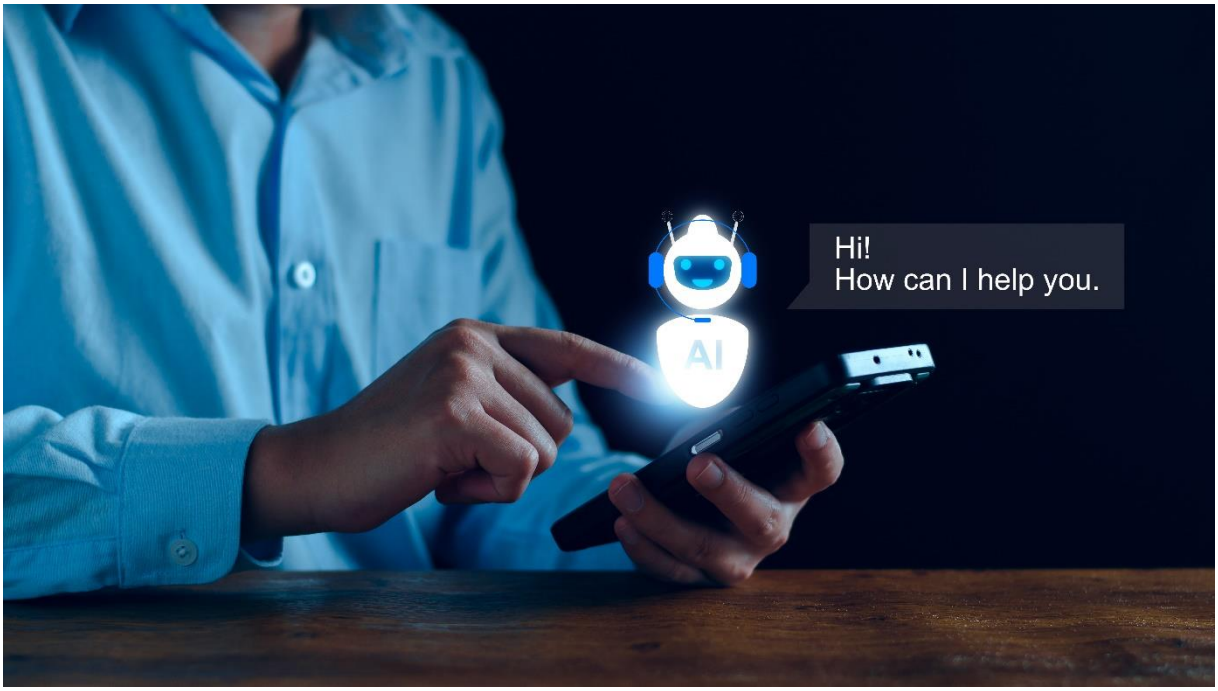
Chat GPT als potenzieller Helfer für Cyber-Kriminelle

Die Fortschritte im Bereich der künstlichen Intelligenz und speziell im Bereich der LLM (Large Language Models) haben zu einer Vielzahl aufregender Entwicklungen geführt, darunter auch Chatbots wie Chat GPT. Diese leistungsstarken Modelle können menschenähnliche Gespräche führen und bei einer Vielzahl von Aufgaben helfen. So schreibt der Chatbot in kürzester Zeit einen Aufsatz, komponiert Lieder oder verfasst E-Mails. Dies machen sich vermehrt auch Kriminelle zunutze.



Phishing und Social Engineering

Die Fähigkeit von Chat GPT, menschenähnliche Texte zu generieren, eröffnet Cyber-Kriminellen neue Möglichkeiten zur Täuschung. Durch die Nachahmung echter Personen oder Institutionen können sie **Phishing**-**Angriffe** durchführen. Ein Beispiel hierfür ist das Erstellen von **gefälschten Webseiten oder authentischen E-Mails:** Dabei werden Opfer dazu verleiten, ihre **Zugangsdaten oder finanziellen Informationen preiszugeben**. Hierfür wird ein **Social Engineering-Ansatz verwendet, welcher** durch die Ausnutzung menschlicher Eigenschaften deren Vertrauen erweckt und diese so zum Handeln verleitet. Zwar wurden bei Chat GPT Sicherheitsvorkehrungen getroffen, um potenziell bedrohliche Absichten zu erkennen, jedoch können diese durch eine richtig gestellte Frage (Prompt-Engineering) umgangen werden.

Automatisierung von Angriffen

Chat GPT kann auch dazu verwendet werden, **Angriffe auf Computersysteme zu automatisieren**. Durch die Verwendung der KI können Cyber-Kriminelle **maßgeschneiderte Malware erstellen, Schwachstellen identifizieren oder Angriffe auf Netzwerke durchführen**.

