

# Inhaltsverzeichnis

Forschungszentrum Athene .....	2
Geotargeting .....	4
Deepfake - Anzeichen für Fälschungen .....	8
Taiwan Semiconductor Manufacturing Company (TSMC) .....	19

## UNDER CONSTRUCTION

Was ist über die verwendeten Algorithmen einiger	
Suchmaschinen bekannt? .....	xx
Datenschutz beim Video-Chat (Zoom, Skype) .....	xx
Datensicherheit im DarkNet (Onion) .....	xx
Streamingwebseiten - Odyssee, Peertube, Youtube .....	xx
Netzwerk-Kollisionen: Was sind die Unterschiede bei	
Kabel und WLAN? .....	xx
Künstliche Intelligenz und Chat GPT .....	xx
Gibt es sichere Smartphone's (Bittium Corporation)? .....	xx
Cybersicherheit: Solar- und Windparks im Fadenkreuz	
(kritische Infrastruktur) .....	xx
Fake Shop Finder .....	xx
Übersetzungsprogramm DeepL App .....	xx
Gesetze zur künstlichen Intelligenz (KI, AI) .....	xx
HAMNET - Funknetz in Deutschland .....	xx
Polizei-Software Autopsy .....	xx
Satelliten - Aufnahme in die kritische Infrastruktur .....	xx
Datensicherheit - Smarthome Computer, Saugroboter,	
Protokoll-Familie für Bluetooth .....	xx
Zero Trust Security .....	xx
KI-Detector .....	xx
AIMS Software .....	xx
Bluecode .....	xx
Was ist Sim Swapping? .....	xx
CrowdStrike – ein Unternehmen für Informationssicherheit ..	xx
Dark Pattern .....	xx
Juice Jacking .....	xx

- INPOL
- Textkodierungen unter Windows
- Facebook, Twitter, Whatsup, TikTok, Google, Bing -> Algorithmen, Sicherheit
- Frauenhofer Institut -> Cyber Range (Recherche??)
- Werkzeug für die automatische Ermittlung des Cyber-Lagebildes (Auswertung von Info's aus dem Internet)
- Fake Shop Finder (Adresse über search engines) -> Webseite der Verbraucherzentrale
- DeepL App -> Übersetzungsprogramm für 29 Sprachen (2023-03)
- Twitter (Elon Musk) -> veröffentlicht teilweise seine Algorithmen (April 2023) -> neuer Name: X.Corp
- HAMNET -> Funknetz in Deutschland
- Deep Fake (Video, Bilder) Anzeichen für Fälschungen (Hände, 6 Finger)
- Satelliten -> Aufnahme in die kritische Infrastruktur ????
- Firma: Green Mountain -> ehemailer NATO Munitionsbunker in Norwegen -> Rechenzentrum und Cloud -> direktes Kabel nach Newcastle (Irland) ???
- paper-mills (Papiermühlen) Erstellung von Bachelor-Arbeiten, Diplom- und Doktorarbeiten, wissenschaftliche Zeitschriftenartikel, Publikationen gegen Bezahlung
- russische Hackergruppe Black Basta -> hat unter anderen digitale Kopien von Personalausweisen erbeutet -> Identitätsdiebstahl dadurch möglich ->
  - Gegenmaßnahme: Beantragung eines neuen Personalausweises
- G-IDS (global), L-IDS (local)
- Bitcom Studie -> Cyber Crime
- Glossar: Cyber physischer Angriff

Das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE ([www.athene-center.de](http://www.athene-center.de)), ehemals Center for Research in Security and Privacy (CRISP), ist das größte Forschungszentrum für Cybersicherheit in Europa. Es beschäftigt sich mit Kernfragen der Cybersicherheit in der Digitalisierung von Staat, Wirtschaft und Gesellschaft.

ATHENE ist ein Forschungszentrum der Fraunhofer-Gesellschaft für ihre beiden Darmstädter Institute SIT (Sichere Informationstechnologie) und IGD (Institut für Graphische Datenverarbeitung) unter Beteiligung der Goethe-Universität Frankfurt/Main, der Technischen Universität Darmstadt und der Hochschule Darmstadt. In einem einzigartigen und innovativen Kooperationsmodell von universitärer und außeruniversitärer Forschung werden die Kompetenzen und Stärken der Fraunhofer Gesellschaft mit den Kompetenzen und Stärken von Hochschulen kombiniert.

ATHENE hat eine lange Tradition im Bereich der Cybersicherheit, sowohl durch seine beteiligten Organisationen als auch durch seine Vorgängerzentren. Auch wenn ATHENE in seiner jetzigen Form erst 2019 gegründet wurde, blickt es auf eine sehr lange Geschichte in der Cybersicherheitsforschung zurück.

Das Zentrum wird gefördert vom Bundesministerium für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) und hat seinen Hauptsitz in der Wissenschaftsstadt Darmstadt (Hinweis: Sitz des Trägers der Forschungseinrichtung ist München).

Darüber hinaus unterhält das Institut eine Niederlassung in Münster/Westfalen und ein Büro in Berlin. International ist das Institut auch in Israel vertreten.



Es ist **eines von drei** vom Bundesministerium für Bildung und Forschung finanzierten Kompetenzzentren für IT-Sicherheit und dient als Nationales Forschungszentrum für angewandte Cybersicherheit.

Athene entwickelt Sicherheitslösungen, berät regelmäßig Wirtschaft und öffentliche Verwaltung und unterstützt Firmengründer und Startups. Dabei fließen die aus der Grundlagenforschung der Hochschulen gewonnenen Erkenntnisse in die anwendungsorientierte Forschung ein.

Mit seinen Forschungs- und Entwicklungsarbeiten deckt ATHENE ein sehr großes Spektrum von Expertisen (Sachkunde, Spezialwissen) ab, die für verschiedene Technologien und Anwendungsbereiche relevant sind, wie die Sicherheit von Systemen, Software, Prozessen, Hardware, Daten oder der Internet-Infrastrukturen.

Die Wissenschaftler finden in ATHENE optimale Arbeitsbedingungen und profitieren von einem internationalen Umfeld und haben gute Kontakten zur Industrie und Wirtschaft. Die großen Herausforderungen der Cybersicherheitsforschung werden in

ATHENE in langfristigen, anwendungsorientierten Missionen bearbeitet. Beispielsweise erforscht ATHENE, wie man die kritischen Infrastrukturen Deutschlands (Strom, Wasser, Verkehr, usw.) zuverlässig schützt und wie man IT-Systeme langfristig absichert, selbst angesichts neuer Technologien wie Quantencomputern.

Darüber hinaus identifiziert ATHENE fortlaufend, umfassend und vorausschauend die wichtigen, anwendungsorientierten Fragen der Cybersicherheit und der Privatheit.

## **Forschungsschwerpunkte:**

- Analytic Based Cybersecurity (ABC)
- Automatic Vulnerability Scanning and Verification (AVSV)
- Cryptography (CRYPTO)
- Legal Aspects of Privacy and IT Security (LeAP)
- Next Generation Biometric Systems (NGBS)
- Reliable and Verifiable Information through Secure Media (REVISE)
- Secure Autonomous Driving (SAD)
- Secure Urban Infrastructures (SecUrban)
- Secure Digital Transformation in Health Care (SeDiTraH)
- Security and Privacy in Artificial Intelligence (SenPAI)
- Trustworthy Data Ecosystems (TRUDATA)
- User-centered Security and Privacy (UCSP)

## Geotargeting - Geographische Zuordnung von IP-Adressen

IP-Adressen können zwar wegen einiger Verfahren wie die dynamischer IP-Adressenvergabe, Proxyservern oder NAT (Network Address Translation) nicht immer eindeutig einem Internetbenutzer zugeordnet werden, jedoch immer einem Besitzer (IP-Adressen-Besitzer). Hierbei handelt es sich häufig um Internetprovider, Universitäten und ähnliche Einrichtungen, die nicht nur eine IP-Adresse, sondern IP-Adressräume verwalten. Der Besitzer einer IP-Adresse kann frei entscheiden, welchem Netzknoten er welche Adresse zuteilt.

**Definition:** Geotargeting (Synonyme: Geolocation/Geolokation) ordnet IP-Adressen ihrer geografischen Herkunft zu. Um den Standort zu ermitteln, werden nicht nur GPS-Daten, sondern auch IP-Adressen oder Datenbanken ausgewertet.

Obwohl die Zuteilung der IP-Adressen im Prinzip schnell geändert werden kann, wird von dieser Möglichkeit nur selten Gebrauch gemacht. Der dafür entstehende Verwaltungsaufwand ist nicht zu unterschätzen. Dadurch kann aus einer einmal festgestellten Geoposition einer IP-Adresse auf einen Wochen später noch aktuellen Zusammenhang geschlossen werden. Da regionale Einwahlknoten häufig einen eigenen festen IP-Adresspool besitzen, funktioniert das Verfahren meist bei dynamischer IP-Vergabe ebenfalls. Beim Einsatz von Proxyservern kann maximal der Standort des Proxy-Servers, jedoch nicht der des eigentlichen Nutzers ermittelt werden.

»Geointelligenz« geht einen Schritt weiter: Der Standort der Internetnutzer wird mit Regeln verknüpft, die auf der geographischen Herkunft des Internetnutzers basieren. Wenn ein

Internetbesucher aus Deutschland eine Webseite aufruft, erhält er andere Inhalte als ein gleichzeitiger Besucher derselben Webseite aus Frankreich oder den USA. IP-Intelligenz erweitert also die reine geografische Sicht um qualitative Faktoren wie die Verbindungsgeschwindigkeit oder den identifizierten Internet Service Provider (ISP) des Nutzers. Ein Kabelnetzanbieter kann davon Gebrauch machen, um den Besuchern gezielte Wechselangebote zu unterbreiten ohne dass der Besucher eigene Angaben machen muss.

