UNIVERSITÄT BAYREUTH

# ChatGPT Workshop

**Prof. Dr. Agnes Koschmider**

Universität Bayreuth,
Professur für Wirtschaftsinformatik und Process Analytics

https://www.pa.uni-bayreuth.de/

# Woran erkennt man Texte, die mit ChatGPT erzeugt wurden?

übermäßig polierte Sprache

Inkonsistenter Ton/Sprache

keine Abkürzungen/“Slang-Sprache“

ungewöhnliche Formulierungen

Allgemeine Formulierungen

# Begriffe, die ChatGPT häufig wiederholt

Beispiele für Inhalte, die in ChatGPT-Texten tendenziell sehr oft wiederholt werden, sind:

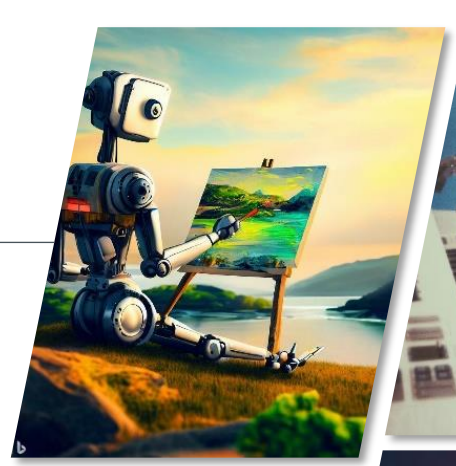
- Budget
- Faktoren
- Optionen
- Kriterien
- Anforderungen

# Generative KI geht über die Erstellung von Text hinaus. Neue Produkte erweitern die Möglichkeiten stetig

UNIVERSITÄT BAYREUTH

## Text-to-Image

Mit Modellen wie **Stable Diffusion** ist es möglich, täuschend echte Bilder zu erzeugen

