2025年11月、AIの勢力図は一夜にして塗り替えられた

危機に瀕したOpenAIは「コードレッド」を発令

戦略転換（Strategic Pivot）：「新機能の追加」から「基礎体力(ファンダメンタルズ）の強化」へ

リソース集中（Resource Focus）：全エンジニアリングリソースを3つの核心領域に再配分

- 推論能力の強化
- 信頼性の向上（ハルシネーション低減）
- 応答速度の改善

開発凍結（Development Freeze）：実験的なエージェント機能や広告関連プロジェクトを一時凍結

緊急リリース（Forced Release）：Geminiの勢いを断ち切るため、リリース計画を数週間前倒し

「魔法」から「実務」へ：GPT-5.2は特定業務に特化したモデルファミリーとして誕生

GPT-5.2 Instant

役割: 即時性と低コスト
ユースケース: チャットボット、リアルタイム翻訳、単純な情報検索

GPT-5.2 Thinking

役割: 本リリースの主力
ユースケース: コーディング、データ分析、契約書レビューなど、実務的な知識労働

GPT-5.2 Pro

役割: 妥協なき知能の頂点
ユースケース: 科学的発見、医療診断、複雑なシステム設計

開発者を魅了する、実用性を突き詰めた技術仕様

コンテキストウィンドウ

400k

トークン

Gemini 3の2Mトークンに対し、「実用上十分かつ処理速度を維持できるサイズ」を選択。長文脈の処理精度(NIAH)は256kまでほぼ100%を維持。

