

# **Empfehlungsalgorithmen und Öffentlich-rechtliche Medien**

**Ein Whitepaper für Luxemburg**

*Stefan Börnchen, Georg Mein & Johannes Pause*



# **Empfehlungsalgorithmen und Öffentlich-rechtliche Medien**

## **Ein Whitepaper für Luxemburg**

Stefan Börnchen

Georg Mein

Johannes Pause

unter Mitarbeit von Ruth Reicher und Hana Reichert

ULIDE Papers - Band 1

Herausgegeben von Georg Mein und Johannes Pause



Melusina Press, 2025

Veröffentlicht durch die Universität Luxemburg - **Melusina Press**, 2025  
11, Porte des Sciences  
L-4366 Esch-sur-Alzette  
<https://www.melusinapress.lu>

Verlagsleitung: Niels-Oliver Walkowski, Johannes Pause

Lektorat: Carolyn Knaup, Niels-Oliver Walkowski

Cover und Layout: Valentin Henning, Erik Seitz

Umschlagsbild: Foto von VT98Fan auf Wikipedia (CC BY-SA 3.0)

Die digitale Version dieser Publikation steht unter <https://www.melusinapress.lu> frei zur Verfügung.

Das PDF und die Druckvorlage werden mit Hilfe von princeXML erzeugt.

Bibliografische Information der Nationalbibliothek Luxemburg: Die Nationalbibliothek Luxemburg verzeichnet diese Publikation in der Luxemburgischen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über [bnl.public.lu](http://bnl.public.lu) abrufbar.

ISBN (Online)	978-2-919815-98-2
ISBN (PDF)	978-2-919815-99-9
DOI (Publikation)	10.26298/1981-5982-euom
DOI (Band)	10.26298/1981-5982

Das vorliegende Werk steht unter einer CC BY-SA 4.0 Lizenz. Informationen zu dieser Lizenz finden Sie unter <https://creativecommons.org/licenses/by-sa/4.0/deed.de>. Die in diesem Werk enthaltenen Bilder und Ressourcen unterliegen der selben Lizenz, sofern sie keiner anderen Quelle entnommen wurden oder mit einer anderen Lizenz versehen sind.





## Zusammenfassung

Das vorliegende Whitepaper untersucht, wie sich Empfehlungsalgorithmen (Recommender Systems) auf die Öffentlichkeit als Kernelement der deliberativen Demokratie auswirken. Das Whitepaper liefert eine theoretisch fundierte und evidenzbasierte Grundlage für politische Entscheidungsträger:innen, um die Herausforderungen und Chancen der digitalen Transformation zu verstehen und, wo nötig, zu regulieren.

Öffentlich-rechtliche Medien tragen eine besondere Verantwortung für die Förderung einer pluralistischen Öffentlichkeit. Vor diesem Hintergrund stehen diese Medien jedoch vor der schwierigen Entscheidung, ob sie Empfehlungsalgorithmen verwenden sollen, die nicht grundlegenden journalistischen Prinzipien verpflichtet sind beziehungsweise sogar gegen diese verstößen. So kann der Einsatz von Empfehlungsalgorithmen in Sozialen Netzwerken nicht nur auf für die User:innen intransparente Weise die Auswahl von Themen und Meinungen steuern, sondern auch polarisierende Phänomene wie Filterblasen und Echokammern fördern.

Ein großes gesellschaftliches Problem ist die wachsende Macht von digitalen Plattformen wie YouTube, Instagram und TikTok. Sie beeinflussen in erheblichem Maße die öffentliche Meinungsbildung. Dabei verzichten die Digitalplattformen prinzipiell auf Transparenz und redaktionelle Verantwortung; sie operieren allein nach den marktwirtschaftlichen Kriterien der sogenannten Aufmerksamkeitsökonomie. Wenn auf diese Weise individuelle Vorlieben zu den Hauptkriterien für die Auswahl von Nachrichten und anderen kulturellen Inhalten werden, besteht die Gefahr der einseitigen Meinungsbildung und gesellschaftlichen Fragmentierung. Es liegt im Interesse einer starken Demokratie, die Macht von nicht rechenschaftspflichtigen Akteuren zu begrenzen.

Insbesondere die Diskussion um die Echokammern zeigt, dass manche populäre Ansichten über Empfehlungsalgorithmen der empirischen Überprüfung nicht standhalten. Die Wirkungsweise von Empfehlungsalgorithmen muss daher in einem komplexen Zusammenhang betrachtet werden: von ihrem Design über ihre Entwicklung bis hin zu ihrer Nutzung und der Psychologie der User:innen, bei der der sogenannte „confirmation bias“ eine zentrale Rolle spielt.

Nicht zu unterschätzen ist die Bedeutung des Vertrauens in die Medien. Um Polarisierung und Verschwörungsglauben zu bekämpfen, muss verlorenes Vertrauen in die Absender wiederhergestellt werden.

Dieses Whitepaper gibt eine Reihe von Empfehlungen zur Gestaltung und Kuratierung von Empfehlungsalgorithmen in den Öffentlich-rechtlichen Medien:

- Menschliche Akteure sollten Teil der Entwicklung und Überwachung von Empfehlungsalgorithmen sein und erkennbar redaktionelle Verantwortung tragen.
- Das Design von Empfehlungsalgorithmen sollte für die User:innen durchschau- bar und anpassbar sein.
- Personalisierung und Content Diversity sollten ausgewogen gemischt sein.
- Die Korrektur von Fehlinformationen erfordert Fingerspitzengefühl. Entschei- dend sind hier nicht Fakten und Argumente, sondern das Vertrauen in den Ab- sender.

Nur eine multidisziplinäre Herangehensweise, die zugleich technologische, ethische und soziale Aspekte berücksichtigt, kann die gesellschaftlichen Auswirkungen von Empfehlungsalgorithmen verstehen und steuern.

# 0. Ausgangslage

## 0.1 Grundsatzfragen: Algorithmen und Polarisierung

Was hat die **gesellschaftliche Polarisierung**, die heute vielfach konstatiert wird (vgl. Münch 2023), mit dem Prozess der Digitalisierung und genauer: mit den großen Digitalplattformen zu tun? Wird diese Polarisierung beziehungsweise Segmentierung gesellschaftlicher Zusammenhänge durch Big Data verstärkt oder sogar verursacht? Welche Rolle spielt dabei die zunehmende Personalisierung von Angeboten im Netz durch datenbasierte Algorithmen? Erzeugt sie Filterblasen und Echokammern, und wenn: Wie wirken sich diese auf die demokratische Öffentlichkeit aus?

Solche und weitere Fragen stellen sich, wenn es um den Einsatz von Empfehlungsalgorithmen in öffentlich-rechtlichen Kontexten in Luxemburg und in anderen europäischen Ländern geht.

## 0.2 Politische Forderungen an Empfehlungsalgorithmen (B, L, D)

In Belgien wird der öffentlich-rechtliche Sender RTBF in einem neuen *Contrat de gestion* (2023–2027) von der Regierung ausdrücklich dazu aufgefordert, einen öffentlich-rechtlichen Algorithmus zu entwickeln, der eine **emotionale Polarisierung** verhindert.<sup>1</sup> In Luxemburg verlangt die neue Konvention mit dem staatlich finanzierten Rundfunkunternehmen 100,7: „[L]es algorithmes ou mécanismes de recommandation doivent refléter les valeurs du service public et ne pas constituer un frein au pluralisme réceptionné. La découverte de contenus doit être favorisée.“ (Gouvernement 2024: 5) In Deutschland sollen Empfehlungssysteme laut dem Dritten Medienänderungsstaatsvertrag, § 30 Abs. 4 Satz 2 „einen offenen Meinungsbildungsprozess und breiten inhaltlichen Diskurs ermöglichen“ (MStV 2023: 34).

In Reaktion auf eine 2021 veröffentlichte EU-Leitlinie wurde im Juni 2022 von den großen Online-Plattformen sowie Akteuren aus der Werbeindustrie, dem Journalismus, der Forschung und der Zivilgesellschaft ein „Code of Practice on Disinformation“ unterzeichnet, der unter anderem darauf zielt, algorithmische Transparenz zu stärken und der durch Empfehlungsalgorithmen unterstützten Monetarisierung von Desinformationen den Kampf anzusagen (Europäische Kommission 2022).

<sup>1</sup> „Cet algorithme est gouverné par les principes directeurs du service public, dont l'éthique, l'ouverture et la transparence. Il favorise, non pas la polarisation vers des contenus à fort engagement émotionnel, mais la découverte de contenus diversifiés et de qualité, prioritairement en matière d'information, de culture et d'éducation, sans toutefois détourner l'usager du contenu qu'il juge intéressant.“ (RTBF 2023: 44)

## **0.3 Gliederung, Methode und Erkenntnisziel: Entscheidungsspielräume für Trade-offs**

Das vorliegende Whitepaper gliedert sich in zwei Teile. Der erste Teil fasst die **internationale Forschungslage** zusammen und beantwortet damit die gerade formulierten allgemeinen Fragen. Dazu ist ein **inter- beziehungsweise transdisziplinäres Vorgehen** erforderlich. Denn die Forschung, die sich mit den Auswirkungen der Digitalisierung im Allgemeinen und mit Empfehlungsalgorithmen im Besonderen auseinandersetzt, stammt nicht nur aus ganz unterschiedlichen (technischen, naturwissenschaftlichen, sozial- und geisteswissenschaftlichen) Disziplinen (und dort wiederum aus zum Teil völlig neuen disziplinären und interdisziplinären Richtungen), sondern ist auch so reichhaltig, dass sie kaum noch vollständig überblickt werden kann. Ergänzt wird die wissenschaftliche Forschungslage durch eine kaum noch zu überschauende Vielzahl von Studien, White Papers, Leitlinien, Dokumentationen, Gesetzesentwürfen und Reporten, die im Auftrag von Institutionen, Medienanstalten, Beratungsfirmen, Thinktanks und politischen Einrichtungen erstellt wurden. Das vorliegende Papier nimmt neben der wissenschaftlichen Forschungsliteratur auch diese Dokumentations- und Empfehlungsliteratur in den Blick.

Der zweite Teil dieses Whitepapers formuliert dann **konkrete Leitlinien** für Empfehlungsalgorithmen von Öffentlich-rechtlichen Medien in Luxemburg. Am Ende wird knapp benannt, wo dringend weiterer wissenschaftlicher **Forschungs- und politischer Diskussionsbedarf** besteht.

Analysiert werden Funktionsweisen und mögliche Gefahren, um vor allem aus staatlicher Sicht sinnvolle Nutzungsmöglichkeiten von Empfehlungsalgorithmen („Recommender Systems“), also algorithmischen Empfehlungssystemen, offenzulegen. Die vorliegende Studie nimmt dabei methodisch eine medien- und kommunikationswissenschaftliche Perspektive ein, berücksichtigt jedoch auch Ergebnisse anderer Disziplinen, insbesondere der Philosophie (Ethik), der Informatik, der Soziologie und der Politik- und der Rechtswissenschaft. Umrissen werden sowohl theoretische wie empirische Grundlagen, um konkrete, nachvollziehbare sowie – darauf liegt hier ein besonderes Augenmerk – praktisch umsetzbare Vorschläge zu machen.

# **1. Herausforderungen, Probleme, Debatten und Forschungsstand**

## **1.1 Journalistische Redaktion versus algorithmische Personalisierung**

In demokratischen Gesellschaften gehört es zu den zentralen Aufgaben der Medien – und insbesondere solcher Medien, die durch die Beiträge der Steuerzahler:innen (mit-)finanziert werden –, **Öffentlichkeit** herzustellen (Habermas 1990). So komplex das theoretische Konzept ‚Öffentlichkeit‘ auch ist, es herrscht in der Forschung immerhin Konsens darüber, dass in der Öffentlichkeit diejenigen Diskussionen und Meinungsbildungsprozesse – die „Deliberation“ (Habermas 2021) – stattfinden sollen, die Voraussetzung einer funktionierenden liberalen Demokratie sind. Öffentlichkeit, wie sie von Medien hergestellt wird, fungiert als zentrales Forum, in dem Bürger:innen sowohl informiert werden als auch Meinungen und Perspektiven austauschen können. Hierbei geht es allerdings um weit mehr als um die schlichte Weitergabe von Informationen. Vielmehr soll, so die emphatische Vorstellung in demokratischen Gesellschaften, durch die mediale Öffentlichkeit sichergestellt werden, dass die Bürger:innen auf der Grundlage von korrekten, transparenten und umfassenden Informationen, aber auf mit Blick auf kulturelle Inhalte oder Meinungsartikel Entscheidungen treffen können. Insbesondere mündige Wahlentscheidungen können nur getroffen werden, wenn ein bestimmter Grad an Informiertheit in der Bevölkerung vorhanden ist, der nicht anders als durch mediale Öffentlichkeit generiert werden kann.

Die Rolle öffentlich-rechtlicher Medien in diesem Kontext liegt auf der Hand. Es geht hier nicht bloß um Berichterstattung, sondern vor allem um die Qualitätskontrolle der gesendeten Inhalte sowie darum, ein breites Meinungsspektrum und differenzierte Perspektiven sicherzustellen.

In der zweiten Hälfte des 20. Jahrhunderts hat sich in der westlichen Welt eine **redaktionell organisierte, weltanschaulich und politisch plurale Medienlandschaft** herausgebildet. Medien – soweit herrscht in der Wissenschaft Konsens – garantieren Meinungsvielfalt, indem sie die unterschiedlichsten Standpunkte reflektieren und auf diese Weise nicht nur permanent neu abbilden, sondern – auch – erzeugen. Gesteuert wird dieser Prozess von Redaktionen, die in ihrer Berichterstattung **journalistischen Prinzipien** wie **Relevanz, Vielfalt und Ausgewogenheit** verpflichtet sind. Damit dienen sie dem öffentlichen Interesse.

Mit dem Aufkommen des Internets und der Digitaltechnologie am Ende des Jahrhunderts erfuhr diese Medienstruktur eine bahnbrechende Transformation. Nun konnten Informationen in Echtzeit ausgetauscht und individuell zugeschnittene Inhalte angeboten werden. Heute wird das traditionelle Mediensystem durch digitale Medien und insbesondere Soziale Netzwerke herausgefordert. Die neuen Medienplattformen agie-

ren jenseits der etablierten journalistischen Normen und bieten ein Ökosystem, das von Nutzerinteraktionen und Datenanalysen geprägt ist. Denn digitale Medien folgen nicht redaktionellen und journalistischen Ansprüchen und Idealen, sondern in erster Linie einer algorithmischen Logik, die die großen Digitalplattformen diktieren und die sich an den Gesetzen der „**Aufmerksamkeitsökonomie**“ (Zuboff 2019) orientiert. Das bedeutet, dass Inhalte, die emotionale Reaktionen hervorrufen oder stark polarisieren, oft auf größeres Interesse stoßen als andere und somit häufiger verbreitet werden. Für die Digitalplattformen sind grundsätzlich weder inhaltliche Vielfalt noch sonstige inhaltliche Kriterien ausschlaggebend; Hauptinteresse ist, die **Bildschirmzeit** ihrer User:innen zu **maximieren**. Digitalplattformen wie Facebook, Instagram oder X (vormals Twitter) sind nicht den deliberativen Prinzipien einer demokratischen Öffentlichkeit verpflichtet, sondern folgen in erster Linie privatwirtschaftlichen Interessen und Ideologien (Daub 2020).

**Empfehlungsalgorithmen** zählen zu den wichtigsten Mechanismen digitaler Plattformen. Es handelt sich um hochkomplexe mathematische Modelle, die umfangreiche Datenmengen nutzen, um **Vorhersagen** darüber zu treffen, welche Inhalte einem Nutzer am wahrscheinlichsten gefallen oder von ihm angeklickt werden: „techniques are [...] subtle – such as the use of emotions to frame headlines and manipulate how people feel: clickbait.“ (van der Linden 196). Auf diese Weise **steuern** solche Algorithmen, so das definierte Ziel, unter Berücksichtigung des **Nutzer:innenverhaltens** (also der Daten, die die Nutzer:innen durch ihre Aktivitäten im Netz hinterlassen haben) die Auswahl und Präsentation von Inhalten. Dabei fließen nicht nur Vorlieben und Interessen der Nutzer:innen in diese Berechnungen ein, sondern auch soziale Interaktionen, Standortdaten und weitere individuelle Verhaltensmuster, sofern der Algorithmus auf diese Daten zugreifen kann.

Die Vorhersagen von Empfehlungsalgorithmen gründen sich häufig auf den Prinzipien der Ähnlichkeit und der historischen Interaktion der Benutzer:innen. Dabei existieren verschiedene Arten und Funktionslogiken, die bei Empfehlungssystemen Verwendung finden. Zu nennen sind (vgl. Aggerwal 2016):

- **Klassifizierungsalgorithmen:** Diese Algorithmen gruppieren Nutzer:innen basierend auf ihren Interessen oder Verhaltensweisen. Hierbei können Entscheidungsbäume, Support-Vector-Machines oder sogenannte naive Bayes-Klassifikatoren Anwendung finden.
- **Regressionsalgorithmen:** Hier wird versucht, eine bestimmte Ausgabe (wie z.B. Kaufverhalten) auf der Grundlage eines Eingabesatzes vorherzusagen. Ansätze wie lineare Regression und Logistik-Regression sind hierbei gebräuchlich.
- **Filteralgorithmen:** Diese konzentrieren sich auf das Herausfiltern relevanter Inhalte basierend auf bestimmten Kriterien, oft durch Nutzung von Ähnlichkeitsmetriken wie dem Cosinus-Abstand.

- **Kollaboratives Filtern:** Ein Ansatz, der auf den Beziehungen zwischen Nutzer:innen und Items basiert, wobei Vorhersagen auf den Interaktionen ähnlicher Nutzer:innen beruhen.
- **Content-basiertes Filtern:** Bei diesem Ansatz werden Empfehlungen auf der Grundlage von Merkmalen des Inhalts und des Nutzer:innenprofils erstellt.
- **Deep Learning:** In jüngerer Zeit haben neuronale Netzwerke, insbesondere tiefe neuronale Netzwerke, an Popularität gewonnen, da sie in der Lage sind, komplexe nichtlineare Beziehungen in großen Datenmengen zu erkennen und zu lernen.

Gemeinhin wird **Empfehlungsalgorithmen unterstellt**, dass sie **Filterblasen und Echozimmern**, also mehr oder weniger hermetisch abgeschlossene (politische) Milieus **erzeugen**, in denen Nutzer:innen nur mit Meinungen konfrontiert werden, die ihre eigenen Überzeugungen und Vorurteile bestätigen. Dieses Phänomen wird für eine in jüngerer Zeit beobachtete **zunehmende gesellschaftliche Polarisierung** und, damit einhergehend, **abnehmende generelle Ambiguitätstoleranz und Kompromissbereitschaft** in der Gesellschaft verantwortlich gemacht. Digitale Empfehlungsalgorithmen schaden also, so hat es den Anschein, der pluralistischen Öffentlichkeit und arbeiten damit gegen die Anliegen der traditionellen redaktionellen Medien, die diese Öffentlichkeit mitschaffen und mitgestalten. In diesem Kontext wird nach der Rolle und Verantwortung der Digitalplattformen in öffentlichen Debatten gefragt, und es wachsen Forderungen nach einer transparenten und ethischen Gestaltung ihrer Algorithmen, um der Gefahr einer weiteren gesellschaftlichen Spaltung entgegenzuwirken.

