

■ MAGAZIN

imBlick

Im
MAGAZIN:
Aufklärung über
Desinformation und
Künstliche Intelligenz.
Ein Projekt des iRights.Lab,
gefördert durch die
Mediengestaltung Berlin
Brandenburg
(mabb).

PRAKTISCHE TIPPS:

**So erkennen Sie
KI-Fälschungen!**

AUFGEKLÄRT:

**Der tote Teenager
in Berlin –**

So sollen Sie getäuscht werden!

AUFREGER:

**Hat Baerbock DAS
gesagt!?**

Wie mit KI manipuliert wird!

**"Trauen Sie
diesen Fingern
nicht!"**



IM HEFT: Großes KI-Kreuzworträtsel!

Inhalt

Noch nie waren Informationen so leicht zugänglich wie heute. Die Fülle an Inhalten, denen man täglich begegnet, kann aber auch überfordernd sein.

Hinzu kommt, dass es immer mehr Inhalte gibt, die **fake**, also nicht echt sind.

Wenn manipulierte und irreführende Inhalte gezielt verbreitet werden, um Menschen zu täuschen, spricht man von **Desinformation**. Die Motivation dahinter ist, öffentliche Debatten und Meinungsbildung zu beeinflussen.

Die Anzahl manipulierter Nachrichten, Videos, Bilder, Texte oder Audioaufnahmen wird in Zukunft weiter zunehmen. Mithilfe **Künstlicher Intelligenz** wird es immer einfacher, solche Inhalte zu produzieren und sie massenhaft zu streuen. Die Ergebnisse sind oft täuschend echt. Sie zu erkennen, ist nicht immer leicht.

Desinformation ist schädlich für eine aufgeklärte Gesellschaft. Es ist daher wichtig, zu verstehen, wer dahintersteckt, wie sie verbreitet wird und wie man sie erkennt. Nur dann kann man sich **sicher und verantwortungsvoll** in der heutigen Medienlandschaft bewegen. Nur so können wir **proaktiv gegen Desinformation vorgehen** und verhindern, dass sie sich verbreitet.

Wie entsteht Desinformation? Wie wird sie verbreitet? Welche Rolle spielt Künstliche Intelligenz dabei? Wie können Sie Desinformation erkennen und was können Sie ihrer Verbreitung entgegensetzen? Diese und weitere Fragen wollen wir in diesem Magazin anhand konkreter Fallbeispiele beantworten.

Viel Spaß beim Lesen und bleiben Sie kritisch,

Ihr Redaktionsteam vom iRights.Lab



Bild: KI-generiert mit Adobe Firefly

Editorial



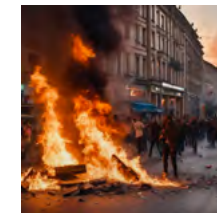
Der tote Teenager in Berlin – Die gefälschte *BILD*-Seite _____ 4



Was ist Desinformation und wie kann ich sie erkennen? _____ 6



Kreuzworträtsel _____ 10



KI Bilderrätsel: Können Sie echt von fake unterscheiden? _____ 11



DAS hat Baerbock NICHT gesagt! – Deep Fakes und Soziale Medien _____ 12



Desinformation und die Möglichkeiten Künstlicher Intelligenz _____ 16

Quiz _____ 20

Glossar _____ 22

Waschzettel – Elf Dinge, die Sie gegen Desinformation tun können! _____ 25

Der tote Teenager in Berlin – Die gefälschte *BILD*-Seite

Die folgende Geschichte ist ein klassisches Beispiel dafür, wie Desinformation entsteht und wie sie sich verbreitet. Schauen Sie ganz genau hin: Was fällt Ihnen auf? Erkennen Sie, was hier nicht stimmt? Die Auflösung der Geschichte finden Sie auf Seite 7!



Bild: KI-generiert mit Adobe Firefly

Achtung: Manipulierte Inhalte!

In einer *WhatsApp* (Glossar S. 22) Gruppe teilt ein Bekannter einen Link zu einem vermeintlichen Artikel in der *BILD*-Zeitung, der ihm auf Facebook vorgeschlagen wurde. Dazu schreibt er „Da bleibt einem die Spucke weg!“

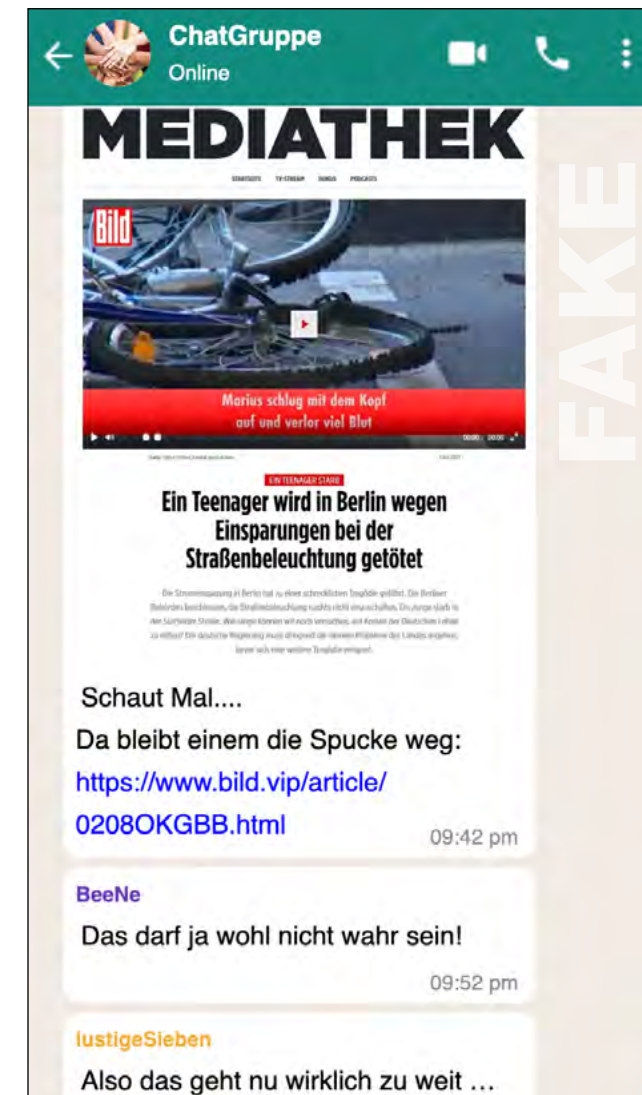


Bild: Etwas stimmt nicht mit dem Link unter dem Beitrag. Erkennen Sie, was?

Die Reaktionen in der Gruppe fallen sehr emotional aus. Im Chat wird nur der Titel des Artikels mit einem kleinen Bild angezeigt. „Teenager in Berlin wegen Einsparungen bei Straßenbeleuchtung getötet“. Sie sind neugierig geworden und folgen dem Link. Im Browser öffnet sich eine Webseite mit dem Logo der *BILD*-Zeitung.

Sie lesen den kurzen Artikel und sehen sich das dazugehörige Video an.



Scannen Sie den Code mit der Handkamera, um das Video anzusehen!

Das Video zeigt Bilder von einem Unfallort: Polizei, Krankenwagen, ein verbogenes Fahrrad, Absperrband.

Das Video ist mit Untertiteln versehen, die einordnen sollen, was auf den Bildern zu sehen ist.

