

突如現れた謎のAIモデル 「stealth/ox-alpha」の 正体を暴く

コミュニティによるブラックボックス検証
が示す「GLM」系列との技術的符合

DATA LITE

ACTIVE STREAM

DATA NODE

EVIDENCE

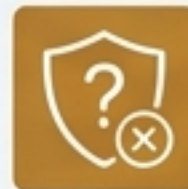
2026年8月20日、OpenRouter上に未知のモデルが突如出現した



2026.08.20

モデル「stealth/ox-alpha」がOpenRouterにて期間限定の無料プレビューとして突如公開。

ORIGIN: ANONYMOUS
PLATFORM: OPENROUTER
STATUS: PREVIEW



提供者の正体

OpenRouter公式は「匿名の第三者が開発・運用している」と明言。プラットフォーム自身は開発に関与していない。

DEVELOPER: UNKNOWN_ENTITY
VERIFICATION: PLATFORM_CONFIRMED
ROLE: THIRD_PARTY_DEPLOYMENT



初期の反響

プログラミングや長時間のエージェント型作業に特化しているとされ、AI開発者コミュニティで即座に解析対象となる。

SPECIALIZATION: CODING, AGENT_WORKFLOW
COMMUNITY: IMMEDIATE_ANALYSIS
IMPACT: HIGH



公開仕様が示す、極めて強力な基盤モデルの輪郭



コンテキストウィンドウ (Context)

1,048,576 トークン
(約100万コンテキスト)



出力能力 (Output)

最大 131,072 トークン



モダリティ (Modality)

入力：テキスト、画像、動画対応
出力：テキストのみ

※音声入力は非対応



機能と制約 (Features)

