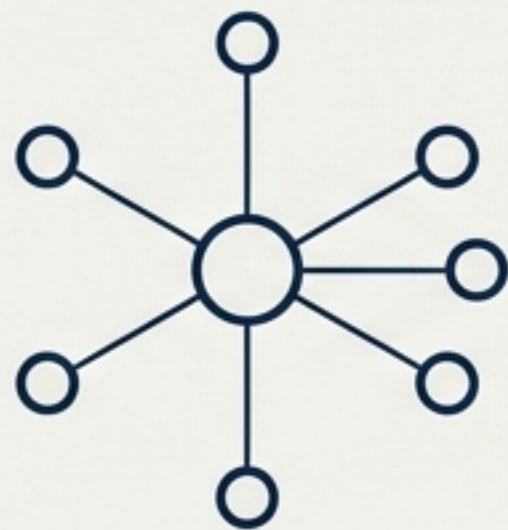


Sakana AI 「Fugu Max」 「Ultra v2」徹底解剖

パレートフロンティアの
オーケストレーション戦略と
企業導入のリアル





The Reality (技術の正体)

新規の巨大基盤モデルではなく
「学習されたオーケストレー
ター」。複数のオープンモデルへ
タスクを動的にルーティングす
る群知能。



The Catch (知財・実務リスク)

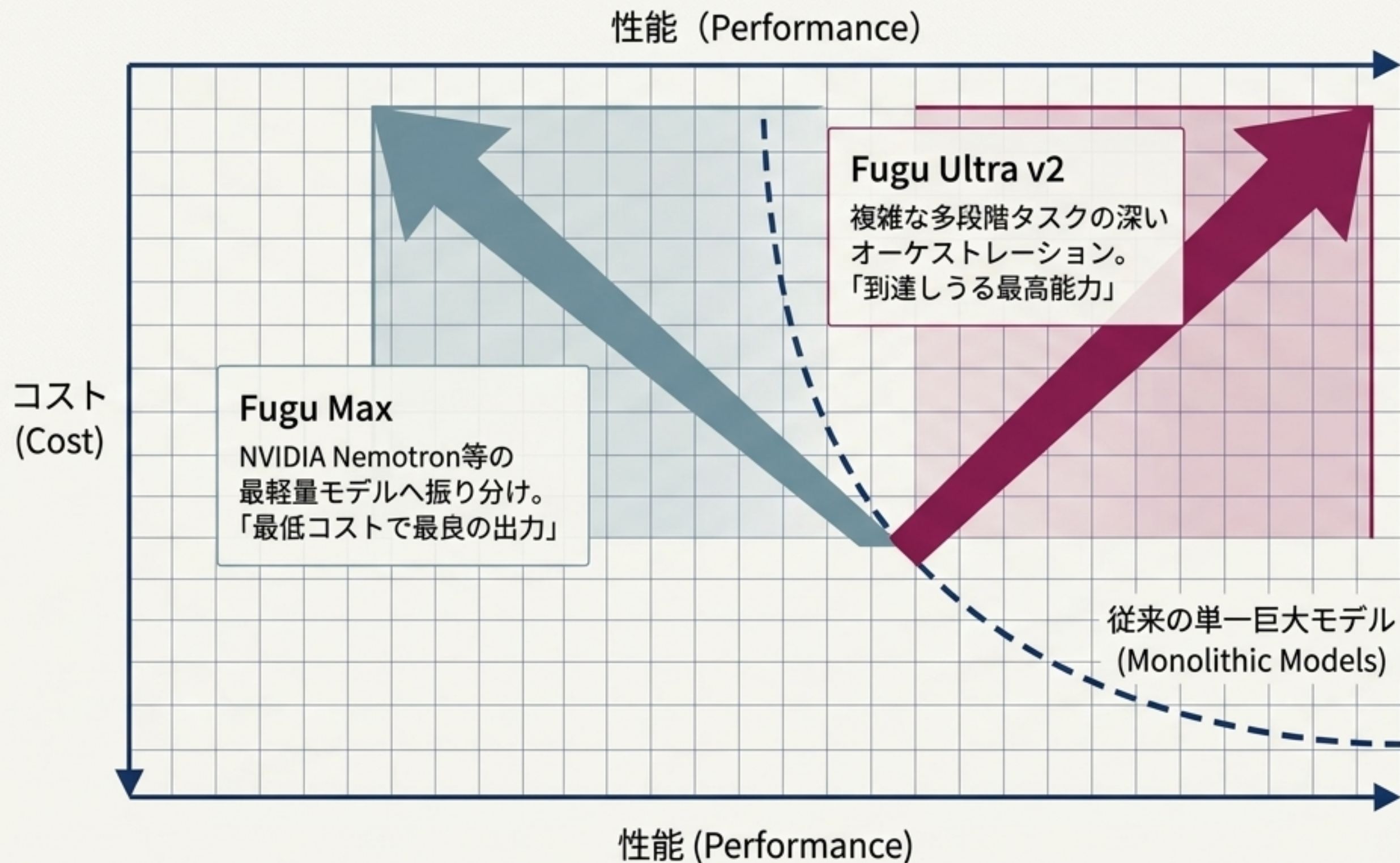
API経由で利用する裏側で、他
社モデル（Kimi等）の規約が
「連鎖」するリスク。実務上のレ
イテンシと不透明が最大の課題。



