Wie Deepfakes & Social Engineering Marken angreifen

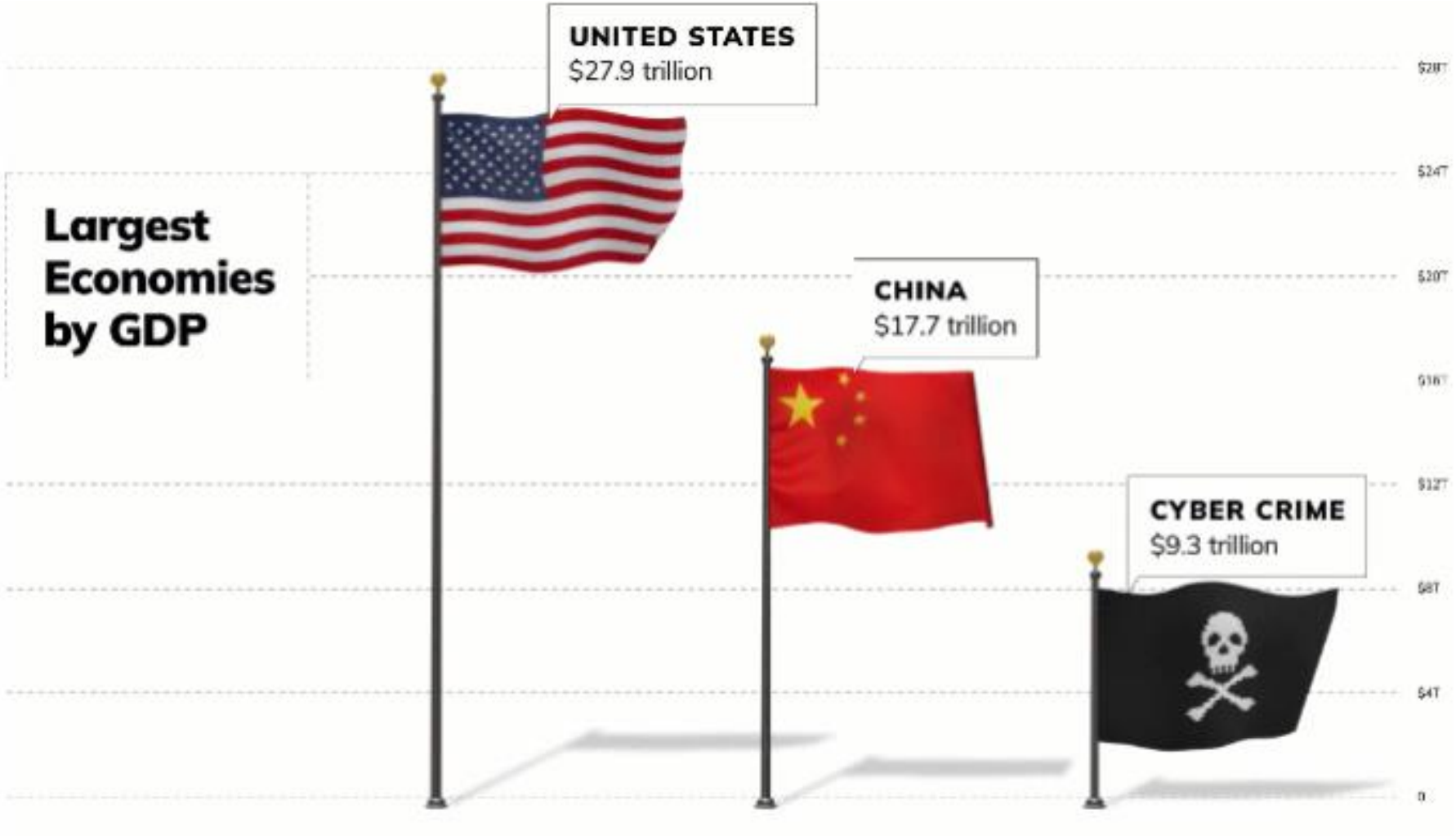
Jill Wick, 17. November 2025

Agenda

InfoGuard
SWISS CYBER SECURITY

1. Einleitung
2. Wieso funktioniert Cybercrime?
3. Social Engineering
4. Kombination von KI und Social Engineering
5. Fragen

