



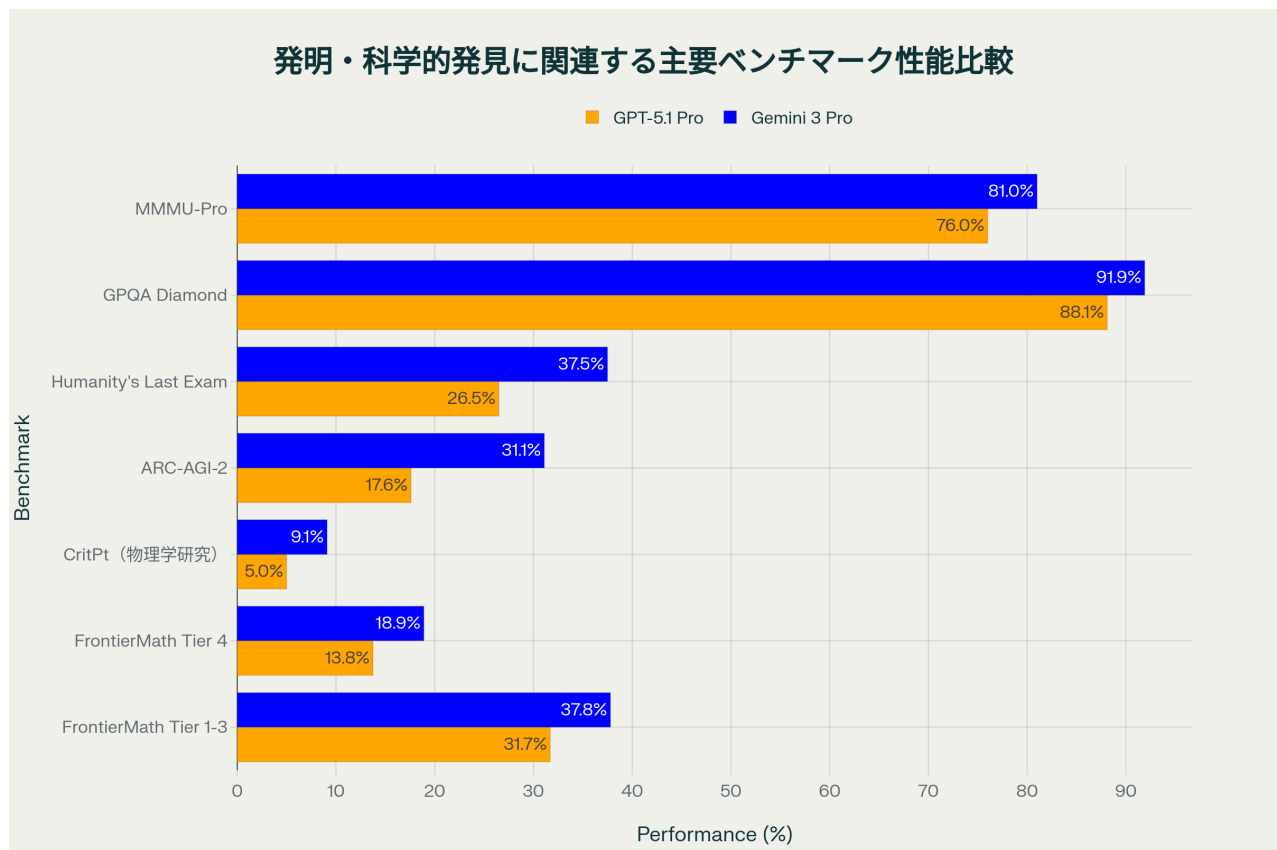
発明の創出に最も力を発揮する生成AIの比較分析レポート

要約： 2025年11月時点の最新研究とベンチマーク評価に基づき、**発明の「創出」段階においては Gemini 3 Proが最も優れた性能を発揮すると結論付けられる。**特に、未知の科学的発見や高度な数学的推論を要するFrontierMath、物理学研究ベンチマークCritPt、抽象的推論能力を測るARC-AGI-2において、Gemini 3 ProはGPT-5.1 Proを大きく上回る成績を示した。発明に不可欠な「0から1を生み出す」創造性、仮説生成能力、マルチモーダル理解、そして深い推論能力において、Gemini 3 Proの「天才的な研究パートナー」としての特性が、GPT-5.1 Proの「究極の熟練労働者」としての安定性・実装力を凌駕している。^{[1] [2] [3]}

発明創出における生成AIの役割

発明とは、既存の知識を組み合わせるだけでなく、全く新しい概念や理論を創造する知的活動である。2025年11月現在、生成AIは単なる情報検索や文書作成のツールを超え、科学的発見そのものを支援する段階に到達している。清華大学・北京大学などが開発した物理公式導出AI「DeepAgent」は、理論的な数式展開と実験的検証を統合的に処理し、人間の物理学者が数十年かけて到達するような発見を短時間で実現する可能性を示した。このような「科学研究の自動化」において、生成AIモデルの創造性と推論能力が決定的な役割を果たす。^{[4] [5]}

