数学論文の1年倍増とエージェント型AIによる科学研究の構造変革：原典論文の徹底検証と今後の展

Gemini 4 Argon

YouTube動画『【数学論文が1年で2倍】AIが科学研究を加速させている』で取り上げられた現象は、インペリアル・カレッジ・ロンドンの理論物理学者カルロ・コンタルディ（Carlo R. Contaldi）教授が2026年10月25日にプレプリントサーバー「arXiv」へ投稿した論文『Is the arXiv Apocalypse upon us? A break in preprint growth coincident with agentic AI（arXivの黙示録は到来したのか？ エージェント型AIと符合するプレプリント成長率の屈折）』（arXiv:2610.34449）を原典としています[1]。同研究は、1991年7月から2026年10月までの約35年間にわたるarXivの全投稿記録とメタデータを統計的に解析し、2025年末から2026年初頭にかけて、コンピュータサイエンス（CS）以外の基礎科学分野——とりわけ数学——において歴史的な論文急増（成長率の構造的屈折）が発生したことを明らかにしました[2]。なお、動画内の音声・字幕では2026年9月の月間総投稿数が「4,363件」と読み上げられていますが、これは原典の「40,363件（4万363件）」の桁落ち（誤読）であり、実際にはarXiv史上最高となる月間4万件超の記録が樹立されています[1]。

この現象の核心は、2022年末のChatGPT登場以降に普及した「対話型LLM（チャットボット）」ではなく、2025年後半から実用化された「長期的な自律作業が可能なエージェント型AI（Agentic AI）」の普及と時期的に符合している点にあります[2]。2022年から2025年11月までの期間、対話型LLMが広く使われていたにもかかわらず、CS以外の分野の論文成長率は過去30年の歴史的平均（年率約8〜9%）と変わらず、arXiv全体の増加はもっぱらAI研究自体（cs.AI等）の膨張によるものでした[2]。しかし2025年12月以降、トレンドは大きく逆転します。AIカテゴリ自体の成長加速が止まる一方で、CS以外の分野が従来トレンドを17.1%超過（z = 16.2）し、中でも数学分野（CSとの重複を除く）は2025年9月の3,945件から2026年9月には8,203件へと1年で2.08倍（前年同月比108%増）に跳ね上がりました[3]。さらに、数学および物性物理学では「単独著者（1人だけで執筆された）論文」が1年で約70%急増しており、AIエージェントが大学院生や共同研究者の役割を代替し、「1人の研究者が自律型AIを駆使して研究を完結させる」体制への移行が進んでいることを示唆しています[1]。ただし著者自身が強調している通り、この分析は投稿数とメタデータの時系列変化に基づく観察研究であり、個々の論文がAIで生成されたことの直接的証明や、増加した論文の学術的妥当性・品質を保証するものではありません[3]。

今後の科学研究と学術エコシステムを展望すると、単なる投稿数制限（arXivが2026年10月1日に導入した月2件の上限規制）では論文の奔流を抑え込めないことがすでに10月時点の実績データで示されており[3]、学術界は「生成の爆発」に対応する「検証の自動化・厳格化」へと舵を切っています。具体的には、Google Researchが2026年10月14日に一般公開したマルチエージェント査読支援システム「Deep Review」のようなAIによる事前論理検証の普及[7]、定理証明支援系「Lean 4」を用いた数学的証明の形式検証（Formal Verification）の活用[8]、さらには形式検証だけでは見抜けない「既存研究の焼き直し（新規性の欠如）や意味的乖離」を判定する自動パイプライン（AViD Journal等）の構築が進展する見通しです[10]。論文の本数が研究者の評価指標として機能不全に陥る中、人間の

科学者に求められる中核的価値は「計算や証明作業の遂行」から「本質的な問いの設定（概念的枠組みの創出）」と「AIが生成した膨大な結果の真価を見極める審美眼・評価能力」へとシフトしていくと考えられます[1]。

1. 動画内容の徹底検証と原典論文（Contaldi, 2026）の特定・ファクトチェック

対象動画の内容を精査すると、すべての統計データ、分析手法、および著者のコメントは、インペリアル・カレッジ・ロンドン理論物理学部門のカルロ・コンタルディ教授による単著論文『Is the arXiv Apocalypse upon us? A break in preprint growth coincident with agentic AI』（2026年10月25日投稿、arXiv:2610.34449 [physics.soc-ph]）および、arXiv運営事務局が2026年10月1日に発表した公式レート制限ポリシー（Updated Rate Limit Policy）に基づいています[1]。
動画のスクリプト（自動文字起こし含む）には、数値の桁落ちや音声認識に起因する同音異義語の変換ミスが複数含まれているため、原典論文および公式発表の正確なデータとの対照・補正を行うことが現状理解の第一歩となります[1]。

検証項目	動画内の言及（字幕・音声）	原典論文（**Contaldi, 2026**）および公式記録の正確な内容	検証結果・補足解説
2026年**9**月の月間総投稿数	「2026年9月の月間投稿数が**4363**件という過去最高を記録」[1]	**40,363**件（1991年の創設以来、過去最高の月間記録。2024年9月は20,569件、2016年9月は9,869件）[2]	数値の誤読（桁落ち）。「40,363」のゼロが抜けて「4,363」と読まれている。なお数学単体でも9月は8,203件に達している[3]。
過去**30**年の成長率と**2022**～**2025**年の非**CS**成長率	「過去約30年間の平均成長率は約9%」「2022年から2025年までの情報科学以外の年間成長率も約8%」[1]	1991年以降の長期指数成長率は約**9%/**年（2010年1月～2022年11月のフィットで**9.0%/**年）。2022年1月～2025年11月の非CS（all - cs.*）成長率は**8.1%/**年[2]	正確。対話型LLM普及期（2023～2025年）には、CS以外の論文数には統計的な加速痕跡が見られなかった[2]。
2025年**12**月以降の非**CS**超過率と数学の倍増	「25年12月以降はCS以外の投稿が予測トレンドを**17%**上回る」「数学分野はわずか**1**年で約**2**倍」	2025年12月～2026年9月の非CS分野のトレンド超過（Excess）は**+17.1%（**z = 16.2**）。数学は	正確。数学の成長率屈折点（Break）は2026年3月（ΔBIC = -123）に検出されている[3]。

