

Datenbasis mit Quellenangaben

Gerold Brägger/ Milena Giordano (2025). Beitrag „Trolls, Bots und Fake News“, PÄDAGOGIK, Heft 5/2025

Der Beitrag beschreibt skizzenhaft, wie Algorithmus-gesteuerte und KI-gestützte Desinformation in sozialen Medien funktioniert und welche Auswirkungen diese auf Nutzer:innen und Gesellschaft haben kann. Dabei werden Zahlen und Prozentangaben als Orientierungsgrößen angegeben. Aufgrund der begrenzten Zeichenzahl wird im Beitrag mehrheitlich auf Quellenangaben verzichtet. Im Folgenden wird die Datenbasis des Beitrags ausführlich dargestellt und kommentiert.

Anmerkungen zur Quellenlage

Die dynamische Natur von Online-Inhalten und die sich ständig weiterentwickelnden Algorithmen und KI-Modelle der Plattformen bedeuten, dass Aufrufstatistiken sehr volatil sein können und ohne kontinuierliche Datenerhebung möglicherweise keine langfristigen Trends widerspiegeln. Die Quellenlage ist aufgrund folgender Faktoren unübersichtlich und wechselhaft:

- Die Algorithmen und eingesetzten generativen Machine-Learning-Systeme (LMS)¹ der Plattformen ändern sich häufig. Es ist wichtig zu betonen, dass jede soziale Medienplattform ihren eigenen Algorithmen verwendet, die spezifische Regeln und Prioritäten verfolgt.
- Die Bemühungen zur Inhaltsmoderation variieren je nach politischen Vorgaben. Seit Beginn der Trump-Regierung in den USA werden die rechtlichen Vorgaben für die Tech-Konzerne gelockert oder ganz zerschlagen (Urheberrecht, Kinder- und Datenschutz etc.). So hat der Metakonzern im Januar 2025 bekannt gegeben, die Kooperation mit Fakten-Checkern zu beenden, dem Beispiel von Elon Musks Plattform zu folgen und auf „Community Notes“ zu setzen. Nutzer:innen von Facebook, Threads und Instagram sollen Online-Inhalte selbst moderieren². Die Folge davon ist, dass wieder mehr Desinformation, radikalisierte und diskriminierende Inhalte ungefiltert massenhaft verbreitet werden³.
- Die Betreiber der Social-Media-Plattformen haben selbst kein Interesse daran offenzulegen, wie viele Trolls, Bots und Fake News auf ihren Kanälen distribuiert werden – geschweige denn die Funktionsweise der eingesetzten Algorithmen und LMS transparent zu machen. Das KI-Gesetz und die Social-Media-Vorgaben der EU sehen empfindliche Strafen vor, was die Bereitschaft zur Transparenz mindert und die Datenlage schmälert.
- Daher sollten alle Zahlen als zeitgebundene Schätzungen verstanden werden, die die rechtlichen, technischen und gesellschaftlichen Bedingungen zum Zeitpunkt der Datenerhebung widerspiegeln.
- Die Vergleichbarkeit der Daten wird zudem weiter durch den Umstand beeinträchtigt, dass die Definition und Kategorisierung von „Fake News“, „Hassreden“, „Verschwörungstheorien“ und „Radikalisierung“ zwischen Studien variieren können. Diese definitorische Unschärfe muss bei der Interpretation und dem Vergleich von Ergebnissen aus verschiedenen Quellen berücksichtigt werden.

Diese und weitere Faktoren (wie länderspezifische Kontexte, Alter und Bildung der Nutzergruppen) führen dazu, dass kaum Metastudien mit generalisierbaren Daten verfügbar sind. Im Beitrag haben wir deshalb versucht, aus mehreren unterschiedlichen Quellen **zusammenfassende quantitative Einschätzungen über die Wirkungen und Wirkungsweise Sozialer Medien** zu treffen. Diese Schätzungen wurden in der Absicht gemacht, den Leser:innen eine gewisse Orientierung über das Ausmaß an Manipulation in sozialen Medien zu geben. Auch wenn die angegebenen Zahlen aufrüttelnd sind – sie wurden sehr zurückhaltend eingeschätzt. Die im Folgenden beschriebenen Quellen zeigen, dass **alle Zahlenangaben weit übertroffen** werden und die verfügbaren Forschungsergebnisse Anlass zur Sorge geben.

¹ Döbeli-Honegger, B. (2024, 3., aktualisierte und erweiterte Auflage). Generativen Machine-Learning-Systeme – die nächste Herausforderung des digitalen Leitmedienwechsel. In: G. Brägger/H.-G. Rolff (Hrsg.): Lernen mit digitalen Medien. (S. 13-36). Weinheim: Beltz.

