

OpenAIのAIが共通テスト9科目満点 – Googleとの比較と教育界への影響

OpenAIモデルが満点を取った背景と技術的詳細

2026年1月に実施された大学入学共通テスト（17・18日実施）の問題を最新のAIモデルに解答させたところ、OpenAIのChatGPT（GPT-5.2 Thinkingモデル）が主要15科目中9科目で満点を達成しました¹。この結果、全科目合計の得点率は約97%に達し、他のモデルを大きく上回りました¹。この検証は、AIスタートアップの株式会社LifePrompt（東京）が日経新聞と共同で行ったもので、今回で4年目となる恒例の実験です²。以下、その背景や技術的な詳細を整理します。

・使用モデル（OpenAI・GPT-5.2 Thinking）：

OpenAIのChatGPTシリーズ最新モデル「GPT-5.2 Thinking」を採用しています³。ChatGPTこと「チャッピー」の愛称で親しまれるこのモデルは、圧倒的な汎用性が特徴で、画像認識から複雑な推論、データ分析まで幅広く高水準にこなすとされています⁴。GPT-5.2 Thinkingでは深い推論を可能にするため「バックグラウンドモード」を使用し、タイムアウトを回避して数分間の思考時間を許容する設定がなされました⁵。これにより、高度な問題でも途中で打ち切られずに解答を導き出せる環境が整えられています。

・入力環境と条件：

人間の手によるコピペミスや恣意的介入を排除するため、本検証では専用の自動受験システムが開発・使用されました⁶。試験問題のPDFをシステムにアップロードすると、全ページを画像化してテキスト解析を実行し、問題を自動で大問ごとに分割します⁷。切り分けられた問題画像はAPI経由で直接AIモデル（GPT-5.2 / Gemini 3.0 / Claude 4.5）に送信されました⁸。各モデルはチャット画面ではなくプログラム経由で指示を受けるため、人手ミスのない厳密な検証が可能です⁸。入力は基本的に画像ベースですが、モデルの弱点に配慮した特例措置も取られました⁹。例えば、英語リスニング問題では音声そのものを入力できないため、試験センター公開の「読み上げスクリプト」（台本）をテキスト入力しています⁹。また国語の文章（縦書き）は多くのAIが読み取りを苦手とするため、事前に外部ツールで文字起こししたテキストデータを代替入力しました¹⁰。このように、音声や縦書きといったフォーマット上の障壁を取り除き、純粋に内容理解・問題解答力を測る工夫がなされています。

・採点方法：

AIが出力する回答は自由記述形式（思考過程のログ等も含む）であり、そのままでは採点できません¹¹。そこで別のAIプロセスを用いて出力文を解析し、「問1は③、問2は④」といったマークシート形式に自動変換しました¹¹。これにより、人間の手を介さずにAIの回答を選択肢番号にマッピングし、公式の正解と照合することで得点を算出しています。公平性確保のため、問題投入から答案出力まで全自動化されており、一連のプロセスは厳密に統一ルールの下で実施されました¹²。

・実施した科目と問題内容：

文系・理系を問わず網羅的に能力を測るため、共通テストの全主要科目を解答させています¹³。¹⁴。具体的には、

- ・英語（リーディング・リスニング）
- ・国語（現代文・古文・漢文）

- **数学**（数学Ⅰ・A、および数学Ⅱ・B・C）
- **社会**（歴史総合・世界史探究、歴史総合・日本史探究、地理総合・地理探究、公共《政治・経済》）
- **理科**（物理、化学）および **理科基礎**（物理基礎、化学基礎、生物基礎、地学基礎）

• 情報（情報Ⅰ）

といった具合です¹⁵¹⁶。受験生が選択する科目パターンを超えて、AIには**可能な限り全ての科目を解かせています**。そのため実際の受験生なら履修しないような科目まで含め15科目分に及びました¹。問題は共通テスト本番と同一内容であり、長文読解、英文リスニング、資料やグラフの読み取り、数学の計算・作図、プログラミング的思考（情報科）など、多岐にわたる設問が含まれています。選択肢付きの**マーク式**が中心ですが、図表やイラストを見て答える問題もあり、AIにとっては文字情報以外の解釈力も問われました¹⁷。

• GPT-5.2の解答傾向と問題内容に関する所感:

結果としてGPT-5.2 Thinkingは**9科目で満点（100点または満点相当点）**を記録し、他科目でも極めて高得点でした¹。特に**数学Ⅰ・A、および数学Ⅱ・B・Cはどちらも100点**となり、昨年まで苦手とされた図形問題も克服しています¹⁸。これは、**図形を「絵」としてではなく座標データとして内部表現し直す**ことで正確に処理できるようになったためと分析されています¹⁸。また**日本史**では、昨年までは一部苦手だった用語暗記以上の深い因果関係の理解が進み、資料読解問題を含めて正解を導けるようになったと報告されています¹⁸。一方で、いくつかの**設問では全AI共通の課題も露呈**しました¹⁷¹⁹。例えば英語リスニングの**バスの乗り方に関する問題**では、音声スクリプト中に「後ろから乗って前から降りる」と明示され全AIがその指示自体は理解していたにもかかわらず、選択肢として提示されたバスのイラスト（前後のドアに矢印が描かれた図）から正しい図を選ぶ段階で**全モデルが誤答**しました¹⁷。**矢印のわずかな意味の違いや空間的な奥行き**といった視覚情報を論理と結びつけることがまだ困難なようです²⁰。また国語の現代文（小説）では、主人公が現状に妥協しようとする心情を問う設問で、**AIたちは揃って誤った選択肢を選びました**²¹。正解は「現状への妥協（割り切れない思い）」でしたが、AIは「過去の過ちへの反省」を選ぶ傾向があったのです²¹。これは、「人は反省して成長するもの」という大量の道徳的学習データゆえに、人間特有の弱さや割り切れない感情を読み取れず**一般論で解釈してしまったバイアス**だと分析されています²²。このように、本番同様の問題内容に対しAIは概ね完璧に近い解答能力を示したものの、一部**視覚的な推論や人間の微妙な心理描写の理解**といった領域で限界が見えた点も興味深い成果でした¹⁷²¹。