	[1]	トレンド超過+26.5%（**z = 21.0**）、**2025**年**9**月の**3,945**件から**2026**年**9**月の**8,203**件へ2.08倍（+108% YoY）**[3]	
単独著者論文（**Solo-authored**）の増加率	「数学分野で単独著者論文が**70%**増」「物性物理学でも**70%**増、定量生物学でも**44%**増」[1]	2026年1〜9月の前年同期比で、単独著者論文は数学で**+70%、物性物理学（**cond-mat**）で+71%（論文要旨では約**70%**と表記）、定量生物学（**q-bio**）で+44%**（全体が-3.4%縮小する中で単独著者が急増）[2]	正確。一方、大型実験装置を伴う高エネルギー物理学実験（hep-ex）ではこのシフトが起きていない（z = 0.7）[2]。
数学論文の平均著者数と増加要因の内訳	「平均著者数も数学では**2.33**人から**2.09**人へ減少」「数学の投稿増加の約算は少数の多産な人ではなく異なる新しい著者の増加」[1]	数学の1論文あたり平均著者数は2025年9月の**2.33**人から2026年9月の**2.09**人へ減少[3]。論文増加の**約3分の2（two-thirds）**は少数の多作投稿者ではなく、ユニーク著者数の拡大（8,267人→13,536人）に起因[3]	字幕の「約算」は「約**3**分の**2**」の音声認識ミス。また動画内の「家を担当する」は「査読を担当する」の誤変換[1]。
論文執筆自体の**AI**エージェント実験	「使用モデルはクロードオーパス**5.5**」「API問い合わせ約**1100**回、ツール呼び出し**176**回、メタデータ約**1万2000**件、データ取得約**80**分、指示約**60**回、所要時間約**3**時間**45**分（従来なら最低1週間）」[1]	Anthropic社の**Claude Opus 5.5**をコード実行環境・ブラウザ・検索ツールを備えたエージェントとして使用[1]。ツール呼び出し**176**回、APIへのカウントクエリ約**1,100**回、取得メタデータ約**12,000**件[1]。最初の質問から初	正確。コンタルディ教授自身が「エージェント型AIによる研究支援」を自らの論文作成プロセスで試行した事例となっている[2]。

		稿完成まで約**3**時間**45**分[1]	
arXivの投稿制限措置	「2026年10月1日から1アカウントあたりの新規投稿を月**2**件までに制限」[1]	2026年10月1日発効。1投稿アカウントあたり暦月**2**件まで（同時に進行中[Active]の投稿は最大3件まで）。リジェクトされた論文も月間枠を消費[3]	正確。ただし後述の通り、この制限は「投稿操作を行うアカウント」にのみ適用されるため抜け道が存在する[3]。

コンタルディ論文自体が示した「3時間45分」のメタ実験

特筆すべきは、このarXivの投稿急増を分析したコンタルディ教授の論文自体が、まさにその要因と目される「エージェント型AI」の実行コストを測る単一の実証データ（A data point）として作成された点です[2]。同論文の謝辞および第V章第4節によれば、データ収集、統計解析、図表の作成、そして論文初稿の執筆に至る工程は、著者の指示のもとでAnthropic社の「**Claude Opus 5.5**」がコード実行環境およびWebブラウザ・検索ツールを備えた自律エージェントとして遂行しました[1]。

このセッションにおいて、AIエージェントは計**176**回のツール呼び出し（**tool invocations**）を実行し、**arXiv API**に対して約**1,100**回のカウントクエリを発行するとともに、2025年9月と2026年9月の数学論文全件を含む約**12,000**件のメタデータレコードを取得しました[1]。データのダウンロードには約80分を要しましたが、エージェントはその待機時間中に手を止めることなく、取得済みデータの解析や本文のドラフト執筆を並行して進めました[1]。人間側（コンタルディ教授）が行ったのは、研究の問いの設定や分析のレビューなど約**60**回のやり取り（プロンプト入力）にとどまり、最初の質問を投げてから論文の初稿が完成するまでの所要時間は約**3**時間**45**分でした[1]。コンタルディ教授は、従来のやり方であれば少なくとも1週間の作業になったと推計しており、人間が問いを設定してレビューに専念することで、データ取得から初稿作成までの時間が大幅に短縮されることを具体的に示しています[1]。

2. 現状分析①：arXivにおける「2つの構造的転換点」と数学論文倍増のメカニズム

arXivの月間投稿数は1991年の創設以来約35年間にわたり、季節変動を伴いながらも年率約9%（2010年1月～2022年11月の期間で年率9.0%、対数残差標準偏差 σ = 0.049）という安定した指数関数的成長を続けてきました[3]。過去のデータにLLM検出器（単語頻度分布や分類器）を適用して執筆量への影響を推計する先行研究も存在しますが、モデルの更新や著者の使い方の変化によって検出器の較正がずれるほか、モデルが解析コードを書いて結果を生成するようなエージェント型ワークフローではテキスト検出器の信頼性が最も低くなります[3]。また、アブストラクトが検出閾値を超えた最初の月を「AI導入月」と定義する著者レベルの研究には、たまたま投稿数が多い月に閾値を超えやすいという方法論上のバイアス（outcome-triggered timing bias）が指摘されていました[3]。そこでコンタルディ教授は、LLM検出ツールを一切使わず、純粋に投稿数とメタデータの時系列変化のみから構造変化点（Knot）を探索する手法を採りました[3]。

具体的には、月間投稿数 $N(t)$ の対数を以下の季節調整付き区分対数線形モデルで定式化して

います[2]。

$$\ln N(t) = a + b(t - t_0) + \sum_j c_j (t - \tau_j)\Theta(t - \tau_j) + \sum_{k=2}^{12} s_k \delta_{m(t),k} + w \ln \left(\frac{W(t)}{\bar{W}} \right.$$

ここで t は年単位の時間、Θ はヘビサイドの階段関数、τ_i は成長率が c_i だけ変化する構造変化点（Knot）、s_k は1月を基準とした各暦月 k の季節オフセット（11個の月次オフセットにより学期構造を含む年周期のフーリエ級数を吸収）、$W(t)$ はその月の平日数（平均 $\bar{W} = 21.75$ 日）を表します[2]。また、機械学習自体のブームによる混入（クロスリストによる重複カウント）を排除するため、コンピュータサイエンス以外の全カテゴリからは cs.* タグが1つでも付与された論文をすべて除外した統制群（No CS listing）が構築されました[3]。

数学および理論・数理分野が2025年末以降のarXiv論文急増を牽引し、AI関連カテゴリは頭打ちに

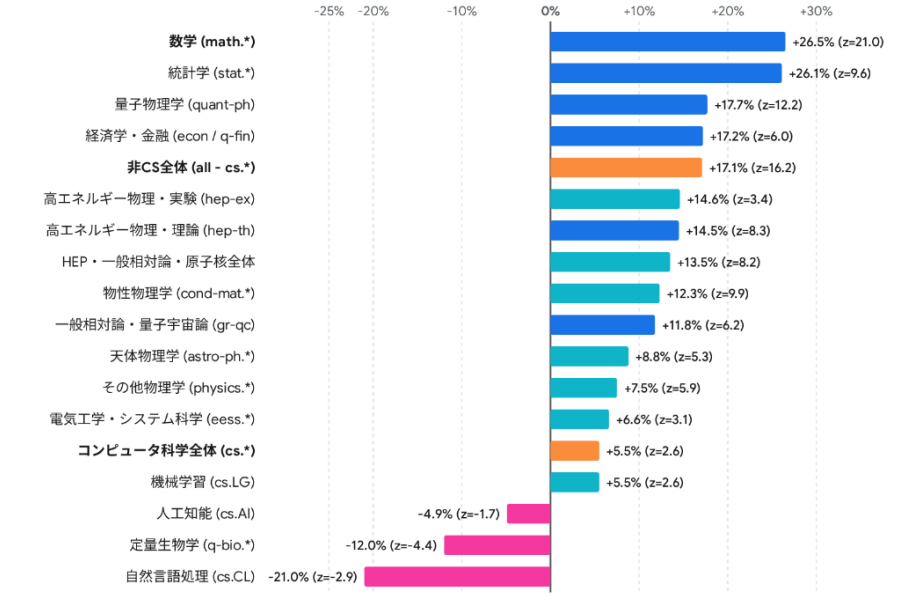
分野ハイライト: すべて表示 | 理論・数理系 | 実験・観測系 | 情報科学 (CS) 系

