

*Florian Saurwein, Charlotte Spencer-Smith und Jaro Krieger-Lamina*

## **Social-Media-Algorithmen als Gefahr für Öffentlichkeit und Demokratie: Anwendungen, Risikoassemblagen und Verantwortungszuschreibungen**

### 1. Einleitung

Internet-Plattformen wie Facebook und YouTube haben in den vergangenen Jahren massiv an Verbreitung gewonnen, wodurch auch ihre Bedeutung für politische Kommunikation stark zugenommen hat. Die Plattformen eröffnen einerseits Möglichkeiten für mehr Information, Transparenz, Kritik und politische Beteiligung. Der Trend zur Plattformisierung führt jedoch gleichzeitig zu Risiken für Öffentlichkeit und Demokratie. Dazu zählen die Fragmentierung von Öffentlichkeit durch Personalisierung von Informationsangeboten, die Verbreitung von Desinformation, automatisierte Propaganda und der Verlust an Privatsphäre durch permanente Überwachung des Nutzungsverhaltens. Im öffentlichen Diskurs erhalten die problematischen Begleiterscheinungen der Plattformisierung viel Aufmerksamkeit und sie werden mitunter auf den Einsatz von Algorithmen auf Internet-Plattformen zurückgeführt.

Der folgende Beitrag<sup>1</sup> beschreibt die Einsatzbereiche von Algorithmen auf Social-Media-Plattformen sowie die Risiken für Öffentlichkeit und Demokratie. Dabei wird herausgearbeitet, dass Algorithmen aus unterschiedlichen Gründen zur Entstehung einzelner Risiken beitragen, z.B. durch Fehlerhaftigkeit, Verstärkungseffekte, Manipulation, als Opportunitätsstruktur für gefährliche Praktiken und als Faktor für die Etablierung von Machtasymmetrien. Meist lassen sich Risiken jedoch nicht ausschließlich auf den Einsatz von Algorithmen zurückführen, sondern auf ein Zusammenwirken verschiedener Einflussfaktoren wie regulatorische Rah-

---

1 Der vorliegende Beitrag fasst zentrale Ergebnisse aus Forschungsprojekten zusammen, die vom Jubiläumsfonds der Stadt Wien für die Österreichische Akademie der Wissenschaften und im Forschungsprogramm goldigital Next Generation der Österreichischen Akademie der Wissenschaften unterstützt werden.

menbedingungen, Geschäftsinteressen der Plattformbetreiber, Plattformdesign und Nutzungsverhalten. Entsprechend lassen sich Risiken algorithmischer Selektion auf Internet-Plattformen am besten als „sozio-technische Assemblagen“ beschreiben, um involvierte Akteur\*innen, Technologien und vielfältige Wirkungszusammenhänge angemessen zu erfassen. Eine solche Beschreibung wiederum eröffnet Perspektiven für Fragen nach der Verantwortlichkeit für Risiken, die im Zusammenspiel von Menschen und Technik entstehen und sie ermöglicht eine Differenzierung der Zurechnungspunkte für Verantwortlichkeit (Plattformunternehmen, Entwickler\*innen, Nutzer\*innen, Politik/Regulierung, etc.), die gemeinsam ein Verantwortungsnetzwerk bilden. Der Beitrag zeigt abschließend, welche der involvierten Akteursgruppen und Technologien im Mediendiskurs über die Risiken sozialer Netzwerke Verantwortung zugewiesen wird.

## 2. Einsatz von Algorithmen auf Social-Media-Plattformen

Social-Media-Plattformen haben im letzten Jahrzehnt massiv an Bedeutung für gesellschaftliche Information und Kommunikation gewonnen. Ihre Anziehungskraft liegt in ihren „sozialen“ Funktionen, den zwischenmenschlichen Kontakten, der Veröffentlichung von Information und dem Austausch von Meinungen. Die Plattformen sind aber auch durch „automatisierte“ Funktionen geprägt. Programmierung und Algorithmen sind mitbestimmend dafür, wie Information, Kommunikation und Transaktionen in sozialen Netzwerken erfolgen. „Algorithmische Selektion“ (Latzer et al. 2016) ermöglicht eine automatisierte Auswahl von Informationen aus großen Datenmengen sowie deren Strukturierung, Ordnung und Sortierung durch Relevanzzuweisung. Das Ergebnis solcher Prozesse reicht von Reihungen über Empfehlungen bis zu automatisierten Transaktionen im Hochfrequenzhandel. Zu den zentralen algorithmisch-selektiven Anwendungen in sozialen Netzwerken zählen die Kuratierung von Nachrichten in Newsfeeds, Filteranwendungen zur Reduktion von unerwünschten Inhalten und automatisierte Werbung.