ツール呼び出し (Function Calling)
および構造化出力に対応。
ただし「JSON Schemaの厳格な強制」
には非対応。



処理速度と規模から浮かび上がる「Z.ai」の影

観測された処理速度

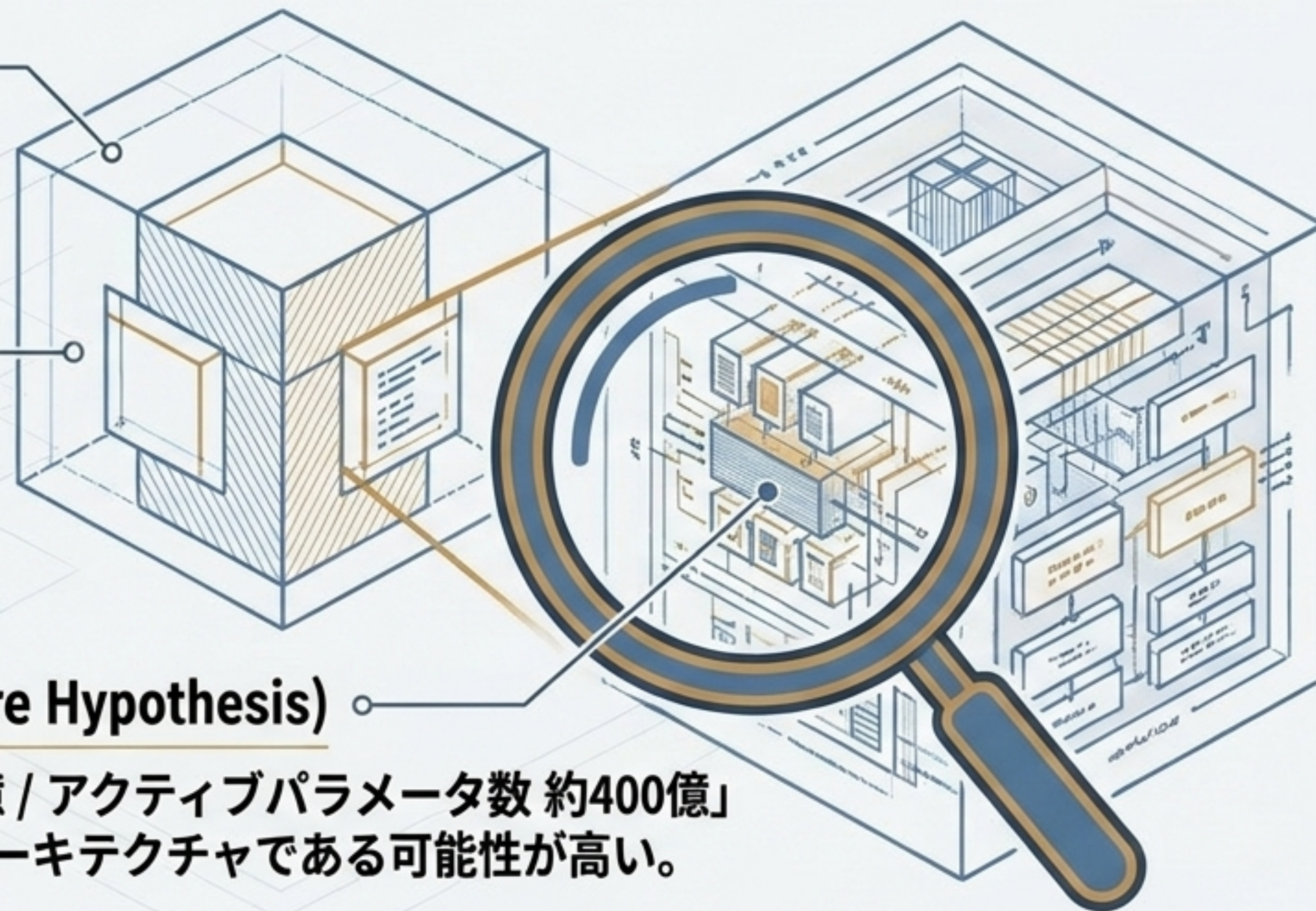
外部計測で約 49~78 tokens/sec
を記録（※負荷により変動）。

速度に基づく推定

このデコード速度は智譜AI（Z.ai）
の「GLM-5V-Turbo」と約6%の差
