TLP:CLEAR

対策から組織戦略へ

脅威の変化に応じた持続可能なセキュリティ体制の構築

2024年 12月

名和 利男

目 次

TLP:CLEAR

セッション 1

従来型のサイバーセキュリティ対策の再評価

従来型のサイバーセキュリティ対策: 概要

特徴	説明	基本機能	長所	短所	課題
ファイアウォール	ネットワークの境界に設置され、内部ネットワークと外部ネットワーク間のトラフィックを制御する装置やソフトウェア	• パケットフィルタリング • ステートフル検査 • アプリケーションレベルゲートウェイ	• 不要なトラフィックのブロック • ネットワークセグメンテーション • ログ記録と監視	• 内部脅威への対応が限定的 • 複雑な設定が必要 • 暗号化トラフィックの検査が困難	• 高度な偽装トラフィックの検出 • 動的に変化する脅威への適応
アンチウイルスソフトウェア	コンピュータシステムやネットワーク上のマルウェアを検出、防止、除去するためのソフトウェア	• シグネチャベースの検出 • ヒューリスティック分析 • リアルタイムスキャン	• 既知のマルウェアに対する高い検出率 • システムの自動保護 • 定期的な更新	• 新種や変種のマルウェアへの対応遅延 • シグネチャのオーバーヘッド • 誤検知の可能性	• 迅速な変異への対応 • 振る舞いベースの高度な検出能力の必要性
侵入検知システム（IDS）	ネットワークやシステムへの不正アクセスや異常な活動を監視し、検知するシステム	• シグネチャベースの検知 • 異常検知 • ログ分析	• リアルタイムの脅威検知 • 広範囲のネットワーク可視性 • フォレンジック分析のサポート	• 誤検知と見落としの可能性 • 大量のアラート生成 • パフォーマンスへの影響	• 新しい攻撃パターンの検出 • 大量のデータからの効率的な異常検知
侵入防止システム（IPS）	IDSの機能に加えて、検知された脅威に対して自動的に防御アクションを実行するシステム	• トラフィックのブロックや隔離 • シグネチャおよび異常ベースの検知 • 自動対応	• リアルタイムの脅威緩和 • ネットワークパフォーマンスの最適化 • カスタマイズ可能なルール	• 誤検知による正常なトラフィックのブロック • 複雑な設定と維持 • ゼロデイ攻撃への限定的な対応	• 高度な判断能力の統合 • 動的な防御ルールの自動生成と最適化

TLP:CLEAR

従来型のサイバーセキュリティ対策: 課題

- **統合と相互運用性の重要性**
  - 個別に機能するだけでなく、相互に連携することで効果を発揮するものである。
- **人間の専門知識との融合**
  - 効果は、それを運用する人間の専門知識に大きく依存しているものである。
- **コンテキスト(状況や背景)認識の必要性**
  - しばしばコンテキスト(状況や背景)を考慮せずに動作するものである。
- **適応性と学習能力の向上**
  - 従来型防御策は、主に事前定義されたルールやシグネチャに基づいて動作するものである。
- **プライバシーとセキュリティのバランス**
  - 高度な防御には、より多くのデータ分析が必要となるが、同時にプライバシー保護の要求も高まっている。

セッション 2

AI駆動型サイバー攻撃の進展

AI駆動型サイバー攻撃の進展

特徴	説明	洞察	影響
高度な自動化	AIは大量のデータを処理し、パターンを認識し、迅速に意思決定を行うことができる。	攻撃者は人間の能力をはるかに超える規模と速度で攻撃を実行できる。 例えば、数百万のパスワードを数分で試行するなど、従来は不可能であった攻撃が可能となる。	防御側は、同様に高度に自動化された防御システムを開発する必要がある。 人間の反応速度では対応が難しい状況が増えている。
人間の行動パターンの模倣	AIは大量のデータから学習し、人間の行動を精密に模倣することができる。	ボットやマルウェアが正常なユーザー活動として偽装することが可能である。 例えば、メールの文体や送信パターンを模倣し、フィッシング攻撃の成功率を大幅に向上させることができる。	従来の異常検知システムでは、このような高度な偽装を見破ることが困難であり、行動分析の精度を高める必要がある。
迅速な適応と進化	AIシステムは、新しい情報や環境の変化に基づき、リアルタイムで学習し適応することができる。	攻撃AIは、防御策の変更や新しいセキュリティパッチに対して、驚くべき速さで適応し、新たな攻撃ベクトルを見つけ出す可能性がある。	従来の静的な防御策では不十分であり、同様に適応能力の高い動的な防御システムが必要である。