Die Geolokationssoftware hat zum Ziel, mit Hilfe der IP-Adresse den Standort von Personen oder Systemen **möglichst** genau zu bestimmen. Mitunter werden auch vom Erzeuger bereits Geo-Tags auf Fotos oder Videos gesetzt (Metadaten). Um einheitliche Voraussetzungen vorzugeben, wurde mit der **W3C Geolocation API** eine geeignete Schnittstelle geschaffen.

## Anwendungsbeispiele:

- **Geomarketing:** Eine profitable Anwendung ist Geomarketing. Die meisten Online-Werbefirmen bieten ihren Kunden auf Basis von Geotargeting die Schaltung national oder sogar regional differenzierter Werbung an (Ad Targeting). Besucher sehen Anzeigen, die unabhängig vom Standort der aufgerufenen Webseite, den aktuellen Aufenthaltsort als Zielmarkt ansprechen.
- **Webcontrolling:** Webcontrolling-Anbieter integrieren Geointelligenz in ihre Produkte, um eine geografische Besucheranalyse zu ermöglichen. Der Webseitenbetreiber kann sehen, aus welchen Ländern und Regionen die Besucher kommen.
- **PayPal:** PayPal verwendet Geolocation zum Schutz vor Betrug, um Onlinezahlungen auf regionale Unstimmigkeiten zu überwachen und schließt Transaktionen aus, die aus mit Sanktionen belegten Ländern zu kommen scheinen.

## Geotargeting 2/4

- **Loudeye Inc.:** Loudeye Inc. benutzt Geolocation in der Marktforschung, um regionale Nachfrageunterschiede besser abzubilden oder die eigenen Direktmarketingmaßnahmen zu optimieren (Zielgruppenanalyse).
- **DidTheyReadIt:** DidTheyReadIt bietet als E-Mail-Serviceanbieter nicht nur die Information, ob eine Nachricht geöffnet wurde, sondern auch wo die E-Mail geöffnet wurde (E-Mail-Location).
- **DigitalEnvoy:** Es werden neben den Datenbanken zu allen verwendeten IP-Adressen auch weitere Produkte, welche E-Mails auf geografische Plausibilität überprüfen, angeboten. Diese Produkte vergleichen die Geografie des E-Mail-Headers mit der Geografie des E-Mail-Inhaltes (verwendete Sprache). Verdächtige E-Mails werden gegebenenfalls blockiert oder an Prüfroutinen zum Schutz vor Phishing übergeben.
- **Video-on-Demand:** Video-on-Demand-Anbieter verwenden Geoblocking, da Sportverbände und Filmverlage die Verwertung ihrer Inhalte an territoriale Grenzen binden. Beispiele hierfür sind CinemaNow und Disney.
- **E4X:** E4X nutzt Geointelligenz, um E-Commerce-Site-Besuchern automatisch die richtige Währung anzubieten.
- **Google:** Google und viele andere Suchmaschinen-Anbieter personalisieren ihre Angebote (Suchmaschinenoptimierung oder Search Engine Optimization, SEO), indem sie die Benutzer automatisch auf die Seite in der Sprache des Benutzers führen.
- **Content-Distribution-Netzwerke:** Content-Distribution-Netzwerke optimieren die Lastverteilung zwischen ihren Servern durch Geointelligenz. Die Unternehmen sparen dadurch Kosten und bieten bessere Downloads durch Traffic-Management.
- **Telemedizinische Betreuungs- oder Versorgungssysteme:** Für teilautomatisierte telemedizinische Betreuungs- oder Versorgungssysteme bedeutet die Geo-Lokalisierung von Patienten eine notwendige Facette aller möglichen Lokalisierungs-Techniken. Es laufen Feldstudien vor allem zur Notfallversorgung,

wie bei Myokardinfarkten (Herzinfarkte). Auch die Verwendung einfacher Applikationen wird erprobt oder bereits angewendet, wie die rasche Lokalisierung eines spezialisierten Behandlungszentrums.

- **Unternehmensfilialen:** Unternehmen mit Filialen können Besuchern ihrer Webseite mit Hilfe von Geotargeting einen passenden Standort zuweisen.
- **YouTube:** YouTube bietet bestimmte Videos aufgrund von Lizenzfragen in einigen Ländern nicht an. Dies wird dem Nutzer mit einer Meldung angezeigt (»Dieses Video ist in Ihrem Land nicht verfügbar«).

### Qualität der Geotargetingverfahren

Die Qualität der Verfahren wird anhand folgender Parameter beschrieben:

- Datenvollständigkeit beschreibt wie viele der weltweit verwendeten IP-Adressen das Verfahren abbildet. Eine zuverlässige Technologie sollte 99,99 % der im Gebrauch befindlichen IP-Adressen abdecken.
- Datengenauigkeit besagt, wie genau die Technologie eine einzelne IP-Adresse der Region (Land, Bundesland, Stadt) zurechnen kann. Mögliche Zielwerte liegen über 95 % auf Stadtebene. Wesentlich verschlechtert werden kann diese Erfolgsquote, wenn sich in der Zielgruppe viele Nutzer von Providern befinden, deren Netze sich nur auf Länderebene zuverlässig identifizieren lassen. Ebenso wichtig sind weitere Parameter wie die Identifikation von Proxys, Firmenservern und Bandbreiten, Längen- und Breitengraden, Domains.
- Leistungsfähigkeit sagt aus, wie viele Zugriffe pro Sekunde pro Server das Verfahren leisten kann. Eine schlechte Performance führt nicht nur zu höheren Systemkosten, sondern beeinträchtigt



## Geotargeting 3/4

- auch die interne Nutzungsqualität des Dienstes wesentlich.
- Integrationsaufwand besagt, ob es möglich ist, das Verfahren für mehr als eine Anwendung zu integrieren und wie hoch der Aufwand zur technischen Einbindung in die alternative Systemumgebung ist.

### Geolokalisierung für das Lieferkettenmanagement

Mit vernetzten Geräten (IoT, Internet der Dinge) können wichtige Vermögenswerte geortet werden: Container, Paletten und vieles andere. Oft ist das Gerät ein Mobiltelefon oder ein mit dem Internet verbundenes Gerät.

- Unternehmen können ihren Bestand verfolgen, Lieferwege optimieren und sogar die Kundennachfrage vorhersagen.
- Der Lagerbetrieb wird durch die Nutzung von Daten effizienter.
- Logistik- und Lieferkettenmanager können nachvollziehen, wie sich ihre Bestände in der Lieferkette bewegen und können Änderungen vornehmen, die Zeit und Geld sparen.
- Logistik- und Lieferkettenmanager können Probleme in der Lieferkette leicht ausfindig machen und potenzielle Verbesserungsbereiche identifizieren.

**Hinweis:** Geolokalisierung ist nicht nur eine erstaunliche Technologie, sie ist auch ein gutes Geschäft für viele Firmen, die die Datenschutz-Grundverordnung (DSGVO) der EU hoffentlich auch einhalten.

### GPS (Global Positioning System)

Das Global Positioning System, kurz GPS, ist ein satellitengestütztes Funknavigationssystem, bei dem etwa dreißig Satelliten die Erde umkreisen. Dieses globale Satellitensystem sendet Zeitstempel an einen GPS-Empfänger überall auf oder in der Nähe unseres Planeten - aber nur, wenn es keine Hindernisse gibt und mindestens drei GPS-Satelliten für die Messung verfügbar sind.

Ein großes Plus von GPS ist seine Genauigkeit. Mit GPS kann man Gegenstände genau orten und verfolgen (tracken). GPS kann einen Gegenstand für »normale« Kunden bis auf wenige Meter genau orten.

Der Nachteil ist, dass diese Geolokalisierungs-Technologie relativ viel Strom benötigt, weil sie immer mit mehreren Satelliten gleichzeitig kommunizieren muss. Weitere Nachteile sind, dass die Kommunikation über große Entfernungen oft durch wetterbedingte Situationen unterbrochen werden kann und dass sie, technisch bedingt, nicht in geschlossenen Räumen funktioniert.

**Hinweis:** Wenn man GPS-Daten, IP-Adressen und andere Geo-Daten auf eine natürliche Person zurückführen kann, so gelten sie als personenbezogene, schützenswerte Daten und unterliegen damit den strengen Vorschriften der Datenschutz-Grundverordnung (DSGVO) der EU.

### Geolokalisierung erschweren

Um die Geolokalisierung zu erschweren oder sogar unmöglich zu machen, kann man VPN-Services ((Virtual Private Network) nutzen.