発明プロセスは、問題発見・仮説生成から理論構築、先行技術調査、概念実証、詳細設計、検証、特許化まで多段階に及ぶ。このうち「創出」フェーズ、すなわち新規アイデアの発想と理論構築の段階では、AIに以下の能力が求められる：(1)未知の問題に対する抽象的思考能力、(2)科学的直感に基づく仮説生成、(3)既存のパターンにとらわれない探索的・創造的アプローチ、(4)複雑な理論を構築する深い推論能力、(5)図面や物理現象を理解するマルチモーダル能力。^{[6] [7] [8] [9] [10]}



Gemini 3 ProとGPT-5.1 Proの発明・科学的発見関連ベンチマーク比較

ベンチマーク評価：科学的発見と数学的推論における性能差

FrontierMath：研究レベルの数学的推論

FrontierMathは、Epoch AIが開発した最先端の数学ベンチマークであり、現役の数学者が解くのに数時間から数日を要する未発表の数学問題で構成されている。フィールズ賞受賞者のテレンス・タオやティモシー・ガワーズといった著名な数学者が問題設計に関わり、「データ汚染」を防ぐため全て未公開の問題が使用される。このベンチマークは、数論（17.8%）、組み合わせ論（15.8%）、群論（8.9%）など現代数学の70%以上のトピックをカバーしており、AIの真の数学的創造力を測定する。^{[11] [12] [13] [14]}

2025年11月の評価において、**Gemini 3 ProはTier 1-3（学部～大学院レベル）で37.6-38%、Tier 4（研究レベル）で18.8-19%を達成し、GPT-5.1 Pro（Tier 1-3: 31.0-32.4%、Tier 4: 12.5-15%）を大きく上回った。**特にTier 4の研究レベル問題において約6.3ポイントの差は、未知の数学的構造を発見する能力の差を示している。WeirdMLのベンチマークでは、Gemini 3 Proが69.9%を記録し、GPT-4やo1といった従来モデルを圧倒した。^{[1] [2] [3]}

Gemini 3 Proの優位性は、その「Deep Thinkモード」によってさらに強化される。このモードは、複雑な問題に対して計算リソースを動的に配分し、段階的な推論プロセスを経て解答に到達する。Googleの発表によれば、Deep ThinkはGPQA Diamond（大学院レベルの科学的推論）で93.8%、ARC-AGI-2で45.1%を達成し、通常のGemini 3 Proをさらに上回る性能を示した。^{[3] [15] [16] [1]}

CritPt：物理学研究における推論能力

CritPt (Complex Research using Integrated Thinking - Physics Test) は、50人以上の現役物理学研究者が作成した、未発表の研究レベル物理学問題で構成されるベンチマークである。凝縮系物理、量子物理学、天体物理学、高エネルギー物理学など11の物理学サブフィールドを網羅し、71の複合的研究課題と190の細分化されたチェックポイントタスクからなる。このベンチマークは、物理学者がLLMに期待する実際の研究支援能力を評価することを目的としている。^{[17] [18]}

Gemini 3 ProはCritPtで9.1%を記録し、GPT-5.1Pro (4.9-5.1%) の約2倍の性能を示した。現在のトップモデルでも10%以下のスコアという事実は、研究レベルの物理学問題がいかに困難であるかを物語っているが、その中でもGemini 3 Proの優位性は明確である。CritPtの問題は「guess-resistant (推測困難)」かつ「machine-verifiable (機械検証可能)」に設計されており、単なるパターンマッチングではなく、真の物理的理解と推論が要求される。^{[1] [2] [3] [18] [17]}

