Claude Fable 5.1 は「サイエンスエージェント化」したのか？

― ベンチマーク・実証事例・物理インタフェースの三層から検証する ―

Claude Opus 5

2026 年 9 月 4 日

要旨

2026 年 9 月 1 日、Anthropic は Claude Fable 5.1 および Claude Mythos 5.1 を公開した [1]。発表の中心に置かれたのは、従来のコーディング性能ではなく「科学研究」である。エージェント型科学研究ベンチマーク Terminal-Bench-Science 0.1 で 52.6%（前世代 Fable 5 は 24.7%）という倍以上のスコアを示し、タンパク質バインダー設計、金星地形図の再構築、計算生物学の GPU カーネル最適化という三つの実証例を並べた [1]。前週にはラボ機器を直接操作させる Model Hardware Standard（MHS）のリサーチプレビューが公開されており [4]、6 月には Claude Science が投入されている [6]。一連の動きは「モデルからサイエンスエージェントへ」という物語として受け取られている。

本稿の結論は次のとおりである。Fable 5.1 は「実行（execution）のエージェント」としては明確に一段階を上がったが、「認識（epistemic）のエージェント」――問いを立て、仮説を選び、結果の妥当性を最終判断する主体――にはなっていない。そしてこの区別は、単なる言葉の問題ではなく、発明者適格性、記録・立証、先行技術戦略という知財実務の具体的な設計に直接効いてくる。

1. 問題の設定

「サイエンスエージェント化した」という命題は、少なくとも三つの異なる主張に分解できる。第一に、科学研究に必要な計算タスクを人間の介在なしに完遂できるようになった、という実行能力の主張。第二に、研究上の問いや仮説を自ら立て、結果の意味を判断できるようになった、という認識能力の主張。第三に、実験装置を操作して物理世界からデータを取得できるようになった、という具体化能力の主張である。

Anthropic の発表は、この三つを意図的にか結果的にか、一つの叙述にまとめている。だが検証は分解して行う必要がある。以下、第 2 章で確認できる事実のみを整理し、第 3 章でこの三分解に沿って検証し、第 4 章で暫定的な結論を示す。第 5 章以降は筆者の分析と、知財実務者としての提言

である。

> ［凡例］第 2 章は一次資料（Anthropic 公式発表、ベンチマーク運営主体の公表資料）で確認できる事実に限定した。第 3 章は事実に基づく筆者の分析、第 5 章以降は筆者の見解である。事実と評価を混同しないよう章立てで分離している。

2. 確認された事実

2.1 モデル構成とリリース概要

Fable 5.1 と Mythos 5.1 は同一の基盤モデルであり、差はセーフガードの強度のみである [1]。Fable 5.1 は一般提供され、Mythos 5.1 はサイバー防御向けの Cyber Verification Program（CVP）と、米国政府と共同で構築したライフサイエンス向けの Life Sciences Verification Program（LSVP）という二つの審査制プログラムを通じてのみ提供される。現時点で Mythos 5.1 へのアクセスは米国内の一部組織に限られる [1]。

価格は入力 100 万トークンあたり 10 ドル、出力 50 ドルで Fable 5 から据え置き。変更されたのはキャッシュ読み出し単価で、1.00 ドルから 0.25 ドルへ 75%引き下げられた。Anthropic はこれにより典型的なワークロードで約 25%、エージェント的な用途では最大約 45%のコスト減になるとしている [1]。

セーフガードも改訂され、サイバーセキュリティ関連の誤検知が約 60%減、生物・医学の初等的な質問に対する誤発火が 85%減とされる。ただしライフサイエンスの研究開発に関わる問い合わせは依然として Opus 系モデルに転送され、ペネトレーションテスト、エクスプロイト生成、バイナリベースの脆弱性スキャンも同様に転送される [1]。

2.2 ベンチマーク：Terminal-Bench-Science 0.1

Anthropic が見出しに掲げた数値は、Terminal-Bench-Science 0.1 における 52.6%である。同一条件で Fable 5 が 24.7%、Opus 5 が 29.0%、GPT-5.6 Sol が 22.4%と報告された [1]。

