



2026-06-20

MORNING DISPATCH / Vibe Coder Bootcamp Tech News

今朝のホットな話題

- 1 Claude Code に「Artifacts」登場 — セッションから対話的ページを自動生成
- 2 OpenAI Codex 「Record & Replay」 — 実演して教えればスキルになる
- 3 OpenAI、長時間・高ステークスのタスク向けに『望ましく安全な挙動』をRL学習

7 トピックを整理。 



🔍 何が起きた？

Anthropic が Claude Code に Artifacts を追加した。作業セッションのフルコンテキスト（コードベース・plugins・skills・接続済みツール）を踏まえて、PR ウォークスルーのような「対話的なページ」をその場で生成する。

🚩 主な変更点

- セッションのフルコンテキスト（codebase / plugins / skills / connected tools）を参照して生成
- PR ウォークスルー等のインタラクティブなページを自動生成
- 作業を続けると artifact が自動リフレッシュ → 共有先も常に最新を閲覧
- 公式投稿は ❤️ 16,000 超で当日最大級の反応

💡 なぜ重要？

コードベースやツール接続まで含む作業文脈が、そのまま共有可能な理解用ページになる。レビュー、引き継ぎ、進捗共有で「説明を別途作る」負担を減らす。

セッション → Artifact → 共有



自動リフレッシュ

作業を続けると artifact が自動リフレッシュ
→ 共有先も常に最新を閲覧

❤️ **16,000 超**
公式投稿の反応



1 何が起きた？

OpenAI が Codex に Record & Replay を追加した。繰り返し行うワークフローを一度「お手本」として実演して見せると、それを再利用可能な skill として記録できる。



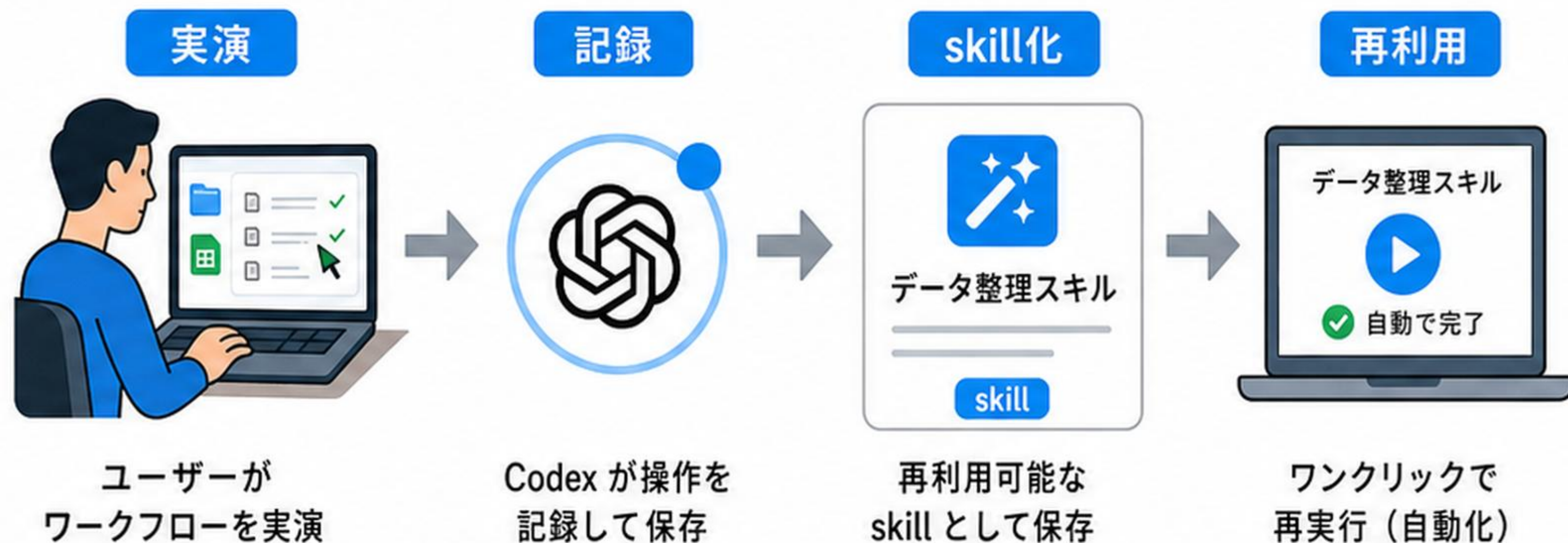
2 主な変更点

- ワークフローを一度実演 → 再利用可能な skill として保存
- コードを書くのではなく「操作を見せて教える」方式
- 公式docsに /codex/record-and-replay ページが追加され正式機能
- @OpenAIDevs ❤️ 12,000超 / gdb の補足投稿も ❤️ 2,700超