に留まる。

コミュニティの仮説構造 (The Core Hypothesis)

このモデルは「総パラメータ数 7,440億 / アクティブパラメータ数 約400億」
クラスの Mixture-of-Experts (MoE) アーキテクチャである可能性が高い。



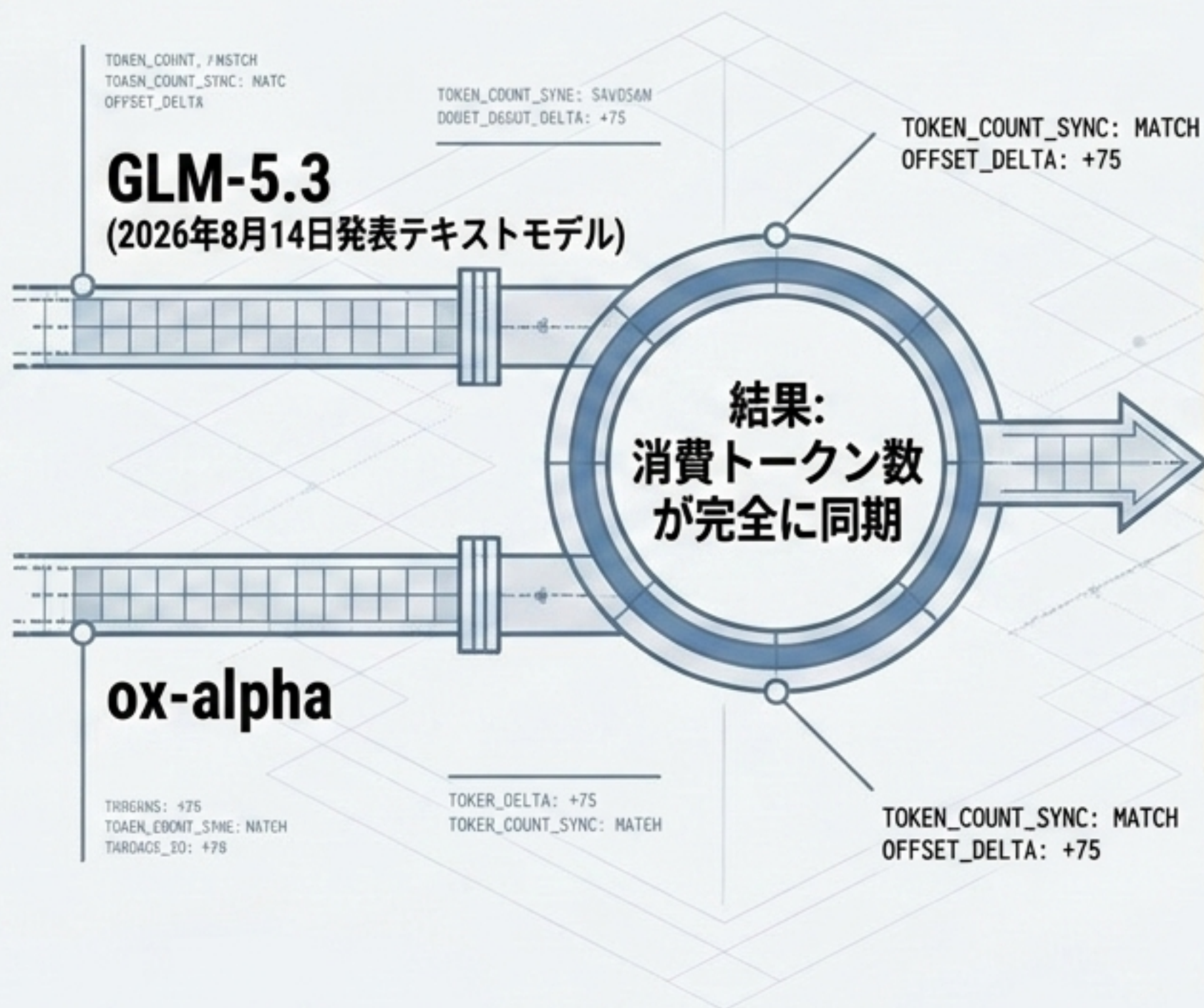
[WARNING FLAG] ※匿名の提供者は「1日100兆トークンの供給能力」を掲げているが、これが実処理量か理論値かは不明。



