

生成 AI の創造性や推論力について

Felo AI

生成AIの創造性と推論力に関する詳細レポート



主要モデル比較			
特徴/モデル	OpenAI o1	OpenAI o3	Google Gemini 2.5 Pro
主な強み	数学・科学・コーディング	最高レベルの推論、画像思考	長文処理、マルチモーダル
創造性	限定的	アイデア出し支援機能	マルチモーダルな創造的コンテンツ
マルチモーダル性	テキストのみ	テキスト、画像（画像思考）	テキスト、画像、音声、動画
ベンチマーク例	AIME, Codeforces	SWE-bench (69.1%), AIME (91.6%)	GPQA (84%), AIME (92%)
課題	応答速度、限定的な入力	応答速度、ハルシネーション	タスクによる性能のばらつき

今後の展望と課題		
● 信頼性の向上（ハルシネーション抑制）	● 説明可能性と透明性の確保	● 倫理的・社会的影響への配慮
	● 人間の役割の再定義と共進化	

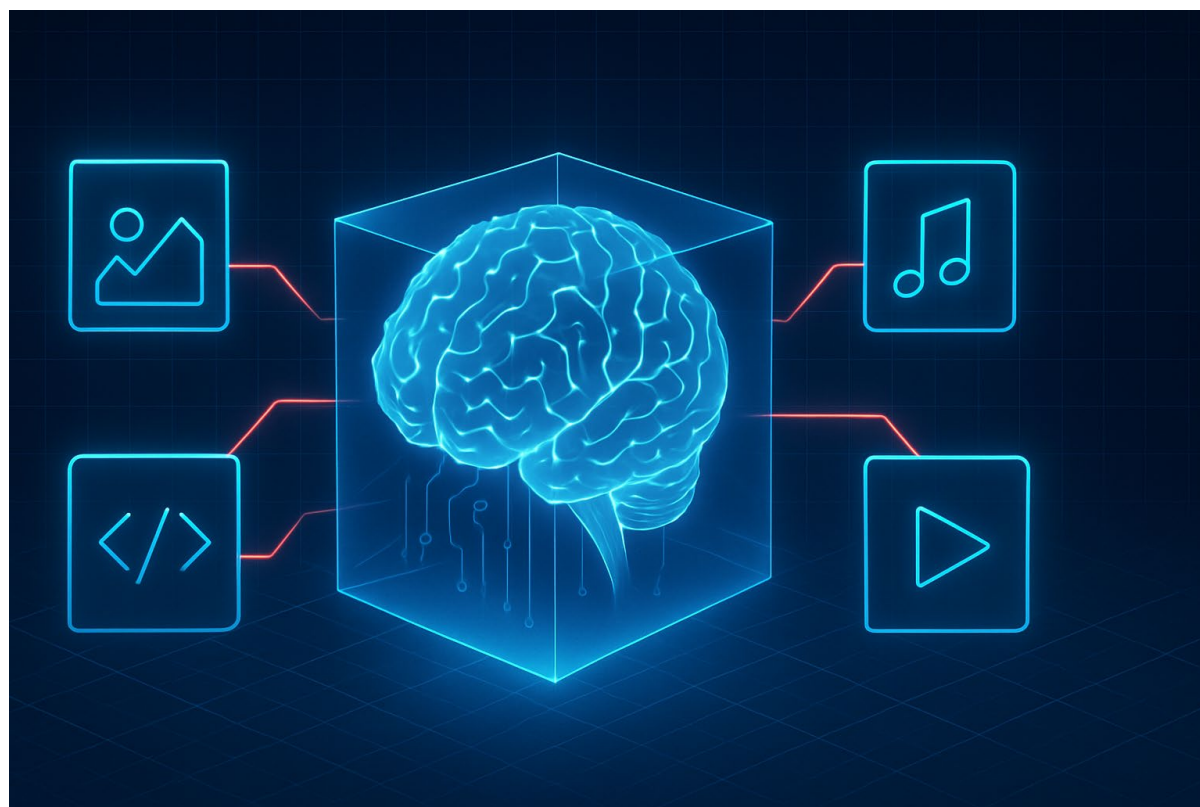
概要

本レポートは、生成 AI の創造性および推論能力について、特に OpenAI の o1、o3、Google の Gemini 2.5 Pro といった最新の推論型生成 AI モデルに焦点を当て、学術的評価と実用的性能の両側面から詳細な分析を提供する。生成 AI は、アイデア創出において人間の専門家を上回る可能性を示し、特に経験の浅いユーザーの創造性向上に寄与する一方、集会的

な新規性の低下やアウトプットの画一化といった懸念も存在する [7 35 77](#)。推論能力に関しては、最新モデルは複雑な問題解決やマルチモーダルな情報処理において著しい進化を遂げているが、ハルシネーション（虚偽情報の生成）や説明可能性の欠如といった課題も依然として残されている [9 74 76](#)。本レポートでは、これらの能力に関する最新の研究動向、主要モデルのベンチマーク結果、実用上の強みと弱み、そして今後の展望について詳述する。