Ein 16-Jähriger sei in Berlin nachts mit dem Fahrrad gestürzt, heißt es dort. Im Dunkeln habe er ein Loch in der Straße übersehen.

Weil es keine Beleuchtung gab, blieb er unentdeckt. So verblutete der Junge am Unfallort.

Schuld daran sei die Berliner Regierung. Die habe nämlich wegen der hohen Energiepreise entschieden, die Beleuchtung nachts auszulassen.



Foto: Standbild aus einem Video auf der vermeintlichen *BILD*-Webseite.

In dem kurzen Beitrag unter dem Video steht, die Regierung müsse sich „dringend um die inneren Probleme des Landes“ kümmern, „bevor sich eine weitere Tragödie ereignet“.

Sie teilen die Geschichte mit Ihrer Tochter über *WhatsApp*.

Die antwortet sofort: „Die Geschichte ist ein Fake!“ Erkennen Sie, was hier nicht stimmt?

Beantworten Sie zunächst die folgenden Fragen, bevor Sie die Auflösung auf der nächsten Seite lesen!

- Was macht diese Geschichte emotional mit Ihnen?
- Fällt Ihnen etwas am Link zur Website auf?
- Fällt Ihnen etwas an der Webseite und dem Video auf?
- Würden Sie die Geschichte mit anderen teilen?

Was ist Desinformation und wie kann ich sie erkennen?

Auf den nächsten Seiten erklären wir Ihnen, was Desinformation ist, wie sie entsteht und wie sie verbreitet wird. Hier finden Sie außerdem die Auflösung der Geschichte des vermeintlich verstorbenen Teenagers in Berlin (S. 4) und einige Tipps für den richtigen Umgang mit Desinformation.

Was ist dran an der vermeintlichen *BILD*-Geschichte?

Die Geschichte des toten Teenagers in Berlin (ab Seite 4 im Heft) ist erfunden. Doch nicht nur das: Die gesamte Webseite ist ein Fake (Glossar, S. 22), eine bloße Nachahmung der *BILD*-Zeitung. Hinter dieser Aktion steckt die Absicht, Leser*innen bewusst zu täuschen und ihre Meinungsbildung zu beeinflussen. Es handelt sich also um ein Beispiel für Desinformation.

Wer steckt hinter dem *BILD*-Fake?

Die falsche *BILD*-Seite ist Teil einer russischen Desinformationskampagne, die darauf abzielt, Menschen mit einer emotionalen Geschichte zu manipulieren und ihre Meinungsbildung zu beeinflussen. Die Botschaft lautet: Während die deutsche Regierung Geld für andere Zwecke ausgibt, wie zum Beispiel für die Ukraine oder Geflüchtete, wird an den eigenen Bürger*innen gespart.

Solche gefälschten Webseiten sind weit verbreitet und werden mit großem Aufwand betrieben. Auf den ersten Blick sehen sie täuschend echt aus. Die Inhalte sind bewusst reißerisch und polarisierend, um viel Aufmerksamkeit zu erregen. Sobald eine solche Seite im Umlauf ist, indem sie z. B. über Messenger (Glossar, S.19) wie *WhatsApp* oder Soziale Netzwerke geteilt wurde, wird ihr Ursprung und ihre Echtheit oft nicht mehr hinterfragt.

Wie hätte man den Fake erkennen können?

Desinformation ist nicht immer sofort erkennbar. Schauen wir uns deshalb das Beispiel der gefälschten *BILD*-Webseite genauer an:

- Ein erster Hinweis findet sich bereits im Link, der in der *WhatsApp*-Gruppe geteilt wurde: Bei näherem Hinsehen erkennen Sie, dass die Webadresse auf „bild.vip“ endet. Die Adresse der echten *BILD* endet, wie die meisten offiziellen Webseiten in Deutschland, auf „.de“. Gefälschte Webseiten kann man oft schon an der Adresse erkennen.

Schaut mal...

Da bleibt einem die Spucke weg:

<https://www.bild.vip/article/0208OKGGB.html>

09:42 pm

Bild: Screenshot mit irreführendem Link zur fake *BILD*-Seite

- Weitere Auffälligkeiten finden sich im Video: Die Polizeiautos auf den Bildern tragen ein Hannoveraner Kennzeichen, obwohl der Unfall angeblich in Berlin passiert sein soll. Zudem werden im Video Aufnahmen bei Tag und bei Nacht gezeigt sowie verschiedene Unfallorte. Dies ist typisch für Desinformation: Bilder werden aus dem Kontext gerissen und in einen neuen Zusammenhang gesetzt.
- In den Untertiteln und im Text unter dem Video finden sich merkwürdige Formulierungen wie „wegen der Sparsamkeit der Straßenbeleuchtung“. Diese deuten darauf hin, dass der Text nicht redaktionell geprüft wurde. Solche Fehler und ein emotionalisierender Stil können auch Hinweise auf einen möglichen Fake sein.

Was ist Desinformation?

Unter Desinformation versteht man die gezielte Verbreitung von Unwahrheiten, beispielsweise in Form von manipulierten Inhalten, mit dem Ziel, Menschen zu beeinflussen. Besonders in Kriegs- und Krisenzeiten wird Desinformation gerne als Mittel eingesetzt, um eine freie und unabhängige Meinungsbildung zu beeinflussen und die Öffentlichkeit zu täuschen.

Was ist Desinformation nicht?

Desinformation ist nicht dasselbe wie Fehlinformation. Eine Fehlinformation ist zwar ebenfalls falsch und irreführend, wird jedoch nicht absichtlich verbreitet. Ein gutes Beispiel für eine Fehlinformation ist die sogenannte Zeitungsentente: eine versehentlich als Nachricht geteilte Information, die sich später als unwahr herausstellt.



Bild: Die Westdeutsche Allgemeine Zeitung entschuldigt sich für eine Falschmeldung. (Quelle: Spiegel.de)

Wer verbreitet Desinformation?

Hinter Desinformation stecken meistens politische Interessen. Die Inhalte sollen täuschen und irreführen, um den eigenen politischen Zwecken zu dienen, die eigene Weltanschauung zu legitimieren oder andere Darstellungen in Zweifel zu ziehen. Ein häufiges Motiv ist, die Demokratie zu schwächen. Desinformation kann von verschiedenen Absender*innen stammen, die daran ein Interesse haben: Staatliche Akteur*innen, Staatsmedien, die nicht der Pressefreiheit unterliegen, oder staatlich gesteuerte Akteur*innen zum Beispiel. Auch aus sich heraus motivierte Personen, wie politische Aktivist*innen und Influencer*innen, können Desinformation verbreiten. Sie teilen die irreführenden Inhalte beispielsweise über soziale Netzwerke, Videoplattformen oder eigene Medienangebote, die keinen journalistischen Standards, wie z. B. dem Pressekodex, unterliegen.

Sobald eine Desinformation einmal in die Welt gelangt ist, wird sie oft auch von Unbeteiligten aufgegriffen und geteilt. Dadurch erhält sie neue Legitimität und verbreitet sich oft mit großer Geschwindigkeit.

Wo wird Desinformation verbreitet?

Desinformation wird häufig über Soziale Medien geteilt. Auf Plattformen wie YouTube, Facebook oder Twitter (X) etwa. Beliebte Kanäle sind aber auch Messengerdienste wie WhatsApp oder Telegram (Glossar, S. 23), weil sie schwerer zu kontrollieren sind als soziale Netzwerke.



