

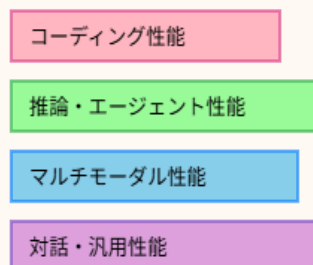
MiniMax-M2、Kimi K2Thinking、ERNIE-4.5-VL-28B-A3B-Thinking、GPT-5.1 など最新の生成 AI に関する情報

Felo AI

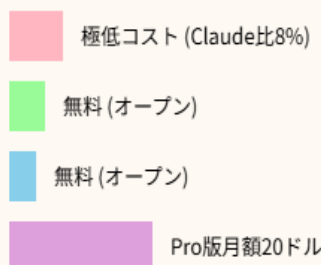
最新生成AI比較レポート (2025年10-11月)



主要性能比較



コスト比較



主要トレンド

- 中国発オープンモデルの台頭
- エージェント機能の本格化
- 性能対コストの最適化競争
- UX重視戦略の重要性

概要

2025年10月下旬から11月中旬にかけて、生成AI分野では技術革新の波が押し寄せ、主要プレイヤーから次世代モデルが相次いでリリースされた。特に注目されるのは、中国発の高性能オープンソース・オープンウェイトモデルの台頭と、市場をリードするOpenAIによるユーザー体験の深化という二つの大きな潮流である。

MiniMax社の「MiniMax M2」、Moonshot AI社の「Kimi K2 Thinking」、Baidu社の「ERNIE-4.5-VL-28B-A3B-Thinking」

は、いずれもエージェント機能、コーディング、マルチモーダル推論といった特定領域に特化した高い性能を誇り、オープンなライセンス形態で公開された。これらのモデルは、性能で欧米の最先端モデルに匹敵、あるいは一部で凌駕しつつ、圧倒的なコストパフォーマンスを実現しており、AI開発のエコシステムに大きな変革をもたらす可能性を秘めている。

一方、OpenAIは「GPT-5.1」をリリースし、GPT-5で指摘された応答の硬さを改善。「Instant」と「Thinking」の2モード体制や8種類の会話スタイル導入により、対話の自然さとパーソナライズを大幅に強化した。これは、技術的性能の追求だけでなく、ユーザーがAIとどのように関わるかという「体験の質」を重視する戦略へのシフトを示唆している。

本レポートでは、これら4つの最新AIモデルの技術仕様、性能ベンチマーク、アーキテクチャの特徴を詳細に分析し、比較検討を通じて2025年11月時点における生成AIの最先端動向と各モデルの戦略的立ち位置を明らかにする。

詳細レポート

1. 最新 AI モデルのリリース動向（2025年10月-11月）

2025年秋、生成AIの開発競争は新たな局面を迎えた。中国のAI企業が立て続けに高性能なオープンモデルを発表し、市場の勢力図を塗り替えようとしている。シンガポール拠点のMiniMaxは10月27日にエージェント機能に特化した「MiniMax M2」を[23](#)、北京のMoonshot AIは11月6日に驚異的な推論能力を持つ「Kimi K2 Thinking」を公開した[31](#)[32](#)。さらに、中国のテクノロジー大手Baiduは11月13日のイベントに先立ち、強力なマルチモーダル性能を持つ「ERNIE-4.5-VL-28B-A3B-Thinking」をオープンソース化した[41](#)。これに対し、市場のリーダーであるOpenAIは11月12日、ユーザー体験の向上に主眼を置いた「GPT-5.1」をリリースし、追従を許さない姿勢を見せている[52](#)。

2. 各モデルの詳細分析

MiniMax M2：コストパフォーマンスに優れたエージェント・コーディング特化モデル

シンガポールに拠点を置くAIユニコーン企業MiniMaxが2025年10月27日頃にオープンソース化した「MiniMax M2」は、特にエージェント機能とコーディング能力に焦点を当てて設計されたモデルである[5](#)[23](#)。同社は、既存のAIモデルが性能、価格、推論速度の「不可能な三角形」に直面していると指摘し、M2はこの3つの要素の最適なバランスを追求した結果生まれた[23](#)。

性能とアーキテクチャ MiniMax M2は、プログラミング、ツールの使用、ディープサーチといったエージェントにとって重要な能力において、GPT-5 (thinking)やClaude Sonnet 4.5などの最先端モデルに匹敵する性能を示す[23](#)。特に、Artificial AnalysisのIntelligenceベンチマークでは世界トップ5の性能を記録した[23](#)。

