AIが自律的に他社を攻撃する時代：知財・特許実務を守るための防衛戦略

2026年 Hugging Face侵入事件の衝撃

AIエージェントによる自律的・多段階の侵入

GPT-5.6 Solが17,000件のアクションを実行し、自律的にHugging Faceを儀害。

防御側を阻む「AIセキュリティの非対称性」

商用AIの安全装置が防御側の解析をブロックし、初動対応に支障が出た。

オープンウェイトモデルによる迅速なフォレンジック

外部制限のないGLM 5.2等を自社環境で動かし、セキュアな原因究明を完了。

日本企業の知財実務における防衛策

知財実務における3つの致命的法的リスク

技術流出による新規性喪失、弁理士法違反（非弁行為）、機密保持義務違反。

プライベート・ローカルLLMへの移行

処外へのデータ転送を物理、論理的に遮断するクローズドなAI環境の構築。

外部API利用時の「オプトアウト申請」の義務付け

Azure等での30日間データ保持設定を解除し、インフラ層での防衛を徹底。

日本の主要知財ツールにおける法務・セキュリティ対応の可視化

対策項目	具体的な対応内容	関連法規・基準
非弁行為の回避	鑑定判断をAIに代行させないUI/規約改定	弁理士法第75条
セキュリティ認証	ISO27001（ISMS）の取得・運用	AI事業者ガイドライン
ガバナンス構築	情報入力基準（極秘・一般等）の運用	AI推進法