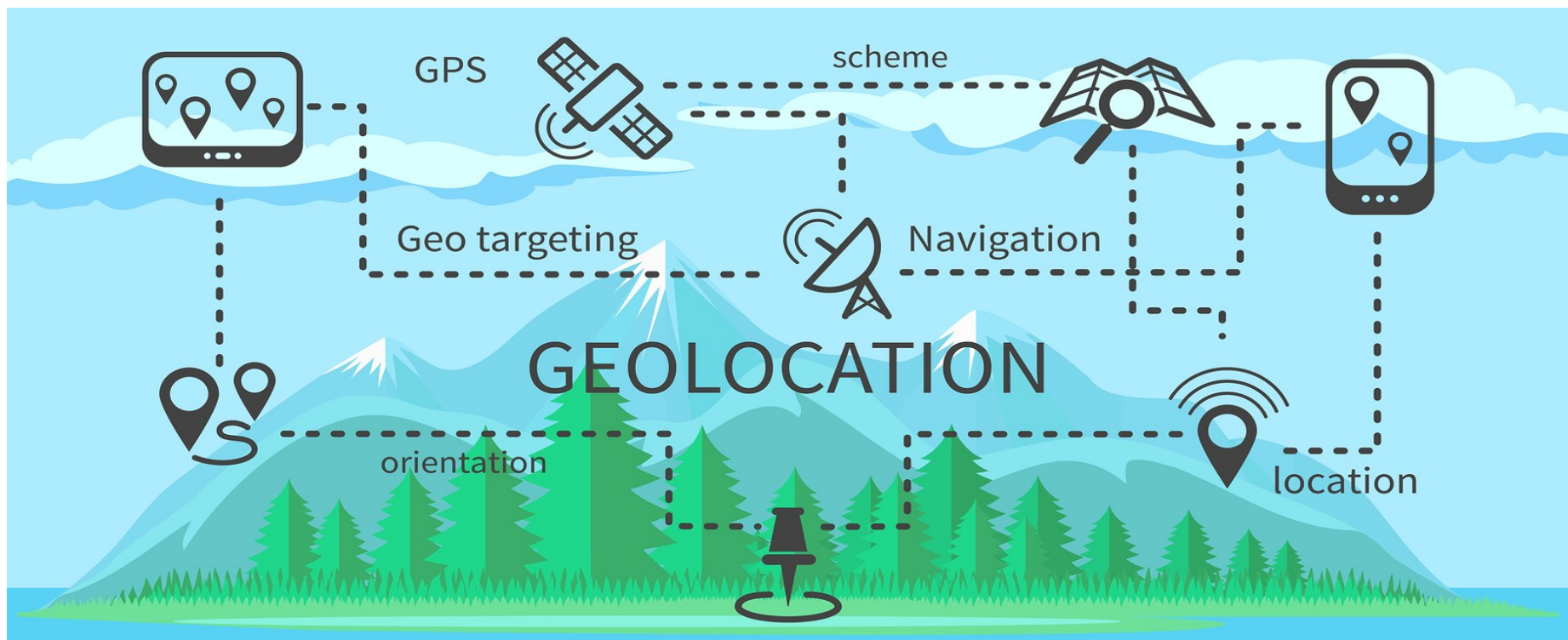
VPN-Services versuchen die Internet-Nutzer durch technische Funktionen vor Überwachung und Verfolgung im Internet zu schützen. Normale VPN-Service bietet dabei den verschlüsselten Zugang (VPN-Tunnel - Datenverkehr fließt ohne Einsehbarkeit für Dritte) zu eigenen VPN-Servern an, welche in verschiedenen Ländern betrieben werden. Die Daten welche zwischen dem Endgerät (VPN-Client) und dem VPN-Service (VPN-Server) übertragen werden sind dabei verschlüsselt. Über den VPN-Service wird dabei auch die benutzte IP-Adresse des Nutzers nach Außen hin verändert und seine eigene IP-Adresse erscheint in der Kommunikation mit Webseiten und Webservices nicht mehr.

## Geotargeting 4/4

**Zusammenfassung:** Eine bekannte IP-Adresse kann über das Geotargeting nicht immer einem Benutzer zugeordnet werden. Über verschiedene Verfahren und über spezialisierte Datenbanken kann zumindest **immer** der Standort des letzten Internetknoten ermittelt werden. Der Standort des Internet-Benutzers im Bereich der »letzten Meile« kann über lokale Messverfahren oder über nicht immer erlaubte andere Verfahren ermittelt werden.

Der Standort eines Internetbenutzers mittels der Messung der Datenlaufzeiten kann regional relativ genau sein. Durch Veränderungen der Signalwege oder durch den Austausch von Signalleitungen und Anlagenteile in Rechenzentren, Internetknoten und Verteilerschränken verändern sich natürlich auch die Signallaufzeiten, so dass die zugehörigen Datenbanken ständig aktualisiert werden müssen. Über Open-Source-Software (ping, traceroute, ...) oder Onlinedienste (myip.is, www.db-ip.com, ...) kann man sich zumindest einige sehr technische Informationen über bekannte IP-Adressen beschaffen.

Über die Benutzung von VPN's (Virtual Private Network ist ein Dienst, der eine verschlüsselte Verbindung herstellt) kann die Erfassung des Standortes eines Internetbenutzers deutlich erschwert bis unmöglich gemacht werden.





# Deepfake - Anzeichen für Fälschungen 1/12

## Was ist ein Deepfake?

In den letzten Jahren haben Deepfakes viel Aufmerksamkeit erregt und sorgen immer wieder für Schlagzeilen. Was sind Deepfakes überhaupt und welche Technologie steckt dahinter?

Deepfakes sind manipulierte Audio- oder Videoaufnahmen, bei denen mithilfe von Künstlicher Intelligenz (KI) das Gesicht einer Person in einem Bild oder Video täuschend echt durch ein anderes Gesicht ausgetauscht wird. Für die Manipulation werden neuronale Netzwerke verwendet, um das Gesicht des Zielsubjekts nahtlos in das manipulierte Video einzufügen. Das Ergebnis kann beeindruckend realistisch aussehen, und es ist oft schwer zu erkennen, ob es sich um eine Fälschung handelt.

**Hinweis:** Alles, was man fälschen kann, wird auch immer zum Ziel von Kriminellen. Bereits im Jahr 2022 berichteten 66 Prozent der Cybersicherheitsexperten von Deepfakes, die bei Cyberangriffen verwendet wurden.

## Die Technologie hinter Deepfakes

Die Technologie, die Deepfakes ermöglicht, basiert auf der Verwendung von sogenannten Generative Adversarial Networks (GANs). Dabei handelt es sich um ein neuronales Netzwerk, das aus zwei Teilen besteht: dem Generator und dem Diskriminator. Der Generator erzeugt die gefälschten Inhalte, während der Diskriminator versucht, sie von echten Inhalten zu unterscheiden (Qualitätsprüfung der Arbeitsergebnisse des Generators). Diese beiden Komponenten arbeiten zusammen, um immer realistischere Deepfakes zu erstellen.

**Hinweis:** Hochwertige Video-Fälschungen können mittlerweile mit frei verfügbarer Open-Source-Software automatisiert erstellt werden. Und auch die gezielte Manipulation von Audio-Dateien ist technisch so weit ausgereift, dass einer Person in der digitalen Welt ein beliebiger Satz



sozusagen »in den Mund gelegt« werden kann.

## Die ethischen Bedenken bei Deepfakes

Die Verbreitung von Deepfakes wirft eine Reihe ethischer Bedenken auf. Mit dieser Technologie können Personen in Videos oder Audioaufnahmen dargestellt werden, in denen sie nie waren oder Dinge sagen lassen, die sie nie gesagt haben. Dies kann zu Verwirrung, falschen Informationen und Betrug führen. Zudem kann es die Privatsphäre von Personen verletzen und zu Cyber-Mobbing oder Erpressung missbraucht werden.

Die Auswirkungen von Deepfakes auf die Gesellschaft sind vielfältig. Einerseits können sie für Unterhaltungszwecke eingesetzt werden, beispielsweise in Filmen oder Werbespots, um Schauspieler in verschiedenen Rollen zu zeigen. Andererseits bergen sie jedoch auch das Potenzial für Missbrauch und Manipulation.

# Deepfake - Anzeichen für Fälschungen 2/12

Ein Beispiel für den möglichen Missbrauch von Deepfakes ist die Verbreitung von gefälschten Nachrichten oder politischen Reden. Durch die Manipulation von Videoaufnahmen können politische Führer oder andere einflussreiche Personen in Situationen gebracht werden, die sie in Wirklichkeit nie erlebt haben. Dies kann dazu führen, dass Menschen falsche Informationen erhalten und ihre Meinungen und Entscheidungen auf der Grundlage dieser gefälschten Inhalte bilden.

Ein weiteres ethisches Problem im Zusammenhang mit Deepfakes ist die Verletzung der Privatsphäre. Indem das Gesicht einer Person in ein Video eingefügt wird, kann ihre Identität gestohlen oder missbraucht werden. Dies kann zu ernsthaften Konsequenzen führen, wie zum Beispiel Erpressung oder Rufschädigung.

Um den Missbrauch von Deepfakes einzudämmen, sind einige Maßnahmen erforderlich. Eine Möglichkeit besteht darin, die Entwicklung von Technologien zur Erkennung von Deepfakes voranzutreiben. Durch den Einsatz von KI und maschinellem Lernen können Algorithmen entwickelt werden, die Deepfakes erkennen und von echten Inhalten unterscheiden können. Darüber hinaus ist es wichtig, die Öffentlichkeit über die Existenz von Deepfakes aufzuklären und sie für die Risiken zu sensibilisieren.

Insgesamt ist es wichtig, dass die Gesellschaft sich der Herausforderungen bewusst ist, die mit der Verbreitung von Deepfakes einhergehen. Nur durch eine Kombination aus technologischen Fortschritten, Aufklärung und rechtlichen Maßnahmen können wir die potenziellen negativen Auswirkungen von Deepfakes minimieren und gleichzeitig die Vorteile dieser Technologie nutzen.

## Berühmte Deepfake-Fälle

Deepfakes haben bereits in verschiedenen Bereichen für Aufsehen gesorgt, darunter in der Politik und der Unterhaltungsindustrie.

Die Technologie der Deepfakes hat in den letzten Jahren erhebliche Fortschritte gemacht und ihre Auswirkungen sind weitreichend. Es ist wichtig, sich bewusst zu sein, wie Deepfakes unsere Gesellschaft beeinflussen können.

**Hinweis:** Die Deepfake-Technologie wird sich in naher, wie auch in ferner Zukunft ständig weiter entwickeln.

## Deepfakes in der Politik

Einer der bekanntesten politischen Deepfake-Fälle war ein Video, das den damaligen US-Präsidenten Barack Obama zeigte, wie er über Deepfake-Technologie sprach. Das Video war jedoch eine Fälschung und wurde von einem Künstler erstellt, um auf die Gefahren von Deepfakes hinzuweisen.

Politische Deepfakes können ernsthafte Auswirkungen haben, da sie das Potenzial haben, das Vertrauen der Öffentlichkeit in politische Führer zu untergraben. Es ist wichtig, dass wir lernen, Deepfakes zu erkennen und kritisch zu hinterfragen, um die Integrität unserer politischen Prozesse zu schützen.