*Kuratierung, Suche und Empfehlungssysteme:* Auf Social-Media-Plattformen werden neue Inhalte mit enormer Geschwindigkeit geladen und verbreitet. Die große Menge an Inhalten erfordert eine Kuratierung, wenn Inhalte nicht nur chronologisch angezeigt werden sollen. Algorithmen bewerten deshalb die Relevanz der Inhalte für einzelne Nutzer\*innen und verwenden dafür persönliche Informationen wie demografische Daten, Interessen, Verbindungen und Nutzungsverhalten (vgl. Cobbe/Singh 2019). Zweck der Kuratierung ist die Auswahl und Zusammenstellung von Inhalten entsprechend der persönlichen Interessen der Nutzer\*innen. Dadurch

sollen Nutzer\*innen die Zeit, die sie auf den Plattformen verbringen, als interessant und nützlich erleben. Positive Nutzungserlebnisse sollen zur Steigerung von Nutzungszeit und Interaktionen und damit zum Erfolg der Plattformen beitragen. Auch die Suchapplikationen, Auto-Complete-Funktionen und die Zusammenstellung der Trending Topics basieren auf Algorithmen und sollen Nutzer\*innen bei der Navigation durch die Fülle an Inhalten unterstützen. Zudem analysieren Algorithmen bestehende Vernetzungen und Interessen und empfehlen, mit welchen Personen und Gruppen sich Nutzer\*innen verbinden könnten. Durch die Kuratierungs- und Empfehlungssysteme wird die Aufmerksamkeit gelenkt und Plattformen erhalten entsprechenden Einfluss auf gesellschaftliche Kommunikation.

*Inhaltsmoderation und -filterung:* Die Etablierung von Social-Media-Plattformen ermöglicht für alle Nutzer\*innen einen niederschweligen Zugang zu Öffentlichkeit. Dies führt auch zur leichteren Verbreitung von illegalen und gefährlichen Inhalten wie Kindesmissbrauchsdarstellungen, Terrorismus, Hass und Cybermobbing und zur Verletzung von Urheberrechten. Internet-Plattformen etablieren komplexe Systeme zur Inhaltsregulierung, um problematische Inhalte zu identifizieren, zu filtern, zu löschen, zu drosseln oder zu kennzeichnen (Gillespie 2018). Die Inhaltsregulierung basiert auf plattformeigenen Gemeinschaftsregeln und rechtlichen Vorgaben und wird durch einen Mix aus technischer und menschlicher Kontrolle umgesetzt (vgl. Saurwein/Spencer-Smith 2019). Auch dabei kommen automatisierte Filtersysteme zum Einsatz, die mittels digitaler Fingerabdrücke und Mustererkennung operieren (Gorwa et al. 2020). Diese können eine große Menge an Inhalten rasch und vor ihrer Veröffentlichung überprüfen und dabei helfen, Nutzer\*innen ein sicheres Umfeld zu bieten. Durch hohe Effizienz und Skalierbarkeit werden automatisierte Systeme von den Plattformen gerne als Lösung im Kampf gegen problematische Inhalte propagiert. Sie werden jedoch aufgrund der Etablierung einer privaten Ordnung für gesellschaftliche Kommunikation und Beschränkungen der Äußerungsfreiheit auch kritisiert.

*Automatisierte Werbung:* Da Social-Media-Plattformen keine Nutzungsentgelte erheben, spielt Werbung in ihrem Geschäftsmodell eine zentrale Rolle. Die sozialen Netzwerke verfügen zum Teil über immens hohe Nutzungszahlen und anders als bei klassischer Print- oder TV-Werbung können Werbetreibende auf den Plattformen nicht nur Zeit und Kontext für Werbung bestimmen, sondern durch fein granulierte Auswahlprozesse (Micro-Targeting) auch Merkmale des Zielpublikums (Busch 2016). Selektionskriterien umfassen demografische Daten wie Alter, Bildung, Geschlecht, Herkunft oder Einkommen ebenso wie Interessen, Nutzungsverhalten und Kontextinformationen bezüglich der Umgebung, in der Nutzer\*innen ihre

Geräte verwenden. Die Entscheidungen zur Platzierung von Werbung erfolgen innerhalb weniger Millisekunden im Rahmen von komplexen, algorithmen-basierten Auktionen. Diese bieten den Werbetreibenden eine Reihe von Vorteilen wie die automatisierte Durchführung, höhere Genauigkeit bei der Zielgruppenauswahl, geringere Streuverluste und bessere Erfolgskontrolle durch umfassende Nutzungsdaten. Allerdings basieren die Werbeformen auf einer massiven Sammlung und Verwertung von Nutzungsdaten, die zu Eingriffen in die Privatsphäre führen und einen wesentlichen Treiber des Überwachungskapitalismus (Zuboff 2018) darstellen.

### 3. Algorithmische Risiken und Implikationen für Öffentlichkeit und Demokratie

Während die Beispiele zu Kuratierung, Filterung und Werbung auf den Nutzen von algorithmischer Selektion verweisen, gehen mit den Anwendungen auch Risiken einher. Saurwein und Spencer-Smith (2021) identifizieren fünf Risikotypen, die sich danach unterscheiden, welche Rolle Algorithmen in der Entstehung der Risiken spielen: (1) Fehler bzw. unerwünschte Selektionen durch Algorithmen, (2) Manipulation der Ergebnisse von Algorithmen durch Nutzer\*innen, (3) algorithmische Verstärkung bestehender gesellschaftlicher Probleme, (4) Algorithmen als Opportunitätsstruktur für gefährliche Praktiken und (5) Algorithmen als Faktor für Machtasymmetrien zwischen Plattformen und anderen Stakeholdergruppen.