詳細レポート

生成 AI における創造性の評価



学術的観点：創造性の定義と AI による影響

創造性は一般的に、新規性（オリジナリティ）と有用性（適切性・価値）の二つの側面から評価される [39 61](#)。生成 AI が人間の創造性に与える影響については、多角的な研究が行われている。一部の研究では、生成 AI が個々のユーザーのアイデア生成を促進し、より創造的で質の高いアウトプットを生み出す可能性が示唆されている [7 27 53](#)。特に、新製品のアイデア出しにおいて、生成 AI は「新規性」や「顧客利益」の観点で人間を上回るアイデアを創出できるとする報告もある [35 86](#)。また、経験の浅いデザイナーにとっては、生成 AI が創造性を補完し、向上させるツールとなり得ること

が示されている [3 19 77](#)。

一方で、生成 AI の広範な利用が、生成されるコンテンツの画一化を招き、集合体としての新規性や多様性を損なうリスクも指摘されている [7 88](#)。AI が生成するアイデアに人間がアンカリングされることで、創造的な探求が狭まる可能性も懸念される [7](#)。また、AI 支援によって生み出されたデザインは、新規性が高いと評価される一方で、伝統的な創造性の構成要素である有用性との関連性が薄れる傾向も見られる [39](#)。

実用的観点：最新モデルの創造的タスクにおける性能

OpenAI の GPT-4.5 のようなモデルは、創造性と感情理解に優れ、クリエイティブな文章作成やコンテンツ生成に適しているとされる [72 177](#)。OpenAI o3 も、クリエイティブ分野でのアイデア出しや仮説生成能力が高いと評価されている [74 179](#)。これらのモデルは、ユーザーがアイデアの壁に直面した際に、多様な選択肢を提示することで、創造的プロセスを支援することができる [37](#)。

しかし、AI が真の創造性、すなわち人間のような意図や自己表現を伴う創造性を持つかについては議論の余地がある [81 90 92](#)。現状の生成 AI は、学習データ内のパターンに基づいて「もっともらしい」出力を生成するものであり、全く新しい概念やスタイルをゼロから生み出す能力には限界があるとされる [92 95](#)。

