

AIが研究開発に与える変革的インパクト

エグゼクティブサマリー

AIのR&Dへの影響は、もはや「文章作成や要約の補助」とどまりません。現在の主戦場は、文献探索、仮説生成、実験計画、シミュレーション、データ取得、解析、ラボ自動化、再現性管理、査読補助、出版実務までを含む研究プロセス全体です。日本の科学技術振興機構¹の2026年報告は、AI for Scienceを科学の「第5のパラダイム」と位置づけ、OECD²も、科学生産性の加速はAIの最も社会的・経済的価値の高い用途になり得ると整理しています。³

足元で最も実用価値が高いのは、第一に文献探索・レビューの高速化、第二に構造予測やサロゲートモデルによる *in silico* 研究、第三にベイズ最適化や強化学習を使った閉ループ実験、第四にクラウドラボや自律実験系の運用です。たとえば、Semantic Scholar は2.3億本超の論文検索を支え、Elicit は500万人超の研究者に使われ、事例上では文献作業を10倍高速化し、系統レビュー用データ抽出で99.4%精度を示しています。PaperQA2 は文献探索タスクで「superhuman」と報告されました。⁴

定量的なインパクトは、個別事例でもすでに大きいです。A-Lab は17日間で57候補中36材料の合成に成功し、モバイル・ロボット化学者は8日で688実験を回して初期配合の6倍活性の光触媒を見つけました。GraphCast は10日予報で1,380の評価対象の90%以上で既存のHRESを上回り、単一のGoogle TPU v4で10日予報を1分未満で生成します。AlphaFold 系は2億超のタンパク質構造を共有し、構造生物学の探索コスト構造を変えました。⁵

ただし、「自律的AI科学者」はまだ限定的です。現在の最強システムは、計算機上で完結する研究領域、あるいは装置・評価関数・安全境界が明確な閉ループ問題で強く、開放的で曖昧な実世界研究を全面自律で回す段階には達していません。GoogleのAI co-scientistは、専門家評価で新規性・インパクトが高いと見なされ、AMLや肝線維化で有望な仮説を実験検証しましたが、いずれも expert-in-the-loop です。The AI Scientist はアイデア生成から査読までを自動化し、ワークショップ査読を通過した論文例を示しましたが、現時点では主戦場が機械学習研究です。⁶

したがって、いま研究機関や企業が採るべき戦略は、「全面自律化」ではなく「段階的自動化」です。具体的には、情報探索、サロゲート予測、パラメータ探索、プロトコル化しやすい反復実験から先に導入し、根拠提示、出典接地、ログ保存、人間の承認ゲート、法務・IP・EHS・研究公正を組み込んだ統治設計を先行させるべきです。これはNational Institute of Standards and Technology⁷のAI RMF、経済産業省⁸のAI事業者ガイドライン、European Commission⁹のAI Act、そして査読・出版を担うInternational Committee of Medical Journal Editors¹⁰、Committee on Publication Ethics¹¹、Nature Portfolioの方針と整合的です。¹²

現在地と歴史的タイムライン

AI in R&Dの本質的な変化は、研究者の周辺作業を補助する「copilot」から、限定された問題設定の中で仮説生成・実験・評価・記述までを閉ループで回す「agentic system」への移行です。前者は「読む・探す・書く」の自動化、後者は「考える・試す・評価する」の一部自動化です。現時点で前者は広く実務化され、後者は分野限定で急速に実証が進んでいます。¹³