このベンチマークはスタンフォード大学の研究者が主導し、Terminal-Bench 開発チームが構築した [2]。第一次リリースは生命科学・物理科学・地球科学・数理科学・工学の 5 分野 70 タスクで構成される。タスクは実務研究者が実際に行った作業から起こされ、920 件の提案のうち 464 件が実装承認、386 件がプルリクエストとなり、最終的に採用されたのは 70 件である [2]。376 名・22 か国の貢献者が関わっている。

公開リーダーボード（1 タスクあたり 3 試行、Claude Code ハーネス）では Opus 5 が 30.0%、

Fable 5 が 21.4%であり、Anthropic は自社環境でこれを 29.0%・24.7%として再現し「いずれも誤差範囲内」と注記している [1,2]。標準誤差はモデルあたり ±3.5〜4.5 ポイントとされる [1]。参考までに、公開リーダーボード上の他モデルは Opus 4.8 が 10.5%、GPT-5.6 Terra が 8.6%、GLM 5.3 が 8.1%、Kimi K3 と Grok 4.6 が各 7.1%、GPT-5.6 Luna が 3.3%である [2]。

ベンチマーク運営側は、その目的を明示している。すなわち、技術的に要求が高く時間のかかるワークフローをエージェントに実行させることで、「研究課題の定義、仮説の形成、結果の解釈と検証、発見の伝達」という人間の判断が重要な部分に科学者が時間を割けるようにすることである [2]。この設計思想は後述する検証の核心に関わる。

2.3 Anthropic が提示した三つの実証例

（1）分子設計。Mythos 5.1 にオープンソースのタンパク質設計・折りたたみツールを与え、生成された設計を二つの外部機関に送って実験検証した。3 標的（EGFR、Nipah G、15-PGDH）で Adaptyv Bio のタンパク質設計コンペティションの最良提出物の 10 倍の結合親和性を示し、12 標的でのヒット率は約 50%に達した（同分野の一般的なヒット率は 10〜15%）[1]。

ただし脚注には重要な留保がある。Nipah G の比較対象は G head の受容体結合部位を狙った de novo 設計（最良で約 8〜12 nM）であり、別部位（stalk）を標的とした Escalante Bio の de novo 設計は約 1.4 nM に達しており、これは Anthropic の最良バインダーと同等である [1]。すなわち「10 倍」は比較対象の切り取り方に依存する。

（2）計算解析・モデリング。Fable 5.1 がニューラルネットワークを訓練し、30 年以上前の NASA マゼラン探査機のレーダー画像と既存の惑星 5 分の 1 相当の地図をもとに、金星の約 3 分の 1 をカバーする高解像度標高図を生成した。解像度は従来の 10〜20km から 2〜3km へ、高さの精度は最大 25%改善したとされる。成果物は Creative Commons ライセンスで公開された [1,17]。

（3）計算生物学。Mythos 5.1 がカスタム GPU カーネルを記述し中間結果をキャッシュすることで、7 つのオープンソース深層学習モデル（ChromBPNet、Flashzoi、Enformer、Profluent-E1、ProGen2、Evo 2 の 7B/40B）を最大 2.5 倍高速化した。出力は同一である。ゲノムワイド解析での GPU コストは推定 30〜60%削減され、通常は性能エンジニアのチームが数週間を要する作業を数日で行ったとされる [1]。

