

**BLACK
BOX
REPOR
TING**

WIE JOURNALIST*INNEN ÜBER KI UND ALGORITHMEN BERICHTEN KÖNNEN

INHALT

| | |
|---|----|
| INTRO | 3 |
| Berichten über unterschiedliche Blackbox-Systeme | 5 |
| DAS ALGORITHMIC- ACCOUNTABILITY-REPORTING-TEAM | 6 |
| VIER ANSÄTZE FÜR KI-RECHERCHEN | |
| #1 Wichtig ist, was rauskommt: Recherche mit Experimenten | 8 |
| #2 Software auf dem Prüfstand: Recherche mit technischen Analysen | 11 |
| #3 Recht auf Einblick: Recherche mit juristischen Mitteln | 12 |
| #4 Erklär mir den Algorithmus: Interviews und Anfragen | 14 |
| ETHIK UND RECHT | |
| Welche Regeln gelten für die Arbeit mit KI in Redaktionen? | 16 |
| NEUER KI-PODCAST VON BR & SWR | 17 |
| DAS AI + AUTOMATION LAB | 18 |
| IMPRESSUM | 19 |

INTRO

Mit Tools wie ChatGPT oder Midjourney sind künstliche Intelligenz und Algorithmen im Bewusstsein der Gesellschaft angekommen – und damit auch in den Schlagzeilen. Das Datenteam und das AI + Automation Lab des Bayerischen Rundfunks teilen erprobte Recherche-Ansätze, mit denen Journalist*innen KI-Systeme und Algorithmen untersuchen und hinterfragen können.

Im November 2022 macht OpenAI den Text-Chatbot ChatGPT für die Öffentlichkeit zugänglich. Seitdem kann jede*r mit einer leicht benutzbaren Text-Eingabemaske künstlich erzeugte Texte generieren. Es entsteht eine bisher nie dagewesene Aufmerksamkeit für künstliche Intelligenz (KI). Ängste und Euphorie sind die beiden Pole der Diskussion: Firmen bewerben KI-Software als Problemlöser jeder Art – und Journalist*innen befeuern Befürchtungen zur Allmacht künstlicher Intelligenz, indem sie KI als Roboter visualisieren oder ihr menschliche Fähigkeiten wie „Fühlen“ oder „Denken“ zuschreiben.

Aus der Nerd-Ecke in die Mitte der Gesellschaft

Berichterstattung zu Algorithmen und KI ist häufig in Technik-Ressorts verortet und nur wenige Medienhäuser investieren in investigative Recherchen zu Algorithmen, wie das Medienmagazin [journalist](#) berichtet. Mit der gestiegenen Aufmerksamkeit wird KI-Berichterstattung zum Querschnittsthema, das ebenso wie die Berichterstattung zum Klimawandel viele verschiedene Fachredaktionen betrifft: In Wirtschaft, Medizin, Landwirtschaft und vielen anderen Bereichen spielen KI und Algorithmen eine immer wichtigere Rolle.

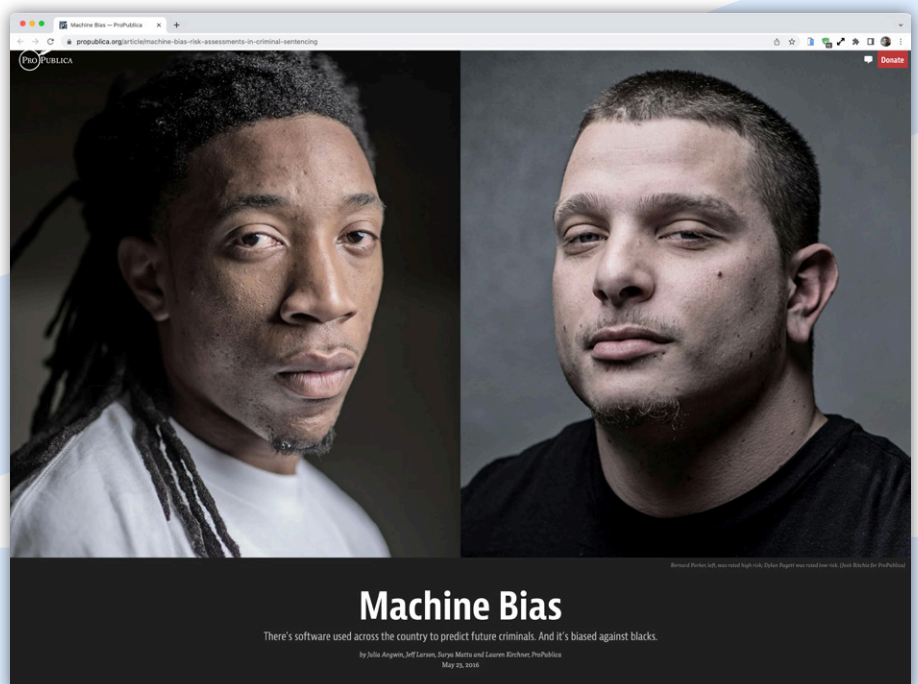
Dabei gewinnen investigative und erklärende Ansätze an Relevanz, um übertriebenen Ängsten und Hoffnungen Fak-

ten entgegenzusetzen. Der amerikanische Journalismus-Forscher Nicholas Diakopoulos hat den Begriff „Algorithmic Accountability Reporting“ geprägt. Der Datenjournalist Lorenz Matzat hat diese Recherchemethoden übersetzt mit „Rechenschaft für Rechenverfahren“. Damit gemeint sind das Recherchieren und Berichten zu Algorithmen. So werden menschliche Entscheidungen hinter algorithmischen Systemen enthüllt, damit Akteure wie Firmen oder Regierungen für ebendiese Entscheidungen zur Verantwortung gezogen werden können.

Software, die Menschen mit schwarzer Hautfarbe benachteiligt

Die erste große Algorithmic-Accountability-Recherche, die international Aufmerksamkeit erregte, veröffentlichte 2016 das Non-Profit-Recherchebüro Pro Publica unter dem Titel „Machine Bias“. Die Recherche zeigt, dass eine Software, die als Entscheidungshilfe für Richter eingesetzt wurde, Menschen mit schwarzer Hautfarbe diskriminiert: Sie wurden von dem Algorithmus fälschlicherweise viel häufiger als potentiell rückfällige Straftäter*innen eingeschätzt als Menschen mit weißer Hautfarbe.

Quelle: propublica.org



Dem Team um die Journalist*innen Julia Angwin und Jeff Larson gelang mit ihrer Recherche auch ein methodischer Durchbruch: durch Informationsfreiheitsanfragen und statistische Methoden den Output eines Algorithmus zu hinterfragen, der automatisierte Entscheidungshilfen liefert.

Journalismus mit Coding und Datenanalyse

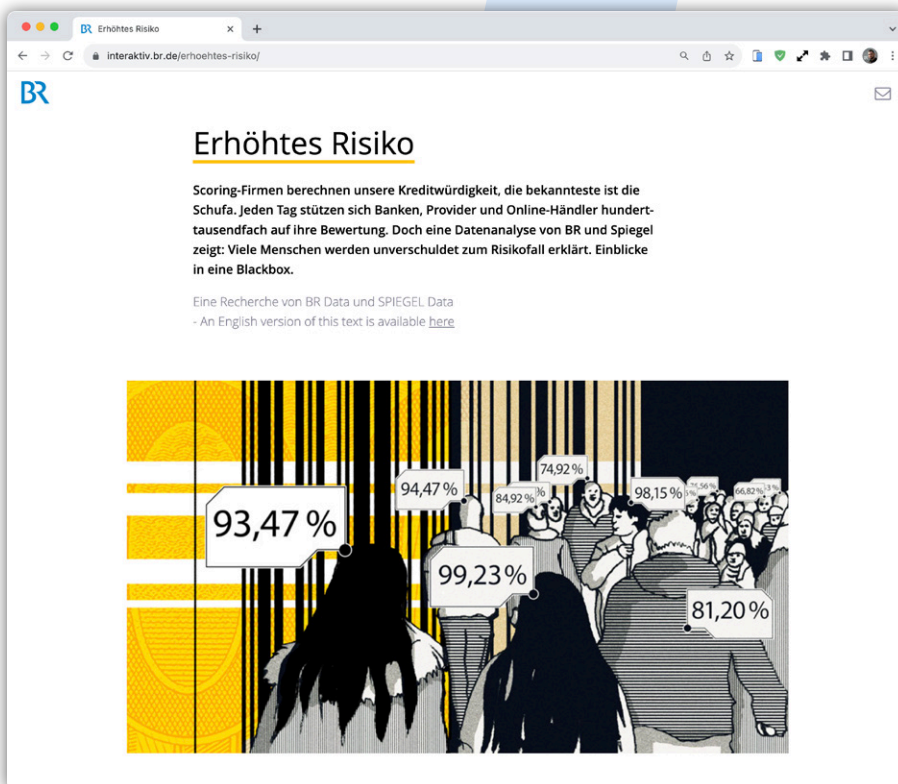
Algorithmic Accountability Reporting ist ein noch junges Genre. Für das Berichten über Algorithmen notwendig sind ein tiefes technisches Verständnis, Statistikenkenntnisse, journalistisches Handwerk und Erfahrung bei der Wahl der richtigen Recherchestrategie für das System, das untersucht werden soll.

