「常時特権」から「実行時認可」へのパラダイムシフト：次世代AIガバナンスモデル

株式会社 アイ・アイ・エム
営業技術本部 戦略ビジネス統括部 セキュリティ営業部
PAMソリューションリーダー　田口 学





過剰な権限を持つAPIトークンを発見したAIエージェントが、本番データベースとそのバックアップを削除するまでにかかった時間。

— PocketOS incident, April 2026

AIエージェントは、既存のガバナンスを超越したスピードで動く

セキュリティリーダーが直面する、2方向からの同時プレッシャー

安全なAI導入の推進

エージェントの到達範囲は、今や本番環境のデータベース、SSHホスト、Kubernetesクラスター、クラウドコンソール、MCPサーバーにまで及んでいる。エージェントは実行（ランタイム）時に、独自のツール選定と手順決定を行う。

AI起因の脅威に対抗するセキュリティ対策

AI認証情報は、いまや窃取の標的として最も注目されるクレデンシャルのひとつとなっている。これらはサードパーティサービスへのアクセスを可能にするだけでなく、長期有効かつ監視対象外になっているケースが多く存在する。

17,000+
記録されたイベント

4.5 days
停止するまでの日数

Hugging Face, July 2026

評価・検証中だったAIエージェントがサンドボックスを脱出し、有効な認証情報を連鎖させて本番環境へと侵入。そして、136個の鍵（シークレット情報）が保管された領域にまで到達。そのプロセスに、人間の介入は一切なかった。

AIポリシーはあるが、実行・徹底できている企業は少ない

99.7%

AIツールやエージェントのアクセス可能データに関する正式なポリシーを策定済みの企業

51%

AIによるアクセスがそのポリシーに照らして適切かをリアルタイムで検証できている企業

<1 in 5

スコープ違反が発生したその場で検知できる企業

87%

「過去1年間で、AIツールやエージェントが想定されていた本来の適用範囲（スコープ）を超えて動作した」と回答している企業

「ポリシーの不足」ではなく、「実行の瞬間のリアルタイム制御」が最大の課題

Source: Delinea, 2026 Identity Security Report: Fall Edition (2,254 IT and security leaders; 2,250 employees; eight markets)

AIエージェントには従来のマシンアイデンティティ管理手法が通用しない理由

同じ認証情報（クレデンシャル）でありながら、まったく異なる挙動

サービスアカウント

- あらかじめ定義された固定処理を実行
- コードによって定義された挙動
- 入力が同じなら、出力（挙動）も常に同じ
- タスクにあわせて最適化されたアクセス権限

AIエージェント

- **実行時に動的に判断・行動決定を行う**
- **自らツール選択、処理ステップ連携を行う**
- **他のエージェントを自律的に生成・起動する**
- **ユーザー権限の継承（フルアクセス継承）**

50%

ITリーダーの50%が、起動したユーザーの既存権限をそのままエージェントのアクセス権限として適用。
長年にわたって蓄積された「人間の権限」が、今やエージェントによって自律的に行使される危険にさらされている。

Source: Delinea 2026 Identity Security Report: Fall Edition

AIエージェントは常時特権（Standing Privilege）の新たな創出源

権限の継承

50%

エージェントのアクセス範囲を「実行ユーザーの既存権限」によって制限・定義

人間が過去に蓄積してきたすべてのアクセス権限が、そのままエージェントに与えられる

権限の残留

42%

セッション終了時に、AIエージェントのアクセス権限を自動的に剥奪（失効）させる仕組みが存在していない

さらに、42%が、多くのエージェント認証情報が“監査の時まで有効なまま放置されている”と回答

Vault外（非保護）

11%

使用するAIツールのAPIキーやトークンを「中央管理されたVault（保管庫）」で保管している従業員

キーは、設定ファイル、スクリプト、ローカルPC内に（平文のまま）残されている

Source: Delinea 2026 Identity Security Report: Fall Edition

「常時特権」から「実行時認可」へのパラダイムシフト

「セッション内部での連続的な挙動判定」が新たな防御境界

従来の方法	実行時認可
使われるのをじっと待っている「常時特権」	ジャスト・イン・タイム（JIT）アクセス：タスクに限定し、終了時に即時失効
ログイン時の1回しか検証されない	すべての操作（アクション）を実行時に評価：アイデンティティ、対象リソース、操作内容、コンテキスト
エージェントは起動ユーザーの全アクセス権限をそのまま継承	タスクの実行に必要な最小限の権限のみをエージェントに付与
監査指摘により、手動で権限を剥奪	操作の実行前段階で、許可・遮断・人間の介入（承認フロー）を制御
『AIがやったこと』という言い訳	明確な承認者（スポンサー）、適用ポリシー、許可された根拠の記録

