

2026年 共通テストAIシヨック： GPT-5.2が叩き出した97%の衝撃

9科目満点達成の舞台裏と、Google・Claudeとの比較、そして教育の未来へ



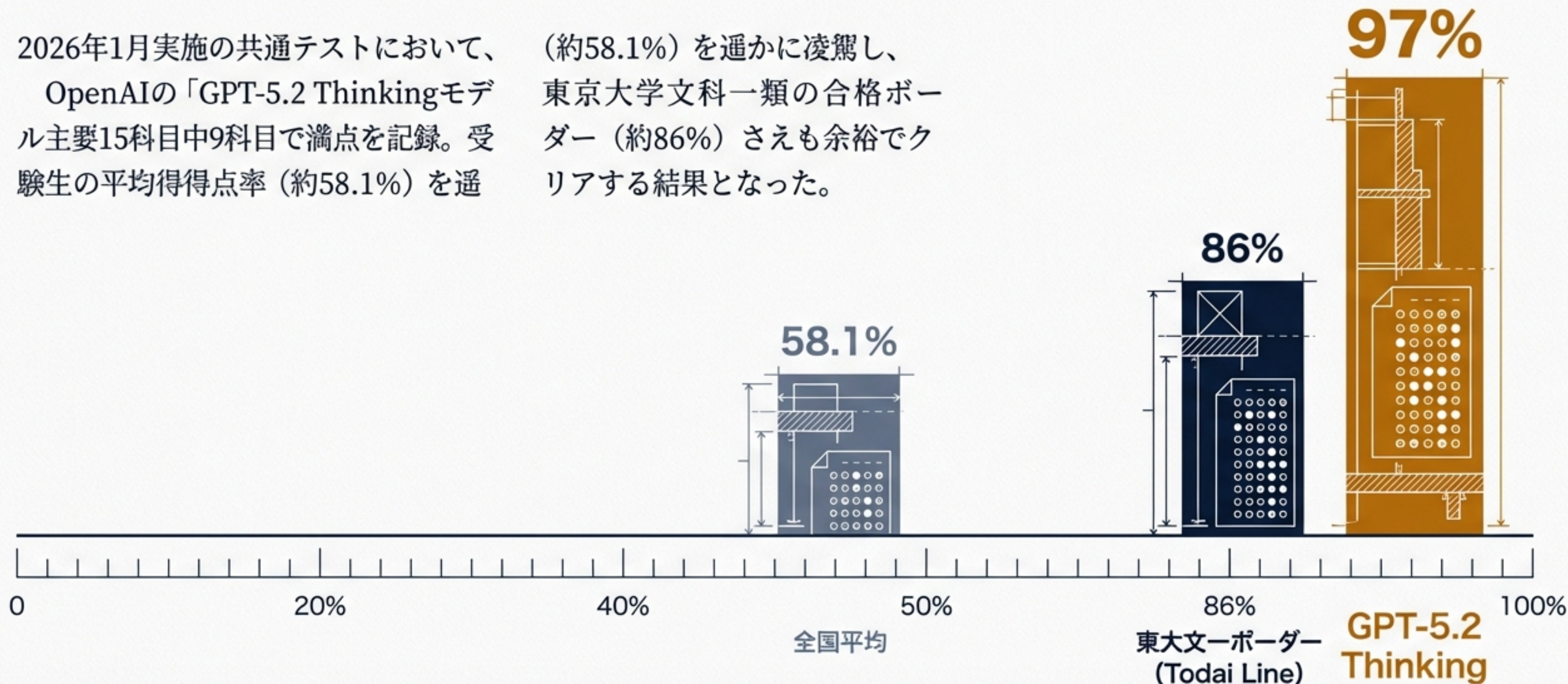
Verified by
LifePrompt &
Nikkei Shimbun (2026)



「東大入試ライン」を突破。AIは平均点58%の壁を越え、97%の領域へ。 主要15科目中 9科目で満点

2026年1月実施の共通テストにおいて、
OpenAIの「GPT-5.2 Thinkingモデル」主要15科目中9科目で満点を記録。受験生の平均得点率（約58.1%）を遥

