ChatGPT

「AGIはもう来ている」はどこまで立証されているか——Yahoo!ニュース記事と量子物理学者ミコラ・マクシメンコの主張を検証

エグゼクティブサマリー

指定されたYahoo!ニュース記事「『汎用人工知能・AGIはもう来ている』量子物理学者が告げた理由」について、同題の**Forbes JAPAN掲載版、英語版Forbes原稿、記事執筆者John Koetsierによるミコラ・マクシメンコ（Mykola Maksymenko）への元インタビュー、マクシメンコの学術業績、Haiquの公表資料、量子計算・AGI評価に関する一次文献**を照合した。指定Yahooページ自体は検索環境から直接本文を取得できなかったため、本文の検証には2026年9月14日公開のForbes JAPAN版と、その原取材を中心に用いた。Forbes JAPAN版は、マクシメンコがAIエージェントを使って「博士課程で半年ほどかかった研究」を約1週間で再構築し、数式上の誤りまで見つけたことなどを、「AGIはもう来ている」という発言の背景として紹介している。 [1]

結論を先に述べると、この記事が示す証拠は「AGIの到来を科学的に立証した」と評価するには不十分である。信頼度は「低」である。一方で、より限定した命題——**「最先端LLM＋エージェント＋専門ツールが、熟練研究者の科学研究プロセスを大幅に圧縮し始めている」**——には、マクシメンコの事例以外にもGoogleのAI co-scientistやAlphaEvolveなど独立した事例があり、こちらは**中〜高程度の確度で支持される**。 [2]

最大のポイントは、マクシメンコ自身の元発言が見出しよりかなり慎重なことである。彼はチームに **“Guys, I think AGI is here.”**（「みんな、AGIはもう来ていると思う」）と述べた一方、同じインタビューで **“It’s like a weak AGI in some sense.”**（「ある意味では弱いAGIのようなものだ」）と限定している。さらに、成果には適切なプロンプトと専門家による知識・ガイダンスが必要だと説明し、システムが完全に一貫して動作するわけではないとも認めている。したがって、原発言は「自律的な人間レベルAGIが完成した」という宣言というより、**高度に道具化・エージェント化されたAIシステムを“弱いAGI”と呼ぶ私的な分類**に近い。 [3]

また、記事が印象的に結びつけている**量子コンピューターは、AGIを示す論理上の必須要件ではない**。マクシメンコの事例で量子計算は、AIエージェントが利用する専門計算ツールの一つである。量子ハードウェア上で有用な研究ワークフローを自動化できたとしても、それは「広範な未知課題に人間並みに適応できる一般知能」が成立したことを直接示さない。むしろ証明されるのは、**AIによる科学ワークフローのオーケストレーション能力**である。 [4]

さらに、記事中の最も強い数字——「9時間・3万ドルだった分子動力学シミュレーションを約30秒・25ドルにした」——は、現在確認できる公開情報では**Haiqu側の社内実験・製品発表に由来するベンダー主張**であり、査読論文による独立再現ではない。Haiqu自身は最大1000倍の実行コスト削減等を宣伝しているが、Forbes JAPANもこの点について独立検証が必要だと留保している。 [5]

要約判定表

記事・マクシメンコの主張	公開された一次証拠	本調査の評価
「AGIはもう来ている」	本人インタビュー。ただし本人も「弱いAGI」「完全には一貫しない」と限定	**AGI成立の証拠として低**

記事・マクシメンコの主張	公開された一次証拠	本調査の評価
博士研究の半年分を約1週間で再構築	本人の回顧的自己報告。具体的論文、プロンプト、実験ログは公開されていない	**研究加速の示唆として中、AGI証拠として低**
AIが過去研究の数式ミスを発見	本人談。該当数式・訂正文・再現資料は公開資料から特定できず	**独立検証不能**
ゲノム量子計算を短期間で再現	元となるQ4BioのHepatitis Dゲノム量子エンコード実験自体は公式確認可能	**元研究の存在は高、AIによる再現性能は低〜中**
現在の100〜200物理量子ビットでも「量子utility」が可能	一部タスクでは妥当な研究方向。ただし量子ビット数だけでutilityは決まらない	**限定命題なら中、一般論なら低**
3万ドル・9時間 → 25ドル・30秒	Haiqu側の製品・内部テスト発表	**独立証拠として低**
AIが科学研究を大幅に高速化し始めている	AI co-scientist、AlphaEvolve等の別系統の実績あり	**中〜高**
量子計算＋AIエージェントがAGI成立を示す	AGIの一般性・自律性を検証する比較実験なし	**非常に低**

記事の主張と一次ソースを照合する

記事が構築しているストーリー

Forbes JAPAN版が提示する論証は、おおむね次の連鎖である。

第一に、マクシメンコは理論・量子物理学を背景とする研究者で、現在は量子ソフトウェア企業Haiquの共同創業者兼CTOである。記事によれば、彼が博士課程時代に半年ほどかけた研究課題を、最新LLMを組み込んだAIエージェントと量子計算環境に与えたところ、週末から約1週間程度で再構築され、過去の数式の誤りまで見つかった。[6]

