DeepFakes, ChatGPT & Co.

Wie KI-Tools Phishing auf ein neues Level heben

Jochen Kurz
Referent Informationssicherheit

bwInfoSec

REAL

OR

FAKE?

Which face Is it real?








Which face Is it real?

https://www.whichfaceisreal.com/

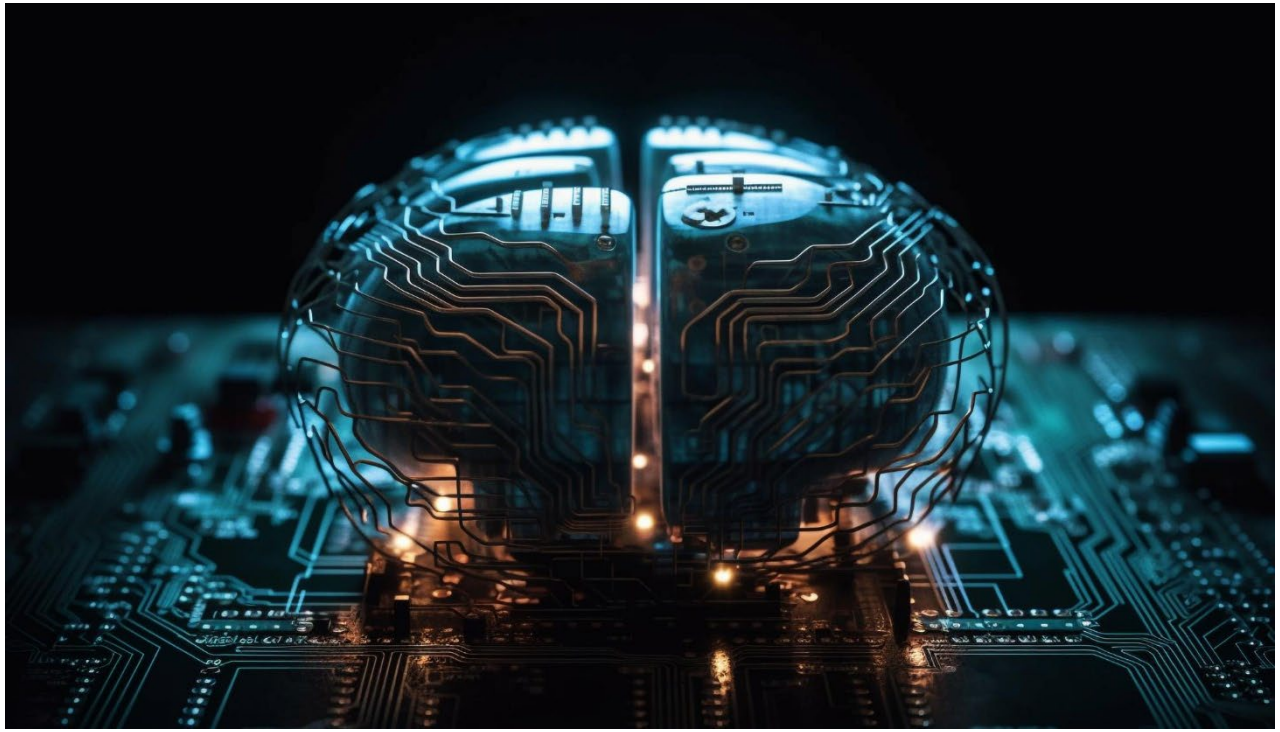
KI
Grundlagen & Methoden

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Was heißt eigentlich KI?

Künstliche Intelligenz (KI) ist ein faszinierendes und sich rasch entwickelndes Feld der Informatik, das darauf abzielt, **Maschinen und Computer** so zu gestalten, dass sie **menschenähnliche Denk- und Lernfähigkeiten** entwickeln.

Im Kern geht es bei KI darum, Algorithmen und Modelle zu entwerfen, die es Computern ermöglichen, Daten zu analysieren, Muster zu erkennen, Schlüsse zu ziehen und Aufgaben zu erledigen, die normalerweise menschliche Intelligenz erfordern.



Wie KI-Tools Phishing auf ein neues Level heben

KI Modelle

KI Systeme unterscheiden sich in Funktionsweise und Lernansatz und werden entsprechend Kategorisiert

Symbolverarbeitung

basiert auf symbolischen Modellen, die auf menschlicher Logik und Regeln beruhen. Sie verwenden Symbole und Manipulationen von Symbolen, um Aufgaben zu lösen

Maschinelles Lernen

Systeme lernen Daten und erkennen Muster, um Vorhersagen zu treffen oder Aufgaben zu erledigen. Beispiele sind neuronale Netze und Entscheidungsbäume.

Deep Learning

ist eine Unterkategorie des maschinellen Lernens, bei der neuronale Netzwerke mit vielen Schichten verwendet werden. Hat zu großen Fortschritten in Bereichen wie Bilderkennung und natürlicher Sprachverarbeitung geführt.

Reinforcement Learning

eine Art des maschinellen Lernens, bei der ein Agent durch Interaktion mit einer Umgebung belohnt oder bestraft wird, um optimale Handlungsstrategien zu erlernen. Diese Technik findet in Anwendungen wie selbstfahrenden Autos Anwendung.

Wie KI-Tools Phishing auf ein neues Level heben

Deep Lerning / ChatGTP

Was ist ChatGPT?

