

2026-08-07

1. Agent Plugins 1.0.0
2. Cursor Router の中身が公開
3. Qwen3.8-Max が Artificial Analysis 総合56点で世界5位

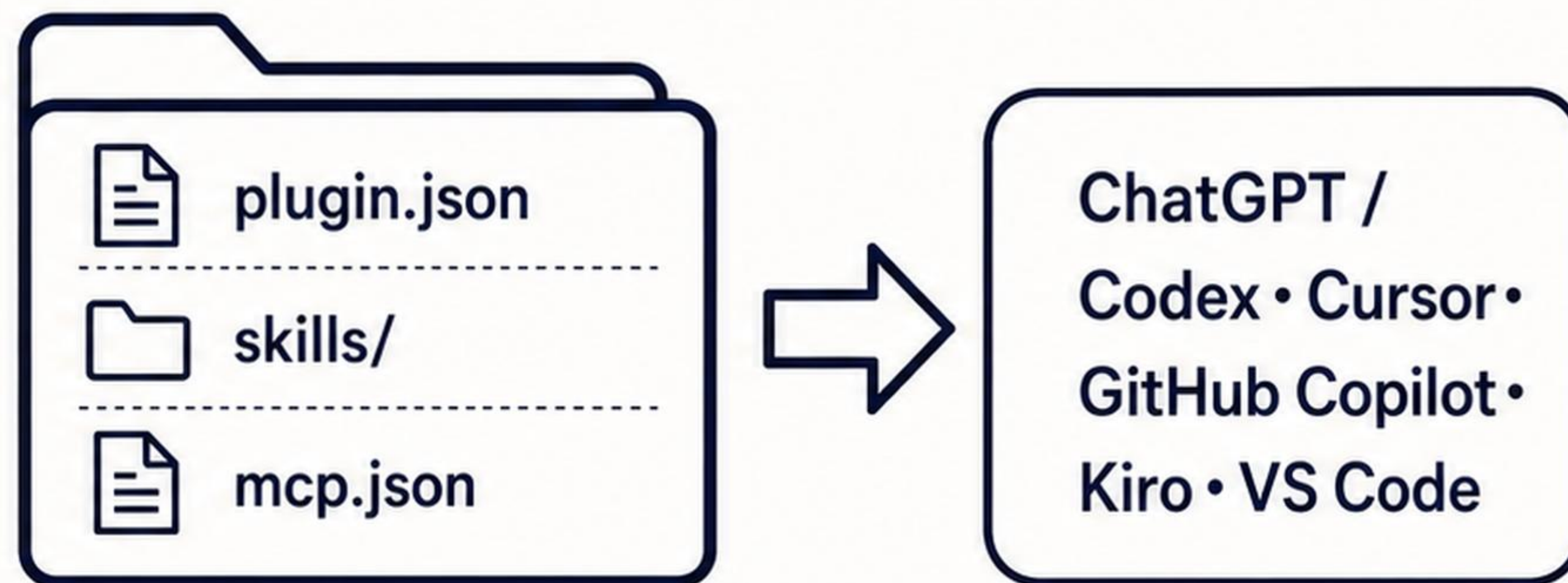
8トピック

1. Agent Plugins 1.0.0

OpenAI・AWS・Microsoft・Cursor・Vercel が共同でベンダー中立の標準を策定

Agent Skills と MCP サーバーを1つのフォルダにまとめて持ち運べるようにする、オープンかつベンダー中立の仕様「Agent Plugins 1.0.0」が 8月6日に公開された。

- ルートに `plugin.json` マニフェストを置き、`skills/` (Agent Skills) と `mcp.json` (MCP サーバー定義…
- クライアント固有の拡張は `com.example.client/` のような逆ドメイン名前空間に隔離し…
- 対応クライアント (ローンチ時) : ChatGPT / Codex・Cursor・GitHub Copilot・Kiro・VS Code