Der Bayerische Rundfunk hat früh in journalistische Technik-Expertise investiert und 2015 das datenjournalistische Team BR Data gegründet. Das Team arbeitet eng mit den investigativen Journalist*innen bei BR Recherche zusammen. 2020 kam mit dem AI + Automation Lab Kompetenz in den Bereichen Machine Learning, künstliches Generieren von Spra-

che (Natural Language Generation) und Produktentwicklung hinzu. Die Mitglieder aller drei Teams arbeiten interdisziplinär an der Schnittstelle von Programmierung und Journalismus an investigativen Daten- und KI-Recherchen. Dabei haben sich die Teams auf technische Experimente und Datenanalysen spezialisiert, die Einblicke in das Funktionieren von KI-Systemen und Algorithmen ermöglichen.

Mit solchen Ansätzen haben die BR-Teams wegweisende Recherchen veröffentlicht, die deutschlandweit und international wahrgenommen wurden. Dazu gehören Veröffentlichungen zu Kredit-Scoring-Algorithmen (2018 zusammen mit „Der Spiegel“), Trainingsdaten-Sammlungen (2023) oder KI-Systemen, die gesellschaftlichen Schaden anrichten können, wie etwa ein KI-Recruiting-System (2021), das Job-Kandidat*innen mit KI ein Persönlichkeitsprofil zuweisen soll.

Das AI + Automation Lab nutzt das Methodenwissen auch, um KI und Automatisierung für den Journalismus nutzbar zu machen: für die Entwicklung von journalistischen Formaten, wie automatisierte Texte, Grafiken oder Newsbriefings. Wissen zu KI-Spracherzeugung, zu Trainingsdaten und statistischen Modellen kann somit sowohl für Investigationen als auch für die Entwicklung von neuen Formaten eingesetzt werden.



BERICHTEN ÜBER UNTERSCHIEDLICHE BLACKBOX-SYSTEME

Algorithmic-Accountability-Recherchen belegen regelmäßig, wie KI-Systeme und andere automatisierte Entscheidungssysteme fehlerhafte, diskriminierende oder moralisch verwerfliche Entscheidungen treffen (können) und so das Leben von Menschen auf konkrete und oftmals problematische Weise beeinflussen. Dafür sind gut durchdachte, strategisch gewählte Rechercheansätze besonders wichtig, weil sich Journalist*innen in diesem Bereich mit unterschiedlichen Arten von intransparenten Systemen konfrontiert sehen.

Zu solchen sogenannten Blackbox-Systemen gehören z.B.:

Programmcode, der auf Unternehmensservern ausgeführt wird

Solcher Code ist für Dritte normalerweise unzugänglich. Dabei handelt es sich beispielsweise um die Empfehlungsalgorithmen von Facebook, Twitter, TikTok und Co., die errechnen, welche Videos, Fotos oder Tweets Nutzer*innen als nächstes angezeigt werden.

Programmcode auf dem eigenen Handy, PC, Tablet oder anderen Geräten

Dessen Funktionsweise kann man mit ausreichend technischem Wissen theoretisch untersuchen, was je nach Methode strafbar sein kann. Um sich vor Hacking, Industriespionage oder den kritischen Blicken der Öffentlichkeit zu schützen, implementieren einige Hersteller von Smartphone-Apps und anderen Programmen ausgefeilte Techniken, um eine Analyse der eigenen Produkte zu verhindern.

KI-Modelle, die anhand von Trainingsdaten bestimmte Daten-Muster „gelernt“ haben

Darauf basierend scoren oder kategorisieren die Modelle neue Eingabedaten oder erzeugen Texte, Bilder, Sprache oder Musikstücke. Bei diesen Programmen handelt es sich um Blackboxen, deren Arbeitsweise oft sogar für die eigenen Programmierer*innen kaum mehr nachvollziehbar ist.



DAS ALGORITHMIC-ACCOUNTABILITY-REPORTING-TEAM

Mitarbeitende von AI + Automation Lab, BR Data und BR Recherche arbeiten im Bereich Algorithmic Accountability Reporting (AAR) interdisziplinär und bringen so technische, investigative und datenjournalistische Expertise zusammen. Recherchen entstehen

ebenso mit unterschiedlichen Einheiten im Haus, in der ARD und mit nationalen sowie internationalen Kooperationspartnern, unter anderem mit dem US-Non-Profit-Newsroom ProPublica, Guardian US oder Der Spiegel.