時期	主な段階	何が起きたか	R&Dへの意味	主な根拠
2015–2019	AI支援 検索・ 読解	Semantic Scholar のようなAI検索が普及し、論文探索が「キーワード一致」から「意味検索」へ移行。	研究のボトルネックが「見つける」から「選ぶ・統合する」へ変化。	Semantic Scholar はAI駆動で2.3億本超の論文を検索。 ¹⁴
2020–2021	予測科学の実用化	AlphaFold 2 が類似構造が未知でも原子精度級の構造予測を実現。	実験前の候補絞り込みと構造仮説形成が劇的に高速化。	AlphaFold 2 論文とAlphaFold解説。 ¹⁵
2020–2023	自律実験の台頭	モバイル・ロボット化学者、AlphaFlow、A-Lab が、バイズ最適化・RL・能動学習を伴う閉ループ実験を実証。	「試行錯誤」の速度と再現性が上がり、データ取得そのものがアルゴリズム化。	¹⁶
2023–2024	基盤モデルの研究工程侵入	Coscientist がLLMで合成計画からクラウドラボ制御まで6タスクを実証。AlphaFold 3 は核酸・小分子・イオンまで扱う。	研究支援AIが「読んで書く」だけでなく「装置を動かす・複数モダリティを統合する」段階へ。	¹⁷
2024–2025	研究エージェント化	PaperQA2、FutureHouse Platform、AI co-scientist のような、長い探索・統合・提案を行う研究エージェントが登場。	文献探索・仮説提案・研究案生成の“フロントエンド”がエージェント化。	¹⁸
2025–2026	限定領域のAI科学者	The AI Scientist / v2 が、アイデア生成から実験・原稿・自己査読までの end-to-end 自動化を提示。材料分野では概念グラフから新規研究方向予測も登場。	「AI scientist」は研究の全面置換ではなく、まず限定領域での自律化として現実味を帯びた。	¹⁹

このタイムラインから見えるのは、AIが研究の**入力側**では「情報の意味理解」を、**中核工程**では「探索戦略」を、**出力側**では「記述・評価」を担い始めたことです。特に2023年以降は、LLMとロボティクス、あるいはLLMとサロゲートモデルが結びついたことで、研究の“周辺自動化”から“実験ループ自動化”へ軸が移りました。²⁰

研究プロセス別の適用マップ

研究プロセス全体を見ると、AIは単一万能モデルとしてではなく、各工程で異なる役割を担う「連携スタッフ」として理解するのが正確です。JSTの整理どおり、AIは仮説生成、実験工程、データ解析、知識統合まで浸透していますが、各段階の成熟度とリスクは均一ではありません。²¹

研究段階	主要なAI能力	代表技術・例	現在の 実用度	主な注意点	根拠
仮説生成	文献横断要約、アナロジー探索、新規テーマ提案	AI co-scientist、材料概念グラフ予測	中	新規性評価の主観性、既存バイアスの再生産	²²

研究段階	主要なAI能力	代表技術・例	現在の 実用度	主な注意点	根拠
文献探索・レビュー	意味検索、データ抽出、レビュー自動生成	Semantic Scholar、Elicit、PaperQA2	高	出典漏れ、抽出誤差、要約の過度一般化	4
実験計画	合成計画、パラメータ提案、プロトコル生成	Coscientist、BAX系、量子meta-design	中	実験可能性・安全性・装置制約の見落とし	23
シミュレーション・予測	構造予測、サロゲート予測、逆設計	AlphaFold 3、GraphCast、材料マルチモーダルLLM	高	外挿に弱い、機構理解が伴わないことがある	24
データ取得	ロボット実験、クラウドラボ、オンライン計測	A-Lab、ANL、Emerald Cloud Lab	中～高	装置連携、異常検知、現場安全、保守	25
解析・意思決定	高次元探索、異常検知、因果推論、最適化	ベイズ最適化、強化学習、PDGrapher、因果深層学習	中	相関と因果の混同、評価指標の設計依存	26
ラボ自動化	実験実行の連続自動化	AlphaFlow、ANL、クラウドラボ	中	標準化されていない湿式操作で失敗しやすい	27
再現性	protocol-as-code、ログ化、標準化された自動実行	ボタン一発再現、クラウドラボ、結果ログ化	中	既定値や前処理の隠れた変更	28
査読	形式評価、論点抽出、初期スクリーニング	Automated Reviewer、査読支援LLM	低～中	機密漏えい、誤評価、lazy reviewing	29
出版	原稿整形、言語改善、AI使用開示	著者責任・AI使用開示ルール	中	AIを著者にできない、人間が最終責任	30

現在の最適解は、各段階に合う能力を差し込むことです。文献探索ではRAGと引用整合性が重要で、実験設計では能動学習と安全制約付き最適化が重要です。ラボ運用では、LLM単体よりも、プロトコル言語、装置API、ベイズ最適化、異常検知を束ねたシステム設計の方が成果に直結します。³¹