ChatGPT ist ein fortschrittliches KI-Modell, das auf der GPT (Generative Pre-trained Transformer) Architektur basiert. Es wurde entwickelt, um **menschenähnliche Textinteraktionen** zu erzeugen.

Funktionsweise von ChatGPT

- **Textverständnis** - ChatGPT versteht und analysiert Texteingaben, indem es den Kontext und die Bedeutung hinter den Wörtern interpretiert.
- **Trainingsdaten** - Es wurde mit einer Vielzahl von Textdaten trainiert, die es ihm ermöglichen, auf breites Wissen und Sprachmuster zuzugreifen.
- **Generierung von Antworten** - Basierend auf dem erlernten Wissen und dem Kontext generiert ChatGPT natürliche und relevante Antworten
- **Kontextbezogene Reaktionen** - Durch die Analyse vorheriger Gesprächsverläufe kann ChatGPT kontextbezogene Antworten generieren, die dem Gesprächsverlauf folgen.

Best Practices:

- Vermeiden Sie die Weitergabe sensibler oder persönlicher Informationen an KI-Modelle.
- Verstehen Sie, dass ChatGPT keine menschliche Intelligenz besitzt und falsch interpretieren kann (oft werden Informationen auch erfunden).

Wie KI-Tools Phishing auf ein neues Level heben

Einsatzbereicht KI (Auszug)

Gesundheitswesen

Diagnoseunterstützung bei medizinischen Bildgebungsverfahren, Vorhersage von Krankheitsausbrüchen und epidemiologischer Überwachung; Personalisierte Medizin und Behandlungspläne auf Grundlage von Patientendaten.

E-Commerce

Produktempfehlungen und Personalisierung von Einkaufserlebnissen. Preisgestaltungsoptimierung und Lagerbestandsverwaltung. Chatbots für den Kundenservice und die Benutzerinteraktion.

Bild- und Sprachverarbeitung

Gesichtserkennung für Sicherheit und Zugangskontrolle. Übersetzungen und Sprachverarbeitung für globale Kommunikation. Automatische Untertitelung und Bilderkennung in sozialen Medien.

Umweltschutz und Nachhaltigkeit

Überwachung von Umweltauswirkungen und -veränderungen. Ressourceneffizienz und Energiemanagement. Frühzeitige Warnsysteme für Naturkatastrophen.

Forschung und Wissenschaft

Analyse großer Datensätze und Mustererkennung. Genomik und Arzneimittelforschung. Klimamodellierung und simulationsgestützte Forschung.

Wie KI-Tools Phishing auf ein neues Level heben

Neue Angriffsszenarien mit KI

Automatisierte Phishing-Angriffe
Cyberkriminelle verwenden KI, um personalisierte Phishing-E-Mails und Nachrichten zu generieren, die schwerer zu erkennen sind.
KI kann E-Mail-Inhalte und Absenderinformationen so anpassen, dass sie vertrauenswürdig erscheinen.
KI kann gefälschte Websites erstellen, die echten Websites ähneln, um Anmeldedaten zu stehlen.

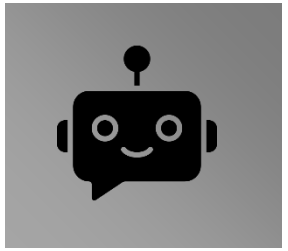
Erfassung und Analyse großer Datenmengen
KI kann große Datenmengen analysieren, um persönliche Informationen und Verhaltensmuster von Anwendern zu extrahieren.
Diese Informationen können für gezielte Angriffe und Identitätsdiebstahl verwendet werden.

Automatisierte Social-Engineering-Angriffe
KI kann soziale Profile und öffentlich verfügbare Informationen nutzen, um überzeugende Phishing-Nachrichten zu erstellen (und E-Mail, Passwortlisten, etc.). Diese Nachrichten können auf persönlichen Informationen oder Interessen der Anwender basieren.

Wie KI-Tools Phishing auf ein neues Level heben

Neue Angriffsszenarien mit KI

Deep-Fakes und gefälschte Medien
KI ermöglicht die Erstellung realistischer gefälschter Audio- und Videodateien, um Anwender zu täuschen. Cyberkriminelle können gefälschte Stimmen oder Videos von vertrauenswürdigen Personen verwenden, um Betrug zu begehen.