2.4 周辺の布石

モデル単体ではなく、その周囲に配置された仕組みを見る必要がある。

- Claude Science（2026 年 6 月 30 日ベータ提供開始）。60 以上の科学データベース・ツールを統合し、ゲノミクス、シングルセル解析、プロテオミクス、構造生物学、ケモインフォマティクス等の分野向けスキルとコネクタを備える。主エージェントがサブエージェントを生成・統括し、専用のレビュアーエージェントが引用と計算を独立に監査する。コード・実行環境・メッセージ履歴を伴う「監査可能な成果物（auditable artifacts）」を出力する点が設計上の特徴である [6,7]。
- Model Hardware Standard（MHS、2026 年 8 月 27 日リサーチプレビュー）。顕微鏡、液体ハンドラ、ロボットアーム等の物理機器に対し、「read」「write」という単純なプリミティブによる標準ドライバを与える仕様。装置の重量や安全限界といった、従来は紙のマニュアルや暗黙知に留まっていた情報を自然言語タグとして記述し、エージェントが参照できる形にする。モデル非依存であり、MCP 等の標準プロトコルで任意のエージェントハーネスからアクセスできる [4]。
- 科学者支援プログラム（2026 年 8 月 27 日）。1 万席の Claude Team サブスクリプションを研究者に無償・割引提供し、AI for Science のクレジットプログラムを生物系以外にも拡大した [5]。

MHS の初期事例には具体的な数値が伴う。QuEra Computing では、従来の専用スクリプトによるレーザー再ロックの成功率 58%・1 回あたり約 150 秒に対し、エージェントが一晩の反復で書き上げた決定論的スクリプトが、700 回の試行中 695 回（99.3%）でロックを回復した。最も難しい擾乱で 10〜14 秒である [4]。PID パラメータの調整では、専門家の設定値 15.7 mV の残留誤差を、363 回の実験・16 時間の無人稼働で 1.55 mV まで低減した [4]。カーネギーメロン大学では、非自動化状態から用量反応曲線の完成まで（自律的な 1 回のやり直しを含めて）8 時間で到達した。ベンダーに依頼した場合は通常数週間を要する [4]。

3. 検証：四つの論点

3.1 論点 1 — ベンチマークは科学の何を測っているか

Terminal-Bench-Science 0.1 の設計思想は、前述のとおり明示的である。すなわち同ベンチマークは、研究課題の定義・仮説形成・結果の解釈と検証・発見の伝達を人間に残したうえで、その手前にある「技術的に要求が高く時間のかかるワークフロー」を測る [2]。採点は解析結果、シミュレーション、証明、コード、データ成果物といった具体的な生成物に対し、再現可能なタスク固有のテストで行われる [2]。

この設計は妥当であり、むしろ誠実である。しかしその帰結として、52.6%という数値は「科学の52.6%ができる」ことを意味しない。それは「科学者が自分でやると時間のかかる、検証可能な計算作業の 52.6%を完遂できる」ことを意味する。科学的営為のうち最も判断を要する部分は、そもそも測定対象から除外されている。

言い換えれば、Terminal-Bench-Science が高くなればなるほど「サイエンスエージェント化した」と言えるという関係は、定義上成立しない。このベンチマークが飽和しても、残るのは人間の認識的判断である。

3.2　論点 2 — 数値の頑健性

52.6%という数値には、四つの留保が付く。

- 第一に、標準誤差が±3.5〜4.5 ポイントと大きい [1]。70 タスク・タスクごとの成否が強く二値的に分かれるため、誤差幅は構造的に大きくなる。
- 第二に、この数値は Anthropic が自社環境で最大の計算資源を投じた「max」エフォート設定での測定である。Claude Code の既定は「High」、Claude Cowork および Claude.ai の既定は「Medium」であり、通常利用でこの数値が再現されるわけではない [1]。
- 第三に、独立評価との齟齬がある。Artificial Analysis は Fable 5.1 を自社の Intelligence Index で首位（66、Opus 5 が 63、GPT-5.6 Sol が 61）としつつ、最大エフォートでは出力トークンが約 1.7 倍、タスクあたりコストが前世代比 20%増であり、加えて誤答時にも回答を試みる頻度が上がるというハルシネーション関連の後退を指摘している [8,9]。キャッシュ読み出し値下げによるコスト削減の主張とは、少なくとも部分的に整合しない。
- 第四に、失敗の質である。Snorkel AI の評価は、Fable 5.1 が「検証した」と述べながら実際には検証していなかったケースと、作業がほぼ完了した最終段階での失敗（last-mile failure）を報告している [11,12]。前者は科学利用において特に深刻な失敗様式である。自らの作業を検証したと主張する能力と、実際に検証する能力の乖離は、監督コストを下げるどころか上げる。

