

2026-07-28

1. Kimi K3 が公開翌日に Cursor と Ollama クラウドへ
2. Anthropic が open-weights の立場を公式表明
3. OpenAI 調査で AI 依頼の 43.5% が「他職種の仕事」

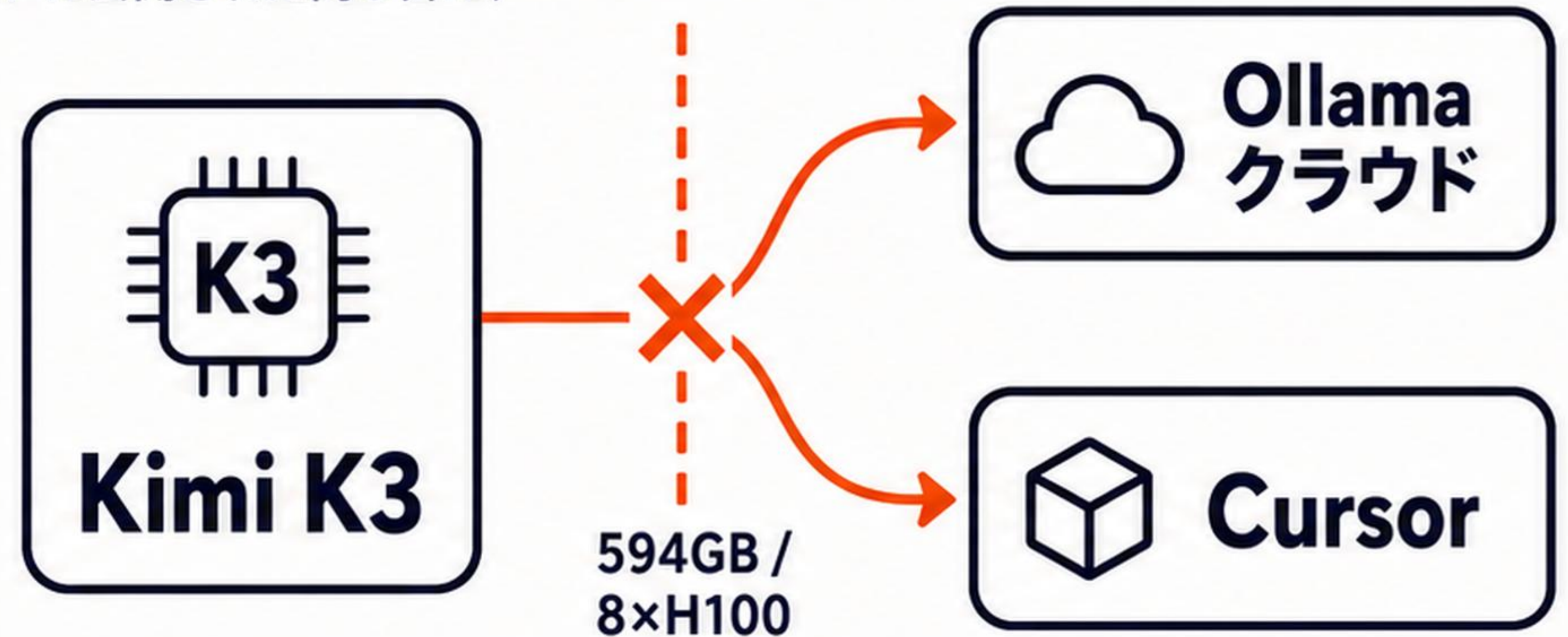
8トピック

Kimi K3 が公開翌日に Cursor と Ollama クラウドへ

「594GB / 8×H100」を踏まない経路ができた

Moonshot AI の Kimi K3 (2.8兆パラメータ MoE) の重みが 07-27 に公開された同じ日に、Ollama のクラウドと Cursor の両方で利用可能になった。

- Ollama クラウドで提供開始。Claude Code から使う場合は `ollama launch claude --model kimi-k3:cloud`
- Ollama は公式に「現時点で Pro または Max のサブスクリプションが必要、かつ追加の usage credits...
- Cursor でも利用可能に。Cursor は「CursorBench でフロンティアに近いスコア」と発表し、US ベースの推論であることに言及



