Proofpoint AIモデルカード

Threat Protection Group (TPG) 検知システム

バージョン: 1.0 | 日付: 2026年6月

---

1. モデル概要

項目	内容
製品グループ	Threat Protection Group (TPG)
機能領域	電子メール脅威検知、コンテンツ分析、URLおよびウェブコンテンツ分析
モデル管理責任者	Proofpoint Threat Protection エンジニアリング部門責任者
経営層による監督	Chief AI & Data Officer、Assistant General Counsel

---

2. 意図された用途

主たる目的: マルウェア、ランサムウェア、フィッシング、ビジネスメール詐欺（BEC）およびスパムを含むサイバーセキュリティ上の脅威を、ProofpointのThreat Protection製品ポートフォリオ全体にわたって検知し、防止すること。

対象利用者: Proofpointの電子メールセキュリティ製品およびコラボレーションセキュリティ製品を利用する企業等のお客様。

適用範囲外の用途: Proofpointの検知システムは、サイバーセキュリティ上の脅威検知の用途に限定して設計および検証されています。本範囲を超える使用はサポートの対象外です。

---

3. モデルの機能

ProofpointのThreat Protection検知システムは、電子メールのメタデータ、メッセージ本文、ならびにURLおよびウェブコンテンツにわたるシグナルを分析し、サイバーセキュリティ上の脅威を特定し遮断します。各検知システムは相互に補完する機能領域にわたって稼働し、これらが一体となって多層的な保護を提供します。

---

4. 学習データ

Proofpointの検知モデルは、以下を組み合わせて学習されています。

- Proofpointの脅威インテリジェンスデータ
- ラベル付けされた悪意のあるメッセージおよびURLのサンプル
- URLおよびウェブコンテンツのデータ
- 送信者およびURLのレピュテーションを含む挙動シグナル

顧客データは、検知サービスの提供および改善のために必要な場合に限り、かつ適切な管理措置および承認が講じられている場合に限り、モデルの学習に使用されます。

---

5. 学習方法論

弊社独自の検知手法を保護するため、公開していません。

---

6. 評価

モデルは、導入に先立ち、ラベル付けされたデータセットに対する適合率・再現率分析を含む標準的な機械学習の評価方法論を用いて評価されます。モデルが本番環境へ昇格されるのは、定められた品質基準に照らして現行モデルを明確に上回る性能を示した場合に限られます。

---

7. 検証方法論

すべてのモデルは、本番リリース前にシャドーデプロイメントによる検証プロセスを経ます。

- **シャドーデプロイメント:** 新規モデルは、実際の処理結果に影響を与えることなく、現行の本番モデルと並行して実運用トラフィックを処理し、実環境下での性能の直接比較を可能にします。
- **昇格基準:** モデルが本番環境へ昇格されるのは、現行モデルおよび既知のラベル付きデータセットに照らして評価される所定の性能基準を満たした場合に限られます。
- **ロールバック機能:** モデルのバージョン管理はすべての検知システムにおいて維持されており、必要に応じて過去のバージョンへロールバックすることが可能です。
- **変更管理:** 新しいモデルバージョンは、既存モデルと並行して導入されます。従前のバージョンが廃止されるのは、新バージョンが継続的に安定した性能を示した後に限られます。

---

8. バイアスおよび公平性

Proofpointの[ISO 42001認証]の対象範囲に含まれます。

---

9. 安全性およびリスク評価

Proofpointの[ISO 42001認証]の対象範囲に含まれます。

---

10. ガバナンスおよび監督

ガバナンス体制

Proofpointは、モデルリスクに関する意思決定の責任の所在を明確にしたAIガバナンス体制を整備しています。モデルリスクに関する権限は、各製品部門において指名された技術リードおよびエンジニアリング責任者に付与され、Chief AI & Data OfficerおよびAssistant General

Counselによる経営層の監督が行われます。

規制との整合

ProofpointのAIガバナンスプログラムは、以下の主要なフレームワークを踏まえて策定されています。

フレームワーク	ステータス
ISO/IEC 42001:2023 — AIマネジメントシステム	https://www.proofpoint.com/us/legal/trust/product-certifications

フレームワーク	ステータス
NIST AI Risk Management Framework (AI RMF)	ガバナンスプログラムに反映
SR 11-7 — モデルリスク管理に関する監督指針	ガバナンスプログラムに反映
EU AI Act	Proofpointの利用は低リスクに該当