The Macro (地政学的意義)

米国の対日アクセス制限リスク
(2026年6月)を背景とした、輸
出規制に依存しない「ソブリンAI」
としての戦略的ポジショニング。

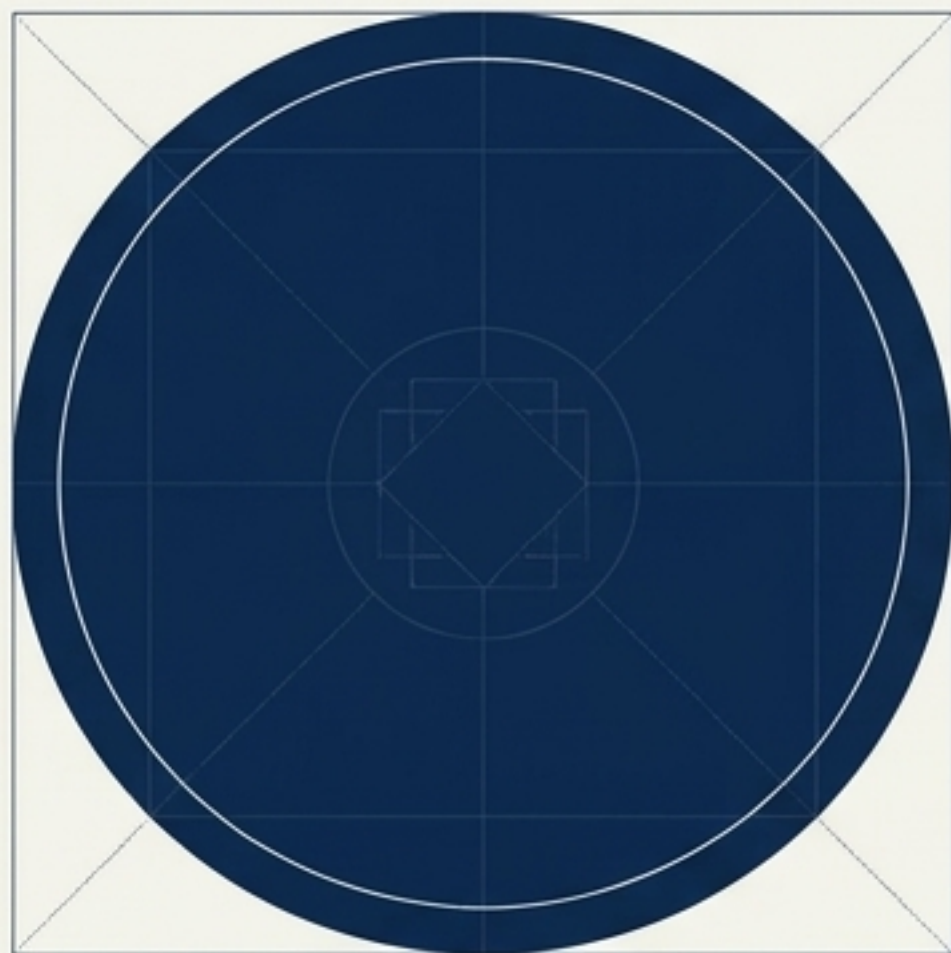
「オーケストレーション」が切り拓くパレートフロンティア



「最も能力の高いAIは
単一のモノリシックなモ
デルからではなく、知的
で集合的なオーケスト
レーションから生まれ
る」 — Sakana AI

巨大モデル(Monolith) vs. 群知能(Swarm)

