2026年末のAGI誕生は現実的か？：サム・アルトマン発言の真意と検証

サム・アルトマン氏の「2026年末までにAGI」という発言の背景にある「定義の相違」と「技術的な実現可能性」を明らかにし、ニュースの真実性を正しく伝える。

二つの「AGI」：定義による達成確度の違い

アルトマン氏の「主観的AGI」

≠

OpenAI公式 (Charter) の「真のAGI」

アルトマン氏の「主観的AGI」

達成確度：中

2026年末までに、彼自身がAGIとみなす「内部システム」が成立する可能性。

OpenAI公式（Charter）の「真のAGI」

達成確度：低

ほとんどの経済的業務で人間を凌駕し、高度な自律性を備えたシステム。

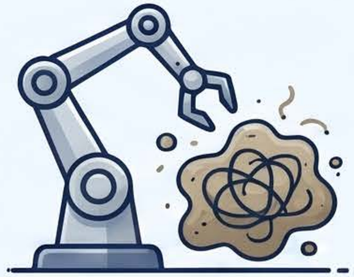
2026年末時点での目標達成の可能性比較

判定対象: アルトマン個人が呼ぶAGI	評価: 中	主な理由: 未公開モデルAstraの数学・研究能力が急進展しているため
判定対象: OpenAI Charter基準のAGI	評価: 低	主な理由: 汎用性・自律性・信頼性において、依然として大きな残差があるため
判定対象: 安全に一般展開できるAGI	評価: 低	主な理由: サイバー能力への懸念から、開発速度の制約（安全封じ込め）が必要なため

Anthropicは2026-27年を支持

GoogleやMetaは5-10年以上先と予測

実現を阻む「技術的ギャップ」と「安全性の壁」



自律性における顕著な弱点

最新のGPT-5.6でもAutomationBench（自律性）は**18.1%**の成功率に留まる。

科学的発見

現実業務の乖離

数学の未解決問題は解決したが、曖昧な要求が伴う現実世界の「汚いタスク」では性能が低下。

安全性を優先した開発の減速

サイバー攻撃能力の懸念やインシデントを受け、大規模な学習工程が保留・停止されている。