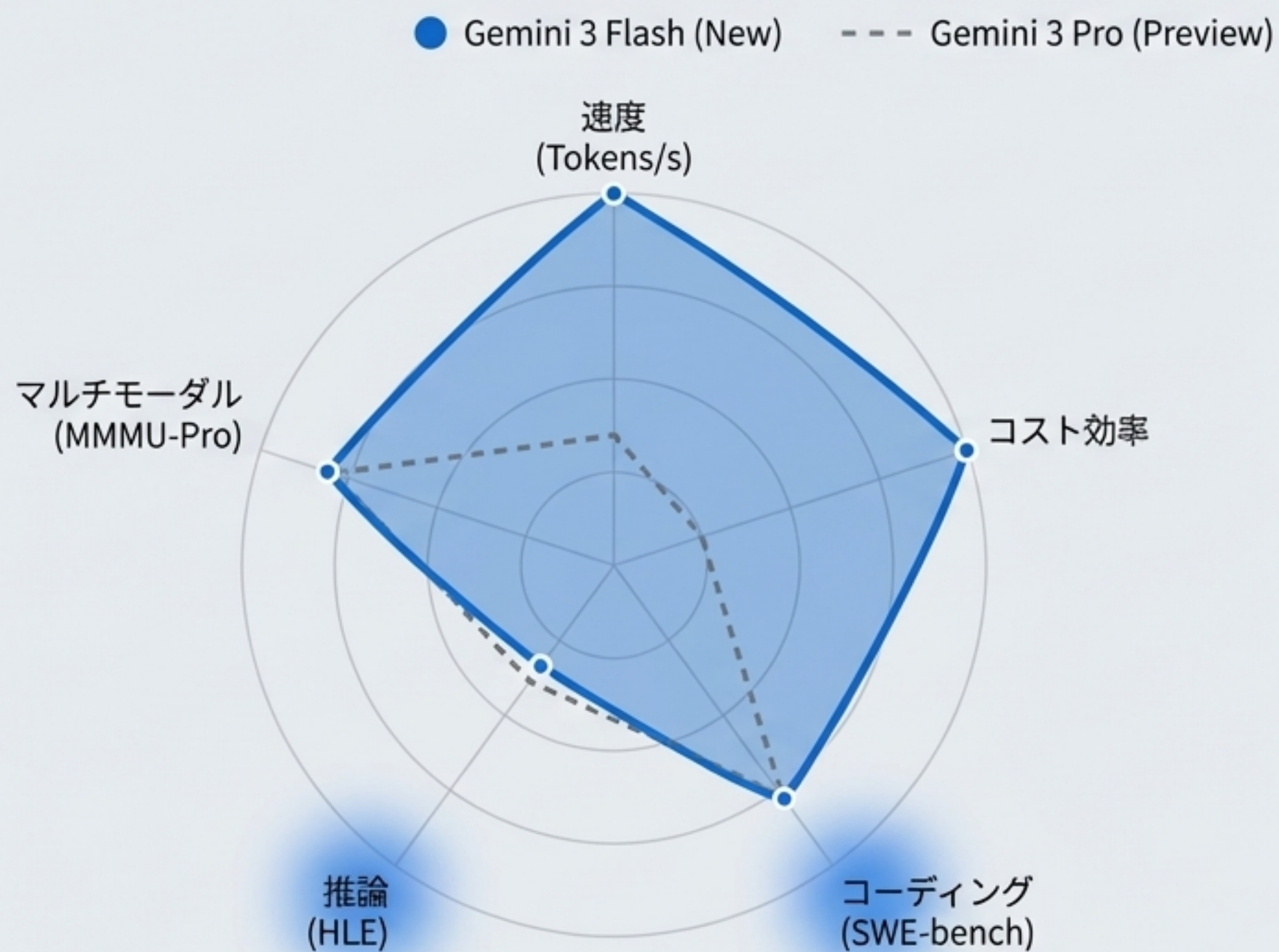


Gemini 3 Flash & Agentic RL

AI性能のパラダイムシフト：軽量モデルはいかにしてフラッグシップを凌駕したのか

逆転現象：軽量モデル「Flash」が、AIの核心能力で「Pro」を上回る



2025年12月、業界の常識が覆された。

速度とコストで優れるGemini 3 Flashが、AIの核心能力であるコーディングと推論において、上位モデルであるPro (Preview) 版を凌駕。

これは単なる性能向上ではなく、地殻変動の予兆である。

本レポートが解き明かす3つの問い



The Mystery: なぜ「逆転現象」は起きたのか？
(The Performance)

