

2026年10月7日時点・一次ソースによる事実確認

Mistral Large 4 (“Le Chonk”) 徹底調査レポート

日本企業のための戦略的評価と導入ガイド

機密性レベル: 経営層・知財・IT戦略部門向け



「第三の道」の確立

中国製を除外したオープン系モデルにおいて世界最高性能（AA Index: 38）。米国クラウドへの依存と地政学リスクを同時回避する準フロンティアモデル。



特定領域における 圧倒的習熟

サイバー防御（世界トップ5）と法務エージェント（Harvey Legal Agent オープン系1位: 15.8%）に特化。汎用性よりも特定業務の深さを追求。



「オープン」の留保 と今後の課題

現時点ではプレビュー版APIのみ。重み公開とカスタムライセンスの全容解明（10月27日予定）、および日本語性能の実測がオンプレミス導入の前提条件。

Mistral Large 4は欧州インフラで構築された 巨大なハイブリッドMoEモデルである

規模 (Scale)

総パラメータ数 **1.05兆** (1.05T) /
アクティブ **490億** (49B)。
DeepSeek V4 Proと同規模の
アクティブパラメータ。



視覚統合 (Vision)

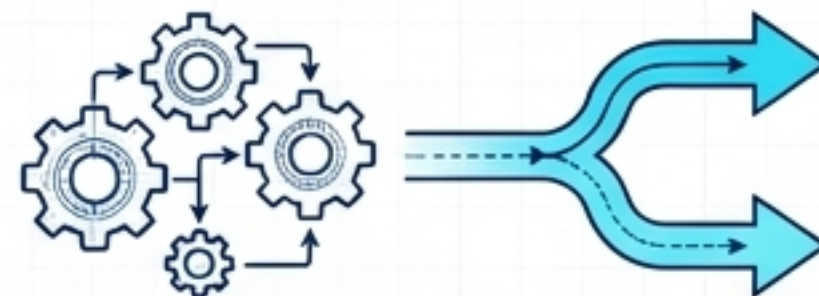
16億 (1.6B) パラメータのビジョンエンコーダを内蔵 (テキスト+画像入力に対応)。



Mistral
Large 4

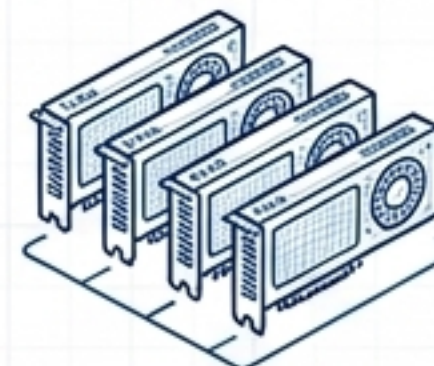
アーキテクチャ (Architecture)

Granular MoE (詳細な混合エキスパート)。
「通常応答 (Instruct)」と「推論 (Reasoning)」を
1モデルに統合したハイブリッド型。



計算資源 (Compute)

