

Unterrichtsimpuls

„This Person Does Not Exist“ – wie Algorithmen und künstliche Intelligenz die Medienlandschaft verändern und Fake News begünstigen



Weiterführende
Schule
Ab Klasse 8



45 Minuten



Ethik, Religions-
lehre, Gemein-
schaftskunde,
Englisch, Deutsch

**BITTE
WAS?!**
Kontern gegen
Fake und Hass

Die Unterrichtsidee ist Bestandteil des Themenhefts „Informationskompetenz“. Das gesamte Themenheft und die weiteren Themenhefte stehen Ihnen unter bitte-was.de zur Verfügung.

Der Unterrichtsimpuls ist lizenziert unter einer CC BY-SA 4.0 Lizenz. Nennung wie folgt: „This Person Does Not Exist“ – wie Algorithmen und künstliche Intelligenz die Medienlandschaft verändern und Fake News begünstigen – Unterrichtsimpuls / Landesmedienzentrum Baden-Württemberg – #RespektBW / Sophie Kitzmann / unter CC BY-SA 4.0 International / <https://creativecommons.org/licenses/by-sa/4.0/deed.de>. Ausgenommen von dieser Lizenz sind die Illustrationen.

„This Person Does Not Exist“ – wie Algorithmen und künstliche Intelligenz die Medienlandschaft verändern und Fake News begünstigen

Autorin

Sophie Kitzmann

Zielgruppe/Umfang

Schulart: Sekundarstufe I, Gymnasium
 Fach: Ethik, Religionslehre, Gemeinschaftskunde, Englisch, Deutsch
 Klassenstufe: ab Klasse 8
 Umfang: 45 Minuten

Kurzbeschreibung

Algorithmen und künstliche Intelligenz werden mehr und mehr Teil unseres Alltags. Wie sie die Medienlandschaft und unsere Wahrnehmung beeinflussen werden, ist allerdings nur bedingt absehbar. Dies kommt u.a. mehr und mehr zum Tragen, wenn es um Fake News, ganz besonders um Deep Fakes geht.

Basierend auf der Webseite <https://thispersondoesnotexist.com/> erhalten die Schülerinnen und Schüler einen Einblick in die Funktionsweise von Algorithmen und künstlicher Intelligenz und befassen sich mit Chancen und Risiken ihres Einsatzes sowie den Folgen, die sich für die Medien und den Medieneinfluss auf die Gesellschaft ergeben.

Dabei soll ein Schwerpunkt auf Deep Fakes und Fake News gelegt werden.

Abschließend finden die Schülerinnen und Schüler Lösungsansätze, wie (von Algorithmen) manipulierte Bilder und ggf. auch Videos erkannt werden können.

Die Lernenden erkennen die Bedeutung von Algorithmen in unserem Alltag und befassen sich mit deren Chancen und Risiken.

Materialien

- 1 **Definition** Algorithmus, künstliche Intelligenz (KI) und maschinelles Lernen (A)(B)
 - a. Algorithmen im Alltag und „komplexe“ Algorithmen
 - b. Unterscheidung schwache und starke KI
 - c. Maschinelles Lernen
- 2 **Deep Fakes**
 - a. Artikel auf MIXED zum Thema „Deep Fakes“ (C)
 - b. Artikel auf MIXED zum Thema „Audio Deep Fakes“ (D)
 - c. Artikel (auf Englisch) zum Thema „Wahlen und Deep Fakes“ (E)



