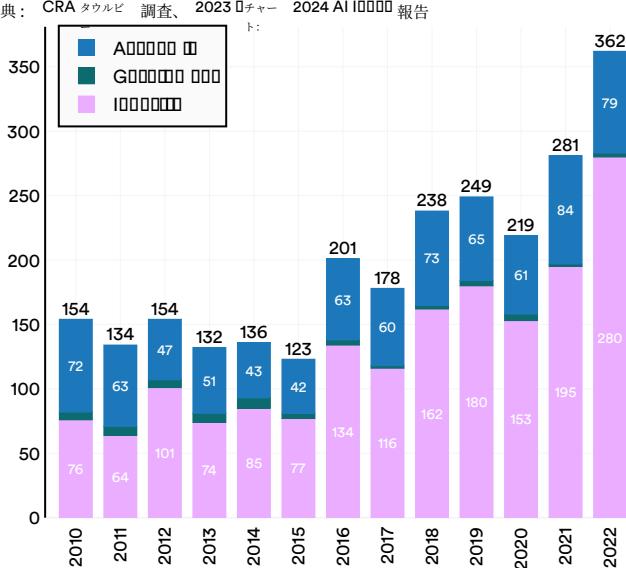


Figure 6.1.8

² The sums in Figure 6.1.7 do not add up to 100, as there is a subset of new AI PhDs each year who become self-employed, unemployed, or report an "other" employment status in the CRA survey. These students are not included in the chart.

² The sums in Figure 6.1.7 do not add up to 100. 每年新たに取得したAI博士号の中には、自営業、失業、またはCRAチャートで「その他」の雇用状況を報告する人々のサブセットが存在します。These students are not included in the survey.

CS, CE, and Information Faculty

To better understand trends in CS and AI education, it is helpful to examine data on CS faculty. Last year, the total number of CS, CE, and information faculty in American and Canadian universities increased 7.2% (Figure 6.1.9). Since 2011, the increase is 42.4%.

CS、CE、および情報学部

CSおよびAI教育のトレンドをよりよく理解するためには、CS教員に関するデータを調べることが役立ちます。昨年、アメリカおよびカナダの大学におけるCS、CE、および情報学部の教員数は7.2%増加しました（図6.1.9）。2011年以降の増加率は42.4%です。

Number of CS, CE, and information faculty in the United States and Canada, 2011–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

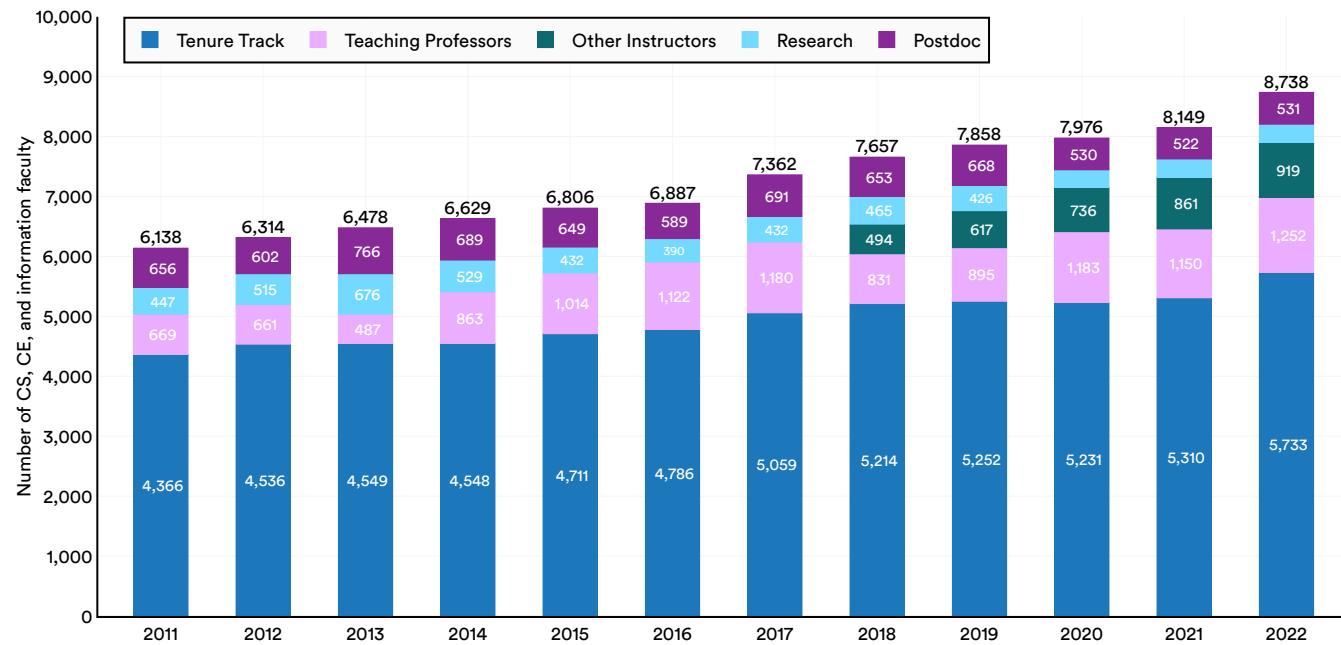


Figure 6.1.9

N アメリカとカナダのCS、CE、および情報学部の教員数、2011–22 出典:CRAタウルビー調査、2023 チャート:2024 AIインデックスレポート 2

SD

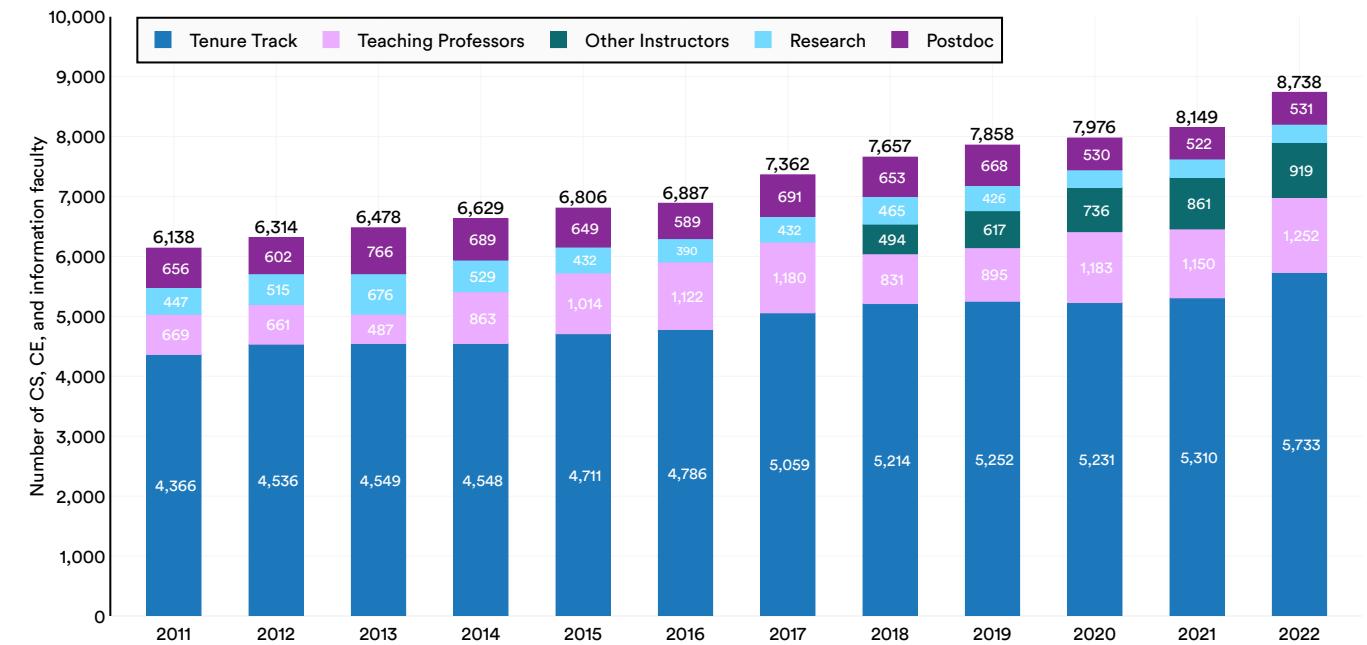


Figure 6.1.9

In 2022, the United States had 7,084 CS faculty members, with the majority (65.7%) on the tenure track (Figure 6.1.10). The total number of American CS faculty has risen 4.4% since 2021 and 45.0% since 2011.

