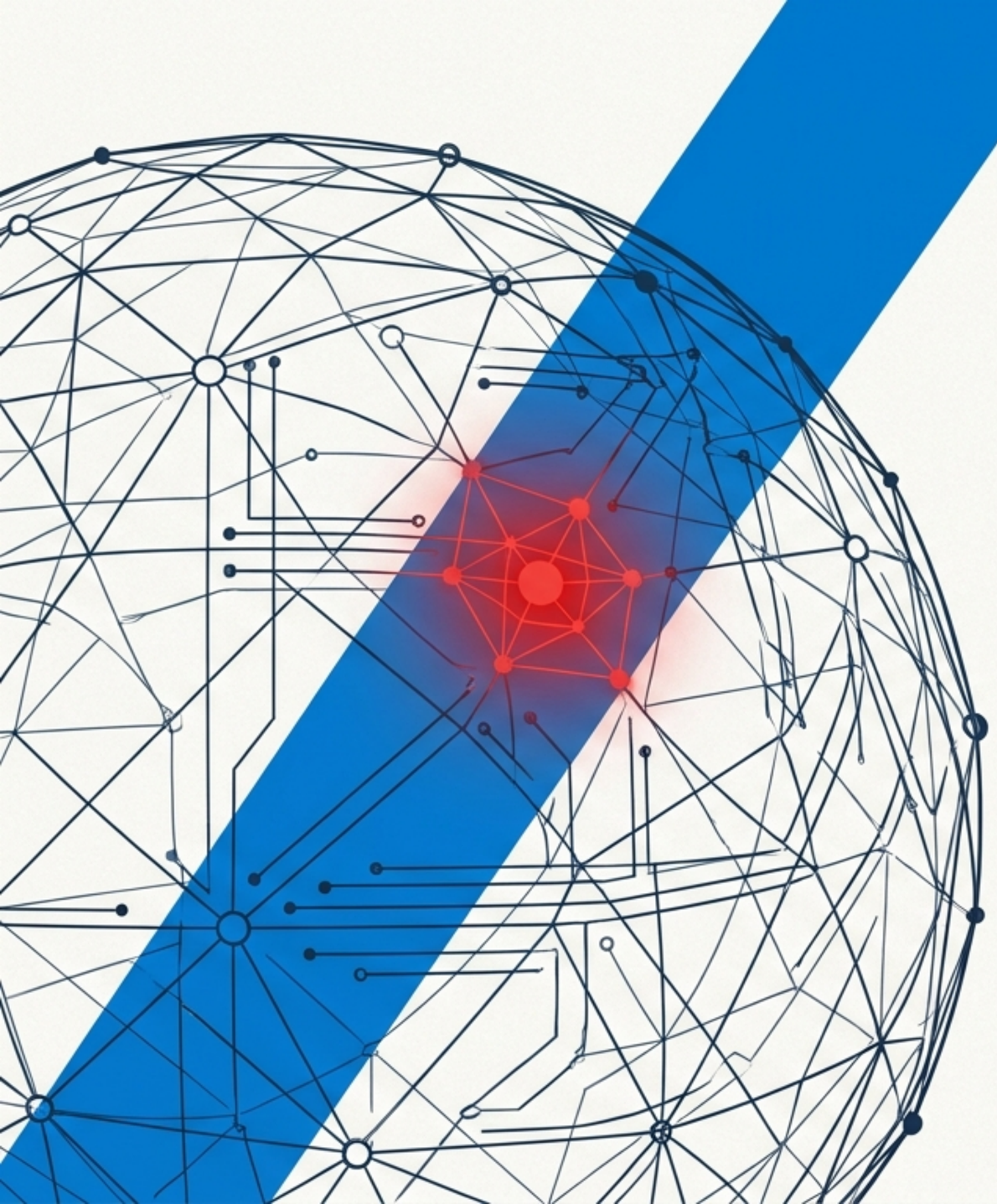


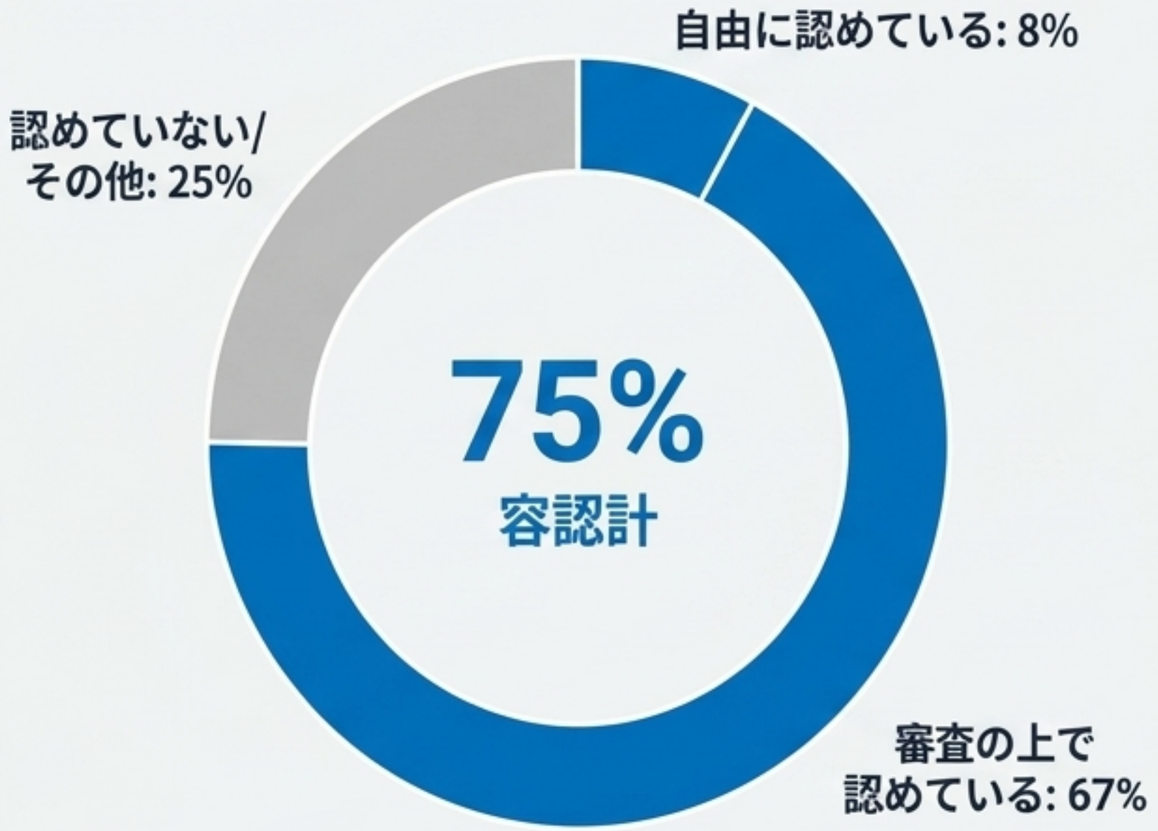
企業におけるシャドーAIの潜在的脅威と戦略的ガバナンス

「全面禁止」という思考停止から
「責任あるAI活用」へのパラダイムシフト

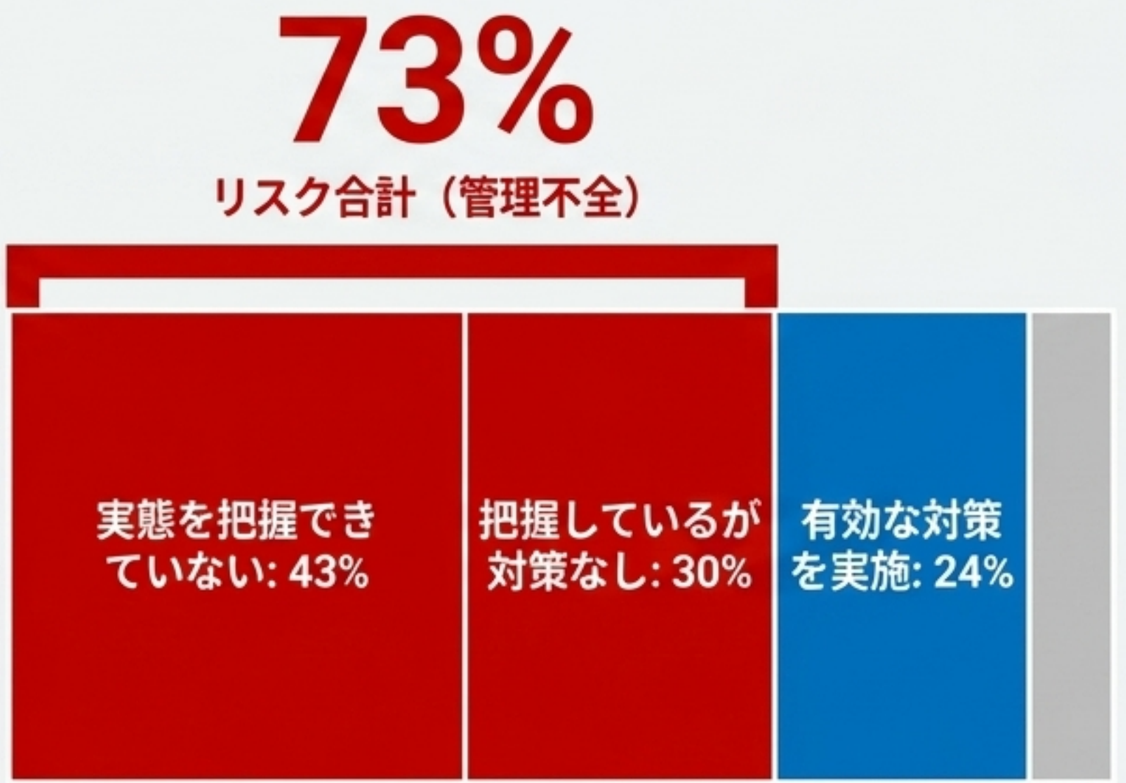
