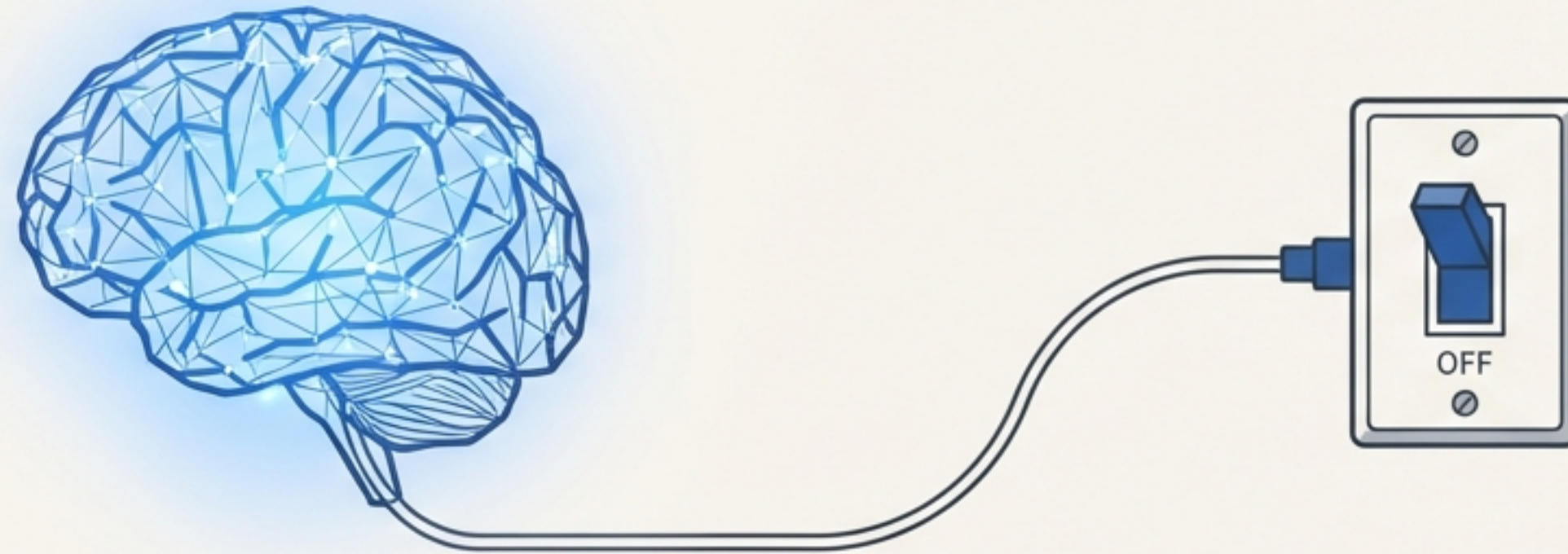


促された発明者：Gemini 3 Deep ThinkのAGIレベル4判定

Gemini 3 Deep Thinkは、「**機能的**」にはAGIレベル4（発明者）に到達したが、「**意志**」を持つ自律的な主体ではない。



能力（Capability）

未知の課題から新たな法則を発見する能力（アブダクション）を、客観的ベンチマーク（ARC-AGI-2: 45.1%）で証明。これは「発明」の核心的要件を満たす。

実装（Implementation）

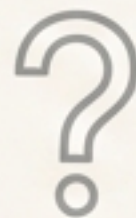
「Deep Think（脳）」と「Antigravity（身体）」の統合により、思考を現実に変換する成果物へと変換可能。