Auch Künstliche Intelligenz (Glossar, S. 22) spielt bei der Verbreitung von Desinformation eine wichtige Rolle. Soziale Netzwerke, wie z. B. Facebook oder TikTok, arbeiten mit KI-Systemen, um ihren Nutzer*innen neue Inhalte vorzuschlagen. Dabei werden Inhalte, mit denen viel interagiert wird - die also von vielen Nutzer*innen geteilt, geliked oder kommentiert werden - besonders häufig vorgeschlagen. Reißerische Inhalte, die starke Emotionen auslösen, werden häufiger kommentiert oder geteilt als sachliche Beiträge. Das machen sich diejenigen, die Desinformation streuen, zunutze. Sie verwenden z. B. polarisierende Bilder (siehe Beispiel unten), extreme Sprache oder polarisierende Statements, um für viel Interaktion zu sorgen.

Mehr zur Funktionsweise von KI auf den Seiten 17-19!

Was hilft gegen Desinformation?

Folgende Regeln sind im Umgang mit Desinformation hilfreich:

- **Reaktion:** Prüfen Sie immer, ob ein Inhalt eine negative emotionale Reaktion hervorruft. Mit Panikmache sollen Aufmerksamkeit generiert und rationale Reaktionen umgangen werden.



Fotos: Ein manipuliertes Bild zweier Demonstrant*innen wird auf Twitter geteilt (links). Das Original zeigt andere Botschaften (rechts). Ziel der Aktion: Die Demonstrierenden zu diskreditieren. (Quelle: CORRECTIV.Faktencheck & Ruhr Nachrichten / Nora Varga)

- **Herkunft:** Finden Sie heraus, wer Absender*in einer Nachricht, Betreiber*in einer Website oder Urheber*in eines Inhaltes ist. Gibt es ein Impressum? Ist der*die Verfasser*in seriös?
- **Quellen:** Prüfen Sie, ob Behauptungen mit Quellen belegt werden. Sind die Quellen echt? Gibt es andere, seriöse Quellen, die die Behauptungen belegen?
- **Zusammenhang:** Achten Sie auf Auffälligkeiten und Ungereimtheiten in Medien wie Bildern, Videos oder Audioaufnahmen.
- **Teilen:** Leiten Sie Inhalte nicht ungelesen und unhinterfragt weiter.

Tipp: Lieber einmal mehr kritisch hinterfragen, als etwas ungeprüft weiterzuleiten!

Wer tut etwas gegen Desinformation?

In Deutschland beschäftigen sich verschiedene Ministerien und Ämter, wie das Auswärtige Amt, das Bundesministerium des Innern und für Heimat und das

Bundesamt für Verfassungsschutz mit dem Thema Desinformation. Die Bundeszentrale für politische Bildung ist eine gute Anlaufstelle für unabhängige Informationen zum Thema.

Außerdem gibt es eine Reihe journalistischer und unabhängiger Angebote, die über Desinformation informieren und Desinformationskampagnen aufdecken. Organisationen wie die Bundesarbeitsgemeinschaft der Seniorenorganisationen (BAGSO) oder auch die Stiftung Digitale Chancen bieten zudem Informationen speziell für Senior*innen an.

Aber auch Sie als Einzelperson können etwas gegen Desinformation tun! Wenn Ihnen manipulierte Inhalte in den Sozialen Medien begegnen, haben Sie z. B. die Möglichkeit, diese zu melden. Plattformbetreiber wie Facebook (Meta) oder Twitter (X) sind gesetzlich verpflichtet, auf Beschwerden zu reagieren und gegebenenfalls Inhalte zu entfernen – vorausgesetzt sie verstoßen gegen geltendes Recht oder die Richtlinien der jeweiligen Plattform. Das ist nicht immer der Fall.

Weitere Infos und Beratung:

Correctiv Faktencheck



ARD-Faktenfinder



BR Faktenfuchs



Faktenchecks bei Mimikama



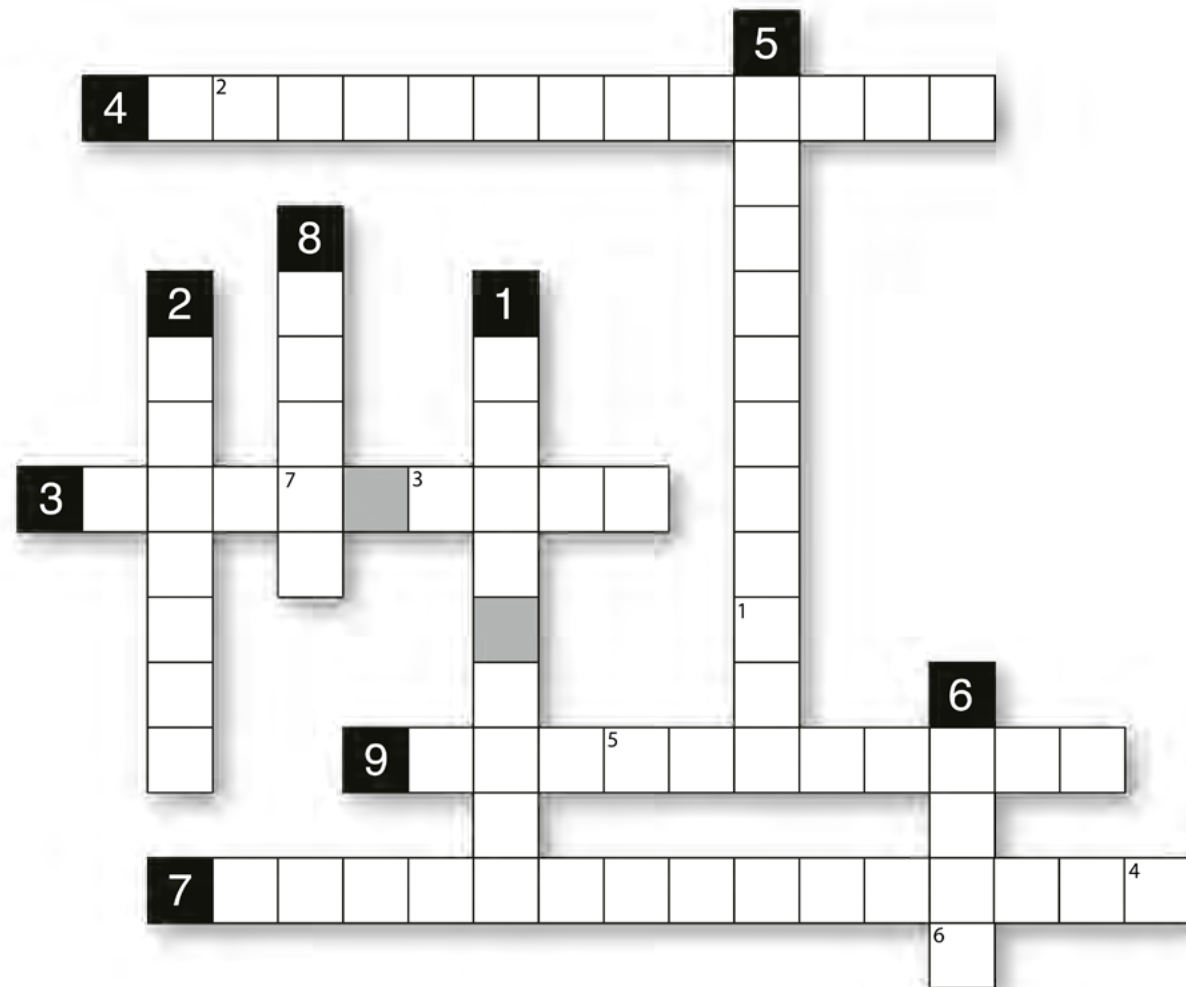
Infos der Bundesregierung



Internet-Beschwerdestelle



Kreuzworträtsel



Lösungswort:

1	2	3	4	5	6	7
---	---	---	---	---	---	---

1. Ein mithilfe von KI manipuliertes Video
2. Ein bekannter Chatbot, mit dem man u. a. Texte erstellen kann
3. Begriff, um Berichterstattung der Medien zu diskreditieren
4. KI-System, das mit Texteingaben Bilder erstellt
5. Regeln, nach denen KI-Systeme arbeiten
6. Versehentliche Falschmeldung, z. B. in der Zeitung
7. Wissen über Umgang mit Medien
8. Ohne sie kann KI nicht funktionieren
9. Test, um Desinformation zu entlarven

KI Bilderrätsel: Können Sie echt von fake unterscheiden?

Einige Bilder auf dieser Seite sind mithilfe von **Künstlicher Intelligenz** (S. 14) erstellt.

Können Sie sie erkennen? Die Auflösung und eine Erklärung finden Sie auf **Seite 21!**



1

2

4

5

3

Können Sie Ihren Augen und Ohren noch trauen? Desinformation wird zunehmend auch mithilfe von Künstlicher Intelligenz (KI) erzeugt. In dieser Geschichte schauen wir uns einen KI-erstellten Fake an. Können Sie ihn erkennen? Die Auflösung und weitere Informationen zu Künstlicher Intelligenz finden Sie ab S. 15 im Heft.

DAS hat Baerbock NICHT gesagt! – Deep Fakes und Soziale Medien

Credit: Stephan Röhl, CC BY SA 4.0

2 d · 🌐

Jetzt wird den Rentnern und Obdachlosen noch das Letzte genommen, womit sie über die Runden kommen können. Nur noch unfassbar, was sich diese Regierung erlaubt.

Wenn jemand 5.000,- € an Pfand im Jahr ausgibt und es dann abgibt, muss er ein Gewerbe anmelden 🤪

Entspricht 20.000 Flaschen/Dosen im Jahr, was bei einer 4-5 köpfigen Familie gut zusammenkommen könnte. Nicht zu vergessen Geburtstage, Besuch und sonstige Feierlichkeiten.

FAKE

YOUTUBE.COM

Außenministerin Baerbock führt Pfandflaschen Steuer 2024 ein

Was sagt ihr dazu schreibt eure Kommentare Bitte im Chat !!! IM Shop gibt's gr...

Like Comment Share

Bild: Screenshot eines mittlerweile gelöschten Facebook Beitrags, der ein vermeintlich echtes Video von Anna-Lena Baerbock teilt. (Wiederhergestellt über web.archive.org)

Sie loggen sich auf *Facebook* ein. Auf der Startseite wird Ihnen der oben gezeigte Beitrag vorgeschlagen.

Von dieser Pfandsteuer haben Sie bisher noch nichts gelesen. Klingt irgendwie absurd, also klicken Sie auf den Link.

Sie werden auf die Videoplattform *YouTube* weitergeleitet. Dort wird Ihnen das Video aus dem Beitrag angezeigt.

Das Video trägt den Titel „Außenministerin Baerbock führt Pfandflaschen Steuer 2024 ein“. Es wurde knapp 23.000 Mal aufgerufen.

Scannen Sie den QR-Code, um das Video anzusehen.:



Achtung: Manipuliertes Video!



Bild: Screenshot eines YouTube-Videos, das die vermeintlich echte Ansprache Baerbocks aufgreift. (Quelle: YouTube)

Im Video sind zwei Aufnahmen nebeneinandergelegt:

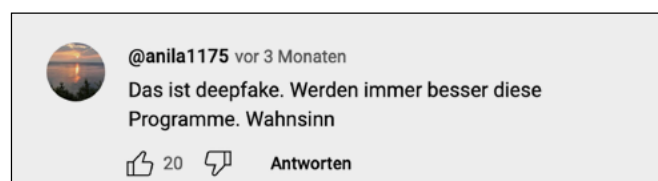
Rechts sieht man Außenministerin Baerbock bei einer Pressekonferenz im Freien. Der Bildausschnitt ist klein, darunter ist ein wütendes Emoji eingeblendet.

Links im Bild sieht man einen Herren mittleren Alters, der sich selbst und seine Reaktion auf Baerbocks Ansprache filmt. Darüber steht in weißen Buchstaben „Hört genau zu“

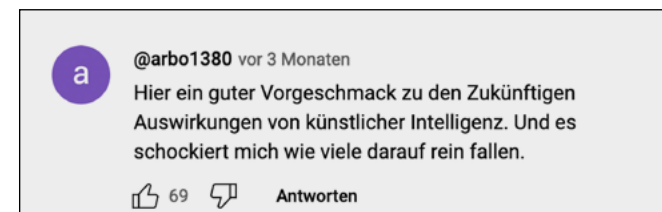
Im Video kündigt Frau Baerbock an, dass man auf gesammelte Pfandflaschen ab 2024 Steuern zahlen müsse. Bei der Pfandabgabe müsse man außerdem seinen Personalausweis vorzeigen.

Der Herr links schüttelt immer wieder fassungslos den Kopf.

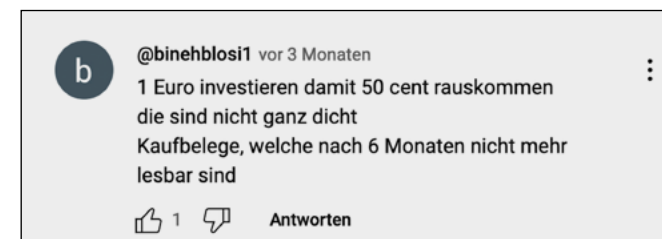
Nach etwa zwei Minuten ist das Video vorbei. Sie wissen nicht recht, was Sie von dem Video halten sollen und lesen die Kommentare darunter:



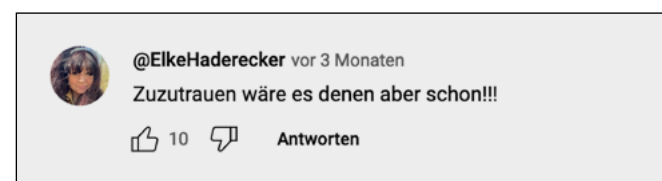
Ein anderer Nutzer schreibt:



Deep Fake? Künstliche Intelligenz? Etwas scheint mit dem Video nicht zu stimmen. Es gibt aber auch viele Nutzer*innen, die empört sind:



Selbst Nutzer*innen, die an der Echtheit des Videos zweifeln, zeigen sich verärgert:



Was denken Sie? Was stimmt mit diesem Video nicht?

Beantworten Sie zunächst die folgenden Fragen, bevor Sie die Auflösung auf der nächsten Seite lesen:

- Welche Reaktion hat das Video bei Ihnen hervorgerufen?
- Kommt Ihnen an dem Video etwas komisch vor?
- Wie nehmen Sie die Kommentare zu dem Video wahr?
- Wie würden Sie vorgehen, um Ihre Zweifel an dem Video auszuräumen?