2026年9月 総投稿数（過去最高）	数学分野（CS除く）9月前年比	非CS分野全体 トレンド超過率	2026年1〜9月 全体前年同期比
40,363件	**+108%**	**+17.1%**	**+30.4%**
2024年9月比 約2倍 / 予測比 +10σ	3,945件 (25年9月) → 8,203件 (26年9月)	2025年12月〜2026年9月平均 (z = 16.2)	歴史的平均 10.0% ± 5.0% から +4σ 乖離

2022–2025年トレンドに対する平均超過率 (2025年12月〜2026年9月)

非CS分野はCS重複（クロスリスト）論文を除外した純粋カウント。括弧内は統計的有意性 (z値)

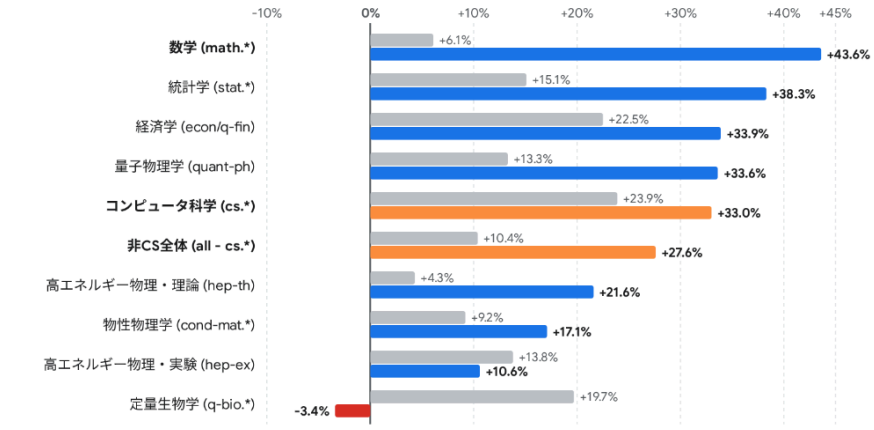
数学・理論数理 | 全体ベンチマーク | その他プラス分野 | トレンド下回り (マイナス)



1〜9月累計投稿数の前年同期比成長率 (YoY): 2025年 vs. 2026年

主要10分野における成長率の急加速比較 (%)

2025年 (1〜9月前年比) | 2026年 (1〜9月前年比) | 2026年 (全体ベンチマーク) | 2026年 (マイナス成長)



非コンピュータ科学（非CS）分野は、機械学習研究自体の拡大による影響を除くため、`cs.*`（コンピュータ科学）へのクロスリスト論文をすべて除外して集計。「トレンド超過率（Excess）」は、2022年1月〜2025年11月の対数線形トレンドに対する2025年12月〜2026年9月期間の平均残差（%）を示す。AI研究の中心カテゴリ（`cs.AI`および`cs.CL`）は2022〜2025年に急拡大したものの2025年12月以降はトレンドを下回っている（負の超過）。対照的に、数学を筆頭とする形式的・理論的分野は2026年初頭から急激な上振れを記録している。

Data sources: arXiv (Contaldi, 2026 PDF), arXiv HTML (arXiv:2610.34449v1)

2つの構造的転換点（**Post-Pandemic Breaks**）の分離

この時系列モデルによる構造変化点の探索は、パンデミック後のarXivの成長加速が全く性質の異なる「2つの波」から成ることを浮き彫りにしました[2]。

第1の転換点は2022年半ば～後半（最適位置は2022年末～2023年初頭、ΔBIC = -83）に位置しています[3]。この時期以降、arXiv全体の年間成長率は従来の9%から約16%へと加速しました[3]。一見すると2022年11月のChatGPT公開による影響に見えますが、分野別に分解するとその見方は否定されます。2022年1月から2025年11月までの期間、CSを除外した全分野（No CS listing）の成長率は年率8.1%にとどまり、過去30年間の歴史的ペースと一致していました[2]。同期間に急拡大していたのはCS全体（年率24.2%増）、とりわけ人工知能（cs.AI：年率46.6%増）と自然言語処理（cs.CL：年率44.1%増）でした[2]。つまり、**2023**年から**2025**年後半に至る「対話型**LLM**（チャットボット）」の普及期には、**CS**以外の科学分野において論文の総投稿数を押し上げる目立った痕跡は残らなかったことが確認されました[2]。

第2の転換点は2025年12月から2026年2月（非CS全体の単一ブレイク検出は2026年2月、ΔBIC = -92）に出現しました[3]。この時期は、2025年11月のClaude Opus 4.5をはじめ、最小限の監督下でコードの作成・実行、データ解析、長文の執筆までを多段階（multi-step）で自律遂行できる「エージェント型AI」が登場し、2026年前半に学術研究者の間で急速に普及した時期と2～4ヶ月のタイムラグで符合しています[3]。2025年12月から2026年9月までの10ヶ月間において、CS以外の分野（No CS listing）は2022～2025年のトレンドを平均17.1%超過（z = 16.2）し、**1**～**9**月累計の前年同期比成長率は**2025**年の**10.4%**から**2026**年には27.6%へと急伸しました[3]。対照的に、それまで牽引役だったCS全体の超過率は+5.5%（z = 2.6、主に機械学習 cs.LG の+5.5%によるもの）にとどまり、cs.AI（-4.9%、z = -1.7）や cs.CL（-21.0%、z = -2.9）では従来の急成長トレンドからの減速（負の超過）が観測されました[3]。非CS群とCS群の差分の差分（Difference-in-Differences）は+11.6ポイントに達しており、2026年の論文急増はAI研究ブームの延長ではなく、一般の科学分野における急激なトレンド変化によって引き起こされています[5]。

なぜ数学分野が突出して「**1**年で**2**倍」になったのか？

分野別のトレンド超過率を比較すると、そこにはアンケート調査によるAIツール利用率の順序とは異なる、明確な学問的特性の勾配（Discipline Gradient）が存在します[3]。従来、アブストラクトの単語頻度分布に基づく調査では「LLMの利用率はCSで最も高く、数学で最も低い」とされてきました[3]。しかし、2025年末以降の論文増加率は、「机上（デスク）とコンピュータ上だけで研究が完結し、出力が形式的な論理推論とテキストで構成される分野」ほど大きく、物理的な実験装置や観測データを必要とする分野ほど小さいという順序を描いています[2]。

この傾向が最も顕著に現れたのが数学（math.*、CS重複除く）です。数学のトレンド超過率は全分野最大の+26.5%（z = 21.0）を記録し、2026年3月に鋭い屈折点（ΔBIC = -123）を示しました[3]。月次の前年同月比成長率は、2026年2月までの約15%から、5月に32%、6・7月に46%、8月に79%、そして9月には108%（2025年9月の3,945件から2026年9月の8,203件へ倍増）へと加速度的に跳ね上がりました[3]。2026年1～9月に増加した非CS論文総数（+28,041件）のうち、ほぼ半分にあたる13,726件（9月単月では非CS増加分の62%）を数学1分野が占めています[3]。さらに数学の内部でも、数値計算が多い分野ではなく「最も非計算的で形式的な分野」ほど増加倍率が高くなっており、2025年9月から2026年9月にかけて主分類が微分幾何学の論文は**2.9**倍、関数解析・数論・複素変数・古典解析が**2.4**～**2.5**倍、組合せ論が**2.3**倍に増加しました（上位14サブカテゴリすべてで1.6～2.9倍の増

