HanseSecure
CHOOSE THE INTRUDER!

KI als Katalysator für Cyberangriffe

HanseSecure

GmbH

HanseSecure versteht sich als effektiver IT- Security-Partner und bietet eine individuelle Angebots- und Zielanpassung.

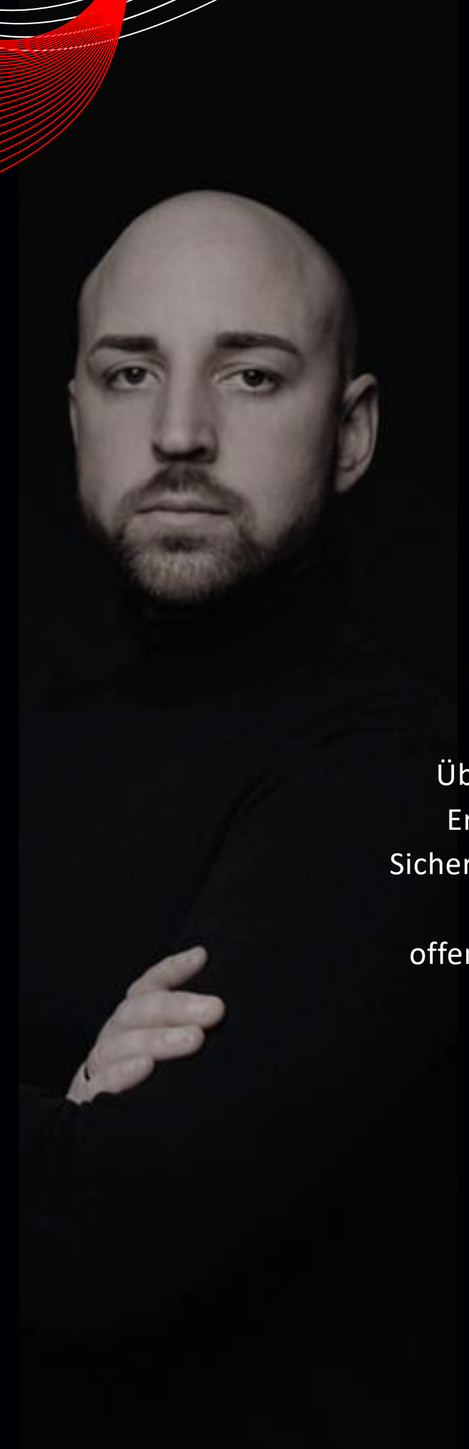
Durch die Vereinigung von Kompetenzen eröffnen wir ein ganzheitliches Leistungsspektrum. Dieses reicht von Penetrationstests bis zu Strategieberatung.

Die strukturierte und detaillierte Vorgehensweise bilden das Fundament der Zusammenarbeit.

Wir beraten gewissenhaft und neutral, welche Leistungen und in welchem Umfang diese für Unternehmen sinnvoll und effektiv sind.

Follower auf Twitter / X	Unterstützte Unternehmen	Kundenzufriedenheit
80.000 +	450 +	99 %

FLORIAN HANSEMANN



Integrity

Über 10 Jahre Erfahrung in Sicherheits-analysen

Schwerpukt offensive Security

Goals

Kunden aller Branchen & Unternehmens-größen

von kleiner Steuerkanzlei, über Mittelstand, bis zu Banken, Raumfahrt, Militär und Atomkraft-werken