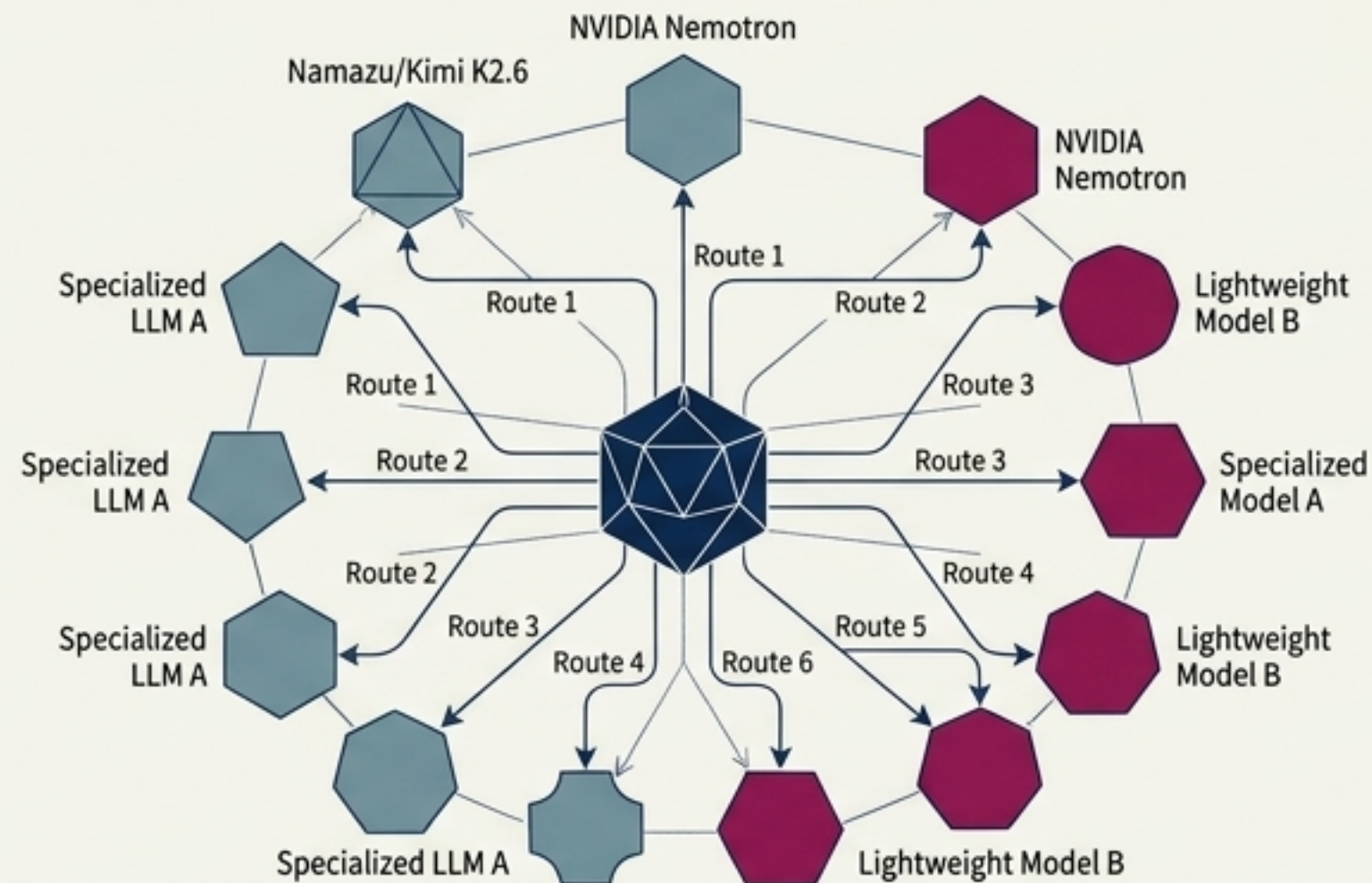
The Old Way



Monolithic Model
(e.g., GPT-5.6 Terra)

膨大な事前学習による全知全能アプローチ。

The Sakana Way



Fugu Orchestrator
(進化的アルゴリズム/強化学習)

