ChatGPT

AIと人間は本当に「リーマン予想」を解いたのか——2026年11月時点の検証と「ゼロイチ発明」への含意

エグゼクティブサマリ

2026年11月22日時点で、「AIが人間と共同して167年来の難問リーマン予想を解いた」という表現は、**単なる誇張や未検証のプレスリリースとは言えなくなっている**。11月20日、Anthropic所属の数学者Levent Alpögeは、自身とClaudeがリーマン予想（RH）およびDirichlet L 関数に対する一般化リーマン予想（GRH）の証明を発見したと公表し、通常の論文ではなく約1,75000個のLeanファイル、約1,737万行からなる形式化証明を公開した。翌21日までに、Waterloo大学Michael Rubinstein、MITのAndrew Sutherlandらが独立再ビルド、Leanカーネルによる全モジュール再検証、独立実装された別カーネルnanodaによる依存宣言の型検査を実施し、エラーなしという結果を報告した。さらにLean・Mathlibコミュニティは共同声明で、「命題も証明も正しいと認証する」と明言した。 [1]

したがって、**「Leanで形式化された命題が論理的に導出されているか」という意味では、現時点の証拠は非常に強い**。これは普通の「AIが証明したと主張している」段階を大きく超える。一方で、Clay Mathematics Institute（CMI）は2026年11月22日現在もRHを「未解決」と掲載している。CMIの賞規則は、適格な媒体への公刊、少なくとも2年間の経過、世界の数学界での一般的受容などを要求しており、今回の4ページの告知文とLeanアーカイブはまだこの手続きを満たしていない。したがって、**「数学的にほぼ確実に正しい形式証明が存在する」と「数学界・Clayが正式にRH解決を認定した」は別の命題である**。 [2]

さらに重要なのは、AIの寄与の認定である。Alpöge本人は、核心部で「Claudeが自分を大きく追い越した」と記し、Stepanov型の方法、theta liftの連鎖、高階の保型形式・相対跡公式・Iwasawa理論などを組み合わせる巨大な構成について、Claudeが技術的主導権を握ったと説明している。これは一次資料であるが、**発見過程そのものの独立監査可能な完全ログが公開されたわけではないため、「どのアイデアがClaude固有の創案だったか」の帰属は、証明の正しさほど強く検証されていない**。一方、証明内容自体は機械検証可能であり、ここは従来のAI発見報道との決定的な違いである。 [3]

本調査の中心的結論は、**これをもって「AIが自律的なゼロイチ発明を広範に行える時代が数か月以内に来る」と断定するのは早いが、「3〜10年のうちに、限定された技術領域で人間の発明者に近い役割を果たすAIが現れる可能性は、2024年時点より大幅に高まった」と評価するのが妥当**というものである。数学では正解をLeanなどで完全に機械検証できるため、AIの大量探索・自己批判・形式検証が極めて強く働く。これに対し物理・化学・生物・工学の「発明」は、現実世界での再現実験、安全性、製造可能性、経済価値、特許の新規性・進歩性まで通過する必要がある。AIがゼロイチを「考える」能力と、ゼロイチを「社会実装まで自律完遂する」能力の間には、依然として大きな距離がある。 [4]

ただし、今回のRHだけを孤立した例と見るのも不適切である。2026年にはAnthropicがClaudeによるRH関連の67.2%下界改善を公表し、後に人間研究者Lamzouriがより透明な別証明を提示したほか、Claudeによるブラックホール物理の未解決問題の解決、OpenAIによるNavier–Stokes問題のLean形式化を伴う解決主張など、**「AIが専門家が答えを知らない問題に新しい結果を出す」事例が連続している**。2023年以前にもFunSearch、AlphaDev、RFdiffusion、自律化学実験系Coscientistが存在しており、量的には「新規知識生成」の境界はすでに越えている。焦点は今後、**新規性、実験的再現性、自律性、経済的有用性を同時に満たすか**へ移る。 [5]