2022年、アメリカ合衆国には7,084人のCS教員がおり、その大多数（65.7%）が tenure trackにあります（図 6.1.10）。アメリカのCS教員の総数は2021年から4.4%、2011年から45.0%増加しました。

Number of CS faculty in the United States, 2011–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

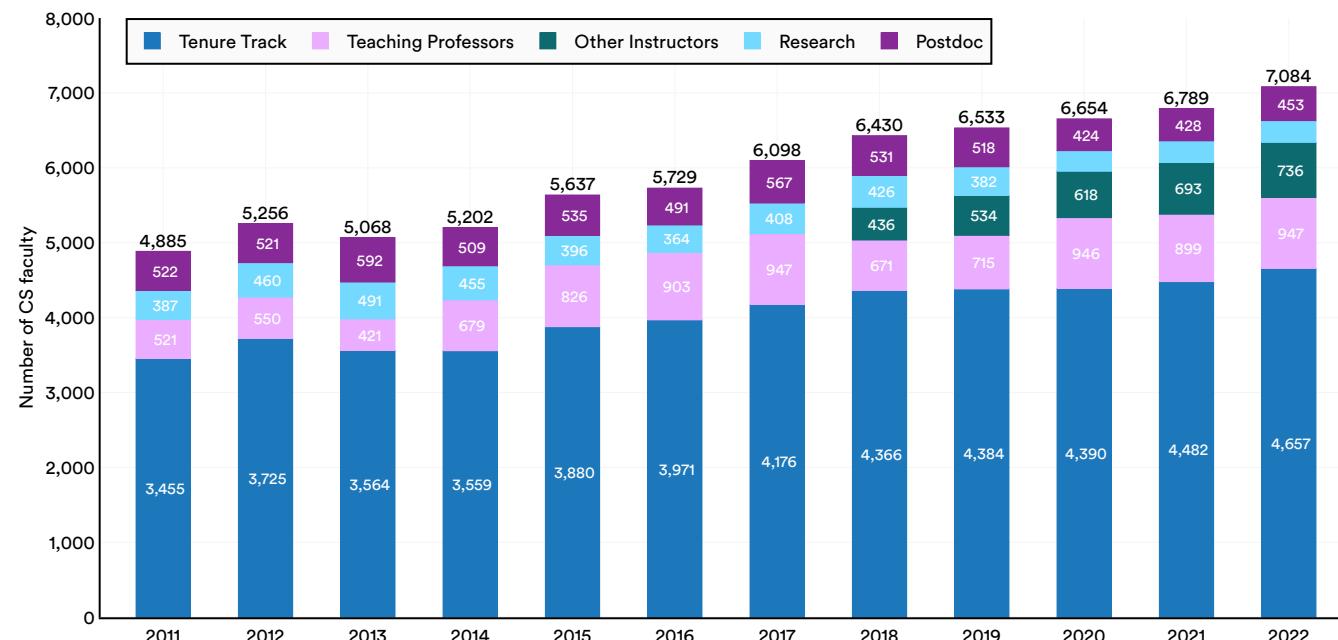


Figure 6.1.10

アメリカ合衆国のCS教員数、2011–22

出典: CRA タウルビ 調査、2023 Uチヤー 2024 AI Index 報告

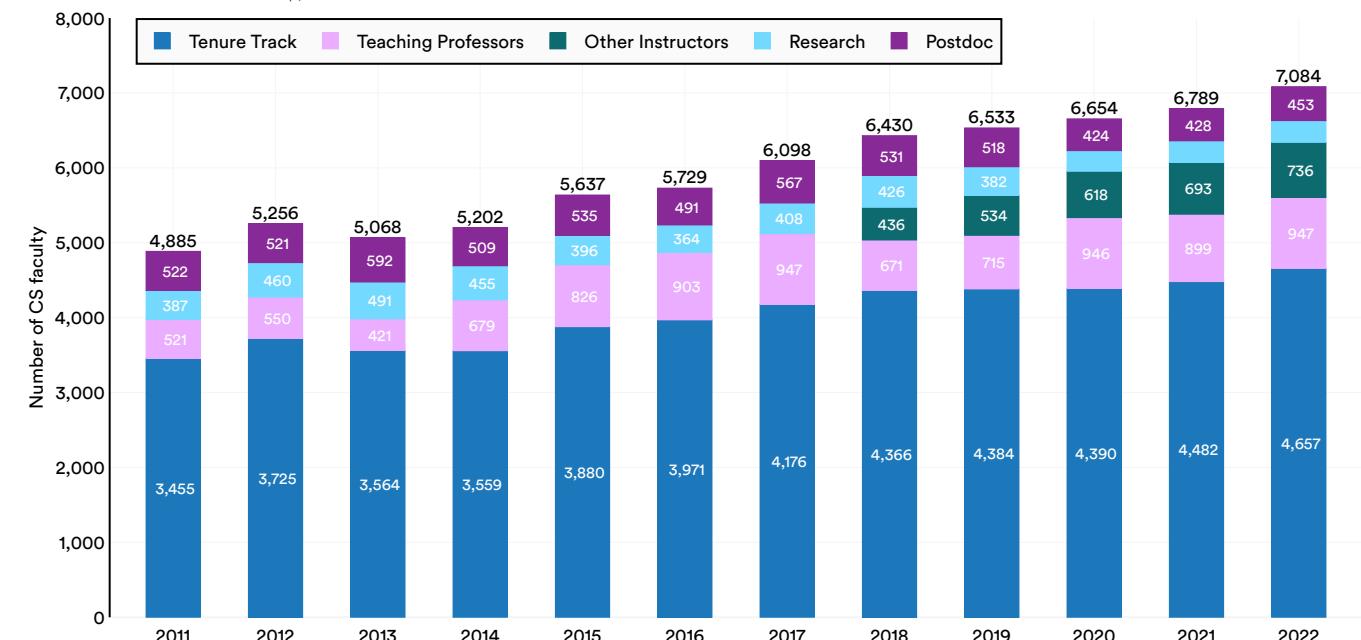


Figure 6.1.10

Last year, 915 new faculty were hired across CS, CE, and information disciplines in North America, a decade high. 455 of these positions were tenure track. (Figure 6.1.11).

昨年、北米でCS、CE、および情報分野において915人の新しい教員が採用され、10年ぶりの高水準となりました。これらのポジションのうち455は tenure track でした。（図 6.1.11）

New CS, CE, and information faculty hires in the United States and Canada, 2011–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