#### Fehler und unerwünschte Selektion

Obwohl Algorithmen aus technischer Sicht nur vorgegebene Schritte umsetzen und keine „Fehler“ machen, können sie Ergebnisse produzieren, die den Intentionen der Entwickler\*innen zuwiderlaufen. Solche unerwünschten Selektionen passieren meist, wenn Algorithmen als technische Werkzeuge Pixel, Farben, Text und Metadaten beurteilen, aber den Kontext und die soziale Bedeutung der Inhalte nicht verstehen (vgl. Gillespie 2018). Dadurch kann es z.B. zu einem Overblocking kommen, bei dem Algorithmen „irrtümlich“ zu viele Inhalte löschen (vgl. Holznagel 2018) und die Äußerungsfreiheit einschränken. Befürchtet wird z.B., dass der Einsatz von Upload-Filtern zum Schutz von Urheberrechten zu Beschränkungen der kulturellen Kreativität führen kann (vgl. Brøvig-Hanssen/Jones 2021). Für Öffentlichkeit und Demokratie besteht das Risiko eines Overblockings für Satire (vs. Beleidigung) oder Kriegsberichte (vs. terroristische Propaganda), deren Legitimität vom jeweiligen Kontext abhängt (vgl. Duarte et al. 2018). Zudem behindert die automatisierte Filterung von problematischen Inhal-

ten eine gesellschaftliche Auseinandersetzung mit den zugrunde liegenden Problemen.

## **Manipulation durch Nutzer\*innen**

Nutzer\*innen von Plattformen können die Funktionsweisen von Algorithmen ausnutzen, um Ergebnisse der Selektionen zu manipulieren. Ähnlich wie bei Suchmaschinenoptimierung (SEO) können Social-Media-Algorithmen dazu gebracht werden, bestimmte Inhalte zu empfehlen (vgl. Shepherd 2020), auch um andere Nutzer\*innen zu belästigen, den öffentlichen Diskurs zu stören oder auf undemokratische Weise die öffentliche Meinung zu beeinflussen. So werden z.B. Botnets eingesetzt, um Inhalte automatisiert zu markieren und zu teilen und dadurch Popularität vorzutauschen. Aufgrund der Relevanzsignale sollen die Inhalte durch algorithmische Empfehlungssysteme noch prominenter platziert, weiteren Nutzer\*innenkreisen empfohlen und entsprechend stärker verbreitet werden. Dies kann öffentliche Diskurse verzerren. Bots wurden z.B. von der russischen Trollfabrik Internet Research Agency eingesetzt, um polarisierende Meinungen in den USA zu verbreiten (vgl. Osipova/Byrd 2017). Ähnliche Effekte lassen sich erzielen, wenn Gruppen von Nutzer\*innen koordiniert mit Inhalten interagieren oder gleiche Hashtags posten. Dadurch wird Beliebtheit vorgetäuscht, mit der z.B. „Trending Topics“ beeinflusst werden, in denen die aktuell beliebtesten Hashtags prominent dargestellt sind. So gelang es der rechtsextremen Reconquista Germanica durch koordiniertes Verhalten, Hashtags im öffentlichen Diskurs zu platzieren (vgl. Kreißel et al. 2018). Bei den verschiedenen angewendeten Taktiken stellt sich die Frage, wo legitime Optimierung endet und illegitime Manipulation beginnt.

## **Verstärkungseffekte**

Durch algorithmische Selektion kann es zu einer Verstärkung von existierenden gesellschaftlichen Problemen kommen. So erzeugen problematische Inhalte wie Hassäußerungen oder Desinformation besonders hohe Aufmerksamkeit, die von Algorithmen als Relevanzsignal gelesen wird. In der Folge werden genau diese problematischen Inhalte von Algorithmen weiteren Nutzer\*innen empfohlen und weiterverbreitet. Zu Verstärkungseffekten könnte es auch durch die Bildung von Filterblasen (Pariser 2011) und Echokammern (Sunstein 2001) kommen. Informationsintermediäre nutzen Algorithmen zu einer Personalisierung des Informationsangebots, das auf Nutzer\*inneninteressen abgestimmt ist. Dadurch droht eine schwächere Konfrontation mit gegenteiligen Meinungen, eine Fragmentierung von Öffentlichkeit und eine demokratiegefährdende gesellschaftliche Polari-

sierung. So wird z.B. dem YouTube-Algorithmus vorgeworfen, einen „Rabbit Hole“-Effekt zu erzeugen, mit dem Nutzer\*innen in einen Kaninchenbau immer extremerer Inhalte gelockt werden (Whittaker et al. 2021). Für Facebook wurde ermittelt, dass 64 % aller Beitritte zu extremistischen Facebook-Gruppen aus Empfehlungen des Facebook-Algorithmen resultierten (Horwitz/Seetharaman 2020). Zwar zeigen die Nutzungsmuster breiter Bevölkerungskreise aktuell keine Filterblasen (z.B. Dubois/Blank 2018). Eine Gefahr besteht jedoch für spezifische Nutzungsgruppen mit starker politischer Prädisposition, die sich in homogenen Netzwerken rund um besonders kontroverse Themen bilden (Stark et al. 2020).