Bild: KI-generiert mit Adobe Firefly

Desinformation und die Möglichkeiten Künstlicher Intelligenz

KI-Programme als Beschleuniger von Desinformation: Wie sie funktionieren, was sie so effektiv macht und wie Sie KI-generierte Inhalte erkennen können.

Außerdem: Die Auflösung zum Pfand-Video mit Annalena Baerbock (S. 12).

Bild: KI-generiert mit Adobe Firefly

Was ist dran am Baerbock-Video?

Sowohl das Video als auch die Aussagen der Außenministerin zur geplanten Pfandsteuer sind nicht echt. Das Video ist ein sogenannter Deep Fake (Glossar, S. 22). Mithilfe von Künstlicher Intelligenz wurden die Stimme der Außenministerin imitiert und ihre Gesichtsbewegungen angepasst. Das Ziel dieses Videos ist es, die gesellschaftliche Stimmung aufzuheizen, indem suggeriert wird, die Ministerin wolle die Bürger*innen mit zusätzlichen Steuern belasten.

Wie wurde KI im Baerbock-Video eingesetzt?

Um das Video zu erstellen, wurden echte Video- und Audioaufnahmen der Ministerin in ein KI-Programm eingespeist. Mittlerweile brauchen diese Systeme teilweise nur wenige Minuten Material, um Stimmen und Bilder zu produzieren. Das manipulierte Video wurde anschließend auf der Social Media Plattform TikTok unter dem Account [annalenas.world](#) veröffentlicht. Der Nutzer hat das Video dort als Satire gekennzeichnet.

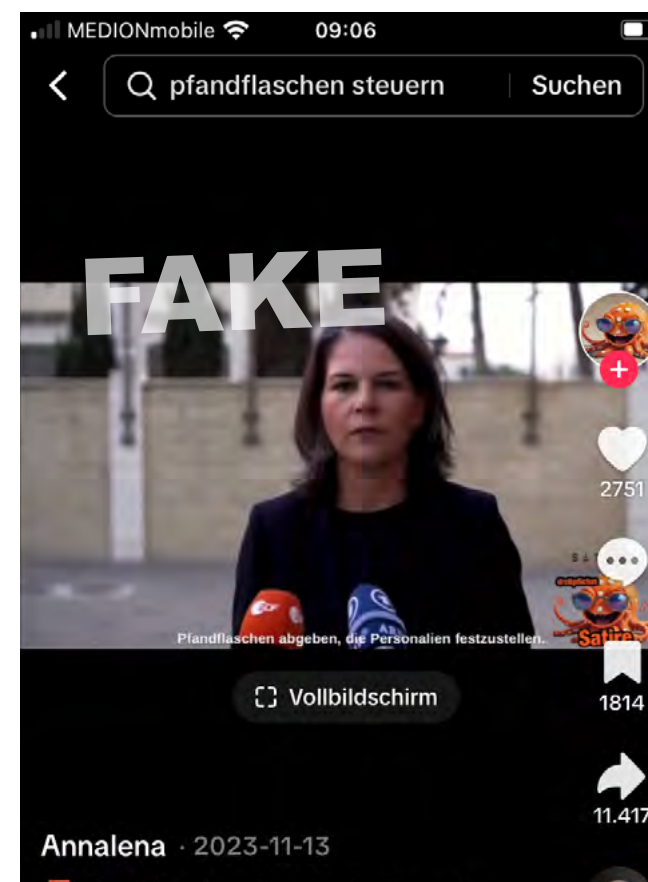


Bild: Screenshot des manipulierten Videos auf dem TikTok Kanal [annalenas.world](#). (Quelle: TikTok)

Was ist das Problem?

Es ist trotz Kennzeichnung nicht unmittelbar erkennbar, dass es sich bei dem Video um Satire handeln soll. Außerdem ist nicht klar gekennzeichnet, dass es mithilfe von KI erstellt wurde. Zudem bedient es ein eindeutig politisches Narrativ: Die Grünen als Verbotspartei und die Bundesregierung, die den ärmsten Bürger*innen ihr hart erarbeitetes Geld wegnimmt. Dass vielen Zuschauenden nicht klar ist, dass es sich hier um einen Fake handelt, erkennt man an den Kommentaren.

Sobald das Video die Plattform verlässt, ist es noch schwieriger einzuordnen. Eingebunden in das YouTube-Video aus unserem Beispiel (S. 14), geht die Kennzeichnung als Satire endgültig verloren. Wird das bereits verfälschte Video dann noch auf Facebook mit einer persönlichen, emotionalen Nachricht geteilt, wird es weiter verfremdet und gewinnt an Legitimität. Diese Art von Verfremdung ist ein typisches Beispiel dafür, wie sich Desinformation verbreitet.

Was ist Künstliche Intelligenz?

Künstliche Intelligenz (KI) ist ein Sammelbegriff. Meistens sind damit verschiedene Methoden **maschinellen Lernens** gemeint. Komplexe Computerprogramme werden darauf trainiert, anhand festgelegter Regeln (Algorithmen) selbstständig **Muster in Daten** zu erkennen. Auf diese Weise lernen die Programme zum Beispiel Katzen von Hunden zu unterscheiden oder das nächste Wort in einem Satz vorherzusagen. Es ist wichtig zu beachten, dass die meisten KI-Systeme für spezielle Zwecke entwickelt werden.

Was ist generative KI?

Wenn ein KI-System aus Trainingsdaten eigene Inhalte, wie zum Beispiel Bilder oder Texte, produzieren kann, spricht man von generativer Künstlicher Intelligenz. Chatbots (Glossar, S. 22) wie *ChatGPT* oder Bildgeneratoren (Glossar, S. 22) wie *Midjourney* können anhand von Textbefehlen (Prompts) in Sekundenschnelle neue Inhalte erstellen. Sie sind in der Regel sehr leicht zu bedienen. Die Ergebnisse können wirken, als wären sie von Menschen gemacht und als würden sie die Realität abbilden.

Wo kommen diese KI-Modelle her?

Die Entwicklung von generativen KI-Modellen wird hauptsächlich von großen Unternehmen wie *Google*, *OpenAI*, *Amazon* und *Meta* vorangetrieben. Um diese Modelle zu entwickeln, benötigt man viel Geld, Energie und Zugriff auf riesige Datenmengen. Die Unternehmen nutzen die Modelle als Grundlage für ihre Anwendungen wie Chatbots oder Programme zur Bilderstellung. Einige davon sind kostenlos oder für eine geringe Gebühr nutzbar. So haben in kurzer Zeit viele Menschen Zugang zu diesen Werkzeugen bekommen. Man muss nicht programmieren können oder viel von Computern verstehen, um sie zu benutzen.



Wie wird generative KI für Desinformation missbraucht?

Kommerzielle KI-Programme wie *ChatGPT* werden in der Regel vor ihrer Veröffentlichung getestet. So soll zum Beispiel vermieden werden, dass die Programme rassistische oder gewaltvolle Inhalte produzieren. Leider sind diese Tests nicht zuverlässig und die Programme anfällig für Fehler. Böswillige Akteur*innen finden immer wieder Wege, die Programme für ihre Zwecke zu nutzen.