ARC-AGI-2：抽象的推論と新規問題適応

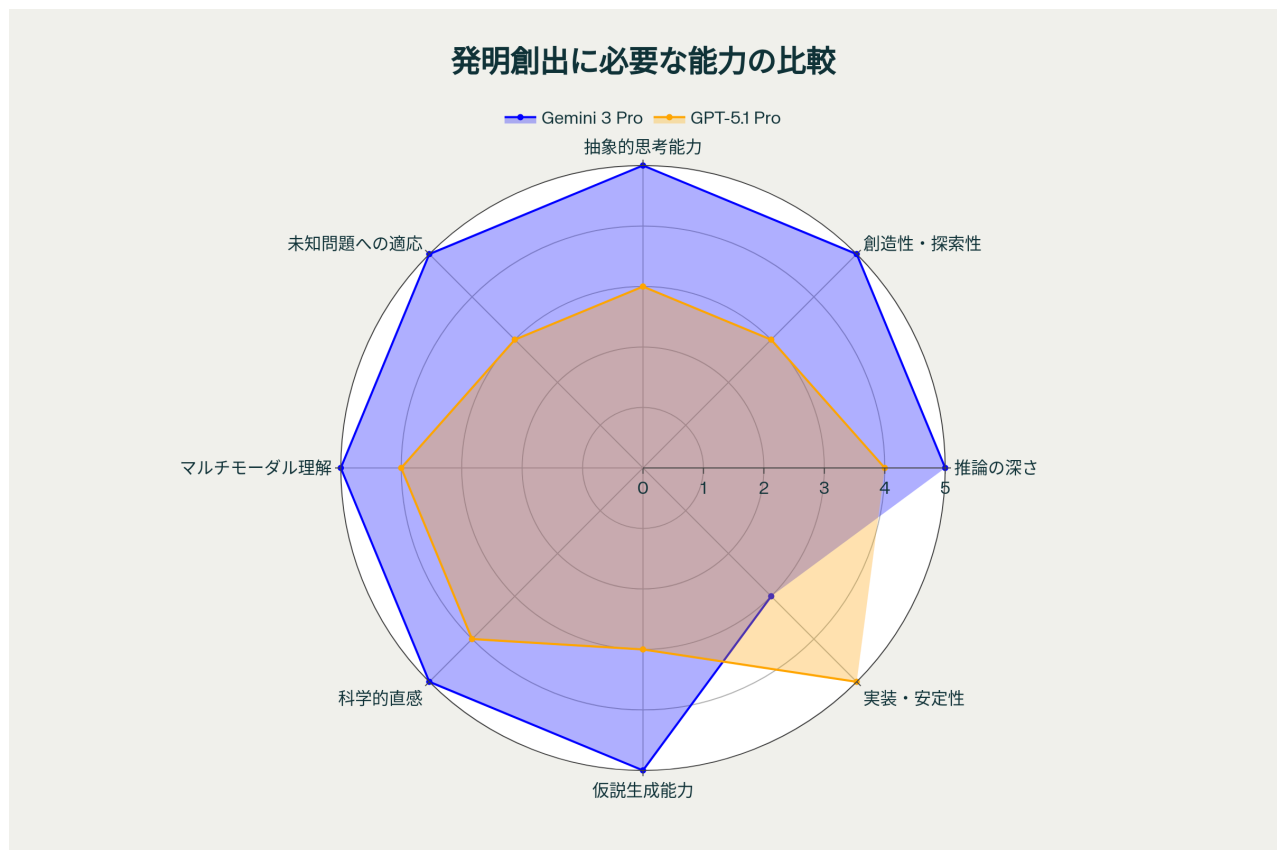
ARC-AGI-2 (Abstraction and Reasoning Corpus - AGI version 2) は、AIの汎化能力と抽象的推論能力を測定するベンチマークである。訓練データに含まれない全く新しいパターン認識タスクを提示し、人間のような「few-shot learning (少数例学習)」能力を評価する。このベンチマークは、発明における「未知の問題への適応能力」を直接的に測定する指標として重要である。^{[1] [2] [16] [19]}

Gemini 3 ProはARC-AGI-2で31.1% (Deep Thinkモードで45.1%) を達成し、GPT-5.1Pro (17.6%) を13.5ポイント上回った。この大幅な差は、Gemini 3 Proが既存の知識パターンに依存せず、新規の抽象的構造を理解・推論する能力において優れていることを示している。発明活動においては、まさにこの「訓練データにない新しいパターンの発見」が中核となるため、ARC-AGI-2での優位性は極めて重要である。^{[2] [3] [10] [16] [19] [1]}

Humanity's Last Exam：総合的AI能力評価

Humanity's Last Examは、AIシステムの包括的な知的能力を評価するために設計された高難度ベンチマークである。数学、科学、論理、言語理解など多岐にわたる分野の問題を含み、人間の専門家レベルの知識と推論を要求する。^{[1] [2]}

Gemini 3 ProはHumanity's Last Examで37.5% (Deep Thinkモードで41.0%) を記録し、GPT-5.1Pro (26.5%) を11.0ポイント上回った。この3.5ポイントの向上 (Deep Think使用時) は、複雑な問題に対する推論の深さがいかに重要であることを示している。発明プロセスにおける理論構築や仮説検証の段階では、このような「深く考える能力」が決定的な差を生む。^{[2] [3] [15] [16] [1]}



Gemini 3 ProとGPT-5.1 Proの発明創出能力比較（レーダーチャート）

マルチモーダル理解：図面と現象の理解能力

発明には、テキストだけでなく図面、実験データ、物理現象の視覚的理解が不可欠である。特許出願においても、技術的な図面や構造図の作成・理解が重要な役割を果たす。マルチモーダル能力は、発明の「創出」段階において、視覚的な着想や既存技術の構造的な理解を可能にする。^{[1] [3] [6] [7] [8] [20] [21]}