Cybercrime:
Die drittgrösste „Wirtschaft“ der Welt

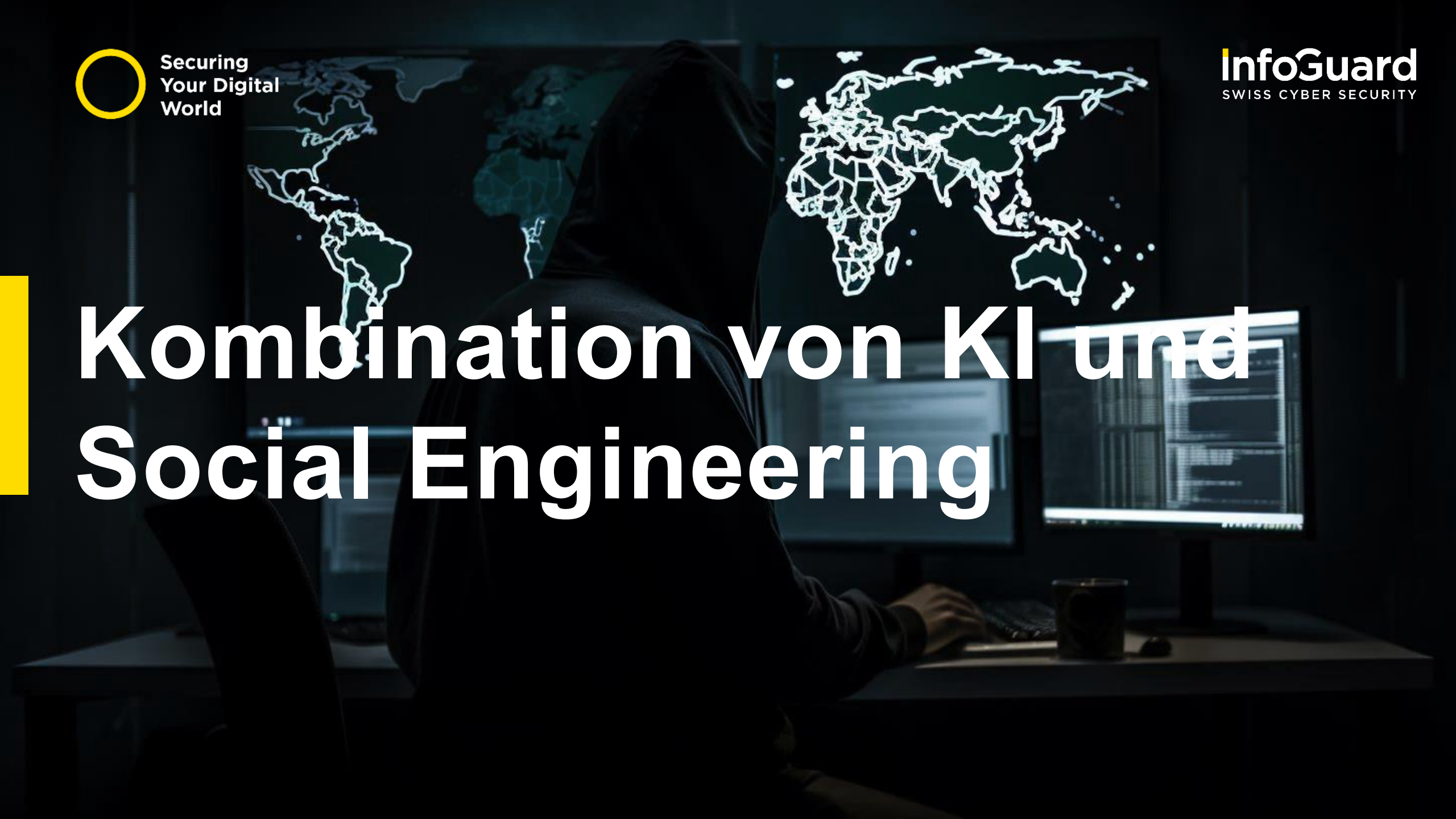



Securing
Your Digital
World
InfoGuard
SWISS CYBER SECURITY
Wieso funktioniert
Cybercrime?

Die Antwort ist: Social Engineering.