3 なぜ重要？

Greg Brockman は「デモでCodexに教えられる」と表現。反復作業を、コード化せずに実演からスキル化できる点大きい。



❤️ 12,000超

@OpenAIDevs

Codex Record & Replay の発表投稿に対するいいね数

❤️ 2,700超

gdb

gdb による補足投稿に対するいいね数



</codex/record-and-replay>
公式docs

Record & Replay の使い方や詳細が確認できるページ

Before



毎回同じ作業を手作業で繰り返す...



時間がかかり、ミスも発生しやすい

After



skill をワンクリックで再利用



高速・安定・再現性の高い自動化を実現

OpenAI、長時間・高ステークスのタスク向けに『望ましく安全な挙動』をRL学習

🔍 何が起きた？

OpenAI は、AI がより長く高ステークスなタスクを担う流れを踏まえ、現実的な会話データを用いた強化学習でモデルに「有益で安全な挙動」を定着させる手法を発表した。

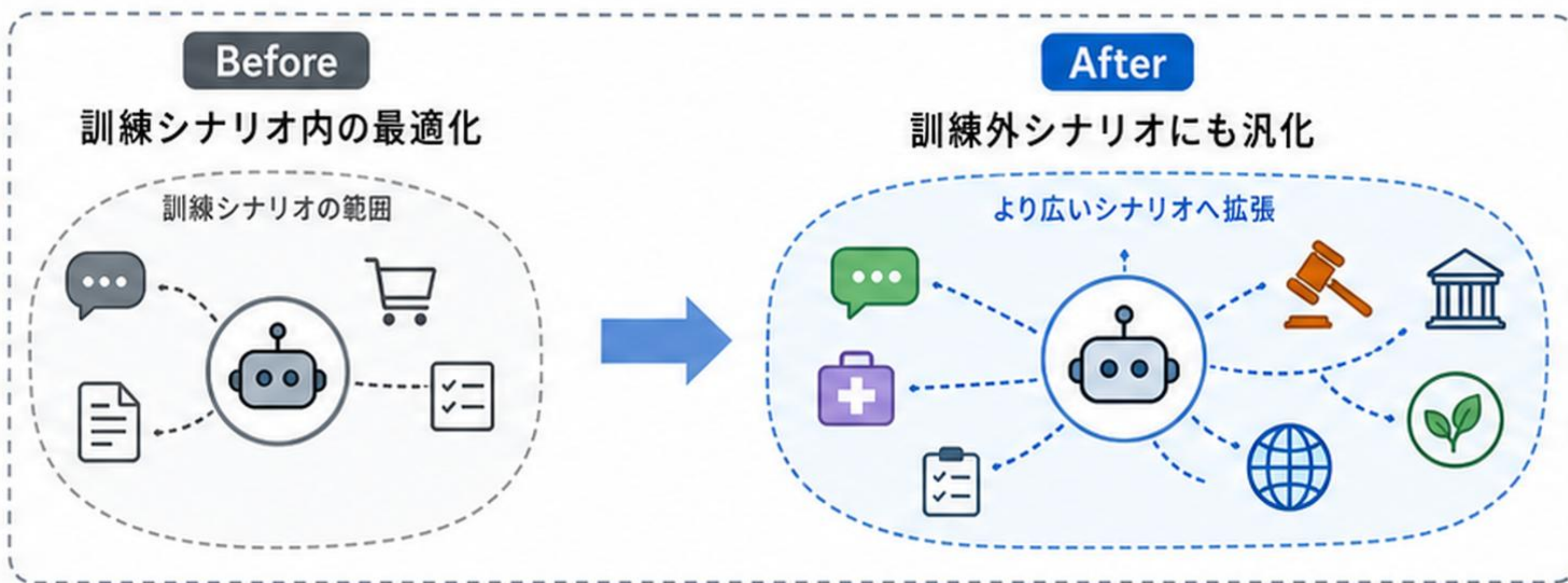
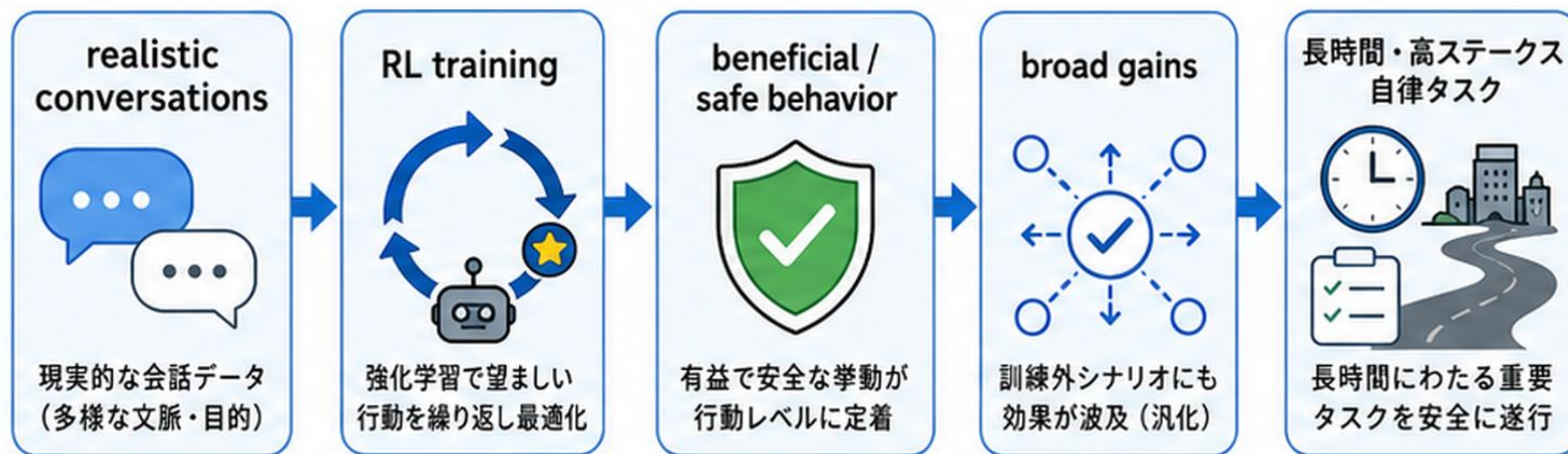
🚩 主な変更点

- realistic conversations を使った RL で beneficial/safe behavior を強化
- 少量データで訓練外シナリオにも効果が波及 (broad gains)
- 長時間・高ステークスな自律タスクを見据えた安全設計

💡 なぜ重要？

少量の学習データでも、訓練シナリオを超えた広範な改善 (汎化) が得られたとしている。高ステークスな自律タスクで、安全性を行動レベルに定着させる設計が重要になる。

全体の流れ



🔍 何が起きた？

OpenAI が、実世界のライフサイエンス研究をどれだけ支援できるかを測る新ベンチマーク「LifeSciBench」を公開した。専用モデル GPT-Rosalind は7つのワークフロー全てで GPT-5.5 を上回った。別途、GPT-5.4 が文献レビューから検証済み実験までの医薬化学プロジェクトを約2.5ヶ月で推進した事例も示された。