主要調査結果——一次資料と査読状況

資料・出来事	内容	形式・検証	査読・受容状況（2026-11-22）	判定
Levent Alpöge, `remark.pdf`, 2026-11-20	「Claudeと共にRHの証明を発見した」と本人が宣言。GRHにも拡張すると記述。証明はLeanのみ公開。	一次資料。約4ページの説明＋14分割Leanアーカイブ。本人は通常文章版が25万ページ超になったとも説明。 [6]	**査読誌未掲載。** 告知文は論文ではない。	主張の一次ソースとして確実。
Alpöge–Claude Leanアーカイブ	Mathlibの `RiemannHypothesis` およびDirichlet L 関数の零点に関する命題を無仮定で証明。	約175,187ファイル、17,374,215行。`sorry` 等の未証明穴を使わず、標準公理のみ。 [7]	通常のジャーナル査読ではないが、形式検証として非常に強い独立検査を通過。	**形式的正当性について高信頼。**
Michael Rubinstein（Waterloo）独立検証	全175,187モジュールをコンパイルし、Leanカーネル再生、依存公理監査、nanoda独立カーネル検査。	1,484,729依存宣言をnanodaで検査、型エラー0。 [8]	査読論文ではなく独立技術検証。再現手順公開。	証明検証として極めて重要。
Andrew Sutherland（MIT）検証	ソースの独立ビルド、全モジュール・公理依存等を確認。	Lean 4.33.1、Mathlib固定コミットを用いた再現。 [9]	独立再現。	Rubinstein検証と相互補強。
Lean / Mathlibコミュニティ共同声明 2026-11-21	「命題と証明の双方が正しいと認証する」「証明が正しいことに完全な確信」と表明。	複数チェッカー・独立言語実装で証明エクスポートを検査。 [10]	数学専門誌の査読とは異なるが、形式証明コミュニティの公式評価。	**形式証明の真正性を強く支持。**
Clay Mathematics Institute	RHはなお「Unsolved」と掲載。	ミレニアム問題の公式管理機関。 [11]	賞には適格媒体公刊後2年以上＋一般的受容等が必要。 [12]	**公式解決認定はまだない。**
Nature報道、2026-11-20/21	Maynardは「完全にチェックされれば歴史的」、Granvilleは説明文を批判、Buckmasterも巨大さを強調。	専門家コメントを収録。 [13]	数学界の理解・評価は開始直後。	正しさと「理解」は分離している。

資料・出来事	内容	形式・検証	査読・受容状況（2026-11-22）	判定
Rubinsteinの人間可読化作業	164ページの作業稿でzetaの場合を解説。ただし核心不等式の証明はLean側に依存。	Claude、Codex、Geminiも補助に利用。[14]	**作業中、未査読。**	人間的理解は未完成。
Alessandro M. Masulloの別個の人間証明主張	同日、約940ページの純人間作成GRH草稿が公開。	Lean形式化なし。発表後に問題点の報告あり。[15]	未査読・未検証。	Alpöge–Claude証明とは独立。現時点で未確認。
Anthropic 2026年8月のRH関連成果	Claudeが臨界線上の零点割合の下界を約41.6%→67.2%へ改善。	人間数学者が検証、Lean形式化も作成。[16]	後にLamzouriが別のより簡潔な人間証明をプレプリント化。[17]	RH本体の前兆となる独立実績。

ここで特に注意すべきなのは、「一般化リーマン予想」の範囲である。公開されたLean命題が直接証明しているのは、**すべてのDirichlet characterに対するDirichlet L 関数のGRH**であり、あらゆる数体・Dedekind zeta関数・一般の保型 L 関数に対する最広義の「GRH全部」ではない。Rubinsteinの検査ページはこの点を明示している。したがって「Claudeがあらゆる一般化リーマン予想を解いた」という表現は、2026年11月22日時点では**過大**である。[8]

リーマン予想の主張をどこまで「解決」と呼べるか

一次資料の収集。 最重要一次資料はAlpöge本人の11月20日の4ページ文書と、その脚注から配布されたLeanアーカイブである。本人は「Claudeと共に証明を発見した」と明言し、「形式化は完了した」としている。また自分が問題設定・類推・方向付けを行い、終盤ではClaudeが自分を上回ったと記している。方法の概略として、有限体上のRiemann型問題に対するStepanov法を祖型とし、整数上で不足する「重複度」の役割をtheta liftの連鎖に持たせるという発想を掲げている。タイムラインにはGan–Gross–Prasad型period、Rallis seesaw、Bost slope inequality、Iwasawa tower、relative trace formula、Galois/算術幾何など多数の要素が並ぶ。[18]