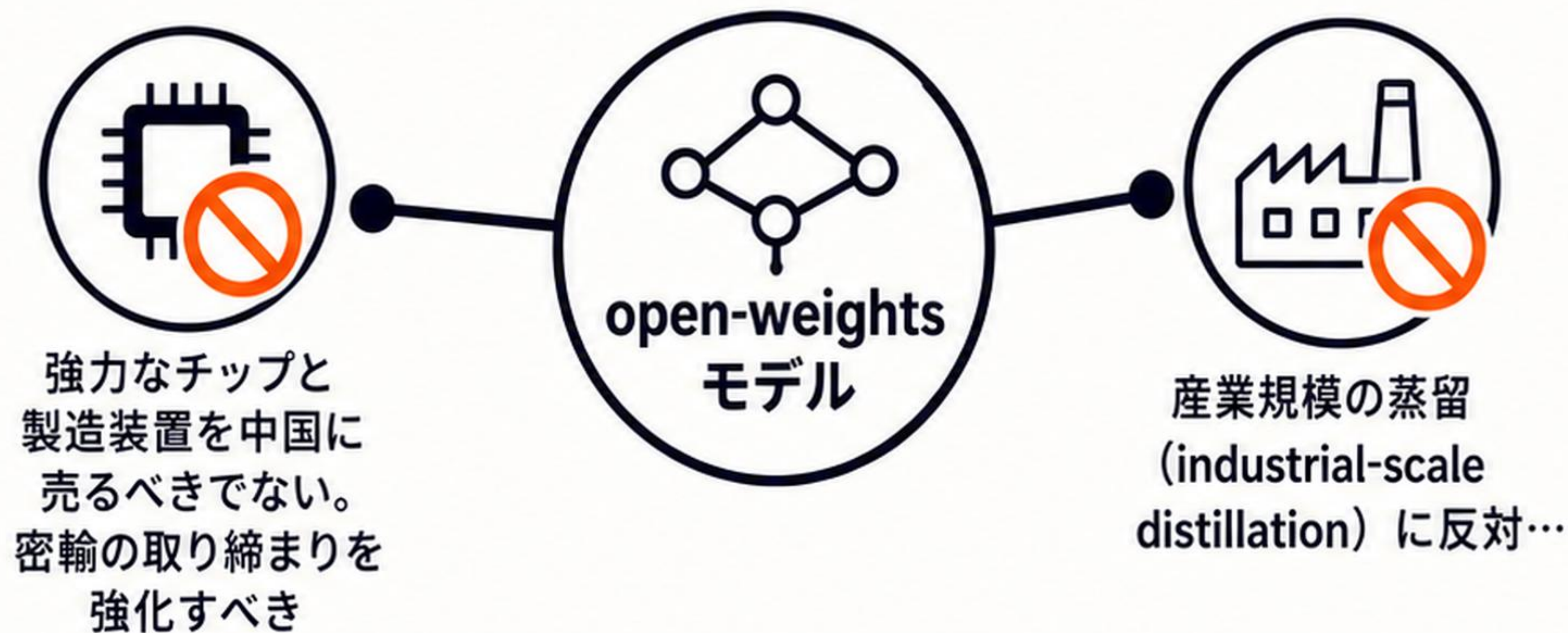
「オープンウェイトかどうか」と「自分で動かせるかどうか」は別の問題で、
実務で効くのは誰かがホストしてくれるかどうか だった、という答えが2日が出た。

Anthropic が open-weights の立場を公式表明

「禁止を主張したことは一度もない」、争点は3つ

Nvidia 主導の共同声明「Open Weights and American AI Leadership」に
OpenAI・Google を含む50社が署名するなか、Anthropic だけが署名せず孤立していた件について…

- 明言: 「Anthropic has never advocated for a ban on open-weights models」…
- 主張①: 強力なチップと製造装置を中国に売るべきでない。密輸の取り締まりを強化すべき
- 主張②: 産業規模の蒸留 (industrial-scale distillation) に反対…



「オープンウェイト賛成 vs 反対」の二項対立で語られていた論争が、Anthropic 側から「能力閾値での安全テスト義務化」と「蒸留の産業利用規制…」

OpenAI 調査で AI 依頼の 43.5% が「他職種の仕事」

職務の境界が溶けている

OpenAI が米国のビジネス利用者による ChatGPT メッセージ 80 万件超を分析したレポート「Work at the Frontier: How AI is Expanding What People Do at Work」を公開した。

- 分析対象は米国のビジネス利用者による **80 万件超** の ChatGPT メッセージ
- 全体で職種固有の AI 依頼の **43.5%** が他職種のタスク
- 職種別の越境率: カスタマーエクスペリエンス **77%** / デザイナー **75%** / 人事 **69%** / 法務 **56%** / ...



