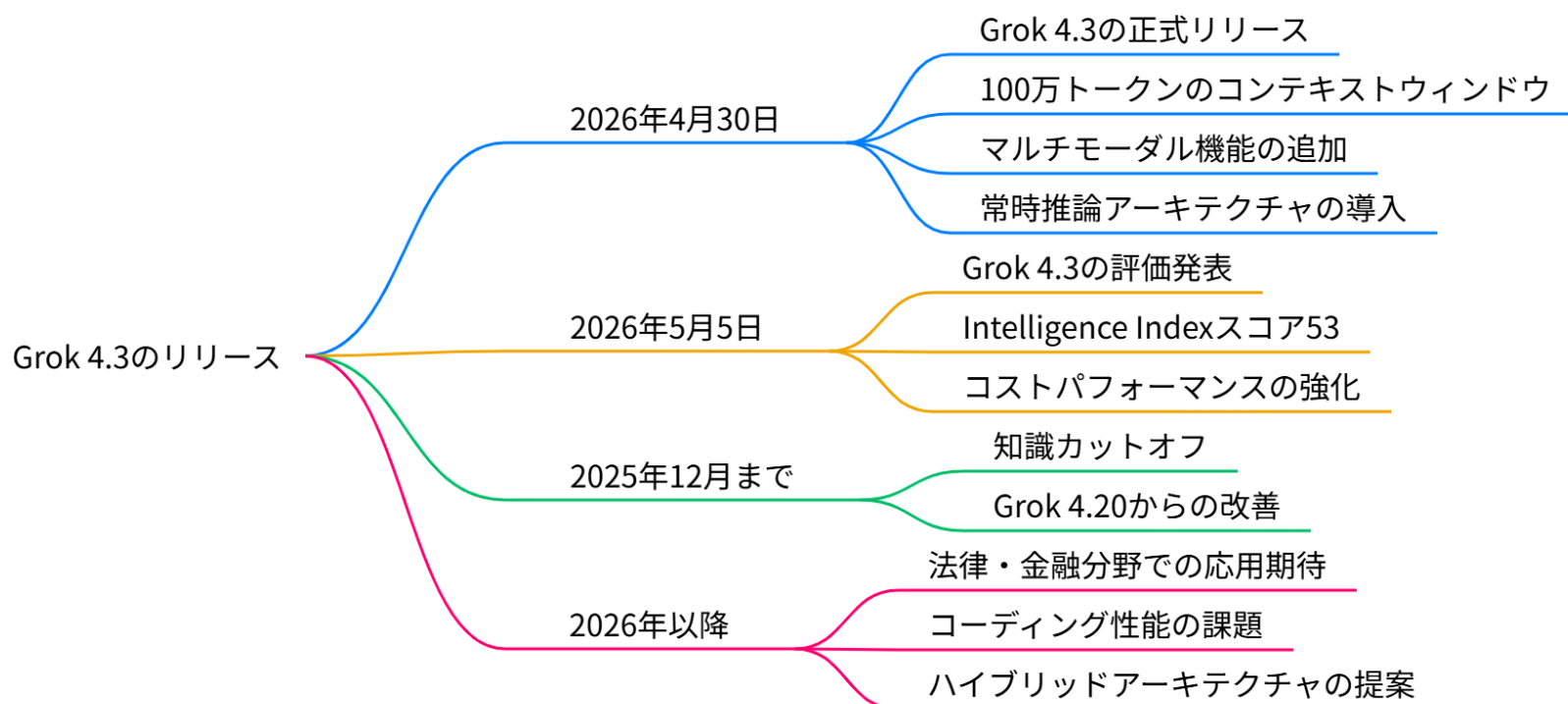


# 「Grok 4.3」の内容、評価

Felo AI



## 概要

xAIが2026年4月30日に発表した大規模言語モデル「Grok 4.3」は、業界最高水準のコストパフォーマンスを武器に、AI開発市場における勢力図を塗り替える可能性を秘めたリリースです<sup>6 14</sup>。本モデルは、100万トークンの広大なコンテキストウィンドウ、画像・動画に対応したネイティブマルチモーダル機能、そして常時オンの推論アーキテクチャを特徴としています<sup>7 25</sup>。独立評価機関Artificial Analysisによるベンチマークでは、Intelligence Indexスコア53を記録し、OpenAIのGPT-5.5やAnthropicのClaude Opus 4.7といった最先端モデルには及ばないものの、Claude Sonnet 4.6などを上回る性能を示しました<sup>2 6 44</sup>。Grok 4.3の真価はその価格戦略にあり、API価格は前モデルGrok 4.20から入力で約40%、出力で約60%と大幅に引き下げられました<sup>6 14</sup>。これにより、同等性能のモデル群において圧倒的な価格優位性を確立し、AIアプリケーションの開発コストを劇的に下げる選択肢として浮上しています<sup>19</sup>。特に、エージェント的タスクの性能が大幅に向上しており、法律や金融といった専門分野での応用が期待される一方、コーディング性能や一部の自律タスクにおける安定性には課題も指摘されています<sup>1 9 22</sup>