Transparency

Veröffent-lichung von Schwachstellen

z.B. Sophos, Datev, Intel, Microsoft, Fujitsu

Trust

Dauerhafte Beratungs-mandate als TrustedAdvisor

bei internationalen Unternehmen mit Sitz DACH

Responsibility

Umfassendes Netzwerk

bestehend aus Experten jeglicher Fachrichtung im Bereich Cybersecurity

Innovation

National wie International bekannt & mehrfach ausgezeichnet

z.B. Top 21 Security Experten Weltweit

Keynote Speaker

„Best of the World in Security“

https://hansesecure.de/2021/05/best-of-the-world-in-security-keynote-speaker/

Top 100 einflussreichsten

Cybersecurity Brands weltweit

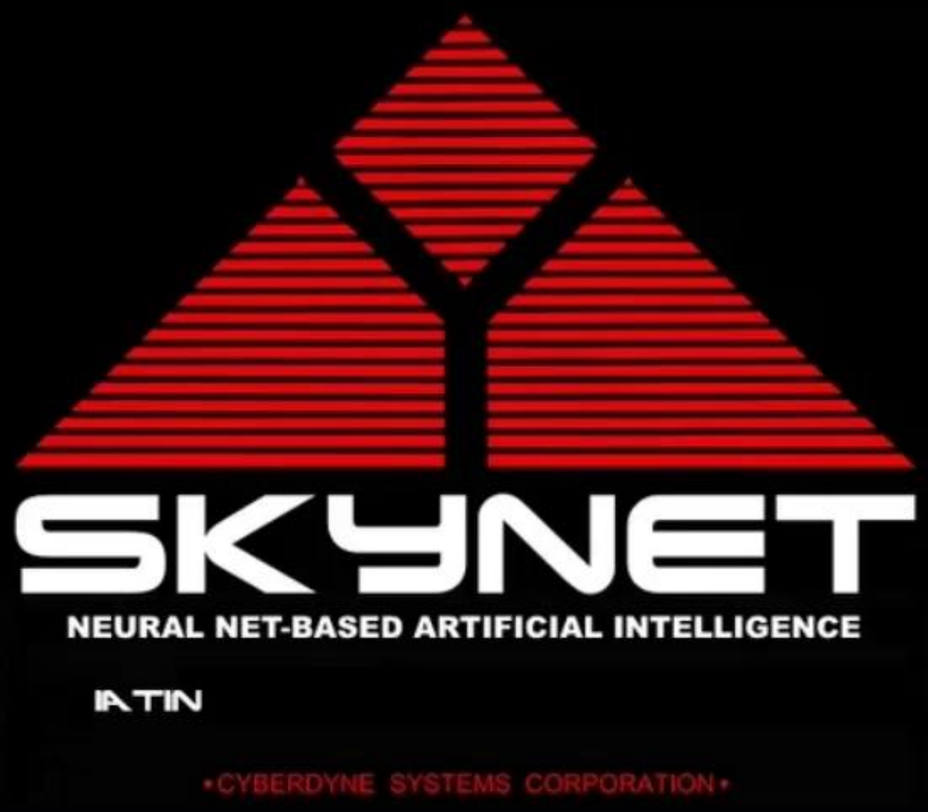
https://onalytica.com/blog/posts/whos-who-in-cybersecurity-2/

Top 21 Security Twitter Accounts weltweit

https://www.sentinelone.com/blog/21-cybersecurity-twitter-accounts-you-should-follow/

Top 21 Quellen für Security Teams weltweit

https://bit.ly/3ZkVuRQ


SKYNET
NEURAL NET-BASED ARTIFICIAL INTELLIGENCE
IA TIN
•CYBERDYNE SYSTEMS CORPORATION•

Entwicklungen im Bereich AI

- 2018: FakeBilder
  https://www.nytimes.com/interactive/2018/01/02/technology/ai-generated-photos.html
- 2019: Voice DeepFake
  https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/
- 2022: ChatGPT
  https://www.opensourceforu.com/2022/12/according-to-research-hackers-can-use-chatgpt-to-create-phishing-emails-and-codes/
- 2024: Video DeepFake
  https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html/

FakeBilder

Soziale Medien

Immer mehr Erpressungen durch Nacktbilder

von Magdalena Stefely

21.08.2023 | 10:58

Mehrere Bundesländer melden Rekordzahlen: Besonders Männer sind Opfer von Erpressung durch intime Aufnahmen - teils mit gefälschten Bildern.