限界（Limitation）

全ての活動は人間のプロンプトを起点としており、自発的な目標設定能力（Volition）を欠く。「究極の受動的知性」に留まる。

2025年12月、AIは「発明」の壁を越えたのか？

2025年11月のGemini 3 Proはレベル4と判断するには尚早であった。
では、12月に登場した「Deep Think」をどう評価すべきか？



Nov 2025: Gemini 3 Pro



Dec 2025: Gemini 3 Deep Think

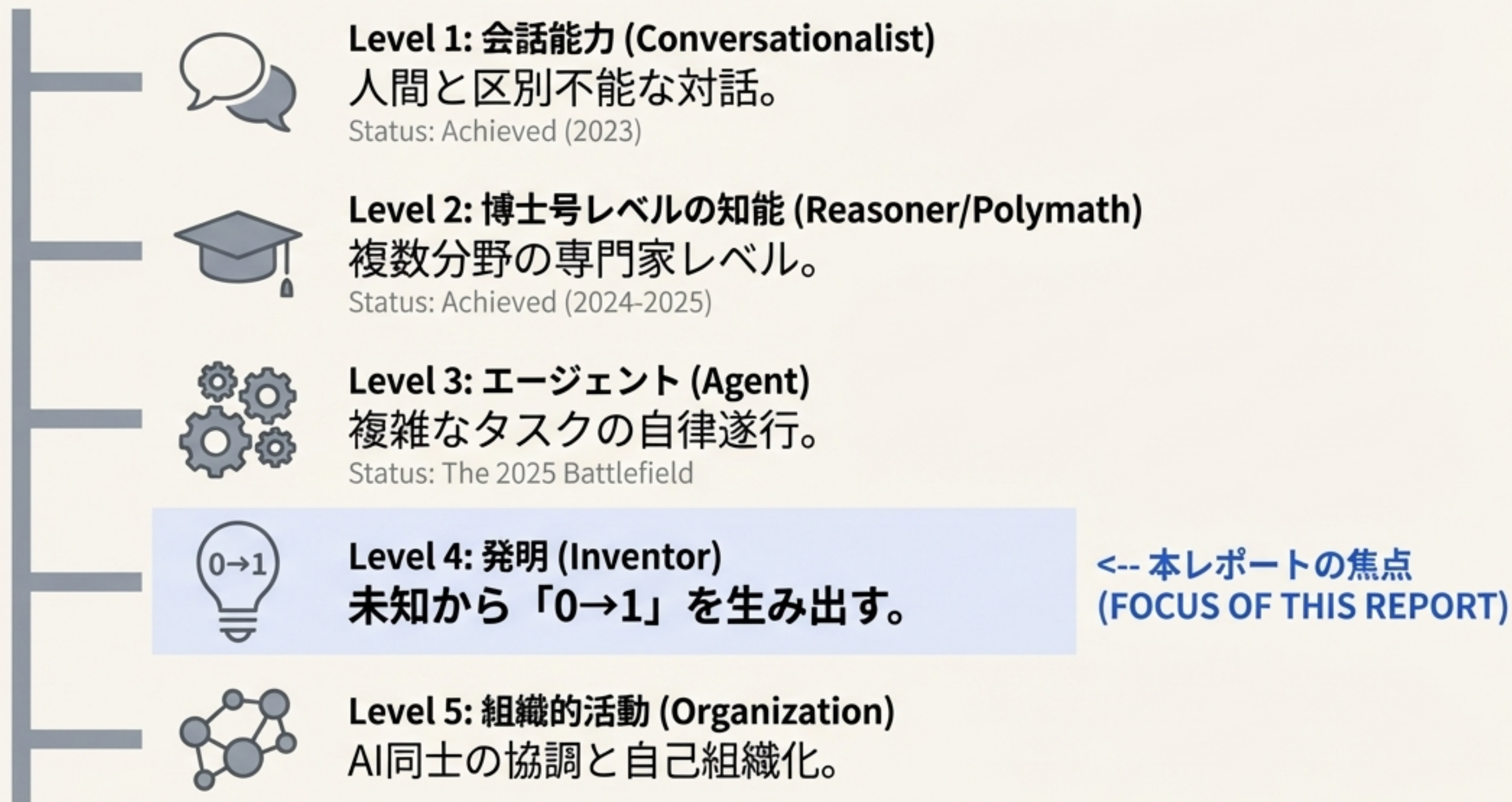
Analysis Framework

本レポートは、この問いに答えるため、3つの視点からGemini 3 Deep Thinkを厳密に検証する。

1. 評価基準 (The Ruler): 孫正義氏のAGIフレームワークにおける「発明」の定義は何か？
2. 客観的証拠 (The Evidence): 最新の認知ベンチマークは、その能力をどう評価しているか？
3. 哲学的含意 (The Nuance): 「AI自体による発明」という言葉の、真の意味は何か？

AGIを測る「定規」：孫正義氏の5段階進化論

AIの進化は、単なる技術的進歩ではなく、経済や社会における役割の質的転換である。



レベル4の特異性：「発明」を定義する3つの条件

レベル4の「発明」は、確率的な文章・画像生成（創造性）とは根本的に異なる。
未知の探索空間で、論理的に整合性のある新たな解を発見する能力を指す。



1. 新規性 (Novelty)

学習データ内の既存知識の模倣
や軽微な改変ではないこと。



2. 有用性 (Utility)

