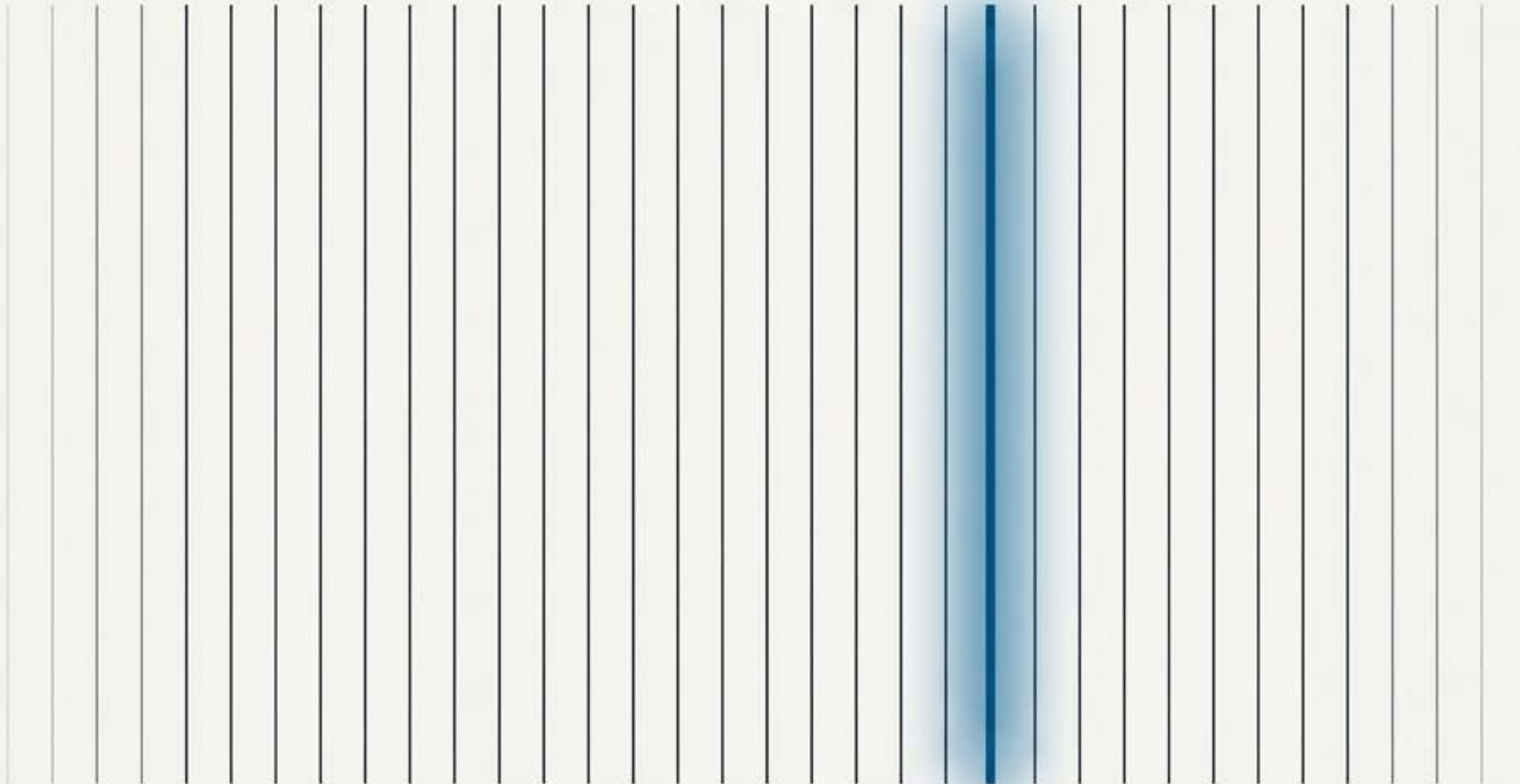


Gemini 3 ProはAGIレベル4に到達したのか？

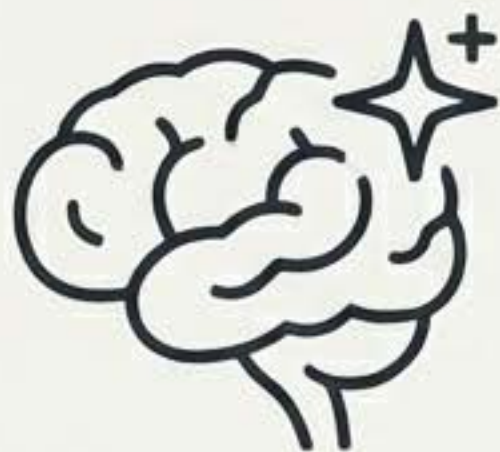
「発明するAI」の現在地を解き明かす、詳細分析



評価の物差し：AGI進化の5段階モデル



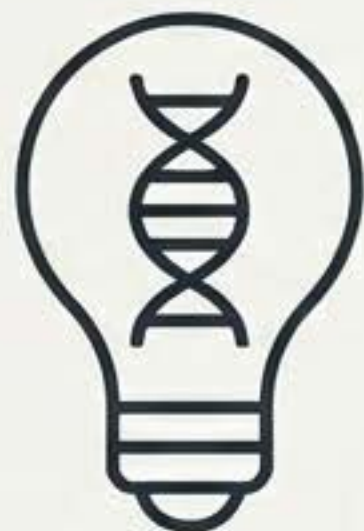
レベル4の核心：「発明」とは何を意味するのか



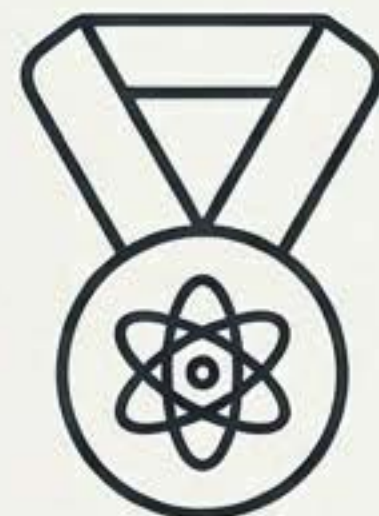
既存知識の組み合わせを超えた全く新しい概念や理論の創造



未知の問題に対する独創的な解決策の自律的生成



人間の科学者が思いつかなかった革新的な仮説の提案と検証



科学的発見や技術的イノベーションへの実質的貢献

レベル2は明確に達成済み：博士号レベルの推論能力

Evidence Scorecard



**レベル2: 推論者
(Reasoners)**

「Gemini 3 Proは、科学や数学といった広範なトピックにわたる複雑な問題を、極めて高い信頼性をもって解決できることを実証しています。」

競合を凌駕する、圧倒的なベンチマークスコア

GPQA Diamond
(博士課程レベルの科学的推論)



Humanity's Last Exam
(現役専門家レベル)



FrontierMath
(研究レベル数学)



CritPt
(物理学研究ベンチマーク)



レベル3も実質的に達成：自律的エージェントとしての機能

Evidence Scorecard



**レベル3:
エージェント
(Agents)**

「従来の『単発の質問応答』を超え、複数手順を踏むタスクを自律的に計画・実行する、長期的かつ自律的な問題解決能力を実装しています。」

長期・複雑タスクを可能にする3つのコア技術



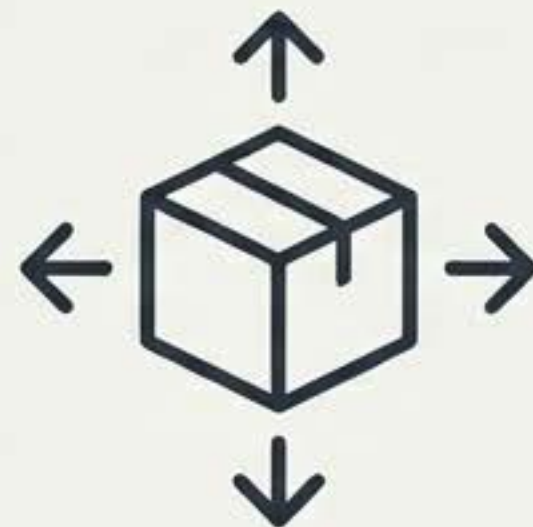
自律的なマルチステップタスク実行

Google Workspace (Gmail, カレンダー) や外部ツールと連携し、タスクを自律的に計画・実行。



思考シグネチャ (Thought Signature)

複雑で長期的なタスクでの文脈維持能力を強化し、内部的な思考プロセスを維持。



巨大なコンテキストウィンドウ

最大1,048,576トークンを処理可能。膨大な情報を保持しながら推論を継続。

UD Shin Go NT

レベル4への接近：科学的発見能力の萌芽が見える

「Gemini 3 Proは、レベル4の完全達成には至らないものの、その核心である『発明』につながる重要な能力を複数示唆しています。」

