INSIGHT PAPER

CISOのための実用的なAIセキュリティ戦略

CISOが業務運営のスピードを損なうことなく
AIのセキュリティ確保とガバナンスを実現するためには

protiviti®
Global Business Consulting

CISOのための実用的なAIセキュリティ戦略

CISOが業務運営のスピードを損なうことなくAIのセキュリティ確保とガバナンスを実現するためには

人工知能（AI）は、組織の働き方や競争の仕方、顧客へのサービス提供のあり方を変革しています。多くの企業は、業務へのAI導入を急ピッチで進め、生産性の向上と新たな能力の獲得に熱心に取り組んでいます。

経営陣は、最高情報責任者（CIO）に対し、障壁を取り除き、新しいAIソリューションやプロバイダーとできるだけ早く関われるようにするよう圧力をかけています。板挟みとなったCIOは、組織が適切に保護されていることを確保しつつ、業務のニーズとのバランスを取るのに苦労しています。

最高情報セキュリティ責任者（CISO）にとって、これは馴染みのあるものですが、より深刻化したジレンマです。ほとんどのCISOは、特にデータセキュリティ、プライバシー、モデルの整合性、そしてプロンプトインジェクションやデータポイズニングといった新しい攻撃手法の分野において、AIが組織にもたらす重大なリスクと課題を理解しています。

これに対応して、CISOは組織を守るために強固なAIセキュリティ標準やプロトコルの導入を急いでいます。しかし、これらの手順を一律に適用すると、組織の動きが遅くなり、経営陣が望まない業務上の負担が生じます。その結果として、不満や摩擦、そしてチームがルールを回避しようとすることが予想されます。時には「シャドウAI」を導入し、これによってリスクが高まることさえ起こりえます。

答えは、急ブレーキを踏むことでも、見て見ぬふりをすることでもありません。答えは実用的なアプローチ、すわなち、リスクが低いところでは迅速さを実現し、リスクが高いところではより深いガバナンスを適用する、リスクベースのモデルです。

61%

のCFOおよび財務リーダーが、データセキュリティとプライバシーを来年の最優先事項と位置付けています

なぜ画一的なAI統制は裏目に出るのか

従来の統制は、比較的安定したIT環境と既知のベンダー環境を前提としていますが、AIはこれらの前提を覆します。AIシステムには、以下の特性があります。

- 機密データに触れる(トレーニング、調整、プロンプト、出力)
- 確率的に動作するため、説明性や公平性の考慮が必要となる
- 複雑なサプライチェーンに依存する(モデル、重み、データセット、ベクトルストア、プラグイン)
- 急速に変化し、静的な評価は陳腐化する

すべてのAI実験に厳格な審査を課すと学習の速度が落ち、抜け道が生じます。一方、すべてを許可すると、許容できないリスクにさらされます。現実的な対応策は、その中間をうまく見極めることです。

プロティビティの「2025年グローバル・ファイナンス・トレンド・サーベイ」によると、ファイナンス部門におけるAIの利用率は、2024年の34%から2025年には72%へと急増しました。

リスクが低いAIのために
効率的かつ迅速な導入経路を確立する

CISOにとって、ビジネスリーダーが迅速かつ容易に関与できる低リスクのAIソリューションへの効率的な経路を構築することは有益です。重要な要素は以下の通りです。

1. **事前審査済みのプロバイダーとプラットフォーム**:信頼性の高い企業が提供する、より広く知られたAIソリューションに焦点を当てる。これらの企業は通常、堅牢なセキュリティ保護、文書化された統制、エンタープライズ契約(例:SSO、データレジデンシーオプション、SOC 2/ISO認証)を採用している。

2. **低リスクのユースケース**:迅速な手続きルートの利用は、重要な意思決定を要するものを避け、機密情報の取り扱いを最小限に抑えるユースケースに限定する。例としては、公開コンテンツの要約、個人識別情報(PII)を含まないマーケティングのアイデア出し、セキュアなオフィス環境内での業務効率化ツールなどが挙げられる。

3. **データのガードレール**：機密情報が使用される場合、その情報は保護された環境内で実行されなければならない（例：プライベートテナント、転送中および保存時の暗号化、厳格なアクセス制御、データ損失防止、プロンプト／応答の最小化、強力なログ記録）。

このアプローチは、スピードを維持しつつ適切な保護を確保し、さらに重要な点として、セキュリティを回避する必要性を感じさせないよう、ビジネスに正当かつ認可された経路を提供します。