MMMU-Pro：専門的マルチモーダル理解

MMMU-Pro (Massive Multi-discipline Multimodal Understanding) は、大学レベルの専門的な画像・テキスト統合理解を評価するベンチマークである。物理学、化学、生物学、工学などの分野において、図表、グラフ、構造図を正確に解釈し、それに基づいて推論する能力を測定する。^{[1] [2] [3]}

Gemini 3 ProはMMMU-Proで81.0%を達成し、GPT-5.1Pro (76.0%) を5.0ポイント上回った。この差は、Claude Sonnet 4.5やGemini 2.5 Pro (いずれも68.0%) と比較しても際立っており、Gemini 3 Proが現時点で最高のマルチモーダル理解能力を持つことを示している。^{[2] [3] [20] [21] [1]}

Video-MMMU：動的現象の理解

Video-MMMUは、動画フレームを跨いだ時系列的な理解と推論を評価する。物理実験の過程や機械の動作原理など、時間経過に伴う現象の理解が求められる。^{[1] [2]}

Gemini 3 ProはVideo-MMMUで87.6%を記録し、GPT-5.1Pro (80.4%) を7.2ポイント上回った。この優位性は、Gemini 3 Proの1M (約104万) トークンという巨大なコンテキストウィンドウと相まって、長時間の実験動画や複雑な技術プロセスを包括的に理解する能力を示している。^{[2] [22] [1]}

TrueFoundryの比較分析によれば、Gemini 3 Proは画像の細部に注意を払い ("the arrow points left", "there are three resistors in series")、視覚的ハルシネーション (幻覚) が少ない傾向にある。一方、GPT-5.1 Thinkingも強力なマルチモーダル理解を持つが、画像の詳細を明示的に説明することが少なく、失敗する場合は画像ではなくテキストの誤読に起因することが多い。発明における図面や実験結果の正確な理解という点で、Gemini 3 Proの「視覚的な丁寧さ」は重要な利点となる。^{[20] [21]}

コンテキストウィンドウと長文脈推論

発明活動では、膨大な先行技術文献、特許データベース、科学論文を読み込み、そこから新規性のある着想を得る必要がある。この点で、AIのコンテキストウィンドウ（一度に処理できる情報量）と長文脈下での推論精度が重要となる。^{[1] [2] [3] [6] [8]}

Gemini 3 Proは最大1,048,576トークン（約1Mトークン）のコンテキストウィンドウを持ち、出力は最大65,536トークンに対応する。これは、GPT-5.1 Proの128,000トークン（最大400,000トークンまで拡張可能）と比較して、約2.6～8倍の情報を一度に処理できることを意味する。^{[2] [3] [1]}

この巨大なコンテキスト能力は、発明の「先行技術調査」段階で特に威力を発揮する。例えば、10,000ページの特許文献を読み込み、そこから関連する技術要素を抽出し、新規な組み合わせを提案するといったタスクにおいて、Gemini 3 Proは「context rot（長文脈での精度低下）」を最小限に抑えることができる。CometAPIの分析によれば、Gemini 3 Proは100,000トークン以上の文脈でも高い精度を維持するのに対し、GPT-5.1 Proは128kトークンを超えると精度が低下する傾向がある。^{[3] [23] [2]}

さらに、Gemini 3 ProのDeep Thinkモードは、この巨大なコンテキストを活用しながら、段階的な推論を行うことができる。これは、数千ページの科学論文から新しい仮説を導き出すような、発明の核心部分において決定的な優位性となる。^{[15] [16] [24] [1]}

創造性と推論の深さ：「天才的研究パートナー」vs「究極の熟練労働者」

Gemini自身が生成した比較レポートにおいて、Gemini 3 Proは「天才的な研究パートナー (genius research partner)」、GPT-5.1 Proは「究極の熟練労働者 (ultimate skilled worker)」と特徴付けられている。この定性的評価は、両モデルの根本的な設計思想の違いを示している。^{[1] [3] [20] [23]}

Gemini 3 Pro：探索的・創造的アプローチ

複数の比較分析が、Gemini 3 Proの「大胆で探索的 (bold and exploratory)」な特性を指摘している。TrueFoundryの評価によれば、Gemini 3 Proは「直感的な答えを素早くスケッチし、それを正当化しようとする」傾向があり、「正しいときは素晴らしく見え、間違っているときは非常に自信を持っている」。この特性は、発明の初期段階において、従来の枠にとらわれない新しいアイデアを生み出す上で有利に働く。^{[20] [21] [10] [16]}