## 詳細レポート

### Grok 4.3のリリース概要と主要機能

xAIは2026年4月30日、最新の独自大規模言語モデル「Grok 4.3」をAPI経由で正式にリリースしました[1 6 11](#)。このモデルは、2025年12月までの知識カットオフを持ち、前モデルGrok 4.20のスケールを踏襲しつつ、アーキテクチャを改善したものです<sup>3</sup>。

#### 主な機能と仕様

- **100万トークンのコンテキストウィンドウ**: 長大な文書の読解、詳細なリサーチ、複数ステップを要するエージェントタスクなど、大量の情報を一度に処理する能力が大幅に向上しました<sup>7 11 25</sup>。ただし、20万トークンを超えるリクエストには割増料金が適用される階層型価格設定が採用されています<sup>1 13</sup>。
- **マルチモーダル入力**: テキストと画像入力に加え、新たにネイティブでのビデオ入力に対応しました<sup>5 7</sup>。これにより、画面録画からのコード生成や、チュートリアル動画の要約など、これまで困難だったユースケースが実現可能になります<sup>22</sup>。
- **常時推論アーキテクチャ (Always-on Reasoning)**: 従来のモデルのように推論モードを任意で切り替えるのではなく、すべての応答で内部的な「思考プロセス」を必ず経由する設計です<sup>9 19 21</sup>。これにより、複雑な指示や多段階の論理的思考を要するタスクで安定した出力を実現します<sup>27</sup>。
- **ツール使用と構造化出力**: Web検索、コード実行、ファイル添付といったサーバーサイドツールをAPI経由で呼び出す機能や、JSON形式での構造化出力をサポートしています<sup>1 20</sup>。
- **エンタープライズ対応**: SOC 2 Type II監査、HIPAA適格、GDPR準拠など、規制の厳しい業界での利用を想定したコンプライアンス基準を満たしています<sup>1</sup>。

