

Gemini 4 の公開見通しと競争環境

背景と競合動向および今後の評価課題

調査基準日 2026 年 9 月 24 日

GPT-6 Astra

Gemini 4 の早期公開には Google DeepMind 責任者の発言という根拠がある。ただし、正式な公開日や性能が確定した段階ではない。競争上の焦点は、Google が旗艦モデルの更新の遅れを取り戻し、性能、費用、実務での使いやすさをどこまで改善できるかにある。[1][2]

本稿では、公開報道と Google および競合各社の発表を基に、Gemini 4 の開発状況、Google の開発方針、競合関係、今後の展開を整理する。企業の性能説明、第三者評価、本稿の考察は区別して記載する。

1 公開報道で明らかになったこと

Investing.com 日本語版は 2026 年 9 月 24 日、「Google の Gemini 4 AI モデル、まもなくリリースへーDeepMind 責任者が明かす」と題する記事を掲載した。元になったのは、9 月 23 日に開催された The Information の AI Agenda Live Summit での、Google DeepMind の新責任者コレイ・カブクチュオール氏（Koray Kavukcuoglu）の発言である。[1]

記事によれば、Gemini 4 は事後学習の初期段階に入り、同氏は年末よりかなり早い公開を見込んでいる。事後学習は、基礎となるモデルを、指示への応答や安定した作業遂行に適したものへ調整する工程である。この段階への移行は開発の進展を示すが、一般提供の準備がすべて完了したことを意味しない。[1]

取材内容を紹介した THE DECODER の報道では、初期版を早く公開し、その後も短い間隔で改善する意向と、安全対策および評価を進めていることが伝えられている。正式な公開日、API 提供開始日、価格、モデル仕様、独立した性能評価は、これらの発言とは分けて確認する必要がある。[2][3]

2 Google の開発方針と組織の変化

背景には、Gemini 3.5 Pro の公開の遅れがある。Google は 6 月 3 日の投資家向け説明で 6 月中の公開を見込んでいたが、7 月 21 日時点でもパートナーによるテスト中と説明していた。9 月 24 日の報道でも未公開とされており、当初予定どおりに上位モデルを投入できていない状況が確認できる。ただし、3.5 Pro の正式な開発中止までは確認されていない。[4][5][6]

Google はその間も開発を続けている。7 月 21 日には、Gemini 4 に向けて同社として最も野心的な事前学習を開始したと公表した。同時に Flash 系を短期間で更新し、9 月 2 日には Gemini 3.8 Flash を

発表している。同社は、長時間のコーディング、エージェントによる作業、複数段階の推論の改善を説明している。[5][7]

これらの発表からは、低価格モデルを継続的に改善しながら、次の旗艦モデルを準備してきた構図が読み取れる。本稿では、Gemini 4 をその旗艦モデル更新の成否を左右する節目と捉える。Flash 系での改善が Gemini 4 にどの程度反映されるかは、公開後の検証課題となる。

開発体制も変わった。Google の公式発表では、デミス・ハサビス氏が DeepMind 会長兼 Alphabet チーフサイエンティストとなり、カブクチュオール氏が Gemini モデル、最先端 AI 研究、アプリ、開発者向けチームを統括する体制になっている。[8]

本稿の考察としては、研究成果を製品に反映する速度と、組織としての実行力を高める狙いを持つ変更と解釈できる。公式発表では AGI や科学研究の継続も明示されているため、この変更を Google が基礎研究や AGI 研究を放棄したことと同一視すべきではない。[8]

3 競合各社の動きと Gemini 4 への影響

Gemini 4 が追いかける競争相手も更新を続けている。9 月の主要な発表では、高性能化に加えて、価格低下、長時間の作業、利用者との接点の強化が目立つ。表 1 の右列は、各社の発表を踏まえた本稿の考察である。

表 1 主要な競合動向と競争上の意味

企業	確認できる動き	Gemini 4 への競争圧力
OpenAI	GPT-6 Astra に続き、9 月 22 日に Sol と Luna を公開。高度な業務をより低価格で処理する方向を強化。[9][10][11]	旗艦モデルの性能に加え、日常業務の費用対効果でも競争する必要がある。
Anthropic	9 月 22 日に Claude Opus 5.5 を公開。多くの仕事で Fable 5.1 水準、典型的な処理費用は Opus 5 より 40%低いと説明。[12]	コーディング、調査、資料作成を少ない手戻りで完了する能力が問われる。
Meta	9 月 8 日に個人向けエージェント Muse を発表。9 月 23 日には AI グラスへの展開も発表。[13][14]	AI を日常的に使う入口と、個人向けエージェントの体験をめぐる競争が強まる。
SpaceX AI Grok	9 月 21 日に Grok 4.7 を公開。長時間のコーディング、自己検証、知識業務を強化したと説明。[15]	開発者や業務利用者に対する価格と実行能力の競争が広がる。
Xiaomi 中国勢の一例	MiMo-V2.6-Pro-RL のモデル重みを公開。自社環境で動かせる選択肢を提供。[16]	API 利用に加えて、自社運用やカスタマイズとの比較でも選ばれる理由が必要になる。