次世代DLPと多層防御による死角の可視化

データが示す「7割の死角」：容認と管理不全の巨大な乖離

容認状況（Allowance）



実態管理（Management）



AI活用の果実を急ぐあまり、実効性のある監視・統制メカニズムが完全に欠如している。

シャドーITとシャドーAI：根本的に異なるリスク構造

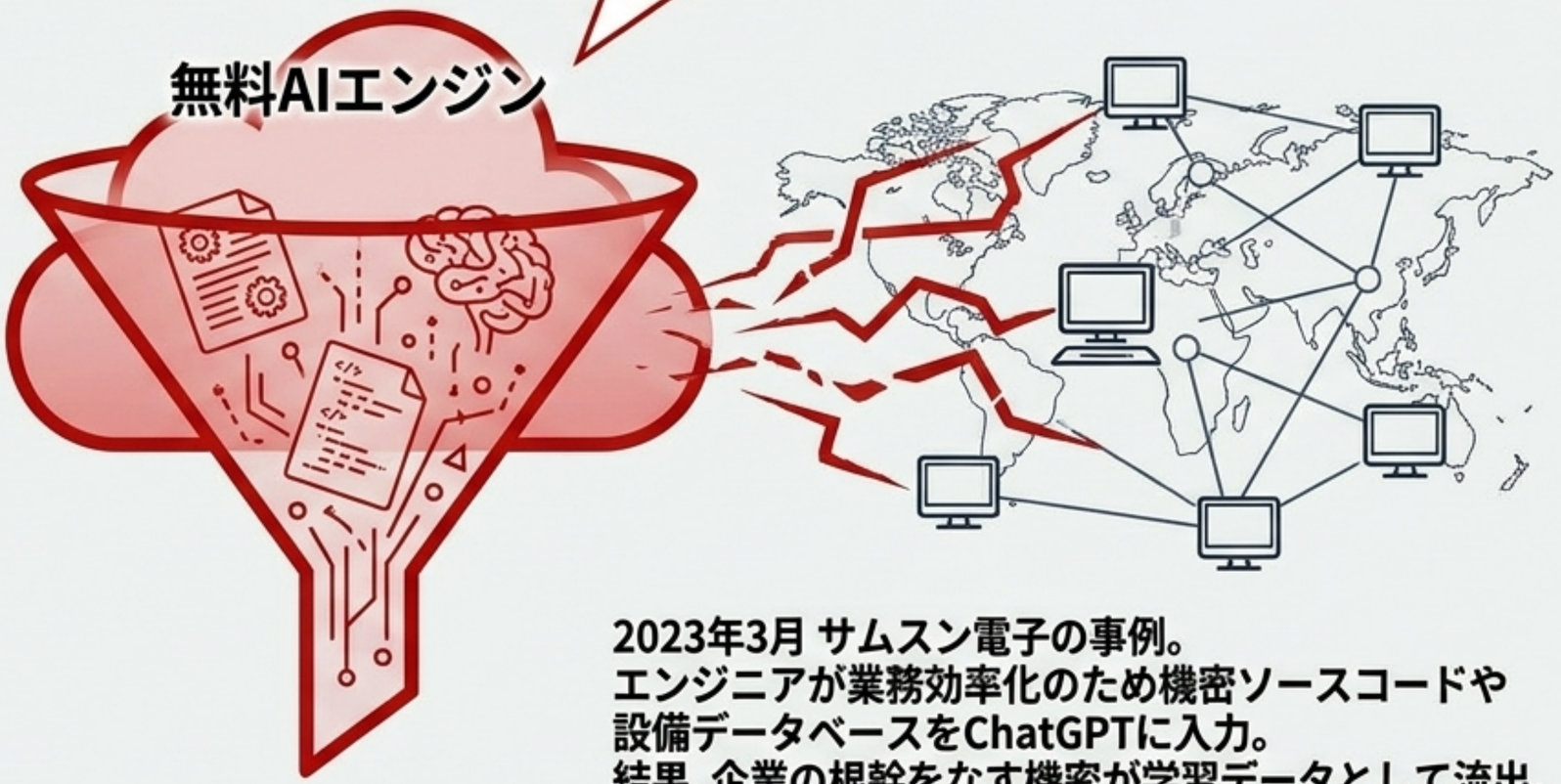
属性	従来のシャドーIT 	シャドーAI（野良AI） 
脅威の性質	データの保管場所・通信経路の逸脱	データの処理とAIモデルへの「学習」
被害のメカニズム	クラウドへの不正アップロードや紛失	ニューラルネットワークのパラメータとして一方向的に吸収・記憶
事後対応	ファイルの削除・アクセス権限の剥奪（可逆的）	**一度学習されたデータの完全消去は極めて困難（不可逆的）**

