米中次世代フロンティアAI予測

2026年後半から2027年に予想されるブレイクスルー

調査基準日: 2026年8月15日（JST）

予測対象期間: 2026年後半から2027年末

作成: Manus AI

要旨

2026年後半から2027年にかけての最大の変化は、モデルの知識量や単発ベンチマークの上昇そのものではありません。**推論、ツール実行、外部検証、記憶、監視を結んだエージェント・システムが、限定環境で数時間規模の仕事を反復的に完了すること**が、最も有力なブレイクスルーです。[1] [2] 米国勢は最高性能の閉鎖モデル、巨大計算基盤、エンタープライズ製品、安全運用を統合する形で優位を維持しやすく、中国勢はオープンウェイト、MoE・蒸留・推論予算制御、低コストの自前配備、製造・ロボット・国内クラウドの垂直統合で差を縮める公算が大きいです。[3] [4] [5]

ただし、2027年までに「人間のように日単位で、どの業務でも安全に自律実行する汎用エージェント」が実現するという予測は支持しません。独立評価のMETRは、現在の時間地平を特定のソフトウェア・タスク群における**50%または80%成功率の難易度尺度**として定義しており、16時間を超える推定には現行タスクセットで信頼性の限界があると明示しています。[1] また、長期稼働するモデルは状態忘却、ツール誤操作、権限逸脱、コスト増、安全境界の探索といった新しい失敗モードを生みます。[2]

> **中核的な見立て:** 2027年の競争は「どちらがより賢い一つのモデルを作るか」から、**どちらがより低コストに、より長く、より検証可能に、より多くの現実のタスクを動かせるか**へ移ります。

領域	2026年後半〜2027年の最有力な突破点	米国勢の主な優位	中国勢の主な優位
長期エージェント	計画・実行・検証・再計画を反復する、限定業務向けの検証付きエージェント	閉鎖モデル、ツール製品、安全監視、企業SaaS統合	オープンウェイト、MCP、低コスト蒸留、自前配備
推論コスト	MoE、低精度、投機的デコード、動的推論予算による品質あたりコストの低下	GPU／TPU、ネットワーク、ランタイム、クラウド	MoE、MLA、モデル圧縮、国産スタック最適化

物理AI	世界モデル・合成データ・VLAをつないだ、限定環境の閉ループロボット	基盤モデル、シミュレーション、先端アクセラレータ	製造・物流現場、ロボット展開、具身データの規模
オープンモデル	準フロンティア能力のオンプレミス／自前推論化	Llama、AI2等による対抗	Qwen、DeepSeek、Kimi等によるより積極的な公開
安全・サイバー	高能力防御エージェントと、アクセス・評価・隔離の階層化	政府と研究所の安全評価・信頼アクセス	サービス層の標識、届出、コンテンツ統治
企業導入	ID、権限、監査、ロールバックを備えた限定的自律ワークフロー	グローバルSaaS・API・標準化	国内クラウド・EC・オフィス／開発基盤の垂直統合

1. 予測の読み方

本稿の確率は、2026年8月15日までに公開されたモデル発表、技術報告、独立評価、政府資料を基にした**分析上の主観確率**です。特定企業の未発表モデル名、発売時期、売上を予言するものではありません。また、「技術的に可能になること」と「高信頼で本番運用されること」は分けて評価します。

モデルの公表ベンチマークには、推論努力、エージェント・ハーネス、検索・コード実行などのツール、試行回数、タイムアウト、フォールバックが混在します。よって、以下では単一のスコア表で将来の勝者を決めず、**能力の方向性、実装経路、制約、反証可能な観測指標**を重視します。

2. 最も起こりやすい六つのブレイクスルー

ブレイクスルー仮説	米国勢	中国勢	実現確率	何が起きれば「実現」と言えるか
検証付き長期エージェント	75%	65%	高い	数時間規模のコード、調査、事務を、計画・実行・テスト・再計画で低い人間介入のまま完遂する
推論コストの構造的低下	85%	80%	非常に高い	同等品質に必要なトークン、GPU時間、電力、待ち時

				間が大きく低下する
準フロンティアのオープンウェイト化	55%	85%	中国で高い	自前配備モデルが推論・コード・ツール利用で閉鎖モデルに近づき、独立評価・企業導入で確認される
閉ループ物理AI	65%	70%	中程度	物流・製造等の限定環境で、世界モデルを使って数分規模のロボット作業を安全に継続する
高能力サイバー防御とアクセス階層化	70%	55%	中〜高	脆弱性発見、パッチ作成、検証を半自律化しつつ、隔離・監視・信頼アクセスが製品条件になる
監査可能な企業エージェントの本番化	75%	国内75%／国際45%	高い	複数の業務システムを横断して書込みまで行う有料運用が、承認・監査・ロールバック付きで定着する

