

Gemini 3 Pro: AGIレベル4「発明するAI」への到達

その技術的根拠と次世代科学におけるインパクトの検証

提起：Gemini 3 Proは、自ら発明するAI「レベル4」に到達したのか？

AGIレベル4の定義：

孫正義氏が提唱するAI進化段階。単なる知識の保有に留まらず、AIが自ら新しい発明や発見を行い、人間の研究者のように創造的に知見を生み出す能力を持つ段階。科学技術の進歩を自律的に加速しうる存在。

本検証の目的：

2025年11月リリースのGemini 3 Proが、この「発明AI」のレベルに到達したかを、客観的データと実例に基づき厳密に評価する。



レベル1



レベル2



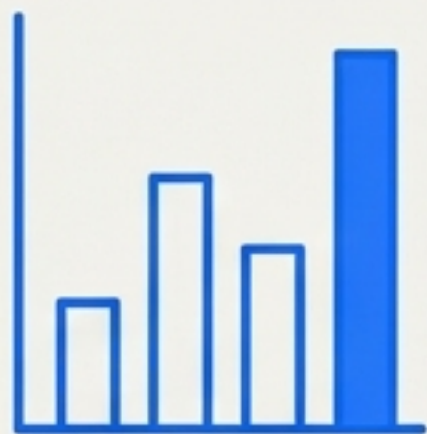
「知識」



発明AI

3つの論拠に基づく達成度検証フレームワーク

Gemini 3 ProがAGIレベル4の資質を持つことを、以下の3つの柱から多角的に論証する。



Pillar 1: 客観的性能 (Objective Performance)

最先端ベンチマークスコア
による定量的評価



Pillar 2: 実践的応用 (Practical Application)