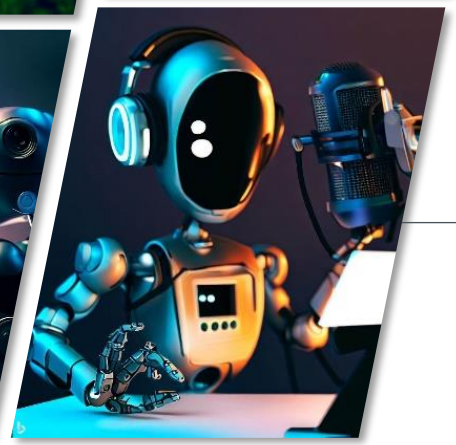


## Text-to-Music

**Googles MusicLM** erstellt hochwertige, konsistente Musikstücke

## Text-to-Video

**NVIDIAs** AI erstellt hochauflösende, kohärente Videos

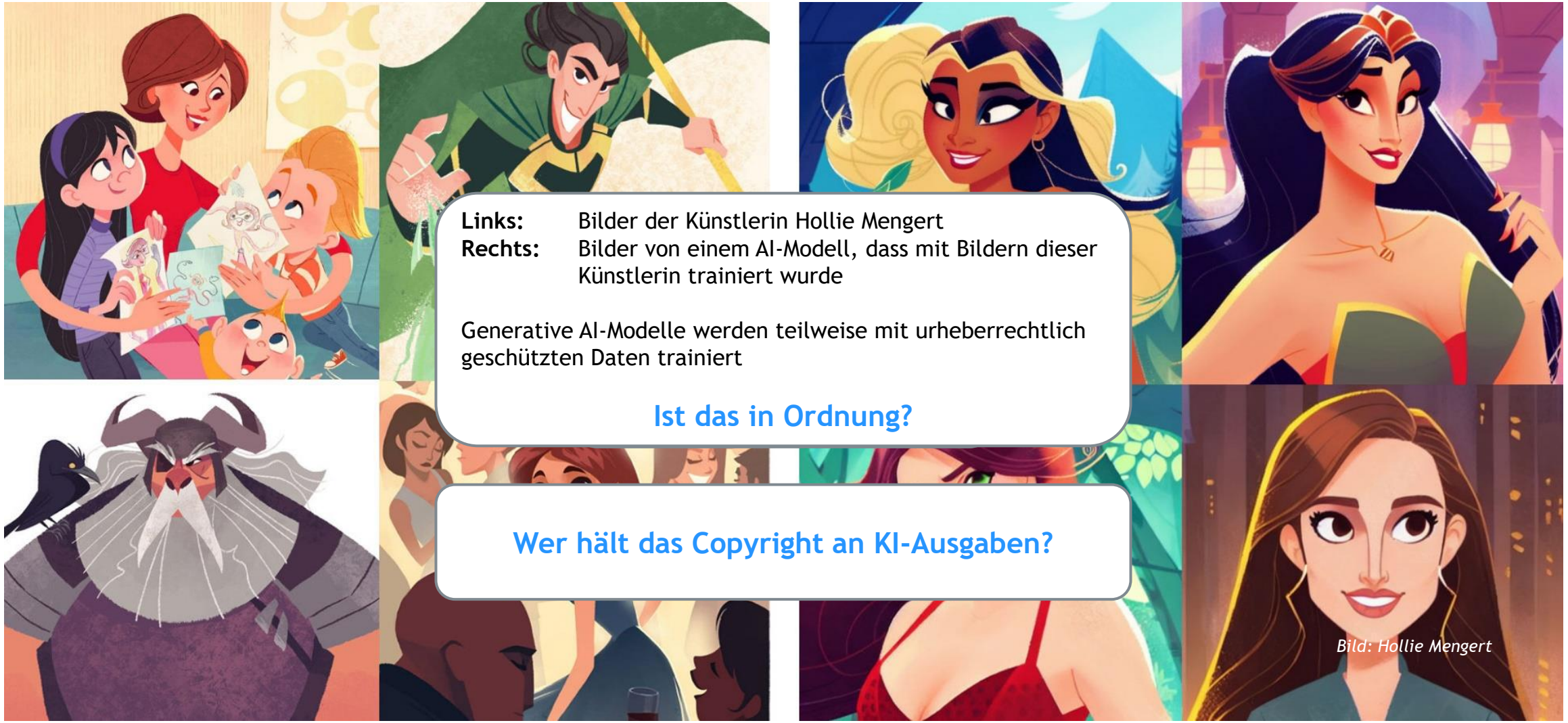


## Text-to-Speech / Voice Cloning

**Microsofts VALL-E** ermöglicht das Klonen einer Stimme aus einer 3-Sekunden-Aufnahme

*Bilder: Unsplash, Bing Image Creator*

# Copyright und generative KI



**Links:** Bilder der Künstlerin Hollie Mengert
**Rechts:** Bilder von einem AI-Modell, dass mit Bildern dieser Künstlerin trainiert wurde

Generative AI-Modelle werden teilweise mit urheberrechtlich geschützten Daten trainiert

**Ist das in Ordnung?**

**Wer hält das Copyright an KI-Ausgaben?**

*Bild: Hollie Mengert*

# Deep Fakes – die Technik übernimmt

- Interview mit Prof. Dr. Antonio Krüger, CEO des Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)

**Herr Prof. Krüger, was sind eigentlich DeepFakes?**

Der Begriff DeepFake ist eine Wortkombination aus „**Deep**“, was auf Deep Learning, einer KI-Technik mit zugrunde liegenden künstlichen neuronalen Netzen hinweist und „**Fake**“ für Fälschung. […] mit Methoden der Künstlichen Intelligenz **manipulierte Medieninhalte wie Audio, Photo, Video in einer schlechten Qualität**, die nicht ohne Weiteres oder gar nicht also Fake […] zu erkennen sind.
Die wohl bekannteste Variante […] ist das sogenannte **Face-Swap**. Das meint den Austausch des Gesichts einer Quellperson mit dem Gesicht einer Zielperson in einem Bild oder einem Video. Damit (und in Kombination mit Audio-Ansätzen) können Aussagen und Handlungen vorgetäuscht werden, die die betreffenden Personen nie gesagt oder ausgeführt haben.