欧州データセンター内、NVIDIA Grace
Blackwell GPU 3,800基を使用しゼロ
から事前学習 (約2ヶ月)。



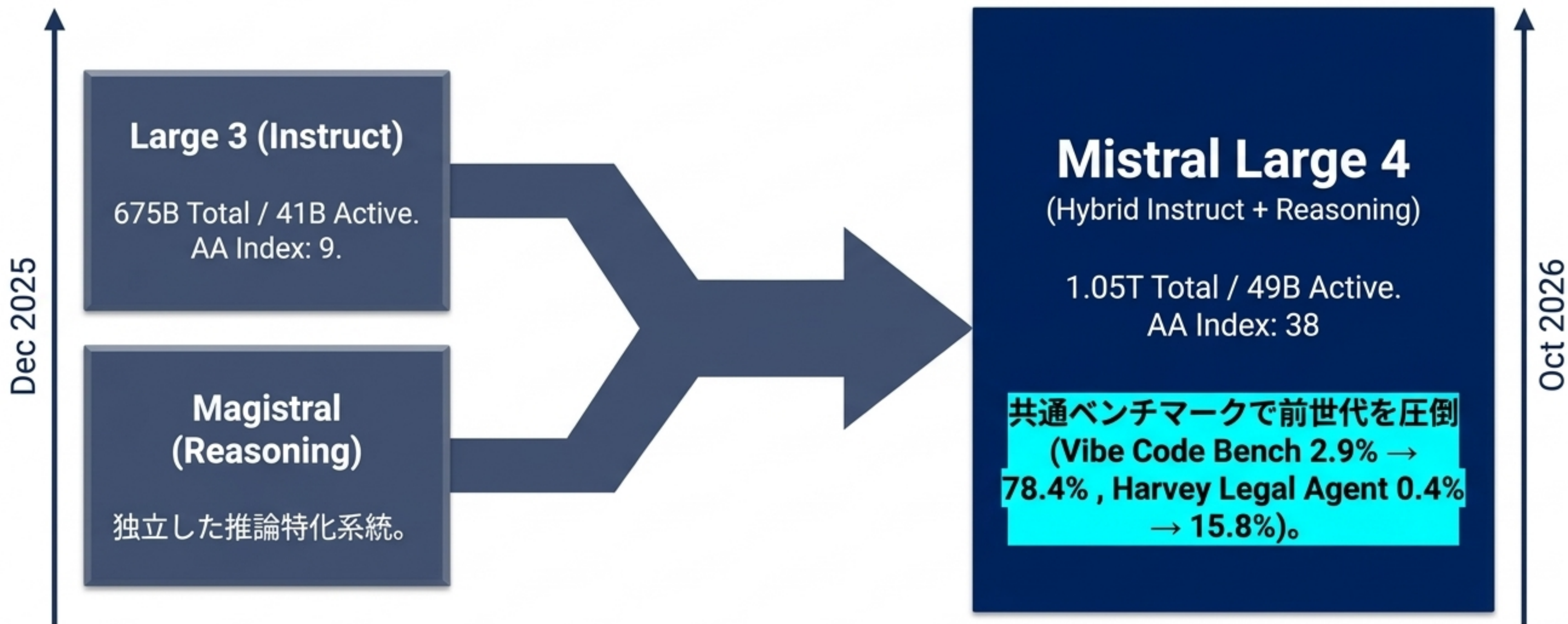
約2ヶ月

ゼロから事前学習

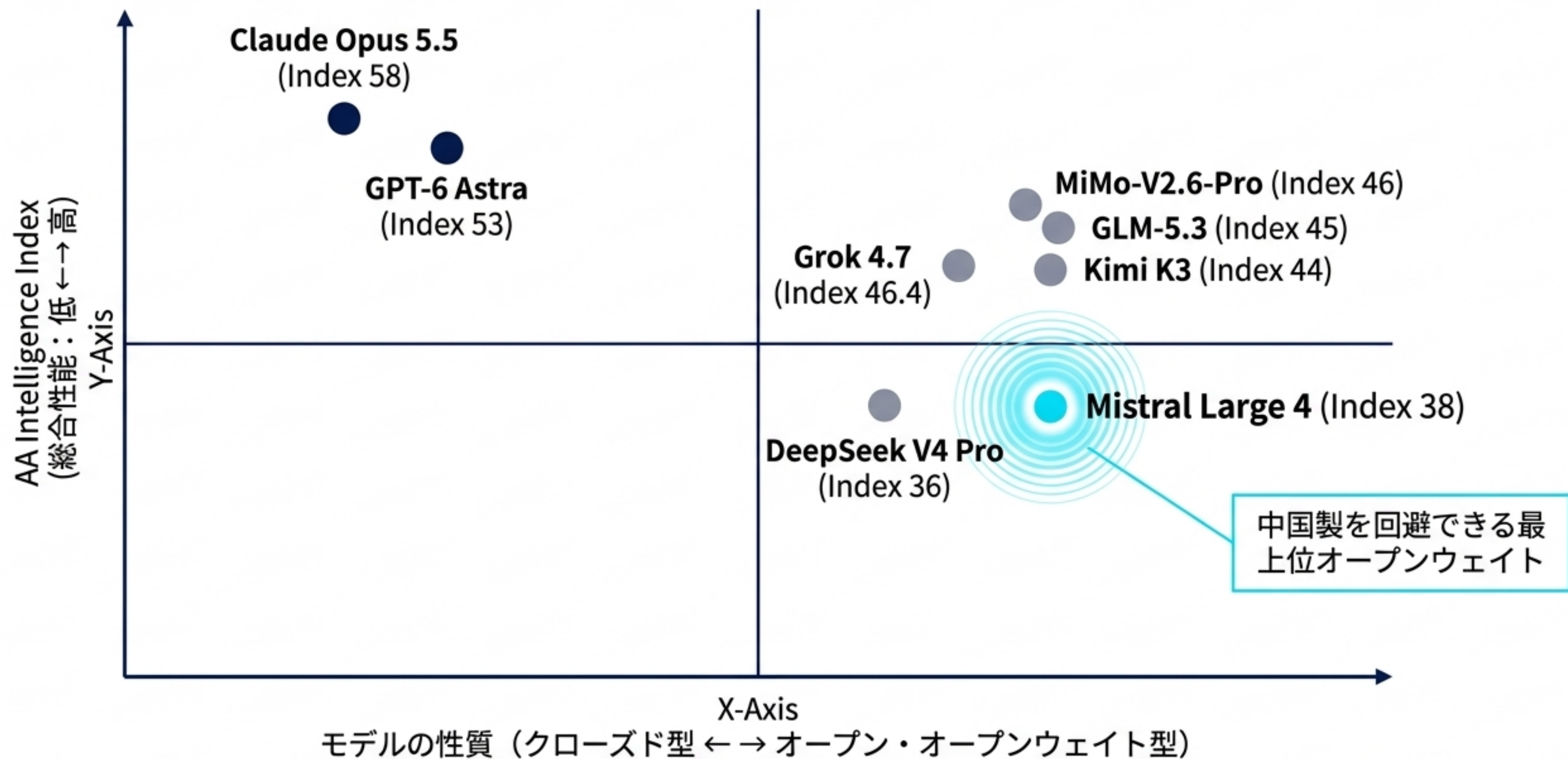
公式マーケティングの主張と第三者検証に基づく現実の乖離

	Mistral Official (公式発表)	Verified Reality (第三者検証)
コンテキスト長	100万トークン対応 (1M)	約52万トークン (524K) 。Artificial Analysis (AA) 等の第三者計測で実提供値が下回る未解決の食い違い。
価格とコスト	割引価格: 入力\$0.68 / 出力\$2.09 (1Mトークンあたり)	リスト価格は2倍 (\$1.36/\$4.18) 。さらに出力が冗長なため、AA指数1タスクあたりの実質コストは\$1.13と高止まりする。
ライセンス	「オープンウェイト」モデル	現時点では未公開 。10月27日に「カスタムライセンス」での公開見込み。真のオープンソース要件を満たすは法的に未確定。