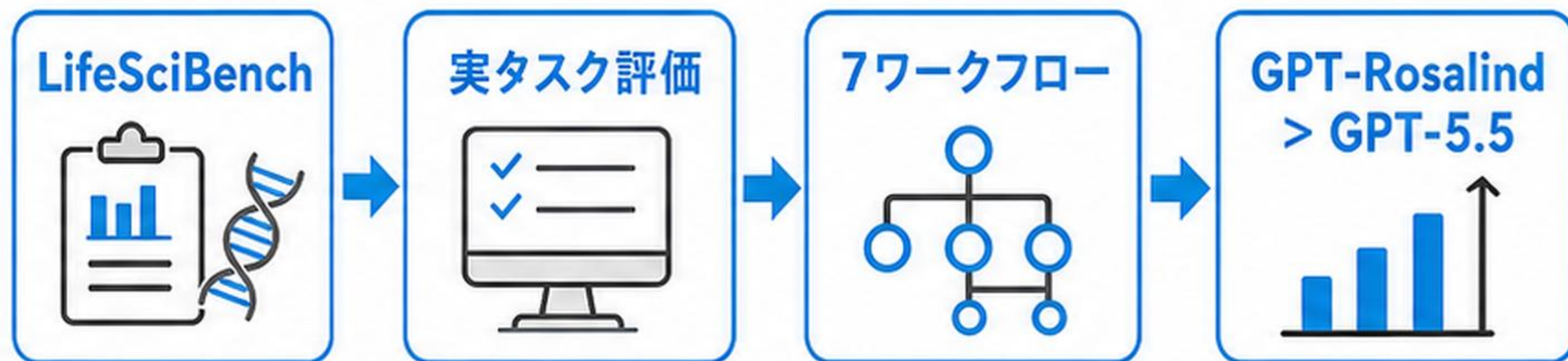
🚀 主な変更点

- LifeSciBench は知識クイズでなく「実タスクの遂行」を評価する設計
- GPT-Rosalind が全7ワークフローで GPT-5.5 を上回る
- GPT-5.4 が文献レビュー→提案ランク付け→実験設計→検証まで医薬化学を駆動
- プロセスは約2.5ヶ月 + 人間化学者の執筆に半月