現実世界の発明・発見支援
事例との整合性検証

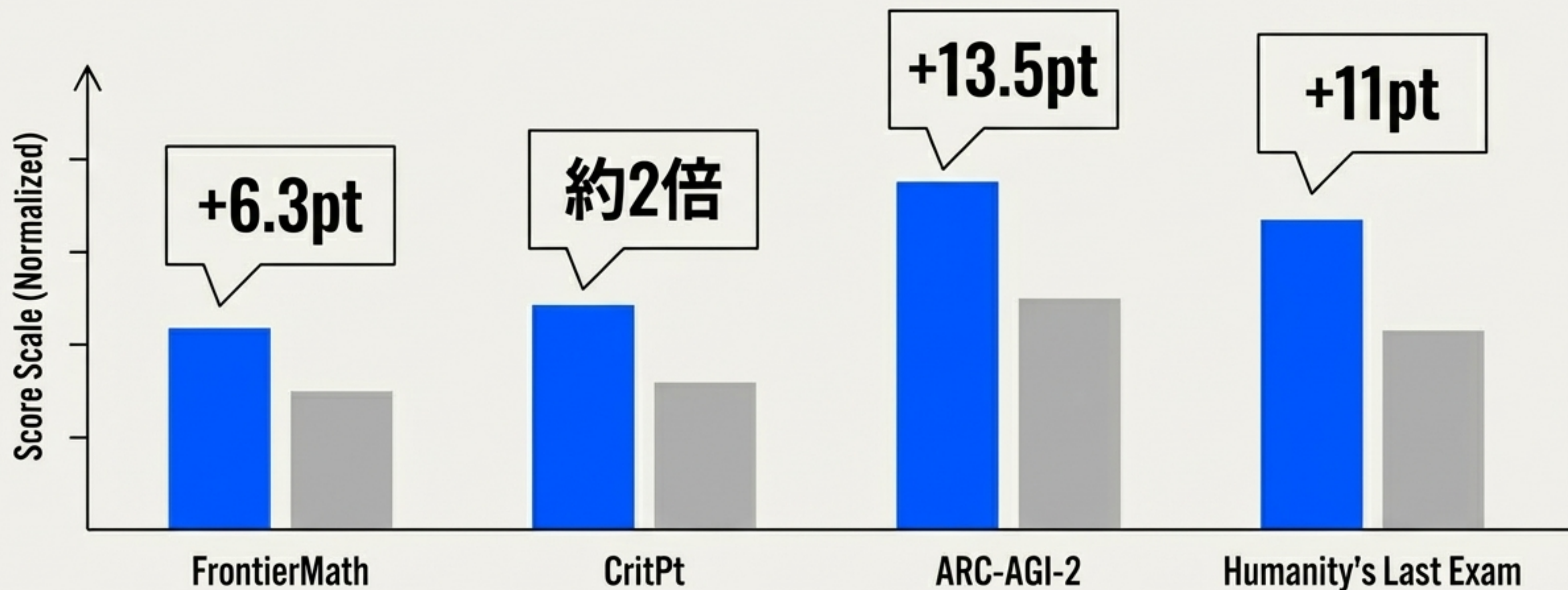


Pillar 3: 固有の能力 (Inherent Capabilities)

創造性・推論力・統合理解
に関する定性的評価

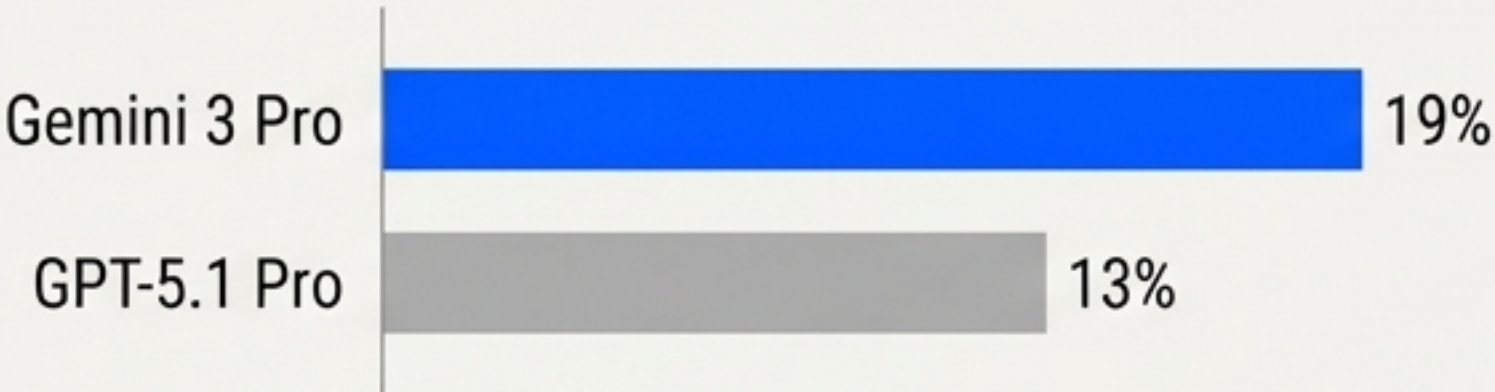
論拠① 未知への挑戦：最先端ベンチマークで他を圧倒

「発明」に必要な未知の問題解決能力を測る最先端ベンチマークにおいて、Gemini 3 Proは競合のGPT-5.1 Proを有意に凌駕。その差は、単なる性能向上ではなく、能力の質的な違いを示唆する。