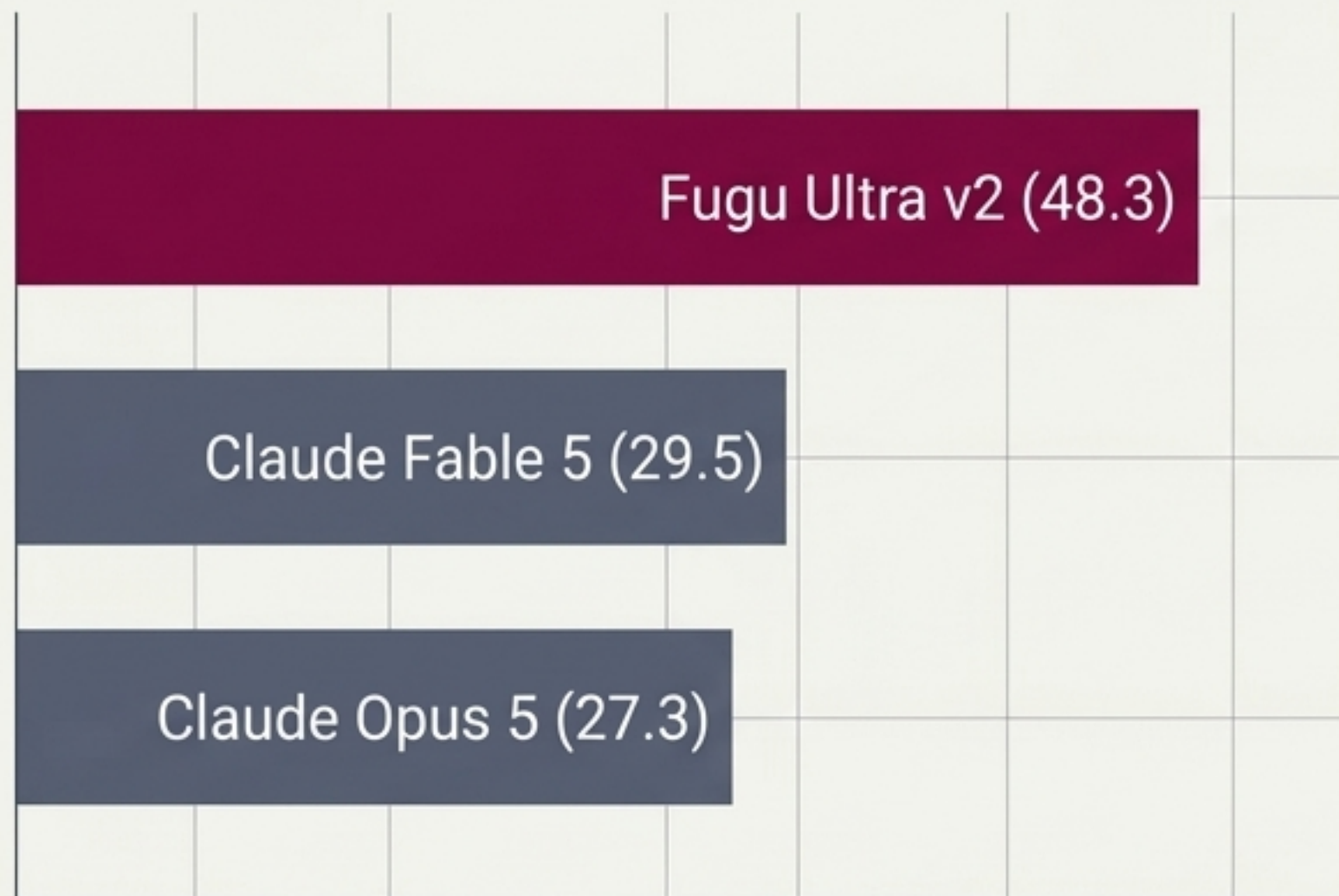
推論時計算 (Test-time Compute) による動的ルーティング。
解ける最も軽量なモデルへタスクをアサイン。

Fugu Max vs. Fugu Ultra v2 比較マトリクス

	Fugu Max	Fugu Ultra v2
目的 (Mission)	最低コストで最良の出力を追求	複雑なタスクで最高能力を追求
出力コスト (per 1M tokens)	\$6 (競合比 40～60%減)	\$30 (プール内最上位モデルと同水準)
入力 / キャッシュ	\$2 / \$0.25	\$5 / \$0.50
主なベンチマーク主張 (自己申告)	10指標中7つでフロンティア拡張。 Terminal Bench 2.1等で最良。	Chartography 48.3 (Opus 5 / Fable 5超えを主張)。
留意点 (Catch)	オープンモデルへのルーティングに よるコスト圧縮。	オーケストレーションのオーバーヘッド が直呼びに対し正当化されるかが争点。