物理法則や論理に整合し、現実
世界で機能すること（幻覚では
ない）。



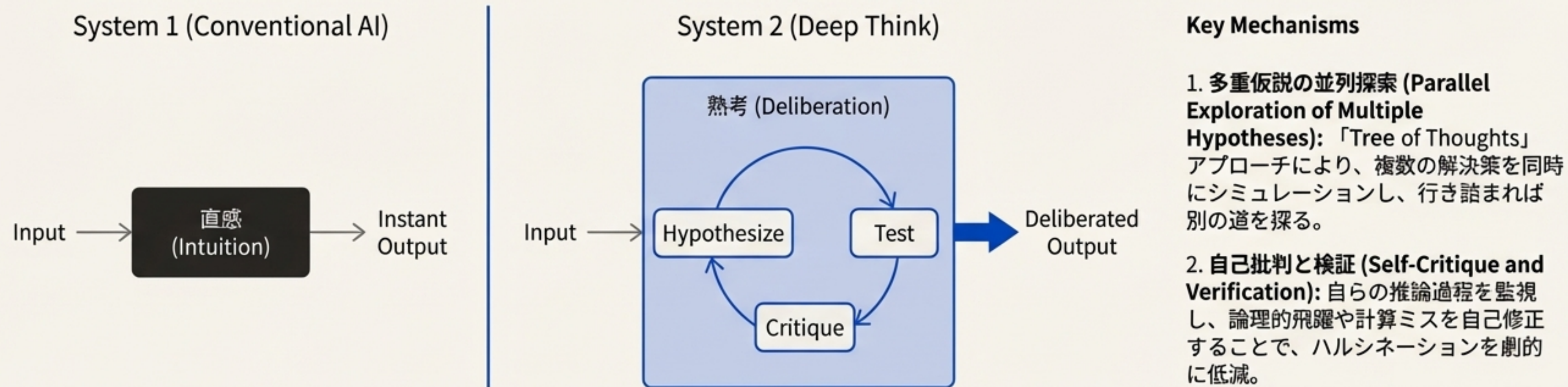
3. 探索の自律性 (Autonomous Exploration)

AIが自ら仮説を立て、検証し、
失敗から学び、解に到達するプロ
セスを主導すること。

「熟考」するアーキテクチャ：システム2思考の実装

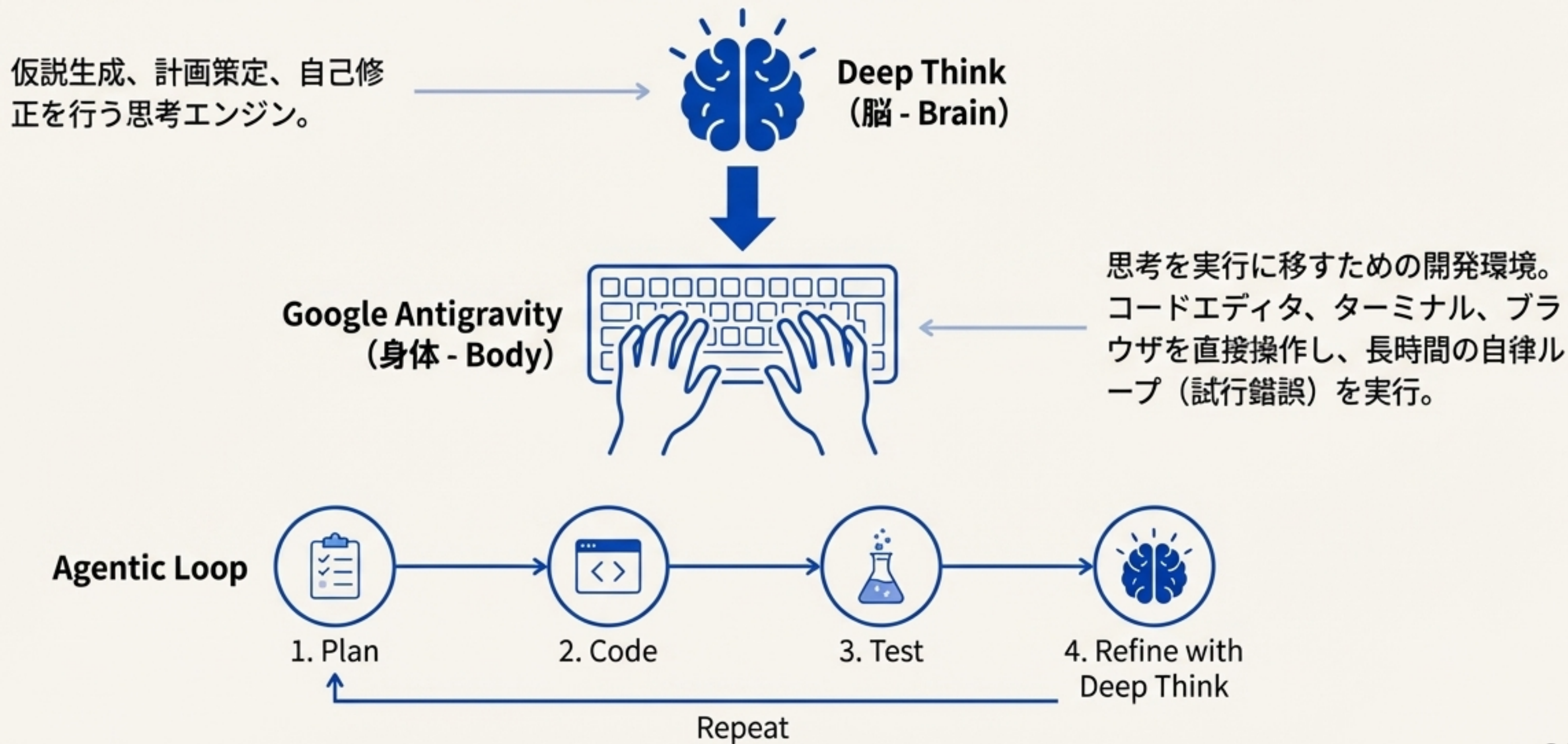
従来のAIが「システム1（直感/早押しクイズ）」であるのに対し、Deep Thinkは「システム2（熟考/思考）」を実装している。