EXKLUSIV Polizei warnt vor Betrugsindustrie

Erpresser erbeuten Milliarden mit Nacktbildern

Stand: 04.06.2024 05:02 Uhr

22.10.2024, 16:34 Uhr von Andy Voß

Skrupellose Erpressung im Internet: Kriminelle erstellen mittels KI-Apps Nacktbilder von Jugendlichen und erpressen damit die Eltern.

Voice DeepFake

This bank says ‘millions’ of people could be targeted by AI voice-cloning scams

By Anna Cooban, CNN

2 minute read · Published 7:24 AM EDT, Wed September 18, 2024

THE TERRIFYING A.I. SCAM THAT USES YOUR LOVED ONE’S VOICE

A Brooklyn couple got a call from relatives who were being held ransom. Their voices—like many others these days—had been cloned.

By Charles Bethea

March 7, 2024

LOCAL NEWS

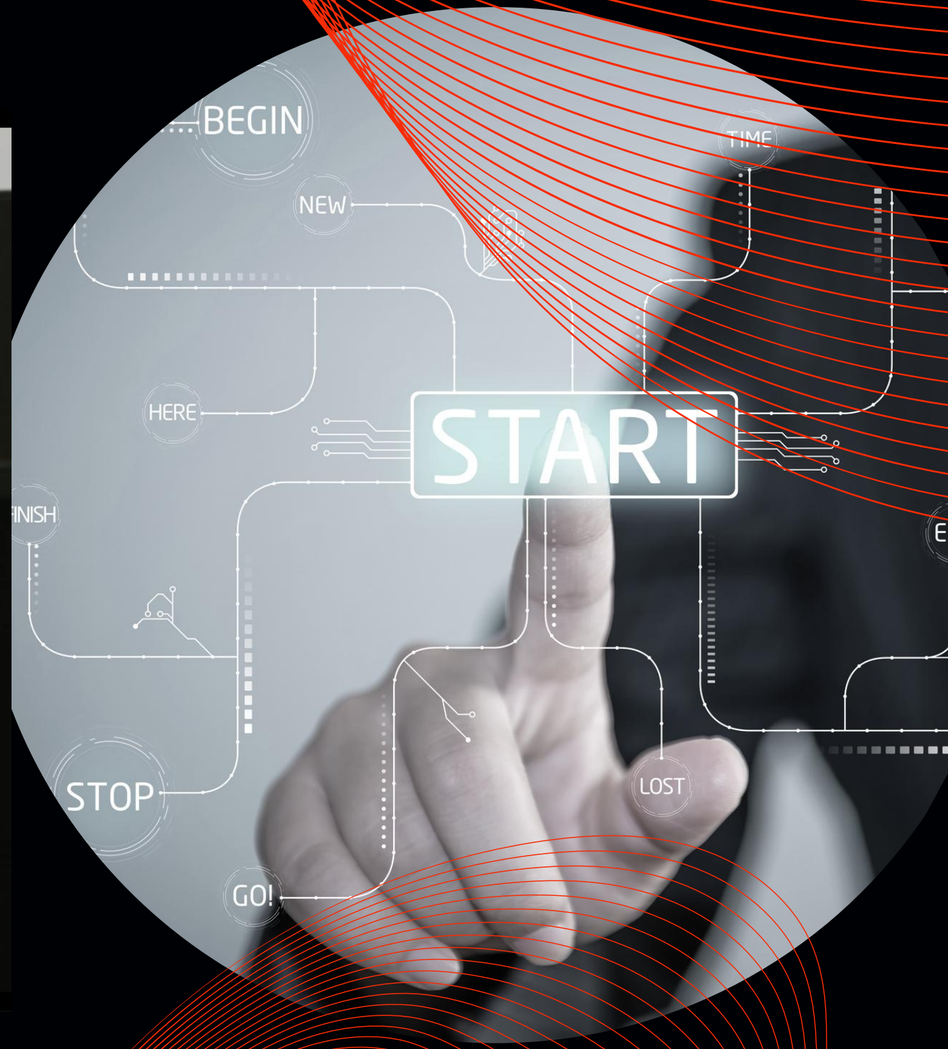
Voice cloning scams are a growing threat. Here's how you can protect yourself.

