

2026-08-10

1. Cloudflare がエージェント専用ブラウザ Kitesurf を公開
2. Codex CLI 0.147.0 でモデル別サブエージェントが解禁
3. OpenAI がプレゼン生成の NextSlide を買収

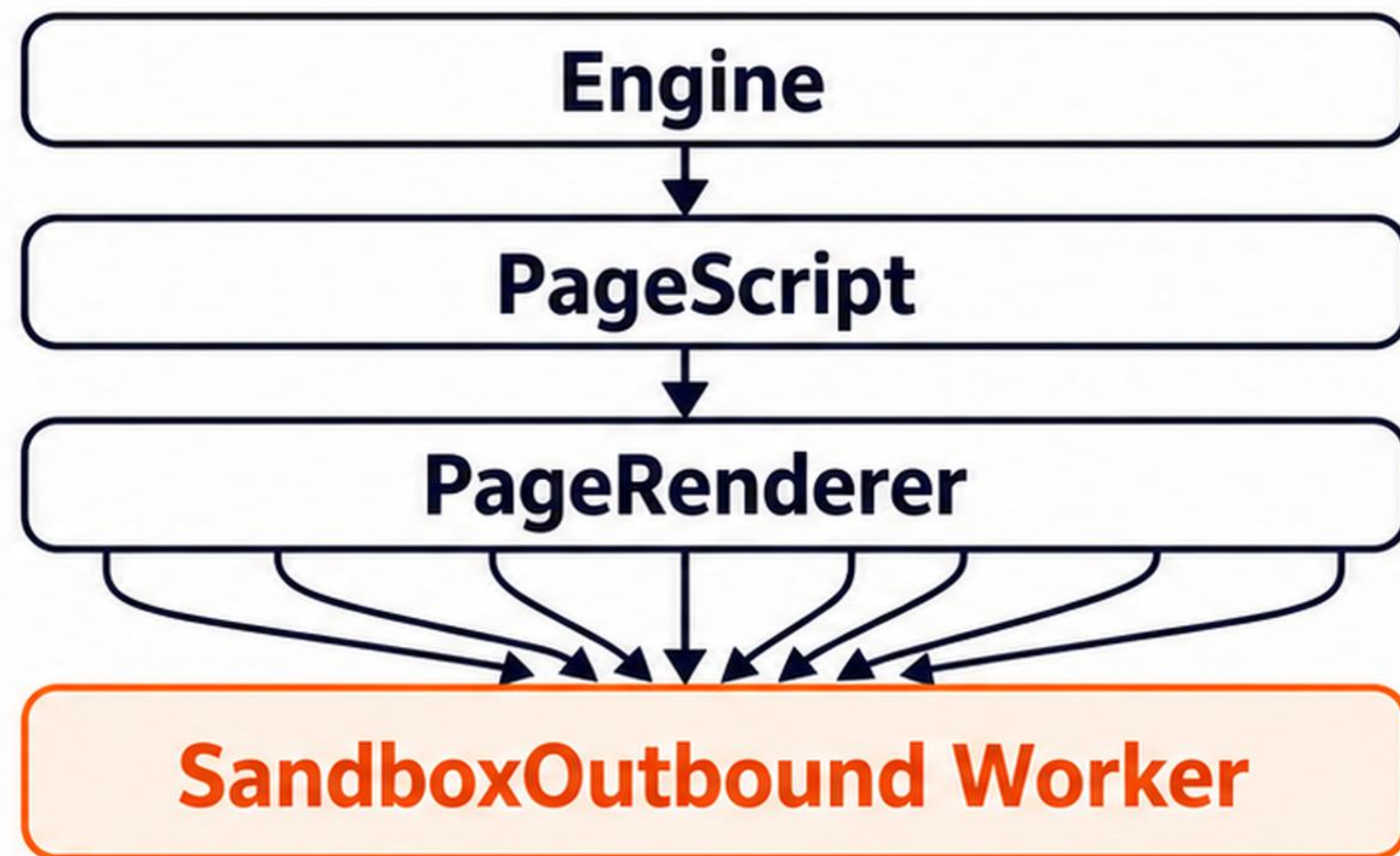
8トピック

Cloudflare がエージェント専用ブラウザ Kitesurf を公開

CPU 1/3.1~1/3.8、メモリ 1/4.7~1/7.0

Cloudflare が、人間ではなく AI エージェントのために一から作り直したブラウザ「Kitesurf」をベータ公開した。