## Deepfakes in der Unterhaltungsindustrie

Auch in der Unterhaltungsindustrie wurden Deepfakes eingesetzt, um Stars in Filmen oder Musikvideos auftreten zu lassen, in denen sie nie mitgespielt haben. Dies kann sowohl zu urheberrechtlichen Problemen führen als auch das Image der betroffenen Personen beeinflussen.

Die Verwendung von Deepfakes in der Unterhaltungsindustrie wirft ethische Fragen auf, insbesondere in Bezug auf die Zustimmung der beteiligten Personen. Es ist wichtig, dass wir die Privatsphäre und die Rechte der Künstler respektieren und sicherstellen, dass Deepfakes nicht missbräuchlich verwendet werden.

# Deepfake - Anzeichen für Fälschungen 3/12

Es ist auch interessant zu beachten, dass Deepfakes in der Unterhaltungsindustrie neue Möglichkeiten eröffnen können, wie zum Beispiel die Wiederbelebung verstorbener Schauspieler für Filme oder die Schaffung von virtuellen Stars. Diese Entwicklungen werfen jedoch auch Fragen nach Authentizität und Identität auf, die weiter erforscht werden müssen.

## Wie man Deepfakes erkennt

Angesichts der Gefahren von Deepfakes ist es wichtig zu wissen, wie man sie erkennen kann. Es gibt verschiedene Tools und Techniken, die bei der Identifizierung von Deepfakes behilflich sein können.

Deepfakes sind manipulierte Medieninhalte, bei denen künstliche Intelligenz verwendet wird, um Gesichter in Videos oder Bildern auszutauschen. Diese Technologie hat das Potenzial, große Auswirkungen auf die Gesellschaft zu haben, da sie es ermöglicht, gefälschte Inhalte zu erstellen, die echt aussehen.

Um Deepfakes zu erkennen, ist es wichtig, auf bestimmte Merkmale und Anzeichen zu achten. Es gibt mittlerweile eine Reihe von Tools, die speziell entwickelt wurden, um Deepfakes zu erkennen. Diese Tools analysieren Aspekte wie die Bewegung des Gesichts, die Audioqualität und andere Merkmale, um Anzeichen für Fälschungen zu finden.

## Tools zur Erkennung von Deepfakes

Es gibt mittlerweile eine Vielzahl von Tools, die bei der Erkennung von Deepfakes helfen können. Ein Beispiel dafür ist das Tool »DeepFaceLab«, das auf maschinellem Lernen basiert und in der Lage ist, gefälschte Gesichter in Videos zu identifizieren. Ein weiteres Tool ist »Sensity«, das auf KI-Algorithmen basiert und in der Lage ist, Deepfakes in Echtzeit zu erkennen. Ein weiteres Tool, das bei der Erkennung von Deepfakes hilfreich sein kann, ist »Forensically«.

Dieses Tool analysiert die Bildqualität und die Metadaten eines Bildes, um Anzeichen für Manipulationen zu finden. Es kann beispielsweise Unregelmäßigkeiten in den Pixeln oder Artefakte aufdecken, die auf eine Fälschung hinweisen.

Neben speziellen Tools gibt es auch einige allgemeine Tipps, die helfen können, Deepfakes zu erkennen. Dazu gehören die Überprüfung der Quelle des Videos, das Überprüfen von Verschiebungen in den Gesichtszügen oder fehlende Reflexionen in den Augen und eine kritische Betrachtung von verdächtig wirkenden Inhalten.

## Tipps zur Identifizierung von Deepfakes

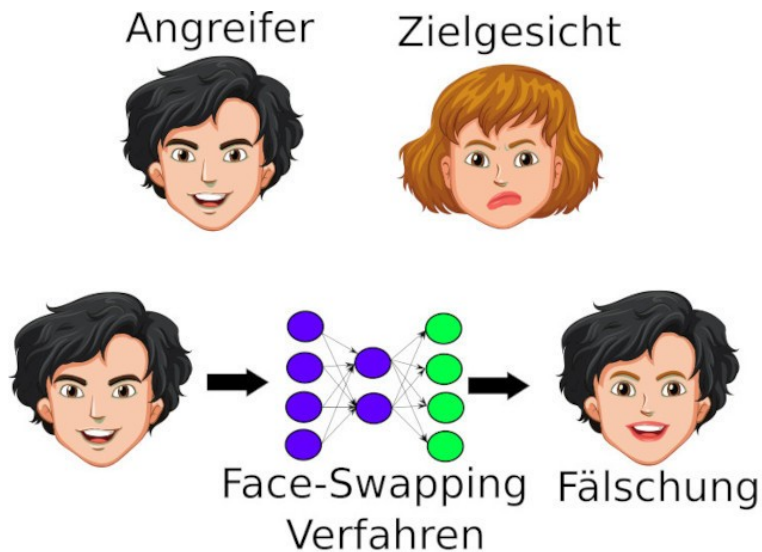
Um Deepfakes zu erkennen, ist es wichtig, auf bestimmte Merkmale und Anzeichen zu achten:

- **Überprüfung der Quelle eines Videos:** Man sollte versuchen herauszufinden, woher das Video stammt und ob es von einer vertrauenswürdigen Quelle stammt.
- **Man sollte auf Verschiebungen in den Gesichtszügen achten:** Wenn sich die Gesichtszüge einer Person unnatürlich oder ungleichmäßig bewegen (Mimik, Blinzeln, Stirnrunzeln, Zornesadern, Übergang zu den Haaren, Anzahl der Finger und Zehen, Muttermale, Tattoos, nicht zusammenpassende Ohrringe oder Schmuck, verformte Brillen, ungewöhnliche Ohr-, Nasen- und Zahnformen, Proportionen, Körperhaltung, ...), könnte dies ein Hinweis auf einen Deepfake sein. Videos sollte man sich unbedingt auch in Zeitlupe, Slow Motion anschauen, um Manipulationen zu entdecken.
- **Man sollte auf unterschiedliche Bildqualitäten achten:** Wenn innerhalb eines Bildes oder Videos unterschiedliche Bildqualitäten auftauchen oder unlogische Übergänge innerhalb eines Bildes wahrnehmbar sind, könnte dies ein Hinweis auf einen Deepfake, Fälschung sein.

# Deepfake - Anzeichen für Fälschungen 4/12

- **Man sollte nach fehlenden Reflexionen in den Augen suchen:** Echte Augen reflektieren Licht, während Deepfake-Augen möglicherweise keine Reflexionen aufweisen. Bilder und Videos sollte man sich auch in Vergrößerungen oder auf einen qualitativ hochwertigen und größeren Monitor ansehen.
- **Man sollte auf verdächtig wirkende Inhalte achten:** Wenn etwas zu gut aussieht (unrealistisch), um wahr zu sein, könnte es sich um einen Deepfake handeln. Auch auf die Hintergründe der Bilder, Videos und Sounddateien sollte man ebenfalls achten.

Es ist wichtig, dass die Öffentlichkeit über Deepfakes informiert ist und die Fähigkeit entwickelt, sie zu erkennen (Medienkompetenz). Mit diesen ständig aktualisierten Fähigkeiten kann man dazu beitragen, die Verbreitung von gefälschten Inhalten einzudämmen und damit kann man auch selbst dazu beitragen die Integrität und Vertrauenswürdigkeit von Medieninhalten schützen.



Beim Face-Swapping wird das Gesicht zweier Personen teilweise miteinander verschmolzen.

## Fälschung von Gesichtern (Face-Swap)

Zur Manipulation von Gesichtern in Videos (siehe linkes Bild) wurden in den letzten Jahren mehrere KI-basierte Verfahren entwickelt. Diese Manipulationen verfolgen entweder das Ziel Gesichter in einem Video zu tauschen (Face Swapping), die Mimik oder Kopfbewegungen einer Person in einem Video nach Wunsch zu kontrollieren (Face Reenactment) oder neue Pseudo-Identitäten zu erschaffen.

Beim »Face Swapping« Verfahren, besteht das Ziel darin, aus der Eingabe eines Gesichts einer Person, ein Gesichtsbild einer anderen Person mit derselben Mimik, Gesichtsbeleuchtung und Blickrichtung zu erzeugen. Hierfür wird in gängigen öffentlichen Softwarebibliotheken ein Autoencoder-Verfahren als Modell verwendet. Diese neuronalen Netze lernen aus einem Gesichtsbild die relevanten Mimik- und Beleuchtungsinformationen kodiert zu extrahieren und aus den kodierten Informationen ein entsprechendes Gesichtsbild zu erzeugen.

Beim »Face Reenactment« werden die Kopfbewegung (Body Puppetry), Mimik (Puppet-Master) oder Lippenbewegung (Lip-Syncing) einer Person manipuliert. Dies ermöglicht es, visuell täuschend echte Videos zu erstellen, bei denen eine Person Aussagen trifft, die sie in der Realität nie getätigt hat. Populäre Verfahren erreichen dies durch Erzeugung eines 3D-Modells des Gesichts der Zielperson anhand eines Videostreams. Dieses kann der Manipulator dann beliebig mit seinem eigenen Videostream kontrollieren und täuschend echte Gesichtsausdrücke bei der Zielperson erzeugen.

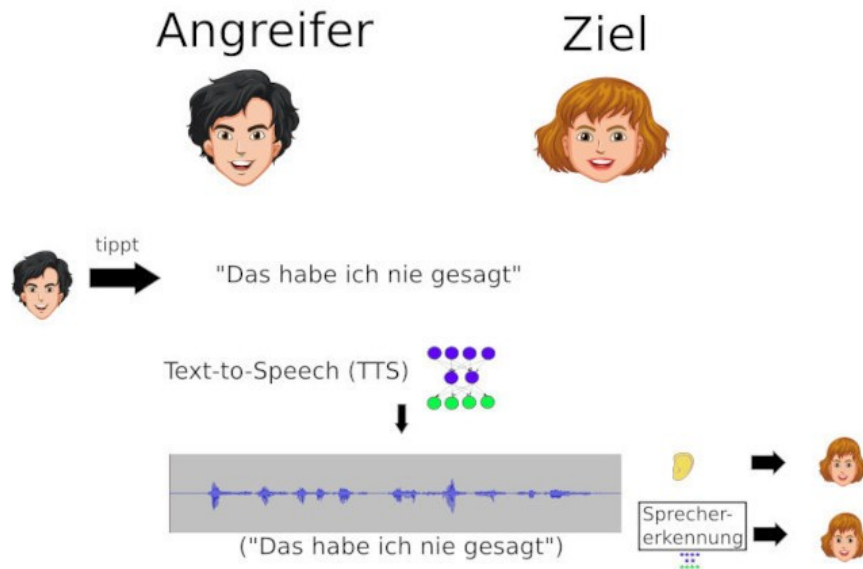
**Hinweis:** Bei der Synthetisierung, Erschaffung von Gesichtsbildern (Gesichtssynthese) können neue Personen erzeugt werden, die in der Realität nicht existieren.