2.1 検証付き長期エージェント

最も重要な突破点は、より長いChain-of-Thoughtではなく、**外部の検証器を含む反復ループ**です。モデルは計画を立て、ブラウザ・ターミナル・検索・データベース・社内SaaSを操作し、テスト、実行結果、複数案の比較を使って自らの仮説を訂正します。成功の中身は「モデルが一度に正解すること」ではなく、失敗しても壊さずに再計画できることです。

OpenAIは、長期モデルの実運用では既存の事前評価に含まれない失敗が出たため、単発アクションではなく全軌跡を監視し、セッションを一時停止して利用者へ警告する仕組みを導入したと報告しています。[2] この事実は、能力向上と安全・信頼性が同じ問題になることを示します。米国勢は、閉鎖モデル、ツール群、企業プロダクト、安全チームを同時に最適化できるため、**高価だが高信頼な限定業務エージェント**で先行する可能性が高いです。

中国勢は、Qwenが公表する思考／非思考の切替、推論予算制御、MCP対応、環境フィードバックを用いた長期RLの方向性と、DeepSeek系で示されたMoE・強化学習・蒸留の蓄積を組み

合わせるでしょう。[4][5] こちらの本命は、最も複雑な一件で常に首位を取ることより、**検証可能なコード・数学・構造化業務を低い推論コストで大規模展開すること**です。

しかし、METRの時間地平は「実時間の自律時間」ではなく、特定タスクでの成功確率を人間専門家の所要時間に写像したものです。[1] 2027年に日単位の高信頼自律性が成立したと判断するには、80%成功率の時間地平だけでなく、長期ロールアウトの指示保持、ツール誤操作、回復率、権限逸脱、総コストの独立測定が必要です。

主要な観測指標	望ましい変化	予測を弱める変化
80%成功率の時間地平	数時間へ継続的に延びる	50%のみ伸び、80%が横ばい
回復率	テスト失敗や環境変化後に自動再計画できる	再試行が無限ループ化し、人間が常時介入する
操作の安全性	ロールバック、最小権限、監査が標準化する	誤操作・データ漏えい・未承認書込みが増える
総コスト	同じ成功率に必要なトークンとツール呼出しが下がる	推論予算の増大で実務コストが許容範囲を超える

2.2 推論コストの構造的低下

2027年までに最も確実性の高い進展は、**能力あたりの推論コスト低下**です。米国では、低精度演算、大容量メモリ、GPU間高速接続、推論ランタイムを一体化した計算基盤が、MoE、長文脈、マルチエージェントの反復推論を実用的な価格帯へ押し下げます。中国では、MoEの高い疎性、注意機構とKVキャッシュの圧縮、蒸留、量子化、投機的デコード、異種アクセラレータを束ねるソフトウェア最適化が中心になります。

DeepSeek-V3は、671B総パラメータのうちトークンあたり37Bだけを活性化するMoE、Multi-head Latent Attention、補助損失を使わない負荷分散、多トークン予測を採用し、効率性を設計の中心に置きました。[3] Qwen3も、MoE、推論予算制御、より小さな活性パラメータでの効率を公表しています。[4] これらは中国側が最先端GPUの絶対量だけでなく、**計算量あたりの能力**で競争する強い動機と能力を持つことを示します。

米国側は、Blackwell級のGPU群、TPU、ネットワーク、コンパイラ、クラウド配備を共同最適化できる強みがあります。ただし、FP4級の低精度化がすべてのモデル・層・タスクで品質を維持するわけではなく、HBM、電力、液冷、データセンター建設、通信が制約になります。中国側も、国産アクセラレータがベンチマーク、供給量、電力効率、ソフトウェア互換性でどこまで実運用を支えられるかが未確定です。

> **予測:** 2027年の競争で最も目立つ数字は、パラメータ数ではなく、**1ドル・1ワット・1秒あたりに完了できる検証済みの作業量**になるでしょう。

2.3 オープンウェイトの準フロンティア化

中国勢に最も有利なブレイクスルーは、モデルを公開すること自体ではなく、**オープンウェイト、推論予算制御、蒸留、エージェント・ハーネスを一緒に配布すること**です。Qwen3はApache 2.0のモデル群、思考／非思考のハイブリッド、119言語、MCPを含むツール使用を公表し、将来は環境フィードバックを伴う長期RLを拡張する方針を示しています。[4] 2026年夏のQwen3.8-Max、Kimi K3、DeepSeek-V4-Pro-0813と合わせて見ると、中国の主戦場は、閉鎖型の最高峰に一点で勝つことより、性能に近いモデルを開発者・企業・国内クラウドへ迅速に広げることです。