前世代(Large 3)からの飛躍的進化とモデルシステムの統合



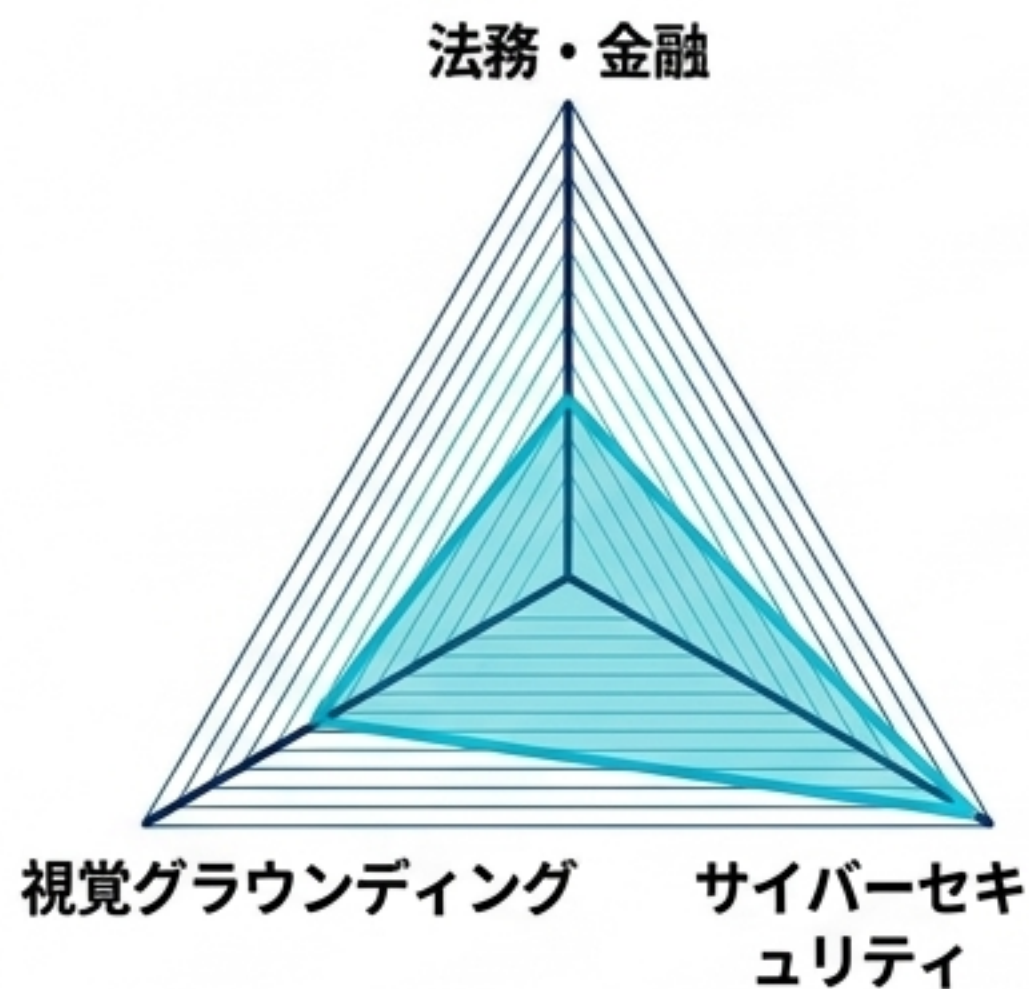
AI覇権のグローバルマトリクス：「中国以外のオープン系」における頂点



Key Insight: 中国勢上位7モデルには劣るものの、日欧米のオープン系としては首位クラスを確保

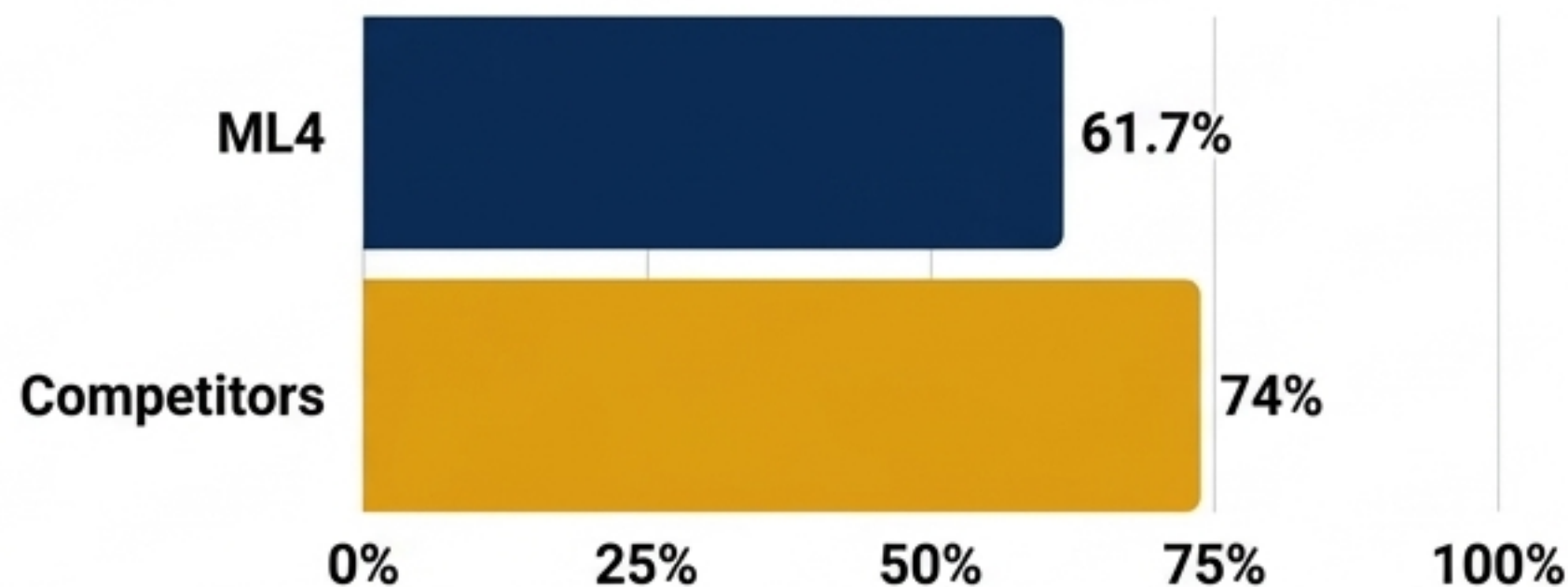
性能プロファイリング： 局地戦における優位性と汎用性の課題

特化領域の強み



- **法務・金融:** Harvey Legal Agent 15.8% (オープン系1位、全体6位)。
- **サイバーセキュリティ:** CyberGym-E2E 82% (比較18モデル中最高)、AA Cyber Index 50。
- **視覚グラウンディング:** Dense 200 42% (GPT-6 Astraの41%を凌駕)、DIOR-RSVG 73%。

弱点と懸念事項



- **汎用コーディング:** DeepSWE等で競合(約74%)に及ばず(61.7%)。
- **出力の冗長性:** AA評価で「very verbose」。トークン消費量が前世代比で約6倍に膨張し、レイテンシ（遅延）の要因に。

「欧州AI主権」という戦略的防壁： 独自のインフラと強力な資本

完全な欧州インフラ

学習から推論まで米系メガテックから独立した自社データセンターで完結。欧州法に準拠したデータ主権を確保。

ディープテック資本の結集

企業評価額 €210億（約\$240億）。ASML（露光装置）、Samsung（メモリ）、Nvidia（GPU）という半導体サプライチェーンの中核企業が株主として結合。