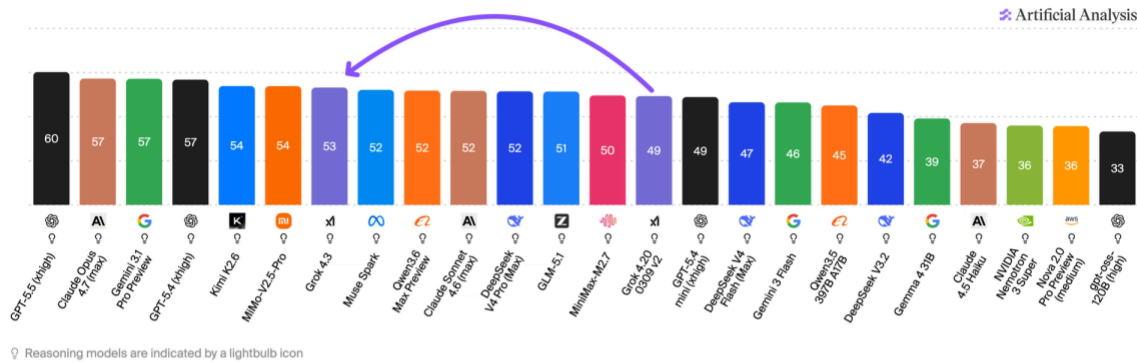
#### 性能ベンチマークと第三者評価

Grok 4.3は、性能とコストのバランスを重視したモデルとして位置づけられています。

**Artificial Analysis**による評価 独立評価機関Artificial Analysisの評価では、Grok 4.3はIntelligence Indexでスコア53を記録しました<sup>2 6</sup>。これは、Grok 4.20(スコア49)から4ポイント向上し、Muse SparkやClaude Sonnet 4.6を上回る結果です<sup>6 26 44</sup>。しかし、市場のリーダーであるOpenAIのGPT-5.5やAnthropicのClaude Opus 4.7には依然として差をつけられています<sup>6</sup>。

## Artificial Analysis Intelligence Index

Artificial Analysis Intelligence Index v4.0 incorporates 10 evaluations: GDPval-AA, r<sup>2</sup>-Bench Telecom, Terminal-Bench Hard, SciCode, AA-LCR, AA-Omniscience, IFBench, Humanity's Last Exam, GPQA Diamond, CritP

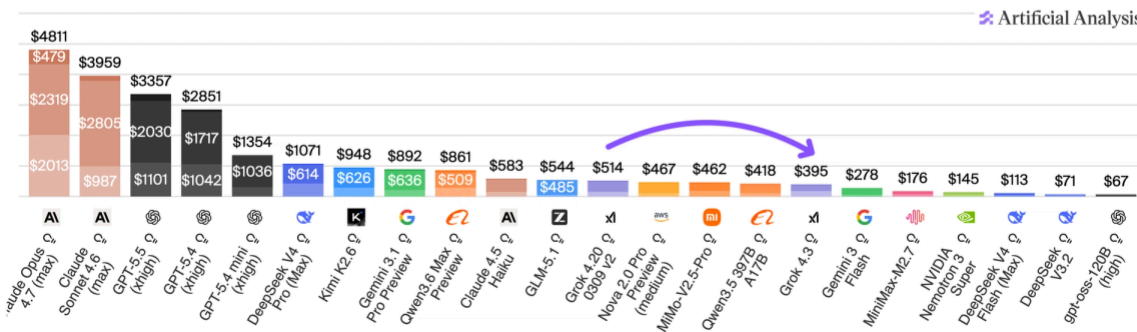


Reasoning models are indicated by a lightbulb icon

## Cost to Run Artificial Analysis Intelligence Index

Cost (USD) to run all evaluations in the Artificial Analysis Intelligence Index

Input Cost Reasoning Cost Output Cost



特に顕著な改善が見られたのは、実世界のタスクをシミュレートするエージェント性能評価ベンチマーク「GDPval-AA」です<sup>1</sup>。Grok 4.3はこのベンチマークでEloレーティング1500を達成し、前モデルの1179から321ポイントという飛躍的な向上を見せました<sup>6</sup>。これにより、Gemini 3.1 ProやGPT-5.4 miniといった競合を上回るエージェント性能を獲得しています<sup>16</sup>。

### 分野別の性能

- 強み(法律・金融): Vals AIなどの独立機関による評価では、判例推論やコーポレートファイナンスといった専門領域でトップクラスの性能を記録しており、契約書レビューや財務分析などの専門業務に適していると評価されています<sup>9 21 27</sup>。この高い性能は、「常時推論アーキテクチャ」が法律や金融の持つ密な論理構造の処理に適していることを示唆しています<sup>1</sup>。
- 弱み(コーディング): コーディング能力に関しては、主要なベンチマークであるSWE-benchにおいて、最先端のClaude Opus 4.7に約14パーセントポイントの差をつけられています<sup>7 22</sup>。このため、Grok 4.3は単純なコード生成には利用できても、複雑で正確性が求められる開発タスクの主役にはなり得ないのが現状です<sup>4</sup>。