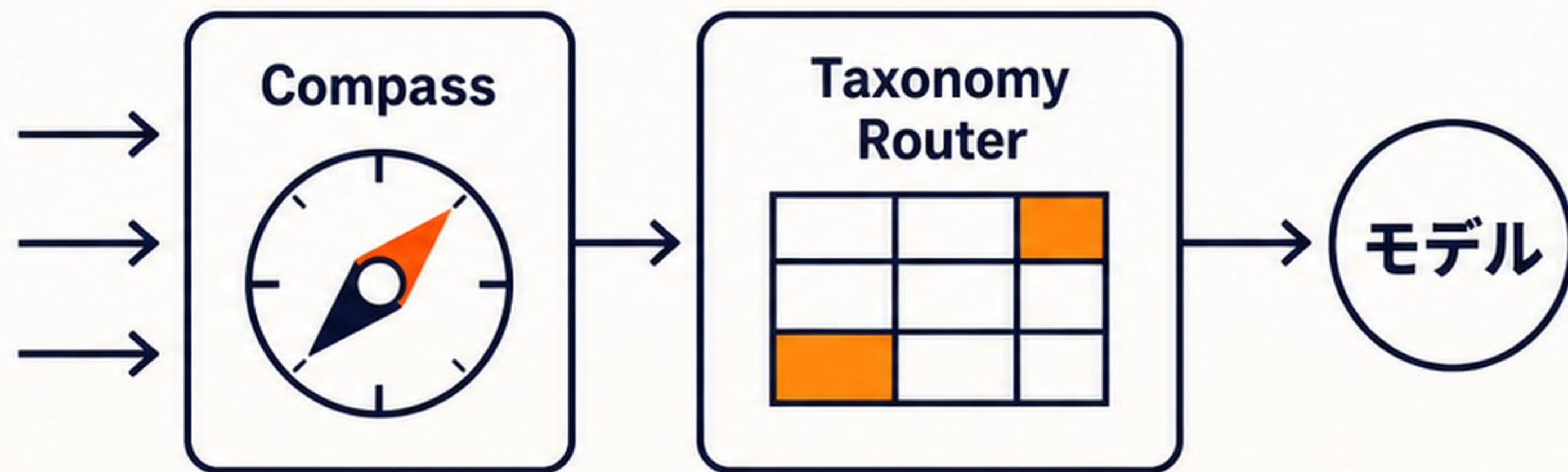
これまで Skill も MCP もクライアントごとに配置場所と書式が違い、同じ中身を何度も詰め直す必要があった。

2. Cursor Router の中身が公開

複雑度スコアとタクソノミの2段構えで、Fable 超えを68%安く

Cursor が自動モデル選択機構「Cursor Router」の内部構造を8月6日に解説記事として公開した。

- 第1段 Compass: 0~1 の複雑度スコア。
高信頼の予測では96%、低信頼では71%の的中率
- 第2段 Taxonomy Router: ドメイン
(バックエンド / フロント / DB) ×
タスク (バグ修正 / テスト / コマンド...)
- モデル別の得意分野を明示 —
Grok 4.5=定型作業を低コストで /
GPT-5.6 Sol=計画とコードベース分析 /...