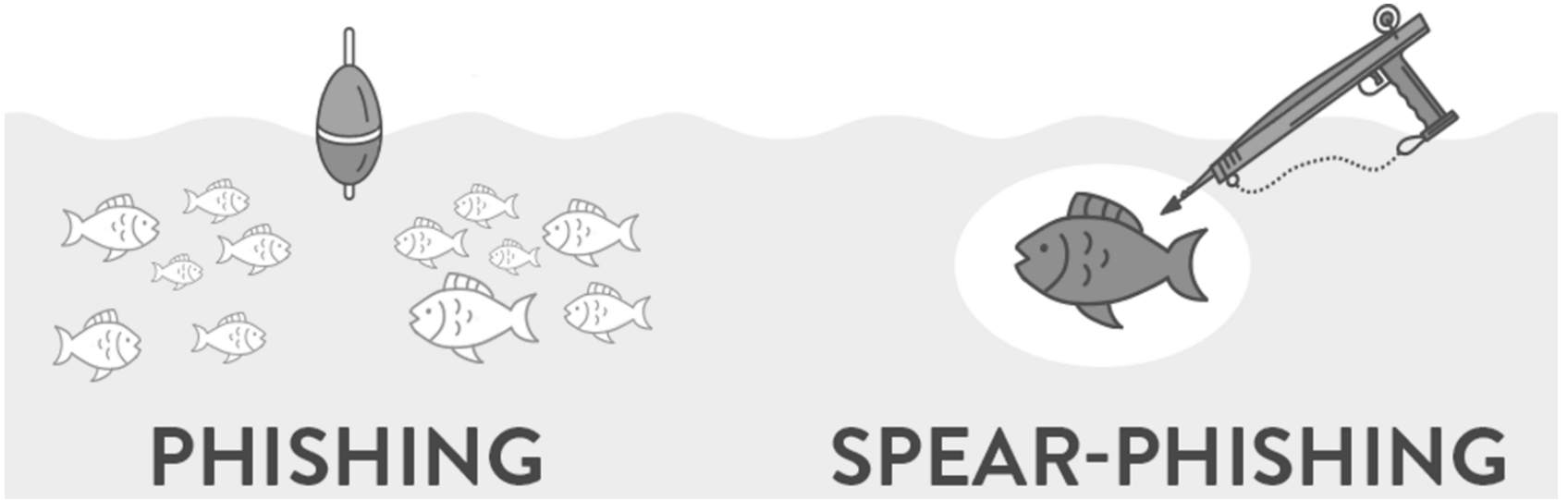
Betrügerische Chatbots
KI-basierte Chatbots können sich als legitime Kundenservice-Agenten oder Kollegen ausgeben, um Informationen von Anwendern zu stehlen.
Sie können auch zur Verbreitung von Malware und zur Verschleierung von Angriffen eingesetzt werden.

Zero-Day-Exploits und Schwachstellenanalyse
KI kann Schwachstellen in Software und Systemen erkennen und automatisch Angriffe entwickeln, bevor Sicherheitspatches verfügbar sind. Dies ermöglicht es Cyberkriminellen, Systeme anzugreifen, bevor Sicherheitsmaßnahmen ergriffen werden können.

Phishing @ Scale

Informationssicherheit im Alltag - Angriffstaktiken

Angriffsmethoden



Spear-Phishing

Spezifischere Form von Phishing.

Hierbei zielt ein Angreifer auf ein einzelnes Opfer oder eine definierte Gruppe von Opfern ab.

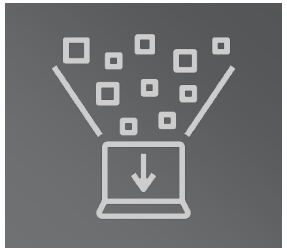
Der Angreifer führt im Vorfeld zumeist eine gründliche Recherche über das Angriffsziel durch, um personalisierte E-Mails oder Nachrichten erstellen zu können, und somit die Glaubwürdigkeit dieser Nachricht stärkt.

- Zumeist über E-Mail
- Nutzung bekannter Absender (Familie, Freunde, Kollegen, Vorgesetzte)
- Sehr oft relevante Inhalte
- Zum Teil mit Inhalten aus alten Daten-Leaks
- Nutzer wird auf Login-Seite geleitet, um dort seine Daten einzugeben

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Phishing mit KI

Spear-Phishing-Angriffen mit KI skalieren und effektiver durchzuführen.



Datensammlung und Profilerstellung
KI kann Informationen über potenzielle Ziele aus öffentlich verfügbaren Quellen, sozialen Medien und anderen Online-Plattformen extrahieren. Diese Informationen können zur Erstellung detaillierter Profile verwendet werden, um personalisierte Phishing-Nachrichten zu erstellen.

E-Mail-Generierung
KI kann verwendet werden, um überzeugende E-Mails zu erstellen, die auf den Interessen, Aktivitäten und Gewohnheiten der Zielpersonen basieren. Dies erhöht die Wahrscheinlichkeit, dass die Opfer auf die Phishing-Nachrichten reagieren.

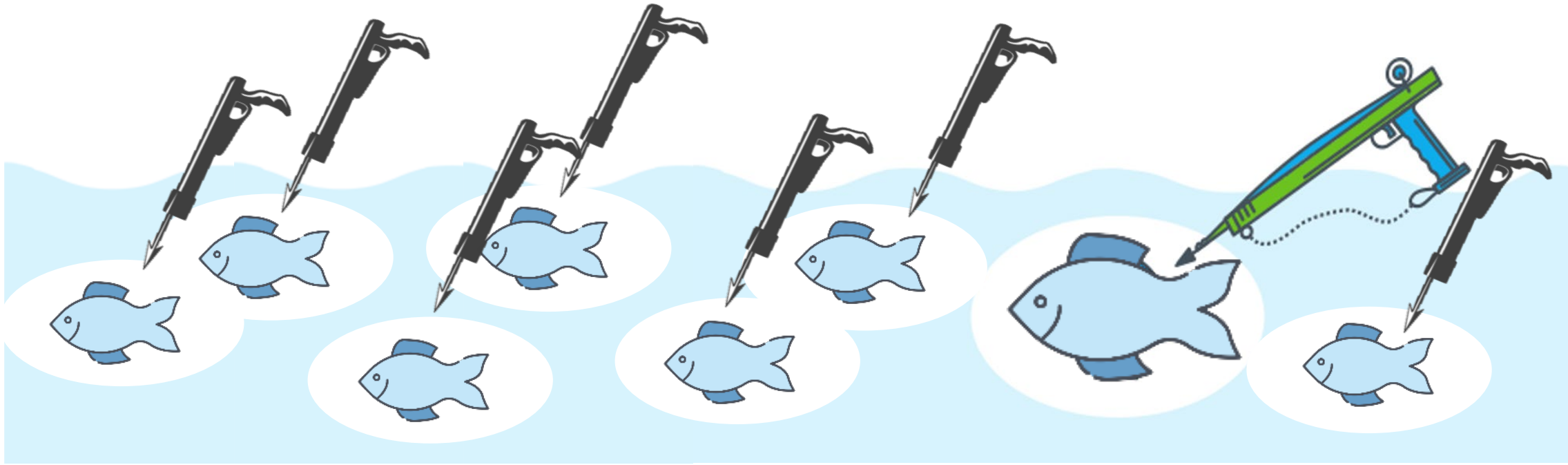
Absenderidentität und Spoofing
KI kann dazu verwendet werden, gefälschte Absenderadressen und Domänen zu erstellen, um den Eindruck zu erwecken, dass die E-Mails von vertrauenswürdigen Quellen stammen.