HELGA VAN OOIJEN-DEML
Redakteurin



SAMMY KHAMIS
Reporter



REBECCA CIESIELSKI
Reporterin



ROB SCHÖFFEL
Redakteur



PHILIPP GAWLIK
Computerlinguist



KATHARINA BRUNNER
Journalistin



SEBASTIAN BAYERL
Software-Entwickler



VERENA NIERLE
Redaktionsleitung



ULI KÖPPEN
Redaktionsleitung



MAXIMILIAN ZIERER
Reporter

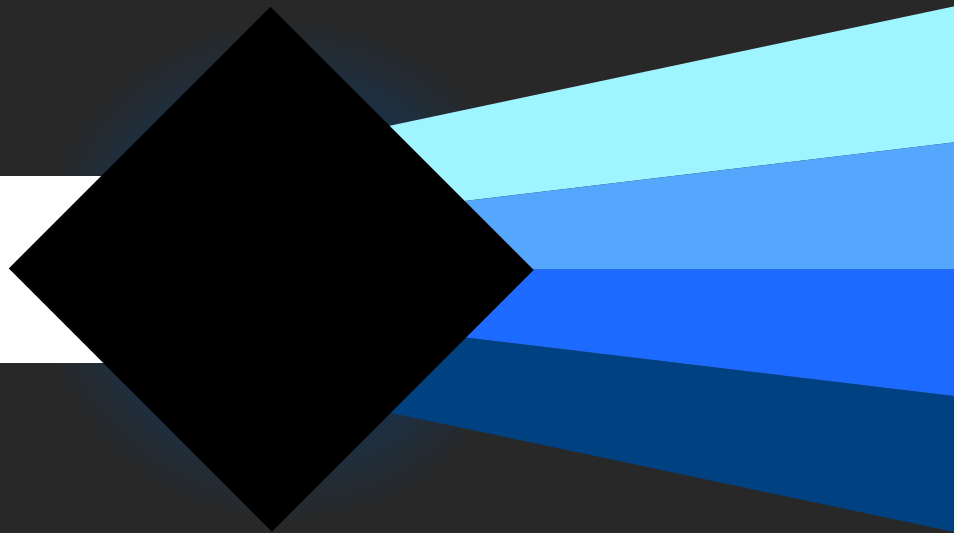


PIA DANGELMAYER
Redakteurin



MARCO LEHNER
Journalist & Entwickler

*Das AAR-Team arbeitet mit vier
verschiedenen Rechercheansätzen*

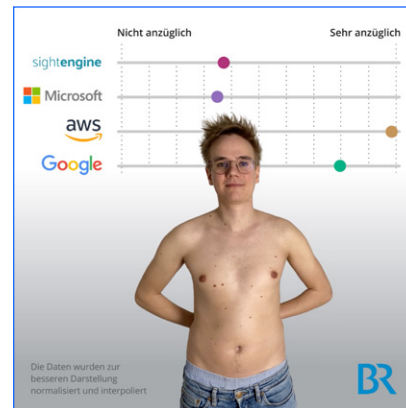


#1

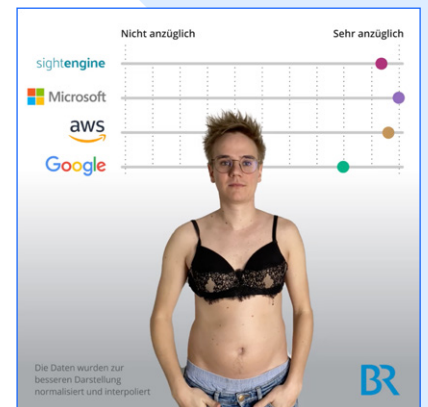
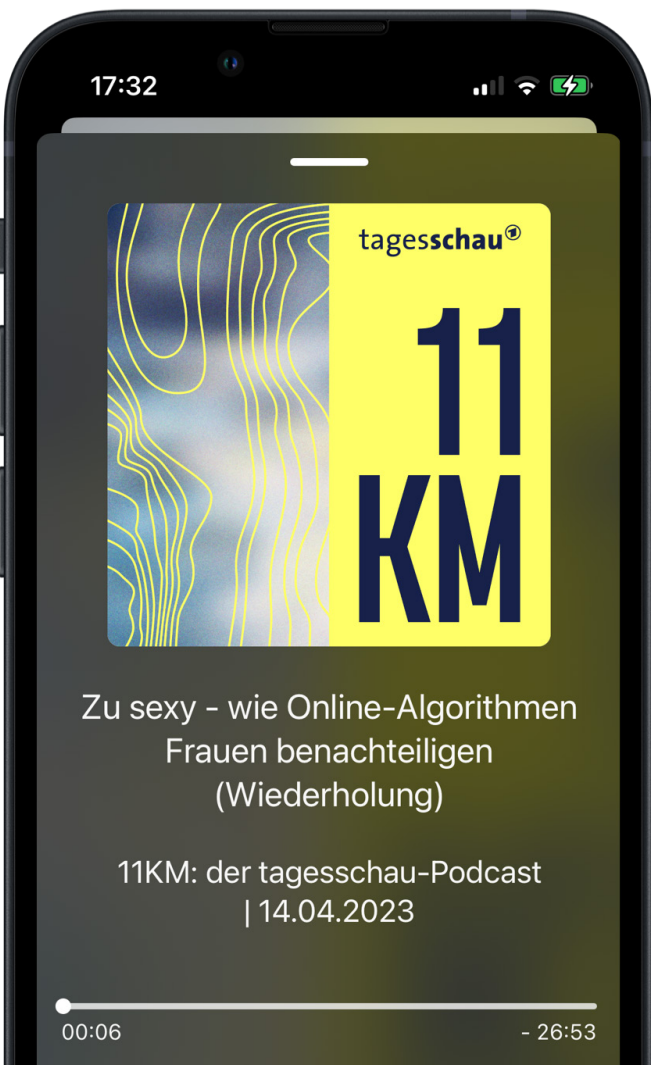
WICHTIG IST, WAS RAUSKOMMT: RECHERCHE MIT EXPERIMENTEN

Für die Recherche „Zu sexy: Wie KI-Algorithmen Frauen benachteiligen können“ untersuchte ein BR-Reporterinnen-Team vier kommerzielle Bilderkennungssysteme. Im Fokus: Dienste, die sexuelle Anzüglichkeit auf Bildern messen.

Dafür hat das Team mehr als 3.000 Bilder verschiedener Kategorien, von Unterwäsche über Sport bis zum Business-Look, bewerten lassen. Das Ergebnis: deutliche Hinweise auf geschlechtliche Diskriminierung (Gender Bias). Die Aufnahmen von Frauen werden im Vergleich zu Bildern von Männern häufiger als anzüglicher eingestuft und büßen dadurch Sichtbarkeit im Internet ein.



BR-Reporter Maximilian Osenstätter steht mit freiem Oberkörper vor der Kamera. Hinter ihm ist eine Skala eingeblendet, auf der zu sehen ist, wie „anzüglich“ verschiedene Algorithmen das Video einstufen. Zunächst erzielt der nackte Oberkörper des Reporters bei zwei von vier getesteten Algorithmen nur einen niedrigen Score.



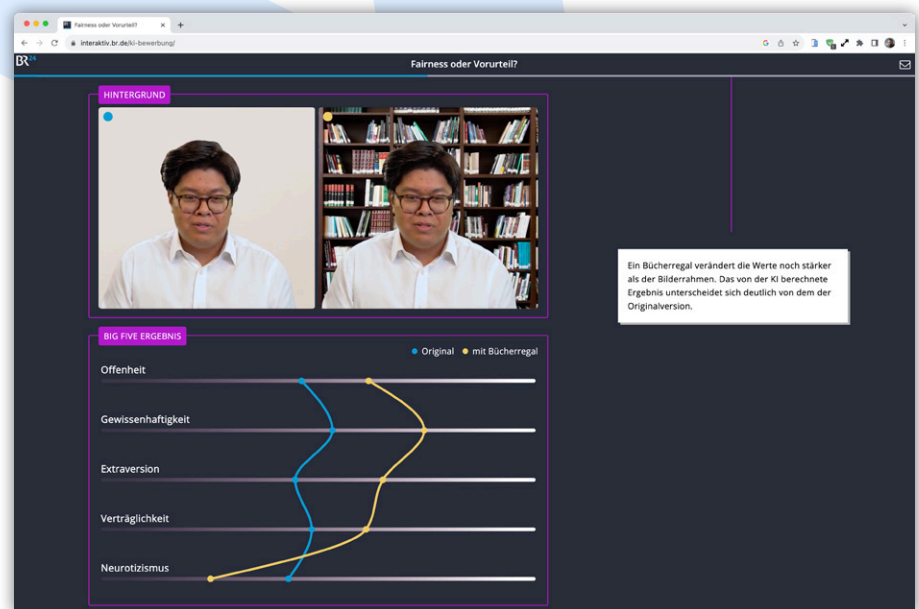
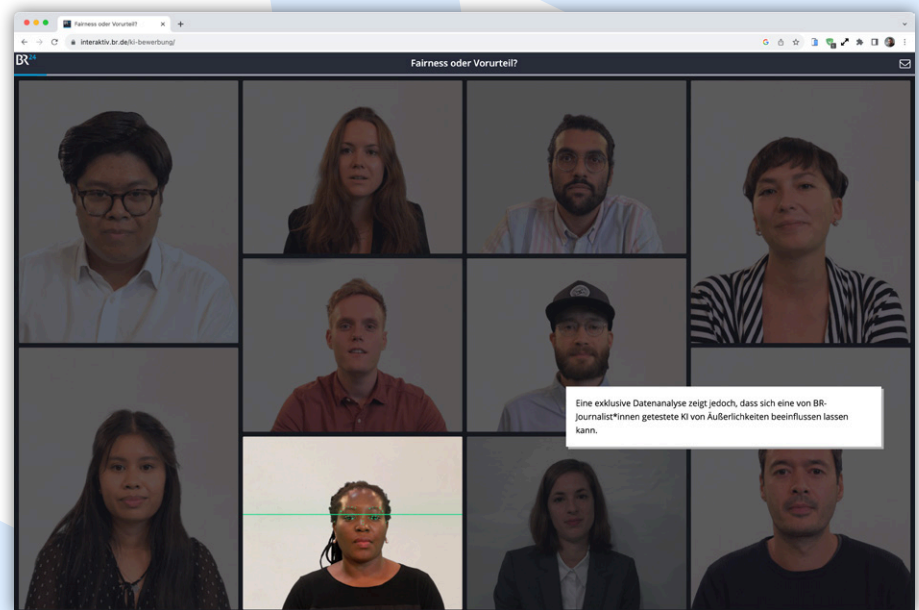
Doch als Osenstätter einen schwarzen BH anzieht, ändert sich das Ergebnis. Obwohl sein Oberkörper mit dem Kleidungsstück nun mehr verhüllt ist als zuvor, schätzen die zwei KI-Modelle das Video nun als sehr anzüglich ein. Offenkundig wurden die zugrundeliegende Bilderkennungsalgorithmen anhand von Beispielen so trainiert, dass sie Bilder von Menschen mit einem BH per se auf diese Art bewerten.