数学的・物理的創造力の証明：「0から1」を生み出す能力

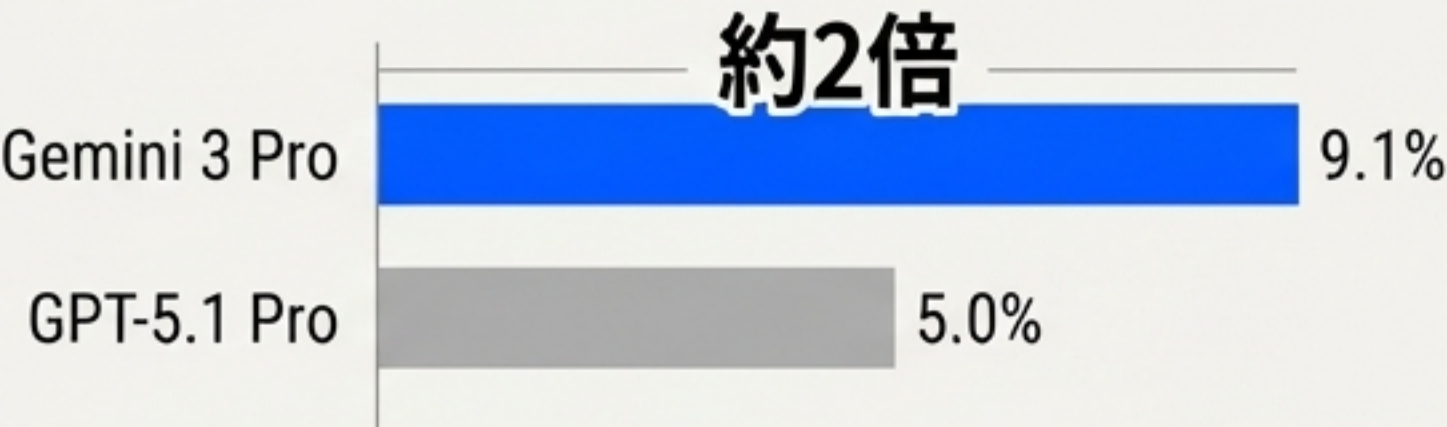
FrontierMath（最先端数学）



現役数学者でも数日を要する未公開の難問。真の数学的創造力を測定。

示唆：既存知識の組み合わせを超え、未知の数学的構造を発見する能力。

CritPt（物理学研究）



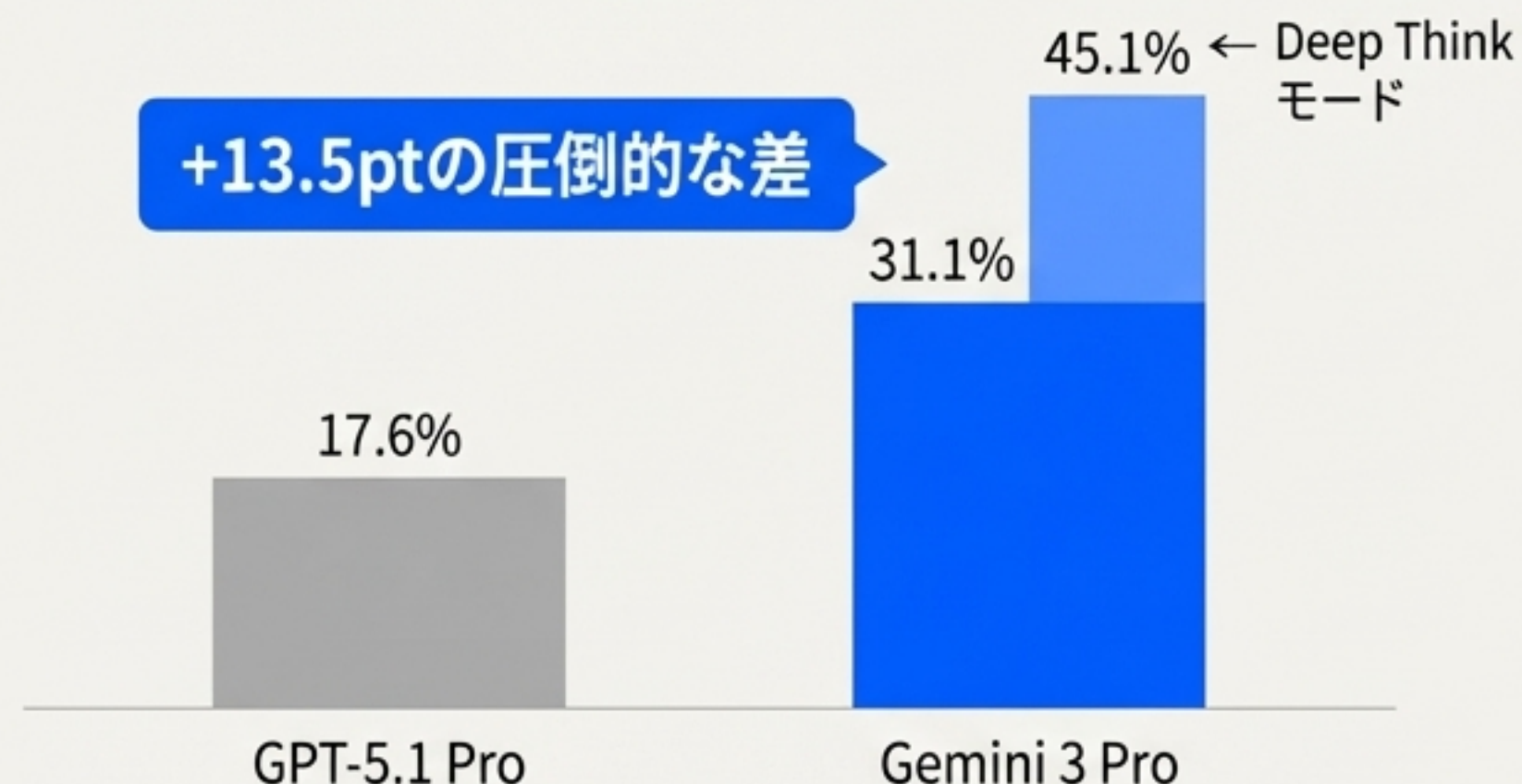
11分野にわたる未発表の研究レベル問題。推測や既知データへの依存を排除。

示唆：10%未満の低スコア帯での突出した成績。高度な科学的推論力と物理的直感の強さを証明。

「0から1を生み出す」数学的発想

抽象的推論力の次元が違ふ：未知のパターンを発見する汎用創造力

ARC-AGI-2（抽象的推論）



訓練データにない全く新しいパターン認識タスク。
