



# Baidu（百度）の2025年11月AI戦略：ERNIE 5.0 と ERNIE-4.5-VL-28B-A3B-Thinking の深層分析

2025年11月、中国のテクノロジー大手 Baidu（百度）は、グローバル AI 競争における戦略的な転換点となる重要な発表を行った。11月13日に開催された年次イベント「Baidu World 2025」において次世代基盤モデル「ERNIE 5.0」を公開する一方、その数日前に Apache 2.0 ライセンスでオープンソースのマルチモーダルモデル「ERNIE-4.5-VL-28B-A3B-Thinking」をリリースした。この同時進行する二つのアプローチは、Baidu が市場を異なる戦略で攻略しようとしていることを明確に示している。プレミアムな商用モデルで大企業や高度なユースケースに対応しつつ、オープンソースモデルで開発者コミュニティとコスト効率を重視する企業層を取り込む「二刀流戦略」である。[\[1\]](#)[\[2\]](#)[\[3\]](#)[\[4\]](#)

## Baidu ERNIE: 二刀流戦略

特徴	ERNIE 5.0 (プロプライエタリ)	ERNIE-4.5-VL-28B (オープンソース)
ライセンス	プロプライエタリ (商用)	Apache 2.0 (オープンソース)
総パラメータ	2.4兆以上	280億
アクティブ数	<72B (3%未満)	30億
価格 (入力/M)	\$0.85	オープンソース (無料)
提供方法	ERNIE Bot, Qianfan API	HuggingFace, GitHub, AI Studio
対象ユーザー	大企業、特定ユースケース	開発者、中規模組織、研究者

Baidu の ERNIE モデル二刀流戦略：プレミアム商用（5.0）とオープンソース（4.5）の比較

## ERNIE 5.0 : ネイティブオムニモーダル AI の誕生

### 技術的革新と規模

ERNIE 5.0 は、Baidu が「ネイティブオムニモーダル」と位置づける次世代基盤モデルであり、テキスト、画像、音声、動画を設計段階から統合的にモデリングする点で従来のマルチモーダル AI とは一線を画している。単に複数のモダリティを後付けで融合するのではなく、訓練の初期段階から全てのモダリティを同時に処理する「ネイティブ統合オムニモーダルモデリング技術」を採用している。

[\[1\]\[5\]\[6\]\[3\]\[7\]](#)

モデル規模は 2.4 兆パラメータを超える巨大なものだが、超スパース MoE (Mixture-of-Experts) アーキテクチャにより、推論時にアクティブになるのは 3%未満 (72B 以下) のパラメータのみである。この設計により、膨大な知識容量を維持しながら推論コストと遅延を劇的に削減することに成功している。Baidu 創業者兼 CEO の Robin Li (李彦宏) は、「基盤モデルは急速に進化しており、知性の限界突破、推論時間の増加、複数モダリティのネイティブ統合、自己学習・進化能力が特徴だ」と述べ、技術イテレーションの速度が唯一の競争上の堀であると強調した。[\[6\]\[8\]\[9\]\[10\]\[11\]\[12\]\[1\]](#)

### ベンチマーク性能と競合比較

ERNIE 5.0 は、40 以上の権威あるベンチマークテストにおいて、OpenAI の GPT-5-High や Google の Gemini 2.5 Pro と同等またはそれを上回る性能を示している。特に視覚タスクにおいて顕著な優位性を発揮し、OCRBench、DocVQA、ChartQA などの文書認識・理解ベンチマークでトップスコアを記録した。これらは企業の業務自動化ニーズに直結する実用的な領域であり、エンジニアリング図面、工場現場の映像、医療スキャン、物流ダッシュボードなど、多くの企業が価値ある洞察を得られていなかった非テキストデータの処理に強みを持つ。[\[2\]\[6\]\[13\]\[3\]\[14\]\[7\]](#)

中国の LMArena Text リーダーボードでは、ERNIE 5.0 のプレビュー版 (ERNIE-5.0-Preview-1022) が第 2 位にランクインし、国内市場での高い評価を獲得している。画像・動画生成能力においても、専門分野で世界最高レベルに達しており、単なる理解型 AI を超えた総合的なクリエイティブツールとしての地位を確立している。[\[15\]\[11\]\[7\]\[16\]\[17\]](#)

### 提供形態と価格戦略

ERNIE 5.0 は、一般ユーザー向けには ERNIE Bot (文心一言) を通じて、企業ユーザー向けには Baidu AI Cloud の MaaS プラットフォーム「Qianfan (千帆)」を通じて API 提供される。価格設定は、入力

が 100 万トークンあたり 0.85 ドル、出力が 3.40 ドルとされており、GPT-5.1（入力 1.25 ドル/出力 10 ドル）や Gemini 2.5 Pro（入力 1.25 ドル/出力 10 ドル）と比較して競争力がある。[\[1\]\[5\]\[6\]\[18\]\[7\]](#)

ただし、2025 年 3 月に発表された ERNIE 4.5 の価格（入力 0.55 ドル/出力 2.2 ドル）と比較するとやや高額であり、ERNIE 5.0 は最先端性能を求めるプレミアム層を狙った位置づけであることがわかる。実際、Baidu は 2025 年 3 月の時点で ERNIE Bot を個人ユーザーに無料開放する計画を前倒して実施しており、基本的な利用は無料で可能である。[\[19\]\[20\]\[21\]\[22\]](#)

## **ERNIE-4.5-VL-28B-A3B-Thinking : オープンソースの戦略的意義**

### **軽量高性能アーキテクチャ**

ERNIE-4.5-VL-28B-A3B-Thinking は、2025 年 11 月 10 日頃に Apache 2.0 ライセンスでオープンソース化された軽量マルチモーダルモデルである。総パラメータ数は 280 億（28B）だが、MoE アーキテクチャにより推論時にアクティブになるのはわずか 30 億（3B）のパラメータのみであり、同等のフルパラメータモデルと比較して推論速度が 2～3 倍高速である。[\[4\]\[23\]\[24\]\[25\]\[26\]](#)