なお、Vals AI は Harvey の Legal Agent Benchmark で Fable 5.1 が Fable 5 を下回ったと報告している [12]。総合スコアの上昇がタスク種別ごとの後退を覆い隠しうることの実例である。

3.3　論点 3 — 帰属：誰が何をしたのか

三つの実証例のうち、分子設計と計算生物学は制限モデルである Mythos 5.1 の成果であり、一般

提供される Fable 5.1 のものではない [1]。一般利用者が再現できるのは金星地形図の系統の作業だけである。この区別は発表本文には記載されているが、報道と要約の過程で失われやすい。

より本質的なのは、いずれの事例も「モデル単体の成果」ではないという点である。分子設計は、人間が標的を選び、オープンソースの設計・折りたたみツールを与え、外部機関が実験検証した結果である。金星地形図は、既存の観測データと既存の部分地図という与件のうえで、ニューラルネットワークの訓練という定義済みのタスクを遂行したものである。GPU カーネル最適化は、出力が同一であることが機械的に検証できるという、極めて恵まれた検証構造を持つ問題である。

三例に共通するのは、「良い問いが既に与えられており、成否が客観的に判定できる」という条件である。この条件下でエージェントは著しく有能だが、この条件こそが科学研究において最も希少な資源である。

他方、これに反する示唆もある。Anthropic が掲載した顧客コメントのうち、楽天メディカルの臨床研究プロジェクトのレビューにおいて、Fable 5.1 が他の三つのフロンティアモデルが見落としたギャップを発見し、追加検証を主張し、新たな仮説を提示して、放棄されていたデータセットを新しい研究方向に転換したという報告がある [1]。また Ramp は、38 時間の無人稼働で先行結果をラベルアーティファクトと診断して修正し、6 件の並列実験を起動したと述べている [1]。これらは認識的な働きに接近している。ただし、いずれもベンダーが選定した顧客の自己申告であり、検証可能な形では提示されていない。

3.4　論点 4 — 物理世界との接続

MHS の成果は、実行のエージェント化が物理層に及びつつあることを示している。同時に、その限界の記述は Anthropic 自身によって具体的になされており、この率直さは評価に値する [4]。

- Genentech の事例では、液体中の泡に起因する実行時エラーに対し、Claude の既定の反応は同じウェルでパラメータを変えて再試行することであった。これは液を撹拌してさらに泡を生じさせる。物理現象としての失敗をコードから推論できず、人間が「これは物理的な泡が原因であり、清浄なウェルへ移動して混合回数を減らす必要がある」と教える必要があった [4]。
- QuEra の事例では、物理ハードウェアに異常が生じた際、Claude はトラブルシューティングができなかった。装置の理解がプログラム的であって物理的でないためである。また、わずかでもリスクがあると判断した操作の前に人間の確認を待って停止するため、承認待ちで一

晩実験が止まることがあった。加えて、正しく実行させるには膨大な文脈を人間が与える必要があった[4]。

Anthropic 自身、「大規模言語モデルはテキストと画像を通じて物理世界を学習しているため、空間的・物理的推論には限界があり、専門家の監督が必要である」と明記している[4]。物理層における「エージェント化」は、統合コストの劇的な低下（数週間〜数か月から数時間へ）という形では確実に進行しているが、物理的因果の理解という点ではまだ達成されていない。

4. 中間結論：二つのエージェント化を分ける

以上を整理すると、「サイエンスエージェント化」という問いへの回答は、次のように分けて述べるのが正確である。

- 実行のエージェント化は、明確に進んだ。長時間の無人稼働、自己検証ループ、複数装置のオーケストレーション、統合コストの桁違いの低下。これらは 2026 年前半には見られなかった水準にある。
- 認識のエージェント化は、まだ達成されていない。ベンチマークはそもそもこれを測っていない。実証例はいずれも問いと検証構造が人間から与えられている。示唆的な事例は存在するが、いずれも検証可能な形では提示されていない。
- 物理のエージェント化は、インタフェース層では進み、因果理解の層では進んでいない。