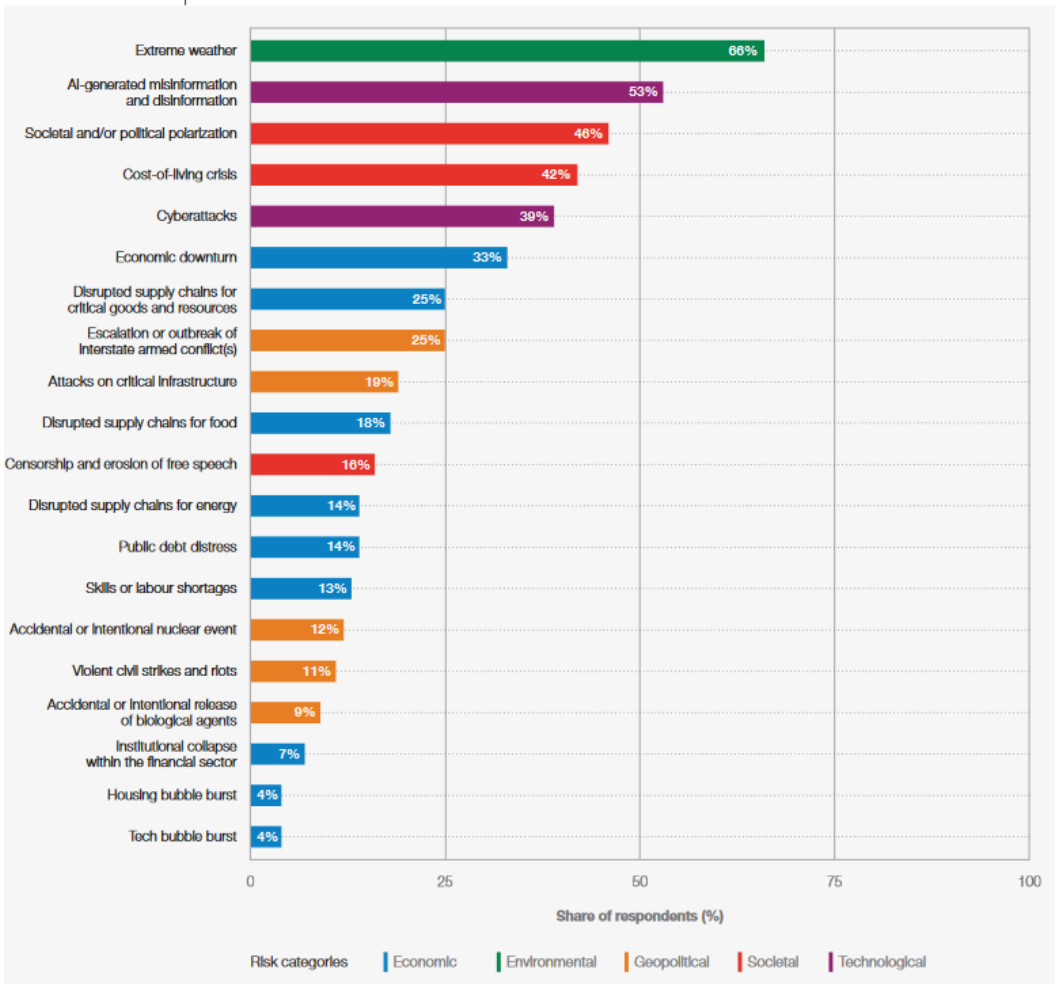
https://www.dfki.de/web/news-media/news-events/interview-was-sind-eigentlich-deepfakes