---

11. 継続的モニタリング

モデルの性能は、以下を通じて継続的に監視されています。

- 大規模環境における有効性の維持を確認するための処理量の追跡
- アナリストが確認した誤検知および検知漏れに照らしたモデル出力の継続的評価
- 変化する脅威環境に対応するための、本番環境からのフィードバックを取り入れた再学習サイクル

継続的改善の取り組み

Proofpointは、以下を含むAIガバナンス実務の成熟度向上に積極的に投資しています。

- **ISO 42001認証** — 包括的なAIマネジメントシステムの導入
- **データ品質評価の自動化** — 学習データに関するオブザーバビリティ指標および異常検知
- **導入前検証の自動化** — 2026年に構築予定のパイプライン
- **データリネージおよび来歴管理** — データの利用状況および来歴の自動追跡

---

12. データの取扱い

処理されるデータ要素

Proofpointの検知システムは、脅威検知サービスの提供に必要な範囲で顧客データを処理します。処理されるデータ要素には以下が含まれます。

- 電子メールのメタデータ（送信者、受信者、件名）
- メッセージ本文（脅威、詐欺、スパムの兆候）
- URLおよび関連するウェブコンテンツ

- 挙動に関する指標（送信者およびURLのレピュテーション、配信シグナル）
- 脅威インテリジェンスデータ

使用される具体的なデータ要素は、検知機能ごとに異なります。データ処理は、Proofpointのセキュリティが確保されたクラウドインフラストラクチャ内で行われます。

顧客データとモデル学習

顧客データは、必要な場合に限り、かつ適切な管理措置の下でモデルの学習に使用されます。

データの保持および削除

- 保持ポリシーは、スケジュールされた処理により自動的に適用されます。
- データの削除は、定められた保持期間に従って体系的に実行されます。

データアクセス制御

AIパイプライン内のデータへのアクセスは、パイプライン、エンジニアリングおよびデータサイエンスの各役割を厳格に分離したロールベースアクセス制御により管理されています。すべての特権アクセスは監査およびログ記録の対象となります。

暗号化

データは、定められた周期で更新される専用の暗号鍵を用いて、保存時に暗号化されます。

顧客データの分離

すべての顧客データは、顧客テナントIDに紐づけて保管されており、これによりデータの管理、削除および将来的な分離対応が容易になります。

---

13. サードパーティおよびクラウドAIサービス

ProofpointのAIシステムは、マネージドクラウド環境（AWS、Azure、GCP）上で稼働します。Proofpointがクラウドホスト型のAI/MLサービスを利用する場合、当該サービスは以下の管理措置の下、Proofpointが管理するクラウドテナント内で稼働します。

- クラウドホスト型のAI/MLサービスを通じて処理されるデータは、Proofpointが管理するクラウド環境内にとどまります。
- クラウド事業者との契約は、当該事業者のモデル学習のために顧客データを使用することを禁止しています。
- クラウドホスト型のAI/MLサービスは、その他のクラウドインフラストラクチャと同一のセキュリティ管理策、アクセス管理およびデータ取扱い条件に服します。
- Proofpointは、クラウドインフラストラクチャ事業者の利用について開示し、当該環境におけるデータセキュリティに関する確約を行います。

オープンソースのモデルアーキテクチャを利用する場合、モデルの重みはProofpoint自身のインフラストラクチャ内で保管および管理されます。

---

14. 人間による監督

- モデルの昇格に関する判断は、すべての検知システムにおいて人間による審査および承認を要します。
- アナリストチームがモデルの出力を継続的に確認し、誤検知および検知漏れのシグナルを生成します。これらは直接モデルの評価および再学習に反映されます。
- 該当する場合には、学習データの品質管理プロセスに人間のラベル付与担当者が関与します。そのアクセス権限はラベル付与作業に限定されます。

---

15. バージョン管理およびリネージ

- モデルのバージョン管理は、すべての検知システムにおいて維持されています。
- 過去のモデルバージョンへのロールバックに対応しています。
- アーキテクチャ仕様、導入済みモデルのメタデータおよびバージョン履歴を含む社内モデル文書は、Proofpointの社内システムにおいて維持管理されています。
- 正式なデータリネージ文書は、ProofpointのISO 42001認証の一環として現在整備中です。

---

16. 継続的改善

Proofpointは、AIガバナンス実務の継続的な改善に取り組んでいます。現在の投資領域には、データ品質評価の自動化、導入前検証の自動化、ならびにデータリネージおよび来歴の追跡が含まれます。

本文書は、2026年6月時点におけるProofpointのAIデータガバナンス実務を記載したものです。Proofpointは継続的な改善に取り組んでおり、実務の進展に応じて本文書を更新します。