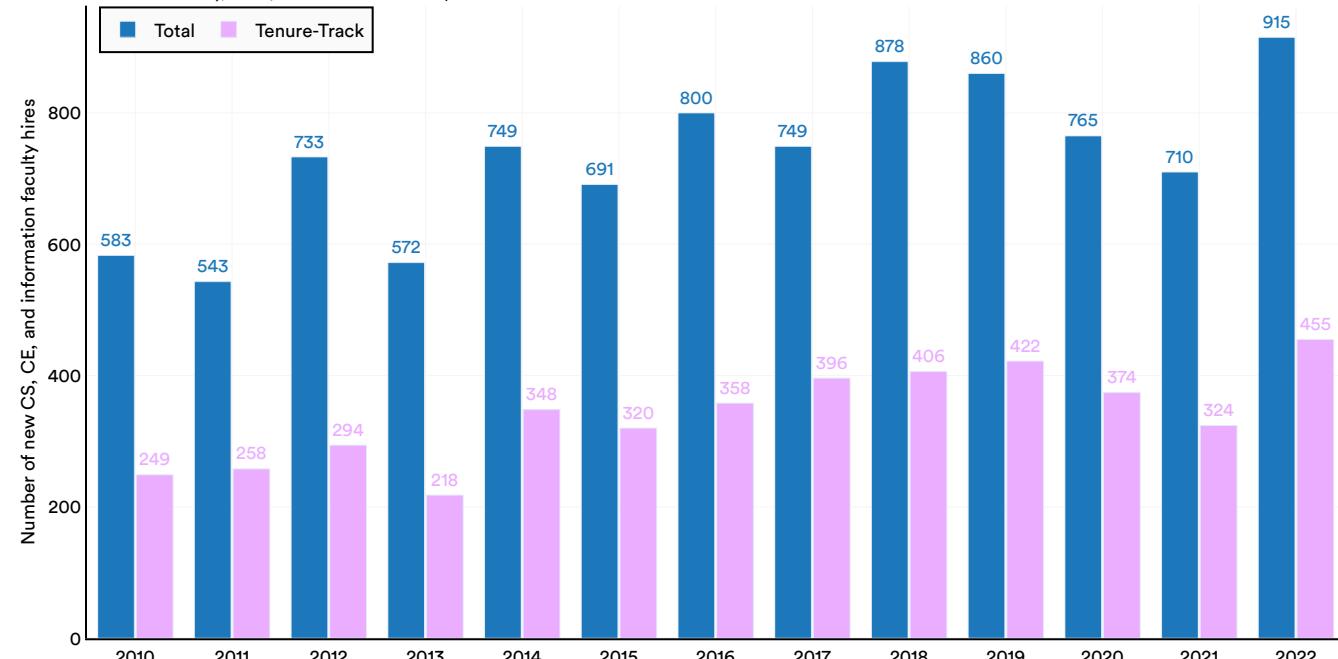


Figure 6.1.11

アメリカとカナダにおける新しいCS、CE、および情報学部の教員採用、2011–22

出典:CRAタウルビー 調査、2023 データ:2024 AIインデックスレポート

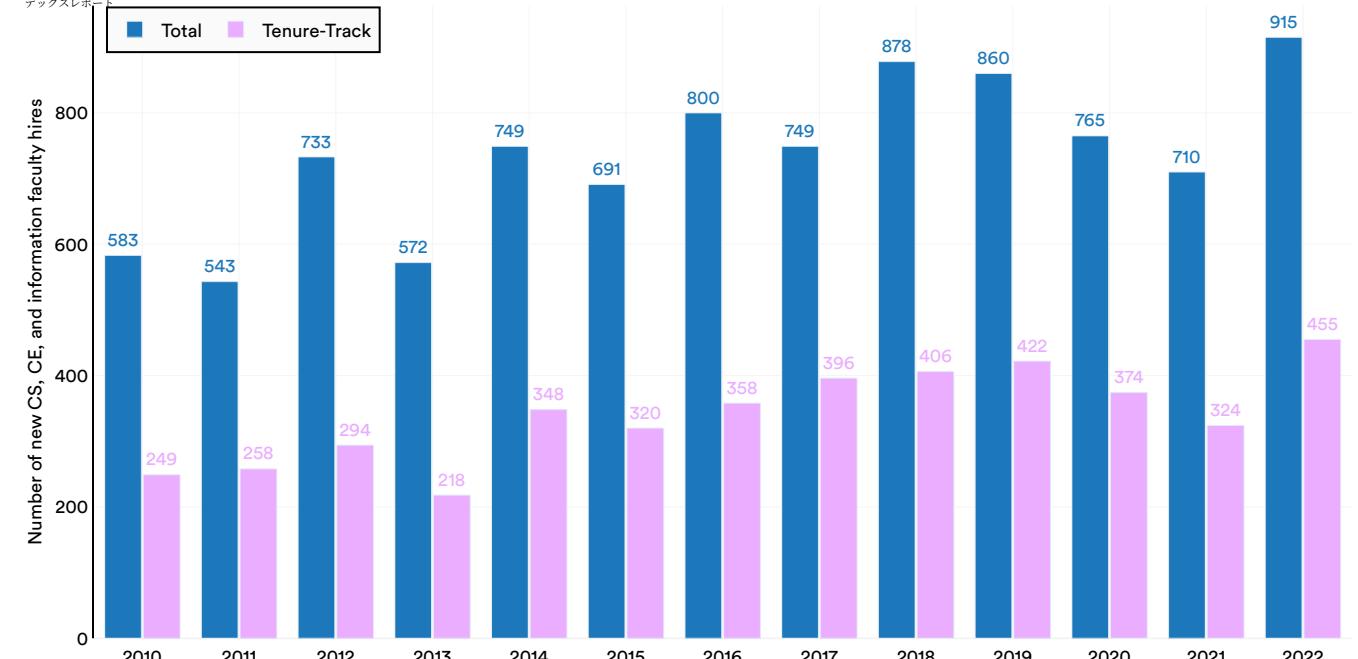


Figure 6.1.11

In 2022, 43% of new faculty appointments came from other academic positions, indicating a “churn” within the academic workforce (Figure 6.1.12). Since these “new” faculty members vacated positions elsewhere, their previous roles will eventually need to be filled. Additionally, the proportion of faculty transitioning from industry in 2022 fell to 7% from 11% in the previous year and 13% in 2019.

2022年には、新しい教員の採用の43%が他の学術職からのものであり、学術労働力内での「入れ替わり」を示しています（図6.1.12）。これらの「新しい」教員が他の場所で職を空けたため、彼らの以前の役割は最終的に埋められる必要があります。さらに、2022年に業界から移行した教員の割合は、前年の11%から7%に、2019年の13%からも減少しました。

Source of new faculty in American and Canadian CS, CE, and information departments, 2018–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