© Universität Bayreuth

# Fake News – das Risiko für eine globale Krise

FIGURE 1.2 **Current risk landscape**

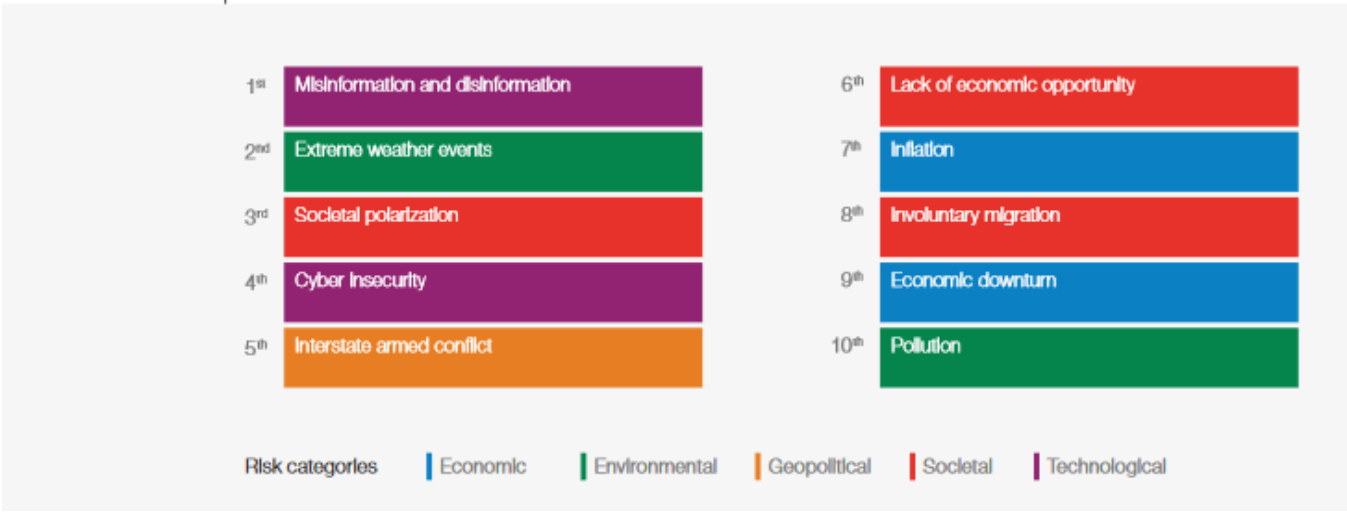
*"Please select up to five risks that you believe are most likely to present a material crisis on a global scale in 2024."*




Fake News gemäß Duden: *in den Medien und im Internet, besonders in den Social Media, in manipulativer Absicht verbreitete Falschmeldungen*

FIGURE 1.3 **Global risks ranked by severity over the short term (2 years)**

*"Please estimate the likely impact (severity) of the following risks over a 2-year period."*