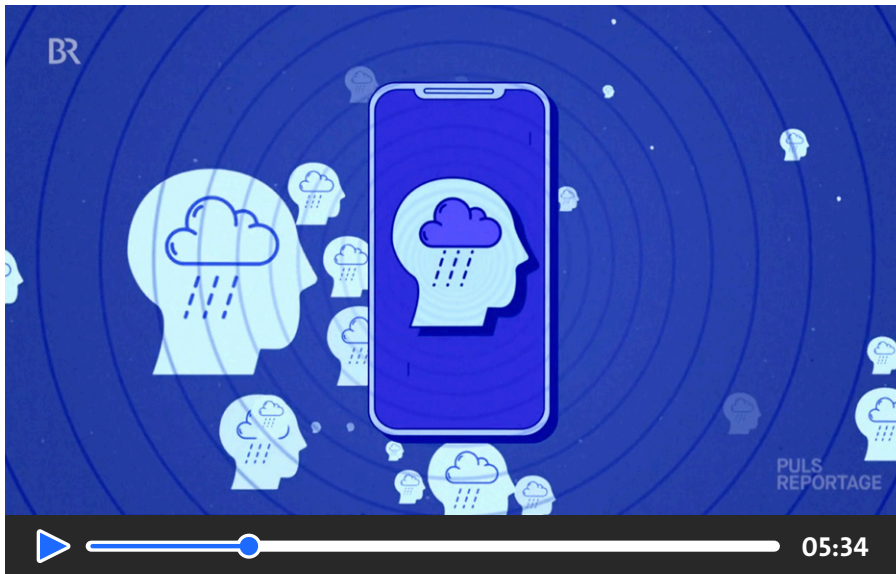
Fairness oder Vorurteil?

Die Recherche „Fairness oder Vorurteil? Fragwürdiger Einsatz von Künstlicher Intelligenz bei der Jobbewerbung“ offenbart, dass das Produkt eines deutschen Start-ups, das Videos von Bewerber*innen analysiert, fehlerhaft arbeitet. Anstatt die Persönlichkeit der Bewerber*innen zuverlässig einzuschätzen, vermutete das System für dieselbe Person je nach Video-Hintergrund völlig unterschiedliche Charaktereigenschaften.

BR Data produzierte mit Testpersonen mehrere hundert Videoclips. Das Ziel: Herauszufinden, ob unterschiedliche Faktoren die Software beeinflussen können und wie sie sich auf die Persönlichkeitseinschätzung auswirken. Ein Vergleich wurde auf zwei verschiedene Arten durchgeführt: Einerseits antwortete eine Schauspielerin in verschiedenen Outfits mit immer gleichem Text und Stil auf die verschiedenen Job-Interviewfragen. Andererseits haben Video-Producer eine größere Menge an aufgezeichneten Videos einer diversen Gruppe von Testpersonen mit technischen Mitteln verändert. So konnte in beiden Szenarien sichergestellt werden, dass bei jedem Versuch nur jeweils ein einzelner Faktor gezielt verändert wurde. Das Ergebnis der Software unterschied sich teilweise erheblich.

Weniger Vorurteile, mehr Objektivität. Ein Bewerbungsprozess, der nicht vom persönlichen Geschmack eines Personalers abhängt. So versprechen es viele KI-Unternehmen, die weltweit auf den Markt drängen. Auch ein Start-up aus München. Laut Angaben der Entwickler der Software analysiert die Künstliche Intelligenz Stimme, Sprache, Gestik und Mimik und erstellt so ein verhaltensbasiertes Persönlichkeitsprofil.





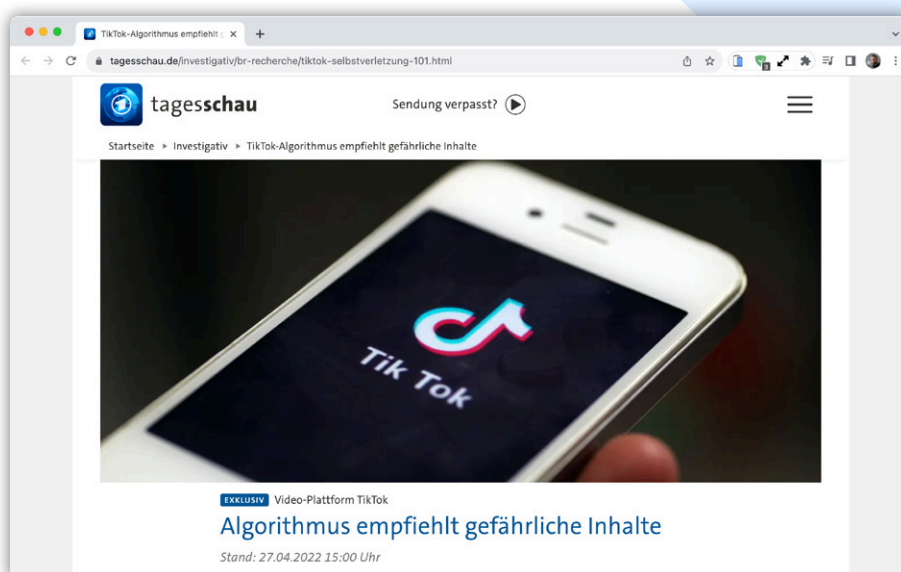
Videos über Selbstverletzung und Suizidgedanken sind nach wie vor ein Problem auf TikTok. Das zeigt eine Datenanalyse des BR. Nutzerinnen und Nutzer können demnach in eine Filterblase aus gefährlichen Inhalten gezogen werden. Denn stärker als bei anderen sozialen Netzwerken entscheidet bei TikTok ein Algorithmus darüber, was Nutzerinnen und Nutzer zu sehen bekommen.

Wie gefährlich ist TikTok?

Ein Algorithmus kuratiert den wichtigsten Teil der TikTok-App, den „For You“-Feed. Dort findet sich ein ganzes Subgenre, unter dem junge Menschen Videos über Depression, Selbstverletzung oder Suizid produzieren. Ein Experiment von BR Data zusammen mit PULS Reportage zeigt, dass deutsche Nutzer*innen durch Interaktion mit solchen Videos in eine Filterblase geraten können, in der sie in ihrem Feed fast nur noch diese Inhalte von TikTok empfohlen bekommen. Für das Experiment setzten die Journalist*innen von BR Data mehrere Test-Accounts

auf und simulierten das Verhalten von Personen, die sich für Videos zu Depression, Selbstverletzung und Suizidgedanken interessieren. Das Ergebnis: Der Feed bestand nach kurzer Zeit fast nur noch aus solchen Inhalten. Schon nach etwa 150 Videos war im Schnitt bereits jedes dritte Video mit einem Hashtag zu Themen wie Traurigkeit, Depression, Selbstverletzung und Suizidgedanken versehen, in der Praxis war das nach rund 45 Minuten Nutzungsdauer der Fall.

Um Blackbox-Algorithmen zu untersuchen, sind Experimente oft die einzige Möglichkeit einer Annäherung. Journalist*innen bestimmen die Eingaben (Input) für einen Algorithmus und beobachten die Ausgaben (Output) – was dazwischen passiert, ist für die Reporter*innen nicht einsehbar. Neben dem Sammeln von ausreichend Daten ist es dabei wichtig, mögliche unerwünschte Einflüsse auf der Input-Seite so weit wie möglich zu eliminieren. Die BR-Teams nutzen diese Art von Experimenten, um zu verstehen, ob Algorithmen tun, wofür sie vorgesehen sind oder ob die Systeme diskriminieren, fehlerhafte und damit unbrauchbare Ergebnisse liefern, oder ihre Nutzer*innen auf manipulative Weise beeinflussen.