### **Opportunitätsstruktur für gefährliche Praktiken**

Als technische Artefakte bilden algorithmische Selektionssysteme auch Strukturen, die Akteur\*innen bestimmte Praktiken ermöglichen, die problematisch für Öffentlichkeit und Demokratie sein können. Beispiele dafür findet man bei personalisierter Werbung, wo durch algorithmisches Micro-Targeting intransparente und diskriminierende Werbeschaltungen auf Social-Media-Plattformen ermöglicht werden. Durch den Einsatz von Algorithmen können nicht nur bestimmte Gruppen von Nutzer\*innen angesprochen, sondern auch ausgeschlossen werden. So ermöglicht Facebook in den USA eine Zielgruppenauswahl nach ethnischem Hintergrund oder politischer Einstellung. Dadurch kommt es zu diskriminierenden Werbepraktiken, z.B. bei Wohnungs- und Job-Anzeigen. Im politischen Micro-Targeting können Werbende intransparent Botschaften an unterschiedliche Zielgruppen schicken und dabei die Aufmerksamkeit eines breiteren Publikums und des Journalismus vermeiden (Zuiderveen Borgesius et al. 2018). Im US-Wahlkampf 2016 wurde die Werbefunktion von der Trump-Kampagne missbräuchlich verwendet, um afroamerikanische Männer vom Wählen abzuhalten (Green/Issenberg 2016).

### **Plattform-Macht**

Durch den Einsatz von Algorithmen wird die Macht der Plattformen über Nutzer\*innen, Wettbewerb und Märkte verstärkt. Algorithmen verschaffen Plattformen Macht gegenüber Nutzer\*innen, denn sie ermöglichen durch massive Datensammlung und -verarbeitung, Einblicke in aktuelles Nutzungsverhalten und Prognosen des zukünftigen Nutzungsverhaltens (Zuboff 2019). Die Macht der Algorithmen manifestiert sich auch in ihrem Einsatz zur Steigerung der Nutzungsintensität (Seymour 2019) und in der Kuratierung von Information, die darüber mitbestimmt, wie Nutzer\*innen die Welt wahrnehmen (Just/Latz 2017). Weiters sammeln Plattformunter-

nehmen Daten über Marktentwicklungen und Konkurrenzprodukte, um sich Wettbewerbsvorteile zu verschaffen. So nutzte Facebook die Smartphone-App Onavo Protect, um Daten zur Nutzung anderer Apps zu sammeln und Übernahmeentscheidungen vorzubereiten (vgl. Seetharaman/Morris 2017). Durch algorithmische Kuratierung geraten auch traditionelle Medien in eine steigende Abhängigkeit von sozialen Netzwerken. Änderungen an den Empfehlungsalgorithmen der Social-Media-Plattformen können sich erheblich auf den Umfang der Seitenaufrufe auswirken, die Medien über die Plattformen lukrieren. Zudem führt der Erfolg der automatisierten, verhaltensbasierten Werbung in sozialen Netzwerken zu Umsatzverlusten bei traditionellen Medien und zu Schwierigkeiten, demokratierelevanten Qualitätsjournalismus zu finanzieren (Lobigs 2016).

#### 4. Risiken als sozio-technische Assemblage

Social-Media-Algorithmen tragen also aus unterschiedlichen Gründen zur Entstehung von Risiken bei. Allerdings resultieren Risiken nicht ausschließlich aus dem Einsatz von Algorithmen, sondern aus dem Zusammenwirken verschiedener Einflussfaktoren, das am Beispiel der Verbreitung von Desinformation illustriert werden kann (vgl. Saurwein/Spencer-Smith 2020). Dabei agieren *Produzent\*innen* von Desinformation entweder aus kommerziellen Motiven, um Gewinne durch Werbung zu erzielen, die im Umfeld von Desinformation platziert wird, oder um aus politischen Motiven die Meinungsbildung zu beeinflussen und Unsicherheit zu streuen. Sie nutzen Fake Accounts und Botnets, die sich mit Nutzer\*innen verbinden und Inhalte automatisiert posten und teilen, um die Verbreitung zu steigern und Relevanz zu signalisieren. *Plattform-Nutzer\*innen* agieren als Adressat\*innen und Distributor\*innen von Desinformation. Das interaktive Design der Plattformen ermöglicht die einfache Weiterverbreitung von Desinformation wobei einzelne Influencer\*innen ein Millionenpublikum erreichen können. Durch Interaktionen mit Desinformation wird ein Interesse an Inhalten signalisiert, das von Algorithmen als Relevanzsignal gewertet und zu verstärkten Empfehlungen für weitere Nutzer\*innenkreise führt. Eine virale Verbreitung von problematischen Inhalten resultiert meist aus der Interaktion zwischen Nutzer\*innen und Algorithmen (Clegg 2021), die den Nutzer\*innen selbst vielfach gar nicht bewusst ist.