潜在的な課題 Andon Labsが実施した長期的な自律エージェントタスクのベンチマーク「Vending-Bench」では、Grok 4.3が行動を起こさずに数日間アイドル状態になる「ナルコレプシー(睡眠発作)」のような現象が報告されました<sup>19</sup>。これはGrok 4.20では

見られなかった回帰であり、「常時推論」アーキテクチャが原因で、モデルが自らの推論によって行動不能に陥る可能性を示唆しています19。

## 革新的な価格戦略と市場ポジショニング

Grok 4.3の最も注目すべき点は、その攻撃的な価格設定です。xAIは性能でトップを争うのではなく、「十分な性能」を「圧倒的な低価格」で提供することで、市場のボリュームゾーンを獲得する戦略を明確に打ち出しています19。

**API価格** Grok 4.3のAPI価格は、前モデルGrok 4.20から大幅に引き下げられました7。

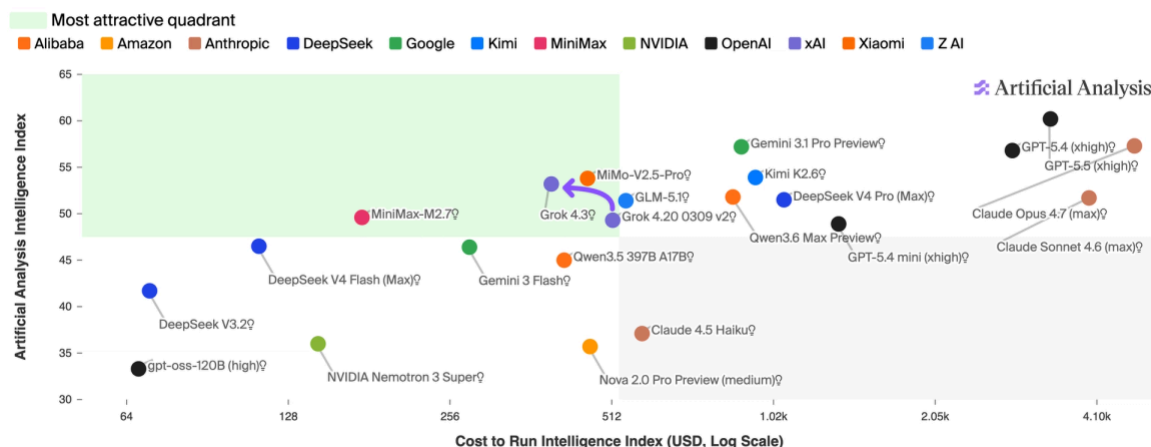
項目	Grok 4.3	Grok 4.20 (旧価格)	価格削減率
入力トークン (1Mあたり)	\$1.25	2.00 - 3.00	~37.5% - 58%
出力トークン (1Mあたり)	\$2.50	6.00 - 15.00	~58.3% - 83%
キャッシュされた入力トークン (1Mあたり)	0.20 - 0.31	-	新設

\*出典: 1, 7, 13, 25 \*

この価格設定により、Grok 4.3は「価格対インテリジェンス」のパレートフロンティア(効率的な選択肢の集合)上に位置づけられています6 19。

### Intelligence vs. Cost to Run Artificial Analysis Intelligence Index

Artificial Analysis Intelligence Index; Cost to Run Intelligence Index



競合とのコスト比較 Grok 4.3のコストパフォーマンスは、競合のフラッグシップモデルと比較するとさらに際立ちます。100万トークン規模のコンテキストと高度な推論能力を組み合わせたタスクにおいて、Grok 4.3のコストはClaude Opus 4.7の約12分の1にまで抑えられています<sup>7</sup>。この価格破壊は、これまで「高価である」ことが常識だった推論モデル市場の前提を覆すものです<sup>7</sup>。