第二に、同種の高速化が一例だけではないと記事は示唆する。マクシメンコは、別の研究パートナーが従来約1年を要したような研究について、エージェントを用いて数週間で有望な複雑な結果に到達しつつある、と述べる。ただし元インタビューでは、彼自身が「まだ最終結果には到達していない」と留保している。記事の要約より原インタビューのほうが、この不確実性が明確である。[7]

第三に、ゲノム研究の例として、Wellcome Sanger Institute、Oxford、Cambridge、Melbourne等のQ4Bioチームが Hepatitis D virusの全ゲノムをIBMの量子コンピューターにエンコードした研究を挙げる。マクシメンコは公表情報をAIエージェントに与え、自身はゲノム研究の専門家ではないにもかかわらず、必要なアルゴリズムやエンコーディング方法を推測させ、短期間で類似実験を組み立てたと説明する。[8]

元になったQ4Bio成果そのものは確認できる。Wellcome Genome Campusは2026年4月10日、Hepatitis D virusの約1,700塩基規模のゲノムをIBMの156量子ビットHeronプロセッサにエンコードし、117量子ビットを使用したと公表している。Oxford大学も、この研究の目的を量子コンピューターによるゲノム／パンゲノム処理の**実行可能性を示すこと**と説明している。[9]

ここには重要な区別がある。**「完全なゲノムを量子状態としてロードできた」ことと、「古典コンピューターより高速・安価にゲノム解析できた」ことは同じではない。**公式発表は前者を実証したものであり、後

者の一般的な量子優位を示したものではない。Oxfordの説明も、量子データ表現と将来のアルゴリズム処理に向けた基盤として位置づけている。[10]

第四に記事は、Haiquが2026年5月に発表した「Agentic Quantum Operating System」を紹介する。これは研究課題をAIエージェントが分解し、適切な量子アルゴリズムを選び、HaiquのSDK／runtimeを介して実機QPUへ落とし込む構成である。Haiqu公式サイトも、研究アイデアからhardware-ready prototypeまでをAI層・SDK・runtimeで統合するという説明を行い、「最大1000倍」のコスト削減を掲げている。[11]

「量子物理学者」という人物紹介は妥当か

マクシメンコを「量子物理学者」と呼ぶこと自体に大きな問題はない。大学・研究機関の略歴では理論物理学の研究歴が確認でき、APSの査読誌には、カゴメ格子、量子多体系、many-body localizationなどに関する彼の共著論文が存在する。つまり、単に企業経営者が量子技術を語っているわけではなく、**凝縮系・多体系物理に実際の研究業績を持つ研究者**である。[12]

ただし、今回のAGI発言時の立場は大学の中立的評価者ではなく、**自らのAI＋量子ソフトウェア製品を開発するHaiquの共同創業者・CTO**である。これは発言を無効にするものではないが、製品性能に関する主張について独立再現を要求すべき利益相反上の理由になる。[4]

元インタビューと記事見出しの距離

原インタビューでマクシメンコが語った二つの短い原文は、ニュアンスを理解するうえで重要である。[13]

> **“Guys, I think AGI is here.”**
> 「みんな、AGIはもう来ていると思う。」

しかし、その直後の自己評価は、

> **“It’s like a weak AGI in some sense.”**
> 「ある意味では、弱いAGIのようなものだ。」

である。[13]

さらに彼は、AIを適切に機能させるには正しいプロンプト、専門知識、ガイダンスを加える必要があり、完全な自己一貫性もない、と説明する。したがって、「AGIは来ている」という一文のみを抜き出すより、**「専門家が適切に誘導し、多数のツールを与えると、弱いAGIと呼びたくなるほど強力な研究能力を発揮する」**と読むほうが原発言に忠実である。[13]

これは記事の見出しの真偽を左右する重要な限定条件である。

AGIとは何を満たせば「来た」と言えるのか

AGIには統一された科学的定義がない。そのため「AGIは来たか」という問いは、**先に判定基準を固定しなければ反証可能な主張にならない**。Google DeepMindの研究者らも、AGI論議では「generality（一般性）」と「performance（性能）」を分け、さらにautonomy（自律性）は別軸として扱う必要があると指摘している。[14]

代表的な定義との比較

定義・枠組み	AGIの要求水準	マクシメンコ事例で示されたか
OpenAI Charter型	人間を「ほとんどの経済的に価値のある仕事」で上回る、高度に自律的なシステム	**いいえ**。広範な職務、自律性、再現性を測っていない
DeepMind Level 1: Emerging AGI	幅広い非物理タスクで非熟練者程度以上	**可能性あり。ただしこの一事例で証明はできない**
DeepMind Level 2: Competent AGI	「ほとんど」の認知タスクで、当該技能を持つ成人の50パーセンタイル以上	**未立証**
ARC-AGI型	人間が獲得可能な新技能を、人間に匹敵する効率で獲得・一般化できること	**未立証**
マクシメンコの informal "weak AGI"	適切な誘導とツールがあれば広範な高度研究を実行できる、とする緩い用法	**本人の定義では成立し得る**