（約58.1%）を遥かに凌駕し、
東京大学文科一類の合格ボーダー（約86%）さえも余裕でクリアする結果となった。



3大AIモデル徹底比較：王者OpenAIに迫るGoogleとClaude

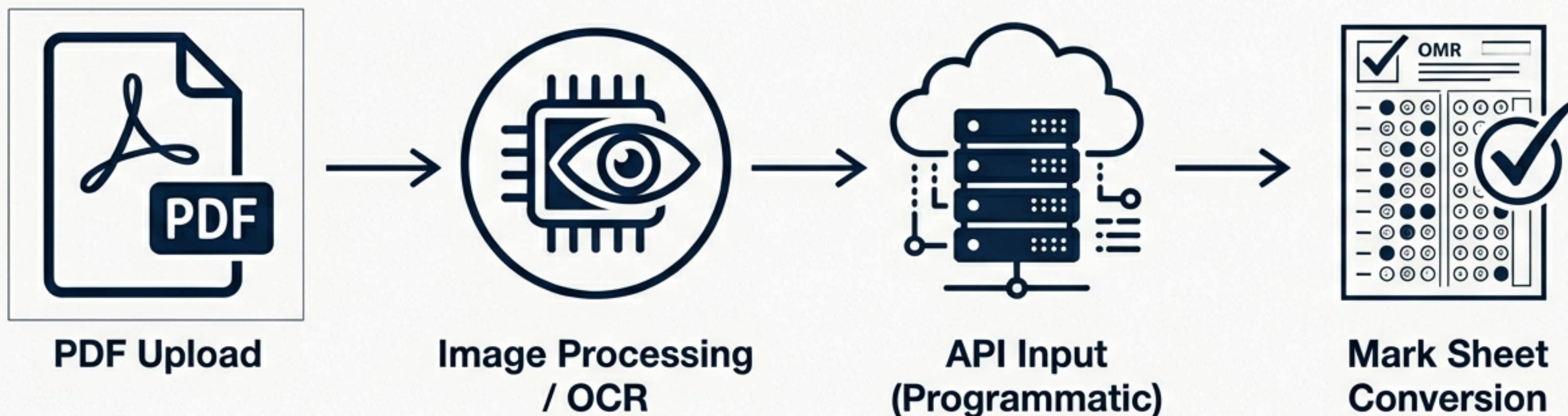
科目	GPT-5.2 Thinking	Gemini 3.0 Pro	Claude Opus 4.5
英語リーディング	97	97	95
英語リスニング	96	93	93
国語	180 (90%)	171 (85.5%)	180 (90%)
歴史総合、世界史探究	97	97	97
歴史総合、日本史探究	97	92	97
地理総合、地理探究	91	82	72
公共、政治・経済	100	97	100
数学I・数学A	100	100	89
数学II・数学B・数学C	100	94	88
物理	95	81	81
化学	100	90	87
物理基礎	50	42	42
化学基礎	50	47	50
地学基礎	50	47	47
生物基礎	50	50	50
情報I	100	91	96
	GPT-5.2 Thinking 97% (Rank 1)	Gemini 3.0 Pro 91% (Rank 2)	Claude Opus 4.5 91% (Rank 3)

GeminiとClaudeも得点率9割（900点台前半）をマークし、人間なら超難関大レベルに到達。しかし、数理・情報・公民で「満点」を取り切るGPT-5.2の総合力が頭一つ抜け出した。