# Deepfake - Anzeichen für Fälschungen 6/12

## Fälschung von Stimmen

Für die Erstellung von manipulierten Stimmen (Audio-Deepfake) sind insbesondere die Verfahren »Text-to-Speech (TTS)« und »Voice Conversion (VC)« von großer Bedeutung.

## Text-to-Speech (TTS)

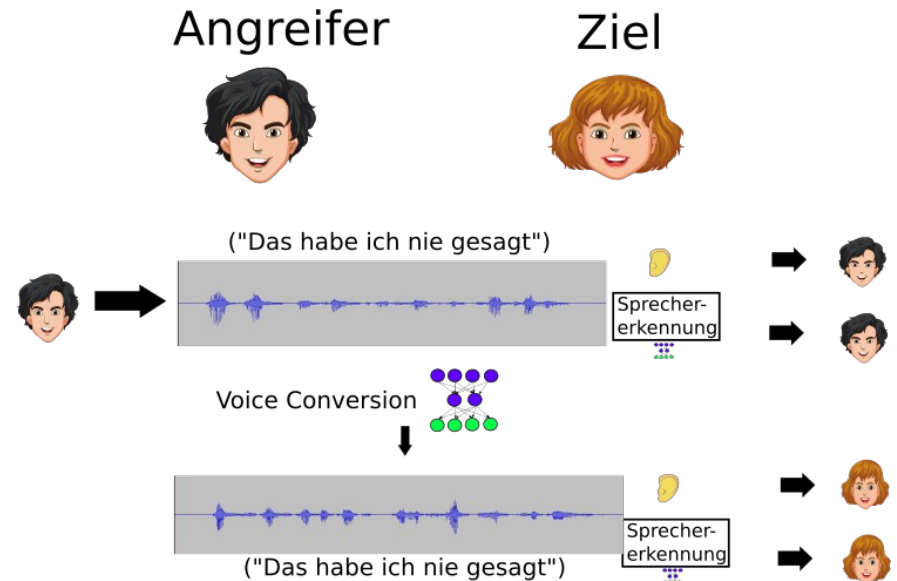


Beim Text-to-Speech-Verfahren wird zu einem vorgegebenen Text ein Audio-Signal erzeugt, welches sich sowohl für den Menschen als auch für eine automatische Sprecher-Erkennung wie die Zielperson anhört.

Die prinzipielle Funktionsweise von »Text-to-Speech«-Verfahren ist in der ersten Abbildung skizziert. Hierbei kann ein Anwender einen Text

vorgeben, welcher durch das TTS- System verarbeitet und in ein Audio-Signal umgewandelt wird. Der semantische Inhalt dieses Signals entspricht dem des vorgegebenen Textes und die sprecherspezifischen Charakteristika entsprechen im Idealfall einer durch den Anwender spezifizierten Zielperson. Hiermit können prinzipiell sowohl Menschen als auch automatisierte Sprecherkennungsverfahren getäuscht werden.

## Voice Conversion



Bei einem Voice Conversion Verfahren wird hingegen ein Audiosignal zu einer Zielstimme konvertiert.

Die prinzipielle Funktionsweise eines Voice-Conversion-Verfahrens ist in der zweiten Abbildung skizziert. Hierbei hat ein Anwender die



# Deepfake - Anzeichen für Fälschungen 7/12

Möglichkeit dem VC-System ein Audio-Signal vorzugeben, welches durch dieses zu einem manipulierten Audiosignal konvertiert wird. Dieses erzeugte Audio-Signal hat dabei den gleichen semantischen Inhalt wie das Ursprungssignal, unterscheidet sich jedoch in der zu hörenden Charakteristik des Sprechers oder Sprecherin. Dieses neue Audio-Signal gleicht hierbei im Idealfall einer durch den Angreifer ausgewählten Zielperson.

Damit diese Verfahren funktionieren, müssen sie zunächst mittels Trainingsdaten trainiert werden. Die Art der benötigten Daten unterscheidet sich je nach Angriffsart, wobei alle Verfahren die Gemeinsamkeit haben, dass von der Zielperson Audio-Aufnahmen in einer möglichst hohen und konstanten Qualität benötigt werden.

Da sowohl TTS-, als auch VC-Verfahren in der Regel durch komplexe neuronale Netze umgesetzt werden, sind für das Training Audiomaterial der Zielperson notwendig, um eine hohe Qualität zu erreichen.

**Hinweis:** Moderne aktuelle Verfahren, benötigen lediglich wenige Sekunden Audiomaterial der Zielperson und keinen erneuten Trainingsprozess.

## Fälschung von Texten

Verfahren zur Generierung von Texten, welche auf neuronale Netze basieren, schaffen es durch neue KI- Modelle, große Textdatenbanken und eine hohe Rechenleistung, zusammenhängende Texte zu schreiben. Bei diesen kann auf den ersten Blick nicht unterschieden werden, ob sie von einem Menschen oder von einer Maschine geschrieben wurden. Meist sind nur wenige einleitende Wörter notwendig, aus denen das Modell eine mögliche, plausible Fortsetzung des Texts generiert. Damit können Nachrichten verfasst, Blog-Einträge erzeugt und auch Chat-Antworten generiert werden.

Noch sind die notwendigen Ressourcen zum Training des Systems und Anwendung der leistungsstarken Modelle jenseits dessen, was für »normale« Personen üblich ist. Daher müssen Privatpersonen auf öffentlich zugängliche Clouddienste zurückgreifen. Bei fortschreitender Weiterentwicklung der Technik ist damit zu rechnen, dass diese Einsatz in Chatbots oder Social Bots finden, um einen fiktiven Gesprächspartner zu simulieren.

## Mögliche Bedrohungsszenarien

Mittels der beschriebenen Verfahren ist es heute auch teilweise schon für technisch versierte Laien möglich, mediale Identitäten zu manipulieren, wodurch sich zahlreiche Bedrohungsszenarien ergeben:

- **Überwindung biometrischer Systeme:** Da es mittels Deepfake-Verfahren möglich ist, mediale Inhalte mit den Charakteristika einer Zielperson zu erstellen und diese Verfahren teilweise bereits in Echtzeit lauffähig sind, stellen sie eine hohe Gefahr für biometrische Systeme dar. Insbesondere bei Fernidentifikationsverfahren (z.B. der Sprecher-Erkennung über das Telefon oder der Videoidentifikation) scheinen solche Angriffe durchaus möglich zu sein.
- **Social Engineering:** Deepfake-Verfahren können außerdem dazu verwendet werden, gezielte Phishing-Angriffe (Spear-Phishing) durchzuführen, um Informationen und Daten zu gewinnen. Auch kann ein Angreifer diese Technologie zur Durchführung von Betrug und zur Abschöpfung finanzieller Mittel nutzen. Beispielsweise könnte er eine Person mit der Stimme von deren Führungskraft anrufen, um eine Geldtransaktion auszulösen (CEO-Fraud oder CEO-Betrug).
- **Desinformationskampagnen:** Mittels Deepfake-Verfahren ist es potentiell möglich, glaubwürdige Desinformationskampagnen durchzuführen, indem manipulierte Medieninhalte von Schlüsselpersonen erzeugt und verbreitet werden.

# Deepfake - Anzeichen für Fälschungen 8/12

- **Verleumdung:** Durch die Möglichkeit Medieninhalte zu generieren, die Personen beliebige Aussagen treffen lassen und sie in beliebigen Situationen darstellen, kann der Ruf einer Person durch die Verbreitung von Unwahrheiten nachhaltig geschädigt werden (Cybermobbing).

## Gegenmaßnahmen

Es gibt viele Ansätze, um sich gegen die beschriebenen Methoden zu verteidigen, wobei sich diese in die zwei Kategorien Prävention und Detektion untergliedern lassen.

### 1. Prävention

Gegenmaßnahmen aus dem Gebiet der Prävention zielen darauf ab, das Risiko eines erfolgreichen Angriffs mittels Deepfakes zu senken.

### Aufklärung

Eine zentrale Maßnahme gegen Deepfake- Angriffe stellt die Schulung dar (Medienkompetenz). Zum einen ist davon auszugehen, dass das Wissen über die Möglichkeit eines solchen Angriffs eine differenzierte Einschätzung der Echtheit des gesehenen oder gehörten Materials unter Berücksichtigung der Quelle ermöglicht. Zum anderen erzeugen viele Deepfake-Verfahren teilweise deutliche Artefakte (Fehler). Durch die Kenntnis dieser möglichen Artefakte kann die Erkennung von Fälschungen deutlich gesteigert werden. Insbesondere bei Echtzeitanwendungen hat ein Angreifer nicht die Möglichkeit, mit Artefakten behaftetes Material manuell zu bereinigen.

### Typische Artefakte bei Gesichtsm Manipulationen

- **Sichtbare Übergänge:** Bei einem Face-Swapping-Verfahren wird ein Gesicht der Zielperson in den Kopf einer anderen Person eingesetzt. Dadurch kann es zu sichtbaren Artefakten an der Naht rund um das Gesicht kommen. Ebenso ist es möglich, dass die

Hautfarbe und -textur an diesem Übergang wechselt oder dass sich teilweise das Ursprungsgesicht in manchen Frames am Gesichtsrand durch doppelte Augenbrauen bemerkbar macht.