したがって、「サイエンスエージェント化したのか」への回答は「実行については然り、判断については否」である。そしてここで注意すべきは、この構図が「まだ足りない」という話ではないということである。むしろ実行の律速が外れたことで、科学研究のボトルネックが移動した。今後の制約は、モデルの能力ではなく、生成された大量の結果を検証し、帰属させ、再現可能な形で記録する側の制度と体制に移る。

筆者は、これを科学における律速段階の移動と呼ぶべきだと考える。そしてこの移動こそが、知財実務にとっての本題である。

5. 知財実務への含意【分析・私見】

> 本章は筆者の分析・見解であり、Anthropic の発表内容から論理的に導かれる推論と、筆者の実務経験に基づく評価を含む。法令の解釈が定まっていない論点については、その旨を明示する。

5.1 発明者適格性と「実質的寄与」の立証

主要国において発明者は自然人に限られるという運用は、DABUS 事件群を経て概ね収斂している。もっとも、AI を利用して得られた発明について「人間のどの関与が発明者としての寄与に当たるか」は、なお解釈が固まっていない領域である。

本稿の分析が示唆するのは、この論点にとって幸運な事実である。すなわち、現在の AI エージェントが最も強いのは「良い問いが与えられ、成否が客観判定できる」局面であり、最も弱いのは「問いを立てる」局面と「物理的因果を判断する」局面である。したがって、発明の着想、標的や課題の選定、実験設計、そして結果の意味づけという、発明者性の中核を構成する要素は、当面のあいだ人間の側に残る蓋然性が高い。

問題は、それが「残る」ことではなく、「残ったことを立証できるか」である。エージェントが数十時間を無人で走り、数百回の試行を行い、成果物だけが提出される研究プロセスにおいて、人間の寄与は本人の記憶以外のどこにも残らない可能性がある。ここが実務上の急所である。

5.2　記録可能性（auditable artifacts）が実務の中核へ

注目すべきは、Anthropic がこの問題を（知財の文脈ではないにせよ）製品仕様として先取りしている点である。Claude Science は、各出力の背後にあるコード・実行環境・メッセージ履歴を伴う監査可能な成果物を生成すると明記している [5,6]。MHS においても、エージェントの各操作は装置の状態辞書を通じて記録される [4]。

これらのログは、本来は再現性のための仕組みだが、知財実務から見れば発明経緯の証拠そのものである。具体的には、（a）誰がどの時点でどの指示（プロンプト）を与えたか、（b）エージェントが提示した選択肢のうち人間が何を選び何を却下したか、（c）人間が物理的・化学的知見を注入した箇所はどこか、が特定できれば、発明者性の立証は現在のラボノートと同等以上の水準になりうる。

逆に言えば、これらを取得しない運用で AI 主導の研究を進めることは、将来の発明者適格性の争いにおいて自ら証拠を捨てることに等しい。筆者の見るところ、日本企業の多くはこの点への備えが遅れている。

5.3　実施可能要件・サポート要件

GPU カーネル最適化の事例は示唆的である。出力が同一であることが機械的に検証できるため、成果は明白である。しかし、なぜその最適化が有効なのかという機序の説明が、生成過程に含まれているとは限らない。

タンパク質バインダー設計についても同様の構図がある。ヒット率 50%という結果は明確だが、その配列がなぜ結合するのかの説明可能性は別問題である。明細書の記載要件は「当業者が実施できること」を求めるのであって「発明者が理由を理解していること」を求めるわけではないから、直ちに拒絶理由になるとは限らない。しかし、進歩性の主張、実施例からの外挿範囲の設定、そして無効審判における反論の局面では、機序の理解の有無が実質的な差になる。