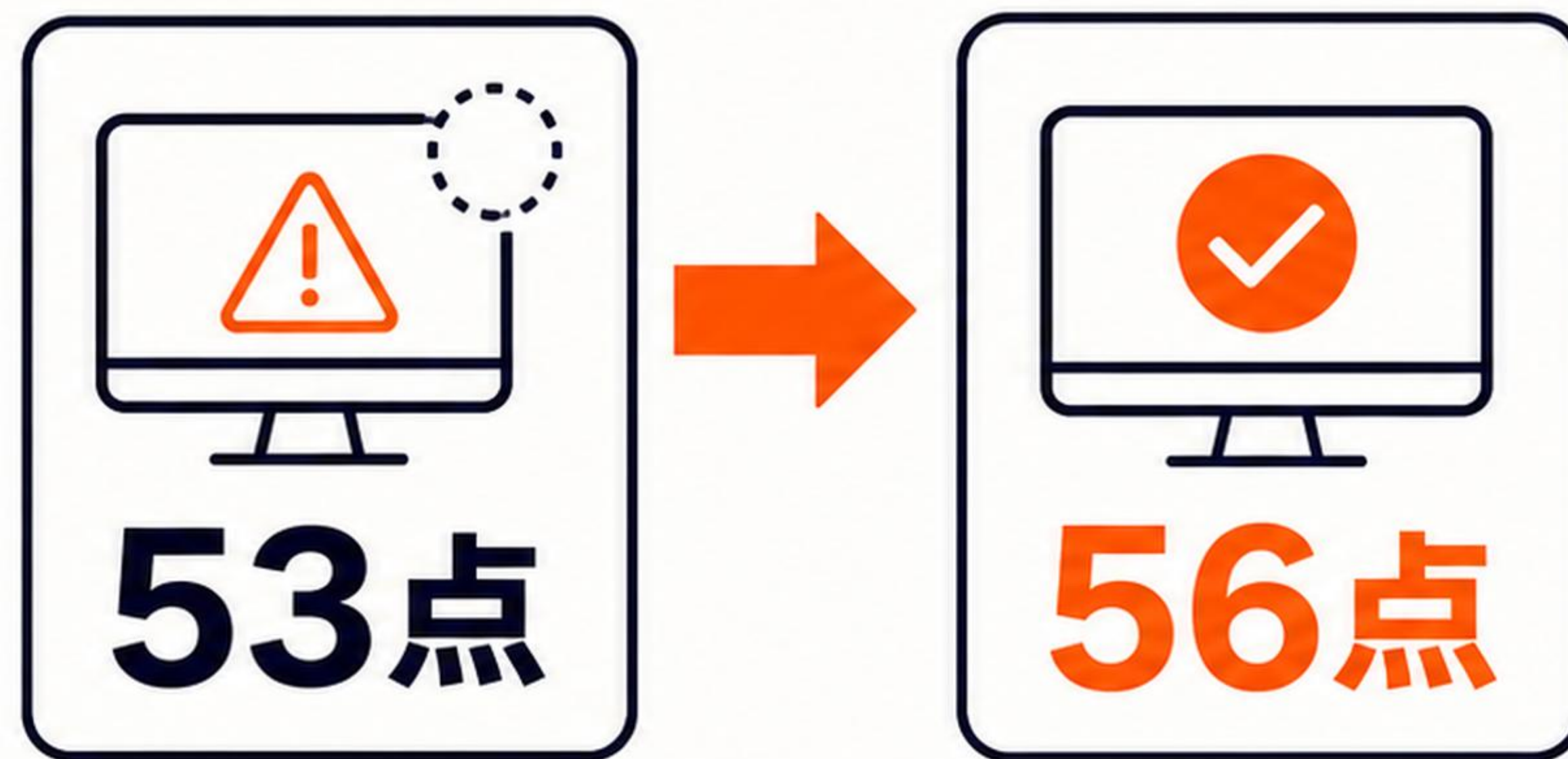
「どのモデルを使うか」を人が固定する前提が崩れつつある。

3. Qwen3.8-Max が Artificial Analysis 総合56点で世界5位

再計測で53から上振れ

Alibaba の Qwen3.8-Max が Artificial Analysis Intelligence Index で56点を記録し、総合5位に入った。

- Artificial Analysis Intelligence Index で **56点・総合5位**
- Claude Opus 4.8 (max) と同水準。
Google・Meta・xAI の全モデルより上
- 先行して公表されていた53点は、
測定対象エンドポイントの断続的な
不具合の影響だった。Alibaba の公開 API...