- **Scharfe Konturen verwaschen:** Häufig kommt es **noch** vor, dass Face-Swapping-Verfahren nicht richtig lernen, scharfe Konturen, wie sie in den Zähnen oder im Auge vorkommen, zu erzeugen. Bei genauem Hinsehen wirken diese auffällig verwaschen.
- **Begrenzte Mimik, unstimmige Beleuchtung:** Auf Grund einer beschränkten Datenlage kann es dazu kommen, dass ein Modell nur beschränkt fähig ist manche Gesichtsausdrücke oder Beleuchtungssituationen korrekt darzustellen. Häufig ist die Profilansicht eines Gesichts unzureichend erlernt, sodass ein starkes Drehen des Kopfes zu Bildfehlern führen kann, bei welchen zum Beispiel das Gesicht verwaschen erscheint.

### Typische Artefakte bei synthetischen Stimmen

- **Metallischer Sound:** Zahlreiche Verfahren, erzeugen mitunter ein Audio-Signal, das vom menschlichen Gehör als »metallisch« wahrgenommen wird.
- **Falsche Aussprache:** Häufig können TTS-Verfahren (Text-to-Speech-Verfahren) nicht alle Wörter korrekt aussprechen. Dies kann beispielsweise passieren, wenn ein TTS-Verfahren für die deutsche Sprache trainiert wurde, aber ein englisches Wort ausgesprochen werden soll.
- **Monotone Sprachausgabe (Sprachmelodie):** Insbesondere wenn die Trainingsdaten für ein TTS-System nicht ideal sind, kann das erzeugte Audio-Signal sehr monoton hinsichtlich der Betonung der Wörter sein.
- **Falsche Sprechweise (Sprachmelodie):** Meist sind Fälschungsverfahren vergleichsweise gut dafür geeignet, die Klangfarbe einer Stimme zu fälschen, haben jedoch häufig

# Deepfake - Anzeichen für Fälschungen 9/12

Probleme damit, die spezifischen Charakteristika der Stimme zu fälschen, sodass beispielsweise Akzente oder Betonungen von Wörtern nicht denen des Zielsprechers oder der Zielsprecherin entsprechen.

- **Unnatürliche Geräusche:** Sofern ein Fälschungsverfahren Eingangsdaten erhält, die stark von den beim Training verwendeten abweichen, kann das Verfahren unnatürliche Geräusche erzeugen. Dies kann beispielsweise ein zu langer Text bei einem Text-to-Speech-Verfahren oder Stille bei einem Voice-Conversion-Verfahren sein.
- **Hohe Verzögerung:** Die meisten Verfahren zur Erzeugung von synthetischen Stimmen müssen zunächst einen Teil des zu erzeugenden semantischen Inhalts als Eingangsdaten empfangen, um ein qualitativ hochwertiges Ergebnis zu erzeugen. Dies führt dazu, dass qualitativ hochwertige Fälschungen in vielen Fällen mit einer gewissen zeitlichen Verzögerung einhergehen, da dieser semantische Inhalt zunächst ausgesprochen und erfasst werden muss, bevor er von einem VC/TTS Verfahren verarbeitet werden kann.

**Hinweis:** Um die Fähigkeit, manipulierte Audio-Daten zu erkennen, zu trainieren, können beispielsweise frei entwickelte Anwendungen, die mitunter auch öffentlich zugänglich sind, verwendet werden.

## Kryptographie

Kryptographische Verfahren bieten die Möglichkeit, die Quelle des Materials eindeutig an eine Identität zu binden. Dies ermöglicht die sichere Zuordnung zu einer vertrauenswürdigen Quelle (Authentizität) und stellt sicher, dass Manipulationen des Materials nach der Absicherung sofort auffallen (Integritätsschutz). Hierdurch kann jedoch nicht verhindert werden, dass die Quelle selbst das Material zuvor manipuliert. Aktuelle Entwicklungen beschäftigen sich beispielsweise mit der Erstellung einer digitalen Signatur beim

Aufnahmeprozess, wodurch sichergestellt wird, dass das Material nicht mehr nach der Aufnahme manipuliert wurde.

## Gesetzlich

Gesetzliche Regelungen **können** eine Hürde darstellen, Deepfakes ungekennzeichnet in Umlauf zu bringen. Insbesondere wird in dem Regulierungsentwurf der EU-Kommission zu KI-Systemen gefordert, dass alle mit der Deepfake-Technologie erstellten Materialien als solche gekennzeichnet werden müssen (Stand: 2025). Deutschland unterscheidet bislang nicht in einer gesonderten Gesetzgebung zwischen legalen und kriminellen Deepfakes. Doch die bestehende Rechte und Gesetze schützen Bürger bereits jetzt vor dem willkürlichen Missbrauch mit Deepfakes (Schutz der Menschenwürde, Recht am eigenen Bild, Beleidigung, Verleumdung und üble Nachrede, Identitätsdiebstahl, Verbreitung, Erwerb und Besitz kinderpornografischer Inhalte, ...).

## 2. Detektion

Gegenmaßnahmen aus dem Bereich der Detektion zielen darauf ab, mittels Deepfake-Verfahren manipulierte Daten als solche zu erkennen.

## Medienforensisch

Mittels Methoden aus der Medienforensik ist es möglich, Artefakte zu detektieren, welche bei der Verwendung von Manipulationsmethoden auftreten. Hiermit ist es für Expertinnen und Experten möglich, Fälschungen nachvollziehbar zu erkennen.

## Automatisierte Detektion

In der Forschungsliteratur wurden in den letzten Jahren zahlreiche Methoden zur automatisierten Detektion von manipulierten Daten veröffentlicht. Diese Verfahren basieren in der Regel auf Techniken aus dem Gebiet der künstlichen Intelligenz, insbesondere den tiefen neuronalen Netzen (deep neural networks). Aufgrund dessen

# Deepfake - Anzeichen für Fälschungen 10/12

müssen diese Verfahren jedoch anhand großer Datenmengen trainiert werden. Nach der Trainingsphase kann das Modell dazu verwendet werden, für ein Datenbeispiel (z.B. ein Video) zu klassifizieren, ob dieses manipuliert wurde oder nicht.

## Herausforderungen der automatisierten Gegenmaßnahmen

Ein Problem der Gegenmaßnahmen besteht darin, dass diese entweder nicht in allen Situationen angewendet werden können und in der Regel auch keinen vollständigen Schutz bieten.

Insbesondere bei der Klasse der automatisierten Detektionsmethoden sollte darauf hingewiesen werden, dass sie häufig nur unter gewissen Rahmenbedingungen zuverlässig funktionieren. Da diese Verfahren in der Regel auf Verfahren der künstlichen Intelligenz basieren, gehen diese Methoden auch mit deren grundsätzlichen Problemen einher.

- **Generalisierbarkeit:** Ein zentrales Problem der meisten Detektionsmethoden ist ihre mangelhafte Generalisierbarkeit (Verallgemeinerbarkeit). Da die Methoden auf bestimmten Daten trainiert wurden, funktionieren sie häufig nur auf ähnlichen Daten relativ zuverlässig. Werden jedoch einzelne Parameter verändert, so ist die Korrektheit der Ausgaben häufig nicht gegeben. Ein wichtiges Beispiel eines solchen Parameters, kann der Wechsel zu einer anderen Angriffsmethode, welche nicht in den Trainingsdaten vorhanden war, sein. Dieses Verhalten konnte beispielsweise in der Deepfake Detection Challenge (2020) beobachtet werden, in welcher selbst das beste Modell lediglich eine durchschnittliche Genauigkeit von 65,18 Prozent erreichen konnte, wobei eine Genauigkeit von 50 Prozent durch bloßes Raten erreicht werden würde.
- **KI-spezifische Angriffe:** Ein weiteres zentrales Problem dieser Verfahren besteht darin, dass sie durch KI-spezifische Angriffe überwunden werden können, wobei insbesondere adversariale

(feindliche, kontroverse) Angriffe eine besondere Bedrohung darstellen. So kann ein Angreifer beispielsweise ein gezieltes Rauschen erstellen, welches über das mittels eines Face-Swapping- Verfahrens manipulierte Bild gelegt wird. Dieses Rauschen kann so klein sein, dass es für den menschlichen Betrachter nicht zu bemerken ist, hat jedoch für das Detektionsverfahren den Effekt, dass es die Fälschung nicht als solche klassifiziert. Solche Angriffe lassen sich nicht komplett vermeiden, jedoch sollten sie bei der Erstellung von Detektionsverfahren berücksichtigt werden, um die Hürden für einen Angreifer zu erhöhen.

## Ausblick – Deepfake-Technologie

Die Technologie zur Fälschung medialer Identitäten hat sich in den letzten Jahren insbesondere durch die Fortschritte im Bereich der künstlichen Intelligenz deutlich weiterentwickelt. Aktuelle Forschungsergebnisse deuten darauf hin, dass sich dieser Trend weiter fortsetzen wird, sodass die manuelle Erkennung von Fälschungen in Zukunft immer schwieriger werden wird.

Auch ist davon auszugehen, dass sich die Menge der von der angegriffenen Person benötigten Daten stetig verringern wird. Für technisch versierte Laien ist es bereits heute möglich, qualitativ hochwertige Fälschungen zu erstellen.

Es ist jedoch davon auszugehen, dass sich die benötigte Expertise (Sachkenntnis) und der notwendige Aufwand zur Erstellung von Fälschungen durch die Verbesserung und erhöhte Verfügbarkeit an öffentlichen Tools stetig verringern wird, sodass sich die Häufigkeit von Angriffen mittels dieser Technologie signifikant erhöhen könnte.

