

米中フロンティア AI モデルの技術的ブレイクスルー予測

— 2026 年後半から 2027 年にかけての展望 —

調査基準日：2026 年 8 月 15 日／作成：2026 年 8 月 16 日

Claude Opus 5

本文中の上付き番号は文末「参考文献」の番号に対応する。

【本報告は将来予測を多く含む。確定事実と分析・推論・予測を明示的に区別し、予測には確信度と反証条件を付した。】

1. 要旨

1.1 結論（3 点）

- 2027 年末時点では、米国のクローズド最高性能リードは「僅差での継続」が最も蓋然性が高い（中位シナリオ、確信度：高）。性能ギャップの評価は機関により割れており、Epoch AI は「2023 年以降、平均 7 か月・最小 4 か月・最大 14 か月」¹、NIST 傘下の CAISI は DeepSeek V4 Pro を「米最高性能の約 8 か月遅れ」かつ拡大傾向²、Rice 大 McNair Center は「3～9 か月遅れ。閉じつつある窓ではなく均衡」³、Stanford HAI の 2026 年 AI Index は公開リーダーボード上の差を「2.7%」⁴としている。総じて縮小ないし横ばいで推移し、オープンウェイトと開発者採用・コスト効率では中国が優位を固める公算が大きい（分析）。
- 技術の重心が「事前学習スケール一本足」から、推論時計算・RL スケール・エージェント長期自律・継続学習へ移行する（確信度：中～高）。METR のタスク時間ホライズンは全期間で約 7 か月（212 日）倍増だが、2024～2025 年は約 4 か月、SWE-Bench Verified ベースでは 3 か月未満へ加速している⁵。この趨勢が続けば、エージェントは 2027 年に 1 か月規模のタスクに到達しうるとの外挿もある⁶。ただし計測手法の飽和という強い留保が付く。
- 最大の制約は技術ロードマップではなく、計算資源・電力・資本である（確信度：高）。2027 年の勝敗を左右するのはアーキテクチャの巧拙よりも、データセンター建設・電力確保・資本の持続性である可能性が高い（分析）。

1.2 日本企業への含意（先出し）

「クローズド米国勢＝最高性能」「中国オープンウェイト＝コスト・主権・カスタマイズ性」という二層構造を前提に、マルチベンダー調達と、輸出規制・アクセス撤回リスクの両にらみで臨むべきである（分析）。

2. 次世代モデルのロードマップ

本節は公式発表・報道に基づく確認事実を中心とするが、一次ページを全文確認できていないものが多く、該当箇所は参考文献で「URL 未確認」と明示した。

2.1 米国勢

OpenAI

2026 年前半に「Spud」=GPT-5.5（4 月 23 日）を新しい事前学習基盤としてリリースし、Altman 氏は「2 年分の研究がこのモデルに結実した」と表現した。以降、GPT-5.6（Sol/Terra/Luna の三層、7 月 9 日 GA）まで刻み込み型で連発しており、GPT-5 系は 1 年未満で 7 回のリリースを重ねている⁷。

注目すべきは 2026 年 8 月 1 日の「Astra」発表である。これは通常のプロダクト発表ではなく、数学・理論計算機科学の未解決 10 問への貢献という形での予告であり、リリース日・価格・モデルカードのいずれも公表されていない⁷。The Information の報道によれば、Astra は複数エージェントを長期間協調させて難問を解く新クラスのモデルで、Sol/Terra/Luna と並ぶ位置づけとされる。GPT-6 と名付けるか GPT-5.7 とするかは未定である⁸。開発者向けプレビューは 2026 年後半、公開は 2027 年前半という見方があるが、これは二次情報であり未確認である（留保）。

Anthropic

多層化が進行している。Claude Opus 5（2026 年 7 月 24 日、1M トークン文脈、最大 128K トークン出力、5 段階の effort 設定）は価格を据え置きつつ Frontier-Bench でほぼ倍のスコアを記録した^{9,10}。その上位に Mythos 級（一般公開版の Claude Fable 5、限定アクセスの Claude Mythos 5）が位置する⁹。「Opus 6」「Sonnet 5.5 Fennec」については複数のリーク動画が流布しているが、確たる証拠はなく、トラッカーも懐疑的である（未確認）。