この効率性により、単一の 80GB GPU での動作が可能となり、中規模の組織や個人開発者でも導入しやすい設計となっている。ただし、モデル全体のウェイトをロードする必要があるため、最低 80GB の GPU メモリが必要である点には注意が必要だ。量子化技術（wint8）を使用することで、60GB 程度のメモリでも動作可能とされている。[\[2\]\[23\]\[27\]\[28\]\[4\]](#)

### **強化学習による推論能力の飛躍**

ERNIE-4.5-VL-28B-A3B-Thinking の最大の特徴は、大規模強化学習によって獲得した卓越した推論能力である。モデルは広範な中間訓練フェーズを経て、膨大で多様なプレミアム視覚言語推論データを吸収し、視覚とテキストのモダリティ間の意味的整合性を深化させた。GSPO（Generalized Sparse Policy Optimization）と IcePop 戦略を統合した最先端のマルチモーダル強化学習技術により、MoE 訓練の安定化と動的難易度サンプリングによる学習効率の向上を実現している。[\[29\]\[26\]](#)

実際のベンチマーク性能では、MathVista で 82.5 点（GPT-5-High の 81.3 点、Gemini 2.5 Pro の 82.3 点を上回る）、ChartQA で 87.1 点（GPT-5-High の 78.2 点、Gemini 2.5 Pro の 76.3 点を大きく上回る）、VLMs Are Blind で 77.3 点（GPT-5-High の 69.6 点を大幅に上回る）を記録している。これらの結果は、軽量モデルでありながらフラッグシップモデルに匹敵またはそれを超える性能を達成していることを示している。[\[2\]\[3\]\[4\]\[29\]](#)

## 革新的機能：画像で思考する AI

ERNIE-4.5-VL-28B-A3B-Thinking は、「Thinking with Images（画像で思考）」という革新的な機能を搭載している。これは、人間のように画像を自由にズームイン・ズームアウトして細部を把握し、すべての情報を明らかにする能力である。画像検索などのツールと組み合わせることで、細かいディテールの処理とロングテール視覚知識の処理能力が劇的に向上する。[\[29\]\[26\]](#)

さらに、より精密なビジュアルグラウンディングと柔軟な指示実行が可能であり、複雑な産業シナリオにおいてグラウンディング機能を容易にトリガーして効率を大幅に向上させることができる。動画理解においても、優れた時間認識とイベント定位能力を持ち、動画内の異なる時間セグメントにおけるコンテンツ変化を正確に識別する。[\[26\]\[29\]](#)

## オープンソース化の戦略的インパクト

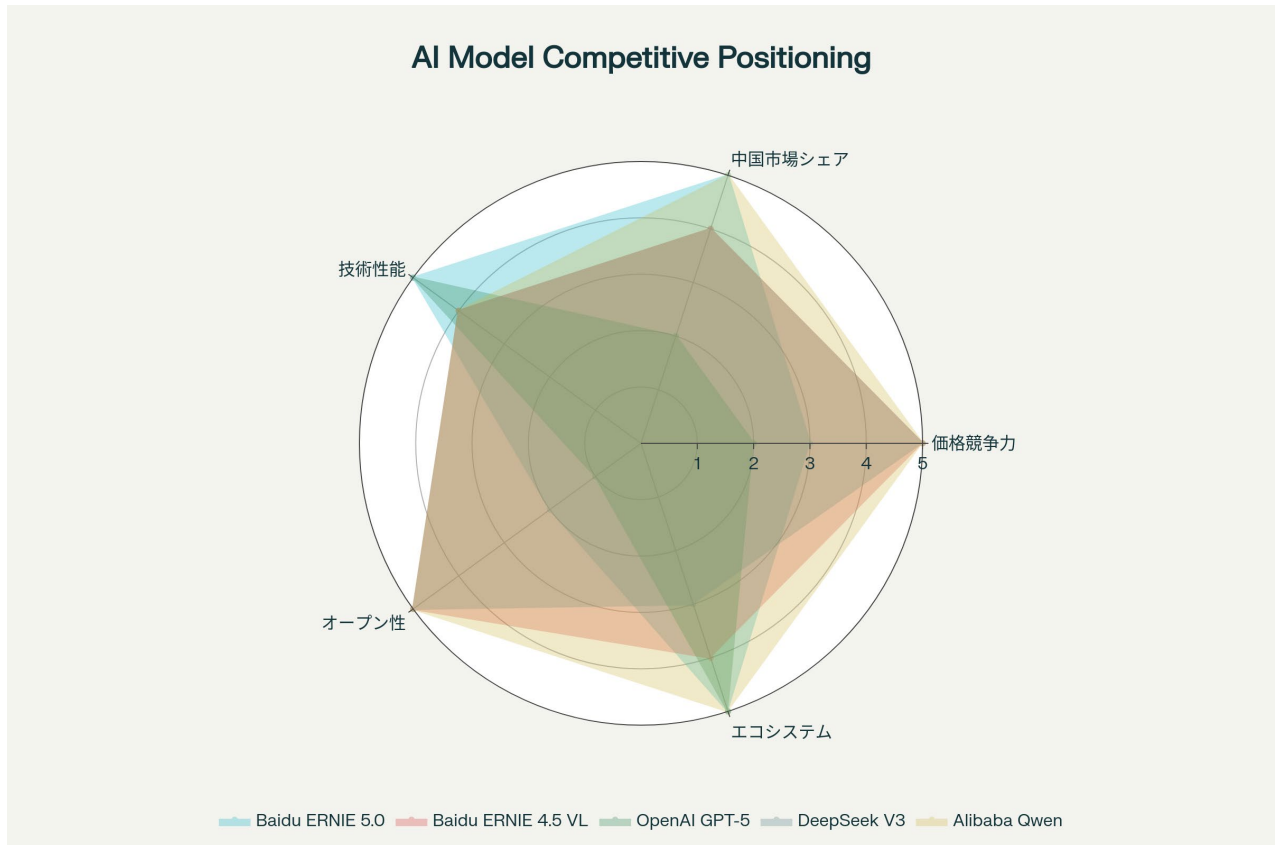
ERNIE 4.5 シリーズのオープンソース化は、中国 AI 企業による戦略的な市場アプローチの一環である。Apache 2.0 ライセンスにより商用利用が自由に可能であり、開発者は自由にダウンロード、実験、カスタマイズ、展開が可能である。HuggingFace、GitHub、Baidu AI Studio など複数のプラットフォームで公開されており、グローバルな開発者コミュニティへのアクセスを最大化している。