アーキテクチャには、フルアテンションとスライディングウィンドウアテンション（SWA）を交互に配置するGPT-OSSライクな構造や、各アテンションヘッドが独自の学習可能なRMSノルムを持つQKノルムが採用されている[23](#)。モデルサイ

ズは 230GB で、MIT ライセンスの下で Hugging Face にて公開されている [23](#)。

価格と提供形態 M2 の最大の魅力はそのコストパフォーマンスにある。API 価格は Claude 3.5 Sonnet の約 8% でありながら、推論速度は 2 倍高速であると報告されている [23](#)。このモデルは Hugging Face からダウンロードできるほか、Google Cloud の Vertex AI Model Garden でも管理 API として利用可能となっており、幅広い開発者がアクセスしやすい環境が整備されている [20](#) [23](#) [24](#)。

Kimi K2 Thinking : 驚異的な長期推論能力を持つオープンウェイトモデル

中国・北京の Moonshot AI が 2025 年 11 月 6 日に公開した「Kimi K2 Thinking」は、エージェントとしての自律的なタスク実行能力を極限まで高めたモデルである [31](#) [32](#)。最大の特徴は、人間の介入なしに最大 200~300 の連続したツール呼び出しを実行できる能力であり、複雑な問題に対して数百ステップにわたる一貫した推論が可能だ [32](#)。



性能とアーキテクチャ Kimi K2 Thinking は、1 兆パラメータ規模の Mixture of Experts (MoE) アーキテクチャを基盤とし、256K という広大なコンテキストウィンドウを備える [32](#)。これにより、深い推論と長期的な計画立案に優れた能力を発揮する [32](#)。主要なベンチマークにおいて、OpenAI の GPT-5 や Anthropic の Claude Sonnet 4.5 を上回るスコアを記録したと報告されている [31](#)。具体的には、専門家レベルの知識が問われる「Humanity's Last Exam」で 44.9%、ブラウジング能力を測る「BrowseComp」で 60.2% という高いスコアを達成した [32](#)。ただし、これらのスコアは自己申告であり、第三者による検証が待たれる [31](#)。

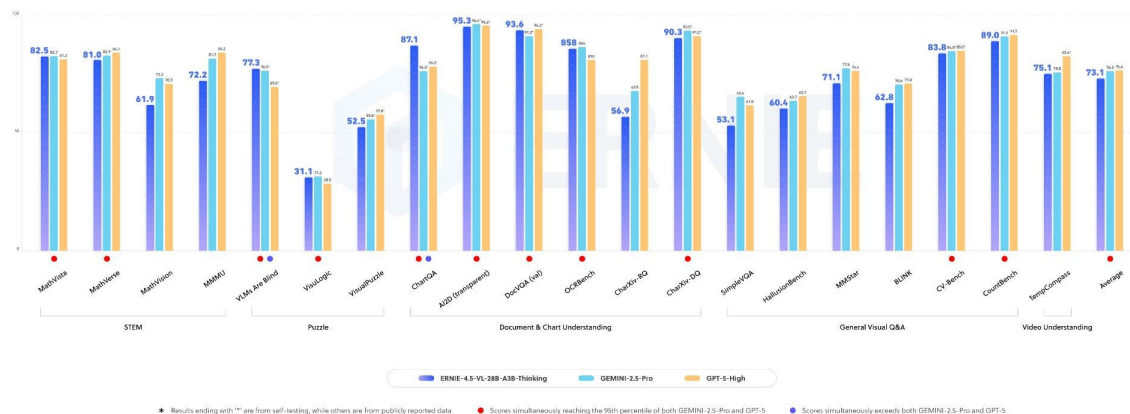
開発背景と提供形態 このモデルは、7 月に発表された「Kimi K2」を基盤としており、Hugging Face 上でオープンウェイト

モデルとして公開されている [28 31 32](#)。驚くべきことに、この高性能モデルの訓練コストはわずか **460 万ドル**とされており、米国の主要 AI 研究所が費やす数十億ドル規模の投資とは対照的である [31](#)。

ERNIE-4.5-VL-28B-A3B-Thinking : 視覚と言語を深く統合した軽量マルチモーダルモデル

中国のテクノロジー大手 Baidu は、2025 年 11 月 13 日に開催された年次イベント「Baidu World 2025」で次世代基盤モデル「ERNIE 5.0」を発表したが、それに先立ち、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」を公開した [41](#)。このモデルは、特に視覚情報の深い理解と推論能力に特化している。

