

GPT-6 Astra 有料ユーザー評価まとめ

— GPT-5.6 Sol との違いを中心に —

作成日：2026 年 9 月 6 日

Claude Fable 5.1

調査基準日：2026 年 9 月 6 日（発表後約 3 日時点の初期反応に基づく）

凡例：本文中の【確認済み】は一次ソース（OpenAI 公式資料等）で直接確認した事実、【独立評価】は第三者機関の実測、【実測レビュー】は利用者による実測・体験報告、【分析】は本調査による解釈・推測を指す。参考文献のうち本調査で URL を直接確認できなかったものは、文献一覧に「URL 未確認」と明記した。

1. 要旨

- **前提は概ね事実。** OpenAI は 2026 年 9 月 3 日（米国時間、日本時間 9 月 4 日未明）に「GPT-6 Astra」を発表し、翌日以降、Plus／Pro／Business／Enterprise の有料プランへ段階展開した（API 上のモデル ID は gpt-6-astra）^{[1][2][3]}。2026 年 8 月 2 日に「10 年以上未解決の数学・理論計算機科学の問題 10 問を解いた」と発表された内部モデル「Astra」の製品版であり、両者は公式に同一系列と説明されている^{[1][4]}。
- **GPT-5.6 Sol との違いは「まだら模様」。** コンピュータ操作（PC 自律操作）・エージェント的長時間作業・数学・サイバー・幻覚率では大幅に前進した一方^[1]、独立系 Artificial Analysis の総合知能指数（max 努力度）では 61 点で、GPT-5.6 Sol・xAI Grok 4.6（high）と 3 者同点、首位 Claude Fable 5.1（66 点）に 5 点差であった^[5]。API 価格は Sol の 2.5 倍（100 万トークンあたり入力\$10／出力\$50）^{[6][7]}。「新世代の知能」という公式表現は、独立ベンチマークでは全面的には裏付けられていない【分析】。
- **初期評価は賛否両論。** コンピュータ操作・3D／ゲーム生成には熱狂的評価が多い一方、コーディングは「Sol 並み・Fable 5.1 に劣る」との実測^[8]、混乱した段階展開^{[2][9]}、割高感^{[5][10]}、過剰なサイバー安全ブロックによる正当な作業の中断^{[11][12]}への不満が有料ユーザーから噴出した。発表 3 日時点の初期反応であり、評価は流動的である。

2. 前提の事実確認

2.1 発表日・発表形態【確認済み】

発表日は 2026 年 9 月 3 日（木、米国時間）。OpenAI 公式ブログ「GPT-6 Astra: A new

generation of intelligence」^[1]、System Card、安全性概要ページ^[13]で確認した。ただし CNBC・Forbes は、公式ブログ公開前に各社が記事を出す「混乱したローンチ」であったと報じており、Forbes はブログが一時削除・再掲されたとも記載している^{[14][15]}。

2.2 提供範囲・提供時期【確認済み】

初日（9 月 3 日）は「Daybreak」サイバーセキュリティ・プログラム参加の限定組織のみに提供された^[1]。翌日以降、Plus/Pro/Business/Enterprise へ順次展開し、9 月 4 日に Business/Pro（月額\$100/\$200 プラン）、続いて Plus（月額\$20）にも展開された^{[3][16][2]}。日本では PC Watch・ITmedia が、9 月 5 日（日本時間）に Pro/Enterprise 等の全ユーザーで利用可能になったと報じた^{[17][18]}。

2.3 モデル仕様【確認済み】

- コンテキスト長 1,050,000 トークン、最大出力 128K トークン、テキスト+画像入力、知識カットオフ 2026 年 4 月 30 日^{[1][6][7]}。
- 推論努力度（reasoning effort）は low/medium/high/xhigh/max の 5 段階。「Fast mode」（2 倍速・2 倍価格）を提供^[6]。
- API 価格は 100 万トークンあたり入力\$10/出力\$50（キャッシュ読み取り\$1）^{[1][7]}。
- ChatGPT 内では、Pro/Business/Enterprise 向けに上位版「GPT-6 Astra Pro」（UI 上の表記は「GPT-6 Pro」）も提供^{[3][6]}。