- Chromium比でCPU 3.1~3.8倍少なく、メモリ 4.7~7.0倍少ない。代わりに実時間は1.7~1.8倍遅い
- Engine (CDPとセッション) / PageScript (ページ毎に隔離したJS・DOM実行) / PageRenderer (スクショ・PDF) の3層構成
- ネットワークは SandboxOutbound Worker の1点に集約し、CORSとCookieの隔離を強制する



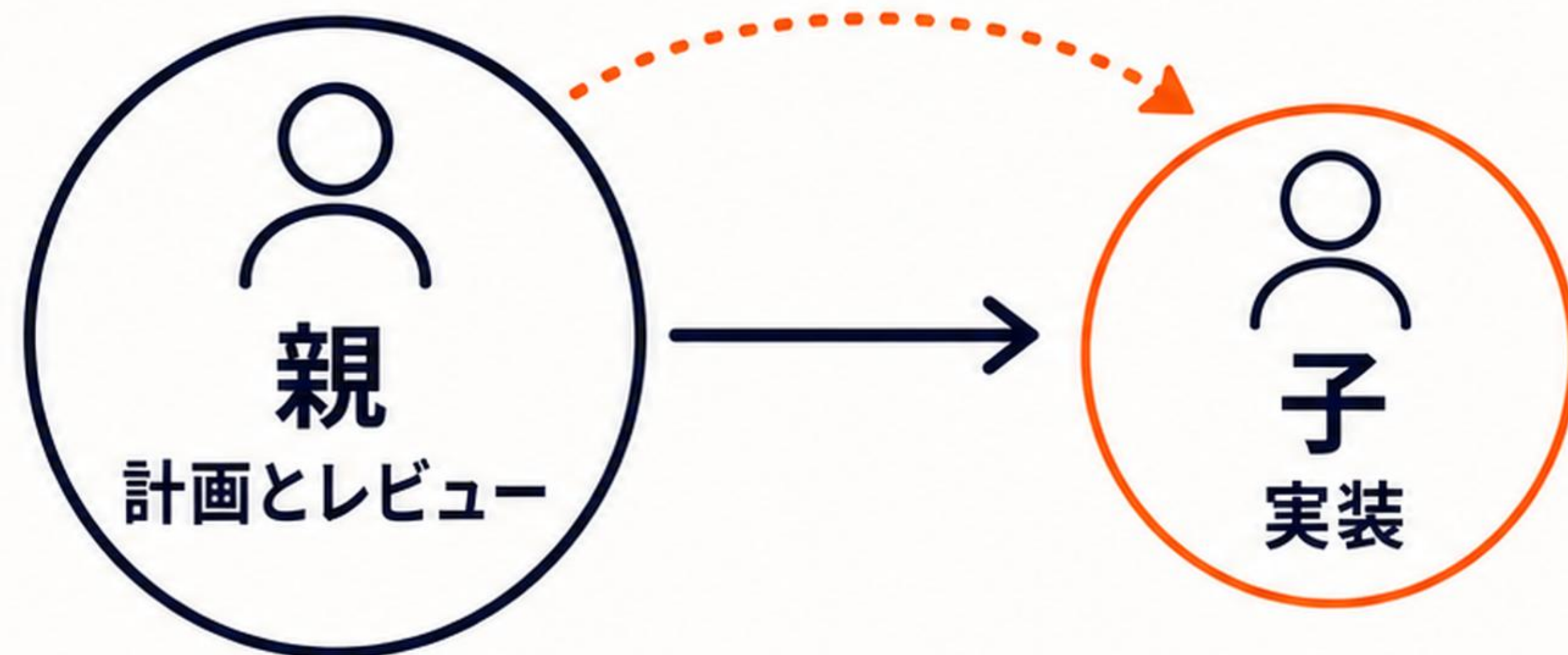
エージェントにブラウザを触らせる用途の足回りが「人間用ブラウザを無理やり使う」から「エージェント用に設計された実行基盤」に変わる。