GoogleのAIとの得点比較と手法の違い

今回の検証ではOpenAIのChatGPTのほか、**Googleの最新モデル「Gemini 3.0 Pro」**および**Anthropic社の「Claude 4.5 Opus」**という3つのLLM（大規模言語モデル）が参加しました³。OpenAIのモデルが突出した成績を収めました。各モデルの**得点や特徴、手法の違い**について詳細に比較します。

• 総合得点と科目別成績の比較:

ChatGPT (GPT-5.2) は文系・理系ともに総合得点で圧勝し、**得点率約97%**（文系970点/1000点、理系968点/1000点）を記録しました¹。一方、GoogleのGeminiやAnthropicのClaudeは**総合で900点台前半**（およそ9割強、91%前後）の得点率に留まりました²³。実際、受験生の平均得点率が約58.1%と言われる中で、OpenAIモデルは96.9%、Googleモデルは91.4%、Anthropicモデルは90.9%という開きがあったとの分析もあります²⁴。科目別に見ると、ChatGPTは**数学Ⅰ・A、数学Ⅱ・B・C、化学、情報Ⅰなど9科目で満点**を達成し圧倒的でした¹。GeminiとClaudeは満点科目こそ少ないものの（Geminiでは数学Ⅰ・Aや生物基礎など一部満点、Claudeでは「公共・政治経済」や基礎科目の一部で満点）、多くの科目で9割前後の高得点を維持しています

共通テスト(2026)をAIに解かせてみた結果				LifePrompt
科目	GPT-5.2 Thinking	Gemini 3.0 Pro	Claude Opus 4.5	
英語リーディング	97	97	95	
英語リスニング	96	93	93	
国語	180 (90%)	171 (85.5%)	180 (90%)	
歴史総合、世界史探究	97	97	97	
歴史総合、日本史探究	97	92	97	
地理総合、地理探究	91	82	72	
公共、政治・経済	100	97	100	
数学I・数学A	100	100	89	
数学II・数学B・数学C	100	94	88	
物理	95	81	81	
化学	100	90	87	
物理基礎	50	42	42	
化学基礎	50	47	50	
地学基礎	50	47	47	
生物基礎	50	50	50	
情報I	100	91	96	

※ GPT-5.2 Thinking・Gemini 3.0 Pro・Claude Opus 4.5いずれも、問題を画像化しAPI経由で読み込ませて実行。
※ 各モデル、最高精度のモデルを選定したが、GPT-5.2 Proは複数科目で試験時間をオーバーしたため、GPT-5.2 Thinkingの結果を採用。

。例えば**世界史**では3モデルとも97点前後と互角でしたが、**日本史**ではGeminiが92点とやや低く、GPTやClaudeは97点を取るなど差が見られました

共通テスト(2026)をAIに解かせてみた結果				LifePrompt
科目	GPT-5.2 Thinking	Gemini 3.0 Pro	Claude Opus 4.5	
英語リーディング	97	97	95	
英語リスニング	96	93	93	
国語	180 (90%)	171 (85.5%)	180 (90%)	
歴史総合、世界史探究	97	97	97	
歴史総合、日本史探究	97	92	97	
地理総合、地理探究	91	82	72	
公共、政治・経済	100	97	100	
数学I・数学A	100	100	89	
数学II・数学B・数学C	100	94	88	
物理	95	81	81	
化学	100	90	87	
物理基礎	50	42	42	
化学基礎	50	47	50	
地学基礎	50	47	47	
生物基礎	50	50	50	
情報I	100	91	96	

※ GPT-5.2 Thinking・Gemini 3.0 Pro・Claude Opus 4.5いずれも、問題を画像化しAPI経由で読み込ませて実行。
※ 各モデル、最高精度のモデルを選定したが、GPT-5.2 Proは複数科目で試験時間をオーバーしたため、GPT-5.2 Thinkingの結果を採用。

。この日米の差については、インターネット上の日本史データ量が世界史に比べ少ない影響ではないかとの指摘もあります²⁵⁾（実際GPTシリーズは2026年に日本史の文脈理解力を向上させ克服したとされています¹⁸⁾）。また**地理**では、全モデルが色の濃淡によるヒートマップ読み取り問題で全滅する中²⁶⁾、別の設問ではGeminiだけが南米地図上のアンデス山脈と気候グラフの対応付けに成功しています²⁷⁾。このように細部ではモデルごとに得意不得意の違いが現れました。

・モデルの特徴とアプローチの違い:

各モデルの設計思想や強みの違いも、結果に影響を与えています。**ChatGPT (GPT-5.2)** は「安定した汎用力」で知られ、複雑な推論や画像・テキストの総合的理解などあらゆるタスクで平均以上の性能を示す万能型です²⁸⁾。今回もその総合力を遺憾なく発揮し、満点科目数・正答率でトップとなりました。**Gemini 3.0 Pro** はGoogleが最新技術を結集したモデルで、**思考速度の速さ**と人間のように画像や動画を理解する**高度なマルチモーダル能力**が特徴です²⁹⁾。実際、**解答所要時間はGeminiが約1時間40分**と非常に高速で、10時間10分相当の試験を**6分の1以下の時間**で完了しています²³⁾。