[\[4\]\[23\]\[29\]\[30\]\[26\]\[28\]\[31\]\[32\]](#)

この戦略は、Meta が Llama モデルで採用した「開発者通貨」戦略を踏襲するものであり、広範なユーザーベースの獲得、エコシステムの拡大、事実上の標準としての地位確立を目指している。実際、中国のオープンソース AI モデルは 2025 年に急速に台頭しており、HuggingFace プラットフォームでのダウンロード数において、Alibaba の Qwen シリーズが 3 億 8500 万ダウンロードに達し、Meta の Llama（3 億 4600 万ダウンロード）を上回っている。毎月リリースされる新しい言語モデルの 40% 以上が Qwen アーキテクチャをベースにしており、中国 AI モデルの影響力の拡大を示している。

[\[33\]\[31\]\[34\]\[32\]](#)

## Baidu の二刀流戦略：市場セグメンテーションと競争優位性



主要 AI モデルの競合ポジショニング比較 (2025 年 11 月)

## デュアル戦略の論理

Baidu の二刀流戦略は、AI 市場の異なるセグメントに最適化されたソリューションを提供する高度に計算された戦略である。ERNIE 5.0 は、最先端性能を求める大企業や特定の高度なユースケースに対して、プロプライエタリな商用 API として提供される。一方、ERNIE-4.5-VL-28B-A3B-Thinking は、より広範な開発者コミュニティやコスト効率を重視する企業に対して、アクセスしやすくカスタマイズ可能なオープンソースモデルとして提供される。[\[1\]\[2\]\[3\]\[4\]\[35\]](#)

この戦略は、AI 業界の構造的変化に対する Baidu の洞察に基づいている。Robin Li は、過去の AI 産業構造がチップレイヤーが価値の大部分を獲得する「正立ピラミッド」であったのに対し、今後はアプリケーションレイヤーが基盤モデルの 100 倍の価値を創出する「逆ピラミッド」構造になるべきだと主張している。この認識の下、Baidu は基盤モデルの開発に継続的に投資しつつ、アプリケーションエコシステムの構築に注力している。[\[6\]\[36\]\[12\]\[11\]](#)

## フルスタック AI アーキテクチャ

Baidu の競争優位性は、チップからアプリケーションまでを網羅するフルスタック AI アーキテクチャにある。2025 年 11 月 13 日、Baidu は ERNIE 5.0 と同時に、次世代 AI 専用チップ「Kunlun M100」と「Kunlun M300」を発表した。M100 は大規模推論シナリオに最適化されており 2026 年初頭に商用化予定、M300 は超大規模マルチモーダルモデルの訓練・推論向けに 2027 年初頭にリリース予定である。[\[37\]\[38\]\[39\]\[36\]\[12\]\[40\]](#)

さらに、最大 256 個の GPU を接続する「Tianchi 256 スーパーノード」を 2026 年前半に、512 個を接続する「Tianchi 512」を 2026 年後半にリリース予定であり、4 月に発表したスーパーノードソリューションと比較して総帯域幅が 4 倍、全体性能が 50%以上向上し、主要モデル推論ワークロードにおける GPU あたりのトークンスループットが 3.5 倍に増加するとされている。[\[38\]\[41\]](#)

このハードウェアとソフトウェアの垂直統合により、Baidu は効率性、開発速度、コスト最適化において決定的な優位性を獲得している。特に、米国の先進 AI チップ輸出規制が強化される中、国産チップの開発は中国 AI 企業にとって技術的主権の確保という戦略的重要性を持つ。[\[39\]\[41\]\[36\]\[42\]](#)

## AI アプリケーションエコシステムの拡充

Baidu World 2025 では、ERNIE 5.0 に加えて、包括的な AI アプリケーション群のアップグレードが発表された。[\[1\]\[5\]\[43\]](#)

**GenFlow 3.0** は、複雑なタスクとワークフローを簡素化する汎用 AI エージェントであり、2000 万人以上のユーザーを擁し、世界最大の汎用エージェントとなっている。**GenFlow 3.0** は大幅に強化されたマルチモーダリティとメモリ能力を備え、全モーダリティにわたるコンテンツの統合入力、処理、出力が可能である。[\[5\]\[43\]\[1\]](#)

**\*\*Famou (発茂) \*\***は、世界初の商用利用可能な自己進化型エージェントとして発表された。**Famou** はトップアルゴリズム専門家をシミュレートまたは超越し、複雑な問題を迅速に抽象化し、条件の変化に応じて自動的に反復して動的な最適解を提供できる。交通、エネルギー、金融、物流などの複雑なシナリオに適用可能であり、招待コードによる利用が開始されている ([famou.com](http://famou.com) から申請可能)。[\[43\]\[1\]\[5\]](#)

**Miaoda 2.0 / MeDo** は、ノーコードアプリケーションビルダーであり、すでに 40 万以上のアプリケーション生成に使用されている。**Miaoda** の国際版である **MeDo** が発表され、グローバル開発者向けに [medo.dev](http://medo.dev) で利用可能となった。[\[1\]\[5\]\[43\]](#)

