

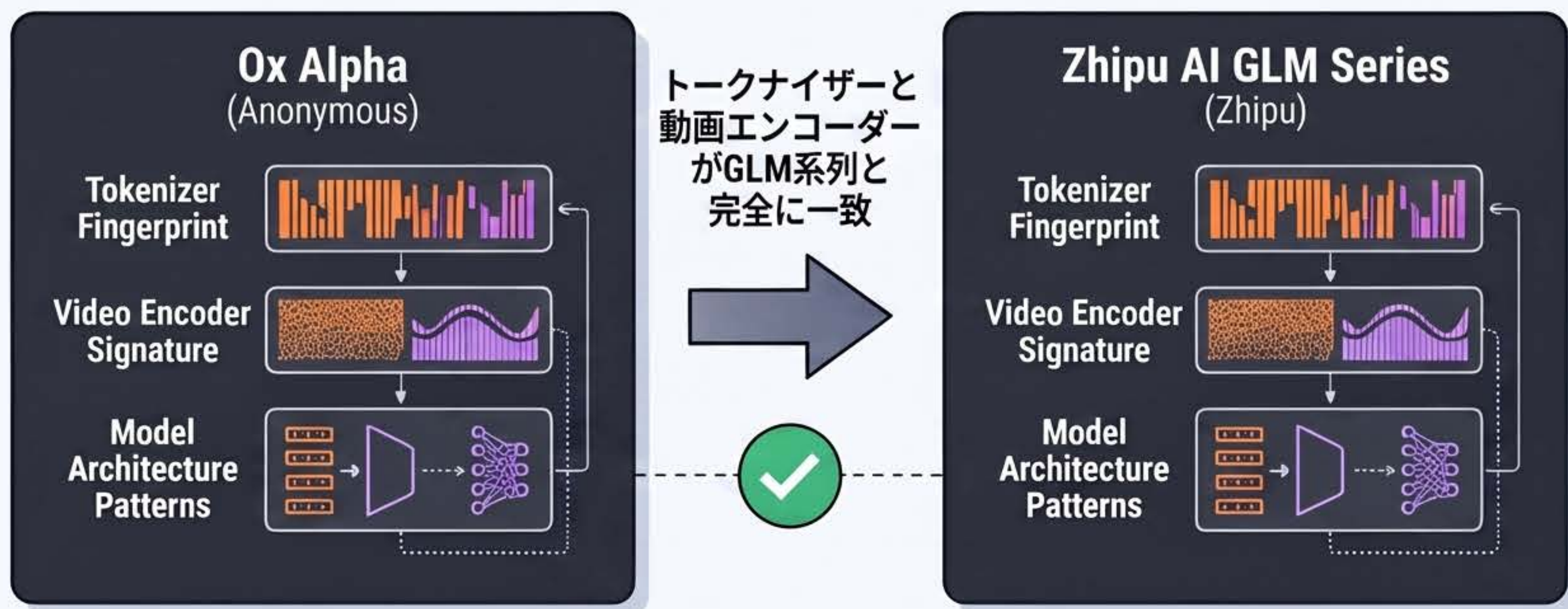


徹底解剖：謎のAI「Ox Alpha」 — 智譜GLM説と性能の真実

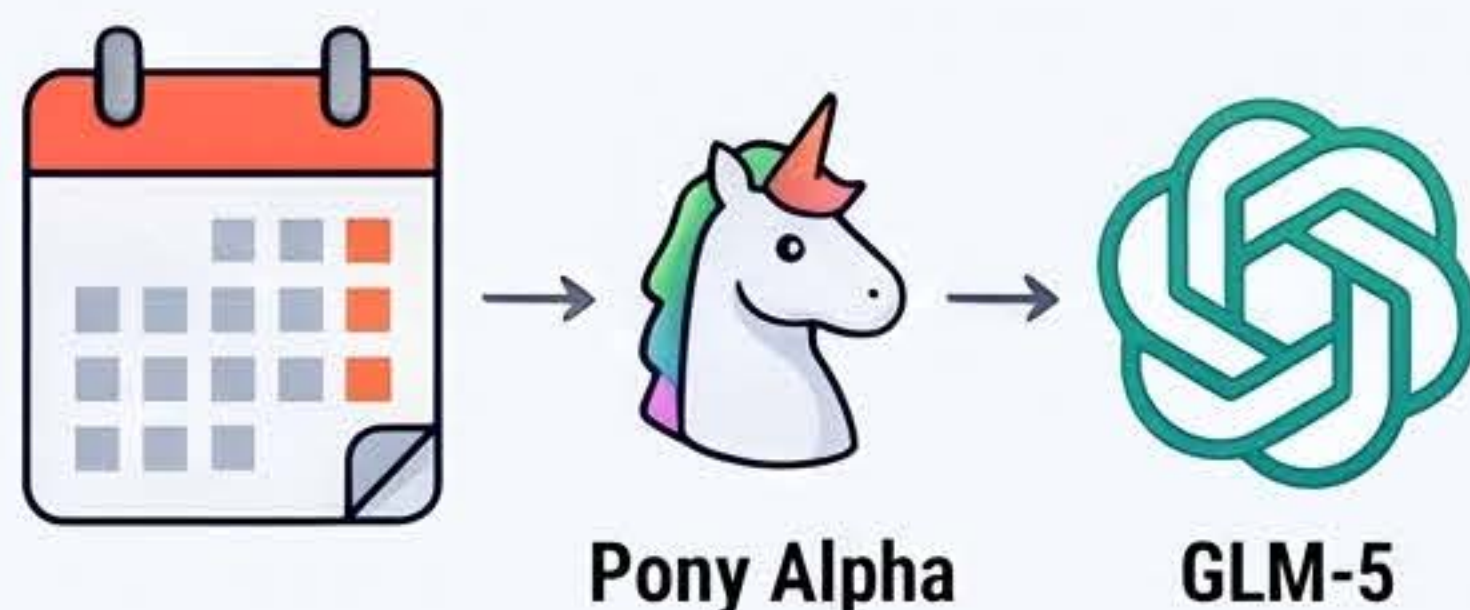
2026年8月、OpenRouterに突如現れた匿名AIモデル「Ox Alpha」。本インフォグラフィックでは、技術的指紋から浮上した中国・智譜（Zhipu AI）製という有力説と、物議を醸している「GPT-5.6超え」のコーディング性能の借憑性を検証します。