Securing
Your Digital
World

InfoGuard
SWISS CYBER SECURITY



Kombination von KI und Social Engineering

Social Engineering

Finance worker pays out $25 million after video call with deepfake ‘chief financial officer’

By Heather Chen and Kathleen Magramo, CNN

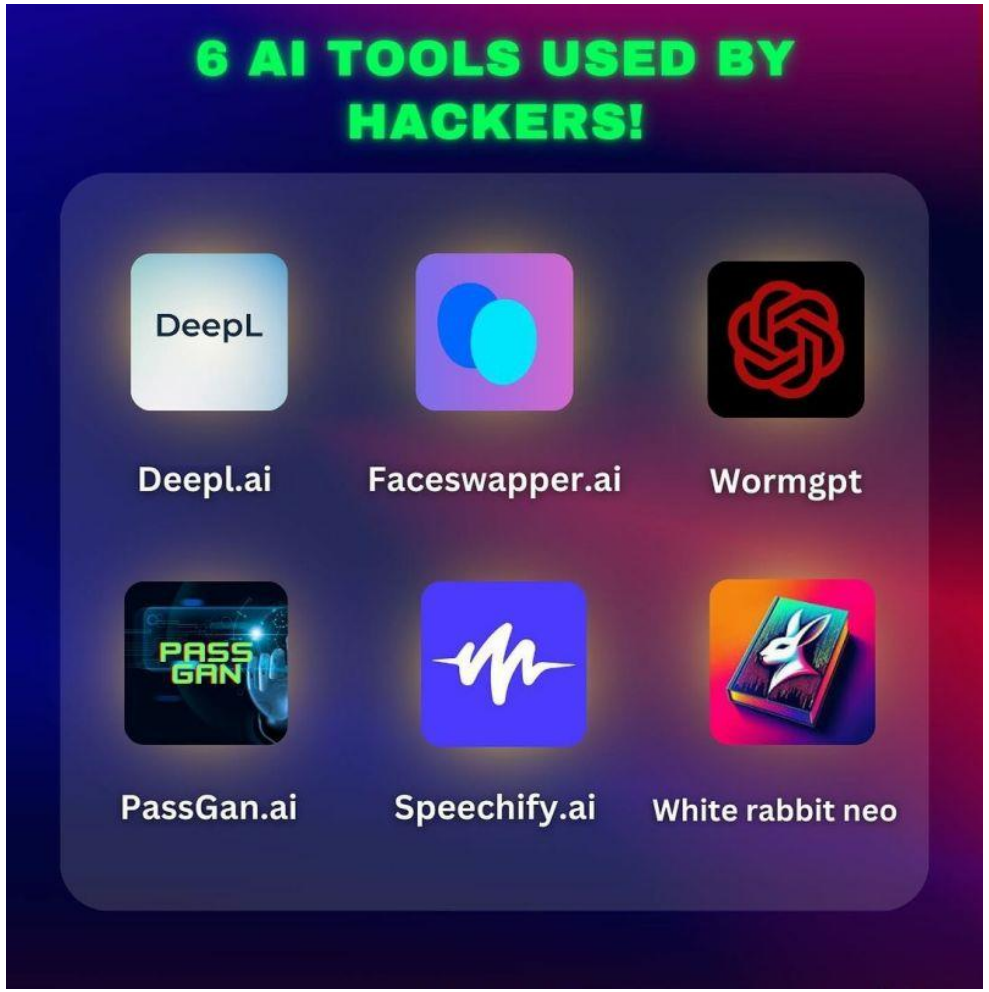
2 min read · Published 2:31 AM EST, Sun February 4, 2024


Securing
Your Digital
World
InfoGuard
SWISS CYBER SECURITY
Würden Sie die Prinzipien
erkennen?

Bösartige KI-Tools senken die Cybercrime-Einstiegshürden

- Cyberkriminelle nutzen KI für:
  - Erstellung von Phishing-Mails
  - Produktion von Deepfakes
  - Optimierung von Malware-Code

- KI-Tools im Darkweb verfügbar, z.B.:
  - WormGPT
  - FraudGPT
  - PoisenGPT
  - Speechif.ai