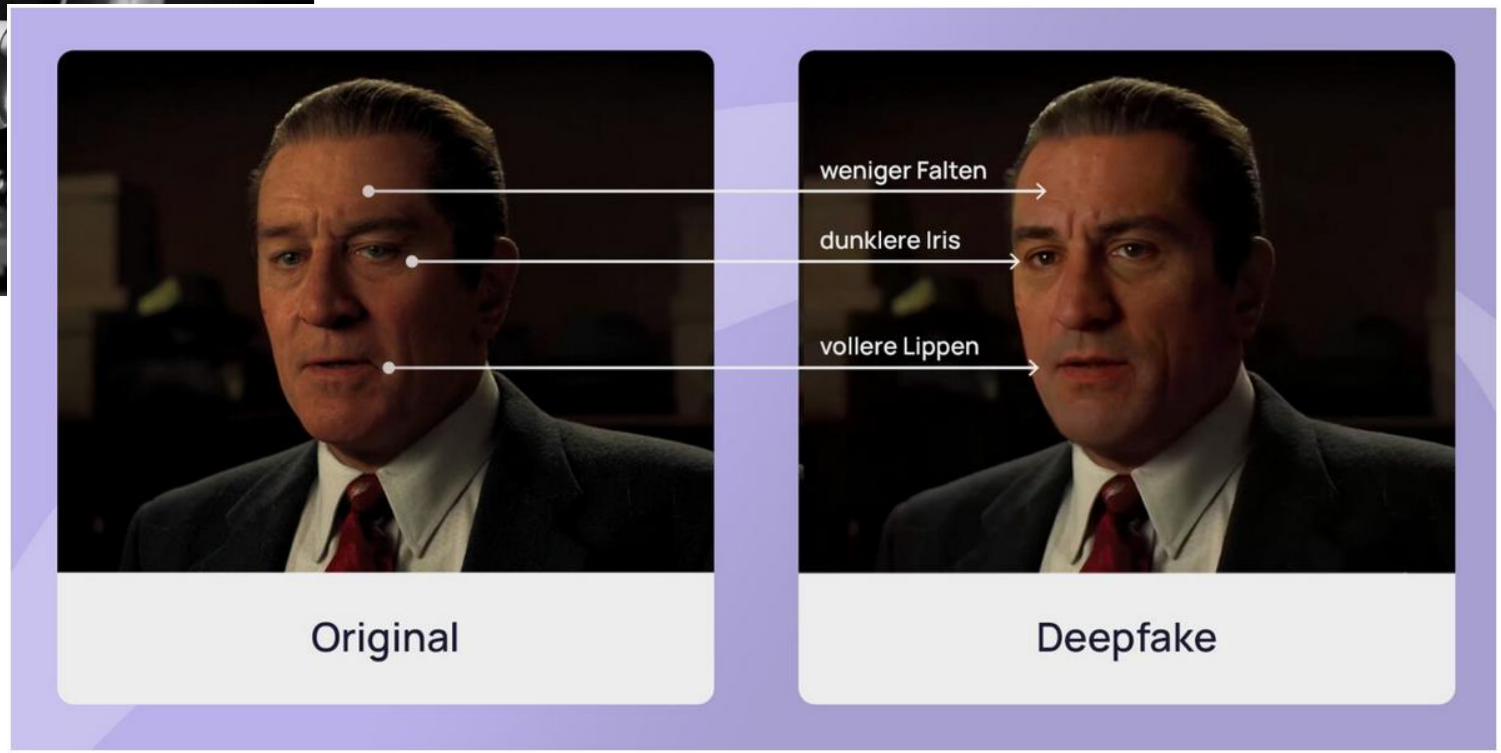



https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf



© Universität Bayreuth

Papst Franziskus im römischen Tanzclub „Magic Fly“ (2017) | © Deepfake: Pavel Sokolov




Quelle: Shamook (2020). De-aging Robert Deniro in The Irishman [Deepfake]



© Universität Bayreuth

# Gefahr von Deep Fakes

- Marktverfügbare Apps ermöglichen eine massenhafte, schnelle und kostengünstige Anfertigung von Deep Fakes. Das erlaubt es neben Staaten auch ressourcenarmen Gruppierungen und Individuen, eigene Desinformationskampagnen im großen Maßstab durchzuführen