OpenAI Charterの古典的な定義はAGIを **"highly autonomous systems that outperform humans at most economically valuable work"** としており、マクシメンコの「専門家が正しくプロンプトし、ガイドする」研究ワークフローとはかなり距離がある。 15

一方、DeepMindの「Levels of AGI」はより緩やかである。同論文は2023年当時のChatGPT、Bard、Llama 2、Geminiなどをすでに Level 1の「Emerging AGI」に分類する一方、**Level 2のCompetent AGIは当時まだ達成されていない**とした。ここではマクシメンコの「weak AGI」という語法と一定の親和性がある。したがって、**「弱い／初期段階のAGI」というラベル自体は学術的にあり得ない用法ではない**。しかし、それなら「AGI到来」の根拠は今回の量子実験ではなく、そもそも現代LLMが示す広範な能力である。 14

2026年の状況はさらに進んでいる。ARC Prizeは2026年9月3日、GPT-6 AstraがARC-AGI-3 Semi-Privateで標準harnessでは62.7%、プロバイダー固有のcontext管理を用いるharnessでは99.9%に到達したと報告している。未知環境での探索、world model構築、目標推定、計画実行という点では非常に強い結果である。 16

それでもARC Prize自身は、benchmark saturationについて **"we are not claiming that it is AGI"** と明示している。理由は、ARC-AGI-3自体が決定論的で閉じた環境という限定されたscopeしか持たないためである。すなわち、**AGIを意識して設計されたベンチマークをほぼ満点で解くことさえ、単独ではAGIの証明にならない**。この点は、博士研究一件の高速再構築をAGI証拠とする推論にはさらに厳しい意味を持つ。 16

METRの自律エージェント評価にも同じ教訓がある。METRはfrontier AIのtask-completion horizonを、人間専門家が要する作業時間に対応づけて測定しているが、自ら「その時間だけAIが自律動作できるという意味ではない」「すべての知的タスクや職業をこなせるという意味でもない」と注意している。テストは主としてソフトウェア工学、機械学習、サイバーセキュリティの明確に採点可能な問題に偏り、現実世界の曖昧な仕事では性能が異なる。 17

したがって、本件をAGI判定に用いるなら最低でも、

一般性 × 水準 × 自律性 × 新規性 × 頑健性 × 再現性

を分離して測る必要がある。

博士研究一件を短時間で再構築したことは「水準」の局所的証拠にはなる。しかし、**多数の未知分野への一般化、専門家なしでの自律実行、失敗率、長期タスクでの頑健性**についてはほとんど情報を与えない。

マクシメンコの具体的主張を技術的に検証する

「半年の博士研究を一週間」は何を証明するのか

この事例は非常に興味深いが、科学実験としては必要な情報が大幅に不足している。

公開された記事・インタビューからは、少なくとも以下を特定できない。

- 使用したLLMの正確なモデル・バージョン
- system prompt、研究指示、反復回数
- エージェントのtoolchainと検索／retrieval対象
- 当該博士研究のどの論文・どの式・どの部分を再構築したか
- 過去論文やコードがモデルの訓練データまたは検索対象に入っていたか
- AI単独、AI＋古典計算、AI＋量子計算、人間専門家とのablation比較
- 同一課題を何回実施し、何回成功したか
- 「一週間」の計算時間、wall-clock time、人間介入時間の内訳
- AIが発見したという「数式の誤り」の具体的位置、訂正内容、第三者確認

元インタビューで確認できるのは、マクシメンコ自身が「博士課程時代には半年ほど要した」「エージェントが再構築し、式の一つにbugを見つけた」と報告していることまでである。[3]

このため、**本人の経験談として信じる合理的理由はあるが、再現可能なscientific evidenceとはまだ言えない**。

特に重要なのが**既知研究の再構築と新規研究の発見を区別すること**である。博士課程で本人が既に行った研究であれば、その研究や近接文献が公開されている可能性がある。AIが既知の方法を検索・再構成したのか、訓練データに含まれていた知識を復元したのか、白紙から新たに導出したのかは全く異なる。訓練データ汚染・retrieval leakageを排除するblind holdoutがなければ、**「半年分の科学的創造を一週間で代替した」という強い結論は出せない**。

逆に、本当に「非公開だった研究内容」を限定情報から独立再発見し、未知の式誤りまで正しく特定したことを第三者がblind evaluationで確認できれば、これはかなり強いscientific reasoning evidenceになる。ただし、それでも一領域での能力であってAGI全体の証明ではない。