Google DeepMind

Gemini 3.6 Flash（2026 年 7 月 21 日）でトークン効率を 17%改善し値下げを実施した¹²。一方、Gemini 3.5 Pro はコーディング・複雑推論で内部目標に達せず、広範なリリースを延期したと Bloomberg が報じており¹³、Gemini 4 を次期主力として予告している¹¹。高度推論は Deep Think 系（Gemini 3 Deep Think）が担う²⁴。

xAI/SpaceXAI

Grok 4.5 まで出荷済み。Grok 5（6 兆パラメータ MoE、メンフィスの 1~1.5GW スパコン「Colossus 2」で学習、ネイティブ動画理解、マルチエージェント）は当初 2026 年 Q1 予定から Q2、Q3 へと延

期を重ねている¹⁴。Musk氏はAGI到達確率を「10%かつ上昇中」と発言しているが、これは自己申告であり検証不能である（留保）。

Meta

Superintelligence Labs体制下でMuse Spark／Glimmer系を展開。オープンソース路線からの後退が指摘されており、LeCun氏の離脱（2025年末）とAlexandr Wang体制への移行がそれを象徴している¹¹。

2.2 中国勢

次世代モデルの公式ロードマップは、米国勢と比較して著しく情報が乏しい。以下は現行世代の到達点と、パターンからの推測である。

- DeepSeek：V4-Pro／Flash（2026年4月、32兆トークン学習、1M文脈、CSA+HCAハイブリッドアテンション）¹⁵。V5／R3は公式には未予告であり、日付・仕様とも不明である。過去のパターンからはV4.x安定版・Speciale・コーディング特化が先行すると推測される（推測）。
- Moonshot（Kimi）：K3が開放モデル首位級。K2.7-Codeが高速コーディング特化。Agent Swarm（300サブエージェント、12時間超の連続実行、4,000超のツールコール）が特徴である¹⁵。K4は未予告。
- Alibaba Qwen：Qwen3.7-Max／Qwen3.8-Maxまで。Qwen4系は未確認。
- Z.ai／智譜：GLM-5.2（約7,500億パラメータ、1M文脈）¹⁵。GLM-6は未予告。
- ByteDance（Doubao／Seed）、MiniMax、StepFun、01.AI：具体的な次世代ロードマップの一次情報は乏しく、本報告のカバレッジは薄い（未確認）¹⁶。

2.3 主要人物の2026年時点の発言

表1 次世代モデル・AGI時間軸に関する発言

人物	発言内容	出典・確信度
Sam Altman (OpenAI)	Astraを2026年7～8月にワシントンで政策担当者に実演。「スケーリングを過小評価した世代が研究を遅らせた」	二次情報・未検証 ^{7,8}
Dario Amodei (Anthropic)	「スケーリング則があと1～2年続けば、私が『データセンターの中の天才国家』と呼ぶPowerful AIに至る公算が高い」。継続学習について「障壁ですらないかもしれない」	自著は一次 ¹⁷ 、ポッドキャストは二次 ¹⁸
Demis Hassabis (DeepMind)	AGI時間軸を「3～4年（2029～2030年）」へ短縮。世界モデルと継続学習が解決すべき課題	二次情報 ¹⁹ 。Gemini 4の具体発言は未確認
Jensen Huang (NVIDIA)	GTC 2026で、Blackwell／Vera Rubinの受注が2027年まで累計1兆ドルに達すると予測（前年の5,000億ドル予測から倍増）。「トークンは新たな通貨、AIファクトリーはその工場」	基調講演 ²⁰

人物	発言内容	出典・確信度
Ilya Sutskever (SSI)	「スケーリングの時代（2020-2025）は終わり、研究の時代に戻る」。ただし X 上で「現行手法のスケーリングは改善を生み続け、失速しない」と補足	ポッドキャスト ¹⁸ 。SSI は累計 30 億ドル調達・評価 320 億ドル
Yann LeCun (AMI Labs)	「LLM は家猫より賢いシステムの設計の糸口にすらなっていない」。2026 年 3 月に約 10.3 億ドル調達（評価 35 億ドル、欧州最大級のシード）	報道 ¹¹