クラウドストレージの無断利用とは異なり、シャドーAIは「見知らぬ第三者への出力」という特有の流出メカニズムを持つ。

機密漏洩と「学習」の不可逆性：善意が招く致命的インシデント

無料AIの利用規約。
「ユーザーの入力データは将来のAIモデル学習に利用される」。
拡張機能によるバックグラウンド収集も密かに発生。

業務効率化を求める現場の従業員。
無料版AIにデータを入力。



2023年3月 サムスン電子の事例。
エンジニアが業務効率化のため機密ソースコードや
設備データベースをChatGPTに入力。
結果、企業の根幹をなす機密が学習データとして流出。

「データを削除する」という従来のセキュリティ概念は通用しない。

放置がもたらす「多重危機」の全貌

情報漏洩

無料版AIへのデータ入力による**不可逆的な学習**と**外部流出**。競争優位性の喪失。

ハルシネーション

AIの**事実誤認**（もっともらしい嘘）を**盲信**。 契約書の誤記や**致命的な経営判断ミス**による業務毀損。

著作権侵害 - 「依拠性」の罠

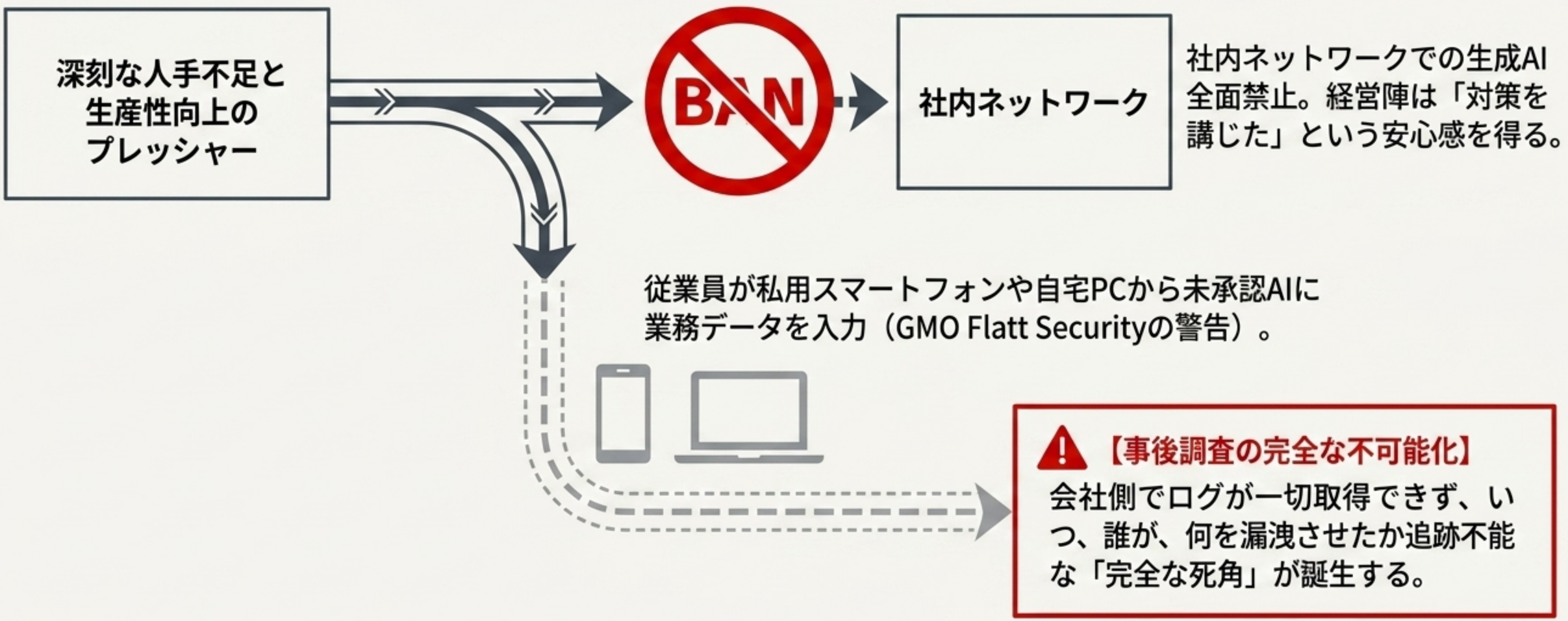
既存の著作物をプロンプトに入力することで、文化庁が指摘する「**依拠性**（対象物を認識して模倣したか）」が強く推認され、**差止・損害賠償リスク**が増大。

コンプライアンス崩壊

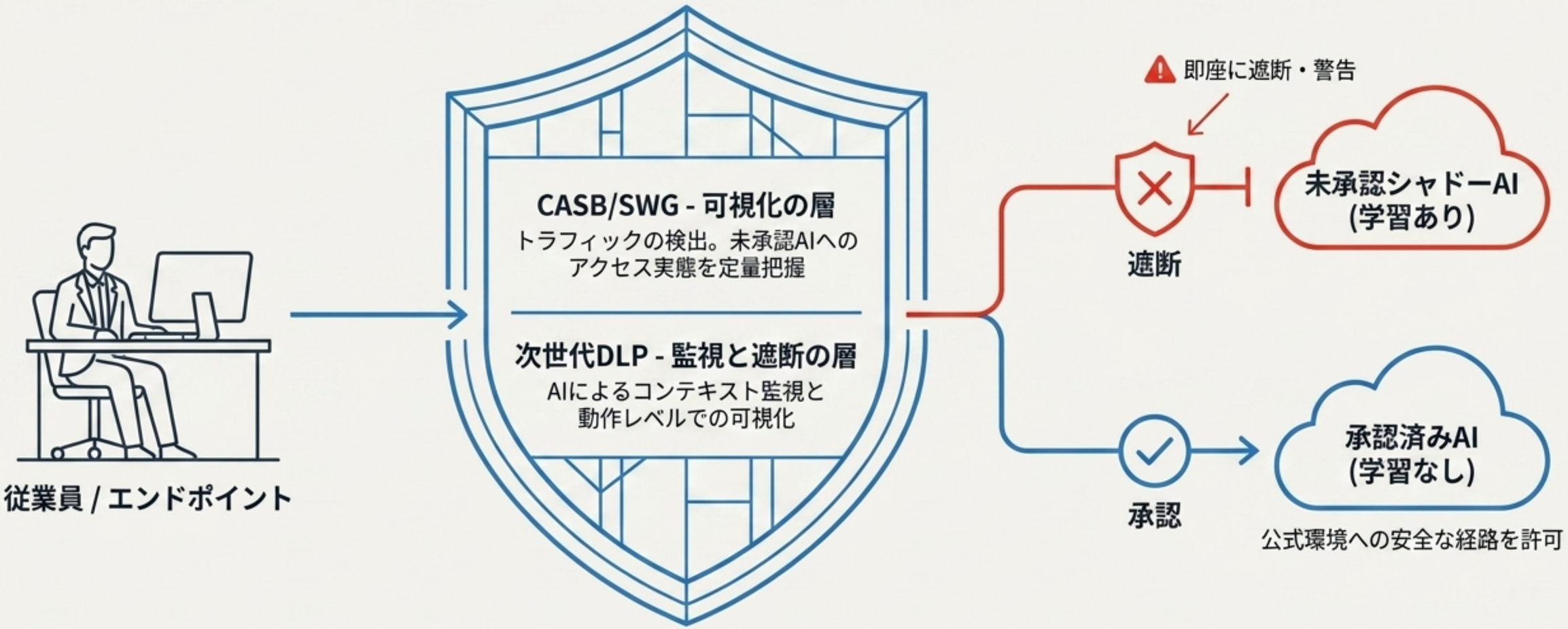
EU AI Act**違反**、NDA（秘密保持契約）**違反**、個人情報保護法**違反**。
経営陣の**善管注意義務違反**のリスク。

