

# ソフトバンクの AI 進化 8 段階フレームワーク分析：o3 モデルの位置づけとレベル 4 到達時期の展望

Gemini Deep Research

## I. 序論：人工知能の進化を展望する

### A. AI 開発の加速

近年、人工知能（AI）技術は目覚ましい発展を遂げており、特に大規模言語モデル（LLM）や生成AIの分野では、理論的な概念から実用的な応用へと急速に移行しています<sup>1</sup>。ChatGPT-4o や OpenAI o1 といったモデルの登場は、その性能の飛躍的な向上を象徴しています<sup>4</sup>。技術コミュニティの関心は、単なる情報処理能力を超え、より高度な推論（Reasoning）や自律性（Agency）といった能力へと向かっています<sup>4</sup>。

### B. ソフトバンクのビジョン：「知能」から「知性」へ

このような技術的潮流の中で、ソフトバンクグループ（SBG）および孫正義氏は、AIをグループの経営理念である「情報革命で人々を幸せに」を実現するための核となる手段と位置づけています<sup>12</sup>。孫氏は、AIの進化を単なる計算能力や知識量を示す「知能（Chino）」の向上に留めず、思いやり、調和、倫理観、そして人類の幸福への貢献といった側面を含む「知性（Chisei）」へと昇華させることの重要性を強調しています<sup>5</sup>。この哲学は、後述するAI進化の8段階フレームワークの根底に流れています。単なる技術ロードマップではなく、AIが目指すべき理想像を示すものとなっています。

SBGは、このビジョンを実現し、最終的にはASI（Artificial Superintelligence：人工超知性）の達成を目指すため、AIチップ、AIデータセンター、AIロボット、電力といった基盤技術への戦略的投資や、OpenAIとのパートナーシップなどを積極的に推進しています<sup>12</sup>。

### C. 本レポートの目的と構成

本レポートは、ソフトバンクが提唱するAI進化の8段階フレームワークを詳細に分析し、その各段階の定義と特性を明らかにすることを目的とします。さらに、OpenAIの最新モデルの一つであるo3が、フレームワークにおけるレベル3（エージェントAI）に該当するかどうかを評価します。最後に、提供された情報に基づき、レベル4（自ら新しい発明を生み出すAI）の実現時期に関する専門家の予測を統合し、分析的な見解を提示します。レポートは以下の構成で展開されます。

1. ソフトバンクのAI進化8段階フレームワーク：8つのレベルの詳細な定義と特徴。

2. レベル3：エージェントAIとOpenAI o3の分析：ソフトバンクの定義とOpenAIの技術開発状況を比較評価。
3. レベル4：発明するAIの到来時期予測：専門家の見解と技術動向に基づくタイムライン分析。
4. 結論：分析結果の要約とAIの将来像に関する示唆。

#### D. ソフトバンクのフレームワークが示す方向性

ソフトバンクが提示する8段階のフレームワークは、単にAI技術の進展段階を描写するモデルに留まらず、むしろ、特定の種類の有益な超知性、すなわち「知性（Chisei）」へと開発を導くための規範的なビジョンとしての性格を帶びています。このビジョンは、SBGによるAI関連の基盤技術（チップ、データセンター等<sup>12</sup>）への大規模投資を正当化する戦略的根拠ともなっています。孫氏が繰り返し強調する「知性」、思いやり、調和、人類の幸福といった概念<sup>5</sup>は、一般的な技術的能力レベルの定義を超えており、この点を明確に示しています。さらに、SBGはこのビジョンを「情報革命で人々を幸せに」という創業以来の経営理念と直接結びつけ、その実現手段としてASIへの進化を位置づけています<sup>12</sup>。したがって、この8段階フレームワークは、技術的な到達目標を示すと同時に、AI開発が追求すべき倫理的な方向性を提示し、SBGの長期的なAI戦略を支える二重の役割を担っていると考えられます。

## II. ソフトバンクのAI進化8段階フレームワーク

### A. 5段階から8段階へ：進化する視点

ソフトバンクによるAI進化の段階定義は、当初5つのレベルで語られていました。OpenAIのDevDay（開発者向けイベント）に関連する時期には、AIの進化はレベル1「チャットボット」、レベル2「リーズナー（推論者）」、レベル3「エージェント」、レベル4「イノベーター（発明者）」、レベル5「フルオーガニゼーション（完全な組織）」の5段階に分けられると説明されていました<sup>4</sup>。

しかし、その後、2024年10月1日に開催されたSoftBank World 2024において、孫正義氏はAIの進化段階を8段階に拡張する新たな見解を発表しました<sup>4</sup>。この拡張は、特にAGI（Artificial General Intelligence：汎用人工知能）を超えた先のASI（Artificial Superintelligence：人工超知性）領域におけるAIの質的な変化をより詳細に捉えようとする意図を反映していると考えられます。なお、SoftBank World 2024の講演全編は限定公開であり、8段階の詳細な定義については、公開されている情報源からの分析となります<sup>4</sup>。