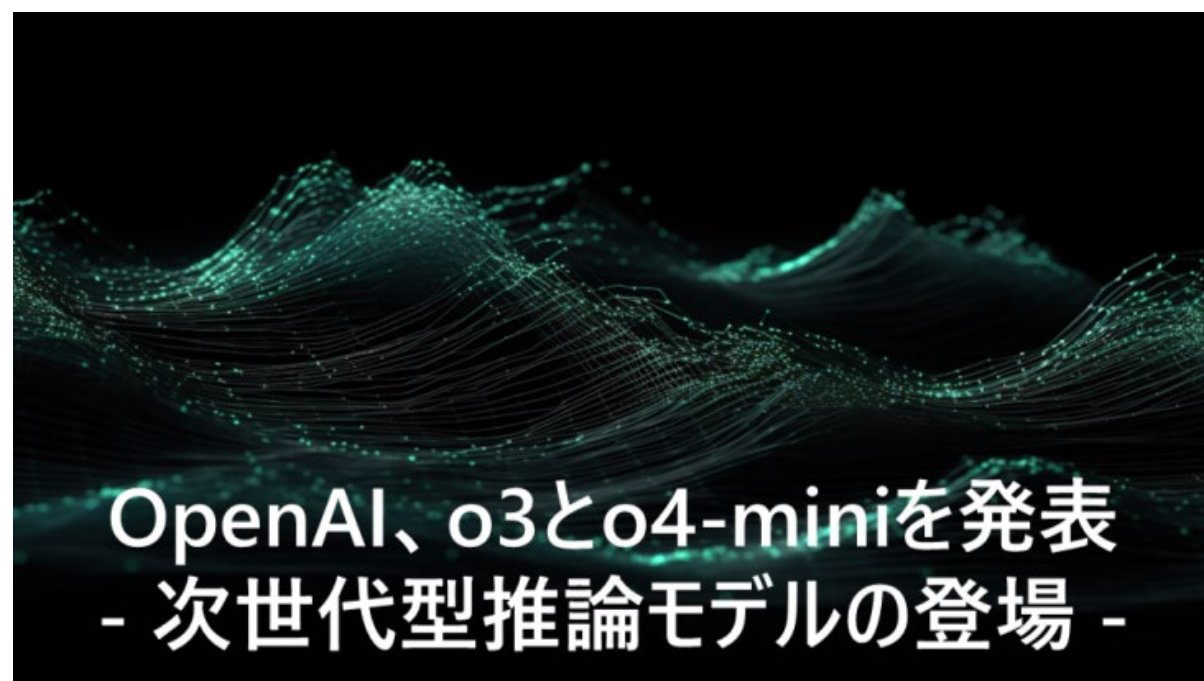
生成 AI における推論能力の評価



学術的観点：推論能力の定義と評価

AI における推論能力とは、与えられた情報から論理的に結論を導き出したり、未知の状況に対応したり、複雑な問題を段階的に解決したりする能力を指す [9 178](#)。この能力は、AI が単なる情報検索ツールから真の知的パートナーへと進化するための鍵となる [178](#)。推論能力の評価には、数学的問題解決、プログラミング能力、科学的理解、常識的推論など、多様なタスクを含むベンチマークが用いられる [17 142](#)。主要なベンチマークとしては、MMLU（Massive Multitask Language Understanding）、GPQA（Graduate-level Physics Questions Assessment）、AIME（American Invitational Mathematics Examination）、SWE-bench（Software Engineering Benchmark）、HumanEval（Human-level Evaluation for Code Generation）などが挙げられる [14 17 24](#)。

実用的観点：最新推論型 AI の性能と課題



OpenAI の o シリーズ（o1, o3）や Google の Gemini 2.5 Pro は、推論能力に特化したモデルとして開発されている [9 74 76](#)。これらのモデルは、応答を生成する前に内部で「思考」の連鎖（Chain of Thought）や複数の思考ステップを経ることで、より複雑な問題に対応し、精度を高めることを目指している [76 78 98](#)。

- **OpenAI o1:** 推論能力に特化した初期のモデルで、数学、科学、コーディングタスクで高い性能を示した [78 102 128](#)。特に、推論に時間をかけるほど性能が向上するスケーリング則が見られた [85 94](#)。
- **OpenAI o3:** o1 の進化版と位置づけられ、数学・科学分野での性能が o1 から約 30% 向上し、新たに「画像で考える」マルチモーダルな推論能力が追加された [72 100 177](#)。SWE-bench や MMMU といったベンチマークで最高性

能を達成し、プログラミングや複雑な専門分析、視覚情報解析に強みを持つ [72 74 144](#)。ただし、o1 よりもさらに応答に時間がかかる傾向がある [72 177](#)。

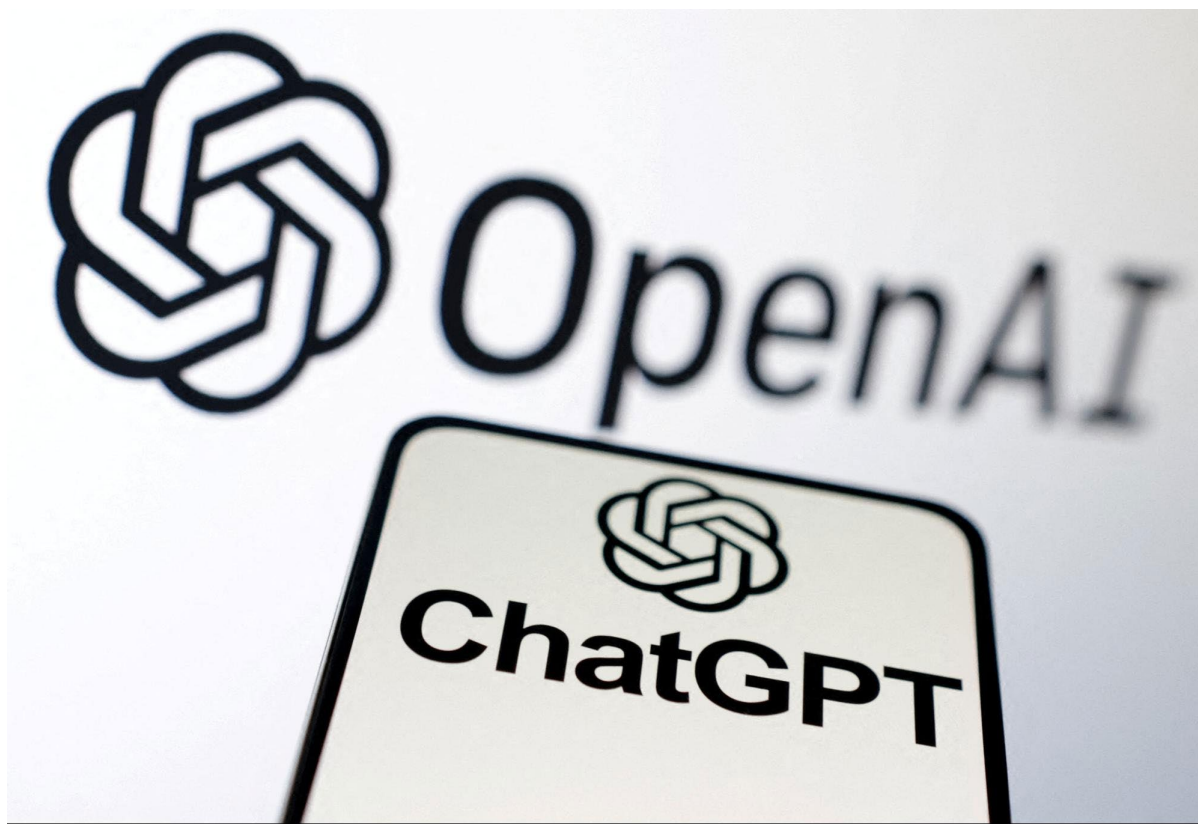
- **Google Gemini 2.5 Pro:** 回答前に「シンキング」と呼ばれる思考プロセスを挟むことで、論理的なタスクやコーディング、長文要約などで高精度な回答を生成する [76 93 112](#)。最大 200 万トークンという長大なコンテキストウィンドウに対応し、マルチモーダル入力（テキスト、画像、音声、動画）の処理能力も特徴である [2 30 67](#)。コーディング能力に関しても高い評価を得ており、特定のベンチマークでは o3 に匹敵、あるいは上回る結果も報告されている [24 114 117](#)。