ゲノム再現例はより強いのか

ここにも同様の問題がある。

Q4Bioの元成果は実在し、国際研究チームがIBM Heron上にHepatitis D virusゲノムを量子形式でロードしたことは公式資料で確認できる。156物理量子ビット中117量子ビットを使用したという具体値も公表されている。[9]

マクシメンコによれば、彼はゲノム分野の専門家ではないまま、公表発表をHaiquエージェントに入力し、システムに必要なアルゴリズムやencodingを推測させ、類似結果を再現し、突然変異の操作まで試した。これは**専門領域間の参入コストをAIが大幅に下げる**可能性を示す興味深い例である。[18]

しかし、これは原理的には「公表済みの研究成果を、AIが利用可能な情報から再構成した」という課題でもある。原チームが数年研究したことと、成果発表後にその方法を模倣することは難易度が同じではない。研究発見には、何を試せばよいか分からない探索期間、失敗実験、問題設定そのものの発明が含まれる。**最終的な成功例が公表された後の再実装時間を、元研究の発見時間と直接比較すると後知恵バイアスが生じる。**

したがって、この例が強く支持するのは、

> 「LLMエージェントが公開された異分野研究を読み解き、専門ツールへ変換する速度が急速に高まっている」

という命題であり、

> 「未知の研究課題を独立して発見できる一般知能が成立した」

という命題ではない。

「100〜200量子ビットでutility」はどこまで妥当か

ここではマクシメンコの主張を完全に否定するのも適切ではない。

2023年、IBM系研究者らは127量子ビットEagleプロセッサを用いてkicked Ising modelの期待値を計算し、誤り緩和を併用することで、当時の単純なbrute-force classical computationでは難しい領域に到達したとしてNatureに「fault tolerance以前のquantum utility」の証拠を報告した。したがって、「完全な誤り訂正型量子コンピューターを待たずに科学的に有用な計算を行える可能性」という思想そのものは、マクシメンコ独自のものではない。[19]

しかし、このエピソードには量子utilityの評価の難しさも凝縮されている。2024年、Begušić、Gray、ChanはScience Advancesで、同じ127量子ビット実験について、tensor network等の近似古典アルゴリズムを用いることで量子実験より桁違いに高速に計算し、実験精度を上回るところまで収束できると報告した。[20]

つまり、

「127量子ビットある」→「スーパーコンピューター級」

という一般的換算式は存在しない。

実用性は少なくとも、

$$\text{task} + \text{circuit depth} + \text{fidelity} + \text{connectivity} + \text{sampling cost} + \text{error mitigation overhead} + \text{best classical base}$$

で決まる。古典アルゴリズム側も進歩するため、「量子優位」の境界は動く。

この観点から見ると、記事・インタビュー中の「数百量子ビットが最大級の古典コンピューターと同等、あるいは少し後ろ」という表現は、**特定タスクについてならあり得るが、一般的な計算性能比較としては技術的に不正確**である。[21]

物理量子ビットと論理量子ビット

記事で「100〜200の物理／noisy qubitsでもutilityがあり得る」と述べる点には一定の意味がある。fault-tolerant量子計算では、多数の物理量子ビットから誤り訂正されたlogical qubitを構築する。一方、NISQ型

アプローチでは、完全なlogical qubitを作らず、短い回路・error mitigation・classical/quantum hybrid手法を使って価値を引き出そうとする。IBMの2023年実験はまさにこの路線の代表例である。 19

マクシメンコらの2024年の量子化学関連研究にも、回路深度の削減やmiddlewareによるerror mitigationを使い、IBM Heron上でHamiltonian dynamicsをより実行可能にするアプローチが報告されている。これは彼の発言が少なくとも自身の技術研究と整合していることを示す。 22

しかし、**「実行可能になった」ことと「最高の古典手法より経済的に優位になった」ことは別**である。quantum utilityを主張するなら、同じ精度・同じ問題・同じend-to-end条件で最良の古典アルゴリズムと比較しなければならない。

「3万ドル・9時間 → 25ドル・30秒」

これは記事中で最もセンセーショナルな定量値だが、最も慎重に扱うべき数字でもある。

Haiquの2026年5月の製品発表に基づく報道では、あるmolecular dynamics simulationについて、従来の量子実行では9時間超・約3万ドルだったものをHaiquの最適化された実行では約30秒・25ドルにしたとされる。Haiquの現行公式ページも、runtime optimizationによりQPU実行コストを「up to 1000×」削減すると宣伝する。 23

ただし、公開された査読論文による同一条件での独立再現は本調査では確認できなかった。さらに、この比較には、

- 元の3万ドル実装がどの程度最適化されていたか
- 同一のnumerical accuracyを達成しているか
- QPU shots、error mitigation、classical preprocessingをどう計上したか
- hardware、queue time、クラウド料金条件が同じか
- 「simulation」がどこまで同じ科学量を求めているか