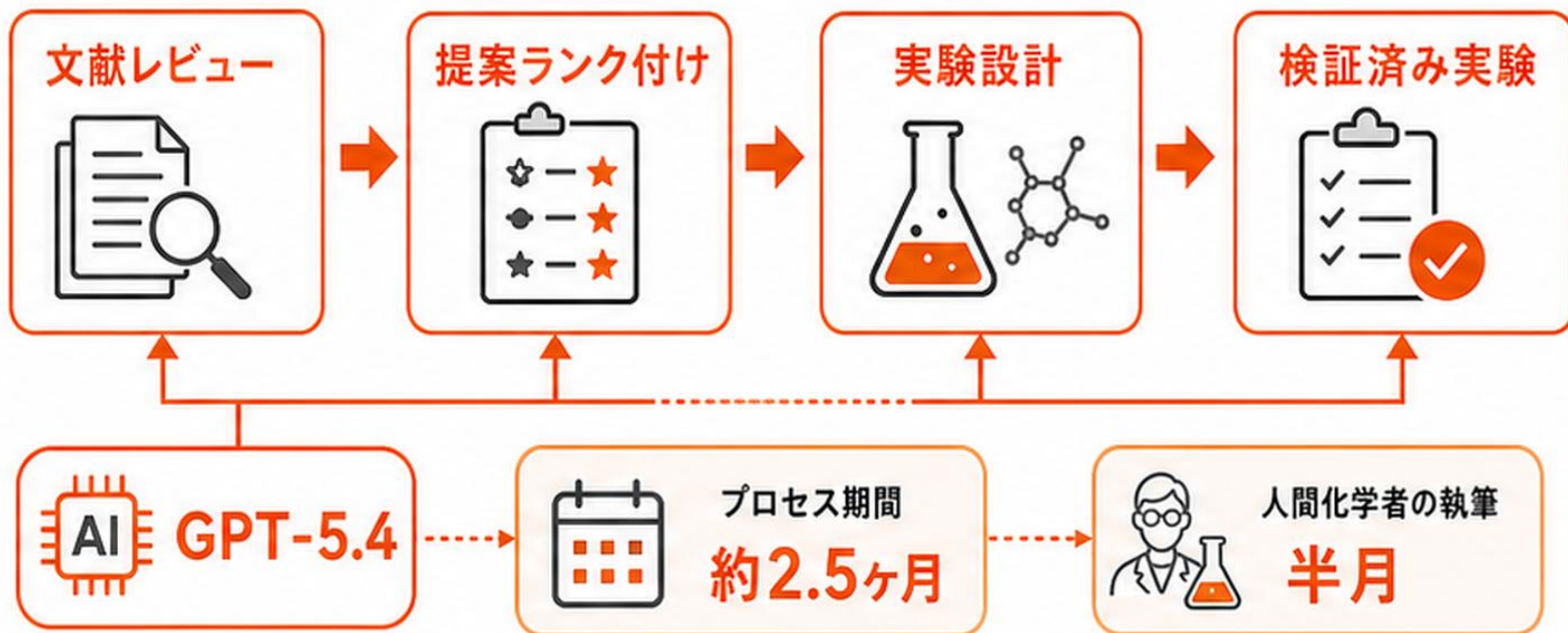
💡 なぜ重要？

研究AIの評価軸が、暗記問題から実験に接続するワークフロー遂行へ移る。創薬研究では、AIが文献調査、仮説生成、候補選定、実験設計、検証までを人間研究者と連携して進める可能性が示された。

LifeSciBench による評価パイプライン



GPT-5.4 が駆動する医薬化学プロジェクトフロー





