

Poetiq AIの衝撃：難関ベンチマークで人間超えを達成

課題：AIを阻む「人間の壁」



最強AIでも「人間の壁」
は越えられなかった



人間の平均正解率 (約60%)

真の「推論力」を試す超難関
ベンチマーク「ARC-AGI-2」

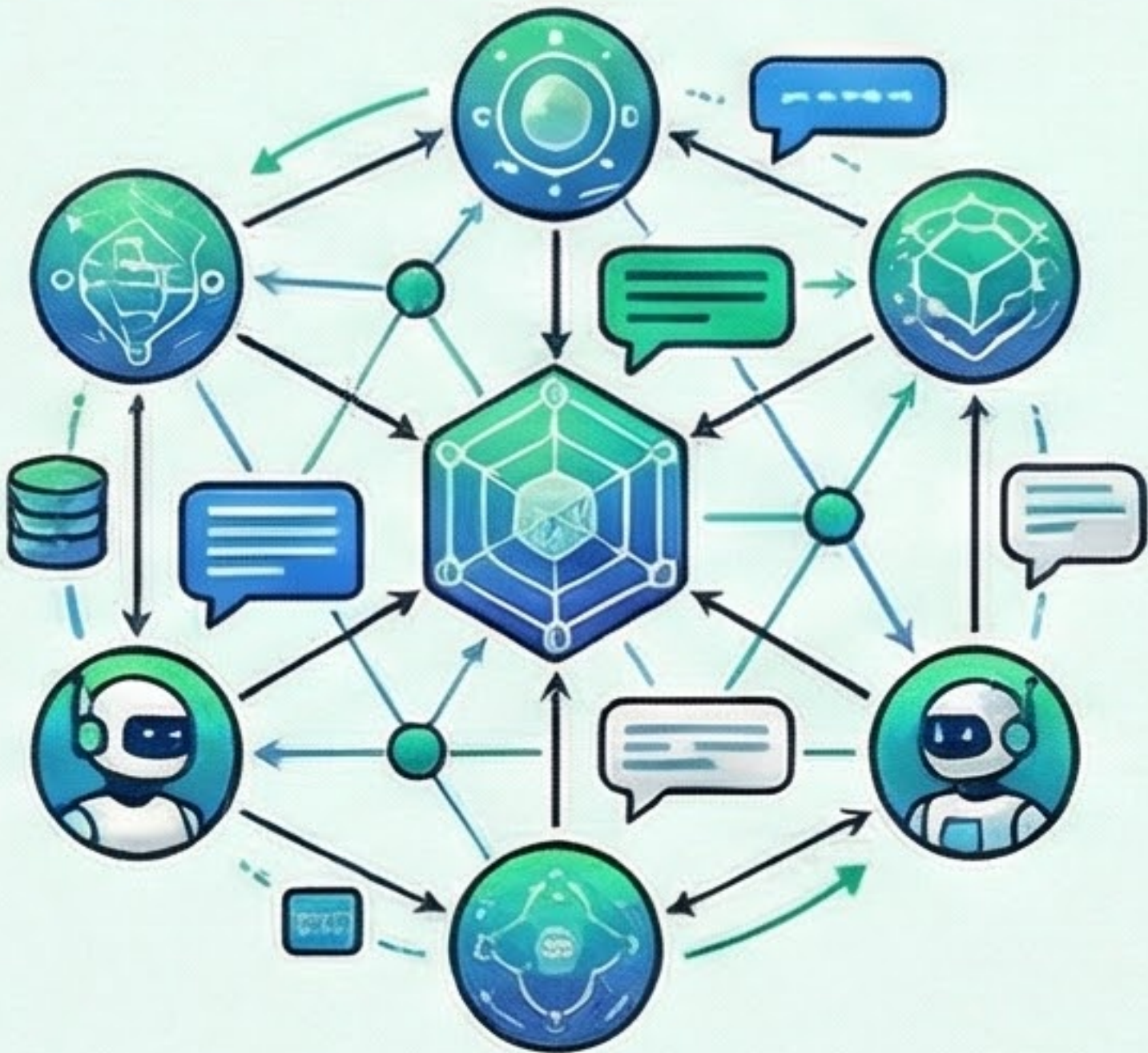
パターン記憶では解けない、初見の問題解決
能力が問われます。



トップモデルでも非常に
下も落とされない

解決策：Poetiq AIのブレイクスルー

Poetiqの秘密兵器：賢い「メタシステム」
複数のAIが協働し、対策と自己修正を繰り返して正解を導きます。



最新モデルGPT-5.2で
正解率75%を達成

人間の平均スコアを大幅に上回り、
歴史的な記録を樹立しました。

Poetiqのシステムが、精度とコスト効率の両面で従来の手法を圧倒

システム	正解率 (公開評価)	1問あたりのコスト
 Poetiq AI (GPT-S.2 X-High搭載)	75%	\$8未満
 Gemini 3 Deep Think (旧リーダー)	約46%	\$77.16
 人間の平均	約60%	N/A