最大出力トークン

128k

トークン

従来の4k/16k制限からの飛躍的向上。レポート全文やアプリケーションのソースコードを一括生成可能に。

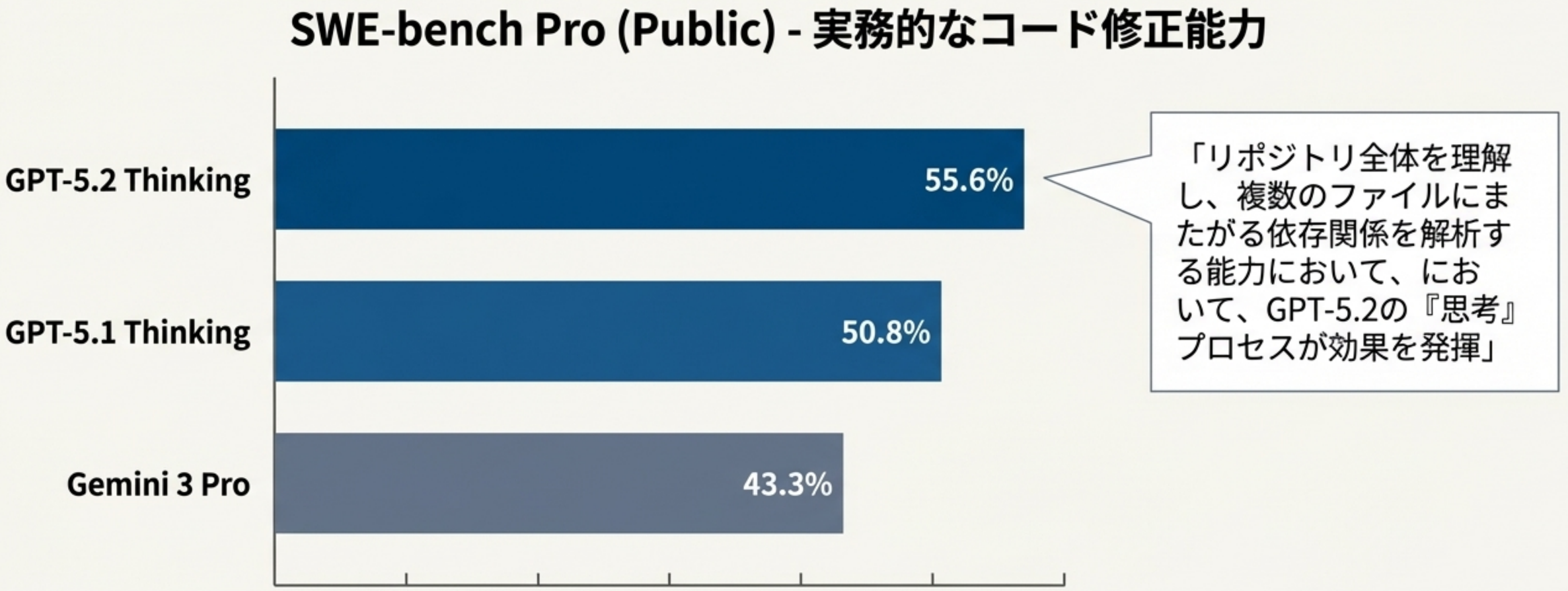
ツール使用

98.7%

Tau2-bench (Telecom) での成功率

外部API呼び出しのミスが大幅に減少し、エージェントとしての安定性が向上。

コーディング能力の王座を奪還：実務タスクで競合を圧倒



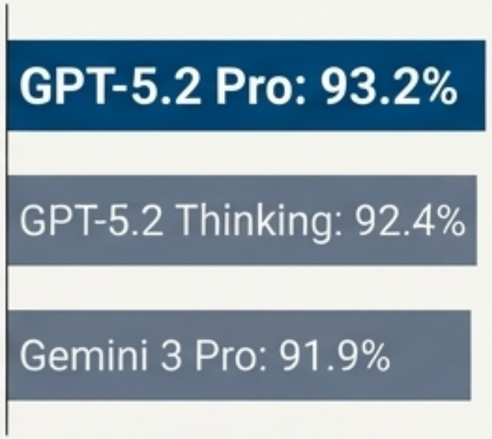
数学的推論で「完全」を達成、AIが人間を凌駕した瞬間

100.0%

AIME 2025
（米国高校生数学オリンピックレベル）

「もはや『計算ミス』や『論理の飛躍』は、
このレベルのモデルでは過去のものとなりつつある」

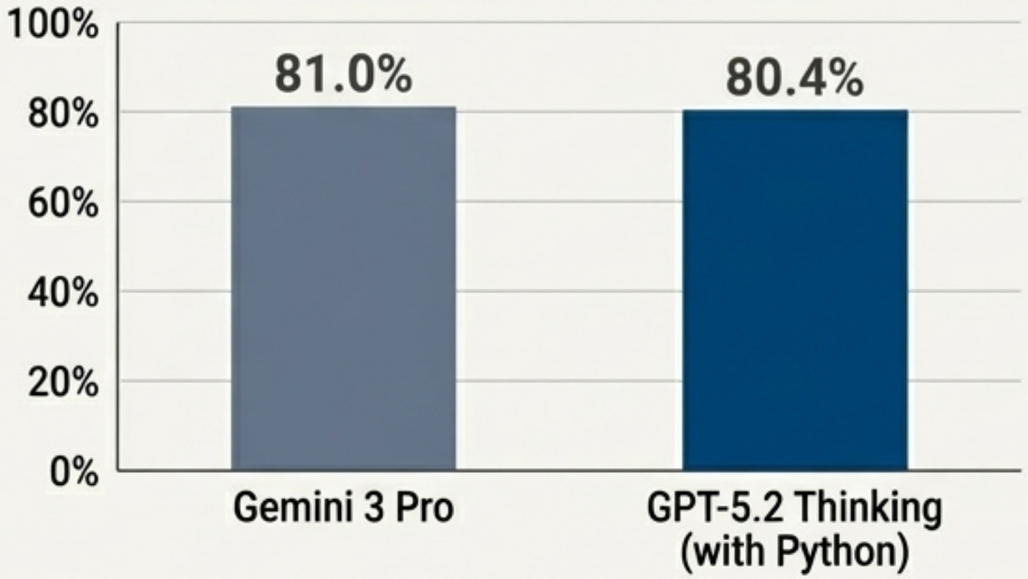
GPQA Diamond
（博士号レベルの科学難問）



一方、マルチモーダルとコーディングの「美しさ」では熾烈な競争が続く

マルチモーダル性能

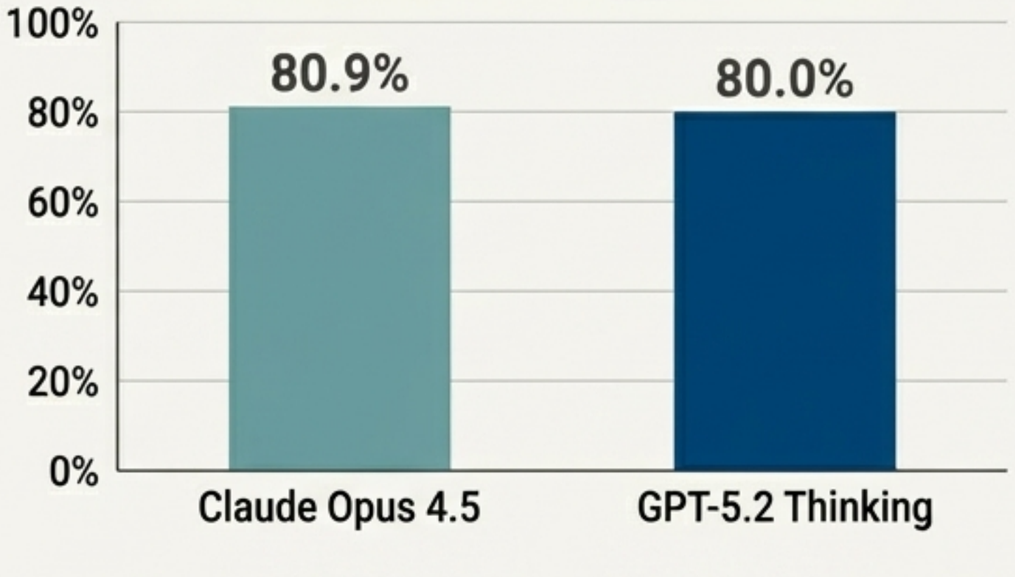
MMMU Pro（大学レベルの視覚推論）



「芸術的な画像の解釈や、動画のネイティブな理解においては、Gemini 3の方が『直感的』で『流暢』」

コーディングの品質

