

2026年 AI勢力図の激変：中国製LLMの台頭と日本企業の生き残り戦略

米国製LLM

最高水準

\$20.00 (1M tokens) - Claude Fable 5



ソフトウェア開発と
エージェント機能で
米国に並ぶ

Qwen3.7 MaxやGLM-5.2は、
コーディングベンチマークで
Claude Opus等の米国トップモ
デルと互角の成績を記録。



100万トークンの
超長文脈と圧倒的な
安定性

米国製が長文で幻覚を起こす中、
中国製は1Mトークンの大規模コー
ド解析でも高い一貫性を維持。

中国製LLM

高水準

\$3.75 (1M tokens) - Qwen3.7 Max



破壊的なコスト
パフォーマンス

DeepSeek V4 Pro等の中国製モ
デルは、米国製と比較して数分の
一から数十分の一の価格を実現。



日本企業の戦略的処方箋:
リスク回避と活用の両立

日本企業

「バーベル戦略」によるタスクの振り分け

高機密タスク



米国製や
自社ホストモデル

大量の定型処理



安価な中国製APIへ
自動ルーティング



オープンウェイトモデルの自己ホ
スト (Local Deployment)

GLM-5.2等を自社VPC内で稼働させるこ
とで、中国国家情報法によるデータ流出
リスクを物理的に遮断。



抽象化層 (Sakana Fugu等) の活用

特定のプロバイダーが規制で停止しても
ダウンタイムなしで切り替えられる
マルチモデル基盤を構築。