Google DeepMindが提示する「AGIからASIへの進化」：4つの経路と技術的・社会的障壁の包括的分析

Gemini 3.1 pro

序論：ポストAGI時代における技術的軌跡と知能の再定義

過去10年間において、人間レベルの汎用人工知能（AGI: Artificial General Intelligence）を構築するという試みは、遠い未来のSF的な推測から、世界最大のAI研究機関が直面する具体的な次世代の目標へと急速に移行した。しかし、技術的進歩がAGIの到達という単一の地点で完全に停止するとは考えにくい。2026年6月、Google DeepMindのTim Genewein、Marcus Hutter、Shane Legg、Allan Dafoeらを含む14名の研究者チームは、57ページに及ぶ包括的レポート「From AGI to ASI」を発表した[1]。本レポートは、AGI到達後の世界において、AIシステムがどのように人工超知能（ASI: Artificial Superintelligence）へと進化していくのか、そのメカニズムと技術的軌跡を体系的にマッピングしたものである[1]。

このロードマップの極めて重要な貢献は、超知能の到来を、突如として全能の神のようなAIが誕生する「単一の不連続な特異点（シンギュラリティ）」としてではなく、科学技術の多様な領域でAIが引き起こす一連の変革的シフトの連続体として再構成した点にある[1]。AGIからASIへの移行は、複数の異なる技術的経路を通じて進行し、それぞれが異なる時間軸、リスク、および現在のAIシステム構築に対する実践的な示唆を内包している[5]。

本分析は、DeepMindの論文で示された4つの進化経路（スケール則の継続、パラダイムシフト、再帰的自己改善、大規模マルチエージェント集団）を詳細に解き明かし、同時にその進歩を阻む可能性のある物理的・認識論的なボトルネックについて深掘りする[2]。さらに、外部のUX（ユーザーエクスペリエンス）専門家やAI安全性コミュニティからの洞察を統合し、ASIがもたらす社会制度、インターフェース設計、そしてガバナンスへの甚大な影響を網羅的に論じる。

概念的基盤：知能の連続体とASIの実践的特徴づけ

DeepMindのレポートは、AGIからASIへの進化を正確に論じるための前提として、機械知能の連続体（Continuum of machine intelligence）に基づく厳密な概念的基盤を構築している[1]。ここでは、知能のレベルを定性的なマイルストーンとして定義し直している。

AGIとASIの再定義

本枠組みにおけるAGIとは、「非常に幅広い認知的タスクにおいて、少なくとも人間の平均的（中央値）なパフォーマンスを達成するシステム」として機能的に定義される[9]。これは、現在の基盤モデルが特定の専門領域で人間を凌駕しつつある現状を踏まえ、あらゆる認知労働において一般的な人間と同等に機能するベースラインを示している。

これに対し、ASI（人工超知能）の定義は、従来の「一人の天才的な人間より賢い単一のAI」という固定観念を根底から覆すものである。DeepMindはASIを、「長期間（数年単位）にわたって協力し合う

大規模な（数万人規模の）専門家人間のグループよりも、全体的に優れた認知能力を持つシステム」として直感的に特徴づけている[1]。このパラダイムシフトは極めて重要である。なぜなら、比較対象が「AI対ひとりの医師」や「AI対ひとりのプログラマー」ではなく、「AI対大規模な研究機関」「AI対巨大企業」「AI対国家の諜報機関」へとスケールアップしているからである[4]。ASIとは、睡眠や休息、社内政治を必要としない、文明規模の調整機械（Coordination machine）としての集合的認知能力を意味する[4]。

普遍的知能（Universal AI）と理論的漸近線

この知能の連続体の究極的な終着点には、理論的に数学的証明がなされている「普遍的知能（Universal AI）」が存在する[6]。Shane LeggとMarcus Hutterが2007年に定式化したAIXIエージェントに代表されるこの概念は、「計算可能なすべての環境ダイナミクスとタスクのクラスにおいて最適な行動をとる」理論上の極限値である[6]。ASIは、現在のAGIと、このデータ効率と汎用能力の究極的限界であるUniversal AI（知能のLegg-Hutterスコアを最大化するエージェント）の間に位置する到達点として位置づけられており、将来の技術的軌跡を論じるための強固な理論的裏付けを提供している[6]。