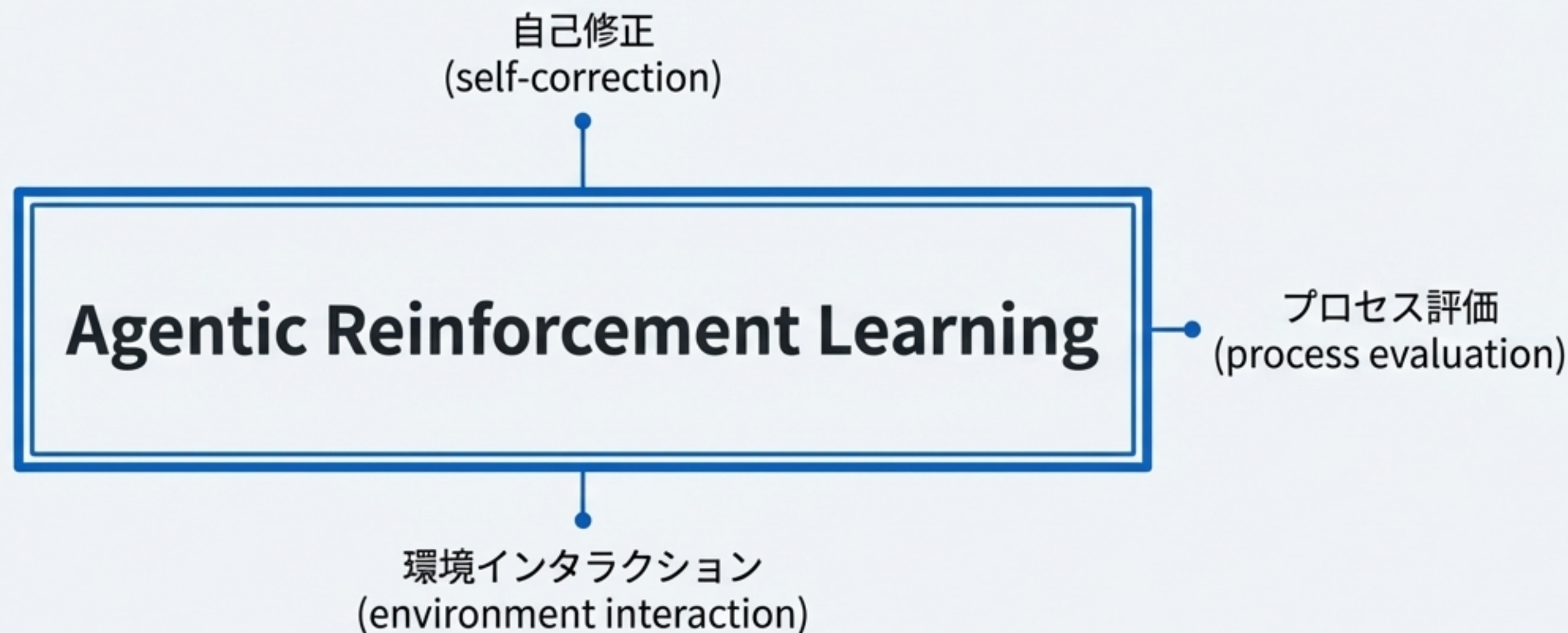


The Revelation: それを可能にした核心技術とは何か？
(The Technology: Agentic RL)



The Consequence: この地殻変動は何を意味するのか？
(The Future)

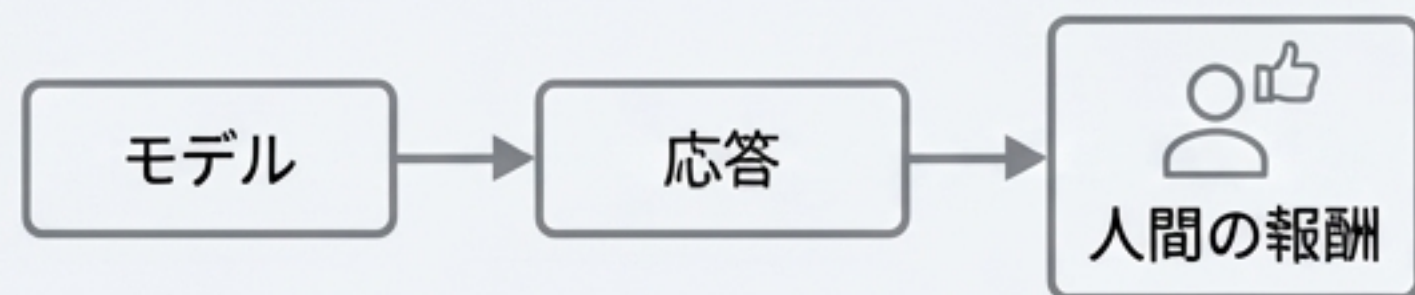
答えは「Agentic RL」：AIの学習パラダイムが根本から進化した



Flashの飛躍の裏には、モデルを単なる「テキスト生成器」から自律的な「エージェント」へと変貌させる新しい学習手法、Agentic RLが存在する。これはAIが自律的に思考し行動する歴史的な転換点である。