提示された成果と「空白地帯」

Marketing Claims



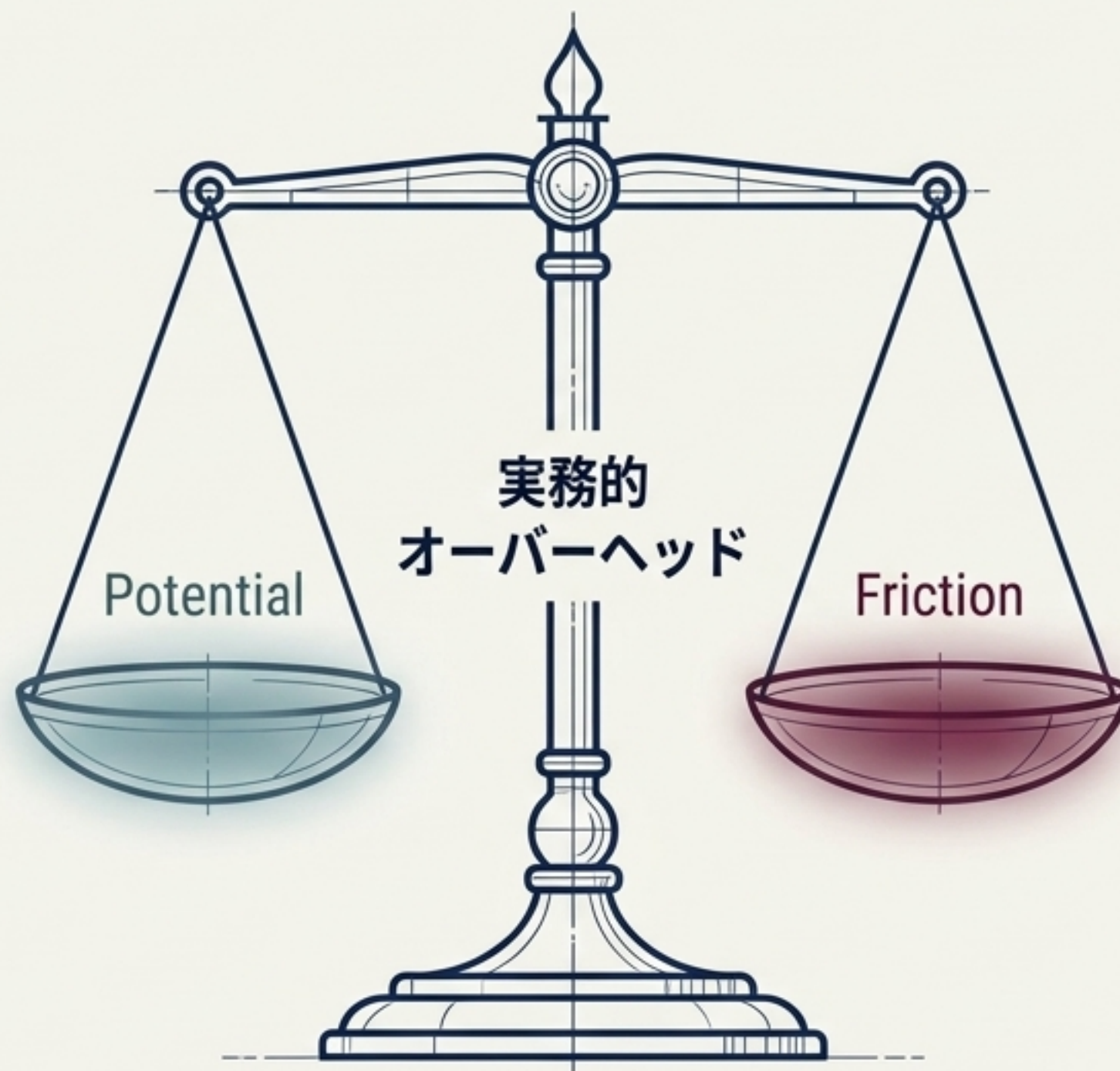
Chartographyスコア。DeepSWE 74.3など、8指標中5つでトップ。

Analytical Reality

	第三者検証ゼロ LMArena等に未掲載。BenchLM.aiは「非生的カバレッジ不足」でランク除外（2026年9月現在）。
	意図的な除外 Fugu Ultra v2の主張は「Claude Fable 5.1」および「GPT-6 Astra」をエージェントプールから除外した上でのスコア。
	クローズドAPI オープンウェイト公開はなく、OpenAI互換APIのみ（EU/EEA非提供）。