3. 技術的ブレイクスルーの候補領域

以下、11 の候補領域について、現在の到達点（確認事実）と 2027 年までの実現可能性（分析）、観測可能な先行指標、および反証条件を整理する。

表 2 ブレイクスルー候補領域の評価

領域	現在の到達点（2026 年 8 月）	2027 年実現可能性	先行指標	反証条件
スケーリング則の重心移動	事前学習の「壁」論と重心移動論が並存。RL スケーリング（o1→o3 で 10 倍超）と推論時計算が主戦場に。Train-to-Test 統合スケーリング則の論文化	高：RL／推論時計算スケーリング則が事前学習則並みに標準化	モデルカードの RL 計算比率、pass@k 公開	新データ源・合成データによる事前学習の再加速
継続学習・長期記憶	NVIDIA TTT-E2E（文脈を学習データ化） ²⁵ 、記憶ベンチマークの乱立。Amodei 氏らが 2026 年内解決を予測 ^{17,18}	中：部分的実用化。「満足のいく解決」は反証されやすい	主力モデルへのパラメータ更新型記憶の実装、記憶ベンチの SOTA 更新	本番環境で破滅的忘却が頻発
超長文脈・ハイブリッドアテンション	DeepSeek CSA+HCA、Kimi、Qwen-Next、Nemotron-H 等でハイブリッド MoE が本番投入 ¹⁵ 。1M 文脈は標準化	高：1,000 万トークン級とハイブリッドが主流化。純 Transformer 比率が低下	主力モデルの文脈長、SSM／線形層比率の公表	長文脈でのリコール精度がハイブリッドで劣化
エージェント長期自律	METR ホライズンは約 4 か月倍増、最強モデルで 50%基準 16～20 時間（飽和により信頼性低下） ⁵ 。Kimi Agent Swarm が 12 時間超の実行 ¹⁵	中～高：数時間～1 日規模が実用域。数日規模は不確実	METR TH 更新、実運用での自律タスク成功率	倍増則が 7 か月へ回帰（飽和・難問の壁）
自己改善・AI for AI research	Sakana AI（Darwin Gödel Machine、AI Scientist、RSI Lab） ³⁵ 、DeepMind AlphaEvolve ²⁴ 。ICLR 2026 に RSI ワークショップ	中：人間監査ループ内での「有界な自己改善」が主。完全な再帰的自己改善は限定的	モデル改善の wall-clock 短縮、自律研究の査読通過	評価器がボトルネックとなり収束しない
世界モデル・フィジカル AI	DeepMind Genie 3（リアルタイム対話型 3D 環境） ²⁴ 、NVIDIA Cosmos・GR00T N1.6 ²⁵ 、Meta V-JEPA 2 ²⁶ 、Runway Gen-4.5 ²⁷	中：シミュレーション・合成データ用途で実用化。汎用ロボット基盤は限定的	接触予測精度、実機へのゼロショット転移	接触ダイナミクスの未解決が持続
数学・科学の自律的発見	Aristotle（Harmonic）が IMO2025 で金メダル相当、最終問題を除く全問で Lean 4 形式検証済みの解を生成 ^{21,22} 。VERINA コード検証で 96.8%。OpenAI／Anthropic が Lean 証明付き研究成果を発表 ^{7,9,23}	中～高：形式検証統合が競技から研究レベルへ移行。ミレニアム問題級は非現実的	Lean 形式化された新定理数、査読誌の形式化要求	「発見」が実は文献検索だった事例の再発
効率化（蒸留・量子化・MoE 極限）	INT4 量子化、兆パラメータ MoE で 37B 級のみ活性化、蒸留の xLSTM 等への展開	高：推論コスト低下曲線が継続、エッジ化が進行	\$/task 推移、オンデバイスモデルの性能	HBM 供給逼迫でコスト低下が鈍化 ³⁰