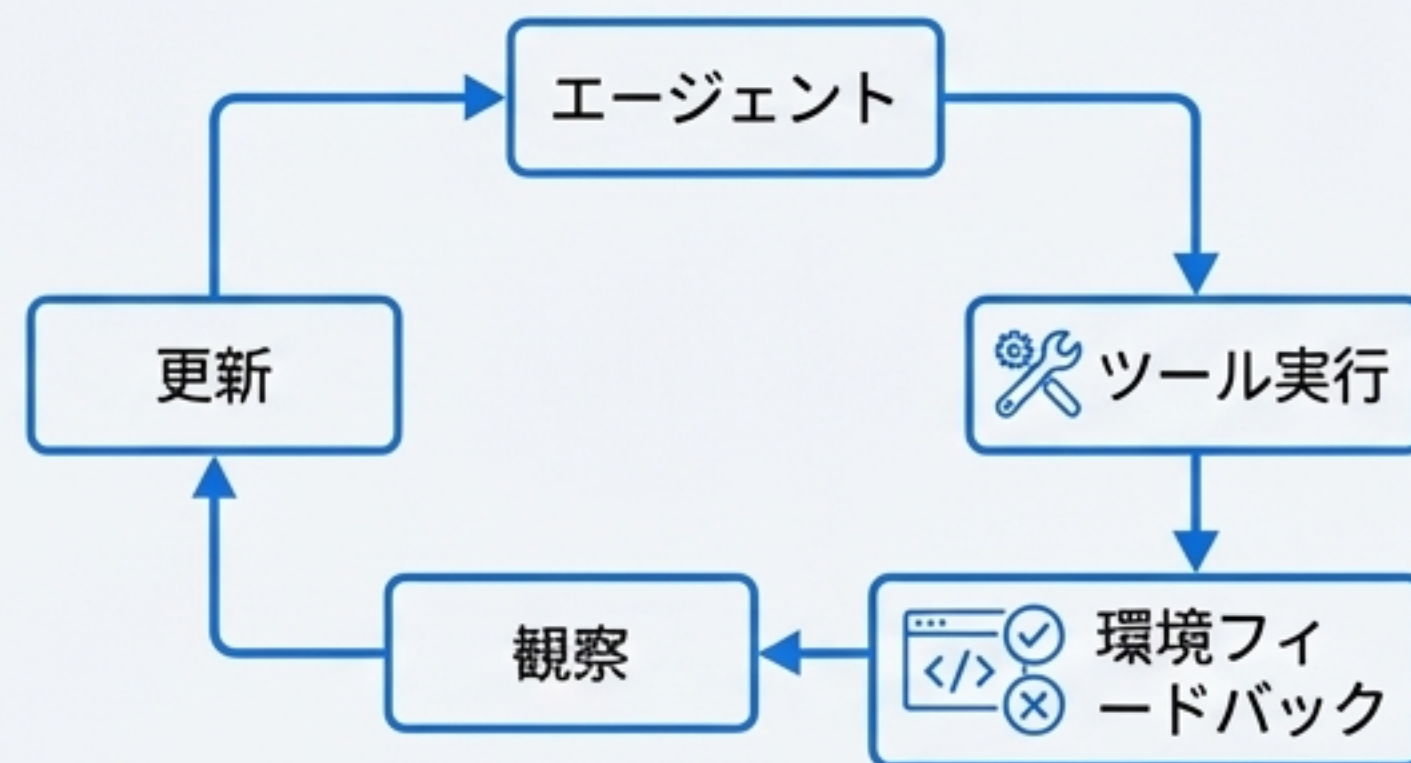
学習パラダイムの進化：静的な「RLHF」から動的な「Agentic RL」へ

従来のRLHF



- 結果主義 (result-oriented)
- 静的なフィードバック (static feedback)