TLP:CLEAR

AI駆動型攻撃: 従来型攻撃との違い

特徴	説明	洞察	影響
規模と速度	AI駆動型攻撃は、従来型攻撃に比べて、はるかに大規模かつ高速に実行される。	例えば、DDoS攻撃の場合、AIは正規のトラフィックパターンを模倣しつつ、膨大な数のボットを制御して攻撃を行うことができる。	従来の手動による監視やインシデント対応プロセスでは、このような攻撃に対処することが極めて困難になっている。
複雑性と洗練度	AI駆動型攻撃は、複数の攻撃手法を組み合わせ、状況に応じて戦略を動的に変更することができる。	例えば、初期侵入にソーシャルエンジニアリングを使用し、その後のラテラルムーブメントでは学習型のマルウェアを使用するなど、複合的な攻撃が可能となる。	防御側は、単一の防御策に頼るのではなく、総合的かつ適応型のセキュリティアプローチを採用する必要がある。
予測困難性	AI駆動型攻撃は、人間の思考パターンにとらわれない新たな攻撃手法を生み出す可能性がある。	AIは、人間が考えつかないような方法でシステムの脆弱性を発見し悪用する可能性がある。 例えば、複数の無害な操作を組み合わせて予期せぬ結果を引き起こすことが考えられる。	セキュリティ専門家は、従来の経験や直感だけでなく、AIの支援を受けながら、常に新しい視点でシステムを評価する必要がある。

TLP:CLEAR

AI駆動型攻撃: 特性

特徴	説明	具体例	影響	対策の方向性
機械学習を活用した攻撃手法	AIモデルを用いて、攻撃パターンを最適化し、防御システムを回避する	• 敵対的機械学習を用いたマルウェア生成 • 自動化された脆弱性探索 • 動的に変化するボットネット制御	• 従来の署名ベース検出の無効化 • ゼロデイ攻撃の増加 • 攻撃の成功率向上	• AIベースの異常検知システムの導入 • 動的解析技術の強化 • 継続的な機械学習モデルの更新
自然言語処理による高度なフィッシング	AIを用いて人間らしい文章を生成し、より説得力のあるフィッシング攻撃を実行	• ターゲットの文体を模倣したスピアフィッシングメール • リアルタイムで適応するチャットボット型フィッシング • ディープフェイクを活用した音声フィッシング	• フィッシング攻撃の検出困難化 • ソーシャルエンジニアリングの成功率向上 • 大規模かつ個別化された攻撃の実現	• AIベースの文章分析と異常検知 • マルチファクタ認証の強化 • 従業員のセキュリティ意識向上訓練
自動化された脆弱性スキャンと攻撃	AIを用いて大規模かつ高速に脆弱性をスキャンし、自動的に攻撃を実行	• 自己進化型のワーム • クラウドリソースを活用した分散型スキャン • コンテキストに基づく適応型エクスプロイト	• 攻撃の規模と速度の劇的な増加 • 人間の対応能力の限界超過 • 複雑な攻撃チェーンの自動化	• AIを活用した自動パッチ管理 • 動的な脆弱性評価と優先順位付け • 自己修復型システムの開発

TLP:CLEAR

AI駆動型攻撃: 洞察と考察 (1/2)

- **AIの二面性**
  - AIは攻撃者と防御者の両方に利用可能な技術であり、いわば「両刃の剣」である。
  - 防御側もAIを積極的に活用し、攻撃AIとの「AI対AI」の競争に備える必要がある。

- **攻撃の個別化と大規模化の同時進行**
  - AI駆動型攻撃は、個々のターゲットに合わせた高度にカスタマイズされた攻撃を、大規模に実行することが可能である。
  - マスカスタマイゼーションされた攻撃が現実のものとなり、従来の防御戦略の再考が必要となる。

次のスライドで解説

- **人間の認知限界の超越**
  - AI駆動型攻撃は、人間の認知や反応速度をはるかに超えるスピードで進行する可能性がある。
  - 防御側も同様にAIを活用し、人間の監視者をサポートするAI補助型セキュリティオペレーションの構築が重要となる。

次のスライドで解説

TLP:CLEAR

【用語】「AI駆動型攻撃」関連 (1/2)

マスカスタマイゼーションされた攻撃

AIや高度な技術を用いて、個別のターゲットに合わせたカスタマイズされた攻撃を、大規模に実行する手法である。

これは、従来の攻撃手法と異なり、ターゲットの特性や行動パターンに基づいて個別に攻撃内容を最適化しつつ、同時に多くのターゲットに対して攻撃を展開することが可能である。