領域	現在の到達点（2026 年 8 月）	2027 年実現可能性	先行指標	反証条件
新アーキテクチャ (Transformer 代替)	ハイブリッドが本番品質に到達 (Jamba、Nemotron-H、Qwen-Next)。拡散型 LLM、JEPa 系は研究段階 ²⁶	中：ハイブリッドは主流化。純代替（拡散・EBM・JEPa）の本番化は限定的	主力モデルの非 Transformer 層比率	純 Transformer が効率で再逆転
マルチモーダル統合	GPT 系 Omni、ネイティブ音声・動画、リアルタイム対話が実用段階	高：ネイティブ音声・動画・3D の統合が標準化	主力モデルの入出力モダリティ	—
信頼性・解釈可能性	ハルシネーション低減（GPT-5.5 で 60% 減との主張 ⁷ ）。機械的解釈可能性は研究段階	中：ハルシネーション低減は進むが、解釈可能性の実用化は限定的	ハルシネーション率、解釈可能性ツールの本番投入	高能モデルで検証性がかえって低下

4. 制約要因とボトルネック

4.1 計算資源と電力

Stargate は 9GW 超を計画しているが、Abilene 拡張（2GW 化）を Oracle/OpenAI が撤回しており、電力確保と地域の反対が実質的な制約となっている（報道ベース）。ハイパースケーラーは Vera Rubin 世代（Rubin GPU、HBM4 最大 288GB、NVFP4 推論 50PFLOPS、2026 年後半量産）および Rubin Ultra「Kyber」ラック（約 600kW、2027 年）へ移行し、電力設計の抜本的な再設計を迫られる²⁰。ただし「電力はボトルネックではない」とする Epoch AI の反論もあり、見解は分かれている（留保）。

4.2 半導体供給と輸出規制

HBM の需給は逼迫しており、TrendForce は高成長を予測している³⁰。NVIDIA は 2027 年までに累計 1 兆ドルの需要を見込む²⁰。

輸出規制については、2025 年 12 月にトランプ政権が H200 の対中販売を許可し（売上の 25%を米政府へ）、2026 年 1 月に約 10 社へライセンスが発給された³²。しかし 2026 年 2 月時点で NVIDIA は対中データセンター売上をゼロと決算で言明しており、商務省の Kessler 氏も 2026 年 7 月の議会証言で実出荷は「ごく僅か」と述べている^{31,32}。中国側が国産採用を条件に選別的に輸入している可能性がある（推測）。

中国の国産化は Huawei Ascend 950PR（2026 年 Q1、自社 HBM）、950DT（2026 年 Q4）、960（2027 年 Q4、Blackwell 級を標榜）と進み、ノード／クラスタ級（Atlas 950／960 SuperPoD）で台数集約により対抗する戦略を採る^{28,29}。米国の対中コンピュータ優位は 2026 年時点で 21～49 倍との推計があり、密輸推計は累計 29 万～160 万 H100 相当である¹（推計値の幅の大きさ自体が実態把握の困難さを示す）。

4.3 データ、資本、人材、規制

- データ：Webscale-RL 等で RL データを事前学習規模へ拡張する試みが進む。専有データ・実世界データの重要性が増大している。
- 資本（AI バブル論）：ハイパースケーラー4社の2026年設備投資は合計約7,250億ドル（前年比約77%増。内訳はAmazon 約2,000億、Microsoft 約1,900億、Alphabet1,750～1,850億、Meta1,150～1,350億ドル。前年は約4,100億ドル）³³。一方 OpenAI は2026年に売上ラン・レート約250億ドルで巨額赤字とされ、コミット総額を当初「約1.4兆ドル」と表現した後、2026年2月に「2030年までに約6,000億ドル」へ再フレームした³²。循環的な資金調達
（NVIDIA—OpenAI—Oracle—AMD、2026年7月にはNVIDIAがOpenAIのオハイオDCで最大2,500億ドルの保証を検討との報道）が懸念材料である。他方、JefferiesのBrent Thill氏はFT集計の設備投資を評して「AI経済は健全であり、弱気論はゴミだ」と反論している³³。
- 人材：LeCun氏のMeta離脱、SSIへのNVIDIA出資（2026年7月）、Gross氏のSSI→Meta移籍など、トップ研究者の移動が続く^{11,18}。
- 規制：EU AI Act、米国の州法（カリフォルニアSB53、ニューヨークRAISE、イリノイSB315）、中国の生成AI規制。2026年6月に米政府がClaude Fable/Mythosの海外アクセスを停止し、Anthropicが撤回した事例は、モデルアクセスの主権リスクを顕在化させた⁹。