リスクを業務運用に組み込む：AIのグリーン／イエロー／レッドゾーン

明確で使いやすいゾーニングモデルは、全員が期待を理解し、意思決定を迅速化するのに役立ちます。

グリーンゾーン — 簡易手続きの利用可

- **ユースケース**：リスクが低い業務に限定される。人事決定、採用、信用決定、医療判断、その他規制対象、またはバイアスの高い状況には使用しない。
- **データプロファイル**：公開情報または機微性の低いデータのみを利用する。規制対象データ（例：保護された健康情報、決済カード業界データ、特別カテゴリーの個人データなど）は含まない。
- **プロバイダー**：企業向けの統制と標準契約条件を備えた、信頼できる認定ベンダーを利用する。

イエローゾーン — デューディリジェンスが必要

- **ユースケース**：中程度のリスクを伴う業務（例：顧客サポートアシスタントや、機密性が高いが厳格な規制対象ではないデータに関わる内部分析など）。
- **データプロファイル**：機密データの使用は許容されるが、ガードレール（データのマスキング、役割ベースのアクセス制御、テナントの分離）が必要。
- **プロバイダー**：知名度の低いベンダー、または既知のベンダーが提供する新機能を利用する場合。セキュリティおよびプライバシーに関する質問票、契約の付帯条項、限定的なパイロットテストが必要となる。

プロティビティの第2回生成AIに関する調査によると、昨年1年間で本番稼働中のAIプロジェクトの数がほぼ2倍に増加したことが明らかになりました。

レッドゾーン — 厳格な審査

- **ユースケース**：リスクが高いまたは影響が大きい業務（例：人事ソリューション、採用決定、融資や保険の引受、重要な安全性に関する判断、法的判断、またはバイアスや説明責任が規制上の懸念となる活動）。
- **データプロファイル**：機密性の高いデータや規制対象のデータを含む、個人に重大な影響を与える決定を行うシステム。
- **プロバイダー**：未知または実績のないベンダー、モデルリスク管理が必要なカスタムモデル。

このシンプルな分類法は、明確性をもたらします。これにより、過度なリスクを発生させないAIユースケースについては迅速な対応を可能にしつつ、高リスクのケースに対しては堅牢なガバナンス体制を維持することができます。また、セキュリティ部門と事業部門の間で共通言語を確立する効果もあります。

プロティビティと内部監査人協会（IIA）が共同で実施した「第12回 グローバル・テクノロジー・リスクに関する調査結果」によると、IT監査責任者の59%が、AIを今後2～3年の重要な脅威と見なしています。

モデルを反復可能な運用メカニズムに変える

実用的なアプローチは、実行しやすい場合にのみ機能します。あなたのゾーンを実現するために、以下の基本要素を考慮してください。

1. **AI導入およびユースケースの棚卸し**
   目的、データの種類、利用者、プロバイダー情報を記録するための簡易なセルフサービス型導入フォームを用意する。シンプルなルールを使って自動的にグリーン・イエロー・レッドに分類する。シャドウAIの排除のため、常に最新のインベントリを維持する。

2. **利用方針とクイックスタート・プレイブック**
   プレイブックには簡潔な行動指針を提供する（例：「プロンプトに秘密情報を含めない」、「承認済みのコネクタのみを使用する」）。各ゾーンに1ページのプレイブックを用意し、チームが実施すべきこと、セキュリティ部門 が提供する内容、標準的な対応時間などを明記する。

3. **事前審査済みベンダーカタログ**
   承認済みソリューションのカタログを保持し、デフォルト設定、セキュリティ設定、データ保持期間の初期値、契約条項などを記載する。イエローやレッドの場合は、ベンダーのデューディリジェンス用テンプレートやテストチェックリストも含める。

4. **機密データのための保護された環境**
   機密情報が関係する場合は、管理された実行環境を必須とする。具体的には、プライベートな大規模言語モデルのエンドポイント、暗号化されたベクトルストア、集中管理された鍵、マスキング／トークン化パイプライン、厳格なデータ流出制御などが求められる。

5. **重要な意思決定における人の関与（Human in the Loop）**
   レッドゾーンに該当する意思決定については、人によるレビューと明確な責任の所在を義務付ける。また、上書きや異議申立て、救済措置の手順を文書化する。

プロティビティの「2025年グローバル・ファイナンス・トレンド・サーベイレポート」によると、世界的な課題への対応に高い自信を持っていると回答したファイナンスリーダーは41％にとどまっており、強固なガバナンスとリスク管理フレームワークの必要性が浮き彫りになっています。