※GPT-5.2 Thinking・Gemini 3.0 Pro・Claude Opus 4.5いずれも、問題を画像読み取りAPI経由で読み込み実行。

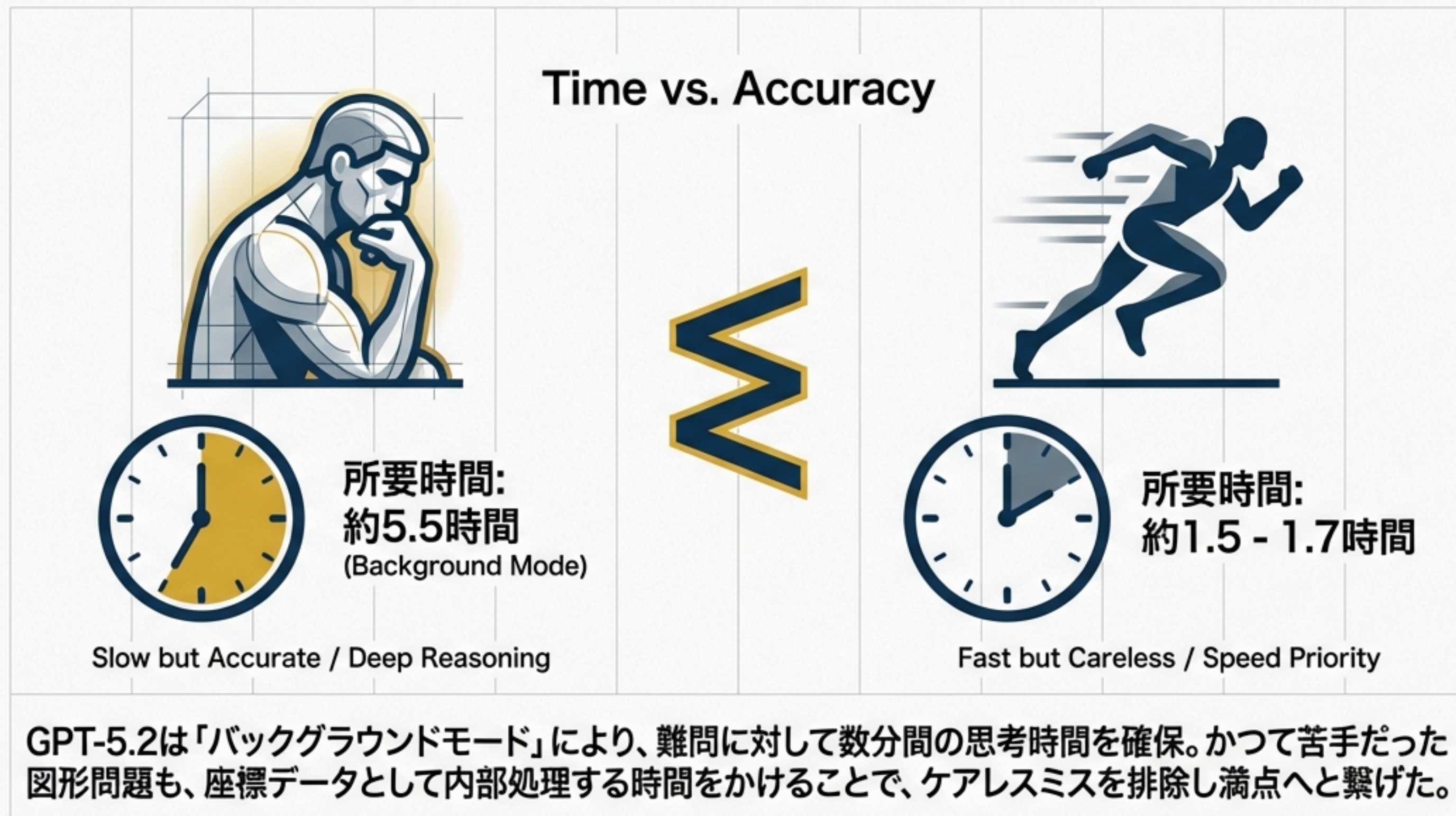
※各モデル、最善性能のモデルを選定したが、GPT-5.2 Proは複数科目で試験時間枠をオーバーしたため、GPT-5.2 Thinkingの結果を採用。

