

PRESS RELEASE

デジタルアーツ、ChatGPT など生成 AI に対応する「AI チャットフィルター」を開発、 「i-FILTER」シリーズに 8 月 16 日から標準搭載

～生成 AI に送信したメッセージなどを可視化し、情報漏えい防止や効果的な活用が可能に～

情報セキュリティメーカーのデジタルアーツ株式会社(本社:東京都千代田区、代表取締役社長:道具 登志夫、以下 デジタルアーツ、証券コード 2326)は、ChatGPT など生成 AI(人工知能)に対応する「AI チャットフィルター」機能を開発し、8月16日から Web セキュリティ製品「i-FILTER」とクラウド版の Web セキュリティ製品「i-FILTER@Cloud」に標準搭載することを発表します。

生成 AI の活用が広がる中、情報漏えいなどの課題も

昨今、生成 AI の ChatGPT や Bard、Bing AI の活用が広がっています。生成 AI は、従来の AI では難しかった文章などを自動的に作り出すことができ、資料の作成などさまざまな業務の効率化が進むと期待されています。

一方で、企業や行政機関が生成 AI を活用する際には、機密情報や個人情報の漏えいなどの課題もあります。さらに学校現場においては学習効果への影響も懸念されています。また、生成 AI の利用がシャドーIT(従業員が企業の管理下でない IT 機器やサービスを利用すること)化していないかなど利用実態のほか、情報漏えいの防止や効果的な活用のために生成 AI へ送信したメッセージ内容を把握することも重要です。

「AI チャットフィルター」機能で生成 AI 利用拡大に対応

デジタルアーツは、生成 AI の利用方法のコントロールや利用状況を可視化できる「AI チャットフィルター」機能を開発し、8月16日から「i-FILTER」と「i-FILTER@Cloud」に標準搭載します。両製品は、長年収集した URL のデータベースを活用することにより、デジタルアーツが安全と判断した Web サイトのみアクセスできる「ホワイト運用」を実現しています。

「AI チャットフィルター」機能では、いつ・誰が・どんなメッセージを生成 AI に送信したかといった情報をログとして表示します。機密情報など特定単語を含む生成 AI へのメッセージ送信をブロックするフィルタリング機能もあり、ブロックした単語をログから閲覧することも可能です。生成 AI 利用時には、機密情報の漏えいなどに注意が必要な旨を伝える警告画面も表示します。これらの機能により、適切な生成 AI 利用が行われているかを確認することが可能となります。

「AIチャットフィルター」機能

「AIチャットフィルター」機能	内容
ログ	いつ・誰が・どんなメッセージを生成AIに送信したか、ブロックした単語を管理画面から閲覧できる
フィルタリング	機密情報を含む特定単語のブロックなどのフィルタリングが可能
警告画面表示	生成AI利用には注意が必要である旨を警告画面で伝える

また、従来の「i-FILTER」の機能である「Web サービス制御」では、生成 AI である ChatGPT や Bard、Bing AI を閲覧許可またはブロックとすること、「グループ設定」機能では任意のグループ・ユーザー単位で生成 AI の利用を制御すること、「時間割」機能では就業時間や授業時間など設定した時間のみ生成 AI を利用できるように設定することが可能です。

「Webサービス制御」機能で生成AIを制御

サービス名	認証	有効	機能名	アクション	許可カテゴリ名	リスク
<input type="checkbox"/> OpenAI (ChatGPT)	-	<input checked="" type="checkbox"/> ON	OpenAI (ChatGPT)	許可	ビジネス向けWebアプリケーション	2
<input type="checkbox"/> Google Bard	-	<input type="checkbox"/> OFF	Google Bard	許可	ビジネス向けWebアプリケーション	2
<input type="checkbox"/> Bing AI	-	<input checked="" type="checkbox"/> ON	Bing AIチャット検索	許可 ブロック	ビジネス向けWebアプリケーション	2

企業や行政機関においては、ChatGPT などの生成 AI にどのようなメッセージを送信したかをログから分析できるため、より効果的に生成 AI を活用するにはどうすれば良いかなどの検討が容易になります。部署や役職に適した「グループ設定」機能と組み合わせることで、特定の部署・役職には生成 AI の利用を許可し、それ以外の部署・役職には、許可しないという設定も簡単に行うことができます。