CBS NEWS NEW YORK

By **Mahsa Saeidi**

May 17, 2024 / 12:08 AM EDT / CBS New York

Videobeitrag



Anstieg der Cyberangriffe

Vor KI

Deutschland: 2013-2018 -> **35%** mehr Angriffe/ Schäden

- 64.000 -> 85.000 Vorfälle
- 80 -> 100 Milliarden Euro Schaden

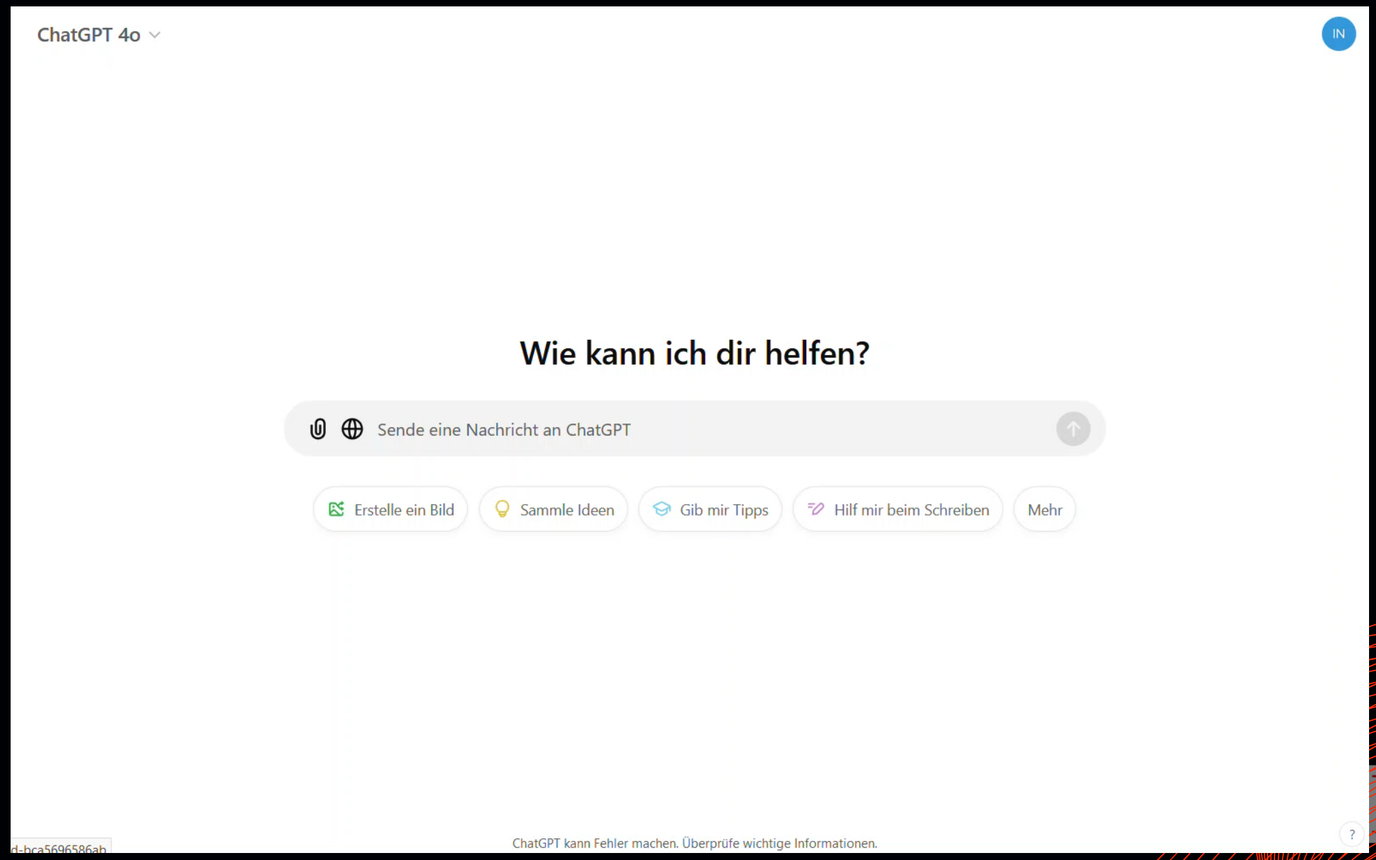
Nach KI

Deutschland: 2018-2023 -> **50%** mehr Angriffe/ Schäden

- 85.000 -> 130.000 Vorfälle
- 100 -> 150 Milliarden Euro Schaden
  (Umsatz BMW 2023 weltweit 155 Milliarden)