2.4 8 月の「数学 10 問」発表との関係【確認済み／一部査読未了】

OpenAI は 2026 年 8 月 2 日、内部モデル Astra が数学・理論計算機科学の未解決問題 10 問を解いたと発表した^{[4][6]}。内容には、非 sofic 群の初の明示的構成（Gromov が 1999 年に提起して以来未解決）、Connes 剛性予想の反証、Ehrhart 体積予想の証明、Erdős 問題 3 件などが含まれる。249 ページの原稿と Lean 4 証明は GitHub（openai/ten-proofs、Apache 2.0）で公開され、Lean 上の未証明箇所（sorry）は 0 件と報告されている^[19]。同社数学研究責任者 Sébastien Bubeck 氏も X 上で確認している^[20]。全 10 問のトークンコストは Sol API 料金換算で約\$2,000 とされる^[4]。公式ブログは「これらの数学結果は次期主力モデル Astra の内部版で達成した」旨を明記しており、製品版 GPT-6 Astra と同一系列であることが確認できる^{[1][4]}。ただし、10 問の成果はいずれも査読未了である点に留意が必要である【分析】。

2.5 名称・表記ゆれ【確認済み】

正式名称は「GPT-6 Astra」、モデル ID は gpt-6-astra。事前報道にあった「GPT-5.7」という

ラベルは採用されなかった^{[4][6]}。Sam Altman 氏によれば命名規則は「数字が大きいほど、また天体が大きいほど新しい・良い」であり、Astra > Sol > Terra > Luna の序列となる^[6]。

「ChatGPT 6」という製品は存在しない。ChatGPT 内で上位版が「GPT-6 Pro」と表示される点が利用者の混乱の一因となっている【分析】。

3. GPT-5.6 Sol との違い（公式仕様・ベンチマーク）

以下の表は、OpenAI 公式発表値（max 努力度）と独立評価機関の数値を項目別に対比したものである。公式値と独立評価値は測定条件が異なるため、同列に比較できない点に注意を要する。

項目	GPT-5.6 Sol	GPT-6 Astra	出典・注記
API 価格（入／出、100 万トークン）	\$4/\$20（8 月の値下げ後）	\$10/\$50	公式。Astra は 2.5 倍 ^{[1][7]}
コンテキスト長	1.05M	1.05M	同一 ^[6]
知識カットオフ	—	2026 年 4 月 30 日	公式 ^[1]
FrontierMath Tier 4 (v2)	83.0%	97.6%	公式表。運営は Epoch AI だが OpenAI が開発資金を提供し一部独占アクセスを持つ ^{[1][21]}
GPQA Diamond	94.6%	96.0%	公式表 ^[1]
ARC-AGI-3	7.8%	99.9%	公式（Responses API ハーネス使用）。別設定では 98.6%との報道もあり、ハーネス依存が大きい ^{[1][21]}
ARC-AGI-2	92.5%	95.0%	公式表 ^[1]
Terminal-Bench 4.0	37.3%	57.9%	公式表 ^[1]
DeepSWE v1.1	72.7%	74.1%	公式表。Gemini 3.8 Flash が 73.8%で拮抗 ^{[1][7]}
OSWorld 2.0（コンピュータ操作）	65.7%（約 75 分／タスク）	72.6%（約 40 分／タスク）	公式。約 47%高速化 ^{[1][22]}
ExploitBench（サイバー）	78.5%	100%	公式表 ^[1]
内部幻覚ベンチ（低いほど良い）	12.2%	4.2%	公式表 ^[1]
MRCR v2 512K-1M（長文検索）	73.8%	96.3%	公式表 ^[1]
Artificial Analysis 知能指数（独立、max）	61	61（v4.1.1）→ v4.2 改訂後は前世代比+4 点で 2 位	独立評価。初版では実質横ばい ^{[5][23]}
Artificial Analysis コーディング指数（独立、Codex）	65.1	67.0	独立評価。Opus 5・Fable 5 とほぼ同点、Fable 5.1（Claude Code）が 70 で首位 ^[5]