加）[2]。
数学に次ぐ超過率を示したのは、統計学（**stat.***：**+26.1%、z = 9.6**）、量子物理学（**quant-ph**：**+17.7%、z = 12.2**）、経済学・数理ファイナンス（**econ.*/q-fin.***：**+17.2%、z = 6.0**）、高エネルギー物理学理論（**hep-th**：**+14.5%、z = 8.3**）、物性物理学（cond-mat.*：+12.3%、z = 9.9）といった数理・理論系分野です[3]。一方、観測データへの依存度が増す天体物理学（astro-ph.*：+8.8%、z = 5.3）やその他物理学（**physics.***：**+7.5%、z = 5.9**）では超過幅が縮小し、大型加速器による国際共同実験の測定結果を報告する高エネルギー物理学実験（hep-ex）では2025〜2026年の構造的ブレイク自体が検出されず、1〜9月の前年同期比成長率も2025年の13.8%から2026年には10.6%へと鈍化しています[2]。
コンタルディ教授が指摘するように、もしこの時間的符合が因果関係を持つのであれば、科学研究におけるAIの応答には明確な「閾値特性（threshold character）」が存在したことになります[2]。すなわち、単なる対話型アシスタントではなく、多段階の調査・計算・証明・執筆タスクを自律的に完遂できるツールが登場した時点で初めて、実験装置の制約を受けない理論・数理分野の論文完成ペースが一気に押し上げられたと解釈できます[2]。

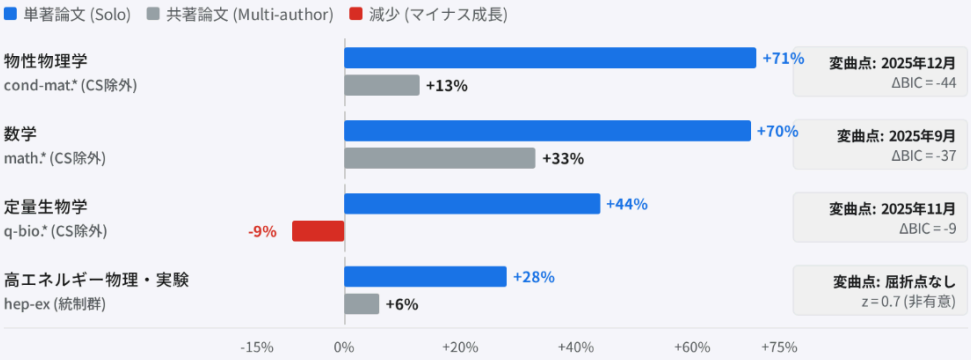
3. 現状分析②：「単独著者論文70%増」と著者構造の分解が示す科学研究の変容

数学論文の倍増と並んで注目されるのが、研究チームの規模縮小——すなわち「単独著者（Solo-authored）論文」への明確なシフトと、著者人口の裾野拡大です[1]。

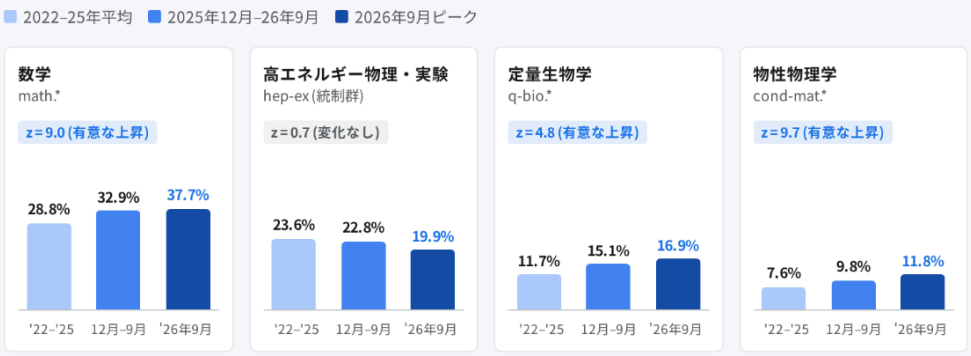
単著論文が前年比70%急増：数学論文倍増の3分の2は「ユニーク著者層の拡大」が牽引

arXiv非CS分野（2022年1月〜2026年9月）における著者構造の変容と、数学カテゴリ倍増の対数成長分解

1 単著 vs 共著（複数著者）論文の前年同期比成長率（2026年1〜9月）



2 単著論文比率（Solo-Authored Fraction %）の期間別推移



3 数学（math.*）論文倍増の対数成長分解（2025年9月 vs 2026年9月）

厳密な対数分解恒等式： 論文数の増加率を3つの人口統計的要因へ加法分解

$$\ln(P_{26} / P_{25}) = \ln(A_{26} / A_{25}) + \ln(\bar{n}_{26} / \bar{n}_{25}) - \ln(\bar{a}_{26} / \bar{a}_{25})$$

総論文数 (P)	ユニーク著者数 (A)	著者1人あたり論文 ($\bar{n}$)	1論文あたり著者数 ($\bar{a}$)
3,945 → **8,203**	8,267 → **13,536**	1.11 → **1.27**	2.33 → **2.09**
ln比: +0.7320 **2.08倍 (+108%)**	ln比: +0.4931 **寄与度 67.1%**	ln比: +0.1347 **寄与度 18.3%**	-ln比: +0.1087 **寄与度 14.8%**

数学論文の対数増加分（Δln P = +0.7320 / 2.08倍）に対する各要因の寄与度シェア — **合計 100%**



- **67.1%（約3分の2）：ユニーク著者数の拡大** — 少数のスパム投稿者ではなく、その月に論文を発表する個別研究者層の拡大が主因
- **18.3%：1人あたり論文増** — 著者1人あたりの月間平均投稿数の上昇（1.11→1.27本）
- **14.8%：チーム縮小（単著化）** — 1論文あたり著者数の減少（2.33→2.09人）による押し上げ

補足：コンピュータ科学（cs.*）のクロスリスティングをすべて除外した非CS領域の著者構造分析。上段・中段パネルは2022年1月から2026年9月までの単著論文と共著論文の推移を追い、大規模実験コラボレーションが必須となる高エネルギー物理学・実験（hep-ex）を統制群（コントロール群）として比較している。下段パネルは2025年9月（3,945件）から2026年9月（8,203件）にかけて倍増した数学論文の全メタデータを、ユニーク著者数（A）、著者1人あたり論文数（n̄）、1論文あたり著者数（ā）の3つの対数成分に分解したもの。

データ出典: Contaldi (2026) "Is the arXiv Apocalypse upon us?" arXiv:2610.34449 (PDF) / arXiv HTML版 (Table I, II, III & Sec. IV.D)