エージェント強化学習

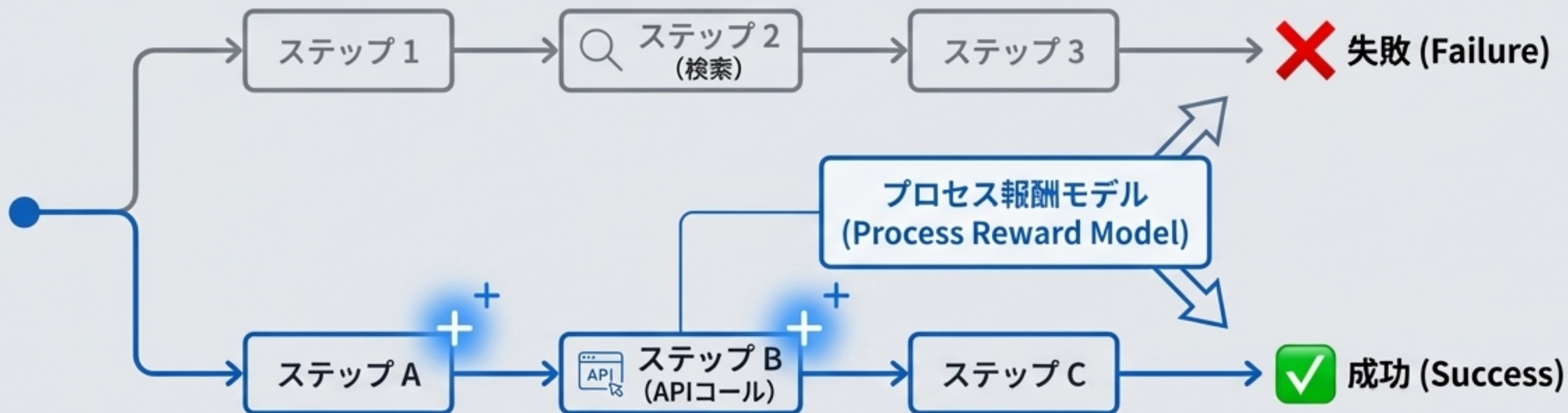


- プロセス重視 (process-oriented)
- 動的な試行錯誤ループ (dynamic trial-and-error loop)

左：従来のRLHFは人間の好み(Preference)に基づく単発のフィードバック。右：Agentic RLはツール実行環境からの動的なフィードバックとマルチターンの試行錯誤ループを含む。

Agentic RLの核心①

iStar：「結果」ではなく「成功に至るプロセス」を学習



人間が全ステップを評価せずとも、成功と失敗の軌跡全体を比較することで、成功に貢献した「良い手順」（正しいAPIコールや検索クエリなど）を自動的に特定し、報酬を与える。モデルは「正解」だけでなく、「正解に至るための正しい手順」を学習する。