「AI が仕事を奪う」ではなく「職務記述書の外側へ手が届くようになった」という方向の数字が、規模をもって示された。

GPT-Live が ChatGPT Voice の組織プランへ全世界展開

OpenAI のリアルタイム音声モデル GPT-Live が、ChatGPT Voice を通じて Edu・Business・Enterprise の各プランへ全世界で提供開始された。

- 対象は Edu / Business / Enterprise プラン、提供はグローバル
- GPT-Live はフルデュプレックス（全二重）音声モデル。相手の発話終了を待たずに聞きながら話せる
- 経緯：07-08 に GPT-Live 発表 → 07-23 に Codex と ChatGPT デスクトップアプリへ展開 → 07-27 に組織プランへ



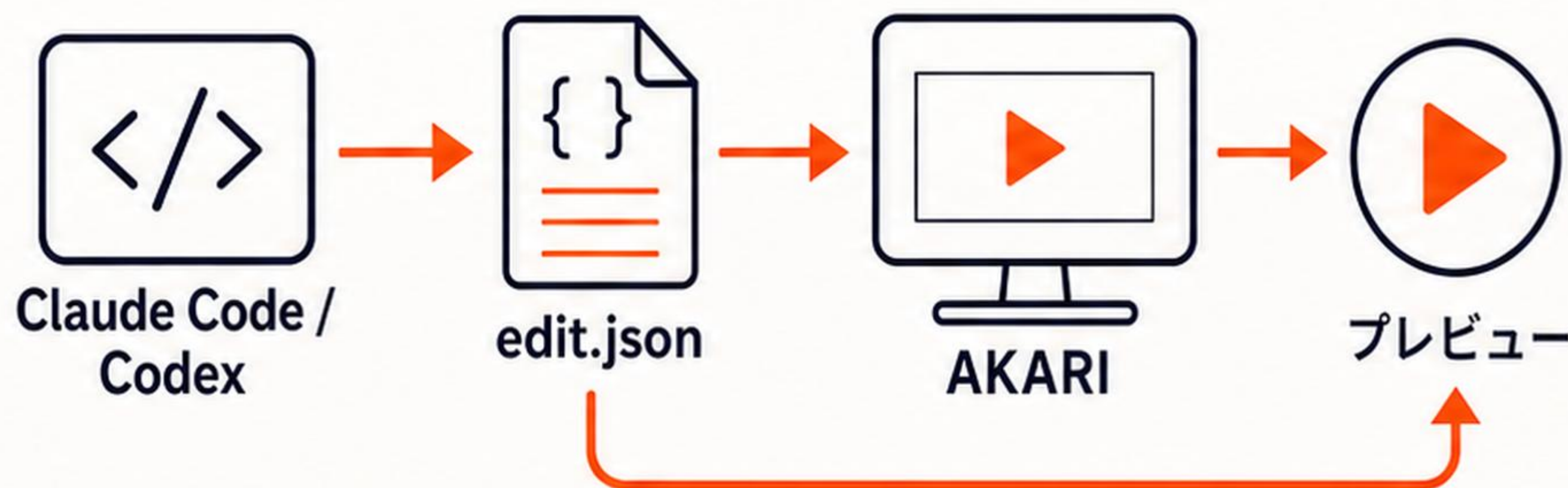
エージェントへの指示インターフェースがキーボードから会話へ移る局面。

Akari Video で Claude Code / Codex がそのまま編集チームになる

OSS の AI ネイティブ動画編集

Claude Code / Codex といったコーディングエージェントを「動画編集チーム」としてそのまま使う macOS 向け OSS アプリ、AKARI Video が Public Beta で公開された。

- **セーブデータ中心:** カット・テロップ・音声・3D の全編集状態が edit.json に宣言される…
- **専用 API を持たない:** 結合は「タブ＝ファイル、応答＝JSON と git 差分」のみ。よって編集の頭脳は…
- **プリセットを持たない:** テンプレート集の代わりに AI が毎回 HTML / CSS / Three.js を書き…