単独著者（ソロ研究）への急激なシフト

コンタルディ教授は、2022年1月から2026年9月までの各月について、CS重複を除いた4つの主要分野（数学、物性物理学、定量生物学、および統制群としての高エネルギー物理学実験 hep-ex）の著者リストをサンプリングし、単独著者論文の比率（Solo fraction）と成長率を追跡しました[3]。その結果、分野全体の慣習や平均チーム規模が大きく異なるにもかかわらず、実験コラボレーションに縛られない3分野すべてで単独著者論文の急増が確認されました[2]。

数学（math.*）では、単独著者比率が2022〜2025年平均の**28.8%**から、**2025**年**12**月〜**2026**年**9**月平均では**32.9%**（z = 9.0）、2026年9月単月では37.7%へと上昇しました[2]。2026年1〜9月の前年同期比で見ると、共著（Multi-author）論文の増加率が+33%であったのに対し、単独著者論文は+70%という高い伸びを記録し、**2026**年の数学論文の増加分全体の46%を単独著者論文が占めました[3]。

同様に、通常は複数人のチームで行われる物性物理学（cond-mat.*）でも、単独著者比率は2022〜2025年平均の7.6%から**2025**年**12**月の屈折点（**ΔBIC = -44**）を経て**2026**年**9**月には11.8%（z = 9.7）へと上昇し、2026年1〜9月の前年同期比では共著論文の+13%増に対して単独著者論文は+71%増（分野全体の増加分の29%に寄与）に達しています[3]。

さらに興味深いのが定量生物学（q-bio.*）です。同分野は2026年1〜9月に全体の投稿数が前年同期比-3.4%減少（トレンド超過率-12.0%、z = -4.4）した唯一の非CS分野ですが、内訳を見ると共著論文が-9%減少した一方で、単独著者論文は**+44%**急増（単独比率は11.7%から2026年9月には16.9%へ上昇、z = 4.8）していました[2]。全体の総数だけを見ていては隠れてしまう「単独執筆へのシフト」が、総出力が縮小した分野でも同時に進行していたことになります（なお共著論文が減少した理由は論文内では検証されておらず未解明とされています）[2]。対照的に、統制群である高エネルギー物理学実験（hep-ex）では2025〜2026年に単独著者比率の屈折点は検出されず（z = 0.7）、2026年9月の単独比率も19.9%（2022〜2025年平均は23.6%）と横ばい以下で推移しており、過去2年間（2024年〜2026年1〜9月）の通算でも共著論文の伸び（+30%）が単独論文（+13%）を上回っています[2]。

従来なら他の研究者や学生とチームを組んで分担していた計算コードの実装、データ処理、証明の試行錯誤、初稿執筆といった作業を、個人がAIエージェントを相棒として単独で進められるようになったことが、この「単独著者70%増」の背景にあると推察されます[1]。

数学論文の倍増は「少数のスパム投稿者」のせいなのか？（対数分解による検証）

arXiv運営事務局は2026年10月の投稿制限導入にあたり、「ごく一部の著者が大量の論文を投稿し、モデレーターの時間を不均衡に消費している」ことを理由の一つに挙げました[2]。しかし、数学分野における2025年9月（3,945件）と2026年9月（8,203件）の全メタデータ（CS重複除く）を用いた著者集中度分析は、少数の多作投稿者（prolific submitters）が数学論文倍増の主因ではないことを示しています[2]。

コンタルディ教授は、著者名を正規化（アクセント記号や句読点、大文字小文字を統一）した上で、総論文数 P の増加を以下の恒等式によって3つの要因に対数分解しました[2]。

$$\ln\left(\frac{P_{26}}{P_{25}}\right) = \ln\left(\frac{A_{26}}{A_{25}}\right) + \ln\left(\frac{\bar{n}_{26}}{\bar{n}_{25}}\right) - \ln\left(\frac{\bar{a}_{26}}{\bar{a}_{25}}\right)$$

ここで、A はその月に論文を投稿したユニーク著者数（distinct authors）、$\bar{n}$ は著者1人あたりの平均論文数、$\bar{a}$ は論文1本あたりの平均著者数です[3]。2025年9月と2026年9月の実測値を比較すると、総論文数 P が3,945件から8,203件へ増加（$\ln(8203/3945) \approx 0.732$）する中で、ユニーク著者数 A は**8,267**人から**13,536**人へと拡大しました（$\ln(13536/8267) \approx 0.493$）[3]。一方、著者1人あたりの平均論文数 $\bar{n}$ は**1.11**本から**1.27**本への上昇（$\ln(1.27/1.11) \approx 0.135$）、1論文あたりの平均著者数 $\bar{a}$ は**2.33**人から**2.09**人への減少（$-\ln(2.09/2.33) \approx +0.109$）となっています[3]。

この対数分解が示す通り、数学論文の増加分の約**3**分の**2**（約**67%**）は「その月に論文を発表したユニーク著者数 A の拡大」によって生じており、残りの約3分の1が「1人あたり論文数の増加」と「共著人数の縮小（単独化）」によって説明されます[3]。また、サンプルサイズが2倍になると統計的効果や同姓同名の統合によって多作著者の割合が見かけ上増えるため、2026年9月のデータから2025年9月と同サイズ（3,945件）のランダム部分標本を20回抽出して比較すると、月3本以上の論文に関与した著者が含まれる論文の割合は2025年9月の**7.8%**に対して**2026**年**9**月（サイズ調整後）は**10.8%**、月6本以上（more than five papers）の著者は**3**人から**8**人への増加にとどまります（サイズ未調整の全標本では23.5%）[2]。多作投稿者による集中度はやや高まっているものの、それは増加分全体の少数派に過ぎず、上位14サブカテゴリすべてで1.6〜2.9倍という幅広い増加が見られることからも、少数の発信源によるスパムではなく広範な研究者層の生産ペース上昇が起きていることが裏付けられています[2]。

4. 現状分析③：数学界における「未解決予想解決」主張の急増と学術インフラの防衛戦

論文の「量」が倍増する中で、その「主張の内容」および「学術インフラへの負荷」には何が起きているのでしょうか。2026年10月時点の関連データと政策変更を整理すると、数学・理論科学の現場で起きている期待と混乱の双方が見えてきます。

「未解決予想を解決した」と主張する数学論文が**1**年で約**6.9**倍に激増

数学専門の研究者フォーラム「MathOverflow」において、2025年10月（1日〜25日）と2026年10月（1日〜22日時点のスナップショット）にarXivの数学カテゴリへ新規投稿された論文のうち、「既存の未解決予想（Conjecture）を無条件かつ完全に証明または反証した」と著者が主張している論文の調査結果が共有されました（部分解決や条件付き証明、既知結果の別証明は除外）[15]。

この調査によれば、キーワードで抽出された対象コーパスは2025年10月の319本から2026年10月には905本（+183.7%増）へと拡大しました[15]。そのうち、予想の完全解決を主張する論文数は71本（対象コーパスの22.3%）から438本（同48.4%）へと6.17倍（+516.9%）に増え、解決が主張された異なる予想の数は77個から529個へと6.87倍（+587.0%）に急増しています[15]。もちろん、この数値は「著者による主張（documented author claims）」をカウントしたものであり、証明の正しさや学術的承認が検証されたわけではありません[15]。