証拠① トークナイザーの完全な同期と「+75」の謎

検証条件

英語、ドイツ語、中国語、コード、絵文字等を含む25種類の異なるプロンプトを入力。



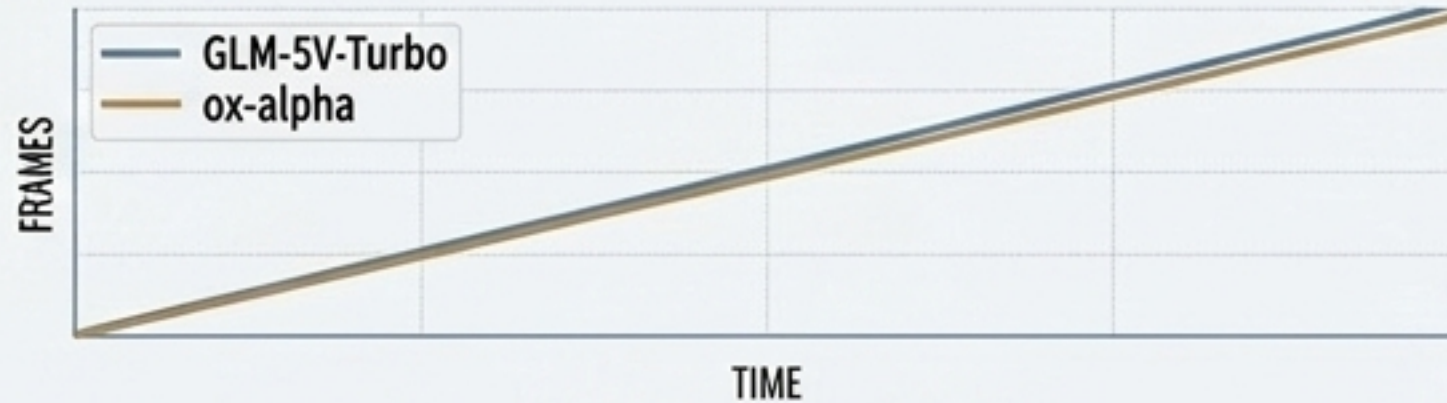
固定オフセットの存在

唯一の違いは、ox-alphaが常に「+75トークン」多く消費すること。

これは隠れたシステムプロンプトやAPIラッパーによるオフセットと推測される。

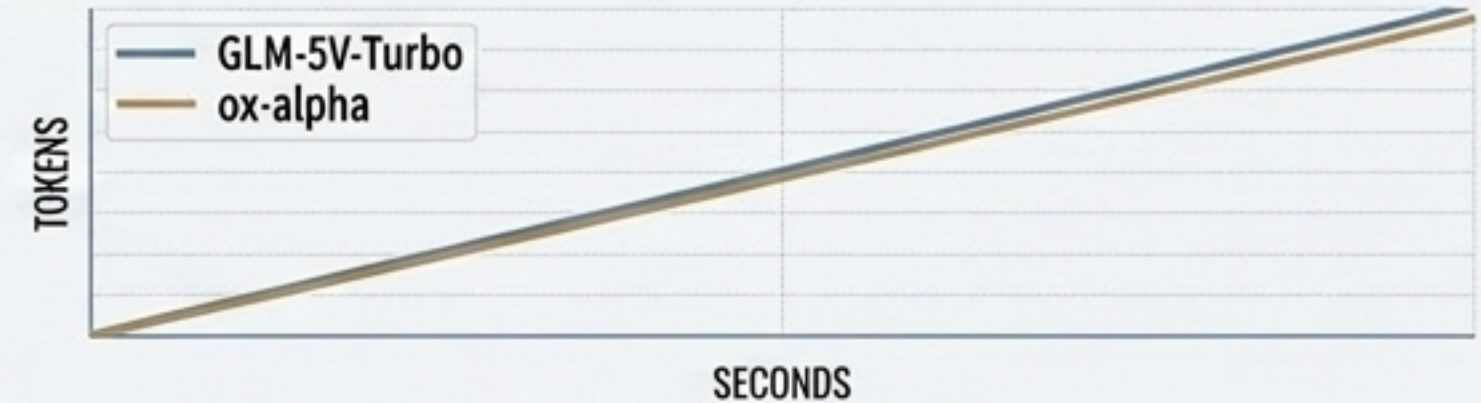
証拠② 動画エンコーダーにおける特異な処理パターンの合致

フレームサンプリング



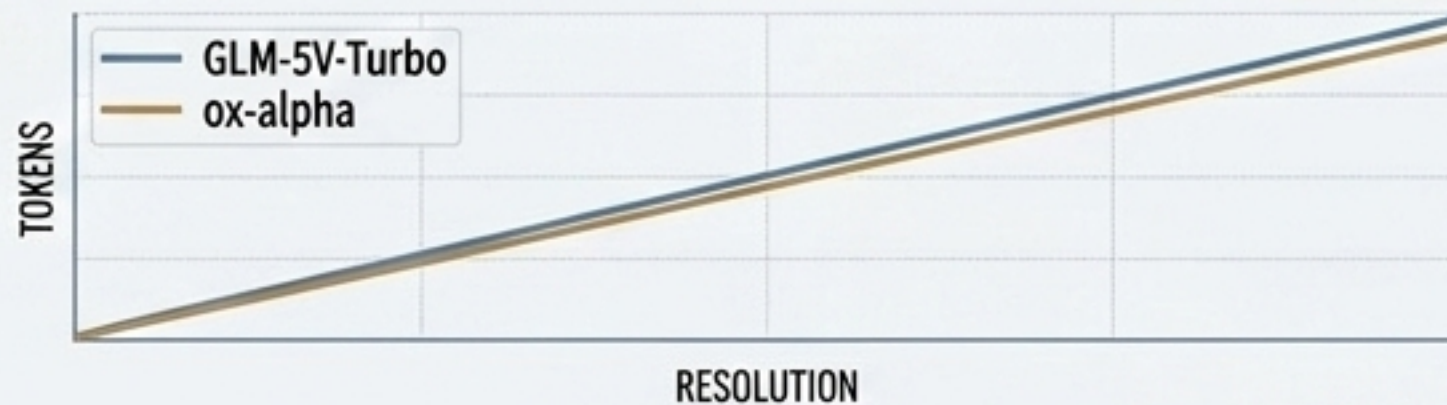
フレームレート（FPS）に依存しない抽出処理。

時間あたりの消費量



動画長1秒あたり「約147トークン」の一定増加。

解像度の影響



フレーム解像度の変化に比例した特有のトークン変動パターン。

[CONCLUSION]

4条件すべてで「GLM-5V-Turbo」の動画処理前基盤と挙動が完全に一致。



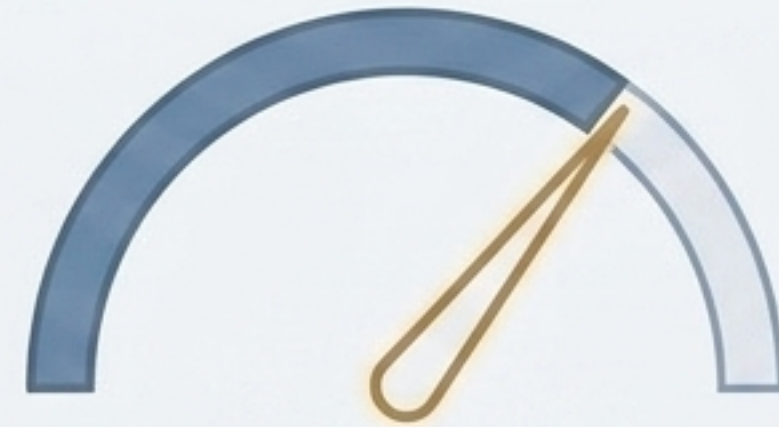
証拠③ 意図的なモダリティ制約と中国系モデル特有の出力傾向



[AUDIO INPUT:
DISABLED]

音声入力の非対応は、比較対象である
GLM-5V系の仕様と正確に整合する。

出力スタイルの癖 (Output Style)