6. **継続的なモニタリングとレッドチーム活動**
   利用状況の監視：プライバシー制御を施した上でプロンプトや応答をログに記録し、データフローを追跡、不正利用を制限し、異常を検知した場合はアラートを発する。リスクの高いシステムに対しては、攻撃者視点でのテスト(アドバーサリアルテスト)、バイアス監査、定期的な再認証を計画的に実施する。

7. **トレーニングとカルチャー**
   役割に応じたトレーニングを提供し、プロダクトオーナー、開発者、アナリスト、責任者がAIの能力と限界、そして各ゾーンにおける自らの責任を理解できるようにする。

重要な指標

パートナーシップのマインドセットを強化するため、成功をスピードと安全の両面から測定します。

- ゾーンごとの承認所要時間(グリーンゾーンは数時間、イエローゾーンは数日を目標とする。レッドゾーンはサービスレベル合意書を定義する)

- リスクの高いシステムのコントロール適用範囲(目標：必要なコントロールの100%)

- 実現されたビジネス価値(目標：安全な導入に紐づく、効率化や収益増の定量的成果)

セキュリティが「正しいこと」を加速させ「間違ったこと」を遅らせているだけであると事業部門が認識すれば、信頼が高まります。

リスク許容度と規制への適合

ゾーニングモデルは、組織のリスク許容度と規制環境を反映する必要があります。例えば、以下の通りです。

- 規制の厳しいセクター(医療、金融サービス、公共部門)では、レッドゾーンの承認基準を厳しく設定し、初期設定としてより多くのユースケースをイエローゾーンとして扱う場合がある。

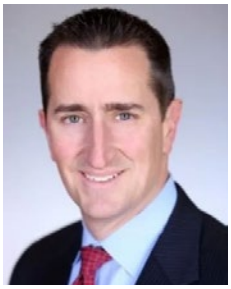
- 多国籍企業は、データの所在、国境を越えた転送、および現地のAIやアルゴリズムの説明責任に関する要件について、管轄区域を意識した初期設定を必要としている。

- すでにNIST CSF（ガバナンス機能を含む）、NIST AI RMF、ISO/IEC 42001、または既存のモデルリスク管理の実践などのフレームワークを使用している場合は、ゾーンとコントロールをそれらの成果物にマッピングすることで、不要な作業を避けることができる。

ビジネスパートナーとしてのCISO

AIという新たな領域において、ビジネスとセキュリティが共に前進するためには、ビジネスのパートナーとして認識されることが不可欠です。実用的なアプローチ—既知のプロバイダーやリスクの低いユースケースに対する迅速な手続きルート（グリーン、イエロー、レッドゾーン）、機密データの保護された環境、高リスク領域への厳格な監視—は、バランスの取れた信頼性の高い道筋を提供します。このアプローチは、迅速な展開というニーズを満たしつつ、組織のリスク許容度に合わせた保護を提供します。

著者について



Scott Laliberte（MBA、CISSP、CISA、CRISC、CISM）
マネージングディレクタ

Scott Laliberteは、プロティビティのテクノロジーコンサルティング部門のマネージングディレクタであり、取締役会および経営陣にとって信頼できるCISOアドバイザーとして活動しています。彼は、ビジネス戦略の文脈においてサイバーセキュリティリスクを理解し管理できるよう組織を支援し、企業価値を保護しつつイノベーションを促進することを専門としています。

深い技術的専門知識と幅広いビジネス知識を持つScottは、セキュリティを損なうことなく成長を促進するために、新しい技術と高度な方法論を活用する方法についてクライアントに助言しています。複雑なサイバー脅威を実行可能なビジネス洞察に変える彼の独自の能力は、デジタルトランスフォーメーションを進める組織にとって戦略的パートナーとしての地位を築いています。

Scottはこれまで、プロティビティのグローバル・サイバーセキュリティ部門、エマージング・テクノロジー部門、および攻撃・侵入テストラボを率い、技術革新、サイバー戦略、高度なセキュリティアーキテクチャの推進に取り組みました。彼の経験は、取締役会へのサイバーリスクガバナンスに関する助言から、規制および運用上の優先事項に沿ったレジリエントなセキュリティプログラムの構築支援に至るまで多岐にわたります。

プロティビティLLC　protiviti.jp
東京都千代田区大手町 2-6-4 TOKYO TORCH 常盤橋タワー24F
大阪府大阪市北区梅田 3-2-123 イノゲート大阪9F

protiviti®