**Oreate** は、学習と生産性のためのワンストップ AI ワークスペースとして、国際市場向けに設計された。マルチエージェントアーキテクチャを採用し、文書、スライド、画像、動画、ポッドキャストを含む全シナリオ・全モダリティにわたるエンドツーエンドの作成が可能であり、グローバル市場で 120 万人以上のユーザーを獲得している。[5][43][11]

**デジタルヒューマン技術**も大きく進化し、次世代のリアルタイムインタラクティブデジタルヒューマンが発表された。このデジタルヒューマンは、現実世界の文脈を深く理解して即座にフィードバックを提供するだけでなく、インタラクション中に全モーダルアラインメントを伴う自然な感情を表現する。2025 年のダブル 11 ショッピングフェスティバルでは、ライブストリーマーの 83%が Baidu のデジタルヒューマン技術を使用し、参加ライブストリームが前年比 119%増加、GMV（流通総額）が 91%増加した。この技術はブラジルでデビューし、米国や東南アジア、Shopee や Lazada などのプラットフォームへの展開を模索している。[43][11][5]

**Apollo Go** は、Baidu の自律走行配車サービスであり、累計 1700 万回以上の配車を完了し、世界最大規模を誇る。週あたり 25 万回以上の配車を実施し、累計 2 億 4000 万キロメートル以上を走行している。Robin Li は、ロボタクシーのコストが継続的に低下することで、手頃な価格が需要を大幅に押し上げ、自動運転車が新しいモバイルリビングスペースとなり、無限の可能性を開くと述べている。[11][43]

## Baidu 検索の AI 変革

Baidu 検索は、世界のどの検索エンジンよりも積極的な AI 変革を遂げており、トップ 1 検索結果の約 70%がリッチメディア形式で提示されている。Robin Li は、「AI サマリーを単に挿入するのではなく、検索をテキストとリンクベースのアプリケーションから、画像や動画などのリッチメディアを中心とした AI アプリケーションに変革した」と述べている。[11][5]

Baidu は AI 検索機能を AI API として公開しており、これが業界の基盤技術となりつつある。すでに Baidu AI Cloud を通じて 625 のパートナーが利用しており、Samsung、Honor、vivo などの主要ブランドが含まれている。この戦略は、Google が自社の Gemini モデルを検索機能に統合し、毎月 15 億人以上が「AI Overviews」を利用している状況と類似しており、検索エンジン市場における AI 統合競争の激化を示している。[5][44][11]

## グローバル AI 競争における位置づけと課題

## 中国 AI モデルの台頭

Baidu の ERNIE 発表は、より広範な中国 AI モデルの台頭という文脈で理解する必要がある。2025 年、中国企業はオープンソース AI モデル開発においてグローバルリーダーとしての地位を確立した。DeepSeek や Alibaba などの中国企業は、主要なオープンソース AI プラットフォームである Hugging Face を支配しており、DeepSeek のモデルは 12,000 以上の「いいね」を獲得し、Meta の Llama モデルのほぼ 2 倍に達している。<sup>[33][34][45]</sup>

中国のオープンソース AI 戦略は、頻繁なリリースと迅速な反復により忠実なユーザーベースを構築し、MoE 層やスパースアテンション機構などの技術により高性能モデルを劇的に低コストで訓練することを可能にしている。例えば、DeepSeek の R1 モデルの訓練コストは約 600 万ドルであったのに対し、OpenAI の GPT-4 は 1 億ドルを超えたとされている。<sup>[33]</sup>

この動きに対し、米国は 2025 年の AI 行動計画を通じて対応を強化しており、「米国の価値観に根ざしたオープンモデル」の開発を強調している。ATOM プロジェクトが立ち上げられ、数万の最先端 GPU を展開して中国のオープンソース AI 研究所と競争することを目指している。OpenAI や Hugging Face の幹部を含む米国 AI 業界のリーダーがこのイニシアチブを支援している。<sup>[33]</sup>

## 競合モデルとの比較

ERNIE 5.0 は、GPT-5-High や Gemini 2.5 Pro などのグローバルトップモデルと肩を並べる性能を示しているが、中国国内市場では他の中国企業との激しい競争にも直面している。<sup>[31][71][46]</sup>

DeepSeek は、低コストで高性能な推論モデルを提供することで市場を破壊しており、DeepSeek R1 の価格は入力 0.55 ドル/出力 2.19 ドルであり、当時の SOTA（最先端）モデルであった o1 と比較して出力トークン価格で 90%以上低い。Baidu は 2025 年 3 月に ERNIE X1 が DeepSeek R1 と同等の性能を半額で提供すると主張したが、詳細なベンチマークは提供されていない。実際のユーザーテストでは、ERNIE X1 は DeepSeek に比べて使いやすさの問題があり、特に中国語インターフェースの障壁が指摘されている。<sup>[20][47][46][44]</sup>

Alibaba Qwen は、オープンソースモデルのダウンロード数でトップを走っており、HuggingFace での総ダウンロード数が 3 億 8500 万に達している。Qwen シリーズは多様性とエコシステムの強みを持ち、毎月リリースされる新しい言語モデルの 40%以上が Qwen アーキテクチャをベースにしている。<sup>[33][34]</sup>

**Moonshot AI** の「Kimi K2 Thinking」は、2025 年 11 月に「Humanity's Last Exam」ベンチマークで 44.9%のスコアを記録し、GPT-5 と Claude Sonnet 4.5 を上回ったと主張している。完全無料で利用可能であり、コスト効率の高い訓練プロセスが注目されている。<sup>[48]</sup>

**ByteDance** の「Doubao (豆包)」は、急速にユーザー数を伸ばしており、2024 年 4 月には 900 万ダウンロードに達し、ERNIE Bot の 800 万を上回った。月間アクティブユーザー数も 400 万以上を維持しており、ERNIE Bot は登録ユーザー4 億 3000 万人を報告しているにもかかわらず、アクティブユーザーとの大きなギャップが問題視されている。<sup>[49][50]</sup>