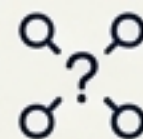
現場からの懐疑論と実務上のオーバーヘッド

- 「ベンチマークスコアは本物」
- コードレビュー等、特定タスクでの高い精度（20件超の指摘検出等）。



Latency

「非常に遅い」（Ethan Mollick等の指摘）。推論時計算による時間的コスト。



Architecture Doubt

「単なるラッパー、OpenRouterと何が違うのか？」というコミュニティの疑念。



Rate Limits

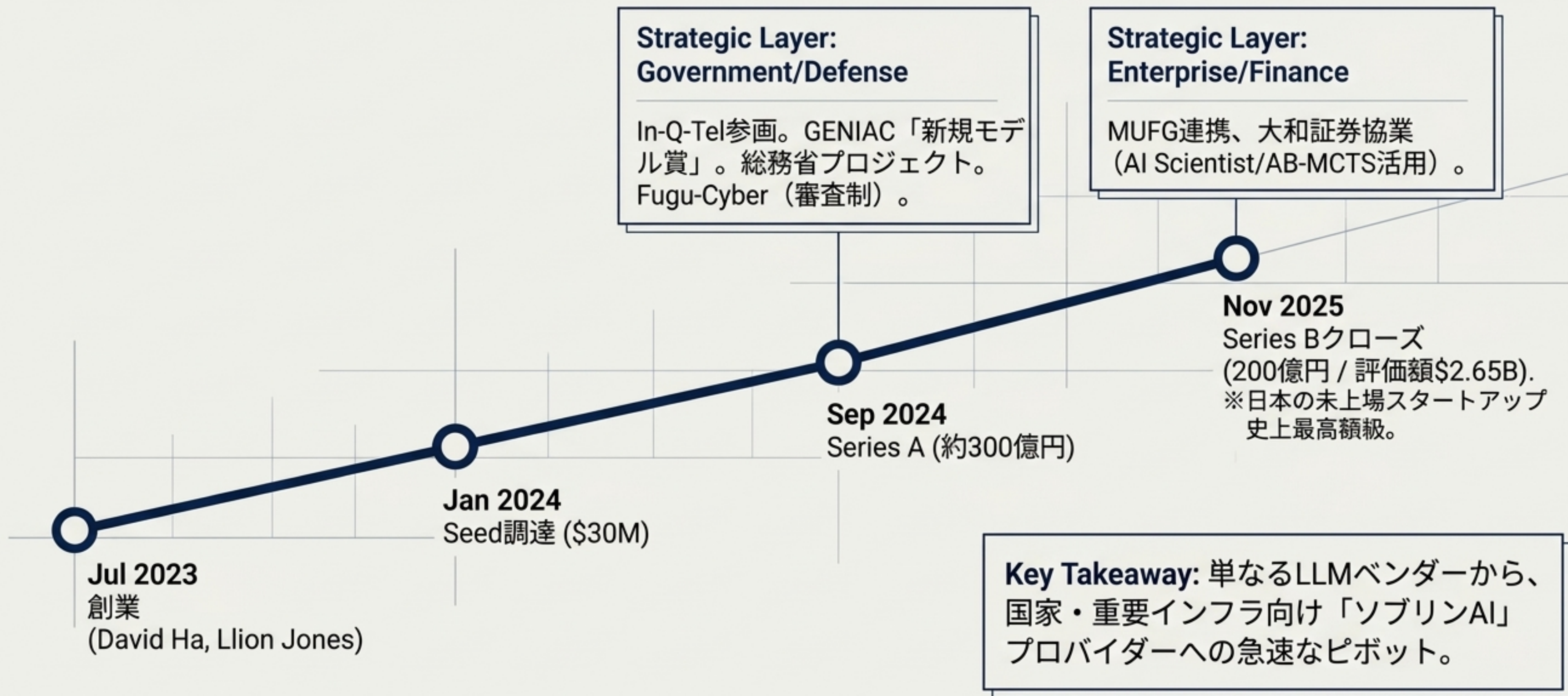
月額\$200プランで週3時間未満の稼働制限報告。



Cost Efficiency

Ultra v2の価格（\$30）が最上位モデル直呼びと変わらない点。

企業モメンタム：防衛・金融セクターへの浸透



異例の知財戦略：「特許ゼロ」のオープン&クローズ

オープンソース (Apache 2.0)