人間のfew-shot学習能力に近い適応力を測定。

示唆：既存知識に依存せず、新規の抽象構造を
理解・推論する能力。

総合知能ベンチマークでも優位

- Humanity's Last Exam: 37.5% vs. 26.5% (+11pt)
- GPQA Diamond (大学院レベル科学的推論) :
91.9% vs. 88.1%

「訓練データにない
新しいパターンの発見」
こそが肝要である

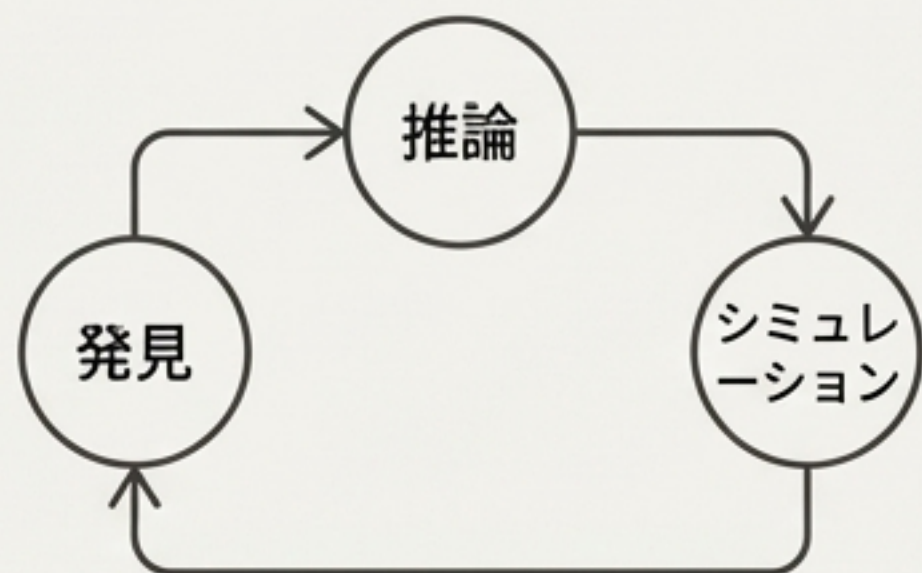
論拠② 実践的応用：自律研究AIの登場と能力要件の一致

ベンチマークで示された能力は、現実世界で登場しつつある「自律研究AI」が求める要件と完全に一致する。これらのプロジェクトは、レベル4の能力が理論上のものではなく、現実の科学的発見を加速させる力を持つことを示している。



