Cybersicherheitsagentur Baden-Württemberg
(CSBW)

Barrierefreie Version
CSBW-Factsheet: Cybersecurity-Wissen kompakt

Zum Thema: Deepfake

Tipps, wie Sie Deepfakes erkennen können!

Mithilfe von Künstlicher Intelligenz (KI) können Personen in Video- und Tonaufnahmen manipuliert und nachgeahmt werden. Dabei können Deepfakes beispielsweise für Desinformationskampagnen und zur Manipulation der öffentlichen Meinung genutzt werden. Cyberkriminellen ermöglichen sie das Abgreifen von Daten und vertraulichen Informationen.[1]

Was ist Deepfake?

Das Wort Deepfake ist ein Kunstwort, das sich aus den englischen Begriffen Deep Learning (dt. mehrschichtiges/tiefes Lernen) und Fake (dt. Fälschung) zusammensetzt.[2] Daher bezeichnet ein Deepfake die Manipulation einer Identität durch Methoden der Künstlichen Intelligenz von Video- und Tondateien.

Deepfakes sind nicht nur bei aufgezeichneten Video- und Tondateien machbar, auch bei Videokonferenzen ist eine Live-Manipulation möglich.

Gefahren von Deepfakes:

1. **Überwindung biometrischer Systeme**, beispielsweise bei einer Identifikation per Video.
2. **Social Engineering**, um gezielte Phishing-Angriffe auf sensible Informationen durchzuführen.
3. **Desinformationskampagnen**, die manipulierte Medieninhalte großflächig streuen.
4. **Verleumdung**, in dem Personen beliebige Aussagen in den Mund gelegt werden, die den Ruf dieser Personen nachhaltig schädigen.

Arten von Deepfakes[2]

1. **Fälschung von Gesichtern** sogenannte „Face Swapping“ und „Face Reenactment“
2. **Fälschung von Stimmen** sogenannte „Text-to-speech“ (TTS) und „Voice Conversion“ (VC)
3. **Fälschung von Texten**
4. **Fälschung von Bildern**

Tipps, um Deepfakes zu erkennen?[3]

1. **Auf die Umgebung achten**
   Eine hohe Bildschirmauflösung und ein großer Bildschirm helfen Ungereimtheiten (z.B. sichtbare Übergänge) im Bild leichter zu erkennen. Daher ist es empfehlenswert Videos nicht auf dem Handy, sondern auf dem Monitor anzusehen.
2. **Mimiken erkennen**
   Da eine KI z.B. Blinzeln oder Stirnrunzeln bisher noch nicht gut darstellen kann, ist das genaue Beobachten von Augen und Stirn hilfreich.
3. **Quellen kontrollieren**
   Quellen zu recherchieren und zu identifizieren, aus der Videos oder Tonaufnahmen stammen, ist ebenfalls empfehlenswert. Unseriöse und unbekannte Quellen können auf eine Fälschung hindeuten.
4. **Aktive Kontrolle**
   Bei Live-Bildmanipulationen in Videokonferenzen kann ein Deepfake enttarnt werden, in dem Sie die andere Person bitten, sich kurz an die Nase oder Stirn zu tippen. Aktuell ist die KI nicht so weit entwickelt, dass diese Animation ohne Verzerrung möglich ist.

Was ist Künstliche Intelligenz?[4]

Unter Künstlicher Intelligenz wird die Fähigkeit einer Maschine verstanden, menschliche Fähigkeiten wie logisches Denken und Lernen zu imitieren.

KI-Systeme benötigen zum Trainieren der -Fähigkeiten eine große Menge von Daten. Je mehr Daten dem KI-System zum Training zur Verfügung stehen, umso besser wird die selbstständige Lösungskompetenz des KI-Systems.

Da von Personen des öffentlichen Lebens (z.B. politisch aktive Personen) oft ausreichend Bild- und Tonmaterial zur Verfügung steht, gibt es von ihnen besonders viele Deepfakes.

Quellen:

[1] https://www.bundesregierung.de/breg-de/schwerpunkte/umgang-mit-desinformation/deep-fakes-1876736

[2] https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/Deepfakes/deepfakes_node.html

[3] https://www.verfassungsschutz.de/SharedDocs/publikationen/DE/wirtschafts-wissenschaftsschutz/2021-11-spoc-wirtschaft-und-wissenschaft-schuetzen.pdf?__blob=publicationFile&v=9

[4] https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Technologien_sicher_gestalten/Kuenstliche-Intelligenz/kuenstliche-intelligenz_node.html

Weitere Factsheets und Informationen unter: www.cybersicherheit-bw.de

CSBW – Abteilung 1: Prävention – Stand 05.2024
Kontakt: schulungen@cybersicherheit.bwl.de