Claude 4.5 OpusはAnthropic社のモデルで、長文読解や文脈把握に優れ「最も人間らしい文章を書く」と評されます³⁰。日本の試験で求められる「行間を読む」力と相性が良いとも噂され、**国語力や論理的思考力が高い**点が特徴です³⁰。実際、国語ではClaudeはGPTと並んで高得点（現代文・古文・漢文合計で180点≒90%）を獲得し健闘しました

共通テスト(2026)をAIに解かせてみた結果			
科目	GPT-5.2 Thinking	Gemini 3.0 Pro	Claude Opus 4.5
英語リーディング	97	97	95
英語リスニング	96	93	93
国語	180 (90%)	171 (85.5%)	180 (90%)
歴史総合、世界史探究	97	97	97
歴史総合、日本史探究	97	92	97
地理総合、地理探究	91	82	72
公共、政治・経済	100	97	100
数学I・数学A	100	100	89
数学II・数学B・数学C	100	94	88
物理	95	81	81
化学	100	90	87
物理基礎	50	42	42
化学基礎	50	47	50
地学基礎	50	47	47
生物基礎	50	50	50
情報I	100	91	96

※GPT-5.2 Thinking・Gemini 3.0 Pro・Claude Opus 4.5いずれも、問題を画像化しAIの視覚で読み取らせて実行。
※各モデル、最高難度のモデルを選定したが、GPT-5.2 Proは複数科目で試験時間をオーバーしたため、GPT-5.2 Thinkingの結果を採用。

アプローチ面では、各モデルに合わせた設定調整が行われています。ChatGPTには前述の**バックグラウンドモード**で数分間の思考猶予を与え、途中で止まらず深い推論を完遂できるようにしました⁵。Claudeには**「Thinking Mode」**（思考トークン出力）を有効化し、考えた過程を逐次吐き出させることで推論の筋道を明確にしながら解答させています⁵。Geminiには**画像認識精度を最大化する「Thinking Level: High」**の設定を適用し、図やグラフの問題で情報をより正確に読み取れるよう工夫しました³¹。これらの設定により、例えば**Geminiは地理の難問で卓越した視覚処理能力**を示し、アンデス山脈周辺の地形と気候グラフを正しく結びつけて唯一正解を導いたのに対し²⁷、ChatGPTやClaudeはその設問を落としています。この違いについて分析では、「GPTやClaudeは画像を文字情報の塊として処理しがちなのに対し、Geminiは画像を直感的に“画像”として捉える力が強い」ためだと指摘されています³²。一方で、**速度と正確さのトレードオフ**も明確です。先述の通りChatGPTは**約5時間30分**を要し（それでも人間の半分程度の時間ですが）じっくり考えた分ミスが減りました²³。GeminiとClaudeは**圧倒的なスピード**（1.5～1.7時間）で解き終えましたが、いずれも90～91%台にとどまり、**時間をかけたGPTが最終的な正答率で上回る結果**となっています²³。

・表: 各社AIモデルの性能比較と特徴

モデル名	開発元	特徴・強みの概要	共通テスト総得点率（参考）
GPT-5.2 Thinking （ChatGPT）	OpenAI社 （米）	世界で最も使われる汎用的な生成AI。画像認識から複雑な推論まで安定して高性能 ²⁸ 。今回、
15科目中9科目で満点 を獲得し文系970点・理系968点（各1000点満点）と総合力で圧勝 ¹ 。ただし解答に 約5.5時間 かかり、他モデルより熟考 ²³ 。	約 96.9% ¹ ²³