² Tagesschau (7.1.25): Meta beendet in den USA Kooperation mit Faktenprüfern, <https://www.tagesschau.de/wirtschaft/unternehmen/meta-faktencheck-moderation-100.html> (Zugriff am 8.01.25)

³ Klicksafe (8.1.25). Instagram, Facebook und Threads lassen wieder mehr Falschinformationen und Hassrede zu. <https://www.klicksafe.de/news/instagram-facebook-und-threads-lassen-wieder-mehr-falschinformationen-und-hassrede-zu> (Zugriff am 8.01.25)

Anmerkungen zum Abschnitt

KI-gestützte Desinformation in sozialen Medien

Seit der breiten Verfügbarkeit generativer KI-Tools Ende 2022 hat die Menge der KI-generierten Inhalte in sozialen Medien stark zugenommen⁴. NewsGuard verzeichnete 2023 eine **zehnfache Zunahme von KI-gestützten Nachrichtenseiten**⁵. Der Anteil von **KI-generierten Langform-Posts auf Facebook** stieg von 5,34 % vor 2023 auf **41,18 %** im November 2024. Dies entspricht einem **4,3-fachen Anstieg** seit der Einführung von ChatGPT. Auch bei Deepfakes gab es eine deutliche Zunahme, mit einer **Verdreifachung bei Videos und einer Verachtfachung bei Sprach-Deepfakes** von 2022 bis 2023⁶.

Einer der Hauptvorteile von KI bei der Desinformationsverbreitung ist die Möglichkeit, **Inhalte in großem Maßstab und zu geringen Kosten zu generieren**. Dies ermöglicht es Akteuren, eine große Menge an Desinformation schnell und kostengünstig zu vertreiben.⁷ Studien zeigen, dass KI-Tools die Produktivität bei der Inhaltserstellung erheblich steigern können, indem sie beispielsweise das Umschreiben von Artikeln oder die Erstellung von Variationen erleichtern.⁸ NewsGuard identifizierte über 1.200 unzuverlässige, KI-generierte Nachrichtenseiten in 16 Sprachen, die mit geringer oder keiner menschlichen Aufsicht betrieben werden und Dutzende bis Hunderte von Artikeln produzieren.⁹

Untersuchungen (z.B. Chang et al. 2025¹⁰) belegen, dass moderne Sprachmodelle leicht aufgefordert werden können, falsche und politisch suggestive Inhalte zu generieren. Chang et al. berichteten, dass bereits Vorgängermodelle **hohe Erfolgsraten bei der Generierung von Desinformation** erreichten (z.B. 86 % bei Davinci-003) und GPT-4 nahezu 99 % der Anfragen in schädlichem Inhalt beantwortet. Weiter zeigen die Autoren, dass „freundliche“ Aufforderungen an KI-Sprachmodelle sogar die Erzeugung falscher Inhalte erhöhen.

Bei der Frage, wie viele KI-generierte Fake News erkannt bzw. geteilt werden, gibt es eine Vielzahl von Studien mit unterschiedlichen Fragestellungen, die nur bedingt vergleichbar sind: Gemäß Key Statistics on Fake News (2024) haben zwischen 2022 und 2023 in den USA **38% der Social Media-News-Nutzer:innen unbemerkt Fake News weitergeteilt**¹¹. Chang et al. zeigen, dass KI-generierte Texte gegenüber von Menschen erstellten Text nicht zu unterscheiden sind und im Falle von KI-generierter Desinformation zugleich meist überzeugender sind.¹² Eine zentrale Herausforderung besteht darin, dass Menschen KI-generierte Inhalte oft nicht von menschengemachten unterscheiden können. Eine Studie mit rund 3.000 Teilnehmenden aus Deutschland, China und den USA zeigte, **dass KI-generierte Medien mehrheitlich als menschengemacht klassifiziert wurden**¹³. Selbst Faktoren wie Alter, Bildungshintergrund oder Medienkompetenz hatten keinen signifikanten Einfluss auf die Erkennungsfähigkeit (ebenda).