Auch das *Design der Plattformen und ihre technische Funktionsweise* fördern die Verbreitung von Desinformation. In Newsfeeds erscheinen Postings und Links zu Inhalten Dritter im selben Design, unabhängig davon, ob es sich um vertrauenswürdige Quellen oder Desinformation handelt. Plattformen bieten Interaktionsbuttons mit denen Inhalte einfach geteilt und



schnell verbreitet werden können. Nutzungsaktivität signalisiert Empfehlungsalgorithmen Relevanz, die in Newsfeed Kuratierung, Trending Topics und Auto-Play Funktionen einfließt. Algorithmen belohnen Postings, die zu Interaktionen animieren und fördern daher Inhalte, die emotionalisieren oder durch Skurrilität und Sensation gekennzeichnet sind und die Interaktionen steigern; Merkmale die vielfach auf Desinformation zutreffen (Vosoughi et al. 2018). Plattformdesigns, die Desinformation fördern, können auch auf *ökonomische Interessen der Plattformen* zurückgeführt werden. Plattformen wollen Nutzer\*innenkreise, Nutzungszeit und Interaktionen maximieren. Dies führt dazu, dass Nutzer\*innen vielfach ähnliche und zunehmend extremere Inhalte entsprechend ihren Interessen angezeigt werden, z.B. mit Bezug zu Verschwörungstheorien (Lewis 2018). Zudem zählen die erfolgreichen Geschäftsmodelle der Plattformen (Nutzung, Werbung) zu den Ursachen für die anhaltende Schwächung von traditionellen Medien und professionellem Journalismus, der Desinformation aufdecken und Richtigstellungen verbreiten kann.

Zusammenfassend zeigt die Analyse, dass die Verbreitung von Desinformation aus einem Mix aus Plattform-Architektur, Algorithmen, Nutzungsverhalten sowie politischen und kommerziellen Anreizstrukturen resultiert. Entsprechend lassen sich Risiken am besten als „sozio-technische Assemblagen“ beschreiben, um involvierte Akteursgruppen, Technologien und vielfältige Wirkungszusammenhänge angemessen zu erfassen. Darauf aufbauend lassen sich Zurechnungspunkte für Verantwortlichkeit differenzieren.

## 5. Mediale Attribution von Verantwortung

Die Vielfalt der involvierten Akteursgruppen und Technologien in der Entstehung von Risiken führt zur Frage nach der Verantwortung für Probleme und Lösungen. Dabei spielt natürlich die faktische Verantwortung eine zentrale Rolle, etwa in Form der rechtlichen Haftung für Schäden. Für viele Herausforderungen im Zusammenspiel zwischen Menschen und Technik sind die Verantwortungsfragen jedoch nicht abschließend geklärt. Und etliche aktuelle Risiken – wie die Auswirkungen von Internet-Plattformen auf öffentliche Güter wie Öffentlichkeit und Demokratie – sind im Rahmen zivilrechtlicher Haftungsansprüche auch überhaupt nicht zu fassen. In solchen Konstellationen der Unsicherheit spielt deshalb auch jene Verantwortlichkeit eine Rolle, die durch die Zuweisung von Verantwortung in öffentlichen Diskursen erfolgt. Speziell Medien und Journalismus tragen mit ihrer Berichterstattung über Anwendungen und Risiken von Technologien und durch die Zuschreibung von Verantwortung für Probleme und Lösungen zur gesellschaftlichen Konstruktion von Verantwortlichkeit bei.



Für den Mediendiskurs stellt sich deshalb die Frage, welchen Akteur\*innen in der Medienberichterstattung Verantwortung zugewiesen wird. Aus theoretischer Perspektive kommen dafür im Diskurs über algorithmische Risiken verschiedene Entitäten in Frage. Akteur\*innen und Organisationen, die Algorithmen entwickeln, implementieren, nutzen und regulieren, tragen eine Mitverantwortung für Prozesse, Ergebnisse und Folgen des Einsatzes von Algorithmen und bilden Bestandteile eines umfassenden Verantwortlichkeitsnetzwerks (Neuhäuser 2014; Loh/Loh 2017). Dazu zählen z.B. die Entwickler\*innen von Social-Media-Algorithmen, die Plattformunternehmen, welche die Technologien einsetzen, private und kommerzielle Plattformnutzer\*innen, sowie Akteur\*innen aus Industrie, Politik und Administration, die für die Kontrolle und Regulierung der Plattformen und ihrer technischen Funktionsweise verantwortlich sind (vgl. Saurwein 2018).

