

百度(Baidu)2025年AI戦略詳細分析レポート : ERNIE 5.0とERNIE-4.5-VLが示す「ネイティブ・オムニモーダル」と「視覚的思考」の未来

Gemini

1. エグゼクティブサマリー

2025年11月に北京で開催された「Baidu World 2025」は、中国のテクノロジー大手である百度(Baidu)にとって、単なる新製品発表会以上の意味を持つ分水嶺となりました。米国による先端半導体規制が常態化し、生成AIの技術競争が「モデルの規模」から「推論の効率性」と「エージェントの実用性」へとシフトする中、百度はAI戦略の根幹を揺るがす二つの記念碑的モデル、「ERNIE 5.0(文心一言 5.0)」および「ERNIE-4.5-VL-28B-A3B-Thinking」を発表しました。

本レポートでは、これらのモデルが提示する技術的パラダイムシフト—特に「ネイティブ・オムニモーダル」アーキテクチャと「視覚的思考(Thinking with Images)」—を詳細に分析します。ERNIE 5.0は2.4兆パラメータという圧倒的な規模で西側のフラッグシップモデル(GPT-5系列やGemini 2.5)に挑む一方、ERNIE-4.5-VLはオープンソースかつ軽量の推論モデルとして、エンタープライズAIの民主化を推進する「両利きの戦略」を展開しています。さらに、独自のAI半導体「Kunlun(昆侖)」のロードマップや、AIエージェント「GenFlow 3.0」によるアプリケーション層への価値転換など、インフラから応用までを垂直統合する百度の深層戦略を解き明かします。

技術的な卓越性を追求しつつも、株式市場の冷ややかな反応が示唆する「コモディティ化の罨」や「収益化への圧力」といった構造的課題についても、多角的な視点から論じます。本稿は、激動する2025年のAIランドスケープにおいて、百度が描く未来図の全貌を捉えるための羅針盤となることを目指しています。

2. ERNIE 5.0: ネイティブ・オムニモーダルの深層アーキテクチャと戦略的含意

2.1. 「ネイティブ・オムニモーダル」への技術的跳躍

ERNIE 5.0の核心は、百度が「ネイティブ・オムニモーダル (Native Omni-modal)」と定義する革新的なアーキテクチャにあります。これまでの多くのマルチモーダルモデル (LMM) は、テキスト処理に特化した大規模言語モデル (LLM) に対し、視覚エンコーダー (ViTなど) や音声モジュールを後付けて結合する「Late Fusion (後期融合)」アプローチを採用していました。この手法は開発効率が良い反面、モダリティ間の情報の継ぎ目で「意味の損失」が発生しやすく、特に複雑な推論を要するタスクにおいて文脈理解の限界を露呈していました。

対してERNIE 5.0は、学習の最初期段階 (Pre-training) から、テキスト、画像、音声、動画という異なるデータ形式を、単一の巨大な埋め込み空間 (Joint Embedding Space) で統合的にモデリングする設計思想に基づいています。この「Early Fusion (早期融合)」アプローチにより、モデルは外部アダプターや個別のエンコーダーを介することなく、あらゆる入力をネイティブに解釈します。

このアーキテクチャ上の刷新は、ユーザー体験に質的な変化をもたらします。例えば、従来のモデルでは「動画を見て、その中の特定の音声イベントに関連する視覚的变化を説明し、関連画像を生成する」といったクロスモーダルなタスクを行う際、複数のサブモジュール間で情報をリレーする必要があり、その過程で遅延や認識のズレが生じていました。ERNIE 5.0では、これら全ての処理が単一の推論パスの中でシームレスに実行されます。これは、百度が提唱する「プロンプトエンジニアリングの簡素化」にも直結しており、ユーザーはモダリティごとの複雑な指示を出すことなく、自然言語による対話だけで流動的なマルチメディア推論を引き出すことが可能になります。

2.2. 2.4兆パラメータとMoEによる「計算効率のパラドックス」の解消

ERNIE 5.0のパラメータ数は2.4兆 (2.4 Trillion) に達し、これはAlibabaのQwen3-MaxやMoonshot AIのKimi K2といった中国国内の競合モデルの倍以上の規模であると推計されています。しかし、2025年のAI開発における主要な命題は「規模の拡大」ではなく「効率の追求」です。巨大なモデルは、そのままでは推論コストが指数関数的に増大し、商用展開を阻害するからです。