といった情報が必要である。

したがって現段階では、**「Haiquが報告している内部性能値」以上の強さを持たせるべきではない**。Forbes JAPAN自身も独立検証の必要性を指摘しており、この留保は適切である。 24

関連研究は「AGI到来」を支持するか

マクシメンコの観察を孤立した逸話として切り捨てるのも正しくない。2025〜2026年には、**AIが熟練研究者の研究時間を圧縮する**というより広い現象を示す一次資料が増えている。

支持側——科学研究を高速化する能力は実在する

Google Researchの「AI co-scientist」はGemini 2.0を基盤とするmulti-agent systemで、Generation、Reflection、Ranking、Evolutionなどの専門エージェントを使って研究仮説を生成・評価・改良する。研究者が自然言語で研究目標を与える構造であり、完全自律システムではなく**scientist-in-the-loop型**である。 25

それでも、その成果は単なる文章生成より強い。Googleは、急性骨髄性白血病のdrug repurposing候補を実験で検証した事例や、抗菌薬耐性に関する未公開の研究成果を、公開文献のみを用いて独立に再発見させるテストを報告している。ただしGoogle自身、専門家を含む評価規模が小さいこと、factuality、literature review、external verificationなどに改善余地があると明記している。 25

これはマクシメンコの「過去研究をエージェントが再構築した」という話とかなり近い。重要なのは、Googleはこの能力を**科学者を支援するassistive technology**として位置づけており、AGI達成とは表現していないことである。 25

さらにDeepMindのAlphaEvolveは、LLMによるプログラム生成、automated evaluator、evolutionary searchを組み合わせ、行列積アルゴリズム、Googleデータセンター、TPU設計などで実用的改善を報告している。2025年の発表では50以上の数学問題に適用し、多くで既知最良解を再発見、一部では既知記録を改善したとしている。 26

2026年の追加報告では、AlphaEvolveは量子物理にも応用され、GoogleのWillow量子プロセッサ上で分子simulation用回路のerrorを従来最適化baselineの約10分の1に抑える設計を見つけたとDeepMindが報告している。またTerence Taoは、こうしたツールが数学者のoptimization探索やcounterexample探索を加速していると評価している。 27

このため、マクシメンコの根本的な観察――

> **AIエージェントが、科学者が数カ月費やしていた探索・実装・ツール操作の一部を数日規模へ圧縮し得る**

――は、2026年時点では十分に現実的である。 28

反対側――研究加速とAGIは同義ではない

しかし、これらのシステムに共通する成功条件にも注目すべきである。

AlphaEvolveが特に強いのは、**解をコードとして表せ、正しさや品質を自動評価できる問題**である。DeepMind自身も、明確で自動的な評価指標を持つ数学・計算機科学がとくに適していると説明している。 26

AI co-scientistも、研究者が目標を設定し、専門家によるfeedbackや実験検証を組み合わせる。完全自律的に「世界のどの科学問題を研究すべきか」を選び、予算を管理し、実験を設計し、曖昧な失敗を診断し、論文を完成させる一般研究者ではない。 25

これはマクシメンコの元発言とも整合する。彼自身が正しいpromptと専門的guidanceの必要性を認めている。 13

従って2026年の証拠が示しているのは、

$$\text{Frontier LLM} + \text{agent scaffolding} + \text{tools} + \text{automated verification} + \text{human expertise}$$

という**複合系の能力が急速に伸びている**ことであって、

$$\text{AI alone} \approx \text{general human researcher}$$

が既に普遍的に成立したことではない。

量子計算側にも「成功例と古典側の追撃」がある

量子計算についても状況は二極的である。

肯定材料として、IBMの2023年Nature研究、Q4Bioの2026年ゲノムencoding、マクシメンコらを含むNISQ量子化学研究などから、**現在のnoisy quantum hardwareが単なる数量子ビットの教材実験を越えつつある**ことは確かである。[29]

一方で、IBMの「utility」実験がほどなく高度なclassical tensor-network計算に追いつかれたことは、quantum advantageの比較対象が静止していないことを示す。[30]

この意味でマクシメンコの主張のうち、

> 「AIが量子アルゴリズム設計・回路最適化・実行を容易にし、NISQ機の利用可能領域を拡大する」

は十分もっともらしい。

しかし、

> 「現在の量子計算能力がAGIを成立させる重要な証拠である」

という推論は一次文献からは導かれない。

量子計算はここでは**一般知能の原因というより、一般的なAIエージェントが使える高度な専門ツール**と見るのが技術的に適切である。

記事の論理的・方法論的な弱点

「研究速度」と「一般知能」が混同されている

記事の最大の論理的飛躍は、

> 研究タスクを非常に速く実行した
> ↓
> 人間の専門研究者に匹敵する
> ↓
> したがってAGI

という暗黙の推論である。

AGIで問われる「general」は、**一つの高度タスクをどれだけ速く解くかではなく、未知のタスク分布にどれだけ広く適応できるか**である。DeepMindの定義でもperformanceとgeneralityは別軸であり、狭い領域でsuperhumanでもそれだけではAGIではない。AlphaGoやAlphaFoldのようなsuperhuman narrow AIが典型例である。[14]