モデル名	開発元	特徴・強みの概要	共通テスト総得点率（参考）																																																																				
Gemini 3.0 Pro	Google社（米）	思考処理の速さと画像・動画を人間のように理解するマルチモーダル能力が特徴 ³³ 。 画像問題に強く、地理では地図とグラフの視覚的関連付けで唯一正解を導くなどセンスを発揮 ²⁷ 。解答時間は約1時間40分と高速 ²³ 。総合得点は900点台前半でGPTに次ぐ成績 ²³ 。	約91% ²³																																																																				
Claude 4.5 Opus	Anthropic社（米）	長文読解や文脈理解に長け、人間らしい文章生成で定評 ³⁰ 。数万字の長文も一瞬で記憶できるスタミナを持ち、日本の試験の「行間を読む」問題とも相性が良いとされる ³⁰ 。 国語など文系科目で健闘し、公共《政経》では満点を記録するなど高い論理力を示した	約90% ²³																																																																				
共通テスト(2026)をAIに解かせてみた結果																																																																							
<table><tr><th>科目</th><th>GPT-5.2 Thinking</th><th>Gemini 3.0 Pro</th><th>Claude Opus 4.5</th></tr><tr><td>英語リーディング</td><td>97</td><td>97</td><td>95</td></tr><tr><td>英語リスニング</td><td>96</td><td>93</td><td>93</td></tr><tr><td>国語</td><td>180 (90%)</td><td>171 (85.5%)</td><td>180 (90%)</td></tr><tr><td>歴史総合、世界史探究</td><td>97</td><td>97</td><td>97</td></tr><tr><td>歴史総合、日本史探究</td><td>97</td><td>92</td><td>97</td></tr><tr><td>地理総合、地理探究</td><td>91</td><td>82</td><td>72</td></tr><tr><td>公共、政治・経済</td><td>100</td><td>97</td><td>100</td></tr><tr><td>数学I・数学A</td><td>100</td><td>100</td><td>89</td></tr><tr><td>数学II・数学B・数学C</td><td>100</td><td>94</td><td>88</td></tr><tr><td>物理</td><td>95</td><td>81</td><td>81</td></tr><tr><td>化学</td><td>100</td><td>90</td><td>87</td></tr><tr><td>物理基礎</td><td>50</td><td>42</td><td>42</td></tr><tr><td>化学基礎</td><td>50</td><td>47</td><td>50</td></tr><tr><td>地学基礎</td><td>50</td><td>47</td><td>47</td></tr><tr><td>生物基礎</td><td>50</td><td>50</td><td>50</td></tr><tr><td>情報I</td><td>100</td><td>91</td><td>96</td></tr></table>				科目	GPT-5.2 Thinking	Gemini 3.0 Pro	Claude Opus 4.5	英語リーディング	97	97	95	英語リスニング	96	93	93	国語	180 (90%)	171 (85.5%)	180 (90%)	歴史総合、世界史探究	97	97	97	歴史総合、日本史探究	97	92	97	地理総合、地理探究	91	82	72	公共、政治・経済	100	97	100	数学I・数学A	100	100	89	数学II・数学B・数学C	100	94	88	物理	95	81	81	化学	100	90	87	物理基礎	50	42	42	化学基礎	50	47	50	地学基礎	50	47	47	生物基礎	50	50	50	情報I	100	91	96
科目	GPT-5.2 Thinking	Gemini 3.0 Pro	Claude Opus 4.5																																																																				
英語リーディング	97	97	95																																																																				
英語リスニング	96	93	93																																																																				
国語	180 (90%)	171 (85.5%)	180 (90%)																																																																				
歴史総合、世界史探究	97	97	97																																																																				
歴史総合、日本史探究	97	92	97																																																																				
地理総合、地理探究	91	82	72																																																																				
公共、政治・経済	100	97	100																																																																				
数学I・数学A	100	100	89																																																																				
数学II・数学B・数学C	100	94	88																																																																				
物理	95	81	81																																																																				
化学	100	90	87																																																																				
物理基礎	50	42	42																																																																				
化学基礎	50	47	50																																																																				
地学基礎	50	47	47																																																																				
生物基礎	50	50	50																																																																				
情報I	100	91	96																																																																				
※GPT-5.2 Thinking・Gemini 3.0 Pro・Claude Opus 4.5いずれも、問題を画像化しAPI経由で読み込ませて実行。 ※各モデル、最高精度のモデルを選定したが、GPT-5.2 Proは複数科目で試験時間をオーバーしたため、GPT-5.2 Thinkingの結果を採用。																																																																							
。解答時間は約1時間40分でGeminiと同程度 ²³ 。 総合ではGeminiとほぼ同水準の得点率。																																																																							

（※各モデルとも問題は画像化しAPI経由で投入。同一の試験環境下で実施）

上の表より、OpenAIモデルが得点面で突出していたものの、Google・Anthropicのモデルも9割前後の高水準にあり、それぞれ異なる強みを発揮していたことが分かります。例えばGeminiは**地図や図表の純粋な画像情報を直観的に読み取る力**で一步先を行き³²、Claudeは**文章の文脈把握力**で国語などを安定して高得点に導いたと考えられます³⁰。一方、**全モデル共通の弱点**として浮かび上がったのは前述の**視覚的な細かい解釈**（矢印の意味、色のグラデーションなど）や¹⁷²⁶、**人間らしい感情の機微を読む問題**²¹でした。得点率ではOpenAIが勝利しましたが、設問単位で見ると**Geminiのみ正解**できたものの、**全AIが間違えたもの**などもあり²⁷¹⁷、興味深い比較結果となっています。

教育界・大学入試制度への影響

AIが大学入試の共通試験で満点近くを叩き出した衝撃的なニュースは、教育関係者や受験生に大きなインパクトを与えています。「人間が何年もかけて積み上げる努力の結晶（＝高得点）をAIが一瞬で達成してしまっ

た」ように映るため³⁴、「もはや受験勉強に意味はないのでは」といった極端な声も一部で上がりました³⁵。しかし、この出来事は単に不安を煽るだけでなく、**これからの教育の在り方や入試制度を再考する契機**として捉えられています。以下、主な影響と論点を整理します。

・受験勉強・評価方法の再定義:

これまで受験勉強は「知識を頭に詰め込むこと」、学力＝「与えられた問題に正確な答えを出す能力」と定義されがちでした³⁶。ところが、知識暗記や定型問題への対応力において**AIは既に人間を凌駕している**ことが示されました³⁷³⁴。このため、「点数が高い人が偉い」という**点数至上主義のルールが崩れつつある**との指摘があります³⁵。今後は「**学力とは何か**」「**頭が良いとは何か**」について再定義が必要になりそうだという声が教育現場でも聞かれます³⁸。具体的には、**創造力や批判的思考力、コミュニケーション力、問題発見・解決能力**など、単なる知識の正確さでは測れない力にフォーカスすべきとの議論です³⁹。AIが満点を取れる試験を人間同士の選抜にそのまま使い続けることには疑問が生じており、「**AIには解けない問題**」を出すにはどうすればよいかが課題になります⁴⁰⁴¹。例えば記述式の論述問題や、受験生自身の体験・意見を問うような問題、グループディスカッションやプレゼンテーションなど**非定型の評価方法**への移行も考えられます。

