



# Gemini 3 : Googleが掲げる「知性の新時代」を徹底検証

2025年11月19日（日本時間）、Googleは次世代AIモデル「**Gemini 3**」を正式発表し、AI業界に大きな衝撃を与えた。スンダー・ピチャイCEOが「あらゆるアイデアを現実にする、これまでで最も高性能なモデル」と豪語するGemini 3は、博士号レベルの推論能力と圧倒的なマルチモーダル性能を武器に、競合モデルを大きく引き離すベンチマーク結果を提示している。本レポートでは、Gemini 3の技術仕様、性能評価、ユーザー反応、競合比較を通じて、この新モデルの真価を多角的に分析する。<sup>[1]</sup><sup>[2]</sup><sup>[3]</sup>

## Gemini 3の全体像：リリース概要と戦略的位置づけ

### 発表の背景とタイミング

Gemini 3の発表は、GoogleのAI戦略における重要な転換点を象徴している。前モデルGemini 2.5 Proは6ヶ月以上にわたりLMArenaリーダーボードで首位を維持し続けたが、OpenAIのGPT-5.1やAnthropicのClaude Sonnet 4.5といった強力な競合の登場により、Googleは技術的優位性の再確立を迫られていた。<sup>[1]</sup><sup>[3]</sup><sup>[4]</sup>

特筆すべきは、**Geminiモデルとして初めてリリース初日にGoogle検索のAIモードに統合された点**である。これはGoogleがGemini 3を単なる研究プロトタイプではなく、実用段階の中核技術として位置づけていることを示している。<sup>[5]</sup><sup>[1]</sup>

### 提供形態とアクセス方法

Gemini 3は以下の形態で段階的に展開されている：<sup>[1]</sup><sup>[6]</sup><sup>[7]</sup>

#### 一般ユーザー向け

- Geminiアプリ（無料版・有料版）
- Google検索のAIモード（Google AI Pro/Ultraサブスクリプション）

#### 開発者向け

- Google AI Studio
- Vertex AI
- Gemini CLI
- Google Antigravity（新エージェント開発プラットフォーム）
- サードパーティ統合（Cursor、GitHub、JetBrains、Replit等）

#### エンタープライズ向け

- Vertex AI
- Gemini Enterprise

## 技術仕様と主要機能

### コンテキストウィンドウとマルチモーダル処理

Gemini 3 Proは**100万トークンのコンテキストウィンドウ**を標準装備しており、これは約1,500ページのテキストまたは3万行のコードに相当する。この長大なコンテキストは、学术论文のデータベース全体を一度に読み込んで分析するような、従来のAIでは不可能だった複雑なタスクを可能にする。<sup>[1]</sup><sup>[2]</sup><sup>[8]</sup><sup>[9]</sup><sup>[10]</sup><sup>[11]</sup>

マルチモーダル機能では、**テキスト、画像、動画、音声、コード**をシームレスに統合処理できる。具体的な応用例として、外国語の手書きレシピの解読・翻訳、動画分析によるスポーツフォーム改善プランの作成、長時間講義動画からのインタラクティブな学習教材生成などが挙げられている。<sup>[2]</sup><sup>[1]</sup>

### Deep Thinkモード：強化推論の新次元

Gemini 3の最大の技術的革新は「**Deep Think**」モードの導入である。このモードは、AIが段階的に深く思考するプロセスを経ることで、複雑な問題解決能力を飛躍的に向上させる。<sup>[1]</sup><sup>[3]</sup><sup>[5]</sup><sup>[12]</sup>

Deep Thinkモードの性能は通常のGemini 3 Proを大きく上回る：<sup>[13]</sup><sup>[1]</sup>

- **Humanity's Last Exam**（ツール不使用）：37.5% → **41.0%**
- **GPQA Diamond**（博士号レベル推論）：91.9% → **93.8%**
- **ARC-AGI**（コード実行あり）：**45.1%**（前例のない記録）

特にARC-AGI-2における45.1%という数値は、未知の課題への適応力を測る指標として業界で注目されており、汎用人工知能（AGI）実現への重要な一歩と評価されている。<sup>[14]</sup><sup>[15]</sup><sup>[16]</sup><sup>[1]</sup>