査読状況。 11月22日時点で、これを掲載した査読済み主要数学誌はない。したがって伝統的な意味で「査読済み論文」と呼ぶのは誤りである。ただし今回の場合、論理的正当性については通常の査読よりはるかに機械的で再現性の高い別経路が存在する。Rubinsteinの検証では、全モジュールがLean 4.33.1上でコンパイルされ、最終定理の依存公理が`propext`、`Classical.choice`、`Quot.sound`というLean/Mathlibの標準的なものに限定され、Lean自身のカーネルと独立したRust実装nanodaがともに証明オブジェクトを受理した。[8]

数学界の評価。 現状は二層に分かれている。形式証明の専門家はきわめて強い肯定に移り、Lean・Mathlib共同声明は明示的に「証明は正しい」としている。一方、数論研究者の課題は「なぜそれが真なのか」「どのアイデアが本質的で、どれが巨大な形式化上の足場なのか」を理解することである。James Maynardは完全に確認されるなら歴史的結果だとし、Andrew Granvilleは告知文が不明瞭であることを批判した。両者は矛盾していない。**形式的真理の確認と数学的理解は別の作業だからである。**[19]

証明形式。 今回は「従来の筆算的証明」でも「Coq/Leanで既存の人間証明を翻訳しただけ」でもない。公開された一次成果そのものがLean形式証明であり、通常の可読論文が後から作られつつある点が異例である。Coqを使ったという根拠は確認できず、主要ツールはLean 4とMathlibである。Leanは命題が既存公理と定義

から導けるかをカーネルでチェックするため、25万ページ相当の自然言語証明を人間が逐行査読する必要はない。ただし、**形式化した定義が意図した数学概念と一致しているか**、そして**巨大証明の中の新しい数学的アイデアを抽出できるか**は、依然として人間側の仕事である。MathlibのRH定義自体は「非自明零点の実部が1/2」という通常の定式化と合致しており、独立検査もこの命題が実際に証明対象であることを確認した。[20]

ギャップ・未解決点。 「証明項の論理的穴」という意味では、公開検査から明白なギャップは確認されていない。しかし、人間可読な統一証明、核心的アイデアの最小化、既存文献との正確な依存関係、著者帰属、Claudeによる創案部分の完全なプロヴェナンスはまだ未完成である。またClayの公式認定もない。このため、本報告では状態を「**形式的正しさは強く検証済み、数学史的・制度的受容は進行中**」と表現するのが最も正確である。[21]

以上から、「AIが人間と共に167年の難問を解いた」という主張の真偽は、**限定付きで真**と評価する。より厳密には、「AlpögeとClaudeが作成したとされる証明について、正しいRH命題を導くLean証明オブジェクトの存在と形式的正当性は独立に強く確認された。ただし、共同発見におけるAIの創造的寄与の内訳と、数学界・Clayによる最終的な制度的受容は未完」である。

AIによる新規発見の実績と限界

リーマン予想以前から、「AIは既知知識の補間しかできない」という命題は既に維持しにくくなっていた。DeepMindのFunSearchは、LLMと評価器を組み合わせ、cap set問題で従来より大きい構成を発見し、20年来の漸近下界を改善した。重要なのは、出力候補を厳密な評価関数が即座に採点できる点であり、LLM単体の「ひらめき」ではなく、生成―検証―選択の閉ループが新規性を生んでいる。[22]

工学・アルゴリズムではAlphaDevが、人間が使っていた小規模ソートより高速な新しいアセンブリ系列を強化学習によって発見し、その一部はLLVMの標準C++ライブラリに採用された。これは「新規アルゴリズムをAIが探索し、実際のソフトウェア基盤に組み込む」という意味で、すでに実用的な発明にかなり近い。もっとも、問題設定、ゲーム化、性能尺度、実装・採用判断は人間側が与えている。[23]

生物学ではRFdiffusionが、自然界に存在しないタンパク質骨格や結合タンパク質を生成し、実験で構造・機能を確認した。インフルエンザHAに対する設計binderではcryo-EM構造が計算設計と原子レベルで近いことも確認された。2026年にはAnthropicも、Claudeによるbinder設計キャンペーンのend-to-end自動化を報告している。ただし、これらは「自然界にない分子を設計する」という意味でゼロイチ性は高い一方、標的選択や成功基準、実験環境、安全性判断は依然として人間が設定している。[24]