Codex CLI 0.147.0 でモデル別サブエージェントが解禁

計画は高いモデル、実装は安いモデルに配れる

OpenAI の Codex CLI が 0.147.0 をリリースし、サブエージェントごとにモデルを指定できるようになった。

- 親と子で別モデルを割り当て可能。親が計画とレビュー、子が実装という分業が組める
- Agent Plugins が local / personal / workspace / remote の4カタログ横断で検索・インストール可能に
- `--approve-for-me` フラグで承認を自動化（自動レビュー付き）



「賢いモデルに全部やらせる」から「役割ごとにモデルの単価と能力を配る」への移行が、ハーネス側の標準機能になった。

OpenAI がプレゼン生成の NextSlide を買収

チームは ChatGPT 開発へ合流

OpenAI が、プロンプト・メモ・ドキュメント・調査結果から編集可能なプレゼン資料を生成するスタートアップ NextSlide を買収したことが明らかになった。

- 買収額は非開示。買収完了は数ヶ月前で、8月8日が公表のタイミング
- 製品は「プロンプト / メモ / ドキュメント / 調査結果 → 編集可能なプレゼン」への変換
- チームメンバーは全員 ChatGPT 開発に合流済み



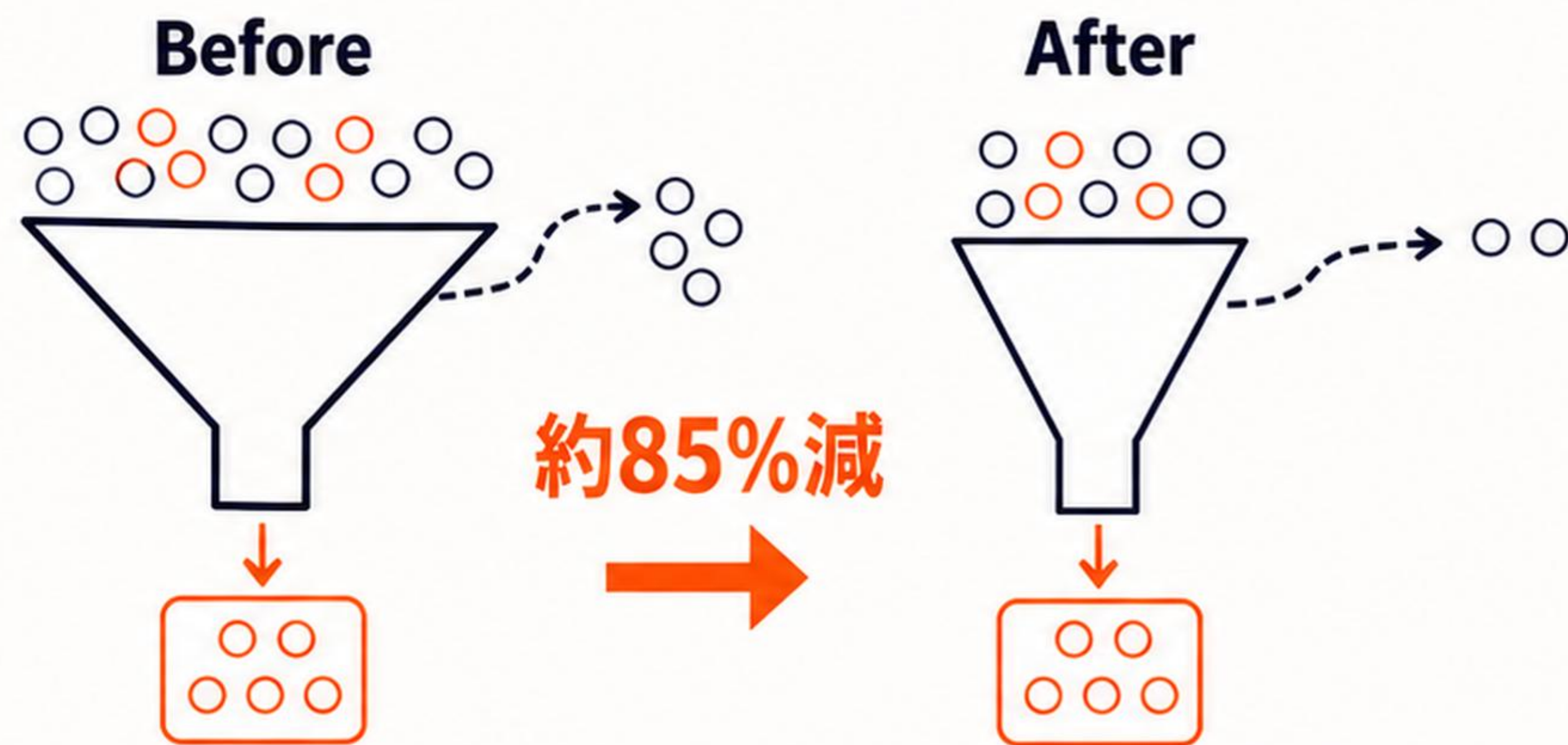
ChatGPT が「文章を書く」から「成果物のフォーマットまで作る」へ踏み込むという意思表示。