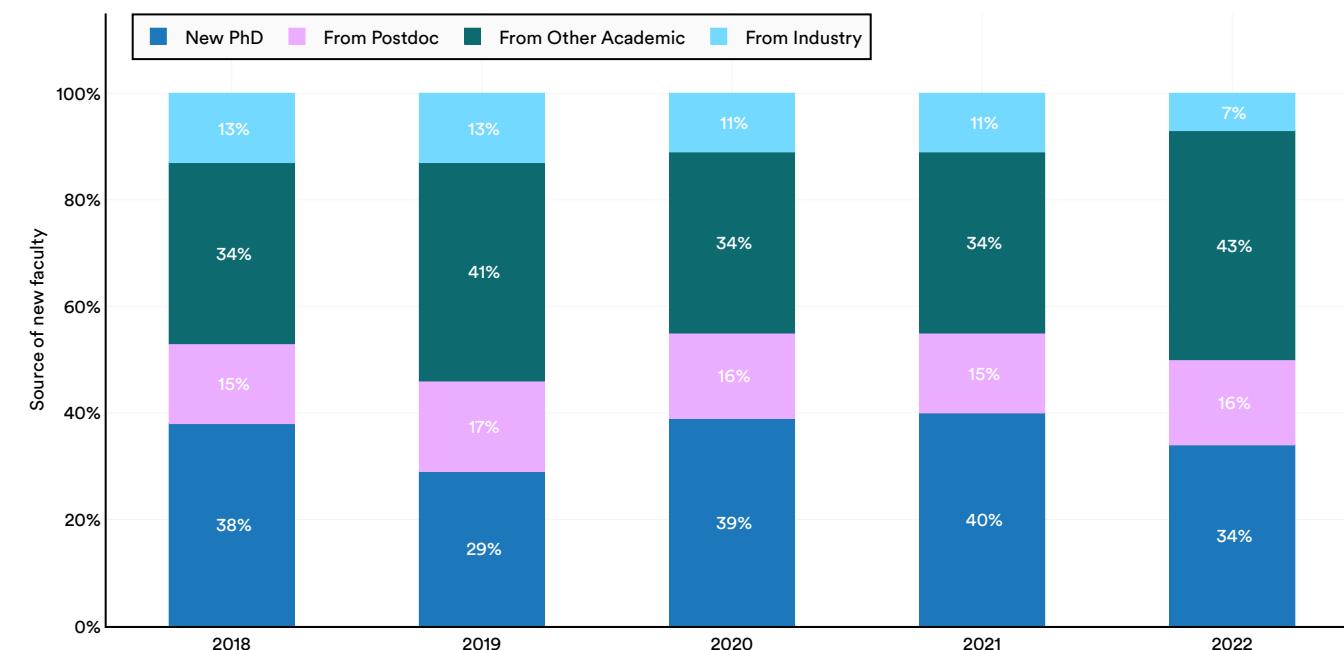


Figure 6.1.12

アメリカおよびカナダのCS、CE、情報学部における新しい教員の出所、2018–22

出典:CRA Taulbee 調査、2023 | 国:2024 年 AI インデックスレポート

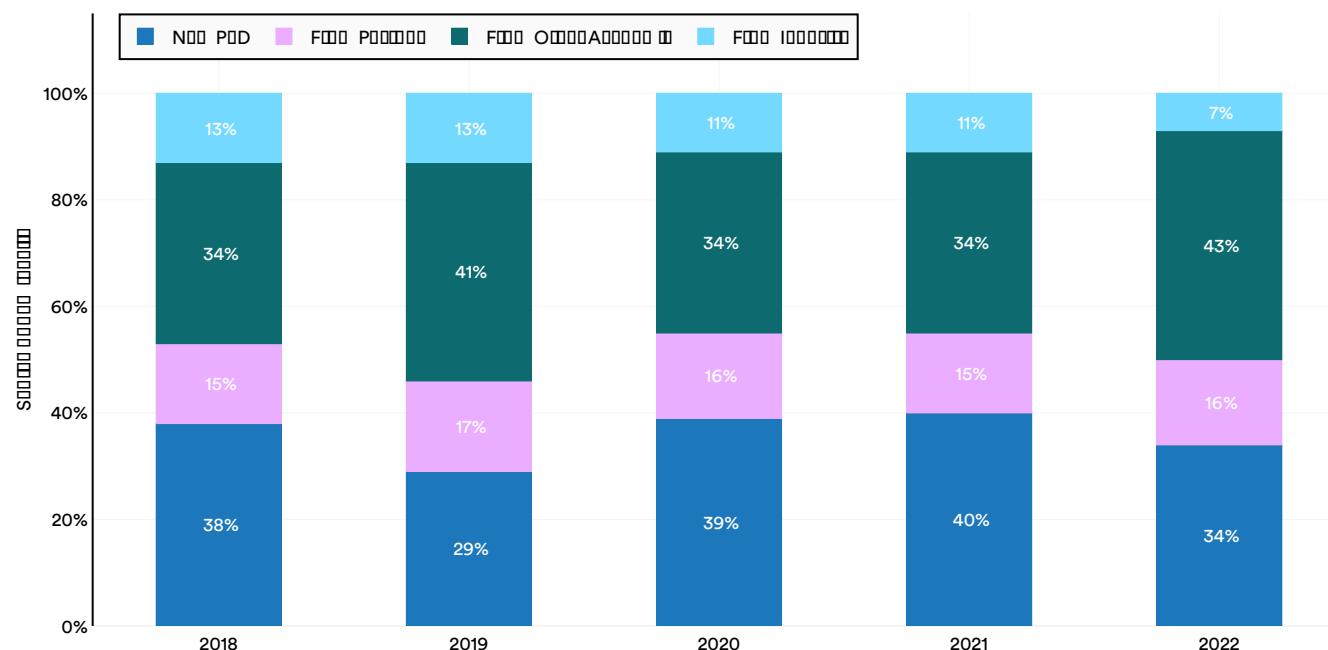


Figure 6.1.12

The reasons for faculty positions remaining unfilled have varied over the past decade. In 2011, 37% of failed searches were due to no offer being made, while 34% were because the offer made was declined (Figure 6.1.13). In contrast, in 2022, only 15% ended with no offer being made, while 55% involved offers that

were turned down. This trend appears to reflect an increasingly competitive market for new CS faculty. However, it remains unclear whether this indicates heightened competition with other academic positions or with industry positions.

Reason why new CS, CE, and information faculty positions remained unfilled (% of total), 2011–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

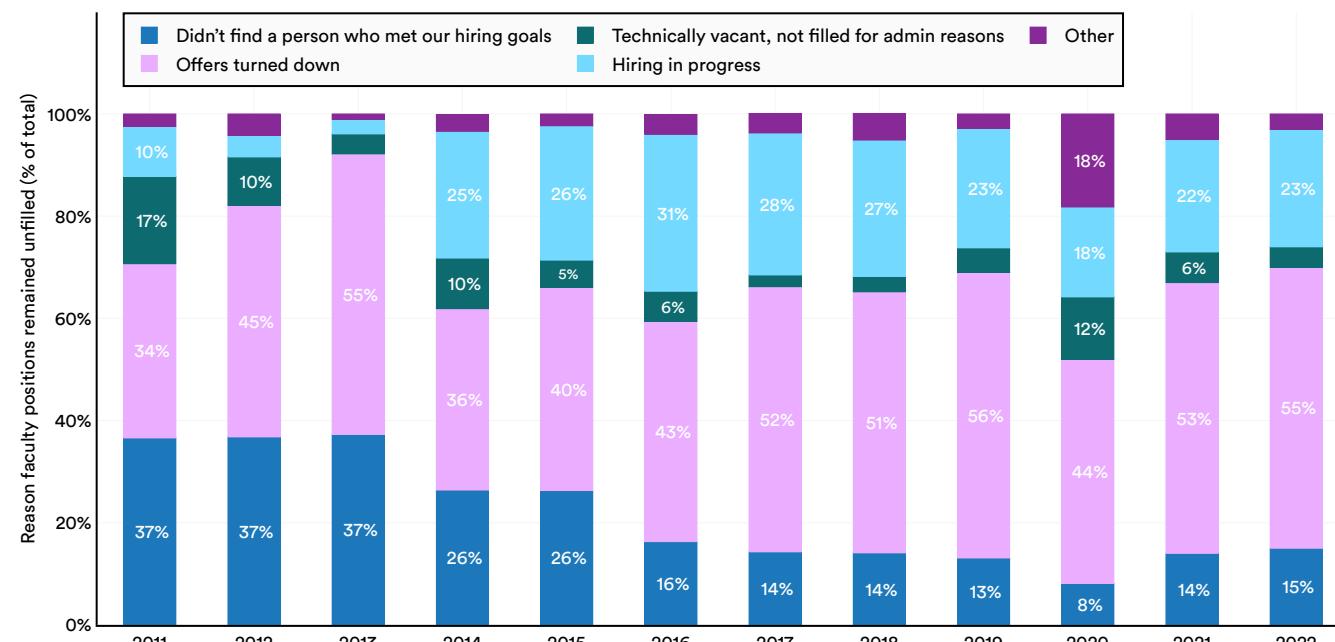


Figure 6.1.13

教員 ポジションが埋まらない理由は、過去10年間で様々でした。2011年には、失敗した採用の37%がオファーが出されなかったことによるものであり、34%は出されたオファーが辞退されたことによるものでした（図6.1.13）。対照的に、2022年には、オファーが出されなかったのはわずか15%であり、55%はオファーが

辞退されました。この傾向は、新しいCS教員のための競争がますます激化している市場を反映しているようです。しかし、これは他の学術ポジションとの競争が激化しているのか、業界ポジションとの競争が激化しているのかは不明です。

新しいCS、CE、および情報学部の教員ポジションが未充実の理由（総数の%）、2011–22 出典：CRAタウルビー調査、2023年チャート：2024 AIインデックスレポート

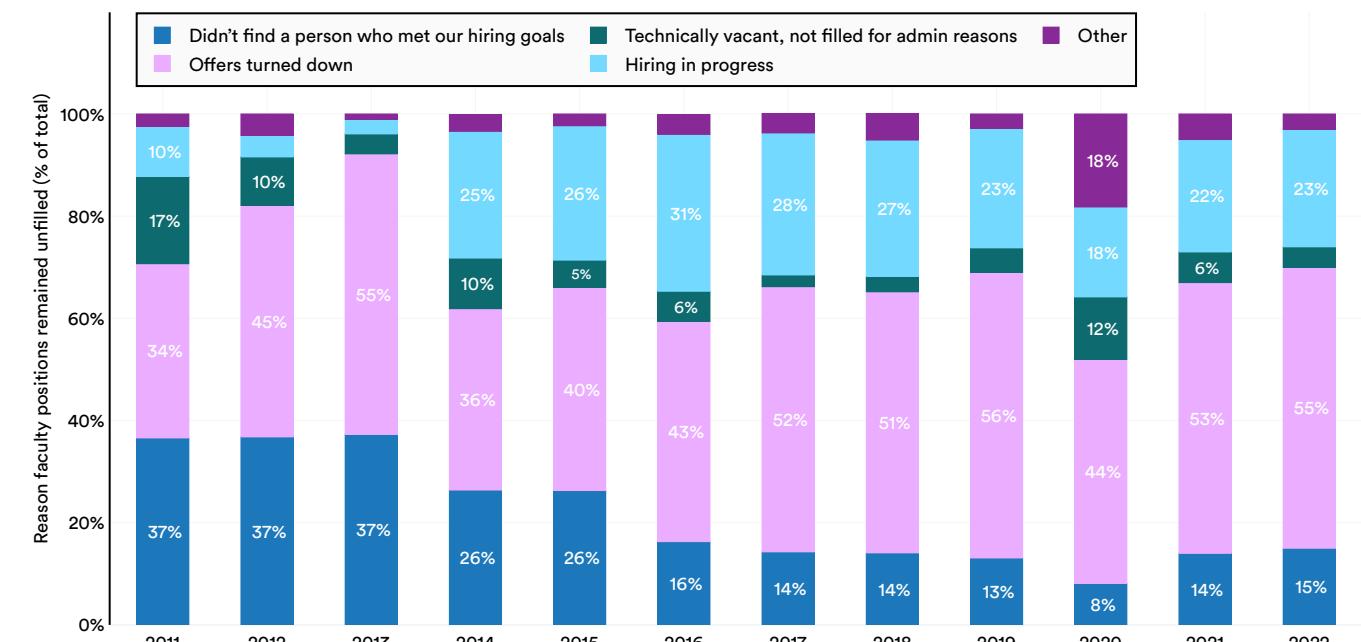


Figure 6.1.13

In 2022, North American departments in CS, CE, and information disciplines experienced a significant increase in faculty departures, totaling 405, compared to 303 in 2021 (Figure 6.1.14). Of these losses, 38.5% left for other academic positions, while 16.3% moved to nonacademic roles, maintaining a trend consistent with previous years.