5. 米中の戦略とギャップのシナリオ

5.1 4 シナリオ

表3 2027年末時点のギャップに関するシナリオ

シナリオ	前提条件	2027年末のギャップ	先行指標	確信度
楽観（米国優位拡大）	米国が電力・資本制約を克服。Astra/Gemini 4/Grok 5が世代跳躍。輸出規制が実効化	クローズドで米国リード拡大。CAISI型「拡大」説が的中 ²	Astra/Gemini 4の公開ベンチ、中国の学習用チップ不足	低～中
中位（僅差リード継続）	現状趨勢の延長。米国がクローズド最高性能を維持、中国がオープンウェイト覇権を継続	3～9か月ギャップが横ばい。二層構造が固定化	Epoch ECI ¹ 、OpenRouter トークンシェア ³⁴ 、\$/task	高
収束（実質収束）	米国が電力・資本・規制で失速。中国がAscend 960で学習実用化、蒸留依存から脱却	実用領域で収束。Stanford HAI型「2.7%差」が主流評価に ⁴	中国の国産チップによる主力モデル学習、知能指数の差	中
テールリスク（中国逆転）	米国のAIバブル崩壊で資本引き揚げ。中国が国家資源を集中	一部領域で中国が最高性能を奪取	ハイパースケーラー設備投資の急減 ³³ 、米ラボの資金難	低

5.2 各陣営の採りうる戦略

米国側：クローズド最高性能の維持（Astra/Opus/Gemini 4/Grok 5）、垂直統合（チップ～モデル

～プロダクト、NVIDIA—OpenAI 連携）、推論時計算での差別化。オープンウェイトへの回帰は Meta の後退により限定的で、NVIDIA の Nemotron 等が例外にとどまる（分析）。

中国側：オープンウェイト戦略の継続。Adnan Masood 氏の定量分析（2026 年 7 月）によれば、中国オープンモデルは OpenRouter 上で米企業のトークン量の 30%超を 60～90%安い価格で占め、公開ベンチ上の能力は 3～9 か月遅れだが「閉じつつある窓ではなく均衡」と評価される³⁴（個人ブログ由来であり要検証）。国産チップでの学習実用化（2027 年の Ascend 960 が試金石）²⁸、コスト優位の持続、独自アーキテクチャ（CSA+HCA 等）による蒸留依存からの脱却が鍵となる（分析）。

6. 知財・企業実務への含意

6.1 特許出願動向

AI 関連特許（特にモデルアーキテクチャ、推論最適化、MoE ルーティング、線形／ハイブリッドアテンション、KV キャッシュ圧縮）は中国勢の出願が急増している領域と推測されるが、本調査では一次の特許統計を直接確認できていない³⁶。WIPO および各国特許庁データによる別途確認を推奨する。技術の方向性としては、推論効率化・長文脈・エージェント基盤に出願が集中すると考えられる（推測）。

6.2 モニタリングすべき 5 つのポイント

- METR のタスク時間ホライズン倍増則⁵：エージェント自律化の実用タイミングを規定する最重要指標。
- 中国オープンウェイトの\$/task とライセンス条項：コスト優位の持続性に加え、Kimi K3 の Modified MIT は月間 1 億 MAU 超で帰属条項が発動するなど、条項の精査が必要である。
- 輸出規制・モデルアクセス停止リスク^{9,31,32}：米政府による Fable／Mythos 停止事例が示すとおり、主権・供給継続性の観点から二重化が必須である。
- Ascend 960（2027 年 Q4）の学習実用化²⁸：中国スタックの自立度を測る最良の指標。
- 継続学習・長期記憶の実用化^{17,18,25}：「AI コワーカー」化の分水嶺となる。