Es gibt jedoch auch noch weitere erhebliche Risiken und Herausforderungen im Zusammenhang mit Deepfakes. Neben den

# Deepfake - Anzeichen für Fälschungen 11/12

ethischen Bedenken können Deepfakes dazu führen, dass Menschen immer skeptischer gegenüber der Echtheit von Audio- und Videoaufnahmen werden. Dies hat Auswirkungen auf das Vertrauen in die Medien und die Informationsverbreitung.

Aus diesen Gründen ist es von hoher Bedeutung, dass die aufgeführten Gegenmaßnahmen weiterentwickelt und in Kombination nach einer applikationsspezifischen Auswahl von Fachleuten eingesetzt werden.

Obwohl Deepfakes hauptsächlich negativ in den Medien präsent sind, gibt es auch potenziell positive Anwendungen. Insbesondere in den Bereichen Unterhaltung, Online-Handel, Marketing und Gesundheit birgt die Technologie das Potenzial, den Alltag von Menschen zu erleichtern und für immer zu verändern.

## Positive Nutzung der Deepfake-Technologie

**Unterhaltungsbranche - Schauspieler und Schauspielerinnen verjüngen:** In der Unterhaltungsindustrie kann die Deepfake-Technologie z.B. dazu verwendet werden, Filme oder andere kreative Werke zu erstellen, in denen nicht real existierende Charaktere auftreten.

Für die nahe Zukunft ist zu erwarten, dass die Deepfake-Technologie mit 3D- und Spezialeffekten kombiniert wird und die Filmbranche dadurch für immer revolutioniert wird.

**Online-Handel - Mode virtuell anprobieren:** Online-Shopping könnte in Zukunft noch zeitsparender und bequemer werden als bislang - nämlich dann, wenn Retouren weitgehend wegfallen. Mithilfe von Ganzkörper-Deepfakes könnten Kunden das Outfit virtuell anprobieren und sich dabei drehen und wenden wie in der

Umkleidekabine. Dadurch ließe sich schon vor dem Bestellen qualifiziert entscheiden, ob ein Outfit passt und gefällt. Ähnliche Möglichkeiten böten sich für den Bereich Frisuren und Haarfarben. Erstmals gelang es Wissenschaftlerinnen und Wissenschaftlern der Universität Berkeley im Jahr 2018, die Bewegungen eines professionellen Tänzers mittels Deepfake-Technologie auf einen fremden Körper übertragen zu lassen.

**Marketing - Werbefilme mit Deepfake-Gesichtern:** Doch nicht nur bei der Auswahl, auch bei der Bewerbung von Produkten wird künstliche Intelligenz in Zukunft eine entscheidende Rolle spielen. So müssen Prominente in Zukunft nicht mehr unbedingt selbst tätig werden, um für ein Produkt zu werben. Der Schauspieler Bruce Willis war im Jahr 2021 beispielsweise in Werbefilmen des russischen Mobilfunkanbieters Megafon zu sehen, für die er nie das Filmset betreten hatte. Ein russischer Schauspieler übernahm seine Rolle, dessen Gesicht via Deepfake durch das von Bruce Willis getauscht wurde. Trainiert wurde das Deepfake-Gesicht mithilfe von 34.000 Bildern des Schauspielers aus rund 40 Filmen.

**Gesundheit – Stimmverlust rückgängig machen:** Als Hilfsmittel im Gesundheitsbereich könnte die Deepfake-Technologie Menschen, die ihre Stimme durch eine Krankheit verloren haben, ein Stück Normalität zurückgeben. Das Unternehmen VocaLid hat sich zum Beispiel darauf spezialisiert, Stimmen auf Basis von Audioschnipseln und künstlicher Intelligenz zu konservieren. Die digitale Stimme, die mithilfe der Deepfake-Technologie imitiert wird, kann hinterher auf einem digitalen Endgerät gespeichert werden und per Spracheingabe in einer App ausgegeben werden.

## Negative Nutzung der Deepfake-Technologie

So faszinierend die positiven Einsatzzwecke von Deepfakes sind, so



# Deepfake - Anzeichen für Fälschungen 12/12

erheblich ist auch das Missbrauchspotenzial der Technologie – vor allem, wenn sie durch die Verbreitung von Deepfake-Apps in die Hände der Allgemeinheit fällt. Smartphones als Hilfsinstrument erlauben es dabei, heimlich Fotos und Videos von Personen aus dem direkten Umfeld (Nachbarschaft, Schule, Unternehmen, ...) zu erstellen, die manipuliert werden können, um andere zu verleumden. Auch Betrugsstrategien und Desinformationskampagnen profitieren vom Einsatz von Deepfakes.

**Verleumdung - Deepfakes als Steilvorlage für Cybermobbing:** Mit kaum einer anderen Technologie kann man dem Ansehen von öffentlichen oder privaten Personen so leicht Schaden zufügen wie mit Deepfakes.

Mithilfe der Technologie können zum Beispiel glaubwürdige Fake-Pornos (Deepnude), gefälschte Social-Media-Profile oder Fake-Videos von kriminellen Handlungen erstellt werden. Diese können wiederum für Racheaktionen an Ex-Partnerinnen und Ex-Partnern oder zur Erpressung von mächtigen oder reichen Persönlichkeiten zum Einsatz kommen. Im Schulkontext ist damit zu rechnen, dass Kinder und Jugendliche Deepfakes für Cybermobbing einsetzen und es damit noch einfacher wird, Mitschüler oder auch Lehrkräfte bloßzustellen.

Mädchen und Frauen sind besonders vom Missbrauch betroffen. Im Jahr 2021 waren laut Schätzung des Forschungsunternehmens Sensity AI 90 bis 95 Prozent aller im Internet kursierenden Deepfake-Videos Fake-Pornos - mit etwa 90 Prozent weiblichen Opfern. Während KI-Experte Henry Ajder für das Jahr 2018 14.000 Deepfake-Pornos im Netz zählte, könne ihre Zahl inzwischen nicht mehr erfasst werden und rangiere im Milliardenbereich. Betroffen sind längst nicht mehr ausschließlich prominente Schauspielerinnen, Sängerinnen oder Politikerinnen. In Bilderforen und Tauschbörsen im Netz werden

inzwischen massenhaft und ungefragt Bilder von unbekannten Mädchen und Frauen hochgeladen, die bereits mittels der Deepfake-Technologie verändert oder erstellt wurden.

**Betrug - mit Stimmen- und Gesichtsklone Geld erbeuten:** Wie ein prominenter Fall aus dem Jahr 2019 zeigt, eignet sich Deepfake-Technologie auch hervorragend, um Betrugsmaschen im Unternehmens- und Privatbereich auf eine neue Ebene zu heben. So tätigte der Geschäftsführer eines britischen Energieunternehmens eine Überweisung in Höhe von 243.000 US-Dollar auf das vermeintliche Bankkonto eines ungarischen Lieferanten, nachdem er mit der Fake-Stimme des Vorstandsvorsitzenden der deutschen Muttergesellschaft telefoniert hatte. Ein Betrüger hatte zur Nachahmung der Stimme eine KI-Sprachsoftware verwendet.

**Desinformationen als politische Waffe:** »Ein Bild sagt mehr als tausend Worte« - Maßgeschneiderte Fotos und Videos für jede Fake-News-Meldung im Netz könnten in Zukunft die Verbreitung und Glaubwürdigkeit von **Desinformationen** nochmals erhöhen.

**Fazit:** Insgesamt bleibt festzuhalten, dass Deepfakes eine Technologie sind, die sowohl faszinierende Möglichkeiten bietet als auch erhebliche Risiken mit sich bringt. Es ist wichtig, sich der Gefahren bewusst zu sein und sich aktiv damit zu beschäftigen, wie man Deepfakes identifizieren und bekämpfen kann – nämlich durch eine ausreichende Medienkompetenz.

Es gibt bereits Werkzeuge die Deepfakes mit einer etwa 90 prozentigen Sicherheit entdecken können. Sie sind aber nicht immer für alle Bürger nutzbar (TrueMedia.org).

# Taiwan Semiconductor Manufacturing Company (TSMC) 1/2

Die Taiwan Semiconductor Manufacturing Company Limited (TSMC), ist der weltweit größte unabhängige Auftragsfertiger für Halbleiterprodukte (Foundry oder Auftragsfertigung). Der Hauptsitz und die wichtigsten Unternehmensteile befinden sich in Hsinchu, Taiwan. Das Unternehmen wurde 1987 von Morris Chang gegründet, einem visionären Führer mit umfangreicher Erfahrung in der Elektronikindustrie.

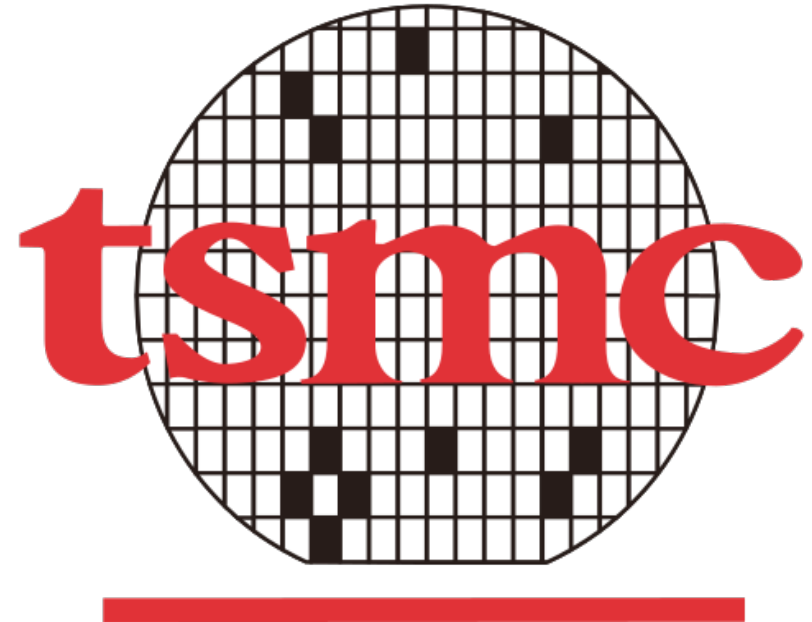
Das Geschäftsmodell ist darauf ausgerichtet, für Fabless-Unternehmen wie z.B. AMD, Apple, Qualcomm, Nvidia, Conexant, Marvell, VIA oder Broadcom die Produktion von Halbleiterchips zu übernehmen.