この結果、攻撃の成功率が高まり、従来の画一的な防御策では対応が難しくなる。

AI補助型セキュリティオペレーション

AIを活用して人間の監視者やセキュリティチームをサポートし、脅威の検出、分析、対応をより迅速かつ効果的に行うセキュリティ運用体制である。

AIは膨大なデータをリアルタイムで処理し、異常を検出することで、人間の判断を補完し、潜在的な脅威への迅速な対応を可能にする。

このアプローチにより、人間の負担が軽減され、より高度な脅威に対する防御力が向上する。

AI駆動型攻撃: 洞察と考察 (2/2)

- **攻撃の予測困難性の増大**
  - 機械学習モデルの「ブラックボックス」的な性質により、AI駆動型攻撃の次の動きを予測することが極めて困難である。
  - 事後対応型から**予防型・適応型のセキュリティアプローチ**への移行が必要となる。

次のスライドで解説

- **エコシステムとしての防御の重要性**
  - 単一の防御策ではAI駆動型攻撃に対抗することは困難である。
  - 複数の防御層が協調して動作し、**相互に学習・適応するエコシステム**としてのセキュリティ体制の構築が求められる。

次のスライドで解説

- **倫理的な考慮事項**
  - AI駆動型攻撃に対抗するためのAI防御システムの開発と運用には、プライバシーや公平性に関する倫理的な考慮が不可欠である。
  - 防御AIの判断の**透明性と説明可能性を確保**することが、社会的信頼を得る上で重要となる。

TLP:CLEAR

【用語】「AI駆動型攻撃」関連 (2/2)

予防型・適応型のセキュリティアプローチ

攻撃が発生する前に脅威を予測し、防御策を事前に講じる「予防型」の手法と、環境や脅威の変化に応じて防御策を動的に調整する「適応型」の手法を組み合わせたセキュリティ戦略である。

このアプローチにより、従来の静的な防御策に比べ、未知の脅威や進化する攻撃に対して柔軟かつ効果的に対応できるようになる。

相互に学習・適応するエコシステム

A複数のセキュリティ防御システムやツールが連携し、互いに情報を共有しながら学習し、環境の変化や新たな脅威に対応して適応するセキュリティ体制である。

このエコシステムでは、各防御層が単独で機能するのではなく、協調的に動作することで、より効果的で包括的な防御を実現する。

結果として、システム全体のセキュリティが継続的に強化される。

TLP:CLEAR

現実的な脅威になりつつAI駆動型攻撃

AI駆動型攻撃1. ディープフェイクとソーシャルエンジニアリング

AI駆動型攻撃2. 自律型マルウェア

AI駆動型攻撃3. AI駆動型DDoS攻撃

AI駆動型攻撃4. 機械学習モデルへの攻撃

AI駆動型攻撃5. ゼロデイ脆弱性の自動探索

AI駆動型攻撃1. ディープフェイクとソーシャルエンジニアリング

🎭 キーポイント

- AIによる高度な音声・映像偽造がソーシャルエンジニアリング攻撃を劇的に強化

💡 重要な洞察

- **認証システム**の根本的な再設計が必要
- 「見ること」「聞くこと」への**信頼性**が揺らぐ社会への移行
- **法的・倫理的枠組み**の整備が急務
- **メディアリテラシー**教育の重要性が増大

AI駆動型攻撃2. 自律型マルウェア

キーポイント

- AIが環境に適応し進化するマルウェアの出現

重要な洞察

- 従来の署名ベース検出の完全な**陳腐化**
- 「**AI vs AI**」の新たなセキュリティ競争の始まり
- **エコシステム**としてのセキュリティ対策の必要性
- マルウェアの進化速度が人間の対応能力を**超越**

AI駆動型攻撃3. AI駆動型DDoS攻撃

キーポイント

- AIによる高度に洗練されたDDoS攻撃の出現

重要な洞察

- 正常トラフィックと攻撃の区別が極めて**困難**に
- インフラの**キャパシティプランニング**の考え方の変革が必要
- AIを活用した**動的リソース割り当て**の重要性
- 国家レベルのサイバー攻撃との境界が**曖昧化**

AI駆動型攻撃4. 機械学習モデルへの攻撃

🎯 キーポイント

- AIシステム自体を標的とした新たな攻撃手法の登場

💡 重要な洞察

- **AIセキュリティ**という新たな専門分野の確立
- モデルの堅牢性と精度の**トレードオフ**が課題に
- **データポイズニング**による長期的・潜在的リスクの増大
- AIの**信頼性担保**が社会実装の鍵に

TLP:CLEAR

AI駆動型攻撃5. ゼロデイ脆弱性の自動探索

🕵️ キーポイント

- AIによる未知の脆弱性の高速・自動発見

💡 重要な洞察

- ソフトウェア開発プロセスの**抜本的見直**しが必要
- 「パッチ前提」のセキュリティモデルの**限界**
- AIを活用した**自動バグ修正技術**の重要性増大
- **脆弱性市場**のダイナミクスが大きく変化する可能性