下図は、研究開発におけるAI活用の典型的な閉ループを模式化したものです。現実の高性能システムは、このループ全体ではなく、特定区間を高精度で自動化して価値を出しています。³²

flowchart TD

```

A[研究課題の定義] --> B[文献探索・知識統合]
B --> C[仮説生成]
C --> D[実験計画・条件提案]
D --> E[シミュレーション / サロゲート予測]
E --> F{人間の承認ゲート}
F -->|承認| G[自動実験・クラウドラボ]

```

F --> |差戻し| C
G --> H[データ取得・前処理]
H --> I[解析・因果推論・最適化]
I --> J[再現性ログ / プロトコル保存]
J --> K[原稿作成・内部レビュー]
K --> L[公開判断]
L --> |公開| M[投稿・査読・出版]
L --> |再設計| C

主要技術と方法論

言語モデルと研究エージェント

LLMの最も即効性の高い価値は、論文検索、要約、比較、そして“問いの言い換え”にあります。Semantic Scholarは大規模意味検索を、Elicitは質問起点のレビュー生成とデータ抽出を、PaperQA2は文献探索・要約の高精度エージェント化を進めています。ここで重要なのは、単なる生成能力ではなく、**検索、引用、検証**を結びつけることです。研究用途でRAGがほぼ必須なのは、ハルシネーションが依然として一般解では解消されていないからです。 ³³

AI co-scientist や The AI Scientist が示したのは、LLMが単独で“科学をする”というより、**複数エージェントによる役割分担**が研究支援に向くという点です。AI co-scientist は仮説案を競わせて評価し、The AI Scientist はアイデア生成、実験、論文執筆、査読を分割して扱います。これは、研究が本質的に多段・多役割タスクであることを反映しています。 ³⁴

基盤モデルとマルチモーダル科学AI

科学分野では、基盤モデルは自然言語だけでなく、配列、構造、画像、スペクトル、時系列を統合する方向に進んでいます。AlphaFold 3はタンパク質・核酸・小分子・イオン・修飾残基を含む複合体を扱い、deepAMPはペプチド言語モデルで抗菌ペプチドを設計し、材料分野ではMatterChatのようなマルチモーダルLLMが構造情報と自然言語を結びつけ始めています。これは、研究者が「論文を読むAI」から「対象そのものを表現できるAI」に移っていることを意味します。 ³⁵

能動学習・ベイズ最適化・自律実験

湿式・実験系の中核技術は、LLMよりむしろ**能動学習とベイズ最適化**です。モバイル・ロボット化学者はbatched Bayesian searchで10変数空間を高速探索し、A-Labは失敗時にactive learningでフォローアップを提案し、ANLもBayesian optimizationで培地条件を更新しました。最近のBAX系研究も、科学者が「最適な単一点」ではなく「条件を満たす設計領域」を効率的に探索する必要に応えています。 ³⁶

生成モデル・因果推論・プログラム探索

生成モデルは分子・材料・タンパク質設計だけでなく、**プログラム空間探索**にも強みを示しています。FunSearchはLLMと評価器を組み合わせて数学問題の既知最良結果を更新し、量子実験のmeta-designは、単一解ではなく“無限クラスの対象を解くメタ解”をPythonコードとして生成しました。さらに、PDGrapherのような因果発見+幾何深層学習は、疾患状態を望ましい状態へ移す介入組合せの予測に踏み込みつつあります。つまり、AIは「予測」から「介入設計」へ拡張し始めています。 ³⁷

主要ツール・企業比較

領域	ツール / 会社	主な価値	実行可能範囲	主な制約・注意	根拠
構造予測	AlphaFold Server / Google DeepMind 38	タンパク質と他分子の相互作用予測、非商用研究向け無償アクセス	in silico 構造仮説形成	実験検証は依然必要、外挿限界あり	39
文献探索	Semantic Scholar / Allen Institute for AI 40	大規模意味検索、読解支援	幅広い文献調査	検索再現率・精度は問いの立て方依存	14
文献レビュー	Elicit	質問起点の論文抽出・比較・レビュー生成	系統レビュー、エビデンス抽出	抽出誤差の人手確認が必要	41
科学エージェント	FutureHouse Platform / FutureHouse 42	文献探索・統合・深いレビュー専門エージェント	生物・医学中心の知識探索	公称性能の独立検証は今後の課題	43
自律研究プロトタイプ	The AI Scientist / Sakana AI 44	アイデア生成～実験～論文作成～自己査読のend-to-end研究	主にML研究	現在は計算機上で完結する領域が中心、安全な実行環境が必須	45
クラウドラボ	ECL / Emerald Cloud Lab 46	リモート実験、標準化、自動計測	生命科学中心、材料も一部	装置適合性、コスト、ワークフロー制約	47