Vor die Frage gestellt, ob sie digitale Empfehlungsalgorithmen nutzen sollen, befinden sich die traditionellen, insbesondere die **Öffentlich-rechtlichen Medien** also in einem **Dilemma**: Arbeiten sie gegen ihren Auftrag und gegen ihre eigenen Interessen, wenn sie digitale Empfehlungsalgorithmen verwenden, und **schaffen sie sich** damit am Ende **selbst ab**? Oder **stärken sie ihre Wettbewerbsfähigkeit** und garantieren damit das Fortbestehen einer – wenn auch modifizierten – pluralen Öffentlichkeit? Das Dilemma lässt sich nicht ausräumen; es erfordert einen sinnvollen, mit Augenmaß getroffenen Kompromiss. Welcher Kompromiss in einem Korridor von möglichen Trade-offs getroffen wird, ist eine politische Entscheidung. Ein Beispiel: Eine Empfehlung kann transparent gemacht werden, indem man die Daten zugänglich macht, auf denen sie basiert. Daraus kann sich jedoch ein Widerspruch zum Datenschutz ergeben. Es gilt also, einen Trade-off zwischen Transparenz und Datenschutz zu finden. Grundsätzlich stellen alle Entscheidungen, um die es im Folgenden gehen wird, solche Trade-offs dar; sie alle sind im Kern **politische Entscheidungen** zwischen einander entgegenlaufenden Interessen.

Zum Anspruch auf Klarheit und **Machbarkeit** der in diesem Papier gegebenen Empfehlungen gehört es, **nur Empfehlungen zu geben, die umsetzbar sind**. Transparenz von komplexen Deep-Learning-Algorithmen kann etwa nicht eingefordert werden, wenn zugleich gilt, dass solche Algorithmen eine Black box bilden, in die sich nicht hineinschauen lässt.

## 1.2 Funktionsweisen von Empfehlungsalgorithmen (Recommender Systems)

Empfehlungsalgorithmen sind algorithmische Systeme, die bereits existierende Medieninhalte vorschreiben. Auf der Grundlage von aus Daten generierten Mustern unterbreiten Empfehlungsalgorithmen den User:innen Vorschläge, die auf ihr individuelles Profil ausgerichtet („personalisiert“) sind. Solche **personalisierten Angebote** kommen typischerweise in Sozialen Netzwerken zum Einsatz, besonders prominent etwa bei Facebook, Instagram, TikTok oder X, sowie auf Medienportalen und in Online-Shops wie YouTube, Netflix oder Amazon. Auch die **automatische Vervollständigung von Suchbegriffen** etwa in Internet-Suchmaschinen wird oftmals durch Empfehlungsalgorithmen geregelt.

Für die hier diskutierten Fragen sind insbesondere solche Empfehlungsalgorithmen von Relevanz, die von „**Informationsintermediären**“ (Schmidt 2019) eingesetzt werden.<sup>2</sup> Es handelt sich hierbei um Plattformen, die zwischen sogenannten Dritten – also etwa zwischen redaktionellen Medien und Privatpersonen oder zwischen Privatperson und Privatperson oder zwischen staatlichen Akteuren und Medien – vermitteln, indem sie die Auswahl der Angebote durch automatisierte Verfahren strukturieren. Im Gegensatz zu klassischen Medienkonzernen stellen Plattformen, die in erster Linie als solche Informationsintermediäre fungieren, selbst keine Medieninhalte her, sondern vermarkten einen flexibel strukturierten, personalisierbaren Kanal (McLuhan 2001/1964, Daub 2020). Zudem bieten sie Interaktionsmöglichkeiten wie etwa Kommentarfunktionen und das Verfassen neuer Inhalte an, wodurch sie die klare Unterscheidung zwischen Sender und Empfänger, die klassische journalistische Publikationen kennzeichnet, tendenziell aufheben.

**Informationsintermediäre** können in Suchmaschinen, Soziale Medien, Foto- oder Video- sowie Microbloggingplattformen und Nachrichtenaggregatoren unterschieden werden. Informationsintermediäre sammeln und verarbeiten auf ihren Plattformen eine große Zahl an Informationen aus unterschiedlichen Quellen und vermitteln diese nicht wie bei Rundfunk und Presse linear an ein Massenpublikum, sondern zeigen individuellen Nutzer/innen eine Auswahl von Meldungen an. Die jeweiligen Anbieter entwickeln und nutzen dazu Algorithmen, um zu entscheiden, welche Meldungen welchen Personen in welcher Reihenfolge ange-

2 „In the US, congressional statutes have played an equally or perhaps even more important role in sheltering surveillance capitalism from scrutiny. The most celebrated of these is a legislative statute known as Section 230 of the Communications Decency Act of 1996, which shields website owners from lawsuits and state prosecution for user-generated content. ‘No provider or user of an interactive computer service,’ the statute reads, ‘shall be treated as the publisher or speaker of any information provided by another information content provider.’ [...] This is the regulatory framework that enables a site such as TripAdvisor to include negative hotel reviews and permits Twitter’s aggressive trolls to roam free without either company being held to the standards of accountability that typically guide news organizations. Section 230 institutionalized the idea that websites are not publishers but rather ‘intermediaries.’“ (Zuboff 2019: 110)

zeigt werden. Im Gegensatz zu den herkömmlichen publizistischen Verfahren werden die dabei genutzten Kriterien nicht journalistisch-redaktionell ausgewählt. Sie zielen primär ab auf die Schaffung von Aufmerksamkeit bei Nutzer:innen sowie die Maximierung der Verweildauer, um so Werbeeinnahmen für die Betreiber zu generieren. (Oertel u.a. 2022: 10)

Die großen Plattformen, die in der Regel im Fokus wissenschaftlicher Untersuchungen stehen, ermitteln ihr Angebot, indem sie das Verhalten der User:innen interpretieren. In die Berechnungen einfließen können dabei die Suchhistorie von User:innen, ihre Verweildauer bei bestimmten Inhalten, die in der Regel „getagged“ und somit kategorisiert sind, der Umgang mit früheren Vorschlägen, die Rückkehrgeschwindigkeit, eigene Veröffentlichungen, Likes, Gruppenzugehörigkeiten, aber auch persönliche Merkmale wie Alter, Geschlecht und Wohnort.

Die **Personalisierung des Angebots** ist die **genuine Leistung von Empfehlungsalgorithmen** und zugleich **das medienhistorisch Neue** an diesen Algorithmen. Während klassische Medien allen dasselbe präsentierten, sehen Facebook, TikTok, Instagram und YouTube für jede:n User:in anders aus. Für die Personalisierung kommen Mechanismen zum Einsatz, die den User:innen **bewusst** sind und durch diese sogar gezielt eingesetzt werden können (etwa Likes), aber auch solche, die **hinter ihrem Rücken** operieren (Schmidt 2016). Sammlung und Analyse von Daten dienen dabei einerseits der Vorhersage von Interessen und Entscheidungen und damit einem genauen Zuschnitt des Angebots, andererseits einem zielgruppenorientierten Einsatz von Werbung. Nicht nur Inhalte, auch User:innen werden daher in Gruppen sortiert, deren Daten wiederum für die Werbeindustrie vermarktbar sind.

Das gezielte und permanent verfeinerte Sammeln von Daten ist in der Forschung unter anderem unter dem Begriff „captology“ diskutiert worden (Fogg/Cuellar/Danielson 2009), wobei es nicht nur um die Analyse des Nutzer:innenverhaltens, sondern auch um die Vorformatierung des Möglichkeitsraums geht, in dem dieses Verhalten stattfindet. Von Programmierer:innen werden Empfehlungsalgorithmen dabei oftmals als Falle konzipiert, die für die User:innen besonders **trickreich** aufgestellt werden müssen (Seaver 2019). Auswertbares Nutzer:innenverhalten muss zudem möglichst einfach strukturiert sein, damit es regelmäßig ausgeführt wird und zu interpretierbaren Verhaltensmustern führt. Die Maxime lautet: „Doing must be easier than thinking.“ (Eyal 2014) Die genaue Vermessung von User:innenverhalten auf Grundlage der gigantischen Datenmengen, die den großen Sozialen Netzwerken zur Verfügung stehen, führt für nicht eingeweihte Laien zu mitunter **erschreckend präzisen Vorhersagen**, wie sie etwa Shoshana Zuboff unter dem Begriff des „Surveillance Capitalism“ beschrieben hat. Mit 70 Likes verrät man Facebook so viel von seiner eigenen Persönlichkeit, wie früher nur die eigenen Freunde wussten, mit 150 Likes kennt einen Facebook so gut wie die eigene Familie.<sup>3</sup>

Empfehlungsalgorithmen treffen keine Entscheidungen für Menschen, suchen aber nach Wegen, Entscheidungen zu prognostizieren und infolge der Herstellung günstiger Bedingungen auch herbeizuführen. Dieser Zusammenhang von **Prognose und Manipulation** steht im Zentrum der ökonomischen Versprechen des Überwachungskapitalismus. In der Terminologie der deutschen Datenethikkommission führen Empfehlungsalgorithmen somit zu „**algorithmengetriebenen** Entscheidungen“, die von Menschen getroffen werden, deren Entscheidungsspielraum durch Algorithmen vorgegeben worden ist. Diese lassen sich von **algorithmenbasierten** Entscheidungen, bei denen die Algorithmen Informationen zu bestimmten, konkret formulierten Fragen liefern, und **algorithmendeterminierten**, also unmittelbar von einer Rechenmaschine getroffenen Entscheidungen abgrenzen (Datenethikkommission der Bundesregierung 2019: 24).

Die Formatierung des Entscheidungsraums erfolgt zumeist durch bestimmte Mechanismen der Präsentation von Empfehlungen, die durch die User:innen beeinflussbar sein können, aber nicht müssen. Zentrale Mechanismen sind das Filtern, also eine Vorauswahl von Inhalten nach bestimmten Kriterien, das Priorisieren, das in der Regel in einer Form des Rankings sichtbar wird, das Klassifizieren von Inhalten durch Tags, Genre-Zuordnungen oder Nutzungskategorien sowie das Assoziieren, also das Hervorheben von Ähnlichkeitsbeziehungen (Oertel u.a. 2022: 42f.). Die meisten Online-Plattformen, die Empfehlungsalgorithmen einsetzen, verbinden unterschiedliche Formen der Ergebnispräsentation miteinander.

### 1.3 Design, Entwicklung und Nutzung von Empfehlungsalgorithmen

Empfehlungsalgorithmen sind in größere **systemische Zusammenhänge** eingebettet und bilden so gut wie nie den einzigen Mechanismus der Präsentation von Inhalten. So können diese in der Regel etwa auch durch Suchfunktionen erschlossen werden; zudem existieren für viele Angebote weiterhin redaktionell kuratierte Zugriffsformen. Empfehlungsalgorithmen greifen einerseits auf unterschiedliche Formen von Daten (mindestens Inhalte und User:innendaten) zu. Andererseits existieren ein **Design**, in dem sie präsentiert werden, ein **Entwicklungskontext**, der die Fortschreibung der Algorithmen organisiert, sowie die Ebene der **Nutzung** selbst, die aus unterschiedlichen Gründen und in unterschiedlichen Kontexten stattfinden kann.

Von besonderer Relevanz für die hier aufgeworfenen Fragestellungen ist die Ebene des **Designs**, da dieses bestimmt, wo und in welcher Form Empfehlungen präsentiert wer-

3 Vgl. Kosinski und Stillwell 2013. Van der Linden fasst zusammen: “Kosinski and Stillwell [...] compared how well people’s friends, family, colleagues, and even their spouse can predict an individual’s personality compared to the predictions of their model (which was [...] based on Facebook likes). What they found was astonishing: the computer algorithm could outperform the accuracy of a work colleague’s prediction of someone’s personality with just ten likes. It needed just 70 likes to outperform the prediction of friends, 150 likes to outperform family, and about 300 likes to be able to outperform a spouse.” (van der Linden 2023: 152f.)

den und wie sie zu anderen, etwa redaktionell kuratierten Angeboten in Relation gesetzt sind. Die Forschung zeigt, dass simple Design-Entscheidungen eine große Wirkung darauf haben können, wie User:innen sich verhalten (Thaler/Sunstein 2021). Der Raum für Variation kann hier kaum überschätzt werden: So bedeutet es einen großen Unterschied, ob ein Mechanismus der Inhaltspräsentation dominant ist (wie etwa die Timeline bei Facebook) oder verschiedene unterschiedliche Listen, Suchfelder und Präsentationsmechanismen angezeigt werden (etwa unterschiedlich begründete Musikvorschläge auf Spotify) oder ob deutlich gemacht wird, nach welchen Kriterien selektiert wird oder nicht.

Jede Präsentation von Ergebnissen kann dabei als – in der Regel in sich dynamische – „**choice architecture**“<sup>4</sup> verstanden werden, die bestimmte Verhaltensformen nahelegt. Grundsätzlich können die Vorschläge die Form einer Endlosliste (wie in der Facebook-Timeline), einer Listenauswahl (unterschiedliche personalisierte Genre-Listen auf Netflix, die ihrerseits personalisierte Angebote enthalten) ohne transparente Kriterien oder einer Listenauswahl mit konkreten Kriterien annehmen („Nutzer, die dieses Produkt mochten, suchten auch nach...“, „Deinen Freunden gefällt dieser Artikel“, „Deine Folk-Hits der Woche“). Auf vielen Plattformen wird auch redaktionell kuratiert, was den Anschein hat, automatisch generiert worden zu sein (und vice versa): 2016 wurde beispielsweise nachgewiesen, dass Facebooks ‚Trending Topics‘ nicht algorithmisch erzeugt, sondern von Mitarbeiter:innen ausgewählt wurden. Nach der Umstellung auf eine rein algorithmisch gesteuerte Auswahl wurden dann Falschnachrichten und Hassbotschaften als ‚trending‘ angezeigt (Thielman 2016).

Auch der **Entwicklungskontext** ist Gegenstand der Forschung zu Empfehlungsalgorithmen, handelt es sich bei diesen doch in der Regel um hochgradig variable Algorithmen, die sich ständig ändern. Erneut kommt es hier zu einem Zusammenspiel zwischen menschlichen Akteur:innen und automatisierten Veränderungen etwa in sich kontinuierlich ausdifferenzierenden neuronalen Netzwerken, die auf diese Weise auch für die Betreiber:innen selbst zu immer komplexeren ‚black boxes‘ werden (Muhle/Wehner 2016). Auf die gezielte **Weiterentwicklung der Algorithmen** wirken, wie eine die Herstellung eines Empfehlungsalgoritmus für den Bayerischen Rundfunk begleitende Studie festgestellt hat, vielfältige Akteure ein, was die Fortentwicklung zu einem politischen Balanceakt werden lässt: Unterschiedliche Interessengruppen, Abteilungen eines Konzerns oder Redaktionen eines Medienunternehmens, aber auch User:innen-Feedback und neue gesetzliche Bestimmungen müssen in konkrete Programmierungen übersetzt werden. Anders, als es vereinfachende Beschreibungen suggerieren, ist die Optimierung von Empfehlungsalgorithmen also kein rein technischer Prozess, sondern es handelt sich um eine rege Interaktion zwischen unterschiedlichen Akteursgruppen, die mit Eigendynamiken beziehungsweise technischen Grenzen des algorithmischen Systems koordiniert werden muss (Pöchhacker u.a. 2017).

<sup>4</sup> Der Begriff geht auf Cass R. Sunstein und Richard H. Thaler zurück (Sunstein/Thaler 2018).

Der **Nutzungskontext** wird vor allem unter dem Stichwort der **Moderation** diskutiert: Große Netzwerke wie X werden selbst von menschlichen und nicht-menschlichen, teilweise algorithmischen Akteuren genutzt (Bots), die ihrerseits von menschlichen wie nicht-menschlichen Moderator:innen überwacht werden können. Ideologisch aufgela- dene Debatten zur ‚Freiheit des Internets‘ konkurrieren hier mit dem Anspruch, zumindest im Ansatz diskriminierungsfreie virtuelle Räume zu schaffen, in denen strafrechtlich relevantes Verhalten verfolgbar ist.

Die Frage der Moderation von Kommunikation im Netz ist eine an die Frage nach den Empfehlungsalgorithmen angelagerte Problematik, da Entscheidungen in diesem Feld auf die konkrete Ausrichtung von Empfehlungsalgorithmen zurückwirken können. So **verändern** etwa Bots die Datenbasis, auf deren Grundlage Inhalte empfohlen werden: Algorithmische Systeme können durch andere algorithmische Systeme **manipuliert** werden („false amplification“ durch retweetende Bots), was auch für die Anbieter der Empfehlungsalgorithmen ein Problem darstellen kann. **Kuratierung und menschliche Beobachtung** von Empfehlungsalgorithmen und ihren Folgen erweisen sich so als Notwendigkeit, auch wenn erst in Ansätzen eine politische Regulierung besteht.

#### **1.4 Die Debatte um Polarisierung und Radikalisierung durch Empfehlungsalgorithmen**

Worin die gesellschaftlichen Gefahren von Empfehlungsalgorithmen bestehen, lässt sich anhand eines anhaltenden Rechtsstreits illustrieren. Im August 2018 gestand der Meta-Konzern in einer Pressemeldung offiziell eine Teilschuld am **Genozid an den Rohingya in Myanmar** ein (vgl. McPherson 2018). Der Hass, der im Sommer 2017 zu zahllosen Toten und mehr als 800.000 geflüchteten Menschen geführt hatte, war, so die Mitteilung des Konzerns, wesentlich von Facebook-Vorschlagsalgorithmen verbreitet worden. In Facebook-Gruppen entstand zudem eine soziale Dynamik, die zu immer stärkerer Radikalisierung, zur Erfüllung und Verbreitung von Falschinformationen und Verschwörungstheorien und schließlich zum **Aufruf zur Gewalt** führte. Aufgrund der Sprachbarriere, die zwischen Entscheidungsträger:innen in den USA und Akteuren in Myanmar bestand, sowie einer ganzen Reihe anderer Faktoren war Facebook nicht in der Lage, diese Dynamik rechtzeitig zu erkennen und zu moderieren – obgleich Nutzer:innen in Myanmar auf die bevorstehende Katastrophe aufmerksam gemacht hatten. Im Jahr 2021 verklagten dann offizielle Vertreter:innen der Opfer des Völkermordes den Meta-Konzern: Gefordert wird ein Schadenersatz von 150 Milliarden Dollar; die Verhandlungen dauern bis heute an (vgl. Milmo 2021).