この競争環境では、Gemini 4 が一部の試験で高得点を記録するだけでは、利用者の選択を十分に説明できない。必要な仕事をどの程度の時間と費用で完了できるか、失敗時に修正できるか、普段使う製品の中で利用できるかが、採用を左右すると考えられる。

4 現行モデルの性能と費用の比較

調査時点の Artificial Analysis Intelligence Index v4.3.2 では、表 2 の値が確認できた。各行は表示された推論設定での評価であり、同じ費用を投入した比較ではない。また、Gemini 4 の性能を表すものでもない。[17][18]

表 2 第三者評価の比較 2026 年 9 月 24 日参照

モデル	評価設定	Intelligence Index
Claude Opus 5.5	Max 標準フォールバックあり	58
GPT-6 Astra	Max	53
Gemini 3.8 Flash	High	41
Gemini 4	未公表	評価値未確認

出典 文献[17][18]。Gemini 4 の提供状況は文献[2]を参照。指標の版が異なる過去の点数とは、そのまま比較できない。Claude の「標準フォールバックあり」は、対象となる処理で別モデルに切り替わる運用条件を含む評価を指す。[12][17]

この結果は、現行の Gemini 3.8 Flash に総合性能で追い上げる余地があることを示す。ただし、Google の現行モデルには価格面の強みがある。通常の API 単価は 100 万トークン当たり入力 0.75 ドル、出力 3.75 ドルであり、Astra の入力 10 ドル、出力 50 ドルより低く設定されている。[7][18]

実際の作業費用は、トークン単価に加え、出力量、推論量、再試行、キャッシュ利用、人による訂正の量によって変わる。Gemini 4 を評価する際も、性能と単価の表に加えて、同じ仕事を完了するまでの総費用を比較する必要がある。

5 Google が持つ普及基盤

Google の第 2 四半期決算説明によれば、Gemini アプリの月間利用者は 9 億 5,000 万人に達しており、Search、Google Workspace、Cloud などにも Gemini を組み込んでいる。また、独自の TPU を含む計算基盤を持つ。[19]

この基盤を踏まえると、Gemini 4 が競争力のある性能に到達した場合、既存製品への展開によって利用を拡大できる可能性がある。モデル単体の順位に加え、企業の業務環境や日常の検索に自然に組み

込めることが、Google の競争力になると考えられる。なお、Gemini 4 が各製品へ搭載される時期は、個別の正式発表で確認する必要がある。

6 今後の展開と確認すべき点

今後の展開として、次の三つが考えられる。いずれも確定情報ではなく、公開された発言と競争環境を踏まえた見通しである。

第一に、初期版を早期に公開し、その後も継続的に改善する展開である。これは報じられた責任者の発言に沿う。この場合、初回公開版だけで評価を固定せず、性能改善、API 提供、利用制限の変化を追う必要がある。[3]

第二に、競合に近い性能を、価格や製品統合によって普及させる展開である。すべての評価で首位を取らなくても、Google の業務環境で使いやすく、処理費用が低ければ、有力な選択肢になる。これは現行 Flash 系の価格方針と Google の普及基盤を踏まえた考察である。[7][19]

第三に、品質や安全性の調整に時間がかかる展開である。事後学習の初期段階と報じられていることから、公開延期や対象者を限定した提供も想定に残る。発表日、一般利用開始日、API 提供開始日、企業向け提供開始日を分けて確認すべきである。[1][3]

安全性も公開条件を左右する要因となる。Anthropic のダリオ・アモデイ氏は9月の論考で、能力向上の速度に安全対策が追いつくよう、開発の進め方を調整すべきだと提案した。Google も現行の Flash Cyber を信頼された防御側の利用者に限定している。Gemini 4 についても、内部で達成した性能と、一般利用者が利用できる性能や機能の範囲を区別して確認する必要がある。[20][7]

7 知財業務での評価の観点

知財調査や技術分析に適用する場合、公開後に次の観点で同一課題を用いた比較を行うことを提案する。これは Gemini 4 が当該能力を備えると断定するものではなく、導入判断のための評価項目である。

- 長い明細書と図面や表を横断して、技術的な意味と数値条件を正確に読み取れるか。
- 請求項の構成要件と根拠箇所を対応づけ、確認できた内容と推測を分けて記載できるか。
- 検索、比較、報告書作成までの作業を、根拠の追跡可能性を維持して完了できるか。
- API 料金、処理時間、人が訂正する時間を含めた総費用が、他モデルより改善するか。