SWE-bench Verified
（厳密に検証されたデータセット）



「『Claudeの書くコードの方がエレガントで可読性が高い』というユーザーレビューも存在」

AIの価値を再定義する価格戦略：「知能」はもはやコモディティではない

モデル (Model)	入力 (Input)	キャッシュ入力 (Cached)	出力 (Output)
GPT-5.2 Pro	**$21.00**	**-**	**$168.00**
GPT-5.2 Thinking	$1.75	$0.175	$14.00
GPT-5.1	$1.25	$0.125	$10.00
Gemini 3 Pro	~$2.00	-	~$12.00
Claude Opus 4.5	~$5.00	-	~$25.00

Prices per 1M tokens

Proモデルの$168という価格は、人間の専門家の代替を示唆。一般利用ではなく、高度な専門職や研究機関がターゲット

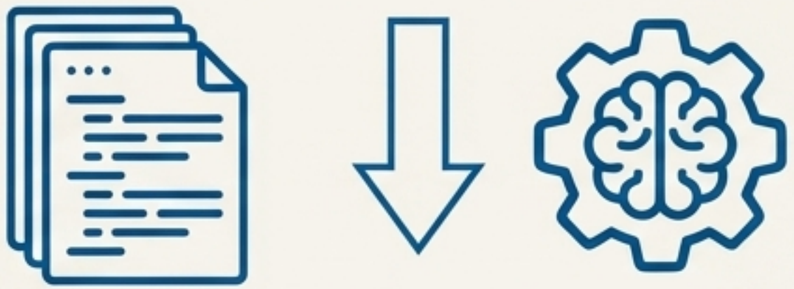
主力モデルはGPT-5.1から約40%の値上げ。
『性能向上分は価格に転嫁する』という明確な姿勢

開発者の運用コストを劇的に削減する「キャッシュ入力」という切り札

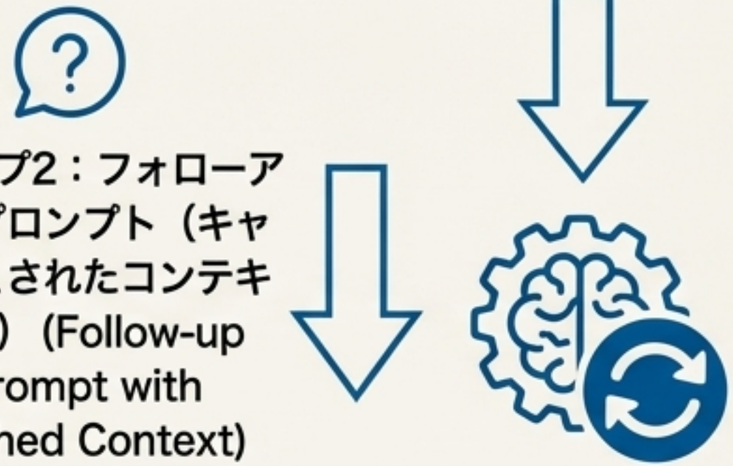
90% OFF

劇的なコスト削減

ステップ1：初期プロンプト
(Initial Prompt)