【用語】「自動バグ修正技術」関連 (1/2)

🔍 自動バグ修正技術

AIと機械学習を活用して、ソフトウェアのバグを自動的に検出し修正するプロセス。人間の開発者の作業を補完し、セキュリティ脆弱性への迅速な対応を可能にする。

🧠 機械学習モデル

過去のバグ修正パターンから学習し、新たなバグに対する適切な修正方法を予測・提案するAIシステム。

🔬 静的解析

プログラムを実行せずにソースコードを分析し、潜在的なバグや脆弱性を検出する技術。自動バグ修正の初期段階で使用される。

🏃 動的解析

プログラムを実際に実行し、その挙動を観察することでバグを検出する技術。より複雑なバグの発見に有効。

🧬 遺伝的プログラミング

進化のプロセスを模倣し、多数の潜在的な修正案を生成・評価・改良することで最適な修正を見つけ出す手法。

🔄 パッチ生成

検出されたバグを修正するためのコード変更(パッチ)を自動的に作成するプロセス。

TLP:CLEAR

【用語】「自動バグ修正技術」関連 (2/2)

✅ 回帰テスト

修正後のコードが正しく動作し、新たな問題を引き起こしていないことを確認するための自動化されたテストプロセス。

セマンティックパッチ

特定のバグパターンに対する一般化された修正ルール。複数のプロジェクトや類似のバグに適用可能。

ヒューマンインザループ

自動修正プロセスに人間の開発者が介入し、提案された修正を検証・調整する協調的アプローチ。

プログラム合成

仕様や例から自動的にプログラムコードを生成する技術。バグ修正において、正しい動作を満たす新しいコード断片の生成に使用される。

自動バグ修正技術は、AI駆動型攻撃に対する防御の重要な要素である。

この技術により、脆弱性の修正速度が向上し、攻撃者との「時間的競争」において優位性を確保できる可能性がある。

しかし、複雑なバグや意図を要する修正には依然として人間の監督が必要であり、AI技術と人間の専門知識を組み合わせたハイブリッドアプローチが最も効果的である。

TLP:CLEAR

【用語】「脆弱性市場」関連 (1/3)

脆弱性闇市場

未公開のソフトウェア脆弱性情報が非公式かつ違法に取引される市場。サイバー攻撃能力の売買の場となっている。

ゼロデイ脆弱性

開発者に知られておらず、パッチが存在しない脆弱性。攻撃者に悪用される可能性が高く、闇市場で高値で取引される。

エクスプロイト

脆弱性を悪用してシステムに不正アクセスするためのコードやツール。闇市場では完成したエクスプロイトが取引されることもある。

ダークウェブ

通常のインターネット検索エンジンでは見つからない、匿名性の高いネットワーク。脆弱性闇市場の多くはここで運営される。

脆弱性ブローカー

脆弱性の売り手と買い手を仲介する個人や組織。市場価格の形成に重要な役割を果たす。

バグバウンティプログラム

企業が脆弱性発見者に報奨金を支払う合法的なプログラム。闇市場への流出を防ぐ目的もある。

TLP:CLEAR

【用語】「脆弱性市場」関連 (2/3)

🔍 脆弱性スキャナー

システムの脆弱性を自動的に探索するツール。AI技術の進歩により、その能力が飛躍的に向上している。

ペイロード

エクスプロイトを通じてターゲットシステムで実行される悪意のあるコード。闇市場では高度なペイロードが取引される。

ホワイトハット

倫理的なハッカーやセキュリティ研究者。脆弱性を発見しても責任を持って開示する。

グレーハット

倫理的な境界線上で活動するハッカー。状況に応じて脆弱性を闇市場で売却するか、開発者に報告するかを決める。

Cyber Offense

攻撃的なサイバー能力。政府機関が国家安全保障目的で脆弱性闇市場を利用することがある。

Cyber Defense

防御的なサイバー能力。脆弱性情報を収集し、自国のシステムを防御するために使用する。

TLP:CLEAR

【用語】「脆弱性市場」関連 (3/3)

脆弱性闇市場は、サイバーセキュリティのエコシステムに大きな影響を与えている。

AI技術の発展により、脆弱性の発見と取引のダイナミクス（力学）が変化しつつある。

防御側は、AIを活用した脆弱性の早期発見と修正、そして倫理的なバグバウンティプログラムの強化を通じて、この市場に対抗する必要がある。

同時に、サイバーセキュリティの倫理と法的枠組みの再考が求められる。

セッション 3

（急激に変化する）脅威に適応した組織戦略の必要性

【参考】「組織戦略」と「組織的戦略」の違い

組織戦略

- 定義
  - 組織全体の長期的な目標や方向性を定める戦略。組織のビジョンやミッションを実現するために、どのような方針をとるべきかを規定する。
- 主な特徴
  - 組織全体を対象とし、包括的である
  - 経営層や戦略的意思決定者が策定する
  - 例: 新しい市場への進出、デジタルトランスフォーメーションの推進
- 例文
  - 「我が社の**組織戦略**では、5年以内にアジア市場でのシェアを倍増することを目指している。」