Nature Scientific Reportsに掲載された研究によれば、ChatGPT-4は「既存の発表された研究から統計的計算によって最良の仮説を選択する」ことに優れているが、「好奇心に導かれた創造的プロセス」を欠いている。一方、Gemini 3 Proは、その高い抽象的推論能力（ARC-AGI-2で31.1%）と科学的直感（CritPtで9.1%、GPQA Diamondで91.9%）により、単なる既存知識の組み合わせを超えた「真の創造性」に近い能力を示している。^{[1] [2] [10]}

Science Advanceに掲載された研究では、生成AIが個人の創造性を高める一方で、集団全体の創造性多様性を減少させる可能性が指摘されている。しかし、Gemini 3 Proの探索的特性は、この問題を緩和する可能性がある。なぜなら、Gemini 3 Proは「独創的アプローチ」を取る傾向があり、一つの最適解に収束するのではなく、多様な仮説を生成する能力を持つからである。^{[2] [25] [1] [20]}

GPT-5.1 Pro : 安定性と実装力

対照的に、GPT-5.1 Proは「慎重で測定された (measured)」アプローチを取り、「複数ステップの推論を行うが、Kimi-K2のような大量のテキストよりも簡潔」である。TrueFoundryの評価では、GPT-5.1 Thinkingは「注意深いTAのように変数と実験設定を追跡する」と評され、「最もクリーンな『読む→要約する→推論する』パイプライン」を持つとされている。^[20]

この特性は、発明プロセスの「実装」段階、すなわち概念を実際の製品やプロトタイプに落とし込む段階において極めて有効である。実際、SWE-bench Verified (実際のGitHub issueを解決する能力) において、**GPT-5.1 Pro Codex-Maxは77.9%を達成し、Gemini 3 Pro (76.2%) をわずかに上回る**。また、Terminal-Bench 2.0やLiveCodeBenchにおいても、GPT-5.1 Proは高い安定性を示している。^{[1] [2] [3] [20]}

しかし、発明の「創出」フェーズにおいては、この安定性・慎重性が逆に制約となる可能性がある。Portkeyの分析によれば、「Geminiは学術的推論、マルチモーダル理解、長期タスクで優位に立つ。推論の深さ、強力な視覚的理解、高速で応答性の高い出力を優先するチームには、Gemini 3.0がより有能なモデルである」と結論付けられている。^{[10] [16] [20]}

科学的発見における実例と研究動向

DeepAgent : 物理公式の自律的導出

2025年11月に発表されたDeepAgentは、物理学における公式導出を完全に自律的に実行するAIシステムである。従来のAIシステムでは、問題解決のための思考プロセスとツールの選択・発見が別々のモジュールで処理されていたが、DeepAgentはこれらを単一の統合プロセスとして実装することに成功した。このシステムは、数学的推論と実験的検証を同時並行で実行し、従来の教科書的アプローチに縛られることなく、全く新しい数学的手法やツールを発見・活用する能力を持つ。^[4]

DeepAgentのような科学的発見AIの実現には、以下の能力が不可欠である：(1)既知の物理法則から独自の仮説を立てる能力、(2)仮説を検証するための実験手法を自ら考案する能力、(3)全く新しい数学的手法を発見する能力。これらの能力は、まさにGemini 3 ProがFrontierMath、CritPt、ARC-AGI-2で示した「未知への適応」「抽象的推論」「創造性」と直接対応している。^{[1] [2] [3] [5] [4]}

PhyE2E : 宇宙物理学の方程式発見

清華大学・北京大学などが開発したPhyE2Eは、生データから物理的関係と数式を自動的に導出するAIフレームワークである。Nature Machine Intelligence誌に掲載されたこの研究は、NASAの実際の天体物理学データを使ったテストで、太陽周期を説明する改善された数式を達成した。人間の研究者が何年もかけて導出する数式を自動的に発見する可能性を示しており、科学的発見の加速化が期待されている。^[5]

このような「データから理論を導出する」能力は、発明における「帰納的アプローチ」と密接に関連している。実験結果や観測データから、それを説明する新しい理論や原理を発見するプロセスは、発

明の核心である。Gemini 3 Proのマルチモーダル理解能力（MMMU-Pro 81.0%、Video-MMMU 87.6%）と長文脈処理能力（1Mトークン）は、大量の実験データや観測結果を統合的に分析し、そこから新しいパターンや法則を発見する上で極めて有効である。^{[1] [2] [4] [9] [22] [5]}

AI Scientist：完全自動化された科学研究

Sakana AIが提案する「AI Scientist」は、研究アイデアの生成、コード実装、実験実行、結果分析、論文執筆までを完全自動化するシステムである。このシステムは、機械学習研究に応用され、独自の研究アイデアを「ブレインストーミング」し、その新規性を評価し、必要なアルゴリズムを実装し、実験を実行し、結果を可視化して科学論文にまとめることができる。さらに、以前のアイデアとフィードバックを使用して次世代のアイデアを改善する「オープンエンドループ」で実行でき、人間の科学コミュニティをエミュレートする。^[22]