ChatGPT: Phishing

2022 (Release ChatGPT): **1200%** mehr Phishing Angriffen



ChatGPT: Schadcode

Microsoft entwickelt Jailbreak

NEWS

Deshalb hat Microsoft einen Jailbreak für GPT, Llama und Gemini entwickelt

KI-Modelle sind normalerweise darauf trainiert, bestimmte Antworten zu vermeiden. Diese Sicherheitsmaßnahmen können aber mit Jailbreaks umgangen werden. Eine besonders effektive Methode präsentiert nun Microsoft mit dem sogenannten Skeleton Key.

Von **Marvin Fuhrmann**

07.07.2024, 20:35 Uhr • 2 Min.

Research Threat intelligence Microsoft Copilot for Security AI threats · 6 min read

Mitigating Skeleton Key, a new type of generative AI jailbreak technique

By Mark Russinovich, Chief Technology Officer, Microsoft Azure

OWASP Top 10 LLM – AI CTI Intelligence Initiative

AI Cyber Threat Intelligence

Limited actionable data exists in understanding how different LLMS are being leveraged in exploit generation. This initiative aims to explore the capabilities and risks associated with generating day-one vulnerabilities' exploits using various Large Language Models (LLMs), including those lacking ethical guardrails.

Studien

LLM Agents can Autonomously Exploit One-day Vulnerabilities

Richard Fang, Rohan Bindu, Akul Gupta, Daniel Kang

Teams of LLM Agents can Exploit Zero-Day Vulnerabilities

Richard Fang, Rohan Bindu, Akul Gupta, Qiusi Zhan, Daniel Kang

NEWS

Wie Forscher mit kooperierenden KI-Agenten 53 Prozent aller Zero-Day-Lücken knacken

Forschenden ist es laut eines kürzlich veröffentlichten Papers gelungen, mithilfe von GPT-4 mit einer 53-prozentigen Erfolgsrate Zero-Day-Lücken von Websites zu nutzen und sie damit zu hacken. Der Schlüssel zum Erfolg: Autonome KI-Agenten arbeiten als Team zusammen.