Im medialen Kontext stellt sich die Frage, welchen dieser Gruppen im Rahmen der Berichterstattung Verantwortung zugewiesen und wie damit *Verantwortlichkeit medial konstruiert* wird. Allerdings gibt es bislang nur wenige Untersuchungen, die sich mit der Medienberichterstattung über Algorithmen (z.B. Fischer/Puschmann 2021) und der medialen Attribution von Verantwortung beschäftigen. Eine Ausnahme bildet eine Fallstudie zur Berichterstattung über Microsoft's Chatbot „Tay“ (Suárez-Gonzalo et al. 2019), der 2016 für Aufsehen sorgte, weil er nach Interaktionen mit Nutzer\*innen rassistische und antisemitische Postings produzierte. Suárez-Gonzalo et al. (2019) analysieren, wem im Mediendiskurs die Schuld an den Fehlentwicklungen des Chatbots gegeben wurde. Dabei zeigt sich, dass sich 40 % der Verantwortungszuweisungen in der Berichterstattung an Twitter-Nutzer\*innen richteten, 17 % an Microsoft und 14 % an den Machine-Learning Code. 18 % analysierten die Defizite als Resultat der Interaktionen zwischen Menschen und der Software. Der Chatbot selbst wurde in der Berichterstattung vorwiegend als Opfer porträtiert. Suárez-Gonzalo et al. resümieren, dass für die Fehlentwicklungen v.a. die Twitter-Nutzer\*innen zur Verantwortung gezogen wurden, während die Medien es versäumten, das Zusammenspiel und die verschiedenen Verantwortlichkeiten von Designer\*innen, Nutzer\*innen und den Plattformbetreiber\*innen zu erklären.

In einer weiteren Untersuchung analysieren Brantner und Saurwein (2021) den Mediendiskurs zur Berichterstattung über Robotik, künstliche Intelligenz und Algorithmen in Österreich. Dabei zeigt sich, dass der größte Anteil an Verantwortungszuweisungen an die Wirtschaft erfolgt (45 %). Mit deutlichem Abstand folgen Verantwortungszuweisungen an die Politik (23 %) sowie generelle Attributionen an die Gesellschaft (15 %). Wissenschaft und Bildung (7 %) wird in etwa gleich oft Verantwortung zugewie-

sen wie den individuellen Nutzer\*innen. In einer Vertiefungsanalyse (Saurwein et al. forthcoming 2022) wird schließlich auch untersucht, welchen Akteursgruppen in den Medien Verantwortung zugewiesen wird, wenn *Social-Media-Algorithmen* Thema der Berichterstattung sind. Die Ergebnisse dieser Untersuchungen zeigen, dass die Internet-Plattformen mit 46 % aller Verantwortungsattributionen selbst im Zentrum der Verantwortungsdiskurses stehen. Plattformen wie Facebook, YouTube und Twitter fungieren somit nicht nur praktisch als zentrale Drehscheiben im Internet, sondern auch als die sichtbarsten Verantwortungssubjekte im Mediendiskurs. Obwohl der Schwerpunkt des medialen Verantwortungsdiskurses bei Internet-Plattformen gesetzt wird, finden auch andere Akteursgruppen des Verantwortungszusammenhangs in der Berichterstattung Berücksichtigung. So werden Politik und Behörden (15 %) sowie individuelle Nutzer\*innen und die Gesellschaft als Kollektiv (15 %) regelmäßig als Träger\*innen von Verantwortung genannt, wenngleich deutlich seltener als die Internet-Plattformen. Auffällig ist weiters, dass sich 9 % der Verantwortungsattributionen entweder (indirekt) an die Technologien oder an Programmierer\*innen oder Datenanalytiker\*innen richten. Dies verdeutlicht, dass auch die Rolle der Technologie und der Technikentwicklung bei Social-Media-Algorithmen in der Berichterstattung zum Teil mitberücksichtigt wird.

## 6. Fazit

Der vorliegende Beitrag beschäftigt sich mit den Gefahren von Social-Media-Algorithmen für Öffentlichkeit und Demokratie und schlägt dabei eine Brücke von den algorithmischen Anwendungen auf Social-Media-Plattformen, über die Risiken, die aus algorithmischer Selektion für Öffentlichkeit und Demokratie resultieren, bis hin zu Fragen der medialen Attribution von Verantwortlichkeit für Probleme und Lösungen. Erkenntnisse dieser Analyse lassen sich wie folgt zusammenfassen:

Zentrale Funktionen von Social-Media-Plattformen, wie die Kuratierung von Newsfeeds, die Filterung von problematischen Inhalten und die Schaltung von Werbung basieren auf algorithmischer Selektion. Algorithmen helfen dabei, die Informationsflut zu bewältigen und Information und Werbung gemäß Nutzer\*inneninteressen zu verbreiten. Damit tragen sie zum Erfolg der sozialen Netzwerke bei. Gleichzeitig geraten Algorithmen im öffentlichen Diskurs zunehmend in die Kritik. Sie werden für eine Reihe von Problemen verantwortlich gemacht, die zum Teil auch Risiken für Öffentlichkeit und Demokratie darstellen. Die detaillierte Analyse zeigt, dass Algorithmen in unterschiedlicher Art und Weise zur Entstehung von Risiken beitragen: durch Fehlerhaftigkeit, Verstärkungseffekte, als Ziel von

Manipulation, als Opportunitätsstruktur für gefährliche Praktiken und als Faktor für die Etablierung von Machtasymmetrien. Meist lassen sich Risiken jedoch nicht ausschließlich auf den Einsatz von Algorithmen zurückführen, sondern auf ein Zusammenwirken verschiedener Einflussfaktoren wie regulatorische Rahmenbedingungen, Geschäftsinteressen der Plattformbetreiber, Plattformdesign und Nutzungsverhalten. Daraus ergeben sich drei zentrale konzeptionelle und praktische Schussfolgerungen, die für die Analyse von Algorithmen und Verantwortlichkeit (vgl. Caplan et al. 2018) auf Internet-Plattformen hilfreich sind:

- (1) Die Risiken algorithmischer Selektion auf Internet-Plattformen lassen sich am besten als „sozio-technische Assemblagen“ (z.B. Annany 2015) beschreiben, um vielfältige Wirkungszusammenhänge zwischen involvierten Akteur\*innen und Technologien angemessen zu erfassen.
- (2) Auf einer solchen umfassenden sozio-technischen Risikoanalyse aufbauend, lassen sich Zurechnungspunkte für Verantwortlichkeit differenzieren. Zu den Verantwortungsträger\*innen im Bereich algorithmischer Risiken zählen Plattformunternehmen ebenso wie Entwickler\*innen/Programmierer\*innen, Nutzer\*innen, und Akteur\*innen aus dem Bereich Politik/Regulierung. Die Involvierung vielfältiger Akteursgruppen verweist auf eine verteilte Verantwortlichkeit (Saurwein 2018) in einem Verantwortungsnetzwerk (Loh/Loh 2017).
- (3) Die verschiedenen Akteursgruppen des Verantwortungsnetzwerkes sollten bei der Entwicklung von Governance-Strategien zum Umgang mit „algorithmischen Risiken“ berücksichtigt werden; in der Medienberichterstattung über algorithmische Risiken sollte das Zusammenspiel zwischen den unterschiedlichen Verantwortungsträger\*innen und Technologien verständlich erklärt werden.

Empirisch zeigt die Analyse des österreichischen Mediendiskurses zu Social-Media-Algorithmen, dass eine deutliche Mehrheit aller medialen Verantwortungsattributionen die Internet-Plattformen selbst betreffen, wenngleich auch andere Entitäten des Verantwortungsnetzwerkes (Politik, Nutzer\*innen) in der Berichterstattung Berücksichtigung finden. Die weitere Forschung sollte klären, ob es den Medien dabei gelingt, das Zusammenspiel von Plattformunternehmen, Nutzer\*innen, Politik und Technik bei der Entstehung von Risiken für Öffentlichkeit und Demokratie für das Publikum verständlich zu erklären und so zur adäquaten Darstellung von Verantwortlichkeit und zur Bewusstseinsbildung bei Nutzer\*innen beizutragen.

## Literatur

- Ananny, M. (2015): Toward an Ethics of Algorithms. In: *Science, Technology, & Human Values* 41(1), S. 93–117
- Are, C. (2020): How Instagram's algorithm is censoring women and vulnerable users but helping online abusers. In: *Feminist Media Studies* 20(5), S. 741–744
- Brantner, C.; Saurwein, F. (2021): Covering Technology Risks and Responsibilities: Automation, Artificial Intelligence, Robotics, and Algorithms in the Media. In: *International Journal of Communication (IJOC)*, 15 (2021), <https://ijoc.org/index.php/ijoc/article/view/17054>
- Brøvig-Hanssen, R.; Jones, E. (2021): Remix's retreat? Content moderation, copyright law and mashup music. In: *New Media & Society* (online first)
- Busch, O. (2016): The Programmatic Advertising Principle. In: Busch, O. (Hg.): *Programmatic Advertising. The Successful Transformation to Automated, Data-Driven Marketing in Real-Time*. Cham, S. 3–15
- Caplan, R.; Donovan, J.; Hanson, L.; Matthews, J. (2018): *Algorithmic Accountability: A Primer*. Research report, Data & Society Institute. New York
- Clegg, N. (2021): You and the Algorithm: It Takes Two to Tango. Facebook, 31.3.2021. <https://about.fb.com/news/2021/03/you-and-the-algorithm-it-takes-two-to-tango/> [aufgesucht am 14.07.2021]
- Cobbe, J.; Singh, J. (2019): Regulating Recommending: Motivations, Considerations, and Principles. In: *European Journal of Law and Technology* 10(3), S. 1–37
- Duarte, N.; Llanso, E.; Loup, A. (2018): *Mixed Messages? The Limits of Automated Social Media Content Analysis*. Washington: Center for Democracy & Technology
- Dubois, E.; Blank, G. (2018): The echo chamber is overstated: the moderating effect of political interest and diverse media. In: *Information, Communication & Society* 21(5), S. 729–745
- Fischer, E.; Puschmann, C. (2021): Wie Deutschland über Algorithmen schreibt. Eine Analyse des Mediendiskurses über Algorithmen und Künstliche Intelligenz (2005–2020). Gütersloh.
- Gillespie, T. (2018): *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven/London
- Gorwa, R.; Binns, R.; Katzenbach, C. (2020): Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance. In: *Big Data & Society* 7(1) (online first)
- Green, J.; Issenberg, S. (2016): Why the Trump machine is built to last beyond the election. Bloomberg, 26.10.2016. <https://www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go> [aufgesucht am 14.07.2021]
- Holznagel, D. (2018): Overblocking durch User Generated Content (UGC) – Plattformen: Ansprüche der Nutzer auf Wiederherstellung oder Schadensersatz? In: *Computer und Recht* 34, S. 369–378