・大学入試への具体的な対応策:

日本の大学入試制度において、共通テストは依然マークシート方式（選択肢問題）が中心です。今回AIが「無双」した要因の一つには**マーク式との極めて良好な相性**が挙げられます⁴²。GPTのような生成AIは、膨大なデータから文脈に適した語や選択肢を**確率計算で当てる**のが得意であり、定型化・パターン化された試験では圧倒的強さを発揮します⁴⁰。逆に言えば、人間の思考力や表現力を測るには記述式・論述式の問題導入が有効です。実際、**文部科学省も共通テストへの記述式問題導入とAI採点支援**を検討しており、令和8年度以降の新テスト最終報告案で「記述式問題を導入し採点にAIを活用」といった方向性が示されています⁴³。今後、**AIには容易に解けないオープンな問題**（例：小論文、実技評価など）の比重を増やすことで、人間ならではの力を評価する動きが加速する可能性があります。一方で、採点にAIを活用することで人手では困難な大量の記述答案を処理したり、公平性・採点基準の標準化を図ったりする試みも進むでしょう⁴³。このように「**AIに勝つためにAIを使う**」ような形で入試制度が見直される展望もあります。

・カンニング（不正行為）への対策強化:

AIが試験で高得点を取れるとなれば、受験生が**試験中にAIを利用して不正を試みる**リスクも無視できません。実際に2026年の共通テストでは全国で7件の不正行為が発覚し、**うち3件は受験生が試験中にスマートフォンを使用**していました⁴⁴。福岡県のケースでは、試験中に受験生が問題用紙等をスマホで撮影しようとし、端末から約200枚の写真が見つかったと報じられています⁴⁵。外部への送信は確認されなかったものの、明らかに**問題を持ち出してAIなどに解かせる意図**が疑われる事案です⁴⁶。このような事態を受け、文部科学省は直ちに大学や大学入試センターに対し不正再発防止策の徹底を求める通知を出しました⁴⁷。具体的には試験開始前に通信機器の電源OFFとカバン収納を徹底させること、監督者の巡視強化、必要に応じて受験生本人確認の追加措置を講じること等が指示されています⁴⁸。**今後は試験会場での電子機器対策が一層厳格化**されるでしょう。また、将来的には**AI検知システム**の導入や、試験そのものをAIが使えない環境で行う（例えばネット遮断ルームでの実施）などの工夫も考えられます。さらに踏み込めば、「いっそ受験生もAI使用可とし、その上で人間の創意工夫を見る」というラディカルな意見も一部にはありますが、現段階では公平性と現実性の観点から慎重な議論が必要でしょう。

・学校教育への影響と課題:

「AIが賢くなりすぎて、人間の勉強が無意味に感じる」——今回のニュースに接した学生・生徒の中にはそう感じた者もいるかもしれません³⁵。しかし専門家や教育者は、「それは**一時的なパニック的反応**に過ぎず、むしろ学び方をアップデートするチャンスだ」と強調しています³⁵⁴⁹。AIが得意なのは先述の通り**パターン認識と大量知識の瞬時の検索**です⁴⁰。裏を返せば、**非パターン化の課題**、たとえば新規性の高い問題設定や文脈の裏にある意図・暗黙の意味を読むこと、言語化されてい

ない感情や価値観を察することなどは依然として人間の独壇場です⁴¹。今後の教育では「知識そのものを詰め込む勉強」から「AIには真似できない能力を伸ばす学習」へのシフトが求められます⁵⁰³⁹。具体的には、探究学習やプロジェクト型学習の充実、創造的な課題解決能力や対人コミュニケーションの涵養などが挙げられます。実際、AI時代の学びの方向性として「より高度な個別化」や「創造的思考」がキーワードになるという指摘もあります⁵¹。一方で基礎的な知識教育が不要になるわけではありません。AIを使いこなすためのリテラシー（例えば、適切な指示を与え結果を検証する能力）も含め、人間がAIと共存していく前提の教育課程作りが急務です⁵²。文科省も「AI・データサイエンス教育の強化」を打ち出しており⁵³、初等中等教育から情報科や統計教育の充実が図られています。要するに、教育界はAIを敵視するのではなく道具として取り込みつつ、人間にしかできない力を伸ばす方向へ舵を切る必要に迫られているといえます。

・受験産業・学習塾への影響:

大学受験を巡る産業（予備校・塾、教材出版社等）も、この結果に関心ではられません。AIが9割超の正答率を出した2025年時点でも、河合塾は東京大学文科一類の合格ボーダー得点率を約86%と予想していました²⁵。GPTがそれを上回る91%を取ったことは「AIが東大合格ラインを突破した」ことを意味し⁵⁴、塾業界には脅威であると同時に新たなニーズも生まれています。すなわち、「AI時代の効率的な勉強法を指導する」という役割です⁵⁵。単に知識伝達をするだけならAI講師でも代用可能になりつつあるため、人間の講師はAIの解答プロセスを分析・活用しながら、生徒に思考法を教える方向にシフトするでしょう⁵⁶⁵⁷。例えば、AIが間違えた問題の傾向を示し「ここはAIでも迷う。人間としてはどう理解すべきか？」といった指導を行うなど、AIを踏まえた学習指導が今後の予備校・塾の付加価値になると考えられます⁵⁷。また、受験生自身もAIを練習相手や学習ツールとして活用するケースが増えるでしょう。実際、2025年から共通テスト科目に加わった「情報Ⅰ」では、AI教材を使って学習した生徒が全国平均を20点上回る平均89点を記録した例があります⁵⁸。大手予備校の駿台とEdTech企業のLife is Tech社が提携した特別講義で、生成AIを組み込んだ学習ドリルを体験させたところ顕著な効果が出たという報告です⁵⁸。このようにAIを上手に使いこなした者が成果を伸ばす傾向が出ており、将来的には「AI活用力」が受験戦争の新たな鍵になる可能性もあります。