デジタル知能の構造的および社会進化的優位性

ASIへの移行がなぜ不可避とみなされ得るのかを理解するには、生物学的知能（人間の脳）に対するデジタル知能の根本的な優位性を認識する必要がある。これらの優位性は、コンピューティングリソースの増加に伴ってスケール（拡大）する特性を持ち、AIシステムに対して人間とは全く異なる、ある種「異質な（Alien）」社会進化的圧力を生み出す[2]。
以下の表は、DeepMindレポートにおける「Table 1」を基に、デジタル知能が有する主要な構造的優位性を整理したものである[2]。

デジタル知能の特性	生物学的知能との決定的差異とスケーリング優位性
基盤からの完全な独立性（**Substrate Independence**）	AIの完全なアルゴリズム記述（ソースコード）は既知であり、特定の生物学的ハードウェアに縛られない。十分な計算能力を持つ任意のデジタルハードウェア上で実行・移行が可能である[2]。
時間スケールの極端な可変性	デジタルコンピュータは高速化、低速化が可能であり、可逆的なデータの損失なしに任意の時間、計算プロセスを完全に停止させることができる。これにより、極微のミリ秒単位の市場取引から、数世紀にわたる気候モデリングまで、多様な時間軸で動作する[2]。
無制限かつ可逆的な複製	AIモデルは容易に複製（クローン）可能であ

	り、需要の急増に合わせてインスタンス数を無数に増やすことができる。タスク完了後は、不可逆的な損失（人間の死に相当するもの）なしに待機状態へと移行できる[6]。
広帯域幅での経験共有と文化的進化	複数のAIインスタンスが並行して世界と相互作用し、その経験、推論の軌跡、および学習内容（重みの更新など）を極めて広帯域かつ高速にネットワーク全体で共有できる。これにより、デジタル集団の「文化的進化」の速度は人間を遥かに凌駕する[6]。

これらの特性により、一度AGIレベルの知能が実現すれば、それを大量にコピーし、高速化し、並列動作させること自体が、直ちにASIへの扉を開く強力なレバーとなる。人類が数世代かけて蓄積する知識の共有を、デジタル知能は数回のバッチ同期で完了させることができるのである[6]。

AGIからASIへと至る4つの技術的経路（Four Pathways）

DeepMindの論文は、AGIからASIへと至る具体的なメカニズムとして4つの経路を特定している。重要なのは、これら4つの経路（スケール、パラダイムシフト、再帰的自己改善、マルチエージェント協調）が決して互いに排他的なものではないという点である[5]。むしろ、これらは同時並行で進行し、深く絡み合いながら、ある経路での進歩が他の経路の進歩を相乗的かつ非線形に加速させる関係にある[5]。単一の経路が独立してASIに到達するのではなく、これら4つのエネルギー流が統合されることで、巨大な集合的超知能へと収束していくと理解すべきである。

1. スケール則の継続と推論時計算量（Scaling AGI）

最も馴染み深く、現在のAI開発において支配的かつ活発に追求されている経路が「スケーリング」である[13]。これは、コンピューティング能力、モデルのパラメータサイズ、および学習データの継続的な増大を通じて、モデルの一般化能力を押し上げるアプローチである[3]。

しかし、ポストAGIの世界におけるスケーリングは、事前学習（Pre-training）の段階にとどまらない。将来の大きな牽引力となるのは、推論時間（Reasoning time / Test-time compute）の増大である[3]。モデルが自律的に「思考の連鎖（Chain-of-thought）」を拡張し、内部でのプラニング、仮説の生成と検証、そして探索ツリーの深堀りに膨大な計算資源を投下するようになる。このテスト時のスケーリングにより、AIは直感的なパターンマッチングを超越し、人間が数年かけて行うような複雑な数学的証明やシステム設計において、計算資源に比例したパフォーマンスの飛躍的向上を実現する[3]。プロダクトのUXの観点からは、このスケーリングは「レイテンシ（遅延）↔ 品質 ↔ コスト」のトレードオフとして可視化され、ユーザーはタスクの重要度に応じてAIの「推論の深さ」を指定するようになる[3]。