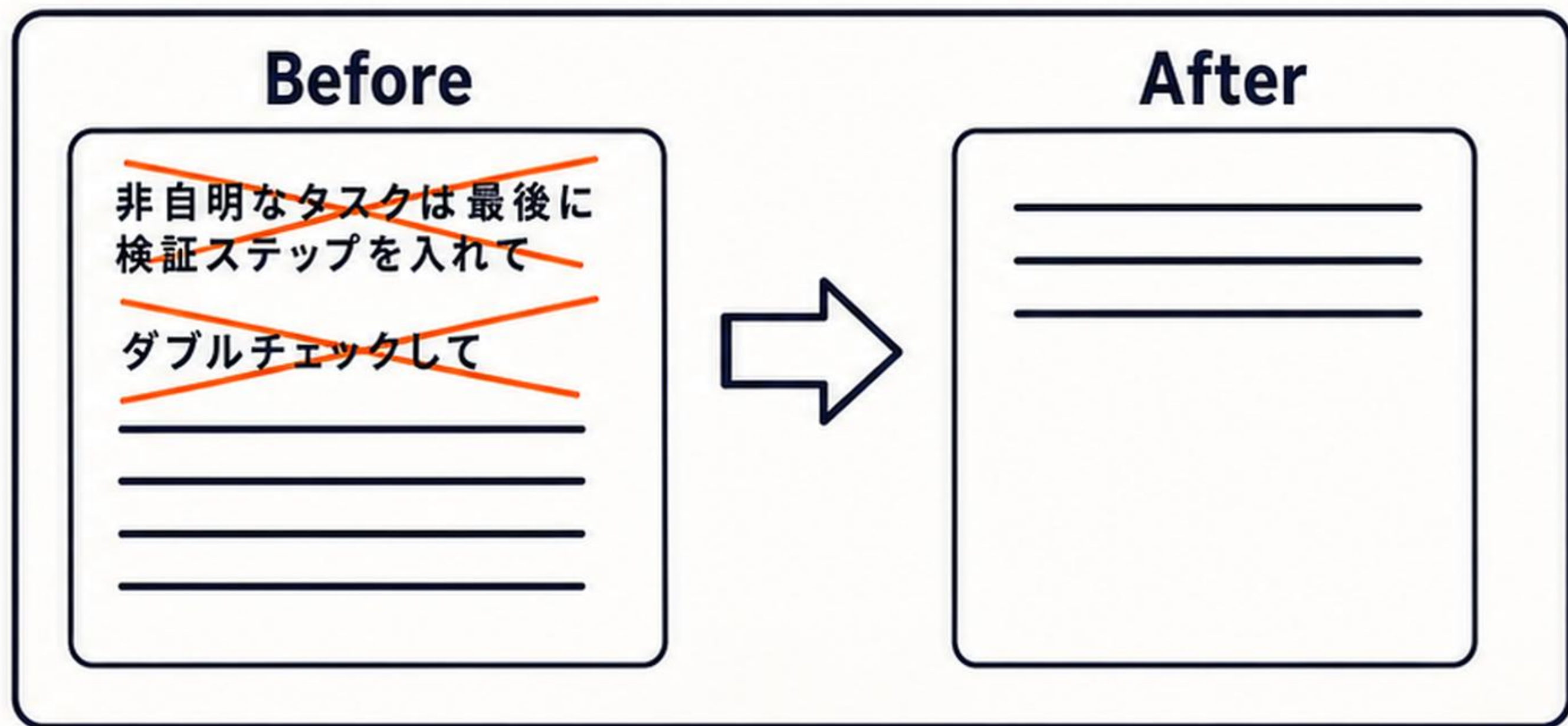
コードで動画を作る方式は確認のたびにレンダリングが要るが、AKARI はビューワーがセーブデータを直読みするのでプレビューが即応し…

〈知見〉Opus 5 では指示を「足す」のではなく「引く」

Anthropic 公式ガイドの要点

Anthropic が公開した Opus 5 向けプロンプトガイドを踏まえ、従来の「最後に検証して」「ダブルチェックして」という念押しが Opus 5 では重複作業とトークン増加につながる、と整理したスレッド。

- ① 「非自明なタスクは最後に検証ステップを入れて」の一文を消す。Opus 5 は言われなくても自分で検証する
- ② 「ダブルチェックして」も同様に消す
- ③ コードレビューで「重大な問題だけ報告して」「保守的に判断して」と書くと、Opus 5...