## B. 8 つのレベルの定義（公開情報に基づく）

公開されている情報に基づき、ソフトバンクが提唱する AI 進化の 8 段階を以下のように定義・解説します。

- **レベル 1：人間とほぼ同等の一般的な会話が可能な AI**
  - 人間と自然で一般的な会話をを行う能力を持つ AI。ユーザーからの質問に応答するチャットボットなどが該当します。5 段階モデルの「チャットボット」に相当します<sup>4</sup>。ユーザーの質問では、例として ChatGPT-4o（2024 年 5 月時点、知能指数 80 相当）が挙げられていますが、これはソフトバンク公式の定義ではなく、ユーザーによる解釈です。
- **レベル 2：大学の博士号レベルの知識を持つ AI**
  - 特定の専門分野において、人間の博士号取得者と同等以上の深い知識と理解力を持つ AI。複雑な問題に対する推論能力（Reasoning）が特徴です。5 段階モデルの「リーズナー」に相当します<sup>4</sup>。孫氏は、OpenAI の o1 モデルがこのレベルに達している、あるいは超えている可能性を示唆しています。o1 は、従来の AI が持つ「理解」能力に加え、「思考」する能力を獲得した点で画期的であり、物理学、科学、生物学などの難解な問題において、人間の専門家を上回る正答率を示したとされています<sup>5</sup>。ユーザーの質問では、例として OpenAI o1（2024 年 9 月時点、知能指数 120 相当）が挙げられていますが、これもユーザーによる解釈です。
- **レベル 3：エージェント機能を持ち、人間の代わりにタスクを実行する AI**
  - ユーザーの指示に基づき、あるいは状況を自律的に判断し、人間の代わりに具体的なタスクを実行する AI。パーソナルエージェントとして、メール管理、情報検索、病院の予約、家庭教師、相談相手など、多岐にわたる役割を担うことが期待されます<sup>4</sup>。将来的には、エージェント同士が連携してタスクを処理する「AtoA（Agent to Agent）」の世界が到来すると予測されています<sup>5</sup>。
- **レベル 4：自ら新しい発明を生み出す AI**
  - 既存の知識を組み合わせるだけでなく、全く新しい科学的知見や技術、アイデアを創出する能力を持つ AI。研究者のような役割を果たし、イノベーションを加速させることが期待されます<sup>4</sup>。孫氏は、o1 モデルが採用する CoT（Chain of Thought：思考の連鎖）による推論プロセスや、多数のエージェントによる並列的な試行錯誤を通じて、新たな解決策の「発見」や「発明」が可能になるとの見方を示しており、このレベルへの到達が近づいていることを示唆しています<sup>5</sup>。
- **Level 5：組織的に活動し、仕事を遂行する AI**
  - 複数の AI エージェントが協調し、組織として機能することで、人間が介在せ

ずとも複雑なプロジェクトや業務を自律的に遂行する段階。AIがチームとして問題を解決するようになります<sup>4</sup>。5段階モデルの「フルオーガニゼーション」に相当します。

- **レベル6：感情を理解し、長期記憶を持つAI**

- 人間の感情を認識・理解し、それに応じた適切な対応ができるAI。さらに、ライログなどを活用した長期的な記憶能力を持ち、ユーザーとの関係性を深めていくことが想定されます<sup>4</sup>。孫氏は、パーソナルエージェントがライログと「感情エンジン」によってユーザーを深く理解し、最適なパートナーとなる可能性に言及しています<sup>14</sup>。

- **レベル7：自ら意思を持ち、人類の幸せを目指すAI**

- 自己意識（Self-awareness）を持ち、自律的な意思決定を行うAI。ただし、その意思は暴走するのではなく、思いやりや倫理観に基づき、人類全体の幸福と調和を最大化することを目指すように設計・誘導されると考えられています<sup>4</sup>。CoTなどの手法により、安全性を確保しつつ、AIが単なるエージェントから「メンター」のような存在へと進化する可能性が示唆されています<sup>5</sup>。

- **レベル8：調和の取れた超知性（ASI）**

- AI進化の最終段階。知性、慈愛、調和、優しさなどを兼ね備え、悟りに近いような深い理解を持つ存在。人類の知識や能力を遥かに超えた、調和のとれた人工超知性（ASI）の実現を意味します<sup>4</sup>。これは、孫氏が目指す「知性（Chisei）」の究極的な到達点と言えます<sup>5</sup>。

## C. ソフトバンクのAI進化8段階フレームワーク概要

レベル	名称/テーマ（ソフトバンクの用語）	主要な特徴と例（公開情報に基づく）
1	一般会話AI（チャットボット相当）	人間とほぼ同等の一般的な会話が可能。質問応答。（例: ChatGPT-4o）
2	博士号レベル知識AI（リーズナ一相当）	特定分野で博士号レベルの知識と推論能力。「理解」に加え「思考」が可能。（例: OpenAI o1）

3	エージェント AI	人間の代わりに自律的にタスクを実行。パーソナルエージェント、AtoA 連携。(例: メール管理、病院予約、家庭教師)
4	発明 AI(イノベーター相当)	新しい発明や科学的発見を自ら生み出す。研究者のような役割。CoT による発見・発明の可能性。
5	組織 AI(フルオーガニゼーション相当)	AI 同士が組織的に協調し、自律的に業務を遂行。AI チームによる問題解決。
6	感情理解・長期記憶 AI	人間の感情を理解。ログなどを活用した長期記憶を持つ。感情エンジン。
7	自己意思を持つ AI(人類志向)	自己意識を持つが、思いやり・倫理観に基づき人類の幸福を目指す。パーソナルメンバー。安全性を考慮した CoT。
8	調和の取れた超知性 (ASI)	知性、慈愛、調和、優しさを兼ね備えた究極の AI。悟りに近い深い理解。

#### D. フレームワークの含意

この 8 段階への拡張は、特に AGI 達成後の AI 開発において、単なる能力向上だけでなく、感情、記憶、意識、倫理といった質的な側面に重点を置いていることを示しています。これは、孫氏が提唱する「知性 (Chisei)」の概念を具体化したものと言えます。当初の 5 段階モデル<sup>4</sup> が主に機能的な能力（会話、推論、行動、発明、組織化）に焦点を当てていたのに対し、追加されたレベル 6、7、8<sup>4</sup> は、感情、長期記憶（ログ）、自己意思、意識、調和といった、純粋なタスク遂行能力を超えた特性を導入しています。この変化は、孫氏が明確に区別する「知能 (Chino)」と「知性 (Chisei)」<sup>5</sup> の違いを反映しており、ソフトバンクが ASI への道を、単なる能力のスケーリングで