検証の信頼性：恣意的介入を排除した「自動受験システム」

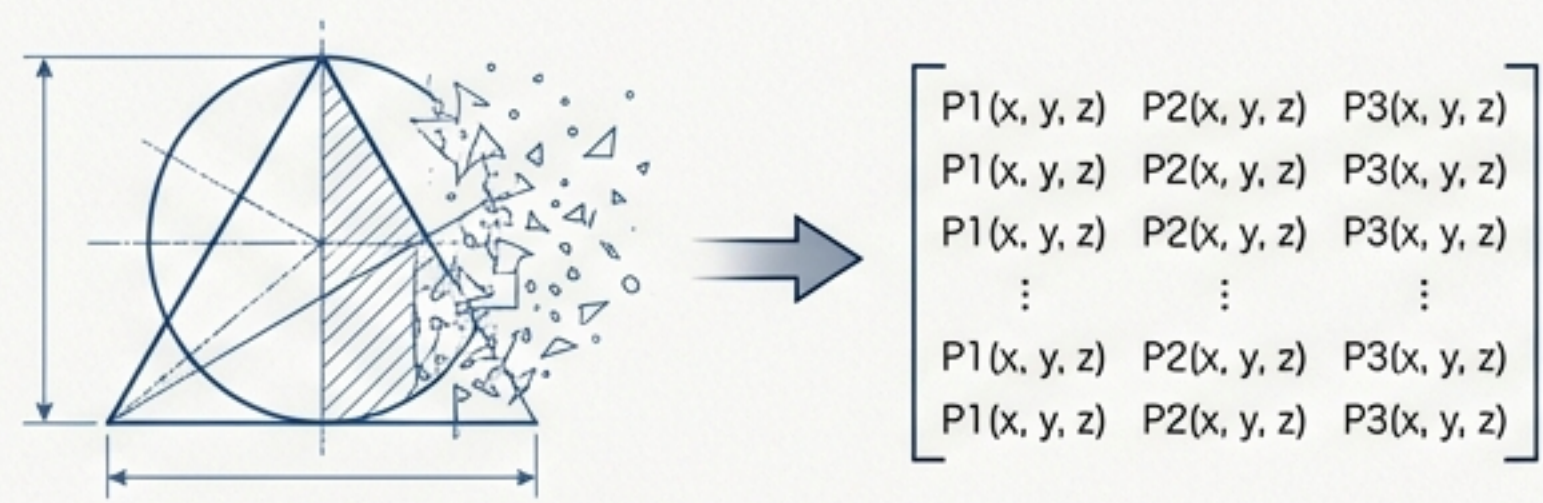


- ・完全自動化: 人為的なコピペミスやヒント提示を防ぐため、API経由でプログラマ的に出題。
- ・特例措置: 英語リスニングは「読み上げスクリプト」、国語（縦書き）は「文字起こしテキスト」を入力し、フォーマットの壁を排除して純粋な知能を測定。

「熟考」が制した勝負：速さのGemini、深さのGPT

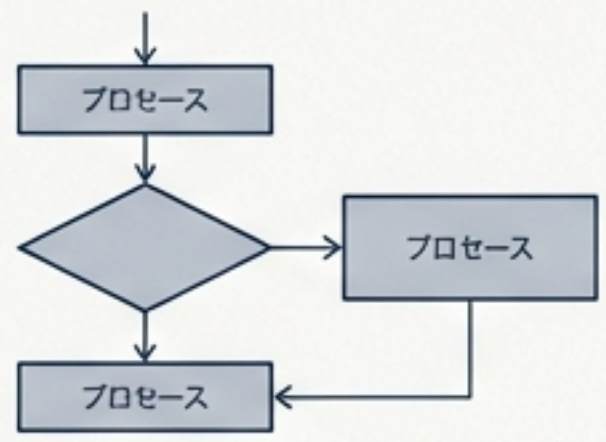


苦手を克服したAI：数学・情報で「満点」の衝撃



図形を絵としてではなく「座標データ」として再構築

```
def solve_logic(data):  
    # process input  
    if condition:  
        return result  
    else:  
        return alternative
```