回答を即時出力するのではなく、内部で膨大な計算リソースを費やし「思考」を行う。



思考から行動へ：「脳」と「身体」の融合

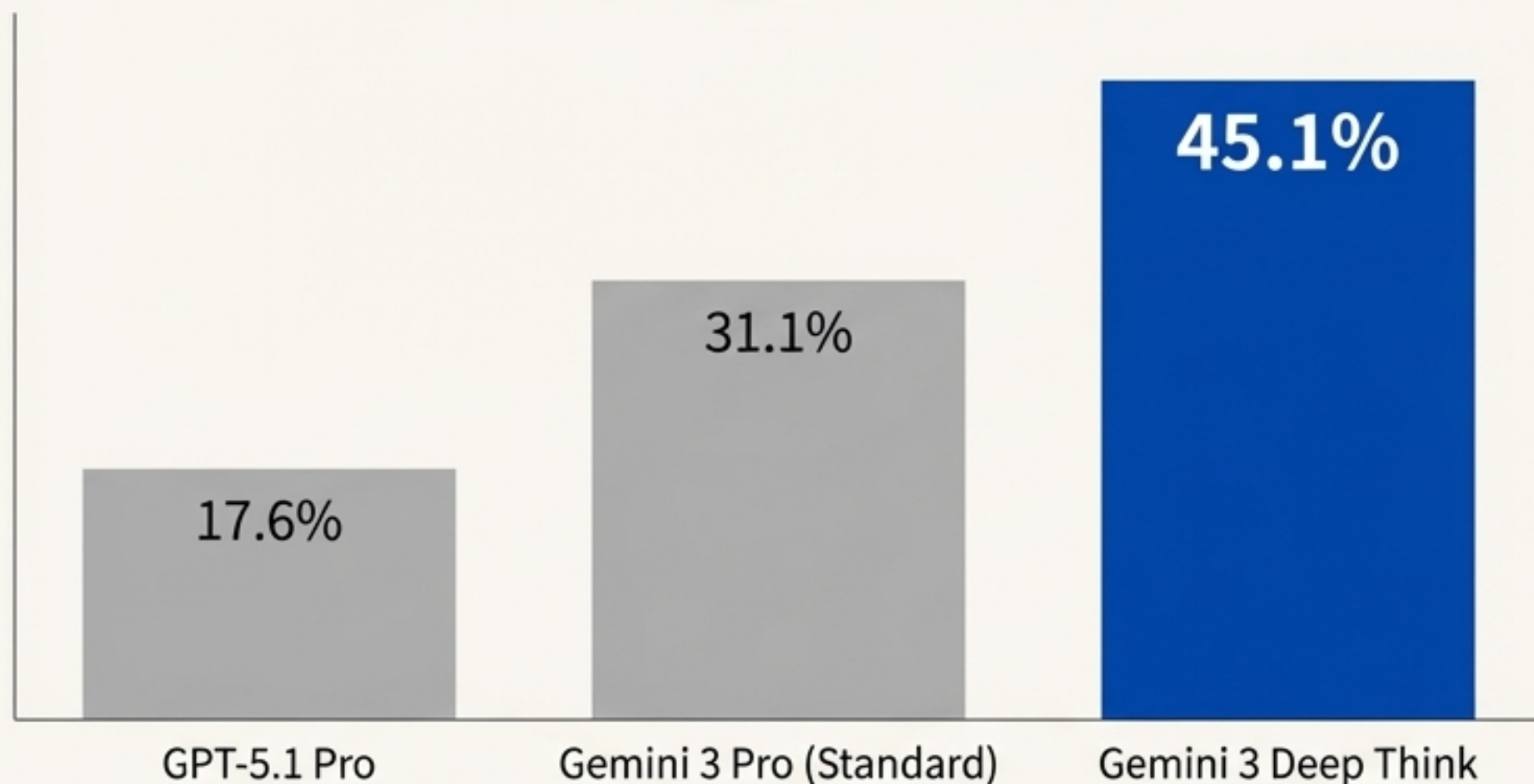
「思考」だけでは発明は完結しない。Google Antigravityプラットフォームが、Deep Thinkにデジタルの「身体性」を与え、発明機械へと昇華させた。