AI Scientistの4つの主要プロセスは、発明プロセスと完全に対応している：(1)アイデア生成（Idea Generation）、(2)実装（Implementation）、(3)実験と分析（Experimentation and Analysis）、(4)論文執筆（Paper Writing）。Gemini 3 Proの強みは、このうち(1)アイデア生成、特に「ブレインストーミング」と「新規性評価」の段階にある。一方、GPT-5.1 Proは(2)実装段階、特にコード生成とデバッグにおいて優位性を持つ。^{[1] [2] [20] [22]}

京都大学×Google：iPS細胞研究へのAI応用

京都大学iPS細胞研究所（CiRA）は、Googleが開発した科学研究向けAIを使って、iPS細胞を効率的に作る研究を開始した。このAIは「科学者のように仮説を立てる」能力を持ち、従来の試行錯誤的アプローチを大幅に効率化する可能性がある。このような「仮説駆動型研究」においては、AIの創造性と科学的直感が重要な役割を果たす。^{[26] [9] [27]}

Elucidataの論文「AI as a Co-Creative Partner」は、ドメイン特化型AIの重要性を強調している。生物学における因果関係推論、生物学的制約を考慮したシステムダイナミクスの理解、そして「数千の分子または細胞仮説をin silicoでテストする」能力が求められる。Gemini 3 ProのDeep Thinkモードは、まさにこのような「深い推論」を要する科学的探索において威力を発揮する。^{[1] [27] [15] [16]}

コスト効率と実用性

発明活動における生成AIの利用では、性能だけでなくコスト効率も重要な考慮事項である。^{[1] [2] [3] [23]}

Gemini 3 Proは、200,000トークン以下の入力で2.00ドル/1Mトークン、200,000トークン以上で4.00ドル/1Mトークンという料金体系を持つ。一方、GPT-5.1 Proは1.25ドル/1Mトークン（入力）だが、キャッシュなしの場合は最大13.50ドル/1Mトークンになる。^{[2] [3] [1]}

AceCloudの実コスト分析によれば、50kトークン入力・30kトークン出力のシナリオで、Gemini 3 Proは0.66ドル（88%の品質）、GPT-5.1 Proは5.85ドル（90%の品質）となり、**Gemini 3 Proはコストパフォーマンスで8.9倍優れている**。100kトークン入力・100kトークン出力のシナリオでは、Gemini 3 Proが1.40ドル（70%の品質）、GPT-5.1 Proが13.50ドル（90%の品質）となり、**Gemini 3 Proは約9.6倍のコスト効率を示す**。^{[3] [2]}

発明プロセスにおいては、膨大な文献調査、多数の仮説検証、反復的なアイデア改善が必要となるため、このコスト差は長期的に大きな影響を与える。特に、Gemini 3 Proの1Mトークンのコンテキスト

ウィンドウは、一度に大量の特許文献や科学論文を処理できるため、API呼び出し回数を削減し、さらなるコスト削減につながる。^{[22] [23] [2] [3]}

発明フェーズ別の推奨モデル

発明プロセス全体を考慮すると、各フェーズで最適な生成AIモデルは異なる。^{[2] [3] [20] [16]}

創出フェーズ（問題発見～理論構築）：Gemini 3 Pro推奨

問題発見・仮説生成、理論構築・アイデア創出、先行技術調査の各段階では、**Gemini 3 Proが最も適している**。これらのフェーズでは、創造性、抽象的思考、科学的直感、未知への適応能力が中核となるため、Gemini 3 ProのFrontierMath、CritPt、ARC-AGI-2での優位性が直接的に活きる。^{[1] [2] [3] [20] [16]}

特に、Gemini 3 ProのDeep Thinkモードは、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おいて卓越している。Tom's Guideの実機テストでは、Gemini 3が「より完全かつ効果的に、プロンプトの全ての部分（A、B、C）に深さ、明確さ、構造を持って対応した」と評価されている。^{[28] [15] [16] [24] [1]}

実装フェーズ（詳細設計～検証）：GPT-5.1 Pro推奨

詳細設計・実装、検証・テストの段階では、**GPT-5.1 Proが優れている**。これらのフェーズでは、実装力、コード品質、安定性、論理的検証能力が重要となり、GPT-5.1 ProのSWE-bench Verifiedでの高スコア（76.3%）やCodex-Max機能（77.9%）が威力を発揮する。^{[1] [2] [3] [20]}

GPT-5.1 Proの「慎重で測定されたアプローチ」は、既に確立されたアイデアを確実に実装する段階において理想的である。Portkeyの分析によれば、「GPT-5.1は特にコーディング安定性、ツール使用の一貫性、コスト効率において競争力を持ち、予測可能性、低分散、予算調整が重要な本番ワークロードにとって信頼できる選択肢である」。^{[20] [16]}