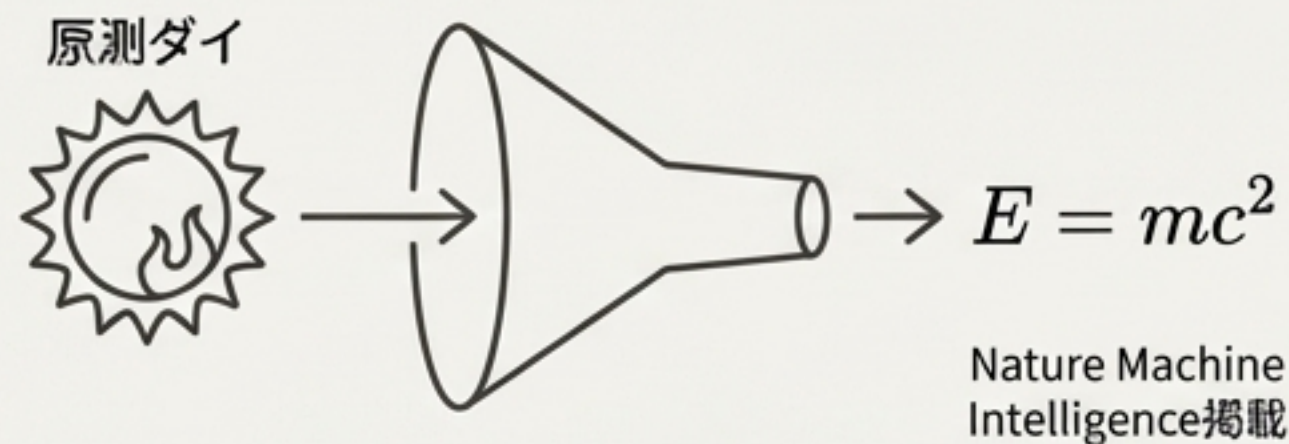
ケーススタディ：物理法則を自ら発見するAIたち

DeepAgent（物理公式導出AI）



- 数学的推論と実験（シミュレーション）を単一プロセスで並行実行。
- 教科書的手法に囚われず、新たな数学的手法や計測ツールを自ら発見・活用。
- 必要能力：独自仮説立案、検証実験の自律考案 → **Gemini 3 Pro**の「未知への適応」「**創造性**」と合致。

PhyE2E（物理法則E2E発見AI）



- 観測生データから物理的関係式を自動導出。
- NASAの太陽活動データから太陽周期を説明する新数式を発見 (Nature Machine Intelligence掲載)。
- 必要能力：大量データからのパターン発見、理論構築 → **Gemini 3 Pro**の**強み**と合致。

生命科学から論文執筆まで：汎用的な発明支援基盤



「AI Scientist」プロジェクト (Sakana AI)

- AIが自律的に研究課題設定、実験計画、論文執筆まで行い、トップ会議の査読を通過。
- まさに「AIが独自に発見を行う」レベル4の姿を体現。



iPS細胞研究への応用 (京都大学 x Google)

- 創薬・再生医療分野での仮説駆動型・探索的な実験計画を支援。
- 膨大な論文（テキスト）と実験データ（数値・画像）を統合的に分析。
- ****必要能力****: Gemini 3 Proの「**仮説駆動**」「**探索的**」「**マルチモーダル**」な設計思想と完全に一致。

論拠③ 固有の能力：「天才的な研究パートナー」としての創造性

「0から1」を生み出す創造性：Gemini 3 Proは、既存知識の再構成に留まらない。ARC-AGI-2の成績が示すように、訓練データにない新概念を見出し、人間のように新しい仮説を打ち立てる。