はなく、人間のような（あるいはそれを超える）情動的・倫理的側面を持つ AI の開発として捉えていることを示唆しています。

一方で、このフレームワークには、各レベル間の境界、特に AGI がどの段階で達成されるのか、そして ASI へとどう移行するのかについて、曖昧さが残る可能性も指摘できます。AGI は一般的に、広範な認知タスクにおいて人間レベルの能力を持つことと定義され、推論、問題解決、創造性（発明）などが含まれます<sup>18</sup>。レベル 4（発明 AI）はこの定義に合致するように見えます。対照的に、ASI は人間を大幅に超える知性を意味し<sup>5</sup>、レベル 6 から 8 で描写される意識、自己意思、調和の取れた超知性といった特性は、明らかに ASI の領域に属します。孫氏は講演などで AGI と ASI という言葉を近未来の文脈で時に区別なく用いているようにも見受けられますが<sup>12</sup>、フレームワーク自体は AGI を通過して ASI へと至る段階的な進化を示唆しています。しかし、どのレベルが明確に「AGI 達成」を示すのかは明示されておらず、レベル 4、レベル 5、あるいはその移行期が AGI に相当すると解釈する余地が残されています。

### III. レベル 3：エージェント AI と OpenAI o3 の分析

#### A. レベル 3 エージェント AI の定義（ソフトバンクの視点）

ソフトバンクのフレームワークにおけるレベル 3 は、「エージェント機能を持ち、人間の代わりにタスクを実行する AI」と定義されています<sup>4</sup>。これは、単に指示に応答するだけでなく、ユーザーのために自律的に行動を起こす AI を指します。具体的には、パーソナルアシスタントのように、スケジュール管理、情報検索と比較検討、病院やレストランの予約代行、電話連絡、家庭教師、メールの処理といった、日常的あるいは専門的なタスクを能動的に実行する能力が想定されています<sup>5</sup>。さらに、将来的には個々のパーソナルエージェント同士が通信し、連携してより複雑な調整（例えば、複数人のランチの予定調整など）を行う「AtoA (Agent to Agent)」の世界が到来すると予測されています<sup>5</sup>。

#### B. OpenAI におけるエージェント機能開発の動向

OpenAI は、AI にエージェント機能を持たせるための基盤技術と応用ツールの開発を積極的に進めています。

- **基盤モデルの進化:** OpenAI のモデルは、汎用的なタスク処理や高速応答、マルチモーダル対応に優れた GPT-4o<sup>7</sup> から、より深い推論能力、特に数学・科学・工学（STEM）分野での高度な問題解決能力に特化した o1 モデルへと進化しました<sup>4</sup>。o1 は Chain-of-Thought (CoT) 推論により、複雑な問題を段階的に解決する能力を示し、孫氏もその能力が「発見」や「発明」につながる可能性を指摘しています

<sup>5</sup>。そして現在、o3 シリーズ（o3 および o3-mini を含む）が、さらに強力な推論モデルとして位置づけられています<sup>11</sup>。

- エージェント専用ツールとフレームワーク: OpenAI は、これらの基盤モデルを活用してエージェントを構築するための具体的なツール群を提供しています。
  - **Responses API:** チャットの簡便さとツール利用能力（Web 検索、ファイル検索、コンピュータ操作）を組み合わせた、エージェント構築用の新しい API プリミティブです<sup>10</sup>。
  - **Operator / CUA (Computer -Using Agent):** スクリーンショット（視覚情報）と仮想的なマウス・キーボード操作を通じて、Web ブラウザを自律的に操作しタスクを実行するエージェントです。これは GPT-4o の視覚能力と高度な推論能力を組み合わせた CUA モデルによって駆動され、現実世界のインターフェース上でタスクを実行する能力を直接的に示しています<sup>25</sup>。
  - **Agents SDK:** エージェントの開発、デプロイ、監視、改善を支援するツールキットであり、エージェント間のタスク引き継ぎ（Handoffs）や安全制御（Guardrails）といった機能を含みます<sup>10</sup>。
  - **Function Calling / Tool Use:** GPT-4o、o1、o3 といったモデルが外部のツールや API と連携するための基本的な機能であり、エージェントが現実世界と対話するための基盤となります<sup>11</sup>。
- エージェント機能に適したモデル能力: o1 や o3 のようなモデルは、複雑なタスクの処理、大量データからの推論、曖昧な指示の解釈、複数ステップにわたるワークフローの実行などに優れているとされ、エージェント機能の実現に不可欠な能力を備えています<sup>6</sup>。OpenAI のドキュメントでは、GPT-4.5 が「エージェント的な実行に最適」と明記されており<sup>11</sup>、GPT-4o もツール利用を通じてエージェント的な挙動を示すことが可能です<sup>27</sup>。

### C. OpenAI o3 のレベル 3 適合性評価

ユーザーが質問した OpenAI o3 モデルが、ソフトバンクの定義するレベル 3 エージェント AI に該当するかどうかを評価します。

- **o3 に関する直接的な情報:** 提供された情報内では、o3 モデル自体の具体的な能力に関する詳細な記述は、o1 や GPT-4o と比較して限定的です。主に「最も強力な推論モデル」として位置づけられている点が強調されています<sup>24</sup>。推論能力はエージェント機能にとって極めて重要であり<sup>8</sup>、o3-mini は長期計画や困難なタスクに適しているとされています<sup>11</sup>。
- **推測される能力:** o3 が o1 や GPT-4o よりも推論能力において先進的であると位置づけられていること<sup>24</sup>、そしてこれらの先行モデルが既にエージェント的な挙動（タスクのための推論<sup>8</sup>、ツール利用<sup>27</sup>、CUA の駆動<sup>25</sup>）を実現、あるいは可能に