*Künstliche Intelligenz (KI) kann aus öffentlichen Informationen wie **Projektbeschreibungen im öffentlichen Dienst** passende Spear-Phishing-Kampagnen erstellen.*

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Spear Phsing mit KI

KI ermöglicht die Skalierung von Spear-Phishing Attacken ohne Aufwand für den Angreifen



DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Erkennen von KI-Spear-Phishing Kampagnen

Das Erkennen von KI-basierten Spear-Phishing-Kampagnen erfordert Achtsamkeit und das Befolgen bewährter Praktiken.

- **Überprüfen Sie die Absenderadresse**
  Schauen Sie sich die E-Mail- oder Nachrichtenadresse des Absenders genau an. Achten Sie auf Rechtschreibfehler oder abweichende Domainnamen.
  Nutzt Ihre Organisation eine Kennung von Externen Mails, so darf diese niemals bei Mails von internen Absendern erscheinen!

- **Kritische Prüfung von E-Mails**
  Seien Sie skeptisch gegenüber E-Mails, die ungewöhnliche oder dringende Anfragen stellen. Überprüfen Sie den Inhalt sorgfältig, insbesondere wenn es um Geldtransfers, persönliche Informationen oder vertrauliche Daten geht.

- **Überprüfen Sie den Inhalt**
  Achten Sie auf Rechtschreib- und Grammatikfehler, die in einer echten Kommunikation ungewöhnlich wären. KI kann manchmal Schwierigkeiten bei der Erstellung fehlerfreier Texte haben.

- **Prüfen Sie URLs und Links**
  Vorsicht bei Links in E-Mails. Überprüfen Sie die URL, indem Sie den Mauszeiger darüber bewegen, ohne darauf zu klicken. Achten Sie auf Abweichungen von der erwarteten URL.

- **Verifizieren Sie Anfragen**
  Wenn Sie aufgefordert werden, Geld zu überweisen oder vertrauliche Informationen preiszugeben, nehmen Sie Kontakt zur betreffenden Person oder Organisation auf, um die Anfrage zu verifizieren. Verwenden Sie dazu die offiziellen Kontaktinformationen, nicht die in der fraglichen Nachricht angegebenen.

Social Engineering

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Social Engineering

Social Engineering mit KI

01 Daten-aggregation

Sammeln von öffentlich verfügbare Informationen aus sozialen Medien, Foren, Blogs und anderen Online-Quellen. Z.B. persönliche Informationen wie Namen, Geburtstage, Interessen, Hobbys und vieles mehr sein.

02 Profil-bildung

Gesammelte Informationen werden genutzt, um detaillierte Profile von potenziellen Zielpersonen zu erstellen. Dies kann dazu beitragen, deren Interessen, Gewohnheiten und Verhaltensweisen besser zu verstehen.