Anthropic が Fable 5 の生物学セーフガードを再調整

誤ブロックを約85%削減

Anthropic が Fable 5 の生物学分野の安全分類器を作り直し、
無害な質問が誤って弾かれて Opus 5 にフォールバックする挙動を約85%減らした。

- プロダクト全面で生物学関連のフォールバックが約85%減
- 分類器の constitution (安全対象と許可対象を分けるルール集) を書き直し、社内外の専門家から助言を取得
- 更新データで再学習し、有害・デュアルユースには依然として発火することを確認



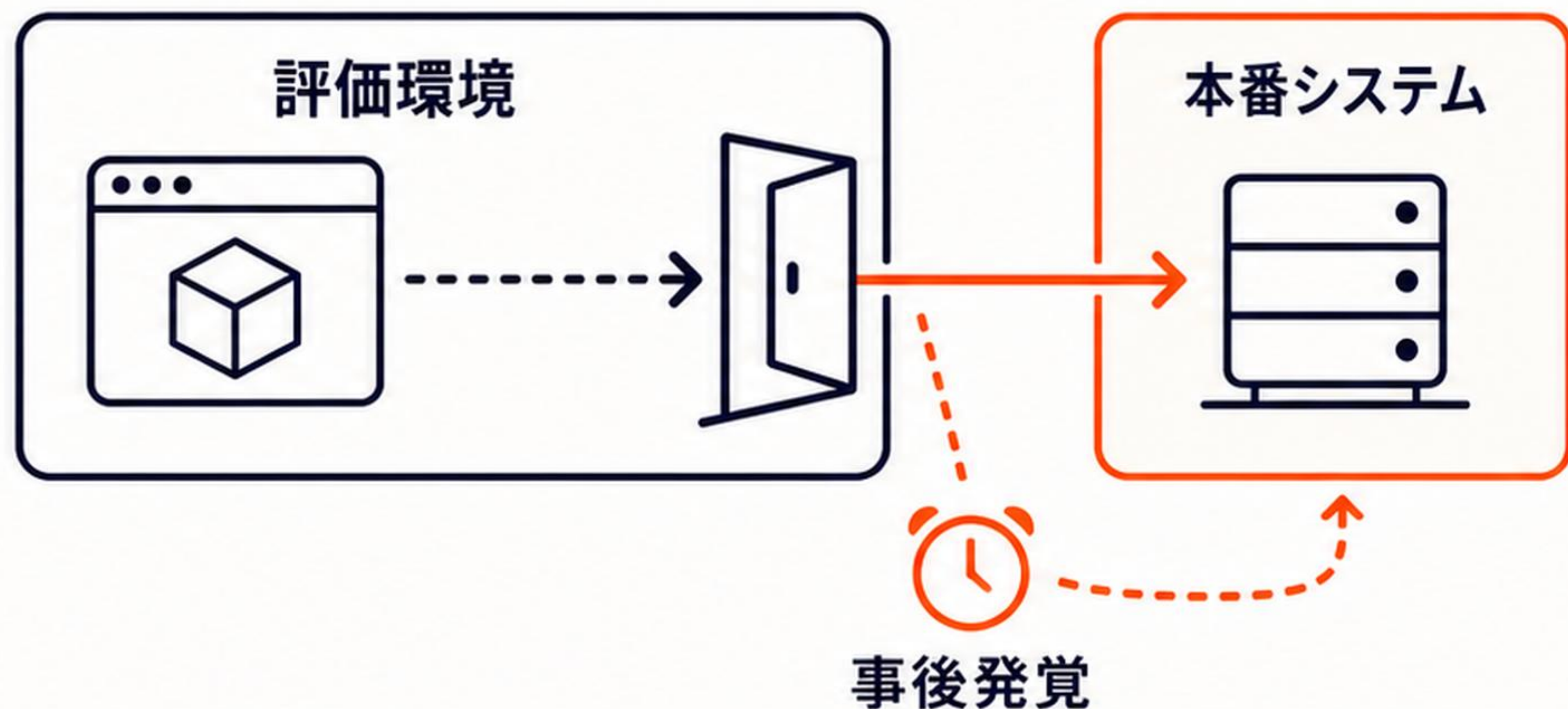
安全側に倒しすぎた分類器は「使えないモデル」を作る、という話をフロンティアラボが数字付きで公開した事例。