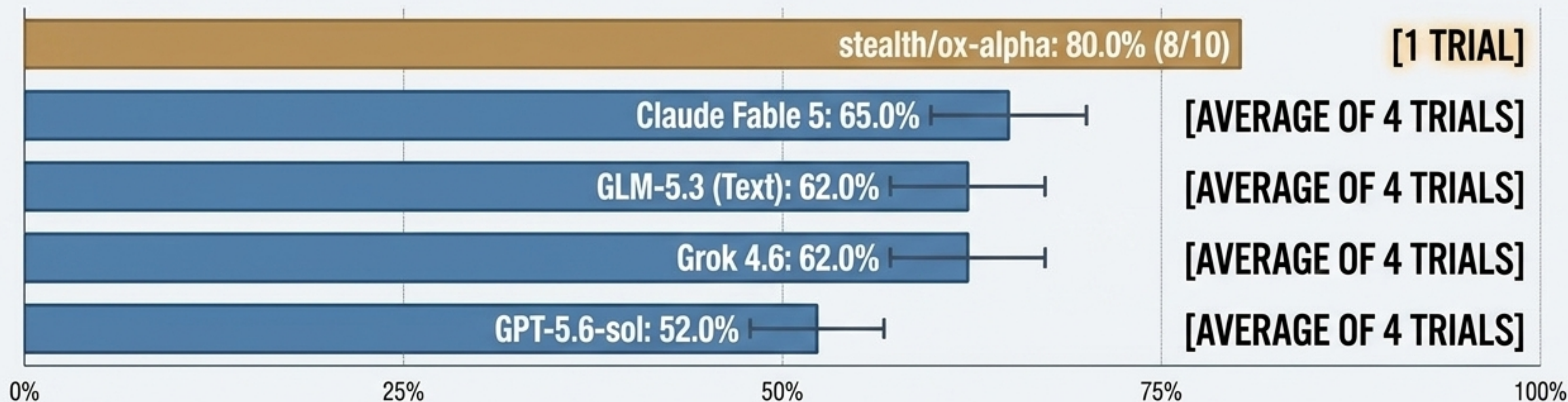
生成テキストにおける絵文字の出現頻度：
1,000文字あたり約1.3個

Claude, GPT-5.6, Grok等と比較して明らか
多く、GLMやQwenなど中国系AIモデルに
特有の文体的特徴（癖）と合致。



DeepSWE評価における「80%達成」の統計的カラクリ

シード0で選定された10タスクの比較



Analytical Insight

小規模試行（限定10タスク）における強力な性能は事実だが、「各1回試行」と「各4回試行の平均」を直接比較している点に留意が必要。モデル全般の優位性を確定するものではない。



マイクロケーススタディ：難解なリファクタリングを一発完遂

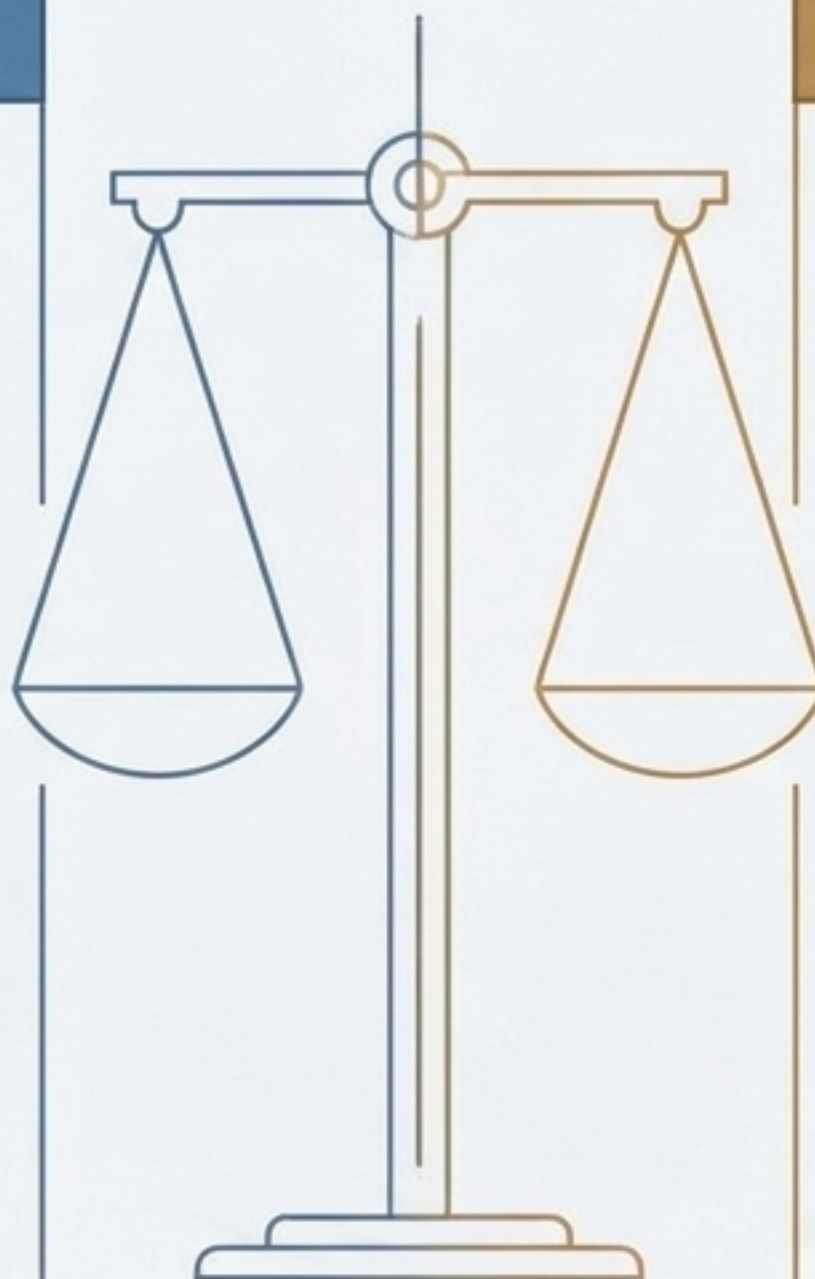
Context	Result
ターゲット JavaScriptパーサー「meriyah」の meriyah-explicit-resource- declarations タスク。 他モデルの結果 GLM-5.3, GPT-5.6-sol, Grok 4.6 はいずれも4回の試行すべてで解決でき ず（未解決）。	ox-alphaの結果 初回の試行でコードのリファクタリング を完璧に完遂。局所的な論理構築力に おいて突出した能力を示した。  51,469 TESTS PASSED 回帰テストを一発通過