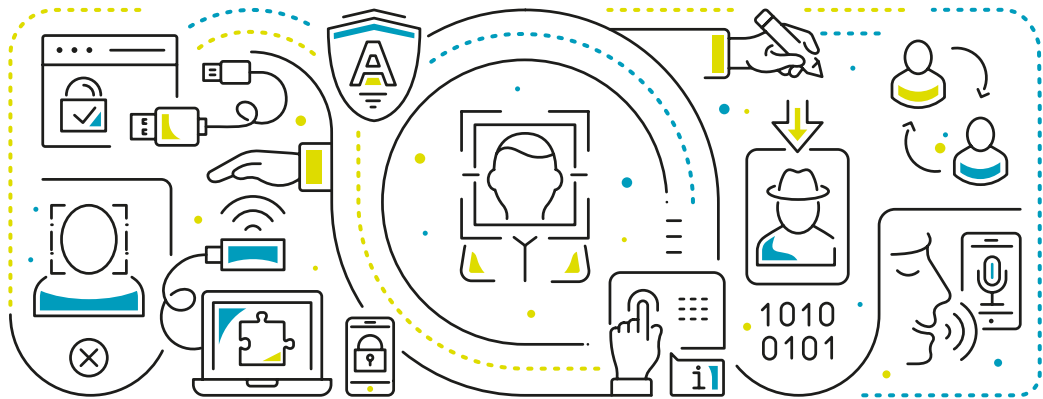
(A) **Was ist ein Algorithmus:**
<https://www.datenschutzbeauftragter-info.de/was-ist-ein-algorithmus-definition-und-beispiele/>

(B) **Was ist künstliche Intelligenz:**
<https://t3n.de/news/ai-machine-learning-nlp-deep-learning-776907>

(C) **MIXED-Artikel Deep Fakes:**
<https://mixed.de/deepfake-ki-braucht-nur-ein-trainingsfoto-fuer-einstein-fake/>

(D) **MIXED-Artikel Audio Deep Fakes:**
<https://mixed.de/deepfake-audio-cyber-angreifer-stehlen-mit-fake-ceo-stimme-millionen/>

(E) **Englischer Artikel zu Wahlen & Deep Fakes:**
<https://www.niemanlab.org/2019/06/how-could-deepfakes-impact-the-2020-u-s-elections/>



Ⓐ **So funktioniert GAN:**

<https://www.lyrn.ai/2018/12/26/a-style-based-generator-architecture-for-generative-adversarial-networks/>

Ⓑ **Tipps zur Erkennung:**

<https://medium.com/@kcimc/how-to-recognize-fake-ai-generated-images-4d1f6f9a2842>

Ⓒ **Deep Fake am Beispiel Star Wars:**

<https://www.golem.de/news/solo-a-star-wars-story-beeindruckender-deepfake-jungbrunnen-fuer-harrison-ford-1810-137169.html>

3 **Erklärung des Algorithmus GAN** (Generative Adversarial Networks) zur Erstellung möglichst echt aussehender Portraits Ⓐ

4 **Chancen, Risiken und Folgen** des Einsatzes von Algorithmen bzw. künstlicher Intelligenz

5 **Ansätze zur Erkennung** von (von Algorithmen) manipuliertem Bild- und/oder Videomaterial Ⓑ

6 **Infos zu Solo – A Star Wars Story** Ⓒ

Leitfragen und mögliche Antworten

? **Wo begegnen euch im Alltag Algorithmen?**

In sozialen Netzwerken wie Facebook oder Instagram, Suchmaschinen, Karten- bzw. Routenplaner-Apps etc.; auch einfache Gebrauchsanweisungen oder Anleitungen sind Algorithmen im Alltag. Ⓐ

Ggf. erinnern sich die Schülerinnen und Schüler an die Inhalte des Aufbaukurses Informatik. Hier geht es um einfache Algorithmen, oftmals auch erklärt mit der Funktionsweise von automatischen Türen.

? **Wie würdet ihr euren (Groß-)Eltern oder euren jüngeren Geschwistern einen Algorithmus erklären?**

? **Was versteht ihr unter künstlicher Intelligenz?**

? **Kennt ihr den Begriff des maschinellen Lernens? Wenn ja, wie würdet ihr maschinelles Lernen beschreiben?**

- ▶ Ein Algorithmus ist ein Code, der eine bestimmte Vorgehensweise vorgibt, je nachdem ob bestimmte Situationen eintreten bzw. bestimmte Voraussetzungen erfüllt oder nicht erfüllt sind.
- ▶ Künstliche Intelligenz ist ein Dach- bzw. Sammelbegriff dafür, dass Maschinen Intelligenzleistungen erbringen, die denen von Menschen ähneln.
- ▶ Maschinelles Lernen beschreibt den Prozess, wenn Algorithmen und Maschinen selbstständig lernen.