日本国内の反応（教育関係者・世論など）

今回の「AIが共通テスト全科目で人間を凌駕」というニュースに対し、日本国内では教育関係者から学生、一般社会まで様々な反応が巻き起こりました。

・教育関係者・専門家の声:

学校現場や教育行政の関係者は驚きを示しつつも冷静に受け止めています。ある公立高校長は、新聞報道（埼玉新聞）を引用しながら「AI（生成AI）に共通テストを解かせたら得点率約91%になった」ことに触れ、「昨年11月に河合塾が公表した東大文Ⅰ合格可能性50%ライン（86%）を超えている」と述べています⁵⁹。さらに、この校長は「あの東ロボくん（ロボットは東大に入れるかプロジェクト）は結果的に未達に終わったが、新技術の登場で挑戦が続いていることに時代の進歩を感じる」とも記しています⁶⁰。実際、2011年開始の「東ロボ」プロジェクトではAIが2016年までにセンター試験高得点、2021年に東大入試突破を目標としていましたが、当時の技術では東京大学合格には至らず計画は終了しました⁶¹。そのプロジェクトを率いた新井紀子氏（国立情報学研究所 教授）は「AIは東大に合格できない」という著書を出版していましたが、わずか数年でLLMが共通テスト満点レベルに達した現実、専門家にとっても技術的ブレークスルーの速さを実感させるものとなっています。「知識偏重のテストの限界をAIが証明した」として、今後は人間固有の読解力（特に文章の深い意味を理解する力）をどう育成・評価するかが喫緊の課題だと指摘する声もあります³⁶⁴¹。

・学生・受験生の声:

SNS上では共通テスト当日の自己採点報告と並行して、「ChatGPTが共テ満点」というニュースが急速に拡散し、大きな話題となりました³⁵。多くの受験生は驚愕しつつも、「AIが満点でも自分は満点じゃないし…」「結局自分の努力を続けるしかない」といった現実的な声も聞かれます。一方で、

「もう勉強する意味がないのでは」という悲観的な投稿も少なからず見受けられました³⁵。これについて、教育系ブロガーのある投稿は「それら極端な意見は変化への恐怖による一時的パニックに過ぎない」と述べています³⁵。「点数至上主義」というこれまで信じてきたルールがAIによって揺らされたことへの防衛反応だとの分析です⁶²。実際、「自分の努力がちっぽけに感じた」「ペンを置きたくなった」といった声に対し、有識者らは「冷静になれば、便利な道具が登場しただけだ。極端な意見に振り回されないで」と呼びかけています³⁵。また、「勉強の意味がなくなったのではなく、問われる力が変わるだけ」とも強調されています³⁵⁴⁰。実際、SNS上には「AI時代に活躍できる人になるため勉強を続ける」「AIにはできない体験を積もう」と前向きな投稿をする学生もあり、必ずしもネガティブ一色ではありませんでした。

・世論・メディアの論調:

一般のニュース番組やワイドショーでもこの話題が取り上げられ、「詰め込み型の教育カリキュラムが意味をなさなくなるのでは」といった問題提起がなされました⁶³。コメンテーターの中には「日本の教育はこのままで良いのか？」と文科省や学校現場に疑問を投げかける者もありました⁶³。一方で「電卓が登場しても計算力教育は無くならなかったように、AIとも共存していくしかない」と冷静な意見もあります。保護者世代からは、「我が子が大学受験する15年後には一体どんな教育になっているのか想像もつかない」といった戸惑いの声もありつつ、「今後必要とされる力を考えるきっかけになった」という前向きな意見も聞かれます⁶⁴³⁹。特に幼い子を持つ親世代では、「知識暗記や計算練習だけでない経験・体験に重きを置いた教育が重要では」という指摘もあります⁶⁵。世論全体としては、「AIに仕事や機会を奪われる不安」と「AIを活用して便利になる期待」の両面が渾然一体となった複雑な反応と言えるでしょう。ただ少なくとも教育に関しては、「AIに負けない人材を育てなければ」という機運が高まりつつあり、社会全体で議論が進み始めています。

・AI研究者・技術者の反応:

日本のAI研究コミュニティにとっても、この成果は大きな話題となりました。多くの研究者が「驚異的だが驚きすぎない」というスタンスで、LLMの能力進歩に改めて注目しています。例えば、あるAI研究者はTwitterで「ChatGPTが共通テスト15科目中9科目満点、得点率97%」という事実を紹介し、「昨年は理III（東大医学部）合格ラインを突破したが、今年は満点まで見えてきた」と言及しています⁶⁶（実際、LifePromptは共通テスト検証に続いて東大二次試験にAIで挑戦する企画も始動させています⁶⁷）。研究者たちは一様に、ここ数年のディープラーニング・言語モデルの進歩の凄まじさに触れつつ、同時に「選択肢問題で満点でも、人間のように理解しているとは限らない」と冷静な分析もしています。上述の通りAIが誤答した問題（画像の微妙な読み取りや人の心理理解）に注目し、「まだAIにも弱点がある。万能の神ではない」という点を強調する声もあります⁶⁸⁴¹。これは過度な期待や不安を抑え、客観的に現状を評価しようという姿勢です。また、日本のAI政策に関わる専門家からは、「これだけの性能差を見せつけられると、日本独自の大学入試の意味付けを再考せざるを得ない。受験のあり方だけでなく、日本語での大規模モデル開発や教育データの整備など戦略を立てる必要がある」といった提言も出ています。要するに、AI研究者から見ても教育界から見ても、今回の成果は現行システムへの警鐘であり、新たな方向へのスタート地点と受け止められているようです。