⁴ Schweizerische Eidgenossenschaft (2024): Beeinflussungsaktivitäten und Desinformation Bericht des Bundesrates. Bern 19.06.2024, S. 7ff, <https://www.news.admin.ch/news/message/attachments/88281.pdf> (Zugriff am 15.09.24)

⁵ Jack Brewster, Eric Effron, Zack Fishman, Sam Howard, Eva Maitland, McKenzie Sadeghi, and Jim Warren (2023). The Year AI Supercharged Misinformation. NewsGuard's 2023 in Review, Abschnitt "Growth of Unreliable AI-generated News Sites in 2023", <https://www.newsguardtech.com/misinformation-monitor/december-2023/> (Zugriff am 05.02.24)

⁶ Key Statistics on Fake News & Misinformation in Media in 2024 - Redline Digital, <https://redline.digital/fake-news-statistics/> (Zugriff am 24.08.24)

⁷ Yan, H. Y. et al (2025). The origin of public concerns over AI supercharging misinformation in the 2024 U.S. presidential election. <https://misinfoeview.hks.harvard.edu/article/the-origin-of-public-concerns-over-ai-supercharging-misinformation-in-the-2024-u-s-presidential-election/> (Zugriff am 12.02.25)

⁸ NewsGuard (2025). Tracking AI-enabled Misinformation: 1,271 'Unreliable AI-Generated News' Websites (and Counting), Plus the Top False Narratives Generated by Artificial Intelligence Tools <https://www.newsguardtech.com/special-reports/ai-tracking-center/> (Zugriff am 12.02.25)

⁹ ebenda

¹⁰ Chang, X., et al. (USA, Frontiers in AI 2025, S. 1–13): "Emotional prompting amplifies disinformation generation in AI large language models". frontiersin.org. (Zugriff am 15.04.25)

¹¹ Key Statistics on Fake News & Misinformation in Media in 2024 - Redline Digital, <https://redline.digital/fake-news-statistics/> (Zugriff am 24.08.24)

¹² Chang, X., et al. (USA, Frontiers in AI 2025, S. 1–13): "Emotional prompting amplifies disinformation generation in AI large language models". frontiersin.org. (Zugriff am 15.04.25)

¹³ Forschung & Lehre (2024). Menschen erkennen KI-generierte Medien kaum. <https://www.forschung-und-lehre.de/zeitfragen/menschen-erkennen-ki-generierte-medien-kaum-6430>. (Zugriff am 28.08.24)

Internationale Reports (EDMO 2023¹⁴) dokumentieren, dass generative KI zunehmend für gefälschte Bilder, Videos oder manipulierte Nachrichten eingesetzt werden (z.B. Fake-Werbeanzeigen oder Deepfake-Videos). KI-Algorithmen steigern zudem nachweislich Engagement – manipulierte Posts werden stärker geklickt oder geteilt. Die aufgeführten Studien unterstreichen, dass KI-basierte Desinformation bereits massiv verbreitet wird und von Menschen kaum als künstlich erkennbar ist. So zeigt eine Studie, dass **Deepfakes zunehmend raffinierter werden und von den Menschen kaum mehr korrekt identifizieren** können¹⁵. Eine besondere Herausforderung stellt dabei der Umstand dar, dass die Menschen sich selbst zugeschriebene Fähigkeit, falsche Inhalte online zu erkennen, nicht mit der tatsächlich gemessenen Fähigkeit korreliert. In einer **OECD-Studie konnten die Befragten wahre und falsche Inhalte nur zu 60 % korrekt identifizieren**¹⁶. Dies deutet darauf hin, dass das Vertrauen in die eigene Urteilsfähigkeit oft stark überschätzt wird: Eine Studie offenbart, dass **nur 0.1% der Menschen KI-generierte Deepfakes sicher erkennen** können.¹⁷

Künstliche Intelligenz spielt eine zunehmend wichtige Rolle in der algorithmischen Inhalts-Kuratierung und Personalisierung. KI-gestützte Empfehlungssysteme analysieren das Nutzerverhalten und schlagen Nutzer:innen Inhalte, Produkte, Konten und andere Nutzer vor, die ihren bereits bestehenden Präferenzen ähneln. KI hat die Fähigkeit, große Datenmengen zu analysieren und Muster und Erkenntnisse zu identifizieren, die Menschen möglicherweise übersehen würden. Dies ermöglicht eine effizientere und effektivere Erstellung und Ausrichtung von Inhalten. So kann KI beispielsweise dabei helfen, die günstigsten Zeitpunkte für die Veröffentlichung von Inhalten zu ermitteln, die Beitragsformate zu identifizieren, die das höchste Maß an Publikumsbeteiligung hervorrufen, und die effektivsten Ansätze für die Interaktion mit verschiedenen Zielgruppensegmenten zu bestimmen¹⁸.