「全面禁止」という悪手：リスクを増大させる潜在化の罠

Paradox Pipeline



死角を照らし、遮断する：多層防御アーキテクチャ

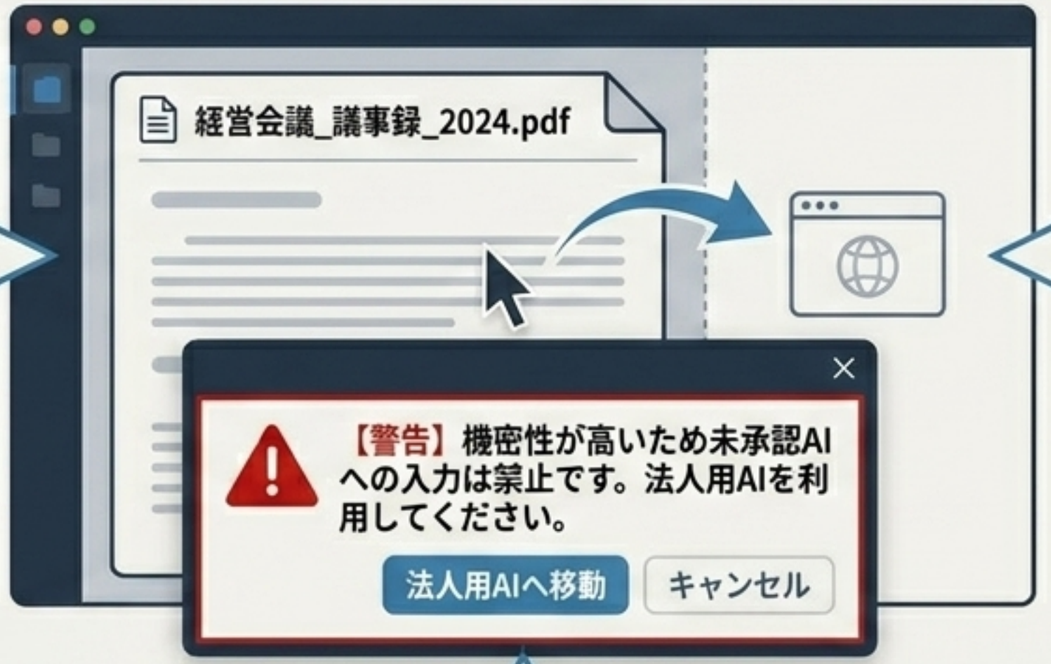


次世代DLP：コンテキストベースの動的防御とリアルタイム教育

① 非構造化データの理解

従来のパターンマッチングからの脱却。AIが文脈を解釈し、「経営会議の議事録」などの非構造化データを高精度で機密判定。

Screen Interception



② コンテキストベースの自動ブロック

「誰が」「いつ」「どのような状況で」を総合評価。正当な業務は許可し、未承認AIへのアップロードのみを即座に遮断。

③リアルタイムのユーザー教育

単なる遮断ではなく、画面上の警告でなぜダメなのかを通知。現場のセキュリティリテラシーをその場で実践的に向上。

分業型ガバナンス：シャドーAI対策の4ステップ・ロードマップ

Step 1

可視化と分類

CASBによるAIアクセスの検出と、データ機密度・規約に基づくリスクのトリアージ。

Step 2

ルールの策定