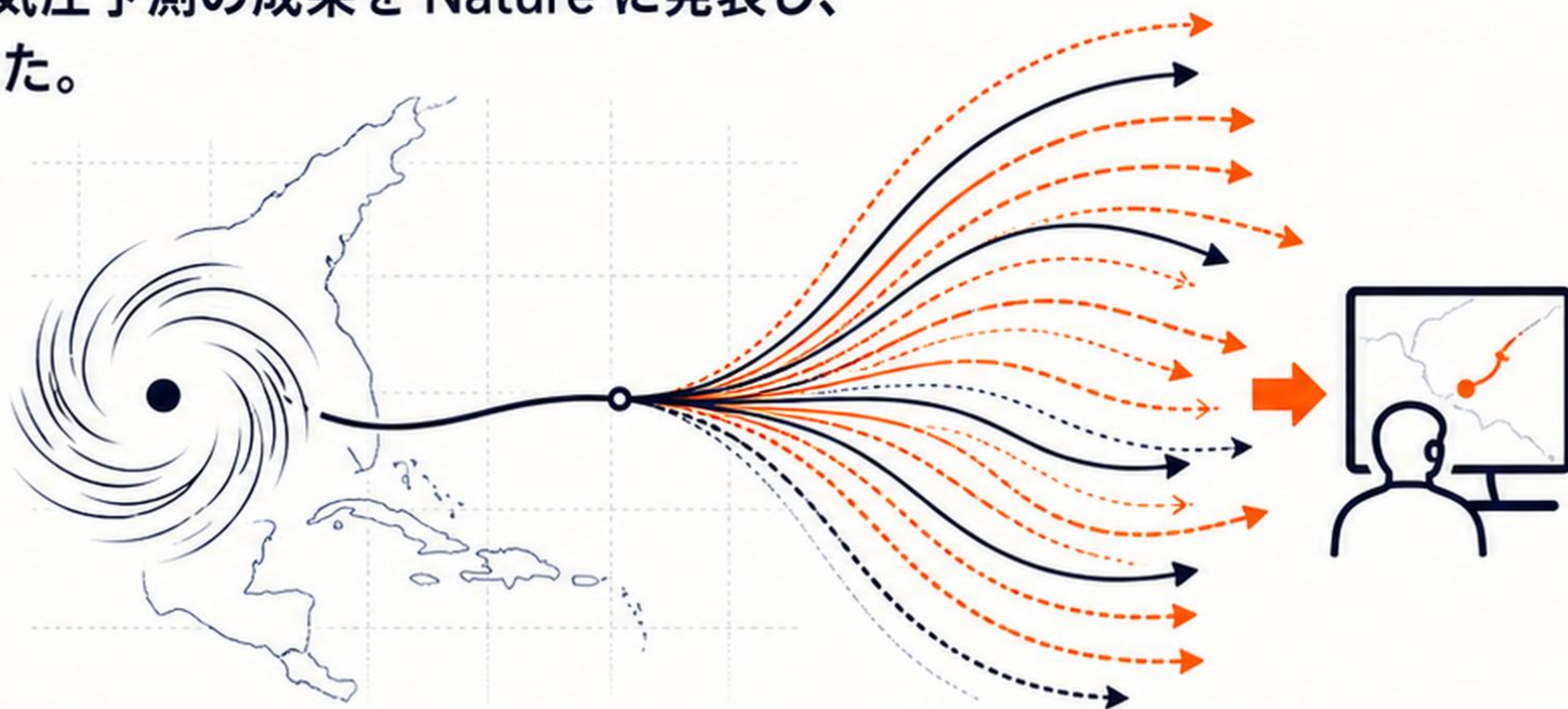


スコアそのものより、「53点」が測定環境の不具合由来で、再実行したら56点になったという経緯のほうが実務的に重い。

4. Google の天気モデル WeatherNext 2 が Nature 掲載、コードと重みを OSS 公開

Google DeepMind が WeatherNext 2 の熱帯低気圧予測の成果を Nature に発表し、同時にコードとモデル重みをオープンソース公開した。