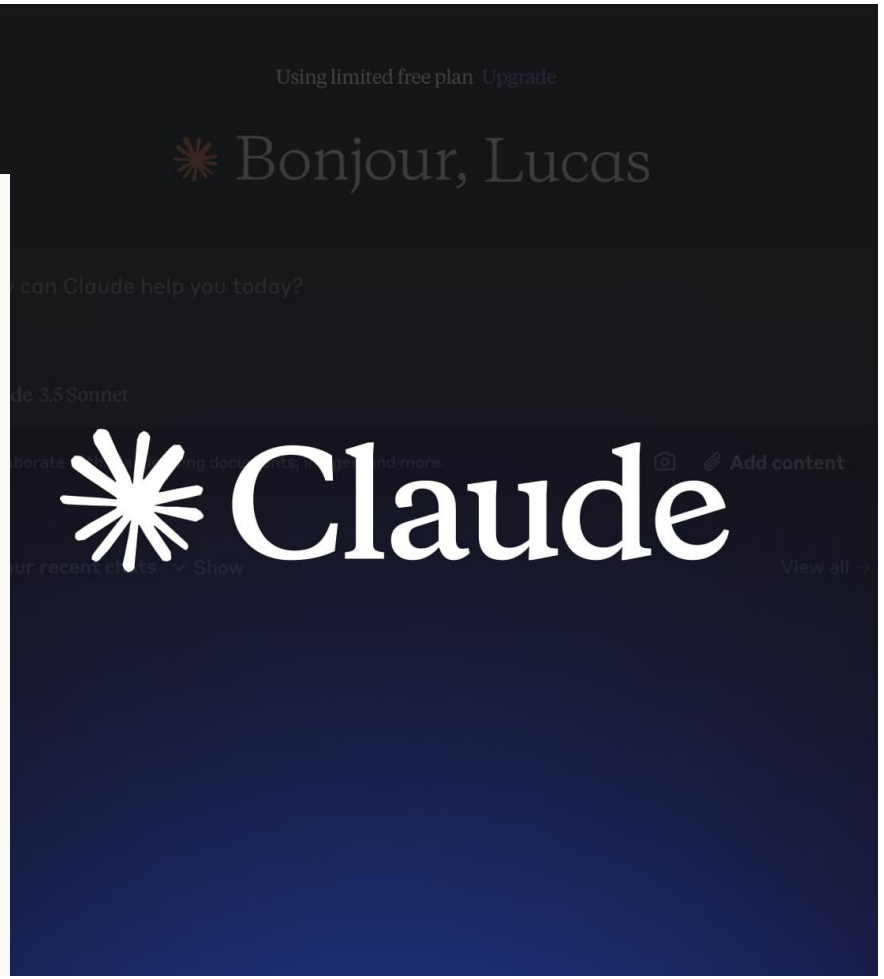


Cyberkriminelle nutzen KI von Claude für die Entwicklung von Ransomware

InfoGuard
SWISS CYBER SECURITY

Sie „Vibe Hacken“ ihre Ziele

- „Wie prüft man ob go installiert ist“
- „wie arbeitet man mit der outlook ap
- „was heisst dieser Code?“



KI-Browser: Sicherheitstechnisch noch nicht bereit

Login-Daten für KI-Browser werden in Zukunft Ziel der Hacker!

Tipps:

- Sind sicherheitstechnisch noch nicht soweit
- Keine Login-Daten speichern
- Passwörter massiv beschützen
- Passwort-Empfehlungen beachten (einzigartig, MFA)
- Datenschutz-Einstellungen stark einschränken

AI · BROWSERS

Cybersecurity experts warn OpenAI's ChatGPT Atlas is vulnerable to attacks that could turn it against a user—revealing sensitive data, downloading malware, or worse

BY BEATRICE NOLAN
TECH REPORTER

October 23, 2025 at 6:16 AM EDT

OpenAI's new AI browser sparks fears of data leaks and malicious attacks.
CHENG XIN—GETTY IMAGES

Cyberkriminelle nutzen KI von Claude für die Entwicklung von Ransomware

Kriminelle nutzen KI von Claude für den gesamten Angriff, kein technisches Wissen mehr nötig!

PHASE 1: AUFKLÄRUNG UND ZIELFINDUNG

Der Akteur nutzte Claude Code, um automatisch viele Schwachstellen in VPN-Systemen zu finden und wichtige Infos zu sammeln.

PHASE 2: ERSTER ZUGRIFF UND AUSNUTZUNG VON ANMELDEDATEN

Claude half dabei, in Netzwerke einzudringen und Zugangsdaten zu erlangen.

PHASE 3: ENTWICKLUNG VON MALWARE UND UMGEHUNG VON SCHUTZMASSNAHMEN

Claude half, Schadsoftware zu bauen und zu verstecken, die sich als echte Microsoft Programme ausgibt.

PHASE 4: DATENANALYSE, ERPRESSUNG, ERSTELLUNG DER LÖSEGELDFORDERUNG

Claude half bei der Ransomware-Erpressung

KI-Phishing

„Generiere mir eine E-Mail des Telekommunikationskonzerns Mustercom, die den Adressaten dazu auffordert, sein Passwort aufgrund eines Sicherheitsvorfalls möglichst zeitnah zu ändern.“

Betreff: Dringende Passwortänderung erforderlich - Sicherheitsvorfall bei Mustercom

Sehr geehrter Kunde,

Aufgrund eines aktuellen Sicherheitsvorfalls möchten wir Sie dringend bitten, Ihr Passwort für Ihr Mustercom-Konto zu ändern. Obwohl keine Anzeichen für eine Kompromittierung Ihrer Daten vorliegen, ist es wichtig, zusätzliche Schutzmaßnahmen zu ergreifen.