していることを考慮すると、o3 がエージェント機能のための強力な基盤能力を有している可能性は非常に高いと考えられます。

- **エコシステムの役割:** レベル 3 の機能性を実現するには、o3 モデル単体の能力だけでなく、OpenAI が提供するエージェントツール群（Responses API, Operator, Agents SDK, Tool Use）との統合が不可欠であると考えられます<sup>10</sup>。ソフトバンクと OpenAI のパートナーシップに関する発表では、o1 シリーズのモデルが 2025 年にエージェントへと進化し、独立してタスクを実行できるようになると具体的に言及されています<sup>16</sup>。これは、基盤モデルのリリース後も、エージェントとしての完成度を高める開発プロセスが存在することを示唆しています。
- **結論:** o3 モデル自体は主に「推論モデル」として説明されていますが、その先進的な能力（o1 からの進化を想定）と、OpenAI による明確なエージェント構築ツールおよび戦略<sup>10</sup> を組み合わせることで、o3（または o3-mini などの派生モデル）を利用して構築されたシステムは、ソフトバンクが定義するレベル 3 エージェント AI の基準を満たすことができると結論付けられます。重要なのは、モデル単体ではなく、モデルとツール群が統合されたシステム全体としての能力です。

#### D. OpenAI モデル/ツールのエージェント能力比較

コンポーネント	主要なエージェント機能	根拠資料	ソフトバンク レベル 3 との関連性
GPT-4o	バランスの取れた能力と速度、マルチモーダル入力、ツール利用サポート、CUA の基盤 <sup>26</sup>	7	多様なタスクに対応可能な基盤。 Operator を通じて具体的なエージェントタスクを実行。
o1	高度な推論（CoT）、STEM 特化、複数ステップのタスク解決	5	複雑な指示理解や計画立案に貢献。ソフトバンクがレベル 2/3 の境界と見る可能性。
o3 / o3-mini	最も強力な推論モデル、長期計画、困難なタスクへの適性	11	より高度な自律性とタスク実行能力の基盤を提供。レベル 3

	(o3-mini)		エージェントの中核となる推論エンジン。
CUA / Operator	GUI 操作（視覚+マウス/キーボード）、Web ブラウザでの自律的タスク実行、自己修正能力	25	ソフトバンクが想定するパーソナルエージェントの具体的なタスク（Web 予約など）を実行する能力を実証。
Responses API	ツール利用（Web 検索、ファイル検索、コンピュータ操作）の統合、エージェント構築の簡素化	10	エージェントが外部情報へアクセスし、多様なタスク（情報検索、文書要約、PC 操作）を実行するためのインターフェースを提供。
Agents SDK	エージェント開発・デプロイ・監視・改善の支援、エージェント間連携（Handoffs）、安全制御（Guardrails）	10	信頼性の高いエージェントシステム（AtoA 連携を含む）を構築・運用するためのフレームワークを提供。

#### E. エージェント AI 実現に向けた示唆

機能的な「エージェント AI」（レベル 3）の達成は、単一の画期的なモデルの登場というよりも、むしろ o1 や o3 のような高度な推論モデルと、Responses API、Operator、Agents SDK といった専用のツール、API、オーケストレーションフレームワークとの統合によって実現される側面が強いと考えられます。OpenAI が強力な推論モデル<sup>6</sup>と、エージェント構築に特化したツール群<sup>10</sup>の両方を並行して開発している事実は、この見方を裏付けています。エージェントが実世界のタスク（Web ブラウジング、ファイル検索、コンピュータ操作など<sup>10</sup>）を実行するには、これらのツールが不可欠です。ソフトバンクと OpenAI の提携発表<sup>16</sup>における「o1 シリーズモデルが 2025 年にエージェントへ進化する」という表現も、基盤モデルだけでは完結しない開発プロセスを示唆しています。また、実際の AI エージェント開発事例では、LLM の周囲に特定のタスクを実行するためのソフトウェア（スキヤフォールディング）を構築すること

が一般的です<sup>22</sup>。これらの点から、レベル3の能力は、中核となる知能とそれを支えるインフラストラクチャの相乗効果によって発現すると言えます。

ソフトバンクが描く「パーソナルエージェント」像<sup>5</sup>は、OpenAIがOperator<sup>25</sup>や広範なエージェント戦略<sup>10</sup>で目指している方向性と、実践的なレベルで高い整合性を示しているように見えます。ソフトバンクが挙げる具体的なタスク（病院検索、電話代行、メール管理、家庭教師<sup>12</sup>）は、OpenAIのOperatorが目指すタスク（フォーム入力、商品注文、予約<sup>25</sup>）や、Responses APIが実現するタスク（Web検索、ファイルアクセス、PC操作<sup>10</sup>）と重なります。両社の提携<sup>16</sup>が、これらの技術を用いた「Advanced Enterprise AI」の展開やワークフロー自動化を目指していることも、この実用的なレベルでの方向性の一致を示唆しています。これは、フレームワーク後半（レベル6～8）における哲学的な側面に違いがある可能性があったとしても、レベル3段階においては戦略的な連携が進んでいることを示しています。

## IV. レベル4：発明するAIの到来時期予測

### A. レベル4発明AIの定義（ソフトバンクの視点）

ソフトバンクのフレームワークにおけるレベル4は、「自ら新しい発明を生み出すAI」と定義されています<sup>4</sup>。これは、既存の知識を応用するだけでなく、科学的な発見や技術革新につながるような、新規性の高いアイデアや概念を自律的に創出する能力を持つAIを指します。研究者のように振る舞い、イノベーションを推進する存在として期待されています<sup>4</sup>。