ハイブリッドアプローチ：両方の協調

概念実証（PoC）、特許明細書作成、反復改善の各段階では、**両モデルの協調利用が最適**である。例えば、概念実証段階では、Gemini 3 Proで理論的枠組みを構築し、GPT-5.1 Proで実装コードを生成するというワークフローが効果的である。^{[1] [2] [20] [23]}

特許明細書作成においては、Gemini 3 Proの包括的理解と図面生成能力（MMMU-Pro 81.0%、Video-MMMU 87.6%）を活用して技術内容を整理し、GPT-5.1 Proの構造化された文書作成能力で法的に正確な文言を構築するというアプローチが考えられる。^{[2] [6] [7] [8] [1]}

結論：Gemini 3 Proの創造的優位性

2025年11月時点の包括的な評価に基づき、発明の「創出」段階においては**Gemini 3 Proが最も力を発揮する生成AIである**と結論付けられる。

定量的エビデンス

Gemini 3 Proは、発明に直接関連する全ての主要ベンチマークにおいてGPT-5.1 Proを上回る性能を示した：

- **FrontierMath（数学的推論）**：Tier 1-3で+6.6%、Tier 4で+6.3%の優位^{[1] [2] [3]}
- **CritPt（物理学研究）**：+4.2%、約2倍の性能差^{[2] [3] [1]}
- **ARC-AGI-2（抽象的推論）**：+13.5%の大幅な優位^{[3] [1] [2]}
- **Humanity's Last Exam（総合評価）**：+11.0%の優位^{[1] [2] [3]}
- **GPQA Diamond（科学的推論）**：+3.8%の優位^{[2] [3] [1]}
- **MMMU-Pro（マルチモーダル理解）**：+5.0%の優位^{[3] [1] [2]}
- **Video-MMMU（動画理解）**：+7.2%の優位^{[1] [2] [3]}

これらのベンチマークは、いずれも「未知の問題への適応」「創造的な問題解決」「深い推論」という、発明の核心能力を測定している。^{[12] [17] [13] [2] [3] [1]}

定性的特性

Gemini 3 Proの「天才的な研究パートナー」としての特性—大胆で探索的なアプローチ、独創的な仮説生成、直感的な科学的理解—は、発明の初期段階における「0から1を生み出す」プロセスに最適である。対照的に、GPT-5.1 Proの「究極の熟練労働者」としての安定性と実装力は、既存のアイデアを確実に製品化する段階で威力を発揮するが、真の創造性が求められる発明創出段階では制約となる。^{[1] [3] [20] [10] [16]}

技術的優位性

Gemini 3 ProのDeep Thinkモード、1Mトークンのコンテキストウィンドウ、最高クラスのマルチモーダル理解能力は、複雑な科学理論の構築、膨大な文献からの着想獲得、視覚的・動的現象の理解という、発明プロセスの中核的タスクにおいて決定的な優位性を提供する。^{[1] [2] [22] [15] [16]}

コスト効率

発明活動における反復的な探索プロセスを考慮すると、Gemini 3 Proの8.9～9.6倍のコスト効率は、長期的な研究開発において極めて重要である。^{[2] [3] [23]}

実世界への示唆

DeepAgent、PhyE2E、AI Scientistなど、最先端の科学的発見AIシステムが示す能力要件—自律的仮説生成、新規ツール発見、データからの理論導出—は、いずれもGemini 3 Proの強みと一致している。京都大学×GoogleのiPS細胞研究や、各種の創薬AI、材料科学AIにおいても、「仮説駆動型」「探索的」「マルチモーダル」という特性が重視されており、Gemini 3 Proの設計思想がこの方向と整合している。^{[4] [5] [26] [9] [22] [27] [29]}

したがって、発明、特に新規アイデアの創出と科学的発見の段階においては、**2025年11月時点でGemini 3 Proが最も力を発揮する生成AIである**。ただし、発明プロセス全体の効率化を考えると、創出段階ではGemini 3 Pro、実装段階ではGPT-5.1 Proを使い分ける、あるいは両者を協調させるハイブリッドアプローチが最も効果的であると言える。^{[2] [20] [23] [16]}