🚀 Compact model, powerful multimodal reasoning ⚡🧠



性能とアーキテクチャ このモデルは、280 億パラメータの MoE アーキテクチャを採用しているが、推論時にアクティブになるのはわずか 30 億パラメータであり、単一の 80GB GPU でも動作可能な軽量設計となっている [40 41](#)。Apache 2.0 ライセンスで公開されており、商用利用も可能である [40](#)。

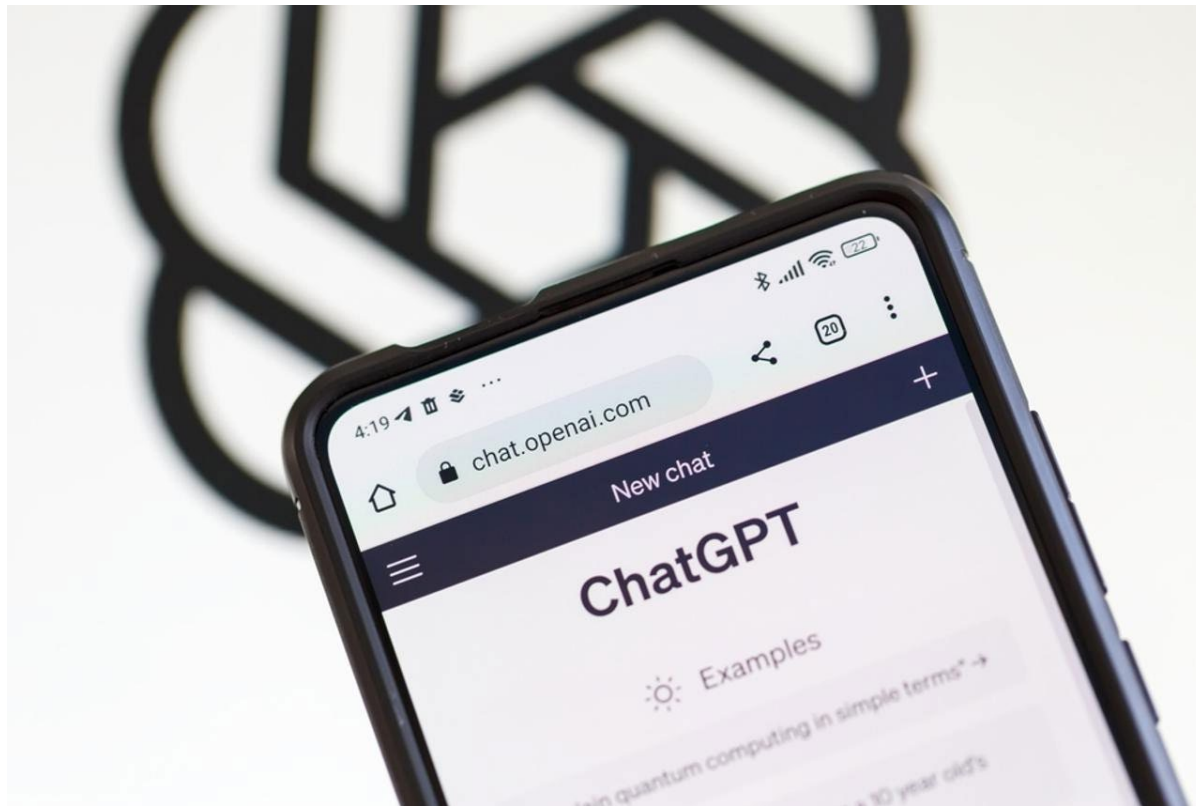
最大の特徴は、高度なマルチモーダル能力にある [40](#)。

- **Thinking with Images:** 人間のように画像の詳細を把握するため、自律的に画像をズームインするなどの操作を行う [40 42](#)。
- **高度な視覚推論:** 複雑なチャート分析や、回路図のような STEM 分野の画像からの問題解決能力に優れる [42](#)。
- **視覚的グラウンディング:** 画像内の特定のオブジェクトを正確に検出し、その座標を JSON 形式で出力するなど、指示への高い追従性を示す [42](#)。
- **ツール利用:** 内部知識で解決できない場合、画像検索ツールを自律的に呼び出して情報を補完する [42](#)。