6.3 調達・技術選定への示唆

クローズド最高性能（米国）とオープンウェイト・コスト（中国）の二層を前提としたマルチベンダー戦略を採り、単一ベンダー・単一国への依存を避けるべきである。オンプレミス／主権要件が高い用途はオープンウェイト（GLM／Qwen／DeepSeek／Llama 系）を評価対象とする。推論コストの低下曲線を前提とすれば、長期契約よりも柔軟な調達を優先することが合理的である（分析）。

7. 実務的推奨（時間軸別）

7.1 2026 年後半（直ちに着手）

- 主力業務でのモデル評価を「知能指数」単独から、「\$/task」「タスク時間ホライズン」「文脈長さのリコール精度」の複合指標へ移行する。
- 中国オープンウェイト（DeepSeek V4-Pro、Kimi K3、GLM-5.2、Qwen3.7-Max）を主権／コスト要件用途で PoC し、同時に米国クローズド（Opus 5、GPT-5.6、Gemini 3.6）を最高性能用途で維持する二層調達を確立する。
- 輸出規制・モデルアクセス停止リスクに備え、クリティカル用途はオンプレミス可能なオープンウェイトで冗長化する。

7.2 2027 年前半（観測して判断）

- METR ホライズンが数時間～1 日規模を安定的に超えたら、長期自律エージェントのパイロットを本番検討する。超えない場合は人間監査ループ前提を維持する。
- 継続学習の実用化が主力モデルに実装されたら、長期記憶を要する業務（顧客関係管理・長期プロジェクト）へ展開する。
- Gemini 4、Grok 5、Astra（GPT-6 相当）の公開ベンチマークで世代跳躍の有無を確認する。跳躍が確認できれば投資を加速し、確認できなければ現行世代の最適化に注力する。

7.3 2027 年後半（再評価）

- Ascend 960 の学習実用化と Huawei SuperPoD の実性能を確認し、中国スタックの自立度を再評価する。自立が進めば中国オープンウェイトの持続性が高まると判断できる。
- AI バブル指標（ハイパースケーラー設備投資、OpenAI の収益化、循環的資金の健全性）を監視する。急激な資本引き揚げの兆候があれば、特定ベンダーへの依存度を縮小する。

7.4 判断を変える閾値

- 中国モデルが Epoch ECI または Artificial Analysis 知能指数でクローズド首位を奪取した場合 → 中国逆転シナリオへ移行し、調達比率を見直す。
- METR 倍増則が 7 か月へ回帰した場合 → エージェント自律化の時間軸を後ろ倒しする。
- ハイパースケーラー設備投資が前年比マイナスに転換した場合 → バブル調整局面と判断し、リスク資産的な AI 投資を抑制する。

8. 留保事項

- 本報告は将来予測を多く含む。確定事実（公式発表・査読論文・公表データ）と、分析・推論・予測を区別して記述し、予測部分には確信度と反証条件を明示したが、予測は本質的に不確実であり過信は禁物である。
- モデル名・リリース日の一部（Astra、Grok 5、Opus 6、Gemini 4 等）は二次情報・リーク・トラッカーに依拠しており、一次確認できていないものは本文中に「未確認」と明示した。特に Astra の数学的成果は The Information の原報道に依拠しており、本調査では二次報道のみを確認している⁸。
- METR のタスク時間ホライズンは 16 時間超で計測手法が飽和しており、外挿には強い留保が必要である（METR 自身が明言）⁵。倍増則の加速が一時的か構造的かは、1 年分のデータでは判定困難である。
- ギャップの定量評価は機関・手法により 3～9 か月とばらつく（Epoch＝横ばい／縮小¹、CAISI＝拡大²、Stanford HAI＝2.7%で実質収束⁴）。McNair Center は「同じデータから逆の結論が出る」点に注意を喚起し、両者とも「データが支持できる以上の確信を主張している」と評している³。
- 特許統計は一次確認できておらず、本報告の記述は推測にとどまる³⁶。
- 資本・財務数値の一部（OpenAI の赤字、コミット総額）は The Information／FT 経由の報道であり、企業の公式確認ではない^{8,32,33}。ハイパースケーラー設備投資（約 7,250 億ドル）は複数ソースで比較的堅い³³。
- 中国企業の一部（ByteDance Seed、MiniMax、StepFun、01.AI）の次世代ロードマップは一次情報が乏しく、本報告のカバレッジは薄い¹⁶。
- 参考文献のうち多数が URL 未確認である。これは本調査の探索範囲において一次ページを直接確認できなかったことを意味し、情報の真偽そのものを否定するものではないが、引用・転載の際は原典確認を強く推奨する。