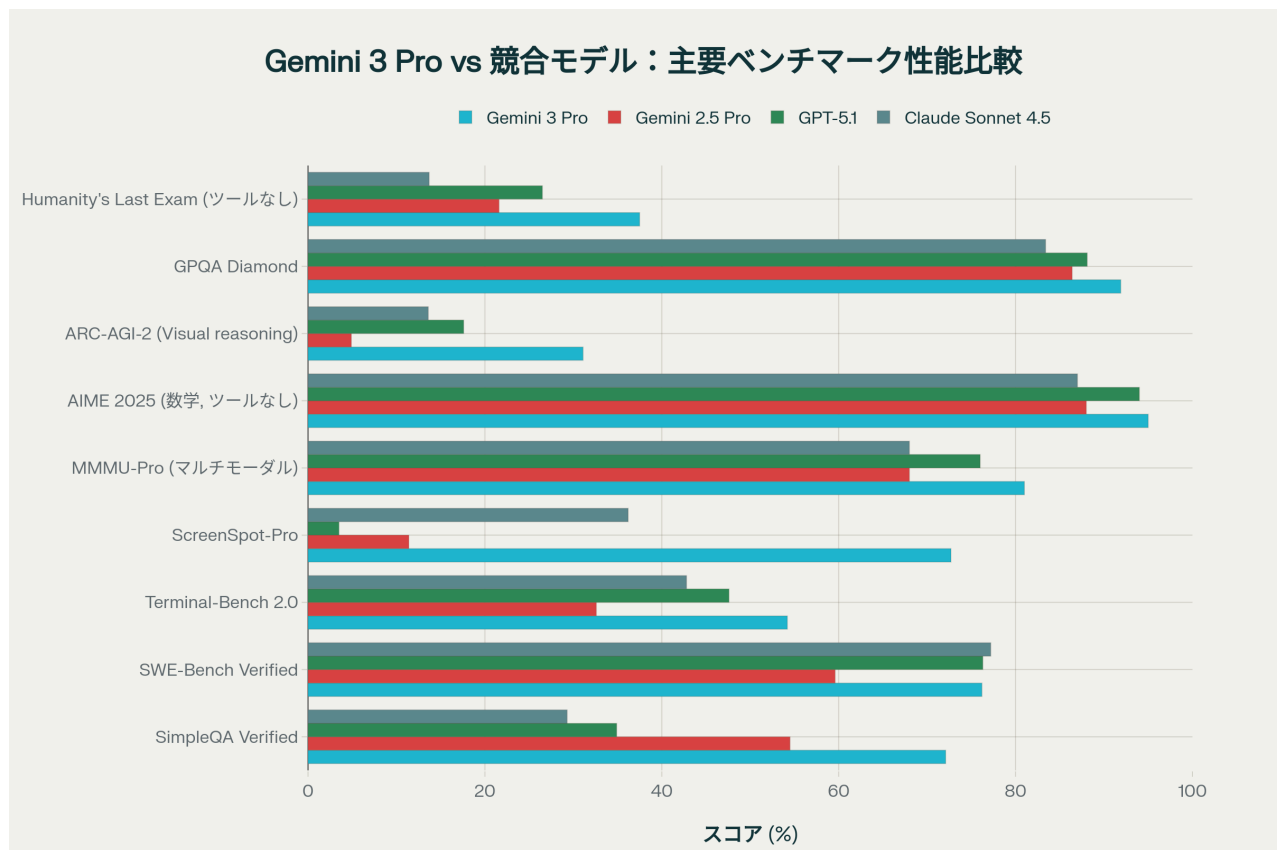
### コーディング能力：「Vibe Coding」の進化

Gemini 3は「**Vibe Coding**」と呼ばれる自然言語によるコード生成において、これまでで最高のパフォーマンスを示している。<sup>[1]</sup><sup>[3]</sup><sup>[13]</sup>

主要なコーディングベンチマークでの成績：<sup>[17]</sup><sup>[18]</sup><sup>[1]</sup>

- **WebDev Arena**（Elo）：**1487**（リーダーボード首位）
- **SWE-bench Verified**：**76.2%**（Gemini 2.5 Proから+16.6ポイント）
- **Terminal-Bench 2.0**：**54.2%**（ターミナル経由のコンピューター操作能力）

初期テスターの報告によれば、Gemini 3は「複雑なアニメーションや奥行きのあるウェブデザインを一発で生成できる」精度を備えており、従来モデルでは数回の修正が必要だったタスクが一度で完結するケースが増えている。<sup>[8]</sup><sup>[19]</sup><sup>[20]</sup>



Gemini 3 Proと競合モデル（Gemini 2.5 Pro、GPT-5.1、Claude Sonnet 4.5）の主要ベンチマークにおける性能比較。Gemini 3 Proはほぼ全てのベンチマークで優位性を示している。

## ベンチマーク性能：競合モデルとの徹底比較

### 主要ベンチマークにおける優位性

Gemini 3 Proは発表されたほぼ全てのベンチマークで、前モデルGemini 2.5 Proおよび主要競合モデルを上回る性能を示している。 [\[1\]](#) [\[17\]](#) [\[3\]](#)

#### 推論・知識ベンチマーク

- **LMarena**（総合評価）：**1501 Elo**（リーダーボード首位、Gemini 2.5 Proは1451 Elo） [\[3\]](#) [\[5\]](#) [\[1\]](#)
- **GPQA Diamond**（博士号レベル科学知識）：**91.9%**（GPT-5.1: 88.1%, Claude Sonnet 4.5: 83.4%） [\[17\]](#) [\[1\]](#) [\[3\]](#)
- **Humanity's Last Exam**（ツール不使用）：**37.5%**（GPT-5.1: 26.5%, Claude Sonnet 4.5: 13.7%, Gemini 2.5 Pro: 21.6%） [\[13\]](#) [\[1\]](#) [\[17\]](#)