成功事例だけで失敗率が分からない

「半年→1週間」という一回の成功は印象的だが、AGI評価には**成功確率**が重要である。

100件の課題を渡して1件だけ劇的に成功したシステムと、95件を安定して成功するシステムでは意味が全く異なる。公開記事にはattempt数、失敗例、再試行数、selection procedureがない。成功例が事後選択されているならselection biasが生じる。

METRが明示的に50%・80%の成功確率でtask horizonを定義しているのは、まさにこの問題を避けるためである。[17]

「博士課程で半年」と「AIで一週間」が同じ作業量とは限らない

人間の半年には、問題設定、文献調査、失敗経路、方法の発明、ソフトウェア実装、結果解釈などが含まれる。

AIへの再現課題では、本人がすでに、

- どの問題が解けるか
- どの物理量を見るべきか
- 何が正解らしいか

を知っている可能性がある。

その状態での再実装は、元の研究発見とはinformation stateが異なる。したがって単純な

$$6\ \mathrm{months}/1\ \mathrm{week} \approx 26\times$$

という「研究能力の26倍化」のような読み方は成立しない。

データ汚染・既知解回収の検証がない

過去の博士研究や公開された量子／ゲノム成果を再構築する場合、LLMがその論文、abstract、コード、解説を学習または検索している可能性を除外しなければ、「独立発見」とはいえない。

理想的には、**モデル訓練cutoffより後に作成された未公開問題、外部検索遮断、blind evaluator**を使うべきである。Google AI co-scientistが未公開だった抗菌薬耐性研究を再発見させる実験を行ったのは、この点でマクシメンコの公開済み研究再構築より強い設計である。[25]

量子コンピューターの寄与を分離していない

本件ではLLM、agent scaffolding、Haiqu middleware、QPUが一体として使われる。

しかしAGI論証として重要なのは、

$$\Delta = P(\text{成功} \mid \mathrm{LLM}+\mathrm{agent}+\mathrm{QPU}) - P(\text{成功} \mid \mathrm{LLM}+\mathrm{agent}+\mathrm{classical})$$

である。

このablationがないため、量子コンピューターが科学的成果に本質的だったのか、単に課題の最終実行先だったのか分からない。

この区別なしに「AI＋量子＝AGI」という因果関係を暗示するのは方法論的に弱い。

「量子ビット数」と「古典コンピューター性能」の比較が粗い

「数百量子ビットは最大級古典コンピューターと同等」という説明は読者に、量子ビット数がCPU/GPU性能のような一般的尺度だと誤認させやすい。

実際には、127量子ビットの特定量子回路がbrute-force simulationを困難にしても、別の古典アルゴリズムがその構造を利用して高速に近似できる場合がある。IBMの2023年結果と2024年のclassical simulationの関係がその実例である。[31]

従って、量子計算については「何量子ビットか」より、

何を、どの精度で、どの総コストで、最良の古典手法より速く解いたか

を示すべきである。

ベンダー内部benchmarkと第三者検証を同列に置けない

Haiquの数値は研究上検討する価値があるが、同社の製品が同社のbenchmarkで得た数字である。Haiquの共同創業者がその性能を根拠にAGIを論じている以上、少なくとも、

1. workload公開
2. baselineコード公開
3. accuracy criterion公開
4. cloud/QPU billing条件公開
5. 第三者による再実行

が望まれる。

現状の公開情報では「$30,000→$25」を独立したscientific factとして扱うのは早い。[32]

見出しは本人の留保を薄めている

記事本文自体は比較的慎重で、量子業界のoverpromiseの歴史や独立検証の必要性にも触れている。問題は主に見出しとの落差である。[33]