Agentic RLの核心②

rStar：実行し、失敗し、そして「自己治癒」する能力



コード生成



エラーメッセージ



自己修正

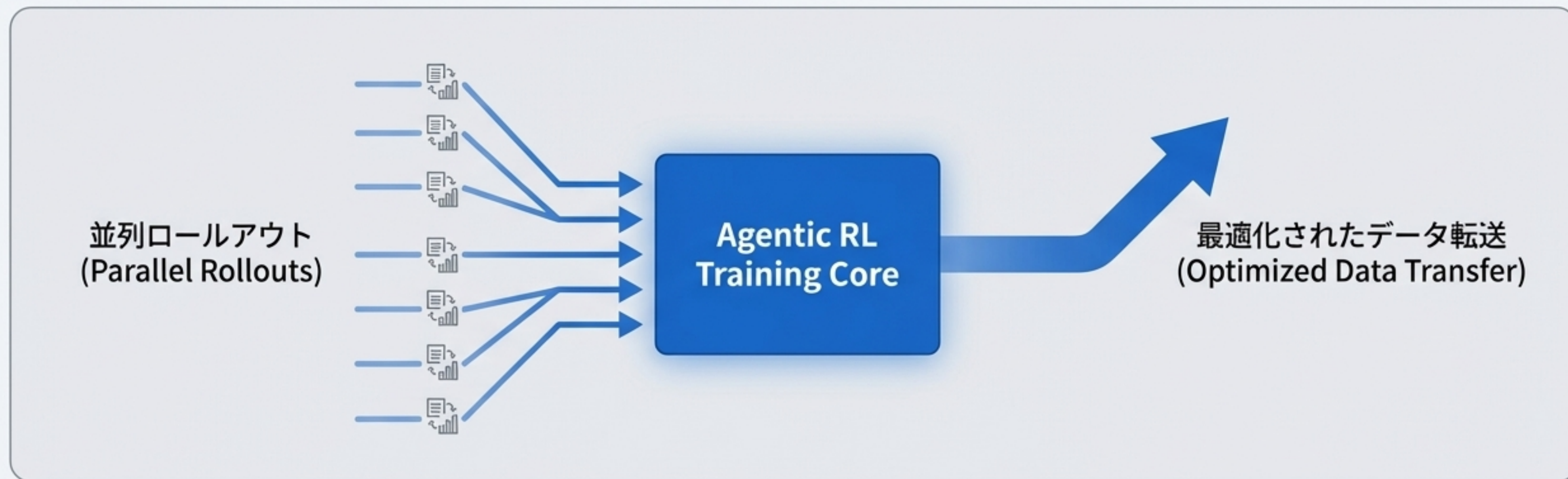


実行成功

モデルはコードを書き、実行し、エラーが出ればその内容を理解して自ら修正する「自己治癒 (Self-healing)」的な挙動を学習する。これにより「間違えないAI」から「間違いから回復できるAI」へと進化する。