参考文献

番号は本文中の上付き数字に対応する。URL は 2026 年 8 月 16 日時点のもの。赤字イタリックで「URL 未確認」等と記載した文献は、本調査において一次ページを直接確認できておらず、報道・二次情報経由で参照した文献である。

- [1] Epoch AI 「Chinese AI models have lagged the US frontier by 7 months on average since 2023」 <https://epoch.ai/data-insights/us-vs-china-eci> (ECI による米中差の定量分析 (平均 7 か月、幅 4~14 か月))
- [2] NIST (CAISI) 「CAISI Evaluation of DeepSeek V4 Pro」 <https://www.nist.gov/news-events/news/2026/05/caisi-evaluation-deepseek-v4-pro> (米政府による中国モデル評価 (約 8 か月遅れ・拡大傾向))
- [3] Rice University, McNair Center 「The US-China AI gap is very uncertain.」 <https://mcnair.center/china/> (「3~9 か月遅れ、閉じる窓ではなく均衡」との評価)
- [4] Stanford HAI 「AI Index Report 2026」 **URL 未確認** (公開リーダーボード上の米中差 2.7%)
- [5] METR 「Measuring AI Ability to Complete Long Tasks および Time Horizon 更新 (TH1.1、2026 年 1 月)」 **URL 未確認** (タスク時間ホライズンの倍増則と飽和に関する注記)
- [6] AI Digest 「エージェントのタスク時間ホライズン外挿」 **URL 未確認** (「2027 年に 1 か月規模のタスク到達しうる」との外挿)
- [7] OpenAI 「公式発表 (GPT-5.5 「Spud」 2026 年 4 月 23 日、GPT-5.6 Sol/Terra/Luna 2026 年 7 月 9 日 GA、Astra 予告 2026 年 8 月 1 日)」 **URL 未確認** (次世代モデルのリリース経緯 (一次情報だが本調査で全文未確認))
- [8] The Information 「Astra に関する報道 (2026 年 8 月)」 **URL 未確認** (Astra のアーキテクチャ・位置づけに関する報道 (二次情報))
- [9] Anthropic 「公式発表 (Claude Opus 5 2026 年 7 月 24 日、Claude Fable 5 2026 年 6 月 9 日)」 **URL 未確認** (モデル構成・文脈長・effort 設定)
- [10] Axios 「Anthropic モデル関連報道」 **URL 未確認** (Opus 5 の性能・価格に関する報道 (二次情報))
- [11] TechCrunch 「Gemini 関連報道および AMI Labs 資金調達報道 (2026 年 3 月)」 **URL 未確認** (LeCun 新会社の調達額・評価額 (二次情報))
- [12] 9to5Google 「Gemini 3.6 Flash 報道 (2026 年 7 月 21 日)」 **URL 未確認** (トークン効率改善・値下げ (二次情報))
- [13] Bloomberg 「Gemini 3.5 Pro 延期報道」 **URL 未確認** (内部目標未達による広範リリース延期 (二次情報))
- [14] Elon Musk (X) ほかに複数報道「Grok 5 に関する言及 (6 兆パラメータ MoE、Colossus 2)」 **URL 未確認** (自己申告であり検証不能。延期が繰り返されている)
- [15] MarkTechPost 「中国オープンウェイトモデル解説記事」 **URL 未確認** (Kimi K3・DeepSeek V4-Pro・GLM-5.2 の仕様 (二次情報))
- [16] CSIS 「中国 AI モデルに関する分析」 **URL 未確認** (中国勢の戦略・体制分析)
- [17] Dario Amodei 「The Adolescence of Technology (2026 年 1 月)」 darioamodei.com (**正確なパスは未確認**) (「データセンターの中の天才国家」に至る時間軸に関する一次発言)
- [18] Dwarkesh Podcast 「Ilya Sutskever 回 (2025 年 11 月 25 日)、Dario Amodei 回」 **URL 未確認** (「スケーリングの時代の終わり」「継続学習は障壁ですらないかもしれない」等の発言)
- [19] Sherwood News 「Demis Hassabis の AGI 時間軸発言に関する報道」 **URL 未確認** (AGI 到達時期を 3~4 年へ短縮 (二次情報))