🔍 何が起きた？

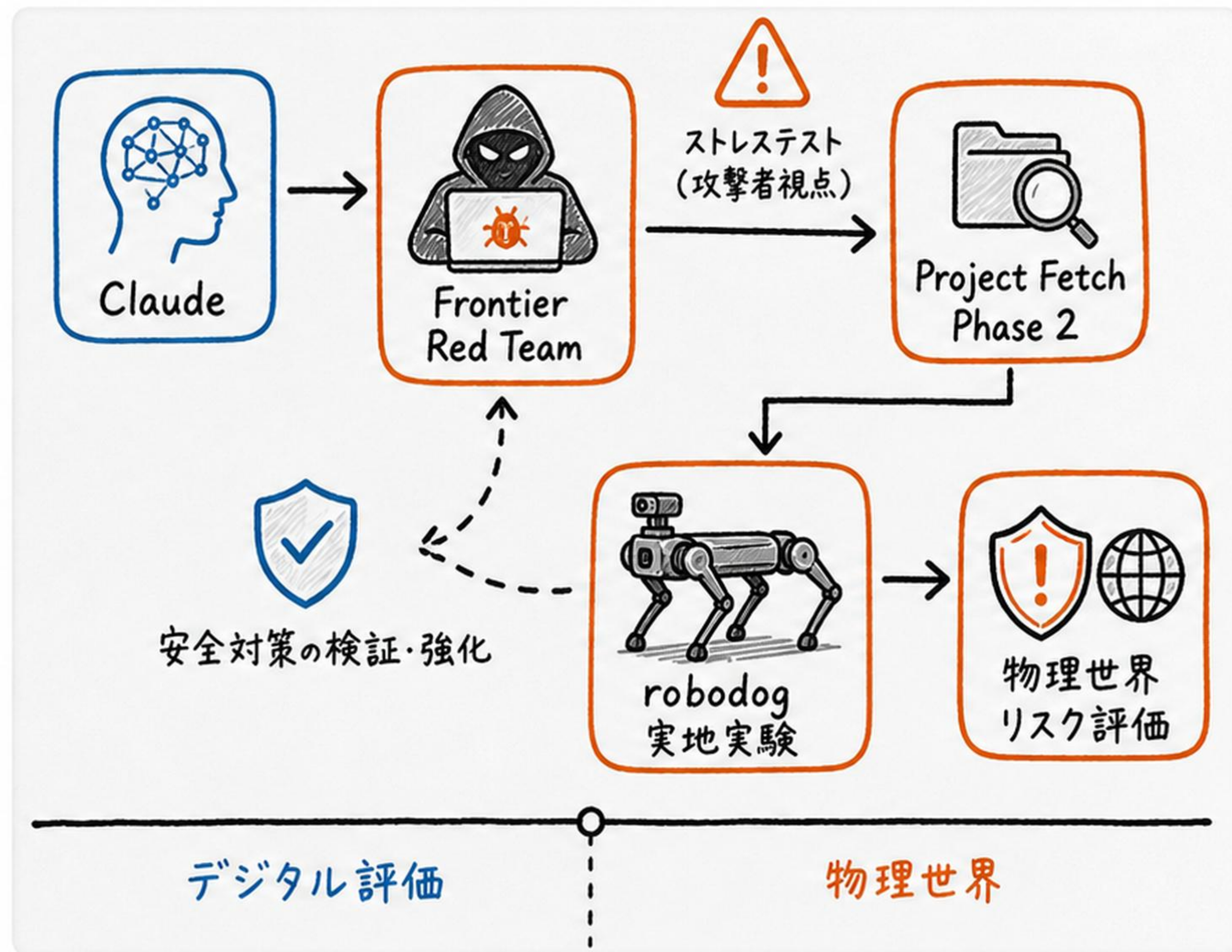
Anthropic の Frontier Red Team が Project Fetch の Phase 2 を公開。Claude が物理世界のタスクでどこまで能力を発揮しうるかを、赤チーム（攻撃者視点）でストレステストする安全研究の続報。

