

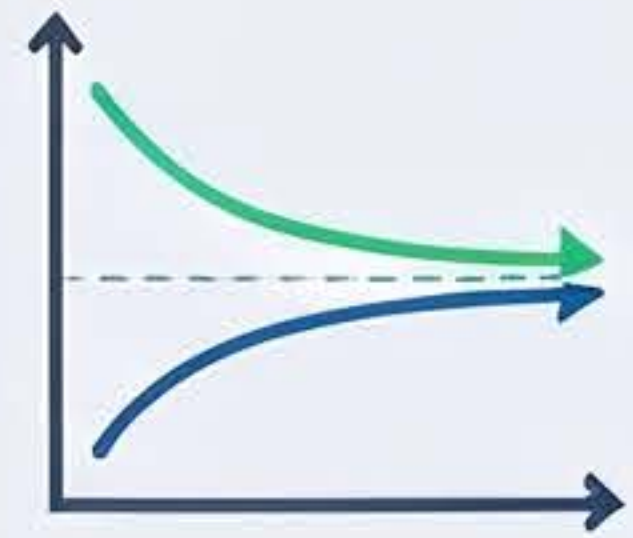
AIエージェント・ハーネスのパラダイムシフト：知能の「補助輪」から「実行OS」へ

なぜ「ハーネス不要論」
が生まれたのか



モデル自身による
「思考の補助輪」の吸収

事後学習や強化学習によりLLM自身が
計画・ツール利用能力を獲得しました。



定型タスクでのハーネス効果の急減
(差分+0.6ptへ収束)

最先端モデルでは複雑な制約を追加しても
スコア向上幅は極めて微小です。



6.2x

Old

New

同等精度でも発生する
最大6.2倍の実行コスト格差

長大な単一セッションはキャッシュ再設定コスト
を爆発させ、推論費用を激増させます。

旧世代（知能補完型）と次世代（実行基盤型）
ハーネスの設計思想の違いを比較する



旧世代 (思考の補助輪)



主な役割
推論・計画・ツール選択
など知能の補助



ツール環境
縦分化された事前定義
ツール群



主な実行場所
ローカルPC・
エディタ画面内

OLD PARADIGM
(INTELLIGENT AUXILIARY)



次世代 (自律実行OS)



主な役割
セキュリティ・状態永続化・
コスト制御



ツール環境
未統的なBash環境 +
オンデマンドスキル



主な実行場所
クラウドVM・本番サーバー群
(タスクの90%)

NEW PARADIGM
(AUTONOMOUS EXECUTION OS)

PARADIGM
SHIFT

次世代アーキテクチャ
「自律実行OS」の3大要件



「Bash中心」の最小限ツール
定義によるコスト削減

細小ツールを廃しBash操作へ集約することで、
実行コストを最大53%削減できます。



タスクの90%が「PCの外」へ：
クラウドVM内密駐化

エージェントをVM内プロセスとして帯びさせ、
高速な状態フォークと永続化を実現します。



ステートフル権限制御と
厳格な自動検証ゲート

言葉による自己申告の完了を認めず、
テスト成功のエビデンスで厳格に判定します。