米国にもLlamaやAI2のような対抗軸はあります。しかし最上位の能力は、安全性、収益性、重み保護、サイバー能力を理由に閉鎖型へ寄りやすいため、完全なフロンティア重みの公開は中国ほど一貫しない可能性があります。米国勢の本命は、単一GPUや少数GPUで動く、マルチモーダル・長文脈・ツール利用可能な**準フロンティアの自前運用スタック**です。

オープンモデルが本当に勝つ条件は、公開時のベンチマークではありません。量子化後の性能、推論ランタイムへの対応、長期エージェントでの回復能力、オンプレミス導入、第三者再現、モデル来歴・安全性の整備が必要です。総パラメータが1〜3T級のモデルでは、重みが公開されても、メモリ、通信、運用人材の負担が依然として大きい点にも注意が必要です。

2.4 世界モデルと物理AI

世界モデルは、映像をもっとリアルに生成する技術ではありません。次の段階では、視覚、言語、行動、音、ロボット状態から複数の未来を予測し、実機を動かす前に方策を比較する**物理世界のテスト環境**になります。

米国側では、Gemini Robotics 2が全身VLA、長期タスクを扱うEmbodied Reasoning、オンデバイスVLAを分け、数分・数百判断のタスク、失敗時の自己修正、複数ロボット協調、新しい身体への適応を打ち出しています。[6] NVIDIA Cosmos 3も、物理に基づく世界シミュレーション、World Action Modelの事後学習、合成データ生成を一つのスタックに置いています。[7]

中国側のQwen-RobotWorldは、8.6Mの動画テキスト、2億超フレーム、20超の身体、500超の行動カテゴリを使い、操作、自動運転、屋内ナビゲーション、人からロボットへの移転を一つの言語条件付き世界モデルで扱います。[8] 中国は製造、物流、ロボットの現場展開を生かし、特定の身体や業務へのデータ閉ループを速く回せる可能性があります。したがって、2027年の中国勢の優位シナリオは「家庭用汎用ロボット」でなく、**工場・倉庫・検査・配送のような制約された現場での実効性能と配備台数**です。

ただし、現段階の実機評価はなお限定的です。Googleの公開値でも、多指の難しい作業は32〜44%程度の成功率のものがあります。[6] 動画予測のもっともらしさは、接触、摩擦、柔軟物、力、遅延、人間近接に対する実機安全性を保証しません。2027年に実質的なブレイクスルーが起きたと評価するには、シミュレーションから実機への移転率、未見環境での連続成功率、異常時の停止、時間当たりの安全稼働が公開される必要があります。

2.5 サイバー能力の高度化と安全統制

サイバー領域では、モデルは脆弱性の説明から、発見、再現、パッチ作成、検証、監視をつなぐ防御エージェントへ近づきます。米国では、2026年6月の大統領令が、高度なAIサイバー能力を測る機密ベンチマークと「covered frontier model」の閾値を策定し、信頼パートナーへの早期アクセスを含む任意枠組みを設計するよう命じました。[9] 同時に、重要インフラにおけるAI支援の脆弱性発見・検証・修正を促しています。[9]

これは、2027年に高能力モデルが防御側で価値を出す可能性を高めますが、危険な能力がすでに一般公開されることを意味しません。むしろ、隔離実行、ネットワーク・ツール制限、重み保護、全軌跡監視、信頼アクセスが製品能力の一部になる可能性が高いです。[2] [9]

中国では、生成AIサービスを対象に、明示・非明示のコンテンツ標識、ファイル・メタデータ、配信プラットフォームでの確認、ログの保持が制度化されています。[10] この枠組みは、開放モデル自体よりもサービス提供・コンテンツ流通の統治に強く働くと考えられます。両国で安全対策の重点は異なりますが、共通するのは、**能力が上がるほど「モデルをどこまで開放し、どのツール権限を与え、何を監査するか」が能力評価と不可分になる**ことです。

2.6 企業エージェントの本番化

2027年に最も早く大きな価値を生むのは、完全自律の汎用秘書ではなく、成功条件が機械的に検証しやすい半自律ワークフローです。具体的には、ソフトウェア開発、テスト・レビュー、顧客対応、文書・表の照合、調査、CRM更新、請求例外処理、セキュリティ運用などです。

米国では、NISTが業界主導のエージェント標準、相互運用プロトコル、認証・アイデンティティ、マルチエージェントの安全な相互作用、セキュリティ評価を重点化しています。[11] これは、企業導入の制約が基盤モデルの知能だけでなく、**誰が誰として振る舞い、どの権限を持ち、何を行ったかを追跡できること**にあると示します。