Faculty losses in American and Canadian CS, CE, and information departments, 2011–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

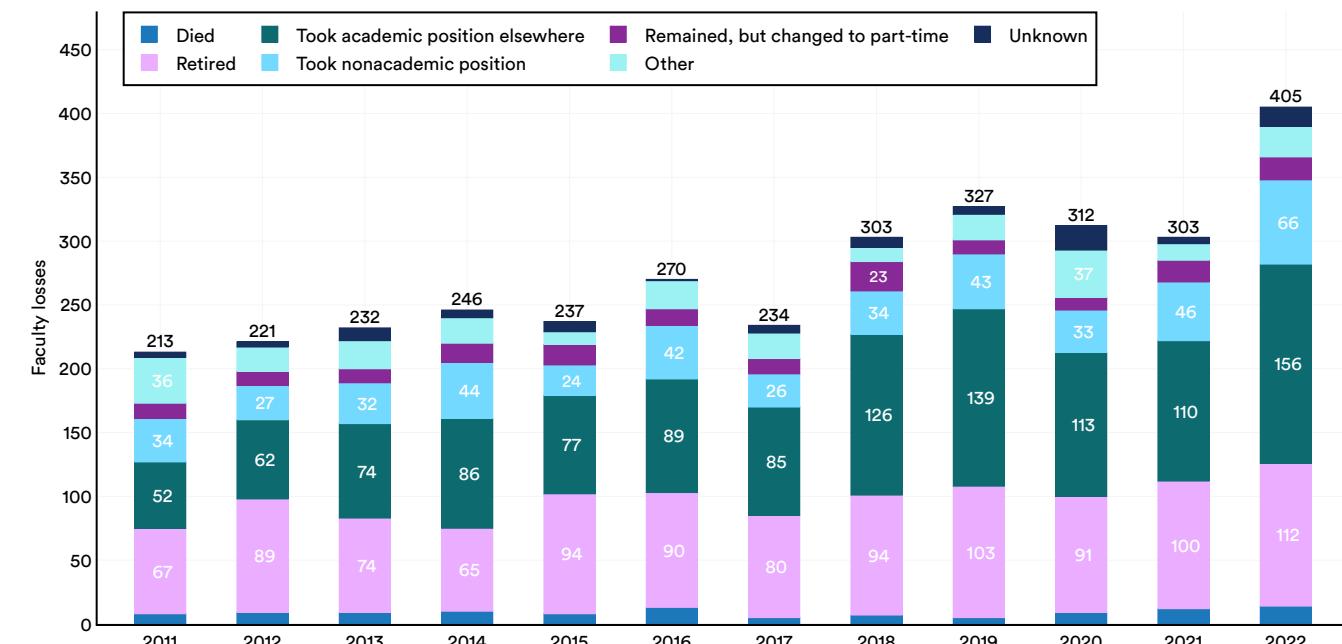


Figure 6.1.14

2022年、北米のCS、CE、および情報分野の学部は、2021年の303人と比較して、教員の離職が405人に達し、著しい増加を経験しました（図6.1.14）。これらの損失のうち、38.5%が他の学術職に移動し、16.3%が非学術的な役割に移ったことは、前年と一貫した傾向を維持しています。

アメリカおよびカナダのCS、CE、情報学部における教員の損失、2011–22

出典：CRAタウルビー調査、2023 | チャート：2024 AIインデックスレポート

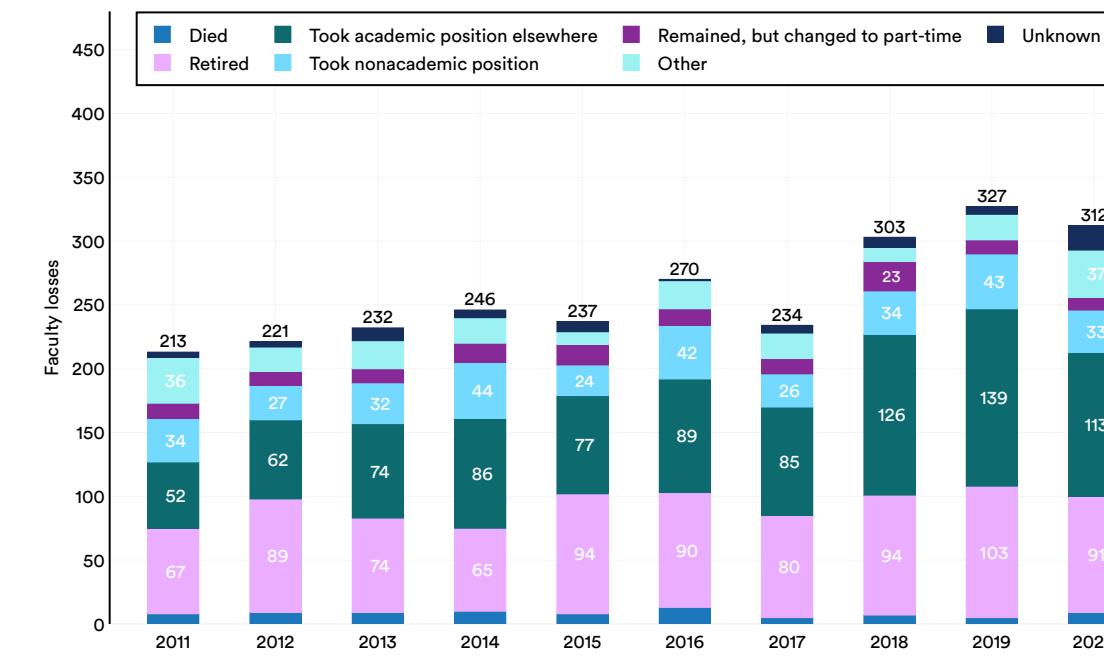


Figure 6.1.14

Since 2015, the increase in median nine-month salaries for full professors has slightly fallen below U.S. inflation rates, whereas median salaries for assistant and associate professors have seen slight increases above inflation. In 2022, a full professor's salary was 3.2% higher than in 2021, which did not keep pace with the 7% U.S. inflation rate, and 16.4% higher than in 2015, still below the 19% inflation increase over those years (Figure 6.1.15).

2015年以降、フルプロフェッサーの中央値の9か月間の給与の増加は、米国のインフレ率をわずかに下回っていますが、アシスタントおよびアソシエイトプロフェッサーの中央値の給与はインフレを上回るわずかな増加を見ています。2022年には、フルプロフェッサーの給与は2021年より3.2%高く、7%の米国インフレ率には追いついておらず、2015年よりも16.4%高いですが、その期間の19%のインフレ増加にはまだ及んでいません（図6.1.15）。

Median nine-month salary of CS faculty in the United States, 2015–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