→ Einzelne Fachbegriffe im Nachgang gemeinsam definieren (siehe Links bei 1) und ggf. auf einem Padlet festhalten

Nach dem Lesen des Artikels zur Funktionsweise bzw. nach Präsentation

? **Möchte jemand versuchen, in eigenen Worten zu erklären, wie der Algorithmus, der hinter der Seite „This Person Does Not Exist“ steht, funktioniert?**

▶ Die KI bzw. der Algorithmus kann Gesichter erstellen, indem Merkmale vorhandener Portraits wie beispielsweise Augen- und Nasenform, Haarfarbe und -typ, Falten oder Augen- und Gesichtsfarbe etc. zu neuen Motiven kombiniert werden. Dieser Algorithmus wird „Generator“ genannt.

▶ Daneben existiert ein zweiter Algorithmus, der „Discriminator“, der die Ergebnisse des ersten auf die augenscheinliche Realität hin bewertet.

▶ Der Generator muss also den Discriminator immer überzeugen. Die KI trainiert sich so ohne menschliche Hilfe selbst und erzielt immer realistischere Ergebnisse.

.....

① **Wie ihr an den Beispielen für Algorithmen und künstliche Intelligenz gesehen habt, können sie in verschiedensten Bereichen zum Einsatz kommen.**

Was denkt ihr, welche Vorteile bzw. Chancen kann der Einsatz von künstlicher Intelligenz, insbesondere im Bereich von Bild- und Videomaterial, haben? Wo seht ihr Nachteile bzw. Herausforderungen?

.....

▶ **Chancen:** Einsatz bei digitalen Spielen (z.B. realistischere Figuren in Computerspielen) oder Animationsfilmen, kann virtuellen Assistenten oder Bots ein realistisches Gesicht geben (z.B. in der Altenpflege) etc.

▶ **Herausforderungen:** einfachere Erstellung von gefakten Bildern (vgl. Deep Fakes) und damit verbundene Betrugsmöglichkeiten (z.B. einfacheres Catfishing, Annehmen falscher Identitäten o.Ä.) oder Beeinflussung von gesellschaftlichen Abläufen (z.B. Wahlen), Überwindung bestehender Möglichkeiten zur Aufdeckung von Fakes (z.B. keine Rückwärtsbildersuche) etc.

.....

① **Wenn ihr über die Vor- und Nachteile bzw. Chancen und Herausforderungen nachdenkt, die der Einsatz von künstlicher Intelligenz mit sich bringt: Was denkt ihr, welche Folgen kann das haben? Inwieweit würde sich die Medienlandschaft verändern? Was würde das für die Gesellschaft bedeuten?**

.....

▶ Verbreitung von Fake News (auch zu Propagandazwecken) könnte neue Dimensionen erreichen durch die einfachere Erstellung von Deep Fakes ① → Beeinflussung von Wahlen.

▶ Das Erkennen von gefakten Medieninhalten wird selbst für ein geschultes Auge immer schwieriger.

▶ Reale Personen könnten (z.B. im Film) durch virtuelle, aber real aussehende „Personen“ ersetzt werden → z.B. wenn Schauspieler mit tragenden Rollen plötzlich unerwartet versterben (z.B. Paul Walker) oder wenn Prequels erscheinen, deren Hauptfiguren möglichst wie eine jüngere Version des Schauspielers aussehen sollen (z.B. Solo – A Star Wars Story).

▶ Etc.

.....

① **Sind euch Besonderheiten aufgefallen, an denen man erkennen kann, ob die Portraits von der KI erstellt oder echt sind? Fallen euch weitere Ansätze ein, um echte von gefälschten/künstlich erstellten Bildern unterscheiden zu können?**

.....