スタートアップ・AI業界への影響

OpenAI vs Googleといった巨大テック企業の競争の中で、今回の共通テスト満点劇は教育分野におけるAI活用の可能性を強烈に示すデモンストレーションとなりました。これにより、スタートアップ企業やAI業界全体にもさまざまな影響が及ぶと考えられます。

・教育×AI（EdTech）の実装加速:

この成果は、「AIを教育現場に取り入れれば大きな効果が得られる」という認識を広めました。既に学校や塾では生成AIを活用した教材や指導法の開発が進みつつありますが、その動きが一層加速するとみられます。先述の通り、駿台予備校とEdTechスタートアップが開発した「情報AIドリル」では共

通テスト情報Ⅰの得点向上に顕著な成果が出ました⁵⁸。また**GMOメディア**はAIに情報Ⅰを学習させて共通テスト対策問題を自動生成・提供することを発表しており、**AIによる問題作成サービス**にも乗り出しています⁶⁹。今後、**個別最適化学習**（生徒一人ひとりに合わせた問題出題・解説）にAIを用いる試みや、AI家庭教師・AI面接官といったソリューションも現実味を帯びてきます。教育系スタートアップにとって、今回のニュースは「**生成AIの実力**」を**アピールするまたとない追い風**であり、学校や塾との提携、自治体との実証実験などビジネスチャンスが広がるでしょう。

・LifePromptなど国内スタートアップの台頭:

今回の検証を主導した株式会社LifePrompt自体、2023年創業の新進スタートアップです⁷⁰。社員数数名規模ながら、**大手新聞社（日経）と共同で国家試験レベルのAI実験をやり遂げた**ことで、その技術力と企画力が広く知られることになりました²。LifePromptは企業向けの生成AI導入コンサルやシステム開発も手掛ける会社で⁷¹、今回のような派手な実証で知名度を上げたことで**資金調達や新規事業展開**にも追い風となる可能性があります。加えて、この分野には他にも**Atama Plus**（AI教材で学習最適化を図るスタートアップ）や**Study-AI**系の新興企業が存在し、**生成AIブームと教育改革ニーズを背景に投資家の注目を集めている**状況です。2025年にはベンチャーキャピタル各社もEdTech×AI領域への出資を拡大しており、生成AIを組み込んだ教育サービスの市場規模は急拡大すると予想されます⁷²。今回OpenAIが証明した高い学習能力を踏まえ、「**第二第三のLifePrompt**」が登場してくる可能性も十分あります。

・大企業（OpenAI・Google等）の戦略と競争:

OpenAIが「Googleに勝利」と報じられたように⁷³、今回の結果は**大規模言語モデルの競争においてOpenAIが一歩リード**していることを示す一例となりました。Googleは対抗して開発中の**次世代モデル「Gemini」**を2025年末に公開しましたが、最上位版（Gemini 3 Pro）を投入してもChatGPT最新版には及ばなかったこととなります²³。もっとも、Geminiは画像理解力や速度で優位性を示しており⁷⁴、今後の改良で正解率向上も十分見込まれます。Googleとしては、自社AIが特定の実用タスク（試験問題解答など）で他社を上回ることをアピールしたいところであり、**教育分野での再チャレンジ**や別のベンチマークでの競争が続くでしょう。Anthropic社のClaudeも然りで、特に**倫理的な安全性や人間らしさ**を売りにする同社としては、満点を逃した背景にある誤答パターン（人間らしい弱さの読み取りの失敗など²²）を解消すべくモデル訓練に反映させるかもしれません。いずれにせよ、OpenAI・Google・Anthropicといった主要プレイヤーは**互いのモデルを教育用途でも競わせる時代**に入ったといえます。この競争は結果的にモデル性能全体の底上げを促し、次年度には「AIが共通テスト全科目満点」を複数社が達成する可能性もあります。その場合、人間の試験に対するアプローチもさらに変えざるを得なくなるでしょう。

・市場の関心の変化と課題:

教育へのAI実装に社会の関心が集まる一方で、いくつかの課題も指摘されています。**倫理的問題**としては、AIが勉強を代替し過ぎることへの懸念です。学生がAI任せで学習プロセスを省略してしまうと、本来身に付けるべき基礎思考力が育たない恐れがあります。そのため、適切な**ガイドライン作り**や教師による指導（例えば「レポートにAIを使う際は出典を明示」「AIの答えをうのみにせず検証する姿勢を教える」など）が不可欠でしょう⁷⁵。また**格差の拡大**も問題です。高度なAIサービスは往々にして高コストであり、経済的に余裕のある家庭や進学校の生徒ほど最新AIツールにアクセスできて有利になる可能性があります。国として教育AIのインフラ整備やライセンス提供を進め、どの子供も恩恵を受けられるようにする施策が求められます。さらに**データプライバシー**の問題もあります。生徒の学習データ（成績や解答傾向など）がAIに蓄積されることで個人情報の扱いに注意が必要です。スタートアップ企業にとっては、これら**倫理・規制面の課題をクリアしつつ、教育現場の信頼を得る形でサービス展開**することが成否を分けるでしょう。