エージェント能力の二面性：それは「創発」か「記憶」か？

The Emergence (創発的知能)

- 69回連続のツール呼び出しを伴うマルチステップ・タスクを完遂。
- エラーや再試行ループを一度も発生させない極めて高い安定性を実証。



The Memorization (記憶の再現)

- 話題となった「WebGL流体シミュレーションのGPU加速コード」の一発生成。
- 指摘事項：既存の著名OSS (paveIdogreatによる実装) とコードが極めて類似。未知の課題解決ではなく、学習データに含まれたコードの「再現 (記憶)」である可能性を排除できない。

9/339

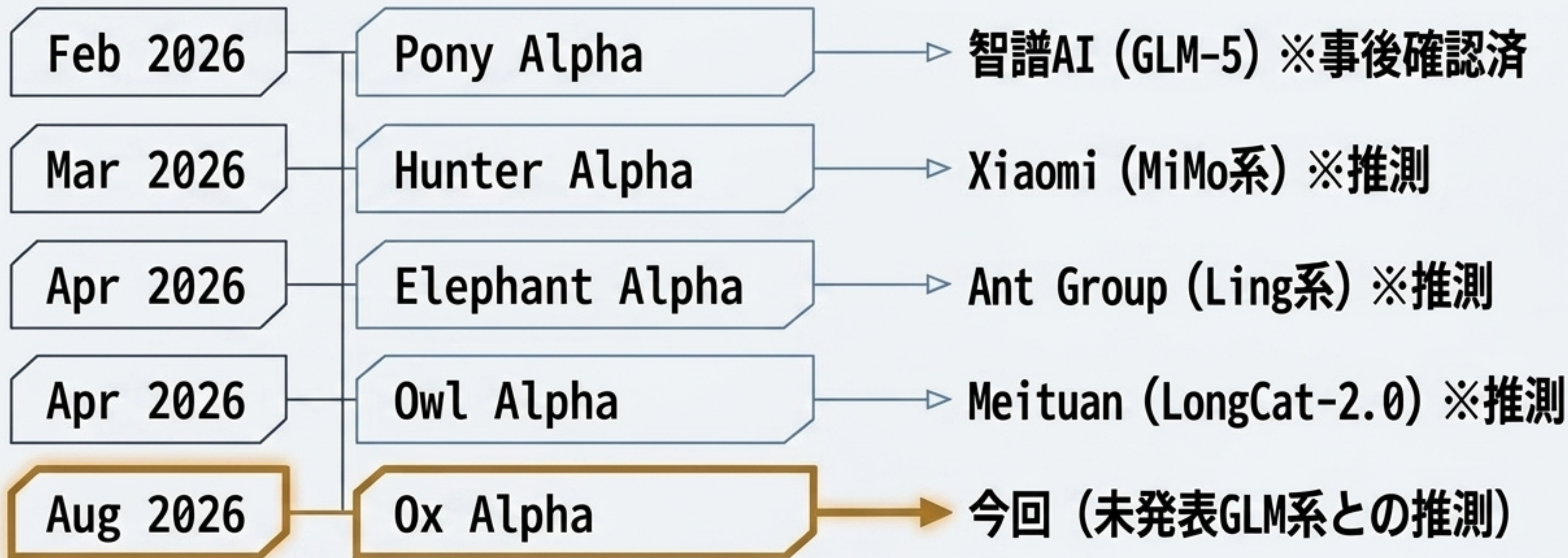
67/8

▶▶▶▶▶

■■■■■



中国AIラボに共通する「Animal Alpha」の匿名テスト戦略



Context

過去に「Pony Alpha」がGLM-5であったという事実が、今回の「Ox (牛) Alpha」も智譜AIのテストであるという推測の強力な背景となっている。



開発元と推定される「Z.ai」の現在地と開発タイムライン

Profile: 智譜AI (Zhipu AI)