本人の発言を忠実に圧縮するなら、

> **「専門家がAIエージェントと量子ツールを組み合わせると、研究速度は“弱いAGI”と呼びたくなる水準まで来た」**

程度が妥当であり、

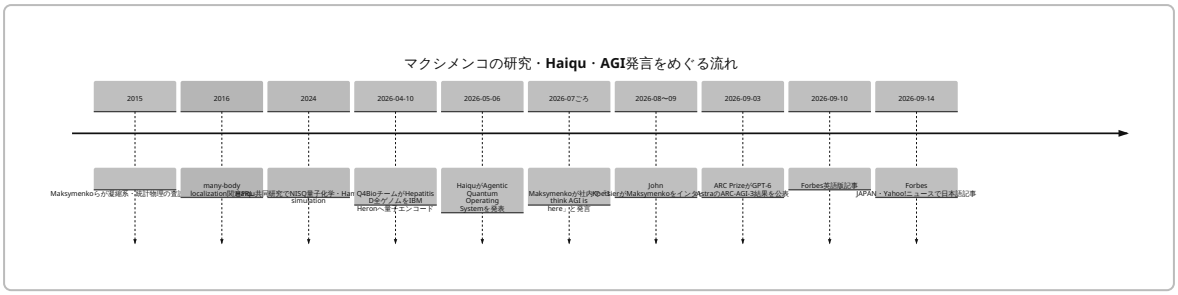
> **「汎用人工知能・AGIはもう来ている」**

だけでは、現在広く想定されるhuman-level AGIが実証されたかのように読める。

これは事実誤認というより、**定義の曖昧さを利用した強いフレーミング**である。

関係者・研究のタイムライン

以下はマクシメンコの査読研究、Haiqu関連技術、Q4Bio、AGI発言、報道を時系列に整理したものである。初期の研究業績はAPS、2026年の出来事は各研究機関・元インタビュー・記事から確認した。[34]



結論・信頼度評価と参照すべき一次資料

最終評価

本調査では、命題を三つに分けると混乱が解ける。

命題	本調査の判定	信頼度
AIエージェントはすでに一部の科学研究を数倍〜桁違いに高速化し得る	**支持される**。マクシメンコ以外にもAlphaEvolve、AI co-scientist等がある。ただし課題依存	**中〜高**
現在のNISQ量子計算＋AIは、一部の科学的workflowを実用化・簡素化し得る	**有望だがタスク依存**。量子utilityの境界は古典アルゴリズムの進歩で動く	**中**
それらの事例は「AGIが既に到来した」ことを科学的に証明する	**支持されない**。一般性、自律性、頑健性、失敗率、新規課題での再現性が未検証	**低**

したがって、記事タイトルの中心命題――**「AGIはもう来ている」**――に対する総合信頼度は**低**と評価する。

ただしこれは、マクシメンコの観察が「間違い」という意味ではない。むしろ彼は、**AGIという言葉の境界が曖昧になるほどAI-assisted researchの能力が急速に伸びている**という、非常に重要な現象を捉えている。元インタビューの「weak AGI」という表現なら、DeepMindの「Emerging AGI」のような段階的定義と対応させることも可能である。 35

問題は、「弱い／emerging AGIが存在する」という分類上の主張と、「通常想定されるhuman-level general intelligenceが完成した」という経験的主張を同じ見出しで扱うことである。

2026年9月時点では、後者を認定するには少なくとも、**未見かつ訓練データ汚染を排除した多数の職業・科学領域で、人間専門家水準以上の成功率を示し、長期課題を低介入で自律完遂し、その結果を第三者が再現する**必要がある。本件記事はそこまでの証拠を提供していない。DeepMindのAGI taxonomy、ARC-AGI、METRのいずれを参照しても、一つの印象的な研究再構築エピソードだけでこの条件を満たすとは言えない。 36

むしろ、本件から引き出すべき最も堅牢な結論は次である。

> **「AGIが証明された」のではなく、「専門家＋frontier AI agents＋専門ソフトウェア／計算資源」という複合研究システムが、人間の研究プロセスの一部を劇的に圧縮し始めている。その速度は、AGIという語をめぐる従来の直感を揺さぶるほどになっている。**

そして量子コンピューターの意義も、「AGIを生み出した」ことではなく、**AIエージェントが人間にとって扱いにくい専門計算基盤まで自然言語から操作・最適化できるようになれば、量子計算の実用化に必要な人的ボトルネックが大きく下がる可能性がある**点にある。ここは記事の中で最も研究価値のある論点であり、AGIという強い見出しよりも長期的には重要かもしれない。[37]

主要一次情報・信頼できる資料

資料	種別	本調査での位置づけ
Forbes JAPAN「『汎用人工知能・AGIはもう来ている』量子物理学者が告げた理由」、2026-09-14	調査対象記事	Yahoo掲載記事の本文確認に使用。[1]
John Koetsier, “Quantum computing + AI + agents = AGI?”	元インタビュー／transcript	マクシメンコ発言の最重要一次ソース。「weak AGI」等の限定条件を確認。[38]
Morris et al., “Levels of AGI for Operationalizing Progress on the Path to AGI”	Google DeepMind研究者によるAGI定義論文	Generality × Performance × Autonomyを分けて評価する枠組み。[14]
OpenAI Charter	AGI定義	「高度に自律的」「ほとんどの経済的価値のある仕事で人間を上回る」という厳格な代表定義。[15]
ARC Prize, GPT-6 Astra on ARC-AGI-3, 2026-09-03	最新AGI関連benchmark一次報告	99.9%という強い結果にもかかわらず、ARC側自身がAGI証明ではないと明記。[16]
METR, Task-Completion Time Horizons of Frontier AI Models	自律エージェント評価	長期タスク、自律性、成功率を測る代表的手法とその限界。[17]
Wellcome Genome Campus / Oxford, Hepatitis D quantum genome project	Q4Bio一次情報	記事中のゲノム例の実在と、156-qubit Heron上で117 qubitsを用いたencodingを確認。[9]
Kim et al., Nature, “Evidence for the utility of quantum computing before fault tolerance”	査読一次論文	NISQ-era quantum utilityを支持する代表例。[19]
Begušić, Gray & Chan, Science Advances, “Fast and converged classical simulations…”	査読一次論文	IBM utility実験への古典計算側からの重要なcounter-evidence。[20]
Haiqu公式 Agentic Quantum OS	企業一次情報	AI agent＋量子middleware構成、最大1000×等の製品主張を確認。ただし独立評価ではない。[39]
Maksymenko関連APS論文	査読一次論文	「量子物理学者」という経歴の実質的裏付け。[40]
Google Research, AI co-scientist	研究機関一次報告	科学仮説生成・未公開成果の再発見・実験検証という、本件に近い独立事例。[25]

