AI事業者ガイドライン第1.2版：一般企業に求められる「自律型AI」への実務対応

2026年3月の改訂は、AIが自律的に行動する「AIエージェント」時代に対応し、高度なガバナンスと「人間の介入」が義務化されます。

旧ガイドライン（第1.1版・現行）：受動的な「ツール」利用

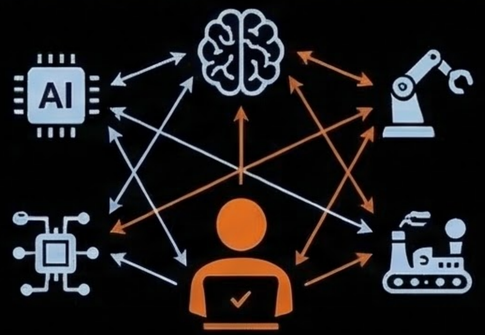
対象AI：受動的なチャットボット

統制メカニズム：入力時の注意（リスク管理）

利用者の責任：免責に近い「利用者」

新ガイドライン（第1.2版・2026年3月～）：能動的な「遂行主体」と責任ある利用

対象AI：自律的エージェント・物理システム



対象AI：自律的エージェント・物理システム

統制メカニズム：出力・実行前の人間介入（HITL）

Human-in-the-Loop（HITL）の実装義務化

AIが外部へ影響を与える操作（メール送信、決済等）を行う際、人間の承認を必須とする設計が求められます。

70% 国内事業者の約70%が導入・検討中

AIエージェントの普及に伴い、ガバナンスはリスク管理から「イノベーションの加速装置」へ再定義されます。

利用者の責任：運用・統合管理者としての重い責任

企業が直ちに着手すべき「３つのアクションプラン」

STEP 1 全社AI利用状況の棚卸しとリスクマッピング

シャドーAI（承許可判用）を特定し、EU AI法などのグローバル規制への抵触リスクを評価します。

STEP 2 HITL（人間の判断必須）のワークフロー構築

AIが下書きし、人間が最終確認・送信ボタンを押す「承認ゲート」をシステム的に実装します。

HITLの要件：
内部完結（低リスク）：事後・任意認認可
外部影響（高リスク）：システム的「事前承認」必須 ⚠

STEP 3 データ処理委託契約（DPA）の刷新

ベンダーに対し、入力データがAIの学習（機械学習）に流用されないことを契約で担保します。