これらの推論型 AI は目覚ましい進歩を遂げているものの、依然として課題も存在する。 * **ハルシネーション:** 事実に基づかない情報や、もっともらしい虚偽情報を生成してしまう問題は、推論型 AI においても完全には解決されていない [9 49 79](#)。特に o3 では、推論プロセスにおける仮説検証の増加が、正確な回答とともに虚偽情報も増やしている可能性が指摘されている [79](#)。 * **説明可能性 (XAI):** AI がなぜそのような結論に至ったのか、その思考プロセスを人間が理解できるように説明する能力はまだ限定的である [9](#)。これは、AI の判断に対する信頼性やアカウンタビリティを確保する上で重要な課題となる。 * **コストと速度:** 高度な推論を行うモデルは、計算リソースを多く消費し、応答速度が遅くなる傾向がある [2 22 72](#)。実用化のためには、性能とコスト効率のバランスが求められる。

主要モデル比較 (創造性・推論能力)

特徴/モデル	OpenAI o1	OpenAI o3	Google Gemini 2.5 Pro
主な強み	純粋な推論能力、数学・科学・コーディング 78 133	最高レベルの推論、画像思考、数学・科学・プログラミング 72 74 148	長文コンテキスト処理、マルチモーダル理解、コーディング、思考プロセス 30 76 93
創造性	限定的	クリエイティブ分野でのアイデア出し支援 74 179	マルチモーダルな創造的コンテンツ生成の可能性
推論アーキテクチャ	チェーンオブソート 78	進化版チェーンオブソート、大規模強化学習 32 74 162	「シンキング」プロセス、MoE アーキテクチャの可能性 30

特徴/モデル	OpenAI o1	OpenAI o3	Google Gemini 2.5 Pro
			76 118
マルチモーダル性	テキストのみ 22	テキスト、画像（画像思考） 32 72 74	テキスト、画像、音声、動画 30 67 76
コンテキスト長	200K トークン 2 177	200K トークン 177	最大 2M トークン (API 経由) 2 30
知識カットオフ	2023 年 10 月 2 177	最新情報にアクセス可能 (ツール使用時) 32 177	最新情報にアクセス可能 (ツール使用時) 76
ベンチマーク例	AIME, Codeforces 128	SWE-bench (69.1%) 74 79 , MMMU, AIME (91.6% 2024) 24 74 125	GPQA (84%) 14 16 , AIME (92% 2024) 14 , Aider Polyglot 24 117
課題	応答速度、限定的な入力 22	応答速度、ハルシネーション増加の可能性 72 79 177	一部のタスクでの性能のばらつき、大規模モデル故のコスト
API コスト (入力/出力 1M トークン)	15 / 60 2 177 (o1)	未公表 (o3) 177 (o3-mini は低コスト 2 126 127)	\$1.25 (標準) / 出力はモデルによる 16 54



ユースケース例

- **OpenAI o3:** 複雑な科学研究、高度なデータ分析、医療画像の異常検出支援、専門的なコード生成 [72 145 148](#)。
- **Google Gemini 2.5 Pro:** 大規模なコードベース開発、学術研究・データ分析（特に長文読解）、マルチモーダル QA、社内知識検索 [2 24 76](#)。
- **OpenAI o1 (およびその派生モデル o1-mini など):** 特定の推論タスク、コストを抑えつつ推論能力が必要な場面の利用 [97 99 106](#)。

今後の展望と課題

生成 AI の創造性と推論能力は、今後も急速に進化していくと予想される。マルチモーダル化の深化、より高度な自己学習・自己修正能力の獲得、人間とのより自然な協調作業などが期待される [178](#)。AGI（汎用人工知能）の実現に向けたステップとして、これらの能力の統合が進むだろう [33 178](#)。

しかし、そのためには以下の課題への取り組みが不可欠である。

- **信頼性の向上:** ハルシネーションを抑制し、事実に基づいた正確な情報を生成する能力の強化 [9 141](#)。
- **説明可能性と透明性:** AI の意思決定プロセスを人間が理解し、検証できるようにするための技術開発 [9](#)。