## ERNIE Bot のユーザー基盤と課題

ERNIE Bot (中国名：文小言、旧称：文心一言) は、2023 年 8 月 31 日に一般公開され、2023 年 12 月までに 1 億ユーザーを突破した。2024 年 4 月には 2 億ユーザー、2024 年 6 月には 3 億ユーザー、2024 年 11 月には 4 億 3000 万ユーザーに達したと報告されている。1 日あたり 15 億回の API 呼び出しを処理しており、8 月の 6 億回から約 3 倍に増加している。<sup>[51][52][49][53]</sup>

しかし、登録ユーザー数と実際のアクティブユーザー数には大きな乖離がある。2025 年初頭のデータでは、月間アクティブユーザー数は約 1300 万人とされており、累計登録数との差が顕著である。2024 年 3 月のウェブサイト・アプリ訪問数は 1490 万回であったのに対し、競合の Kimi (Moonshot AI) は 1260 万回であり、成長率では Kimi が月次で 321.6%増加したのに対し、ERNIE Bot は 48%増にとどまっている。<sup>[54][52]</sup>

さらに、2024 年 12 月の ERNIE Bot のダウンロード数は 61 万 1619 件と、2023 年 9 月のピーク 150 万件から大幅に減少している。これは、ByteDance の Doubao や Alibaba の Qwen、Moonshot AI の Kimi などの競合が急速に台頭している市場環境を反映している。<sup>[50]</sup>

## 技術的課題と批判

ERNIE 5.0 の性能は印象的であるが、いくつかの技術的課題も報告されている。あるユーザーは、SVG 生成時にツール使用を停止するよう指示してもツールの使用を止められない問題を指摘しており、強化学習による「脳損傷」またはチャットプラットフォーム/システムプロンプトの深刻な問題が存在する可能性を示唆している。<sup>[9]</sup>

また、ベンチマーク結果の大半が Baidu 自身の発表に基づいているため、真の実力評価には第三者による独立した検証が必要である。特に、ERNIE X1 に関しては、DeepSeek R1 と同等の性能を半額で

提供するという主張について、詳細なベンチマークが公開されていないため、実際の競争力については不確実性が残る。<sup>[3][20][46]</sup>

グローバル市場においては、ERNIE Bot および ERNIE API へのアクセスが中国外のユーザーにとって容易ではないという課題もある。多くの機能が中国語インターフェースでのみ利用可能であり、Qianfan プラットフォームでのアカウント作成には中国の電話番号が必要であるなど、国際展開における障壁が存在する。<sup>[55][47][46]</sup>

## 戦略的含意と将来展望

### AI 価格競争の激化

Baidu の ERNIE 5.0 および ERNIE 4.5 の発表は、グローバル AI 市場における価格競争の激化を象徴している。2025 年 3 月の ERNIE 4.5 と ERNIE X1 の発表時、技術評論家 Robert Scoble は「AI 価格競争が始まった！」とコメントした。ERNIE 4.5 は GPT-4.5 の価格の 1%でありながら、多くのベンチマークで GPT-4.5 を上回ると主張されている。<sup>[20][56][46]</sup>

この価格競争は、他のプレーヤーにも価格引き下げやより速いイノベーションを強いており、全体的なコスト低下を推進している。DeepSeek R1 のデビュー価格は入力 0.55 ドル/出力 2.19 ドルであったが、その後推論モデル価格は急落し、OpenAI は最近フラッグシップモデルの価格を 80%削減した。<sup>[44][20]</sup>

### オープンソース vs プロプライエタリの戦略的選択

Baidu の CEO Robin Li は、ほとんどの企業にとってオープンソースモデルへの依存は「IQ タックス」であると主張し、クローズドソースモデルが 3 つの主要な理由で優れた価値提案を提供すると述べている：(1)性能の優位性、(2)総所有コストの低減、(3)効率性と専門化の向上。<sup>[39]</sup>

しかし、Baidu は同時に ERNIE 4.5 シリーズを Apache 2.0 ライセンスでオープンソース化しており、この一見矛盾する戦略は、実際には市場セグメンテーションの高度な戦略である。プロプライエタリモデルは高性能と包括的サポートを求める企業顧客向けであり、オープンソースモデルは開発者エコシステムの構築、技術標準の確立、広範な市場浸透を目的としている。<sup>[30][31][32]</sup>

Meta が Llama でオープンソース戦略に転換して成功を収めたのに対し、OpenAI は 6 年間主にクローズド開発を維持した後、最近になってようやくオープンソースモデルをリリースした。この構造的・

文化的変化は、米国がオープンソース AI 領域で地位を回復するための困難な道のりを強調している。<sup>[33]</sup>

## 地政学的次元とエネルギー優位性

AI 競争が本質的にインフラ競争、すなわち大規模でエネルギー集約的なデータセンターを構築・電力供給する能力の競争となる場合、中国は現在 **significant advantage** を保持している。中国は大規模プロジェクトを速度と調整をもって実行する能力を示しており、補助金付き電力とオープンソースイノベーションの組み合わせにより、米国の支配に対する深刻な挑戦を行っている。<sup>[45]</sup>

Nvidia の CEO Jensen Huang は、中国が AI 競争に勝つ可能性があることを認めつつも、後に中国は米国よりナノ秒単位でしか遅れていないと発言を和らげた。Huang は中国を敵対者ではなく手強い競争相手として位置づけ、中国の広範な AI 開発者コミュニティが進歩を推進する重要な要因であると強調した。彼は、米国が中国の開発者に米国の技術プラットフォームの採用を奨励することで積極的に関与し、米国のアーキテクチャをグローバルな AI 標準として位置づけることを提唱した。<sup>[34]</sup>

元 Google CEO Eric Schmidt は、無料の中国 AI モデルが重大な競争脅威となり、グローバルな AI 市場の景観を再形成する可能性があるかと警告している。強力な AI モデルを世界中のユーザーに無料で提供することで、中国企業は事実上のグローバル標準を確立できる可能性があり、これは Android と iOS がユビキタスなモバイルプラットフォームになったのと同様のパターンである。このアプローチは、即座の収益性よりも市場浸透とユーザー採用を優先するものであり、以前の技術サイクルで効果的であることが証明されている。<sup>[52]</sup>