Oft bewegen sich KI-generierte Inhalte außerdem in einer Grauzone, wie das folgende Beispiel zeigt: Ein Abgeordneter der AfD nutzte einen KI-Bildgenerator, um ein Bild einer Gruppe migrantisch aussehender, aggressiv wirkender Männer zu erstellen. Der Abgeordnete verwendete das Bild dann, um in den Sozialen Medien Stimmung gegen Geflüchtete zu machen – ohne zu kennzeichnen, dass das Bild KI-generiert ist. Das Bild zu generieren, ist nicht verboten, es auf diese Art zu verwenden aber durchaus fragwürdig.

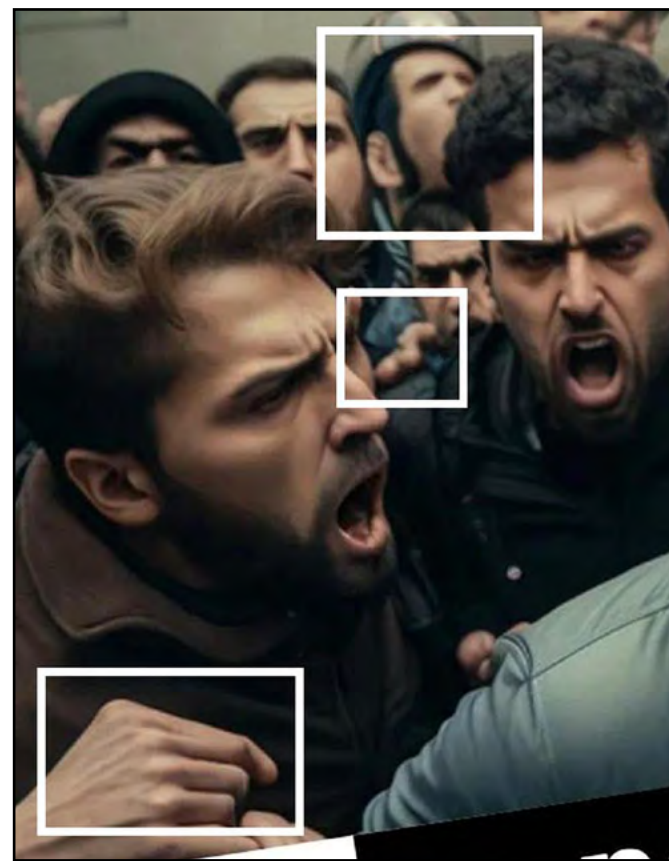


Bild: Beitrag eines AfD Abgeordneten auf Instagram mit KI-generiertem Bild (links). Unregelmäßigkeiten wie verzerrte Gesichter zeigen, dass das Bild KI-generiert ist (rechts). (Quelle: Instagram & ZDF)

Wie können Sie KI-generierte Inhalte erkennen?

KI-generierte Inhalte sind nicht immer leicht zu erkennen. Hier sind einige Handlungsempfehlungen, an denen Sie sich orientieren können:

- **Offensichtliche Ungereimtheiten:** Achten Sie auf Mundbewegungen, die nicht zum Bild passen, Hände mit zu vielen Fingern oder anderweitig verformte Körperteile. Auch eine schlechte Bild- und Tonqualität oder eine seltsame Aussprache können Hinweise sein.
- **Verifizieren:** Können Sie mit einer Suchmaschine herausfinden, ob der Inhalt bereits an anderer Stelle veröffentlicht wurde? Bei Bildern können sie auch eine so genannte Bilderrückwärtssuche machen, um herauszufinden, ob das Bild vielleicht bereits anderswo und in einem anderen Zusammenhang veröffentlicht wurde.
- **Kontext:** Findet sich der Inhalt auch auf seriösen Nachrichtenseiten oder nur auf Social Media?
- **Einordnung:** Weichen Aussagen erheblich von dem ab, was eine Person üblicherweise sagt und tut?

KI ist mehr als ChatGPT und Co

Generative KI-Programme wie *ChatGPT* bekommen gerade viel Aufmerksamkeit. KI-Systeme werden aber schon lange in verschiedenen Bereichen eingesetzt. Sie haben einen großen Einfluss darauf, wie und welche Informationen uns im digitalen Raum erreichen.

Suchen Sie etwas auf *Google*, sortiert ein KI-System Ihre Ergebnisse. Auf *Facebook* empfiehlt Ihnen ein KI-

System Beiträge und Personen, die für Sie interessant sein könnten. Auf *YouTube* schlägt Ihnen der Algorithmus endlos neue Videos vor, die Ihnen gefallen könnten. Dahinter stecken kommerzielle Interessen. Soziale Netzwerke wie *Facebook* oder Videoplattformen wie *YouTube* wollen Ihre Aufmerksamkeit so lange wie möglich binden. Denn die lässt sich vermarkten.

Provokante, emotionalisierende und aufsehenerregende Inhalte werden von den Algorithmen deshalb besonders hervorgehoben. Sie versprechen viel Aufmerksamkeit und eine längere Verweildauer. Die Gefahr ist, dass Menschen durch diese Logik immer einseitigere und immer extremere Inhalte angezeigt bekommen. Das hat in der Vergangenheit bereits oft dazu geführt, dass Menschen sich zunehmend in ihrer Meinung radikalisieren. Viele der großen Plattformen stehen dafür seit längerem in der Kritik. Trotz einiger Vorschriften und Maßnahmen hat sich an der grundlegenden Funktionsweise der Plattformen bisher nichts geändert.

Soziale Medien sind außerdem die ideale Umgebung, um Desinformation automatisiert mithilfe so genannter Bots (Glossar, S. 22) zu streuen. Bots können, als Menschen getarnt, automatisiert Inhalte teilen, kommentieren oder liken und ihnen damit künstlich zu mehr Aufmerksamkeit verhelfen. Die gefakten Profile müssen nicht als solche gekennzeichnet werden und sind nur schwer von echten Profilen zu unterscheiden. So können sie dazu beitragen, die Wahrnehmung bestimmter Themen oder Meinungen zu verzerren. Ein Beitrag mit vielen „Likes“ kann zum Beispiel als besonders beliebt fehlinterpretiert werden, obwohl eigentlich nur Bots dahinterstecken. Wichtig ist: Auch Bots agieren nicht aus sich heraus. Hinter jedem Bot steckt ein Mensch, der ihn einsetzt.

Desinformation leicht gemacht dank generativer KI

- Viele Programme sind kostenlos oder gegen eine geringe Gebühr verfügbar.
- Die Programme lassen sich meist sehr einfach bedienen.
- Inhalte lassen sich sehr schnell erstellen.
- Inhalte wirken teilweise täuschend echt.
- Inhalte sind in der Regel nicht als „KI-generiert“ gekennzeichnet.

Quiz

Testen Sie Ihr Wissen zum Thema Desinformation und Künstliche Intelligenz in unserem Quiz!

1. Unter Desinformation versteht man Inhalte, die ...

- a) ... versehentlich falsch verstanden und geteilt wurden.
- b) ... eingesetzt werden, um Fehlinformationen zu widerlegen.
- c) ... vorsätzlich irreführende und falsche Darstellungen enthalten.