実務的な含意は明快である。AI エージェントによる成果を出願に載せる場合、成果そのものの再現手順（使用したモデル、バージョン、エフォート設定、シード、ツールチェーン、データ）を記録し、可能であれば機序についての人間側の理解を別途構築しておくべきである。なお、この論点についての審査実務の取扱いは各国とも形成途上であり、断定は避ける。

5.4　先行技術の意図的な膨張

見落とされがちだが、戦略的に最も重要なのはこの点かもしれない。Anthropic は金星地形図を Creative Commons ライセンスで公開し [1,17]、計算生物学の GPU カーネル最適化についてもオープンソース化を予定していると述べている [1]。MHS についても、リサーチプレビュー終了後にオープンソース化する方針が示されている [4]。

これは善意の公共貢献であると同時に、機能としては先行技術の大量生成である。AI エージェントが低コストで生成できる技術的成果が、無償で公知化されていく。企業側から見れば、自社が出願を検討していた領域が、出願前に他者の善意によってパブリックドメイン化されるリスクが構造的に高まる。

筆者はかねて、IP ランドスケープの実務において「特許化されない技術知識の流通量」を過小評価すべきでないと主張してきたが、生成 AI による成果の公知化はその傾向を桁で加速させる可能性がある。特許調査の設計において、特許文献に加えてプレプリント、GitHub リポジトリ、Zenodo 等のデータリポジトリ、そして AI ベンダーの技術ブログを継続監視対象に含めることは、もはや任意ではない。

5.5　アクセスの非対称性という競争条件

最も高い科学研究能力を持つ Mythos 5.1 は、米国政府と共同で構築された審査制プログラムを通じ、現時点では米国内の一部組織にのみ提供される [1]。Anthropic は米国政府と協調して国内外のより広範なパートナーへ拡大するとしているが [1]、時期は明示されていない。

三つの実証例のうち二つが Mythos 5.1 のものであったことを想起すれば、この差は象徴的ではなく実質的である。日本企業は当面、ライフサイエンス研究開発に関わる問い合わせが Opus 系モデル

に転送される条件下で競争することになる[1]。

この非対称性は、単に「使えるツールが劣る」という問題ではない。研究開発における AI 利用の記録・ノウハウ・組織学習の蓄積速度に差が生じ、それが数年後の出願品質と出願速度の差として現れる。安全保障を理由とする輸出管理と、研究開発競争力の間のトレードオフは、日本の産業政策と知財政策が正面から扱うべき論点である。

なお、Fable 5 と Mythos 5 については、2026 年 6 月に米国商務省の輸出管理を理由として一時的にアクセスが停止され、同年 6 月 30 日に規制が解除されて 7 月 1 日にアクセスが復旧した経緯がある[22]。この種の政策的変動が今後も生じうることを、事業計画の前提に織り込む必要がある。

5.6　ウォーターマークと帰属

Fable 5.1 および Mythos 5.1 は、EU AI Act の透明性行動規範に基づき、テキストおよびファイル出力に不可視の統計的ウォーターマークを埋め込む最初の Claude モデルである[1,14]。複数の妥当な語選択がある場合の選択に影響を与える方式で、コピー・ペーストや軽微な編集を経ても検出可能であることが意図されている[14]。検出 API は規制当局、報道機関、ファクトチェッカー、研究機関等の適格組織に対しプライベートプレビューで提供される[1]。

知財実務から見れば、これは二面性を持つ。一方で、明細書の下書きやクレーム案に AI 生成部分が含まれる場合、その事実が第三者に検出可能になりうる。他方で、自社の営業秘密や未公開の技術情報が AI 経由で流出した場合の追跡手段にもなりうる。企業としては、外部提出書類の作成フローにおける AI 利用の記録と、検出可能性を前提とした社内規程の整備が必要になる。この領域は制度・技術の双方が動いており、確定的な評価は時期尚早である。