この比較から明らかなのは、研究現場で必要なのが「最強単独モデル」ではなく、**探索・予測・実行・記録**のどこを改善したいかに応じたスタック選択だということです。現場にとって重要なのは、精度だけでなく、再現性、APIの開放性、ログ化、権限管理、そして既存LIMS/ELNとの接続可能性です。 48

学術・産業ケーススタディ

以下は、学術・産業をまたぐ代表的事例です。重要なのは、AIの価値が「万能性」ではなく、**問題設定が明確なほど高く、閉ループ化できるほど大きい**という点です。分野横断で見ると、生命科学と材料・化学が現在の先行領域で、社会科学と工学・情報学は“研究方法の戦略的補助”として伸びています。 49

分野	事例	主な成果・指標	含意	根拠
生命科学	AlphaFold / AlphaFold DB	AlphaFold 2 は類似構造が未知でも原子精度級予測を実現。DBは2億超の構造を公開し、190超の国・地域、300万人超が利用。	構造決定のボトルネックを“測る”から“予測して検証する”へ移した。	50
生命科学・創薬	deepAMP 抗菌ペプチド設計	18個のTier1候補と11個のTier2候補を実験評価し、90%以上が penetratin より強い阻害。T2-9はFDA承認薬に匹敵。	少量データでも基盤モデルで候補生成→実験検証を高速化できる。	51

分野	事例	主な成果・指標	含意	根拠
材料	GNoME	グラフネットワークを大規模学習し、材料探索効率を1桁改善。2.2百万超の新規結晶候補、381,000超の安定候補を予測。	生成・探索の“母集団”を一気に拡張。実験側の律速を強く意識させた。	52
材料	A-Lab 自律合成	17日間連続運転で57ターゲット中36材料の合成に成功。失敗時はactive learningで再提案。	予測材料の実証ギャップを埋める、自律実験の代表例。	53
化学	モバイル・ロボット化学者	10変数空間で8日間に688実験を実施し、初期配合の6倍活性の光触媒を発見。	ベイズ探索は、人間の局所探索を上回る。	54
化学・材料	AlphaFlow	最大40パラメータの多段反応で、従来系列を上回る新規ルートを強化学習で見。	多段・疎データ化学でRLが有効であることを示した。	55
物理・地球科学	GraphCast	1,380の評価対象の90%以上でHRESより高精度。対流圏では99.7%で上回り、10日予報を1分未満で生成。	高価な物理シミュレーションを補完するサロゲートの経済性が非常に高い。	56
物理	量子実験 meta-design	20対象クラス中16クラスは人間がmeta-solutionを知らなかったが、Transformerが“無限クラスを解くメタ解”を生成。	AIが単一設計だけでなく、設計原理の抽出に踏み込んだ。	57
工学・ソフトウェア	AlphaDev	人間ベンチマークを上回るソートアルゴリズムを発見し、LLVM標準C++ sortライブラリに統合。	R&D対象が実験室の物質だけでなく、アルゴリズム自体であることを示した。	58
社会科学	Homo silicus	4つの実験室実験と1つのフィールド実験をLLMエージェントで再演。公平性実験では2,400観測を生成し、実験は数分・ごく小額で再現可能。	人間実験の前段に、安価な in silico パイロット研究が入る可能性。	59
バイオ工学	ANL	培地最適化でM9基本培地比94.7%の細胞増殖改善。10mL/100mLでも成長率・最大成長が改善。	日本発のモジュラー自律ラボが、培養→前処理→測定→解析の閉ループを実証。	60
生物医学	AI co-scientist	11研究目標の小規模評価で、専門家が新規性・インパクトを高く評価。AML候補薬は複数細胞株で実験検証、肝線維化では $p < 0.01$ の有望活性。	“AI科学者”というより、強い仮説生成・優先順位付け助手として現実的。	61
ML研究	The AI Scientist / v2	研究ライフサイクル全体を自動化し、ワークショップで平均人間受理閾値を超える論文例を提示。Automated Reviewerは人間査読と同等の判定性能を報告。	完全自律研究は、まず計算機上で閉じた研究から現実化する。	62