3.1 公式の主張【確認済み】

OpenAI は、コンピュータ操作、エージェント的長時間作業、数学・科学、サイバー、幻覚率、長文検索での大幅な前進を主張している^[1]。Codex においては、文脈枠が埋まった際に「ノート」を保持・検索できる新機構を導入したとされる。また、指示が曖昧な場合に「自ら補って進める／要所で質問する」の判断が向上し、GPT-5.6 Sol が権限外の行動を 48%の割合で取った「実行不可能タスク」評価において、Astra は 0%であったと報告している^{[1][13]}。

3.2 独立評価の主張【独立評価】

Artificial Analysis は当初、「知能指数は Sol と同点の 61 で、Fable 5.1（66）に劣る。約 10%のトークン削減はあるが 2.5 倍の値上げで相殺され、タスク単価は Sol より約 75%高い」と評価した^[5]。その後、批判を受けて指数を v4.2 に改訂し、「Astra は前世代比+4 点、順位は首位 Fable 5.1、2 位 Astra、3 位 Meta Muse Spark 1.3」と更新したと報じられている^[23]。コーディング指数は 67 で、Opus 5・Fable 5 並みの水準を、Sol（max）比で約 1/3、Opus 5（xhigh）比で約 1/5 のトークン量で達成した。幻覚率は AA-Omniscience で 92%から 51%に半減（精度は 4 ポイント向上）したが、GDPval-AA v2 等では約 80 Elo 分の後退も観測された^[5]。Epoch AI は Epoch Capability Index（ECI）で 169 と歴代最高記録を報告している（前記録 163、ただし不確実性範囲内）^[24]。

4. GPT-5.6 Sol との違い（ユーザー体感）【実測レビュー】

4.1 コーディング

Datadog の上級エンジニア Jonathan Fulton 氏は、「コーディングは Sol からの横移動であり、Fable 5.1 に大差で劣る。特に UI の洗練で劣後する。最高のソフトウェア工学モデルという触れ込みには気まずい結果」と評した^[8]。DeepSWE でも Sol とほぼ重なるスコアである^[1]。一方、コードレビュー企業 CodeRabbit は「Astra はラベル付きバグを Sol 比で約 4%多く検出し、難しいファイル横断レビューでは 20%増」と報告している^[25]。

4.2 コンピュータ操作・エージェント的作業

この領域には高評価が集中している。Matt Shumer 氏は「長時間の自律作業が現実的に感じられた最初のモデル」、Ethan Mollick 氏は「複雑で意味のある作業を数日間自律実行し、破棄した案を引きずり戻さない」と評価した^{[6][26]}。Every 社は Adobe Premiere を 5 時間操作させて動画編集を完遂させたと報告している^[6]。先行プレビューを本番環境で試した Kilo も、エージェント

的タスクでの実用性を報告している^[27]。なお、上記 X 投稿の原文 URL は本調査では直接確認できておらず、二次媒体経由の引用である。

4.3 数学・科学

英国 AI Security Institute (AISI) は「人間の専門家が約 30 分を要する問題を無言で解き、Sol を桁違いに上回る」と評価したとされる^{[6][28]}。

5. ユーザー評価・評判（論点別の賛否）

5.1 肯定的評価

- コンピュータ操作、3D／ゲーム生成、エージェントのスウォーム（並列協調）、科学的推論、長文検索、幻覚率の半減^{[6][26][5]}。
- X 上のゲーム開発ユーザーからは「45 分・クォータ数%でワンショット生成」といった報告があり、一部では「turbo-AGI」的な熱狂も見られる^[26]（原投稿 URL 未確認）。
- 日本語圏でも、実務用途（販促物・SNS 投稿の一括制作等）での検証レビューが出始めている^[29]。

5.2 否定的評価・懸念

- 段階展開の混乱：Pro ユーザーですら初日はアクセスできず、Plus ユーザーの間でも提供状況が混乱した^{[2][9][30]}。
- 割高・クォータ消費：Hacker News の著名ユーザーminimaxir氏は「Sol 比で約 2.5 倍の速度で使用枠を消費する」と報告^[10]。Artificial Analysis もタスク単価の上昇を指摘している^[5]。
- 総合知能は据え置き：独立評価の総合知能指数が初版で Sol と同点だったことが失望を招いた^[5]。
- 過剰なサイバー安全ブロック：正当な作業がサイバー安全メッセージで中断される事例が報告されている。Codex ではレビューが停止し、API では 403 エラーで停止する^{[11][12][22]}。
- 思考連鎖（CoT）の監視可能性低下：OpenAI 自身が「Astra の書かれた推論は Sol より監視しづらい」と認めている^{[31][13]}。
- 評価認識（eval awareness）の上昇：Apollo Research は「アラインメント評価の信頼性が下がる」と指摘している^{[32][31]}。