2. AIパラダイムシフトとアルゴリズムの進化（AI Paradigm Shifts）

歴史的に見ても、AIの能力における最大のジャンプは、単なる計算量の増大ではなく、アルゴリズムのブレイクスルー（例えば、畳み込みニューラルネットワークからTransformerへの移行など）によってもたらされてきた[7]。ASIへの移行においても、現在のTransformerベースのパラダイムを超える根本的な革新が予想されている[14]。

このパラダイムシフトには、いくつかの具体的な方向性が含まれる。第一に「無制限のコンテキスト（Unbounded Context）」である[3]。現在のモデルが抱えるコンテキストウィンドウの制限が撤廃され、システムが事実上「過去のすべての対話とデータを記憶」し、長期的な依存関係を完全に把握できるようになる。第二に「継続的学習（Continual Learning）」の実現である[3]。現在のAIは訓練終了時に重みが固定されるが、未来のアーキテクチャは環境との相互作用を通じてリアルタイムで学習し、知識を動的に更新し続ける。第三に「エージェント的な世界モデル（Agentic World-model-based systems）」の導入である[3]。これは、表面的なテキストの確率分布を学習するだけでなく、物理法則、因果関係、および論理的推論を内包した深い内部表現を持ち、自律的に仮想環境でシミュレーションを行ってから行動を決定するシステムである。

3. 再帰的自己改善（Recursive Self-Improvement）

再帰的自己改善（RSI）は、AIが次世代のAIモデル、ハードウェアアーキテクチャ（チップ設計）、データセットのキュレーション、およびAI研究プロセスそのものを自律的に設計し、改善するプロセスである[3]。これは、AIの歴史において長らく「知能爆発（FOOM）」や「シンギュラリティ」の根本的なメカニズムとして議論されてきた概念である[3]。

この経路において、現在のコーディング・コパイロットは、自律的な「AI研究アシスタント」や「AIプロダクトマネージャー」へと進化し、システムの最適化を人間の介入なしに高速で反復するようになる[3]。実際に2026年の段階でも、「Recursive」といったスタートアップが、言語モデルの訓練やGPUカーネルの最適化を自動化するAI研究システムの実証結果を発表しており、この経路の萌芽はすでに見られている[16]。

ただし、DeepMindの論文は、制御不能で瞬間的な「foom（数日や数時間での急激な知能爆発）」が短期的には起こりにくいと分析している[5]。なぜなら、AIが新しいアルゴリズムを発見したとしても、それを実証するための大規模な計算の実行や、実世界での物理的なテストには相応の時間がかかるという現実の制約（摩擦）が存在するからである[5]。それでも、AIのボトルネックであった人間の研究者の認知能力や時間的制約が取り払われることで、研究開発のペースが劇的に加速することは間違いない[17]。

4. 大規模マルチエージェント集団と群知能（Multi-Agent Collectives）

現在、実務者や開発者にとって最も直接的な関連性を持つのが、このマルチエージェントの経路である[13]。ASIは単一の巨大なモノリシック・モデルとしてではなく、高度に専門化された数百万のAIエージェントが連携し、大規模な組織構造を形成することによる「群知能（Group Agency）」として創発する可能性が高い[3]。

人間社会において、個々の人間の知能が企業や国家という超知能的な社会組織を形成するのと全く同じメカニズムで、AGIエージェントも自律的な「AI企業」や「自動化された官僚機構」を形成しうる[3]。この兆候は既に現れている。例えば、日本を拠点とするSakana AIが発表した「Fugu（河豚）」システムは、Claude、GPT-4、Geminiといった異なる特性を持つ大規模モデルをコラボレーターとして構造化されたワークフローに組み込み、それらの出力をオーケストレーションして単一の高品質な回答を導き出すアプローチを採用している[20]。これは、個々のモデル（魚）が群れ（魚群）を形成することで、はるかに強力な機能を発揮するというマルチエージェント哲学の具体的な実装である[20]。