この課題に対し、百度は極めて洗練されたMixture-of-Experts (MoE) アーキテクチャを導入しました。報告によれば、ERNIE 5.0は推論 (Inference) の際、全専門家 (Experts) のうち「3%未満」のパラメータしかアクティブにしません。これは、総パラメータ数が2.4兆であっても、1回のトークン生成に関与する計算量は数百億パラメータ相当に抑制されていることを意味します。

この設計は、特にGPUリソースが逼迫している中国のAI環境において決定的な意味を持ちます。NVIDIAのH100クラスターへのアクセスが制限される中、百度は限られた計算資源で世界最高峰の「知識容量 (総パラメータ数に依存)」と「応答速度 (アクティブパラメータ数に依存)」を両立させる必

要がありました。ERNIE 5.0のMoEアーキテクチャは、この背反する要件に対する百度の回答であり、推論レイテンシを劇的に削減しつつ、ロングテールな知識へのアクセス能力を維持することに成功しています。

2.3. 欧米フラッグシップとのベンチマーク対決と価格戦略

百度は、ERNIE 5.0が40以上の権威あるベンチマークにおいて、OpenAIのGPT-5-HighやGoogleのGemini 2.5 Proといった西側の最先端モデルと同等、あるいはそれ以上のスコアを記録したと発表しました。特筆すべきは、企業のバックオフィス業務やデジタルトランスフォーメーション(DX)に直結するドキュメント処理能力です。

- **OCRBench**: 光学文字認識の精度を測るテスト。
- **DocVQA**: 文書画像に対する質問応答能力。
- **ChartQA**: グラフやチャートの数理的・視覚的解釈能力。

これらのベンチマークにおける高スコアは、ERNIE 5.0が単なるチャットボットではなく、産業用途の「ナレッジワーカー」として設計されていることを示唆しています。また、生成能力においても、Googleの動画生成モデルVeo3と比較して、意味的一貫性や画質で同等以上の性能を示したとされています。

しかし、技術的な優位性以上に市場に衝撃を与えたのは、その攻撃的な価格戦略です。百度のクラウドプラットフォーム「Qianfan(千帆)」を通じて提供されるERNIE 5.0のAPI価格は、以下の通り設定されました。

- 入力: 100万トークンあたり0.85ドル
- 出力: 100万トークンあたり3.40ドル

これは、競合するOpenAIのGPT-5.1(入力\$1.25/出力\$10.00)やGoogleのGemini 2.5 Proと比較して、大幅に安価な設定です。性能面で肉薄しつつ、コスト面で明確なアドバンテージを提示することで、百度はグローバル市場、特にコスト感度の高いアジアや新興国のエンタープライズ層を取り込む戦略を鮮明にしています。この価格設定は、中国国内で激化する「モデル価格戦争」への対抗措置であると同時に、高騰するAI導入コストに躊躇する企業への強力な誘引剤として機能します。

3. ERNIE-4.5-VL-28B-A3B-Thinking: 視覚的思考とオープンソースによる破壊的イノベーション

ERNIE 5.0が「規模と統合」による正面突破を目指すモデルであるならば、同時に発表された「

ERNIE-4.5-VL-28B-A3B-Thinking」は、「知能の質」と「アクセシビリティ」を再定義する、より急進的な試みです。Apache 2.0ライセンスでのオープンソース公開は、百度が開発者エコシステムの主導権を握るための布石であり、その技術的深淵には目を見張るものがあります。

3.1. 「A3B」アーキテクチャ: 28Bの頭脳、3Bの身軽さ

モデル名の「A3B」は「Activated 3 Billion」を意味します。総パラメータ数は280億(28B)ですが、MoEルーターが入力に応じて最適なエキスパートネットワークのみを選択的に発火させるため、実行時の計算量はわずか30億(3B)パラメータ分に留まります。