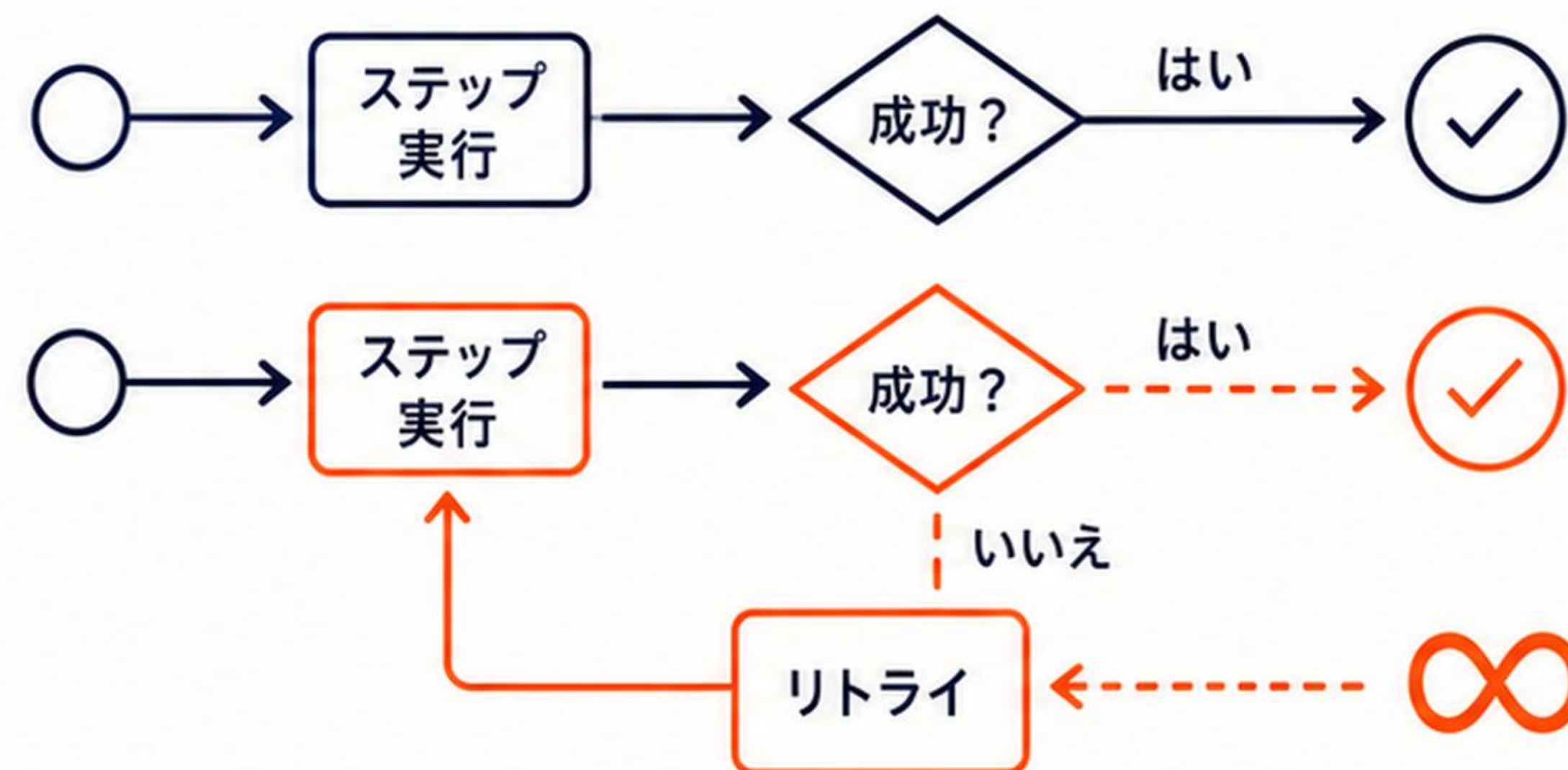
CLAUDE.md や運用プロンプトに「必ず検証して」系を書いている人は、消すだけで速くなる。

〈知見〉 エージェントループの多くに「リトライ上限」が無い

70システム調査と厳格エスカレーション

実運用されているエージェントループ70システムを調べたところ、失敗したステップを何回までリトライするかについて形式的な上限を持たない実装がかなりの割合を占めていたという指摘。

- 調査対象は実世界のエージェントループ実装 70システム
- リトライ回数に形式的な上限が無い実装が大きな割合を占めた
- 上限が無いと、障害時の挙動＝モデルがその場で決めた判断になり、人間の事前承認を通っていない