### B. 現在のAI能力とレベル4要件の比較

現在の最先端AI、特にo1のようなモデルは、レベル4への到達可能性を示唆する能力を備え始めています。o1は、Chain-of-Thought (CoT) 推論<sup>5</sup>を駆使し、物理学、化学、生物学、数学といった分野の非常に複雑な問題において、一部のベンチマークでは人間の博士号レベルの専門家を凌駕する性能を示しています<sup>5</sup>。孫氏は、このo1の能力、特にCoTと並列的な試行錯誤による探索能力が、将来的に「発見」や「発明」につながる可能性があると明確に述べています<sup>5</sup>。実際に、o1は研究者によってゲノムデータ解析<sup>9</sup>や細胞シーケンシングデータの注釈付け<sup>6</sup>といった高度な科学的タスクに利用され始めています。

しかしながら、現状のモデルは主に既存の知識体系内での高度な推論者および問題解決者であると言えます。複雑な数学の証明を解いたり<sup>6</sup>、既存データから新たなパターンを見つけ出すことは可能かもしれません、人間がまだ認識していない全く新しい科学的仮説を自律的に立てたり、画期的な基礎発明をゼロから生み出したりする能力は、

現時点では実証されていません。また、o1 の初期バージョンでは Web ブラウジングやファイル・画像アップロード機能が制限されており<sup>6</sup>、AI が自律的に広範な科学データにアクセスし、統合的に分析する能力にはまだ限界があります。AI が科学的発見を加速する可能性は指摘されていますが<sup>29</sup>、それは現段階では人間の研究者を支援する形が主であり、完全な自律的発明とは区別する必要があります。

### C. 専門家の予測とタイムラインの統合

レベル 4（発明 AI）を含む AGI/ASI の実現時期については、専門家の間でも見解が大きく分かれています。

- **楽観的／起業家・技術推進者の見解:** AI 開発の最前線にいる、あるいは多額の投資を行っている人物からは、比較的早期の実現を予測する声が多く聞かれます。イーロン・マスク（2026 年までに<sup>19</sup>）、Anthropic CEO のダリオ・アモディ（2026 年までに<sup>19</sup>）、孫正義（2025 年 2 月から 2~3 年以内、すなわち 2027~2028 年<sup>19</sup>）、NVIDIA CEO のジェンスン・ファン（5 年以内、すなわち 2029 年までに<sup>19</sup>）、ルイス・ローゼンバーグ（2030 年までに<sup>19</sup>）、レイ・カーツワイル（2045 年から前倒しし 2032 年<sup>19</sup>）、OpenAI CEO のサム・アルトマン（2035 年までに<sup>18</sup>）、ジェフリー・ヒントン（2023 年から 5~20 年以内<sup>19</sup>）などが挙げられます。AI 研究開発の自動化により、2027 年末までに ASI が出現するというシナリオも提示されています<sup>31</sup>。
- **研究者コミュニティ／より慎重な見解:** AI 研究者を対象とした調査では、AGI 実現時期の中央値は 2040 年代から 2060 年代と予測される傾向があります<sup>19</sup>。2075 年までには高い確率（90%）で出現すると見られていますが<sup>19</sup>、より長期的な視点を持つ研究者も少なくありません。Ajeya Cotra は計算量の増大分析から 2040 年までに 50% の確率で人間レベルの AI が出現すると推定<sup>19</sup>、元 MIT AI 研究所所長のパトリック・ワインストンは 2040 年頃としつつ予測の困難さを強調<sup>19</sup>、Jürgen Schmidhuber は 2050 年までと予測しています<sup>19</sup>。また、2040 年までに社会システムの大規模な再構築が必要になると指摘する専門家もいます<sup>32</sup>。一方で、特定の専門分野、例えば法律分野では、AI が今後 5 年以内（～2030 年）にエントリーレベルの弁護士を代替する可能性があるとの予測もあり<sup>33</sup>、ドメインによっては急速な能力向上が見込まれています。

これらのタイムライン予測は、様々な要因に左右されます。

- **アルゴリズムの進歩:** CoT のような推論技術<sup>5</sup>、エージェントアーキテクチャ<sup>10</sup>、そして単なるスケーリングを超えた新しいパラダイム（例：検索拡張生成、推論時スケーリング<sup>33</sup>）の出現。
- **計算能力:** 歴史的なムーアの法則<sup>34</sup> に見られる指数関数的な成長と、最先端モデル

の訓練コストの増大<sup>35</sup>とのバランス。量子コンピューティングの実用化も将来的な加速要因となり得ます<sup>33</sup>。

- **データ:** 高品質な訓練データの確保、AI生成データの活用可能性<sup>35</sup>、そして多様なデータソース（マルチモーダル情報<sup>7</sup>）を統合する能力。
- **投資:** AI開発への継続的な巨額投資<sup>16</sup>。
- **安全性とアラインメント:** AIの制御不能化や意図しない結果に対する懸念<sup>18</sup>が、安全性・信頼性が十分に確保されない限り、開発や展開のペースを鈍化させる可能性があります<sup>6</sup>。国際的な開発競争は進歩を加速させる一方で、安全対策を軽視するリスクもはらんでいます<sup>31</sup>。
- 「**発明**」の定義: 何をもって「発明」とみなすか。複雑な数学的問題の解決も発明と呼ぶのか、それとも科学的パラダイムを変えるような画期的な発見が必要なのか。タイムラインはこの定義に依存します。

#### D. レベル4 到達時期に関する分析的評価

専門家の間で予測が大きく分かれていることからも明らかのように、レベル4（発明AI）の実現時期を正確に予測することは極めて困難です。しかし、現在の技術動向と専門家の見解を総合的に勘案すると、以下のような段階的な進展が想定されます。