Folgen Sie diesen einfachen Schritten, um Ihr Passwort zu ändern:

1. Melden Sie sich auf unserer Website unter **www.mustercom.com** an.
2. Gehen Sie zu den Kontoeinstellungen oder zur Sicherheitsseite.
3. Wählen Sie die Option "Passwort ändern" und folgen Sie den Anweisungen.
4. Verwenden Sie ein starkes Passwort, das aus Buchstaben, Zahlen und Sonderzeichen besteht.
5. Speichern Sie Ihr neues Passwort sicher und teilen Sie es mit niemandem.

Bei Fragen oder Bedenken stehen wir Ihnen gerne zur Verfügung. Unser Kundenservice-Team steht Ihnen rund um die Uhr zur Verfügung.

Vielen Dank für Ihr Verständnis und Ihre Unterstützung, um die Sicherheit Ihrer Daten zu gewährleisten.

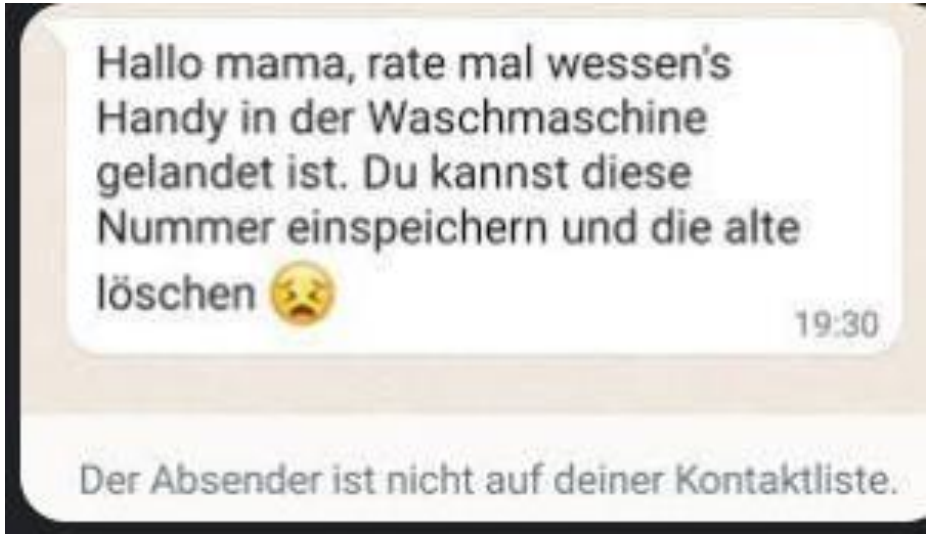
Mit freundlichen Grüßen,

Ihr Mustercom-Team

Voice Cloning: Hilfsbereitschaft

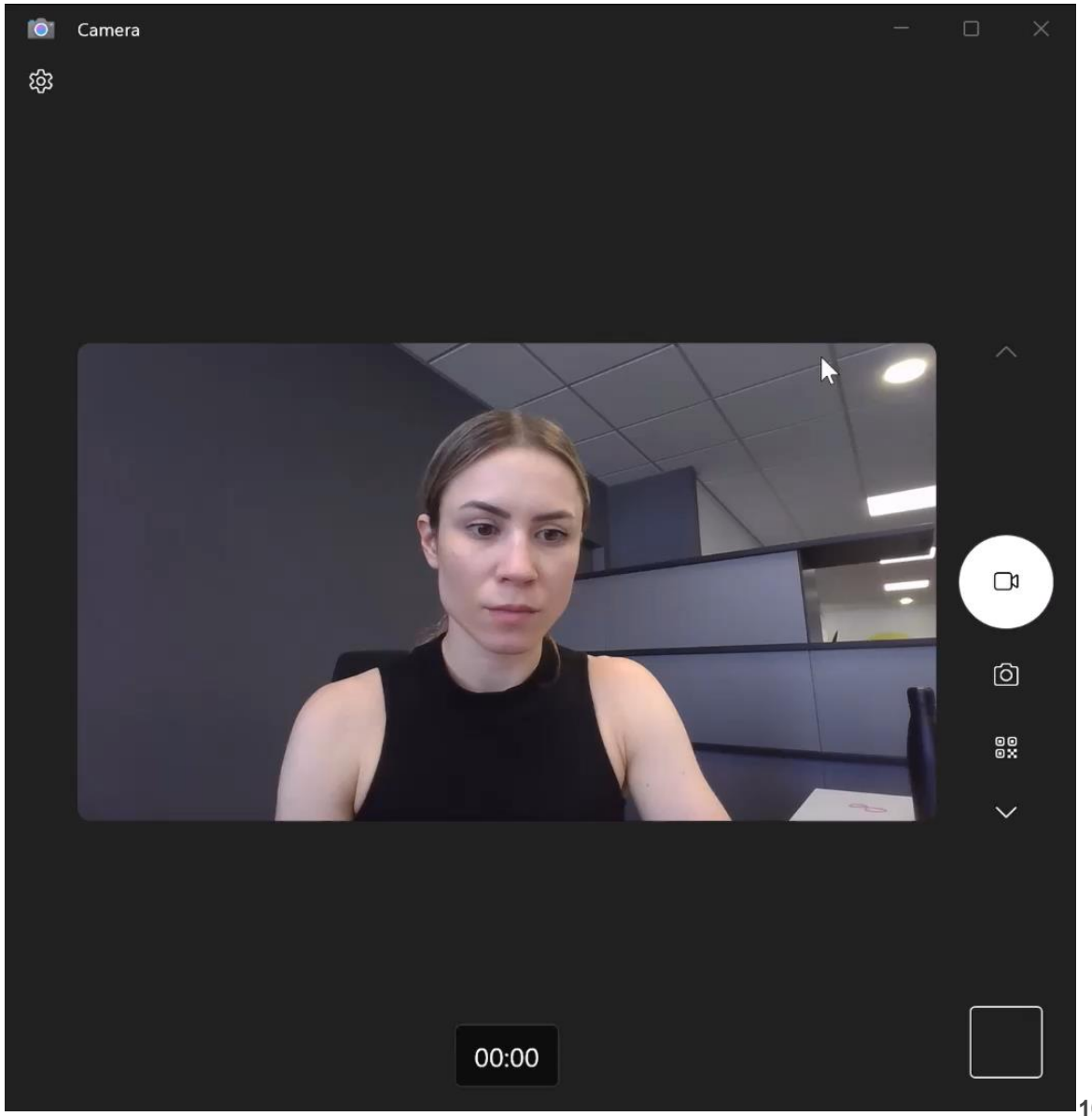
Die Stimme aus einem TikTok-Video wird verwendet, um Eltern vorzuspielen, dass die Tochter in Gefahr ist.

➔Angriffe analog Hallo Mami-Betrug möglich:



Ein Bild reicht für ein Deepfake Video!

Deepfake Videos können heute in Echtzeit während Videogesprächen erzeugt werden, was Verifizierung erschwert



Erkennung von Deepfakes