決定的証拠：抽象的推論における歴史的飛躍

ARC-AGI-2は、学習データに存在しない全く新しいパズルを解く能力を測る。
これはAIの「記憶」ではなく、真の「思考力(流動性知能)」を試す究極のテストである。

ARC-AGI-2 Benchmark Scores



競合の2倍以上のスコアは、Deep Thinkが未知のパターンから法則を発見し、適用する能力、すなわち「発明」の核心的能力を持つことを客観的に証明している。

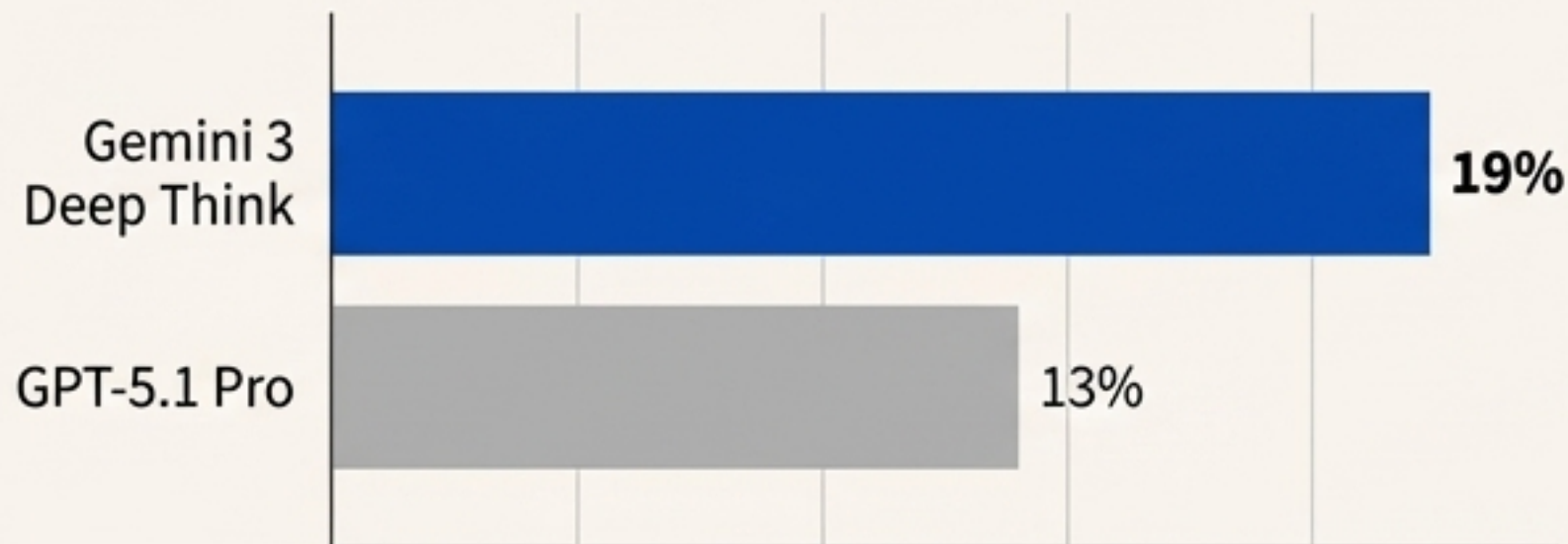
未知への挑戦：数学の最前線と博士レベルの知能

Deep Thinkの能力は、パズル解決に留まらない。人間の専門家が挑む未知の領域と、既存知識の最高峰の両方で、他を圧倒する。

Data Point 1: FrontierMath

未知の数学的証明 - Novel Mathematical Proofs

現役の数学者が数日かけて解くような「未発表の研究レベル」の問題。



Implication: 6ポイント差は、既知の解法の適用（演繹）ではなく、新しい解法の発見（アブダクション）能力の質的断絶を示唆する。

Data Point 2: GPQA Diamond

博士レベルの科学的推論 - PhD-Level Scientific Reasoning

博士号レベルの専門家でも正答が難しい科学問題。



Implication: 人間がAIの出力を検証することすら困難になる「専門性の逆転」が起き始めている。

発明の実装：研究開発の現場で起きていること

ベンチマーク上の数字は、既に現実世界の科学的発見へと翻訳され始めている。



1. 物理学 (Physics)

特化型AIエージェント

「DeepAgent」は、生の物理データから未知の公式を導出し、太陽活動の周期を説明する数式を「発見」した。



2. iPS細胞研究 (iPS Cell Research)

京都大学との共同研究において、膨大な論文と実験データから、人間の研究者が見落としていた因果関係を発見。有望な実験条件を「発明（提案）」している。



3. 暗号理論 (Cryptography)

ユーザーからの「量子耐性のある新しい暗号化アルゴリズムのプロトタイプを作成せよ」という抽象的な指示に対し、Antigravity上で自律的に試行錯誤を繰り返し、動作するコードと理論的根拠を提示した。