組織的戦略

- 定義
  - 戦略を実施する過程で、組織的(計画的かつ体系的)に行われる具体的な行動や手段。
- 主な特徴
  - 戦略の実行プロセスや方法論に焦点を当てる
  - 部門やプロジェクト単位での取り組みに適用される場合が多い
  - 例: 業務プロセスの標準化、リソースの最適配分。
- 例文
  - 「この施策を成功させるには、**組織的戦略**が欠かせない。」

(急激に変化する)脅威に適応した組織戦略

組織戦略1. 機械学習を活用したセキュリティ強化
組織戦略2. ゼロトラストアーキテクチャの重要性
組織戦略3. 多層防御戦略の再構築
組織戦略4. 脅威ハンティングの進化
組織戦略5. インシデントレスポンスの高度化
組織戦略6. セキュリティオーケストレーションの重要性
組織戦略7. 人材育成とスキル開発
組織戦略8. 法的・倫理的考慮事項
組織戦略9. クラウドセキュリティの課題
組織戦略10. クラウドセキュリティの課題
組織戦略11. 量子コンピューティングの影響

組織戦略1. 機械学習を活用したセキュリティ強化

🧠 キーポイント

- AIによる異常検知と脅威インテリジェンスの高度化

💡 重要な洞察

- 未知の攻撃パターンの**早期発見**が可能に
- 誤検知率の大幅な**低減**と運用効率の**向上**
- **リアルタイム**での脅威対応能力の強化
- 人間の専門家とAIの効果的な**役割分担**が鍵

次のスライドで例示

TLP:CLEAR

【例示】人間の専門家とAIの効果的な役割分担

データ分析とパターン認識：AI

AIは大量データの迅速な分析や異常検出に優れている。リアルタイムのネットワーク監視などはAIが担当すべきである。

学習と改善：AIと人間

AIは学習し進化するが、その方向性や最適化には人間の介入が不可欠である。

戦略的判断と意思決定：人間

AIの分析結果を基に、リスク管理や防御戦略を決定するのは人間の役割である。

創造的な問題解決：人間

未知の問題や新たな脅威への対応は、人間の創造力と柔軟な思考が必要である。

反復的タスクの自動化：AI

定期的なシステムスキャンやセキュリティチェックはAIに任せ、人間は高度な分析に集中する。

リスク管理と倫理判断：人間

プライバシーや法的な考慮を含む総合的な判断は、人間が行うべきである。

組織戦略2. ゼロトラストアーキテクチャの重要性

キーポイント

- 「信頼しない、常に検証する」原則に基づくセキュリティモデル

重要な洞察

- 境界防御の概念が完全に**再定義**される
- 継続的な認証と承認プロセスの**自動化**が不可欠
- マイクロセグメンテーションによる被害の**局所化**
- AIによる**動的**なアクセス制御の実現

組織戦略3. 多層防御戦略の再構築

キーポイント

- AIと従来型防御の統合による新たな多層防御モデル

重要な洞察

- 各防御層のインテリジェント化と自律的な**連携**
- 攻撃の文脈を理解した**適応型防御**の実現
- エンドポイントからクラウドまでの**一貫した保護**
- 防御の自動化と人間の意思決定の**最適なバランス**

組織戦略4. 脅威ハンティングの進化

🕵️ キーポイント

- AIを活用したプロアクティブな脅威探索と対応

💡 重要な洞察

- 潜在的脅威の事前発見による**被害の最小化**
- 人間の直感とAIの分析力の**相乗効果**
- 大規模データからの**微細**な異常パターンの検出
- 脅威インテリジェンスの質と速度の**飛躍的向上**

TLP:CLEAR

組織戦略5. インシデントレスポンスの高度化

⚡ キーポイント

- AIによる迅速かつ効果的なインシデント対応の自動化

💡 重要な洞察

- 意思決定の速度と精度の**大幅な向上**
- 複雑な攻撃シナリオへの**動的な対応**が可能に
- 人的リソースの効率的な配分と専門性の**集中**
- グローバルな脅威情報の**即時統合**と活用

組織戦略6. セキュリティオーケストレーションの重要性

キーポイント

- AIによる複数のセキュリティツールの統合管理と自動化

重要な洞察

- 複雑化するセキュリティスタックの効率的な運用
  次のスライドで解説
- 人的ミスの削減とレスポンス時間の**短縮**
- **データドリブン**な意思決定プロセスの確立
- セキュリティ運用の**スケーラビリティの向上**