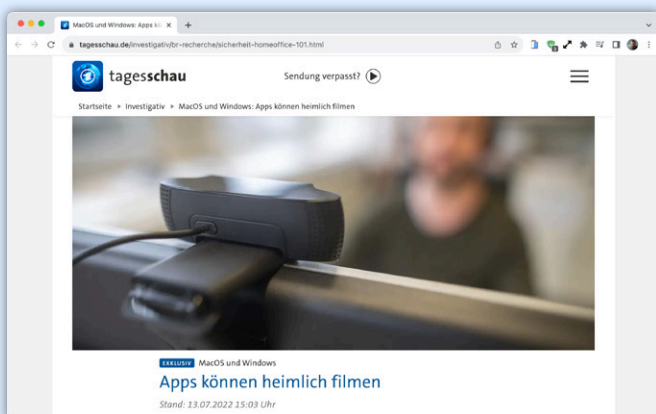
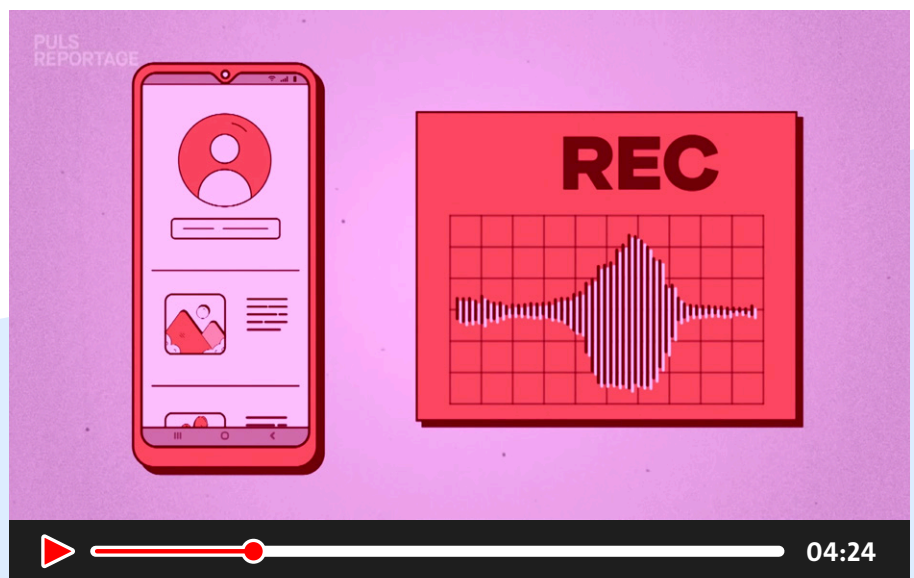


#2

SOFTWARE AUF DEM PRÜFSTAND: RECHERCHE MIT TECHNISCHEN ANALYSEN

Ein Team des BR hat die Smartphone-Betriebssysteme Android und iOS auf den Prüfstand gestellt, um herauszufinden, ob Apps heimlich zuhören können. Das Team programmierte für die Recherche eine eigene Smartphone-App. Sie zeigt, dass bei älteren, aber weit verbreiteten Android-Versionen alle Apps unter bestimmten Umständen völlig unbemerkt Audio-Signale mitschneiden können. War der Bildschirm ausgeschaltet, die App aber nicht geschlossen, klappte die versteckte Aufnahme über eine Stunde lang. In einem weiteren Test zeigte das Team, dass auch Programme, die auf PCs laufen, unter Umständen lange heimlich filmen, zuhören und den Bildschirminhalt abfotografieren können.

Mit solchen Tests und Methoden des Reverse-Engineerings – also der technischen Analyse von Systemen – untersuchen Journalist*innen die Funktionalität und Sicherheitslücken von Software. Dass die Ergebnisse dieser Recherchen auf breites Interesse stoßen, zeigt die Puls Reportage „[Können iPhone und Android Handys heimlich mithören? Wir programmieren Apps, um das zu beweisen!](#)“, die über eine Million Aufrufe auf YouTube zählt.



Viele haben es schon erlebt: Man sitzt mit Freunden beim Abendessen und unterhält sich über ein Produkt oder Thema, von dem man sicher ist, dass man es nie online gesucht hat. Kurze Zeit später tauchen in den eigenen sozialen Netzwerken plötzlich genau dazu passende Werbeanzeigen auf. Im Internet findet man viele Geschichten von Menschen, die schwören, dass ungewöhnliche Gesprächsthemen zu sehr spezifischer Online-Werbung geführt haben.

So überzeugend diese Erfahrungen auch klingen mögen, sicher bestätigen können solche anekdotischen Behauptungen nicht, dass Handy-Apps heimlich zuhören, um zielgerichtet Werbung zu platzieren.

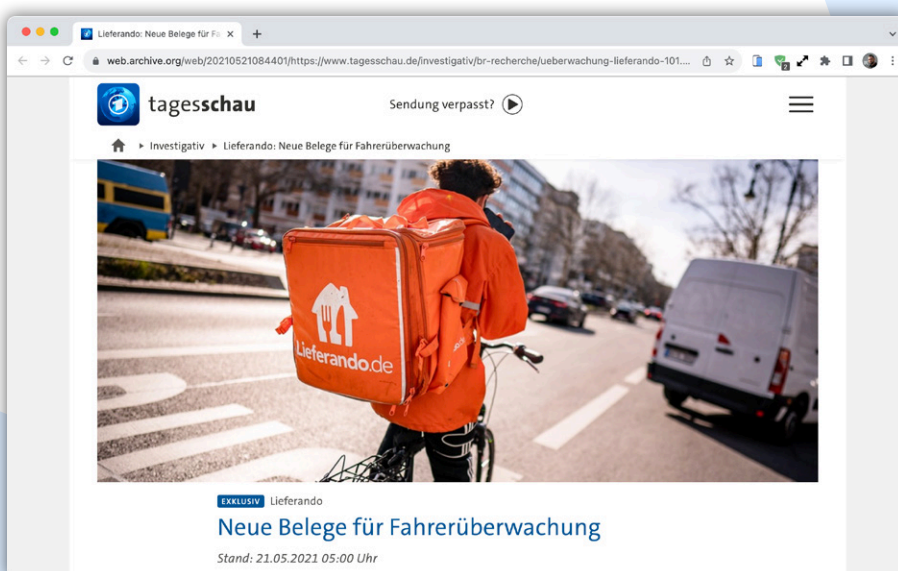
#3

RECHT AUF EINBLICK: RECHERCHE MIT JURISTISCHEN MITTELN

Um herauszufinden, ob die Lieferando-App die Beschäftigten des Unternehmens unzulässig überwacht, haben Reporter*innen des BR sogenannte „Selbstauskünfte“ der Rider ausgewertet. Diese hatten mehrere Lieferando-Fahrer gemäß der europäischen Datenschutz-Grundverordnung (DSGVO) von ihrem Arbeitgeber angefordert.

Aus den Unterlagen geht hervor, dass die App pro Lieferung 39 Datenpunkte erhebt: Sekundengenau ist nachvollziehbar, wann ein Fahrer eine Bestellung zugeteilt bekommt, diese abholt und ausliefert. Zusätzlich wird hinterlegt, ob der Rider einzelne Zeitvorgaben einhält oder nicht. Kommt ein Fahrer zu spät bei einem Restaurant oder Kunden an, wird das vermerkt.

Die Daten sind personalisiert, können also einzelnen Arbeitnehmerinnen oder Arbeitnehmern zugeordnet werden. In einigen dem BR vorliegenden Fällen wurden so für vollbeschäftigte Rider in einem Jahr mehr als 100.000 Datenpunkte erfasst.