- 2023～2025年のサイクロンで、進路・強度・風構造の予測が主要現業モデルより平均1日以上先行（3日先の予測が従…
- 不確実性に対応するため、1つのサイクロンにつき **1,000通りのシナリオ** を生成して予報官の判断を支える
- 2025年の大西洋ハリケーンシーズンで米国立ハリケーンセンターの現業ワークフローと並走。ハリケーン Melissa…



フロンティアモデル競争とは別の軸で、AI が実際に人命に効いていることが査読を通った形で示された数少ない例。

5. Google Maps の Ask Maps がエージェント化

注文・予約・チケット購入まで会話で完結

3月に投入された Gemini ベースの会話機能 Ask Maps に、実際に行動を起こすエージェント機能が 8月6日に追加された。

- フードデリバリー: 自然言語で希望を伝えると候補を出し、カートまで作って Uber Eats・Toast・Square...
- ホテル: 価格・立地・設備、さらに「雰囲気」のような曖昧な条件でも空室確認まで実行する
- チケット: 「今晚、職場の近くでやってるお笑いライブは？」のような聞き方で候補と購入リンクを返す



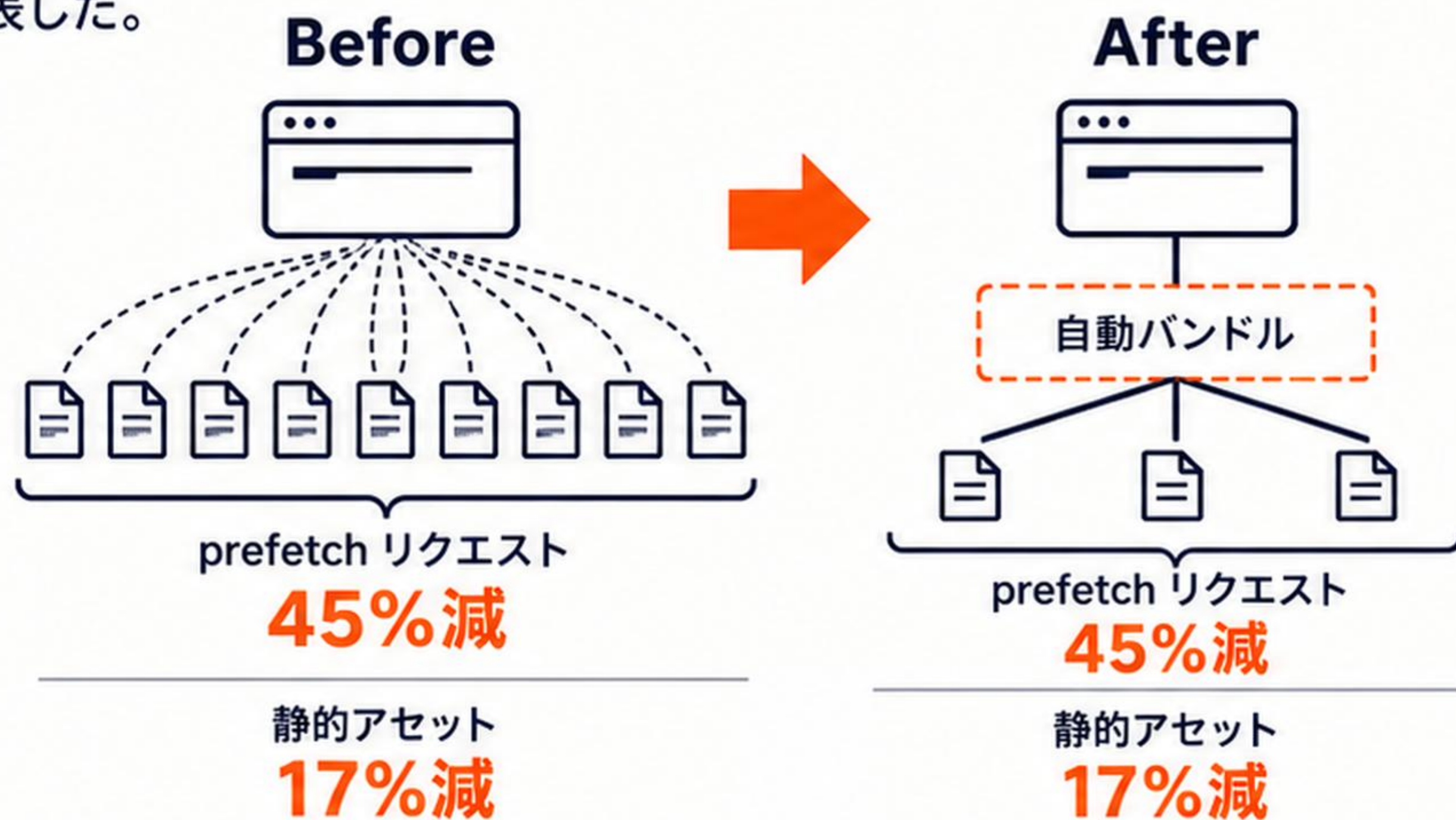
「検索して見せる」から「代わりに手続きまで済ませる」への線を、10億人規模の既存プロダクトが跨いだ。

