

ループエンジニアリング時代の知財実務：自律型AIを制御する新基準

2026年、AI活用の中心は単発の「プロンプト」から、AIが自律的に計画・実行・評価を繰り返す「ループエンジニアリング」へと移行しています。知財業務においては、AIの暴走を防ぎつつ、人間の創作的寄与を立証するための「境界付き自律性」と「証拠としてのログ管理」が不可欠な戦略となります。

プロンプトから「ループ」への進化

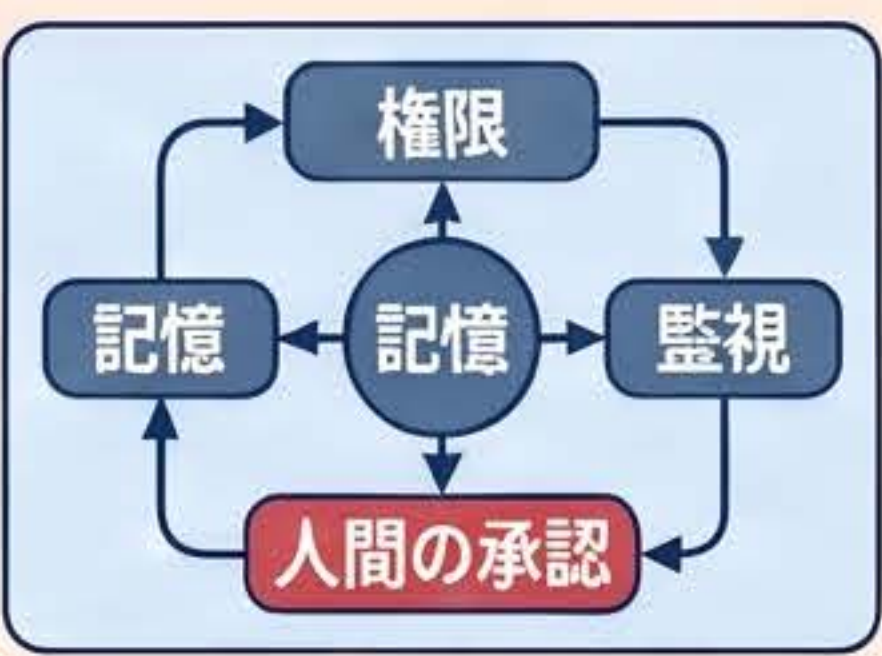
ループエンジニアリングとは



ループエンジニアリングとは
目標に向けAIが「計画・実行・観察・修正」を反復するシステム全体の設計。



上位概念への移行



単発のプロンプト
単なる指示文ではなく、権限、記憶、監視、人間の承認を一体として設計する。

知財業務の「境界付き自律性」
法的効果を伴う地点では必ず人間が承認し、AIの行動を明確な境界内に制限する。



法的効果を伴う地点
法的効果を伴う地点では必ず人間が承認し、AIの行動を明確な境界内に制限する。

知財リスクの統制と証拠管理

ログを知財証拠基盤として位置付ける



ログを知財証拠基盤として位置付ける
誰が、どの技術課題に、どう創作的に関与したかの軌跡を改ざん耐性のある形で記録する。

報酬・目的関数の多目的化



自律レベルに応じた承認ゲートの設置



出願、公開、契約承諾などの「不可逆な行為」は、AI単独で実行させない。

知財業務における自律レベルに応じた権限と承認の基準

自律レベル	AIの権限	知財業務の例	承認プロセス
提案 (レベル2)	案を作るが実行はしない	クレーム案・契約修正案の作成	担当者による郡度承認
条件付き実行 (レベル3)	定型・可逆処理を実行	社内タスク登録、定型通知	事前包括承認+事後監査
高リスク実行 (レベル4)	法的効果・外諺公開を伴う	出願、権利放棄、契約承諾	二人承認+専門家必須