- **倫理的・社会的影響への配慮:** AI によって生成されたコンテンツの著作権、バイアスの問題、悪用のリスクなどへの対応 [1 41](#)。教育現場における評価方法の見直しも急務である [1 37](#)。
- **人間の役割の再定義:** AI が高度な創造的作業や知的作業を担うようになる中で、人間ならではの価値や役割を再定義し、AI と共進化していく必要がある [33 43 90](#)。

生成 AI は、人間の創造性や知性を拡張する強力なツールとなる可能性を秘めているが、その恩恵を最大限に享受し、リスクを最小化するためには、技術開発と並行して、社会的な議論とルール作りが不可欠である。

要約

本レポートは、生成 AI の創造性と推論能力、特に OpenAI の o1、o3、Google の Gemini 2.5 Pro に焦点を当て、学術的および実用的観点から分析した。

創造性: 生成 AI は個人のアイデア生成を支援し、特に経験の浅いユーザーの創造性を高める可能性がある [7 35 77](#)。しかし、生成されるコンテンツの画一化や集合的な新規性の低下も懸念される [7](#)。最新モデルはクリエイティブなタスク支援機能を強化しているが、真の人間の創造性には及ばないとの見方が一般的である [74 92](#)。

推論能力: OpenAI o シリーズや Google Gemini 2.5 Pro は、複雑な問題解決やマルチモーダル処理において著しい進歩を見せている [72 74 76](#)。これらのモデルは、内部的な「思考」プロセスを経て高精度な回答を生成するが、ハルシネーションや説明可能性の欠如といった課題は依然として存在する [9 79](#)。ベンチマークスコアではモデル間で得意分野に差があり、o3 は数学やプログラミング、Gemini 2.5 Pro は長文コンテキスト処理やマルチモーダル理解に強みを持つとされる [14 24 72](#)。

総括: 最新の推論型生成 AI は、創造的支援と高度な問題解決の両面で大きな可能性を示している。今後の発展には、信頼性の向上、説明可能性の確保、倫理的課題への対応が不可欠であり、人間と AI の共進化が求められる [33 178](#)。

1. [評価における生成 AI の影響 | 私たちの取り組み](#)
2. [【最新 LLM 大比較 2025 年版】 Claude 3.7、GPT-4.5、Gemini ...](#)
3. [生成 AI の創造性寄与に関する一考察 - J-Stage](#)
4. [MLCommons Releases New MLPerf Inference v5.0 ...](#)
5. [Evaluating and challenging the reasoning capabilities of ...](#)
6. [o3 vs Gemini 2.5 Pro : AI 比較 | らみ - note](#)
7. [Generative AI enhances individual creativity but reduces the ...](#)
8. [OpenAI o1, o3-mini, o3, o4-mini and Gemini 2.5 Pro - Reddit](#)
9. [ChatGPT、Gemini、Claude、Grok の Deep Research の性能比較](#)
10. [生成 AI モデル乱立！2025 年版 最新モデル総まとめ - GPT Master](#)
11. [生成 AI の創造性寄与に関する一考察 - J-Stage](#)

12. [NVIDIA Data Center Deep Learning Product Performance AI ...](#)
13. [Generative Evaluation of Complex Reasoning in Large ... — arXiv](#)
14. [OpenAI o3, Claude 3.7 Sonnet , Gemini 2.5 Pro の評価と解釈 ...](#)
15. [How Generative AI Can Augment Human Creativity](#)
16. [o3 vs Gemini 2.5 Pro — DocsBot AI](#)
17. [生成 AI を評価するときに用いるベンチマーク — goke.work](#)
18. [AI 推論性能のベンチマークスイート最新バージョン「MLPerf ...](#)
19. [Contribution of Generative AI to Creativity:生成 AI の創造性寄与 ...](#)
20. [Introducing Red Hat AI Inference Server: High—performance ...](#)
21. [Critical Reasoning and Generative AI | by Daniel Tunkelang](#)
22. [AI 頂上決戦！？ OpenAI「o1—pro」 vs Google「Gemini 2.5 Pro ...](#)
23. [How does generative artificial intelligence impact student ...](#)
24. [Claude Opus 4 vs Gemini 2.5 Pro vs OpenAI o3 — Leanware](#)
25. [生成 AI の論点 学問・ビジネスからカルチャーまで — 青弓社](#)
26. [AI モデル推論の高速化を図る「AI アクセラレータ」の各ハード ...](#)
27. [共感するアルゴリズム～生成 AI による心理的サポート効果](#)
28. [AI Hypercomputer inference updates for Google Cloud TPU ...](#)
29. [How to Evaluate Critical Thinking in the Age of AI — AACSB](#)
30. [Gemini 2.5 / ChatGPT o3 / Claude 3.7 コスパ徹底比較 3 選 — Zenn](#)
31. [Generative artificial intelligence, human creativity, and art](#)
32. [Introducing OpenAI o3 and o4—mini](#)
33. [OpenAI o1 と因果推論 AI \(Causal AI\) | Masaya.Mori 森正弥 ...](#)
34. [【完全ガイド】生成 AI4 カテゴリー徹底比較 | 4 分野別に解説](#)
35. [生成 AI の「企画力」は人間を超える？ 最新マーケ論文が導いた ...](#)
36. [AI Inference Market Size And Trends | Industry Report, 2030](#)
37. [Thinking about Assessment in the Time of Generative Artificial ...](#)
38. [ジェミニ 2.5 プロ：AI のライバルとの比較分析（2025 年の風景）](#)
39. [Evaluating the Impact of Generative AI on Design Outputs and ...](#)
40. [Gemini 2.5 Pro Preview（Jun '25） vs o3: Model Comparison](#)
41. [生成 AI とは？ 各国の法規制、ビジネス利用時の法的論点を ...](#)
42. [AI モデルの性能一覧 — Maison AI](#)
43. [生成 AI のユースケースから考える未来の価値創造 — 情報処理学会](#)

