

知財AIエージェントを成功に導く「ハーネス設計」とガバナンス

知財業務におけるハーネス導入の優位性

エージェント・ハーネス
(Agent Harness)

法的リスクを防ぐ二重検証とライフサイクル制御

タスク解決率
20%pt以上向上

モデルの重みを固定したまま制御機構を最適化するだけで精度が大幅改善。




決定論的リンター
構文チェック
(Syntax Check)


QA Judge
(意味論的評価)
高次元意味論的評価
(Semantic Evaluation)

決定論的リンターとQA Judgeによる二重検証ゲート
先行調や符号の構文チェック後にLLMによる高次元意味論的評価を実施。



危険行動率
38%低減
攻撃成功率
42%低減

SafeHarnessにより危険行動率38%・攻撃成功率42%低減
4フェーズの防衛と投機的な権限降格により機密流出や暴走を防止。



自然人による最終承認
(HITL)



イミュータブルな監査ログ

自然人による最終承認(HITL)とイミュータブルな監査ログ
弁護士の善管注意義務と署名責任を担拠する人間介在プロセスを強制。



Patent-MAF
(ローカルOSS /
役割分解ハーネス)

85.4
(総合品質スコア)



GPT-5
(外部クラウド)

85.1

Qwen3-VL-72B
(単体モデル): 01.6

知財業務におけるハーネス導入の優位性
外部API不使用のローカル環境で総合品質スコア85.4を達成。



マネージャー



ドラフター



ポリッシャー

役割分解マルチエージェントによる明細書・クレーム生成
マネージャー・ドラフター・ポリッシャーの分業と画面連携で不整合を抑制。