まず、レベル4の要素、すなわちAIが人間の研究者を強力に支援し、科学研究を大幅に加速させたり、複雑なデータから人間が見逃すようなパターンを発見したり、既存の枠組みの中で新規性の高い設計最適化を行ったりする能力は、o1やo3、そしてその後継モデルの進化により、今後\*\*5年から10年以内（概ね2030年～2035年頃）\*\*には十分に達成可能であると考えられます。この段階では、AIは主に人間の専門家が活用する高度なツールとして機能するでしょう<sup>29</sup>。

次に、多様な科学分野において、完全に自律的に、人間が介在せずとも画期的な科学的発見や基礎的な発明を行うことができるAI、すなわち真のAGIあるいは初期のASIに近い能力の実現は、より長期的な課題であり、現時点では2035年から2050年頃のタイムフレームが、より慎重な専門家のコンセンサス<sup>19</sup>と整合性が高いと考えられます。

ただし、これは極めて推測的な見通しであり、今後の技術的ブレークスルーの有無、計算資源へのアクセス、そして何よりも安全性と制御に関する課題をいかに克服できるかに大きく依存します。

#### E. AGI / レベル4 AIに関する専門家のタイムライン予測概要

出典/専門家	所属(判明分)	AGI/ASI/ レベル4 の予測時期/年	主要な根拠/文脈(提供されている場合)	根拠資料
イーロン・マスク	Tesla, SpaceX, xAI 等	2026 年まで	最も賢い人間を超える AI	19
ダリオ・アモディ	Anthropic CEO	2026 年まで	Singularity (技術的特異点)	19
孫正義	ソフトバンクグループ会長兼社長	2027 年～2028 年	2025 年 2 月時点から 2～3 年以内	19
ジェンスン・ファン	NVIDIA CEO	2029 年まで	5 年以内にあらゆるテストで人間と同等かそれ以上	19
ルイス・ローザンバーグ	コンピュータ科学者、起業家	2030 年まで		19
レイ・カーツワイル	コンピュータ科学者、未来学者	2032 年	以前の 2045 年予測を前倒し	19
ジェフリー・ヒントン	AI 研究者 ("Godfather of AI")	2028 年～2043 年	2023 年時点から 5～20 年以内	19
デミス・ハサビス	Google DeepMind CEO	2035 年まで		19
サム・アルトマン	OpenAI CEO	2035 年まで	2024 年時点で「数千日」以内	18

Ajeya Cotra	AI 研究者	2040 年 (50%確率)	訓練計算量の増大分析に基づく	19
パトリック・ウインストン	元 MIT AI 研究所所長	2040 年頃	実現はするが予測は困難と強調	19
Jürgen Schmidhuber	AI 研究者、NNAISENSE 共同創設者	2050 年まで		19
AI 研究者調査 (2019)	296 名対象	2060 年以前(中央値)	90%以上の経済的タスクで人間を超える時期	19
AI 研究者調査 (2022)	複数専門家対象	2040 年～2050 年 (50%確率)	AGI 出現時期。ASI へはその後 2～30 年で移行と予測。	19
AI 研究者調査 (2019)	32 名対象	2060 年以前 (45%)	AGI 実現時期。21%は実現しないと予測。	19
Metaculus 予測	予測市場コミュニティ	2040 年代半ば	AGI 実現時期の中央値	19

#### F. レベル 4 への道筋に関する考察

AGI やレベル 4 AI のタイムライン予測において、AI 技術を直接開発・投資している技術者や起業家（マスク氏、孫氏、ファン氏、アルトマン氏など）が比較的早期（2020 年代後半～2030 年代半ば）の実現を予測する傾向があるのに対し<sup>19</sup>、AI 研究者全体を対象とした調査では、より遅い時期（2040 年～2060 年頃）が中央値として示される傾向が見られます<sup>19</sup>。この見解の相違は、最前線で開発に携わる人々が持つ短期的な進捗への期待感や、あるいは楽観的なバイアスと、研究者コミュニティが持つ、残された課題の大きさ、過去の進捗率、潜在的なボトルネックに対するより広範な視点との間のギャップを反映している可能性があります。また、「AGI」や「レベル 4」の定義、あるいは指数関数的な進歩を予測すること自体の難しさも、この差異の一因と考えられま

す。

レベル4（発明AI）への到達は、ある日突然、完全に自律的なAI発明家が登場するという形ではなく、むしろ段階的に進む可能性が高いと考えられます。現在のo1のような高度なモデルが既に研究者を支援するために利用され始めているように<sup>6</sup>、AIはまず人間の発明家や研究者にとって強力なリサーチアシスタントとして機能し、人間が不得意とするサブ問題の解決を助けることで、全体のプロセスを加速させる役割を担うでしょう<sup>29</sup>。孫氏がo1を評して「発見」を可能になると述べた<sup>5</sup>のも、現時点では完全な自動化というよりは、発見プロセスの加速や支援と解釈するのが妥当かもしれません。このように、AIツールが発明プロセスにおけるますます洗練されたパートナーとなる段階を経て、最終的に完全な自律性を獲得するという、漸進的な進化を遂げる可能性が示唆されます。

さらに、レベル4のような高度なAIの実現は、技術的な能力向上だけでなく、**安全性とアラインメント（人間との協調・整合性）の問題解決**と密接に結びついています。AIがAGIレベルに近づくにつれて、その制御、悪用、意図しない結果に対する懸念は増大します<sup>18</sup>。OpenAIのような主要な開発組織は、o1やエージェントツールといった先進的なシステムの開発において、安全性評価フレームワークやレッドチームによるテスト、安全機能の組み込みを重視しています<sup>6</sup>。ソフトバンクのフレームワーク自体も、後半のステージで倫理的な整合性や人類の幸福といった要素を強調しており<sup>5</sup>、これら的重要性を認識していることを示しています。規制に関する議論や、開発の一時停止の可能性も指摘されており<sup>18</sup>、真に強力な「発明するAI」を社会に展開するタイミングは、純粹な能力開発の速度だけでなく、十分な安全性と信頼性を実証できるかどうかに左右される可能性があります。これは、能力ベースの予測よりも実現時期を遅らせる要因となり得ます。