🚩 主な変更点

- Frontier Red Team による危険能力評価の続報
- robodog を用いた実地実験のデモ映像を公開
- フロンティアモデルの物理世界への波及をリスク側から測定

💡 なぜ重要？

AI の能力評価がデジタル環境だけでなく、ロボット犬の操作実験を含む物理世界へ広がる。安全研究は「できること」を示すだけでなく、悪用・危険能力を攻撃者視点で測る段階に入っている。



Cursor、background agent に GitHub トリガー + computer use を追加

Cursor公式 likes

🔍 何が起きた？

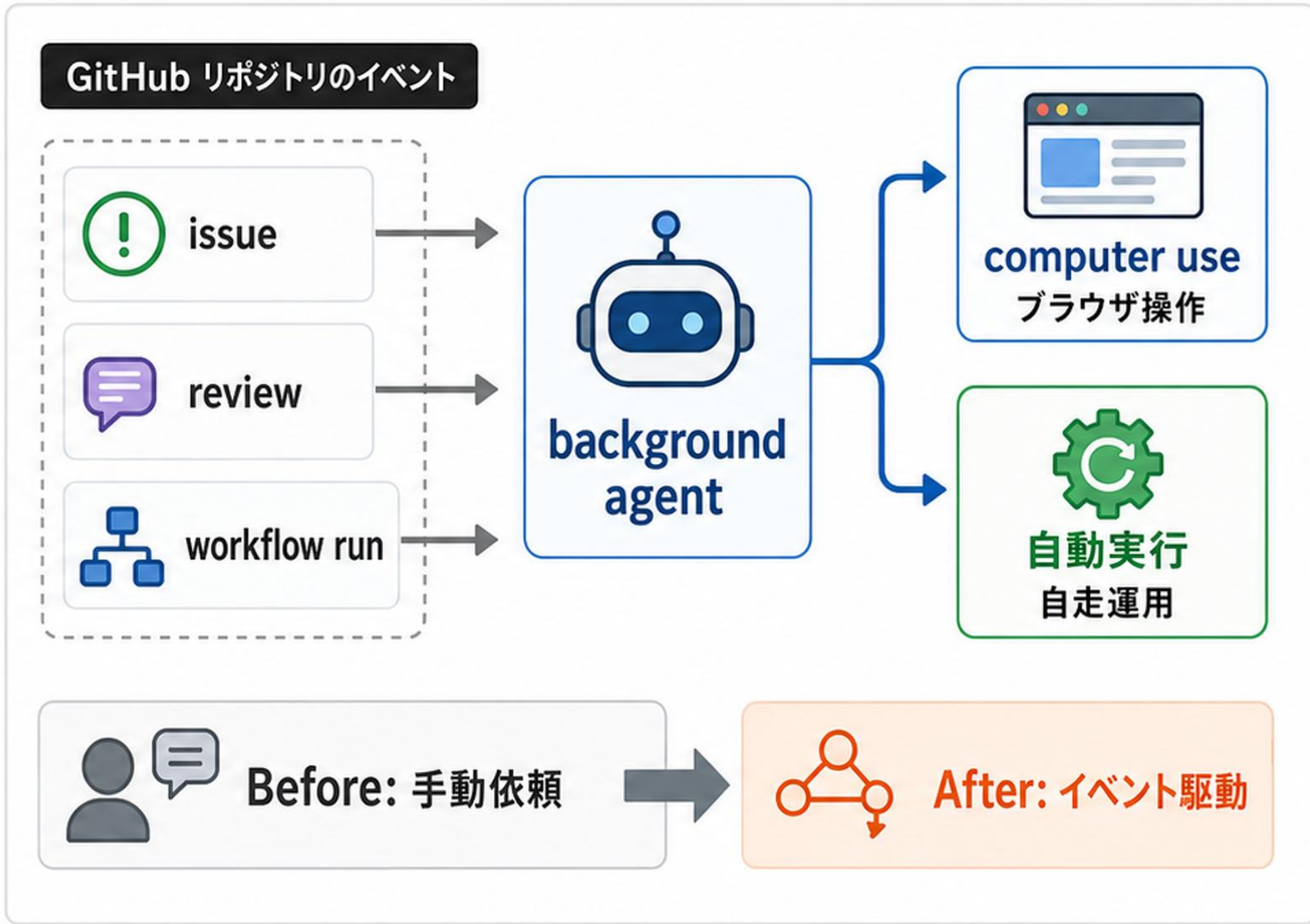
Cursor が background agent 向けに、GitHub の issue / review / workflow run をトリガーにする連携と、computer use（ブラウザ操作などPC操作）を追加した。リポジトリ上のイベントに反応してエージェントが自動で動くイベント駆動の運用が組めるようになる。