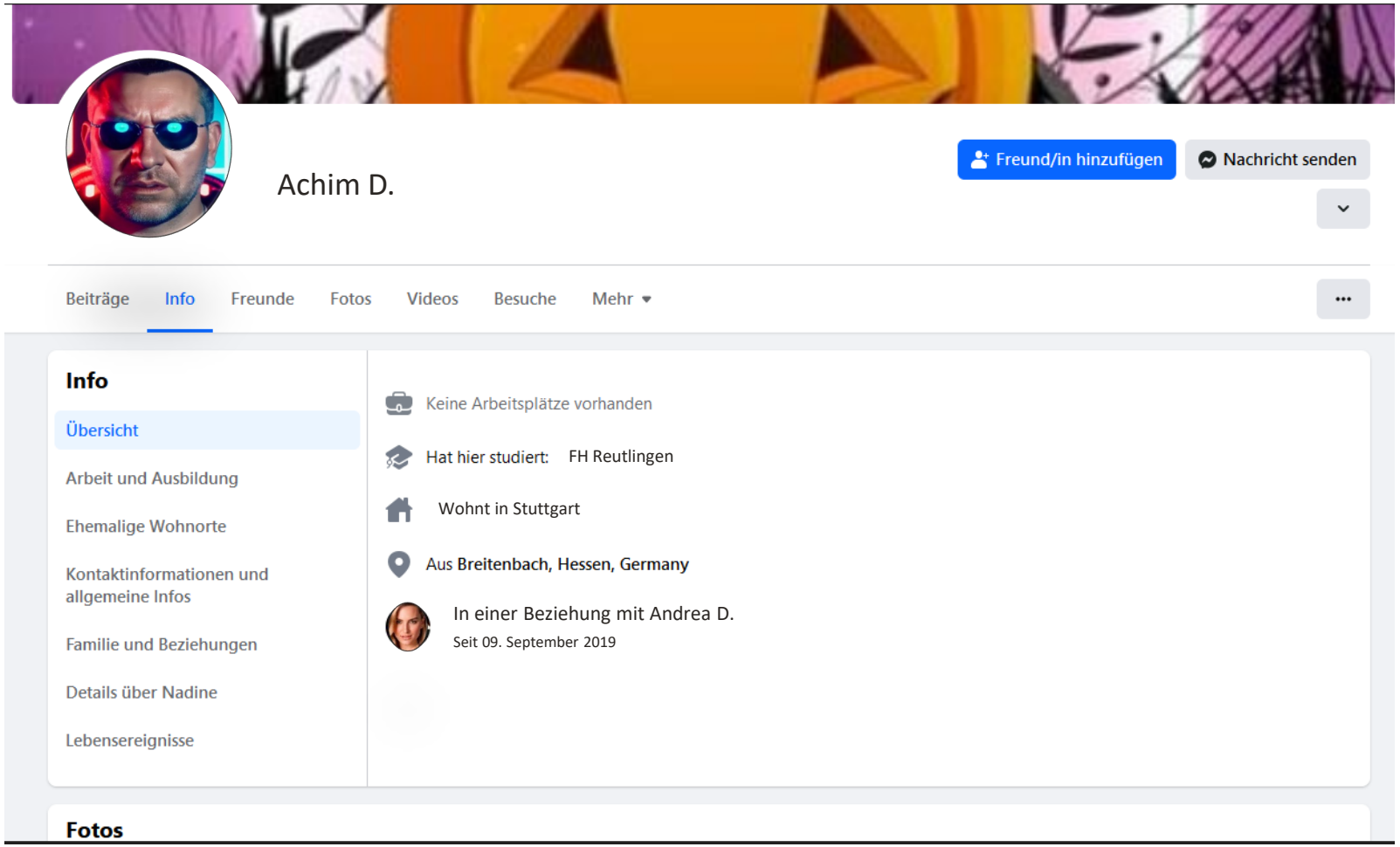
03 Erzeugen von Passwortlisten

Gesammelte Informationen werden genutzt, Liste möglicher E-Mails, User-Namen und Passwörter zu erstellen.

Werden öffentlich verfügbare Informationen für die eigenen Passwörter genutzt, so können diese mittels KI „erraten“ werden

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Social Engineering - Beispiel



Daten-Aggregation

- Achim ist am 25.10.1989 geboren.
- Seine Frau heiß Andrea.
- Sie haben am 09.09.2019 geheiratet.
- Ihre Katzen heißen Hans und Franz und sind in 2015 geboren.
- Er ist Fan des Fc Köln.

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Social Engineering - Beispiel Brute Force Angriffe

Potentielle Passwörter

- Achim2510892
- Andrea0909193
- HansFranz20154
- 25081989Achim
- 09092019Andrea
- HansUndFranz15
- A251089H
- A090919
- A251089H@
- A090919#
- HUndF2015!
- Achim1989$

Potentielle Nutzernamen

- Achim1989Andrea2019
- AchimAndreaCats2015
- FC_KoelnFanAchi
- …

Potentielle Mail-Adressen

- achim.dinder1989 @gmail.com
- andrea.dinder 2019@gmail.com
- Hansfranzcats2015 @gmail.com
- …
- .

Passwörter, Nutzernamen & Co. die aus öffentlich verfügbaren Informationen bestehen können von KI leicht kombiniert werden