6. Next.js 16.3 の実アプリ実測

Vercel 上で prefetch リクエストが45%減

Vercel が 8月5日、16.3 へ上げた実アプリでの計測結果を公表した。

- Vercel 上の Next.js 16.3 アプリで prefetch リクエスト 45%減、静的アセット 17%減
- しきい値以下の prefetch を自動でバンドル。大きい共有セグメントは分けたまま残し、複数ルートで再利用できるようにしている
- 同リリースの他の実測値 — next dev のメモリ最大90%減 (vercel.com のダッシュボードで 21.5GB → 2GB)



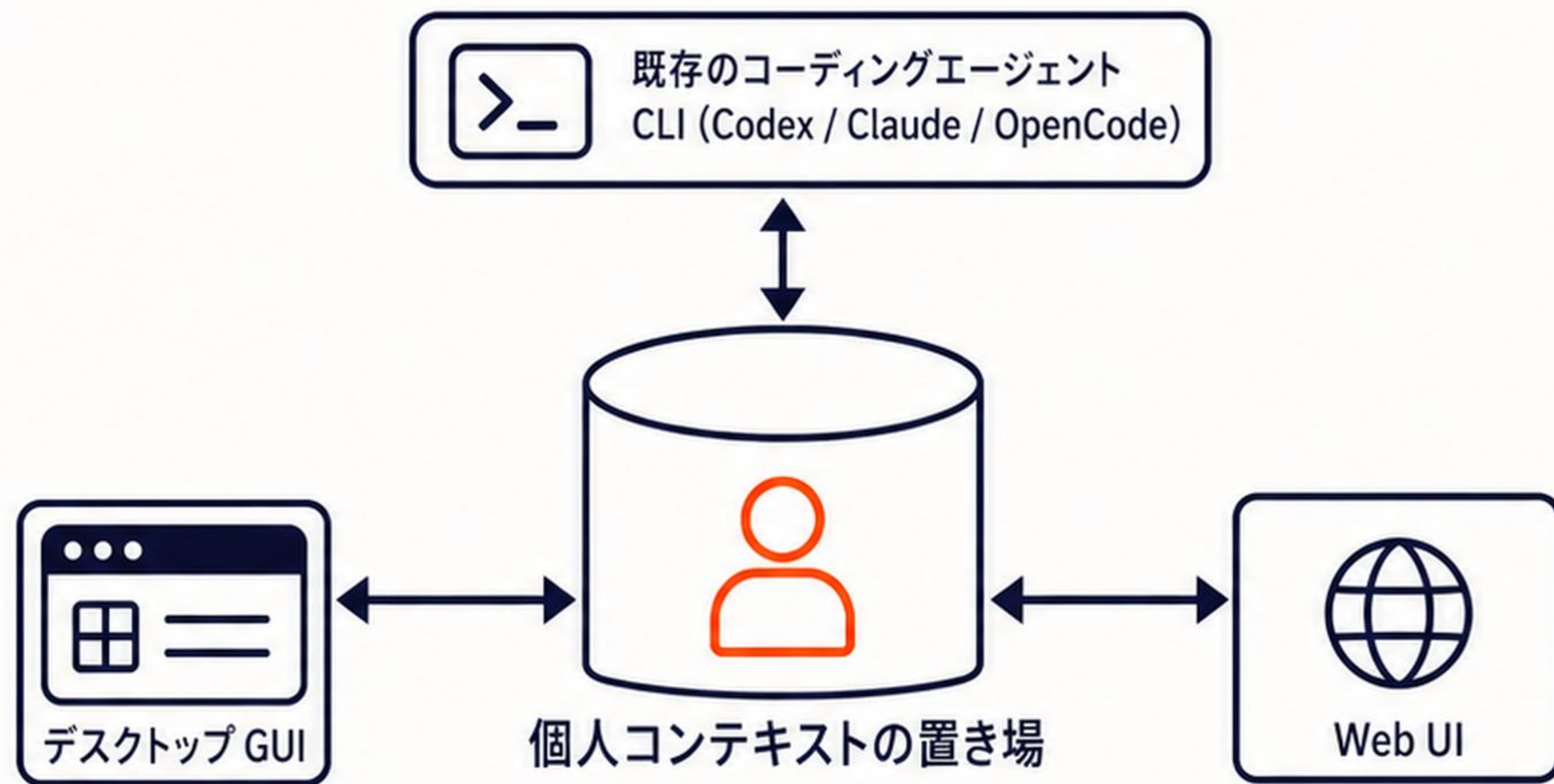
フレームワークのバージョンを上げるだけで効く数字が並んでいる。

7. 〈知見〉 OpenContext

既存の Codex / Claude Code をそのまま使い回す「個人コンテキストの置き場」

エージェント用の個人コンテキストストア。

- 既存のコーディングエージェント CLI (Codex / Claude / OpenCode) を再利用する。追加のエージェント課金がない
- 内蔵の Skills / ツール群と、コンテキストを捕まえて検索・編集するデスクトップ GUI を持つ
- Web UI もあり、インストールなしでローカルのコンテキストを閲覧・編集できる



