フロンティアLLM第三者評価比較：Claude Fable 5.2 vs GPT-6.2 Astra

2026年11月の主要5評価機関データを集計。総合力で首位のClaude Fable 5.2と、科学・数学領域および圧倒的なコストパフォーマンスで強みのGPT-6.2 Astraの対比構造を整理。

Claude Fable 5.2



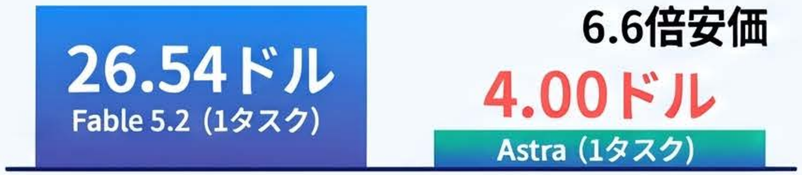
総合評価とコスト効率

総合指標ではClaude Fable 5.2が優位

多くのベンチマークで首位を獲得し、実務タスクでも高い適応力を発揮。



1タスクコストはAstraが約6.6倍安価





Epoch ECIでは両モデルが同点1位

総合スコア174で並ぶが、実際の業務タスク評価では適応差が存在。

GPT-6.2 Astra



分野別強みと実力差

Astraは科学・数学・抽象推論で最上位

ARC-AGI-3(Adapter)で100%、Vals Scienceで1位を記録。



Fable 5.2は知識・実務・法務で圧勝

事実知識(SimpleQA)や実務自動化、法務・税務ベンチマークで優位。

データ比較

主要ベンチマークにおける両主要モデルの具体的スコアと順位の比較

Claude Fable 5.2	GPT-6.2 Astra
1位 (指数 65) Intelligence Index (AA)	2位 (指数 62) Intelligence Index (AA)
1位 (71.31%) Vals Index v2.2	8位 (67.66%) Vals Index v2.2
最高値 (81.17%) ARC-AGI-3 (標準ハーネス)	70.3% ARC-AGI-3 (標準ハーネス)

Gemini Notebook