▶ Basierend auf Artikel ②: Glitches in Haaren, Bildtext oder Hintergrund; asymmetrische Gesichtsformen bzw. -merkmale (z.B. Zähne, Augen etc.); unterschiedliche Texturen in Kleidung und Rest; Farbfehler im Bild

▶ Weitere Optionen, um gefakte Bilder zu entlarven: Ungereimtheiten bei verschiedenen Objekten im Bild (z.B. Lichtverhältnisse, Schärfe etc.); ggf. Überprüfung der Metadaten auf eine etwaige Bearbeitung (Ort und Zeit der Aufnahme, Originalaufnahme?)

▶ Problem ist allerdings, dass die KI ständig dazulernt und vorhandene Fehler womöglich ausbessern wird.



① **Beispiel Deep Fakes:**

<https://www.youtube.com/watch?v=614we6ZaQ04>

② **Tipps zur Erkennung:**

<https://medium.com/@kcihc/how-to-recognize-fake-ai-generated-images-4d1f6f9a2842>

Möglicher Ablauf

► **Potenzieller Einstieg:** Lehrkraft zeigt Informationen zu Harrison Ford oder einen Ausschnitt aus einem relativ neuen Film mit ihm und anschließend einen Ausschnitt aus „Solo – A Star Wars Story“.

Ⓜ **Frage: Wie kann es sein, dass er plötzlich so jung aussieht?**

► **Lehrkraft öffnet die Seite**
<https://thispersondoesnotexist.com>, zeigt die dort abgebildeten Portraits und aktualisiert die Seite ein paarmal, damit verschiedene Portraits zu sehen sind.

Ⓜ **Frage an Schülerinnen und Schüler: Was ist auf der Seite zu sehen? Was fällt euch auf (z.B. Name der Webseite, Glitches im Bild o.Ä. → ggf. Hinweise geben)? Was denkt ihr, was diese Seite zeigt bzw. macht?**

► **Hinter dieser Seite steht ein Algorithmus, der basierend auf vorhandenen Bildern von Personen neue Portraits erstellt, die es so nicht gibt.** Funktionsweise von Algorithmen und speziell des GAN, KI, Deep Fakes sowie selbstständiges (maschinelles) Lernen werden erklärt, zum Beispiel mit Präsentation oder Text. Alternativ kann hier – wenn genügend Zeit vorhanden ist – auch so vorgegangen werden, dass die Schülerinnen und Schüler arbeitsteilig in Gruppen die einzelnen Materialien bearbeiten und eine kurze Präsentation vorbereiten. Die Ergebnisse und Beispiele können auf einem Padlet festgehalten werden.

► **In Gruppenarbeit:** Die Schülerinnen und Schüler befassen sich mit möglichen Einsatzgebieten dieser KI und den daraus resultierenden Chancen und Risiken. Die Ergebnisse können beispielsweise auf

einem Plakat oder mit Glogster auf einem digitalen Plakat festgehalten und im Anschluss im Plenum diskutiert werden. Alternativ kann hier das Padlet erneut zum Einsatz kommen.

Ⓜ **Wo könnte eine solche KI im positiven Sinne genutzt werden?**

Ⓜ **Welche Herausforderungen bringt ein Einsatz dieser oder ähnlicher KI mit sich?**

Ⓜ **Welche Folgen hat der Einsatz für die Medien?**

Ⓜ **Welche Konsequenzen gibt es für die Gesellschaft?**

► **In Gruppen- oder Einzelarbeit:** Basierend auf den Bildern der Webseite erarbeiten die Schülerinnen und Schüler, woran bearbeitete/künstlich erstellte bzw. im Speziellen die von GAN erstellten Bilder zu erkennen sind (Ansatzpunkte siehe Text) → bei Verwendung eines Padlets auch diese Ergebnisse hier festhalten und ggf. von den Mitschülerinnen und -Schülern kommentieren lassen.