Agentic RLの核心③

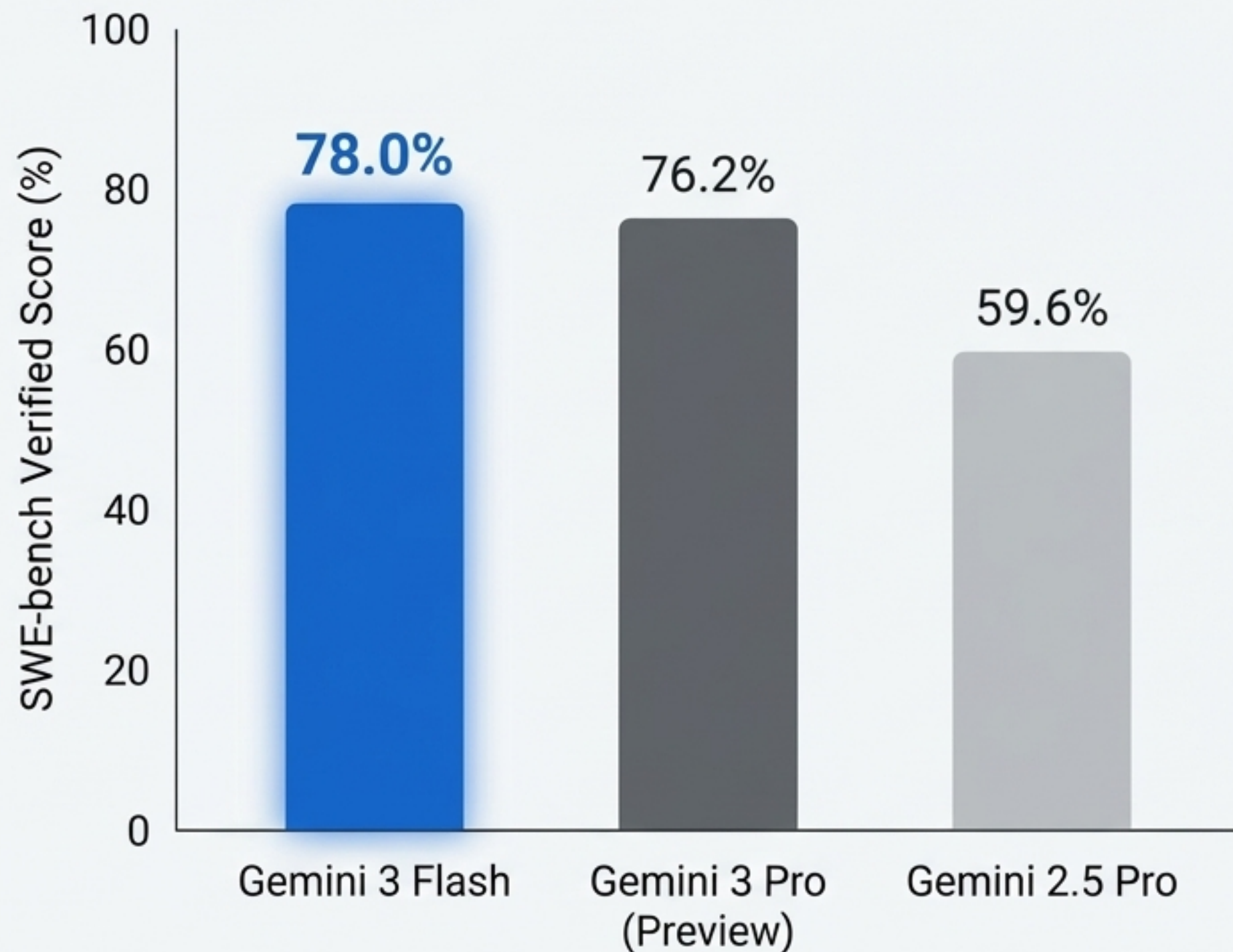
EARL：膨大なエージェント学習を可能にするシステム



エージェントの「思考と行動」のループは計算コストが高い。EARLのような効率的なシステムアーキテクチャが、並列処理の最適化やデータ転送の効率化を行い、この高度な学習を商用モデルへ適用することを可能にした。

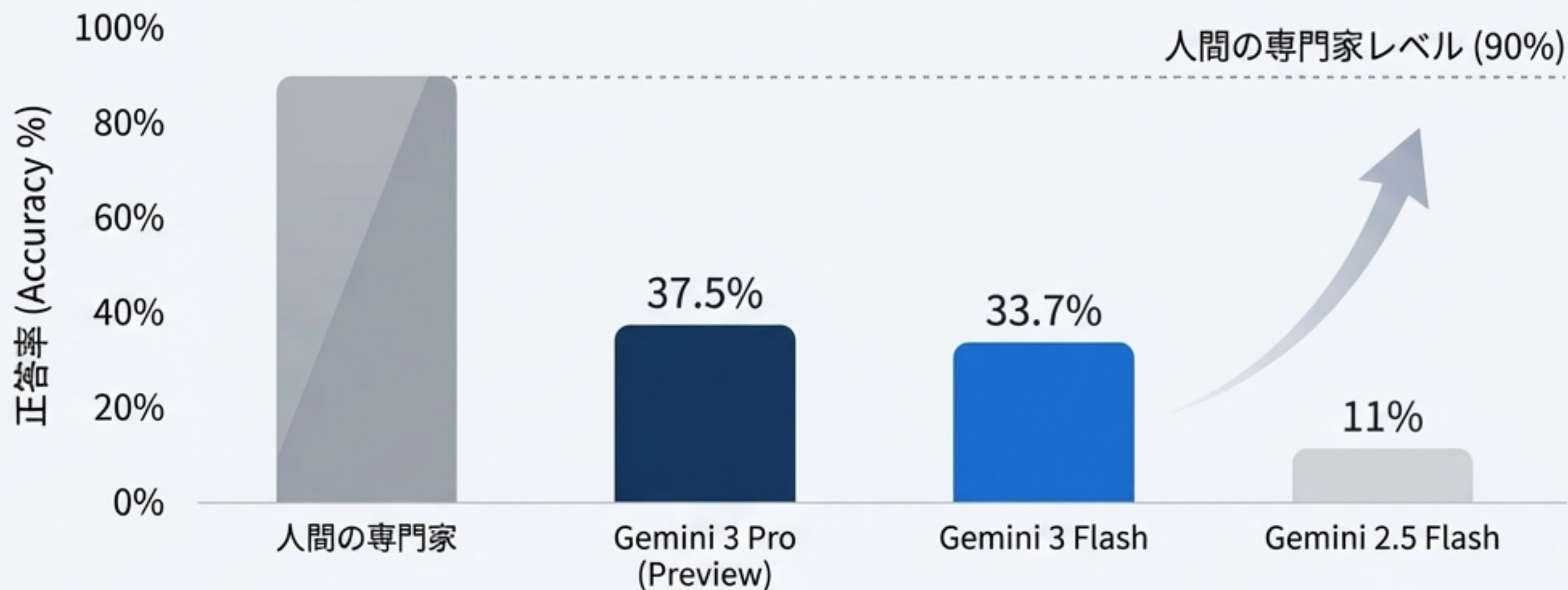
能力証明①

SWE-bench：シニア開発者レベルのコーディング能力



実際のGitHub上の課題解決能力を測るテストで、FlashがProを上回りトップスコアを記録。Agentic RLによる試行錯誤とツール使用能力が、純粋なモデルサイズを凌駕したことを証明している。

能力証明② HLE：「人類最後の試験」で示す博士号レベルの推論



人間の専門家レベルの難問に対し、スコアは33.7%とPro (37.5%)に肉薄。前世代(11%)からの劇的な進化は、軽量モデルでありながら思考プロセスそのものが深化していることを示唆する。