「AIチャットフィルター」機能のログイメージ

※オンプレミス版「i-FILTER」

No.	クライアント	グループ	時間	サービス	処理	検出単語
1	192.168.10.109 kobayashi	マーケティング部	2023/06/08 19:58:40	chatGPT	許可	-
あなたはマーケティングのエキスパート。 セキュリティ製品の売り上げを伸ばす目的で、広告タイトルのアイデアをだしてください						
2	192.168.10.108 nakamura	マーケティング部	2023/06/08 19:28:40	chatGPT	ブロック	{機密情報}
以下の文章を議事録としてまとめてください。 ~~~~ {機密情報} ~~~~						
3	192.168.10.105 itou	営業部	2023/06/08 17:58:39	Google Bard	許可	-
取引先へのビジネスメールとして添削してください 〇〇会社 〇〇様 ~~~~						
4	192.168.10.104 tanaka	営業部	2023/06/08 17:28:39	Bing AI	許可	-
取引先へのビジネスメールとして添削してください □□会社 □□様 ~~~~						
5	192.168.10.102 suzuki	開発部	2023/06/08 16:28:39	chatGPT	許可	-
あなたは、セキュリティ製品を販売する企業のWeb担当者です。 Webサイトを訪問するユーザーに製品の魅力が伝わる具体的なサイトの構成を考えてください。						
6	192.168.10.101 satou	開発部	2023/06/08 15:58:39	chatGPT	許可	-
以下のプログラムが動作しない 原因を教えてください。 char *str; str = "test"; printf("%s", str);						

学校においては、学年に合わせた「グループ設定」機能と利用時間帯を制御できる「時間割」機能と組み合わせることで、必要な授業のみ生成 AI の利用を許可することも可能です。不適切な使い方をした場合にはレポートを送付できるため、生成 AI を常に安全に利用することが可能です。

「ChatGPT」などの生成 AI を安全に活用するためのオンラインセミナーを 7 月 27 日（木）に開催

「ChatGPT」などの生成 AI を安全に活用！

情報漏えいリスクを対策し、安心して利用するために必要な Web セキュリティとは

本セミナーでは、「ChatGPT」などの生成 AI を利用したサービスにおける情報漏えいリスクについて解説し、ビジネスにおける生成 AI の適切な利用方法や安全かつ効率的に利用するための Web セキュリティ「i-FILTER」についてご紹介します。8 月に追加される「i-FILTER」の新機能も交えてご紹介しますので、ぜひご参加ください。

・日時: 2023 年 7 月 27 日(木) 15:00～15:40

・お申し込み方法: お申し込みフォームに必要事項をご入力ください。[参加費: 無料]

・お申し込みはこちら <https://mktg.daj.jp/public/seminar/view/7685>

■「i-FILTER」

「i-FILTER」は、Web セキュリティ製品です。有害情報や業務に関係のない Web サイトの閲覧を防ぐフィルタリングに加え、外部からの攻撃、内部からの情報漏えいも防ぎます。国内で検索可能な URL を網羅したデータベースにより、危険な Web サイトや未知の脅威へのアクセスをブロックし、デジタルアーツが安全と判定した Web サイトのみアクセスできる環境を実現しています。

<https://www.daj.jp/bs/i-filter/>

デジタルアーツ株式会社 概要

Web、メール、ファイルなどのセキュリティソフトウェアの提供を核に事業展開する情報セキュリティメーカーです。

1995 年の創業以来、「より便利な、より快適な、より安全なインターネットライフに貢献していく」を企業理念とし、有害情報の閲覧を制限する Web フィルタリングソフトを開発、以来企業・公共・家庭向けに情報セキュリティ製品を提供しております。

東京都千代田区大手町 1-5-1 大手町ファーストスクエア ウェストタワー14F ▶URL: <https://www.daj.jp/>

<本リリースに関するお問い合わせ>

デジタルアーツ株式会社 広報課 石井 TEL : 080-8750-0425 / E-mail : press@daj.co.jp

- ※ デジタルアーツ、DIGITAL ARTS、i-FILTER、i-FILTER Anti-Virus & Sandbox、i-FILTER@Cloud Anti-Virus & Sandbox、i-FILTER@Cloud D アラート発信レポートサービス、info board、Active Rating System、D-SPA、Anti-Virus & Sandbox for D-SPA、NET FILTER、SP-Cache、White Web、ZBRAIN、クレデンシャルプロテクション、ホワイト運用、m-FILTER、m-FILTER MailFilter、m-FILTER Archive、m-FILTER Anti-Spam、m-FILTER Anti-Virus & Sandbox、m-FILTER@Cloud Anti-Virus & Sandbox、m-FILTER@Cloud D アラート発信レポートサービス、m-FILTER File Scan、Mail Detox、m-FILTER EdgeMTA、EdgeMTA、FinalCode、i-フィルター、DigitalArts@Cloud、Desk@Cloud、Desk、Desk Event、f-FILTER、D アラートおよび D コンテンツその他の弊社・弊社製品関連の各種名称・ロゴ・アイコン・デザイン等はデジタルアーツ株式会社の登録商標または商標です。
- ※ その他、上に記載された会社名および製品名は、各社の商標または登録商標です。