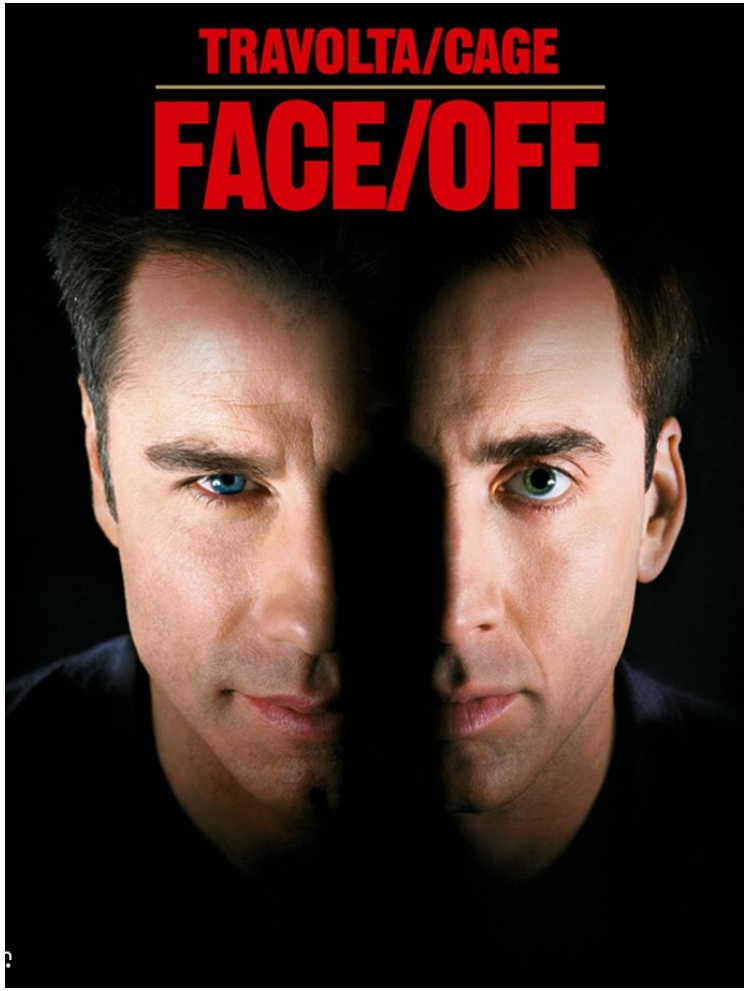
Deep Fake

Wie KI-Tools Phishing auf ein neues Level heben

Deep Fake

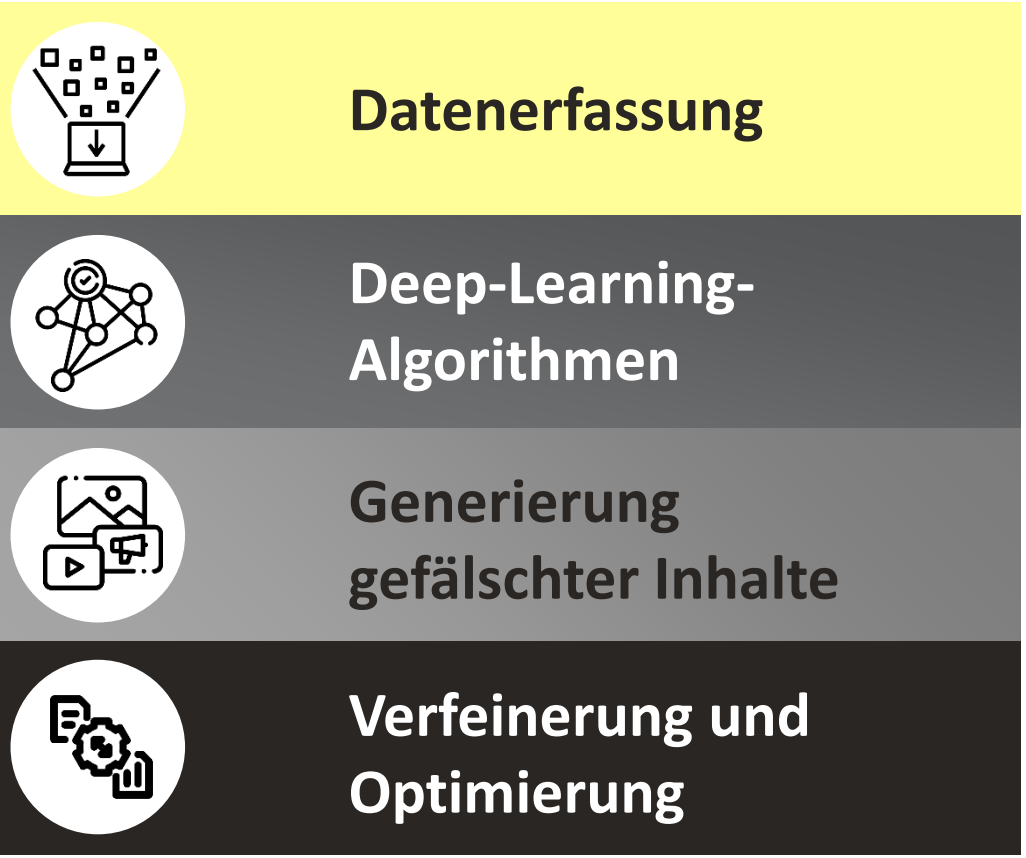
Deep Fake bezieht sich auf die **Manipulation von Medieninhalten** (meistens Videos), indem fortgeschrittene künstliche Intelligenz und Machine-Learning-Algorithmen eingesetzt werden, um irreführende oder gefälschte Inhalte zu erzeugen.

Deep Fakes können für betrügerische oder manipulative Zwecke eingesetzt werden, um Desinformation zu verbreiten oder Personen zu täuschen. Sie können ernsthafte Auswirkungen auf den Ruf und die Glaubwürdigkeit von Personen haben.



Wie KI-Tools Phishing auf ein neues Level heben

Deep Fake – Wie funktioniert es?



Ein (umfangreicher) Datensatz wird benötigt, oft bestehend aus Bildern, Videos und Audioaufnahmen der Zielperson

Mithilfe von DL-Modellen, wie z.B. neuronale Netzwerke, werden diese Daten analysiert und die Merkmale der Zielperson erlernt.

Basierend auf dem erlernten Wissen generiert der Algorithmus täuschend echte Inhalte, indem er das Gesicht und die Stimme der Zielperson in Videos oder Audiodateien einsetzt.

Durch wiederholtes Training und Optimierung des Modells werden die erstellten Inhalte immer realistischer und schwerer von echten Inhalten zu unterscheiden.