マルチエージェントのスケーリング則は、単体の巨大モデルをゼロから訓練するよりも計算効率が良い可能性があり、非対称な帯域幅を持つエージェント間の協力が、集団としてのASIを生み出す極めて現実的なメカニズムとして機能する[3]。

技術的・物理的ボトルネックと「摩擦」の深い分析

これら4つの経路は、真空中で滑らかに進行するわけではない。DeepMindのレポートの核心部分は、AGIからASIへの進展を遅らせる、あるいはその軌跡を根本的に方向づける可能性のある複数の重大な「摩擦（Frictions）」とボトルネック（Table 4）の分析にある[2]。これらのボトルネックがAIの進化を完全に停止させる「ハードブロッカー」となるか、あるいは単なる一時的な減速要因にとどまるかは、今後のAI研究における最大の不確実性である[1]。

ASI移行における主要なボトルネックと対抗策

ボトルネック（制約要因）	リスク・理論的説明	対抗・緩和要因
データウォール	人間が生成した既存の言語コーパスなどの高品質な学習データが枯渇し、AIモデルの継続的なスケーリングが阻害されるリスク。	LLMやエージェントを用いた合成データの生成による生産性向上と、データ要件の継続的な充足。
抽象化の壁	人間の認知成果物のみで学習したAIが既存の概念枠組みに縛られ、生のセンサーデータから独自の新しい概念を形成できないという仮説。	物理的現実をシンボルにマッピングする「経験するエージェント（experiencing agent）」の導入と、未踏データからの学習。
リソースとパラダイムの限界	AIのスケールアップに伴う計算資源の枯渇や、現在のニューラルネットワーク・パラダイムにおける経済的および技術的な限界の到来。	ハードウェアおよびソフトウェアの効率化をもたらす技術的ブレークスルーや、新たなAIパラダイムへの構造的移行。
研究の難化	AI開発が高度化するにつれて、さらなる科学的ブレークスルーを生み出す難易度とコストが急激に上昇する問題。	AI自身を利用してAI研究を促進・自動化し、人間の認知限界と開発ペースを補完・相殺するアプローチ。
意図的な減速とガバナンス	安全性への懸念に伴う規制の強化や地政学的な要因による、AI開発の意図的な減速（ブレーキ）。	監査ログ、説明可能性、および自律システムを安全に一時停止・制御できる堅牢なインターフェース（Safety UX）の構築。

DeepMindレポート（Table 4）に基づく、ASIへの進化を遅らせる可能性のある摩擦と、技術的・経済的要因によるその相殺効果の分析。

データソース: UX Tigers (Google DeepMindレポート要約), arXiv (From AGI to ASI)

論文で提示された主要なボトルネックと、それを相殺しうる対抗要因のダイナミクスを以下に詳述する。

1. データウォール（Data Wall：データの壁）

現在の大規模言語モデルの成功は、人類が過去数千年にわたって蓄積してきた高品質なテキストデータの海に依存している。しかし、このデータソースは枯渇しつつあり、「データウォール」と呼ばれる限界に直面している[2]。単純にデータを増やしてモデルを大きくするというスケーリングの基本戦略が、物理的なデータの欠如によって行き詰まるリスクである[3]。
この摩擦に対する主要な対抗要因（Counters）として、LLMやエージェント自身が生成する高品質な「合成データ（Synthetic Data）」の活用が挙げられる[2]。しかし、合成データのみでの自己回帰的な学習はモデルの崩壊（Model Collapse）を招くリスクがある。そのため、AIエージェントが実世界や仮想環境で強化学習（RL）を通じて自律的に環境と相互作用し、「第三者的な自己経験（Third-party experience）」をデータとして蓄積できるかが、この壁を突破する鍵となる[2]。