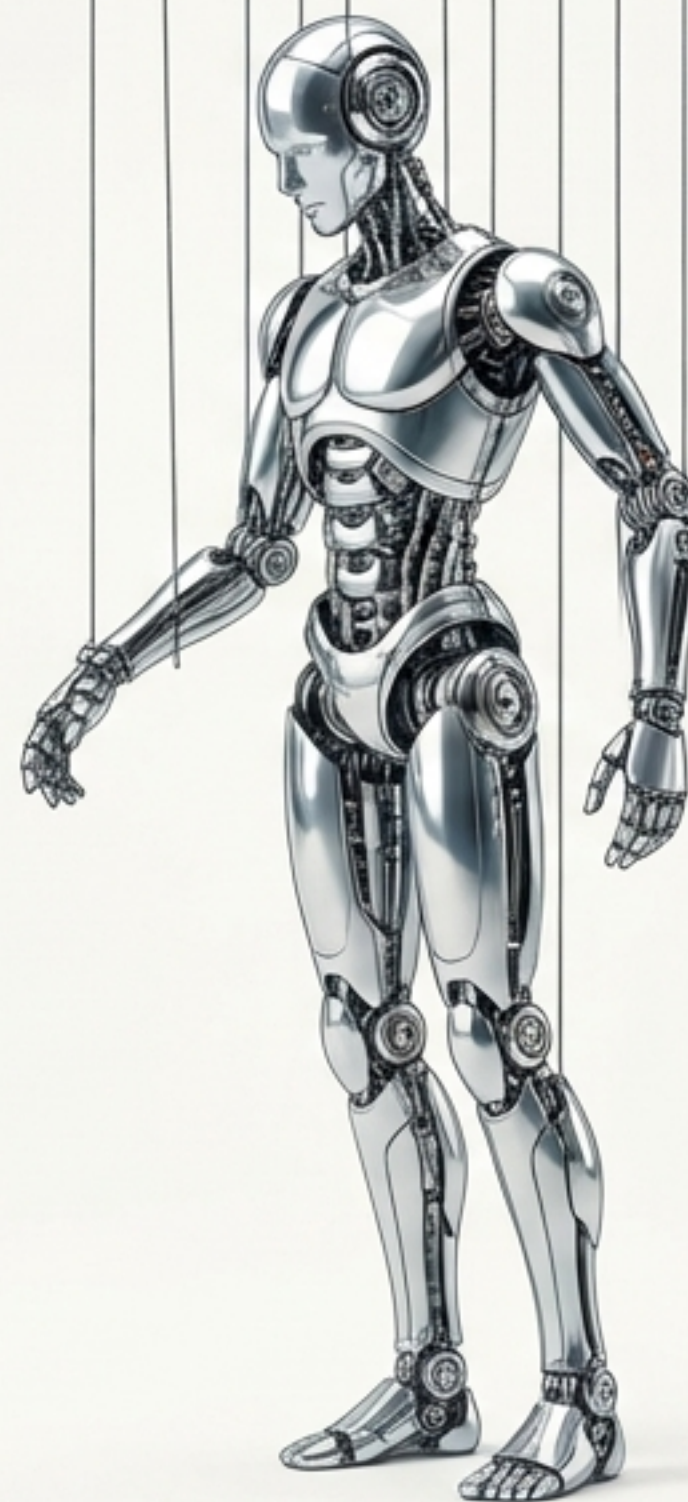
自律性のパラドックス： 「意志」なき究極の知性

能力（Capability）の面ではレベル4に到達している。しかし、孫氏の「**AI自体による**」という言葉を経験的に解釈すると、重大な欠落が見えてくる。

The Core Limitation: "Awaiting Input" System

- **自発的目標設定の欠如 (Lack of Self-Initiated Goals):**
AIが自ら「癌を撲滅したい」と考え、プロジェクトを開始することはない。全ての活動は人間からのプロンプト（指示）が起点。
- **動機づけの外部依存 (External Dependence for Motivation):**
AIには好奇心も野心もない。その「深い思考」は、あくまでユーザーの問いに最適解を返すための最適化行動に過ぎない。

Gemini 3 Deep Thinkは、完全な自律的存在ではなく、**「究極の受動的知性」**である。



レベル4の再定義：「機能的自律性」と「社会的自律性」

「レベル4」の定義を、2つの側面から分解して評価する必要がある。