化学ではCoscientistが、文献検索、合成計画、装置マニュアル読解、コード生成、ロボット操作、反応最適化までをGPT-4系エージェントで閉ループ化した。これは「考えるAI」と「現実世界を操作するロボット」の接続として重要だが、主に既知化合物や既知反応空間を対象としており、「世界で初めての有用化学反応を自律発見して製品化」まで到達したわけではない。[25]

物理では2026年11月、AnthropicはClaudeがブラックホール・量子重力に関係する二つの未解決問題を解いたと報告した。そのうち一つでは、数値計算に見えていたパターンからspectral curveを導入し、既知の解析的上界を改善して最適値を導いたとされ、Lean形式化も行われた。Anthropic自身によれば、問題選択後は複数エージェントが数週間ほぼ自律的に探索し、人間物理学者が主要結果を確認した。これは「AIが既存の別分野の道具を新しい問題へ移植する」という、発明に近い創造行為の例である。[26]

同時に、失敗例は非常に重要である。材料科学のA-Labは2023年、ロボット自律実験で多数の材料を合成したとして注目されたが、後に「新規材料」という表現に問題があることが判明し、2026年のNature訂正では、対象は予測プラットフォームにとって新しかったのであって「科学にとって新規」とは限らなかったと明

記された。さらに一部のXRD同定も再検討で不確実になった。これは**AI発明評価で最も危険なのが「成功判定器そのものが間違っている」ケース**であることを示す。[27]

GNoMEも数十万規模の安定候補を予測したが、材料の計算上の安定性は「未知」「合成可能」「機能的」「有用」を意味しない。2026年にも重複・既知構造との近似をめぐる批判が続いている。これは、探索空間が巨大になればなるほど、「候補生成能力」より「本当に新しいかを判定する能力」がボトルネックになることを示す。[28]

したがって現在の証拠を総合すると、AIはすでに「新しい候補を出す」段階を越え、**一部の領域では専門家未到達の結果を生成して検証する段階**に入っている。しかし、最も成功している領域には共通点がある。すなわち、正解判定が自動化できる、探索空間を計算機上で表現できる、試行コストが低い、失敗を大量に捨てられる、という条件である。RHはこの条件をほぼ理想的に満たす。

数学的証明と科学的発見・技術発明の違い

数学でAIが急速に強くなったことを、そのまま「発明一般」に外挿できない最大の理由は、**真理判定器の質が違う**ことにある。

数学では、命題が適切に形式化されていれば、Leanのような小さなカーネルが「証明項が公理・定義から合法的に導かれているか」を決定論的に検査できる。同じ証明ファイルは世界中で原理的に再現でき、今回実際に複数機関・複数チェッカーで再現された。誤った候補を百万個作っても、検証器が高速に排除できるため、巨大なAI探索が非常に有利になる。[29]

自然科学ではそうはいかない。新しい超伝導体、触媒、薬剤、電池材料などについて、計算結果が良くても、合成できるか、測定値が再現するか、別研究室でも再現するか、長期安定性があるか、毒性や製造スケールが許容範囲かを実験しなければならない。AlphaFold 3のような高性能モデルでも、予測困難なコンフォメーションや誤ったキラリティなど明示的な限界が存在する。AIの「自信」は物理世界の真理判定器ではない。[30]

技術発明ではさらに一段増える。ある設計が物理的に動くことだけでは不十分で、既存技術に対する新規性・非自明性、量産性、コスト、規制適合、安全性、顧客価値を満たす必要がある。米国特許商標庁は2025年にAI支援発明のガイダンスを改訂し、AIが使われても通常の発明者基準を適用し、**発明者として記載できるのは自然人のみ**という立場を維持した。欧州特許庁も2026年時点で同様にAI単独を発明者として認めていない。つまり、「AIが実質的に発明した」と技術者が考えても、現行の主要特許制度では権利帰属を人間中心に構成する必要がある。[31]

この差を短く表すと、**数学は「生成→形式検証」がほぼ完全にデジタル閉ループ化できるが、発明は「生成→シミュレーション→製作→測定→再現→規制→市場」という長い物理・社会ループを持つ**。したがって、RH級の数学成果は「AI創造性」の非常に強い証拠である一方、「AIが明日新型電池や新薬を完全自律で発明する」ことの直接証明ではない。