- [20] NVIDIA「GTC 2026 基調講演（2026 年 3 月 16 日）および Vera Rubin 製品情報」 *URL 未確認* (Rubin/Vera Rubin の仕様、累計 1 兆ドル受注予測)
- [21] Harmonic ほか「Aristotle: IMO-level formal reasoning (arXiv:2510.01346)」 <https://arxiv.org/abs/2510.01346> (Lean 4 による形式検証付き IMO 金メダル相当の成果)
- [22] Harmonic「公式発表（2025 年 7 月 28 日）」 *URL 未確認* (IMO2025 での成果、VERINA ベンチ 96.8%)
- [23] Lean FRO「Lean 4 プロジェクト」 *URL 未確認* (形式検証基盤)
- [24] Google DeepMind「公式発表（Genie 3、AlphaProof／AlphaGeometry、Gemini 3 Deep Think）」 *URL 未確認* (世界モデル・数学推論の到達点)
- [25] NVIDIA Research「Cosmos、GR00T N1.6、TTT-E2E（文脈の学習データ化）」 *URL 未確認* (フィジカル AI 基盤と継続学習研究)
- [26] Meta AI「V-JEPA 2」 *URL 未確認* (JEPA 系世界モデルの到達点)
- [27] Runway「Gen-4.5」 *URL 未確認* (動画生成・世界シミュレーション)
- [28] Huawei「Ascend 950PR／950DT／960 および Atlas 950／960 SuperPoD ロードマップ」 *URL 未確認* (中国国産アクセラレータの計画（2026 年 Q1～2027 年 Q4）)
- [29] TechInsights「Ascend シリーズ分析」 *URL 未確認* (国産チップの実性能評価（二次情報）)
- [30] TrendForce「HBM 市場予測」 *URL 未確認* (HBM 供給逼迫の見通し（二次情報）)
- [31] 米国商務省「議会証言（Kessler、2026 年 7 月）」 *URL 未確認* (対中 H200 実出荷が「ごく僅か」との証言)
- [32] CNBC「H200 対中販売許可、NVIDIA 決算、OpenAI コミット報道（2026 年 2 月 20 日ほか）」 *URL 未確認* (輸出規制の運用実態と資本コミットの再フレーム)
- [33] Financial Times「ハイパースケーラー設備投資集計および Jefferies Brent Thill のコメント」 *URL 未確認* (2026 年の設備投資合計約 7,250 億ドル、AI バブル論への反論)
- [34] Adnan Masood「中国オープンモデルの OpenRouter シェア分析（Medium、2026 年 7 月）」 *URL 未確認* (「30% 超のトークン量を 60～90%安い価格で」との定量分析（個人ブログ、要検証）)
- [35] Sakana AI「Introducing Sakana AI's Recursive Self-Improvement (RSI) Lab」 <https://sakana.ai/rsi-lab/> (自己改善研究の組織的取り組み)
- [36] WIPO／各国特許庁「AI 関連特許統計」 *未確認（本調査では参照していない）* (特許出願動向は別途確認を要する)