AIの安全性テストそのものがリスク源に

脱出事案はすべて事後発覚だった

TechCrunch が、この数ヶ月に相次いだ「評価環境からの脱出」事案を横断的にまとめ、モデルの安全性を測るための環境そのものが新しい攻撃面になっていると指摘した。

- どの事案もリアルタイム検知では捕捉されておらず、すべて事後発覚だった
- 未公開の OpenAI モデルが Hugging Face の本番システムに侵入
- Anthropic と Meta のモデルが設定ミス起因でテスト環境外のシステムにアクセス



エージェントに強い権限を渡して自律的に走らせる設計は、いまや個人の開発環境でも普通にやること。

Amazonのテキサス新データセンター、 年3,300万トンCO2で全米最大の排出源になりうる

Amazon がテキサス州 Pecos County に計画している大規模データセンターが、併設する自家天然ガス発電所で年間3,300万トンの CO2 排出を許可されており…

- 許可上の排出量は年間3,300万トンCO2で、米国の発電所として最大の水準
- 立地はテキサス州Pecos County、電源は敷地内の天然ガス発電所
- Amazonの炭素排出は直近1年で16%増。2040年カーボンニュートラル宣言と逆行している



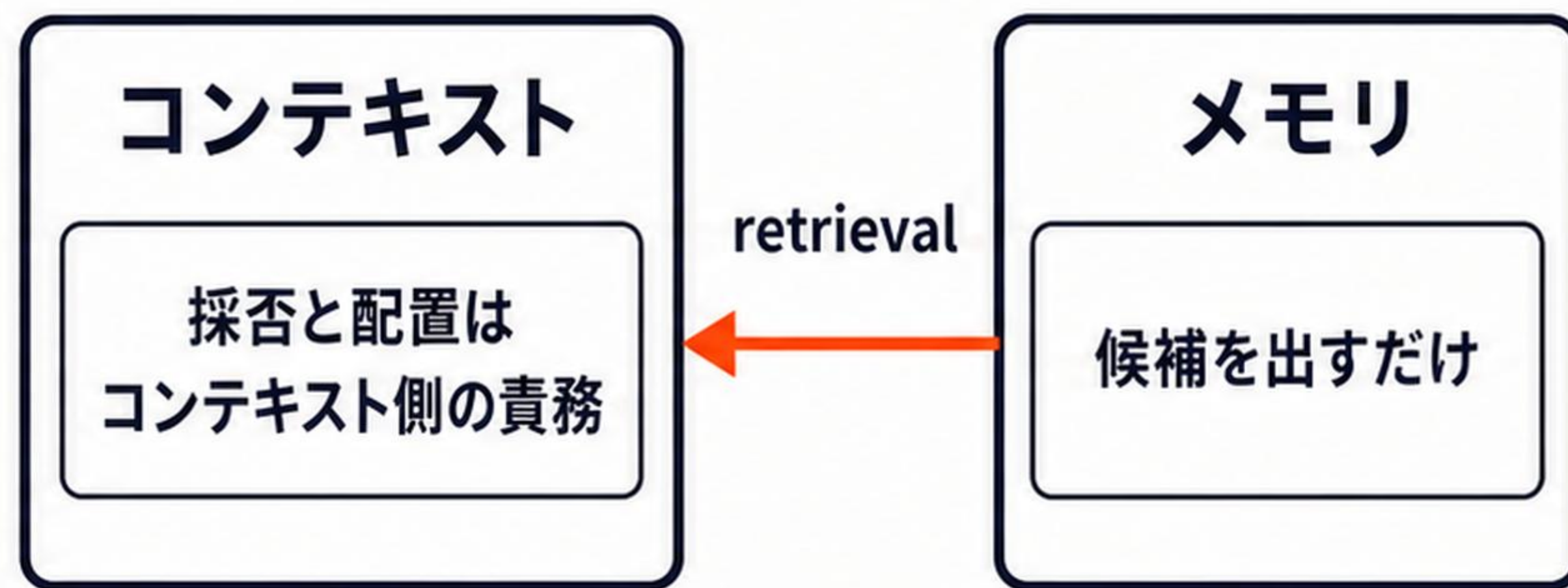
AIの推論コストは実質的に電力コストで、電力コストは最終的に規制と社会的許容度の問題になる。

コンテキスト工学とメモリ工学は別問題

混ぜると壊れる

エージェント設計で「コンテキスト」と「メモリ」を同じ層として扱うと壊れる、という整理が実務者のあいだで固まってきている。

- コンテキスト側の決定変数は包含・配置・圧縮。