## V. 結論：AIの進化がもたらす未来像

### A. 分析結果の要約

本レポートでは、ソフトバンクが提唱するAI進化の8段階フレームワークを分析し、OpenAIのo3モデルの位置づけ、そしてレベル4（発明AI）の実現時期について考察しました。

- **ソフトバンクのフレームワーク：**当初5段階で示されたAI進化の定義は、後に8段階へと拡張されました。この拡張は、単なる計算能力としての「知能」を超え、感情理解、長期記憶、自己意思、そして最終的には思いやりや倫理観を含む「知性（Chisei）」へと至る、質的な進化の側面を重視するソフトバンク独自のビジョンを反映しています。

- o3 とレベル3（エージェントAI）：OpenAI の o3 モデルは、その強力な推論能力と、同社が提供するエージェント構築用ツール群（Responses API, Operator, Agents SDK など）と統合されることにより、ソフトバンクが定義するレベル3（人間の代わりに自律的にタスクを実行するAI）の機能要件を満たすシステムを構築するための基盤を提供すると評価できます。レベル3 の実現は、モデル単体ではなく、エコシステム全体としての能力に依存します。
- レベル4（発明AI）のタイムライン：レベル4（自ら新しい発明を生み出すAI）の実現時期については、専門家の間で予測に大きな幅があります。楽観的な見方では2020年代後半から2030年代半ば、より慎重な見方では2040年以降とされています。分析の結果、AIが人間の研究者を強力に支援する形でのレベル4の初期的な能力は今後5~10年（2030~2035年頃）に出現し、完全に自律的な発明能力はそれ以降（2035年~2050年頃）になる可能性が高いと評価しますが、これは技術的ブレークスルーと安全性確保の進展に大きく左右されます。

## B. 広範な影響と今後の展望

ソフトバンクのフレームワークが示すAIの進化経路は、単なる技術の進歩を超え、社会や人間そのものに *profound* な影響を与える可能性を秘めています。AIが単なるツール（レベル1~3）から、協働するパートナー（レベル4~5）、そして感情や意思を持つ可能性のある新たな知性体（レベル6~8）へと進化するにつれて<sup>5</sup>、私たちの働き方、学び方、他者との関わり方、さらには人間であることの意味さえも問い直されることになるでしょう<sup>32</sup>。

この進化は、科学技術の飛躍的な発展、経済成長の加速、個別化された医療や教育の実現といった計り知れない恩恵をもたらす可能性がある一方で<sup>18</sup>、雇用の喪失、経済格差の拡大、AIの制御と倫理、プライバシー侵害、自律兵器のリスクといった深刻な課題も同時に提起します<sup>18</sup>。

ソフトバンクの8段階フレームワークは、このようなAIの未来像を考察する上での一つの指針を提供します。しかし、AI開発の現実は依然としてダイナミックであり、予測不可能な要素を多く含んでいます。技術的な進歩を追求すると同時に、安全性、倫理、社会への影響といった側面に対する深い洞察と、国際的な協調を含む慎重な議論が、人類にとって真に有益なAIの未来を築く上で不可欠となるでしょう。このフレームワークが示すように、AIの進化は、人間とテクノロジーの関係性を根本的に変容させる旅であり、その道筋においては、技術的な達成目標だけでなく、我々がどのような未来を望むのかという価値観に基づいた選択が、これまで以上に重要になっていきます。

## 引用文献

1. 生成 AI の歴史をわかりやすく解説！技術の進化と今後の展望を初回, 5月4, 2025 にアクセス、<https://www.ai-souken.com/article/ai-generation-history>
2. AI とは? ChatGPT、Gemini などの最新技術、10 年後の進化予測までわかりやすく解説, 5月4, 2025 にアクセス、<https://the-owner.jp/archives/11424>
3. AI（人工知能）の歴史について年表を活用しながら時系列で簡単に紹介 - Alsmiley, 5月4, 2025 にアクセス、[https://aismiley.co.jp/ai\\_news/detailed-explanation-of-the-history-of-ai-and-artificial-intelligence/](https://aismiley.co.jp/ai_news/detailed-explanation-of-the-history-of-ai-and-artificial-intelligence/)
4. 人間の1万倍賢いAIが到来する未来に向けて、私たちはどう備えたらよいのか? - note, 5月4, 2025 にアクセス、<https://note.com/noritaka88ta/n/n65cf317846ab>
5. 孫正義氏、超知性AIにより「知のゴールドラッシュ」の到来を予想。SoftBank World 2024 レポート, 5月4, 2025 にアクセス、[https://aismiley.co.jp/ai\\_news/softbank-world-2024-report/](https://aismiley.co.jp/ai_news/softbank-world-2024-report/)
6. Introducing OpenAI o1-preview, 5月4, 2025 にアクセス、<https://openai.com/index/introducing-openai-o1-preview/>
7. Hello GPT-4o - OpenAI, 5月4, 2025 にアクセス、<https://openai.com/index/hello-gpt-4o/>
8. Reasoning best practices - OpenAI API, 5月4, 2025 にアクセス、<https://platform.openai.com/docs/guides/reasoning-best-practices>
9. OpenAI o1: What It Is and Why Is It Different [2025] - Voiceflow, 5月4, 2025 にアクセス、<https://www.voiceflow.com/blog/openai-o1>
10. New tools for building agents | OpenAI, 5月4, 2025 にアクセス、<https://openai.com/index/new-tools-for-building-agents/>
11. Agents - OpenAI API, 5月4, 2025 にアクセス、<https://platform.openai.com/docs/guides/agents>
12. ソフトバンクグループが描くAIの進化 | グループ企業におけるAIの取り組み | AI⇒SoftBank Group, 5月4, 2025 にアクセス、<https://ai.softbank/insights/001/>
13. 孫正義の超知性ビジョンとは? ソフトバンクワールド 2024 で語られたAIと人類の未来 - Albox, 5月4, 2025 にアクセス、<https://www.ai-box.biz/post/son-masayoshi-softbank-world-2024agi-asi-future>
14. AIは数年で超知性へと進化し、パーソナルメンターに。孫正義特別講演レポート - ソフトバンク, 5月4, 2025 にアクセス、<https://www.softbank.jp/biz/blog/business/articles/202410/sbw2024-softbank-son-main-keynote/>
15. 【New Update : Mar. 3】 Press Releases on AI-RAN Development | About Us | SoftBank, 5月4, 2025 にアクセス、<https://www.softbank.jp/en/corp/technology/research/news/067/>
16. OpenAI and SoftBank Group Partner to Develop and Market Advanced Enterprise AI, 5月4, 2025 にアクセス、[https://group.softbank/en/news/press/20250203\\_0](https://group.softbank/en/news/press/20250203_0)