#### 数学・論理推論

- **AIME 2025**（数学、ツール不使用）：**95.0%**（GPT-5.1: 94.0%とほぼ同等） [\[17\]](#)
- **MathArena Apex**：**23.4%**（フロンティアモデルの新記録） [\[1\]](#) [\[3\]](#)

#### マルチモーダル理解

- **MMMU-Pro**（画像推論）：**81.0%**（GPT-5.1: 76.0%, Claude/Gemini 2.5 Pro: 68.0%） [\[13\]](#) [\[1\]](#) [\[17\]](#)

- **Video-MMMU**（動画理解）： **87.6%** <sup>[3] [1]</sup>
- **ScreenSpot-Pro**（画面理解）： **72.7%**（GPT-5.1: 3.5%, Gemini 2.5 Pro: 11.4%） <sup>[17]</sup>

## 事実精度

- **SimpleQA Verified**（パラメトリック知識）： **72.1%**（GPT-5.1: 34.9%, Claude Sonnet 4.5: 29.3%, Gemini 2.5 Pro: 54.5%） <sup>[1] [17]</sup>

## 特に優れている領域

**視覚的推論・画面理解**：ARC-AGI-2で31.1%（Gemini 2.5 Pro: 4.9%、GPT-5.1: 17.6%）、ScreenSpot-Proで72.7%という圧倒的な差は、Gemini 3がビジュアル情報処理において大きなブレークスルーを達成したことを示している。 <sup>[17]</sup>

**コンテキスト保持**：MRCR v2（8-needle）ベンチマークにおいて、128kトークンでの平均77.0%、100万トークンでのポイントワイズ26.3%という成績は、長大なコンテキストでも情報を正確に保持・検索できる能力を実証している。 <sup>[17]</sup>

**マルチモーダル統合**：Video-MMMUで87.6%という高スコアは、動画内の複雑な文脈理解において他モデルを引き離している。 <sup>[1] [13]</sup>

## 弱点と課題

**SWE-bench Verified**では76.2%とClaude Sonnet 4.5の77.2%に僅差で劣っており、実用的なコーディングエージェント能力では依然としてClaudeが優位な領域も存在する。ただし、この差は誤差範囲内とも解釈でき、開発者コミュニティでは「ベンチマーク特化による過学習（benchmark gaming）」の可能性も指摘されている。 <sup>[17] [21] [22] [23]</sup>

また、**長期プラン実行**を測るVending-Bench 2ではGemini 3 Proがリーダーボードのトップに立っているものの、実際のエージェントタスクにおける30時間以上の自律動作という点では、Claude Sonnet 4.5の実績が依然として優れているとの評価もある。 <sup>[1] [21]</sup>

## API価格とコストパフォーマンス

### 料金体系

Gemini 3 Pro APIは**従量課金制**で提供され、コンテキスト長に応じた階層料金が設定されている：  
<sup>[24] [25]</sup>

### Gemini 3 Pro Preview（2025年11月時点）

- 入力：\$2.00/100万トークン（≤200Kトークン）、\$4.00/100万トークン（>200K）
- 出力：\$12.00/100万トークン（≤200K）、\$18.00/100万トークン（>200K）

### 比較：Gemini 2.5 Pro

- 入力：\$1.25/100万トークン（≤200K）、\$2.50/100万トークン（>200K）
- 出力：\$10.00/100万トークン（≤200K）、\$15.00/100万トークン（>200K）

### 比較：Claude 3.5 Sonnet

- 入力/出力：\$3.00/\$15.00（コンテキスト長による割増なし）<sup>[21] [24]</sup>

## コスト分析

Gemini 3 Proは前モデルから**入力で60%、出力で20%の価格上昇**となっているが、これは性能向上を考慮すれば妥当な範囲と評価されている。Claude 3.5 Sonnetと比較すると、短いコンテキスト（≤200K）ではGemini 3 Proの方が入力コストで33%安く、出力コストで20%安い。<sup>[24] [21]</sup>

ただし、GPT-5.1は「多くの単純なクエリで使用トークンが少ない」と主張されており、実際のコストは使用パターンに大きく依存する。また、Googleは通常プレビューモデルが安定版になる際に20～50%の料金削減を実施するため、2026年初頭には\$1.50/\$10（≤200K）程度に落ち着くと予想されている。<sup>[21] [24]</sup>

## ユーザー評価と実使用における評判

### 肯定的評価

**推論能力の向上**：Hacker NewsやRedditのディスカッションでは、「複雑な問題に対する深い理解」「微妙なニュアンスを捉える能力」が高く評価されている。特に、従来モデルで見落とされがちだった論理的欠陥を指摘できる点が開発者から好評を得ている。<sup>[22] [26]</sup>