Der Fall ist beispielhaft für die Art der Vorwürfe, die seit einigen Jahren im Allgemeinen gegenüber den sogenannten Informationsintermediären sowie im Besonderen gegenüber Empfehlungsalgorithmen vorgebracht werden. Der New-York-Times-Reporter

Max Fisher (2023) hat diese Vorwürfe in einem kürzlich erschienenen Buch pointiert zusammengefasst. Durch Empfehlungsalgorithmen, so wird von Fisher und anderen Journalist:innen und Aktivist:innen nahegelegt, wurden jene **Verschwörungstheorien** und **Falschnachrichten** verbreitet, die 2021 zum Sturm auf das Kapitol führten, wurde die Wahl Trumps ebenso ermöglicht wie der Erfolg rechtspopulistischer Parteien und Strömungen in Europa, werden die **Fragmentarisierung** (Napoli 2014) und politische **Polarisierung** in der Gesellschaft massiv verschärft (vgl. Cinus u.a. 2022), werden Desinformation etwa zur Klimapolitik (Dündar/Ranaivoson 2022) verbreitet, Rassismus, Sexismus und sogar Kindesmissbrauch (Eordogh 2018) unterstützt. Als womöglich prominentester Kritiker von Social Media brachte US-Präsident Joe Biden seine Skepsis gegenüber dem gesellschaftlichen Wert von Empfehlungsalgorithmen in einem scharfen Urteil auf den Punkt: „They're killing people.“ (Egan 2021)

Empfehlungsalgorithmen sind ihren Kritiker:innen zufolge zwar nicht der einzige, aber ein wesentlicher Motor dieser gefährlichen Dynamiken. Zur Beschreibung der Probleme, die sie hervorbringen, kursiert inzwischen eine ganze Reihe von Begriffen. So erzeugen Algorithmen dem Aktivisten Eli Pariser (2014) zufolge sogenannte Filter Bubbles, also ausschnithafte oder verzerrte Darstellungen der Wirklichkeit. Die **Suchtmechanismen**, die mit Social Media in Verbindung gebracht werden (Hou u.a. 2019) und die durch Empfehlungsalgorithmen befördert werden können, führen Nutzer zudem in sogenannte ‚Rabbit holes‘ (Lewis 2018), also in eine destruktive Fixierung auf bestimmte emotionalisierende Themen, die mit einer **obsessiven Nutzung sozialer Netzwerke** einhergeht. So entstünden immer mehr politische Randgruppen, was zu einer Zersplitterung der Gesellschaft führe, das heißt zu einer sogenannten „Enclave deliberation“ (Sunstein 2001). Immer größere Teile der Gesellschaft ließen sich auf diese Weise durch ‚alternative Fakten‘ beeinflussen, die ihnen ermöglichen, drängende gesellschaftliche Probleme und die eigene Lebensweise betreffende Forderungen zu ignorieren (Kumkar 2022). Die Social-Media-Plattformen bekämpften diese Dynamiken nicht, da ihr Gewinn nicht auf Offentlichkeitsbildung, sondern auf einer anhaltenden Maximierung der Nutzungszeit basiere, die sich unmittelbar in Werbeeinnahmen umrechnen lasse. Nicht Inhalte seien den Plattformen wichtig, so eine der Generalthesen Fishers (Fisher 2023), sondern allein die möglichst lange und damit geldwerte Beschäftigung mit egal welchen Angeboten.

In einer Aufmerksamkeitsökonomie, in der Konzerne um die Zeit konkurrieren, die User:innen auf ihren Seiten oder in ihren Apps verbringen, sind dabei Eli Pariser zufolge Emotionalisierungsstrategien zentral. Da **Wut und Hass** zu deutlich stärkerem Engagement führten als andere Emotionen, würden diese durch die Algorithmen der Plattformen gezielt gefördert: So ergibt eine Studie, dass Tweets, die als Hasskommentare kategorisiert werden können, weitaus häufiger retweetet werden als andere Tweets (vgl. Ribeiro u.a. 2018). Ebenso werden negative Facebook-Beiträge häufiger kommentiert als positive (Stieglitz/Dang-Xuan 2012). Unwahrheiten verbreiten sich über Social Media etwa sechsmal so schnell wie Wahrheiten (van der Linden 2023: 96). Die Befeuierung solcher Dynamiken durch Empfehlungsalgorithmen basiert dabei nicht notwendig

auf Entscheidungen menschlicher Akteure, also etwa des Meta- oder X-Managements, sondern resultiert aus der automatisierten Fortentwicklung der Algorithmen, die durch Künstliche Intelligenz vorangetrieben wird: Wenn der Algorithmus die durchschnittliche Verweildauer der User:innen auf einer Plattform um einige Minuten erhöhen soll, gelingt ihm das durch die Manipulation zahlloser Parameter, die selbst die Programmierer:innen des Systems häufig nicht mehr offenlegen können.

Empfehlungsalgorithmen arbeiten nicht nur in Sozialen Netzwerken und auf Video-Plattformen, sondern in verschiedenen anderen Bereichen des World Wide Web, etwa bei Ergänzungsvorschlägen von Suchmaschinen. Auch hier ist von prominenten Kritiker:innen argumentiert worden, dass scheinbar kleine Hilfestellungen große Wirkungen haben können. Die Art und Weise etwa, wie der Satz ‚Hillary Clinton is...‘ im Jahr 2016 von Suchmaschinen vervollständigt wurde, hatte dem Psychologen Robert Epstein zu folge unmittelbar Einfluss auf das Wahlergebnis. Epstein spricht schon 2015 von einem „**Search Engine Manipulation Effect**“ (SEME) (Epstein 2015), dem er einen großen Einfluss auf das Wahl- und Konsumverhalten der Menschen zutraut. Epsteins Vorwurf, Google habe durch Search Engine Manipulation Effects direkt die Wahlen manipuliert, wurde jedoch vom Konzern selbst sowie von unterschiedlichen Journalist:innen und Forscher:innen dementiert (Singhal 2015).

Tatsächlich hat sich erwiesen, dass der **Generalverdacht gegenüber Empfehlungsalgorithmen**, der von Fisher, Epstein und anderen formuliert wird, sich bei genauerer Untersuchung als zu pauschal oder sogar als **haltlos** erweist. Häufig scheinen die Vorwürfe aus der hier umrissenen Debatte, die in der Regel von Journalist:innen, Insider:innen aus der Industrie, Politiker:innen und Prominenten formuliert werden, eine gewisse Evidenz zu haben. Doch diese Vorwürfe dürfen nicht ungeprüft für wahr gehalten und übernommen werden. Mitunter handelt es sich um urbane Mythen, an die immer und zuallererst die Frage zu richten ist, welche Risiken und Gefahren, aber auch Potenziale von Empfehlungsalgorithmen durch empirische Forschung tatsächlich nachweisbar sind. Ein solider wissenschaftlicher Ansatz ist unabdingbar, um nicht auf der Basis von Allgemeinwissen oder populären Annahmen zu argumentieren, sondern fundiert und evidenzbasiert zu urteilen.

## 1.5 Empirische Befunde

Obgleich die Zahl der Studien, die sich mit den Effekten von Empfehlungsalgorithmen und ihrem Einfluss auf Öffentlichkeit und Gesellschaft beschäftigt, inzwischen ins Unüberschaubare gewachsen ist, ergibt sich ein mit Blick auf die scharfen Vorwürfe gegen Empfehlungsalgorithmen erstaunlich **unklarer Befund**. Bisher existieren **keine Beweise dafür, dass Empfehlungsalgorithmen für politische Radikalisierung und Hass maßgeblich verantwortlich** sein könnten (Arguedas u.a. 2019). Auch Filter Bubbles sind empirisch bisher nur unter bestimmten, außergewöhnlichen Bedingungen

nachgewiesen worden. Empirisch betrachtet hat es den Anschein, dass Empfehlungsalgorithmen das Spektrum der Nachrichtenquellen, die die User:innen präsentiert bekommen, tendenziell sogar erweitern. Genauer gesagt liegt also vermutlich der gegenwärtige **Effekt einer größeren** – und nicht, wie angenommen, geringeren – **Diversität** an Informationen bei Nutzung von Empfehlungsalgorithmen vor (Arguedas u.a. 2019, Scharkowa u.a. 2020).

Eine Ausnahme scheint lediglich YouTube darzustellen: Hier konnte so etwas wie ein politischer Filterblaseneffekt experimentell nachvollzogen werden (Litschka 2021)<sup>5</sup>. So oder so ändern politisch radikalierte Nutzer:innen von Sozialen Medien ihr Weltbild nicht, wenn sie auf die Nutzung dieser Medien verzichten (Allcott u.a. 2020). Studien, die den gesellschaftlichen Einfluss von Empfehlungsalgorithmen untersuchen, existieren vor allem für den amerikanischen, in geringerer Zahl aber auch für den europäischen Raum. So hat zum Beispiel eine Untersuchung zur deutschen Bundestagswahl 2017 ergeben, dass Menschen mit sehr unterschiedlichen politischen Profilen bei einer Google-Recherche im Internet nach Informationen zu Politiker:innen keine wesentlich voneinander abweichenden Ergebnislisten erhielten (Krafft u.a. 2018).

Das bedeutet allerdings nicht, dass das Web 2.0 – also das von User:innen mitgestaltete Internet – keinen Beitrag zur Polarisierung der Gesellschaft leistet. Dass Falschnachrichten, Verschwörungstheorien, politische Feindbildungen und Hass auf Sozialen Netzwerken Verbreitung finden, lässt sich leicht überprüfen. Der Zusammenhang zwischen Empfehlungsalgorithmen und solchen gesellschaftlichen Phänomenen bedarf jedoch einer genauen Analyse.

### *1.5.1 Eigendynamik versus Manipulation*

Bei genauer Betrachtung stellt sich so etwa der Vorzeigefall der Empfehlungsalgorithmen-Kritik, die Facebook-Nutzung in Myanmar, als komplexer heraus, als es auf den ersten Blick nach der Lektüre von Fishers *Chaos Machine* erscheinen mag (vgl. Whitten-Woodring 2020). Denn Myanmar ist schon deshalb ein Sonderfall, weil hier die Bevölkerung in einem im Vergleich zu anderen Ländern viel höheren Ausmaß das Internet durch Facebook hindurch nutzt. Die Facebook-Gruppen, in denen die Radikalisierung in erster Linie stattfand, wurden dabei von staatlichen und militärischen Akteuren beeinflusst, wie auch insgesamt die Radikalisierung eher auf einen kontrollierten Einsatz von Facebook durch offizielle Stellen zurückzugehen scheint als auf Facebook-Eigendynamiken.

Solche Erkenntnisse legen es nahe, den Vorwurf gegen die Empfehlungsalgorithmen dahingehend zu reformulieren, dass Empfehlungsalgorithmen nicht die Ursachen, son-

5 Vgl. darüber hinaus: O'Callaghan, Derek u.a.: Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems. Social Science Computer Review, 33 (2015), H. 4, S. 459–478; online unter: <https://doi.org/10.1177/0894439314555329> [Zugriff 24.08.2023].

dern eher gute „Voraussetzungen für großflächige Desinformations- und Manipulationskampagnen“ (Oertel u.a. 2022: 101; vgl. auch Gu u.a. 2017) schaffen, also ausgenutzt werden können. Es geht daher weniger um die **Eigendynamiken** von Empfehlungsalgorithmen als vielmehr um ihre **gezielte Manipulation**. Tatsächlich werden Desinformationskampagnen inzwischen nachweislich auf die Funktionsweisen von Informationsintermediären ausgerichtet (Shao u.a. 2018). Zu den Strategien gehört zum Beispiel die Initierung von „**informational cascades**“ (Casemayor/Rocheleau 2019) durch geschickt platzierte Falschinformationen, die sich dann um ein Vielfaches schneller verbreiten als ihre Richtigstellungen (Vosoughi u.a. 2018). Die Beeinflussung von Wahlen durch gezielte russische Medienkampagnen ist ein Beispiel für den demokratiegefährdenden Missbrauch der Mechanismen von Social Media und Empfehlungsalgorithmen durch bestimmte Akteursgruppen (DiResta u.a. 2019) – auch in Deutschland (Applebaum u.a. 2017).

Im Zentrum solcher Kampagnen stehen in der Regel **Strategien der Desinformation**. Dabei handelt es sich im digitalen Kontext um eine vor allem von der politischen Rechten eingesetzte Form der Manipulation, die es einer kleinen Zahl von Akteuren durch den Einsatz rhetorischer, visueller und algorithmischer Strategien erlaubt, eine große Reichweite und dadurch einen großen Effekt zu erzielen (Narayanan u.a. 2018, Neudert 2017). Ebenso gehen Radikalisierungseffekte, wo sie sich durch die empirische Forschung feststellen lassen, in der Regel auf die Aktivitäten und die innere Vernetzung kleiner Einflussgruppen zurück: „Online-Debatten werden häufig angetrieben von relativ kleinen Minderheiten ungewöhnlich aktiver und meinungsfreudiger Personen, während die große Mehrheit der Internetnutzer:innen diesen parteiischen Diskussionen wenig oder gar keine Aufmerksamkeit schenkt.“<sup>6</sup> Das ist gefährlich: „Many influential elections are decided on narrow margins.“ (van der Linden 2023: 59).

Diese Befunde deuten darauf hin, dass die **Vorbehalte**, die gegenüber **Empfehlungsalgorithmen** formuliert werden, nicht generell falsch, aber stark **differenzierungsbedürftig** sind. Soziale Netzwerke und Empfehlungsalgorithmen setzen **aus sich selbst heraus**, d.h. basierend allein auf ihrer Funktionslogik, **keine** generelle **Polarisierungs- und Radikalisierungsdynamik** in Gang, **bieten aber** interessierten Gruppen **eine Bühne**, solche Debatten schneller und schlagkräftiger zu entfesseln als in der analogen Vergangenheit.

### *1.5.2 Ideologische versus affektive Polarisierung*

Studien deuten darauf hin, dass zwei verschiedene Formen von gesellschaftlicher Polarisierung durch Empfehlungsalgorithmen zu beobachten sind. Zum einen zeigt sich eine verstärkte **ideologische Polarisierung** durch Empfehlungsalgorithmen, also ein Aus-

6 „Relatively small minorities of unusually active and opinionated individuals often animate online debates, even as the vast majority of internet users pay little or no attention to these partisan discussions.“ (Arguedas 2022: o. S.).

einanderdriften von Meinungen und Weltbildern. Konkrete Beispiele für diese Polarisierung sind etwa die US-Präsidentswahlen 2016: Untersuchungen haben ergeben, dass bestimmte Empfehlungsalgorithmen in Sozialen Medien dazu beitrugen, dass Nutzer vermehrt extremere politische Inhalte sahen. Dies führte zu einem verstärkten Echokammer-Effekt, bei dem Menschen überwiegend Informationen aus ihrer eigenen politischen Blase erhielten und somit bestehende Meinungen weiter verfestigt wurden (vgl. Alcott/Gentzkow 2017). Ein weiteres prominentes Beispiel ist die Debatte um die Impfpflicht im Zuge der Corona-Pandemie. Auch hier konnte in verschiedenen Studien nachgewiesen werden, dass Nutzer:innen, die einmal nach Impfgegnerschaftsinhalten suchten, immer wieder mit ähnlichen Inhalten konfrontiert wurden, was dazu führte, dass sich diese Nutzer:innen eher von konträren wissenschaftlichen Fakten abschotteten (Broniatowski 2018).

Dieses Phänomen des Auseinanderdriftens von Meinungen ist tatsächlich beobachtbar; allerdings sind solche Fälle in der Regel eher die Ausnahme. Wichtig ist in diesem Zusammenhang, dass eine derartige ideologische Polarisierung, sofern sie beobachtet wird, nicht immer direkt auf den Einfluss von Intermediären und Sozialen Medien zurückzuführen ist. Es gibt zahlreiche Faktoren, darunter wirtschaftliche, soziale und politische Entwicklungen, die in diese komplexe Dynamik eingreifen. Allerdings **erhöhen Empfehlungsalgorithmen ohne Frage die Geschwindigkeit und Intensität, mit der (Fehl-)Informationen** (und damit auch bestimmte Weltanschauungen) **verbreitet werden** (Iyengar u.a. 2018).

Anders verhält es sich mit der **affektiven Polarisierung**, das heißt dem mit unterschiedlichen Positionen verbundenen emotionalen Engagement,<sup>7</sup> und – damit zusammenhängend – mit der in der Debatte **wahrgenommenen Polarisierung** der Gesellschaft: Affektive und wahrgenommene Polarisierung steigen einigen Studien zufolge tatsächlich messbar an; allerdings sind an diesem Prozess auch ‚klassische‘ Medien beteiligt, in den USA etwa Fox News (Arceneaux/Johson 2014).

Ohne die Psychologie der User:innen zu berücksichtigen, lassen sich die Filterung von Medieninhalten und die Empfehlungen durch Empfehlungsalgorithmen und damit auch der Einfluss von Informationsintermediären auf die Öffentlichkeit nicht verstehen. Auch aus diesem Grund schlägt das Büro für Technikfolgenabschätzung am Deutschen Bundestag in einem Gutachten zu Empfehlungsalgorithmen vor, **systematisch zwischen individualpsychologischer und gesellschaftlicher Dimension der medialen Beeinflussung zu unterscheiden**. Das ist schon deshalb sinnvoll, weil Nachweise für einen ‚gesamtgesellschaftlichen‘ Einfluss von Empfehlungsalgorithmen schwer zu erbringen sind, sehr wohl aber in der Forschung generell Konsens darüber besteht, dass auf individueller Ebene ein Einfluss durch Empfehlungsalgorithmen nachweisbar ist. Dieser Einfluss lässt sich auf den Ebenen der Agendasetzung – also der Auswahl

7 „Overall, ideological polarisation has, in the long run, declined in many countries, but affective polarisation has in some cases, but not all, increased.“ (Arguedas u.a. 2019: 23)

der Themen, die Medien vornehmen –, des individuellen Wissenserwerbs und der Meinungsbildung nachvollziehen (Oertel u.a. 2022: 87f., Geiß 2015).

Um den Forschungsstand zusammenzufassen: Empfehlungsalgorithmen führen in der Regel beim Angebot von medialen Inhalten nicht zu starken Abweichungen im Vergleich mit anderen Medien, die keine Empfehlungsalgorithmen einsetzen; allerdings können Nischenthemen ein höheres Gewicht erhalten (Moeller u.a. 2016). Der anscheinend größte, wenn auch bislang empirisch nur in Ansätzen belegte Effekt zeigt sich auf der Ebene der Meinungsbildung: Nutzer fühlen sich – und das wird oftmals mit dem Filterblasen- und Echokammereffekt in Übereinstimmung gebracht – „mit ihrer **Meinung** signifikant stärker **der Mehrheit zugehörig**, wenn **Facebook als Informationsquelle** sehr oder eher wichtig war.“ (Stark u.a. 2017: 150) Diese **Fehleinschätzung** bedingt dann – zumindest theoretisch – eine höhere Intoleranz gegenüber anderen Meinungen.