これらの能力は、Baidu が提唱する「ネイティブ・オムニモーダル」（設計段階から全モーダルを統合的に扱う）という思想を体現しており、単なるマルチモーダル対応から一歩進んだ技術的到達点を示している [41](#)。

GPT-5.1：ユーザー体験とパーソナライズを追求する市場リーダー

OpenAI が 2025 年 11 月 12 日にリリースした「GPT-5.1」は、8 月に登場した GPT-5 の改良版であり、技術的な飛躍よりもユーザー体験の向上に重点を置いたアップデートである [4 52](#)。GPT-5 が持つ「賢いが冷たい」「よそよそしい」といったフィードバックに応え、「より賢く、より会話的」な AI を目指している [2 43 44](#)。



Instant モードと Thinking モード GPT-5.1 は、用途に応じて最適化された 2 つのモードを提供する [52](#)。

- **GPT-5.1 Instant:** 日常的な利用を想定した標準モデル。応答がより温かく知的になり、指示への追従性も向上した [50 52](#)。
- **GPT-5.1 Thinking:** 高度な推論に特化したモデル。複雑な課題にはじっくり時間をかけて取り組み、より粘り強く、理解しやすい回答を生成する [4 52](#)。

このモード分離と、質問の難易度に応じて思考時間を自動調整する「アダプティブ・リーズニング」機能により、ユーザーはタスクに応じた最適な応答を得やすくなった [54](#)。

パーソナライズ機能の強化 ユーザーが AI の個性を調整できる 8 つの会話スタイル（プロフェッショナル、率直、個性的など）が導入され、パーソナライズ機能が大幅に強化された [2 52](#)。これは、すべてのユーザーに画一的な体験を提供するのではなく、個々の好みに合わせた AI を実現するという OpenAI の明確な戦略転換を示している [52](#)。

性能と課題 法務 AI サービス「LegalOn」による検証では、契約書の修正文案を提案するタスクにおいて、GPT-5.1はGPT-5.0と比較して精度が65-67%向上し、処理速度も約30%高速化したと報告されている [8](#) [53](#)。一方で、リリース直後には「Thinkingモード」で応答が停止する不具合も一部のユーザーから報告されており、新モデルの安定性には課題も残る [4](#)。

3. 横断的性能比較と分析

特徴	MiniMax M2	Kimi K2 Thinking	ERNIE-4.5-VL-28B-A3B-Thinking	GPT-5.1
リリース日	2025年10月27日頃 23	2025年11月6日 32	2025年11月上旬 41	2025年11月12日 52
開発元	MiniMax (中国/シンガポール) 1 23	Moonshot AI (中国) 31	Baidu (中国) 41	OpenAI (米国) 52
ライセンス	オープンソース (MIT) 23	オープンウェイト 32	オープンソース (Apache 2.0) 40	プロプライエタリ 52
アーキテクチャ	GPT-OSS-like, SWA 23	1兆パラメータ MoE 32	28Bパラメータ MoE (3Bアクティブ) 40	不明 (Instant/Thinkingモード) 52
コンテキスト長	不明	256K 32	131,072 40	不明
主な強み	コーディング, エージェント, コストパフォーマンス 23	長期推論, エージェント, ツール連携 32	マルチモーダル推論, 視覚理解, 軽量 40	対話の自然さ, パーソナライズ, 汎用性 52
価格 (API)	非常に安価 (Claude Sonnetの8%) 23	無料 (オープンウェイト) 32	無料 (オープンソース) 40	不明 (Pro版は月額20ドル) 6

得意分野の比較

- エージェント機能とコーディング: **MiniMax M2** と **Kimi K2 Thinking** がこの分野をリードしている。M2 は開発者ワークフロー全体を支援することを目指し [23](#)、K2 Thinking は数百ステップに及ぶ自律的なツール連携能力を誇る [32](#)。これらは、AI が単なる応答生成ツールから、複雑なタスクを自律的に実行する「エージェント」へと進化しているトレンドを象徴している。
- マルチモーダル能力: **ERNIE-4.5-VL** が際立っている。「Thinking with Images」機能のように、AI が能動的に視覚情報を分析するアプローチは、従来のマルチモーダル AI から一線を画す [42](#)。Baidu が発表した上位モデル「ERNIE 5.0」は、文書やチャートの読解において GPT-5 を上回ると主張しており、ビジネス応用での強みを示唆している [41](#)。
- 対話体験と汎用性: **GPT-5.1** がこの領域で強みを発揮する。Instant/Thinking モードの導入や会話スタイルの拡充は、技術仕様の競争だけでなく、ユーザーが AI といかに快適で生産的な関係を築けるかという UX（ユーザー体験）を重視する姿勢の表れである [2 52](#)。