**Google生態系との統合**：Gmail、Googleカレンダー、Docs、スプレッドシートとのシームレスな連携は、Geminiシリーズの大きな強みとして一貫して評価されている。Gemini 3では特にGemini Agentによる受信トレイ整理や旅行計画などの複雑なワークフロー支援が可能になった。<sup>[1] [3] [27]</sup>

**マルチモーダル性能**：動画分析やリアルタイム音声処理における精度向上が実感されており、「スポーツ動画からフォーム改善提案を得られた」「長時間会議の文字起こしと要約が正確」といった実用例が報告されている。<sup>[26] [28] [1]</sup>

### 否定的評価・懸念点

**ハルシネーション（幻覚）問題**：いくつかのユーザー報告では、Gemini 3でも依然として事実と異なる情報を生成するケースがある。ただし、最新のHallucination Leaderboard（Vectara社、2025年4月）では、Gemini-2.0-Flash-001が幻覚率0.7%で首位、Gemini-2.0-Pro-Expが0.8%で2位を獲得しており、Geminiシリーズ全体として事実忠実度は大幅に改善している。<sup>[26] [29] [30] [31] [32]</sup>

**レート制限と可用性**：Hacker Newsでは、無料版ユーザーが「すぐにレート制限に達する」「有料プランでも思ったほど使えない」といった不満を表明している。これは高性能モデルの計算コストが高いことの反映だが、ユーザー体験としてはマイナス要因となっている。<sup>[26]</sup>

**音声文字起こしの精度**：一部のテストでは、Gemini 3 Proが会議の文字起こしで「内容を大幅に変更」「会話全体を捏造」するハルシネーションが報告されており、Gemini 2.5 Proの方が安定していたとの指摘もある。<sup>[26]</sup>

**日本語対応の不十分さ**：日本語対応がまだ発展途上であり、一部に英文表記が残っている点、GPT-4と比較して日本語の自然さでやや劣るという評価が継続している。<sup>[33] [27]</sup>

## コミュニティの総合評価

Reddit、Hacker News、Xなどの技術コミュニティでは、Gemini 3に対する評価は慎重ながらも肯定的である。「ベンチマークは素晴らしいが、実際に使ってみないと分からない」「期待しすぎると失望する可能性がある」といった冷静な意見が多く見られる一方で、「特定タスクでは明らかに従来モデルを超えている」「コーディング支援が実用レベルに達した」といった実感を伴う評価も増えている。<sup>[34] [35] [22] [26]</sup>

## 競合モデルとの戦略的比較

### GPT-5.1との比較

**推論の深さ vs 広さ**：業界アナリストは、GPT-5.1が「複雑な推論と戦略立案」に優れているのに対し、Gemini 3は「マルチモーダル理解と現実世界の状況把握」に強みを持つと分析している。<sup>[4] [28] [36]</sup>

**エコシステム**：GPT-5.1はChatGPTアプリ、Windows連携、Office系統合で成熟しているのに対し、Gemini 3はGoogleの巨大なサービス群（検索、YouTube、Gmail、Maps、Workspace）との統合で優位性を持つ。<sup>[28]</sup>

**コスト効率**：定価ベースではGPT-5.1の方が大幅に安価だが（出力が約2～3倍安い）、実際のコストはキャッシュ、バッチ処理、使用トークン数によって変動する。<sup>[21]</sup>

### Claude Sonnet 4.5との比較

**信頼性重視のClaude**：Claude Sonnet 4.5は「確実に動作するコード」を生成する信頼性で開発者から絶大な支持を得ており、コーディングとエージェント構築において「クラス最高」と評価されている。<sup>[4] [21] [37]</sup>

**30時間の自律動作**：Claudeは約30時間の自律動作を維持できると報告されており、持続的なエンジニアリングタスクやエージェント自動化では依然として優位。<sup>[21]</sup>

**安全性とアライメント**：Anthropicの安全性優先のエンジニアリングは、プロンプトインジェクション耐性や倫理的判断において高く評価されている。<sup>[29] [21]</sup>

## 適材適所の時代

AI専門家は「もはや単一モデルが全てのタスクで最高という時代ではない」と指摘している。推奨される使い分けは以下の通り：<sup>[4] [28] [36]</sup>

- 複雑な思考・戦略立案：GPT-5.1 Thinking
- Web操作自動化・マルチモーダル（動画・3D）：Gemini 3
- 信頼性が求められるコーディング・システム連携：Claude Sonnet 4.5

# Google Antigravity : エージェント開発の新パラダイム

## 概要と特徴

Gemini 3と同時に発表された「**Google Antigravity**」は、エージェント型開発を可能にする新しいIDEプラットフォームである。従来のAIアシスタントが「ツール」だったのに対し、Antigravityは「能動的なパートナー」として複雑なエンドツーエンドのソフトウェアタスクを自律的に計画・実行する。<sup>[1] [3] [38] [39]</sup>