Die **Wirkungsweise von Algorithmen und KI in sozialen Medien** hat erhebliche Auswirkungen auf die Sichtbarkeit von Inhalten und die gesamte Nutzererfahrung. Algorithmen bestimmen, welche Inhalte im Feed eines Nutzers angezeigt werden und in welcher Reihenfolge. Dies kann dazu führen, dass die organische Reichweite von Unternehmensseiten oft reduziert wird, wodurch Marken vermehrt auf bezahlte Werbung angewiesen sind, um ihre Zielgruppen zu erreichen. Gleichzeitig haben Inhalte, die ein hohes Maß an Engagement erzeugen (z.B. durch Likes, Kommentare und Shares), eine größere Chance, organisch eine breitere Zielgruppe zu erreichen. Es ist jedoch wichtig zu beachten, dass Algorithmen auch unbeabsichtigte Folgen haben können. So kann die Priorisierung von Inhalten, die auf maximale Interaktion ausgelegt sind, unbeabsichtigt zu einer Verstärkung von extremen oder kontroversen Inhalten führen, da diese oft mehr Reaktionen hervorrufen.

Die primäre Funktion von Algorithmen in sozialen Medien ist die **Maximierung der Nutzerbindung**, was oft durch die Förderung von Inhalten geschieht, die starke Reaktionen hervorrufen. **KI verstärkt diese Personalisierung**, was zu hochgradig maßgeschneiderten, aber potenziell einschränkenden Informationserfahrungen führt. Das Ziel der Plattformen, die Verweildauer der Nutzer zu maximieren, führt zur Entwicklung von Algorithmen, die Inhalte priorisieren, die hohe Engagement-Raten aufweisen. KI ermöglicht eine präzisere Vorhersage, welche Inhalte Nutzer wahrscheinlich engagieren werden, was die Personalisierung weiter intensiviert. Diese Mechanismen können unbeabsichtigt oder absichtlich die Informationslandschaft der Nutzer formen, was Auswirkungen auf ihre Wahrnehmung der Realität und ihre Fähigkeit zu informierten Entscheidungen hat.

Wie weit die Personalisierung Sozialer Medien mithilfe von KI inzwischen vorangetrieben wird, zeigt der kürzliche Entscheid des Meta-Konzerns, **alle persönlichen Daten, die Nutzer:innen auf Facebook, Instagram und WhatsApp gepostet haben, für das Training der Künstlichen Intelligenz zu nutzen**. «Meta plant, ab dem 27. Mai sämtliche öffentlich zugänglichen Inhalte in Instagram und Facebook, also öffentliche Posts, Storys, Fotos, Videos, Kommentare unter Fotos und in Gruppen und so weiter, dauerhaft in KI-Trainingsdaten zu verwandeln. (...) Metas KI-Produkte sollen nun also auch mit dem trainiert werden, was Milliarden von Menschen weltweit täglich an Inhalten produzieren. Und an persönlicher Kommunikation: Geburtstagswünsche, Gratulationen zum Jobwechsel oder der Geburt eines Kindes, verzückte Kommentare unter Babyfotos, wütendes Gezeter unter journalistischen Posts, Troll-

¹⁴ European Digital Media Observatory (EDMO) 2024 (EU). Generative AI and Disinformation. Generative AI and Disinformation: Recent Advances, Challenges, and Opportunities, <https://edmo.eu/wp-content/uploads/2023/12/Generative-AI-and-Disinformation-White-Paper-v8.pdf> (Zugriff am 17.04.25)

¹⁵ U.S. Government Accountability Office (2024). Combating Deepfakes. <https://www.gao.gov/products/gao-24-107292> (Zugriff am 28.08.24)

¹⁶ OECD (o.J.). Disinformation and misinformation. <https://www.oecd.org/en/topics/disinformation-and-misinformation.html> (Zugriff am 12.06.24)

¹⁷ iProov (2025). Study Reveals Deepfake Blindspot: Only 0.1% of People Can Accurately Detect AI-Generated Deepfakes. <https://www.iproov.com/press/study-reveals-deepfake-blindspot-detect-ai-generated-content> (Zugriff am 12.03.25)

¹⁸ Elearncollege (2025). Social Media Algorithms and the Illusion of Choice. <https://elearncollege.com/technology/social-media-algorithms-and-the-illusion-of-choice/> (Zugriff am 12.03.25)

Provokationen und Beschwichtigungsversuche, massenhaft Desinformation und Menschenhass, den Metas Moderatorinnen und Moderatoren nicht entfernt haben. Alles ein gigantischer Mahlstrom aus Bildern, Videos und Text, der kontinuierlich und in atemberaubendem Tempo immer weiterwächst“ (Stöcker 2025¹⁹).

