

Künstliche Intelligenz

Einführung, Einsatz-Szenarien sowie
Gefahren und Gegenmaßnahmen

Tim Hoffmann

IT-Trends digital & sicher | 10.04.2025

Berater für Datenschutz und Informationssicherheit

- Wirtschaftswissenschaften an der Universität-GH Essen
- Studien-Schwerpunkte; u. a.
 - » Organisation
 - » Informationsmanagement
- Seit 2002 als Berater mit den Schwerpunkten
 - » Datenschutz und
 - » Informationssicherheit / IT-Sicherheit



Datenschutz

Externe Datenschutzbeauftragung, Audits, E-Learning, Dienstleister-Auditierung etc.



Informationssicherheit

Aufbau eines ISMS bis zur Zertifizierung, Informationssicherheitsbeauftragter etc.



Organisation / Strategie

Beratungsleistungen bei konzeptionellen und strategischen Fragestellungen

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

Agenda



KI: Eine kurze Einführung und Einsatz-Szenarien

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

K

*„KI-System“ ein maschinengestütztes System, das für einen in unterschiedlichem Grade **autonomen Betrieb** ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können
(KI-Verordnung)*

- Forschungsgebiet der Informatik: (noch) keine abschließende Definition
- KI kann aus Daten lernen und ihre Leistung verbessern
- bisherige algorithmenbasierte Systeme basieren auf festgelegten Regeln
- Unterscheidung:
 - » Generative KI: erzeugt originelle Daten (Text, Bilder oder Musik)
 - » Diskriminative KI: erkennt Unterschiede zwischen verschiedenen Daten

Szenarien für diskriminative KI

- Qualitätskontrollen (Auswertung großer Datenmengen, Anomalien feststellen)
- Bewertung/Selektion von Kunden & Bewerbern
- Analyse von Programm-Code
- Lieferketten (z. B. Lageroptimierung)
- Unterstützung von Security-Maßnahmen
 - » Erkennung von Cyberangriffen,
 - » Betrugsversuchen (z. B. bei Banken) und
 - » unerwünschten Inhalten (Spam/Phishing)



Szenarien für generative KI

- Erzeugung von Pressemitteilungen / Werbemitteln / Präsentationen / Protokollen
- Individuelle Bilder statt Stock-Fotos
- Übersetzungen
- Erzeugung von Programm-Code
- Chatbots
- Kollaborative Multiagentensysteme



Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

ChatGPT Deutsch

Hier kannst du das [Modell](#) GPT-4o mini von [OpenAI](#) kostenlos und ohne Registrierung nutzen. Der Chat wird über die offizielle [API von OpenAI](#) betrieben, jedoch handelt es sich hierbei nicht um das offizielle Chat-Interface.

[Zum DALL-E 3 Bildgenerator wechseln](#)



Hallo Mensch, ich bin ein KI-Chatbot, Modell GPT-4o mini.



Wie viele Kinder hat die ehemalige Bundeskanzlerin Angela Merkel?



Angela Merkel hat keine leiblichen Kinder. Sie ist jedoch Stiefmutter von zwei Kindern aus der ersten Ehe ihres Mannes, Joachim Sauer. Wenn du weitere Fragen hast, stehe ich dir gerne zur Verfügung!

Schreibe hier deine Nachricht...

Powered by ChatGPT API

Gefahren mittels KI

(eigene Nutzung)

- Schatten-IT durch KI aktuell verstärkt
- Rechtliche Probleme [auch beim Training der KI]
 - » Urheberrechte
 - » Datenschutz
 - » Geschäftsgeheimnisse (eigene / fremde)
 - » Sonstige: Markenrecht, UWG, AGG etc.
- Werden die eingegebenen Daten für Training der KI genutzt?



Gefahren mittels KI

(Nutzung durch Dritte/Fremde)

- Unterstützung bei Schwachstellensuche
- Schnellere Analyse von Patches
(Erhöhung von „Zero Days“?)
- Re-Identifizierung von anonymisierten Daten
- Generierung / Verbesserung /
Individualisierung von Malware
- Manipulation der eingesetzten KI



Gefahren mittels KI

(Nutzung durch Dritte/Fremde)

- Wissenssammlung und -aufbereitung im Rahmen von Cyberangriffen
- Fake-News / Hoaxes
- Professionalisierung bzw. Optimierung von Social Engineering



„soziale Manipulation“

- ist an sich nichts Neues (früher „Trickbetrug“)
- Cyber-Kriminelle verleiten das Opfer auf diese Weise beispielsweise dazu,
 - » vertrauliche Informationen preiszugeben,
 - » Sicherheitsfunktionen auszuhebeln,
 - » Überweisungen zu tätigen oder
 - » Schadsoftware auf dem Gerät zu installieren.