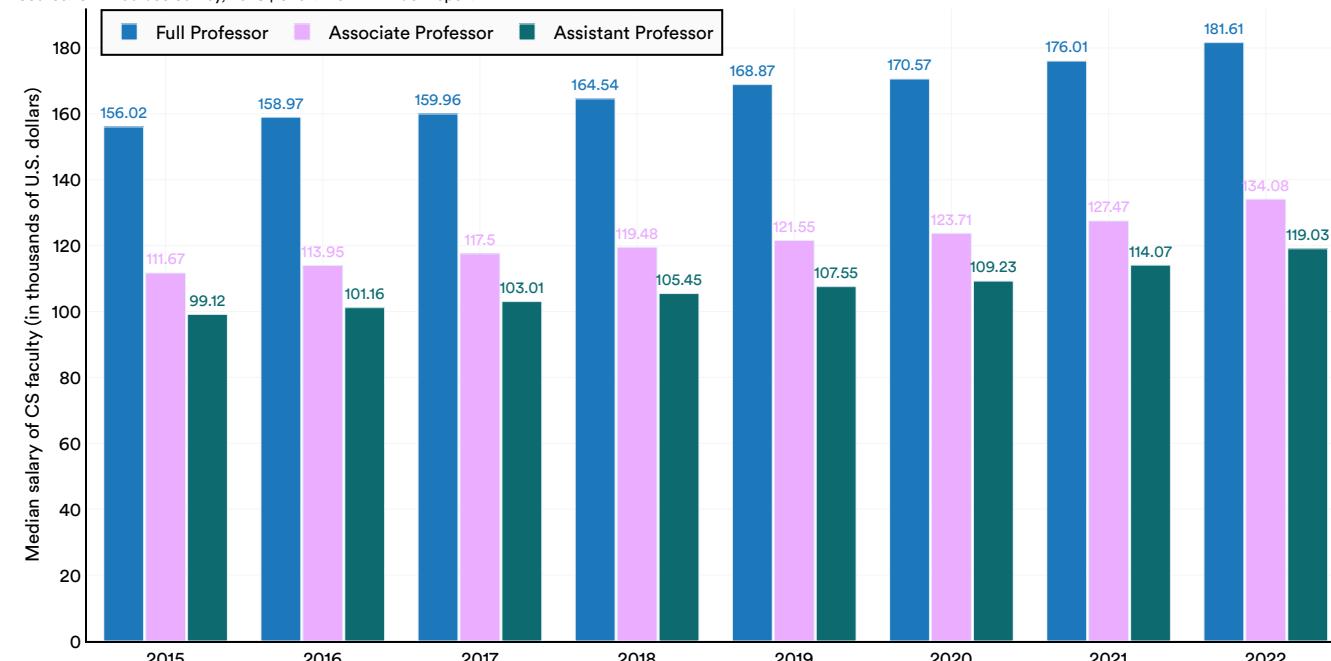


Figure 6.1.15

アメリカのCS教員の9か月間の中央値給与、2015–22 出典:CRAタウルビー調査、2023 チャート:2024 AIイ

ンデックスレポート

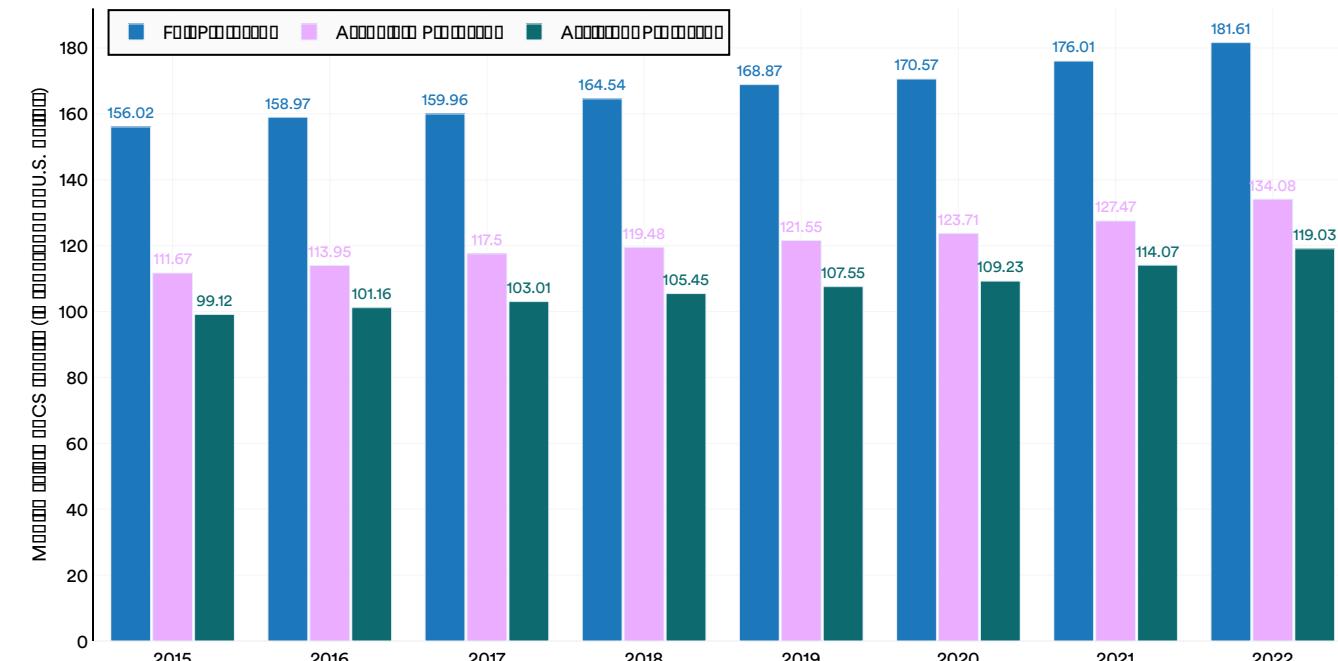


Figure 6.1.15

In 2022, the proportion of international hires among new tenure-track faculty in CS, CE, and information disciplines significantly increased to 19.3% from 13.2% the previous year (Figure 6.1.16). This marked the second-highest percentage recorded in the past decade, only surpassed by 2013.

2022年には、CS、CE、および情報分野の新しい tenure-track 教員の国際的な採用割合が、前年の13.2%から19.3%に大幅に増加しました（図 6.1.16）。これは過去10年間で記録された2番目に高い割合であり、2013年に次ぐものです。

New international CS, CE, and information tenure-track faculty hires (% of total) in the United States and Canada, 2010–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

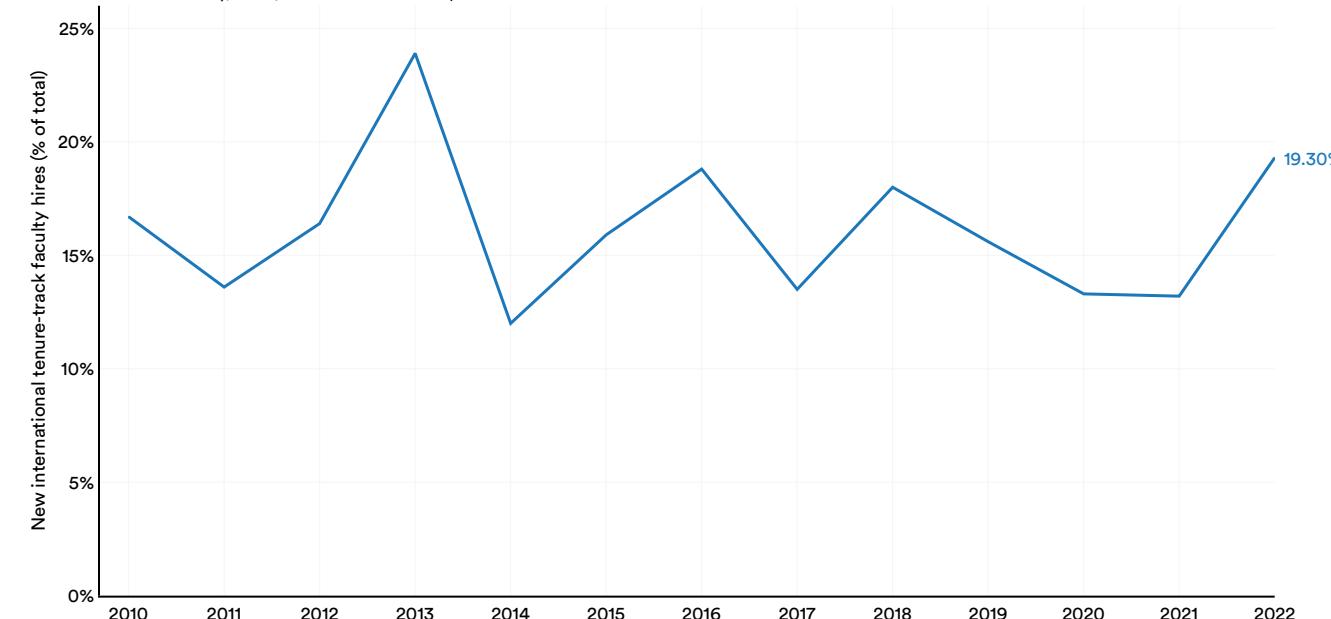


Figure 6.1.16

アメリカとカナダにおける新しい国際的なCS、CE、および情報 tenure-track 教員の採用割合（総数の%）、2010–22

Source: CRA Taulbee Survey, 2023 | Chart: 2024 AI Index report

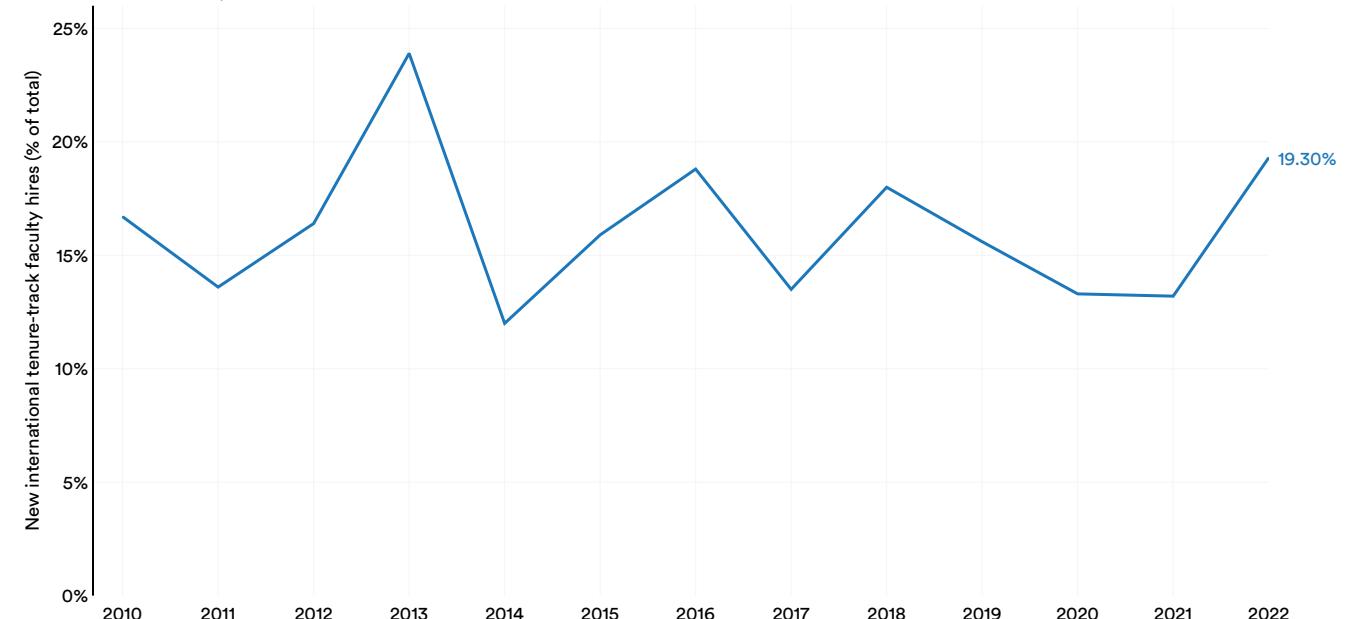


Figure 6.1.16

Europe

Data on European CS graduates comes from Informatics Europe, an academic and research community that, among other goals, monitors the state of informatics education in Europe.³ Informatics Europe gathers data on graduates in informatics, CS, CE, computing, and information technology (IT) disciplines from statistical offices of European governments.⁴