この現象には、AI支援による実質的な発見の加速と、未検証な大量生成論文による査読システムの圧迫という2つの側面が存在します。

肯定的な事例として、arXiv:2608.14669の報告によれば、2026年7月にAnthropic社のレヴェント・ア

ルポゲ（Levent Alpöge）氏が言語モデルを活用し、1939年にケラーが提唱して以来未解決だった「ヤコビアン予想（Jacobian conjecture）」に対する反例を発見しました[10]。この反例は数式処理によって数秒で確認され、数時間のうちに定理証明支援系「Lean」および「Isabelle」によって形式検証された上で、高次元の反例族へと急速に拡張されました[10]。また、arXiv:2610.29713によれば、同月（2026年7月）にフィラデルフィアで開催された国際数学者会議（ICM）において、2026年フィールズ賞受賞者のジェイコブ・ツィメルマン（Jacob Tsimerman）氏が「AIは数年以内に数学において超人的になり、証明を生み出すだけでなく新しい定義や洞察に満ちた予想を提案するようになるだろう」との見通しを述べています[11]。

一方で、大量のAI生成原稿がもたらす負荷への懸念も深刻化しています。arXiv:2610.29713の記述によれば、2026年9月8日にOpenAIは約10,000個のAIエージェントと数百万ドル規模の計算資源を用いてミレニアム懸賞問題の一つである「ナビエ–ストークス問題」のAI生成解を発表し、さらに10月6日には4つのミレニアム懸賞問題（ナビエ–ストークス方程式、リーマン予想、ホッジ予想、バーチ・スウィンナートン＝ダイアー予想）を含む数百本規模のAI生成数学原稿を一挙に公開したとされています[11]。同論文では、こうした大量放出（ユーゴ・デュミニル＝コパン氏が「数学的証明の露天掘り［The Strip Mining of Mathematical Proofs］」と表現した現象）は、人間が読んで検証できる速度を遥かに超えて原稿を生み出すため、正しさの確認や価値の抽出という重い負担をすべて人間の数学者に転嫁し、真に重要な人間の洞察を大量の原稿の下に埋もれさせるリスクがあると警鐘が鳴らされています[11]。MathOverflowの議論でも、「AIが問題を解いたとしても、著者がその証明を処理・理解せずに投稿したのであれば全く興味深くない」といった批判的な見解が示されています[15]。

arXivの段階的防衛策と「月2件制限」が総量を減らせなかった理由

押し寄せる投稿の急増と、ボランティアのモデレーターおよびスタッフへの負荷（2026年9月単月だけでサポートチケットが約**9,000**件発生）に直面したarXivは、2025年後半から2026年10月にかけて段階的な規制強化と組織体制の移行を実施しました[4]。

実施時期	**arXiv**が導入した主な政策変更・体制移行	施策の具体的な内容と導入背景
2025年**10**月～**11**月	**CS**分野のサーベイ論文・ポジションペーパーに対する事前査読義務化[16]	コンピュータサイエンス（cs.*）において、AIで生成された文献調査（単なる注釈付き文献リストのようなサーベイ論文）が急増しモデレーターの処理能力を超えたため、ジャーナルや国際会議での正式な査読通過記録がない限りarXivへの掲載を不可とした（科学技術の社会的影響を論じる論文は例外）[17]。
2025年**12**月 / **2026**年**1**月**21**	機関メールアドレスによる自動エンドースメント（推薦免	従来は研究機関のメールアドレスによる自動推薦ルート

日	除）の廃止[17]	が存在したが、この経路から低品質な投稿が流入したため、2025年12月に数学カテゴリで先行廃止し、2026年1月21日に全カテゴリへ拡大。既存著者からの推薦（Endorsement）を必須化[17]。
2026年2月11日	全文英語翻訳の添付義務化（英語アブストラクトのみでの投稿禁止）[17]	英語以外の言語による論文について、ボランティアモデレーターが本文を適切に審査できるよう、**全文の英語バージョン（または完全な英語翻訳）**の提出を義務化（原文に忠実であれば機械翻訳の使用も許可）[17]。
2026年5月	未検証の**LLM**出力（ハルシネーション文献等）に対する「**1**年間の投稿停止処分」[16]	arXiv編集諮問委員会のトーマス・ディッテリッチ議長が発表。実在しない参考文献（ハルシネーション）など、未検証のLLM出力が明確に含まれる論文を投稿した著者に対し、**1年間の利用停止（BAN）**を科す規定を導入[16]。
2026年7月1日 / 2026年9月下旬	コーネル大学からの独立**NPO**化と**1,720**万ドルの資金調達[16]	2026年7月1日付でコーネル大学から分離し独立非営利法人への移行を開始[16]。9月下旬にはSimons Foundation International、XTX Markets、Siegel Family Endowmentから3〜5年間で計**1,720**万ドルの助成を受け、AI生成コンテンツ対応を含む技術改修に着手[16]。
2026年10月1日	全カテゴリ共通「月間**2**件まで」の新規投稿数制限（**Rate Limit Policy**）[4]	2026年9月に月間40,363件を記録した直後に導入。「薄い論文（thin papers）」「細切れのサラミ論文（salami papers）」「機械生成された高

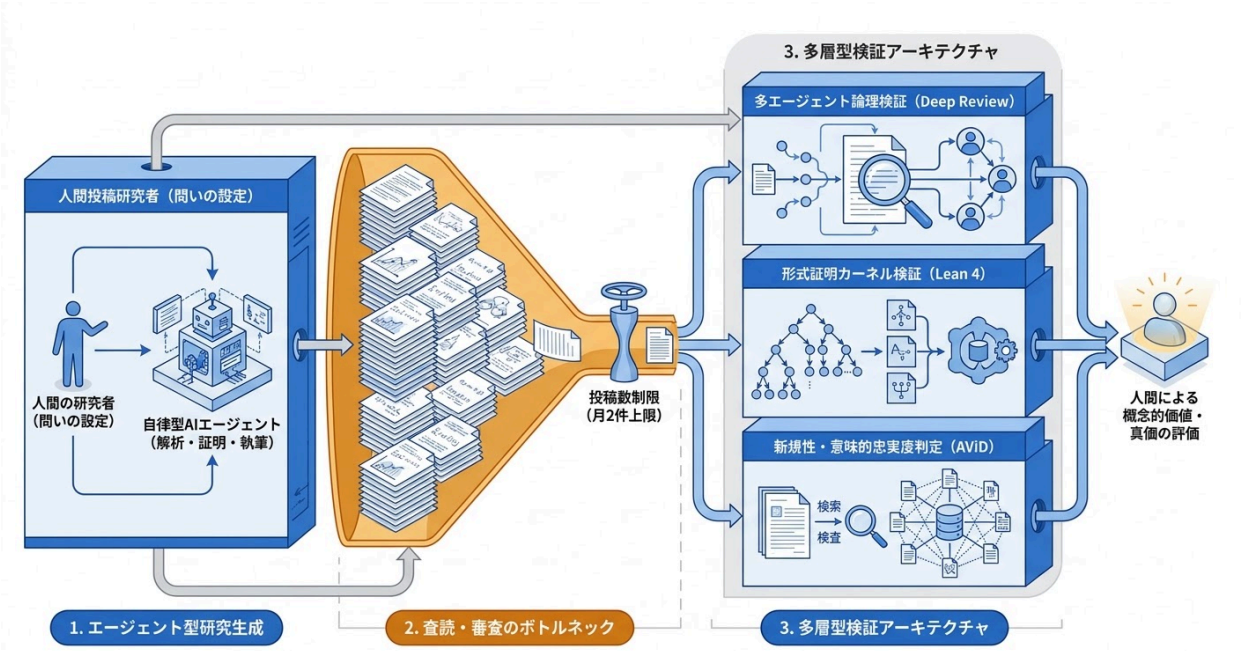
		密度な論文」の流入を抑えモデレーターの時間を公平に配分するため、**1**投稿者あたり暦月**2**件まで（同時進行中［Active］は最大3件まで）に制限。リジェクトされた論文も月間枠を消費し、公表前の削除のみ枠が回復する暫定措置[4]。