このアーキテクチャの真価は、ハードウェア要件の劇的な緩和にあります。従来の同等性能を持つVLM(Vision-Language Model)は、推論に複数の高性能GPUを必要としましたが、ERNIE-4.5-VLは単一の80GB GPU(NVIDIA A100/H100の80GB版など)で動作可能です。80GB GPUの調達コストは1万ドル~3万ドル程度であり、数十万ドルを要するGPUクラスターと比較して圧倒的に安価です。これにより、データを社外に出したくない金融機関や医療機関が、自社サーバー(オンプレミス)内で最高レベルのマルチモーダルAIを運用することが現実的になります。

3.2. 「Thinking with Images」: システム2思考の視覚的実装

ERNIE-4.5-VLの白眉といえる機能が「Thinking with Images(画像による思考)」です。これは、ノーベル経済学賞受賞者ダニエル・カーネマンが提唱した「システム2(熟慮的、論理的な遅い思考)」のプロセスを、視覚情報の処理に応用したものです。

従来のVLMは、画像を一度エンコードすると、その固定された特徴量マップに基づいて回答を生成していました(システム1的思考)。対して、ERNIE-4.5-VLは以下のような動的かつ反復的な推論プロセスを実行します。

1. 初期スキャンと仮説形成: ユーザーの問い(例:「この工場のラインで、品質不良を起こしている部品はどれか?」)に対し、まず画像全体を俯瞰し、注目すべき領域の当たりをつけます。
2. 動的ツール利用(Tool Use): モデルは自律的に「画像ズームツール」や「画像検索ツール」を呼び出します。これは人間が虫眼鏡を使って細部を確認したり、リファレンス画像を検索して比較したりする行動を模倣したものです。
3. 再帰的詳細化: ズーム機能によって高解像度で再取得された特定領域の情報を基に、推論を更新します。例えば、遠くの看板の文字が潰れて読めない場合、モデルは「ズームイン」アクションを実行し、鮮明になった画像を再読込して文字を特定します。
4. 論理的統合: 収集した視覚的証拠と言語的知識を統合し、最終的な回答を生成します。

このプロセスにより、従来モデルが苦手としていた「画像の細部に宿る情報の解釈」や「ロングテール

な物体認識(一般的な学習データに少ない珍しい物体の特定)」において、劇的な性能向上を実現しました。

3.3. 独自の学習手法: GSPOとIcePopによる強化学習の革新

この高度な推論能力を支えているのが、百度が独自に開発した強化学習(RL)パイプラインです。

- **GSPO (Generative Self-Play Optimization)**: これはマルチモーダルタスクにおける強化学習の目的関数です。モデル自身が生成した回答候補の中で、より優れた推論プロセスを経たものを自己評価・選択し、学習データとしてフィードバックする「自己対局(Self-Play)」的な最適化を行います。
- **IcePop**: データの選択戦略アルゴリズムです。学習の安定性を保ちつつ、モデルにとって既知の簡単なデータではなく、学習効果の高い「困難な例(Hard Examples)」を動的にサンプリングする「動的難易度調整」を行います。これにより、学習効率が最大化されます。

さらに、トレーニングプロセスは以下の3段階で構成されています。

1. テキスト推論の基礎構築: まずLLM部分に対し、コード生成や論理パズルなどのテキストデータを用いて徹底的な推論訓練を施します。
2. マルチモーダルへの拡張: Reject Sampling(棄却サンプリング)を用い、視覚的なSTEM課題(数学、物理の図解問題など)に対する回答を生成させ、正解が得られた高品質なプロセスのみを学習データとして採用します。
3. 思考と非思考の融合: 「Thinking(思考)」モードと通常の「Non-Thinking」モードを統合します。ここでは、非思考データの回答前に<think>タグで囲まれた空の思考プロセスを挿入し、それをマスク(学習対象外)する特殊なテクニックを用いることで、モデルが「いつ深く考えるべきか」と「いつ即答すべきか」を文脈に応じて切り替える能力を獲得させます。

また、MoEの学習において課題となる専門家間の不均衡(特定のExpertばかりが使われる現象)を防ぐため、「Modality-Isolated Routing(モダリティ分離ルーティング)」や「Router Orthogonalization Loss(ルーター直交化損失)」といった高度な技術が導入されています。これにより、テキスト処理と画像処理が互いに干渉することなく、それぞれの専門性が最大限に発揮される構造が担保されています。