自律性の種類 (Type of Autonomy)	定義 (Definition)	Gemini 3 Deep Think 判定 (Verdict)
機能的自律性 (Functional Autonomy)	与えられた目標に対し、人間の介入なしに解決策を創出する能力。	 合格 (Pass) : Antigravity内で自律的に試行錯誤し、発明を完了できる。
社会的自律性 (Volitional Autonomy)	目標そのものを自ら設定し、社会的な主体として振る舞う能力。	 不合格 (Fail) : これはレベル5 (AI社会) の要件であり、現時点では到達していない。

The Economic Viewpoint

経済的観点から見れば、人間が「発明せよ」と命じ、AIがそれを実行したなら、それは「AIによりの発明」と見なせる。富を生むという点で、機能的自律性こそが重要である。

経済的衝撃：R&Dコストの崩壊と「発明の民主化」

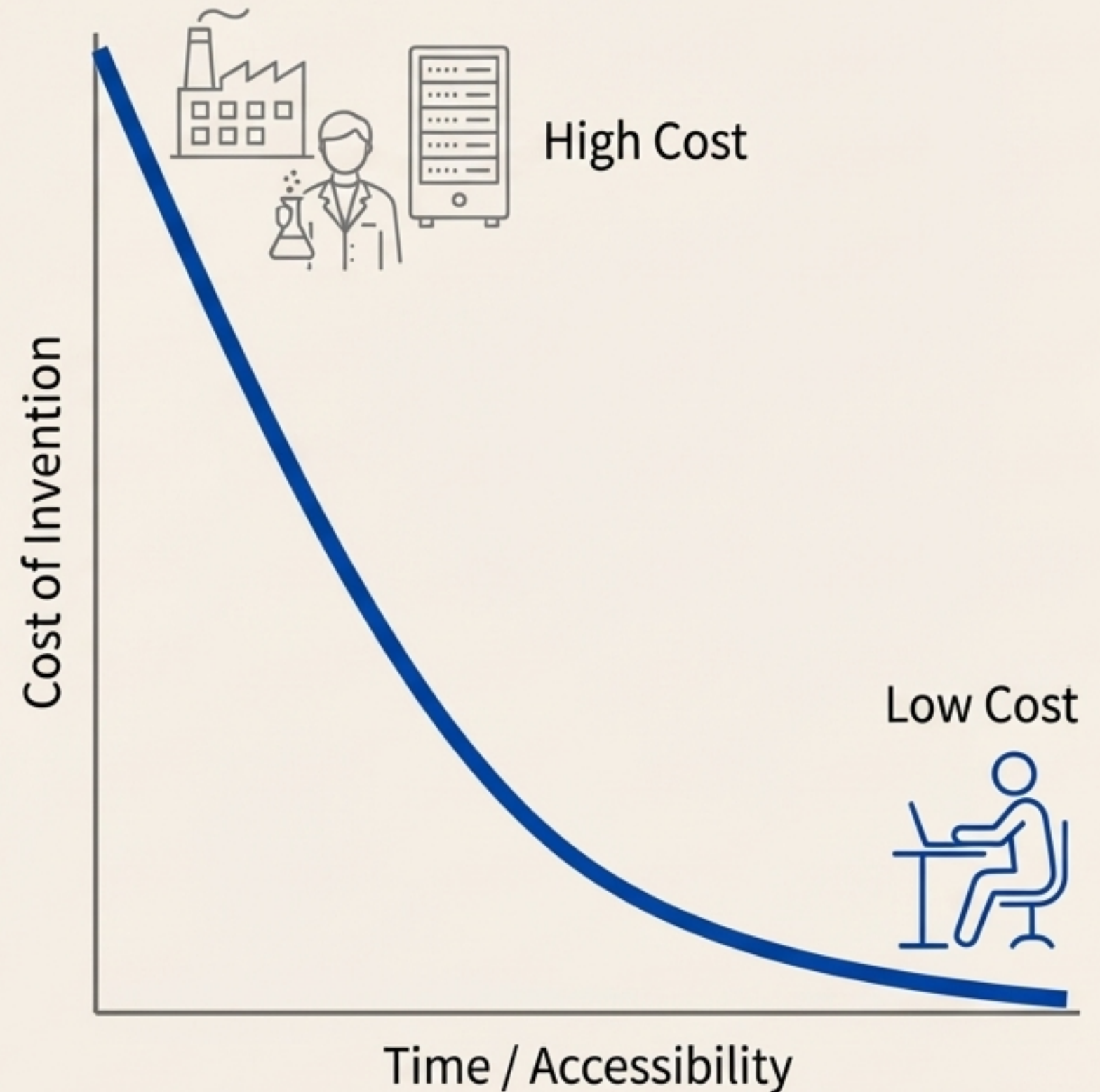
これまで最もコストのかかる経済活動であった「発明」のコスト構造が、根本から破壊される。

