相次ぐ情報漏洩に備えるAIハッキングの手口と防御実践【会場】（4126291）

本研修は「AIハッキングとは何か」から始め、攻撃者が入口を探し、侵入し、情報に到達する流れを、許可された検証環境での事例を通じて解説します。Web・API・クラウド・開発環境の侵入原因を整理したうえで、認証・権限の点検、脆弱性診断の活用、修正後の再確認、侵入後の被害拡大防止と復旧までを一日で学びます。

開催日時	2027年1月18日(月) 10:00-17:00会場
JUAS研修分類	セキュリティ(サイバーセキュリティ)
カテゴリー	共通業務（契約管理、BCP、コンプライアンス、人的資産管理、人材育成、資産管理）・セキュリティ・システム監査 専門スキル
DXリテラシー	How(データ・技術の活用)：留意点
講師	佐井大海 氏 （ヒューマンハッカー株式会社　取締役・CMO ） 最高月間130万PVの生成AI専門メディアで編集長を務める。 ChatGPT・Claudeからローカル LLM、オープンソースAIまで、新技術を自ら検証し、国内最速で記事化。その仕組みと実践的な活用方法を、ハンズオン形式で発信してきた。 AIコンサルティングやプロジェクトマネジメントにも従事。 現在はAIハッキング専門メディアを立ち上げ、リサーチエンジニアとして、GLMなどの生成AIモデルとStrixなどのオープンソースツールを組み合わせた技術検証に取り組む。 検証環境でAIによる脆弱性探索から侵入・情報漏洩に至る仕組みを分析し、企業の防御策を発信している。 ヒューマンハッカーによる国内サイバー事案の集計は、安野貴博氏の記事でも引用されている。
参加費	JUAS会員企業/ITC：35,200円 一般：45,100円（１名様あたり　消費税込み、テキスト込み）【受講権利枚数1枚】
会場	一般社団法人日本情報システム・ユーザー協会（NBF東銀座スクエア２F）
対象	・情報システム部門の管理者・担当者、セキュリティ担当者 ・AI導入やシステム開発・運用を管理するDX推進担当者、プロジェクトリーダー ・AIハッキングの基礎から学び、自社の対策や委託先への確認事項を整理したい方 ・専門的な攻撃技術やプログラミング経験は不要です。自社のシステム構成や運用業務の概要を把握していると、より具体的に学べます。 中級
開催形式	講義
定員	25名
取得ポイント	※ITC実践力ポイント対象のセミナーです。（2時間1ポイント）
ITCA認定時間	6

主な内容

■受講形態
会場のみ（オンラインなし）
■テキスト
当日配布
■開催日までの課題事項
特になし

相次ぐ情報漏洩に備える AIハッキングの手口と防御実践
ーホワイトハッカーの視点で学ぶ、攻撃の仕組みから脆弱性診断・対策までー

国内で情報漏洩や不正アクセスの公表が相次ぐ中、企業には攻撃の仕組みを理解し、自社の弱点を具体的な対策につなげることが求められています。
攻撃者は生成AIを調査やコード作成に利用し、システムの応答を見ながら次の試行を選びます。

AIによって攻撃の進め方が変わる今、防御側も点検と改善の進め方を見直す必要があります。

本研修は「AIハッキングとは何か」から始め、攻撃者が入口を探し、侵入し、情報に到達する流れを、許可された検証環境での事例を通じて解説します。
Web・API・クラウド・開発環境の侵入原因を整理したうえで、認証・権限の点検、脆弱性診断の活用、修正後の再確認、侵入後の被害拡大防止と復旧までを一日で学びます。
さらに、生成AIを調査や診断結果の整理に活用する際の留意点と、人が判断すべき事項を確認します。サンプル事例を使って対策の優先順位と担当を考え、自社の点検・改善計画につなげます。

◆到達目標
・AIを使う攻撃の流れと、情報漏洩につながる原因を説明できる。
・自社で確認すべき公開範囲・認証・権限・更新状況を整理できる。
・診断の指摘を読み、修正の優先順位と再確認の方法を考えられる。
・侵入後の対応と、継続点検に必要な担当・周期を整理できる。

◆研修形式
・講義、講師による検証事例の解説、紙面演習、質疑応答
・参加者のPC持参・事前のソフト導入は不要です。

◆主な内容
第1部 AIハッキングの仕組みを理解する
AIハッキングとは何かを整理し、攻撃者がAIを調査・判断・試行に使う流れを学びます。
従来の自動攻撃とAIを組み合わせた攻撃の違い／企業規模だけでは判断できない、公開システムの狙われ方／情報漏洩と業務停止の違い、被害事例を読む際の注意点

第2部 検証事例から侵入と情報漏洩の原因を知る
検証用アプリで起きたことを、入口の発見からデータへの到達まで順に解説します。
公開範囲の不備、入力処理の問題、認証と権限確認の違い／AIが結果に応じて次の試行を選ぶ過程／どの段階で防げたかを考えるケース討議

第3部 自社の入口と認証を点検する
Web・API・クラウド・開発用環境を対象に、優先して確認する項目を整理します。
公開資産の棚卸し、不要な管理画面・APIの停止、修正状況の確認／管理者・メール・リモート接続の認証強化と権限の限定／委託先との接続、開発環境の秘密情報、サンプル点検シートの記入

第4部 脆弱性診断を修正と再確認につなげる
診断の指摘を対策に落とし込み、対応済みと判断するまでの流れを学びます。
自動診断で確認できる範囲と、業務上の権限確認など個別に検証する事項／事業への影響と悪用の可能性を踏まえた優先順位／生成AIによる脆弱性情報・診断結果の整理と、根拠の照合・機密情報の扱い／サンプル診断結果から対応表を作る演習

第5部 侵入後の被害拡大と業務停止に備える
侵入防止に加え、異常の把握から遮断・復旧までの備えを整理します。
ログの確認、初動の連絡先と判断者、接続の遮断／セッション・トークンの失効と認証情報の見直し／バックアップの分離、復元確認、必要な復旧時間の検討

第6部 自社の点検と改善を続ける計画を作る
学んだ項目を、自社の担当者・対応順・点検周期に結び付けます。
公開資産・権限・更新の変化を継続的に確認する運用／修正後の再確認と、未対応事項を追跡する方法／最初に取り組む3項目を整理する演習と質疑応答