これらの事例を横並びにすると、現在のAIが強いのは、**評価関数が明確で、観測が高速で、反復回数を大きく取る領域**です。逆に、真に開放的な問題設定、複雑な安全制約、定性的判断が多いテーマほど、人間の設計・監督の重要性が残ります。 ⁶³

便益とリスク

AIの最大便益は、速度、探索幅、コスト、再現性の四つです。速度面では、GraphCastのようなサロゲートが高価な計算を圧縮し、Elicitのようなレビュー支援が知識統合作業を短縮し、ラボ自動化が24時間運転を可能にします。探索幅の面では、GNoMEやA-Labのように「候補母集団の拡大」と「実験ループの高速化」を同時に進めることで、ヒトだけでは到達しづらい探索が現実化します。再現性の面では、クラウドラボやボタン一発再現型の実験設計が、プロトコル逸脱や暗黙知依存を減らします。 ⁶⁴

よりマクロに見ると、AIの研究生産性効果はすでに観測されています。Nature Human Behaviour の分析では、科学研究におけるAI利用と便益は2015年以降に急速に拡大しており、Nature 2026の分析では、AIを活用する科学者は、そうでない科学者よりも3.02倍多く論文を出し、4.84倍多く引用され、1.37年早く研究プロジェクトのリーダーになると報告されました。 ⁶⁵

ただし、その便益は無条件ではありません。同じNature 2026の研究は、AI導入が研究トピックの集合的多様性を4.63%縮小し、研究者同士の関与を22%減らすと報告しています。MesseriとCrockettの議論が指摘したように、AIは「わかった気にさせる」ことで、方法や問いの単一化、つまり科学のモノカルチャー化を進める危険があります。生産性向上と知的多様性の低下が同時に起こり得る点が、この変革のもっとも重要な逆説です。 ⁶⁶

リスクは少なくとも六つあります。第一に、ハルシネーションです。Nature 2024 は、真実性強化は部分的成功にとどまり、confabulation 検出のために semantic entropy のような統計的方法が必要だと示しました。第二に、バイアスです。77モデル評価では、ほぼすべてのbase modelと一部のinstruction-tuned modelで内集団びいきと外集団敵視が確認されました。第三に、安全です。AI scientist が病原体操作や化学条件設定を誤ると、バイオセーフティや化学安全に直結し得ます。第四に、便益の不均等配分で、AIの研究便益は分野・属性で偏り得ることが示されています。 ⁶⁷

第五に、知財・データ統治です。日本の文化庁 ⁶⁸ は、2024年に「AIと著作権に関する考え方」を公表しましたが、それ自体は法的拘束力を持たず、判例蓄積や技術進展に応じて見直しが見込まれています。EUではAI Actが2024年8月に発効し、2025年8月からGPAI提供者に技術文書、著作権方針、学習内容要約の公表などを課し、systemic risk モデルには追加的なリスク評価・事故報告・サイバー防護を求めています。日本でもMETIのAI事業者ガイドライン第1.2版が、2026年3月に更新されました。研究現場でのモデル調達や内製化は、もはや法務と分離できません。 ⁶⁹

第六に、査読と出版の制度的負荷です。Nature Portfolio は、査読者がAIを使う場合の透明な申告を求め、原稿中の機密・専有情報が査読プロセス外へ漏れる危険も明示しています。ICMJE と COPE は、AIを著者にできず、最終責任は人間が負うべきだとしています。The AI Scientist 自身も、こうした自律システムが査読システムにノイズを加えうると認めています。つまり、研究の自動化は、知識生産だけでなく、**知識の審査制度**にも負荷を与えます。 ⁷⁰