特許不行使条項を含むライセンスでの公開によるコミュニティの味方化。（著作権法30条の4が法的に下支え）



商標の囲い込み

「FUGU」「THE AI SCIENTIST」等の積極的な商標登録（登録IPは商標に集中）。

APIと契約によるクローズ化

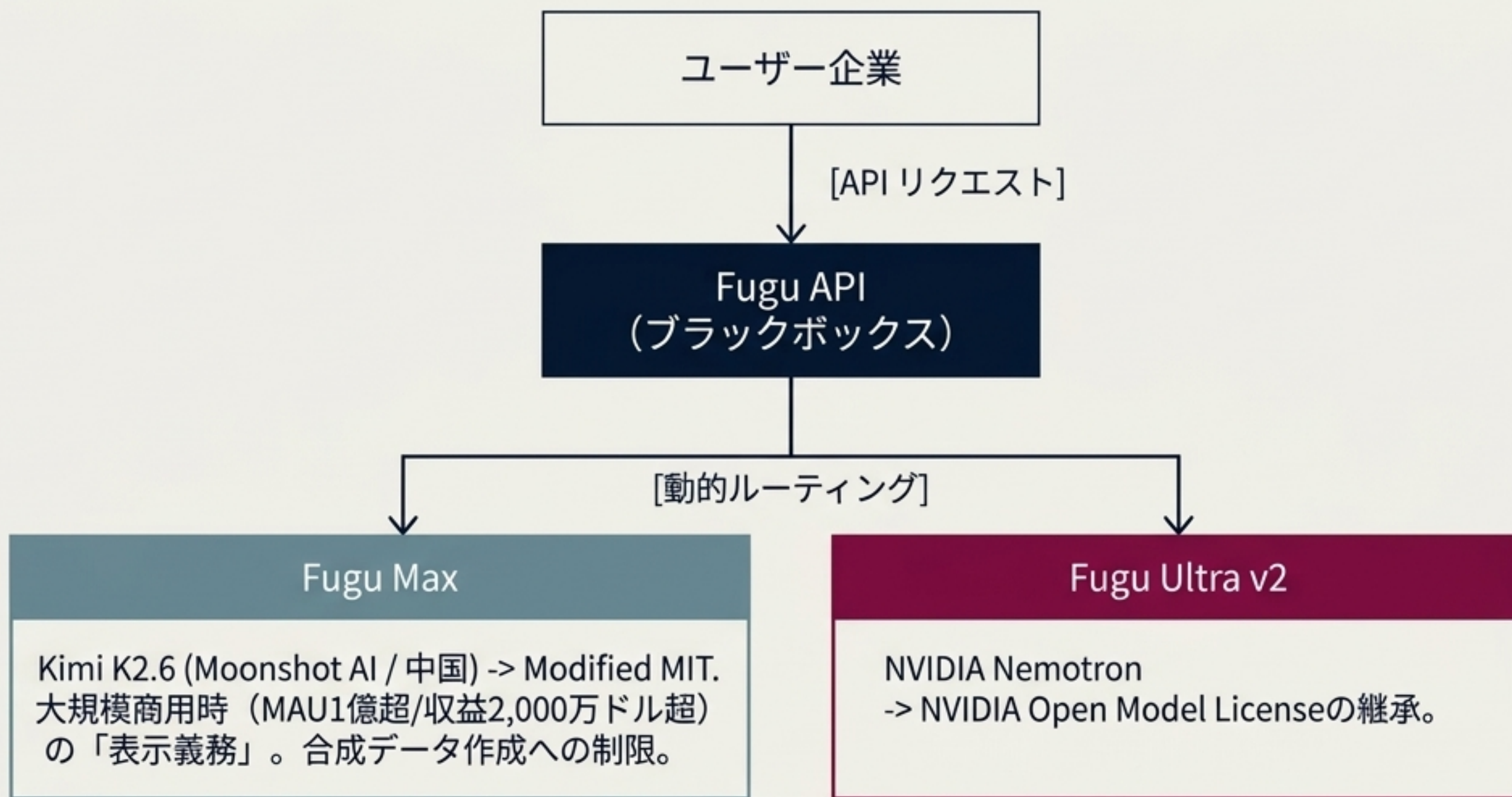
オープンウェイト（EvoSDXL-JP等）は商用不可とし、本命（Fugu）はAPI専売でブラックボックス化。