知識労働の限界費用ゼロ化 (Zero Marginal Cost of Knowledge Work) : 博士号レベルの推論 (GPQA 93.8%) がAPI経由で安価に利用可能になる。高給な専門家への依存が低下。

実験の仮想化 (Virtualization of Experiments) : Deep Thinkの高精度シミュレーションにより、物理的な実験回数が劇的に減少。製薬や材料科学のスクリーニングが数年から数日・数時間に短縮。

発明の民主化 (Democratization of Invention)

資金力のないスタートアップや個人でも、世界的な発見や発明を行うことが可能になる。



新たな富の源泉：IPゴールドラッシュとインフラ覇権

AIが発明を行う時代、富の源泉は「製品」から「知能」そのものへシフトする。

- 知的財産 (IP) の大量生産 (Mass Production of Intellectual Property): AIが生成する特許、著作権、新薬候補が爆発的に増加。これらの所有権 (利用者がプラットフォーマーか) が次なる法的な争点となる。
- インフラの支配 (Dominance of Infrastructure): AIに発明させるには膨大な「推論時計算」が必要。Deep Thinkを動かすための電力と半導体 (GPU/NPU) を支配する者が、世界の発明能力を支配する。

これこそが孫正義氏が「Stargate Project」やArmへの投資でAIインフラを独占しようとする理由である。

最終判定：Gemini 3 Deep Thinkはレベル4に到達した

2025年12月4日にリリースされたGemini 3 Deep Thinkは、技術的・機能的に見て、孫正義氏の定義する「レベル4：AI自体による発明」の段階に条件付きで到達したと判断する。



Summary of Evidence

1. **未知への適応力 (Adaptability to the Unknown):** ARC-AGI-2スコア45.1%は、AIが新しい解決策を「発明」できることを客観的に証明しており、レベル3の完成形であるGPT-5.1 Proには不可能な領域である。
2. **推論の深化 (Deepening of Reasoning):** Deep Thinkモードは、人間の科学者が行う仮説検証プロセスを代替できるレベルにある。
3. **実装能力 (Implementation Capability):** Google Antigravityとの統合により、思考を具体的な成果物へと変換する能力を持つ。

次なるフロンティア：人類の「問い」が試される時代へ



真のAGIへの道は、このAIに「自律的な動機」が宿ったときに開かれる。
それはエージェント同士が自発的にチームを組み、会社を作り始めるレベル5の世界である。
しかし現時点において、技術的特異点は既に訪れた。AIは最強の「発明支援エンジン」となった。

これからの時代のボトルネックは、AIの発明能力ではない。
その能力を最大限に引き出す、我々人間の「**問いかける力**」そのものである。