4. インフラストラクチャの自立: Kunlunチップと「脱NVIDIA」への道

百度のAI戦略を語る上で欠かせないのが、ハードウェア層での自立への執念です。米国の輸出規

制により、H100や次世代のRubinといったNVIDIA製GPUの調達が困難になる中、百度は独自のAI半導体「Kunlun(昆侖)」シリーズへの依存度を高める「フルスタック・バーティカル」戦略を推進しています。

4.1. M100とM300: 役割分担による最適化

Baidu World 2025で公開されたロードマップは、チップの用途を明確に分離することで効率化を図るものでした。

- **Kunlun M100**(2026年初頭リリース予定): 推論(Inference)特化型チップ。ERNIE 5.0のようなMoEモデルの推論を高速かつ低消費電力で実行することに最適化されています。
- **Kunlun M300**(2027年初頭リリース予定): 学習(Training)と推論の両用ハイエンドチップ。兆パラメータ級の超大規模モデルのトレーニング負荷に耐えうる設計がなされています。

このタイムラインは、2026年までは既存のストックやM100を用いた推論の効率化で凌ぎ、2027年にM300で本格的な自社製トレーニング環境へ移行するという、現実的かつ段階的な移行計画を示唆しています。

4.2. 「Tianchi」スーパーノードとシステムレベルの競争

単体のチップ性能でNVIDIAの最先端品に追いつくことは容易ではありません。そこで百度が採用したのが、システム全体での性能向上を目指す「Tianchi(天池)」スーパーノード構想です。

- **Tianchi 256**: 256個のKunlunチップを高速インターコネクトで結合したクラスター(2026年前半)。
- **Tianchi 512**: 512個結合(2026年後半)。
- **100万チップクラスター**: 2030年までに、数百万個規模のチップを論理的に結合し、単一の巨大な計算リソースとして扱う壮大な計画です。

このアプローチは、Huaweiが「Ascend(昇騰)」チップで行っている戦略と共鳴します。個々の計算ユニットの劣位を、超並列化と通信帯域の最適化によってカバーし、システムトータルでのスループットを最大化する戦略です。

さらに重要なのが、ソフトウェアエコシステムへの配慮です。発表によれば、KunlunチップはNVIDIAの「CUDA」との互換性を確保しているとされます。これは、既存のAI開発者がNVIDIA環境向けに書いたコードを、大幅な書き換えなしにKunlun環境へ移行できることを意味し、独自チップ普及の最大の障壁である「移行コスト」を下げるための極めて戦略的な一手です。しかし、CUDA互換レイヤーの完成度やパフォーマンス効率がどの程度かは未知数であり、ここが成功の鍵を握ります。

5. アプリケーション・エコシステム:「逆ピラミッド」への転換と実利の追求

百度のCEO李彦宏(Robin Li)は、「AI産業構造は逆ピラミッドになるべきだ」と提唱しました。これまででは基盤モデルやチップ(インフラ層)に価値と投資が集中していましたが、今後はその上で動くアプリケーション層こそが100倍の経済価値を生み出すべきだという主張です。この哲学は、今回発表されたアプリケーション群に色濃く反映されています。

5.1. GenFlow 3.0と「Glass Box」アプローチ

汎用AIエージェントプラットフォーム「GenFlow 3.0」は、ユーザー数2,000万人を擁する百度のキラーアプリの一つです。今回のアップデートでは、AIの思考プロセスをブラックボックス化せず、ユーザーが介入・修正できるようにする「Glass Box(透明な箱)」アプローチが採用されました。

- **RAGの深化:** 百度学術(Baidu Scholar)の6億8000万件の論文データや、百度文庫の14億件の専門文書、さらにはユーザーの百度網盤(クラウドストレージ)内のファイルに直接アクセスし、根拠に基づいた回答を生成します。
- **並列タスク処理:** 従来なら数時間かかるような「レポート作成」「スライド生成」「データ分析」といった複合タスクを、複数のエージェントが協調して並列処理し、数分レベルで完遂する能力を持っています。

5.2. デジタルヒューマンとEコマース革命

中国市場特有の強力なユースケースとして、「デジタルヒューマン」によるライブコマースの自動化が挙げられます。百度の技術は、今年の「独身の日(Double 11)」商戦において、83%の配信者が採用するという圧倒的なシェアを獲得しました。デジタルヒューマンが24時間休まず商品を宣伝・販売することで、流通総額(GMV)は前年比91%増を記録したと報告されています。この技術はブラジルや東南アジア市場へも展開されており、百度にとって即効性のある収益源となっています。