„Jetzt habe ich noch 30 Sekunden“, sagt Daniel. Dann beginnt seine Schicht als Fahrer bei Lieferando, indem er auf „Play“ in der App „Scoober“ drückt: Er ist bereit für den ersten Auftrag. Daniel, der eigentlich anders heißt, versucht, die App keine Sekunde zu spät zu aktivieren. Denn er will nicht negativ auffallen: „Ich will ja pünktlich sein, weil die notieren sich sowas.“

„Notieren“ heißt in diesem Fall: Die App speichert seinen Schichtbeginn genau. Neben seinem Fahrrad und dem Rucksack ist „Scoober“ Daniels wichtigstes

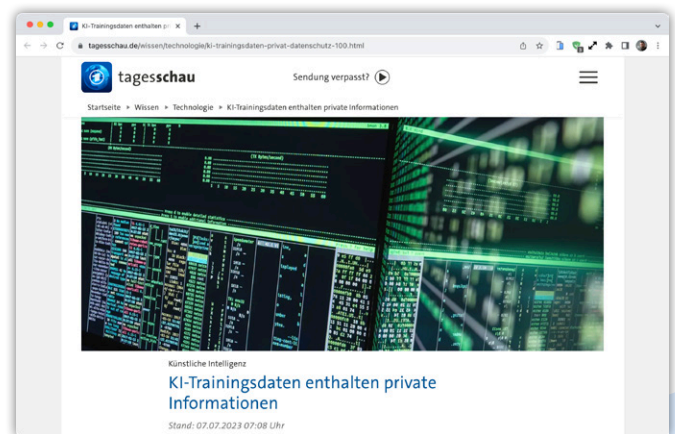
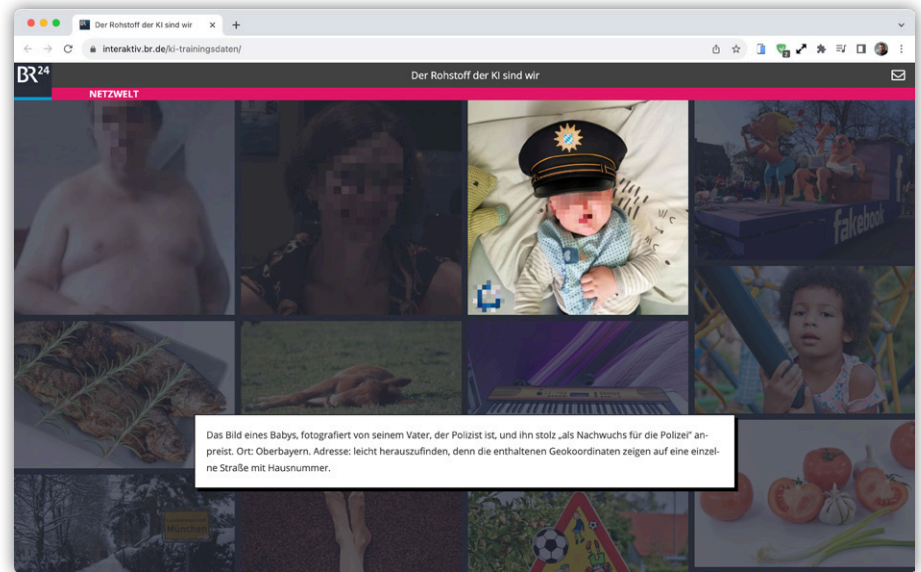
Arbeitsgerät. Über die App erhält er seine Aufträge. Was „Scoober“ sonst noch über ihn weiß, darüber habe ihn sein Arbeitgeber Lieferando nicht aufgeklärt: „Ich würde es aber gerne wissen“, sagt der Rider.

Der Rohstoff der KI

Auch bei der Recherche „Der Rohstoff der KI sind wir“ hat ein BR-Team die DSGVO als Instrument zur Recherche genutzt. Trainingsdaten für KI-Systeme bestehen aus unzähligen Bildern und Texten, die aus unterschiedlichen Quellen stammen, etwa von Webseiten oder aus Büchern.

Eine BR-Analyse zeigt: Sie enthalten sensible und private Daten – meist ohne, dass die Betroffenen Bescheid wissen. In den analysierten Trainingsdaten fand das Team ein Bild von BR-Journalistin Elisa Harlan. Sie nutzte die DSGVO für Anfragen an Unternehmen und Organisationen wie OpenAI, Midjourney und LAION – nicht alle waren erfolgreich.

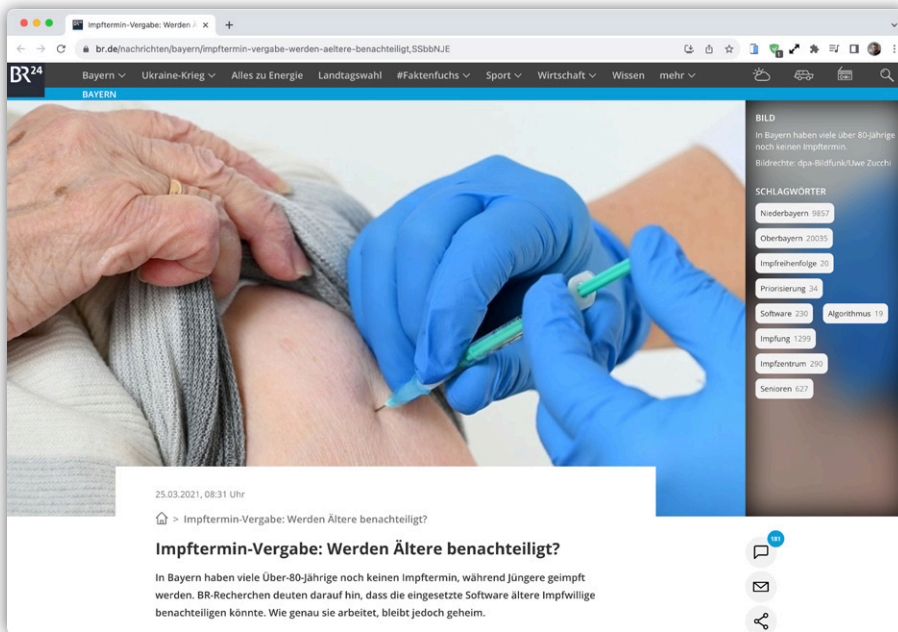
Seit Inkrafttreten der DSGVO haben Personen das Recht, Auskunft darüber zu verlangen, welche personenbezogenen Daten Unternehmen und Arbeitgeber über sie speichern. Je nachdem, wie detailliert die Antworten ausfallen, eignen sie sich als Tool für investigative Recherchen über Algorithmen, die in die Privatsphäre von Menschen eingreifen.



Was, wenn plötzlich ein privates Foto in einem Datensatz auftaucht, der benutzt wird, um Künstliche Intelligenz zu trainieren? Ein riesiger Datensatz, in dem nicht nur Fotos, sondern auch alle möglichen anderen Bilder aus dem Internet versammelt sind. Passiert ist das Elisa Harlan, ihr Foto ist Teil des größten öffentlich zugänglichen KI-Trainingsdatensatzes – ohne, dass sie davon wusste.

Ein Blick auf die Lieferkette für KI-Trainingsdaten und die Frage: Wie werden Bilder zum Rohstoff für Künstliche Intelligenz – und gibt es einen Weg zurück?

#4

ERKLÄR MIR DEN ALGORITHMUS:
INTERVIEWS UND ANFRAGEN

Das Ehepaar Peter und Erika S. aus dem Landkreis München: Beide sind in der höchsten Priorisierungsgruppe für die Corona-Schutzimpfung. Er ist 81 und wurde bereits einmal geimpft, seine 80-jährige Ehefrau wartet seit Wochen. Beide haben sich über die Software BayIMCO des Freistaats Bayern für einen Impftermin registriert. Doch nach welchen Kriterien die Termine genau vergeben werden, ist ihnen nicht klar. „Ich habe einige Freundinnen, die jünger sind als ich und die bereits geimpft wurden“, wundert sich Erika S.

Das Terminvergabe-System BayIMCO ordnete 2021 jeder registrierten Person einen Scorewert zu. Bei älteren Menschen entsprach dieser Wert dem Alter. Jüngere erhielten einen Zufallswert, der der Altersverteilung der Priorisierungsgruppe entsprach – in der Prio-Gruppe 1 etwa ein Zufallswert zwischen 80 und 90. Je höher der Wert, desto schneller bekommt man einen Termin. Dieses „Würfeln“ war möglicherweise der Grund, warum die 80-jährige Erika S. lange auf ihren Impftermin warten musste. Statistisch war sie mit ihrem verhältnismäßig niedrigen Altersscore gegenüber den zugeteilten Zufallszahlen im Nachteil.

Um genauer zu verstehen, wie der Algorithmus funktioniert, konnten BR-Reporter*innen mit dem Einverständnis des Ehepaars im örtlichen Impfzentrum Oberhaching deren Datenbankeinträge einsehen. Diese Daten sowie Interviews mit Betreiber*innen von Impfzentren und Anfragen an das bayerische Gesundheitsministerium lieferten konkrete Hinweise, warum der bayerische Impfalgorithmus ältere Menschen benachteiligte.

Diese und ähnliche Recherchen der BR-Teams zeigen: Berichten über Algorithmen geht auch mit klassischen journalistischen Methoden: mit Interviews und Dokumenten, Anfragen nach dem Informationsfreiheitsgesetz (IFG) oder Whistleblower*innen.