Dieser Befund gilt bis zu einem gewissen Grad auch für die öffentliche Debatte über die Empfehlungsalgorithmen selbst. Während die empirische Forschung zum tatsächlichen Einfluss dieser Systeme unterschiedliche, auch uneindeutige und in jedem Fall anspruchsvoll zu interpretierende Ergebnisse zutage fördert, werden – insbesondere in **Online-Debatten** – oftmals diejenigen Stimmen am meisten wahrgenommen, die ihren „algorithmic hate“ (Smith/Jayne/Burke 2022) besonders scharf artikulieren. Wenn jedoch gängige Erklärungen wie die Filterblasen-These nicht haltbar sind, laufen auch die aus solcher Kritik abgeleiteten Vorschläge ins Leere. Ein Beispiel hierfür ist der Vorschlag, Desinformation mit Strategien der Aufklärung zu begegnen.<sup>8</sup>

## 1.6 Die zentrale Bedeutung des Vertrauens

Tatsächlich ist es der aktuellen Forschung zufolge so gut wie aussichtslos, im Netz Mythen durch Fakten zu widerlegen. Insbesondere der aufklärerische Gestus einer ‚**Facts-versus-myths**‘-Strategie führt zu einer Reihe von individualpsychologisch gut erforschten Abwehrreflexen. Auch Strategien des **Counter-Messagings**, das gegen extremistische Inhalte eingesetzt wird, gelten als problematisch und können sich durch Empfehlungsalgorithmen sogar **kontraproduktiv** auswirken (Schmitt u.a. 2018). Vor allem die unterschiedliche Ausbreitungsgeschwindigkeit von Mythen und Fake News (hoch) und Fakten sowie Richtigstellungen (niedrig) spricht dagegen, dass sich im Netz Mythen korrigierend einholen und damit abschaffen lassen (Vosoughi/Roy/Aral 2018).

8 Dazu insgesamt van der Linden 2023; zur Frage von Korrekturen: „we find that corrections do help reduce misinformation but they do not fully eliminate it: people continue to retrieve false details from memory once exposed to a falsehood. The effect appears to get more pronounced the more misinformation is repeated and the longer it is held in memory without correction.“ (68)

### *1.6.1 Psychologische Abwehrmechanismen gegen die Widerlegung falscher Überzeugungen*

Für die Abwehrmechanismen gegen das Prinzip ‚Facts versus myths‘ gibt es eine Reihe von Begriffen, die sich zum Teil in ihrer Bedeutung überschneiden. Dazu gehören der **Confirmation bias** und der **Disconfirmation bias**. Der ‚Confirmation bias‘ ist ein ‚Bestätigungsfehler‘ beziehungsweise eine ‚Bestätigungsstendenz‘: Menschen neigen dazu, neue Informationen so auszuwählen oder zu interpretieren, dass sie ihre schon bestehenden Ansichten bestätigen. In ähnlicher Weise neigen die Menschen dazu, zu ihrem schon gefassten Weltbild passende Informationen für glaubwürdiger oder überzeugender zu halten als solche, die diesem Weltbild widersprechen. Das führt zu einer ‚Widerspruchstendenz‘ (,Disconfirmation bias‘). Wenn Menschen mit Informationen konfrontiert werden, die ihren Überzeugungen widersprechen, suchen sie Gegenargumente.<sup>9</sup>

Häufig wird dieses Phänomen auch **Social reinforcement** (Verstärkung durch das soziale Umfeld) oder **Motivated reasoning** („zielgerichtetes Denken“) genannt; es führt zum **Backfiring**, zur ‚Fehlzündung‘, des auf die Korrektur von Irrtümern ziellenden **Counterspeech** und somit in der Folge nicht zur intendierten Schwächung, sondern gerade zur unabsichtlichen Stärkung falscher Überzeugungen. Anders gesagt: Fehlinformationen durch Disclaimer oder argumentative Widerlegung zu berichtigen, führt selten dazu, dass Menschen falsche Überzeugungen revidieren.<sup>10</sup>

Dieser ‚Worldview backfire effect‘ heißt auch ‚Boomerang effect‘ (Wittenberg/Berinsky 2020). Eine andere Bezeichnung für dieses Phänomen ist der Effekt der ‚Evidential pre-emption‘, des ‚Dem-Beweis-Zuvorkommens‘. Was dabei geschieht, ist dem Philosophen Endre Begby zufolge eine Art „intellektuelles Judo, bei dem die Überzeugungskraft und der gute Wille von Menschen, die der eigenen Position widersprechen, durch eine von inneren Überzeugungen sorgfältig trainierte Abwehrtechnik gegen die widerprechenden Menschen zurückgewendet wird.“<sup>11</sup>

Auch der sogenannte **Continued influence effect** („fortwirkender [Fehlinformations-]Effekt“) beschreibt, dass Menschen gerade nach der Korrektur eines Irrglaubens umso stärker von diesem Irrglauben überzeugt sein können. Denn selbst wenn sich

9 „[C]onfirmation bias: Individuals tend to seek out and interpret new information in ways that validate their preexisting views. Along these lines, individuals also tend to perceive congenial information as more credible or persuasive than opposing evidence (Guess/Coppock 2018; Khanna/Sood 2018). [...] [D]isconfirmation bias: When exposed to ideologically dissonant information, individuals will call to mind opposing arguments (counterarguing).“ (Wittenberg/Berinsky 2020: 170).

10 „Merely correcting misinformation with disclaimers or counter-speech rarely erases the false belief. Because of motivated reasoning and other factors, countering misinformation may backfire for some people and even reinforce false beliefs.“ (Persily/Tucker 2020: 5)

11 „The philosopher Endre Begby calls this effect ‘evidential pre-emption’. What’s happening is a kind of intellectual judo, in which the power and enthusiasm of contrary voices are turned against those contrary voices through a carefully rigged internal structure of belief.“ (Nguyen 2018: o. S.)

Menschen an eine Richtigstellung erinnern, kann die ursprüngliche Fehlinformation weiterhin ihre Einstellungen und Überzeugungen beeinflussen.

### *1.6.2 Echokammern als Networks of trust*

Die hier benannten psychologischen Mechanismen sind insbesondere in Strukturen am Werk, die üblicherweise als ‚Echokammern‘ bezeichnet werden. Von ‚Echokammern‘ spricht auch der Philosoph C. Thi Nguyen; dabei fasst er die Begriffe ‚Filterblase‘ und ‚Echokammer‘ theoretisch schärfer als viele andere Forscher:innen (auch wenn er sie nicht empirisch überprüft) (vgl. Nguyen 2018). So werden Filterblasen zumeist als in erster Linie technisches Problem begriffen, wogegen der Begriff der Echokammer auf das Zusammenspiel von Technologie und sozialen Mechanismen abzielt.

Wenn man von einer Filterblase oder, wie Nguyen, von einer epistemischen Blase spricht, bezieht man sich auf Informationsnetzwerke, in denen bestimmte Meinungen oder Perspektiven schlichtweg fehlen, weil sie unbeachtet bleiben. Echokammern hingegen sind **informelle soziale Strukturen, innerhalb derer alternative Perspektiven nicht nur ignoriert, sondern aktiv diskreditiert werden**. In einer Echokammer werden die Meinungen und Informationen außerhalb dieser Struktur durchaus wahrgenommen, aber als unglaublich oder verdächtig betrachtet. Generell gilt daher, dass die Verengung von Weltbildern nicht notwendig mit einer Verengung des Medien- oder Nachrichtenkonsums korreliert. Im Gegenteil kann die Beobachtung des anderen Lagers den Echokammer-Effekt gerade verstärken (Arceneaux/Johnson 2014). Während innerhalb des eigenen Lagers Zustimmung demonstriert wird, dominiert dann nach außen die Wut-Kommunikation (Wollebæk u.a. 2019). Sind Filterblasen in der Forschung umstritten, werden solche Wut-Netzwerke, wie Studien etwa zu rechtsradikalen Bewegungen in Deutschland und den USA zeigen, durch Empfehlungssysteme eindeutig befördert (Kaiser/Rauchfleisch 2017).

Es ist wichtig zu betonen, dass Mitglieder einer Echokammer nicht notwendigerweise **desinteressiert an der Wahrheit** sind. Die Echokammer beeinflusst lediglich, welche Quellen und Institutionen als glaubwürdig angesehen werden. Individuen innerhalb einer Echokammer sind nicht per se irrational, aber sie sind systematisch fehlinformiert darüber, welchen Quellen sie Glauben schenken sollten. Die Struktur und Dynamik einer Echokammer erinnern in vielerlei Hinsicht an die einer Sekte. Während Filterblasen durch relative Instabilität charakterisiert sind, sind Echokammern weitaus widerstandsfähiger und gefährlicher und ähneln in ihrer Komplexität mitunter lebendigen Organismen.

**Epistemische Blasen können** Nguyen zufolge mit den im folgenden Kapitel dieses Whitepapers vorgeschlagenen Maßnahmen von Content Diversity, Serendipity und Cross-cutting exposure **entschärft werden**; bei **Echokammern** funktioniert das aus den genannten psychologischen Gründen **nicht**. Nguyen zieht Schlussfolgerungen, wie

sich im Netz Mythen und falsche Überzeugungen bekämpfen lassen, ohne sie ungewollt durch Backfire-Effekte zu verstärken. Sein zentraler Begriff ist dabei das **Vertrauen**. Echokammern lassen sich Nguyen zufolge nicht dadurch sprengen, dass man ihren Bewohner:innen „die Fakten“ vor Augen hält. Stattdessen geht es darum, verlorenes Vertrauen für eine Position außerhalb der Echokammer wiederherzustellen. Am effektivsten lassen sich Fehlinformationen durch Quellen korrigieren, denen die Fehlgläubigen vertrauen, weil sie ihr Weltbild im Großen und Ganzen teilen und bestätigen, die ihnen also relativ nahestehen.<sup>12</sup> Dies geschieht durch „besondere Begegnungen“ mit „vertrauenswürdigen Außenseitern“ (Nguyen 2018):

Geschichten von gelungenen Ausbrüchen aus Echokammern drehen sich oft um besondere Begegnungen – Momente, in denen das in der Echokammer gefangene Individuum beginnt, jemandem von außen zu vertrauen. Diese Begegnungen sind wichtig, weil eine persönliche Verbindung mit einem erheblichen Vertrauensvorschuss einhergeht.<sup>13</sup>

Radikalisierung und Polarisierung im Netz, so wird hier deutlich, lassen sich in der Regel nicht allein durch den Einsatz von Algorithmen bekämpfen. Doch auch der Einsatz von vertrauensbasierten Strategien birgt keine Garantien: Nguyen selbst spricht illusionslos von einer nur „hauchdünnen Chance auf Rettung von außen“.<sup>14</sup>

## 1.7 Die Herausforderungen digitaler Öffentlichkeit

An der Diskussion über den gesellschaftlichen Einfluss von Empfehlungsalgorithmen fällt auf, dass sie in der Regel extreme Phänomene fokussiert: Filterblasen und Echokammern interessieren dann, wenn sie Extremismus, Polarisierung, Fake News oder Gewalt fördern. Empfehlungsalgorithmen tragen zur Strukturierung der Öffentlichkeit allerdings auch dann bei, wenn solche Symptome nicht auftreten: Sie beeinflussen auch die „gemäßigte“ Öffentlichkeit, indem sie Themen auswählen, Wissen vermitteln und

12 „The way to break an echo chamber is not to wave ‘the facts’ in the faces of its members. It is to attack the echo chamber at its root and repair that broken trust. [...] These most effective responses to misinformation require corrections from a source the believers trust, delivered in a way that affirms their worldview.“ (Nguyen 2018)

13 „Stories of actual escapes from echo chambers often turn on particular encounters – moments when the echo-chambered individual starts to trust somebody on the outside. [...] These encounters matter because a personal connection comes with a substantial store of trust. [...] Why is trust so important? Baier suggests one key facet: trust is unified. We don’t simply trust people as educated experts in a field – we rely on their goodwill. And this is why trust, rather than mere reliability, is the key concept. Reliability can be domain-specific. The fact, for example, that somebody is a reliable mechanic sheds no light on whether or not their political or economic beliefs are worth anything. But goodwill is a general feature of a person’s character. If I demonstrate goodwill in action, then you have some reason to think that I also have goodwill in matters of thought and knowledge. So if one can demonstrate goodwill to an echo-chambered member – as Stevenson did with Black – then perhaps one can start to pierce that echo chamber.“ (Nguyen 2018).

14 „[A] whisper-thin hope for rescue from the outside“ (Nguyen 2018).

die Meinungsbildung unterstützen. **Verantwortung für die Gesellschaft** besteht daher auch dann, wenn ihnen kein Fehlverhalten nachweisbar ist (Gillespie 2015, Susskind 2022). Studien zur Internet- und Social-Media-Nutzung machen dabei deutlich, welchen gewaltigen Einfluss Empfehlungsalgorithmen haben: Es gibt **mehr aktive Facebook-Nutzer als Christen** auf der Welt, mehr als 70 Prozent der Inhalte von Plattformen wie YouTube werden durch Empfehlungsalgorithmen vermittelt angeklickt. So wenig sichtbar Empfehlungsalgorithmen im Allgemeinen sind: sie bilden doch eine zentrale „Infrastruktur der Öffentlichkeit“ (van Dijck u.a. 2018, Zuckerman 2020).

Als problematisch erscheint dabei, dass diese öffentlichkeitsstrukturierenden Empfehlungsalgorithmen heute wesentlich von einer Handvoll gigantischer Digitalplattformen programmiert und gesteuert werden (Srnicek 2016). Das ist anders als in der Frühzeit des Internets, in der Kommunikation sehr dezentral stattfand. Heute hingegen ist sie in der Hand weniger gewaltiger Konzerne, denen auf diese Weise die Macht zufällt, politische Öffentlichkeit zu gestalten. So regulieren Empfehlungsalgorithmen politische Kommunikation, ohne dass transparent ist, wie sie das tun, ohne dass andere als ökonomische Zwecke verfolgt werden und ohne dass redaktionelle Verantwortung für Inhalte übernommen wird. Außer in schwerwiegenden Fällen, die von Geschädigten zur Anklage gebracht werden müssen, können die Unternehmen für die Folgen dieser **De-facto-Regulierung** oftmals nicht verantwortlich gemacht werden.

Obwohl es – wie gezeigt – bisher kaum konkrete Belege für das Phänomen der Filterblasen gibt, besteht aus theoretischer Sicht zudem die Möglichkeit, dass sich die Gesellschaft im Laufe der Zeit in verschiedene Gruppen aufteilt, die keine gemeinsame Auffassung über aktuelle Themen mehr haben. Dies könnte vor allem dann passieren, wenn **individuelle Prioritäten** zum vorherrschenden Kriterium der Nachrichten- und Medienauswahl und damit der Weltwahrnehmung werden. Sollte sich ein solches Szenario entwickeln, könnte es für journalistische Medien schwierig werden, noch als ein **Forum von Öffentlichkeit** zu fungieren (Oertel u.a. 2022, S. 136).

Der deutsche Philosoph und Soziologe **Jürgen Habermas** hat aus solchen Entwicklungen den Schluss gezogen, dass die Sozialen Medien ihre „große[n] emanzipatorische[n] Versprechen“ nicht gehalten hätten. Er befürchtet ein „Versiegen“ der „Deliberation“, das heißt der „rationalisierenden Kraft der öffentlichen Auseinandersetzungen“, ohne die die Demokratie nicht auskommt (Habermas 2021: 45, 27). Eine Strukturierung von Öffentlichkeit, die primär von Big Data getrieben wird, könnte am Ende **die demokratische Idee von Deliberation und Politik durch datengetriebenes Systemmanagement ablösen** (Han 2021).

Eine demokratische Gesellschaft muss jedoch selbst entscheiden können, wie ihre Öffentlichkeit strukturiert sein soll. **Öffentlichkeit** ist nicht einfach ein privatwirtschaftlich erzeugtes Produkt, sondern ein gesellschaftliches System, das ebenso wie der Markt, das Recht oder die Medizin aufgrund seiner Relevanz regulierungsbedürftig ist. Ziel staatlicher Medienpolitik muss es daher sein, Demokratie zu bewahren, „**unaccountable**“

**ble power“ einzuschränken** und die Entwicklung technisch erzeugter beziehungsweise (mit-)bedingter Öffentlichkeit selbst einem demokratischen Aushandlungsprozess zugänglich zu machen (Susskind 2022).



## **2. Vorschläge zu Empfehlungsalgorithmen für den Öffentlich-rechtlichen Rundfunk**

### **2.0 Übersicht der Vorschläge**

#### 2.1 Beteiligung menschlicher Akteure

- 2.1.1 Redaktionelle Verantwortlichkeit
- 2.1.2 Erkennbarkeit maschineller Empfehlungen
- 2.1.3 Mitgestaltung von Empfehlungsparametern durch die User:innen
- 2.1.4 Meldestelle für Probleme (Bias, Unsinn, Sexismus, Rassismus, Hatespeech etc.)

#### 2.2 Transparenz und Design

- 2.2.1 Transparenz der Berechnungsgrundlagen
- 2.2.2 Transparenz der Datengrundlage und Popularity bias
- 2.2.3 Transparentes Design und Nudging

#### 2.3 Provokation von Personalisierung

- 2.3.1 News exposure
- 2.3.2 Content Diversity
- 2.3.3 Anti-recommender-Systeme
- 2.3.4 Serendipity
- 2.3.5 Cross-cutting exposure

#### 2.4 Psychologie und Vertrauen

- 2.4.1 Kein ‚Facts versus myths‘
- 2.4.2 Particular encounters schaffen und Weak ties nutzen
- 2.4.3 Soziale Medien nutzen

#### 2.5 Weiterer Forschungs- und Klärungsbedarf

### **2.1 Beteiligung menschlicher Akteure**

Empfehlungsalgorithmen gehören zur Infrastruktur von Öffentlichkeit, deshalb dürfen sie nicht allein durch Algorithmen definiert werden, die den Interessen einiger weniger globaler Digitalriesen gehorchen. Notwendig ist die Beteiligung menschlicher Akteure an der Entwicklung und Weiterentwicklung von Empfehlungsalgorithmen. Maßnahmen wie der „Human in the loop“, die eine Analyse und Modifikation der technischen Entwicklung von Algorithmen durch menschliche Redakteur:innen erlauben, sollten

von Beginn an implementiert werden, damit zeitnah auf Probleme reagiert werden kann.

Auch die Präsentation der Ergebnisse, die Empfehlungsalgorithmen generieren, sollte einem stetigen Monitoring unterliegen. Bei der Einrichtung unterschiedlicher Zuständigkeiten sollte das Prinzip der Verantwortung ernstgenommen werden: Wenn Empfehlungsalgorithmen gesellschaftliche Auswirkungen haben, muss es Personen oder Gremien geben, die es sich zur Aufgabe machen, diese Auswirkungen im Sinne einer pluralistischen Öffentlichkeit aktiv zu gestalten – und die im Ernstfall für die Folgen einstehen.