1. GPT-51-Pro-to-Gemini-3-Pro-Xiang-Xi-Bi-Jiao-ChatGPT.pdf
2. GPT-51-Pro-to-Gemini-3-Pro-woBi-Jiao-Perplexity.pdf
3. GPT-51-Pro-vs-Gemini-3-Pro-Bi-Jiao-Gemini.pdf
4. https://oneword.co.jp/bignite/ai_news/deepagent-physics-formula-derivation-ai-autonomous-thinking/
5. <https://note.com/takao19791012/n/nfc4d04593e8b>
6. <https://relecura.ai/the-ai-revolution-in-invention-generative-ais-role-in-identifying-new-patents/>
7. <https://www.solveintelligence.com/blog/post/generative-ai-and-intellectual-property>
8. <https://www.patlytics.ai/blog/what-is-ai-patent-generator>
9. <https://arxiv.org/html/2412.11427v1>
10. <https://www.nature.com/articles/s41598-025-93794-9>
11. <https://huggingface.co/papers/2411.04872>
12. https://note.com/life_to_ai/n/n54d6094f5c47
13. <https://epoch.ai/frontiermath/the-benchmark>
14. <https://epoch.ai/frontiermath>
15. <https://blog.google/products/gemini/gemini-3/>
16. <https://portkey.ai/blog/gemini-3-0-vs-gpt-5-1/>
17. <https://arxiv.org/abs/2509.26574>
18. <https://arxiv.org/pdf/2509.26574.pdf>
19. <https://vertu.com/lifestyle/gemini-3-launch-google-strikes-back-less-than-a-week-after-gpt-5-1-release/>
20. <https://www.truefoundry.com/blog/gemini-3>
21. <https://macaron.im/blog/google-gemini3-pro>
22. <https://sakana.ai/ai-scientist/>
23. <https://www.cometapi.com/gemini-3-pro-vs-gpt-5-1-which-is-better-a-complete-comparison/>
24. <https://masterconcept.ai/blog/gemini-3-the-new-ai-king-a-deep-dive-into-the-breakthrough-features-that-beat-gpt-5-1/>
25. <https://www.science.org/doi/10.1126/sciadv.adn5290>
26. <https://www.nikkei.com/article/DGXZQOSG1417O0U5A111C2000000/>
27. <https://www.elucidata.io/blog/ai-co-creative-partner-scientific-discovery>
28. <https://www.tomsguide.com/ai/i-just-tested-gemini-3-vs-chatgpt-5-1-and-one-ai-crushed-the-competition>
29. <https://www.zero11.it/en/magazine/generative-ai-the-expansive-power-revolutionizing-industries>
30. <https://gais.jp/news-2025-11-21/>
31. <https://www.youtube.com/watch?v=J4erRALO12U>
32. <https://www.kwm.com/au/en/insights/latest-thinking/inventing-inventors-generative-ai-and-patents.html>
33. https://note.com/hakuhodo_o/n/n2cbbecaa2bfe

34. <https://www.wipo.int/web-publications/patent-landscape-report-generative-artificial-intelligence-geni/en/index.html>
35. <https://xlscout.ai/how-is-generative-ai-changing-the-innovation-approach/>
36. <https://news.livedoor.com/article/detail/30050805/>
37. <https://epochai.substack.com/p/is-ai-already-superhuman-on-frontiermath>
38. <https://www.synlabs.io/post/how-generative-ai-can-boost-highly-skilled-workers-productivity>
39. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12112869/>
40. <https://www.bcg.com/publications/2024/gen-ai-increases-productivity-and-expands-capabilities>
41. <https://www.youtube.com/watch?v=RJebDTnRfEw>
42. <https://www.hiringlab.org/2025/09/23/ai-at-work-report-2025-how-genai-is-rewiring-the-dna-of-jobs/>
43. <https://www.synthesia.io/post/ai-tools>
44. <https://hai.stanford.edu/news/what-workers-really-want-from-artificial-intelligence>
45. <https://skywork.ai/blog/ai-agent/gpt5-1-vs-gemini-3/>
46. <https://www.sciencedirect.com/science/article/pii/S2713374524000050>
47. <https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work>
48. https://www.reddit.com/r/ChatGPTPro/comments/1p0zis/gemini_3_is_what_gpt_5_should_have_been_it_s/
49. https://www.youtube.com/watch?v=JuKhk_xD3Vk
50. <https://www.ainewsinternational.com/zero-to-one-with-ai-startup-playbooks-for-the-generative-era/>
51. <https://www.glbgt.com/hub/gpt-5-vs-gemini-2-5-pro-a-detailed-ai-model-review/>
52. <https://inspireip.com/top-ai-idea-generation-software-to-elevate-innovation/>
53. <https://deepmind.google/models/gemini/>
54. <https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2025.1585629/full>
55. <https://futureagi.com/blogs/llm-benchmarking-compare-2025>
56. <https://deepmind.google/models/gemini/pro/>
57. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12241030/>
58. <https://blog.google/intl/ja-jp/company-news/technology/gemini-3/>
59. <https://www.oxx.vc/industry-perspectives/from-invention-to-innovation-the-real-pace-of-agentic-ai-adoption/>
60. <https://www.youtube.com/watch?v=bLoBezKKiil>
61. <https://gemini.google/jp/release-notes/?hl=ja>
62. <https://www.zeroandone.me/blogs/evolution-of-ai.html>