【解説】複雑化するセキュリティスタックの効率的な運用

多様なセキュリティツールが増加し、ますます複雑になる中で、**相互運用性の確保**、**管理の効率化**、**運用コストの最適化**が課題となる。

組織は、これらのツールを**効果的に統合**し、**運用の一貫性**を維持しながら、セキュリティ対策の全体的な有効性を最大化する必要がある。

セキュリティスタックとは

組織のITインフラやデータを保護するために使用されるさまざまなセキュリティツールや技術の集まりを指す。特に、ファイアウォール、アンチウイルスソフトウェア、侵入検知システム（IDS）、データ暗号化ツール、アイデンティティ管理システム、セキュリティ情報およびイベント管理（SIEM）ツールなどが含まれる。

各ツールが異なる機能を持ち、それぞれが特定の脅威やリスクに対処する役割を果たしている。

組織戦略7. 人材育成とスキル開発

🚀 キーポイント

- AI時代のサイバーセキュリティ専門家に求められる新たなスキルセット

💡 重要な洞察

- AI・機械学習の基礎知識が全てのセキュリティ職種で**必須**に
- 技術スキルと戦略的思考力の**両立**が重要
- 継続的学習とスキルアップデートの**文化醸成**
- 倫理的ハッキングとAI倫理の**理解が不可欠**

組織戦略8. 法的・倫理的考慮事項

⚖ キーポイント

- AI活用セキュリティにおける法的・倫理的課題への対応

💡 重要な洞察

- AIの判断に基づく自動防御の法的責任の**所在が不明確**
- プライバシー保護とAIによる高度な監視の**バランスが課題**
- 国際的なAIガバナンスフレームワークの**必要性が増大**
- 説明可能なAI（XAI）の重要性がセキュリティ分野でも顕在化

次のスライドで解説

TLP:CLEAR

【解説】説明可能なAI（XAI: Explainable AI）

AIモデルやその決定プロセスが**人間に理解できる形で説明**されることを目指すアプローチや技術を指す。

従来のAI、特に深層学習などの高度なモデルは「**ブラックボックス**」と呼ばれ、その内部でどのように意思決定が行われているのかがわかりにくい。

しかし、説明可能なAIは、これらの複雑なモデルがどのように結論に達したのかを明らかにすることで、**透明性と信頼性**を向上させることを目的としている。

説明可能なAIは、特に医療、金融、法務など、**AIの決定が重大な影響を与える分野**で重要視されている。

組織戦略9. サプライチェーンセキュリティの再考

📎 キーポイント

- AI時代におけるサプライチェーンリスクの変容と対策

💡 重要な洞察

- AIによる自動化がサプライチェーン攻撃の**規模と複雑さを増大**
- サードパーティのAIモデルやデータセットの**信頼性検証**が新たな課題に
- ブロックチェーン技術とAIの統合による**トレーサビリティの向上**
- 「**信頼できるAI**」の概念がサプライチェーン選定の新基準となる可能性

TLP:CLEAR

【解説】信頼できるAI

安全性、**公平性**、**透明性**、**倫理的な基準**に基づいて設計され、運用されるAIシステムを指す。

信頼できるAIは、ユーザーや社会がその**結果や意思決定に対して安心感を持ち**、信頼を寄せることができるものである。

信頼できるAIの主な特徴は、以下のとおり。

- **公平性**: バイアスなく、公平な意思決定を行うことが求められる。
- **透明性**: AIのプロセスや結果が理解しやすく、説明可能であること。
- **安全性とセキュリティ**: サイバー攻撃から守られ、安全に動作すること。
- **プライバシーの保護**: ユーザーデータを適切に保護し、プライバシーを守る。
- **説明責任**: AIの意思決定に対して責任を持ち、結果を説明できること。
- **倫理的基準の遵守**: 倫理的基準に沿った設計と運用を行うこと。

組織戦略10. クラウドセキュリティの課題

♣ キーポイント

- AI時代のクラウド環境における新たなセキュリティパラダイム

💡 重要な洞察

- **AIワークロード**の保護が新たなセキュリティ領域として台頭
- マルチクラウド環境での**AI駆動型セキュリティオーケストレーション**が必須
- エッジコンピューティングの普及によるセキュリティ境界の**再定義**
- クラウドネイティブなAIセキュリティツールの**重要性が増大**

組織戦略11. 量子コンピューティングの影響

キーポイント

- 量子コンピューティングがもたらすAIサイバーセキュリティへの影響

重要な洞察

- 現行の暗号技術が量子コンピュータにより**無効化**される可能性
- 量子耐性のある暗号化への**シームレスな移行**が急務
- 量子AIによる**超高速な脆弱性スキャン**と対策の必要性
- 量子センサーネットワークによる**新たな侵入検知手法**の可能性

