報道関係者各位
PRESS RELEASE
株式会社サイバーセキュリティクラウド

2026 年 7 月 28 日

『AI MONBAN』に続く第二弾！
世界水準の AI Red Teaming と Web セキュリティの知見を融合した『AI アプリケーション脆弱性診断サービス』を提供開始

生成 AI を活用したサービスに潜むセキュリティリスクを攻撃者の視点で検証。発見した問題と具体的な改善方法を提示。

グローバルセキュリティメーカーの株式会社サイバーセキュリティクラウド（本社：東京都品川区、代表取締役社長 兼 CEO：小池 敏弘、以下「当社」）は LLM、RAG、MCP、AI エージェント などを活用した AI アプリケーションを対象に、固有のセキュリティリスクを診断する『AI アプリケーション脆弱性診断サービス』を 2026 年 7 月 28 日より提供開始します。
本サービスは AI や AI エージェントに関わる MCP のデータフローを可視化・制御・監査する『AI MONBAN』に続く、当社グループの AI セキュリティ領域における第二弾のサービスです。
AI 活用時のリスクを「可視化・制御する」領域から、AI アプリケーション自体の安全性を「診断・評価する」領域へとサービスを拡充します。

■ AI アプリケーションにも提供開始前の安全性検証が求められる時代へ

昨今、生成 AI や AI エージェント の活用が広がり、RAG や MCP、Tool Calling などを組み込んだ AI アプリケーションの本番利用が急速に進んでいます。
一方で AI アプリケーションにはプロンプトインジェクションや機密情報の漏洩、不正なツール実行など、従来の Web アプリケーションとは異なるリスクが存在するほか、外部 API や業

本リリースに関する報道関係のお問い合わせ
株式会社サイバーセキュリティクラウド　広報 井野部さりー・川崎・絹本
TEL：080-4486-0029（いのべ さりー）　MAIL：pr@cscloud.co.jp

務システム、各種ツールとの連携によってアタックサーフェイスも拡大しており、従来型の脆弱性診断だけでは十分に評価できないケースもあります。
こうした背景から当社は AI・LLM、RAG、MCP、AI エージェント 固有のリスクを攻撃者の視点から検証し、具体的な改善・緩和策まで提示する『AI アプリケーション脆弱性診断サービス』を提供します。

■『AI アプリケーション脆弱性診断サービス』について

『AI アプリケーション脆弱性診断サービス』は LLM を利用した Web アプリケーション、生成 AI サービス、RAG システム、AI エージェント、MCP を利用したシステムなどを対象に、AI アプリケーションの安全性を評価するプロフェッショナルサービスです。
「OWASP Top 10 for Large Language Model Applications」に基づく診断項目をベースに、対象システムの用途、利用モデル、参照データ、権限、外部サービスとの連携状況などに応じて、診断範囲と攻撃シナリオを個別に設計します。
自動診断ツールによる体系的な検査と、専門家による AI Red Teaming を組み合わせることで、機械的なチェックだけでは発見が難しい、システムの仕様、権限設計、業務フロー、AI エージェント の自律的な挙動、外部連携などに起因するリスクを検証します。
診断後には検出した脆弱性の内容、重要度、想定される影響、確認方法に加え、具体的な改善・緩和策を診断報告書として提示します。

■ 主な診断対象

- LLM を利用した Web アプリケーション・生成 AI サービス
- RAG を利用したシステム
- AI エージェント・マルチエージェントシステム
- MCP・Tool Calling を利用したシステム
- LLM API を利用した業務システム
- マルチモーダル AI システム

※ システム構成や利用モデルに応じて、診断範囲を個別に設計します。

本リリースに関する報道関係のお問い合わせ
株式会社サイバーセキュリティクラウド　広報 井野部さりー・川崎・絹本
TEL：080-4486-0029（いのべ さりー）　MAIL：pr@cscloud.co.jp

■ 主な診断観点

- LLM・生成 AI：プロンプトインジェクション、情報漏洩、出力制御、サプライチェーンリスクなど。
- RAG：データ汚染、アクセス制御、情報漏洩など。
- AI エージェント：権限管理、不正なツール実行、外部システムとの連携など。
- MCP・Tool Calling：認証・認可、MCP サーバ、クレデンシャル管理、外部サービスとの連携など。

※ システム構成に応じて、Web アプリケーション、API、認証・認可、データベース、クラウド基盤などとの接続面も確認します。

■ 『AI アプリケーション脆弱性診断サービス』の特徴

1. Web・クラウドセキュリティで培った知見を AI セキュリティ へ展開

当社は Web アプリケーションをサイバー攻撃から守るセキュリティサービスを提供し、WAF の運用、防御、脆弱性情報の収集・管理、クラウド環境のマネージドセキュリティなど、Web・クラウドセキュリティ領域における知見を蓄積してきました。
本サービスでは、AI・LLM 固有の脆弱性を検出するだけでなく、AI アプリケーションを構成する Web アプリケーション、API、認証・認可、データベース、クラウド基盤なども含め、実際のシステム構成や運用を踏まえてリスクを評価します。AI 固有の専門性と当社が培ってきた Web・クラウドセキュリティの知見を組み合わせることで、AI アプリケーション全体を対象とした実践的な診断を提供します。