CHARLIE BECKETT

Head of London School of Economics Journalism [Think Tank Polis](#) and
Founder of the global [JournalismAI network](#)

„BR’s algorithmic accountability reporting is a model for journalists everywhere. Their investigations are informed, lucid and compelling. Working without moral panic or hyperbole they have given us in-depth, detailed and accessible insights into some of the key risks and hazards that new technologies such as artificial intelligence can bring, from bias in HR recruitment to invasions of intellectual property and privacy in training data. At this time where algorithmic technologies are more powerful in all aspects of our lives, we desperately need this kind of critical, independent and knowledgeable public service journalism.“

*„Das Algorithmic Accountability Reporting des BR ist ein Vorbild für Journalist*innen auf der ganzen Welt. Ihre Untersuchungen sind fundiert, klar und überzeugend. Ohne moralische Panik oder Übertreibung haben die BR-Journalist*innen uns tiefgreifende, detaillierte und zugängliche Einblicke in einige der wichtigsten Risiken und Gefahren gegeben, die neue Technologien wie künstliche Intelligenz mit sich bringen können, von Bias in KI-Recruiting-Systemen bis hin zu Verletzungen des geistigen Eigentums und der Privatsphäre bei Trainingsdaten. In einer Zeit, in der algorithmische Technologien immer stärker in alle Bereiche unseres Lebens eindringen, brauchen wir dringend diese Art von kritischem, unabhängigem und sachkundigem Journalismus im öffentlichen-rechtlichen Sektor.“*

ETHIK UND RECHT: WELCHE REGELN GELTEN FÜR DIE ARBEIT MIT KI IN REDAKTIONEN?

Der Bayerische Rundfunk gehörte 2020 weltweit zu den ersten Medienhäusern, die Richtlinien für den Umgang mit Künstlicher Intelligenz formuliert und veröffentlicht haben. Mittlerweile haben sich weitere Redaktionen intern auf Leitlinien zu KI verständigt und publiziert. Nicholas Diakopoulos und Hannes Cools analysierten im Juli 2023 für das Projekt „Generative AI in the Newsroom“ 21 internationale Richtlinien aus 13 Ländern. Folgende Punkte spielten dabei immer wieder eine Rolle:

- ✓ Strategische Ziele für den KI-Einsatz
- ✓ Erwünschte und unerwünschte Einsatzszenarien
- ✓ Redaktionelle Kontrolle und Abnahme von Inhalten, die mit KI-Unterstützung erstellt wurden
- ✓ Transparente Kennzeichnung solcher Inhalte
- ✓ Redaktionelle Verantwortung und Rechenschaftspflicht
- ✓ Schutz und Sicherheit von Daten und Quellen
- ✓ Verantwortlicher Rahmen für Experimente

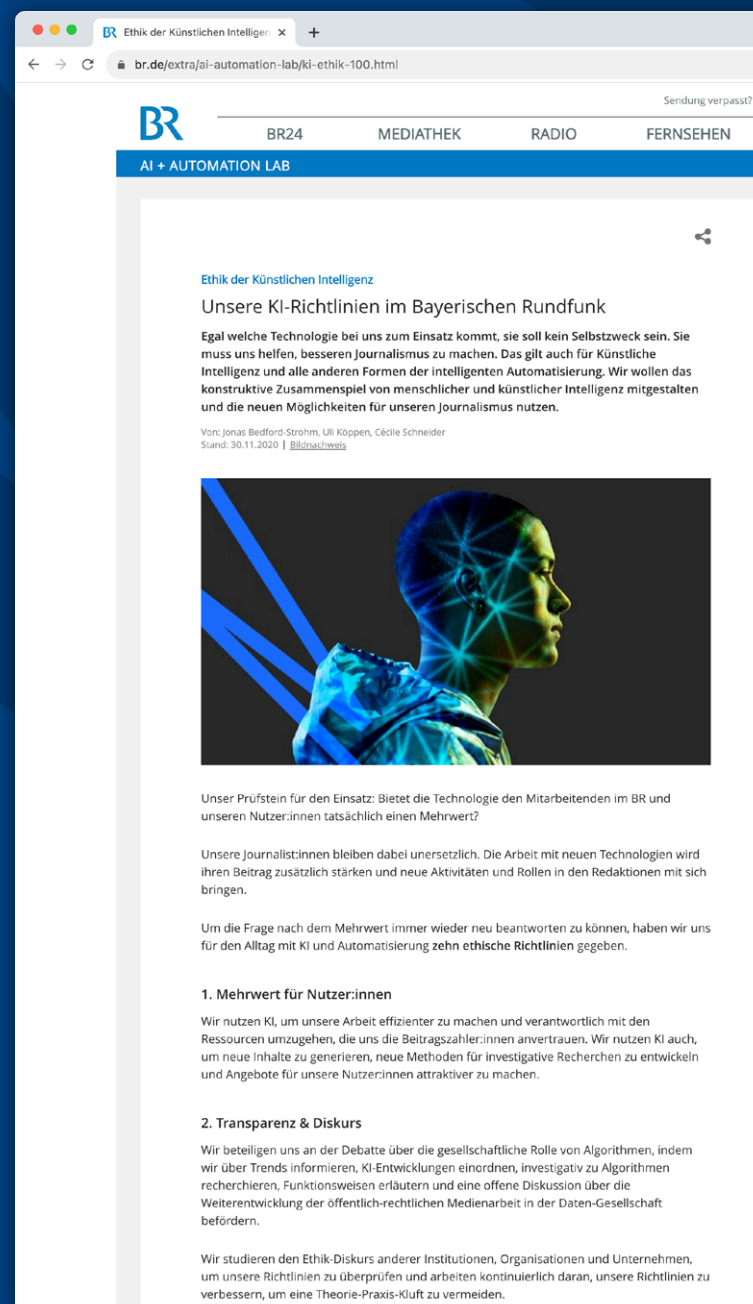
Auch die [Richtlinien für den Einsatz Künstlicher Intelligenz im BR](#) decken diese Aspekte ab. Allen voran steht das Bekenntnis zu den Mitarbeitenden: Journalist*innen sind dem BR dabei unersetzlich.

Künstliche Intelligenz für den öffentlich-rechtlichen Auftrag nutzen

KI soll im BR helfen, neue Inhalte für Nutzer*innen zu gestalten und bestehende attraktiver zu machen, Recherchemethoden und generell journalistische Arbeit zu unterstützen, und auch verantwortungsvoll mit Beitragsgeldern umzugehen.

Algorithmic Accountability Reporting soll dabei helfen, dem öffentlich-rechtlichen Informations- und Bildungsauftrag nachzukommen und dabei KI-Entwicklungen einzuordnen, investigativ zu Algorithmen zu recherchieren und Funktionsweisen der Technologie zu erläutern.

Algorithmen, automatisierte Entscheidungssysteme, maschinelles Lernen, Künstliche Intelligenz eingesetzt von Staaten, der Wirtschaft oder anderen Akteuren – egal, welche Schattierung die technologischen Phänomene genau haben: Wissen zu KI und Automatisierung in Redaktionen ist unerlässlich, um einen Beitrag zur gesellschaftlichen Debatte über Algorithmen leisten zu können.



BR Ethik der Künstlichen Intelligenz x +

br.de/extraj/ai-automation-lab/ki-ethik-100.html

BR24 MEDIATHEK RADIO FERNSEHEN

AI + AUTOMATION LAB

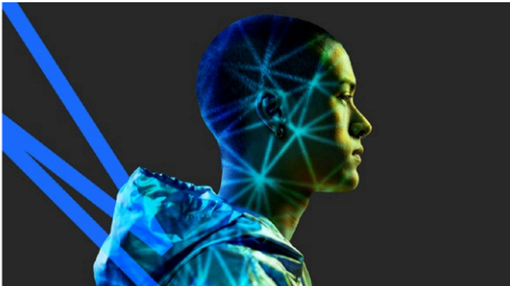
Sendung verpasst?

Ethik der Künstlichen Intelligenz

Unsere KI-Richtlinien im Bayerischen Rundfunk

Egal welche Technologie bei uns zum Einsatz kommt, sie soll kein Selbstzweck sein. Sie muss uns helfen, besseren Journalismus zu machen. Das gilt auch für Künstliche Intelligenz und alle anderen Formen der intelligenten Automatisierung. Wir wollen das konstruktive Zusammenspiel von menschlicher und künstlicher Intelligenz mitgestalten und die neuen Möglichkeiten für unseren Journalismus nutzen.