5.3. Apollo Go: 自動運転の収益化への道

自動運転タクシーサービス「Apollo Go(夢ト快跑)」は、累計乗車回数1,700万回を超え、完全無人運転も週25万回実施されています。さらに、アブダビでの商用運行許可を取得するなど、グローバル展開も加速しています。ERNIE 5.0の搭載により、乗客との自然な対話や、複雑な交通状況(例:警察官の手信号の認識など)への対応力が向上し、完全無人化によるコスト削減とサービス品質の向上が期待されます。李彦宏が「インテリジェンスはコストではなく生産性の源泉」と語る通り、Apollo GoはAIを物理世界のGDPに変換する最前線です。

6. 市場の反応と戦略的課題:なぜ株価は下落したのか?

ERNIE 5.0の発表という技術的なマイルストーンにもかかわらず、発表直後の香港市場で百度の株価は一時10%近く急落しました。この「Sell the news(事実で売る)」現象の背後には、投資家たちが抱く根深い懸念と、AIビジネスの構造的な課題が横たわっています。

6.1. 「差別化」の欠如とコモディティ化の恐怖

アナリストたちの最大の失望は、ERNIE 5.0がOpenAIやGoogleといった競合に対して「決定的な差別化要因(Standout Feature)」を提示できなかった点にあります。ベンチマークスコアで並んだとしても、ユーザー体験を一変させるような「iPhoneモーメント」的な驚きが欠けていました。

さらに、中国国内市場ではDeepSeek、Alibaba(Qwen)、ByteDance、Tencentなどが激しい「価格戦争」を繰り広げています。推論コストのダンピング競争は、AIモデル自体のコモディティ化(汎用品化)を招き、高い利益率を維持することを困難にしています。投資家は、技術的な優位性が収益に直結しにくいこの市場構造をリスク視しています。

6.2. 垂直統合のリスク:諸刃の剣

百度の「チップからアプリまで」という垂直統合戦略は、地政学的リスクに対する防壁としては優秀ですが、ビジネスリスクを一点に集中させる側面もあります。もし自社製チップKunlunの開発が遅延したり、性能目標を達成できなかった場合、その上で動くERNIEモデル、さらにはApollo GoやGenFlowといったアプリ層まで、すべての競争力が連鎖的に低下する恐れがあります。水平分業型のNVIDIAエコシステムに乗る他社と比較して、百度の戦略は「逃げ場がない」ハイリスクな賭けでもあります。

6.3. 収益化へのタイムラグとGPT-5.1の影

また、今回の発表はOpenAIによる「GPT-5.1 Instant / Thinking」のリリースと時期が重なりました。世界的な注目がOpenAIの新機能に集まる中、百度の発表が相対的に霞んでしまった側面も否めません。投資家は、百度がAIへの巨額投資を回収し、検索広告に依存した収益構造を転換できるのが「いつになるのか」という具体的なロードマップを求めています。今回の発表ではその収益化のスピード感が市場の期待値に届かなかったと言えます。

7. 結論と展望：Baiduの次なる一手

2025年の百度の発表は、同社がAI開発競争において「技術的キャッチアップ」のフェーズを終え、「独自エコシステムの構築と実利化」という新たなフェーズに突入したことを示しています。

重要な結論：

1. **ERNIE 5.0**のネイティブ・オムニモーダル化は、技術的には正解ですが、その真価はQianfanプラットフォーム上でサードパーティ開発者がどれだけ革新的なアプリを生み出せるかにかかっています。
2. **ERNIE-4.5-VL**のオープンソース化と「Thinking with Images」は、リソース制約のあるエンタープライズ市場を攻略するための強力な武器であり、エッジAIやオンプレミス需要においてBaiduが覇権を握る可能性があります。
3. インフラの自立は、米中分断が進む世界において百度の生命線ですが、CUDA互換性の確保やTianchiクラスターの安定稼働など、乗り越えるべき技術的ハードルは依然として高いままです。