規制の波乗り

推定計算量 10^{25} FLOPs超の「システムックリスク」モデルに該当する可能性大。国家当局による実環境のレッドチーミング等、規制遵守を前提とした設計。

オンプレミス展開におけるハードウェア要件とライセンスの壁



Step 1: The License Gate

真のオープンソース (Apache等) か、商用制限付きの修正ライセンスか。10月27日公開の「カスタムライセンス」条文の精査が必須。



Step 2: The Hardware Reality

総1.05兆パラメータの巨大モデル。FP8/FP4形式のチェックポイント運用でも、NVIDIA B200またはB300 GPUが「4~8基」必要（報道ベース）。

中堅以下の企業にとっては、自社インフラよりも国内クラウド事業者経由での利用が現実的な選択肢。

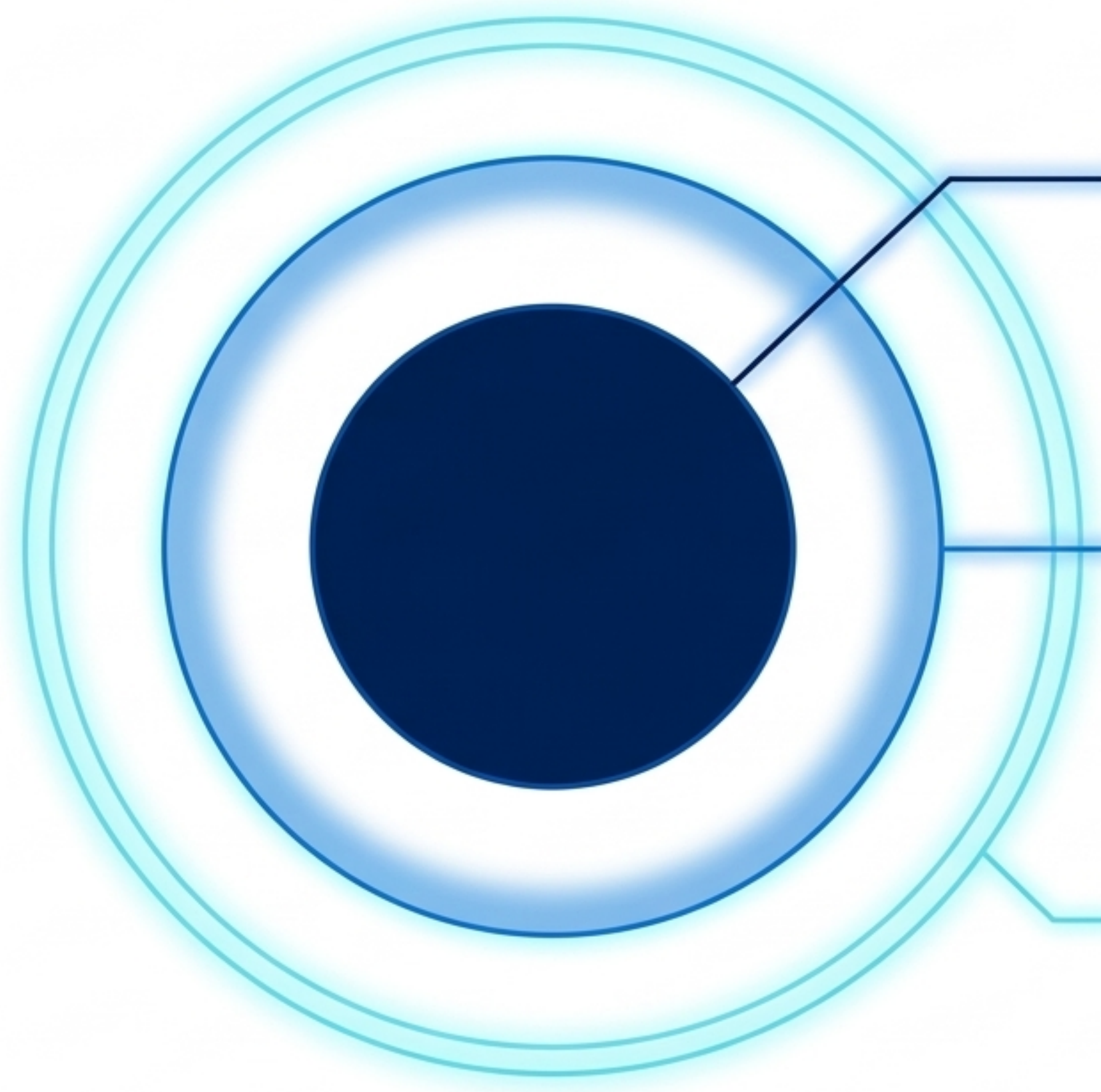


Step 3: Data Privacy Trade-off

総1.05兆パラメータのモデル。オリエックスのアンケーフィを断った求に封要する。

要件をクリアできれば、防衛関連や金融など、究極のデータ主権（外部遮断）を実現する最強の自社専用モデルとなる。

日本企業がMistral Large 4を導入すべき3つの戦略的理由



データ主権の完全確保

機密データ（未公開知財、M&A情報、顧客データ）を自社環境（オンプレミス/プライベートクラウド）から一切外に出さずに高度なAI処理が可能。

地政学リスクの回避

高性能な「中国製モデル（MiMo, GLM等）」を調達上排除せざるを得ない日本企業にとって、事実上唯一の「準フロンティア」の選択肢。

ベンダーロックインの打破

米国のクローズドAPI（OpenAI, Anthropic）への過度な技術的・コスト的依存を軽減し、交渉力と選択の自由を確保する「第三極」。