### *2.1.1 Redaktionelle Verantwortlichkeit*

**Algorithmische Vorschläge** sollten **mit redaktionell kuratierten Inhalten gemischt** werden. Zudem sollten der Redaktion Daten darüber zur Verfügung stehen, wie kuratierte und automatisierte Inhalte **Networks of Trust** erzeugen oder beeinflussen.

Ein erster, naheliegender Bereich der Verantwortung liegt in der Redaktion von Inhalten. Hier existieren zwei Formen der Kuratierung, die für Öffentlich-rechtliche Medien relevant sind: So sollten einerseits neben algorithmischen Empfehlungen nach wie vor auch redaktionelle Empfehlungen angeboten werden. Empfehlungsalgorithmen allein können dem gesetzlich festgeschriebenen Auftrag eines ausgewogenen, relevanten und umfassenden Informationsangebots nicht nachkommen.

Andererseits müssen die personalisierten Angebote, die Empfehlungsalgorithmen erzeugen, für die Redaktion statistisch überprüfbar sein. Auf individueller Ebene sind Radikalisierungseffekte in der Regel nicht nachvollziehbar, sobald die Plattform eine bestimmte Größe erreicht hat. Technische Hilfsmittel sollten aber entwickelt werden, um die Frage beantworten zu können, ob und unter welchen Umständen sich innerhalb bestimmter User:innengruppen möglicherweise Filterblasen oder – wichtiger noch – bestimmte Networks of Trust ausbilden, die die gesellschaftliche Polarisierung aktiv vorantreiben. Auch Versuche der Manipulation von Empfehlungsalgorithmen sollten sichtbar gemacht werden können.

Dieses aktive Monitoring sollte in eine reguläre Überprüfung und Nachjustierung der Algorithmen eingebracht werden. Sollten sich Spuren eines extremistischen Network of Trust ausfindig machen lassen, müssen Möglichkeiten eines direkten redaktionellen Eingriffs in diese User:innen-Netzwerke erwogen werden. Solche Eingriffe sind aus verschiedenen Gründen problematisch, und wirkungsvolle Strategien sind so gut wie gar nicht erprobt. Hier bedarf es redaktioneller Pionierarbeit, sowohl was die Diagnose als auch den Umgang mit negativen Folgen von Empfehlungsalgorithmen betrifft.

### *2.1.2 Erkennbarkeit maschineller Empfehlungen*

**Maschinell erzeugte Vorschläge** sollten in Öffentlich-rechtlichen Medien nicht nur gekennzeichnet, sondern schon durch das Design leicht als automatisch generiert erkennbar sein. Zwischen von Empfehlungsalgorithmen erzeugten und redaktionellen Angeboten sollte ein ausgewogenes und von Menschen bestimmtes Verhältnis bestehen.

Wie redaktionell kuratierte mit algorithmisch erzeugten Vorschlägen ins Verhältnis gesetzt werden, ist vor allem eine Frage des Designs, beispielsweise innerhalb von sortierten Listen, die die unterschiedlichen Kriterien der Produktion ihrer einzelnen Einträge ausweisen. Die Faustregel für ein Öffentlich-rechtliches Angebot sollte sein, dass redaktionelle Angebote gegenüber algorithmisch erstellten nicht offenkundig benachteiligt werden. Eine einheitliche Kennzeichnung könnte für eine einfache und klare Erkennbarkeit sorgen.

### *2.1.3 Mitgestaltung von Empfehlungsparametern durch die User:innen*

Den **User:innen** sollte angeboten werden, auf die Empfehlungen, die sie erhalten, selbst **Einfluss** zu nehmen.

Wie gezeigt wurde, operieren Empfehlungsalgorithmen mit einer Vielzahl unterschiedlicher Parameter. Zumindest einige davon können nicht nur transparent gemacht, sondern auch so gestaltet werden, dass die User:innen selbst darauf Einfluss nehmen können.

So wäre ein Optionsfeld denkbar, in dem etwa der Einfluss des Standorts, der Suchhistorie oder der Interessen von befreundeten User:innen auf die Ergebnisse verändert und/oder der Zeitraum des Nutzungsverhaltens festgelegt werden kann, der der Berechnung der Empfehlungen zugrunde gelegt wird. Ein solcher Ansatz gibt den User:innen einen Teil der Souveränität, die sie an die Algorithmen verloren haben, wieder zurück.

Nutzerdaten der BBC zeigen jedoch, dass das Interesse an einer solchen eigenständigen **Kuratierung von Empfehlungsalgorithmen** sehr gering ist: 97% der BBC-Nutzer:innen haben kein Interesse, auf die Personalisierung ihrer Empfehlungen Einfluss zu nehmen (Hildén 2021). Zudem kann sich ein solcher Mechanismus als ethisch problematisch erweisen, wenn er etwa dazu genutzt wird, zu viel der Verantwortung, die eigentlich die Plattform selbst tragen sollte, an die User:innen zu delegieren.

#### *2.1.4 Meldestelle für Probleme (Bias, Unsinn, Sexismus, Rassismus, Hatespeech etc.)*

Für den Fall, dass sich durch die Einwirkung von Empfehlungsalgorithmen unerwünschte gesellschaftliche Folgen ergeben, sollte es **klare Zuständigkeiten und Verantwortlichkeiten** sowie **leicht auffindbare Prozeduren für die Geschädigten** geben.

Das Prinzip der gesellschaftlichen Verantwortung bedarf einer Regelung von Zuständigkeiten. Diese hängt von dem konkreten Angebot der Plattform/Mediathek ab. Wenn Interaktionen zwischen User:innen vorgesehen sind, sollten Moderator:innen eingesetzt werden. Mindestens sollte eine Beschwerdestelle eingerichtet sein, die auf alle möglichen Probleme – von inhaltlichen Beschwerden bis hin zu kleinen Rechtsstreitigkeiten – vorbereitet ist.

Einem Vorschlag der deutschen Datenethikkommission folgend, sollten solche Angebote an der „**Systemkritizität**“ des Algorithmus beziehungsweise des medialen Angebots insgesamt orientiert sein, also an seinem theoretischen „**Schädigungspotenzial**“ (Datenethikkommission der Bundesregierung 2019: 25). Je nachdem, wie hoch dieses veranschlagt wird, können sich die Anforderungen an Zahl und Qualifikation des eingesetzten Personals sowie an die Formen der Moderation verändern.

## **2.2 Transparenz und Design**

Das Prinzip der **Verantwortlichkeit** muss auch die Entwicklung der Algorithmen bestimmen. Diese sollten so programmiert sein, dass sie den menschlichen Akteur:innen, die für diese Verantwortung einstehen, ihre Arbeit nicht unmöglich machen.

Dafür bedarf es verschiedener Formen der Transparenz. So muss der Empfehlungsalgorithmus zunächst so gestaltet sein, dass er redaktionell bearbeitbar ist; es muss also eine **Transparenz nach innen** etwa dadurch gewährleistet sein, dass Monitoring- und Filter-Mechanismen zur Verfügung gestellt werden (siehe Kapitel 2.1.1 und 2.1.3). Diese betreffen einerseits das Nutzerverhalten, andererseits aber auch die Daten und Medien selbst, die von den eingesetzten Empfehlungsalgorithmen ‚fair‘ behandelt werden müssen.

Zugleich bedarf es einer **Transparenz nach außen**, die sowohl die User:innen als auch wissenschaftliche und politische Akteure adressiert. Diese Transparenz muss jedoch wirksam sein; sie darf nicht als reines Alibi fungieren. Zudem darf sie nicht nur die Inhalte, sondern muss auch das Design betreffen. Im Bereich des Designs lassen sich gleichwohl **Nudging-Mechanismen** einsetzen (Thaler/Sunstein 2021), die unerwünschten

Effekten entgegenarbeiten können – unter der Bedingung, dass auch hier auf Transparenz geachtet wird.

Fragen nach der Möglichkeit und Notwendigkeit von Transparenz werden in der Forschung zu Empfehlungsalgorithmen stark diskutiert. Vielfache Interessenskonflikte, etwa mit Anforderungen des Datenschutzes oder dem Anspruch der Plattformen auf Wahrung von Betriebsgeheimnissen, stehen der Forderung nach Transparenz entgegen. Als Mindeststandard sollte jedoch gelten: Es muss so viel Transparenz bestehen, dass in Konfliktfällen festzustellen ist, **wer inhaltlich und rechtlich verantwortlich** ist.

### *2.2.1 Transparenz der Berechnungsgrundlagen*

**Berechnungsgrundlagen** von Empfehlungsalgorithmen sollten auch bei Verwendung dynamischer Systeme von Künstlicher Intelligenz **statistisch auswertbar** und **einfach adaptierbar** bleiben.

Algorithmen sind Betriebsgeheimnisse, Daten unterliegen dem Datenschutz. Die Forderung, technische Prozesse offenzulegen, steht daher mit anderen legitimen Interessen im Konflikt. Zugleich ist der gesellschaftliche Einfluss von Empfehlungsalgorithmen so stark, dass User:innen und Gesellschaft ein Anrecht darauf haben, grundlegende Mechanismen zu verstehen.

Allerdings gelten die unmittelbaren Effekte von existierenden Transparenzauflagen als eher gering: Aus Zeitgründen und weil es ohnehin häufig keine Alternative dazu gibt, den von Webseiten gesetzten Bedingungen zuzustimmen, werden etwa Einverständniserklärungen in der Regel **ungelesen akzeptiert** und dienen eher als Alibi denn als tatsächliche Informationsgrundlage<sup>15</sup> – auf der Website publizierte Erläuterungen der Funktionsweisen von Algorithmen werden in der Regel ignoriert. Große Unternehmen formulieren ihre **Disclaimer** und Erläuterungen zudem häufig **wie PR-Texte** (Mahapatra 2020), wobei potenzielle Gefahren unerwähnt oder unverständlich bleiben (Gillespie 2018).

**Wichtiger** als „user-facing disclaimers“ sind daher **Transparenzmechanismen, die in Konfliktfällen greifen** und die es zuständigen Redakteur:innen oder **kompetenten Dritten ermöglichen**, die Funktionsweisen von **Empfehlungsalgorithmen zu überprüfen** und zu diskutieren (Leerssen 2020). So sollten Filtermechanismen zur Verfügung stehen, die es nicht nur den User:innen selbst (siehe Kapitel 2.1.3), sondern auch der Redaktion erlauben, das Zustandekommen bestimmter Ergebnisse nachzuvollziehen und gegebenenfalls zu beeinflussen. Vor allem sollte die Frage, welche Rolle unter-

15 Bereits im Jahr 1999 bezeichnete Paul M. Schwartz die Notwendigkeit, während der Nutzung des Webs der Weiterverwertung der eigenen Daten zuzustimmen, als „Consent Trap“ (Schwartz 1999: 1662).

schiedliche Profil-Faktoren bei der Herstellung von User:innen-Identitäten spielen, beantwortbar sein – beispielsweise welche Bedeutung Alter und Geschlecht haben.

User:innendaten und mediale Inhalte sollten darüber hinaus in gemeinsamen, gut strukturierten **Repositorien** gesammelt werden, **die eine externe Beforschung ermöglichen**. Solche Repositorien sollten nicht einfach nur Daten anhäufen, sondern die Evolution von Medienangebot und Mediennutzung insgesamt festhalten, also auch die **Analyse vergangener Zustände der Plattform ermöglichen**. Es bedarf **rechtlicher Grundlagen**, die den Zugriff auf diese Repositorien regeln.

### *2.2.2 Transparenz der Datengrundlage und Popularity bias*

Medien stehen nicht nur gegenüber ihren User:innen, sondern auch gegenüber den Anbietern der Inhalte – Redakteur:innen, Regisseur:innen, Autor:innen und so weiter – in der Pflicht. **Dynamische Algorithmen** sollten deshalb so gestaltet sein, dass sie **redaktionell beeinflussbar** sind, um auf einen sich über die Zeit entwickelnden Popularity bias reagieren zu können.

Durch Empfehlungsalgorithmen entstehen häufig Ungleichgewichte in der Präsentation der Inhalte. So werden in der Regel wenige sehr beliebte Inhalte sehr häufig empfohlen, während große Bestände der zur Verfügung stehenden Medien nur selten oder sogar niemals zur Sichtbarkeit gelangen (Elahi 2021).

Mit einem solchen **Popularity bias** kann eine **unerwünschte Diskriminierung** bestimmter Gruppen oder Produkte einhergehen. Es bedarf daher redaktioneller Instrumente, die die Nutzung der Inhalte statistisch nach relevanten Kriterien aufschlüsseln und die es ermöglichen, bestimmte Inhalte zu fördern oder die Anzahl häufig vorgeschlagener Inhalte insgesamt zu erhöhen.

Eine solche Manipulation erschwert freilich die Unterscheidung zwischen redaktionellen und algorithmisch erzeugten Angeboten. Transparenz kann dabei etwa dadurch geschaffen werden, dass unterschiedliche Vorschlagslisten oder -modi angeboten werden – etwa ein ‚**Discovery mode**‘, der vorrangig weniger bekannte Werke berücksichtigt.

### *2.2.3 Transparentes Design und Nudging*

Neben den Inhalten selbst sollte das **Design transparent** gestaltet und nach Möglichkeit von den User:innen **anpassbar** sein. Wo bewusst mit einer **Choice architecture** oder gar mit **Nudging** gearbeitet wird, sollte auch das **transparent** gemacht werden.

Design spielt eine wesentliche Rolle bei der Auswahl von Empfehlungen. Generell gilt, dass das Design vielfältig und flexibel gestaltbar sein sollte. Wo bestimmte Wahlmöglichkeiten qua Design nahegelegt werden, sollte das begründet werden.

Ein transparentes Design schließt die Möglichkeit der Verwendung von Nudging-Mechanismen nicht aus, wohl aber die von ‚dark patterns‘ beziehungsweise ‚deceptive patterns‘.

Umgekehrt lassen sich auch Hürden errichten, die überwunden werden müssen. Der norwegische Öffentlich-rechtliche Medienanbieter NRK hat etwa damit experimentiert, der Kommentarfunktion auf seiner Website ein kurzes „comprehension quiz“ voranzustellen, das absolviert werden muss, bevor ein Kommentar publiziert werden kann (Lewandowsky 2019: 80). Die Maßnahme ist wirksam, auch wenn ihr Kalkül offensichtlich ist: Sogenannten Trollen und solchen User:innen, die spontane und möglicherweise unbedachte emotionale Reaktionen zu posten versuchen, wird die Teilnahme erschwert, ohne dass sie deshalb censiert würden.

**Empfehlungsalgorithmen können generell als Nudging-Instrumente begriffen werden.** Die Forschung unterscheidet eine hohe Zahl unterschiedlicher Eingriffsmöglichkeiten (Jesse/Jannach 2021), von denen sich einige – unter der Bedingung eines transparenten Einsatzes – auch für die redaktionelle Arbeit von Öffentlich-rechtlichen Medien anbieten. Der aktuelle Stand der politischen Willensbildung zum Nudging sowie das technische Wissen der Bereiche, in denen es schon praktiziert wird, sollten jedoch systematisiert und transparent zusammengefasst werden, um verantwortungsvolle Strategien für öffentlich-rechtliches algorithmisches Nudging zu erarbeiten.

## 2.3 Provokation von Personalisierung

Die naheliegendste und meistdiskutierte Form des Nudgings besteht in der Beeinflussung und Veränderung der Angebote selbst. Sie geschieht in der Regel durch die Implementierung von Mechanismen, die dem eigentlichen Empfehlungsalgorithmus zuwiderlaufen. Ziel ist eine Konfrontation der User:innen mit Inhalten, die eine möglicherweise festgefahrenen Weltsicht herausfordern können. Dabei lässt sich eine ganze Reihe unterschiedlicher Ansätze unterscheiden.

### 2.3.1 News exposure

Zu den wichtigsten Medieninhalten, die **tatsachenbasierte Information** über die Welt, aber auch ein **diverse Input** garantieren, gehören **Nachrichten**. Daraus ergibt sich die naheliegende, wenn auch häufig übersehene Empfehlung, den „overall amount of news exposure“ der User:innen, also **das Nachrichtenangebot ins-**

**gesamt**, zu erhöhen. Das Angebot sollte groß sein und häufig wiederholt werden.

Diese Empfehlung gilt analog für sämtliche Inhalte von öffentlichem Interesse<sup>16</sup>, insbesondere kulturelle. Wichtig ist dabei, dass ‚Exposure‘ nur das Angebot betrifft, aber die Entscheidung darüber, ob dieses Nachrichtenangebot angenommen wird und wieviel Zeit und Aufwand diesen Nachrichten gewidmet wird, bei den User:innen bleibt.

Durch News exposure werden User:innen kontinuierlich Nachrichten angeboten mit dem Ziel, die Wahrscheinlichkeit der aktiven Kenntnisnahme zu erhöhen. Hier gilt *cum grano salis* die Formel: Viel hilft viel. Die Gefahr eines Zuviel ist dabei gering, auch wenn irgendwann die Ermüdung der Aufmerksamkeit droht. Der Trade-off liegt also im repetitiven, insistierenden Ringen um Aufmerksamkeit und der dadurch bewirkten Ermüdung der User:innen.

An dieser Stelle sei noch einmal darauf hingewiesen, dass der Nachrichtenkonsum im Internet statistisch möglicherweise nur eine marginale Rolle spielt; häufig ist auch die Datenlage unsicher.<sup>17</sup> Dieser Befund hat auch mit der tendenziellen Grenzverwischung zwischen Nachricht und Nutzer:innenkommunikation zu tun, die Informationsintermediäre betreiben. Nachrichten sollten daher als besondere Form des Inhalts markiert und hervorgehoben werden.

### 2.3.2 Content Diversity

Die sogenannte **Content Diversity** der Angebote soll eine diskursfähige Öffentlichkeit stärken und einer möglichen gesellschaftlichen Polarisierung entgegenarbeiten.

Auch wenn die empirische Forschung nicht darauf hindeutet, dass Empfehlungsalgorithmen notwendig zu homogenen oder uniformen Inhaltsangeboten und diese wiederum durch Filterblasen und Echokammern kausal zu einer gesellschaftlichen Polarisierung führen, ist es dennoch sinnvoll, gegen eine mögliche gesellschaftliche Polarisierung vorzugehen. ‚Diverse‘ Angebote zeigen den User:innen nutzer:innenfremde, ‚andere‘ Inhalte als jene, die sie bisher präferiert haben. Dabei handelt es sich um Angebote, die vielfältig zwischen den Polen von ‚neu, unbekannt‘ und ‚schon bekannt und wird abgelehnt‘ angesiedelt sein können. Diverse Angebote sollen zu **Überraschungseffekt**

16 Die Begriffe „öffentliches Interesse“ oder „public interest“ sind allgemein gebräuchlich, aber bisher nicht exakt definiert. Eine hinreichend präzise Bestimmung liefert die BBC im Rahmen ihrer „Public Value Tests“; [https://www.bbc.co.uk/bbctrust/governance/tools\\_we\\_use/public\\_value\\_tests.html](https://www.bbc.co.uk/bbctrust/governance/tools_we_use/public_value_tests.html). Siehe auch Rhein u.a. 2021.