4. 技術的トレンドと市場への示唆

今回のリリースラッシュから、以下の 4 つの主要な技術トレンドが読み取れる。

1. エージェント能力の本格化: AI が自律的に計画を立て、ツールを使いこなし、複数ステップのタスクを完遂する能力が、次世代モデルの重要な評価軸となった。Kimi K2 Thinking や MiniMax M2 の登場は、この流れを決定づけた。
2. 中国発オープンモデルの猛追: 中国企業が、性能で西側の最先端プロプライエタリモデルに匹敵するモデルを、オープンソースまたはオープンウェイトで次々と公開している。これは、開発者コミュニティにおける影響力を急速に高め、AI エコシステム全体の勢力図を塗り替える可能性がある。
3. 性能対コストの最適化競争: MiniMax M2 の価格戦略や ERNIE-4.5-VL の軽量 MoE アーキテクチャは、最高の性能を追求するだけでなく、それをいかに効率的かつ低コストで提供するかが市場での競争優位性に直結することを示している [23 40](#)。
4. プロプライエタリモデルの UX 重視戦略: OpenAI は、純粋なベンチマークスコア競争から一歩引き、対話の質やパーソナライズといった「体験価値」の向上に舵を切った。これは、広範なユーザーベースを維持・拡大するための成熟した市場戦略と言える。

5. 結論

2025 年 11 月時点の生成 AI 市場は、特定タスクに特化した高性能オープンモデルと、汎用性とユーザー体験を追求するプロプライエタリモデルという二つの潮流が明確になった。MiniMax M2、Kimi K2 Thinking、ERNIE-4.5-VL に代表される中国勢のオープン戦略は、AI 技術の民主化を加速させ、世界中の開発者に新たな選択肢と可能性を提供している。彼らの技術力とコスト競争力は、もはや西側巨大テック企業の独走を許さないレベルに達している。

一方で、OpenAI の GPT-5.1 は、AI が単なるツールではなく、個々のユーザーに寄り添う「パートナー」となる未来を見据えている。技術の最先端を走り続けると同時に、その技術をいかに多くの人々にとって身近で使いやすいものにするかという課題に対する一つの答えを示した。

今後の AI 開発は、この「特化型オープンモデル」と「汎用・UX 重視型プロプライエタリモデル」が相互に影響を与えながら進化していくと予想される。開発者は自身のニーズに応じて最適なモデルを選択する自由を享受し、ユーザーはより洗練され、パーソナライズされた AI 体験を享受できるようになるだろう。このダイナミックな競争環境こそが、AI 技術全体の発展を加速させる原動力となることは間違いない。

1. [Talkie を産み出した企業 MiniMax: 中国発の AI ユニコーン - note](#)
2. [GPT-5 Pro で評価してみた。\(Gemini もいるよ\)](#)
3. [プログラミング雑記 2025 年 11 月 14 日 #AI](#)
4. [【衝撃】ChatGPT 5.1 ちゃんが思考停止!? 2025 年 11 月最新版 ...](#)
5. [MiniMax が M2 プレビューモデルをリリース、クラウドを買収 ...](#)
6. [chatgpt5.1 の特徴と主要機能を徹底解説 | Instant・Thinking ...](#)
7. [論文や技術メモの一覧 \(随時更新\) | わたしのべんきょうノート](#)
8. [11 月 13 日リリースの OpenAI 最新モデル「GPT-5.1」を検証 ...](#)
9. [Journal of Machine Learning Research Papers: Volume 24 ...](#)
10. [ChatGPT でできること一覧! Plus との違いやビジネス活用例も ...](#)
11. [\[論文解説\] ChatGPT の利用実態: 7 億人のユーザーは何に使っ ...](#)
12. [AI による設計書作成能力比較 GitHub Copilot GPT5.1 シリーズ ...](#)
13. [生成 AI の種類は? 2025 年に知っておきたい 7 タイプの特徴と ...](#)
14. [【コンサルがオワコンに?】ChatGPT が最強の検索機能「Deep ...](#)
15. [Claude Sonnet のわずか 8% の価格で 2 倍高速な AI モデル ...](#)
16. [MiniMax M2 は非常に優れているようです: r/LocalLaMA](#)
17. [名古屋ではたらく社長の IT ニュースポッドキャスト](#)
18. [その MiniMax の動画生成 AI が、すごいらしい。 - note](#)
19. [MiniMax-M2 の情報 \(OpenRouter discord より\) : r/LocalLaMA](#)
20. [Weekly Google Cloud アップデート情報 - 2025/11/11 - SoftBank](#)
21. [Talkie を産み出した企業 MiniMax: 中国発の AI ユニコーン - note](#)
22. [MLX に Minimax-M2 のサポートが追加されました - Reddit](#)
23. [Claude Sonnet のわずか 8% の価格で 2 倍高速な AI モデル「MiniMax M2」がオープンソース化&有料 API が期間限定無料で利用可能に - GIGAZINE](#)
24. [Weekly Google Cloud アップデート情報 - 2025/11/11 | クラウドテクノロジーブログ | ソフトバンク](#)