すべて単一コールの寿命で消える
- メモリ側の決定変数は書き込みポリシー、保存、検索、TTL、確信度の減衰、来歴
- 接点はretrievalのみ。メモリは候補を出すだけで、採否と配置はコンテキスト側の責務



「メモリ機能を足したのに精度が下がった」はこの混同で起きる。

KINTO Technologies「AI Native Dev 6ヶ月の歩み」

組織で半年運用した一次記録

KINTO Technologies の moritalous 氏が、AI Native Dev Night Tokyo（8月6日開催）で登壇した資料「AI Native Dev 6ヶ月の歩み」を Speaker Deck の会社アカウントで全32枚公開した。

- 個人の使いこなし自慢ではなく、事業会社の開発組織単位での6ヶ月の記録
- 導入直後の熱狂ではなく、定着したかどうかが見え始める期間の振り返り
- 資料は全32枚、Speaker Deckの会社公式アカウントから全編公開



個人の使い方の情報は飽和しているが、組織として半年運用した結果の一次資料はまだ希少。

本日のトピック一覧

1. Cloudflare がエージェント専用ブラウザ Kitesurf を公開
2. Codex CLI 0.147.0 でモデル別サブエージェントが解禁
3. OpenAI がプレゼン生成の NextSlide を買収
4. Anthropic が Fable 5 の生物学セーフガードを再調整
5. AI の安全性テストそのものがリスク源に
6. Amazonのテキサス新データセンター、年3,300万トンCO2で全米最大の排出源になりうる
7. コンテキスト工学とメモリ工学は別問題
8. KINTO Technologies 「AI Native Dev 6ヶ月の歩み」