Hört auf, ChatGPT eure Geheimnisse zu verraten!

Text von
Ben Seegatz, Data Scientist LOOPING GROUP

Foto von

02.02.2023

5 MINUTEN

SIGN UP HERE

SHARE

- Amazon und andere Unternehmen warnen ihre Mitarbeiter:innen davor, interne Daten mit ChatGPT zu teilen.
- Welches Risiko stellen KI-Tools wie ChatGPT wirklich dar?
- Und wie können Unternehmen sich schützen?

Wir alle haben früher gelernt: Vertrau keinen Fremden im Internet. Sei vorsichtig, was du über dich online verrätst. Aber es sieht aus, als hätten wir diese einfache Lektion vollkommen vergessen. Wir posten live, wo wir uns gerade aufhalten – und verraten so dem ganzen Netz, wann unsere Wohnung leer steht. Influencer (und solche, die es werden wollen) erzählen für ein paar Likes auf LinkedIn ihre gesamte Mental Health Geschichte. Und immer mehr Menschen verraten Chat GPT Firmengeheimnisse.

Vor wenigen Tagen wurden einige Nachrichten aus den internen Slack-Channels von Amazon öffentlich: Einige Mitarbeiter:innen hatten ChatGPT für ihre Arbeit verwendet, um ihre Produktivität zu erhöhen. So nutzten sie ChatGPT unter anderem als Coding-Assistent, um ihre Codes effizienter zu gestalten – nur sind diese natürlich vertraulich.

Aber warum kann das zum Problem werden?

Kurz zur Erinnerung: ChatGPT ist ein auf künstlicher Intelligenz basierendes Konversations-Tool, das auf eine Vielzahl von Fragen und Aufforderungen schnelle und intelligente Antworten finden kann. Durch die Effizienz der Antworten und die vielfältigen Anwendungsmöglichkeiten, hat das Tool Potential, die Redaktionelle Gesellschaft radikal zu verändern.

Das Problem ist, dass ChatGPT keine sichere Plattform ist. Es ist ein öffentlich zugängliches Tool, das von jedem benutzt werden kann, der Zugang zum Internet hat. Auch wenn bei ChatGPT diverse Maßnahmen implementiert sind, um personenbezogene Daten zu schützen, ist keine Übertragung im Internet jemals vollständig sicher oder fehlerfrei.

Das heißt, dass jede Information, die man in ChatGPT eingibt, potenziell von Dritten eingesehen werden kann. Und dazu gehören auch Hacker, Cyberkriminelle, Wirtschaftsspione und andere mit schlechten Absichten. Das erhöht das Risiko von Identitätsdiebstahl, Betrug und anderen Online Scams.

Das Problem liegt im System

ChatGPT ist zudem mit einer gigantischen Datenmenge trainiert. In den Nutzungsbedingungen steht: Um das Tool weiter zu verbessern, dürfen deine Eingaben gespeichert und weiterverwendet werden. Von dem verbesserten Modell profitieren dann auch andere Akteure und Unternehmen.

Genau das wurde einigen Amazon-Mitarbeiter:innen zum Verhängnis: Wenn ChatGPT Eingaben wie Zeilen eines vertraulichen Codes speichert und weiterverarbeitet, macht das Tool es Hackern leichter, Schwachstellen zu finden. Und dafür können sie bequem einfach ChatGPT benutzen. Eine hochrangige Anwältin von Amazon warnte, dass es bereits Fälle gab, in denen der ChatGPT-Output geheimen Daten von Amazon ähnelte.

Wie können wir uns besser schützen?

Eine Lösung könnten eigene KI-Programme sein, die extra für Firmen entwickelt und nur von diesen intern genutzt werden. Branchengerüchten zufolge arbeitet zum Beispiel Amazon bereits an so einem Projekt. Diese Programme könnten auf den Ton einer Firma trainiert werden und wären so

nicht nur sicherer, sondern auch besser auf die Inhalte, Stil und anderen Bedürfnisse des Unternehmens angepasst.

Doch auch abseits von ungewollt veröffentlichten Daten kann ChatGPT zum Cybersicherheitsrisiko werden: Das Programm kann auch einfach zum Phishing genutzt werden. Noch nie war es so einfach, schnell und überzeugend Mails im Ton einer echten Firma zu schreiben und zu verbreiten, oder Chat-Bots zu faken, um Menschen im Chat Geheimnisse wie Passwörter, Kontodaten oder Adressen zu entlocken.

OpenAI, das Unternehmen hinter ChatGPT, wird darauf wohl sicher bald eine Antwort finden müssen. Bis dahin haben sind hier sechs Tipps, um deinen Computer und Smartphone sicherer zu machen:

1. Nutze **sichere Passwörter** – auch für den Account bei ChatGPT. Grundsätzlich gilt: je länger, desto besser. Minimum: 12-14 Zeichen. Großbuchstaben, Zahlen und Sonderzeichen zu mischen hilft.
2. **Zwei-Faktor-Authentifizierung** gibt Accounts eine zusätzliche Sicherheitsebene. Beim Einloggen werden zwei Formen der Identifizierung verlangt, normalerweise ein Passwort und ein Code, der beispielsweise ans Telefon gesendet wird.
3. Halte deine Betriebssysteme stets auf dem **neusten Stand**: Updates enthalten meistens Sicherheitspatches, die Schwachstellen beheben und den Computer/das Smartphone vor Angriffen schützen. Also, auch wenn es nervt: Updates nie wegklicken, sondern gleich installieren.
4. Öffne nur E-Mails von **vertrauenswürdigen Absendern**, klicke nicht auf Links und downloade keine Attachments von unbekannten Absendern. Wenn euch der Ton einer E-Mail seltsam vorkommt: Ruft die Person erst an, die euch die E-Mail geschickt hat.
5. Unternehmen sollten regelmäßige **Schulungen** zu KI-Bots und Cybersecurity für ihre Mitarbeiter:innen anbieten, um sie auf potentielle Bedrohungen vorzubereiten.
6. Und natürlich: **Gib keine vertraulichen Daten an Bots wie ChatGPT.** Auch wenn die Nutzungsmöglichkeiten beeindruckend sind, gilt nach

wie vor die altbekannte Regel: Vertraue keinem Fremden im Internet – auch keinem KI-Bot.

Zur Person

Ben Seegatz ist seit Herbst 2019 Associate Data Scientist bei der LOOPING Group in Berlin. Er studierte als Jungstudent Klavier und später Gesellschafts- und Wirtschaftskommunikation an der Universität der Künste Berlin. Parallel zur Fortsetzung seiner akademischen Bildung an der Wirtschaftsuniversität Wien und an der University of Richmond war er als Regieassistent an der Universität für

Musik und darstellende Kunst in Wien tätig und arbeitet künstlerisch mit digitalen Medien und künstlicher Intelligenz.

SHARE

BACK TO ALL PINGS

YOUR STORY STARTS HERE

This was some of our work. If you liked it, that's great. Maybe we should work together?

JOIN US

BOOK US