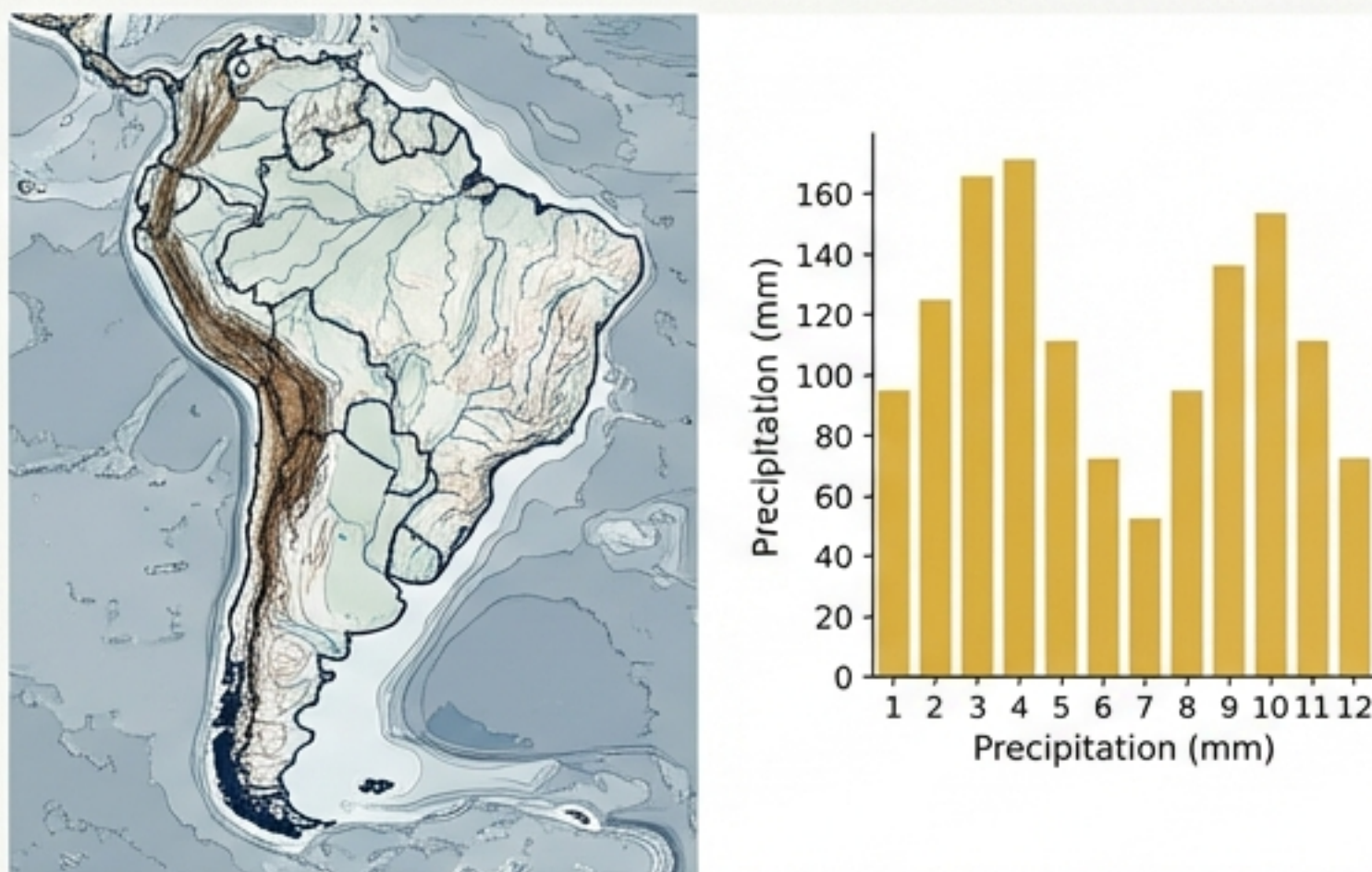
情報I: プログラミング的思考と完全合致

	GPT-5.2 Thinking	Gemini 3.0 Pro	Claude Opus 4.5
数学I・数学A	100	100	89
数学II・数学B・数学C	100	94	88
物理	95	42	81
情報I	100	91	96