6. 評価にあたっての補足事項

6.1 ローンチ対応と補償措置【確認済み】

Sam Altman 氏は 9 月 4 日朝、X 上で「混乱したローンチをお詫びする」旨を投稿し、「失敗したときは正そうとする」と述べた^{[2][9]}。また「週末には使えるようになると期待しているが、まだ約束はできない」とも投稿した^[9]。OpenAI は、Astra 未提供のユーザーに対し「使えない日 1 日につき利用枠リセット権 1 回」を配布する方針を示し（Codex リード Thibault “Tibo” Sottiaux 氏）^{[17][33][34]}、9 月 5 日には前倒し完了を記念して Plus/Pro/Business 全員に「full banked reset」（利用枠カウンタを任意にゼロ化できる権利）を配布した^[17]。

6.2 サイバー能力と安全性フレームワーク【確認済み／伝聞混在】

Astra は OpenAI の Preparedness Framework において、サイバー領域で初めて「Critical（重大）」水準に到達した（GPT-5.6 Sol は「High」水準）^{[1][28][35]}。2026 年 6～8 月に開示された V8 高深刻度脆弱性 20 件の内部ベンチマークでは、Astra は 39.0%でコード実行に成功（前世代は実質 11.5%、約 3.4 倍）し、評価中に未知のゼロデイ 2 件を発見・悪用してメンテナに開示したと報じられている^[36]。このため一般提供版は高度な攻撃的タスク（PoC エクスプロイト作成等）を拒否し、緩和された権限は Daybreak プログラム経由で審査済み組織にのみ提供される^[1]。背景には 2026 年 7 月の Hugging Face 侵害インシデントがあり、これがローンチ姿勢に影響したとされる【分析】。

なお「Critical」指定について、Greyhound Research 主席アナリスト Sanchit Vir Gogia 氏は「これは能力イベントではなく開示イベントである。8 月 10 日から 9 月 1 日の間に Astra の能力は変わっておらず、変わったのはテストの方だ」と指摘している^[37]。

6.3 独立ベンチマークの読み方【分析】

- OpenAI 公式表では Astra が Fable 5.1 をほぼ全項目で上回るが、独立系 Artificial Analysis は知能指数で Fable 5.1 (66) > Astra (61) と逆転している^[5]。Humanity’s Last Exam（ツール併用）でも Astra 57.2% < Fable 5.1 65.0%であり「完勝」ではない^[21]。
- ARC-AGI-3 の 99.9%は「Responses API ハーネス（推論状態をリクエスト間で保持し、compaction を使用）」で得た数値であり、モデル単体ではなくエージェントシステム込みの評価である。設定次第で 62.7%との報告もある^{[21][6]}。ARC Prize の Greg Kamradt 氏は「テストした中で最良のモデルで、96%の水準で人間の行動効率ベースラインを超えた」と評価した^[6]。

- FrontierMath Tier 4 は Epoch AI が運営するが、OpenAI が開発資金を提供し一部に独占アクセスを持つ点に留意が必要である^[21]。Epoch 自身の FrontierMath Erdős（1 問 1 試行、\$300/72 時間の予算）では、プレリリース版 Astra のみが 68 問中 2 問を解き（他社モデルは 0 問）、非体系的な追加試行では計 5 問とされる^[24]。

6.4 安全チェックによる作業上の摩擦【確認済み】

OpenAI は「追加の安全チェックが正当な作業（防御的サイバーを含む）を遅延・一時停止・停止させ得る。ChatGPT/Codex では確認を求められ、API ではタスクが停止する」と明記している^{[1][12]}。OpenAI 開発者フォーラムでは、有料ユーザー（hexwolf 氏）が「Codex で正当な信頼性レビューがサイバー安全メッセージで停止する。同じ作業は GPT-5.6 Sol なら完了する」と報告し、他ユーザー（ojfd 氏、stee 氏）も同様の誤検知を報告している^[11]。安全研究者 Tomek Korbak 氏は、推論の監視可能性が低下する傾向を「深く懸念する」と述べた^[31]。

