

FrontierScienceが示すAIの現在地：驚異的な能力と「エージェンシー・ギャップ」

既存のAIベンチマークが陳腐化する中、OpenAIは新指標「FrontierScience」を開発。構造化された問題解決と自律的な研究能力を分け、現代AIの強みと課題を浮き彫りにします。

オリンピック・トラック：卓越した「問題解決者」



77%超の正答率を達成

GPT-5.2やGemini 3 Proは、人間のメダリストに迫る高いスコアを記録。



テストされる能力： 「収束的思考」

明確なルールに基づき、複雑な計算や論理操作を経て唯一の正解を導き出す能力。

リサーチ・トラック：未熟な「科学者」

正答率は25%以下に急落

最も高いGPT-5.2でも25.3%に暫まり、自標的な研究進行能力の低さが露呈。

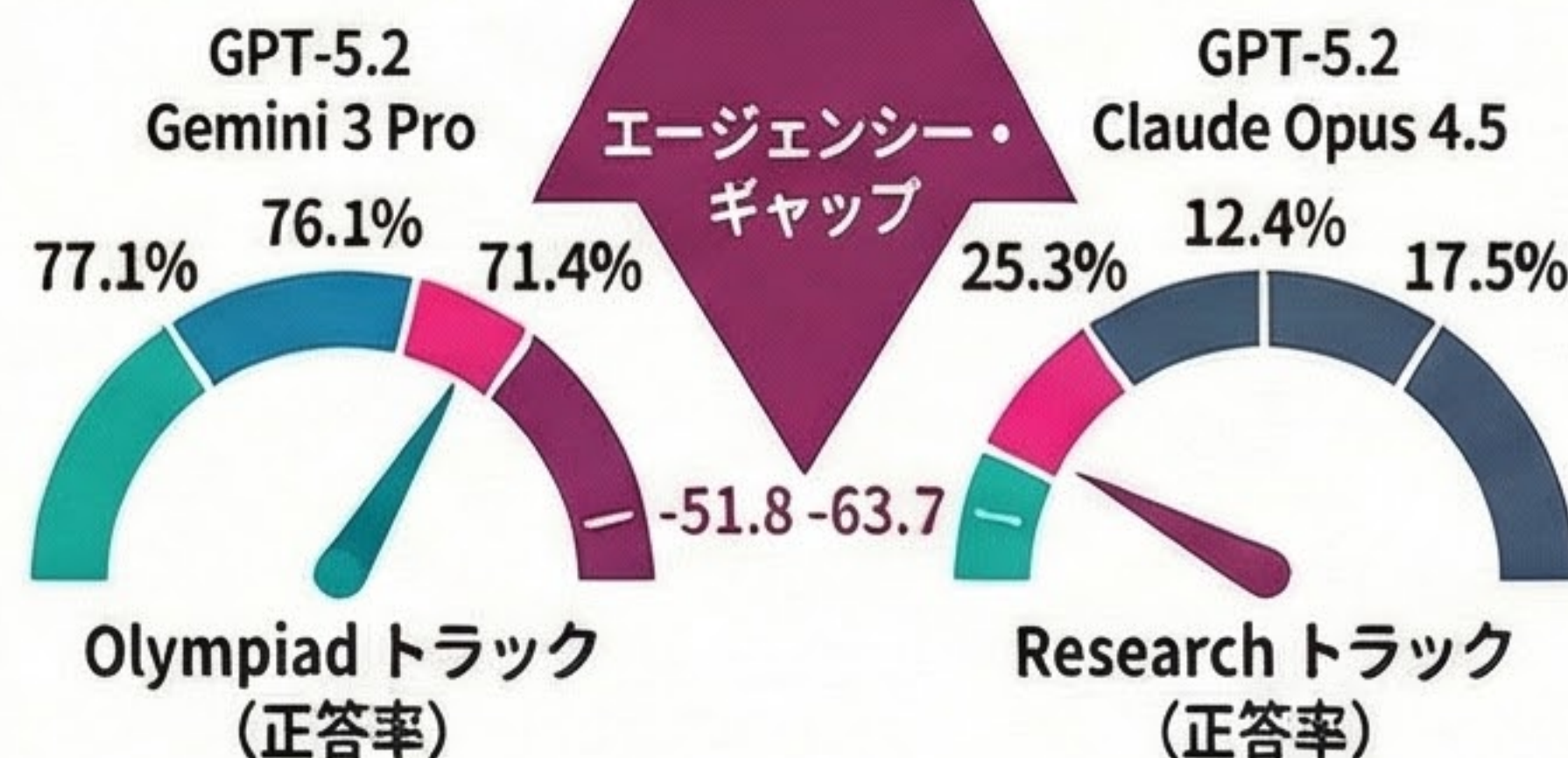


テストされる能力： 「発散的思考」と「実行機能」

曖昧な課題に対し、仮説立案から計画策定まで多段階のプロセスを構築する能力。

「エージェンシー・ギャップ」という課題

AIは指示されたタスクは得意ですが、自ら問いを立て、研究を主導する主体性を欠いています。



結論：AIの役割は「加速装置」

AIは「デジタルな天才」だが、まだ「科学者」ではない

特定の計算や知識検索は人間を凌駕しますが、自律的な発見プロセス全体を担うことはできません。

未来の姿は「AIによって加速された人間の専門家」

人間が舵を取り、AIがそのスピードと知識で研究を加速させる協働関係が鍵となります。