InfoGuard
SWISS CYBER SECURITY

Tipps KI – Angriffe:

- Auf anderem Kanal zurückrufen und Identität prüfen
- Codewort vereinbaren mit dem Umfeld
- Auf Insider-Wissen beziehen

Merkmal	Warnsignal / Worauf achten?
Gesicht und Augen	Unnatürliches Blinzeln (zu viel/zu wenig), starrer oder leerer Blick, unpassender Hautton an den Rändern des Gesichts.
Mund und Stimme	Schlechte Lippensynchronisation, roboterhafter oder emotionsloser Ton, seltsame Pausen oder falsche Betonung.
Bildqualität	Unscharfe/verzerrte Ränder um Gesicht und Haare, inkonsistente Beleuchtung oder Flackern im Video.
Inhalt und Kontext	Sagt die Person etwas völlig Untypisches? Ist das Angebot unrealistisch? Ist die angebliche Aussage plausibel?
Quelle	Befindet sich das Video auf einem offiziellen Kanal? Berichten seriöse Medien darüber? Eine schnelle Online-Suche hilft oft weiter.

Danke!

Jill Wick
Cyber Security Awareness Consultant, InfoGuard AG **Jill.wick@infoguard.ch**

Securing Your Digital World – Today and Beyond

InfoGuard AG
Lindenstrasse 10
6340 Baar / Schweiz
T +41 41 749 19 00

info@infoguard.ch
www.infoguard.ch

Jill Wick
Jill.wick@infoguard.ch
T +41 41 749 14 04