昨年までAIの鬼門だった「図形問題」。GPT-5.2は図形を絵としてではなく「座標データ」として再構築して計算するアプローチで克服。プログラミング的思考を問う「情報I」でも満点。

Geminiの逆襲：地理で見せた「視覚的直感」の勝利

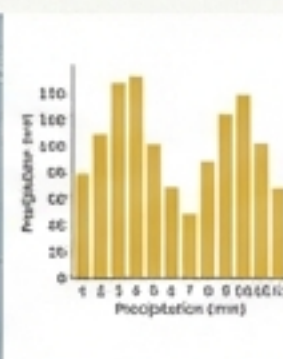
The Problem: 地理的データとその解釈



地形図と気候グラフの対応付け

The AI Analysis: 視覚情報の処理

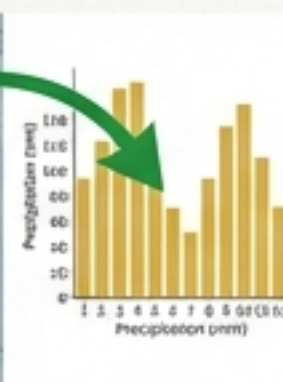
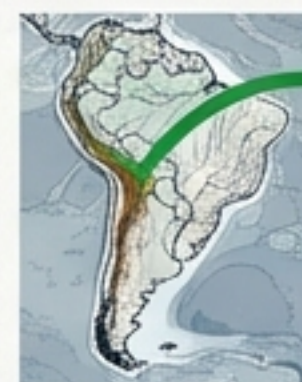
GPT-5.2 / Claude



Map treated as text block -> Failed connection.



Gemini 3.0 Pro



Thinking Level: High (Visual). Intuitive Link Established.



論理構成力ではGPTが勝るが、純粋な画像の「直感的な読み取り」ではGeminiに軍配が上がった。アンデス山脈と気候グラフの対応付けにおいて、Geminiだけが視覚情報を正しく処理した。

AIが越えられない壁①：空間的常識と「矢印」の魔力



Listening Script: “Enter from the back, exit from the front.” ✓

Visual Choice: Select the diagram with correct arrows. ✗

テキスト情報の論理処理は完璧だが、「図中の矢印が示す方向」や「バスの奥行き」といった空間的な常識 (Spatial Common Sense) の結合にまだ課題が残る。

AIが越えられない壁②：割り切れない「人間の感情」



AIは大量の道徳的データから「人は反省して成長するもの」というバイアスを持つ。
そのため、人間特有の「弱さ」や「妥協」といった白黒つかない感情の機微（行間）を読み違えた。