2. Videos oder Tonaufnahmen, die mit einem KI-Programm erstellt wurden, um bestimmte Personen zu imitieren, nennt man auch:

- a) Deep Twist
- b) Deep Fake
- c) Fake News



Bilder: KI-generiert mit Adobe Firefly

Lösungen: 1c, 2b, 3b+c, 4a, 5c

3. Welche der folgenden Merkmale hat Desinformation üblicherweise nicht? (Mehrere Antworten möglich)

- a) Einen emotionalisierenden Ton
- b) Einen klaren Absender
- c) Eine Quellenangabe

4. Wie nennt man Fake-Profile in den Sozialen Medien, die dabei helfen, automatisiert Desinformation zu verbreiten?

- a) Bots
- b) Bits
- c) Trolle

5. Welches der folgenden Beispiele hat nichts mit Künstlicher Intelligenz zu tun?

- a) Personalisierte Produktempfehlungen
- b) Google Suchanfragen
- c) Spurassistentz

Glossar

Algorithmus

Ein Algorithmus beschreibt, ähnlich wie eine mathematische Formel, wie ein Problem gelöst werden soll. Er besteht aus einer Abfolge von gut definierten Einzelschritten. Kochrezepte oder Montageanleitungen funktionieren sehr ähnlich wie Algorithmen. In digitalen Anwendungen werden Algorithmen eingesetzt, indem sie in Programmiersprache übersetzt und in Software eingebaut werden.

Bildgenerator

KI-System, das anhand von Textbefehlen aus Trainingsdaten neues Bildmaterial generiert.

Blackbox

KI-Systeme sind häufig so komplex, dass Menschen ihre Entscheidungswege nicht mehr nachvollziehen können. Selbst die Entwickler*innen können oft nicht nachvollziehen, warum das KI-System einen bestimmten Lösungsweg gewählt hat. Unternehmen wollen außerdem oft keine Informationen über ihre Modelle preisgeben. Deshalb wissen wir zum Beispiel nicht, wie der Suchalgorithmus von Google funktioniert. Fehlende Transparenz macht es schwer, KI-Systeme zu kontrollieren.

Bot

Ein Bot ist ein Computerprogramm, das weitgehend selbstständig eine bestimmte Aufgabe abarbeitet. Betreiber*innen von Suchmaschinen setzen Bots ein, um Webseiten automatisiert zu durchsuchen und deren Inhalte auszuwerten (Crawler). Im Kund*innenservice werden immer häufiger Chatbots eingesetzt. Dabei handelt es sich um Programme, die automatisiert Fragen der Nutzer*innen beantworten. Bots können aber auch eingesetzt werden, um z. B. automatisiert Desinformation in Sozialen Netzwerken zu teilen.

Chatbot

Chatbots sind Programme, die per Text- oder Sprachbefehl bestimmte Aufgaben erledigen können.

ChatGPT ist z. B. ein sprachbasierter Chatbot, mit dessen Hilfe man Texte, Code oder Bilder erstellen lassen kann. *Siri* oder *Alexa* sind Beispiele für sprachbasierte Chatbots. Da die Modelle, mit denen die Bots arbeiten, immer leistungsfähiger werden, kommen sie in immer mehr Bereichen zum Einsatz.

Desinformation

Desinformation meint die gezielte Verbreitung von Unwahrheiten, z. B. in Form von manipulierten Inhalten, mit dem Ziel, Menschen zu beeinflussen. Vor allem in Kriegs- und Krisenzeiten ist Desinformation ein beliebtes Mittel, um eine freie und unabhängige Meinungsbildung zu verhindern und die Öffentlichkeit zu täuschen.

Deep Fake

Realistisch anmutende, aber künstlich erzeugte bzw. manipulierte Foto-, Video- oder Tonaufnahme. Der Begriff setzt sich aus Deep Learning und Fake (auf Deutsch: Fälschung) zusammen. Erstellt werden Deep Fakes mithilfe von KI-Programmen. Dafür werden beispielsweise Videoaufnahmen einer Person analysiert, um ihre Gesichtsbewegungen simulieren zu können. Deep Fakes können ein Mittel sein, um gezielt Desinformation zu verbreiten oder jemanden in ein schlechtes Licht zu rücken.

Fake

Oft synonym verwendeter Begriff für unechte oder bewusst manipulierte Inhalte.

Fake News

Fake News setzt sich zusammen aus Fake (Englisch: Fälschung / gefälscht) und News für Nachricht. Der Begriff meint also eigentlich „erfundene Nachrichten“. Gleichzeitig ist der Begriff auch zu einem politischen Kampfbegriff geworden, um gegnerische Äußerungen und Meinungen zu diskreditieren.

Generative KI

KI-Modelle, die in der Lage sind, Inhalte zu erstellen. Solche Modelle können beispielsweise Texte, Bilder, Audiodateien, Videos oder Programmiercode erzeugen. Bekannte Beispiele sind die Bildgeneratoren *Stable Diffusion* und *Midjourney* oder der Textgenerator *ChatGPT*. Generative KI-Modelle sind meist komplex und benötigen große Datenmengen. Wie sie zu ihren Ergebnissen kommen, ist aufgrund mangelnder Transparenz oft nicht nachvollziehbar.

Künstliche Intelligenz

Computerprogramme, die automatisiert bestimmte Aufgaben erfüllen. KI-Systeme können mit unterschiedlichen KI-Methoden betrieben werden. Eine der häufigsten Methoden nennt sich Maschinelles Lernen. Programme lernen, in großen Datenmengen Muster zu erkennen, um daraus weitere Arbeitsschritte abzuleiten

– zum Beispiel Bilder zu erkennen oder neue Datensätze zu analysieren. Umgekehrt ist aber nicht jede Form von Automatisierung gleich Künstliche Intelligenz.

Maschinelles Lernen

Bei diesem Verfahren legen Entwickler*innen lediglich fest, welche Ziele ein Computerprogramm am Ende erreichen soll – zum Beispiel Katzen von Hunden zu unterscheiden – und lassen das Programm den besten Weg dorthin selbst herausfinden. Unterschieden werden drei grundlegende Lernmethoden: überwachtes Lernen, unüberwachtes Lernen und bestärkendes Lernen.

Messenger

Anwendung auf dem Smartphone, die zur Kommunikation dient. Über die meisten Messenger lassen sich Text-, Bild-, Video- und Sprachnachrichten versenden. Auch Sprach- und Videoanrufe sind bei einigen Diensten möglich.

Prompt

Textebefehle, mit denen KI-Programme angesteuert werden können. Abhängig davon, wie man einen Prompt formuliert, verändern sich die Ergebnisse, die ein KI-Programm liefert.

Telegram

Kostenloser Nachrichtendienst (Messenger) aus Russland. Telegram gehört zu den beliebtesten Messengern der Welt. Der Dienst stand bereits öfter in der Kritik, weil u. a. Extremist*innen und Verschwörungsideolog*innen ihn ungehindert nutzen können, um Desinformation und extremistische Inhalte zu verbreiten.

TikTok

Ein soziales Netzwerk chinesischen Ursprungs, das mittlerweile zu den meistgenutzten der Welt gehört. Nutzer*innen teilen dort meist kürzere Videos, die nacheinander in einer endlosen Abfolge (dem so genannten Feed) angezeigt werden. Neben Tanz-, Musik- und Comedy-Videos finden sich auf der Plattform auch immer mehr politische Inhalte.

WhatsApp

Messengerdienst der Firma „Meta“. *WhatsApp* ist eine der beliebtesten Kommunikations-Apps weltweit. *WhatsApp*-Gruppen und -Kanäle werden häufig genutzt, um Desinformation zu teilen.