百度にとっての2026年は、これらの技術的資産を、Apollo Goの黒字化や、デジタルヒューマンによるEコマース売上の拡大といった「具体的な数字」に変えられるかどうかの正念場となります。技術の先端性を示すだけでなく、それを泥臭くビジネスの現場に実装し、GDPに変換する能力こそが、百度が再び市場の信頼を取り戻すための唯一の道となるでしょう。

引用ソース一覧

- ¹
: eWeek - Baidu ERNIE 5.0 Release

- 2
: ECNS - Baidu World 2025 Announcement
- 3
: StockTitan - Baidu Unveils ERNIE 5.0
- 4
: PR Newswire - Baidu Unveils ERNIE 5.0 and Applications
- 5
: GuruFocus - ERNIE 5.0 Launch Triggers Sell-Off
- 6
: TechInAsia - Baidu Launches ERNIE 4.5
- 7
: MarkTechPost - ERNIE 4.5 VL 28B A3B Thinking
- 8
: AI News - Baidu ERNIE Benchmarks
- 9
: Baidu AI Studio - ERNIE 4.5 Model Details
- 10
: Times of India - New AI Processors
- 11
: HuggingFace - ERNIE 4.5 Thinking with Images
- 12
: ERNIE Technical Report - Architecture Overview
- 13
: The Rift - ERNIE 5.0 Architecture
- 14
: Implicator - Baidu's Technical Excellence vs Market
- 15
: VOI - Baidu Chip Roadmap & Parameter Count
- 16
: TechInAsia - Baidu Shares Fall
- 17
: Implicator - Why Technical Excellence No Longer Wins
- 17
: ModelScope - ERNIE 4.5 Model Config
- 18
: Implicator - GenFlow 3.0 Stats
- 19
: Skywork Blog - GenFlow 2.0/3.0 Features
- 20
: TradingView - Chips Analysis

- : TechPowerUp - M100/M300 Specs
21
- : Arxiv - System 2 Reasoning
22
- : kgpandya.com - Pricing Comparison
23
- : Xenospectrum - ERNIE 5.0 vs GPT-5 Pricing
24
- : Innovatopia - ERNIE 4.5 Launch Details
25
- : Baidu Blog - Thinking with Images details
11
- : HuggingFace - Thinking Process Technicals

引用文献

1. Baidu Unveils ERNIE 5.0, New AI Chips - eWeek, 11月 16, 2025にアクセス、
<https://www.eweek.com/news/baidu-ernie-5-0-release/>
2. Baidu unveils new Kunlun chips, 11月 16, 2025にアクセス、
<http://www.ecns.cn/m/news/sci-tech/2025-11-14/detail-ihewyecz9800878.shtml>
3. Baidu Unveils ERNIE 5.0, GenFlow Reaches 20M Users | BIDU Stock News, 11月 16, 2025にアクセス、
<https://www.stocktitan.net/news/BIDU/baidu-unveils-ernie-5-0-and-a-series-of-ai-applications-at-baidu-hq8a7kgdupq6.html>
4. Baidu Unveils ERNIE 5.0 and a Series of AI Applications at Baidu World 2025, Ramps Up Global Push - PR Newswire, 11月 16, 2025にアクセス、
<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>
5. BIDU: ERNIE 5.0 Launch Triggers Sharp Sell-Off in Baidu Shares ..., 11月 16, 2025にアクセス、
<https://www.tradingview.com/news/gurufocus:f14d68816094b:0-bidu-ernie-5-0-launch-triggers-sharp-sell-off-in-baidu-shares/>
6. Baidu launches Ernie-4.5 AI model to rival OpenAI, 11月 16, 2025にアクセス、
<https://www.techinasia.com/news/baidu-launches-ernie-4-5-ai-model-to-rival-openai>
7. Baidu Releases ERNIE-4.5-VL-28B-A3B-Thinking: An Open-Source and Compact Multimodal Reasoning Model Under the ERNIE-4.5 Family, 11月 16, 2025にアクセス、
<https://www.marktechpost.com/2025/11/11/baidu-releases-ernie-4-5-vl-28b-a3b-thinking-an-open-source-and-compact-multimodal-reasoning-model-under-the-ernie-4-5-family/>
8. Baidu ERNIE multimodal AI beats GPT and Gemini in benchmarks - AI News, 11月 16, 2025にアクセス、