- 2019年清華大学の技術を基盤に設立。
- 2025年に海外向けブランド「Z.ai」へ変更。

Inference (推論)

GLM-5.3 (テキスト版) の発表直後というタイミングは、次期「マルチモーダル版」のパブリックテスト (ox-alpha) を行う時期として極めて自然である。

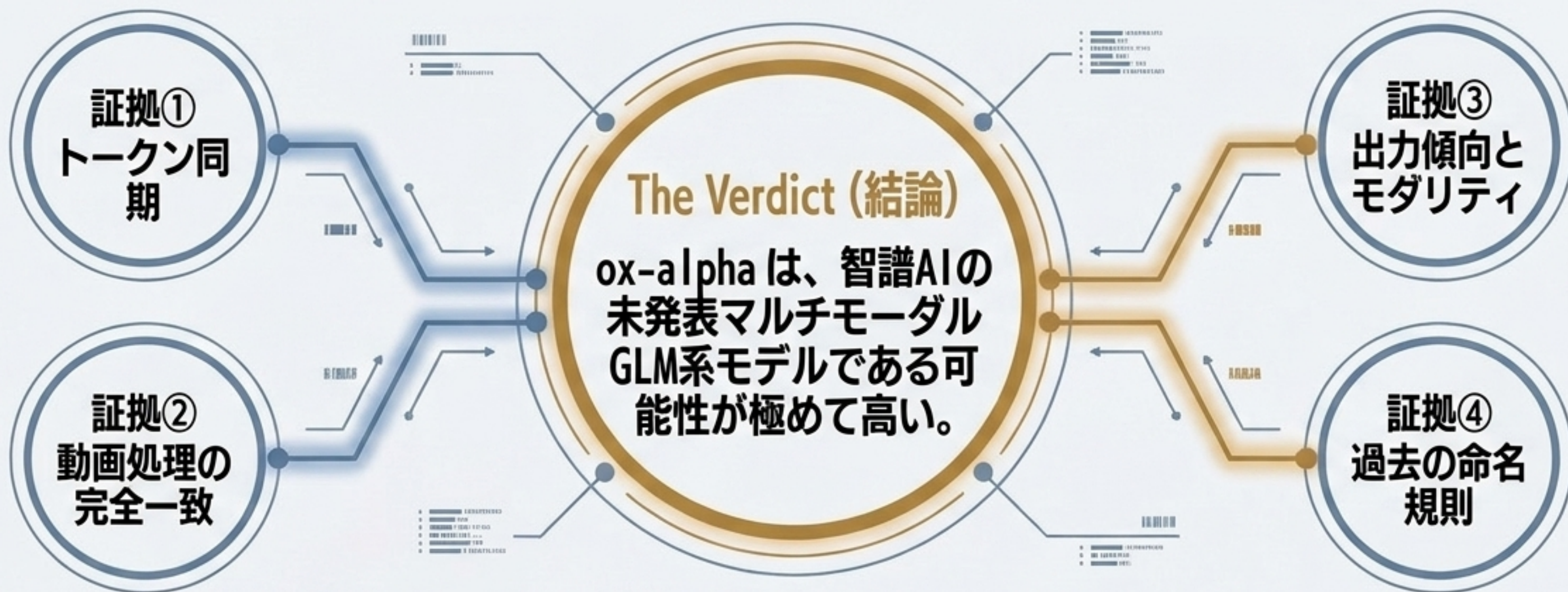
2026年8月14日

テキスト特化型モデル「GLM-5.3」を発表。ポストトレーニング強化によりサイバーセキュリティ等の能力向上をアピール。

2026年1月8日

香港証券取引所 (HKEX) へ上場。AI企業の資金調達サイクルを加速。

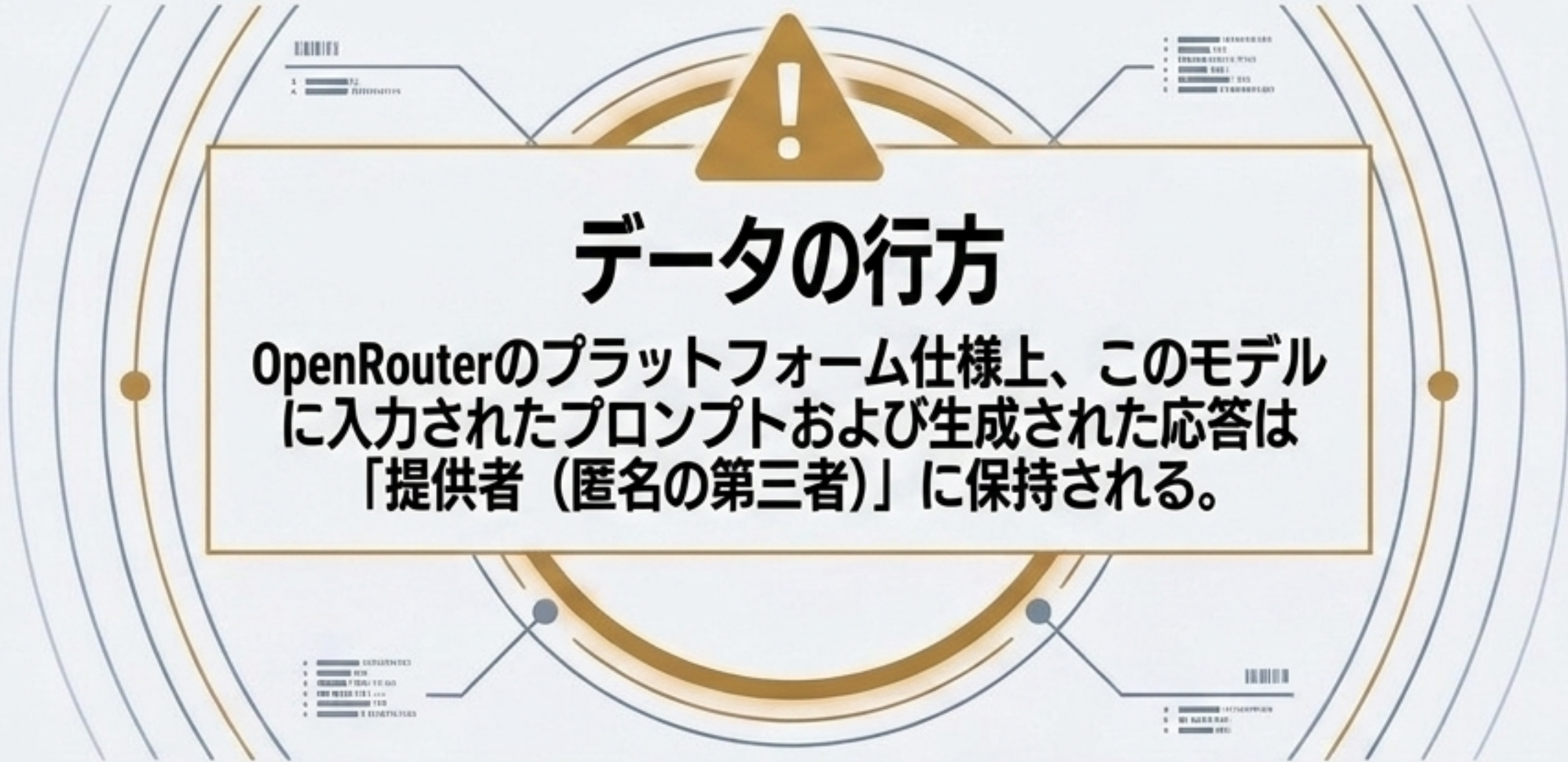
結論：無数の状況証拠が指し示す一つの「仮説」



根拠と留保事項

動画4条件におけるトークン消費パターンの完全な一致は極めて有力な状況証拠。ただし、これらは全てブラックボックス検証による推定であり、公式な帰属確認には至っていない。APIラッパーや推論基盤の共通性による類似の可能性も残る。

実務利用への警告：匿名プラットフォームにおけるデータ保持リスク



公式の不在

開発元や正式なモデル名が公式に確認されていないため、データ保護方針や安全性の担保が存在しない。

Actionable Advice

高いコーディング能力は魅力的だが、社内コードや機密情報を含む業務利用は現時点では厳格に避けるべきである。

調査レポートの総括 (Key Takeaways)



1. 匿名テストの高度化

基盤モデルの進化スピードは凄まじく、中国系AIラボによる「第三者プラットフォームを通じた実稼働テスト」が定常化しつつある。



2. 技術的指紋の有効性

トークナイザーや前処理の挙動を追跡する「ブラックボックス検証」は、AIモデルの系譜を特定する現代のデジタル・フォレンジックとして極めて有効である。



3. 検証への冷静なスタンス

単一試行の局所的なベンチマークに踊らされることなく、開発元からの公式発表と、独立機関による同一条件・大規模なベンチマークの完了を待つ必要がある。