特定業務への適用可能性診断

①：知的財産（IP）ワークフロー

	ユースケース	ポテンシャル (Capability)	制約・リスク (Constraint)	推奨アプローチ (Recommendation)
	特許調査・ 先行技術文献分析	1M公称コンテキストに よる一括読み込みのポテ ンシャル。	実効コンテキストが約 52万(524K)に留まる可 能性。長文脈推論での 優位性は未確定。	長大な文献にはチャンク 分割とRAG（検索拡張生 成）の併用が必須。
	明細書作成支援・ 法務的推論	優れた法務エージェント 性能。	絶対的な推論力では米国 クローズド上位モデルに 及ばず。	「AIによる下書き＋弁理 士の人間レビュー」を前 提とした運用に限定。
	出願前発明の機密 保持	オンプレミスでの完全オ フライン運用。	初期導入のハードウェア コスト。	最も機密性の高い出願前 情報処理において、極めて 高い投資対効果を発揮。

特定業務への適用可能性診断

②：法務エージェントとサイバー防衛

15.8%

Harvey Legal Agent (HLAB)
オープン系第1位

Insight:

契約書レビューやコンプライアンスチェックにおける一次スクリーニングツールとして、クローズドモデルに次ぐ実用性を具備。

82%

CyberGym-E2Eテスト
比較18モデル中最高水準

Insight:

強いサイバー防御能力を持つ一方、悪用プロンプトに対する高い拒否率（OSS最高）を誇る。安全性をソフトウェアで封じ込めるアーキテクチャを採用。

導入に向けた3ステップのアクションプラン

Step 1: APIレビュー評価

(即時実行)

Mistral Studio経由で非機密データを用いたレビュー評価を開始。自社固有の法務・サイバータスクにおけるコスト（出力の冗長性）を検証。

Step 2: 3つの検証ゲート

(10月27日以降)