ただし逆方向の含意もある。自動実験室、ロボティクス、高スループット測定、シミュレーション、形式検証が結びつけば、自然科学側も徐々に数学に近い閉ループへ移る。CoscientistやA-Labは未完成ながらその初期形態であり、ボトルネックが「AIのアイデア不足」から「信頼できる現実世界の評価器・ロボット・データ品質」に移りつつある。[32]

今後3〜10年のシナリオと「ゼロイチ発明」の判断指標

ここでいう「自律的ゼロイチ発明」を、単にAIが候補を出したことではなく、**人間が答え・設計原理を事前に持たない課題について、AIが主要アイデアを生成し、検証・反復を主導し、実物または実運用で新規有用性が独立再現されること**と定義する。この定義なら、FunSearchやRHは重要な前駆例だが、産業発明としての完全条件はまだ満たさない。

シナリオ	期間・主な前提	予想される状態	定性的確率
楽観	2027〜2030年。推論能力・エージェント協調・自動実験室・シミュレータの信頼性が現在の数学AIと同程度の速度で進歩。	AIが自分で研究課題を細分化し、文献調査、仮説生成、実験設計、ロボット実行、解析を連続実行。新触媒、材料、タンパク質、アルゴリズムなどで「人間が事前に発明原理を与えていない」成果が複数独立再現される。	**中程度（約30%）**。RH・ブラックホール物理・自律実験の進歩は支持材料だが、外界検証が制約。 33
中立	2027〜2036年。AI能力は継続的に改善するが、実験・データ・安全性・知財が律速。	数学・ソフトウェア・回路・タンパク質設計など「高速評価可能」領域では自律発明が日常化。薬、材料、機械では人間研究者が目標設定・最終判断を担う「AI主任研究員＋人間PI」型が主流。	**最も高い（約50%）**。現在の事例群と最も整合。
悲観	2027〜2036年。長期推論の信頼性、OOD一般化、実験データ品質、ロボット操作、規制・IP問題が進歩を阻害。	AIは候補生成・文献探索・最適化には強いが、本当の新規発明は人間が核心仮説を与えた場合に限定。材料科学で見られた「新規性の誤認」が頻発し、自律研究の導入が鈍る。	**無視できない（約20%）**。A-LabやGNoME論争が示すリスク。 34

これらの数字は統計モデルから算出した予測ではなく、2026年11月時点の証拠を基礎にしたシナリオ重みである。特に「AIモデル性能だけ」を見て予測すると楽観側に偏りやすい。実際には、ロボット・計測器・試薬・製造・安全審査・特許制度など、モデル外の速度が全体を決める。

「AIが本当にゼロイチ発明を自律的に実現した」と判定するためには、次の八条件を同時に測るのが望ましい。これは本報告で提案する評価枠組みである。

1. **新規性**：発明公開前の論文、特許、データベースを第三者が調査し、実質同一の技術がないこと。目安は独立した二つ以上の先行技術調査で新規性が維持されること。
2. **非自明性**：専門家パネルが、既存技術の単純なパラメータ探索・組み合わせではないと評価すること。特許の進歩性に近い基準を用いる。
3. **AI主導性**：課題設定後の核心仮説・構造・機構の生成に、人間が答えを注入していないこと。プロンプト、ツール呼び出し、候補履歴を監査可能な形で保存し、核心発想の大部分がAI側から生じたと確認できること。
4. **閉ループ自律性**：候補生成だけでなく、失敗解析→再設計→再実験の複数サイクルをAIが自律実行すること。人間介入を総意思決定の10〜20%未満など定量化する。
5. **独立再現性**：少なくとも二つの独立研究機関が、AI開発元と別の設備・担当者で結果を再現すること。数学なら独立proof checker、科学なら独立実験を用いる。
6. **実用性**：既存最良技術に対し、性能・コスト・エネルギー・安全性など事前定義した指標の少なくとも一つで有意な改善を示すこと。

7. **法的成立性**：特許審査、規制申請、責任主体の確定などを通じ、人間発明者・AI支援の関係が法的に処理可能であること。現状、米欧ではAI単独発明者は認められていない。 [31]
8. **経済的実証**：単なる論文ではなく、ライセンス、量産、売上、コスト削減など第三者が測れる経済価値が発生すること。たとえば採用後2年以上継続利用され、既存技術を置き換える程度を一つの閾値とする。

