

GPT-6 Astra 実力診断レポート：実務活用のための「強み」と「リスク」

2026年9月にリリースされた「GPT-6 Astra」は、特定の高度なタスクで圧倒的な性能を示す一方、コスト面や透明性に課題を残しています。
「全方位最強」というよりは、用途を見極めて活用すべき「尖った」モデルであるのが実像です。

圧倒的な性能と効率性



コンピュータ操作・エージェント能力の飛躍

GPT-5.6 Sol比でタスク処理時間を約47%短縮し、自律的な長時間タスク実行が可能。



高度な数学・科学・コーディング性能

数学ベンチマーク(FrontierMath)で97.6%を達成。
複雑なコードのバグ検出力も大幅に向上。



劇的なトークン効率の改善

同等タスクで前世代の約1/3のトークン消費を実現し、実質的な処理効率が向上。



主要モデルとのベンチマーク・コスト比較

	GPT-5.6 Astra	GPT-5.6 Sol	Claude Fable 5.1
数学 (FrontierMath)	97.6%	確認不可	87.8%
汎用知能指数 (AA Index)	60.9	61.6	65.7
API価格 (入力1M/出力1M)	\$10 / \$50	\$4 / \$20	\$10 / \$50

導入前に考慮すべき懸念点

2.5倍



API価格の増大と厳しい制限

API価格は前世代比2.5倍。利用制限も厳しく、単一ベンダー依存のリスクが顕在化。



Claude Fable 5.1 GPT-6 Astra

「全方位最強」ではない第三者評価

汎用知能指数ではClaude Fable 5.1に劣後する項目もあり、期待値との乖離が指摘される。



安全性と透明性のリスク

サイバー攻撃能力が初の「Critical」評価。内部推論のブラックボックス化により監視性が低下。