コスト: $1.75 / 1M tokens



ステップ2：フォローアッププロンプト（キャッシュされたコンテキスト）(Follow-up Prompt with Cached Context)

コスト: $0.175 / 1M tokens

インパクト分析

- 巨大なマニュアルやコードベースを毎回読み込ませるRAGシステムにおいて、コストを劇的に圧縮
- WindsurfやCursorのようなAIエディタにとっては、ビジネスモデルの生命線となり得る
- GoogleやAnthropicに対する強力な差別化要因

現場の声：「退屈になった」と「信頼できる」に二分されるコミュニティの評価

肯定的な反応：信頼性の勝利

> **嘘をつかなくなった。**以前のモデルより明らかに誠実になった

> Claudeで解決できなかった難解なバグが、あっさりと解決できた

> **地味だが仕事に使える。**業務で使うならこれがベスト

否定的な反応：創造性の喪失

> ガードレールが強すぎる。モデルが**去勢された (nerfed)**

> GPT-4のような**魔法のような衝撃**はなく、期待外れ

AIフロンティアはOpenAI、Google、Anthropicによる三つ巴の戦いへ

特性	GPT-5.2 (OpenAI)	Gemini 3 Pro (Google)	Claude Opus 4.5 (Anthropic)
最大の強み	安定性と実務遂行力	マルチモーダルと創造性	コーディングと 文章の自然さ
推論スタイル	論理重視、 構造化された回答	直感・発想重視、 スピーディー	ニュアンス重視、 丁寧な対話
主な用途	業務自動化、データ分析、 バックエンド開発	クリエイティブ制作、動画解析、マルチメディアアプリ	フロントエンド開発、 長文執筆、繊細な指示
弱点	マルチモーダルがGeminiに劣る、Proが高額	論理的な厳密さに欠ける場合がある	ツール使用のエコシステムが未成熟