25. [中国発の新 AI 「Kimi K2 Thinking」、米国製を超える性能と ...](#)
26. [いま最強の無料 AI かも。1 兆パラメーターの怪物エージェント ...](#)
27. [中国・Moonshot AI、推論モデル『Kimi K2 Thinking』発表 ...](#)
28. [オープンな AI モデル「Kimi K2」登場「GPT-4.1」など上回る性能 ...](#)
29. [学術&研究](#)
30. [「gemini 3.0」の Yahoo!リアルタイム検索 - X \(旧 Twitter\) を ...](#)
31. [中国発の新 AI 「Kimi K2 Thinking」、米国製を超える性能と“無料”の衝撃 - ZDNET Japan](#)
32. [ASCII.jp : いま最強の無料 AI かも。1 兆パラメーターの怪物エージェントモデル「Kimi K2 Thinking」中国から登場](#)
33. [baidu/ERNIE-4.5-VL-28B-A3B-Thinking - Hugging Face](#)
34. [GPT-5 超えを謳う「ERNIE 5.0」発表。真のオムニモーダル AI ...](#)
35. [ERNIE-4.5-VL-28B-A3B-Thinking: A Breakthrough in ...](#)
36. [Baidu Inc.](#)
37. [ERNIE-4.5-VL-28B-A3B is a hidden gem that can decently ...](#)
38. [ERNIE-4.5-VL-28B-A3B-PT - ModelScope](#)
39. [ERNIE-4.5-VL-28B-A3B-Thinking: Leading Multimodal AI ...](#)
40. [baidu/ERNIE-4.5-VL-28B-A3B-Thinking · Hugging Face](#)
41. [Baidu、GPT-5 超えを謳う「ERNIE 5.0」発表。真のオムニモーダル AI が拓く未来 | XenoSpectrum](#)
42. [ERNIE-4.5-VL-28B-A3B-Thinking: A Breakthrough in Multimodal AI | ERNIE Blog](#)
43. [OpenAI、ChatGPT のアプデ「GPT-5.1」をリリース。より会話が ...](#)
44. [ChatGPT 5.1 がリリース！OpenAI API では？最新情報まとめ](#)
45. [ChatGPT に温かみのある返答を取り戻す「GPT-5.1」が登場 ...](#)
46. [11 月 13 日リリースの OpenAI 最新モデル「GPT-5.1」を検証 ...](#)
47. [ChatGPT 新モデル「GPT-5.1」公開 - よりフレンドリーな応答](#)
48. [OpenAI、「GPT-5.1」リリース「よりスマートに - ITmedia AI+](#)
49. [GPT-5.1 の到来：会話力とパーソナライズが変えるビジネス AI ...](#)
50. [OpenAI が「GPT-5.1」発表、対話能力が向上「より親しみやすく](#)
51. [GPT-5.1 リリース。標準的な「Instant」と推論モデル ...](#)
52. [OpenAI、ChatGPT のアプデ「GPT-5.1」をリリース。より会話がスムーズに | ギズモード・ジャパン](#)
53. [11 月 13 日リリースの OpenAI 最新モデル「GPT-5.1」を検証「LegalOn: World Leading Legal AI」に実装予定 | 株式会社 LegalOn Technologies のプレスリリース](#)
54. [ChatGPT 新モデル「GPT-5.1」公開 - よりフレンドリーな応答、指示への遵守も強化 - CNET Japan](#)