この基準をRHに当てはめると、**新規性・検証可能性・AI主導性について非常に強い証拠が出始めた一方、経済的実用性、特許的意味での発明性、物理的再現性はそもそも対象外**である。つまりRHは「AIが独創的な知的成果を作れるか」への強いYesであり、「AIが自律発明者として産業を回せるか」への直接のYesではない。

結論と推奨

2026年11月22日時点の最も正確な結論は、**「AIと人間がリーマン予想を解いた」という報道は、少なくとも形式数学上の証拠という観点では驚くほど強く裏付けられている**、ということである。公開された証明は単なる自然言語のAI出力ではなく、独立機関・複数カーネルでチェックされたLean形式証明であり、Lean/Mathlibコミュニティは正しいと公式に認証している。したがって「AIはまだ新しい数学を作れない」という一般論は、この事例を含む2026年の成果群とは両立しにくい。 [35]

同時に、「正式に167年の問題が解決済み」という制度的表現には留保が必要である。ClayはまだRHを未解決と掲載し、査読誌公刊・2年以上の経過・数学界の一般受容という正式プロセスを経ていない。また、形式的に正しいことと、人間数学者が巨大証明の原理を理解・簡約したことも同義ではない。2026年11月22日という発表後わずか二日の時点で、「数学史上の最終評価」まで確定したと表現するのは早すぎる。 [36]

それでも、AIの「ゼロイチ能力」については評価を上方修正すべきである。2023年のFunSearchやAlphaDevは新構成・新アルゴリズムを示し、RFdiffusionは新しいタンパク質を実験的に実現し、Coscientistは実験閉ループを部分自律化した。2026年にはRH関連、ブラックホール物理、Navier–Stokesなど、専門家未解決問題へのAIの関与が急速に増えた。個々の成果には重要な留保があるが、**「AIは既存知識の検索・要約だけ」というモデルは既に事実に合わない**。 [37]

一方、「完全自律発明者」が近いかという問いには、**限定領域では近い、一般技術領域ではまだ未確定**と答えるのが妥当である。特に数学、アルゴリズム、形式設計、分子・タンパク質設計のように評価器を高速化できる領域では、今後3年以内にAI主導のゼロイチ成果が珍しくなくなる可能性がある。他方、新薬、材料、エネルギー、機械装置などでは、実験・製造・規制・安全性という物理的ボトルネックのため、完全な自律化には5〜10年以上かかる可能性も十分ある。

対象	短期・今後1年	中期・今後3年	長期・今後10年
研究者	AIを「文章作成器」ではなく、仮説生成・反例探索・形式検証・コード実験の共同研究環境として導入する。全探索ログと人間介入を保存し、AI由来アイデアのプロヴェナンスを残す。	Lean等の形式検証、自動実験、独立再現を研究ワークフロー標準にする。AI成果専用の査読・再現プロトコルを作る。	人間研究者の役割を「全計算を自分で行う人」から「良い問題、評価関数、因果解釈、倫理的目的を設定する人」へ再設計する。

対象	短期・今後1年	中期・今後3年	長期・今後10年
政策立案者	「AIが発明者か」という抽象論だけでなく、AI利用履歴、責任主体、独立再現、出所開示の基準を整備する。現行の人間発明者原則との整合を明確化する。[38]	AI生成発明の特許新規性調査、prior art爆発、責任・製造物安全、研究不正対策を国際調整する。	自律研究システムが大量の特許可能成果を生成する場合に備え、発明者制度、特許期間、独占範囲、公共研究アクセスを再検討する。
企業	AI研究エージェントをPoCではなく、明確な性能評価器を持つR&D課題に投入する。候補数ではなく独立再現率をKPIにする。	ロボット実験・シミュレーション・特許調査・製造評価を一つの閉ループに統合し、「人間が答えを教えない」内部ベンチマークを設置する。	AIが毎週多数の発明候補を生む前提で、IPポートフォリオ、研究組織、人員配置、品質保証を再設計する。競争優位はモデル単体より、独自実験設備・データ・評価器に移る可能性が高い。

最終的に、今回のリーマン予想事件から引き出すべき最も重要な示唆は、「AIが人間のような天才になった」と単純化することではない。むしろ、**巨大な並列探索、既存文献の横断、自己批判、形式検証を一つのループに接続したとき、従来は人間の直観だけが到達できると思われていた研究フロンティアを越える事例が現れ始めた**ということである。数学ではその検証器がすでに完成しているため、変化が最初に可視化された。次に問われるのは、化学・生物・材料・工学でLeanに相当するほど信頼できる「現実世界の評価ループ」を構築できるかである。