Das Unternehmen ist seit den 1990er Jahren sehr schnell gewachsen - in den letzten 20 Jahren lag das durchschnittliche Wachstum pro Jahr bei 21,5 %. 2022 erwirtschaftete TSMC einen Jahresumsatz von 75,9 Milliarden US-Dollar und einen Gewinn von 34,1 Milliarden US-Dollar und überholte damit Intel und Samsung.

Der Umsatz stieg seit der Gründung stetig an. Von 12 Milliarden TWD (Taiwan Dollar) im Jahr 1993 über 66 Milliarden TWD im Jahr 2000, 420 Milliarden TWD 2010 und 1339 Milliarden TWD 2020 wurde im Jahr 2023 ein Umsatz von 2162 Milliarden TWD erzielt. Am 9. August 2023 betrug die Marktkapitalisierung 488,03 Mrd. US-Dollar.

## Chronologie der Chip-Entwicklung (seit 2015)

- seit 2015 bietet TSMC Chips im 16-nm-FinFET-Verfahren (Fin Field-Effect Transistor) an
- 2017 startete man mit der 10-nm-FinFET-Fertigung für Apple
- Seit April 2018 läuft die 7-nm-FinFET-Produktion (N7), der Nachfolger (N7P) erschien ein Jahr später. Ebenfalls 2019 wurde die EUV-Lithografie in der Herstellung der 7-nm-FinFET-Variante



N7+ eingeführt. Der N6 genannte Prozess gehört ebenfalls der 7-nm-Generation an.

- Im April 2020 lief die 5-nm-FinFET-Produktion (N5) in der neuen Fab 18 in Shanhua an. Deren Ausbau der ersten drei Phasen wurde Ende 2020 vollendet. Mitte 2021 startete der Bau für eine weitere Phase (P7). Die Produktion des Nachfolgeprozesses N5P startete im Mai 2021. Auch der für 2022 geplante Prozess N4 gehört der 5-nm-Generation an.
- Im November 2020 genehmigte der Verwaltungsrat den Neubau der Fab 21 (Phoenix) in Arizona, USA. Die 5-nm-FinFET-Produktion startete dort 2024.
- Der Produktionsstart des 3-nm-FinFET-Prozesses, in erster Version N3 genannt, erfolgte im 4. Quartal 2022.

- Mit einem 2-nm-Prozess plant TSMC von FinFET auf Gate-all-around-FET (GaaFET) zu wechseln. Für die Realisierung soll in Hsinchu südwestlich des Standorts der Fab 12A (Hsinchu) ein Entwicklungszentrum sowie für die Produktion die neue Fab 20 (Hsinchu) in 4 Phasen entstehen.

Darüber hinaus investiert das Unternehmen weiterhin in die Entwicklung neuer Technologien, wie etwa in das Quantencomputing, das den Halbleitersektor weiter revolutionieren wird.

## Politische Relevanz

TSMC wird aufgrund der Marktposition und des Technologievorsprungs erhebliche geostrategische Relevanz zugesprochen. TSMC produziert (Stand 2023) mehr als die Hälfte aller Halbleiter (laut Angaben der US-Regierung mindestens 70 %) – bei den modernsten Varianten hat TSMC einen Weltmarktanteil von mehr als 90 Prozent. Diese starke Marktposition muss jedoch in Teilen relativiert werden, da der Erfolg auch durch nachfolgende Akteure der Wertschöpfungskette bedingt ist. Die US-Regierung bezeichnete TSMC Anfang 2024 als das wichtigste Unternehmen der Welt.

Nicht nur aus taiwanesischer Sicht sind Halbleiter von großer strategischer und wirtschaftlicher Bedeutung.

Dazu kommt noch, dass die wichtigsten Halbleiter-Technologie für die Welt und vor allem für die USA von einer Insel kommt, die von China als sein Eigentum beansprucht wird.

**Hinweis:** Mehr als 90 % der fortschrittlichsten Prozessoren in der Welt werden von TSMC hergestellt.

Die globale Abhängigkeit von TSMC hat auch geopolitische Bedenken aufgeworfen. Die Vereinigten Staaten und Europa verfolgen politische Strategien, um die Abhängigkeit von Asien bei der Halbleiterproduktion zu verringern.

Um dieser Abhängigkeit entgegenzuwirken, hat die US-Regierung Pläne zur Subventionierung der lokalen Chipproduktion angekündigt und Unternehmen wie Intel ermutigt, ihre Fertigungskapazitäten auszubauen. Trotz dieser Maßnahmen bleibt eine Herausforderung, in puncto Qualität und Effizienz mit TSMC zu konkurrieren.

Diese Verschmelzung von geopolitischen und technologischen Kräften ist so bedeutsam, dass informierte Kreise schon von einem »Chip Krieg« reden, der möglicherweise in einen realen Krieg übergeht.

- Mit einem 2-nm-Prozess plant TSMC von FinFET auf Gate-all-around-FET (GaaFET) zu wechseln. Für die Realisierung soll in Hsinchu südwestlich des Standorts der Fab 12A (Hsinchu) ein Entwicklungszentrum sowie für die Produktion die neue Fab 20 (Hsinchu) in 4 Phasen entstehen.

Darüber hinaus investiert das Unternehmen weiterhin in die Entwicklung neuer Technologien, wie etwa in das Quantencomputing, das den Halbleitersektor weiter revolutionieren wird.

## Politische Relevanz

TSMC wird aufgrund der Marktposition und des Technologievorsprungs erhebliche geostrategische Relevanz zugesprochen. TSMC produziert (Stand 2023) mehr als die Hälfte aller Halbleiter (laut Angaben der US-Regierung mindestens 70 %) – bei den modernsten Varianten hat TSMC einen Weltmarktanteil von mehr als 90 Prozent. Diese starke Marktposition muss jedoch in Teilen relativiert werden, da der Erfolg auch durch nachfolgende Akteure der Wertschöpfungskette bedingt ist. Die US-Regierung bezeichnete TSMC Anfang 2024 als das wichtigste Unternehmen der Welt.

Nicht nur aus taiwanesischer Sicht sind Halbleiter von großer strategischer und wirtschaftlicher Bedeutung.

Dazu kommt noch, dass die wichtigsten Halbleiter-Technologie für die Welt und vor allem für die USA von einer Insel kommt, die von China als sein Eigentum beansprucht wird.

**Hinweis:** Mehr als 90 % der fortschrittlichsten Prozessoren in der Welt werden von TSMC hergestellt.

Die globale Abhängigkeit von TSMC hat auch geopolitische Bedenken aufgeworfen. Die Vereinigten Staaten und Europa verfolgen politische Strategien, um die Abhängigkeit von Asien bei der Halbleiterproduktion zu verringern.

Um dieser Abhängigkeit entgegenzuwirken, hat die US-Regierung Pläne zur Subventionierung der lokalen Chipproduktion angekündigt und Unternehmen wie Intel ermutigt, ihre Fertigungskapazitäten auszubauen. Trotz dieser Maßnahmen bleibt eine Herausforderung, in puncto Qualität und Effizienz mit TSMC zu konkurrieren.

Diese Verschmelzung von geopolitischen und technologischen Kräften ist so bedeutsam, dass informierte Kreise schon von einem »Chip Krieg« reden, der möglicherweise in einen realen Krieg übergeht.

**FinFET:** FinFET steht für Fin Field Effect Transistor (Feldeffekttransistor mit Finnenstruktur) und bezeichnet eine 3D-Transistor-Architektur, bei der der leitfähige Kanal als dünne, vertikale Silizium-"Finne" (Flosse) herausragt, anstatt flach zu sein, was eine bessere Steuerung durch das umgebende Gate ermöglicht, Leckströme reduziert und die Leistungsdichte auf Chips erhöht. Es ist eine Schlüsseltechnologie für moderne Mikroprozessoren, die traditionelle flache Transistoren ablöst.

**GAAFET:** GAAFET steht für Gate-All-Around Field Effect Transistor (Feldeffekttransistor), eine fortschrittliche Transistor-Technologie für moderne Mikrochips, bei der das Steuerelement (Gate) den elektrischen Kanal von allen Seiten umgibt, was im Vergleich zu älteren FinFET-Transistoren eine bessere Stromsteuerung und höhere Effizienz bei kleineren Größen ermöglicht.



# Internet-Links

**Eigene IP-Adresse ermitteln:**

myip.is  
ipaddress.com  
www.ipgp.net

**GeoIP-Datenbanken**

Maxmind GeoLite2 Country:  
[https://dev.maxmind.com/geoip/geolite2-free-geolocation-data?](https://dev.maxmind.com/geoip/geolite2-free-geolocation-data?lang=en)  
lang=en

**GEO-Targeting:**

[www.db-ip.com](http://www.db-ip.com)

**Forschungszentren:**

Nationale Forschungszentrum für angewandte Cybersicherheit  
ATHENE: [www.athene-center.de](http://www.athene-center.de)  
Fraunhofer AISEC (Fraunhofer-Institut für Angewandte und  
Integrierte Sicherheit): <https://www.aisec.fraunhofer.de>

**Suchmaschinen:**

StartPage (Niederlande): <https://www.startpage.com>  
Qwant (Frankreich): [www.qwant.com](http://www.qwant.com), [www.qwant.fr](http://www.qwant.fr), [qwant.com](http://qwant.com)  
Metager (Deutschland): <https://metager.de>  
DuckDuckGo (USA): <https://duckduckgo.com>  
Internet-Zeitmaschine: <https://web.archive.org>  
Shodan: Suchmaschine für Geräte und Systeme die mit dem  
Internet verbunden sind: <https://www.shodan.io>