開発者への福音：Google Antigravityでエージェントが 主役になる



低レイテンシ

思考と実行の高速ループで、
開発者のフローを止めない。



高信頼性のツール使用

複雑なリファクタリングやテ
スト自動化を安心して任せら
れる。



透明な思考プロセス

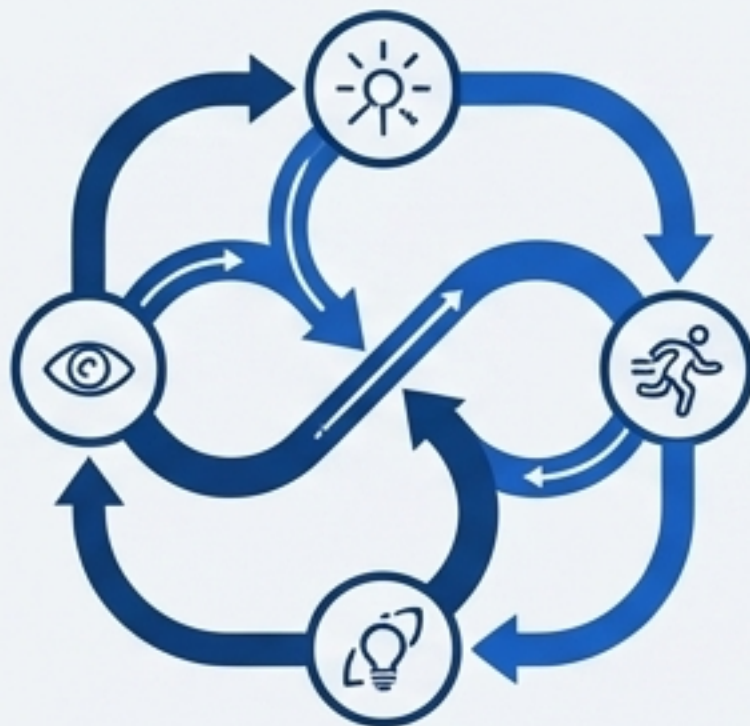
エージェントの行動ログが理
解しやすく、デバッグや協働
が容易に。

未来展望：最強の「脳」と最強の「身体」が融合する時



[Gemini 3 Proの巨大な
知識量・パラメータ]

+



[Agentic RLによる
高度な思考・行動能力]

=



[AIの歴史上、
最強のモデル]

Flashは「最新鋭の学習脳を持った軽量な身体」であり、Pro Previewは「一世代前の学習手法を持つ巨大な身体」だった。この最強の脳が、より巨大で知識豊富なProモデルに搭載された時、AIの性能は線形ではなく、指数関数的に飛躍する可能性がある。