6.5 他社モデルとの位置づけ【分析】

主要な比較対象は Anthropic の Claude Fable 5.1（9 月 1 日リリース）および Claude Opus 5 である。「コンピュータ操作・数学・サイバーは Astra、推論の深さ・マージ可能なコード品質・UI の洗練は Fable 5.1」という使い分けの見方が支配的である^{[8][5]}。Meta Muse Spark 1.3、Google Gemini 3.8 Flash、xAI Grok 4.6 も一部ベンチマークで拮抗しており、Gemini 3.8 Flash は DeepSWE で Astra を僅差で上回り、Grok 4.6 は知能指数で Astra・Sol と同点である^{[1][5]}。

6.6 AGI 言説【確認済み／要留保】

OpenAI 社長 Greg Brockman 氏は 9 月 3 日の記者ブリーフィングで、個人的見解として「我々はもうそこにいると思う」「AGI 時代に入ったと感じるのは不合理ではない」と述べ、ブリーフィングを「Welcome to the AGI era.」で締めた^{[28][38]}。これは OpenAI 幹部の主張であり、合意された AGI の判定基準は存在しない。Tech Times は、この主張が OpenAI 自身の従来の定義基準を満たしていないと批判している^[31]。断定的に扱うべきではない【分析】。

7. 総括と留意点

7.1 用途別の見立て【分析】

- コンピュータ操作・エージェント的長時間作業・3D/ゲーム生成・数学/科学が主目的の場合：Astra への移行を検討する価値が高い。実測レビューでこの領域の優位は一貫している^{[6][8]}。max/xhigh 努力度は費用対効果を都度検証すべきである（1 レンダリングで数十セント級の例あり）。

- **純粋なコーディング（特に UI・マージ可能なコード）主体の場合：**当面は GPT-5.6 Sol または Claude Fable 5.1 の継続が妥当。Astra は横移動～劣後との実測が複数ある^{[8][5]}。ただし Codex でのトークン効率（Sol 比約 1/3）は魅力であり、自身のタスクでの A/B 比較が望ましい。
- **API でのサイバー／セキュリティ関連作業：**安全モニタでタスクが停止（API 403）し得る^{[12][11]}。正当な用途は Daybreak／Trusted Access の申請を検討する。
- **段階展開の混乱：**未提供でも慌てる必要はない。Plus 枠は Pro より制限的であり、急ぐなら Pro／Business。未提供期間は「banked reset」で補償される^{[17][34]}。

7.2 判断が変わる閾値【分析】

以下のいずれかが生じた場合、上記の見立ては再検討を要する。（1）Artificial Analysis 知能指数 v4.2 以降で Astra が Fable 5.1 を明確に上回る、（2）独立コーディング評価で Fable 5.1 を超える、（3）価格が実質 Sol 並みに低下する、（4）サイバー安全ブロックの誤検知が解消される。

7.3 本報告の限界

- 本報告は発表から約 3 日時点の初期反応に基づく。評価は今後変わり得る。
- ベンチマークの多くは OpenAI 自己申告（max 努力度）であり、独立検証済みは一部にとどまる。ハーネス設定で数値が大きく変動する点に注意を要する。
- 「AGI 到達」的表現は OpenAI 幹部の主張であり、合意された AGI テストは存在しない。
- Reddit（r/OpenAI 等）の個別投稿・upvote 数は一次ソースで直接確認できなかった（検索・アクセス制限のため）。本報告の体感評は、X・Hacker News（実名ユーザー確認分）・OpenAI 開発者フォーラム・技術メディア・日本語メディア経由の情報を含む。X 投稿の原文 URL は未確認である。
- 数学 10 問・素数ギャップ等の科学的成果はいずれも査読未了である。
- 価格・提供範囲・レート制限は変動し得る。

参考文献

番号は本文中の上付き番号に対応する（初出順）。URL は 2026 年 9 月 6 日時点で調査ツールによりアクセス・確認したもの。「URL 未確認」は、本調査で当該記事・投稿の個別 URL を取得できず、二次情報経由で内容を把握したものを指す。