Erstellung von Malware, Exploits und Ransomware:

Cyber-Kriminelle könnten Chat GPT dazu nutzen, **maßgeschneiderte Malware** (Schadsoftware) und Exploits (Malware/Befehlsfolge zur Ausnutzung von Sicherheitslücken und Fehlfunktionen) zu entwickeln. Durch die Programmierung des Sprachmodells mit Kenntnissen über Schwachstellen und Angriffsmethoden können die Kriminellen automatisch schädlichen Code generieren, der auf bestimmte Ziele angepasst ist. Dies könnte die Effektivität und Verbreitung von Malware erhöhen und die Entdeckung durch Sicherheitslösungen erschweren.

Bei einer **Ransomware-Attacke** werden Daten des Opfers zunächst verschlüsselt, um anschließend Lösegeld für deren Freigabe zu erhalten. Für die Zahlung des Erpressergeldes nutzen Hacker die KI auch, um Zahlungssysteme für Kryptowährungen zu erstellen. Doch nicht nur für

diese Finanzbewegungen wird die KI verwendet, sondern auch für Geldwäsche. Durch die authentischen Gespräche über Geschäftsaktivitäten werden die Transaktionen von Überwachungssystemen übersehen oder als unauffällig eingestuft.

Fazit

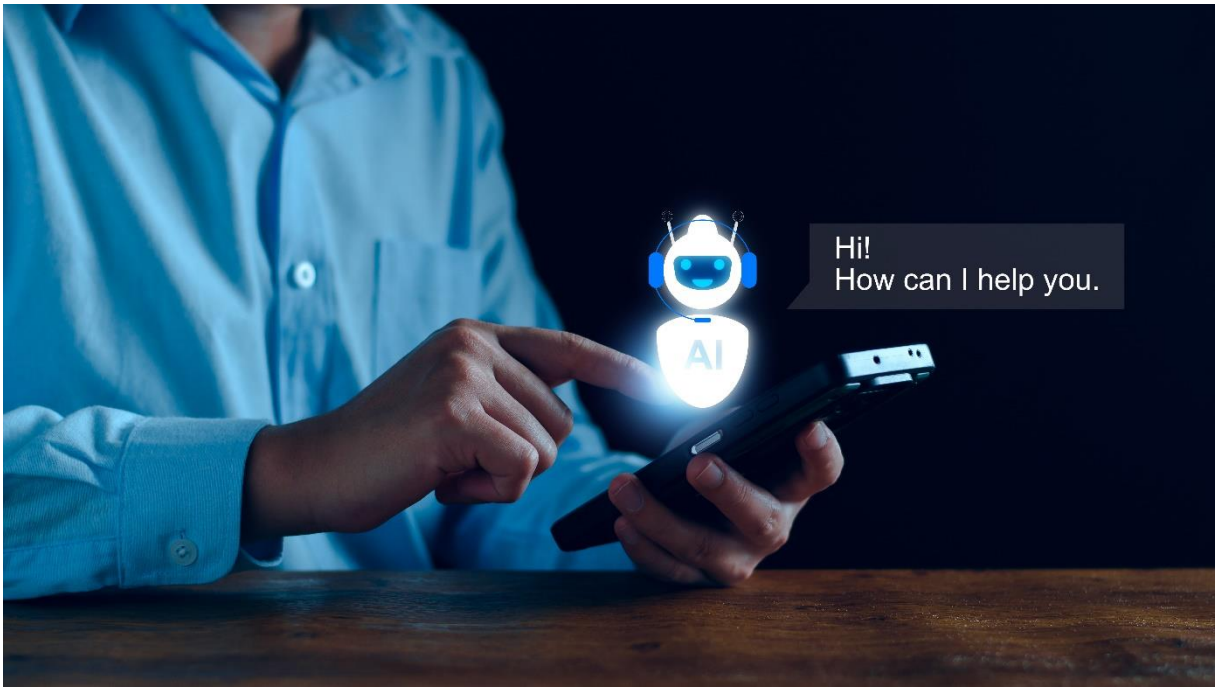
Die potenzielle Verwendung von Chat GPT als Helfer für Cyberkriminelle birgt ernsthafte Risiken. Es ist entscheidend, dass Entwickler, Hersteller und Benutzer dieser Technologie ihre **Verantwortung erkennen und Maßnahmen ergreifen**, um Missbrauch zu verhindern. Die Security-Experten von **KONVERTO** weisen darauf hin, dass durch die Implementierung einer **umfangreichen Sicherheitslösung** sowie kontinuierlicher Überwachung Modelle diese **Risiken minimiert** werden können. Gleichzeitig sollten wir die positiven Eigenschaften von Chat GPT verwenden und **ethische und verantwortungsvolle Nutzung** fördern.