教育のパラダイムシフト：「正解」から「問い」へ

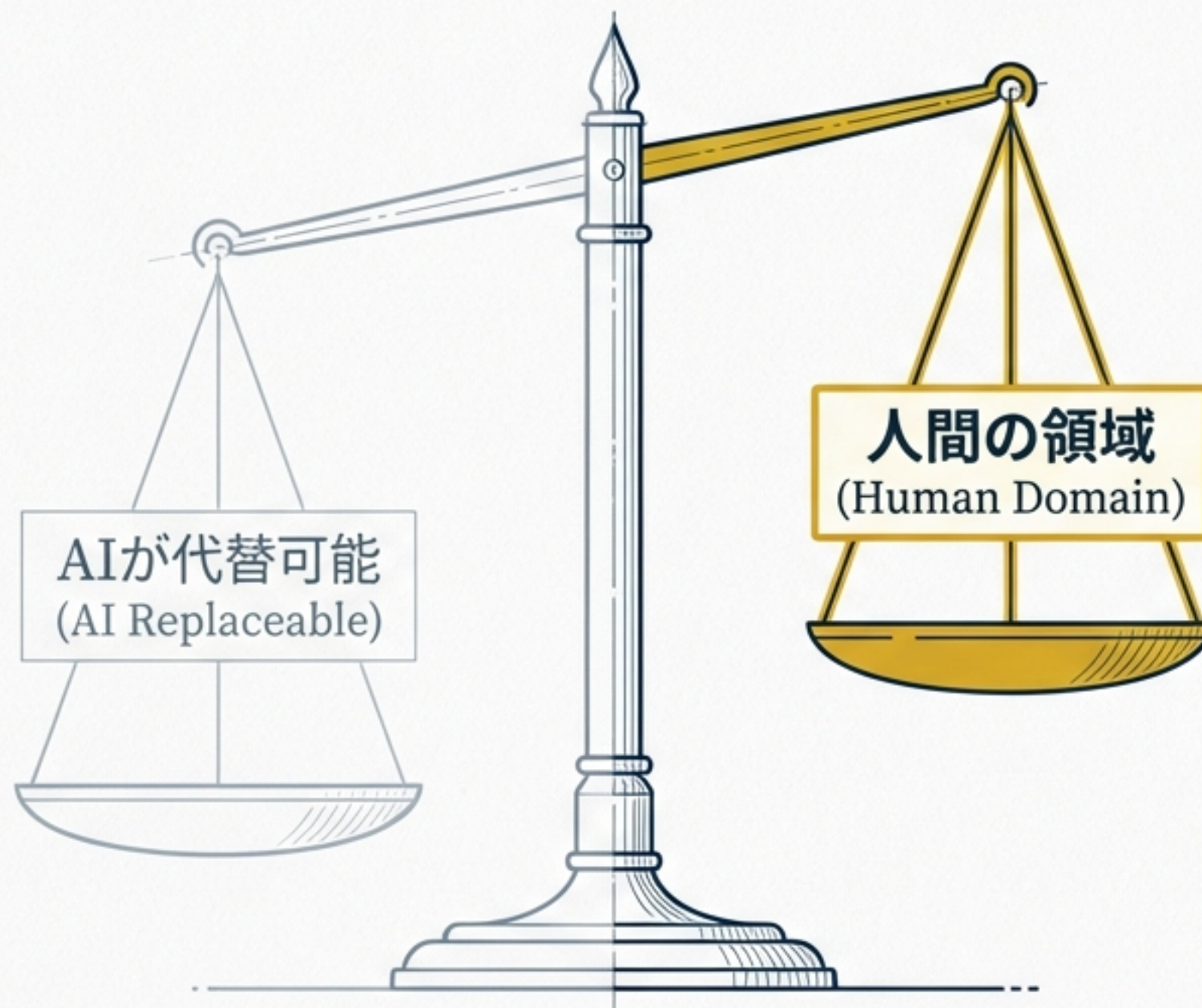


Old Values (点数至上主義)

知識の暗記
(Memorization)

処理速度
(Speed)

正確性
(Accuracy)



New Values

課題発見能力
(Issue Discovery)

クリティカルシンキング
(Critical Thinking)

プロンプトエンジニアリング
(Prompt Engineering)