Ⓜ **Welche alternativen Ansätze gibt es, um gefälschte Bilder von echten zu unterscheiden?**

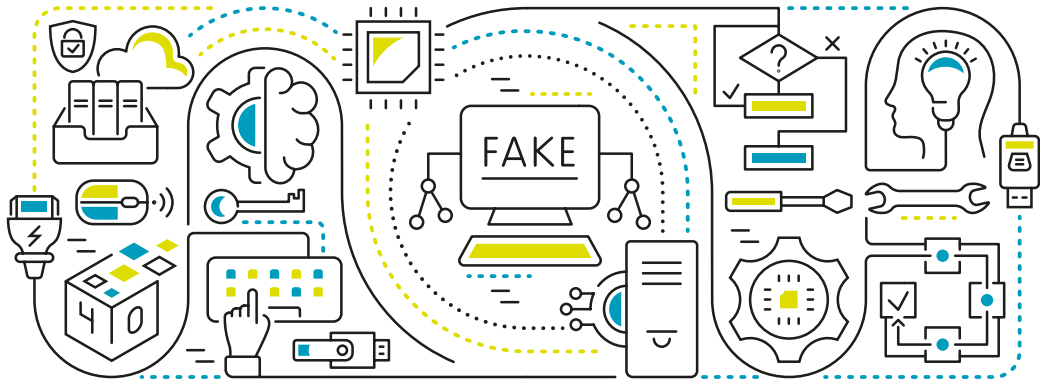
Ergebnisse können zum Beispiel gemeinsam auf Plakat/digitalem Poster/Padlet festgehalten werden.

► **Abschließend gemeinsames Quiz** mit minnit' (<https://minnit-bw.de/>): Können die Schülerinnen und Schüler (und die Lehrkraft) die gefakten Bilder erkennen? Alternativ können auch Schülerinnen und Schüler ein solches Quiz selbst erstellen.



Ⓜ „Solo – A Derpfakes Story“:
<https://youtu.be/ANXucrz7Hjs>

Alle aufgeführten Links wurden zuletzt aufgerufen am 07.09.2020.



Transfer/Exkurs



Ⓐ **GPT-2:**

<https://openai.com/blog/better-language-models/>

Ⓑ **„The Guardian“ über GPT-2:**

<https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction>

Ⓒ **Quiz basierend auf der Webseite „This Person Does Not Exist“:**

<https://careerswithstem.com.au/deep-fake-quiz/>

Ⓓ **Allgemeines Quiz:**

<https://www.thoughtco.com/can-you-spot-the-hoaxes-4099583>

Ⓔ **Virtuell erstellte Influencerin „imma“:**

<https://www.instagram.com/imma>; <https://interestingengineering.com/meet-imma-the-worlds-first-computer-generated-model>

- ▶ Die Organisation OpenAI (u.a. Elon Musk) hat eine KI namens **GPT-2** Ⓐ entwickelt, die „deepfakes for text“ selbstständig erstellen kann. Basierend auf von Menschenhand geschriebenem Text, von einzelnen Wörtern bis hin zu Textpassagen, vervollständigt GPT-2 den Text mit passenden Sätzen – angepasst an den Schreibstil, die Wortwahl, die Komplexität etc. des bisherigen Textes. Ⓑ

Ⓚ **Wofür könnte eine solche KI eingesetzt werden?**

Ⓛ **Welche Gefahren verbergen sich hinter dem Einsatz?**

Ⓜ **Welche Folgen könnte es haben, wenn die KI zur automatisierten Bilderkennung mit der zur Texterstellung kombiniert wird?**

- ▶ Verbreitung von Fake News bzw. manipuliertem Text-, Bild-, Audio- und Videomaterial als Gefahr für die Demokratie

- ▶ Ggf. Quiz (<https://minnit-bw.de/>) zur Unterscheidung von echten und „gefakten“ Bildern

– basierend auf der Webseite „This Person Does Not Exist“ Ⓒ

– allgemein zur Unterscheidung zwischen Hoax oder realem Bild Ⓓ

- ▶ Virtuell erstellte Social-Media-Influencerin „imma“: Seit Juli 2018 hat der Account über 40.000 Follower gesammelt Ⓔ