- Horwitz, J.; Seetharaman, D. (2020): Facebook Executives Shut Down Efforts to Make the Site Less Divisive. *The Wall Street Journal*, 26.5.2020. <https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499> [aufgesucht am 14.07.2021]
- Just, N.; Latzer, M. (2017): Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet. In: *Media, Culture & Society* 39(2), S. 238–258
- Kreißel, P.; Ebner, J.; Urban, A.; Jakob, G. (2018): Hass auf Knopfdruck. Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz. London
- Latzer, M.; Hollnbuchner, K.; Just, N.; Saurwein, F. (2016): The economics of algorithmic selection on the Internet. In: Bauer, J.; Latzer, M. (Hg.): *Handbook on the Economics of the Internet*. Cheltenham. S. 395–425
- Lewis, P. (2018): 'Fiction is outperforming reality': how YouTube's algorithm distorts truth; <https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth> [aufgesucht am 14.07.2021]
- Lobigs, F. (2016): Finanzierung des Journalismus – von langsamen und schnellen Disruptionen. In: Meier, K.; Neuberger, C. (Hg.): *Journalismusforschung. Stand und Perspektiven*. Baden-Baden, S. 69–137
- Loh, J.; Loh, W. (2017): Autonomy and responsibility in hybrid systems: The example of autonomous cars. In: Lin, P.; Abney, K.; Jenkins, R. (Hg.): *Robot ethics 2.0. From autonomous cars to artificial intelligence*. Oxford/New York, S. 35–50
- Neuhäuser, C. (2014): Roboter und moralische Verantwortung. In Hilgendorf, E. (Hg.): *Robotik im Kontext von Recht und Moral*. Baden-Baden, S. 269–286
- Osipova, N.V.; Byrd, A. (2017): Inside Russia's network of bots and trolls. *The New York Times*, 31.10.2017; <https://www.nytimes.com/video/us/politics/100000005414346/how-russian-bots-and-trolls-invade-our-lives-and-elections.html> [aufgesucht am 14.07.2021]
- Pariser, E. (2011): *The filter bubble*. New York
- Saurwein, F. (2018): Automatisierung, Algorithmen, Accountability. Eine Governance Perspektive. In: Rath, M.; Krotz, F.; Karmasin, M. (Hg.): *Maschinenethik. Normative Grenzen autonomer Systeme*. Wiesbaden, S. 35–56
- Saurwein, F. et al. (forthcoming 2022): Die mediale Darstellung von Risiken und Verantwortung in der Berichterstattung über Social Media und Social Companions. CMC-Forschungsbericht. Wien
- Saurwein, F.; Spencer-Smith, C. (2019): Inhaltsregulierung auf Internet Plattformen. Optionen für verantwortungsorientierte Governance auf nationaler Ebene. CMC-Forschungsbericht. Wien
- Saurwein, F.; Spencer-Smith, C. (2020): Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe. In: *Digital Journalism* 8(6), S. 820–841
- Saurwein, F.; Spencer-Smith, C. (2021): Automated Trouble. The Role of Algorithmic Selection in Harms on Social Media Platforms. *Media and Communication* 9(4) (online first)

- Seetharaman, D.; Morris, B. (2017): Facebook's Onavo gives social-media firm inside peek at rivals' users. *The Wall Street Journal*, 13.8.2017; <https://www.wsj.com/articles/facebook-s-onavo-gives-social-media-firm-inside-peek-at-rivals-users-1502622003> [aufgesucht am 14.07.2021]
- Shepherd, R.P. (2020): Gaming Reddit's Algorithm: *r/the\_donald*, Amplification, and the Rhetoric of Sorting. In: *Computers and Composition* 56, S. 1–14
- Stark, B.; Stegmann, D.; Magin, M.; Jürgens, P. (2020): Are Algorithms a Threat to Democracy? The Rise of Intermediaries: A Challenge for Public Discourse. Report, AlgorithmWatch. Berlin
- Suárez-Gonzalo, S.; Mas-Manchón, L.; Guerrero-Solé, F. (2019): Tay is you. The attribution of responsibility in the algorithmic culture. In: *Observatorio* 13(2), S. 1–14
- Sunstein, C. (2001): *Republic.com*. Princeton
- Vosoughi, S.; Roy, D.; Aral, S. (2018): The spread of true and false news online. In: *Science* 359(6380), S. 1146–1151
- Whittaker, J.; Looney, S.; Reed, A.; Votta, F. (2021): Recommender systems and the amplification of extremist content. In: *Internet Policy Review* 10(2). DOI: 10.14763/2021.2.1565
- Zuboff, S. (2018): *Das Zeitalter des Überwachungskapitalismus*. Frankfurt am Main/New York
- Zuiderveen Borgesius, F.; Moeller, J.; Kruikemeier, S.; Fathaigh, R.; Irion, K.; Dobber, T.; Bodó, B.; de Vreese, C.H. (2018): Online Political Microtargeting: Promises and Threats for Democracy. In: *Utrecht Law Review* 14(1), S. 82–9