「未知への適応力」と「天才的パートナー」の片鱗

31.1%

ARC-AGI-2（抽象的推論）

(Deep Think: 45.1%)

「訓練データにない全く新しいパターンの発見」能力を測定し、発明の核心を直接評価。

“天才的な研究パートナー
(genius research partner)”
という評価。

探索的・創造的アプローチ

複数の分析が、Gemini 3 Proの「大胆で探索的 (bold and exploratory)」な特性を指摘。

自律的研究システムとの 統合可能性

AI Researcherのような「完全自律的な研究システム」の中核エンジンとして使用可能な水準。

しかし、真の発明者と呼ぶにはまだ壁がある

残された課題①：現実世界での実績と、人間への依存



実際の発明実績の不足

現時点で、Gemini 3 Pro自体が完全自律的に新規の科学的発見や特許可能な発明を生み出した実証例は公開されていない。



人間の監督と誘導の必要性

完全に自律的に「何を研究すべきか」を決定し、独自に研究課題を設定して解決するレベルには至っていない。

残された課題②：創造性の本質と、信頼性のジレンマ



創造性の限界

現在のAIは「好奇心に導かれた創造的プロセス」を欠いている可能性が指摘されている。「真の創造性」と「既存パターンの高度な組み合わせ」の境界は依然として曖昧。



信頼性とハルシネーション

Omniscience Indexで**88%のハルシネーション率**が報告されている。
「非常に自信を持って間違える」という問題が残っており、自律的発明に必要な信頼性には課題がある。

総合評価：Gemini 3 Proの現在地はレベル3.8



- ✓ レベル2 (推論者): 完全達成
- ✓ レベル3 (エージェント): 実質的達成
- 🟡 レベル4 (イノベーター): 部分的達成

総合評価: **レベル 3.8**

最強の「発明支援ツール」であり、自律的な「発明者」ではない

Gemini 3 Proは「発明の『創出』段階において最も優れた性能」を持ち、「0から1を生み出す創造性、仮説生成能力、深い推論能力」において高い能力を示している。

「これは『発明を支援する最強のツール』であることを意味し、『完全自律的に発明を行うAI』への到達を意味するものではありません。」

次なる飛躍へ：レベル4完全到達は2025-2030年が視野に

業界専門家の見解では、レベル4の実現時期は「2025～2030年頃」と予測されている。Gemini 3 Proはまさにその移行期の最先端に位置する存在と言える。

今後の展望: AI Researcherのようなフレームワークとの統合、そして実世界での自律的発明実績の蓄積が、レベル4への完全到達を現実化する鍵となる。

