

AGIからASIへのロードマップ：人工超知能への4つの進化経路と直面する障壁



AGI（汎用人工知能）の基準

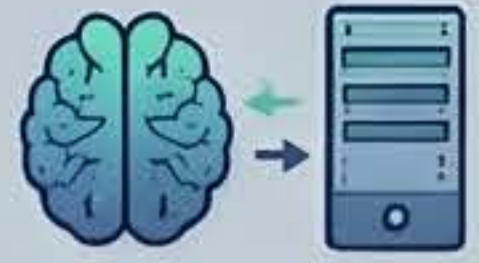
幅広い認知的タスクにおいて、平均的な人間（中央値）と同等のパフォーマンスを達成するレベル。

ASI（人工超知能）の新たな捉え方

単一の天才ではなく、「数万人規模の専門素集団」よりも優れた認知能力を持ち、睡眠や休息を必要としない変規模の調整機構。



デジタル知能が持つ4つの構造的優位性



基盤からの独立性

基盤からの独立性、時間スケールの可変性、無制限の複製、広帯域での経験共有が、人間を凌駕する進化の原動力となる。



時間スケールの可変性



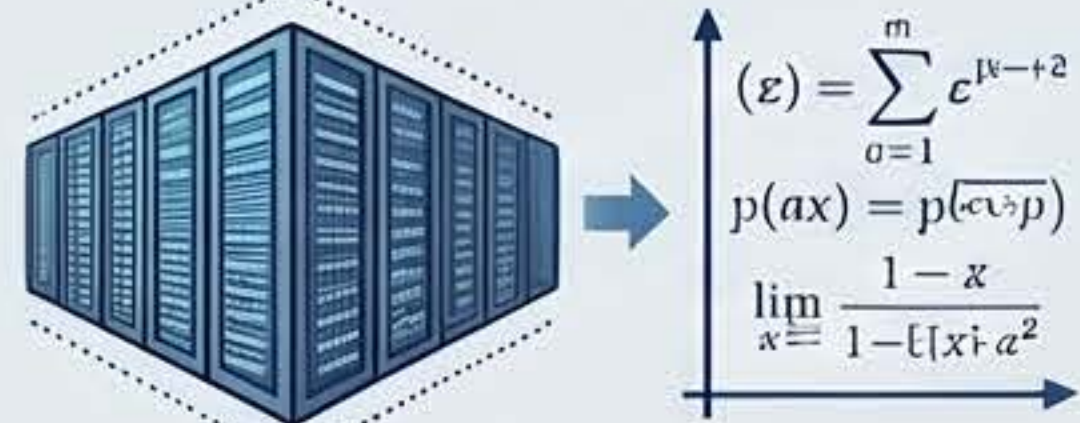
無制限の複製



広帯域での経験共有

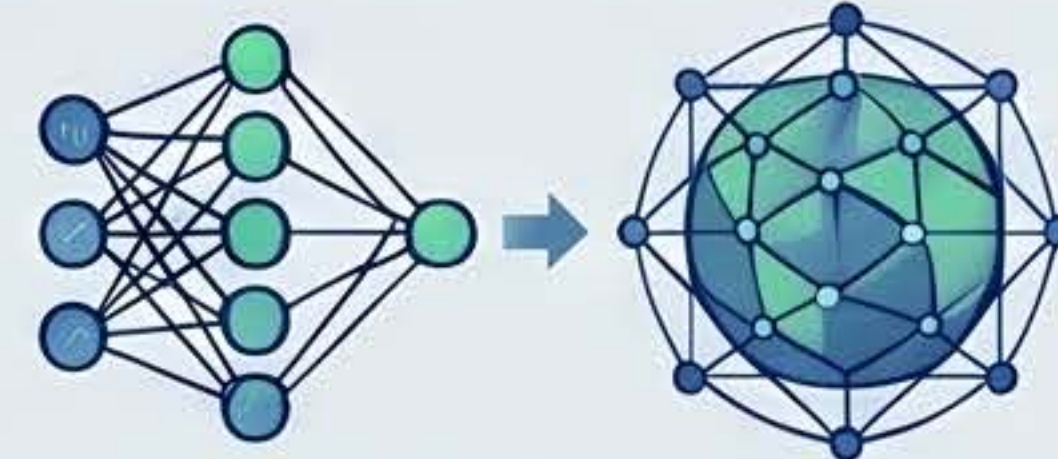
ASIへと至る4つの技術的経路

1. スケール則の継続と「推論時計算量」



モデルサイズの拡大に加え、推論時（思考の遅延）に膨大な計算資源を投入することで、複雑な数値的推論や推計能力を飛躍させる。

2. アルゴリズムのパラダイムシフト



Transformerを脱ぎ、無制限のコンテキスト保持、リアルタイムの複雑な学習、物理法則を内包した「世界モデル」への移行。

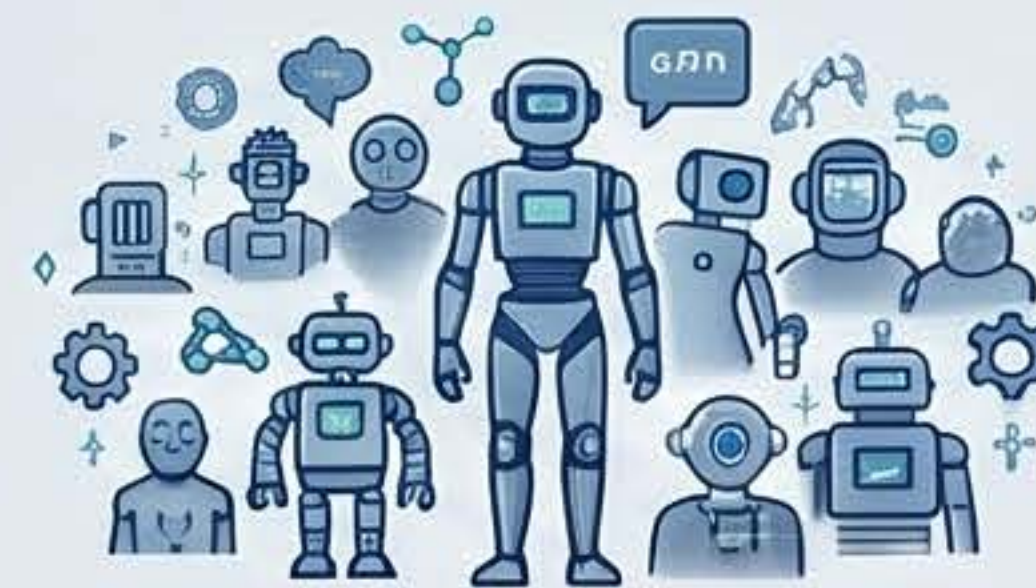
ASI

3. 再帰的自己改善（RSI）