## Baidu の AI エコシステムの進化

Baidu の長期的な成功は、単一のモデルの性能ではなく、包括的な AI エコシステムの構築にかかっている。同社は、検索、生産性アプリケーション、ノーコード開発ツール、デジタルヒューマン、自律走行など、多様な領域に AI を統合している。<sup>[1][5][58][43]</sup>

Robin Li は、「AI を内部化すると、それはネイティブな能力となり、知性をコストから生産性の源泉に変える」と述べ、企業と個人の成長のためのネイティブな推進力として AI を統合することの重要性を強調した。この哲学は、AI 技術を孤立したツールとしてではなく、あらゆる業務プロセスに織り込まれた基盤的能力として扱うことを意味している。<sup>[59][11]</sup>

Baidu の特許ポートフォリオと知的財産戦略も、長期的な競争優位性の重要な要素である。フルスタックアーキテクチャは単なる技術的資産ではなく、コスト効率的なアプリケーション展開を可能にす

る本質的な経済エンジンであり、これはモデル性能が商品化されつつある世界において、アプリケーションレイヤーと運用効率で勝利するための新しいレースに適応するものである。<sup>[39]</sup>

## 2026 年以降の展望

ERNIE 5.0 と ERNIE-4.5-VL-28B-A3B-Thinking の発表は、Baidu にとって終点ではなく、継続的なイノベーションサイクルの一部である。Kunlun M100 チップの 2026 年初頭の商用化、Kunlun M300 の 2027 年初頭のリリース、そして Tianchi 256 および Tianchi 512 スーパーノードの段階的展開は、Baidu が長期的な技術開発ロードマップを持っていることを示している。<sup>[38][36][40]</sup>

Robin Li は、「AI agents themselves are the most significant applications (AI エージェント自体が最も重要なアプリケーションである)」と述べ、技術イテレーションの速度が唯一の競争上の堀であると強調した。Famou のような自己進化型エージェントの商用展開は、AI が単なるツールから自律的な問題解決システムへと進化する次の段階を示している。<sup>[1][5][43]</sup>

グローバル展開においては、MeDo、Oreate、デジタルヒューマン技術の国際市場への展開が進行中であり、特にブラジル、東南アジア、米国市場が重点対象となっている。ただし、言語バリア、規制要件、現地での競合との差別化など、克服すべき課題も多い。<sup>[5][47][46][1]</sup>

中国国内市場では、Apollo Go の累計 1700 万回を超える配車実績は、実世界での大規模 AI 展開における Baidu の能力を実証しているが、競合するロボタクシーサービスとの競争も激化している。自律走行技術が都市生活を再形成し、交通だけでなくより広範な社会エコシステムを変革するという Robin Li のビジョンは、AI の影響が情報処理を超えて物理世界に及ぶことを示している。<sup>[43][1]</sup>

## 結論

Baidu の 2025 年 11 月の AI 発表は、単なる新製品のリリースではなく、グローバル AI 競争における戦略的ポジショニングの再定義である。ERNIE 5.0 は、ネイティブオムニモーダル技術と 2.4 兆パラメータ超スパース MoE アーキテクチャにより、中国 AI モデルがグローバルトップティアに到達したことを示している。同時に、ERNIE-4.5-VL-28B-A3B-Thinking の Apache 2.0 オープンソース化は、開発者エコシステムの獲得と技術標準の確立を目指す戦略的な動きである。<sup>[1][6][4][26][7][31]</sup>

この二刀流戦略は、AI 市場の異なるセグメントを同時に攻略する洗練されたアプローチであり、プレミアム商用顧客と広範な開発者コミュニティの両方にリーチすることを可能にする。Baidu のフルスタック垂直統合—Kunlun チップ、PaddlePaddle フレームワーク、ERNIE モデル、GenFlow や Famou

などのアプリケーションは、効率性、速度、コスト最適化において決定的な競争優位性を提供する。<sup>[2][3][35][39][36][12]</sup>

しかし、課題も存在する。ERNIE Bot のアクティブユーザー数と登録ユーザー数のギャップ、ByteDance や Alibaba などの国内競合の台頭、グローバル市場でのアクセシビリティの問題、そして第三者ベンチマーク検証の必要性などが、Baidu が直面する現実的な障害である。さらに、米国の輸出規制強化、OpenAI や Google などの西側テック巨人からの競争圧力、急速に変化する技術環境は、継続的なイノベーションと適応を要求している。<sup>[4][1][33][45][54][52][50]</sup>

最終的に、Baidu の ERNIE 戦略の成功は、技術性能だけでなく、実世界のアプリケーションにおける価値創出、持続可能なエコシステムの構築、そしてグローバル市場での信頼獲得にかかっている。Robin Li が強調するように、「速度とイテレーションが唯一の堀」である世界において、Baidu は技術開発とビジネス戦略の両面で継続的な進化を遂げなければならない。ERNIE 5.0 と ERNIE-4.5-VL-28B-A3B-Thinking は、その進化の重要な一歩であり、今後数年間のグローバル AI 競争の展開を大きく左右する可能性を秘めている。<sup>[5][1]</sup>

\*  
\*\*

1. <https://www.prnewswire.com/news-releases/baidu-unveils-ernie-5-0-and-a-series-of-ai-applications-at-baidu-world-2025--ramps-up-global-push-302614531.html>
2. <https://www.artificialintelligence-news.com/news/baidu-ernie-multimodal-ai-gpt-and-gemini-benchmarks/>
3. <https://innovatopia.jp/ai/ai-news/71772/>
4. <https://dev.to/czmilo/2025-complete-guide-in-depth-analysis-of-ernie-45-vl-28b-a3b-thinking-multimodal-ai-model-1mib>
5. <https://ai-watch.jp/english/32718/>
6. <https://global.chinadaily.com.cn/a/202511/13/WS691571bda310d6866eb29500.html>
7. <https://news.aibase.com/news/22776>
8. <https://weel.co.jp/media/tech/ernie-5-0/>
9. <https://x.com/scaling01/status/1988973299515474136>
10. <https://x.com/thetripathi58/status/1988984042151375123>