Wie ein cleverer Ferrari-Manager einen Deepfake entlarvte

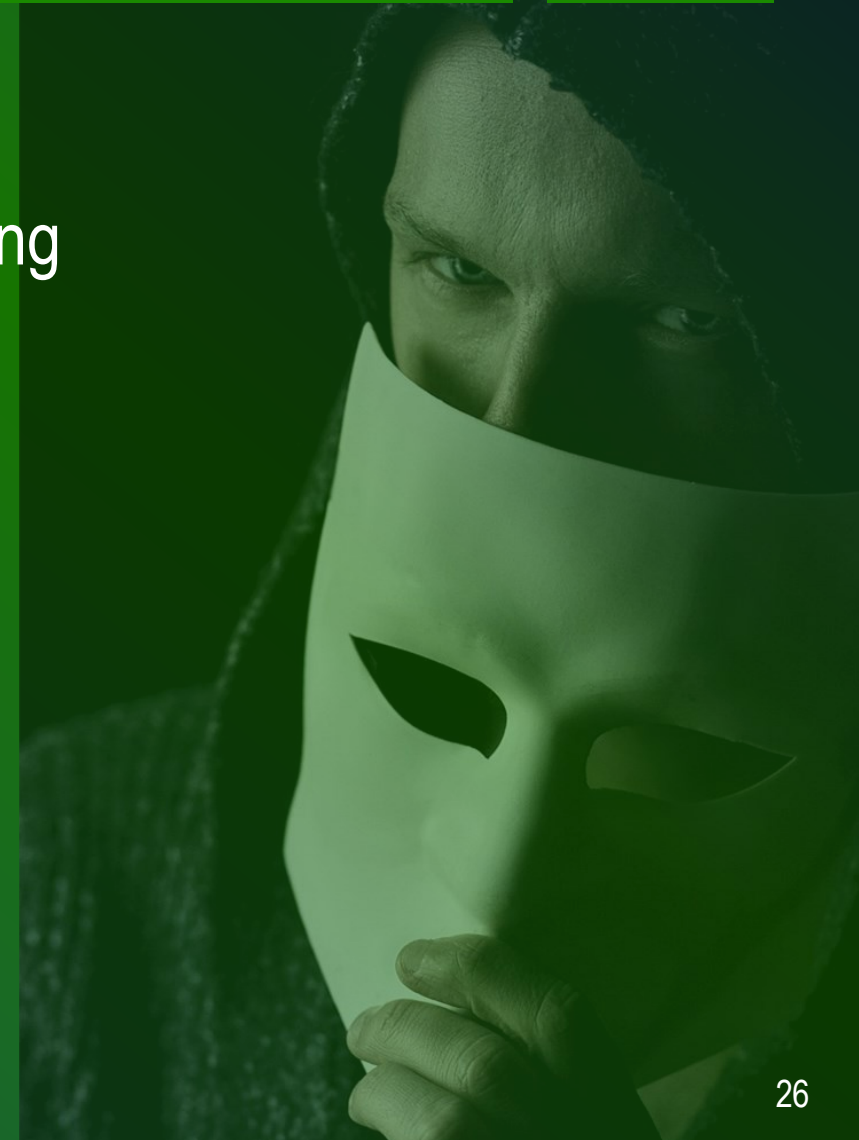
30. Juli, 2024 09:20



<https://www.it-daily.net/shortnews/wie-ein-cleverer-ferrari-manager-einen-deepfake-entlarvte>

KI-Szenarien im Social Engineering

- Text:
 - » Größere Masse bei gleichzeitiger Individualisierung
- Stimme:
 - » Text-zu Sprache (TTS, Text-to-Speech)
 - » Stimmenkonvertierung (VC, Voice Conversion)
- Fotos/Videos
 - » Face Swapping und Face Reenactment



- Text: Verbesserung von Phishing-Mails, CEO-Frauds etc.
- Stimme: Fake-Anrufe
- Foto / Video:
 - » Aufbau von Vertrauen
 - » Fake-Botschaften
 - » Eindringen in Videokonferenzen
 - » Überwindung biometrischer Systeme



Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

■ Drei Dinge

- » Sensibilisierung der Mitarbeiter:innen
- » Sensibilisierung des Managements
- » Sensibilisierung aller

„Die Anbieter und Betreiber von KI-Systemen ergreifen **Maßnahmen**, um nach besten Kräften sicherzustellen, **dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ein ausreichendes Maß an KI-Kompetenz verfügen**, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind.“ (Artikel 4 KI-VO)

■ Drei Dinge

- » Sensibilisierung der Mitarbeiter:innen
- » Sensibilisierung des Managements
- » Sensibilisierung aller

■ Aber auch

- » Risikobewertung (anpassen)
- » Organisation anpassen (z. B. BCM)
- » aufbauende Prozesse (z. B. Codewords, 4-Augen-Prinzip)

■ IT-Sicherheitsmaßnahmen

- » Patchmanagement
- » Aufbau einer resilienten IT-Infrastruktur
- » Verbesserung der Angriffserkennung
- » Multi-Faktor- Authentifizierung
- » eigene Nutzung der KI für Verteidigungsmaßnahmen
(z. B. Erkennung von Bedrohungen und Schwachstellen).

■ Schutz der KI

- » Sicherstellung der Qualität und der Integrität der Trainingsdaten
- » Schutz der Modelle vor Diebstahl und Manipulation
- » Durchführung von umfassenden Tests
- » Auswahl des Modells und des betreibenden Unternehmens
- » etc.

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

- KI betrifft alle Unternehmen, egal ob sie es aktiv nutzen
- KI bietet viele Chancen (insb. im Hinblick auf die Effizienz)
- Vor dem internen Einsatz
 - » Betrachtung rechtlicher Implikationen
 - » Durchführung einer **Risikobewertung**
 - » Umsetzung angemessener Sicherheitsmaßnahmen
- Risikobewertung und **Anpassung des ISMS**
- Belegschaft über die Gefahren **sensibilisieren**



Fragen??

UIMC

*Mancher ertrinkt lieber,
als daß er um Hilfe ruft.
– Wilhelm Busch*

thoffmann@uimc.de

Folgen Sie mir auf LinkedIn:
www.linkedin.com/in/tim-hoffmann-uimc



Diese und weitere Präsentationen unter
<http://update.uimcollege.de>
Gastzugang: web.e#2025*UIMC