TCSP Toyama Prefectural Police Cyber Security Personnel

サイバーセキュリティの置き薬

2023 年
第 11 号

「偽」の対話型生成 AI に注意

～情報窃取やマルウェア感染のおそれ～

1　対話型生成 AI とは？

ChatGPT や BingChat、Bard 等の対話型生成 AI は、あたかも人間と自然に会話をしているかのような受け答えが可能であり、文章作成、翻訳等の素案作成等、多岐にわたる活用が広まりつつあります。

対話型生成 AI は、大規模言語モデルに基づき、与えられたテキストに対し**後続する単語等を予測して、回答を生成**するものです。

与えるテキストを工夫することで、高精度な回答を得られる可能性が上がりますが、**回答は誤りを含む可能性**が常にあり、時には、事実と全く異なる内容や文脈と無関係な内容等が出力されることもあります。

2　偽アプリ等の危険性について

昨今、対話型生成 AI に**偽装**した
アプリケーションや**ウェブサイト**
が発見されており、これらを利用すると**知らない間**に

- パスワード等の**情報が抜き取られる**
- **データが暗号化**される
- 偽のウェブサイトから**マルウェア感染**に誘導される

等の動作が実行されるおそれがあります。

3　被害に遭わないための対策

○　基本的なセキュリティ対策

ぜい弱性に対応するため、**OS やソフトウェアを更新**する
ウィルス感染リスクを減少させるため、**対策ソフト等を導入し、最新の状態を維持**する

○　偽アプリ対策

アプリ導入前に提供元や作成者を確認し、**公式アプリ以外は導入しない**
不要なアクセス権限を求められた場合は**許可せず**、判断に迷う場合は導入しない

○　偽のウェブサイト対策

見た目で偽のサイトを見破ることは困難なため、**URL に不審な箇所がないか等を確認**する
マルウェアに感染するおそれがあるため、サイトのリンクを**安易にクリックしない**

○　その他情報漏えい対策

入力内容が提供元で利用されることが規約に明記されているものもあり、意図しない情報漏えいにつながるおそれがあるため、**機密情報や個人情報は入力しない**

【参考】
・TREND MICRO
ChatGPT や Midjourney などの AI ツールを装う不正な広告から情報窃取ツール「RedlineStealer」が拡散
https://www.trendmicro.com/ja_jp/research/23/e/malicious-ai-tool-ads-used-to-deliver-redline-stealer.html
・警察庁（基本的なセキュリティ対策等）
https://www.npa.go.jp/bureau/cyber/index.html

県警公式 HP
「置き薬」

YouTube
注意喚起動画