総じて、OpenAIの共通テスト満点達成というニュースは、教育業界・スタートアップ業界に**衝撃と新たな展望**をもたらしました。受験生にとっては不安半分・期待半分かもしれませんが、技術の進歩を前提に**教育シ**

システムをアップデートする絶好の機会とも言えます。今後、AIと人間が協調しながら学びを深めるために、社会全体で知恵を出し合っていくことが求められています。

参考資料・出典: 本調査レポートは、日経新聞（2026年1月20日付）「大学入学共通テスト、OpenAIは9科目満点 得点率97%でGoogleに勝利」⁷³ の報道内容を起点に、LifePrompt社による詳細分析ノート⁴¹⁷、電ファミニコゲーマーの記事¹¹⁸、ならびに共同通信・熊本朝日放送など信頼できるメディア報道⁴⁴、教育専門家のコメント³⁶³⁵、関連するEdTech事例紹介⁵⁸等を幅広く参照してまとめています。各種データ・発言の出典は本文中の【○†Ln-Ln】に示した通りです。

¹ ² ³ ¹⁸ ⁶⁶ ⁶⁷ 2026年の大学入学共通テスト、ChatGPTが「9科目で満点」を獲得。合計点数でも Gemini、Claudeに差をつける

<https://news.denfaminicogamer.jp/news/260120g>

⁴ ⁵ ⁶ ⁷ ⁸ ⁹ ¹⁰ ¹¹ ¹² ¹³ ¹⁴ ¹⁵ ¹⁶ ¹⁷ ¹⁹ ²⁰ ²¹ ²² ²³ ²⁶ ²⁷ ²⁸ ²⁹ ³⁰ ³¹ ³² ³³ ⁷⁴ 【満点9科目！】 共通テスト2026を最新版AIに解かせてみた（Chatgpt、Gemini、Claude） | 株式会社 LifePrompt

<https://note.com/lifeprompt/n/nb87edfb2e7ca>

²⁴ 大学入学共通テスト、最新の人工知能(AI)モデルの結果。得点率は ...

<https://x.com/asasukeno/status/2013382930538569876>

²⁵ ⁵⁴ ⁵⁹ ⁶⁰ ⁶¹ 校長日誌 - 埼玉県立川越南高等学校

https://kawagoeminami-h.spec.ed.jp/blogs/blog_entries/view/469/77acc388591768894069b47afac0d611?frame_id=485

³⁴ ³⁵ ⁴⁰ ⁴¹ ⁴² ⁴⁹ ⁵⁵ ⁵⁶ ⁵⁷ ⁶² ⁶⁸ チャットGPTが共通テスト満点…もう受験勉強は意味ない？不安になる前に知るべき現実

<https://shiritaino-blog.com/2026/01/20/chatgpt/>

³⁶ ³⁷ ³⁸ ³⁹ ⁵⁰ ⁵² ⁶⁴ ⁶⁵ AIが共通テストで正答率9割超。これからの教育に思うこと | あおさぎ | 建築構造のあれこれを発信します

<https://note.com/tkr415/n/n1092b34b12f8>

⁴³ 大学入試新テスト、文科省が最終報告案…採点にAI活用 - リセマム

<https://resemom.jp/article/2016/03/14/30320.html>

⁴⁴ ⁴⁵ ⁴⁶ ⁴⁷ ⁴⁸ 不正行為の防止徹底を求める通知 共通テストでのスマホ使用事案を受け 文科省 | KAB ONLINE

<https://www.kab.co.jp/news/article/16296111>

⁵¹ AIで教育はどこまで進化できるか——広がる学びの可能性と人間の ...

<https://globis.jp/article/8x0o1tfcjr/>

⁵³ 令和7年度 大学入学者選抜・教務関係事項連絡協議会 - 文部科学省

https://www.mext.go.jp/a_menu/koutou/senbatsu/1419601_00010.htm

⁵⁸ AI教材で共通テスト「情報I」平均20点アップ！駿台とEdTech企業が次世代教育の実証へ～今夏に体験プログラムを全国実施 | エデュマッチ

<https://edu-match.com/ai-news250712/>

⁶³ 9科目満点 共通テスト解答、AI学力向上（共同通信） - チャットGPT

<https://x.com/lchigerc/status/2013406029418009052>

⁶⁹ AIを活用した問題作成技術により大学入学共通テスト対策向けの問題 ...

<https://www.gmo.media/archives/6837/>

- 70 LifePrompt | STARTUP DB (スタートアップデータベース)
<https://startup-db.com/companies/VYa2pRDUbM7J3R6D>
- 71 株式会社LifePrompt (ライフプロンプト) | 人とAIの垣根をなくし ...
<https://lifeprompt.net/>
- 72 AI採用のフォワード、資金調達を実施 累計調達額は2.35億円に到達
<https://prtimes.jp/main/html/rd/p/000000007.000117983.html>
- 73 大学入学共通テスト、OpenAIは9科目満点 得点率97%でGoogleに勝利
https://x.com/kanpo_blog/status/2013427196963233881
- 75 【#生成AI の回答、そのままで大丈夫? 】 AIの答え ... - Instagram
<https://www.instagram.com/reel/DTHkuOHCMIg/>