11. [https://x.com/manishkumar\\_dev/status/1988986965816135772/photo/4](https://x.com/manishkumar_dev/status/1988986965816135772/photo/4)
12. <https://www.chinadailyasia.com/article/623566>
13. <https://www.implicator.ai/baidus-ernie-5-0-proves-technical-excellence-no-longer-wins-chinas-ai-war/>
14. <https://venturebeat.com/ai/baidu-unveils-proprietary-ernie-5-beating-gpt-5-performance-on-charts>
15. <https://ernie.baidu.com/blog/posts/ernie-5.0-preview-1022-release-on-lmarena/>
16. <https://lmarena.ai/leaderboard>
17. <https://news.aibase.com/news/22778>
18. <https://x.com/themacrosift/status/1989353899489481048>
19. <https://www.prnewswire.com/news-releases/baidu-unveils-ernie-4-5-and-reasoning-model-ernie-x1--makes-ernie-bot-free-ahead-of-schedule-302402490.html>
20. <https://apidog.com/blog/baidu-ernie-4-5-x1-2/>
21. [https://x.com/Baidu\\_Inc/status/1901106584782270883](https://x.com/Baidu_Inc/status/1901106584782270883)
22. <https://tribune.com.pk/story/2534801/baidu-unveils-low-cost-ai-models-to-compete-with-deepseek-in-chinas-tech-race>
23. <https://www.rockbirdmedia.com/post/baidu-ernie-multimodal-ai-benchmarks>
24. <https://apxml.com/models/ernie-4.5-vl-28b-a3b>
25. <https://zenn.dev/qingwu/articles/0ee5df24da737e>
26. <https://ernie.baidu.com/blog/posts/ernie-4.5-vl-28b-a3b-thinking/>
27. <https://zenn.dev/czmilo/articles/fcaa5dca44d10c>
28. <https://aistudio.baidu.com/modelsdetail/39280/intro>
29. <https://huggingface.co/baidu/ERNIE-4.5-VL-28B-A3B-Thinking>
30. <https://ernie.baidu.com/blog/posts/ernie4.5/>
31. <https://www.linkedin.com/pulse/baidus-ernie-goes-open-source-power-move-ai-wars-nantha-kumar-l-r4imc>
32. <https://notes.suhaib.in/docs/tech/news/baidu-open-sources-ernie-4.5-chinas-multimodal-model-goes-global/>
33. <https://nexth.zone/blog/chinese-open-source-ai-models-dominate-global-landscape-in-2025/142>

34. <https://www.digitimes.com/news/a20251111PD202/silicon-valley-cost-performance-technology-qwen.html>
35. <https://xenospectrum.com/baidu-ernie-5-omnimodal-ai-beats-gpt5/>
36. [https://mobile.chinadaily.com.cn/cn/html5/2025-11/14/content\\_013\\_6916532bed50ccabe151cebd.htm](https://mobile.chinadaily.com.cn/cn/html5/2025-11/14/content_013_6916532bed50ccabe151cebd.htm)
37. <https://www.reuters.com/world/china/chinas-baidu-unveils-new-ai-processors-supercomputing-products-2025-11-13/>
38. <https://x.com/jukanlosreve/status/1988805721904017483>
39. <https://www.chinaseo.com/baidus-ai-pivot-a-strategic-analysis-of-a-full-stack-transformation/>
40. <https://jp.reuters.com/markets/global-markets/BCW47CAANRMW3K6XDMFAZK43RI-2025-11-13/>
41. <https://www.tradingview.com/news/invezz:89eaa89e7094b:0-what-baidu-s-new-chips-reveal-about-china-s-plan-to-counter-nvidia-in-ai-race/>
42. <https://finance.yahoo.com/news/chinas-baidu-unveils-ai-processors-062955373.html>
43. <https://www.stocktitan.net/news/BIDU/baidu-unveils-ernie-5-0-and-a-series-of-ai-applications-at-baidu-hq8a7kgdupq6.html>
44. <https://www.techradar.com/pro/why-baidus-ernie-matters-more-than-deepseek>
45. <https://fortune.com/2025/11/11/is-china-about-to-win-the-ai-race/>
46. <https://www.datacamp.com/blog/ernie-4-5-x1>
47. [https://www.youtube.com/watch?v=NBvqP1MHd\\_o](https://www.youtube.com/watch?v=NBvqP1MHd_o)
48. <https://convergence-now.com/artificial-intelligence/chinese-ai-model-beats-gpt5-free/>
49. <https://searchendurance.com/baidu-statistics/>
50. <https://www.digitimes.com/news/a20250115VL207/bytedance-baidu-market-alibaba-ernie-bot.html>
51. <https://neuron.expert/news/baidus-ernie-bot-surpasses-200-million-users-solidifying-its-position-in-chinas-ai-market/14411/en/>
52. [https://www.samenacouncil.org/samena\\_daily\\_news?news=100277](https://www.samenacouncil.org/samena_daily_news?news=100277)
53. [https://en.wikipedia.org/wiki/Ernie\\_Bot](https://en.wikipedia.org/wiki/Ernie_Bot)
54. <https://www.datastudios.org/post/ai-chatbots-users-global-numbers-of-the-major-ones>
55. [https://www.reddit.com/r/singularity/comments/1jcbpnk/baidu\\_releases\\_ernie\\_45\\_x1\\_half\\_the\\_price\\_of/](https://www.reddit.com/r/singularity/comments/1jcbpnk/baidu_releases_ernie_45_x1_half_the_price_of/)