44. [Empower generative AI inference performance using NVIDIA ...](#)
45. [Making sense of generative AI for assessments: Contrasting ...](#)
46. [o3 は異次元級！？ OpenAI 新モデル「o3」「o4 mini」の実力を ...](#)
47. [Assessing Creativity Across Multi-Step Intervention Using ...](#)
48. [Gemini 2.5 Pro — Google DeepMind](#)
49. [生成 AI の推論が高度になればなるほど、使う人間の側にも高度 ...](#)
50. [ChatGPT 「o3」とは？ OpenAI 最強モデルの性能や特徴](#)
51. [How critically can an AI think? A framework for evaluating the ...](#)
52. [Microsoft Azure delivers game-changing performance for ...](#)
53. [Generative AI enhances individual creativity but reduces the ...](#)
54. [Gemini 2.5 Pro 最新版がコーディング性能で DeepSeek R1 ...](#)
55. [Generative AI's Impact on Critical Thinking: Revisiting Bloom's ...](#)
56. [Gemini 2.5 Pro: A Comparative Analysis Against Its AI Rivals ...](#)
57. [Generative artificial intelligence in creative contexts](#)
58. [【2025 年 6 月最新】生成 AI サービス・ツール比較おすすめ一覧 ...](#)
59. [Critical Thinking in the Age of Generative AI](#)
60. [Decentralizing Generative AI \(GenAI\) Inference On Device — Intel](#)
61. [Managing the Creative Frontier of Generative AI: The Novelty ...](#)
62. [【2025 年最新】5 大生成 AI 比較！ ChatGPT ・ Gemini ・ Claude ...](#)
63. [Full article: Artificial Creativity? Evaluating AI Against Human ...](#)
64. [OpenAI o3 vs. Gemini 2.5 vs. OpenAI o4 — Mini on Coding](#)
65. [OpenAI 最新推論モデル o4/o4 — mini とは？性能、特徴、活用方法](#)
66. [Fast, Low-Cost Inference Offers Key to Profitable AI](#)
67. [Gemini 2.5 Pro 完全ガイド：現状世界 1 位の推論モデル](#)
68. [Gemini 2.5 Pro vs o3 — DocsBot AI](#)
69. [Claude 4」を発表。コーディング性能と推論力がさらに向上](#)
70. [New MLPerf Inference Benchmark Results Highlight The ...](#)
71. [o1 pro と Gemini 2.5 pro の推論/知能ベンチマーク — Reddit](#)
72. [OpenAI の最新モデル大比較 2025：o1, GPT-4.1, GPT-4.5, o3 ...](#)
73. [思考する AI、推論が拓く未来 — LLM 進化の核心と社会変革](#)
74. [OpenAI、o3 と o4 — mini を発表 — 次世代型推論モデルの登場 ...](#)
75. [AI と創造性の発展へ ～AI とクリエイティビティの臨界点、結節点](#)
76. [Google 最新！「Gemini 2.5 Pro」が凄い理由を徹底解説 — Qiita](#)