Informatics, CS, CE, and IT Bachelor's Graduates

In 2022, the United Kingdom led with the highest number of new graduates in informatics, CS, CE, and IT at the bachelor's level, totaling approximately 25,000 (Figure 6.1.17).⁵ Germany and Turkey followed closely. Most countries in the sample saw an increase in graduates in these fields compared to a decade ago, though there were exceptions like Poland, Spain, and the Czech Republic (Figure 6.1.18).

New informatics, CS, CE, and IT bachelor's graduates by country in Europe, 2022

Source: Informatics Europe, 2023 | Chart: 2024 AI Index report

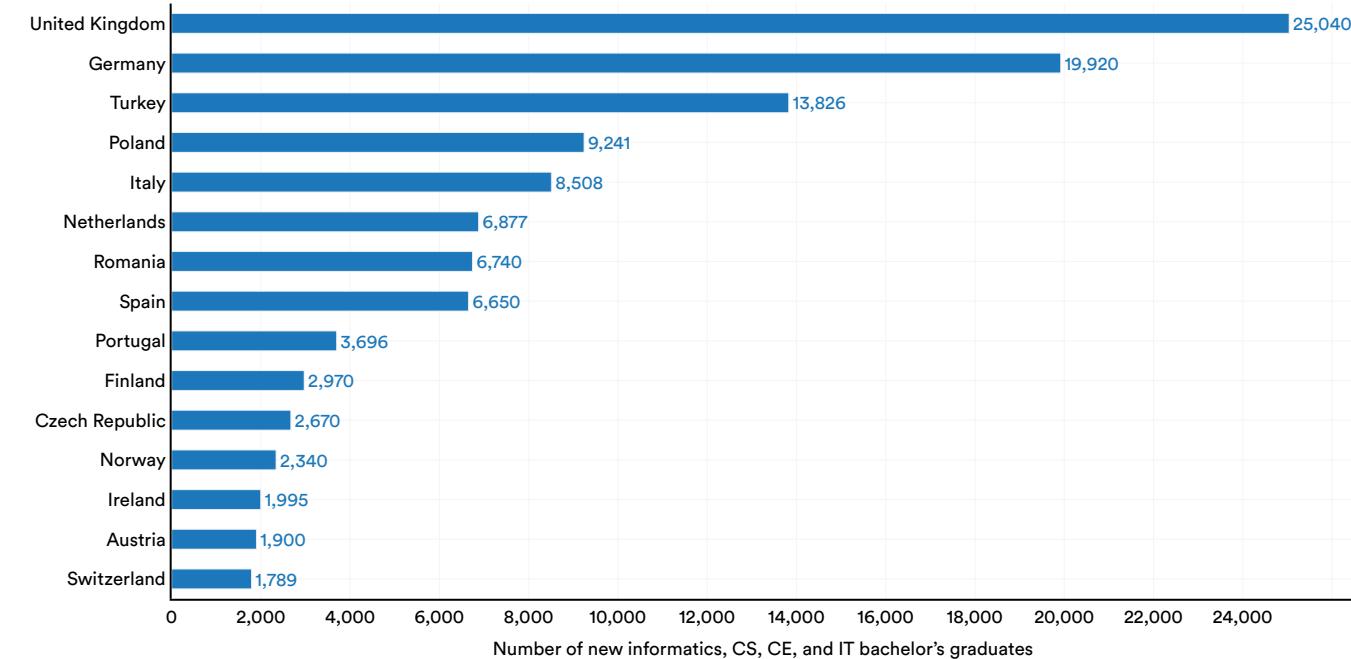


Figure 6.1.17

ヨーロッパ

ヨーロッパのコンピュータサイエンス 卒業生に関するデータは、情報学教育の現状を監視することを含む目標を持つ学術および研究コミュニティであるInformatics Europeから得られています。³ Informatics Europeは、ヨーロッパの政府の統計局から情報学、コンピュータサイエンス、コンピュータ工学、計算、情報技術（IT）分野の卒業生に関するデータを収集しています。⁴

情報学、コンピュータサイエンス（CS）、コンピュータ工学（CE）、および情報技術（IT）の学士卒業生

2022年、イギリスは情報学、CS、CE、およびITの学士レベルで新卒者数が最も多く、約25,000人に達しました（図6.1.17）。ドイツとトルコが続きました。サンプル内のほとんどの国は、10年前と比較してこれらの分野で卒業生が増加しましたが、ポーランド、スペイン、チェコ共和国のような例外もありました（図6.1.18）。

2022年のヨーロッパにおける新しい情報学、コンピュータサイエンス、コンピュータ工学、ITの学士卒業生数

出典: Informatics Europe, 2023 | チャート: 2024 AIインデックスレポート

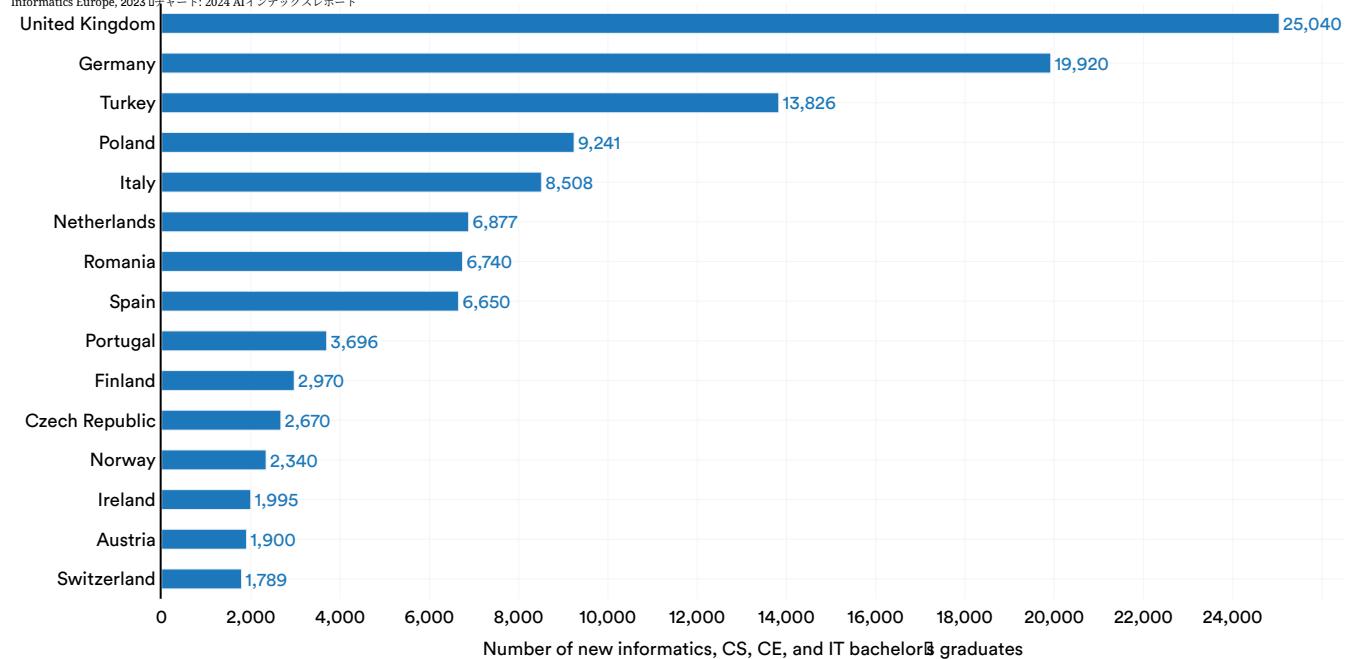


Figure 6.1.17

³ There is no singular term for CS education that is used uniformly across European countries. Across Europe, CS education can be reflected in terms such as informatics, computer science (CS), computer engineering (CE), computing, information technology (IT), information and communication technology (ICT), and information science and technology (IST). The full list of subject names (and English translations) that Informatics Europe uses to identify informatics studies programs can be found at the following link.