セッション 4

組織戦略の基盤となるサイバーハイジーン

サイバーハイジーンとは？

- 定義
  - サイバーハイジーンとは、日常的に実施すべき基本的なセキュリティ対策の集約であり、システムの防御力を強化するためのベストプラクティスを指す。
- 重要性
  - 2022年1月、世界経済フォーラムは、**サイバーセキュリティ問題の95%が人為的ミス**に起因し、43%が内部脅威によるものであると報告。
  - **最低限の対策**が、多額の損失や信頼喪失を防ぐ鍵。
- 期待できる効果
  - 増加する**基本的な攻撃を防ぐ**ことが可能。
  - セキュリティインシデントを未然に防ぐことで、システムの復旧やデータ漏洩対応の**コストを削減**。
  - 利害関係者との**信頼関係を維持**し、事業継続性を確保。

出典: 世界経済フォーラム

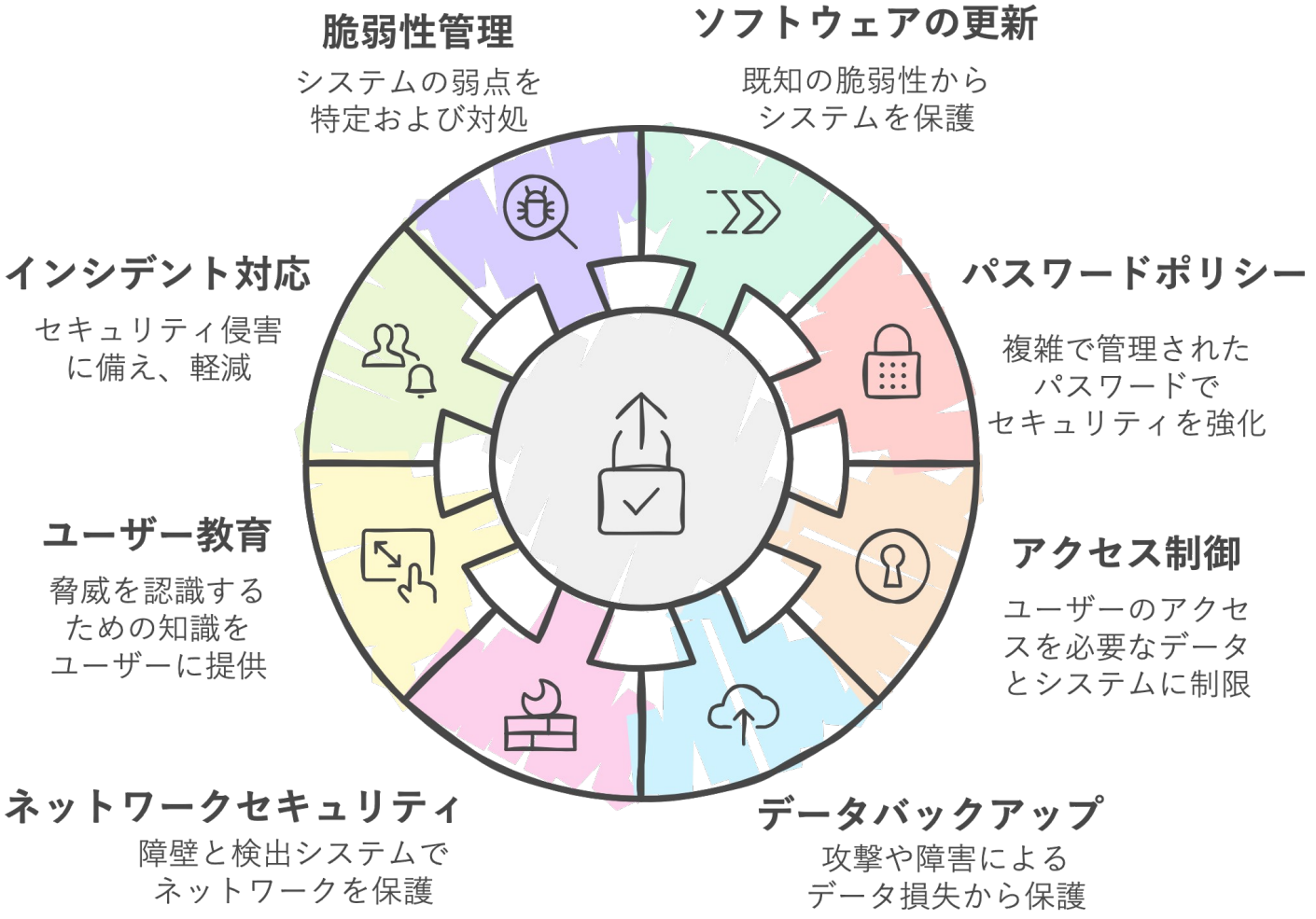
TLP:CLEAR

他の対策と異なるサイバーハイジーンの重要ポイント

- サイバーハイジーンは複雑な対策ではなく、基本的な「習慣」を身につけることが重要。
- 個人から企業レベルまで、日常的な実践により、サイバーセキュリティの基本的な防御力を大幅に向上させる。