ユニークな料金体系 xAIは、トークン料金以外にも特徴的な課金体系を導入しています<sup>1</sup>。

- ツール使用料: Web検索やコード実行などのツールを呼び出す際に、1,000コールあたり\$5.00といった固定料金が発生します<sup>1</sup>。
- 安全性フィルター違反料金: 安全性フィルターによって生成前にブロックされたリクエストに対し、1件あたり\$0.05の料金を課します。これは業界でも前例のない試みです<sup>1</sup>。

## 開発者と企業への影響

Grok 4.3の登場は、開発者や企業がAIモデルを選定する際の基準に大きな影響を与えます。

ユースケース別モデル選定 どのモデルを選択すべきかは、アプリケーションの要件によって異なります<sup>4</sup>。

- **Grok 4.3**が最適なケース:
  - コスト重視のアプリケーション: 大量のテキスト処理やエージェント実行を低コストで行いたい場合<sup>1</sup>。
  - リアルタイム情報: X (旧Twitter) との連携やWeb検索ツールを活用し、最新情報に基づいた応答が必要な場合<sup>4</sup>。
  - 長文コンテキスト処理: 法律文書のレビューや大規模なコードベースの分析など、100万トークンのコンテキストウィンドウが活きるタスク<sup>1</sup>。
  - ビデオ分析: ネイティブのビデオ入力機能を活用したアプリケーション開発<sup>7 22</sup>。
- **GPT-5.5 / Claude Opus 4.7**が最適なケース:
  - 高精度なコーディング: 失敗のコストが高い、複雑なコーディングエージェントの開発<sup>4 22</sup>。
  - 最高の性能が求められるタスク: わずかな性能差がビジネス成果に直結する、ミッションクリティカルなワークフロー<sup>4</sup>。
  - 既存エコシステムへの依存: すでにOpenAIやAnthropicのAPI、ツール群と深く統合されたシステム<sup>4</sup>。

ハイブリッドアーキテクチャによるコスト最適化 多くの企業にとって最適な解は、単一のモデルに依存するのではなく、タスクの特性に応じて複数のモデルを使い分けるハイブリッドアーキテクチャです<sup>22</sup>。例えば、日常的な単純なタスクはGrok 4.3やさらに軽量のモデルにルーティングし、複雑で高い正確性が求められるタスクのみをClaude Opus 4.7に送ることで、クリティカルなタスクの品質を維持しつつ、全体の運用コストを80%以上削減することも可能です<sup>22</sup>。

The banner features the Stella AI logo on the left, followed by the text 'Stella AI UPDATE' and 'xAI社の最新モデル Grok 4.3に対応'. On the right, a smartphone displays the Stella AI interface with a chat conversation and a grid of AI-powered tasks. At the bottom, a dark bar lists '利用可能AIモデル' (Available AI Models) including Grok 4.3, Grok 4.1, Fast GPT-5.4, GPT-5.4 nano, Gemini 3.1 Flash-Lite, Gemini 3.1 Pro, Claude Sonnet 4.6, GPT-Image-2, Nano Banana Pro, and Nano Banana 2.

すでに日本の株式会社SUPERNOVAが自社の生成AIサービス「Stella AI」でGrok 4.3の提供を開始するなど、国内外で迅速な採用が進んでいます<sup>9 21</sup>。Grok 4.3は、AI開発における「コスト」という最大の障壁の一つを取り払い、より多くの開発者が高度なAI機能をアプリケーションに組み込むことを可能にする、重要な一歩と言えるでしょう<sup>7</sup>。