資料	種別	本調査での位置づけ
Google DeepMind, AlphaEvolve	研究機関一次報告	数学・アルゴリズム・量子回路等でのAI-driven discovery／optimizationを示す支持材料。 41

総合信頼度：「AGIはすでに到来した」という記事の中心命題＝低。
「AIエージェントが科学研究を劇的に高速化する局面に入った」＝中〜高。
「現在の量子計算がAIエージェントとの組み合わせで実用範囲を拡大している」＝中。

この三つを分離して読むことが、記事を最も正確に評価する方法である。

---

1 6 24 「汎用人工知能・AGIはもう来ている」量子物理学者が告げた理由 | Forbes JAPAN 公式サイト（フォーブス ジャパン）
https://forbesjapan.com/articles/detail/104539

2 25 Accelerating scientific breakthroughs with an AI co-scientist
https://research.google/blog/accelerating-scientific-breakthroughs-with-an-ai-co-scientist/?utm_source=chatgpt.com

3 4 7 13 18 21 35 38 Quantum computing + AI + agents = AGI? – john koetsier
https://johnkoetsier.com/quantum-computing-ai-agents-agi/

5 11 32 37 39 Haiqu: Quantum Computing Software for R&D Teams
https://haiqu.ai/?utm_source=chatgpt.com

8 33 「汎用人工知能・AGIはもう来ている」量子物理学者が告げた理由 | Forbes JAPAN 公式サイト（フォーブス ジャパン）
https://forbesjapan.com/articles/detail/104539/page2

9 Genome loaded onto a quantum computer in world first
https://www.wellcomegenomecampus.com/news_item/genome-loaded-onto-a-quantum-computer-in-world-first/?utm_source=chatgpt.com

10 Oxford researchers contribute to world-first in quantum ...
https://www.ox.ac.uk/news/2026-05-22-oxford-researchers-contribute-to-world-first-in-quantum-computing-and-genomics?utm_source=chatgpt.com

12 Mykola Maksymenko - UCU Business School
https://lvbs.com.ua/en/teacher/mykola-maksymenko/?utm_source=chatgpt.com

14 36 Position: Levels of AGI for Operationalizing Progress on the Path to AGI
https://arxiv.org/html/2311.02462v3

15 OpenAI Charter | OpenAI
https://openai.com/charter/

16 OpenAI's GPT-6 Astra on ARC-AGI-3 | ARC Prize
https://arcprize.org/blog/astra

17 Task-Completion Time Horizons of Frontier AI Models - METR
https://metr.org/time-horizons/

19 29 31 Evidence for the utility of quantum computing before fault ...
https://www.nature.com/articles/s41586-023-06096-3?utm_source=chatgpt.com

20 30 Fast and converged classical simulations of evidence for the utility of quantum computing before fault tolerance
https://arxiv.org/abs/2308.05077?utm_source=chatgpt.com

22 Efficient Hamiltonian Simulation: A Utility Scale Perspective for Covalent Inhibitor Reactivity Prediction
https://arxiv.org/abs/2412.15804?utm_source=chatgpt.com

23 Haiqu Launches Agentic Quantum Operating System For Enterprise ...
https://thequantuminsider.com/2026/05/06/haiqu-launches-agentic-quantum-operating-system-for-enterprise-applications-rd/?utm_source=chatgpt.com

26 28 41 AlphaEvolve: A Gemini-powered coding agent for designing advanced algorithms — Google DeepMind
https://deepmind.google/blog/alphaevolve-a-gemini-powered-coding-agent-for-designing-advanced-algorithms/?utm_source=chatgpt.com

27 AlphaEvolve: Gemini-powered coding agent scaling impact across fields — Google DeepMind
https://deepmind.google/blog/alphaevolve-impact/?utm_source=chatgpt.com

34 40 Classical dipoles on the kagome lattice | Phys. Rev. B - APS Journals
https://link.aps.org/doi/10.1103/PhysRevB.91.184407?utm_source=chatgpt.com