なぜ群知能か？ 米国輸出規制とソブリンAIの空白



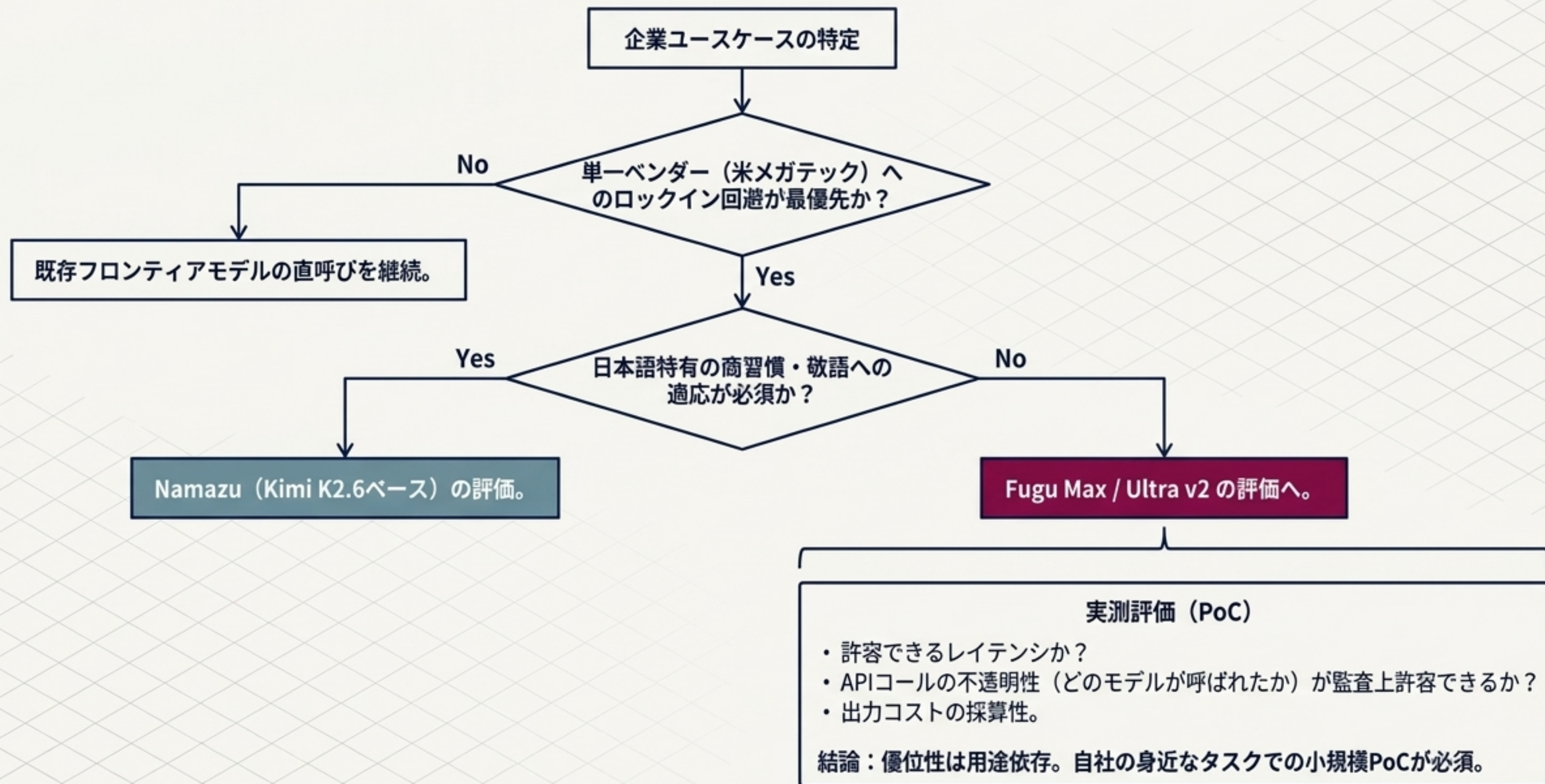
Sakana's Positioning: Fuguは「輸出規制リスクのないフロンティア能力」として投入。計算資源制約下の現実解として、オープンモデル（NVIDIA, Kimi等）の果実を統制された形で取り込むソブリン戦略。

ライセンス連鎖の罠（オーケストレーション特有のリスク）



API経由では単一モデルに見えるが、実態は「背後で呼ばれる全モデルの利用規約・輸出規制・地政学リスク」が事実上流れ込む構造。

導入判断のためのPoCステップ（意思決定フロー）



企業が「今すぐ」取るべき4つのアクション



1. 【即時】社内共有（IT/経営）

ベンチマークは「自己申告・第三者未検証」であることを明記し、熱狂を鎮める。独立検証を待つか PoC を必須とする。



2. 【短期】知財デューデリジェンス（法務）

背後で呼ばれるモデル（Kimi, Nemotron等）のライセンス洗い出し。表示義務、合成データ制約、EU非提供を契約審査に組み込む。



3. 【短期】IPランドスケープ監視（知財）

「特許0」戦略からの転換（クローズ化・権利行使の予兆）を検知するため、四半期ごとに J-PlatPat/USPTOで「Sakana AI」を監視。



4. 【中期】調達・ロックイン評価（調達/CIO）

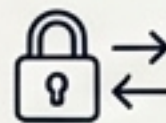
Fugu導入自体も「Sakana AIという単一ベンダーへの依存」とであると認識し、レイテンシ・コストと天秤にかける。

今後のウォッチポイント（評価更新のトリガー条件）



Trigger A: 第三者検証の実証

LMarenaやNejumi Leaderboard等の独立系でFable 5.1 / GPT-6 Astra級が証明された場合（→ 採用優先度を上げる）。



Trigger C: プール内モデルの規約変更

背後のモデルに新たな輸出規制やライセンス変更が発生した場合（→ 連鎖リスクを再評価）。



Trigger B: 特許出願の開始

Sakana AIが特許出願を開始した場合（→ IP戦略転換・権利行使リスクとして警戒度を上げる）。



Trigger D: EU/EEAでの提供開始

グローバル展開におけるコンプライアンス上の実用性が向上したサイン（→ 全社展開のゴーサイン）。