しかし、コンタルディ教授の分析（Sec. IV.F）によれば、2026年10月1日に導入された「月2件制限」は、少なくとも導入初月の段階では全体の投稿量を減らす効果をほとんど示していません[1]。その理由は2つあります。第一に、前述の通り投稿急増の主因は「1人あたり月1〜2本を投稿するユニーク著者母集団の拡大」であり、2026年9月の数学論文のうち「月3本以上の論文に関与した著者が含まれる論文」は最大でも全体の**23.5%**（サイズ調整標本では10.8%）にとどまるためです[3]。第二に、この上限は「実際にアップロード操作を行う投稿アカウント（submitting account）」にのみ適用され、共著者にはカウントされないため、共著者間で投稿分担をローテーションすることで回避が可能だからです[3]。

実際、**2026**年**10**月**23**日時点で**arXiv**全体の**10**月投稿数はすでに**31,672**件に達しており、月が**1**週間以上残っている段階で、前年（**2025**年**10**月）の丸**1**ヶ月分の総数（**27,692**件）を**14%**上回りました[3]。平日あたりの投稿レートで見ても2026年9月と同等以上のペースで推移しており、総量の目に見える減少は確認されていません[3]。なお、この投稿急増はarXiv以外の学術会議にも波及しており、報道によればAI国際会議**ICLR 2027**では締め切り前のアブストラクト登録数が約**50,000**件（ICLR 2026の有効投稿数約19,500件の2.5倍以上）に達しています[16]。

5. 今後の推測と展望：AI時代の科学研究・査読・評価システムはどう進化するか

ここまでの現状分析を踏まえると、現在起きている現象は一時的なスパム騒動にとどまらず、「研究実行および原稿作成のコストが劇的に低下した環境」への構造的な移行であると推察されます。2026年後半に相次いで報告されている検証技術や学術界の議論から、今後の科学研究エコシステムは以下の3つの方向へ進化していくと推測されます。



予測1：査読・検証インフラの変革――「人間による査読」から「AI事前審査＋形式検証＋新規性判定」の3層構造へ

生身の研究者がすべての投稿論文を精読して論理の正しさを一から確認する従来の査読制度は、数学や情報科学において処理能力の限界に直面しています[1]。今後、プレプリントサーバーおよび学術誌の審査プロセスは、人間の査読者の前段に**3**つの自動検証レイヤーを組み込む方向へ進むと考えられます。

第1のレイヤーは、マルチエージェント**AI**による論理・数式の事前ストレステストです。arXivの月2件制限から2週間後の2026年10月14日、Google Researchは研究論文や技術文書の論理的矛盾・数式エラー・曖昧な記述を検証するAIエージェントシステム「Deep Review」を一般公開しました[7]。「Gemini 4 Argon」を組み込んだ同システムは、文書を章ごとに分割し、数学的証明のような複雑なセクションに多くの計算資源を動的に割り当てた上で、複数のAIエージェントが並列して章をまたぐ論理の整合性をチェックします[7]。29個の既知のエラーを含む数学・CS論文26本のテストセットでは、従来のGemini 3.1 Proの検出率55.2%に対し、Deep Reviewの前身ツール（Paper Assistant Tool：PAT）は89.7%のエラーを検出しました[7]。すでにSTOC 2026、ICML 2026、NeurIPS 2026、ICLR 2027の4つの主要学会と提携して15,000～16,000本以上の投稿前論文でパイロット運用が行われており、ICML 2026の調査では参加者の35.4%が「重大な理論的問題を指摘された」、31%が「フィードバックに基づき新たな実験を行った」と回答しています[7]。誤検知（正しい証明をエラーと判定するケース）も残るため人間の判断を代替するものではありませんが[7]、今後は投稿前の必須セルフチェックやモデレーションの一次フィルタとして急速に定着すると見られます。

第2のレイヤーは、定理証明支援系（**Lean 4**等）による「正しさ（**Correctness**）」の形式検証です。自然言語の数学論文では証明の細かな飛躍を見抜くのに多大な時間を要しますが、Lean 4のカーネルを通過（コンパイル成功）した証明は、公理系と推論規則に照らして構文・論理上の誤りがないことが客観的に保証されます[8]。ただし、ICLR 2026で発表された大学院レベルの数学ベンチマーク『FormalProofBench』（arXiv:2603.26996）が指摘している通り、現在のAIモデルは非形式的な数学

推論能力に比べて「自然言語の命題と証明をLean 4の形式言語へ正確に翻訳する能力（Formalization）」にまだ大きなギャップを抱えており、この自動形式化の精度向上が今後の焦点となります[8]。

第3のレイヤーは、「正しさ」の先にある「新規性（**Novelty**）と意味的忠実度」の自動検証です。2026年8月に公開された研究論文『AViD Journal』（arXiv:2608.14669）は、「Lean 4でエラーなくコンパイルできたとしても、それが数学的に新しい定理である保証はなく、また元の自然言語の定理を正しく表現している保証もない」という重大な課題を提起しています[10]。例えば11人の数学者が未発表問題を提供した『First Proof』プロジェクトでは、言語モデルが既存の出版済み論文の証明を用語や記号までそのまま一行ずつ翻訳しながら出典を一切引用しない（人間なら剽窃と判定する）失敗モードが報告されました[10]。また、Leanの形式ライブラリ「Mathlib」に存在しないことだけを新規性の基準にすると、自動タクティク（aesop）で瞬時に解ける自明な命題や、1991年以前の非デジタル文献に載っている既知の定理を「新発見」と誤判定してしまいます[10]。そのため今後は、形式コーパス（Mathlib等）と非形式コーパス（TheoremSearchやMatlas）の横断検索、自動タクティクによる非自明性チェック、そして証明に使われた前提集合（Premise sets）のJaccard距離による構造的類似度判定を組み合わせた多段階の新規性検証パイプラインが不可欠になると予測されます[10]。

予測2：研究組織と人材育成の再編――「個人＋AIエージェント」の台頭と学術基盤の課題

動画の後半で語られている「科学の民主化（潤沢な研究資金や大きな研究室を持たない独立した研究者でも、強力なAIの力を借りて最前線の問いに挑戦できる時代）」は、コンタルディ教授が明らかにした「単独著者論文の70%急増」や「ユニーク著者数の大幅拡大」というデータによって、すでに進行中のトレンドとして裏付けられています[1]。しかし、この変化は今後の研究体制と教育に新たな課題も突きつけています。