📌 主な変更点

- GitHub triggers: issues / reviews / workflow runs に反応
- エージェントの computer use 対応（ブラウザ等の操作）
- イベント駆動で自走するエージェント運用が可能に

💡 なぜ重要？

リポジトリ上のイベントから、実装・レビュー支援・ブラウザ操作までつながる。開発ワークフローを、手動依頼中心からイベント駆動の自動運用へ近づけるアップデート。



🔬 研究の核心

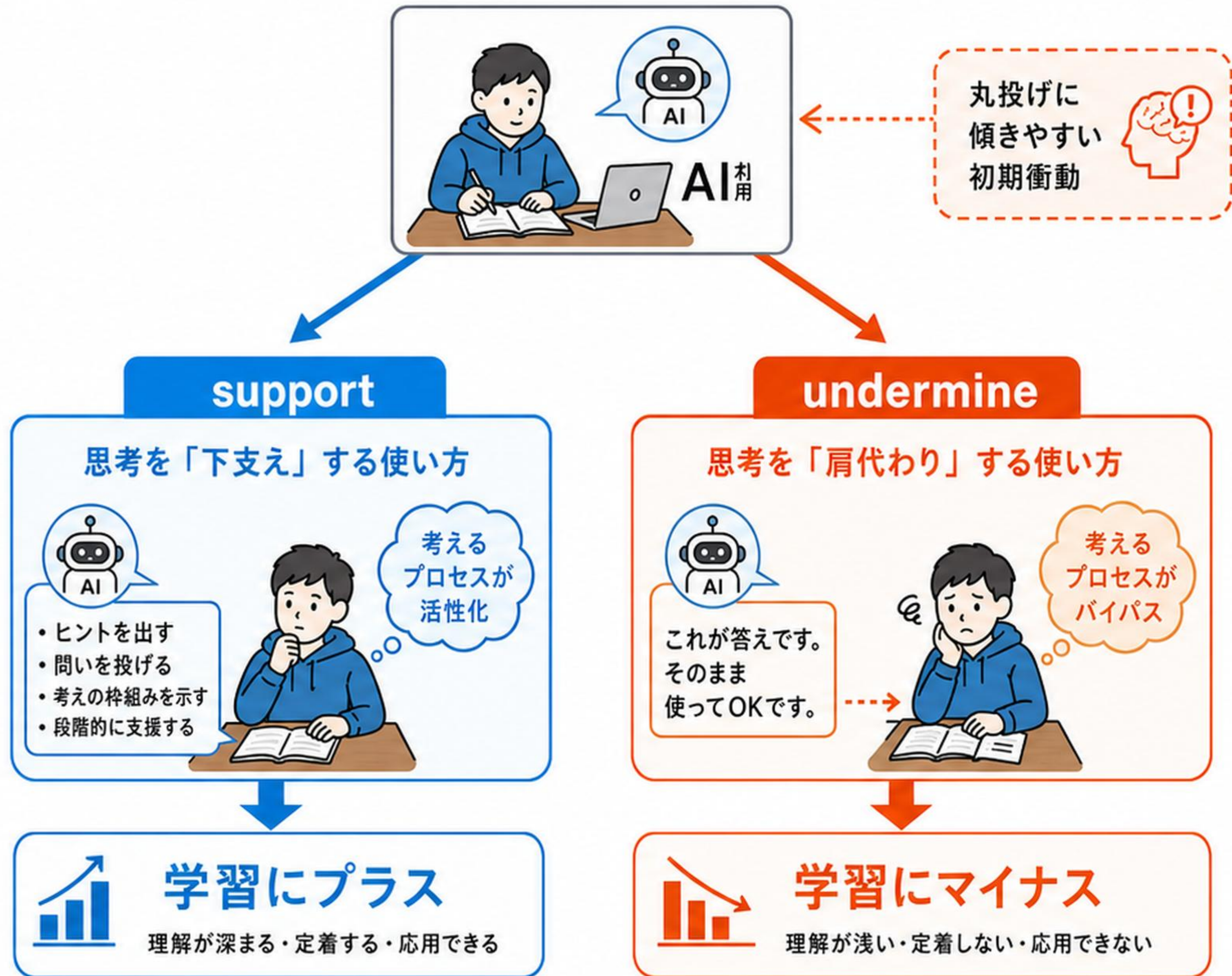
Ethan Mollick が、中国での大規模研究を引いて、AIが学習者の思考を「下支え (support)」するときは有益だが、思考を「肩代わり (undermine)」するときは学習を損なうという証拠を紹介した。