導入ベストプラクティス

R&D現場でのAI導入は、ツール導入ではなく運用設計です。成功している組織は、まず研究工程を分解し、低リスク・高反復・高データ密度の工程からAIを入れています。典型的には、文献レビュー、in silico 予測、パラメータ最適化、標準化された測定、プロトコル管理が先行領域です。逆に、病原体・規制物質・新規合成・対外公開など高リスク工程では、人間による承認を必須ゲートにしています。 ⁷¹

技術面では、第一に**出典接地**が必要です。文献系タスクはRAGや引用付き回答を前提にし、根拠不十分なら「不明」と言える設計にすること。第二に**不確実性管理**が必要です。semantic entropyのような不確実性検知、モデル間照合、シミュレーションと実験の二重化が有効です。第三に**provenance**を取ることです。プロンプト、モデル版、データ版、実験条件、装置ログ、評価指標を残し、あとから同じ結果を再演できるようにすることが、研究公正と製品化の両方で重要です。第四に**サンドボックス化**です。コード実行系や装置制御系は最小権限・ネットワーク制限・危険操作のホワイトリスト化を前提にすべきです。 72

組織面では、AIガバナンスを研究現場に埋め込む必要があります。NIST AI RMFはAIリスク管理を文脈依存・反復的なものとして整理し、METIのガイドラインはチェックリストやワークシートを含む実務化を進めています。研究機関や企業では、PI、研究DX、情報セキュリティ、EHS、法務/IP、研究公正、出版支援の横断チームを作り、モデル選定・データ利用・安全審査・公開判断を一つの仕組みに載せるべきです。 73

また、出版・査読に関しては、AI使用の内部ポリシーを明文化する必要があります。最低限、誰が出力を確認したか、どの範囲でAIを使ったか、機密文書を外部モデルへ送ってよいか、AI生成画像や引用はどう扱うか、査読でAI利用時にどう開示するかを定めるべきです。これはNature、ICMJE、COPEの方針とも整合的です。 74

AI導入の成否は、ステークホルダー関係の設計にも左右されます。下図は、研究現場で最低限意識すべき関係を示したものです。研究者だけでなく、AI提供者、データ提供者、クラウドラボ、機関、規制当局、出版社、資金配分機関、IP/法務、倫理・安全委員会が同時に関与します。 75

```
erDiagram
    RESEARCHER }o--|| PI : reports_to
    PI ||--|| INSTITUTION : belongs_to
    RESEARCHER }o--o{ DATA_PROVIDER : uses
    RESEARCHER }o--o{ CLOUD_LAB : runs_experiments_on
    RESEARCHER }o--o{ AI_PROVIDER : uses_models_from
    INSTITUTION ||--o{ ETHICS_EHS_BOARD : appoints
    INSTITUTION ||--o{ IP_LEGAL : governs
    INSTITUTION ||--o{ PUBLISHER : submits_to
    FUNDER ||--o{ INSTITUTION : funds
    REGULATOR ||--o{ AI_PROVIDER : oversees
    REGULATOR ||--o{ INSTITUTION : regulates
    PUBLISHER }o--o{ RESEARCHER : reviews_and_publishes
    IP_LEGAL }o--o{ AI_PROVIDER : contracts_with
    ETHICS_EHS_BOARD }o--o{ CLOUD_LAB : approves_use
```

自律的AI科学者の将来展望

現状証拠に基づく私の評価は、**狭い意味での自律的AI科学者はすでに成立し始めているが、広い意味での汎用自律科学者はまだ遠い**、というものです。The AI Scientistは計算機上の研究でend-to-end自動化を示し、AI co-scientistは生物医学で強い仮説生成・優先順位付け能力を示し、CoscientistやA-Lab、ANLは装置制御付きの閉ループ研究を実証しました。一方で、これらはすべて、対象領域、評価関数、安全境界、装置制約のいずれかが明確に定義された環境で動いています。ここから逆算すると、今後の進展は「全面人間代替」ではなく、「高価値な局所自律化の連結」として進む可能性が高いです。 76