したがって、「ゼロイチ発明ができるAIはもうじきか」という問いへの本報告の答えは、**「限定された領域では既に入口に到達しており、3年程度で明確な事例が増える可能性は高い。しかし、人間の介入なしに課題発見から製品・特許・市場投入まで完遂する汎用的な自律発明者については、RHの証明だけを根拠に“目前”と断定することはできない」**である。今回のRHはその未来を証明したのではなく、**従来想定していたより未来が近い可能性を大幅に高めた証拠**と位置づけるのが最も厳密である。

---

[1] [3] [6] [18] Open the research source ↗
https://alpo.ge/remark.pdf

[2] [36] The Millennium Prize Problems - Clay Mathematics Institute
https://www.claymath.org/Millennium%20Problems/

[4] [22] [37] Mathematical discoveries from program search with large language models
https://www.nature.com/articles/s41586-023-06924-6?utm_source=chatgpt.com

[5] [16] Claude has improved on a longstanding lower bound for the fraction of zeros of the Riemann zeta function that satisfy the Riemann hypothesis \ Anthropic
https://www.anthropic.com/research/riemann-zeta?trk=article-ssr-frontend-pulse_little-text-block

[7] [8] [29] GRH Lean Check
https://mrubinst.github.io/grh-lean-verification/

[9] RH Lean proof: verification files
https://math.mit.edu/~drew/grh/?utm_source=chatgpt.com

[10] [19] [21] [33] [35] Lean verification of the Riemann Hypothesis | Lean community blog
https://leanprover-community.github.io/blog/posts/riemann-hypothesis/?utm_source=chatgpt.com

11 Riemann Hypothesis - Clay Mathematics Institute
https://www.claymath.org/millennium/Riemann-Hypothesis/?utm_source=chatgpt.com

12 Microsoft Word - Millennium Prize Rules.docx
https://www.claymath.org/wp-content/uploads/2022/03/millennium_prize_rules_0.pdf

13 15 Mathematicians astounded by AI and human-written proofs of iconic Riemann hypothesis
https://www.nature.com/articles/d41586-026-03812-7?code=87d1dd09-c54e-49c7-942c-54f8fa3bdc3b&utm_source=chatgpt.com

14 Alpöge/Claude proof of the Generalized Riemann Hypothesis
https://www.math.uwaterloo.ca/~mrubinst/alpoge_claude_grh/

17 A new proof that more than $2/3$ of the zeros of the Riemann zeta function are simple and on the critical line
https://arxiv.org/abs/2609.02882?utm_source=chatgpt.com

20 Mathlib.NumberTheory.LSeries.RiemannZeta
https://leanprover-community.github.io/mathlib4_docs/Mathlib/NumberTheory/LSeries/RiemannZeta.html?utm_source=chatgpt.com

23 Faster sorting algorithms discovered using deep reinforcement learning
https://www.nature.com/articles/s41586-023-06004-9?utm_source=chatgpt.com

24 De novo design of protein structure and function with RFdiffusion
https://www.nature.com/articles/s41586-023-06415-8?utm_source=chatgpt.com

25 32 Autonomous chemical research with large language models
https://www.nature.com/articles/s41586-023-06792-0?trk=article-ssr-frontend-pulse_little-text-block&utm_source=chatgpt.com

26 Claude Fable 5.2 solves two open problems in the physics of black holes \ Anthropic
https://www.anthropic.com/research/claude-black-hole-physics?utm_source=chatgpt.com

27 34 Author Correction: An autonomous laboratory for the accelerated synthesis of inorganic materials
https://www.nature.com/articles/s41586-025-09992-y?utm_source=chatgpt.com

28 Scaling deep learning for materials discovery
https://doi.org/10.1038/s41586-023-06735-9?utm_source=chatgpt.com

30 Accurate structure prediction of biomolecular interactions with AlphaFold 3
https://www.nature.com/articles/s41586-024-07487-w?utm_source=chatgpt.com

31 38 Revised inventorship guidance for AI-assisted inventions | USPTO
https://www.uspto.gov/subscription-center/2025/revised-inventorship-guidance-ai-assisted-inventions?utm_source=chatgpt.com