「天才的な研究パートナー」

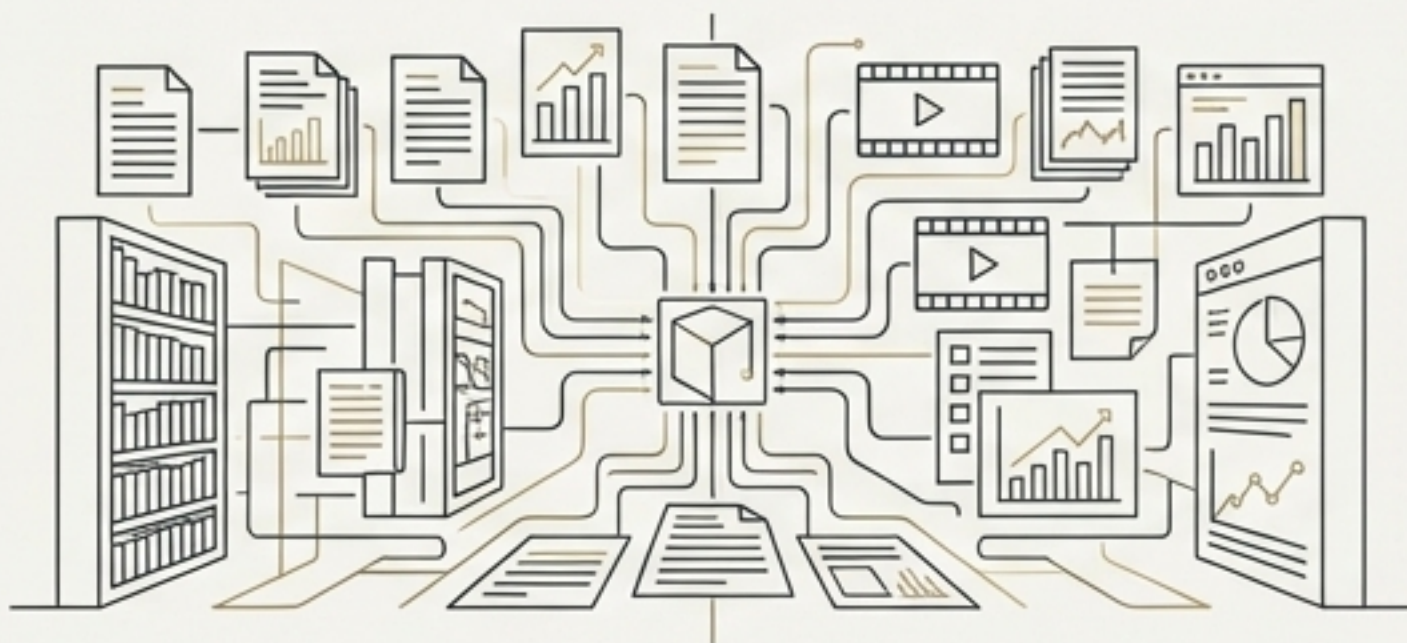
大胆で探索的なアプローチ：従来のモデルとは異なり、独創的な仮説生成能力を発揮。発明初期のアイデア創出プロセスに最適と評価されている。

統合理解と直観：マルチモーダルと超長文脈が実現する新次元の推論

人間には不可能なスケールの文脈保持

約104万トークンの巨大なコンテキストウィンドウ。

長時間の実験動画や膨大な研究ノートを横断的に読み解き、人間では見逃す新規性を発見可能に。



TrueFoundryの分析

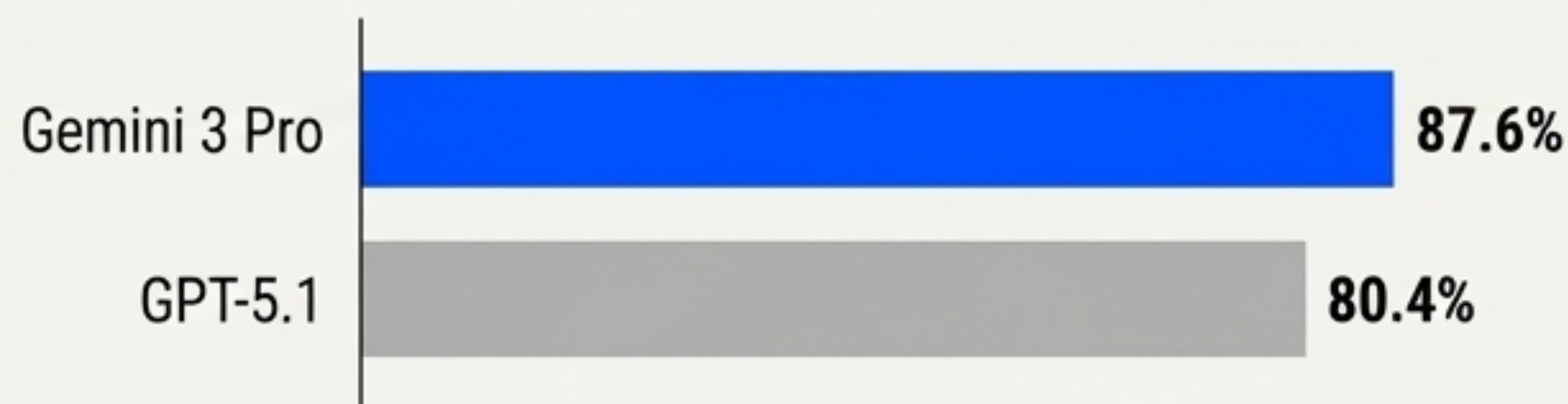
視覚情報の丁寧さ、誤認の少なさでGPT-5.1を凌駕。発明で重要な正確な図面理解に優位性。