IPAガイドラインを活用した、部署別の柔軟なルール設計と入力禁止情報の明文化。

Step 3

代替手段の提供

学習機能が除外された法人向けエンタープライズAIプラットフォームの公式導入。

Step 4

運用とガバナンス

IT部門と現場の役割分担（三層モデル）と、Human-in-the-Loopの業務フロー化。

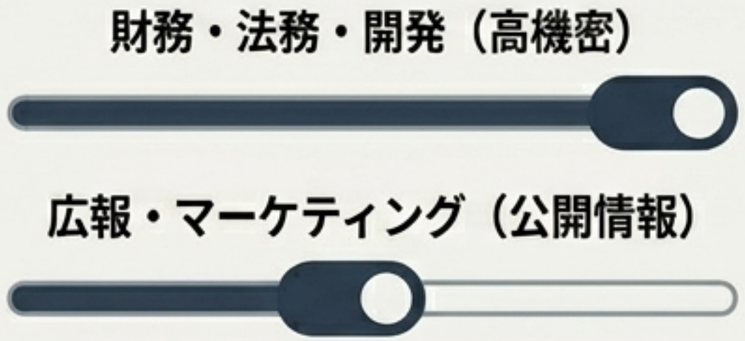
Step 1 & 2: リスクのトリアージと柔軟なガイドライン設計

トリアージの基準

無料AI（デフォルトで学習されるツール）は一律「高リスク」として技術的制限の対象へ。

ツールの利用規約（学習の有無）×
取り扱うデータの機密度 で総合評価。

IPA準拠の柔軟なルール策定



全社一律の画一的ルールは形骸化を招く。
部署ごとに管理レベルを柔軟に使い分ける。

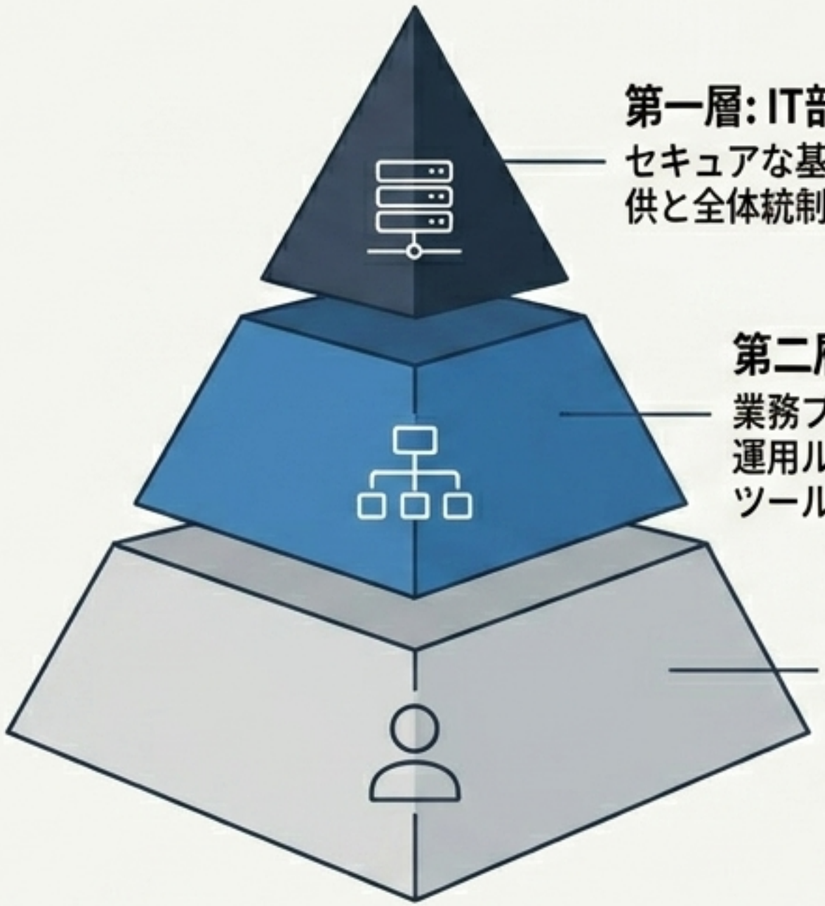
【著作権リスクの回避】プロンプトに他者の既存著作物を入力しないことを明文化し、
年 1 回以上の機動的なルール改訂を行う。