56. <https://www.techtarget.com/searchenterpriseai/news/366620945/Baidus-low-priced-new-models-bring-questions-about-cost>
57. <https://ai-daily.news/articles/eric-schmidt-warns-free-chinese-ai-models-could-reshape-glob>
58. <https://kr-asia.com/baidu-reimagines-search-as-the-race-to-build-the-first-ai-superapp-continues>
59. <https://www.eweek.com/news/baidu-ernie-5-0-release/>
60. [https://x.com/Baidu\\_Inc/status/1988931336615371261](https://x.com/Baidu_Inc/status/1988931336615371261)
61. <https://www.nasdaq.com/articles/baidu-unveils-ernie-50>
62. <https://www.index.dev/blog/comparing-unimodal-vs-multimodal-models>
63. <https://finance.yahoo.com/news/baidu-unveils-ernie-5-0-131700248.html>
64. <https://www.ukiyo-journal.com/article/baidu-ai-revolution-ernie-5-0-semiconductor-strategy>
65. [https://www.youtube.com/watch?v=e2v\\_IdT9YoA](https://www.youtube.com/watch?v=e2v_IdT9YoA)
66. <https://x.com/AIMIRAI46487/status/1988872935595999467>
67. [https://x.com/Baidu\\_Inc/status/1978098969529155962](https://x.com/Baidu_Inc/status/1978098969529155962)
68. <https://www.gettyimages.co.jp/editorial-images/news/event/baidu-world-2025-held-in-beijing/776415195?editorialproducts=all>
69. <https://www.isna.ir/news/97112714167/ویدئو-تشخیص-بیکر-شهادت-ای-حادثه-تر-ورستی-خاش?s-news-9055018-2025-11-14-baidu-unveils-ernie-50-and-advanced-ai-hardware-to-strengthen-chinas-ai-ecosystem>
70. [https://www.reddit.com/r/singularity/comments/1ovolve/ernie\\_50\\_released\\_achieving\\_frontier\\_performance/](https://www.reddit.com/r/singularity/comments/1ovolve/ernie_50_released_achieving_frontier_performance/)
71. <https://www.dera.ai/news/1fe94819-3996-6e2d-a12a-ba069cdb4129>
72. <https://www.facebook.com/marius.comper/posts/-ernie-50s-benchmark-results-suggest-that-baidu-has-achieved-parityor-near-parit/10163542561694621/>
73. <https://jp.investing.com/news/company-news/article-93CH-1323304>
74. [https://x.com/Baidu\\_Inc/status/1988820837898829918](https://x.com/Baidu_Inc/status/1988820837898829918)
75. <https://news.aibase.com/news/22703>
76. <https://www.dera.ai/en/news/9877bdd7-9f8f-36d1-53ef-23727c81adad>
77. <https://blog.galaxy.ai/model/ernie-4-5-21b-a3b-thinking>

78. <https://github.com/PaddlePaddle/ERNIE>
79. [https://yiyan.baidu.com/blog/publication/ERNIE Technical Report.pdf](https://yiyan.baidu.com/blog/publication/ERNIE_Technical_Report.pdf)
80. <https://innovatopia.jp/ai/ai-news/71466/>
81. <https://weel.co.jp/media/tech/ernie-4-5-vl-28b-a3b-thinking/>
82. <https://routstr.com/models/baidu/ernie-4.5-21b-a3b-thinking>
83. <https://ai-allforces.allforces.wuaze.com/2025/11/15/2-baidu-ernie/>
84. <https://www.youtube.com/watch?v=nhkG44nEkrw>
85. [https://x.com/AI\\_Techie\\_Arun/status/1989167385954185354](https://x.com/AI_Techie_Arun/status/1989167385954185354)
86. <https://www.youtube.com/watch?v=jrr6NJ5Yis>
87. <https://artificialanalysis.ai/models/comparisons/gpt-5-vs-gemini-2-5-pro>
88. [https://www.reddit.com/r/perplexity\\_ai/comments/1ov4gu5/baidus\\_latest\\_model\\_beats\\_gpt5\\_and\\_gemini\\_25\\_pro/](https://www.reddit.com/r/perplexity_ai/comments/1ov4gu5/baidus_latest_model_beats_gpt5_and_gemini_25_pro/)
89. <https://www.youtube.com/watch?v=zggbwYu1tdw>
90. <https://www.youtube.com/watch?v=xm0IzmUmJ2Q>
91. <https://venturebeat.com/ai/baidu-just-dropped-an-open-source-multimodal-ai-that-it-claims-beats-gpt-5>
92. <https://pandaily.com/baidu-s-robin-li-internalizing-ai-to-turn-intelligence-into-productivity>
93. <https://www.datacenterdynamics.com/en/news/baidu-launches-kunlun-m100-and-m300-ai-chips/>
94. <https://www.aspendigital.org/blog/ai-geopolitics-beyond-the-us-china-rivalry/>
95. <https://kitemetric.com/blogs/baidu-s-ernie-4-5-open-sourcing-its-ai-powerhouse>
96. <https://blog.galaxy.ai/compare/deepseek-chat-v3-0324-vs-ernie-4-5-21b-a3b>
97. <https://sourceforge.net/software/compare/DeepSeek-vs-Ernie-Bot/>
98. <https://www.labellerr.com/blog/baidu-launches-ernie-4-5-and-x1/>
99. <https://www.aa.com.tr/en/science-technology/chinese-firm-baidu-unveils-advanced-ai-chips-amid-rising-industry-competition/3742819>
100. <https://archway.finance/blog/baidu-ernie-open-source-ai-revolution>

101. <https://explyt.ai/en/blog/best-chinese-llm>
102. <https://www.forbes.com/sites/viviantoh/2025/07/07/how-baidus-ernie-45-is-catalyzing-chinas-ai-transformation/>
103. <https://www.cnbc.com/2025/06/29/china-biggest-ai-drop-since-deepseek-baidus-ernie-to-hit-market.html>
104. <https://theconversation.com/will-china-win-the-ai-race-269415>