Seeing is not believing - anymore

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Deep Fake Trainingsmaterialien

Bilder

(>) 6 Selfies / Portraitfotos
für perfekte KI- Bilder nötig

Video & Voice

Wenige Minuten Ton/Video
für Deep Fakes

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

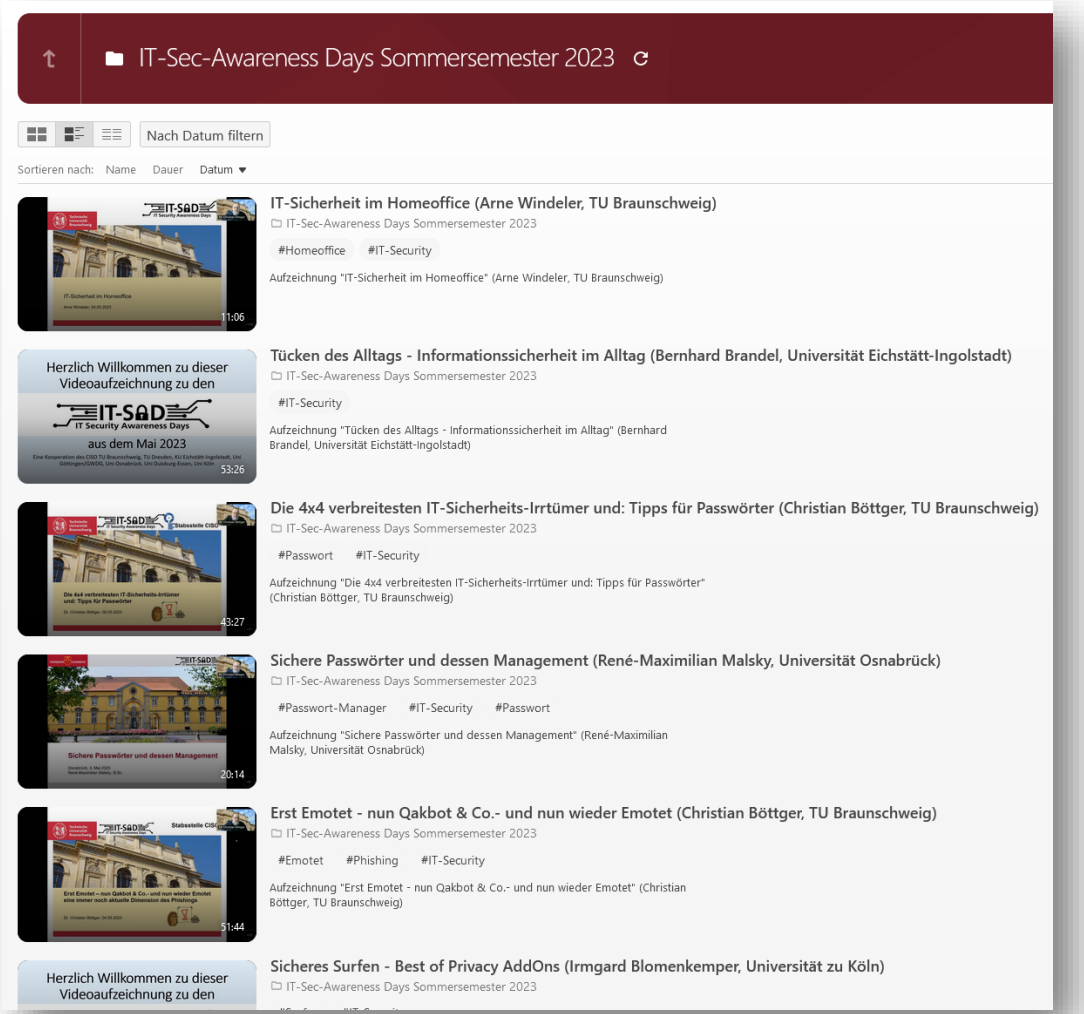
Deep Fake Trainingsmaterialien

Semesteransprachen / Image Videos

Präsident FH-Reutlingen

Kanzlerin FH Furtwangen

Online Konferenzen



Wie KI-Tools Phishing auf ein neues Level heben

Deep Fake – Gefahren und Missbrauch

Deep Fakes können für **betrügerische oder manipulative Zwecke** eingesetzt werden, um Desinformation zu verbreiten oder Personen zu täuschen.
Sie können ernsthafte Auswirkungen auf den Ruf und die Glaubwürdigkeit von Personen haben. ➔ Wahlkampf!

Gefälschte Videos oder Audioaufnahmen von Führungskräften
Angreifer könnten gefälschte Videos oder Audiomaterial erstellen, die vorgeben, von Geschäftsführern oder leitenden Angestellten zu stammen. Diese gefälschten Materialien könnten verwendet werden, um Mitarbeiter zu täuschen und vertrauliche Informationen oder Zugangsdaten preiszugeben.

Imitation von Mitarbeitern oder Vertrauenspersonen
Phisher könnten versuchen, Videos oder Audioaufnahmen zu erstellen, die vorgeben, von bekannten Mitarbeitern oder vertrauenswürdigen Quellen innerhalb eines Unternehmens zu stammen. Diese könnten genutzt werden, um Zugang zu Informationen zu erhalten oder um gefälschte Anfragen für Zahlungen oder vertrauliche Daten zu

Schutzmaßnahmen:
Sensibilisierung für die Existenz von Deep Fakes und die potenziellen Risiken.

Überprüfen von Quellen und Inhalten, insbesondere bei fragwürdigen oder unbestätigten Informationen.

Tipp: Small Talk!