主要機能 : <sup>[39] [1]</sup>

- エディター、ターミナル、ブラウザへの直接アクセス
- 複雑なソフトウェアタスクの自律的な計画と実行
- 自身が生成したコードの自動検証
- 詳細なアーティファクトを通じたユーザーとのコミュニケーション

## 統合モデル

Antigravityには以下のモデルが統合されている : <sup>[1]</sup>

- **Gemini 3 Pro** (主力)
- **Gemini 2.5 Computer Use** (ブラウザ制御)
- **Nano Banana (Gemini 2.5 Image)** (画像編集)

## 評価と課題

開発者コミュニティでは、Antigravityの可能性に期待が集まる一方で、「まだ粗削り」「思った通りに動かない」といった初期段階の課題も指摘されている。特に、AIエージェントが「間違った方向に突き進む」「指示を誤解する」といった問題は、長時間タスクにおいて依然として解決すべき課題である。<sup>[26] [40]</sup>

## 安全性とセキュリティへの取り組み

### 包括的な安全性評価

Googleは、Gemini 3を「これまでで最も安全なモデル」と位置づけ、同社が開発した全てのモデルの中で最も包括的な安全性評価を実施したと発表している。<sup>[1] [3] [41]</sup>

主な安全性強化 : <sup>[3] [1]</sup>

- 迎合性 (sycophancy) の低減
- プロンプトインジェクション耐性の向上
- サイバー攻撃による悪用への防御力強化

## 外部機関による独立評価

Googleは**Frontier Safety Framework**に基づく社内テストに加え、以下の外部評価を実施している：<sup>[1] [41] [42]</sup>

- 英国AI Safety Institute (AISi) への早期アクセス提供
- Apollo、Vaultis、Dreadnodeなど業界専門家による独立評価
- 世界をリードする有識者との協力評価

## プライバシーとデータガバナンス

無料版では入力データが基盤モデルの学習に利用される可能性があり、人間のレビュアーがデータを閲覧することもあるため、機密情報や個人情報の入力には注意が必要である。有料版では入力データが学習に利用されないなど、より高いセキュリティが確保される。<sup>[43] [44] [42]</sup>

## 技術的イノベーションと今後の展望

### Sparse MoEと強化学習

技術解説記事によれば、Gemini 3はSparse Mixture of Experts (MoE) アーキテクチャと強化学習による「思考プロセス」の統合により、専門家レベルの推論精度を実現している。この設計により、指示されなくても自律的に深い推論ループを継承できるようになったとされる。<sup>[45]</sup>

### エージェント能力の進化

Gemini 3のエージェント能力は、単なるツール呼び出しを超えて、**長期的視野での確実な計画立案**にまで進化している。Vending-Bench 2での成績は、シミュレートされた1年間の自動販売機ビジネス運営において、一貫したツール使用と意思決定を維持できることを示している。<sup>[1]</sup>

### 将来のロードマップ

Googleは「Gemini 3時代の始まりに過ぎない」として、今後数ヶ月以内に**Gemini 3シリーズに追加モデル**を公開する予定であることを明らかにしている。これには、さらなる軽量・高速モデルや特定領域に特化したバリエーションが含まれる可能性がある。<sup>[1] [6] [7]</sup>

また、**Google検索のAIモード**における生成UI体験（インタラクティブなツール、シミュレーション）の拡充により、Gemini 3は検索体験そのものを根本的に変革する可能性を秘めている。<sup>[1]</sup>

## 批判的考察：期待と現実のギャップ

### ベンチマーク至上主義への警鐘

技術コミュニティでは、「ベンチマークスコアと実際の使用感は必ずしも一致しない」という指摘が繰り返されている。特にSWE-bench VerifiedのようなPythonに偏ったベンチマークは、実際の開発現場での汎用性を正確に反映していない可能性がある。<sup>[34] [35] [22] [23] [46]</sup>



## 「狼少年」現象

Gemini 3リリース前、X（旧Twitter）では「明日リリース」という投稿が毎日のように繰り返され、期待が過剰に煽られた結果、実際のリリース時には一部で失望感も見られた。この「狼少年」現象は、AI業界におけるマーケティングと実体のバランスについて重要な教訓を提供している。<sup>[8] [19]</sup>

## 過度な期待の危険性

GPT-5リリース時にも「すごいけど、思ったほどじゃない」という声が上がったように、AIの進化は確かに目覚ましいが「魔法のように全ての問題を解決してくれるわけではない」。Gemini 3についても、リーク情報で語られた「GPT-5の3倍の性能」「圧倒的なコストパフォーマンス」といった主張は、実際にリリースされて検証するまで慎重に評価すべきである。<sup>[19]</sup>