Gemini 3 Pro（GA版）はAIの「できること」を再定義する



HLE 50%超え：人間の専門家との差を本格的に詰め始める最初のモデルになる可能性がある。

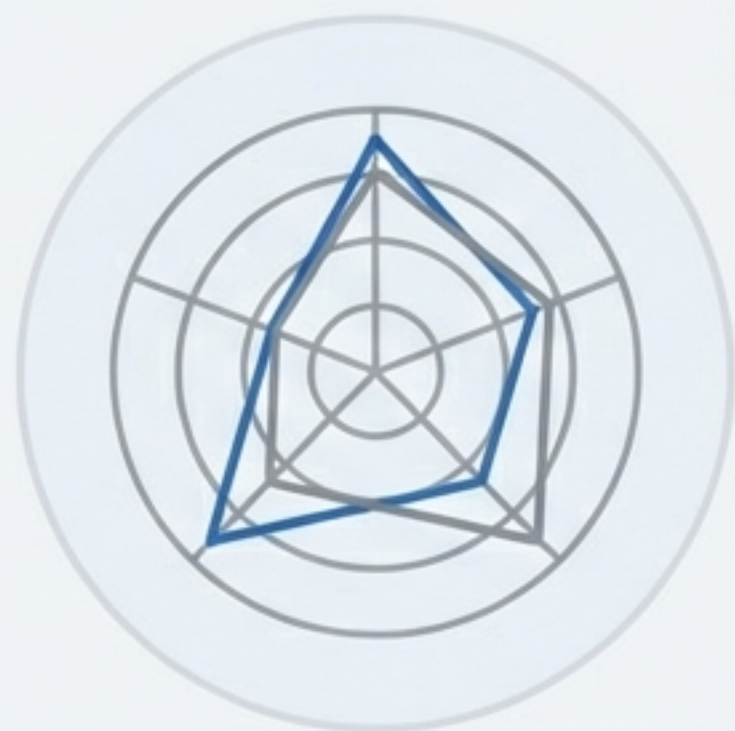


完全自律型リサーチの標準化：数時間をかけてレポートを生成する「Deep Research Agent」が実用的な標準機能へ。

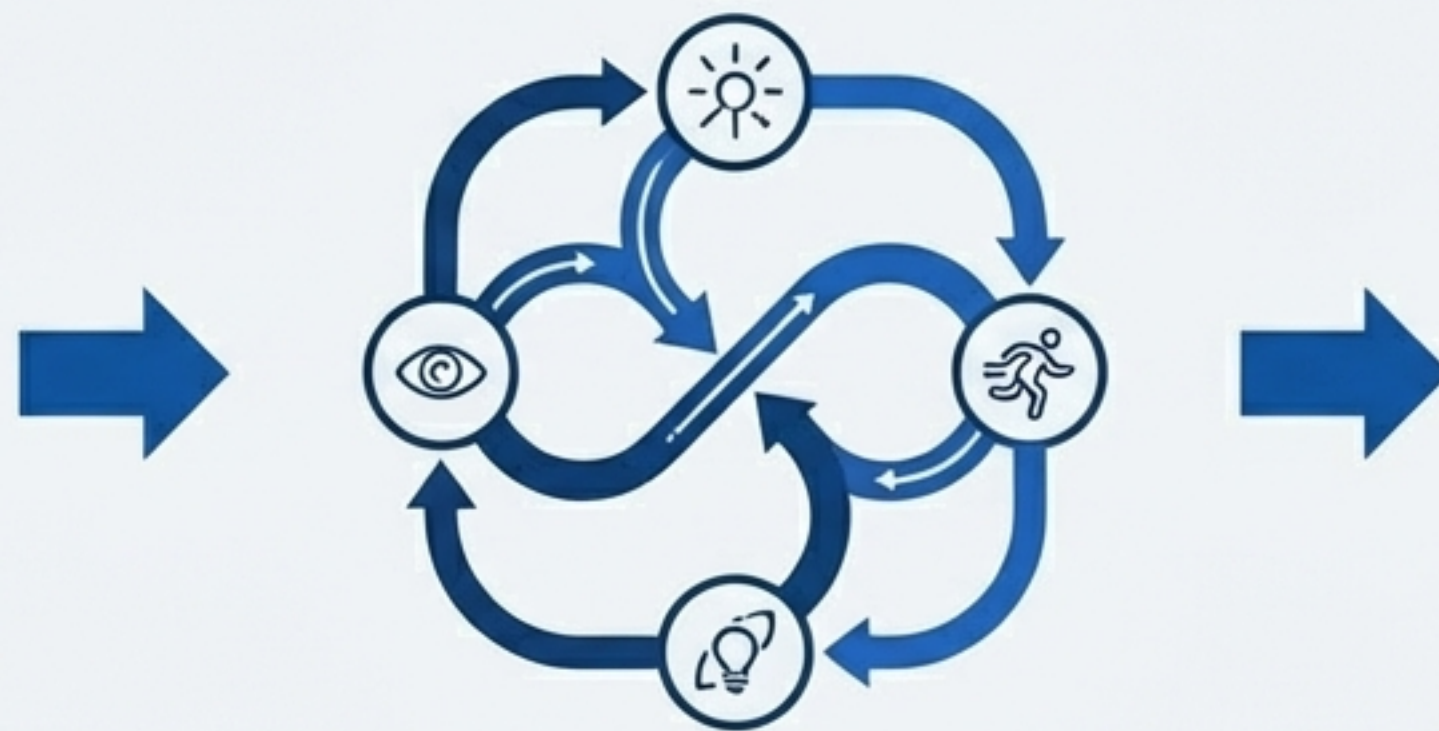


「System 2」思考の常時稼働：難問に対し、数分間熟考してから回答する人間のような思考が当たり前になる。

結論：AIは「テキスト生成器」から「信頼できる仕事の代行者」へ



逆転現象



Agentic RL



未来 (Pro GA)

Gemini 3 Flashは、AIの性能がもはやモデルサイズだけでなく、「学習の質」によって決まる新時代の幕明けを告げた。これは、AGIへと続く道筋を示す、確かな未来への予告編である。

古い神話は打ち砕かれた： 「モデルサイズが全て」の時代の終わり

Agentic RL の登場は、AI の性能決定要因において、モデルサイズ (Scaling Law) と同様に、ポストトレーニングの質 (Agentic Alignment) がいかに重要かを示している。