「半年遅れ」の終焉：米中AIフロンティアモデル最新比較（2026年夏）

米中AIモデルの格差は「一律の時間の遅れ」から「領域別の専門化・競合」へと変化。非一様な競争の時代へ。

米国 (US)

中国 (China)

性能比較：逃げ切る米国

総合性能と複雑な推論は依然として米国優位

長期ソフトウェア工学や深い車用能力では、GPT-5.6 SolやFable 5が依然として基準点。

配備戦略：閉鎖型エコシステム

成熟した閉鎖型API提供。低コストでの高い統合性と安定性を重視し、企業向けサポート環境に強み。

米国勢：成熟した閉鎖型API提供

低コストでの高い統合性と安定性を重視し、企業向けサポート環境に強み。

主要ベンチマークにおける米中代表モデル性能比較

端末コーディング（Terminal Bench）

米国代表	(GPT-5.6 Sol / Fable 5):	88.0 - 88.8
中国代表	(Kimi K3 / GLM-5.3等):	86.6 - 88.3 (ほぼ同等)

長期ソフトウェア工学（DeepSWE）

米国代表	69.7 - 72.7
中国代表	76.6 - 67.5 (米国優位)

深い悪用能力（ExploitBench）

米国代表	76.5 - 78.0
中国代表	28.8 - 54.4 (大きな差)

性能比較：追いつく中国

コーディング・ツール利用での「収束」

Terminal Bench等の街標で、Kimi K3やGLM-5.3が米国最上位モデルと同水準に到達。

配備戦略：オープンウェイト

オープンウェイトによる制御可能性。オンプレミス配備やデータ主権の確保、独自エージェント構築における自由度で優位。

2.7%

米中格差は「2.7%」まで接近。

Stanford HAI報告（2026年3月）

巨大モデルの運用コストが新たな壁。

中国勢の重み公開モデルは1.7T～2.8Tパラメータと巨大で、適用には高度なインフラが必要。