## 結論：Gemini 3がもたらす変革と課題

### 主要な知見のまとめ

Gemini 3は以下の点で明確な技術的進歩を示している：

1. **博士号レベルの推論能力**：GPQA Diamond 91.9%、Humanity's Last Exam 37.5%という成績は、専門的知識を要する複雑な問題解決において人間専門家に迫る能力を実証している。<sup>[1] [17] [13]</sup>
2. **マルチモーダル理解の飛躍**：特に視覚的推論（ARC-AGI-2: 31.1%）、画面理解（ScreenSpot-Pro: 72.7%）、動画解析（Video-MMMU: 87.6%）において、競合モデルを大きく引き離す性能を達成している。<sup>[17] [1]</sup>
3. **実用的なコーディング支援**：WebDev Arena 1487 Elo、SWE-bench Verified 76.2%という成績は、「Vibe Coding」が実用段階に達したことを示している。<sup>[18] [1] [17]</sup>
4. **長大なコンテキスト処理**：100万トークンのコンテキストウィンドウと高精度の情報検索能力（MRCC v2で77%）は、従来不可能だった大規模文書分析を可能にする。<sup>[10] [1] [17]</sup>
5. **Google生態系との深い統合**：検索、Gmail、Workspace等との連携により、日常業務における実用性が大幅に向上している。<sup>[28] [27] [1]</sup>

### 残された課題

一方で、以下の課題も明確である：

1. **ハルシネーションの完全解消には至っていない**：改善傾向にあるものの、依然として事実と異なる情報を生成するケースが報告されている。<sup>[26] [29] [30]</sup>
2. **長時間自律動作の信頼性**：30時間レベルの持続的なエージェントタスクでは、Claude Sonnet 4.5が依然として優位な可能性がある。<sup>[21]</sup>
3. **日本語対応の完全性**：英文表記の残存、自然な日本語生成における改善の余地がある。<sup>[33] [27]</sup>
4. **料金の上昇**：前モデルから60%の入力コスト増は、高頻度利用者にとって無視できない負担である。<sup>[24]</sup>

## 最終評価：「適材適所」の時代の到来

Gemini 3は、GoogleがAI技術競争において再び主導的地位を確立する上で重要な一步を踏み出したことを示している。ベンチマーク性能、マルチモーダル能力、生態系統合の面で明確な優位性を示す一方で、「全てのタスクで最高」という単純な結論には至らない。

むしろ、Gemini 3の登場はAI利用の「適材適所」時代の到来を象徴している。複雑な思考タスクはGPT-5.1 Thinking、信頼性が求められるコーディングはClaude Sonnet 4.5、マルチモーダル理解と現実世界の状況把握はGemini 3、という使い分けが現実的である。[4] [28] [36]

今後、Gemini 3シリーズの追加モデル、Deep Thinkモードの一般公開、Google Antigravityの成熟により、Geminiの実用性はさらに向上すると予想される。同時に、競合各社も継続的にモデルを改善するため、AI性能競争は今後も激化し続けるだろう。この競争は最終的にユーザーに利益をもたらすが、過度な期待や誇大広告に惑わされず、冷静に各モデルの強みと弱みを見極める姿勢が重要である。



1. <https://blog.google/intl/ja-jp/company-news/technology/gemini-3/>
2. <https://k-tai.watch.impress.co.jp/docs/news/2064382.html>
3. <https://japan.zdnet.com/article/35240617/>
4. <https://note.com/sherpayamasaki/n/n59d262be0518>
5. <https://smhn.info/202511-google-gemini-3>
6. [https://note.com/masa\\_wunder/n/n21fe502360e7](https://note.com/masa_wunder/n/n21fe502360e7)
7. <https://helentech.jp/news-78428/>
8. [https://note.com/ai\\_otaku\\_2030/n/ndca811c8ec6c](https://note.com/ai_otaku_2030/n/ndca811c8ec6c)
9. <https://skywork.ai/blog/llm/what-is-gemini-3-full-overview-differences-from-gemini-jp/>
10. <https://docs.cloud.google.com/vertex-ai/generative-ai/docs/long-context?hl=ja>
11. <https://support.google.com/gemini/answer/16275805?hl=ja-JP>
12. [https://note.com/ai\\_otaku\\_2030/n/n831702e29c89](https://note.com/ai_otaku_2030/n/n831702e29c89)
13. <https://9to5google.com/2025/11/18/gemini-3-launch/>
14. <https://news.mynavi.jp/techplus/article/20251119-3686738/>
15. <https://ai.codigital.co.jp/news/google-ai-model-gemini-3/>
16. <https://slashgear.jp/2025/11/19/googlegemini-3/>
17. <https://simonwillison.net/2025/Nov/18/gemini-3/>
18. <https://blog.google/products/gemini/gemini-3/>
19. <https://ai-otaku.jp/463/>
20. <https://note.com/yasuhitoo/n/n7300d6de770b>
21. <https://www.cometapi.com/ja/gpt-5-1-vs-claude-sonnet-4-5/>
22. <https://news.ycombinator.com/item?id=45963836>
23. [https://scale.com/leaderboard/swe\\_bench\\_pro\\_public](https://scale.com/leaderboard/swe_bench_pro_public)
24. <https://apidog.com/jp/blog/gemini-3-0-api-cost/>