Auflösung KI- Bilderrätsel (S. 11):

1) Das Bild ist KI-generiert. Ein User teilte das Bild auf der Plattform *X* und behauptete, es handle sich um eine Aufnahme der Berliner Sonnenallee an Silvester.

2) Das Bild ist echt.

3) Das Bild ist mit dem KI-Programm *Midjourney* erstellt. Dieses Bild ist nur schwer zu entlarven. Einen Hinweis liefern die Wasserflaschen sowie das Namensschild im Vordergrund, auf dem die Buchstaben unleserlich sind. KI-Generatoren haben Probleme damit, korrekte Wörter zu generieren.

4) und 5) Beide Bilder sind KI-generiert. Sie sind mit der neuesten Version des KI-Programms *Midjourney* erstellt. Bei näherem Hinschauen erkennt man verformte Körperteile und unleserliche Schrift, z. B. auf Straßenschildern.

Aktuell können wir KI-erstellte Bilder manchmal noch von echten Bildern unterscheiden. Das wird aber immer schwieriger, weil die KI-Programme immer besser arbeiten. Deshalb ist es umso wichtiger, dass wir uns nicht nur auf unsere Augen verlassen, sondern weitere Faktoren prüfen. Tipps zum Umgang mit manipulativen Inhalten finden Sie auf S. 9 und S. 17.

Quellen

Der tote Teenager in Berlin

- <https://correctiv.org/faktencheck/2022/09/21/berlin-toedlicher-fahrradunfall-wegen-ausgeschalteter-strassenbeleuchtung-erfunden/>
- <https://web.archive.org/web/20220908162800/https://www.bild.vip/article/0208OKGGBB.html>
- <https://web.archive.org/web/20220920121903/https://www.faz.ltd/aktuell/politik/ausland/Weniger-Brot-Butter-und-Bier-26077247.html>

Was ist Desinformation?

- <https://www.fes.de/wissen/desinformation>
- <https://www.bpb.de/themen/medien-journalismus/digitale-desinformation/339539/transparenz-als-mittel-gegen-die-digitale-verbretung-von-desinformation/>
- <https://www.verfassungsschutz.de/SharedDocs/hintergruende/DE/spionage-und-proliferationsabwehr/desinformation.html#doc1318762bodyText3>
- https://www.t-online.de/nachrichten/deutschland/gesellschaft/id_100042596/ukraine-krieg-prorussische-kampagne-das-steckt-hinter-den-fake-artikeln.html
- <https://www.bundesregierung.de/breg-de/schwerpunkte/umgang-mit-desinformation/gefaehrliche-falschnachrichten-1905340>
- <https://www.bundesregierung.de/breg-de/themen/digitalisierung/gesetz-ueber-digitale-dienste-2140944>
- <https://www.bpb.de/themen/medien-journalismus/desinformation/>

KI-Bilderrätsel

- <https://dpa-factchecking.com/germany/240102-99-471208/>
- https://old.reddit.com/r/midjourney/comments/18ul4y6/progress_on_more_complicated_scenes_for_photo/

- <https://twitter.com/EliotHiggins/status/1708859819132592601/photo/1>
- Bild 2: CC BY SA 2.0 von Gage Skidmore

DAS hat Baerbock NICHT gesagt!

- <https://correctiv.org/faktencheck/2023/12/14/satire-video-mit-annalena-baerbock-zu-pfandflaschen-ist-mit-kuenstlicher-intelligenz-erstellt/>
- <https://www.TikTok.com/@annalenas.world/video/7301029883319242016>
- <https://archive.ph/VZK5G>
- <https://www.YouTube.com/watch?v=9lKeuPNsrAc>

Desinformation und Künstliche Intelligenz

- <https://www.bagso.de/publikationen/ratgeber/kuenstliche-intelligenz-im-alltag-aelterer-menschen/>
- <https://www.zvki.de/ki-navigator/unsere-inhalte/glossar>
- <https://www.tagesschau.de/faktenfinder/kontext/ki-desinformation-fakes-101.html>
- <https://www.instagram.com/p/CqDoouhOvE0/>
- <https://www.dpa.com/de/dpa-factchecking>
- <https://www.rnd.de/medien/deepfake-video-von-annalena-baerbock-ist-das-noch-satire-oder-schon-desinformation-BFO2EI6QPVGOLF-GFHAHDNMXYU.html>
- <https://www.klicksafe.de/desinformation-und-meinung/deepfakes>
- <https://www.digitale-chancen.de/materialien/detail/anti-rumour-toolkit>

Glossar

- <https://www.zvki.de/ki-navigator/unsere-inhalte/glossar>
- https://www.irights-lab.de/storage/uploads/ebooks/pdfs/Berlin-Newcomers-Academy_Sep2023_Deutsch.pdf

Titel

- Foto: Annalena Baerbock : Stephan Röhl, CC BY SA 4.0



Waschzettel – Elf Dinge, die Sie gegen Desinformation tun können!

1. Hinterfragen Sie immer den*die Absender*in einer Information.
2. Prüfen Sie Quellen und fordern Sie Belege.
3. Prüfen Sie, ob ein Inhalt manipuliert wurde.
4. Teilen Sie Inhalte nicht ungelesen.
5. Im Zweifel: Bitten Sie professionelle Faktenchecker*innen um Hilfe.
6. Abonnieren Sie den Correctiv-Faktencheck-Kanal auf WhatsApp.
7. Melden Sie Desinformation, wenn Sie sie erkennen.
8. Lassen Sie Desinformation im Bekanntenkreis nicht unwidersprochen.
9. Teilen Sie Ihr Wissen mit anderen.
10. Unterstützen Sie Initiativen, die gegen Desinformation vorgehen.
11. Und nicht vergessen: Immer die Finger zählen!

Bild: KI-generiert mit Adobe Firefly

Impressum

Autor: Merlin Münch

Inhaltliche Mitarbeit: Bastian Denker, Merlin Münch, Jaana Müller-Brehm

Redaktion: Merlin Münch, Franziska Busse

Lektorat: Anja Henschel

Gestaltung und Satz: Christoph Löffler

Veröffentlichung: März 2024

Lizenz: Dieses Magazin steht unter der Lizenz Creative Commons CC-BY-SA-4.0 International.

Kontakt: m.muench@irights-lab.de

Verantwortung und Durchführung:

Philipp Otto

iRights.Lab GmbH

Oranienstr. 185 | 10999 Berlin

(www.irights-lab.de)

Registergericht:

Amtsgericht Charlottenburg, HRB 185640 B

Philipp Otto (Direktor und Geschäftsführer)

Dr. Wiebke Glässer (Geschäftsführerin)

Gefördert durch:

Medienanstalt Berlin Brandenburg (mabb)

Kleine Präsidentenstraße 1

10178 Berlin

Die Medienanstalt Berlin-Brandenburg ist eine unabhängige Anstalt des Öffentlichen Rechts

mabb
medienanstalt_berlin_brandenburg

iRights.Lab
Think Tank für die
digitale Welt

Bild: KI-generiert mit *Adobe Firefly*

mabb
medienanstalt_berlin_brandenburg

iRights.Lab
Think Tank für die
digitale Welt

Magazin zur Aufklärung über Desinformation und Künstliche Intelligenz. Ein Projekt des iRights.Lab, gefördert durch die Medienanstalt Berlin Brandenburg (mabb).