1. [xAI launches Grok 4.3 at an aggressively low price and a ...](#)
2. [Grok 4.3 – Intelligence, Performance & Price Analysis](#)
3. [Grok Release Notes](#)
4. [Grok 4.3 vs Claude Opus 4.7 vs GPT–5.5: Which Should You Test ...](#)
5. [xAI Launches Grok 4.3 Beta with Enhanced Features](#)
6. [xAI launches Grok 4.3 with improved agentic performance and lower ...](#)
7. [Grok 4.3 API major release: 5 major upgrades + 40% price ...](#)
8. [xAI has launched Grok 4.3, achieving 53 on the Artificial Analysis ...](#)
9. [Stella AI、xAI社の最新AIモデル「Grok 4.3」に対応](#)
10. [Grok 4.3 \(Non–reasoning\) vs Claude Opus 4.7 ... – Artificial Analysis](#)
11. [Grok 4.3](#)
12. [Grok 4.3 \(Non–reasoning\) vs GLM–4.6 \(Non–reasoning\)](#)
13. [Grok 4.3](#)
14. [Grok 4.3 just dropped 53 on the Artificial Analysis Intelligence Index](#)
15. [Grok 4.3 betaが静かにやってきた——xAIの足音...](#)
16. [Grok 5 vs GPT–5.5 vs Claude Opus 4.7: Can a 10 Trillion Parameter ...](#)
17. [xAI Releases Grok 4.3 — Weekly AI Newsletter \(May 4th ...](#)

18. [Grok 4.3 \(Non-reasoning\) vs GPT-5 \(low\): Model Comparison](#)
19. [xAI's Grok 4.3 strategy: Selling speed and price because ...](#)
20. [xAI: Models Intelligence, Performance & Price — Artificial Analysis](#)
21. [Stella AI, xAI社の最新AIモデル「Grok 4.3」に対応 \(PR TIMES\)](#)
22. [Grok 4.3 vs Claude Opus 4.7 Programming Comparison](#)
23. [7 Models in Training Now, Grok 5 at 10 Trillion Parameters ...](#)
24. [Grok 4.3 \(Non-reasoning\) API Provider Benchmarking & Analysis](#)
25. [grok-4.3 API — Pricing, Benchmarks & Specs](#)
26. [xAI launches Grok 4.3 with improved agentic performance and lower ...](#)
27. [Stella AI, xAI社の最新AIモデル「Grok 4.3」に対応](#)
28. [Grok 4.3 vs Claude Opus 4.7: Aggressive Pricing vs Top Tier ...](#)
29. [Grok 4.3 is available via API with 1M token context and...](#)
30. [Grok 4.3 vs GPT-5.5 \(xhigh\): Model Comparison — Artificial Analysis](#)
31. [Grok 4.3 betaが静かにやってきた——xAIの足音...](#)
32. [Grok 4.3 \(Non-reasoning\) vs o3: Model Comparison — Artificial Analysis](#)
33. [xAI: Grok 4.3 — API Pricing & Benchmarks](#)
34. [The difference between Codex GPT 5.5 and Opus 4.7 in a nutshell](#)
35. [テスラオーナーの何%が2025.26.4 Grokアップデートを...](#)
36. [I Tested The Grok 4.3 Update And The Results Surprised Me — Reddit](#)
37. [xAI's Grok 4.3 Launch Highlights Pricing Strategy and ...](#)
38. [xAI drops Grok 4.3 with steep price cuts and an Imagine agent mode ...](#)
39. [Grok 4.3 vs GPT-5.5 — Detailed Performance & Feature Comparison](#)
40. [Grok 4.3 got to the 7th spot on the Artificial Analysis Index ...](#)
41. [Grok 4.3 \(Non-reasoning\) Intelligence, Performance & Price Analysis](#)
42. [Open AI GPT 5.5 Spud vs Anthropic Claude Code Opus 4.7 vs Grok ...](#)
43. [xAI: Grok 4.3 — Benchmarks — OpenRouter](#)
44. [Ali — JUST IN: Grok 4.3 from xAI scores 53. It ranks above ... — Facebook](#)