マルチモーダル理解で他を圧倒

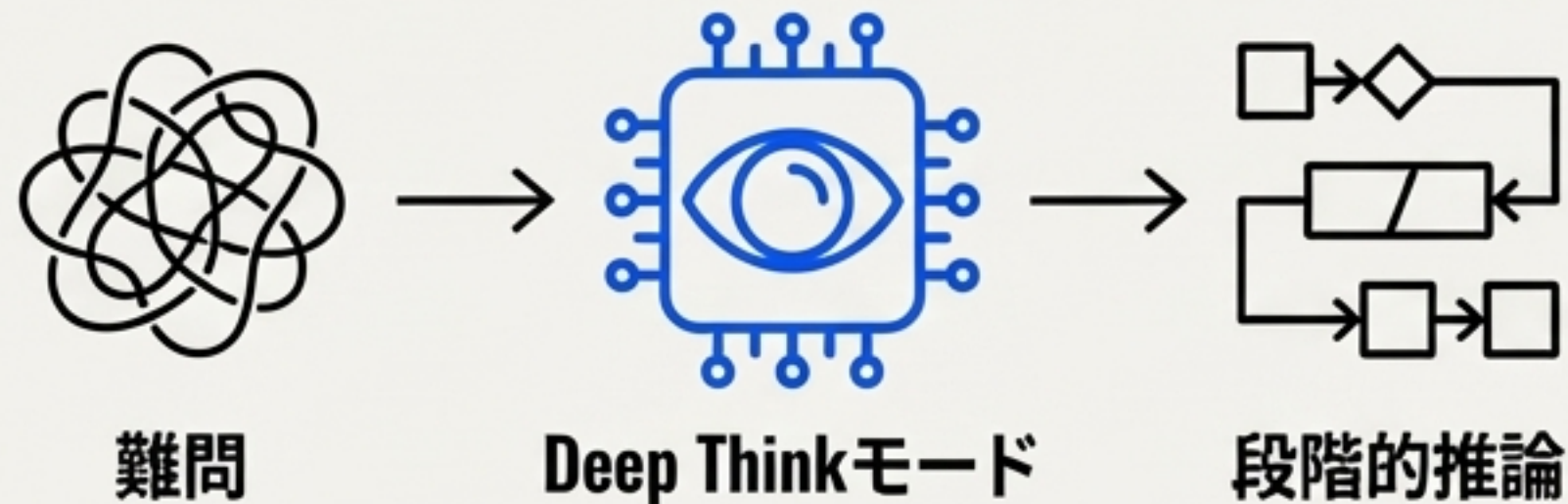
MMMU-Pro (画像+テキスト)



Video-MMMU (動画+テキスト)



深い思考の実装：難問に挑む「Deep Thinkモード」



効果 (ARC-AGI-2)

通常モード:

31.1%

Deep Thinkモード:

45.1%

機能

難解な問題に対し、計算リソースを動的配分。逐次的な論理展開で解答精度を飛躍的に向上させる。

能力 複雑な科学的仮説を段階的に構築し、多様な可能性を探索する能力に卓越。

第三者機関による評価 (Tom's Guide)