⁴ Readers are cautioned against making per capita comparisons between the CRA North American data and the European CS graduate data detailed in subsequent sections, as the European data is collected from national statistical offices and boasts broader coverage.

⁵ Note that not all countries for which the AI Index has data are visualized in the figures in this section. To access the complete data, please view the public data associated with this chapter. Moreover, the year label refers to the year in which an academic year ends. For example, the figures visualizing new graduates for 2022 reflect the number of graduates reported for the 2021/2022 academic year. For the sake of visual simplicity, the Index opts to focus on the year in which students graduated.

³ CS教育に関して、ヨーロッパの国々で一様に使用される単一の用語は存在しません。ヨーロッパ全体で、CS教育は情報学、コンピュータ工学（CE）、コンピューティング、情報技術（IT）、情報通信技術（ICT）、および情報科学と技術（IST）などの用語で表現されることがあります。情報学ヨーロッパが情報学の研究プログラムを特定するために使用する科目名の完全なリスト（および英語の翻訳）は、以下のリンクで確認できます。

⁴ 読者は、CRA北米データと後のセクションで詳述されるヨーロッパのCS卒業生データとの間で一人当たりの比較を行うことに注意してください。ヨーロッパのデータは国統計局から収集され、より広範なカバレッジを誇ります。

⁵ AIインデックスがデータを持つすべての国がこのセクションの図に視覚化されているわけではないことに注意してください。完全なデータにアクセスするには、この章に関連する公開データを参照してください。さらに、年のラベルは学年度が終了する年を指します。たとえば、2022年の新卒者を視覚化する図は、2021/2022学年度に報告された卒業生の数を反映しています。視覚的な単純さのために、インデックスは学生が卒業した年に焦点を当てることを選択しています。

Percentage change of new informatics, CS, CE, and IT bachelor's graduates by country in Europe, 2012 vs. 2022

Source: Informatics Europe, 2023 | Chart: 2024 AI Index report

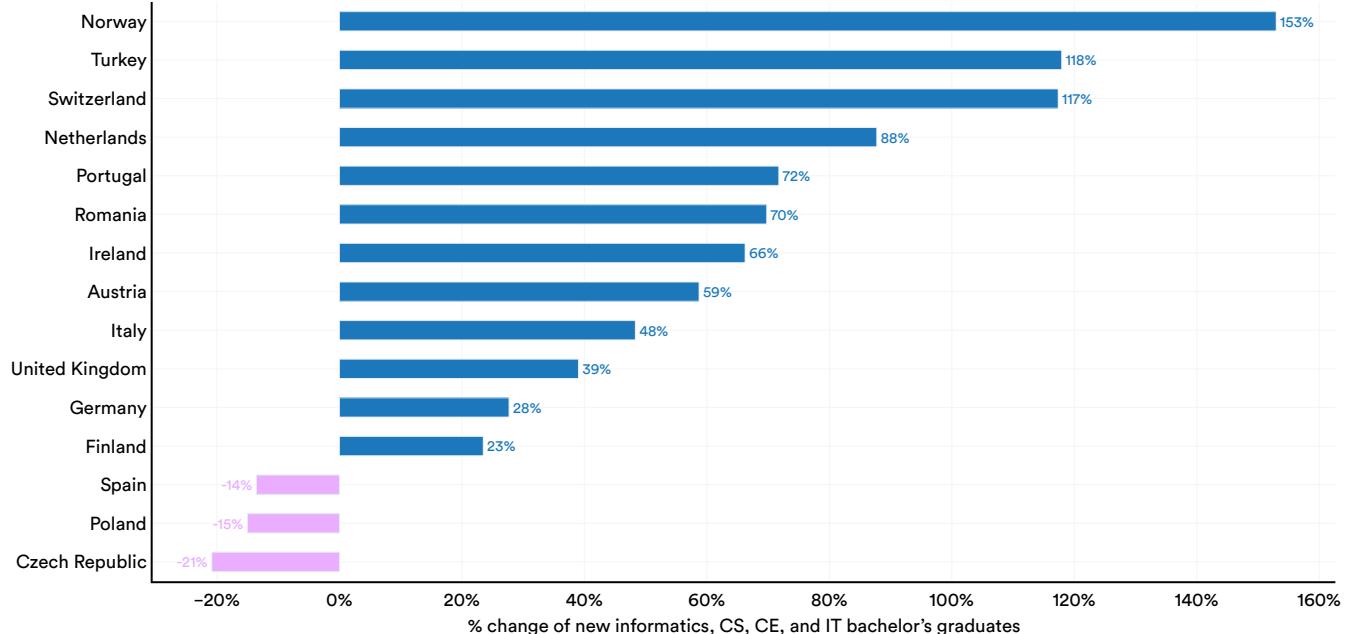


Figure 6.1.18

Finland (53.4), Norway (42.6), and the Netherlands (38.6) lead in the number of new bachelor's graduates in informatics CS, CE, and IT per 100,000 inhabitants (Figure 6.1.19). On a per capita basis, most sampled European countries have seen increases in the total number of informatics, CS, CE, and IT bachelor's graduates (Figure 6.1.20).

国別の新しい情報学、CS、CE、およびITの学士卒業生の割合の変化、2012年対2022年

出典: インフォマティクス・ヨーロッパ, 2023 チャート: 2024 AIインデックスレポート

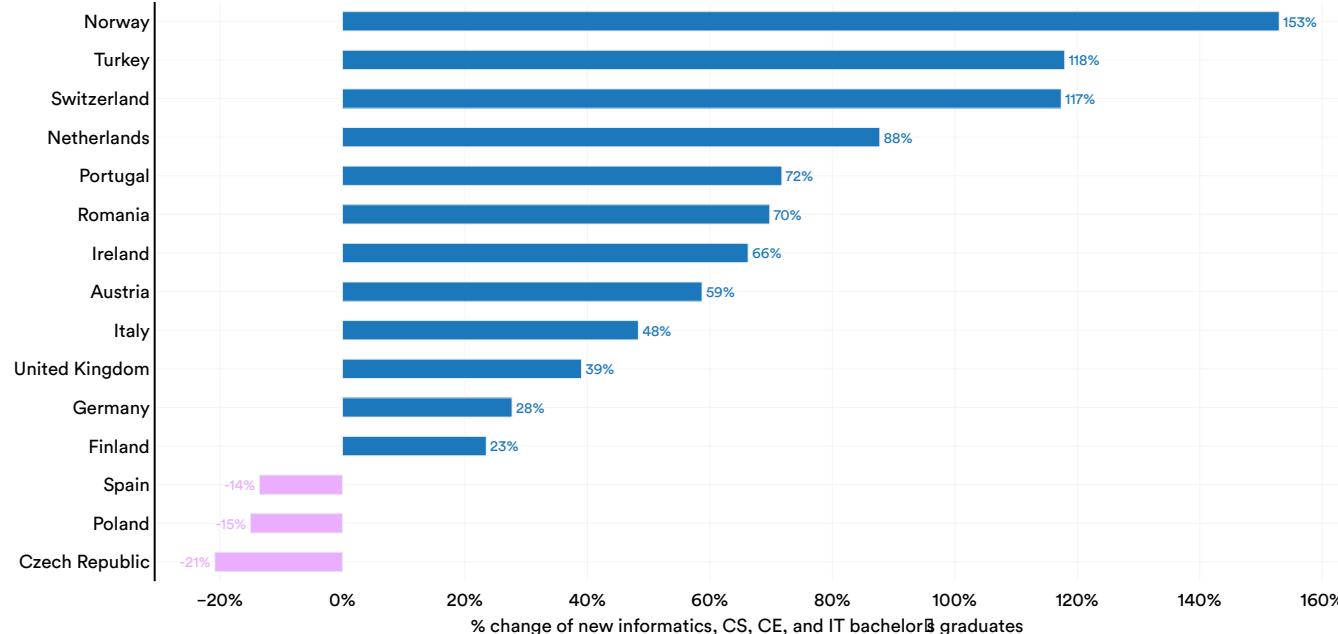


Figure 6.1.18

フィンランド（53.4）、ノルウェー（42.6）、オランダ（38.6）は、10万人あたりの情報学、CS、CE、ITの新しい学士卒業生の数でリードしています（図 6.1.19）。サンプルに含まれるほとんどのヨーロッパ諸国は、情報学、CS、CE、ITの学士卒業生の総数が増加しています（図 6.1.20）。