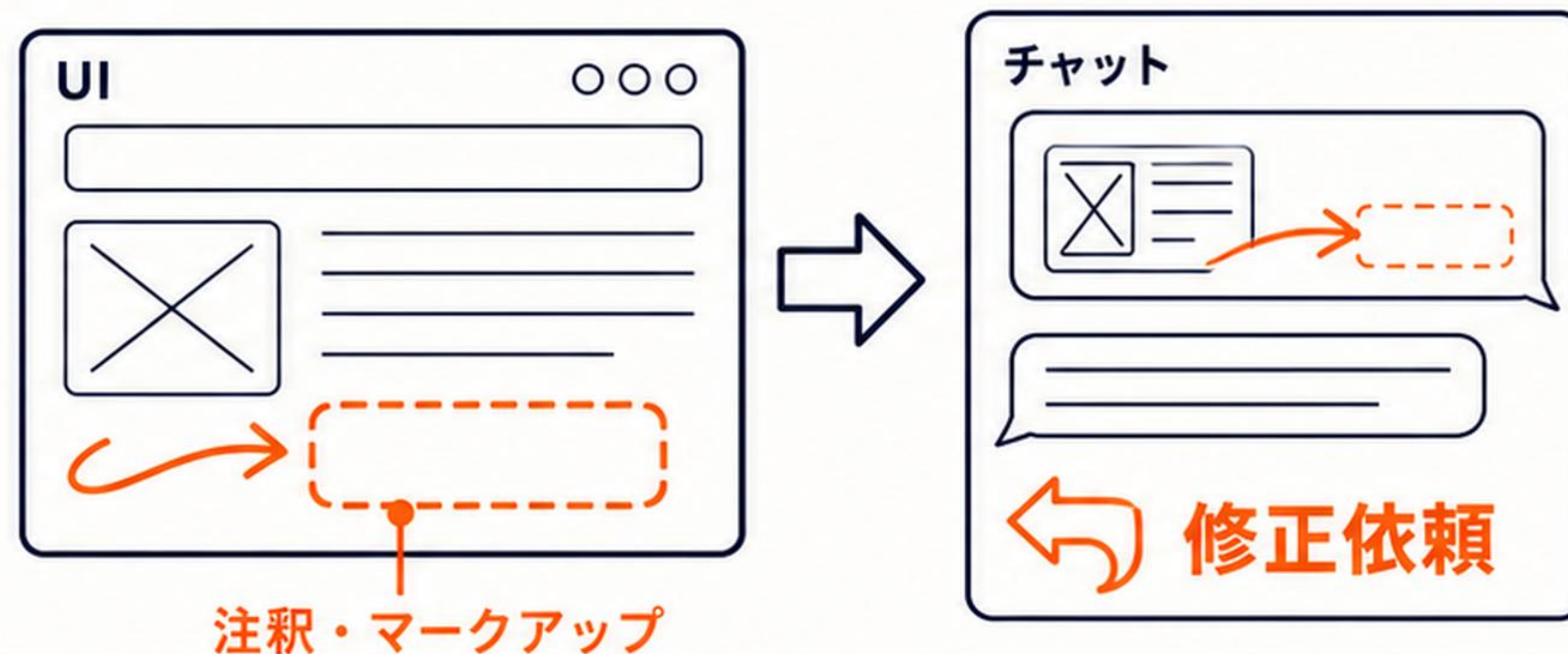


ループを組むときに設計されるのは大抵「うまくいく経路」で、壊れたときの経路は暗黙のまま残る。

〈知見〉 Claude Code デスクトップは UI に直接マークアップして修正指示を出せる

Claude Code のデスクトップアプリでは、開発中のアプリ UI に直接注釈やマークアップを入れて、そのままチャットへ送って修正させられる、という実践報告。

- 対象は Claude Code の **デスクトップアプリ** (CLI ではない)
- アプリの UI に注釈・マークアップを入れ、それを直接チャットへ送って修正依頼にできる
- 位置や見た目に関する指示を言語化せずに済むため、視覚的なフィードバックの往復コストが下がる



UI の手直しは「どこの何を」を言葉にするコストが高く、AI との往復が伸びやすい工程。

本日のトピック一覧

- 01 Kimi K3 が公開翌日に Cursor と Ollama クラウドへ
- 02 Anthropic が open-weights の立場を公式表明
- 03 OpenAI 調査で AI 依頼の 43.5% が「他職種の仕事」
- 04 GPT-Live が ChatGPT Voice の組織プランへ全世界展開
- 05 Akari Video で Claude Code / Codex がそのまま編集チームになる
- 06 〈知見〉 Opus 5 では指示を「足す」のではなく「引く」
- 07 〈知見〉 エージェントループの多くに「リトライ上限」が無い
- 08 〈知見〉 Claude Code デスクトップは UI に直接マークアップして修正指示を出せる