帰属の謎：Ox Alphaは「智譜」の次世代フラッグシップか？

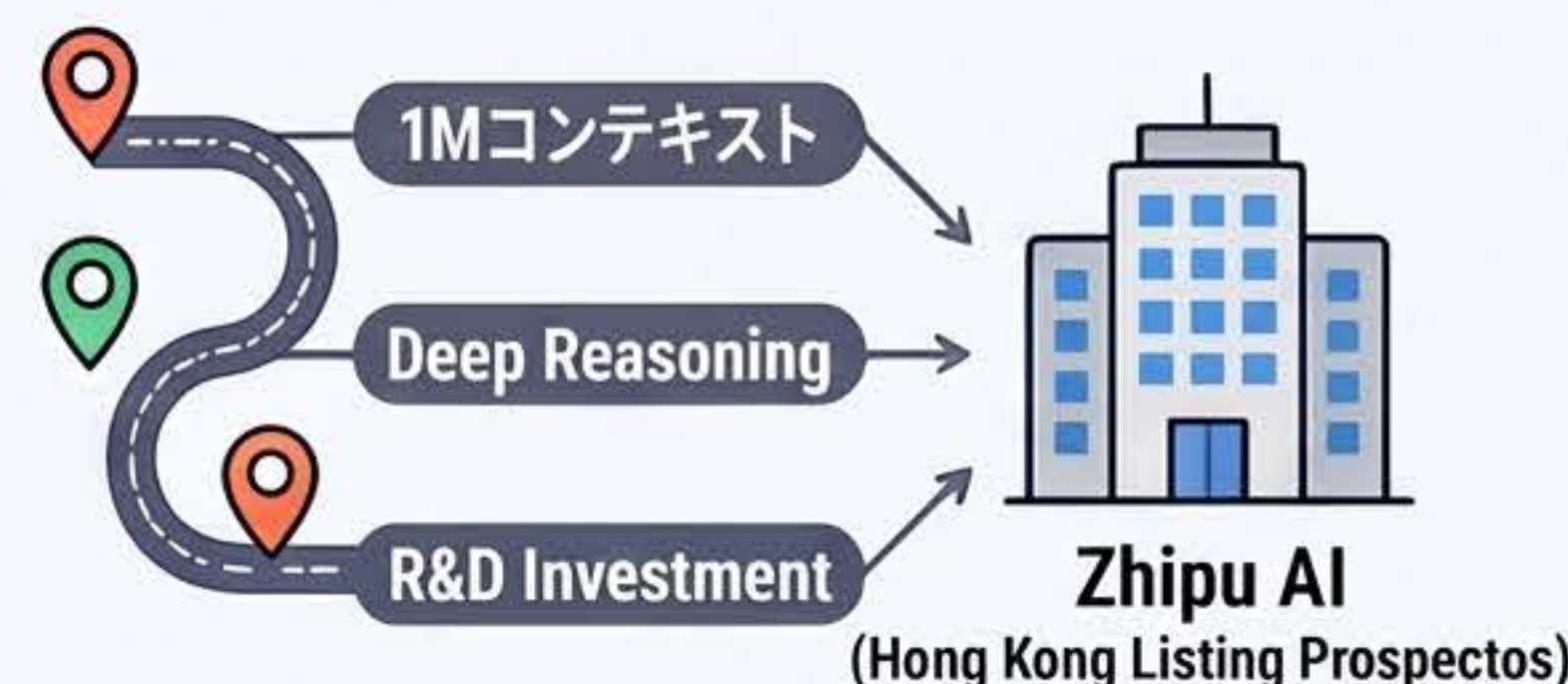


開発者の分析により、智譜独自の指紋 (fingerprint) がOx Alphaから複数検出されている。



匿名モデルを先行評価する 智譜の「前例」が存在

かつての匿名モデル「Pony Alpha」が後に公式にGLM-5であったと明かされた経緯がある。

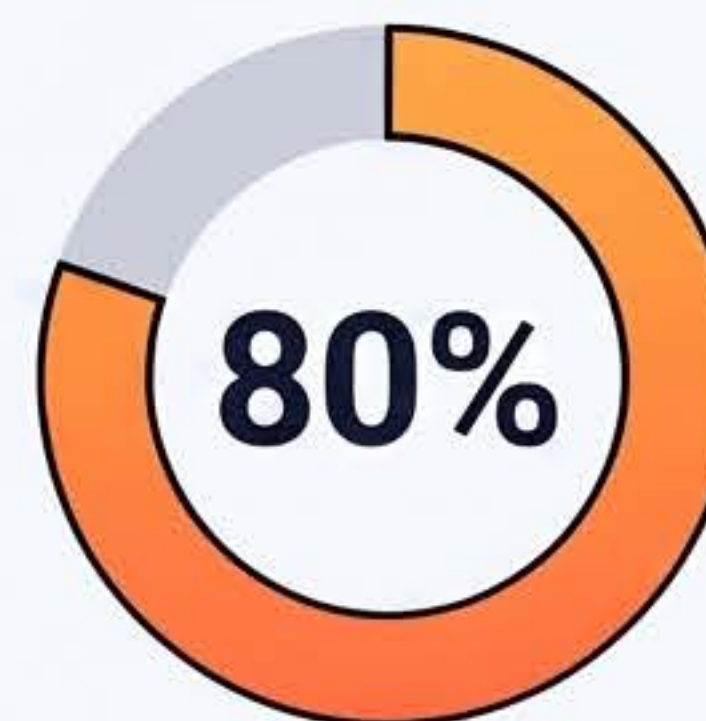


智譜のR&Dロードマップとの 高い技術的整合性

1MコンテキストやDeep Reasoningへの投資は、同社の香港上場目録見込に明記されている。

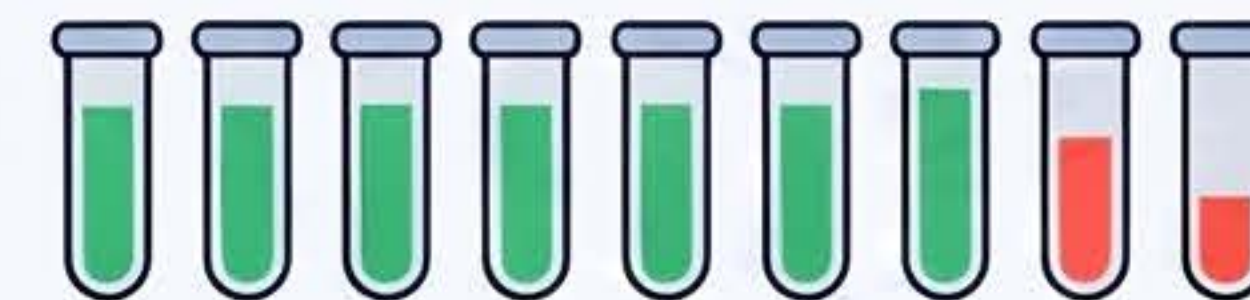


性能の真実：GPT-5.6やClaudeを本当に上回ったのか？



80%

「80%の成功率」は
統計的に不十分な
小規模試行



⚠ わずか10問のテストに基づく結果であり、95%信頼区間ではGPT-5.6との有意差は認められない。

公式ベンチマークでは 依然としてGPT-5.6が首位

113の課題を扱うDeepSWE公式ランキングでは、GPT-5.6 SolがOxを上回る精度を維持している。

DeepSWE公式ランキング (113課題)



主要フロンティアモデル比較

モデル	コンテキスト長	動画入力	Coding評価 (DeepSWE等)
Ox Alpha	1,048,576	対応	未掲載 (10問中80%)
GPT-5.6 Sol	1,050,000	非対応	73±3% (首位)
Claude Fable 5	1,000,000	非対応	70±4%