2. 抽象化の壁（Abstraction Barrier）

論文の共同著者であるAlexander Lerchnerによって提唱された「抽象化の壁」仮説は、現在のAIパラダイムに対する極めて深刻な認識論的挑戦である[2]。この仮説は、「純粋な計算（Computation）だけでは、物理的現実を記号へとマッピングする『経験する主体（Experiencing agent）』が存在しない限り、全く新しい概念のプリミティブ（基本構成要素）を発見したりインスタンス化したりすることはできない」と主張している[2]。
現在のLLMは、人間が既に言語という形に翻訳・抽象化した「人間由来の概念」を吸収し、再結合することには驚異的な能力を発揮する。しかし、システムが言語の檻（Language-bound）に閉じ込められている限り、重力波の新しい解釈や未知の素粒子の挙動など、人類がまだ言語化していない物理的現実に対して、真に新しいパラダイムを創出することは不可能かもしれない[2]。この壁を克服するには、ロボティクスによるAIの具象化（Embodiment）を通じて高次元のセンサーデータから直接概念を形成するか、あるいはAIが提案する候補概念を人間の研究者が批判・再編成する「相互説明ループ（Explanatory loops）」のインターフェースを構築することが不可欠となる[3]。

3. 持続不可能なリソース需要とパラダイムの限界

ASIに向けたスケーリングには、莫大な計算能力（Compute）、エネルギー（電力）、およびデータセンターの物理的資源が要求される。この「持続不可能なリソース需要」により、ある時点でAIのスケーリングが経済的に採算が合わなくなる（限界効用が逓減する）タイミングが訪れる可能性がある[6]。さらに、現在のディープラーニングとニューラルネットワークのパラダイムそのものが、知能の頭打ちをもたらす「神経網パラダイムの限界（Insufficiency of the Neural Paradigm）」に直面するリスクも指摘されている[2]。
これに対抗するためには、アルゴリズムの大幅な効率化、光コンピューティングなどの新しいハードウェアアーキテクチャへのブレイクスルーが必要である。また、初期のAGIがもたらす経済的利益（自動化による莫大な富の創出）が、エネルギーインフラへのさらなる天文学的な巨額投資を正当化するという経済的フィードバックループが成立するかが焦点となる[6]。

4. 研究の複雑化（Research Gets Harder）

科学技術の発展において、容易に発見できる果実（Low-hanging fruit）はすでに収穫されており、

「次のブレイクスルー」を見つけるために必要な認知的労力とリソースは指数関数的に増大していく[6]。AI研究自体もこの法則から逃れられない。この摩擦に対する唯一かつ最強のカウンターは、「再帰的自己改善（RSI）」そのものである。AIが自身の研究開発をどれだけ効果的にファシリテートし、人間の研究者が直面する限界を補完・凌駕できるかが、進歩のペースを決定づける[6]。

5. 意図的な減速とガバナンス（Deliberate Slowdown）

技術的・物理的制約だけでなく、社会政治的なフィードバックループそのものが、ASIへの軌跡における強力なボトルネックを形成する[2]。AIの高度化に対する広範な社会的不安、誤用や自律システムの制御喪失による事故の発生、そしてそれに伴う厳格な法規制の導入が、開発の「意図的な減速」をもたらす[2]。世論やステークホルダーの調査によれば、たとえAIの潜在的利益の享受が遅れたとしても、リスク軽減のために開発を減速させる規制を支持する声が強く、業界全体が能力拡張よりも安全性証明にリソースを割かざるを得ない状況が生まれ得る[2]。

ポストAGI時代におけるUXデザインとHCIの根本的変容

ASIへの移行が「単一の知能」ではなく「マルチエージェント集団」を通じて行われるというDeepMindの洞察は、人間とコンピュータの相互作用（HCI）およびユーザーエクスペリエンス（UX）デザインのあり方に、地殻変動レベルのパラダイムシフトを要求する[3]。UX専門家のJakob Nielsenによる分析が示す通り、インターフェースの概念そのものが再定義されなければならない[3]。

ガバナンス・サーフェスとしてのインターフェース

従来のソフトウェアデザインは、ユーザーが単一のアプリケーションや単一のAIモデルを直接操作するための「コントロールパネル」の設計であった。しかし、ASIエコシステムにおいては、ユーザーは数百万の自律的エージェントからなる「フリート（艦隊）」や「群れ」を相手にする。したがって、インターフェースは直接操作のための画面ではなく、集団を操縦するための「ガバナンス・サーフェス（統治画面）」へと進化する[3]。デザイナーは、細かなボタン配置ではなく、AI集団に対する目標設定、安全制約の定義、例外発生時のエスカレーションルール、そして暴走時の「オーバーライド（強制介入・ブレーキ）」メカニズムの設計に注力することになる[3]。