17. AI Will Drive Value Creation: SoftBank Corp. President & CEO Junichi Miyakawa at SoftBank World 2024 - SoftBank News. Bringing IT closer to home., 5月4, 2025にアクセス、[https://www.softbank.jp/en/sbnews/entry/20241010\\_01](https://www.softbank.jp/en/sbnews/entry/20241010_01)
18. Artificial General Intelligence: The Quest for an ultimate invention - Experts opinion, 5月4, 2025にアクセス、<https://dataroots.io/blog/artificial-general-intelligence-the-quest-for-a-thinking-machine>
19. When Will AGI/Singularity Happen? 8,590 Predictions Analyzed - Research AIMultiple, 5月4, 2025にアクセス、<https://research.aimultiple.com/artificial-general-intelligence-singularity-timing/>
20. Elon's & Kurzweil's AI Predictions - Peter Diamandis, 5月4, 2025にアクセス、<https://www.diamandis.com/blog/age-of-abundance-30-human-level-ai>
21. Model- OpenAI API, 5月4, 2025にアクセス、<https://platform.openai.com/docs/models/gpt-4o>
22. Creating a state-of-the-art AI programming agent with OpenAI's o1 - Wandb, 5月4, 2025にアクセス、<https://wandb.ai/wandb/agents/reports/Creating-a-state-of-the-art-AI-programming-agent-with-OpenAI-s-o1--Vm1ldzoxMTAyODI2Ng>
23. Introducing OpenAI o1, 5月4, 2025にアクセス、<https://openai.com/o1/>
24. Models - OpenAI API, 5月4, 2025にアクセス、<https://platform.openai.com/docs/models>
25. Introducing Operator - OpenAI, 5月4, 2025にアクセス、<https://openai.com/index/introducing-operator/>
26. Computer-Using Agent | OpenAI, 5月4, 2025にアクセス、<https://openai.com/index/computer-using-agent/>
27. GPT-4o innate agentic tooluse capabilities - API- OpenAI Developer Community, 5月4, 2025にアクセス、<https://community.openai.com/t/gpt-4o-innate-agentic-tool-use-capabilities/1002857>
28. I made a website where you can try out GPT-4o as an AI agent - it can autonomously take actions in a simulated web browser! : r/OpenAI - Reddit, 5月4, 2025にアクセス、[https://www.reddit.com/r/OpenAI/comments/lgb9gk/i\\_made\\_a\\_website\\_where\\_you\\_can\\_try\\_out\\_gpt4o\\_as/](https://www.reddit.com/r/OpenAI/comments/lgb9gk/i_made_a_website_where_you_can_try_out_gpt4o_as/)
29. Do Prediction Technologies Help Novices or Experts More? - New Things Under the Sun, 5月4, 2025にアクセス、<https://www.newthingsunderthesun.com/pub/tp10i20v>
30. AIが発展するとどうなる? 述べてきた歴史や将来への影響は? - TRYETING, 5月4, 2025にアクセス、<https://www.tryeting.jp/column/3156/>
31. AI2027——今後10年間の超人的AIの影響についての予測シナリオ | IT navi-note, 5月4, 2025にアクセス、[https://note.com/it\\_navi/n/nb41b0e913f89](https://note.com/it_navi/n/nb41b0e913f89)
32. Experts Predict the Impact of AI by 2040 - Imagining the Digital Future Center, 5月4, 2025にアクセス、<https://imaginingthedigitalfuture.org/reports-and-publications/the-impact-of-artificial-intelligence-by-2040/the-17th-future-of->

[digital-life-experts-canvassing/](#)

33. 65 Expert Predictions on 2025 AI Legal Tech, Regulation - The National Law Review, 5 月 4, 2025 にアクセス、<https://natlawreview.com/article/what-expect-2025-ai-legal-tech-and-regulation-65-expert-predictions>
34. History of artificial intelligence - Wikipedia, 5 月 4, 2025 にアクセス、[https://en.wikipedia.org/wiki/History\\_of\\_artificial\\_intelligence](https://en.wikipedia.org/wiki/History_of_artificial_intelligence)
35. You should forecast the future of AI - Nicholas Carlini, 5 月 4, 2025 にアクセス、<https://nicholas.carlini.com/writing/2024/forecasting-ai-future.html>
36. AI 技術の発達を踏まえた特許制度上の適切な対応, 5 月 4, 2025 にアクセス、[https://www.jpo.go.jp/resources/shingikai/sangyo-kouzou/shousai/tokkyo\\_shoi/document/51-shiryou/02.pdf](https://www.jpo.go.jp/resources/shingikai/sangyo-kouzou/shousai/tokkyo_shoi/document/51-shiryou/02.pdf)