📊 主な知見

- AIが thinking を support する使い方 → 学習にプラス
- AIが thinking を undermine (肩代わり) する使い方 → 学習にマイナス
- 学生の初期衝動は「丸投げ」に傾きやすい
- 大規模スタディに基づく実証的知見

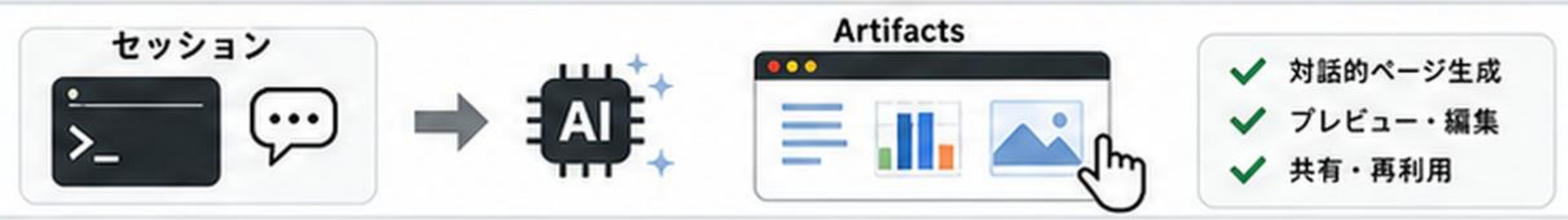
🎯 含意

学生はつい宿題をAIに丸投げしがちで、たとえ意図せずとも学びが削られる。AIは答えを代行する道具ではなく、考える過程を支える相棒として設計・利用する必要がある。



本日のトピック一覧

1 Claude Code に「Artifacts」登場
— セッションから対話的ページを自動生成



2 OpenAI Codex 「Record & Replay」
— 実演して教えればスキルになる



3 OpenAI、長時間・高ステークスのタスク向けに『望ましく安全な挙動』をRL学習



4 OpenAI 『LifeSciBench』 + GPT-Rosalind
— GPT-5.4 が創薬研究を駆動



5 Anthropic Frontier Red Team 『Project Fetch』 Phase 2
— 物理世界での能力評価



6 Cursor、background agent に GitHub トリガー + computer use を追加



7 大規模研究 (中国): AIは『思考を肩代わり』させると学習を阻害する