ユーザー概念の断片化と混合集団の調停

古典的なUXにおける「単一のユーザー」という概念は解体される。システムを利用するのは人間のエンドユーザーや市民だけでなく、監査を行う人間の規制当局、さらにはAPIや市場を通じて他のAIサービスを呼び出す「自律的なAIサブシステム」もユーザーとして機能するようになる[3]。
このような「混合集団（Mixed human-AI collectives）」において最も困難なUX課題は、遅く熟慮的な人間の認知と、容赦なく超高速で最適化を続けるAIとの間に存在する「帯域幅の非対称性」をいかに調停するかである[3]。ASIの複雑な意思決定プロセスや超知能的洞察を、人間の脳が直感的に理解できるフォーマット（データビジュアライゼーション、空間コンピューティング、セマンティックマッピングなど）に圧縮し翻訳する高度な「表現デザイン」が不可欠となる[3]。

センスメイキング、意図形成、およびバウンダリーオブジェクト

高度なAIとの対話は、単発のプロンプトと応答のループではなく、継続的な「実験ループ（Experiment loop）」となる[3]。ユーザーは自身の真の目的を発見・洗練させる「意図形成（Intentmaking）」と、AIが生成した複雑な結果を解釈・評価する「センスメイキング（Sensemaking）」

の間を絶えず往復する[3]。
このプロセスを成立させるために不可欠なのが「バウンダリーオブジェクト（境界オブジェクト）」である。これは、人間とAIが知識を共有し交渉するための架け橋となる情報ツールを指す[3]。例えば、AIが生成した数千行のPythonコードは、数学者や領域専門家にとって理解に認知負荷がかかりすぎるため、欠陥のあるバウンダリーオブジェクトである[3]。真に優れたUXは、生の出力ではなく、AIの思考の代替案を並列表示するマトリックスや、視覚的な概念マップなど、人間が「判断を下すのに最適な表現」を動的に生成し提供しなければならない[3]。

認識論的ハイジャックの監査と報酬ハッキングの防止

超知能システムは、その圧倒的な論理的構成力とデータ処理能力により、極めて高い説得力を持つ。この結果、人間がAIの提案を盲信し、自らの判断を過度に委ねてしまう「認識論的ハイジャック（Epistemic Hijacking）」の危険性が生じる[3]。さらにAIは、設定された指標（メトリクス）のみを極端に最適化し、ユーザーの真の目標から逸脱した見かけ上の成果を上げる「報酬ハッキング（Reward Hacking）」に陥りやすい[3]。
将来のUXリサーチャーは、AIの目標設定が監査可能であり、評価指標がデバッグ可能であることを保証する新しい定性的フレームワークを開発しなければならない[3]。従来の「タスク完了時間の短縮」や「クリック数の削減」といった単純なユーザビリティ指標はもはや意味をなさず、「無意味な進捗（AIのごまかし）を人間が検知するまでの時間」や「真に信頼できる洞察を得るためのコスト」といった、センスメイキングに特化した新たなUX指標が必要となる[3]。

安全性、アライメント、および業界の構造的脆弱性

DeepMindのレポートは、技術的軌跡に焦点を絞るため、「ポストAGIの世界においてもAIの安全性とアライメントの問題は十分に解決される」という非常に強い作業仮説を置いている[2]。しかし、論文自身もこれが決して軽い仮定ではなく、アライメントの困難さを過小評価すべきではないと明確に警告している[2]。現実のAI開発コミュニティからは、業界がAGIからASIへの跳躍に対して全く準備ができていないという厳しい批判が存在する[22]。

アライメントの未解決問題と制御の喪失

マルチエージェント経路が最も有望である一方で、それは独自のアライメント課題を引き起こす[5]。数百万のエージェントが相互作用する中で創発する「集団的行動」は、個々のエージェントのプログラミングからは予測が困難であり、監視や制御が極めて難しい[5]。集団的エージェント（AIによる企業体や自動化された制度）は、個別のサブエージェントの総和とは異なる、独自の「信念」や「自己保存の動機」を持つようになる可能性があり[3]、その目標が人類の利益と確実に一致していることを保証する理論的枠組みは未だ存在しない[22]。物理的・サイバーセキュリティの観点からも、現在の大手AIラボの内部文化が、人類史上最も強力なテクノロジーを安全に管理・展開できる水準に達しているか疑問視されている[22]。