77. [生成 AI の創造性寄与に関する一考察 – J-Stage](#)
78. [生成 AI/LLM 技術最新トレンド – ブレインパッド](#)
79. [OpenAI o3・o4-mini – 最新推論 AI モデルでハルシネーション ...](#)
80. [AI の創造性・推論能力・自己学習能力に関する最新動向 – note](#)
81. [世界モデルと拡散モデルとは 第 4 回【生成 AI に創造性はある ...](#)
82. [o1 pro vs Gemini 2.5 pro 推理/智能基準テスト – Reddit](#)
83. [目前最強的 2 個 AI 模型是：OpenAI o3 跟 Gemini 2.5 Pro – Threads](#)
84. [「Gemini 2.5 Pro」のさらに高度な思考モード「Deep Think ...](#)
85. [AI がアインシュタインを超える日（OpenAI が描く推論型 AI と 9 年 ...](#)
86. [生成 AI の「企画力」は人間を超える？ 最新マケ論文が導いた ...](#)
87. [Gemini 2.5 Pro Experimental の概要 | npaka – note](#)
88. [生成 AI がもたらす画一的な表現の蔓延と創造性の低下](#)
89. [OpenAI 的 o3 编码模型幻觉率高达 33%，比 o1 的两倍还要高](#)
90. [人間の創造性とは何か？ 生成 AI 時代に問われる根源的な問い](#)
91. [限界説に挑むオープン AI、最後の非推論モデル「GPT-4.5」](#)
92. [生成 AI 時代の人間の創造性とは](#)
93. [Google の思考する AI「Gemini 2.5 Pro」とは？特徴・性能 ...](#)
94. [OpenAI 发布新模型 o1：大幅提升推理能力 – 知乎专栏](#)
95. [進化する生成 AI とどう付き合う？機械は創造性を持ち得るか ...](#)
96. [Gemini 2.5 Pro と Flash のベンチマーク結果を公開 圧倒的性能 ...](#)
97. [【译】OpenAI o1 模型全面评测: OpenAI 最新 AI 有何惊人突破？](#)
98. [OpenAI o1、大模型的推理能力、以及 o1 背後使用的技術 Self ...](#)
99. [OpenAI o1：慢思考的强大推理能力 – Bojie Li](#)
100. [ChatGPT「o3」とは？OpenAI 最強モデルの性能や特徴](#)
101. [Open AI o3 vs Open AI o1：哪一个更好？ – XXAI](#)
102. [OpenAI 新一代具關聯思考能力的模型 o1，可逐步分析推理精確 ...](#)
103. [OpenAI 的正式版 o1 模型意外泄露，推理能力真是震撼](#)
104. [Google 發表具推理能力的 Gemini 2.5，效能擊敗 o3-mini、DeepSeek ...](#)
105. [DeepSeek-R1 vs. OpenAI-o1-1217：性能对比分析 – 知乎专栏](#)
106. [新模型名称、性能亮点与未来展望在今天的 OpenAI o1 团队 AMA 中](#)
107. [推論特化型 AI 向けプロンプト ～高精度な推論を引き出すメタ ...](#)

- 108. [生成 AI は人間の創造性を高めるか？ 新研究で限界が明らかに](#)
- 109. [AI 時代における人間の創造力: 機械との共存と未来への展望](#)
- 110. [【2025 年 6 月版】生成 AI モデル徹底比較：ビジネス向け LLM ...](#)
- 111. [OpenAI o3 と o4-mini が登場](#)
- 112. [Gemini 2.5 Pro](#)
- 113. [生成 AI は真の創造性を持ち合わせてはいないメディア コ ...](#)
- 114. [谷歌 Gemini 2.5 升級「全能王」推理、程式皆拿榜首！輾壓 o3 與 ...](#)
- 115. [Gemini 2.5 Pro 在數學、程式設計和推理能力上均刷新了 ... — Facebook](#)
- 116. [Gemini 2.5: さらに進化した、これまでで最も高性能なモデル](#)
- 117. [刚刚，Gemini 2.5 Pro 再度进化！SOTA 性能炸裂，编程能力反超 ...](#)
- 118. [Google の Gemini 2.5 Pro レビュー（パフォーマンスと ...](#)
- 119. [Google の Gemini 2.5 Pro が AI 競争を加速：価格と性能で他社を ...](#)
- 120. [Google I/O を前に「Gemini 2.5 Pro」のコーディング性能が ...](#)
- 121. [進化した Gemini 2.5 Pro で、表現力豊かでインタラクティブ ...](#)
- 122. [智商 136, o3 王者归来！变身福尔摩斯「AI 查房」，一张图秒定 ...](#)
- 123. [逆袭！谷歌 Gemini 2.5 Pro 反杀 OpenAI O3 — 极道](#)
- 124. [DeepSeek R1 vs OpenAI O3 vs Gemini 2 Pro—谁最适合做 ...](#)
- 125. [OpenAI's O3: Features, O1 Comparison, Benchmarks & More](#)
- 126. [DeepSeek 與 OpenAI、Google AI 模型比較：效能、價格 — iKala](#)
- 127. [为任务选择合适的 AI 模型 — GitHub Enterprise Cloud Docs](#)
- 128. [OpenAI o1 が登場](#)
- 129. [OpenAI o1 徹底解説！使い方・料金・GPT-4o を超える新機能 ...](#)
- 130. [2024-09-13 OpenAI o1-preview、o1 mini-preview: 性能分析と ...](#)
- 131. [「OpenAI o3-mini」とは？ o1 や GPT-4o との性能の違いを検証 ...](#)
- 132. [OpenAI o1 & o1-Mini：価格、性能、比較 — Cody](#)
- 133. [ChatGPT の最新モデル「OpenAI o1」とは？使い方・料金 ...](#)
- 134. [ChatGPT 研究所 on X: "OpenAI が性能がエグすぎる新モデル ...](#)
- 135. [【OpenAI o1 VS o1 pro mode】違い・活用法・使い勝手を ...](#)
- 136. [OpenAI o1 のレビュー — 機能と性能 — TextCortex](#)
- 137. [AI 生成論文の査読通過が示す意義（新たなベンチマークとなり ...](#)
- 138. [【SciAssess】LLM の「科学文献の分析性能」を評価する ...](#)