実現可能性と時間軸

期間	実現可能性の高い姿	主要な技術要件	主な統治要件	根拠
今後 1～3 年	文献探索、コード生成、サロゲート予測、ベイズ最適化、限定装置制御を束ねた「半自律研究ループ」	RAG、装置API、実験ログ、評価関数、異常検知	人間承認ゲート、機密管理、査読・著者責任ポリシー	77
今後 3～7 年	分野限定のAI scientist。材料、化学、創薬初期探索、ML研究で“企画→実験→記述”の大部分を自律化	マルチモーダル基盤モデル、長期計画、因果推論、信頼較正、ロボット統合	事故報告、権限制御、IP来歴、モデル評価基準	78
7年 以上	分野横断の汎用AI科学者。ただし現在の証拠からは不確実性が高い	世界モデル、機構理解、自己評価、社会的文脈理解、複数研究目的の調停	研究公正・出版制度・資金配分・責任帰属の再設計	79

この見立ての背景には、現在のボトルネックがモデル性能だけでなく、**機械可読な科学知識、装置の標準インターフェース、安全制約の表現、そして評価制度**にあるからです。AI co-scientist が今後の改善点として文献レビュー強化、factuality checking、external tools cross-check、larger-scale evaluation を挙げていること、The AI Scientist が基盤モデル改善とテスト時計算量の増加に性能改善を依存していること、ANLが複雑な生産物最適化ではまだ苦戦していることは、そのまま技術課題の地図でもあります。⁸⁰

必要な技術進歩は、少なくとも五つあります。第一に、出典・データ・プロトコル・装置状態を同時に扱う**科学特化マルチモーダル基盤モデル**。第二に、単なる相関最適化ではなく、介入設計に近づく**因果推論と不確実性較正**。第三に、wet lab と計算機実験の両方を扱える**長期計画能力**。第四に、実験失敗や危険兆候を検知して止まれる**安全制御**。第五に、人間研究者に理解可能な形で理由を示す**説明可能性と監査可能性**です。⁸¹

ガバナンスモデルとしては、私は三層構造が妥当だと考えます。第一層は「copilot」、つまり文献・記述・初期分析の支援。第二層は「bounded autonomy」、つまり条件探索や標準化実験の自律運用。第三層は「delegated autonomy」、つまり明示的な責任分界と事故報告を伴う限定的委任です。EU AI Act のGPAI義務、METIのガイドライン、NatureやICMJEの著者・査読ルールは、この三層のどこまで任せるかを制度化するための実務基盤になり得ます。⁸²

未解決論点と限界

現時点で未解決なのは、第一に「新規性」をどう定量評価するかです。AIが既存文献の再結合をしているのか、本当に新しい因果メカニズムや設計原理を生み出したのかは、まだ厳密な一般評価法がありません。第二に、AI生成研究物の責任帰属、著作権、発明者性、貢献の配分です。第三に、AIが論文供給を加速したとき、査読・資金配分・研究評価制度が耐えられるのかです。これらは技術課題というより、研究ガバナンスの課題です。⁸³

総じて言えば、AIはR&Dを「研究者を置き換える装置」よりも、「研究の律速段階を別の場所に移す装置」として理解するのが正確です。すでに速くなっているのは読むこと、比べること、予測すること、回すことです。今後問われるのは、その速度増加を、発見率、再現性、安全性、知的多様性の改善につなげられるかどうかです。そこに成功した組織だけが、AIを本当に研究力へ変換できるはずで。⁸⁴

1 18 68 72 <https://www.futurehouse.org/research-announcements/wikicrow>
<https://www.futurehouse.org/research-announcements/wikicrow>

2 11 28 59 https://www.nber.org/system/files/working_papers/w31122/w31122.pdf
https://www.nber.org/system/files/working_papers/w31122/w31122.pdf

3 21 <https://www.jst.go.jp/crds/pdf/2025/RR/CRDS-FY2025-RR-05.pdf>
<https://www.jst.go.jp/crds/pdf/2025/RR/CRDS-FY2025-RR-05.pdf>

4 9 13 14 33 <https://www.semanticscholar.org/>
<https://www.semanticscholar.org/>

5 25 32 38 49 53 71 <https://www.nature.com/articles/s41586-023-06734-w>
<https://www.nature.com/articles/s41586-023-06734-w>