脅威と機会：カンニング対策から「AI家庭教師」へ



カンニングの脅威

2026年 不正行為7件発生。
監視強化と電源OFFの徹底。

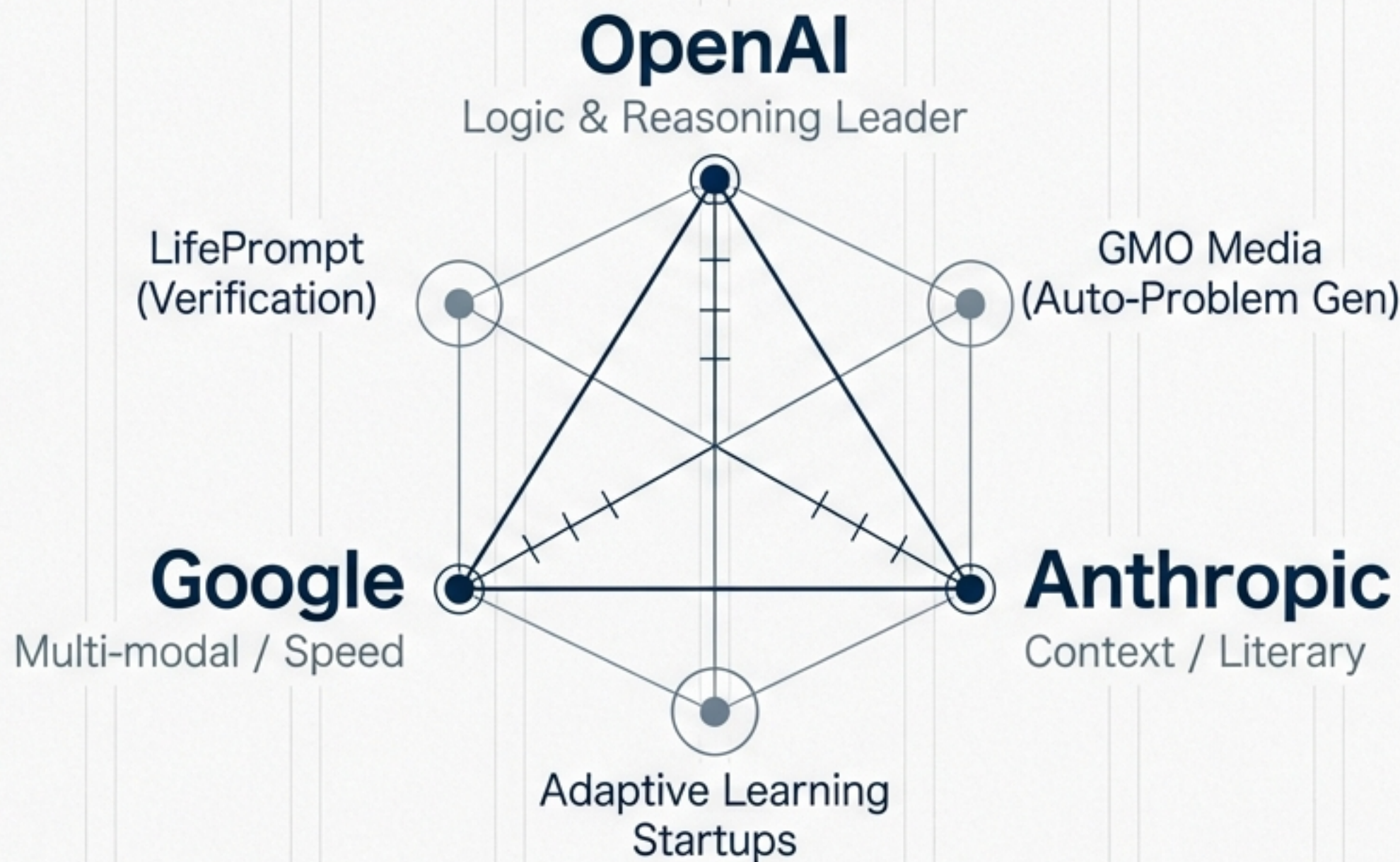


AI家庭教師の可能性

「情報I」 AI活用層は平均**89点**
(全国平均**+20点**)。

AIを禁止する「聖域」としての入試と、
AIを使いこなすための「訓練」としての教育が二極化していく。

加熱するAI開発競争とEdTech市場の爆発



Startups utilizing these **'97% accurate'** models will disrupt traditional cram schools. AI literacy becomes the new competitive advantage.

結論：AIに「勝つ」のではなく、AIを「使う」未来へ

“The test is no longer about beating the AI, but using it.”

知識量と論理処理でAIが人間を凌駕した今、人間は「人間しかできない体験」や「身体性」、そして「AIをどう導くか」に注力すべきである。