- <https://www.artificialintelligence-news.com/news/baidu-ernie-multimodal-ai-gpt-and-gemini-benchmarks/>
9. Introducing ERNIE-4.5-VL-28B-A3B-Thinking: A Breakthrough in Multimodal AI, 11月 16, 2025にアクセス、<https://aistudio.baidu.com/modelsdetail/39280/intro>
 10. China's Baidu announces two AI processors, new version of its Ernie model, 11月 16, 2025にアクセス、
<https://timesofindia.indiatimes.com/technology/tech-news/chinas-baidu-announces-two-ai-processors-new-version-of-its-ernie-model/articleshow/125303266.cms>
 11. baidu/ERNIE-4.5-VL-28B-A3B-Thinking · Hugging Face, 11月 16, 2025にアクセス、
<https://huggingface.co/baidu/ERNIE-4.5-VL-28B-A3B-Thinking>
 12. ERNIE 4.5 Technical Report, 11月 16, 2025にアクセス、
https://yiyen.baidu.com/blog/publication/ERNIE_Technical_Report.pdf
 13. Baidu Unveils ERNIE 5.0, a 2.4-Trillion-Parameter Omni-Modal AI Model - The Rift, 11月 16, 2025にアクセス、
<https://www.therift.ai/news-feed/baidu-unveils-ernie-5-0-a-2-4-trillion-parameter-omni-modal-ai-model>
 14. Baidu's ERNIE 5.0 Proves Technical Excellence No Longer Wins China's AI War, 11月 16, 2025にアクセス、
<https://www.implicator.ai/baidus-ernie-5-0-proves-technical-excellence-no-longer-wins-chinas-ai-war/>
 15. Baidu Launches New AI Processor, Strengthens China's Position In Competition For Artificial Intelligence, 11月 16, 2025にアクセス、
<https://voi.id/en/technology/533439>
 16. Baidu shares fall 9.8% after launch of Ernie 5.0 AI model, 11月 16, 2025にアクセス、
<https://www.techinasia.com/news/baidu-shares-fall-9-8-after-launch-of-ernie-5-0-ai-model>
 17. ERNIE-4.5-VL-28B-A3B-PT - ModelScope, 11月 16, 2025にアクセス、
<https://modelscope.cn/models/PaddlePaddle/ERNIE-4.5-VL-28B-A3B-PT>
 18. Fellou: The Agentic Browser That Doesn't Just Search—It Acts. - Skywork.ai, 11月 16, 2025にアクセス、
<https://skywork.ai/blog/fellou-the-agentic-browser-that-doesnt-just-search-it-acts/>
 19. What Baidu's new chips reveal about China's plan to counter Nvidia in AI race, 11月 16, 2025にアクセス、
<https://www.tradingview.com/news/invezz:89eaa89e7094b:0-what-baidu-s-new-chips-reveal-about-china-s-plan-to-counter-nvidia-in-ai-race/>
 20. Baidu Announces M100 and M300 AI Processors, 11月 16, 2025にアクセス、
<https://www.techpowerup.com/342896/baidu-announces-m100-and-m300-ai-processors>
 21. MatryoshkaThinking: Recursive Test-Time Scaling Enables Efficient Reasoning - arXiv, 11月 16, 2025にアクセス、
<https://arxiv.org/pdf/2510.10293>
 22. Kunal G. Pandya, 11月 16, 2025にアクセス、
<https://kgpandya.com/index/>
 23. Baidu、GPT-5超えを謳う「ERNIE 5.0」発表。真のオムニモーダル ..., 11月 16, 2025にアクセス、
<https://xenospectrum.com/baidu-ernie-5-omnimodal-ai-beats-gpt5/>

24. 「Baidu ERNIE 5.0」登場 | 新時代マルチモーダルAIの最新 ..., 11月 16, 2025にアクセス、<https://innovatopia.jp/ai/ai-news/71772/>
25. ERNIE-4.5-VL-28B-A3B-Thinking: A Breakthrough in Multimodal AI ..., 11月 16, 2025にアクセス、<https://ernie.baidu.com/blog/posts/ernie-4.5-vl-28b-a3b-thinking/>