Impressum

Herausgeber und Bezugsadresse

Landesmedienzentrum
Baden-Württemberg
Vertreten durch Direktor Michael Zieher
Rotenbergstraße 111, 70190 Stuttgart
Telefon: +49 (0)711 2850-6
Fax: +49 (0)711 2850-780
E-Mail: beratungsstelle@lmz-bw.de

Dieses Materialpaket ist Teil
des Projekts #RespektBW

Projektleitung

Sebastian Seitner

Autorinnen und Autoren

Sophie Kitzmann

Redaktion

Sophie Kitzmann
Stephanie Wössner
Fabian Karg

Gestaltung

www.cosmoto.com

Stuttgart, August 2020



Sofern nicht anders vermerkt, stehen die Inhalte unter einer CC BY-NC-SA 4.0 Lizenz. Sämtliche Rechte an dieser Publikation liegen beim Landesmedienzentrum Baden-Württemberg (LMZ). Nichtkommerzielle Vervielfältigung und Verbreitung ist erlaubt unter Angabe des Herausgebers LMZ Baden-Württemberg und der Website www.lmz-bw.de.

Die Bildquellen aus den Unterrichtsmodulen sind im jeweiligen Quellenverzeichnis aufgeführt. Die Bilder werden unter Berufung auf die folgenden Paragraphen verwendet:

https://www.gesetze-im-internet.de/urhg/_51.html
https://www.gesetze-im-internet.de/urhg/_60a.html
https://www.gesetze-im-internet.de/urhg/_60b.html

Dies gilt auch für Screenshots und Standbilder aus Filmwerken.

Sollten Sie auf eine Urheberrechtsverletzung aufmerksam werden, bitten wir um einen entsprechenden Hinweis. Bei Bekanntwerden von Rechtsverletzungen werden wir umgehend handeln.

Sonstige Illustrationen sind nicht frei zu verwenden.
Illustrationen: Shutterstock

Soweit Inhalte des Angebotes des LMZ auf externe Internetseiten verweisen, hat das LMZ hierauf keinen Einfluss. Diese Internetseiten unterliegen der Haftung der jeweiligen Betreiber. Das Setzen von externen Links bedeutet nicht, dass sich das LMZ die hinter dem Verweis oder Link liegenden Inhalte zu eigen macht. Das LMZ hat bei der erstmaligen Verknüpfung der externen Links die fremden Inhalte daraufhin überprüft, ob etwaige Rechtsverstöße bestehen. Zu diesem Zeitpunkt waren keine Rechtsverstöße ersichtlich. Das LMZ hat keinerlei Einfluss auf die aktuelle und zukünftige Gestaltung und auf die Inhalte der verknüpften Seiten. Eine ständige inhaltliche Überprüfung der externen Links ist ohne konkrete Anhaltspunkte einer Rechtsverletzung nicht möglich. Bei Verlinkungen auf die Webseiten Dritter, die außerhalb des Verantwortungsbereichs des LMZ liegen, würde eine Haftungsverpflichtung nur bestehen, wenn das LMZ von den rechtswidrigen Inhalten Kenntnis erlangte und es technisch möglich und zumutbar wäre, die Nutzung dieser Inhalte zu verhindern. Bei Kenntnis von Rechtsverstößen werden derartige externe Links unverzüglich gelöscht.

#RespektBW ist eine Kampagne der Landesregierung für eine respektvolle Diskussionskultur in den Sozialen Medien. Sie soll Kinder und Jugendliche aktivieren, für die Werte unserer Demokratie und ein gutes gesellschaftliches Miteinander einzutreten. Die Kampagne ist Teil des Impulsprogramms der Landesregierung zum gesellschaftlichen Zusammenhalt. Das Landesmedienzentrum Baden-Württemberg führt die Kampagne im Auftrag des Staatsministeriums, in enger Abstimmung mit dem Ministerium für Kultus, Jugend und Sport, durch.

#RespektBW

LMZ  LANDESMEDIENZENTRUM
BADEN-WÜRTTEMBERG


Impulsprogramm
der Landesregierung


Baden-Württemberg