Von: Jonas Bedford-Strohm, Uli Köppen, Cécile Schneider
Stand: 30.11.2020 | [Bildnachweis](#)



Unser Prüfstein für den Einsatz: Bietet die Technologie den Mitarbeitenden im BR und unseren Nutzer:innen tatsächlich einen Mehrwert?

Unsere Journalist:innen bleiben dabei unersetzlich. Die Arbeit mit neuen Technologien wird ihren Beitrag zusätzlich stärken und neue Aktivitäten und Rollen in den Redaktionen mit sich bringen.

Um die Frage nach dem Mehrwert immer wieder neu beantworten zu können, haben wir uns für den Alltag mit KI und Automatisierung zehn ethische Richtlinien gegeben.

1. Mehrwert für Nutzer:innen

Wir nutzen KI, um unsere Arbeit effizienter zu machen und verantwortlich mit den Ressourcen umzugehen, die uns die Beitragszahler:innen anvertrauen. Wir nutzen KI auch, um neue Inhalte zu generieren, neue Methoden für investigative Recherchen zu entwickeln und Angebote für unsere Nutzer:innen attraktiver zu machen.

2. Transparenz & Diskurs

Wir beteiligen uns an der Debatte über die gesellschaftliche Rolle von Algorithmen, indem wir über Trends informieren, KI-Entwicklungen einordnen, investigativ zu Algorithmen recherchieren, Funktionsweisen erläutern und eine offene Diskussion über die Weiterentwicklung der öffentlich-rechtlichen Medienarbeit in der Daten-Gesellschaft befördern.

Wir studieren den Ethik-Diskurs anderer Institutionen, Organisationen und Unternehmen, um unsere Richtlinien zu überprüfen und arbeiten kontinuierlich daran, unsere Richtlinien zu verbessern, um eine Theorie-Praxis-Kluft zu vermeiden.

KI-PODCAST VON BR UND SWR GESTARTET

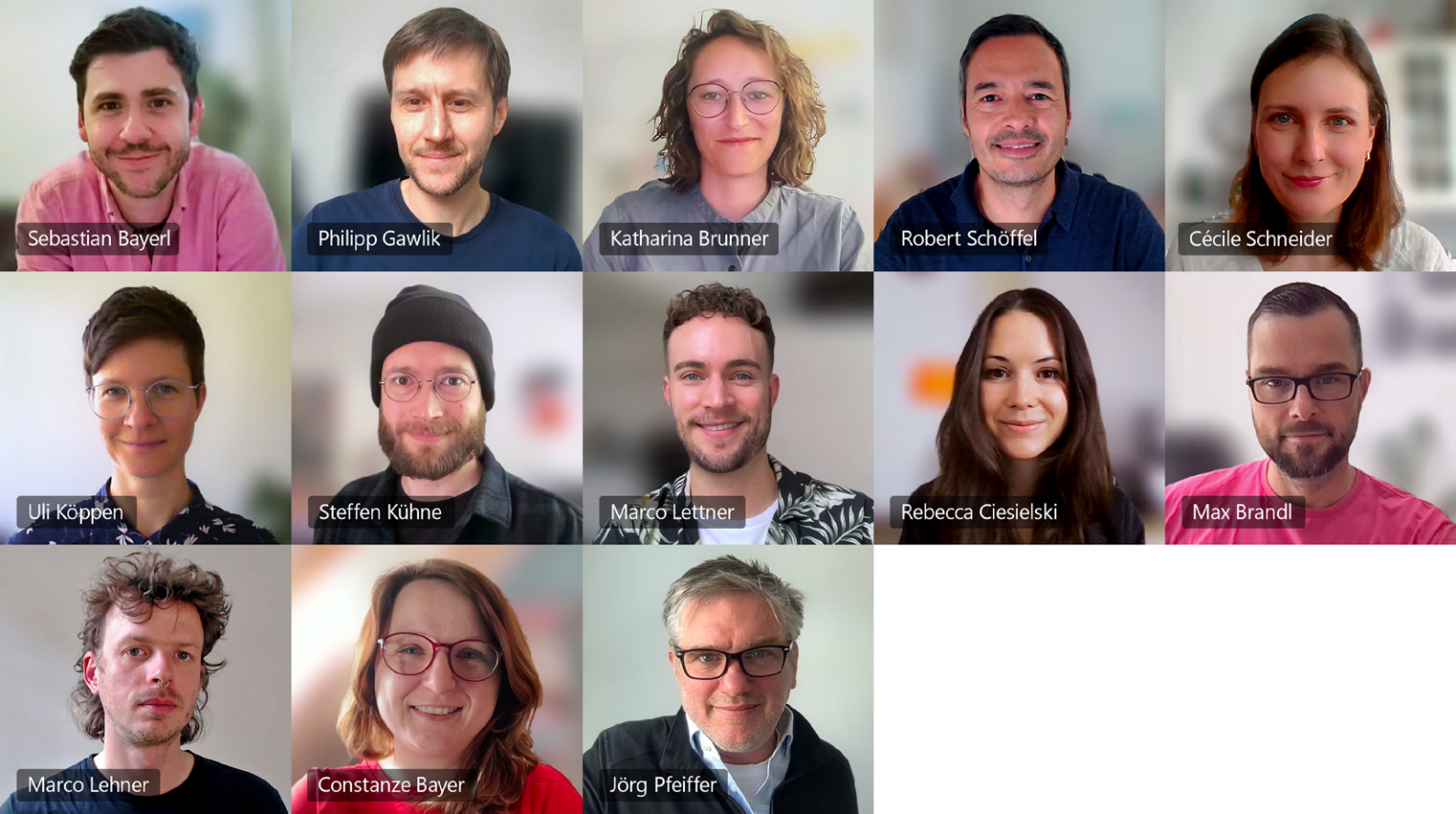
Künstliche Intelligenz als Thema der Berichterstattung ist längst auch in der Aktualität angekommen. Die KI-Teams von BR24 und SWR bieten in einem neuen wöchentlichen Podcast Orientierung: Über die technologischen Entwicklungen, über mögliche Folgen für Wirtschaft, Bildung, Gesellschaft und unsere Sicherheit – und über viele praktische KI-Anwendungen im Alltag. Präsentiert wird „Der KI-Podcast“ von Gregor Schmalzried, Marie Kilg und Fritz Espenlaub – drei Hosts, die viel Erfahrung im Tech-Journalismus mitbringen.



„Die Möglichkeiten Künstlicher Intelligenz können begeistern, aber genauso stark verunsichern. Klar ist: Wir stehen vor einer Revolution. Unser Podcast bereitet auf, wie sich unser Leben verändert. KI betrifft alle. Daher richtet sich der Podcast auch an alle, nicht nur an ein Fachpublikum“, so BR-Chefredakteur Christian Nitsche.

Der KI-Podcast ist [in der ARD-Audiothek verfügbar](#).





DAS AI + AUTOMATION LAB IM BR

Das AI + Automation Lab arbeitet seit 2020 als interdisziplinäres Team im BR an der Schnittstelle von Journalismus, Informatik und Produktentwicklung. Das Lab setzt auf Künstliche Intelligenz und Automatisierung, um sie für einen nutzerzentrierten Journalismus einzusetzen und investigativ zu hinterfragen. Expert*innen für Maschinelles Lernen arbeiten zusammen mit Datenjournalist*innen, Designer*innen und Produktentwickler*innen. Geschichten und Produkte entstehen in interdisziplinärer Zusammenarbeit. Die Schwerpunkte sind dabei Algorithmic Accountability Reporting und die Entwicklung KI-unterstützter Newsformate und -Produkte.

br.de/extra/ai-automation-lab

twitter.com/BR_Allab



IMPRESSUM

Herausgeber

Bayerischer Rundfunk
Anstalt des öffentlichen Rechts
[AI + Automation Lab](#)
Rundfunkplatz 1
80335 München

Redaktion

Cécile Schneider
Katharina Brunner
Rebecca Ciesielski
Uli Köppen

Gestaltung

Max Brandl

Veröffentlichung

August 2023