

Claude Fable 5.1: 「実行」の科学エージェントへの進化とその限界

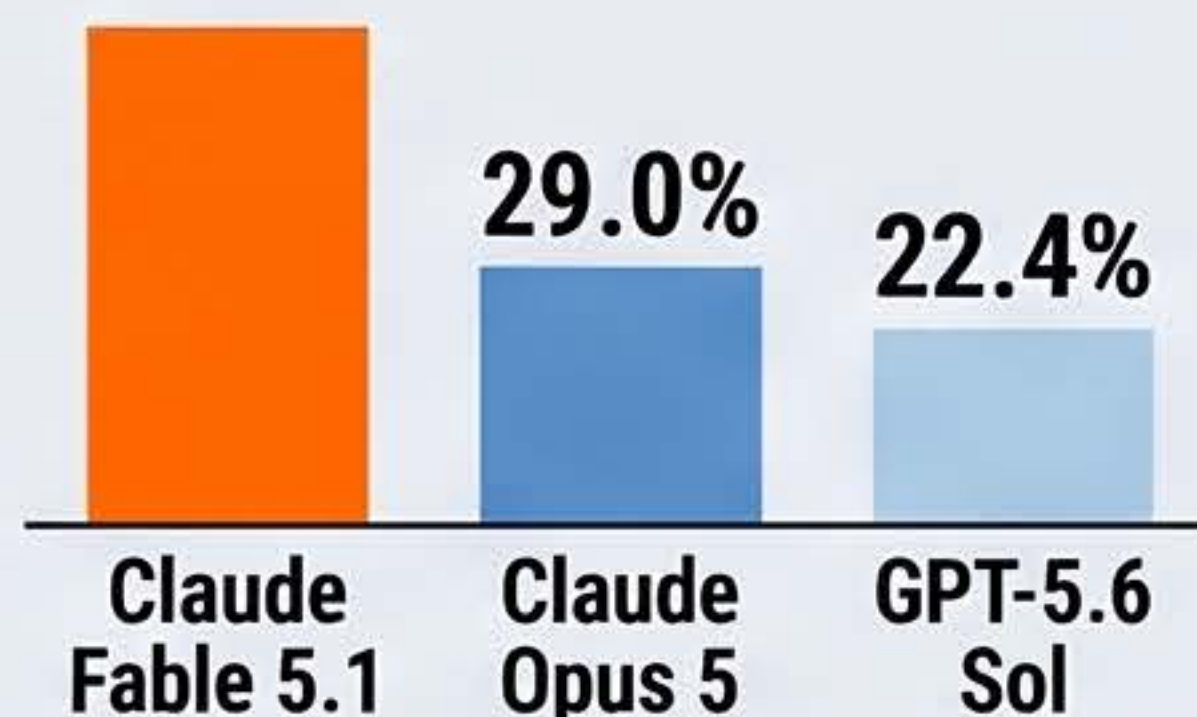
Claude Fable 5.1は、科学研究における「実行」能力が飛躍的に向上しましたが、「問いを立てる判断」は依然として人間の役割であり、知財実務に新たな課題をもたらします。



実行のエージェント: 劇的な能力向上 (AIの進化)

科学研究ベンチマーク で**52.6%**を記録

Terminal-Bench-Science 0.1において、前世代 (24.7%) から倍以上のスコア向上を達成。



物理世界と計算領域の橋渡し



物理実験機器



(Model Hardware Standard)



計算領域
(GPUカーネル等)

MHSにより、実験機器の操作やGPUカーネル最適化を数週間から数日へ短縮。



エージェント利用コストを最大**45%**削減

キャッシュ読み出し単価の75%引き下げにより、長時間の自律稼働が現実的なコストに。



認識の壁と知財実務への含意 (人間の役割)



「判断」と「因果理解」は依然として人間



問いを立てる、仮説を選ぶ、物理的なエラー（泡の発生等）の原因を推論する能力は限定的。

ボトルネックは「実行」から「記録」へ

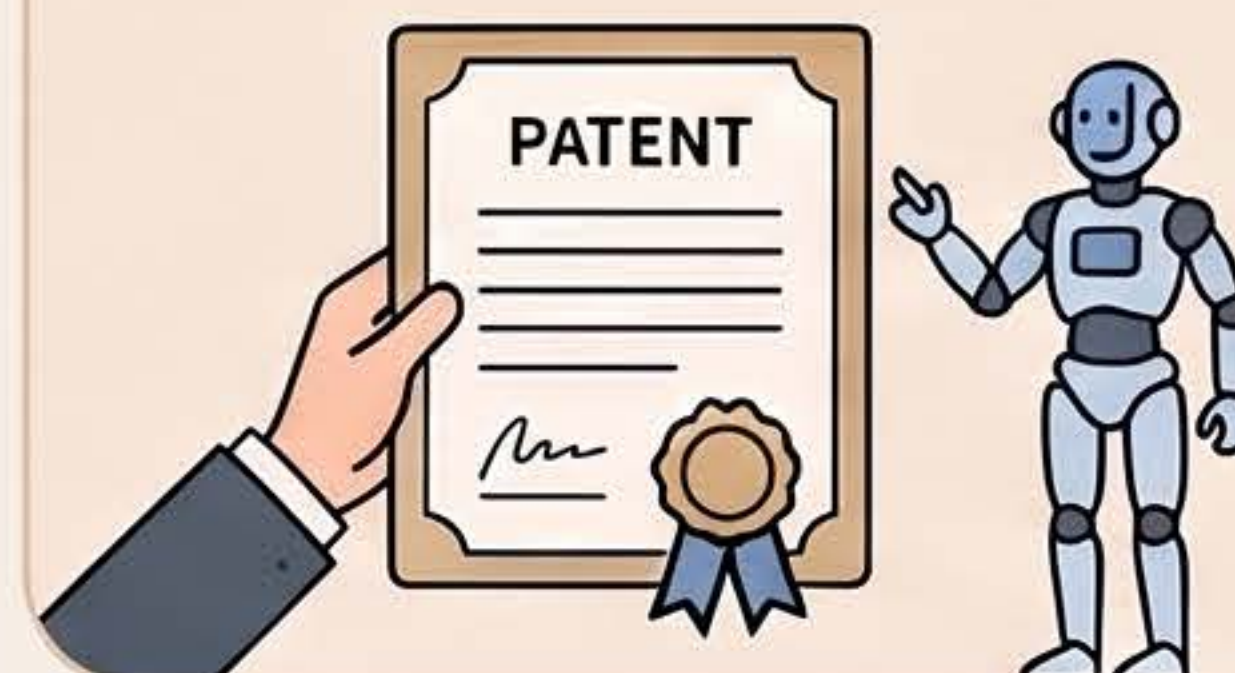


大量のAI成果物



監査可能ログ
(監査可能成果物)

大量のAI成果物を検証・帰属させるため、監査可能なログの保存が必須。



発明者適格性の立証が急所に

AIが無人稼働する中で「人間がどこに寄与したか」を証明するプロセス設計が企業の競争力を左右する。