Was ist Chat GPT?

Chat GPT (Generative Pre-trained Transformer) ist ein leistungsstarkes KI-Modell, welches im November 2022 von OpenAI veröffentlicht wurde. Die Software wurde darauf trainiert, auf eine Vielzahl von Fragen und Anfragen in natürlicher Sprache zu antworten. So kann die KI-Texte generieren, Lieder schreiben, Informationen bereitstellen (manchmal veraltet oder falsch dargestellt) und bei Problemlösung helfen. Seit Anfang Mai ist das Tool nach kurzer Sperre auch in Italien wieder verfügbar.

Chat GPT come potenziale aiuto per i criminali informatici

I progressi nel campo dell'intelligenza artificiale, e in particolare nell'area degli LLM (Large Language Models), sono stati all'origine di una serie di sviluppi interessanti, tra cui i chatbot come Chat GPT. Tali strumenti sono in grado di condurre conversazioni simili a quelle umane e di aiutare a svolgere una serie di compiti. Ad esempio, il chatbot è in grado di scrivere un saggio, comporre canzoni oppure scrivere e-mail in pochissimo tempo. In questo modo, i malintenzionati ne approfittano sempre di più.



Phishing e Social Engineering

La capacità di Chat GPT di generare testi simili a quelli umani apre nuove opportunità per i criminali informatici di ingannare. Impersonando persone o istituzioni reali, possono effettuare **attacchi** di **phishing**. Un esempio di ciò è la creazione di **siti Web falsi** o **e-mail autentiche,** in cui le vittime sono indotte con l'inganno a **rivelare le loro credenziali o** informazioni. A tal fine, **viene utilizzato un approccio di Social Engineering**, che risveglia la loro fiducia sfruttando le caratteristiche umane e quindi li invoglia ad agire. Sebbene siano state prese precauzioni di sicurezza con Chat GPT per rilevare intenzioni potenzialmente minacciose, queste possono essere aggirate ponendo una domanda corretta (Prompt-Engineering).

Automazione degli attacchi

Chat GPT può anche essere utilizzato per **automatizzare gli attacchi ai sistemi informatici**. Utilizzando l'intelligenza artificiale, i criminali informatici possono **creare malware personalizzato, identificare vulnerabilità o eseguire attacchi alle reti**.

Creazione di malware, exploit e ransomware

I criminali informatici potrebbero utilizzare Chat GPT per **sviluppare malware** (software dannoso) ed exploit (sequenza di programmi/comandi dannosi per sfruttare vulnerabilità e malfunzionamenti) personalizzati. Grazie alla programmazione del modello linguistico con la conoscenza delle vulnerabilità e dei metodi di attacco, i criminali possono generare automaticamente codice maligno su misura per obiettivi specifici. Questo potrebbe aumentare l'efficacia e la diffusione del malware e renderne più difficile il rilevamento da parte delle soluzioni di sicurezza.

In un **attacco ransomware**, i dati della vittima vengono prima crittografati e poi viene richiesto un riscatto per il loro rilascio. Per pagare il riscatto, gli hacker utilizzano anche l'intelligenza artificiale per creare sistemi di pagamento in criptovaluta. Tuttavia, l'intelligenza artificiale non viene utilizzata solo per questi movimenti finanziari, ma anche per il riciclaggio di denaro. Mediante

conversazioni autentiche sulle attività commerciali, le transazioni vengono ignorate o classificate come poco appariscenti dai sistemi di sorveglianza.

Conclusione

L'uso potenziale della chat GPT come strumento per i cyber-criminali comporta seri rischi. È fondamentale che gli sviluppatori, i produttori e gli utenti di questa tecnologia **riconoscano le proprie responsabilità e adottino misure per prevenire gli abusi**. Gli esperti di sicurezza di **KONVERTO** sottolineano che questi **rischi** possono essere **ridotti al minimo** implementando una **soluzione di sicurezza esaustiva** oltre a modelli di monitoraggio continuo. Allo stesso tempo, dovremmo sfruttare le caratteristiche positive della chat GPT e promuovere un **uso etico e responsabile.**

Cos'è Chat GPT?

Chat GPT (Generative Pre-trained Transformer) è un potente modello di intelligenza artificiale rilasciato da Open AI nel novembre 2022. È stato istruito per rispondere ad un'ampia gamma di domande e quesiti in linguaggio del tutto naturale. Ciò consente all'IA di generare testi, scrivere canzoni, fornire informazioni (talvolta obsolete o travisate) ed aiutare a risolvere i vari problemi. Dall'inizio di maggio lo strumento è nuovamente disponibile in Italia in seguito ad una breve interruzione