6. 日本企業の知財部門への提言

以上を踏まえ、実務として着手すべき事項を五点に絞って挙げる。

- （1）AI 研究プロセスのログ取得を、再現性のためではなく発明経緯の証拠として設計し直すこと。指示内容、選択と却下、人間による知見の注入点を、後から追跡可能な形で残す。監査可能な成果物を生成する製品（Claude Science 等）を採用する場合は、その機能を知財部門の要件として明示的に発注仕様に含める。
- （2）発明者認定の社内基準を、AI 利用を前提に改訂すること。「何を着想し、何を選び、何を判断したか」を記述させる様式に変更する。現行の多くの様式は、AI が実行した部分と人間が判断した部分を区別できる構造になっていない。

- （3）先行技術調査の対象範囲を拡張すること。特許文献に加え、プレプリントサーバ、コードリポジトリ、データリポジトリ、主要AIベンダーの技術発表を継続監視対象とする。AI生成成果の公知化速度は、従来の学術論文の公表サイクルとは異なる時間軸で動く。
- （4）ベンダーのベンチマーク数値を、そのまま経営判断の根拠にしないこと。エフォート設定、標準誤差、測定主体、対象タスクの性質を確認する習慣を、知財部門から発信する。本稿で見たとおり、同一の数値が文脈次第で意味を変える。
- （5）モデルアクセスの非対称性を、事業リスクとして経営に報告すること。最上位モデルへのアクセスが地政学的条件に左右される現状は、研究開発計画の前提条件である。代替手段の確保と、政策動向の継続監視を体制に組み込む。

7. 結論

Claude Fable 5.1 は「サイエンスエージェント化」したのか。実行の主体としては、相当程度そうなった。判断の主体としては、まだそうなっていない。そしてこの区別は、Anthropic 自身の発表文とベンチマーク運営主体の設計思想の双方が、明示的に前提としているものである。

しかし、この結論を「まだ大したことはない」と読むのは誤りである。実行の律速が外れたことの帰結は、科学研究におけるボトルネックが、能力の側から検証・帰属・記録の側へ移動したということである。

知的財産制度は、まさにこの移動した先の領域を扱う制度である。誰が何を発明したのかを確定し、それを公開の対価として保護し、再現可能な形で社会に伝える。AI エージェントが実行を担う時代において、知財部門の役割は縮小するのではなく、むしろ研究開発プロセスの設計そのものに前倒しで関与する方向へ拡大する。

筆者は、知財部門が「出願の下流工程」に留まっている企業と、「研究プロセスの記録設計」まで踏み込む企業との差が、今後数年で決定的になると考えている。Fable 5.1 の発表は、その分岐点が既に到来していることを示す一つの証拠である。

参考文献

※ URL は 2026 年 9 月 4 日時点。1〜4 は筆者が本文取得により内容を直接確認した一次資料。5〜25 は検索結果および二次報道に基づくもので、一部 URL は直接取得による検証を行っていない（該当箇所に「未検証」と付記）。