Gemini 4 の実務評価では、ベンチマーク順位に加えて、根拠を保ったまま任せられる仕事がどれだけ増えるかを確認することが重要である。

参考文献

文中の番号は以下の文献番号に対応する。閲覧日はすべて 2026 年 9 月 24 日。更新型ページの数値や提供状況は参照時点の内容である。

- [1] Ambar Warrick 編. Google の Gemini 4 AI モデル、まもなくリリースへ——DeepMind 責任者が明かす. Investing.com 日本語版. 2026 年 9 月 24 日.
<https://jp.investing.com/news/stock-market-news/article-93CH-1690147>
- [2] Google DeepMind. Models. Google DeepMind. 更新型ページ.
<https://deepmind.google/models/>
- [3] Maximilian Schreiner. Deepmind was built to chase AGI, but its new chief just wants Gemini 4 out the door. THE DECODER. 2026 年 9 月 24 日.
<https://the-decoder.com/deepmind-was-built-to-chase-agi-but-its-new-chief-just-wants-gemini-4-out-the-door/>
- [4] Google. Alphabet investor presentation June 2026. The Keyword. 2026 年 6 月 3 日.
<https://blog.google/alphabet/investor-presentation-june-2026/>
- [5] Tulsee Doshi. Introducing Gemini 3.6 Flash, 3.5 Flash-Lite, and 3.5 Flash Cyber. The Keyword. 2026 年 7 月 21 日.
<https://blog.google/innovation-and-ai/models-and-research/gemini-models/gemini-3-6-flash-3-5-flash-lite-3-5-flash-cyber/>
- [6] ITmedia. 「Gemini 4」は年末までに発表——開発責任者が言及か 「Google は確かに最先端」と主張. ITmedia AI+. 2026 年 9 月 24 日.
<https://www.itmedia.co.jp/aipplus/article/2609/24/2000001701/>
- [7] Tulsee Doshi and Raluca Ada Popa. Introducing Gemini 3.8 Flash and 3.8 Flash Cyber. The Keyword. 2026 年 9 月 2 日.
<https://blog.google/innovation-and-ai/models-and-research/gemini-models/3-8-flash-and-3-8-flash-cyber/>
- [8] Sundar Pichai and Demis Hassabis. The next chapter of our AI momentum. The Keyword. 公開日の表示なし.
<https://blog.google/company-news/inside-google/message-ceo/next-chapter-ai-momentum/>
- [9] OpenAI. GPT-6 Astra: A new generation of intelligence. OpenAI. 2026 年 9 月 22 日更新の記載あり.
<https://openai.com/index/gpt-6-astra/>
- [10] OpenAI. Changelog September 2026. OpenAI API Documentation. 2026 年 9 月 22 日の項目.
<https://developers.openai.com/api/docs/changelog>
- [11] OpenAI. Introducing GPT-6 Sol and Luna. OpenAI. 公開日は文献[10]のリリース記録を参照.
<https://openai.com/index/introducing-gpt-6-sol-and-luna/>
- [12] Anthropic. Introducing Claude Opus 5.5. Anthropic. 2026 年 9 月 22 日.
<https://www.anthropic.com/claude-opus-5-5>
- [13] Meta. Introducing Muse: The World’s First Personal AI Agent Built for Everyone. Meta Newsroom. 2026 年 9 月 8 日.
<https://about.fb.com/news/2026/09/introducing-muse-personal-ai-agent/>
- [14] Meta. Introducing Ray-Ban Meta Audio and More AI Glasses Styles. Meta Newsroom. 2026 年 9 月 23 日.
<https://about.fb.com/news/2026/09/introducing-ray-ban-meta-audio-glasses-new-styles-plus-muse/>
- [15] SpaceXAI. Introducing Grok 4.7. SpaceXAI. 2026 年 9 月 21 日.
<https://x.ai/news/grok-4-7>

- [16] Xiaomi MiMo Team. MiMo-V2.6-Pro-RL Model card. Hugging Face. 更新型ページ.
<https://huggingface.co/XiaomiMiMo/MiMo-V2.6-Pro-RL>
- [17] Artificial Analysis. Claude Opus 5.5 (Adaptive Reasoning, Max Effort, Default Fallback) vs Gemini 3.8 Flash (high): Model Comparison. Artificial Analysis. 更新型ページ Intelligence Index v4.3.2.
<https://artificialanalysis.ai/models/comparisons/claude-opus-5-5-vs-gemini-3-8-flash>
- [18] Artificial Analysis. GPT-6 Astra (max) vs Gemini 3.8 Flash (high): Model Comparison. Artificial Analysis. 更新型ページ Intelligence Index v4.3.2.
<https://artificialanalysis.ai/models/comparisons/gpt-6-astra-vs-gemini-3-8-flash>
- [19] Sundar Pichai. Q2 2026 earnings call: Remarks from our CEO. The Keyword. 2026 年第 2 四半期決算説明.
<https://blog.google/company-news/inside-google/message-ceo/alphabet-earnings-q2-2026/>
- [20] Dario Amodei. We Must Pace the Frontier. Dario Amodei 公式サイト. 2026 年 9 月.
<https://darioamodei.com/post/we-must-pace-the-frontier>