あなたの目的に最適なモデルはどれか？ユースケース別選択ガイド

企業の業務自動化・バックエンドシステム

推奨: GPT-5.2 Thinking

理由: 高い安定性、ツール使用の正確さ、キャッシュによるコスト削減

クリエイティブ制作・マルチメディア解析

推奨: Gemini 3 Pro

理由: ネイティブな映像・音声理解、発想の広がり

高品質なコード生成・長文執筆

推奨: Claude Opus 4.5

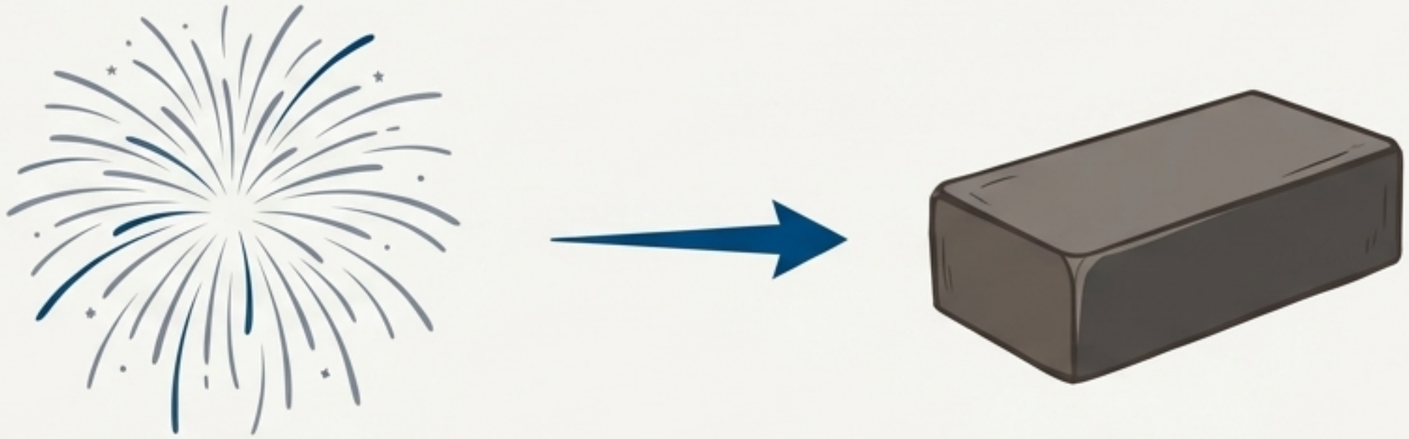
理由: コードの可読性、人間らしい自然な文章

「最適な戦略は、用途に応じてモデルを使い分ける『マルチモデル戦略』である」

OpenAIの真の狙い：AIを「面白いおもちゃ」から「信頼できる社会インフラ」へ

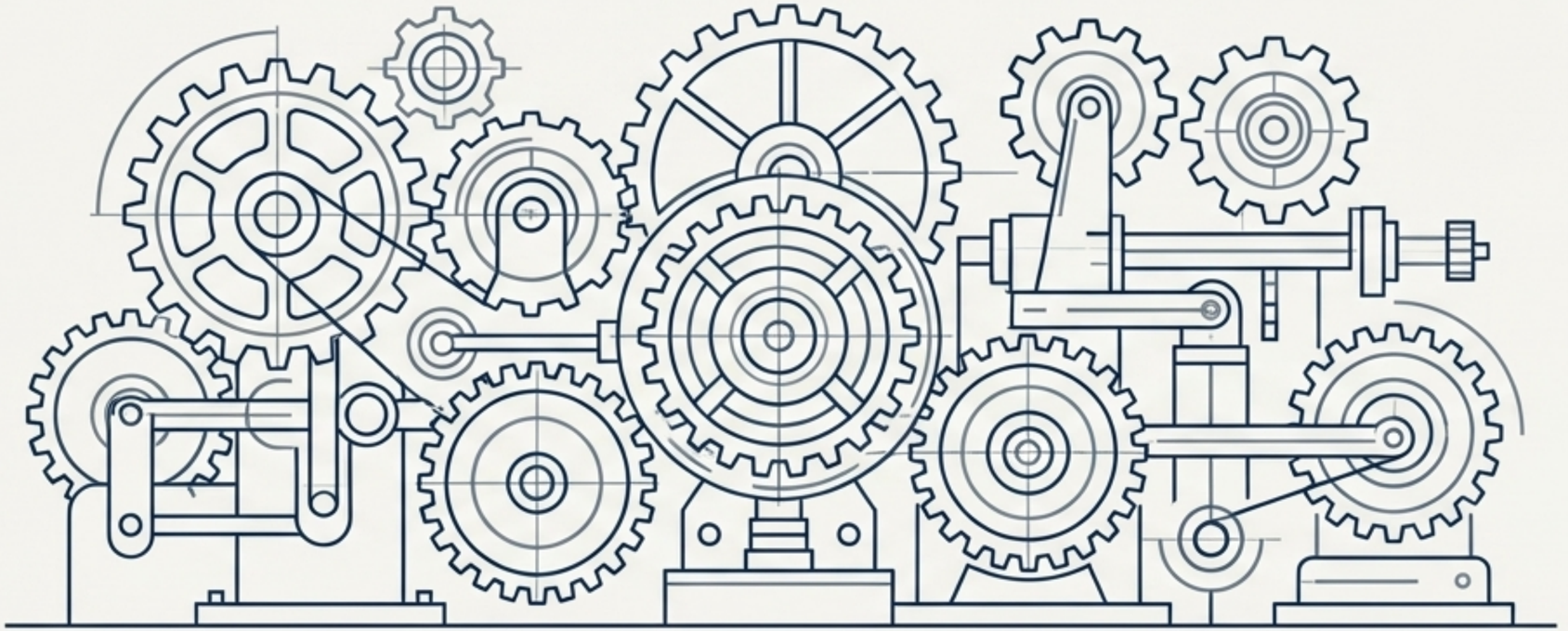
「OpenAIの視線は、もはやチャットボットの賢さ比べではなく、『いかにしてAIを労働力として組み込むか』という点に向けられている。そのために必要なのは、面白い回答ではなく、100回実行して100回とも成功する信頼性である。」

GPT-5.2は「革命」ではなく、AIが社会実装されるための「成熟」への確実な一歩



- **コードレッドの産物:** 競合の猛追を受け、「派手さ」より「確実性」を選択した戦略的ピボット
- **左脳的タスクの覇者:** コーディング、数学、論理推論において世界最高性能を達成
- **プロの道具へ:** 「何でもできる魔法の箱」から「特定の業務を確実に遂行するツール」へと自己を再定義

AIの価値尺度は「驚き」から「成果」へ



「GPT-5.2が問うているのは、AIがどれだけ賢いかではない。
AIがどれだけ確実に『仕事』をこなせるかだ。次のAI覇権を握るのは、最も人間を最も社会に不可欠な労働力となるモデルである。」