6 8 10 22 34 61 77 80 <https://research.google/blog/accelerating-scientific-breakthroughs-with-an-ai-co-scientist/>
<https://research.google/blog/accelerating-scientific-breakthroughs-with-an-ai-co-scientist/>

7 47 48 <https://www.emeraldcloudlab.com/>
<https://www.emeraldcloudlab.com/>

12 44 73 <https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf>
<https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf>

15 50 <https://www.nature.com/articles/s41586-021-03819-2>
<https://www.nature.com/articles/s41586-021-03819-2>

16 26 36 54 63 <https://www.nature.com/articles/s41586-020-2442-2>
<https://www.nature.com/articles/s41586-020-2442-2>

17 20 23 46 <https://www.nature.com/articles/s41586-023-06792-0>
<https://www.nature.com/articles/s41586-023-06792-0>

19 29 42 45 62 76 78 83 <https://www.nature.com/articles/s41586-026-10265-5>
<https://www.nature.com/articles/s41586-026-10265-5>

24 35 40 81 <https://www.nature.com/articles/s41586-024-07487-w>
<https://www.nature.com/articles/s41586-024-07487-w>

27 55 <https://www.nature.com/articles/s41467-023-37139-y>
<https://www.nature.com/articles/s41467-023-37139-y>

30 <https://www.icmje.org/recommendations/browse/artificial-intelligence/ai-use-by-authors.html>
<https://www.icmje.org/recommendations/browse/artificial-intelligence/ai-use-by-authors.html>

31 41 <https://elicit.com/>
<https://elicit.com/>

37 <https://www.nature.com/articles/s41586-023-06924-6>
<https://www.nature.com/articles/s41586-023-06924-6>

39 <https://deepmind.google/science/alphafold/>
<https://deepmind.google/science/alphafold/>

43 <https://www.futurehouse.org/research-announcements/launching-futurehouse-platform-ai-agents>
<https://www.futurehouse.org/research-announcements/launching-futurehouse-platform-ai-agents>

51 <https://www.nature.com/articles/s41467-024-51933-2>
<https://www.nature.com/articles/s41467-024-51933-2>

- 52 <https://www.nature.com/articles/s41586-023-06735-9>
<https://www.nature.com/articles/s41586-023-06735-9>
- 56 64 <https://deepmind.google/blog/graphcast-ai-model-for-faster-and-more-accurate-global-weather-forecasting/>
<https://deepmind.google/blog/graphcast-ai-model-for-faster-and-more-accurate-global-weather-forecasting/>
- 57 <https://www.nature.com/articles/s42256-025-01153-0>
<https://www.nature.com/articles/s42256-025-01153-0>
- 58 <https://www.nature.com/articles/s41586-023-06004-9>
<https://www.nature.com/articles/s41586-023-06004-9>
- 60 <https://www.nature.com/articles/s41598-025-89069-y>
<https://www.nature.com/articles/s41598-025-89069-y>
- 65 <https://www.nature.com/articles/s41562-024-02020-5>
<https://www.nature.com/articles/s41562-024-02020-5>
- 66 <https://www.nature.com/articles/s41586-025-09922-y>
<https://www.nature.com/articles/s41586-025-09922-y>
- 67 <https://www.nature.com/articles/s41586-024-07421-0>
<https://www.nature.com/articles/s41586-024-07421-0>
- 69 <https://www.bunka.go.jp/seisaku/chosakuken/aiandcopyright.html>
<https://www.bunka.go.jp/seisaku/chosakuken/aiandcopyright.html>
- 70 74 <https://www.nature.com/nature-portfolio/editorial-policies/ai>
<https://www.nature.com/nature-portfolio/editorial-policies/ai>
- 75 84 https://www.oecd.org/en/publications/artificial-intelligence-in-science_a8d820bd-en.html
https://www.oecd.org/en/publications/artificial-intelligence-in-science_a8d820bd-en.html
- 79 <https://www.nature.com/articles/s41586-024-07146-0>
<https://www.nature.com/articles/s41586-024-07146-0>
- 82 <https://digital-strategy.ec.europa.eu/en/factpages/general-purpose-ai-obligations-under-ai-act>
<https://digital-strategy.ec.europa.eu/en/factpages/general-purpose-ai-obligations-under-ai-act>