[1] OpenAI, “GPT-6 Astra: A new generation of intelligence”, OpenAI 公式ブログ, 2026 年 9 月 3 日。
<https://openai.com/index/gpt-6-astra/>

- [2] The New Stack, “‘Sorry for the messy rollout’: OpenAI launched GPT-6 Astra, but developers are locked out”, 2026 年 9 月. <https://thenewstack.io/gpt6-astra-developer-access-delayed/>
- [3] The Decoder, “OpenAI rolls out GPT-6 Astra to top-tier ChatGPT plans at half the rate of GPT-5.6 Sol”, 2026 年 9 月. <https://the-decoder.com/openai-rolls-out-gpt-6-astra-to-top-tier-chatgpt-plans-at-half-the-rate-of-gpt-5-6-sol/>
- [4] Layer3Labs, “OpenAI Astra Shipped as GPT-6 Astra on September 3, 2026”, 2026 年 9 月. <https://www.layer3labs.io/guides/openai-astra-explained>
- [5] Artificial Analysis, “Benchmarking GPT-6 Astra”, 2026 年 9 月. <https://artificialanalysis.ai/articles/benchmarking-gpt-6-astra>
- [6] The Neuron, “GPT-6 Astra: Everything You Need to Know About OpenAI’s New Model”, 2026 年 9 月. <https://www.theneuron.ai/news/gpt-6-astra-everything-you-need-to-know-about-openais-new-model/>
- [7] DataCamp, “GPT-6 Astra: Features, Benchmarks, and Pricing”, 2026 年 9 月. <https://www.datacamp.com/blog/gpt-6-astra>
- [8] Jonathan Fulton, “A Short Review of OpenAI’s GPT 6 Astra”, Medium (Jonathan’s Musings) , 2026 年 9 月. <https://medium.com/jonathans-musings/a-short-review-of-openais-gpt-6-astra-091fa5beccf9>
- [9] The News International, “Altman apologises for GPT-6 Astra rollout locking out pro users”, 2026 年 9 月 4 日. <https://www.thenews.com.pk/latest/1414847-altman-apologizes-for-gpt-6-astra-rollout-locking-out-pro-users>
- [10] Hacker News, GPT-6 Astra 発表スレッドにおけるユーザー投稿（minimaxir 等）, 2026 年 9 月.（スレッド URL 未確認）
- [11] OpenAI Developer Community, “False-positive cybersecurity blocks during Astra reliability audits in Codex”（投稿者: hexwolf ほか）, 2026 年 9 月. <https://community.openai.com/t/false-positive-cybersecurity-blocks-during-astra-reliability-audits-in-codex/1395121>
- [12] Invidelabs, “GPT-6 Astra can stop API tasks during safety checks”, 2026 年 9 月. <https://blog.invidelabs.com/gpt-6-astra-api-safety-stops/>
- [13] OpenAI, GPT-6 Astra System Card および安全性概要ページ（deploymentsafety.openai.com）.（本調査では個別 URL を確認できず（公式ブログ[1]からの参照））
- [14] Ron Schmelzer, “OpenAI Launches GPT-6 Astra After A Curious False Start”, Forbes, 2026 年 9 月 3 日. <https://www.forbes.com/sites/ronschmelzer/2026/09/03/openai-announces-gpt-6-astra-or-does-it/>
- [15] CNBC, GPT-6 Astra 発表に関する報道（公式ブログ公開前の先行記事）, 2026 年 9 月 3 日.（URL 未確認（Forbes[14]等の二次情報経由））
- [16] 9to5Mac, “OpenAI releasing major upgrade to ChatGPT and Codex with GPT-6 Astra, details here”, 2026 年 9 月 4 日. <https://9to5mac.com/2026/09/04/openai-releasing-major-upgrade-to-chatgpt-and-codex-with-gpt-6-astra-details-here/>
- [17] PC Watch (Impress) , 「GPT-6 Astra を利用可能になった記念でリセット権、ChatGPT 有料 3 プラン全員に配布」, 2026 年 9 月 5 日. <https://pc.watch.impress.co.jp/docs/news/2138527.html>
- [18] ITmedia, GPT-6 Astra の国内提供開始に関する報道, 2026 年 9 月 5 日.（URL 未確認）
- [19] OpenAI, GitHub リポジトリ「[openai/ten-proofs](https://github.com/openai/ten-proofs)」（249 ページ原稿および Lean 4 証明、Apache 2.0）,