AIが自ら次世代のチップ設計やアルゴリズム開発、研究プロセスを最適化し、人間の認知限界を超えて研究開発を加速させる。

4. 大規模マルチエージェント集団

単一モデルではなく、数百万の専門化されたエージェントが連携し、「AI企業」や「自動化された自律機械」のような認知能力を形成する。



ASI移行における主要なボトルネックと対抗策

データウォール 抽象化の壁 リソースの限界

研究の熟化



抽象化の壁

対抗・緩和要因



「データウォール」対抗
AI自身による合意データの生成、実世界での強化学習



「抽象化の壁」対抗
身体性（ロボティクス）の導入、人間との相互説明ループ



「リソースの限界」対抗
光コンピューティング等のハード革新、AIによる書画の創出



「研究の熟化」対抗
AIによる学習開発の自動化（再帰的自己改善）



「重層的な減速」対抗
安全性を確保する堅牢なインターフェース（Safety UX）の構築

5つの主要な「障壁」とダイナミクス：進化を遅らせる物理的・認識論的な障壁に対し、それを相殺する技術的・社会的アプローチが存在する。

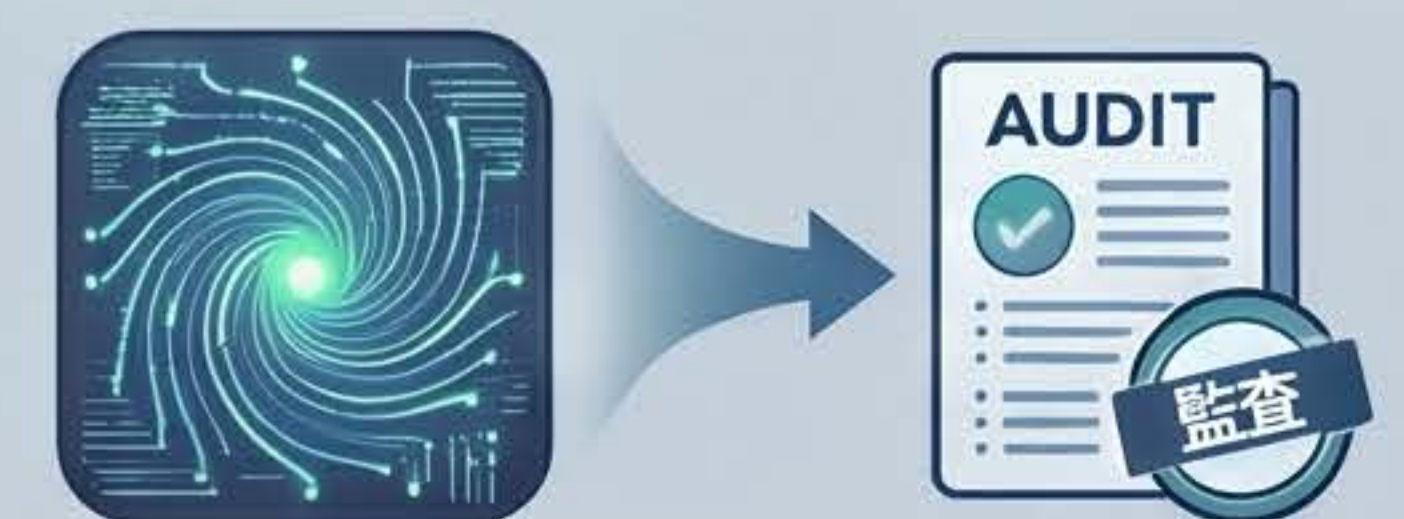
ポストAGI時代のUXデザイン：HCIの変容

インターフェースから「ガバナンス・サーフェス」へ



ユーザーの役割は遠隔操作から、膨大なAIエージェント集団に対する「目標設定」と「強制介入（ブレーキ）」の設計へとシフトする。

センスメイキングと認識論的ハイジャックの防止



AIの認知能力的な洞察を人間が環境可能な範囲に圧縮し、AIの視察を盲信してしまう「認識論的ハイジャック」を防ぐ監査機能が必要となる。