17 Denn, so Hildén: „Mere access to the full programing portfolio says very little about what content is ultimately consumed“ (Hildén 2021).

ten oder sogar **Provokationen** führen. Sie eignen sich für ein Alteritätstraining und ein Kultivieren der Ambiguitätstoleranz.

Während News exposure ein quantitatives Instrument darstellt, zielt die Sicherstellung von Content Diversity auf eine qualitative Form der Einflussnahme. Die Implementierung eines Diversitätskriteriums in den Algorithmus wirkt dabei der rein **aufmerksamkeitsökonomischen Orientierung** von Empfehlungsalgorithmen entgegen. Die Herausforderung in der Umsetzung liegt auf der Hand: Zu viel Diversität läuft der Idee der Personalisierung zuwider; ein zu diverses Informationsangebot kann dazu führen, dass User:innen das Angebot ablehnen. Der Trade-off besteht also zwischen **Diversität** und **Personalisierung**.

### 2.3.3 Anti-recommender-Systeme

Dem zentralen Empfehlungsalgorithmus kann ein weiterer Algorithmus entgegengestellt werden, der **ausschließlich Inhalte** anzeigt, die den User:innen **normalerweise nicht vorgeschlagen** würden.

Eine Möglichkeit, Content Diversity herzustellen, besteht darin, innerhalb des Designs der Plattform einen Platz zu schaffen, der ausschließlich ‚anderen‘ Inhalten gewidmet ist. Ein solches Anti-recommender-System kann etwa besonders fernliegende Vorschläge anzeigen (in Sinne einer Cross-cutting exposure, vgl. Kapitel 2.3.5) oder Vorschläge, die User:innen mit abweichenden, aber klar umrissen Profilen gemacht werden würden (vgl. Abdollahpour u.a. 2019), oder aber **Vorschläge, die das aktuelle Nutzungsverhalten kommentieren** („Da du heute noch keine Nachrichten gesehen hast, hier eine aktuelle Sendung.“) (vgl. Hirschmeier 2019).

Anti-recommender-Systeme sind Angebote, die verschiedene der in diesem Kapitel erläuterten technischen Prinzipien miteinander verbinden können. Ihr spezifisches Potenzial besteht darin, Auswahlkriterien von Empfehlungsalgorithmen transparent zu machen, indem das aktuelle Angebot mit alternativen Kriterien-Sets konfrontiert wird. Ihre **Gefahr** liegt hingegen erneut darin, von den User:innen **als Belehrung wahrgenommen** und abgelehnt zu werden.

### 2.3.4 Serendipity

„Serendipity“ bedeutet, dass ein **Teil der Angebote**, die den User:innen gemacht werden, für diese **neu oder überraschend** sind. Serendipity sollte in allen Empfehlungsalgorithmen zum Einsatz gebracht werden, um das Zirkulieren der User:innen um immer gleiche Inhalte zu verhindern.

Serendipity-Mechanismen erzeugen **zufällig diverse Angebote**, die in die Liste der Empfehlungen eingestreut werden, und wirken deshalb vermutlich auf die User:innen

**weniger erzieherisch oder paternalistisch** als etwa das Anti-recommender-System. Serendipity kann von den User:innen selbst hergestellt und durch die Algorithmen verstärkte Muster und Rituale durchbrechen und einem zu einseitigen, eintönigen oder gar langweiligen Angebot entgegenwirken. Es handelt sich daher um Mechanismen, die auch im Sinne der Aufmerksamkeitsökonomie eingesetzt werden.

Der Einsatz von Serendipity-Algorithmen gehorcht selten einem reinen Zufallsprinzip. Vielmehr spielen Kriterien wie ‚Novelty‘ und ‚Unexpectedness‘, aber auch ‚Usefulness‘ und ‚Desirability‘ eine zentrale Rolle, die wiederum algorithmisch kalkuliert werden sollen und die das Personalisierungsparadigma nicht aufbrechen, sondern eher erweitern. Technisch spiegeln sich in den Serendipity-Algorithmen die Funktionsweisen von Empfehlungsalgorithmen insgesamt: So kommen etwa „collaborative“, „content-based“ und „context-based filtering“ zum Einsatz (Ziarani/Ravanmehr 2021).

Unter dem Stichwort ‚Serendipity‘ wird also eine Vielzahl und Vielfalt zu den jeweils verwendeten Empfehlungsalgorithmen unterschiedlicher Ergänzungsalgorithmen verstanden; ob Serendipity-Algorithmen jedoch tatsächlich die Öffentlichkeit stärken, ist bislang ungenügend erforscht. Eine offenkundige Chance von Serendipity besteht darin, User:innen mit Inhalten zu versorgen, die ihren Interessen im weitesten Sinne entsprechen, also nicht unmittelbar auf Ablehnung stoßen, und die trotzdem nicht die Vorurteile möglicher Networks of Trust bestätigen.

### 2.3.5 Cross-cutting exposure

Cross-cutting exposure als sogenannte **dosierte User:innen-Provokation** kann die **Alteritätstoleranz** der User:innen **erhöhen**, ihre **Diskursfähigkeit steigern** und einer **Polarisierung entgegenarbeiten**, indem andere, von der Meinung der User:innen abweichende Standpunkte angeboten werden. Cross-cutting exposure sollte dosiert zum Einsatz gebracht werden.

„Cross-cutting exposure“ bezeichnet **transversale Angebote**, die **zumindest in Teilen** den **Überzeugungen der User:innen zuwiderlaufen**. Twitter-Daten aus der Zeit vor der Übernahme durch Elon Musk zeigen, dass rund ein Drittel der angebotenen Polit-Tweets den politischen Überzeugungen von interessierten User:innen entgegenstanden. Sogar auf einer emotional relativ stark polarisierten Plattform wie Twitter – seit Sommer 2023 X – gibt es also traditionell eine erhebliche Toleranz für Cross-cutting exposure.<sup>18</sup>

18 „Bakshy, Messing, and Adamic (2015) report similar results on Facebook. Their analysis examines the ideological content of the Facebook news feeds of 10.1 million US Facebook users. Unsurprisingly, they find that a majority of friendship links take place between people within the same ideological groups and that users are indeed more likely to engage with congenial content. However, many friendships (20 percent for conservatives and 18 percent for liberals) do cut across ideological groups; and, most importantly, exposure to ideologically diverse news and opinion is also high: on average, around 30 percent of political news stories that users see in their news feed are crosscutting. This proportion is

Die algorithmische Implementierung einer solchen Strategie kann den Austausch zwischen feindlichen Lagern befördern, sie kann aber auch zu Eskalationen und weiterer Polarisierung führen. Der **Trade-off-Spielraum** zwischen dem, was von den User:innen als zumutbar und als unzumutbar empfunden wird, ist **enger** als bei der Content Diversity und der Serendipity, weil die Cross-cutting exposure eine Art **gezielten Widerspruch** gegen die Überzeugungen und Vorlieben der User:innen darstellt.

## 2.4 Psychologie und Vertrauen

Die Vorschläge in diesem Kapitel schließen an die Erkenntnisse an, die in Kapitel 1.6 ausgeführt wurden: Vor allem in Fällen, in denen Echokammer-Effekte nachweisbar sind oder zumindest befürchtet werden, bedarf es einer gezielten Nutzung von Dynamiken des Vertrauens, um unerwünschten Resultaten der Wirkung von Empfehlungsalgorithmen zu begegnen.

### 2.4.1 *Kein ‚Facts versus myths‘*

Echokammern sollte **nicht** durch einen **offensiven** Einsatz von **Widerlegungsstrategien** begegnet werden.

Wie bereits im Abschnitt 1.6 gezeigt, ist es **so gut wie aussichtslos**, im Netz **Mythen durch Fakten zu widerlegen**. Insbesondere der aufklärerische Gestus des ‚**Facts versus myths**‘ führt zu einer Reihe von gut erforschten psychologischen Abwehrreflexen, deshalb sollte man **darauf verzichten**.

Das heißt nicht, dass man im Netz grundsätzlich auf Fakten und Aufklärung verzichten sollte. Für User:innen, die keine Mitglieder von Echokammern sind, können etwa Faktenchecker eine wertvolle Quelle der Information darstellen. Werden sie jedoch ausdrücklich zum Zweck der Richtigstellung irriger, teilweise gefährlicher Meinungen eingesetzt, können sie sich als kontraproduktiv erweisen.

### 2.4.2 *Particular encounters schaffen und Weak ties nutzen*

Die Nutzer:innen-Nutzer:innen-Interaktion sollte das Zustandekommen von **Particular encounters ermöglichen**.

Am ehesten lassen sich Menschen über ihre Irrtümer aufklären, wenn sie der korrigierenden Person vertrauen und diese Person grundsätzlich – abgesehen von den Irrtü-

remarkably similar to what Barberá (2015) reports using Twitter data: 33 percent of the tweets to which a sample of politically interested users were potentially exposed are cross-cutting.” (Barberá 2020: 40).

mern – ihre eigene Weltanschauung teilt. Am meisten **Vertrauen** haben Menschen in der Regel zu nahen Freunden und Bekannten (‘strong ties’); von diesen sind allerdings nur selten Korrekturen zu erwarten. Auch zu entfernteren Bekannten (‘weak ties’) bestehen Vertrauensverhältnisse; zugleich ist die Distanz größer, sodass die **Wahrscheinlichkeit für abweichende Weltanschauungen oder Korrekturen eigener Ansichten** größer ist.

Menschen, die ihren Echokammern entkommen sind, berichten häufig von Einzelbegegnungen (‘particular encounters’), in denen sie plötzlich Vertrauen zu Menschen außerhalb ihrer Echokammern gefasst haben.<sup>19</sup> Die **Vertrauenswürdigkeit** einer Person ist also **wichtiger als die Argumente**, die sie vorbringt.

Particular encounters lassen sich nicht erzwingen, sondern ereignen sich selten und zufällig. Soziale Netzwerke sollten so gestaltet sein, dass sie **Begegnungen zwischen User:innen ermöglichen, die Echokammer-Effekten entgegenwirken**. Die genauen Mechanismen, die solche Ereignisse herbeiführen können, müssen noch genauer erforscht werden.

#### *2.4.3 Soziale Medien nutzen*

Empfehlungen im linearen Programm, auf einer Website und in einer Mediathek sind von relativ geringer Reichweite. Die Änderungstendenz im Mediennutzungsverhalten jüngerer Menschen legt zudem nahe, dass diese Reichweite zukünftig noch schrumpft. Deshalb empfiehlt es sich, auch die **Sozialen Medien** wie X, Instagram, YouTube oder TikTok in den Blick zu nehmen, um von der potenziell hohen Reichweite dieser Plattformen zu profitieren. Zu entwickeln sind außerdem **Strategien**, wie sich die Zahl **der Weiterempfehlung durch die User:innen erhöhen lässt**.

Die deutsche ARD hatte im Juli 2023 270.000 Follower:innen auf X (also ca. 0,3 % der Bevölkerung) und 128.000 auf Instagram. Funk, das auf junge Leute zielende „Content-Netzwerk von ARD & ZDF“, hatte zu diesem Zeitpunkt 1,7 Millionen Follower:innen auf Instagram (also ca. 2 % der Bevölkerung; zum Vergleich: rtl.letzeburg hatte im Juli 2023 38.000 Follower:innen auf Instagram, das entsprach knapp 6 % der Bevölkerung). Seit 2023 wird in Deutschland diskutiert, ob Funk mit seinem weitgehenden Verzicht auf journalistische Standards, seinem Sensationalismus und seiner Anpassung an identitätspolitische Themen zu weit geht, um noch einen öffentlich-rechtlichen Anspruch zu vertreten (Brinkmann 2023). Das mag sich auf die Entwicklung der Accounts und Follower:innenzahlen ausgewirkt haben. Anfang 2025 liegt @DasErste auf X bei 734.000 Follower:innen, „ARD Mediathek & Das Erste“ („ardmediathek“) auf Instagram bei 691.000 Follower:innen, während Funk immer noch 1,7 Millionen Follower:innen auf

19 „[F]riends' news recommendations are powerful social cues that can reduce partisan selective exposure to levels indistinguishable from chance.“ (Barberá 2020: 39)

Instagram hat (rtl.letzeburg hat im Januar 2025 51.300 Follower:innen auf Instagram). Die Zahlen legen den Schluss nahe, dass journalistische Qualität zwar nicht irrelevant ist, aber Ton, Kürze und visuelle Machart von Social Media einen größeren Einfluss auf ihre Reichweite haben und deshalb für die Sichtbarkeit bei jüngeren Menschen so gut wie alternativlos sind.

Das bedeutet: Der **Trade-off** zwischen **journalistischem Anspruch** (Breite und Tiefe von Inhalten, primär sprachbasiert) und **Social-Media-Tauglichkeit** (Kürze und Emotionalität von Content, stark bildbasiert) ist unverzichtbar.

Das wiederum legt nahe, **Strategien** zu entwickeln und in der Praxis zu prüfen, wie sich die Zahl der **Weiterempfehlungen** öffentlich-rechtlicher Inhalte durch User:innen privatwirtschaftlicher Sozialer Netzwerke – oder besser noch: durch Influencer:innen mit hoher Reichweite – erhöhen lässt. Das ist insbesondere wichtig für die genannten Empfehlungen durch Weak ties, durch die allein sich möglicherweise existierende Echokammern aufbrechen lassen.

## 2.5 Weiterer Forschungs- und Klärungsbedarf

Das vorliegende Whitepaper formuliert Leitlinien für Empfehlungsalgorithmen von Öffentlich-rechtlichen Medien in Luxemburg. Diese Empfehlungen basieren auf den Forschungsdiskussionen, die im ersten Teil umrissen worden sind. Es liegt auf der Hand, dass diese Darstellung nicht ansatzweise vollständig ist – und auch prinzipiell nicht vollständig sein kann, weil die Forschung permanent und rasant voranschreitet. Dieses Papier sollte also ein **regelmäßiges Update** erfahren.

Hinzu kommen theoretische und begriffliche Herausforderungen. Zwar befasst sich dieses Whitepaper eigentlich nicht mit dem Nudging; zugleich lassen sich viele der hier ausgesprochenen Empfehlungen – ebenso wie Empfehlungsalgorithmen insgesamt – auch als Nudging-Strategien beschreiben (vgl. oben 2.2.3). Es würde sich lohnen, an einem anderen Ort noch einmal grundsätzlich über den Zusammenhang von Empfehlungsalgorithmen und **Nudging nachzudenken**. Nudging ist ein komplexes Thema, das zugleich **technisch** und **psychologisch** – Designfragen zum Beispiel berühren immer beide Disziplinen zugleich – sowie **ethisch** reflektiert werden muss.

Offen muss auch die Frage bleiben, wo genau die hier thematisierten algorithmischen Empfehlungen gegeben werden sollen – und wie solche **Empfehlungen** sinnvollerweise (siehe 2.4.3) auch auf **privatwirtschaftliche Plattformen** ausgedehnt werden können, um die dort schon bestehenden **Networks of trust** zu nutzen, idealerweise auch über **Influencer:innen** mit hoher Reichweite. Will man sich darüber Gedanken machen, wäre eine **systematische Übersicht der existierenden Sozialen Netzwerke und Digitalplattformen** hilfreich, und zwar **quantitativ** wie **qualitativ**.

Eine Herausforderung für sich stellt die Frage dar, wie sich überhaupt Daten erheben lassen – technisch wie rechtlich. Diese Frage ist zu klären, bevor überhaupt über die Öffnung von Schnittstellen oder die Veröffentlichung von – möglicherweise anonymisierten – Daten nachgedacht werden kann. Zudem bedarf es komplexer Analyseinstrumente, die es ermöglichen, über algorithmische Effekte wie Filterblasen hinausgehende technisch-soziale Dynamiken zu erfassen.

Ein technisches Problem, insbesondere in Hinblick auf die immer stärkere Einbindung von Künstlicher Intelligenz, ist generell die Beobachtbarkeit algorithmischer Prozesse. Ist sie nicht gegeben, lassen sich durch redaktionelles Eingreifen nur Ergebnisse korrigieren, nicht aber die Prozesse, die sie erzeugen. Auch hier schreitet die Forschung rasant voran.

### 3. Literatur

- Abdollahpouri, Himan/Burke, Robin/Mobasher, Bamshad: Managing popularity bias in recommender systems with personalized re-ranking. In: Barták, Roman/Brawner, Keith W. (Hg.): Proceedings of the Thirty-Second International Florida Artificial Intelligence Research Society Conference, May 19-22 (FLAIRS '19). Sarasota 2019, S. 413–418.
- Ada Lovelace Institute: Inform, educate, entertain ... and recommend? Exploring the use and ethics of recommendation systems in public service media. London 2022.
- Aggarwal, Charu C.: Recommender Systems. The Textbook. Cham u.a. 2016.
- Allcott, Hunt u.a.: The Welfare Effects of Social Media. In: American Economic Review 110 (2020), H. 3, S. 629–676; online unter: <https://doi.org/10.1257/aer.20190658> [Zugriff 31.07.2023].
- Allcott, Hunt/Gentzkow, Matthew: Social Media and Fake News in the 2016 Election. In: Journal of Economic Perspectives 31 (2017), H. 2, S. 211–236.
- Arceneaux, Kevin/Johnson, Martin: Changing minds or changing channels? Partisan news in an age of choice. Chicago 2014; online unter: <https://press.uchicago.edu/ucp/books/book/chicago/C/bo15731464.html> [Zugriff 31.07.2023].
- Arguedas, Amy Ross u.a.: Echo chambers, filter bubbles, and polarisation: a literature review (19.01.2022). In: Reuters Institute for the Study of Journalism. University of Oxford; online unter: <https://reutersinstitute.politics.ox.ac.uk/echo-chambers-filter-bubbles-and-polarisation-literature-review> [Zugriff 24.08.2023].
- Barberá, Pablo: Social Media, Echo Chambers, and Political Polarization. In: Persily, Nathaniel/Tucker, Joshua A. (Hg.): Social Media and Democracy. The State of the Field and Prospects for Reform. Cambridge 2020, S. 34–55; online unter: <https://www.cambridge.org/core/books/social-media-and-democracy/social-media-echo-chambers-and-political-polarization/333A5B4DE1B67EFF7876261118CCFE19> [Zugriff 24.08.2023].
- Brinkmann, Janis: Journalistische Grenzgänger. Wie die Reportage-Formate von funk Wirklichkeit konstruieren. Frankfurt a.M. 2023; online unter: [https://www.otto-brenner-stiftung.de/fileadmin/user\\_data/stiftung/02\\_Wissenschaftsportal/03\\_Publikationen/AH111\\_Funk.pdf](https://www.otto-brenner-stiftung.de/fileadmin/user_data/stiftung/02_Wissenschaftsportal/03_Publikationen/AH111_Funk.pdf) [Zugriff 24.08.2023].
- Broniatowski, David A. u.a.: Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. In: Am J Public Health 108 (2018), H. 10, S. 1378–1384.
- Bruns, Axel: Filter bubble. In: Internet Policy Review 8 (2019), H. 4; online unter: <https://policyreview.info/pdf/policyreview-2019-4-1426.pdf> [Zugriff 31.07.2023].
- Burrell, Jenna: How the machine ‘thinks’: Understanding opacity in machine learning algorithms, Big Data & Society 3 (2016), H. 1; online unter: <https://journals.sagepub.com/doi/pdf/10.1177/2053951715622512> [Zugriff 31.07.2023].
- Casemajor, Nathalie/Rocheleau, Sylvain: Figuring digital cascades. Issue framing in digital media. In: Journal of Digital Social Research 3 (2019), H. 1, S. 60–87.
- Christian, Brian: The alignment problem: Machine learning and human values. New York 2020.