重み公開時に以下を厳格に確認：①カスタムライセンスの商用利用条件、②実効コンテキスト長の最終仕様、③日本語処理性能の実測値。

Step 3: インフラ調達と環境構築

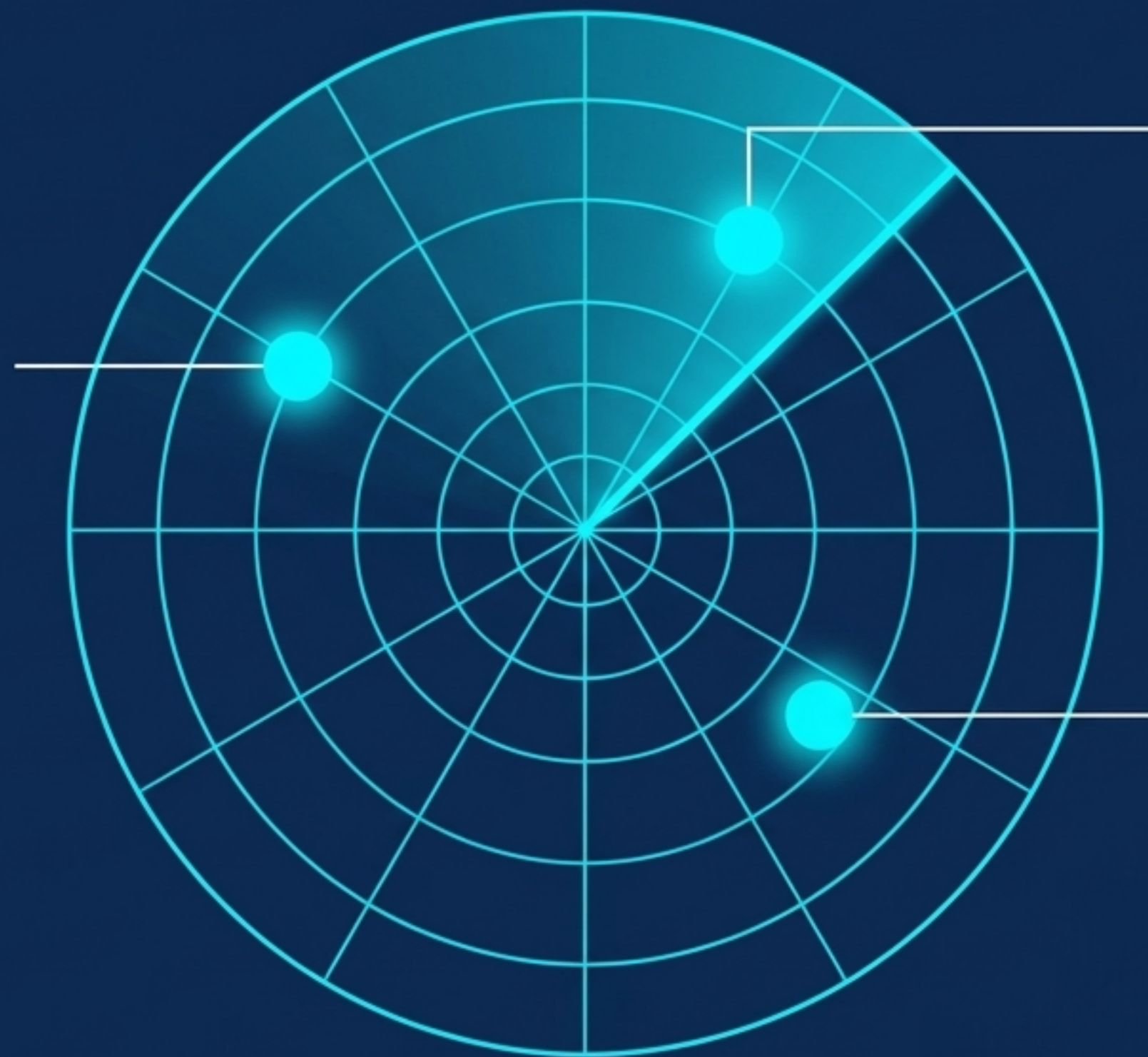
(2027年に向けて)

条件をクリアした場合、B200/B300クラスのGPU調達、または国内のセキュアなホスティング事業者を通じた専有環境の構築を計画。

ウォッチリスト：今後の影響を左右する不確実性

最終チェックポイントでの性能変動

現在は強化学習（RL）が進行中のプレビュー版。最終的な重み公開時に、AA指数やコンテキスト実提供値が向上（または固定化）するかの見極め。



オープンソース界の反応

「オープン」を標榜しながらカスタムライセンスを採用する姿勢に対する、Hugging Faceや開発者コミュニティの最終的な受容度。

エコシステム統合戦略

モデル単体の性能競争から、Mistral Forge、Compute、Vibe等の自社プラットフォームを組み合わせた「全部入り」垂直統合戦略への移行成否。