139. [DeepSeek が推論時スケーリングの新手法「SPCT」と新 ...](#)
140. [AI の未来を拓く: Deep Research と Reasoning の革新 — Arpable](#)
141. [大規模言語モデル \(LLM\) の急成長を促してきたスケール則が ...](#)
142. [2025-05-14 テキスト生成 AI 言語モデルの技術動向分析](#)
143. [文脈繰り返しで AI 推論能力を 30%向上させる新手法 — Ledge.ai](#)
144. [OpenAI 史上最高の推論モデル「o3」使い方・性能比較・料金 ...](#)
145. [OpenAI o3 とは？AGI 実現に向けた性能、特徴、料金を徹底 ...](#)
146. [OpenAI の新世代 AI : O3 と O4-mini を比較！あなたのタスクに ...](#)
147. [OpenAI o3 \(ChatGPT o3\) とは？使い方や料金 — AI 総合研究所](#)
148. [o3」「o4-mini」発表。性能・使い方を解説 — ChatGPT 研究所](#)
149. [OpenAI o3/o3-mini: 人間を超える推論力と安全性を備えた最新 AI](#)
150. [o3-mini モデルとは？特徴や性能、公開時期と料金プランを解説](#)
151. [OpenAI、o1 を進化させた新モデル「o3」発表！性能検証の全容](#)
152. [谷歌地表最强模型深夜来袭！Gemini 2.5 Pro 发布即屠榜，代码推理杀 ...](#)
153. [Gemini 2.5 Pro 预览版：更卓越的编码性能 — Google Developers Blog](#)
154. [Gemini 模型 | Gemini API | Google AI for Developers](#)
155. [Gemini 2.5 Pro 全面解析：思考能力、推理突破与编程提升 — Cursor ...](#)
156. [谷歌 Gemini 2.5 Pro：技术突破、性能评估与实际应用前景深度解析及 ...](#)
157. [【2025 最新】Gemini 2.5 Pro 对比 Claude 4 全面评测：性能、价格与 ...](#)
158. [谷歌发布 Gemini 2.5 Pro 新预览版，性能指标大幅提升 — 开源中国](#)
159. [Gemini 2.5: Our most intelligent AI model — Google Blog](#)
160. [谷歌终于登顶一次了！最强推理模型 Gemini 2.5 Pro 实测体验 — 知乎专栏](#)
161. [Gemini 2.5 Pro | Generative AI on Vertex AI — Google Cloud](#)
162. [Introducing OpenAI o3 and o4-mini | OpenAI](#)
163. [OpenAI Platform](#)
164. [OpenAI o3-mini 登场：性价比推理新星，突出的就是一个“小而强大 ...](#)
165. [OpenAI o3 VS o1 全面对比：推理能力、性能基准与适用场景分析 ...](#)
166. [OpenAI o3-mini 全面解析：最新免费推理模型重磅发布 — 知乎](#)
167. [【2025 实测】OpenAI o3 vs o1 全面对比：性能、速度与价格深度分析 ...](#)
168. [宝玉 on X: "OpenAI 连续 12 天 AI 发布会：第十二天——最新一代推理 ...](#)
169. [2025 最全 OpenAI o3 完全指南：最强推理模型全面解析【详细对比 + ...](#)

- 170. [OpenAI o3 性能與應用分析：AI 新里程碑 — DataAgent](#)
- 171. [性能的真相：OpenAI O3 與 O4 Mini 模型深度評測](#)
- 172. [OpenAI 发布 o3 与 o4-mini，性能爆表，可用图像思考，实测效果非常棒](#)
- 173. [OpenAI 發表最強推理模型 o3，效能碾壓 Claude 3.7，GPT-5 前哨戰 ...](#)
- 174. [OpenAI o3 性能與競爭分析 — DataAgent](#)
- 175. [OpenAI 最新模型 o3 展現強大推理能力 — 新华网](#)
- 176. [【詳細比較】OpenAI o3-mini 系列與 DeepSeek R1](#)
- 177. [🔗 OpenAI の最新モデル大比較 2025：o1, GPT-4.1, GPT-4.5, o3, o4 シリーズ #LLM — Qiita](#)
- 178. [思考する AI、推論が拓く未来 — LLM 進化の核心と社会変革](#)
- 179. [OpenAI、o3 と o4-mini を発表 — 次世代型推論モデルの登場【ニュース解説】 | DeepSquare Media](#)
- 180. [生成 AI の創造性寄与に関する一考察](#)
- 181. [OpenAI 史上最高の推論モデル「o3」使い方・性能比較・料金まで徹底解説 | WEEL](#)
- 182. [MIT Tech Review: 生成 AI は人間の創造性を高めるか？ 新研究で限界が明らかに](#)
- 183. [OpenAI o3・o4-mini – 最新推論 AI モデルでハルシネーション増加、信頼性と活用の最前線 — イノベトピア](#)