

人間のAIに対する「判断移譲」の変遷と未来（2024–2026）

2024–2026年：判断移譲の3段階ステップ

2025年：ファイル編集・
コマンド実行の自動化



作業の実行許可を段階的にAIへ渡し、
承認待ちを減らし始めました。

2026年8月：「承認の要否」
自体をAIが判断（標準化）

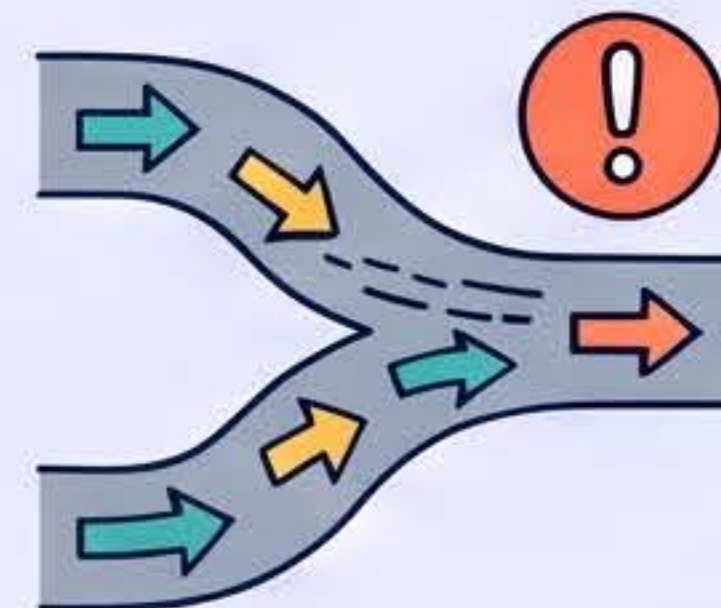


主要開発ツールでAIが危険度を判定し
自動承認する設定が既定使化。

移譲の背景：承認の形骸化（承認率97%）と速度低下



移譲の背景：
承認率97%）



人間の手動承認が開発速度のポ
トルネックとなっていました。

人間の新たな役割：個別承認から「規則設計・監査」へ



毎回の「実行許可」から…



…「ルール・境界の設計」へシフト
人間は個別の操作ではなく、拒否ルール
やサンドボックス境界を設定します。

人間に残る3つのコア領域



目的と価値の設定

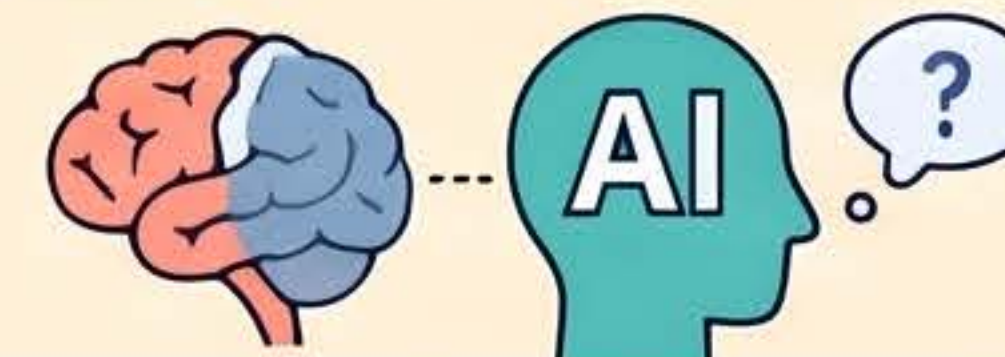


不可逆な操作の境界設計



事後責任と監査

⚠ 課題：過剰依存と人間側の技能低下リスク



AIへの過信による批判的思考の減少やスキル喪失
への対策が必要です。