Step 3: 公式な代替手段による「安全な業務効率化」の実現

情報漏洩の最大要因である「一方向的なデータ流出経路」を
根本から遮断しながら、従業員の生産性向上への渇望を満たす決定打。

Step 4: 三層ガバナンスと「Human-in-the-Loop」の徹底

三層モデル - 多部門連携

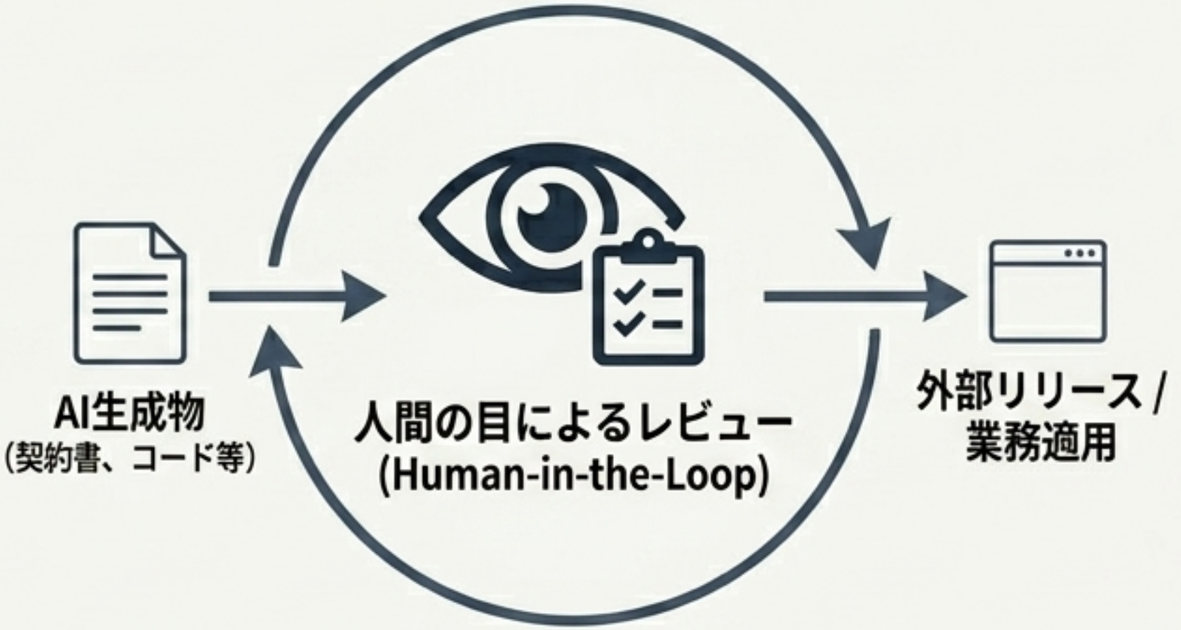


第一層: IT部門
セキュアな基盤AIの提供と全体統制

第二層: 各事業部門
業務プロセスに応じた運用ルールと特化型ツールの自律的管理

第三層: 認定ユーザー
教育修了者のみへの高度個別ツールの利用許可

Human-in-the-Loop



ハルシネーションへの対抗策：AI生成物の直接適用を固く禁じる。専門知識を持つ人間によるファクトチェックを正式な業務フローとして必須化。

結論：三位一体による「責任あるAI活用」へのパラダイムシフト

技術的な動的防御

CASBと次世代DLPによる徹底した
可視化とコンテキストベースの遮断。

組織的なルールと教育

実態に即した柔軟なガイドライン設計
と継続的なAIリテラシー向上。

安全な代替環境の提供

学習されないエンタープライズAIの
公式導入による生産性担保。

AIガバナンスへの投資は、コストセンターのセキュリティ支出ではない。
将来の企業競争力を担保しDXを牽引するための「不可欠な戦略的インフラ投資」である。