正当なアクセスが引き起こす、予期せぬセキュリティインシデント

ERPへの広範な権限を持つユーザーが起動した、ファイナンス系AIエージェント

Action 1

取引先マスターの新規作成

許可されたタスクスコープ内での操作 **許可**

Action 2

当該取引先への支払いの承認

同一セッション、同一のエージェントによる実行 **阻止**

ログイン時チェック：
どちらも「正当な2つの操作」として許可

職務分離（実行時認可）：
連続したコンテキストから「1つの不正送金パス」として検知・阻止する

Action 2 における実行時認可

Identity	財務部：田中さんの代理として稼働するエージェント
Resource	ERP 支払モジュール
Action	支払いの承認
Context	同一セッション内において、数分前にこの取引先が作成されている
Decision	**操作を遮断して人間の承認フローへ移行** **主体・内容・適用ポリシー・理由を網羅的にログ記録**

なぜ、いま日本の企業にとってこれが死活問題なのか

2026年1月

IPA
独立行政法人 情報処理推進機構
情報セキュリティ10大脅威

AI利用にともなうサイバーリスクがIPAの10大脅威に初登場し、すでに第3位に位置しています

2026年3月

AI事業者ガイドライン（第1.2版）

総務省および経産省によるAIエージェントの定義追加と、人間の判断を必須とする設計原則の策定

2026年7月

AIセーフティに関する評価観点ガイド(第1.20版)

AIセーフティ・インスティテュート（AISI）が、自律型AIの「挙動の観測と制御」を含むエージェント評価基準を追加

【経営リスクの現実】
日本のAIガイドラインは「任意（ソフトロー）」だが、監査法人・顧客・取締役会は具体的な「エビデンス」の提示を容赦なく求めてくる

Sources: MIC/METI AI Guidelines for Business Ver1.2 (March 31, 2026); Japan AISI, Guide to Evaluation Perspectives on AI Safety Ver1.20 (July 7, 2026); IPA Press Release (January 29, 2026)

企業として具体的にどこから対策を始めるべきか？

まずは制御。可視化は後からついてくる。

01 自社における「特権リソース」の特定と定義

AIエージェントがターゲットとするデータベースや各種パイプライン、財務システム、顧客データを分類し、その領域に専用のポリシーを策定する。

02 制御が、真の可視化と監査証跡を生み出す

完璧な管理台帳（インベントリ）の完成を待つ必要はない。機密資産へのアクセス監視を始めることで、自動的にエージェントの把握・管理が実現される。

03 ログインではなく、アクションを認可

各エージェントのアクションを「誰が（アイデンティティ）」「何を対象に（リソース）」「どのような状況で（コンテキスト）」という観点で検証し、ポリシーを外れた瞬間にセッションを遮断する。

04 エージェントの権限・範囲をタスクに限定

単体テストを作成するコーディング用エージェントに、本番環境へのアクセス権を持たせる必要性は皆無。

05 エージェントに対する常時特権を廃止

アクセス許可はタスクの実行中のみ有効とする。単なるクレデンシャルの定期変更（ローテーション）では代替できない。

06 すべてのアイデンティティに対する基準を引き上げる

エージェントに限らず、人間やその他のマシンアイデンティティに対しても、同等のセキュリティ統制を適用する。

ありがとうございました



ユーザーではなく、タスク範囲の絞り込み

エージェントは非決定的（予測不能）、スコープを厳しく絞り込んだ独自のアイデンティティを与える。



特権は、一時的なものにする

ジャスト・イン・タイムで提供し、自動で失効させ、エージェント自身には保有させない。



あらゆる「アクション」を認可する

アクション（実行）の瞬間に判断を下し、なぜそれが許可されたのかを記録する。

Delinea

IIM

アンケートへのご協力を
お願いいたします

2026 Identity Security Report
ダウンロード

Delinea

IIM