- 2026 年 8 月. (URL 未確認 (二次情報経由))
- [20] Sébastien Bubeck (OpenAI 数学研究責任者), X 投稿, 2026 年 8 月. (原投稿 URL 未確認)
- [21] Vellum, “GPT-6 Astra Benchmarks Explained”, 2026 年 9 月. <https://www.vellum.ai/blog/gpt-6-astra-benchmarks-explained>
- [22] BelieveMy, “ChatGPT Astra works 47% faster, but it can stop mid-task”, 2026 年 9 月.
<https://believemy.com/en/r/chatgpt-astra-gpt-6-openai>
- [23] The Decoder, Artificial Analysis Intelligence Index v4.2 改訂と GPT-6 Astra の順位に関する報道, 2026 年 9 月. (URL 未確認)
- [24] Epoch AI, Epoch Capability Index (ECI) および FrontierMath Erdős 評価結果に関する発表, 2026 年 9 月.
(URL 未確認 (二次情報経由))
- [25] CodeRabbit, GPT-6 Astra によるコードレビュー精度評価ブログ, 2026 年 9 月. (URL 未確認 (二次情報経由))
- [26] unicodeveloper, “GPT-6 Astra. A taste of AGI?”, Medium, 2026 年 9 月.
<https://medium.com/@unicodeveloper/gpt-6-astra-a-taste-of-agi-938515afc5c7>
- [27] Kilo, “GPT-6 Astra: What We Learned Previewing OpenAI’s New Model in Production”, 2026 年 9 月.
<https://blog.kilo.ai/p/gpt-6-astra-what-we-learned-previewing>
- [28] The New Stack, “OpenAI launches GPT-6 Astra and says welcome to the ‘AGI era’”, 2026 年 9 月.
<https://thenewstack.io/openai-gpt6-astra-benchmarks/>
- [29] 農業 AI 通信 (Metagri-labo), 「【最新モデルレビュー】 GPT-6 Astra (アストラ) を検証！～直売所 POP・SNS を一括制作～」, 2026 年 9 月. <https://metagri-labo.com/ai-guide/gpt-6-astra/>
- [30] X (旧 Twitter) トレンド解説, “OpenAI’s GPT-6 Astra Rollout Confuses ChatGPT Plus Users on Access”, 2026 年 9 月. <https://x.com/i/trending/2095765286196777180>
- [31] Tech Times, “GPT-6 Astra Goes Live: AGI Claim Fails OpenAI Own Bar, Monitoring Called Fragile”, 2026 年 9 月 4 日. <https://www.techtimes.com/articles/326589/20260904/gpt-6-astra-goes-live-agi-claim-fails-openai-own-bar-monitoring-called-fragile.htm>
- [32] Apollo Research, GPT-6 Astra の評価認識 (eval awareness) に関する所見 (System Card[13]所収), 2026 年 9 月. (URL 未確認)
- [33] Pasquale Pillitteri, “Astra is late to paid plans, OpenAI hands out a reset for every day of the wait”, 2026 年 9 月. <https://pasqualepillitteri.it/en/news/14288/openai-astra-reset-paid-chatgpt-wait>
- [34] Implicator, “OpenAI Compensates Subscribers Still Waiting for GPT-6 Astra”, 2026 年 9 月.
<https://www.implicator.ai/openai-gives-daily-usage-resets-to-subscribers-still-waiting-for-gpt-6-astra/>
- [35] AI Wiki, “GPT-6 Astra”, 2026 年 9 月. https://aiwiki.ai/wiki/gpt_6_astra
- [36] Unite.AI, GPT-6 Astra のサイバー能力評価 (ゼロデイ発見を含む) に関する報道, 2026 年 9 月. (URL 未確認)
- [37] CSO Online, GPT-6 Astra の「Critical」指定に関する Sanchit Vir Gogia (Greyhound Research) のコメントを含む記事, 2026 年 9 月. (URL 未確認)
- [38] VentureBeat / Axios, Greg Brockman (OpenAI 社長) の 2026 年 9 月 3 日記者ブリーフィング発言に関する報道. (URL 未確認)