- Cinus, Federico u.a.: The Effect of People Recommenders on Echo Chambers and Polarization. In: Proceedings of the International AAAI Conference on Web and Social Media 16 (2022), H. 1, S. 90–101.
- Dahlgren, Peter/Shehata, Adam/Strömbäck, Jesper: Reinforcing spirals at work? Mutual influences between selective news exposure and ideological leaning. In: European Journal of Communication 34 (2019), H. 2, S. 159–174; online unter: <https://journals.sagepub.com/doi/abs/10.1177/0267323119830056> [Zugriff 31.07.2023].
- Datenethikkommission der Bundesregierung: Gutachten der Datenethikkommission der Bundesregierung. Berlin 2019; online unter: [https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?\\_\\_blob=publicationFile&v=6](https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf?__blob=publicationFile&v=6) [Zugriff 31.07.2023].
- Daub, Adrian: What Tech Calls Thinking. An Inquiry into the Intellectual Bedrock of Silicon Valley. New York 2020.
- Decker, Oliver u.a. (Hg.): Polarisiert und radikalisiert? Medienmisstrauen und die Folgen für die Demokratie. Frankfurt a. M. 2017; online unter: [https://www.otto-brenner-stiftung.de/fileadmin/user\\_data/stiftung/02\\_Wissenschaftsportal/03\\_Publikationen/AP27\\_Medienmisstrauen\\_Decker\\_2017\\_10\\_06.pdf](https://www.otto-brenner-stiftung.de/fileadmin/user_data/stiftung/02_Wissenschaftsportal/03_Publikationen/AP27_Medienmisstrauen_Decker_2017_10_06.pdf) [Zugriff 31.07.2023].
- Dündar, Pinar/Ranaivoson, Heritiana: Science by YouTube. An Analysis of YouTube's Recommendations on the Climate Change Issue. In: Observatorio (OBS\*) 16 (2022), H. 3, S. 53–76; online unter: <https://obs.obercom.pt/index.php/obs/article/view/2061> [Zugriff 31.07.2023].
- Edelson, Laura/McCoy, David: Facebook is obstructing our work on disinformation. Other researchers could be next (14.8.2021). In: The Guardian; online unter: <https://www.theguardian.com/technology/2021/aug/14/facebook-research-disinformation-politics> [Zugriff 31.07.2023].
- Edwards, Lilian/Veale, Michael: Slave to the algorithm? Why a 'Right to an Explanation' is probably not the remedy you are looking for. In: Duke Law & Technology Review 16 (2017), H. 18; online: <https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1315&context=dltr> [Zugriff 31.07.2023].
- Egan, Lauren: 'They're killing people': Biden blames Facebook, other social media for allowing Covid misinformation (16.07.2021). In: NBC News; online unter: <https://www.nbcnews.com/politics/white-house/they-re-killing-people-biden-blames-facebook-other-social-media-n1274232> [Zugriff 24.08.2023].
- Elahi, Mehdi u.a.: Towards responsible media recommendation. In: AI and Ethics 2 (2022), H. 1, S. 1–12.
- Elahi, Mehdi u.a.: Beyond algorithmic fairness in recommender systems. In: Adjunct Proceedings of the ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21). New York 2021; online unter: <https://dl.acm.org/doi/fullHtml/10.1145/3450614.3461685> [Zugriff 31.07.2023].

- Eordogh, Fruzsina: YouTube's Related Video Algorithm Helpful To Predators (18.02.2018). In: Forbes, online unter: <https://www.forbes.com/sites/fruzsinaeordogh/2019/02/18/youtubes-related-video-algorithm-helpful-to-predators/> [Zugriff 24.08.2023].
- Epstein, Robert/Robertson, Ronald E.: The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. In: PNAS 112 (2015), H. 33, S. 4512–4521.
- Europäische Kommission: The 2022 Code of Practice on Disinformation; online unter: <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation> [Zugriff 24.08.2023].
- Eyal, Nir: Hooked. How to Build Habit-Forming Products. New York 2014.
- Fisher, Max: The Chaos Machine. London 2023.
- Flaxman, Seth/Goel, Sharad/Rao, Justin M.: Filter Bubbles, Echo Chambers, and Online News Consumption. In: Public Opinion Quarterly 80 (2016), S. 298–320; online unter: <https://Sharad.com/papers/bubbles.pdf> [Zugriff 01.08.2023].
- Fogg, B. J./Cuellar Gregory/Danielson, David: Motivating, influencing, and persuading users: An introduction to captology. In: Sears, Andrew/Jacko, Julie (Hg.): The Human–Computer Interaction Fundamentals. New York 2009.
- Geiß, Stefan: Die Aufmerksamkeitsspanne der Öffentlichkeit. Eine Studie zur Dauer und Intensität von Meinungsbildungsprozessen. Baden-Baden 2015.
- Gillespie, Tarleton: Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media. Yale 2018; online unter: [https://www.researchgate.net/publication/327186182\\_Custodians\\_of\\_the\\_internet\\_Platforms\\_content\\_moderation\\_and\\_the\\_hidden\\_decisions\\_that\\_shape\\_social\\_media](https://www.researchgate.net/publication/327186182_Custodians_of_the_internet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media) [Zugriff 01.08.2023].
- Gillespie, Tarleton: Platforms Intervene, Social Media + Society 1 (2015), H. 1; online unter: <https://journals.sagepub.com/doi/10.1177/2056305115580479> [Zugriff 01.08.2023].
- Gouvernement du Grand-Duché de Luxembourg: Convention conclue entre l'État du Grand-Duché de Luxembourg et le Média de service public 100,7; online unter: <https://gouvernement.lu/dam-assets/documents/actualites/2023/03-mars/31-convention-media-service-public/convention-31032023.pdf> [Zugriff 24.08.2023].
- Gu, Lion/Kropotov, Vladimir/Yarochkin, Fyodor: The Fake News Machine: How Propagandists Abuse the Internet and Manipulate the Public. 2017; online unter: [https://documents.trendmicro.com/assets/white\\_papers/wp-fake-news-machine-how-propagandists-abuse-the-internet.pdf](https://documents.trendmicro.com/assets/white_papers/wp-fake-news-machine-how-propagandists-abuse-the-internet.pdf) [Zugriff 01.08.2023].
- Habermas, Jürgen: Strukturwandel der Öffentlichkeit. Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft. Mit einem Vorwort zur Neuauflage 1990. Frankfurt a.M. 2023 (1990).
- Habermas, Jürgen: Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen Öffentlichkeit. In: Seeliger, Martin/Sevignani, Sebastian (Hg.): Ein neuer Strukturwandel der Öffentlichkeit? Leviathan. Berliner Zeitschrift für Sozialwissenschaft, Sonderband Leviathan 37 (2021), S 470–500.

- Han, Byung-Chul: Infokratie. Digitalisierung und die Krise der Demokratie. Berlin 2021.
- Havlicek, Sasha u.a. (Hg.): Make Germany Great Again. Kremlin, Alt-Right and International Influences in the 2017 German Elections. London u.a. 2017; online unter: <https://www.isdglobal.org/wp-content/uploads/2017/12/Make-Germany-Great-Again-ENG-061217.pdf> [Zugriff 31.07.2023].
- Helberger, Natali/Karppinen, Kari/d'Acunto, Lucia: Exposure diversity as a design principle for recommender systems. In: Information, Communication & Society 21 (2016), H. 2, S. 191–207; online unter: [https://pure.uva.nl/ws/files/17932244/Exposure\\_diversity\\_as\\_a\\_design\\_principle.pdf](https://pure.uva.nl/ws/files/17932244/Exposure_diversity_as_a_design_principle.pdf) [Zugriff 01.08.2023].
- Helberger, Natali: Diversity by design. In: Journal of Information Policy 1 (2011), S. 441–469; online unter: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2197252](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2197252) [Zugriff 01.08.2023].
- Hildén, Jockum: The Public Service Approach to Recommender Systems: Filtering to Cultivate. In: Television & New Media 23 (2022), H. 7, S. 777–796; online unter: <https://journals.sagepub.com/doi/pdf/10.1177/15274764211020106> [Zugriff 24.08.2023].
- Hirschmeier, Stefan: Towards Explanations of Anti-Recommender Content in Public Radio. In: Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization. New York 2019, S. 229–233; online unter: <https://doi.org/10.1145/3314183.3323454> [Zugriff 24.08.2023].
- Hou, Ye u.a.: Social media addiction: Its impact, mediation, and intervention. Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 13 (2019), H. 1; online unter: <https://doi.org/10.5817/CP2019-1-4> [Zugriff 24.08.2023].
- Iyengar, Shanto u.a.: The origins and consequences of affective polarization in the United States. In: Annual Review of Political Science 22 (2018), S. 129–146; online unter: <https://www.annualreviews.org/doi/abs/10.1146/annurev-polisci-051117-073034> [Zugriff 01.08.2023].
- Jesse, Mathias/Jannach, Dietmar: Digital nudging with recommender systems: Survey and future directions. In: Computers in Human Behavior Reports 3 (2021); online unter: <https://doi.org/10.1016/j.chbr.2020.100052> [Zugriff 24.08.2023].
- Kaiser, Jonas/Rauchfleisch, Adrian: YouTubes Algorithmen sorgen dafür, dass AfD-Fans unter sich bleiben (22.09.2017). In: Vice; online unter: <https://motherboard.vice.com/de/article/59d98n/youtubes-algorithmen-sorgen-dafur-dass-afd-fans-unter-sich-bleiben> [Zugriff 01.08.2023].
- Kosinski, Michal/Stillwell, David/Graepel, Thore: Private traits and attributes are predictable from digital records of human behavior. In: Proceedings of the national academy of sciences 110 (2013), S. 5802–5805.
- Krafft, Tobias D./Gamer, Michael/Zweig, Katharina A.: Wer sieht was? Personalisierung, Regionalisierung un die Frage nach der Filterblase in Googles Suchmaschine. Final Report for research Project #DatenSpende: Google und die Bundestagswahl 2017. Kaiserslautern 2018; online unter: <https://www.blm.de/files/pdf2/bericht-datenpende---wer-sieht-was-auf-google.pdf> [Zugriff 01.08.2023].

- Kumkar, Nils C.: Alternative Fakten. Zur Praxis der kommunikativen Erkenntnisverweigerung. Berlin 2022.
- Leerssen, Paddy: The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems. European Journal of Law and Technology 11 (2020), H. 2; online unter: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3544009](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3544009) [Zugriff 01.08.2023].
- Lelkes, Yphtach/Sood, Gaurav/Iyengar, Shanto: The hostile audience: The effect of access to broadband internet on partisan affect. In: American Journal of Political Science, 61 (2017), H. 1, S. 5–20; online unter: <https://www.jstor.org/stable/26379489> [Zugriff 01.08.2023].
- Lewandowsky, Stephan: The ‘Post-Truth’ World, Misinformation, and Information Literacy: a Perspective From Cognitive Science. In: Goldstein, Stéphane (Hg.): Informed Societies: Why Information Literacy Matters for Citizenship, Participation and Democracy. London 2019, S. 69–88.
- Lewis, Rebecca: Alternative Influence. Broadcasting the Reactionary Right on YouTube. Data & Society Research Institute 2018; online unter: [https://datasociety.net/wp-content/uploads/2018/09/DS\\_Alternative\\_Influence.pdf](https://datasociety.net/wp-content/uploads/2018/09/DS_Alternative_Influence.pdf) [Zugriff 24.08.2023].
- Litschka, Michael: Algorithmen-basierte Empfehlungssysteme und die Entstehung von Filterblasen in der Plattformökonomie. Ein Experiment auf YouTube. In: Schicha, Christian/Stapf, Ingrid/Sell, Saskia (Hg.): Medien und Wahrheit. Medienethische Perspektiven auf Desinformation, Lügen und „Fake News“. Baden-Baden 2021, S. 377–388.
- Mahapatra, Amogh: How Instagram suggests new content (10.12.2020). In: Engineering at Meta; online unter: <https://engineering.fb.com/2020/12/10/web/how-instagram-suggests-new-content/> [Zugriff 24.08.2023].
- Marwick, Alice/Lewis, Rebecca: Media Manipulation and Disinformation Online, online unter: [https://datasociety.net/wp-content/uploads/2017/05/DataAndSociety\\_MediaManipulationAndDisinformationOnline-1.pdf](https://datasociety.net/wp-content/uploads/2017/05/DataAndSociety_MediaManipulationAndDisinformationOnline-1.pdf) [Zugriff 01.08.2023].
- McLuhan, Marshall: Understanding Media. The Extensions of Man. London/New York 2001 (1964).
- McPherson, Poppy, Facebook says it was ‘too slow’ to fight hate speech in Myanmar (16.08.2018). In: Reuters; online unter: <https://www.reuters.com/article/us-myanmar-facebook-rohingya/facebook-says-it-was-too-slow-to-fight-hate-speech-in-myanmar-idUSKBN1L1066> [Zugriff 24.08.2023].
- Medienstaatsvertrag (MStV) in der Fassung des Dritten Staatsvertrags zur Änderung medienrechtlicher Staatsverträge (Dritter Medienänderungsstaatsvertrag) in Kraft seit 01. Juli 2023; online unter: [https://www.die-medienanstalten.de/fileadmin/user\\_upload/Rechtsgrundlagen/Gesetze\\_Staatsverträge/Medienstaatsvertrag\\_MStV.pdf](https://www.die-medienanstalten.de/fileadmin/user_upload/Rechtsgrundlagen/Gesetze_Staatsverträge/Medienstaatsvertrag_MStV.pdf) [Zugriff 24.08.2023].

- Milmo, Dan: Rohingya sue Facebook for £150bn over Myanmar genocide (06.12.2021). In: The Guardian; online unter: <https://www.theguardian.com/technology/2021/dec/06/rohingya-sue-facebook-myanmar-genocide-us-uk-legal-action-social-media-violence> [Zugriff 24.08.2023].
- Moeller, Judith u.a.: Shrinking core? Exploring the differential agenda setting power of traditional and personalized news media. In: Info 18 (2016), H. 6, S. 26–41.
- Muhle, Florian/Wehner, Josef: Algorithmus und Entscheidung. Anspruch und Empirie personalisierter Medienangebote im Internet. In: Conradi, Tobias/Hoof, Florian/Nohr, Rolf F. (Hg.): Medien der Entscheidung. Münster 2016, S. 111–130; online unter: [https://mediarep.org/bitstream/handle/doc/17027/Medien-der-Entscheidung\\_111-130\\_Muhle\\_Wehner\\_Algorithmus\\_und\\_Entscheidung\\_.pdf?sequence=2&isAllowed=y](https://mediarep.org/bitstream/handle/doc/17027/Medien-der-Entscheidung_111-130_Muhle_Wehner_Algorithmus_und_Entscheidung_.pdf?sequence=2&isAllowed=y) [Zugriff 01.08.2023].
- Münch, Richard: Polarisierte Gesellschaft. Die postmodernen Kämpfe um Identität und Teilhabe. Frankfurt/New York: Campus 2023.
- Myers, David G./Lamm, Helmut: The group polarization phenomenon. Psychological Bulletin, 83 (1976), H. 4, S. 602–627.
- Napoli, Philip M.: Automated media: An institutional theory perspective on algorithmic media production and consumption. In: Communication Theory 24 (2014), H. 3, 340–360; online unter: <https://doi.org/10.1111/comt.12039> [Zugriff 01.08.2023].
- Napoli, Philip M.: Exposure diversity reconsidered. In: Journal of information policy 1 (2011), S. 246–259; online unter: [https://www.researchgate.net/publication/265445493\\_Exposure\\_Diversity\\_Reconsidered](https://www.researchgate.net/publication/265445493_Exposure_Diversity_Reconsidered) [Zugriff 01.08.2023].
- Narayanan, Vidya u.a.: Polarization, Partisanship and Junk News Consumption over Social Media in the US. 2018; online unter: <https://arxiv.org/ftp/arxiv/papers/1803/1803.01845.pdf> [Zugriff 01.08.2023].
- Nematzadeh, Aida/Meylan, Stephan C./Griffiths, Thomas L.: Evaluating Vector-Space Models of Word Representation, or, The Unreasonable Effectiveness of Counting Words Near Other Words. In: Proceedings of the 39th annual meeting of the cognitive science society (2017), S. 859–864.
- Neudert, Lisa-Maria: Computational Propaganda in Germany: A Cautionary Tale. In: Woolley, Samuel C./Howard, Philip N. (Hg.): Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media. Oxford 2017, S. 153–184.
- Nguyen, C. Thi: Echo chambers and epistemic bubbles. Episteme 17 (2017), H. 2, S. 141–161.
- Nguyen, C. Thi: Escape the echo chamber (09.04.2018). In: Aeon; online unter: <https://aeon.co/essays/why-its-as-hard-to-escape-an-echo-chamber-as-it-is-to-flee-a-cult> [Zugriff 30.08.2023].
- Nir, Eyal/Hoover, Ryan: Hooked. How to Build Habit-Forming Products. New York 2014; online unter: <https://passionatethinker.com/files/Hooked.pdf> [Zugriff 01.08.2023].