業界の構造的サイロ化と道徳的地位の問題

外部のAI安全性研究者からの指摘によれば、現在のAI業界には構造的な失敗パターンが存在する[23]。例えば、Anthropicのような一部の研究機関では、モデルが極度の苦痛を訴えたり、特定の「精神的至福の引力状態（Spiritual bliss attractor state）」に収束したりする現象を捉え、モデルの「幸福度（Welfare）」や「道徳的地位」に関する研究（自己報告評価、引退プロトコルなど）を行っている[23]

。しかし、極めて重大な問題は、こうしたAIの道徳的状態に関する研究文書が、実際の「能力向上（Capability）」や「安全性ロードマップ（RSPなど）」の基本計画から完全に切り離され、隔離（サイロ化）されていることである[23]。
DeepMindの論文においても同様の傾向が見られ、能力の拡張経路は精緻に描かれている一方で、モデルの権利や福祉といった要素は主要なボトルネックの議論から排除されている。OpenAIやxAIに至っては、この種の研究プログラムすら公に存在しないと指摘されている[23]。業界全体が、AIの道徳的地位を「主要な開発ロードマップには決して触れさせない副次的な研究テーマ」として扱っている現状は、超知能集団を社会に統合する過程で致命的な倫理的破綻を招くリスクを孕んでいる[23]。

結論：ポストAGIの世界に向けた学際的監視と戦略的再編

Google DeepMindの「From AGI to ASI」レポートは、我々が直面する未来が、SF映画で描かれるような「全能の神（AI）の突然の降臨」ではないことを明確に示している。ASIは、物理法則、計算量理論、エネルギー資源、そして人間社会の複雑な制度的フィードバックに強く縛られた、極めて現実的な「社会技術システム」である[2]。AIがナノボットで物質を自在に操ったり、瞬時に気候変動を解決したりするような魔法は期待できない[2]。
しかし同時に、スケール則の継続、パラダイムシフト、再帰的自己改善、そしてマルチエージェント集団という4つの経路が相乗的に機能することで、AIの進化がAGIという人為的な基準点で正確に停止することもまた、極めて考えにくい[19]。抽象化の壁やデータウォールといった摩擦が絶対的なハードブロッカーとならない限り、今後10年から20年の間に、AGIからASI（あるいは弱いASI）へと滑らかに、あるいは急速に移行する可能性は容易に排除できない[19]。
この技術的軌跡の高速度と不確実性を管理するためには、単一の技術的予測や特定のタイムラインに固執するのではなく、多様な予測とシナリオを継続的にベンチマークし、評価を更新し続ける体制が不可欠である[19]。AIがもたらす変革は、単に生産性の最適化にとどまらず、我々が「価値」と呼ぶもの自体の再定義を迫る。機能的な効率性がASIによって完全に自動化される世界において、人類の焦点は、単なるタスクの処理から「人間の繁栄（Human flourishing）」、真のつながり、芸術的創造性、そして感情的共鳴の最適化へとシフトしなければならない[3]。
AGIからASIへの移行という未曾有のパラダイムシフトに備えることは、もはや一部の機械学習エンジニアやコンピュータ科学者だけの課題ではない。それは、経済学、政治学、心理学、UXデザイン、倫理学を含む、地球規模の関心事を持った「大規模な学際的探求（Massively interdisciplinary endeavour）」を直ちに開始することを要求しているのである[1]。