25. <https://ai.google.dev/gemini-api/docs/pricing?hl=ja>
26. <https://news.ycombinator.com/item?id=45967211>
27. <https://yukeblog.org/blog/ai51/>
28. <https://www.reimagines.jp/gpt5-1-vs-gemini3-0-docchi/>
29. <https://craftai.jp/gemini-many-mistakes/>
30. <https://gigazine.net/news/20240329-ai-hallucinations-software-package/>
31. <https://chatgpt-enterprise.jp/blog/hallucination-leaderboard/>
32. <https://note.com/hoboai/n/n15fc4fa11036>
33. <https://tenshoku.mynavi.jp/engineer/guide/articles/ZIW0ExIAACgAPGCp>
34. [https://www.reddit.com/r/Bard/comments/1n6uxau/gemini\\_3\\_will\\_be\\_good\\_in\\_coding\\_and\\_multimodal/](https://www.reddit.com/r/Bard/comments/1n6uxau/gemini_3_will_be_good_in_coding_and_multimodal/)
35. [https://www.reddit.com/r/Bard/comments/1mpceoj/what\\_is\\_your\\_expectation\\_of\\_gemini\\_30/](https://www.reddit.com/r/Bard/comments/1mpceoj/what_is_your_expectation_of_gemini_30/)
36. <https://note.com/tagtag/n/ne82595bdf084>
37. [https://www.gizmodo.jp/2025/10/anthropic\\_claude\\_sonnet\\_4\\_5\\_release.html](https://www.gizmodo.jp/2025/10/anthropic_claude_sonnet_4_5_release.html)
38. [https://search.yahoo.co.jp/realtime/search/matome/2f872f37b04b42abb7b7a48846f8dbb5-1763490012?rkf=1&ifrt=tl\\_matometl](https://search.yahoo.co.jp/realtime/search/matome/2f872f37b04b42abb7b7a48846f8dbb5-1763490012?rkf=1&ifrt=tl_matometl)
39. <https://blog.google/technology/developers/gemini-3-developers/>
40. <https://news.ycombinator.com/item?id=45968104>
41. <https://innovatopia.jp/ai/ai-news/72021/>
42. <https://liginc.co.jp/656995>
43. [https://note.com/norito\\_hiraoka/n/n2f8ca72c0f61](https://note.com/norito_hiraoka/n/n2f8ca72c0f61)
44. <https://shift-ai.co.jp/blog/17419/>
45. <https://www.youtube.com/watch?v=Llj1PorbB6s>
46. [https://www.reddit.com/r/singularity/comments/1p0956s/gemini\\_3\\_benchmarks/](https://www.reddit.com/r/singularity/comments/1p0956s/gemini_3_benchmarks/)
47. <https://www.youtube.com/watch?v=KfCoibaBTnE>
48. <https://www.knowledgeful.ai/plus/rabbit-ai-hallucination/>
49. [https://www.gizmodo.jp/2025/11/google\\_gemini\\_3\\_released.html](https://www.gizmodo.jp/2025/11/google_gemini_3_released.html)
50. <https://forest.watch.impress.co.jp/docs/news/2064403.html>
51. [https://x.com/Kaito\\_\\_\\_AI/status/1990827209570087188](https://x.com/Kaito___AI/status/1990827209570087188)
52. <https://www.techno-edge.net/article/2025/11/19/4721.html>
53. <https://note.com/chankostin/n/nf391f6409a8d>
54. <https://news.yahoo.co.jp/articles/ed7eb319f576a5761c3251680f023463e4cb26e0>
55. <https://news.livedoor.com/topics/detail/30017552/>
56. <https://deepmind.google/models/gemini/pro/>
57. <https://x.com/NewsFromGoogle/status/1990889759565140289>
58. <https://cloud.google.com/blog/ja/products/ai-machine-learning/gemini-3-is-available-for-enterprise/>
59. <https://romptn.com/article/74984>
60. <https://shift-ai.co.jp/blog/20257/>
61. <https://www.taneyats.com/entry/when-comes-gemini-flash-3>
62. <https://news.livedoor.com/article/detail/30017452/>