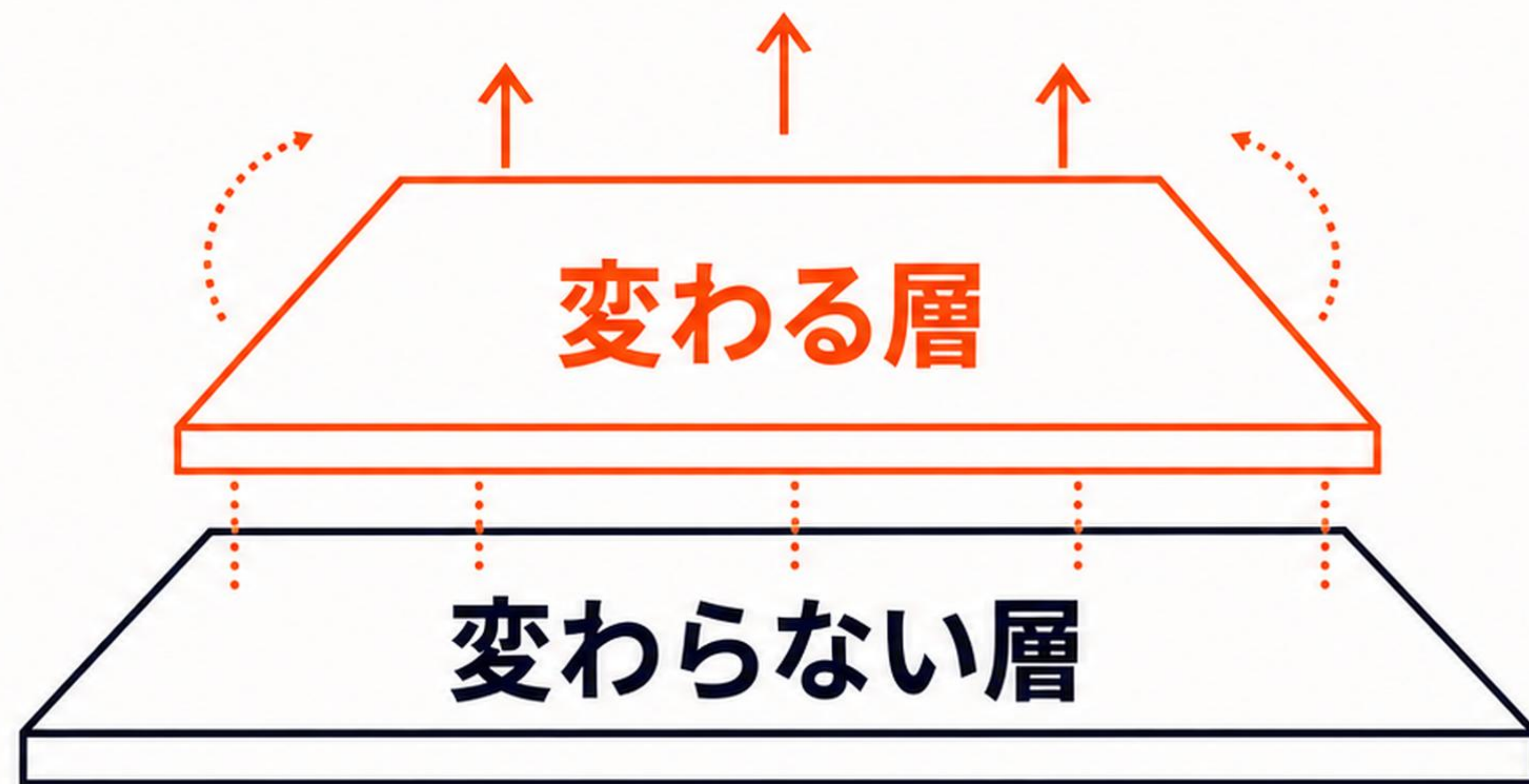
エージェントに記憶を持たせる話は増えているが、多くは「新しいエージェントに乗り換えろ」という形をとる。

8. 〈知見〉変化の早い Claude Code を書籍に落とし込む

資料を「変わらない層」と「変わる層」に割る

Findy 主催の Claude Code イベントでの登壇資料。

- 扱っているのは「変化の早さ」そのもの。執筆中に機能が追加・改名・廃止されるツールをどう書くかの方針論
- 書籍側は寿命の長い層（原理・型・判断基準）に厚みを置き、短い層はサポートサイトや正誤表で追随させる構成になりやすい
- この考え方は書籍に限らず、社内マニュアルや講座カリキュラムを作る側にそのまま当てはまる



AI ツールの手順書を書いたことがある人なら全員が踏んでいる問題に、はっきりした処方箋が出ている。

本日のトピック一覧

1. Agent Plugins 1.0.0
2. Cursor Router の中身が公開
3. Qwen3.8-Max が Artificial Analysis 総合56点で世界5位
4. Google の天気モデル WeatherNext 2 が Nature 掲載、コードと重みを OSS 公開
5. Google Maps の Ask Maps がエージェント化
6. Next.js 16.3 の実アプリ実測
7. 〈知見〉 OpenContext
8. 〈知見〉 変化の早い Claude Code を書籍に落とし込む