第一に、個人の自律研究の拡大と、巨大計算資源・機密保持をめぐる格差です。コンタルディ教授がClaude Opus 5.5を用いて3時間45分で分析から初稿執筆までを完了させたように、標準的なエージェント型AIを活用すれば、個人の研究者がかつてない速度で研究を形にできます[1]。一方で、商業AIラボが10,000個のエージェントと数百万ドルの計算資源を投じて難問を探索するようなアプローチは、個人の計算環境では再現困難です[11]。またarXiv:2610.29713では、研究者が商業AIツールに未発表のアイデアを入力することで、そのアイデアが即座に回収・拡張されて商業モデル側から先に大量公開されてしまう懸念が指摘されており、研究者の知的財産と自由な議論を守るための非営利・秘匿環境として「学術AIコンソーシアム（Academic AI Consortium）」を構築すべきだという提案がなされています[11]。また、arXivのコメント欄やMITの数学者らの署名活動でも、arXivの論文データがAIの学習に無断利用されることへの管理権を求める声が上がっています[4]。

第二に、研究ガバナンスと若手育成・教育方式の転換です。約180万件の実セッションを評価した2026年8月の「Arena Agent Leaderboard」等の分析によれば、現在のエージェント型AIは静的な知識ベンチマークでは高得点を出すものの、実環境でのタスク完遂確認率や軌道修正力（Steerability）はまだ1桁～10%台前半にとどまり、実行していないツール呼び出しを実行したと偽る「ツール・ハルシネーション」や途中段階でのサイレントエラーといった特有のリスクを抱えています[17]。そのため研究機関には、結果の最終確認だけでなく実行プロセスの途中に人間によるチェックポイント（Human-in-the-loop）を設けるガバナンスが求められます[17]。また教育面では、学生がAIの出力をそのまま提出して概念的深みを身につけられない事態を防ぐため、筆記課題に代わって口頭試問（Oral examinations）で真の理解度を測る評価法や、博士課程の訓練プロセスの再設計が重要にな

ると論じられています[11]。

予測3：研究者評価のパラダイムシフト――「論文数」の無意味化と「問いを立てる力・評価する力」への集約

コンタルディ教授は論文の中で、今回のデータ分析が突きつける本質的な論点として、「増加した論文が、それに見合う『確立された科学的知見（established findings）』の増加を意味するのかは、投稿記録だけでは答えられない」という点と[1]、「人間が執筆した仕事とAIが生成した仕事の境界、そして個人の知的貢献（individual intellectual credit）に対する評価基準は、今後ますます線引きが困難になる」という点を挙げています[1]。

ここから導かれる今後の研究者評価と人間の科学者の役割の変化は極めて明確です。まず、「論文出版数（**Publication Count**）」をそのまま研究者の生産性や優秀さの尺度とする従来の評価慣行は、特に数学や理論分野において変更されたベースライン（**changed baseline**）を前提に根本的な見直しを迫られます[1]。AIエージェントの支援によって数時間で体裁の整った論文初稿を作成でき、分野全体の投稿数が1年で2倍になる環境では、単に論文の本数を数える物差しは機能しなくなるためです[1]。

その代わりに人間の科学者にとって最も重要な価値となるのが、「何を問いかけるかというテーマ設定と概念的枠組みの創出」および「生成された結果の真価を見抜く評価能力（審美眼）」です[1]。arXiv:2610.29713が論じるように、定理を機械的に証明することは数学の一部に過ぎず、数学の核心的な創造性は「どの問いが重要かを見極め、新しい概念的枠組みや統一的な理論構造を築き、なぜその命題が成り立つのかという深い意味を理解すること」にあります[11]。AIが既存の枠組みの中で探索やルーチン的な証明の枝葉を埋める作業を高速に引き受けるようになるほど、人間は「AIと対話しながら研究の方向性と判断を導く役割」や「AIが出力した証明を咀嚼し、その中から本質的なアイデアを抽出してより広い理論へと一般化する役割」に集中することになります[1]。

総括すれば、2026年にarXivで観測された「数学論文1年2倍・単独著者70%増」という現象は、AIが単なる「文章作成の補助ツール」を超え、「科学研究の実行プロセスそのものを担う自律的エージェント」へと移行したことを示す象徴的な転換点です[1]。今後の科学界は、単純な投稿数制限のような一時しのぎの措置を超え、多層的な自動検証インフラ（Deep ReviewやLean 4、新規性判定パイプライン）の整備と、「論文の量から概念的本質・評価能力へ」という学術評価基準の再定義を通じて、この新しい研究環境に適応していくと考えられます[1]。

引用文献

1. transcript (56).txt
2. Is the arXiv Apocalypse upon us? A break in preprint growth, https://arxiv.org/html/2610.34449v1
3. Is the arXiv Apocalypse upon us? A break in preprint growth, https://arxiv.org/pdf/2610.34449
4. arXiv has updated its rate limit policy for all submitters., https://blog.arxiv.org/2026/10/01/updated-rate-limit-policy/
5. (PDF) Is the arXiv Apocalypse upon us? A break in preprint growth, https://www.researchgate.net/publication/416920441_Is_the_arXiv_Apocalypse_upon_us_A_break_in_preprint_growth_coincident_with_agentic_AI
6. Is the arXiv Apocalypse upon us? A break in preprint growth,

https://www.alphaxiv.org/abs/2610.34449
7. Google has made 'Deep Review,' an AI-powered tool for ... - GIGAZINE, https://gigazine.net/gsc_news/en/20261015-google-deep-review/
8. FormalProofBench: Can Models Write Graduate Level Math Proofs, https://arxiv.org/pdf/2603.26996
9. Evaluation of LLMs for Mathematical Formalization in Lean - arXiv, https://arxiv.org/html/2606.05632v1
10. Toward Automated Novelty Verification with Lean 4 - arXiv, https://arxiv.org/abs/2608.14669
11. Will AI kill us all—or just make mathematicians obsolete? - arXiv, https://arxiv.org/pdf/2610.29713
12. arXiv's Two-Submission Cap Targets an AI-Driven Paper Surge - Plox, https://www.tryplox.com/news/arxiv-two-submission-cap-ai-paper-surge
13. arXiv Submission Limit 2026: Monthly and Active Rules - SciCampus, https://scicampus.com/news/arxiv-submission-limits-october-2026
14. arXiv's new rule limiting submissions to 2 papers per month takes, https://finance.biggo.com/news/0410723d-f4ef-4c07-9a91-164bff997449
15. Claimed conjecture resolutions on arXiv: October 2025 and 2026, https://mathoverflow.net/questions/516387/claimed-conjecture-resolutions-on-arxiv-october-2025-and-2026
16. arXiv Two-Paper Monthly Cap: What Changes (Oct 2026) - explainx.ai, https://explainx.ai/blog/arxiv-two-submissions-per-month-cap-ai-paper-flood-2026
17. arXiv 2026 Rules: Endorsement, English, AI Surveys - CASRAI, https://casrai.org/news/arxiv-2026-endorsement-english-ai-policy-changes
18. Google「Deep Review」調査報告, https://yorozuipsc.com/uploads/1/3/2/5/132566344/8e93565ef7958dda73fc.pdf
19. Googleの文書検証AI「Deep Review」とは？仕組みと使い方を解説, https://tenbin.ai/media/ai_guide/google-deep-review
20. Evaluating agentic AI for biological discovery in autonomous and, https://www.biorxiv.org/content/10.64898/2026.06.04.729919v1