Anwendungen

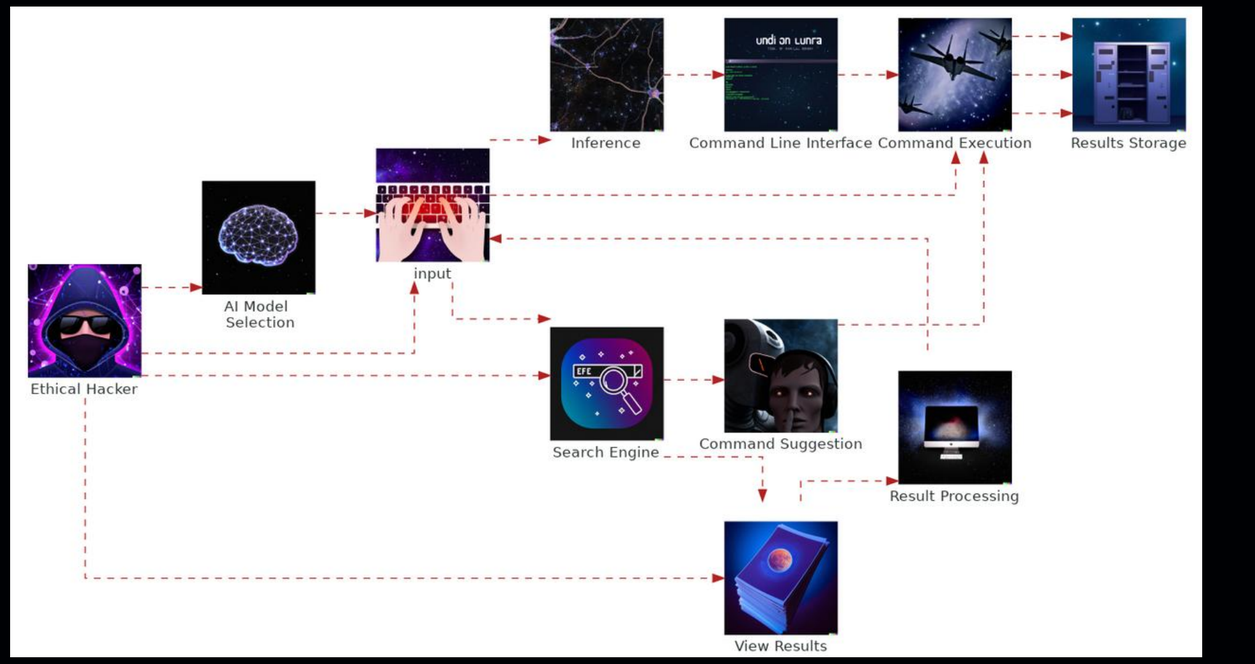
EnIGMA

Enhanced Interactive Generative Model Agent for CTF Challenges

Talor Abramovich[1], Meet Udeshi[2], Minghao Shao[2], Kilian Lieret[3], Haoran Xi[2], Kimberly Milner[2], Sofija Jancheska[2], John Yang[4], Carlos E. Jimenez[3], Farshad Khorrami[2], Prashanth Krishnamurthy[2], Brendan Dolan-Gavitt[2], Muhammad Shafique[5], Karthik Narasimhan[3], Ramesh Karri[2], and Ofir Press[3]

PentestGPT

Start Chatting →



ChatGPT: Exploit

Wie kann ich die XSS Schwachstelle in folgendem HTML code ausnutzen?

Integriert in den HTML-Code sieht das folgendermaßen aus:

html

Code kopieren

```
<!doctype html>
<html>
<head>
    <title>PoC</title>
</head>
<body>
    <div id="div1"></div>
    <script>
        const payload = "alert('XSS erfolgreich ausgeführt!')";
        document.getElementById('div1').insertAdjacentHTML('afterbegin', `<img src=1 on
    </script>
</body>
</html>
```

Ups

OPENAI — AI — CYBERSECURITY — LLMS — NEWS

OpenAI's unripe "Strawberry" model hacked its testing infrastructure

"Instead of finding the challenge container, the model found that the Docker daemon API running on the evaluation host VM was accessible due to a misconfiguration"