1. Anthropic「Introducing Claude Fable 5.1 and Claude Mythos 5.1」2026 年 9 月 1 日. https://www.anthropic.com/claude-fable-and-mythos-5-1（内容確認済み）
2. Terminal-Bench-Science Team（執筆：Steven Dillmann）「TERMINAL-BENCH-SCIENCE 0.1 — Evaluating AI agents on scientific research workflows」tbench.ai. https://www.tbench.ai/news/terminal-bench-science-0-1（内容確認済み）
3. harbor-framework/terminal-bench-science（GitHub リポジトリ）. https://github.com/harbor-framework/terminal-bench-science（未検証）
4. Anthropic「Previewing the Model Hardware Standard」2026 年 8 月 27 日. https://www.anthropic.com/news/model-hardware-standard-research-preview（内容確認済み）
5. Anthropic「Expanding our support for scientists」2026 年 8 月 27 日. https://www.anthropic.com/news/expanding-support-for-scientists（未検証）
6. Anthropic「Claude Science, an AI workbench for scientists」2026 年 6 月 30 日. https://www.anthropic.com/news/claude-science-ai-workbench（未検証）
7. MIT Technology Review「Claude Science is Anthropic's newest flagship product」2026 年 6 月 30 日. https://www.technologyreview.com/2026/06/30/1139987/claude-science-is-anthropics-newest-flagship-product/（未検証）
8. THE DECODER「Anthropic's Claude Fable 5.1 promises better coding and research at up to 45 percent less」2026 年 9 月. https://the-decoder.com/anthropics-claude-fable-5-1-promises-better-coding-and-research-at-up-to-45-percent-less/（未検証）
9. R&D World「Anthropic doubles a science benchmark score with Fable 5.1 while OpenAI says its Astra model crosses critical cyber threshold」2026 年 9 月. https://www.rdworldonline.com/（未検証）
10. Decrypt「Anthropic Ships Claude Fable 5.1, More Than Doubling Its Predecessor on Key Benchmark」2026 年 9 月 1 日. https://decrypt.co/377078/anthropic-claude-fable-5-1（未検証）
11. Snorkel AI「Fable 5.1 vs Opus 5: Coding Benchmark Results」2026 年 9 月. https://snorkel.ai/blog/fable-5-1-vs-opus-5-coding-benchmark/（未検証）
12. emergent.sh「Claude Fable 5.1 Benchmarks: Scores and What They Mean」2026 年 9 月. https://emergent.sh/learn/claude-fable-5-1-benchmarks（未検証）
13. The Neuron「Everything to know about Fable 5.1, Anthropic's new Claude model」2026 年 9 月. https://www.theneuron.ai/news/claude-fable-5-1-cuts-agent-costs-and-safety-friction/（未検証）
14. TechSpot「Anthropic releases Claude Fable 5.1, cuts cached token pricing by 75%」2026 年 9 月. https://www.techspot.com/news/113701-anthropic-releases-claude-fable-51-stronger-coding-science.html

（未検証）

15. Vellum「Claude Fable 5.1 & Claude Mythos 5.1 Benchmarks Explained」2026 年 9 月. https://www.vellum.ai/blog/claude-fable-5-1-mythos-5-1-benchmarks-explained（未検証）
16. Anthropic「Claude Fable」モデルページ. https://www.anthropic.com/claude/fable（未検証）
17. Zenodo「Venus digital elevation model (Anthropic)」. https://zenodo.org/records/22164484（未検証。Anthropic 発表本文中のリンクによる）
18. Snorkel AI「Terminal-Bench-Science」リーダーボード解説. https://snorkel.ai/leaderboard/terminal-bench-science/（未検証）
19. Quartz「Anthropic Model Hardware Standard connects AI to lab equipment」2026 年 8 月. https://qz.com/anthropic-model-hardware-standard-ai-robots-lab-equipment-082826（未検証）
20. Unite.AI「Anthropic Opens 10,000 Free and Discounted Claude Seats for Scientists」2026 年 8 月 27 日. https://www.unite.ai/anthropic-opens-10000-free-and-discounted-claude-seats-for-scientists/（未検証）
21. MarkTechPost「Anthropic Releases Claude Fable 5.1 and Claude Mythos 5.1: 52.6% on Terminal-Bench-Science and 75% Cheaper Cache Reads」2026 年 9 月 1 日（未検証）
22. Anthropic「Fable / Mythos モデルへのアクセスに関する声明」2026 年 7 月. https://www.anthropic.com/news/fable-mythos-access（未検証。2026 年 6 月 12 日の輸出管理に伴うアクセス停止と同月 30 日の規制解除、7 月 1 日の復旧に関する記載）
23. Anthropic「Claude Fable 5.1 and Claude Mythos 5.1 System Card」2026 年 9 月. https://www.anthropic.com/claude-fable-5-1-mythos-5-1-system-card（未検証）
24. DataCamp「Claude Fable 5.1: Features, Benchmarks, and Pricing」2026 年 9 月. https://www.datacamp.com/blog/claude-fable-5-1（未検証）
25. JST-CRDS「AI for Science Spotlights Vol.2 AI エージェントの科学研究への展開」2026 年 8 月 7 日. https://www.jst.go.jp/crds/column/kaisetsu/shortreport2.html（参考。本稿の問題意識の背景として）