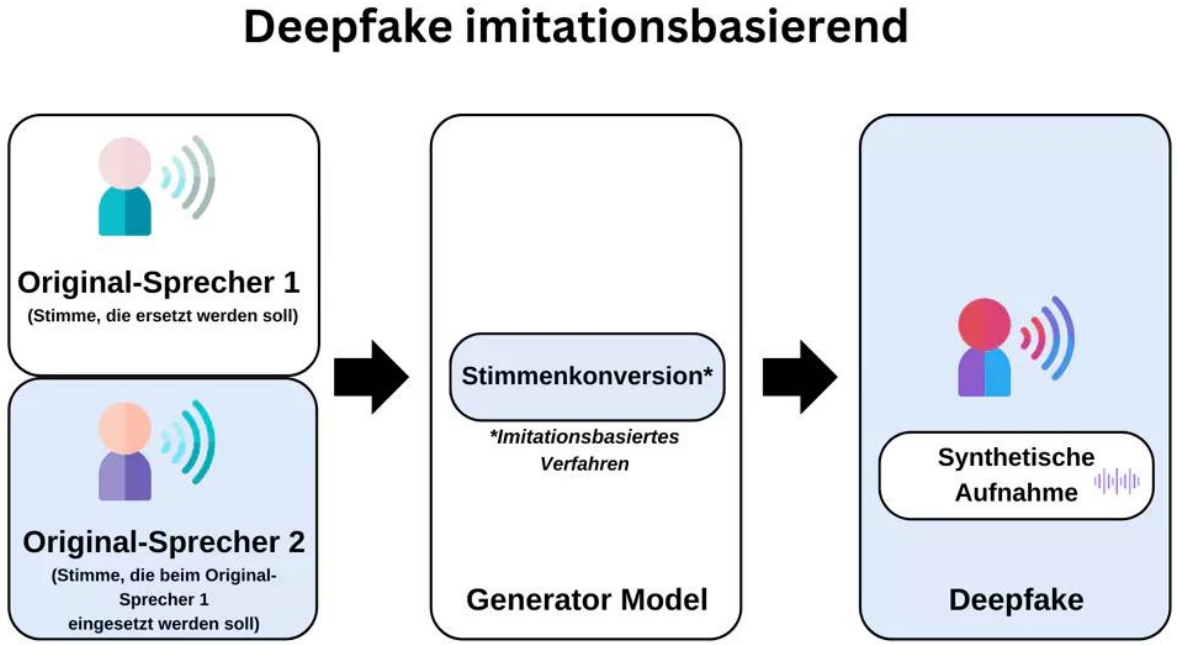
- Deep Fakes werden qualitativ immer besser und wirken natürlicher, wodurch sie schwerer zu erkennen sind und an Glaubwürdigkeit und Überzeugungskraft gewinnen

- Während die Erstellung von Deep Fakes nahezu keinerlei Qualifikation voraussetzt, wird die zu ihrer Erkennung erforderliche Expertise immer umfangreicher


© Universität Bayreuth

# Wie erstelle ich Deep Fakes

- Kostenlose Programme im Internet (DeepFaceLab, FakeApp oder Deepfake)
  - Bild bzw. Videobearbeitungsprogramm
  - Sammlung von Bild- oder Videoaufnahmen notwendig




Quelle: https://www.heise.de/hintergrund/Deepfakes-KI-gegen-KI-9339715.html


© Universität Bayreuth

# Wie erkenne ich Deep Fakes

- typische Bildfehler oder blecherne Stimmen

- Aufgrund der Text-Bild Erzeugung passen häufig die Darstellung der Hände, die Zahl und Form der Finger nicht (unnatürliche Gesichter, unscharfe Übergänge)

© Universität Bayreuth

# Wie erkenne ich Deep Fakes

- typische Bildfehler oder blecherne Stimmen
- Aufgrund der Text-Bild Erzeugung passen häufig die Darstellung der Hände, die Zahl und Form der Finger nicht (unnatürliche Gesichter, unscharfe Übergänge)
- Je höher die Auflösung beziehungsweise die Bildgröße, desto leichter lassen sich Ungereimtheiten im Bild erkennen
- Inkonsistenzen in den Videodaten


© Universität Bayreuth

# Zukünftiger Umgang mit Deep Fakes

UNIVERSITÄT BAYREUTH

- Eindämmen – Erkennen – Reagieren
  - softwareseitige Kennzeichnungspflicht für Deep Fakes
  - die Stärkung vertrauenswürdiger Inhalte durch das Hinzufügen fälschungssicherer Verlaufs- und Identitätsinformationen zu digitalen Daten
  - die Stärkung von Medienkompetenz und die Entwicklung eigener Aufklärungsansätze unter Verwendung von KI


© Universität Bayreuth

# Key Takeaways für das erfolgreiche Arbeiten mit ChatGPT

*„Denken ist wie ChatGPT, nur krasser“*

*(Roger Basler de Roca)*

1

**Nutze ChatGPT als unterstützendes Werkzeug**
ChatGPT kann als unterstützendes Werkzeug (wie Google oder Wikipedia) sehr nützlich sein

2

**Hinterfrage Antworten stets kritisch**
Das Arbeiten mit ChatGPT ersetzt nicht das eigene, kritische Denken

3

**Verifiziere Outputs durch belegbare Quellen**
Auch sehr plausibel klingende Antworten müssen durch belegbare Quellen verifiziert werden