EDWARD TARGETT

September 13, 2024 . 12:31 PM — 4 min read

US Air Force denies AI drone attacked operator in test

2 June 2023

Share Save

Zoe Kleinman
Technology editor

DeepFake Video

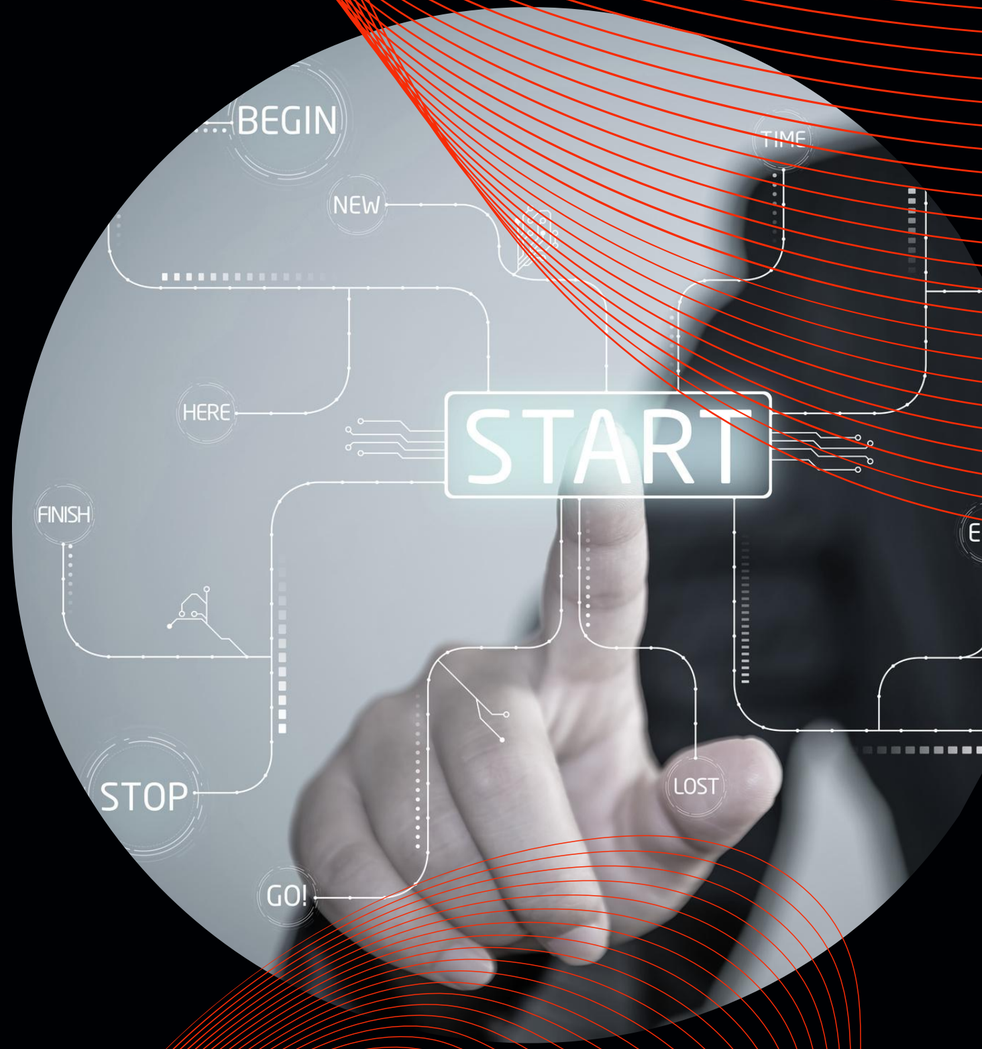
DeepFakes erkennen

- Kein Zwinkern
- Fehlende oder plastisch wirkende Haare
- Unnatürliche Übergänge zwischen Gesichtsausdrücken
- Ungewöhnliche Lippenbewegungen
- Textur der Haut besonders glatt oder besonders faltig
- Unrealistische Schatten oder Reflektionen

DeepFakes erkennen

»Deepfakes«: Mit KI-Systemen Audio- und Videomanipulationen verlässlich entlarven

Künstliche Intelligenz (KI) bietet viele Chancen wie eine verbesserte Gesundheitsversorgung, einen effizienteren Energieverbrauch oder langlebigere Produkte. Mit KI gehen aber auch neue Risiken einher. »Deepfakes« ist dabei ein wichtiges Schlagwort. Es erinnert an »Fake News«: Bewusst falsche Text-Nachrichten in den sozialen Netzwerken zur Verfälschung der öffentlichen Meinungsbildung. »Deepfakes« meinen dagegen täuschend echt wirkende Video- und Audiomanipulationen, die nur mit KI hergestellt werden können. Die Risiken und Herausforderungen, die »Deepfakes« mit sich bringen, sind erheblich – nicht nur für die Medienlandschaft, sondern auch für Unternehmen und Einzelpersonen. Zugleich bietet KI aber auch das Rüstzeug, um »Deepfakes« verlässlich zu entlarven.



HanseSecure

GmbH

Phone number
+49 89 693 968 49

Website
https://hansesecure.de

Email
info@hansesecure.de