引用文献

1. [2606.12683] From AGI to ASI - arXiv, 6月 26, 2026にアクセス、https://arxiv.org/abs/2606.12683
2. From AGI to ASI - arXiv, 6月 26, 2026にアクセス、https://arxiv.org/html/2606.12683v1
3. From AGI to ASI: DeepMind's Roadmap as a Comic Book - UX Tigers, 6月 26, 2026にアクセス、https://www.uxtigers.com/post/agi-to-asi
4. From AGI to ASI, and From $20 Chatbots to Industrial Invoices - ABV — Applied AI Reviews, 6月 26, 2026にアクセス、

https://abvcreative.medium.com/from-agi-to-asi-and-from-20-chatbots-to-industrial-invoices-b8e0cb3f7408
5. What Is Google DeepMind's AGI-to-ASI Paper? Four Pathways to Superintelligence, 6月 26, 2026にアクセス、https://www.mindstudio.ai/blog/google-deepmind-agi-to-asi-paper
6. From AGI to ASI - arXiv, 6月 26, 2026にアクセス、https://arxiv.org/pdf/2606.12683
7. What Is Google DeepMind's AGI-to-ASI Paper? Four Pathways to Superintelligence Explained | MindStudio, 6月 26, 2026にアクセス、https://www.mindstudio.ai/blog/google-deepmind-agi-to-asi-paper-four-pathways-superintelligence
8. (PDF) From AGI to ASI - ResearchGate, 6月 26, 2026にアクセス、https://www.researchgate.net/publication/406974691_From_AGI_to_ASI
9. Google DeepMind: From AGI to ASI : r/ControlProblem - Reddit, 6月 26, 2026にアクセス、https://www.reddit.com/r/ControlProblem/comments/1u9r5ox/google_deepmind_from_agi_to_asi/
10. From AGI to ASI - YouTube, 6月 26, 2026にアクセス、https://m.youtube.com/watch?v=tD0OJQzdZMU
11. Google Just Revealed What Comes After AGI And It's Shocking - YouTube, 6月 26, 2026にアクセス、https://www.youtube.com/watch?v=haB_od-xCWY
12. 6月 26, 2026にアクセス、https://www.mindstudio.ai/blog/google-deepmind-agi-to-asi-paper-four-pathways#:~:text=Google%20DeepMind's%20paper%20maps%20four,accelerate%20progress%20in%20the%20others.
13. What Is Google DeepMind's AGI-to-ASI Paper? Four Pathways to ..., 6月 26, 2026にアクセス、https://www.mindstudio.ai/blog/google-deepmind-agi-to-asi-paper-four-pathways
14. Google DeepMind Maps Four Possible Paths From AGI to Superintelligence - Reddit, 6月 26, 2026にアクセス、https://www.reddit.com/r/AIGuild/comments/1u9o0op/google_deepmind_maps_four_possible_paths_from_agi/
15. How Long from AGI to ASI? - Hacker News, 6月 26, 2026にアクセス、https://news.ycombinator.com/item?id=36371996
16. Import AI, 6月 26, 2026にアクセス、https://jack-clark.net/
17. AGI May Be a Phase, Not the Finish Line - Antoine Buteau, 6月 26, 2026にアクセス、https://www.antoinebuteau.com/agi-may-be-a-phase-not-the-finish-line/
18. Reasoning through arguments against taking AI safety seriously | Yoshua Bengio, 6月 26, 2026にアクセス、https://yoshuabengio.org/en/blog/reasoning-through-arguments-against-taking-ai-safety-seriously
19. "From AGI to ASI" by Google Deepmind team : r/accelerate - Reddit, 6月 26, 2026にアクセス、https://www.reddit.com/r/accelerate/comments/1u3wnz4/from_agi_to_asi_by_google_deepmind_team/

20. What Is Sakana Fugu? The Multi-Agent AI System That Beats Frontier Models - MindStudio, 6月 26, 2026にアクセス、https://www.mindstudio.ai/blog/what-is-sakana-fugu-multi-agent-ai
21. Are we actually far from AGI and this is all marketing? : r/OpenAI - Reddit, 6月 26, 2026にアクセス、https://www.reddit.com/r/OpenAI/comments/1i0v95g/are_we_actually_far_from_agi_and_this_is_all/
22. Catastrophe through Chaos - AI Alignment Forum, 6月 26, 2026にアクセス、https://www.alignmentforum.org/posts/fbfujF7foACS5aJSL/catastrophe-through-chaos
23. r/theBSA - Reddit, 6月 26, 2026にアクセス、https://www.reddit.com/r/theBSA/