2. OWASP Top 10 for Large Language Model Applications に基づく体系的な診断

国際的な AI・LLM セキュリティ基準である「OWASP Top 10 for Large Language Model Applications」に基づき、LLM、RAG、AI エージェント、MCP、Tool Calling など、AI アプリケーション固有のリスクを体系的に評価します。

本リリースに関する報道関係のお問い合わせ
株式会社サイバーセキュリティクラウド　広報 井野部さりー・川崎・絹本
TEL：080-4486-0029（いのべ さりー）　MAIL：pr@cscloud.co.jp

対象システムの用途や構成に応じて、診断項目や攻撃シナリオを個別に設計し、実際の利用環境に即した診断を実施します。

3. 専門家による AI Red Teaming

AI・LLM は入力内容や利用環境によって挙動が変化するため、自動診断だけでは十分に評価できないリスクが存在します。
本サービスでは自動診断ツールによる体系的な検査に加え、専門家が攻撃者の視点から AI Red Teaming を実施します。間接プロンプトインジェクションや AI エージェント の権限・外部連携を悪用した攻撃など、自動診断だけでは発見が難しいリスクまで検証し、AI アプリケーション全体の安全性を評価します。

4. 診断から改善までを見据えた実践的な支援

対象システムの構成や利用状況を踏まえて診断を実施し、検出した脆弱性について、リスク評価、想定される影響、対応優先順位、具体的な改善・緩和策を整理した診断報告書を提供します。
診断結果報告会を通じて内容を説明するとともに、必要に応じて改善コンサルティングや実装・運用支援などの関連サービスも提案し、お客様の継続的な AI セキュリティ対策を支援します。

■ Konoe Intelligence 株式会社との業務提携について

当社は AI・LLM 脆弱性診断及び AI Red Teaming 領域におけるサービス提供体制の強化を目的として、Konoe Intelligence 株式会社と業務提携しました。
Konoe Intelligence は AI・LLM 脆弱性診断及び AI Red Teaming に高い専門性を有する AI セキュリティカンパニーです。同社のチームは、OpenAI 主催の国際的なセキュリティ競技「Red-Teaming Challenge – OpenAI gpt-oss-20b」において、5,000 件を超える応募の中から、日本人として唯一 TOP20 に入賞されました。
本提携により同社が有する AI Red Teaming 及び AI・LLM 脆弱性診断の専門性と、当社が Web アプリケーションセキュリティ、WAF、クラウドセキュリティ、マネージドセキュリティ運用を通じて培ってきた防御・運用の知見を融合します。これにより AI・LLM 固有のリスクに加え、その周辺を構成する Web、API、認証・認可、クラウド基盤までを踏まえた実効性の高い脆弱性診断サービスを提供します。今後も両社の連携を通じて AI セキュリティ 領域におけるサービス及び技術の高度化を推進してまいります。

■ Konoe Intelligence 株式会社について

Konoe Intelligence は「安全な超知能に最速で到達する」をミッションに掲げる AI セキュリティカンパニーです。
企業がセキュアな AI 活用を実現するために AI・LLM 脆弱性診断、AI Red Teaming、シャドー

本リリースに関する報道関係のお問い合わせ
株式会社サイバーセキュリティクラウド　広報 井野部さりー・川崎・絹本
TEL：080-4486-0029（いのべ さりー）　MAIL：pr@cscloud.co.jp

AI の検出、AI ガバナンスの可視化、AI ファイアウォールなど、AI セキュリティに関するサービス・ソリューションを提供しています。

会社名：Konoe Intelligence 株式会社
所在地：京都府京都市上京区甲斐守町 97 西陣産業創造會館
代表者：代表取締役　勘佐 圭吾
設立：2025 年 2 月
事業内容：AI セキュリティ事業、サイバーセキュリティサービス
Web サイト：https://www.konoe-intelligence.net/

■ 株式会社サイバーセキュリティクラウドについて
会社名：株式会社サイバーセキュリティクラウド
所在地：〒141-0021 東京都品川区上大崎 3-1-1 JR 東急目黒ビル 13 階
代表者：代表取締役社長 兼 CEO　小池 敏弘
設立：2010 年 8 月
Web サイト：https://www.cscloud.co.jp/

「世界中の人々が安心安全に使えるサイバー空間を創造する」をミッションに掲げ、世界有数のサイバー脅威インテリジェンスを活用した Web アプリケーションセキュリティサービスを軸に、脆弱性情報収集・管理ツールや、クラウド環境のフルマネージドセキュリティサービスを提供する日本発のセキュリティメーカーです。
当社はサイバーセキュリティにおけるグローバルカンパニーの一つとして、サイバーセキュリティに関する社会課題を解決し、社会への付加価値提供に貢献してまいります。

■ 本サービスに関するお問い合わせ
株式会社サイバーセキュリティクラウド
戦略事業室
TEL：03-6416-9996
Web サイト：https://www.cscloud.co.jp/ai-application-vulnerability-assessment/