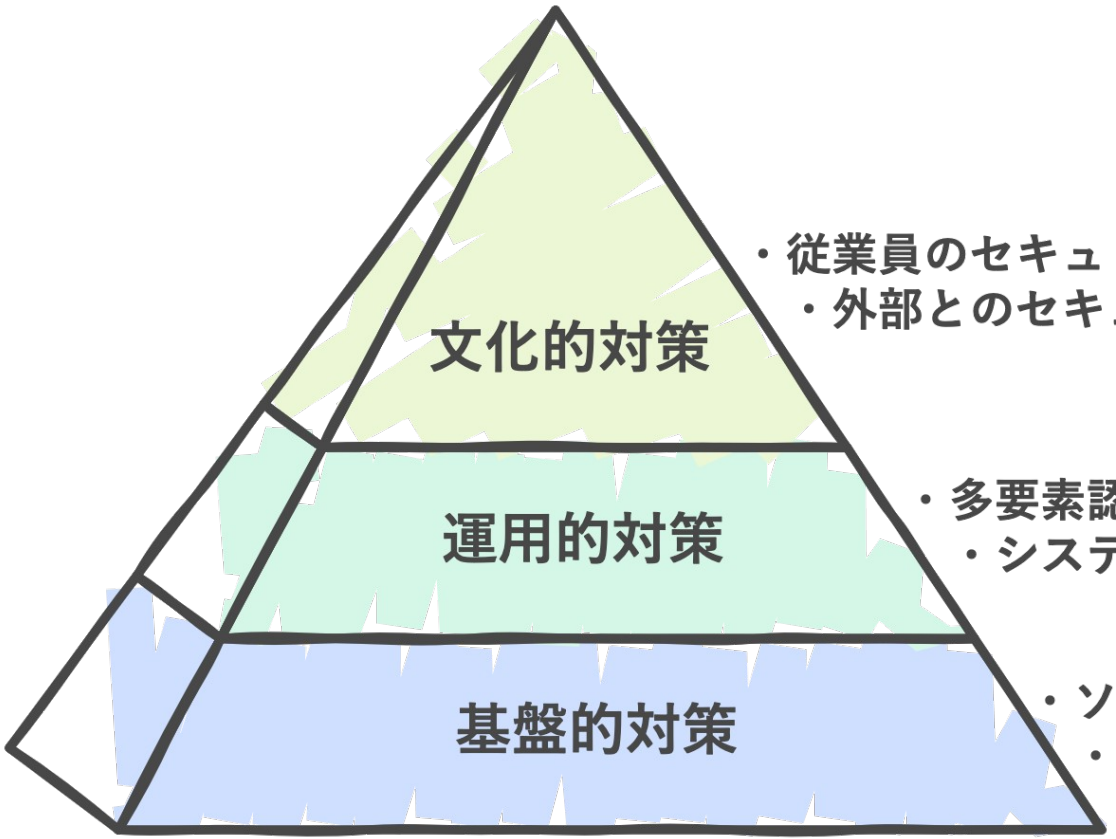
TLP:CLEAR

サイバーハイジーンの基本フレームワーク



TLP:CLEAR

サイバーハイジーンの戦略ピラミッド



・従業員のセキュリティ意識向上
・外部とのセキュリティ契約の明文化

・多要素認証の導入
・システムログの定期確認

・ソフトウェア更新（パッチ適用率向上）
・アクセス制御（最低権限の原則）

TLP:CLEAR

セッション 5

まとめ(主要ポイントの再確認)

主要ポイントの再確認

- **AI駆動型攻撃による従来対策の限界**
  - AI技術により高度化した攻撃に対し、従来の静的防御では不十分。
- **組織戦略とサイバーハイジーンの重要性**
  - 組織戦略は体制の方向性を示し、サイバーハイジーンがその基盤を支える。両者の連携が高度な攻撃への耐性を強化する。
- **セキュリティアプローチの根本的な変革**
  - ゼロトラスト、多層防御の再構築など、動的で適応型の防御へシフト。

TLP:CLEAR

本資料に関する連絡先

名和 利男（Toshio NAWA）

SITE: https://www.nawa.to

PGP: 0xE38B4E01