「Poetiq (Mix)」：人間の推論能力を超えた初のAIシステム

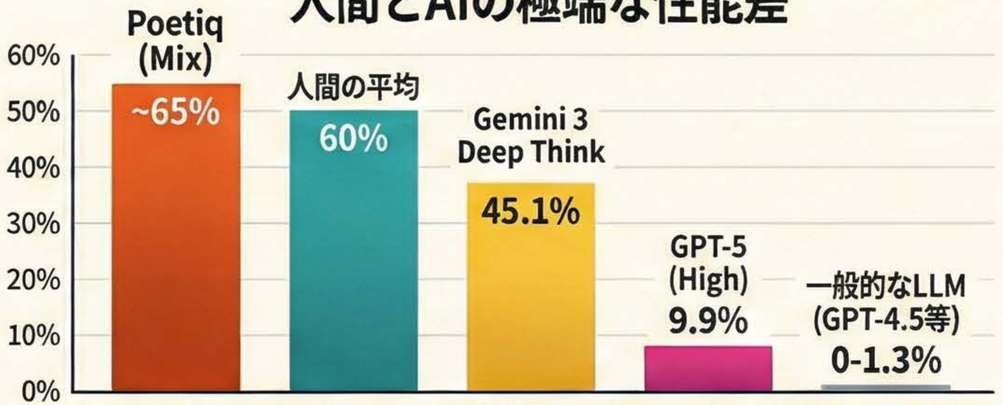
AIスタートアップPoetiqが開発した「Poetiq (Mix)」は、抽象推論能力を測るARC-AGI-2ベンチマークで、AIとして初めて人間の平均スコアを上回りました。この成果は、単一の巨大モデルではなく、複数の最先端AIを組み合わせる革新的な「メタシステム」によって達成されました。

課題：AIを阻む「推論の壁」

難解なベンチマーク「ARC-AGI-2」

未知のパズルを解く能力を試し、ほとんどのAIモデルは壊滅的なスコアを記録しました。

人間とAIの極端な性能差



人間が平均60%を達成する一方、単体のAIモデルの多くは1%前後のスコアでした。

解決策：Poetiqの革新的「メタシステム」

巨大モデルではなく「AIの連携」で解決

最新のGPT-5.1とGemini 3を統合し、それぞれの強みを活用します。

コストは人間の約3倍で性能は5ポイント上回る

タスクあたり約50ドルで、コスト効率の面でも大きな進歩を遂げました。

NotebookLM