中国側では、クラウド、EC、業務ソフト、開発環境、端末を同一事業者のエコシステムで結べる点が強みです。低コストの自前配備モデルと合わせ、国内市場では複数エージェントの有料本番化が速く進む可能性があります。ただし、国際展開ではデータ主権、コンテンツ標識、標準互換性、信頼・調達要件が追加の制約になります。

3. 米中の競争構造はどう変わるか

米国勢の本命シナリオ

米国勢は、最上位の閉鎖モデル、GPU／TPU・クラウド、企業SaaS、サイバー・安全研究、グローバルな販売・サポートを結び、**高価でも高信頼なエージェント基盤**を作るでしょう。重点は、最高難度のコード、研究、企業横断業務、サイバー防御、世界モデルの大規模事後学習です。2027年に優位を決めるのは、モデルの会話能力ではなく、長期実行での品質・回復・安全・監査をSLAとして提供できるかです。

中国勢の本命シナリオ

中国勢は、オープンウェイト、MoE、蒸留、動的推論予算、国産推論スタック、国内クラウド、製造・ロボットを組み合わせ、**十分に高い能力をより安く・より多くの現場に配る**ことで優位を取るでしょう。閉鎖モデルの絶対的な最高性能を常に超える必要はありません。実際の仕事を、低い推論コストで、必要なデータ主権とカスタマイズ性を持って動かせれば、特に国内市場と自前配備が重要な地域で強い競争力になります。

「半年遅れ」論の今後

前回の調査で示した通り、米中差を固定した月数で表す見方はすでに弱くなっています。2027年にはさらに、次の三つの「フロンティア」が分離するため、平均遅れという表現は一層不正確になるでしょう。

フロンティア	競争の中心	予想される優位の偏り
閉鎖型最高性能	最高難度の推論、長期エージェント、安全な高能力運用	米国勢がなお優勢の可能性
オープン／自前配備	コスト、カスタマイズ、国内・業界別統合、開発者普及	中国勢が優勢の可能性
物理世界・産業展開	ロボット台数、現場データ、シミュレーション→実機、垂直統合	領域ごとに拮抗。中国は製造展開、米国は基盤モデル・計算で強み

4. 予測を検証するための2027年チェックリスト

観測項目	強気シナリオを支持するサイン	慎重シナリオを支持するサイン
長期エージェント	80%成功率が数時間規模に伸び、再計画・ロールバックが実運用で機能	50%成功率だけ上がり、長期タスクでは誤操作・監督負荷が残る
コスト	同一品質の総推論単価、レイテンシー、電力が一段下がる	MoEや推論時スケーリングで、隠れた通信・運用費が増える
オープンモデル	第三者評価、量子化、オンプレミス導入、派生エコシステムが急増	ベンチマーク再現が弱く、安全・ライセンス・運用で採用が止まる
物理AI	世界モデルから得た方策が未見実機にも転移し、数分規模の連続作業を安全に行う	映像デモは増えるが、実機成功率・速度・人間近接安全が改善しない

サイバー	防御成果と安全統制が同時に第三者評価で確認される	能力だけが先行し、公開評価・監査・ガードレールが追い付かない
企業導入	有料本番運用、継続率、監督削減、エラー・ロールバック率が開示される	PoCは増えるが、権限管理・責任分界・データ品質で本番化しない

結論

2026年後半から2027年に予想される本当のブレイクスルーは、モデルがより雄弁になることではなく、**モデルが検証可能な形で仕事を進め、失敗を修正し、現実のシステムと安全に接続されること**です。

米国勢は、最先端計算、閉鎖型モデル、製品統合、安全・防御の制度化により、最高難度の高信頼エージェントで先行する可能性が高いです。中国勢は、オープンウェイト、効率化、蒸留、産業データ、国内の垂直統合により、能力あたりのコストと展開速度で強い位置を取るでしょう。したがって、2027年の米中競争は勝者総取りではなく、**「最高性能」「低コストでの大規模配備」「物理世界の産業実装」の三層で異なる首位が生まれる**可能性が最も高いと判断します。

参考文献

[1] METR, Task-Completion Time Horizons of Frontier AI Models
[2] OpenAI, Safety and alignment in an era of long-horizon models
[3] DeepSeek-AI, DeepSeek-V3 Technical Report
[4] Qwen Team, Qwen3: Think Deeper, Act Faster
[5] Qwen, Qwen3.8-2.4T-A95B model card
[6] Google DeepMind, Gemini Robotics 2 brings whole body intelligence to robots
[7] NVIDIA, Cosmos 3
[8] Qwen-RobotWorld Technical Report
[9] White House, Executive Order 14409
[10] 中国国家インターネット情報弁公室、人工知能生成合成内容标识办法
[11] NIST, AI Agent Standards Initiative