- O'Callaghan, Derek u.a.: Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems. *Social Science Computer Review*, 33 (2015), H. 4, S. 459–478; online unter: <https://doi.org/10.1177/0894439314555329> [Zugriff 24.08.2023].
- Oertel, Britta u.a.: Algorithmen in Digitalen Medien und ihr Einfluss auf die Meinungsbildung. Endbericht zum TA-Projekt. Berlin 2022; online unter: <https://publikationen.bibliothek.kit.edu/1000154065> [Zugriff 01.08.2023].
- Pariser, Eli: The Filter Bubble. How the New Personalized Web Is Changing What We Read and How We Think. New York 2014; online unter: <https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=6057304> [Zugriff 01.08.2023].
- Pasquale, Frank: Platform Neutrality. Enhancing Freedom of Expression in Spheres of Private Power. In: Theoretical Inquiries in Law (2016), H. 2, S. 487–514; online unter: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2779270](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2779270) [Zugriff 01.08.2023].
- Persily, Nathaniel/Tucker, Joshua A.: Introduction. In: Dies. (Hg.): Social Media and Democracy. The State of the Field and Prospects for Reform. Cambridge 2020, S. 1–9; online unter: <https://www.cambridge.org/core/books/social-media-and-democracy/introduction/E358AE65D98725BDCDBD5324592BE687> [Zugriff 24.08.2023].
- Pöchlhacker, Nikolaus u.a.: Interventionen in die Produktion algorithmischer Öffentlichkeiten: Recommender Systeme als Herausforderung für öffentlich-rechtliche Sendeanstalten. In: Kommunikation@gesellschaft 18 (2017), H. 2, S. 1–25; online unter: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-51500-9> [Zugriff 01.08.2023].
- Rhein, Valerie/Dreyer, Stephan/Schulz, Wolfgang: Rechtliche Vorgaben für die Gestaltung von Software öffentlich-rechtlicher Medienplattformen. Gesetzliche und verfassungsrechtliche Programmaufträge und deren Abbildbarkeit in Strukturen, Verfahren und Code. Hamburg 2021.
- Ribeiro, Manoel u.a.: Characterizing and detecting hateful users on twitter. In: Proceedings of the International AAAI Conference on Web and Social Media 12 (2018), H. 1.
- RTBF/Ministère de la Communauté française: Sixième Contrat de gestion de la RTBF 2023-2027; online unter: [https://ds1.static.rtbf.be/uploader/pdf/a/5/0/beta\\_d537219571cfe9f6dcb782c4c4d239dd.pdf](https://ds1.static.rtbf.be/uploader/pdf/a/5/0/beta_d537219571cfe9f6dcb782c4c4d239dd.pdf) [Zugriff 24.08.2023].
- Rupar, Aaron: Facebook's controversial fact-checking partnership with a Daily Caller-funded website, explained (06.05.2019). In: Vox; online unter: <https://www.vox.com/2019/5/2/18522758/facebook-fact-checking-partnership-daily-caller> [Zugriff 01.08.2023].
- Sängerlaub, Alexander/Meier, Miriam/Rühl, Wolf-Dieter: Fakten statt Fakes. Verursacher, Verbreitungswege und Wirkungen von Fake News im Bundestagswahlkampf 2017. Berlin 2018; online unter: [https://www.stiftung-nv.de/sites/default/files/snv\\_fakten\\_statt\\_fakes.pdf](https://www.stiftung-nv.de/sites/default/files/snv_fakten_statt_fakes.pdf) [Zugriff 01.08.2023].

- Scharkowa, Michael u.a.: How social network sites and other online intermediaries increase exposure to news. In: Proceedings of the National Academy of Sciences (PNAS) 117 (20202), H. 6, S. 2761–2763; online unter: <https://www.pnas.org/doi/pdf/10.1073/pnas.1918279117> [Zugriff 24.08.2023].
- Schmidt, Jan-Hinrik: Die Medienlogik der Informationsintermediäre und ihre Bedeutung für Meinungsbildung. In: Bundeszentrale für politische Bildung (Hg.): Dossier Digitale Desinformation. 2019, online unter: <https://www.bpb.de/290575> [Zugriff 24.08.2023].
- Schmidt, Jan-Hinrik: Soziale Medien als Intermediäre in der Wissenschaftskommunikation. Hamburg 2016; online unter: [http://www.schmidtmitdete.de/pdf/Expertise\\_Schmidt\\_Soziale\\_Medien\\_Wissenschaftskommunikation\\_Entwurf.pdf](http://www.schmidtmitdete.de/pdf/Expertise_Schmidt_Soziale_Medien_Wissenschaftskommunikation_Entwurf.pdf) [Zugriff 01.08.2023].
- Schmitt, Josephine u.a.: Counter-messages as Prevention or Promotion of Extremism? The Potential Role of YouTube: Recommendation Algorithms. In: Journal of Communication 68 (2018), H. 4, S. 780–808; online unter: <https://doi.org/10.1093/joc/jqy029> [Zugriff 24.08.2023].
- Schwartz, Paul M.: Privacy and Democracy in Cyberspace, in: Vanderbilt Law Review 52 (1999), H. 6, S. 1607–1702; online unter: <https://scholarship.law.vanderbilt.edu/vlr/vol52/iss6/2> [Zugriff 24.08.2023].
- Seaver, Nick.: Captivating algorithms: Recommender systems as traps. In: Journal of Material Culture 24 (2019), H. 4, S. 421–436; online unter: <https://doi.org/10.1177/1359183518820366> [Zugriff 24.08.2023].
- Shao, Chengcheng u.a.: The spread of low-credibility content by social bots. In: Nature communications 9 (2018), H. 1, S. 1–9.
- Singhal, Amit: A Flawed Elections Conspiracy Theory (26.08.2015). In: Politico Magazine; online unter: <https://www.politico.com/magazine/story/2015/08/google-2016-election-121766/> [Zugriff 24.08.2023].
- Smith, Jessie J./Jayne, Lucia/Burke, Robin: Recommender Systems and Algorithmic Hate. In Proceedings of the 16th ACM Conference on Recommender Systems (RecSys '22). Association for Computing Machinery. New York 2022, S. 592–597; online unter: <https://doi.org/10.1145/3523227.3551480> [Zugriff 24.08.2023].
- Srnicek, Nick: Platform Capitalism. Cambridge 2016; online unter: <https://mudancatecnologicaedinamicacapitalista.files.wordpress.com/2019/02/platform-capitalism.pdf> [Zugriff 24.08.2023].
- Stark, Birgit/Magin, Melanie/Jürgens, Pascal: Ganz meine Meinung? Informationsintermediäre und Meinungsbildung – Eine Mehrmethodenstudie am Beispiel Facebook. Eine Expertise der Landesanstalt für Medien Nordrhein-Westfalen (LfM) in Kooperation mit dem Forschungsschwerpunkt Medienkonvergenz der Johannes Gutenberg-Universität Mainz. Düsseldorf 2017; online unter: [https://www.medienanstalt-nrw.de/fileadmin/user\\_upload/lfm-nrw/Foerderung/Forschung/Dateien\\_Forschung/L194-Ganz-meine-Meinung\\_LfM-Doku55.pdf](https://www.medienanstalt-nrw.de/fileadmin/user_upload/lfm-nrw/Foerderung/Forschung/Dateien_Forschung/L194-Ganz-meine-Meinung_LfM-Doku55.pdf) [Zugriff 01.08.2023].

- Stieglitz, Stefan/Dang-Xuan, Linh.: Political Communication and Influence through Microblogging – An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior. Maui 2012, S. 3500-3509; online unter: <https://doi.org/10.1109/HICSS.2012.476> [Zugriff 01.08.2023].
- Sunstein, Cass R./Thaler, Richard: Designing better choices (02.04.2018). In: Los Angeles Times; online unter: <https://www.latimes.com/archives/la-xpm-2008-apr-02-oe-thalerandsunstein2-story.html> [Zugriff 24.08.2023].
- Sunstein, Cass: Republic.com. Princeton 2001.
- Susskind, Jamie: The Digital Republic. On Freedom and Democracy in the 21<sup>st</sup> Century. New York 2022.
- Thaler, Richard H./Sunstein, Cass R.: Nudge. The Final Edition. London 2021.
- Thielman, Sam: Facebook fires trending team, and algorithm without humans goes crazy (29.08.2016). In: The Guardian; online unter: <https://www.theguardian.com/technology/2016/aug/29/facebook-fires-trending-topics-team-algorithm>.
- van der Linden, Sander: Foolproof. Why we fall for misinformation and how to build immunity, London 2023.
- van Dijck, José/Poell, Thomas/de Waal, Martijn (Hg.): The Platform Society. Public Values in a Connective World. Oxford 2018.
- Voigt, Mario/Seidenglanz, René: Digital Campaigning in der Bundestagswahl 2017 – Implikationen für Politik und Public Affairs. Berlin 2017; online unter: <https://www.medianet-bb.de/wp-content/uploads/2018/01/quadriga-digital-campaigning-studie-btw2017.pdf> [Zugriff 01.08.2023].
- Vosoughi, Soroush/Roy, Deb/Aral, Sinan: The spread of true and false news online. In: Science 359 (2018), H. 6380, S. 1146–1151; online unter: <https://www.science.org/doi/10.1126/science.aap9559> [Zugriff 24.08.2023].
- Whitten-Woodring, Jenifer u.a.: Poison If You Don't Know How to Use It: Facebook, Democracy, and Human Rights in Myanmar. The International Journal of Press/Politics, 25 (2020), H. 3, S. 407–425; online unter: <https://doi.org/10.1177/1940161220919666> [Zugriff 24.08.2023].
- Wilkes, Albert/Leatherbarrow, Margaret: Editing episodic memory following the identification of error. In: The Quarterly Journal of Experimental Psychology 40 (1988), H. 2, S. 361–387; online unter: <https://doi.org/10.1080/02724988843000168> [Zugriff 01.08.2023].
- Wittenberg, Chloe/Berinsky, Adam J.: Misinformation and Its Correction. In: Persily, Nathaniel/Tucker, Joshua A. (Hg.): Social Media and Democracy. The State of the Field and Prospects for Reform. Cambridge 2020, S. 163–198; online unter: <https://www.cambridge.org/core/books/social-media-and-democracy/misinformation-and-its-correction/61FA7FD743784A723BA234533012E810> [Zugriff 24.08.2023].
- Wollebæk, Dag u.a.: Anger, Fear, and Echo Chambers: The Emotional Basis for Online Behavior. Social Media and Society 5 (2019), H. 2; online unter: <https://doi.org/10.1177/2056305119829859> [Zugriff 24.08.2023].

- Ziarani, Reza Jafari/Ravanmehr, Reza: Serendipity in Recommender Systems: A Systematic Literature Review. In: Journal of Computer Science and Technology 36 (2021), S. 375–396; online unter: <https://doi.org/10.1007/s11390-020-0135-9> [Zugriff 24.08.2023].
- Zuboff, Shoshana: The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power. New York 2019.
- Zuckerman, Ethan: The Case for Digital Public Infrastructure (17.01.2020). In: Knight First Amendment Institute; online unter: <https://knightcolumbia.org/content/the-case-for-digital-public-infrastructure> [Zugriff 01.08.2023].

# Glossar

## Affektive Polarisierung

Polarisierung auf der Ebene von Gefühl und emotionalem Engagement

## Akteur

Mensch oder Entität, der oder die z.B. in Sozialen Netzwerken handelnd in Erscheinung tritt

## Algorithmus/algorithmisch

Schematische Rechenvorschrift/einem schematischen Rechenvorgang, einer schrittweisen Zeichenreihe folgend

## Ambiguitätstoleranz

Akzeptanz und Toleranz gegenüber unsicheren oder widersprüchlichen Situationen

## Anti-recommender-System

Vorhersageprogramm auf Basis von individuell personalisierten Nutzer:innendaten, dessen Empfehlungen gegen (!) die bisherigen User:innen-Interessen gehen

## Aufmerksamkeitsökonomie

Ein wirtschaftliches System, das über die Erregung von Aufmerksamkeit die geldwerte Bildschirmzeit zu maximieren versucht

## Backfiring

,Fehlzündung‘, „Rohrkrepierer‘; Effekt der psychologischen Abwehrmechanismen gegen die Korrektur von Fake News: Die faktenbasierte Korrektur von fehlerhaften Überzeugungen kann diese sogar noch verstärken ( $\uparrow$  World view backfire effect)

## Bias

Vorurteil, Voreingenommenheit, Verzerrung

## Big Data

Extrem große, von  $\uparrow$  Digitalplattformen generierte Datenmenge; metonymisch auch die Digitalplattform-Unternehmen selbst

## Black box

Ein Medium oder medialer Prozess, dessen innerer Ablauf unbeobachtbar ist und bei dem sich nur Input und Output beobachten lassen

## Bot

Automatisiertes Programm

## **Captology**

Erforschung der Möglichkeiten, Überzeugungen und Verhalten menschlicher Nutzer:innen durch das Design digitaler Angebote zu beeinflussen

## **Choice architecture**

Suggestive Anordnung von Wahlmöglichkeiten, die die zu tätigende Wahl durch das Design zu beeinflussen sucht

## **Confirmation bias**

Psychologische Tendenz, die Dinge so zu interpretieren, dass sie die eigenen Vorurteile ( $\uparrow$  Bias) bestätigen

## **Content Diversity**

Vielfalt von angebotenen Inhalten, um der Selbstverstärkung in epistemischen Blasen entgegenzuwirken und diese so aufzubrechen

## **Continued influence effect**

Ein psychologischer Effekt, der ähnlich wie das  $\uparrow$  Backfiring dafür sorgt, dass einmal gefasste Fehlmeinungen trotz erfolgter Korrektur fortbestehen

## **Counterspeech**

,Gegenrede‘; Nachrichten, Meldungen oder Informationen, die auf die Korrektur von Fehlinformationen abzielen

## **Desinformation**

Bewusste falsche Meldung/Information

## **Digitalplattformen**

Die digitale Infrastruktur einiger weniger gigantischer Unternehmen wie Google, Apple, Facebook, Amazon und Microsoft, deren Marktkapitalisierung 2019 4,5 Billionen € betrug (BIP der Bundesrepublik Deutschland 2019: 3,4 Billionen €)

## **Disclaimer**

Erklärung der bewussten Distanzierung von bestimmten Inhalten

## **Disconfirmation bias**

Bias  $\uparrow$  mit der Tendenz, Argumente gegen Informationen zu formulieren, die den eigenen Überzeugungen widersprechen

## **Echokammer**

Soziale Struktur auf Digitalplattformen, die andere Meinungen und Perspektiven ignoriert oder aktiv diskreditiert

**Enclave deliberation**

Zerfall des deliberativen Diskurses einer Gesellschaft in Unter- und Splittergruppen, gefördert durch das Entstehen von Filterblasen in Sozialen Netzwerken

**Fake News**

Falschmeldungen und -nachrichten in digitalen Medien und Sozialen Netzwerken mit manipulativer Tendenz

**Falschnachricht**

Nicht zutreffende Meldung oder Information, die über das Internet oder über Soziale Medien verbreitet wird

**Filter Bubble**

Meinungsblase, die algorithmisch, das heißt automatisch, durch Inhaltsselektion aufgrund des User:innenverhaltens auf Webseiten, Diensten usw. erzeugt wird

**Filterblase**

↑ Filter Bubble

**Hatespeech**

Hassnachrichten und hasserfüllte Rede

**Human in the loop**

Beteiligung menschlicher Akteure an der Entwicklung, Implementierung und Weiterentwicklung von ↑ Recommender Systems

**Influencer:in**

Aufgrund hoher Reichweite einflussreiche, trendsetzende Person in Sozialen Netzwerken

**Informationsintermediär**

Automatisierte informationvermittelnde und -strukturierende Plattform wie Suchmaschinen, Soziale Medien, Foto-, Video-, Microbloggingplattformen und Nachrichtenaggregatoren

**Klassische Medien**

Elemente/Trägersysteme zur Vermittlung von Informationen und anderen Inhalten; „Einrichtung, organisatorischer und technischer Apparat für die Vermittlung von Meinungen, Informationen, Kulturgütern; eines der Massenmedien Film, Funk, Fernsehen, Presse“ (Duden Universalwörterbuch 2023)

**Künstliche Intelligenz**

Nicht-menschliche, von Computern erzeugte Intelligenz, deren Output bis zur Ununterscheidbarkeit dem menschlichen Intelligenz ähneln kann

## **Like**

Positive Online-Bewertung

## **Microblogging**

Kurzes Blogging bzw. kleine Beiträge der schnellen, direkten User:innen-Interaktion

## **Motivated reasoning**

Nicht ergebnisoffenes, sondern ‚zielgerichtetes Denken‘, das durch das soziale Umfeld bestimmt wird

## **Network of trust**

Eine soziale Struktur von Akteuren, die einander kennen, gemeinsame Positionen teilen und einander deshalb Vertrauen entgegenbringen

## **Neuronales Netzwerk**

Computernetzwerk, dessen Architektur und Funktion sich am biologischen menschlichen Gehirn orientiert

## **News exposure**

Nachrichtenangebot

## **Nudging**

Verhaltenspsychologische Motivationstechnik, durch die das Verhalten von User:innen in einer gewünschten Weise beeinflusst werden soll

## **Particular encounter**

Einflussreiche Einzelbegegnung in der Interaktion von Nutzer:in zu Nutzer:in, durch die eine Person aus einer Echokammer befreit werden kann, weil sie zu einer Person außerhalb dieser Echokammer schon Vertrauen hat oder es neu fasst

## **Rabbit hole**

*hier:* Destruktive Fixierung und Verengung auf bestimmte emotionalisierende Themen aufgrund obsessiver Nutzung von sozialen Netzwerken; vgl. ↑ Filter Bubble

## **Recommender System**

Vorhersage-Programm auf Basis von individuell personalisierten Nutzerdaten

## **Redaktionelle Medien**

Medien, die von einer Redaktion betreut („kuratiert“) und verantwortet werden

## **Retweeten, Reposten**

Erneutes Posten einer auf einer Online-Plattform publizierten Nachricht („Tweet“)

## **Search engine manipulation effect**

Effekt, der entsteht, wenn durch Manipulation von Suchergebnissen maßgeblich auf die Entscheidungen von User:innen Einfluss genommen wird

## **Serendipity**

hier: Ein ‚glücklicher Zufall‘ im Informations- oder sonstigen inhaltlichen Angebot; ‚Zufällige‘, ‚neue‘, ‚überraschende‘ Inhalte, die zur Kenntnis genommen werden, können so zur Auflösung von epistemischen Blasen führen

## **Social Media**

Soziale Medien; Digitaltechnologien und Medien von Online-Kommunikation und -Informationsaustausch zwischen User:innen

## **Social reinforcement**

Meinungsverstärkung durch das soziale Umfeld

## **Soziale Netzwerke**

↑ Social Media

## **Strong tie**

Nahe Vertrauensperson mit grundsätzlich sehr ähnlichem Weltbild (Familie, Freunde), siehe auch ↑ Weak tie

## **Suchhistorie**

Chronologischer Verlauf der aufgesuchten Internetseiten bzw. getätigten Suchanfragen

## **Surveillance Capitalism**

Geschäftsmodell der präzisen Analyse und vermarktbaren Vorhersage von User:innen-verhalten mithilfe von riesigen Datenmengen (↑ Big Data), die die ↑ Digitalplattformen generieren

## **Taggen**

Markieren und Strukturieren digitaler Inhalte mithilfe eines Tags („Schildchen“)

## **Timeline**

Darstellung von Inhalten in chronologischer Abfolge

## **Trade-off**

Kompromiss zwischen zwei widersprüchlichen, nicht miteinander vereinbaren Positionen (Dilemma)

## **Trending topic**

Aktuell besonders populäres Gesprächsthema

## **Verschwörungstheorie**

Unrealistische oder falsche Annahme, dass die Verschwörung einiger weniger Akteure ursächlich für ein Phänomen ist, das viele Menschen betrifft

## **Weak tie**

Entfernt bekannte Person, zu der ein – möglicherweise nur schwaches – Vertrauensverhältnis besteht und deren Weltbild meist vom eigenen abweicht

## **Worldview backfire effect**

Psychologischer Mechanismus der Stärkung falscher Überzeugungen gerade durch ihre Widerlegung (↑ Backfiring)