63. <https://www.adcal-inc.com/column/gemini-api-price-plan-guide/>
64. [https://www.reddit.com/r/Bard/comments/1oci6ir/last\\_update\\_at\\_this\\_point\\_i\\_dont\\_give\\_a\\_f\\_as\\_long/](https://www.reddit.com/r/Bard/comments/1oci6ir/last_update_at_this_point_i_dont_give_a_f_as_long/)
65. [https://note.com/hiro\\_seki/n/n35ce1b91bf87](https://note.com/hiro_seki/n/n35ce1b91bf87)
66. [https://www.reddit.com/r/singularity/comments/1bbeqjx/claude\\_3\\_vs\\_gemini\\_vs\\_gpt4\\_who\\_can\\_make\\_a\\_mazing/](https://www.reddit.com/r/singularity/comments/1bbeqjx/claude_3_vs_gemini_vs_gpt4_who_can_make_a_mazing/)
67. [https://note.com/r1250\\_gs/n/nee9c7c427f7f](https://note.com/r1250_gs/n/nee9c7c427f7f)
68. <https://biz.moneyforward.com/ai/basic/839/>
69. [https://www.reddit.com/r/Bard/comments/1nx0p2v/holy\\_shit\\_i\\_just\\_got\\_maybe\\_the\\_best\\_web\\_design/](https://www.reddit.com/r/Bard/comments/1nx0p2v/holy_shit_i_just_got_maybe_the_best_web_design/)
70. <https://skywork.ai/blog/ai-agent/chatgpt-5-1-vs-claude-vs-gemini-2025-comparison/>
71. <https://ai.google.dev/gemini-api/docs/gemini-3?hl=ja>
72. [https://search.yahoo.co.jp/realtime/search/matome/bd216b01c6c743c5adb8583141ce0ebe-1763489406?rkf=1&iffr=tl\\_matometl](https://search.yahoo.co.jp/realtime/search/matome/bd216b01c6c743c5adb8583141ce0ebe-1763489406?rkf=1&iffr=tl_matometl)
73. [https://search.yahoo.co.jp/realtime/search?ei=UTF-8&rkf=1&iffr=tl\\_unit&p=gemini+3.0](https://search.yahoo.co.jp/realtime/search?ei=UTF-8&rkf=1&iffr=tl_unit&p=gemini+3.0)
74. [https://www.reddit.com/r/hackernews/comments/1p0bc6d/gemini\\_3\\_pro\\_model\\_card/](https://www.reddit.com/r/hackernews/comments/1p0bc6d/gemini_3_pro_model_card/)
75. <https://zenn.dev/hai08320/articles/be4a445aa33f3d>
76. [https://x.com/ygr\\_kabu/status/1989444292574941561](https://x.com/ygr_kabu/status/1989444292574941561)
77. <https://news.ycombinator.com/item?id=45970519>
78. <https://www.technologyreview.jp/s/372628/googles-new-gemini-3-vibe-codes-responses-and-comes-with-its-own-agent/>
79. [https://search.yahoo.co.jp/realtime/search?rkf=1&p=Gemini+3.0&btid=MTk1MzQ5NTE3MTE2NzA4NDc3OA%3D%3D&iffr=tp\\_bz](https://search.yahoo.co.jp/realtime/search?rkf=1&p=Gemini+3.0&btid=MTk1MzQ5NTE3MTE2NzA4NDc3OA%3D%3D&iffr=tp_bz)
80. [https://www.reddit.com/r/hackernews/comments/1p0ffef/gemini\\_3\\_pro\\_preview\\_live\\_in\\_ai\\_studio/](https://www.reddit.com/r/hackernews/comments/1p0ffef/gemini_3_pro_preview_live_in_ai_studio/)
81. <https://biz.moneyforward.com/ai/basic/378/>
82. <https://developers.googleblog.com/en/introducing-gemini-2-5-flash-image/>
83. [https://note.com/ryuya\\_330/n/n8c77cbe63556](https://note.com/ryuya_330/n/n8c77cbe63556)
84. <https://www.swebench.com>
85. <https://x.com/jaguring1/status/1990846927936934031>
86. <https://cloud.google.com/blog/ja/products/ai-machine-learning/vertex-ai-context-caching>
87. <https://news.yahoo.co.jp/articles/efa50309310a1c49a6c75281966a78c5556a9395>
88. <https://gihyo.jp/article/2025/10/vibe-coding-ai-app-using-google-ai-studio-build-mode>
89. <https://note.com/mi6242/n/n249aee70c72f>
90. <https://book.st-hakky.com/data-science/gemini-vibecoding-2025-guide>
91. <https://www.watch.impress.co.jp/docs/news/2064371.html>
92. [https://www.reddit.com/r/Bard/comments/1ocoomv/ai\\_studios\\_new\\_vibe\\_coder\\_is\\_ing\\_crazy\\_it\\_has\\_to/](https://www.reddit.com/r/Bard/comments/1ocoomv/ai_studios_new_vibe_coder_is_ing_crazy_it_has_to/)
93. [https://x.com/m2ai\\_jp/status/1990825858047545394](https://x.com/m2ai_jp/status/1990825858047545394)
94. <https://zenn.dev/gatechnologies/articles/4327115bdc0358>
95. <https://onlyone-mgt.jp/2024/06/10945/>
96. <https://blog.scuti.jp/ai-hallucination-root-causes-openai-research-statistical-truth-solutions/>
97. <https://qiita.com/autotaker1984/items/932a7059239d3445731f>

98. <https://ai-blog.am-bition.jp/column/ai-trends/ai-hallucination/>
99. <https://creators.spotify.com/pod/profile/nagaraai/episodes/818-GPT-5Gemini3-0AI-e36u4e5>
100. <https://gigazine.net/news/20250512-ai-hallucination-short-answer/>
101. <https://momo-gpt.com/column/chatgpt5-gemini/>