Deep Fakes demonstrieren die Macht und Herausforderungen fortschrittlicher Technologien. Ein kritisches Bewusstsein und gesunde Skepsis bei der Medienkonsumption sind entscheidend, um sich vor den potenziellen Auswirkungen zu schützen

Wie KI-Tools Phishing auf ein neues Level heben

Deep Fake (Voice) – Beispiel

KI

Deepfake eines Finanzchefs ermöglicht Millionenbetrug

Ein Angestellter überwies 23 Millionen Euro an Betrüger, die sich mithilfe von Deepfakes in einer Videokonferenz als seine Vorgesetzten ausgegeben haben.

in Pocket speichern | merken

4. Februar 2024, 12:34 Uhr, Andreas Donath

(Bild: KI-generiert mit Midjourney)

Gefaktes Bild einer Fake-Videokonferenz (Symbolbild)

Die Behörden in Hongkong haben einen ausgeklügelten Betrug mit Deepfakes aufgedeckt, bei dem ein Angestellter dazu gebracht wurde, 200 Millionen Hongkong-Dollar (umgerechnet etwa 23 Millionen Euro) an Betrüger zu überweisen, die sich als Führungskräfte des Unternehmens ausgaben, berichtet CNN.

Mehrstufiger Angriff

Nachricht über auszuführende Transaktion

↓

VC mit 3 vertrauten „Kollegen“

↓

Ausführen der Überweisung = 23 Mio €

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

KI kombinierte Angriffe – Bsp. Enkeltrick

Der sogenannte "**Enkeltrick**" und ähnliche Betrugsmaschen können durch den Einsatz von Künstlicher Intelligenz (KI) noch gefährlicher werden.

Stimmenimitation
KI kann verwendet werden, um die Stimmen von Verwandten oder Freunden genau nachzuahmen, wodurch die Betrüger noch überzeugender wirken.

Personalisierung
Mit KI können Betrüger detaillierte Informationen über die Zielpersonen sammeln und personalisierte Betrugsanrufe oder Nachrichten erstellen. Dies kann dazu führen, dass die Opfer glauben, dass der Betrüger tatsächlich ein Verwandter ist.

Echtzeitanalyse von Social-Media-Profilen
KI kann Echtzeitanalysen von Social-Media-Profilen durchführen, um aktuelle Informationen über die Zielpersonen zu erhalten und in den Betrug einzubeziehen.

Effizienzsteigerung
KI kann dazu beitragen, dass Betrüger effizienter arbeiten, indem sie potenzielle Ziele identifizieren, Anrufe tätigen und Informationen sammeln. Dies kann die Anzahl der gleichzeitig durchgeführten Betrugsversuche erhöhen.

Automatisierte Kommunikation
KI-basierte Chatbots könnten eingesetzt werden, um mit den Opfern zu kommunizieren und den Betrug durchzuführen. Dies ermöglicht es Betrügern, mehrere Opfer gleichzeitig anzusprechen.

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Schutz vor KI kombinierte Angriffe

Verifizierung von Identität

Wenn Sie kontaktiert werden und um Geld oder persönliche Informationen gebeten werden, überprüfen Sie die Identität des Anrufers oder Absenders, bevor Sie reagieren. Rufen Sie beispielsweise an, um die Informationen zu verifizieren.

Vertrauenswürdige Kontaktmethoden

Verwenden Sie vertrauenswürdige Kontakt-methoden, um die Identität von Personen zu überprüfen. Wenn Sie Zweifel haben, verwenden Sie die offiziellen Kontaktdaten, um die Person zu erreichen.

Überprüfen Sie ungewöhnliche Anfragen

Seien Sie besonders vorsichtig bei ungewöhnlichen oder dringenden Anfragen, insbesondere wenn es um finanzielle Angelegenheiten geht. Nehmen Sie sich Zeit, um die Anfrage zu prüfen und zu überlegen.

Grenzen setzen

Seien Sie bereit, "Nein" zu sagen und Grenzen zu setzen, wenn Sie sich unwohl oder unsicher fühlen. Vermeiden Sie es, persönliche oder finanzielle Informationen weiterzugeben

Erkennen & Schützen

DeepFakes, ChatGPT & Co – Neue Gefahren durch KI

Final Thoughts: Schutz vor KI-basierten Phishing- und Angriffsgefahren

Sensibilisierung ist der Schlüssel

- **Erkennung und Verifizierung**
  Überprüfen Sie immer Absenderinformationen, Links und verdächtige E-Mails.
  Verifizieren Sie Anfragen, bevor Sie reagieren.
- **Vorsicht bei ungewöhnlichen Anfragen**
  Seien Sie skeptisch bei ungewöhnlichen oder dringenden Anfragen, insbesondere wenn es um Geld oder persönliche Informationen geht.
- **Sicherheitssoftware nutzen**
  Installieren Sie **aktuelle** Sicherheitssoftware, um KI-generierte Phishing-Angriffe zu erkennen und abzuwehren.
- **Melden Sie verdächtige Aktivitäten**
  Wenn Sie einen Verdacht haben, melden Sie verdächtige Phishing-Angriffe an die entsprechenden Stellen.
  Gemeinsam schützen. Jeder kann dazu beitragen, die Sicherheit zu erhöhen und sich vor KI-generierten Angriffen zu schützen.
- **Bleiben Sie auf dem Laufenden**
  Verfolgen Sie die aktuellen Entwicklungen in der KI-basierten Phishing- und Sicherheitslandschaft

Erhöhen Sie Ihr Bewusstsein für KI-generierte Phishing-Angriffe und deren Gefahren.
Seien Sie vorsichtig und achtsam.

Danke für Ihre Aufmerksamkeit.