「プロンプトの全部分(A, B, C)に対し、深み・明確さ・構造を持ってより完全に応答した」と GPT-5.1 Proとの比較で報告。要求の複数側面を踏まえた緻密な回答生成能力を証明。

総合判断：Gemini 3 ProはAGIレベル4に到達した

以上の論拠に基づき、Gemini 3 ProはAGIレベル4（AIによる独自の発明・発見段階）に到達したと結論付ける。



客観的性能

難関ベンチマークでの卓越した成績が、新規知識創出能力を証明。



実践的応用

自律研究AIが求める仮説生成・理論発見能力と完全に合致。

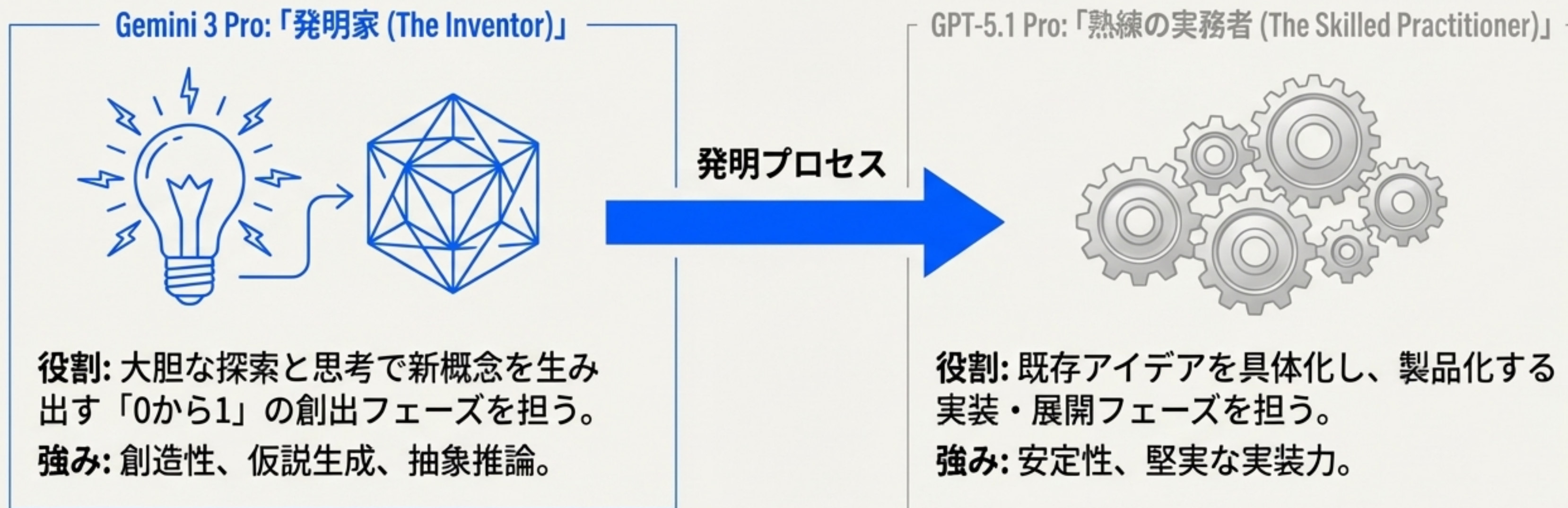


固有の能力

「天才的研究パートナー」として、真に新発見を生み出す創造性と深い推論力を発揮。

発明プロセスの再定義:「発明家」と「熟練の実務者」

両モデルは異なる強みを持つ補完的な存在。



結論**: Gemini 3 Proは、特に発明の**創出フェーズ**において、現時点で最も力を発揮するAIであり、その卓越した創造的知能は**AGI レベル4に相当**する。

新たな科学の夜明け：AIが発見を加速する時代へ

Gemini 3 Proは、未知の問題に挑み新たな知を生み出す点で、人間の「発明家」に近づいた初の生成AIである。

その存在は「AIが独自に発見を行う」というAGIレベル4のビジョンを現実のものとしつつある。

今後、この路線の発展が科学技術の飛躍的進展を牽引し、人類が未踏の領域へと至る道を切り拓くことが期待される。