Anmerkungen zum Abschnitt

Social Bots

Social Bots sind automatisierte Accounts, die Inhalte posten und diskutieren. Social Bots sind ein wichtiges Werkzeug zur algorithmischen Manipulation und Verbreitung von Desinformation. Sie können Inhalte massenhaft verbreiten, Engagement simulieren und Narrative verstärken. Wie die 2024 und 2025 durchgeführten Wahlen in europäischen Ländern und der USA gezeigt hat, beeinflussen sie das gesellschaftliche Klima und die politischen Diskussionen erheblich. Sie manipulieren die öffentliche Meinung, indem sie ein falsches Bild von Popularität oder Konsens erzeugen (Astroturfing). Sie sind besonders effektiv in den ersten Momenten, bevor eine Nachricht viral geht²⁰.

Varol et al. haben in ihrer Studie bereits 2017 geschätzt, dass etwa **9–15 % der aktiven Twitter-Accounts Bot-Verhalten** zeigen²¹. Mit der Übernahme von Twitter durch Elon Musk im Oktober 2022 und der Umwandlung von X als Mittel der politischen Einflussnahme ist aktuell von einer **deutlich höheren Anzahl von Bots von geschätzt 25-30%** auszugehen²². Mehrere Studies des Center for Countering Digital Hate belegen, dass nach Musks Übernahme die Interaktionen mit Inhalten bekannter Desinformationsquellen stark zunahmten und dabei folgende Taktiken betrieben werden: Verstärkung durch algorithmische Empfehlungssysteme, Nutzung von Hashtag-Hijacking, Einsatz von Koordinierungsplattformen und Bots außerhalb von X²³.

Langzeitstudien und Reviews zeigen, dass Bots gezielt politische Debatten dominieren können. Beispielsweise dokumentieren McKenzie Himelein-Wachowiak et al (2021), dass Social Bots Informationen zu Covid massenhaft manipulierten und das Diskussionsklima verschlechterten²⁴. Shao et al. (2018) fanden, dass Social Bots die Verbreitung unglaubwürdiger Nachrichten erheblich verstärken (z.B. überproportional viele Retweets ungültiger Quellen)²⁵. Neueste Forschung (z.B. Lalor et al. 2025) zeigen, dass **Bots die Nutzer-Interaktion zwar steigern, aber Dialoge verkürzen: Bots erhöhen die Klickzahlen und Kommentarhäufigkeit, senken jedoch Tiefgang der Diskussionen**²⁶. In Klimadebatten verhalten sich Bots besonders aggressiv und ziehen negative Gefühle an, wie Li et al. (2024) beschreiben (Bots in Twitter-Klimadebatten verdrängen menschliche Nutzer und verschärfen den Ton)²⁷. Insgesamt konsistent berichten Studien, dass Social Bots viral verbreitete Fehlinformation gezielt pushen und menschliche Debatten „vergiften“. Eine weitere Studie analysiert die überlegene Online-Popularität der AfD im Vergleich zu den anderen deutschen Parteien. Die Beweise würden darauf hindeuten, dass **automatisierte Konten zu dieser Online-Überlegenheit beigetragen haben**.²⁸

¹⁹ Stöcker, Ch. (2025). KI-Trainingsdaten. Verweigern Sie sich Zuckerbergs dystopischen Visionen. SPIEGEL (18.05.25).

https://www.spiegel.de/wissenschaft/mensch/facebook-instagram-whatsapp-und-ki-verweigern-sie-sich-mark-zuckerbergs-dystopischen-visionen-kolumne-a-224b869b-6c7f-437e-ae5b-917ea7622ced?sara_ref=re-so-app-sh (Zugriff am 18.05.25)

²⁰ Observatory on Social Media, Indiana University (219). <https://osome.iu.edu/research/blog/twitter-bots-spread-misinformation> (Zugriff am 03.03.22)

²¹ Varol, O., Ferrara, E., Davis, C., Menczer, F. & Flammini, A. (ICWSM 2017, Palo Alto): “Online Human-Bot Interactions: Detection, Estimation, and Characterization” – Schätzung: 9–15 % aller aktiven Twitter-Accounts sind Social Bots cdn.aaii.org, S. 280 (Zugriff am 06.02.23)

²² Social Bot Detection Techniques and Challenges—A Review (2023).

https://www.researchgate.net/publication/370474695_Social_Bot_Detection_Techniques_and_Challenges-A_Review_Approach (Zugriff am 06.04.24)

²³ Reports on Twitter/ X 2023- 2025: https://counterhate.com/research/?_sft_topic=twitter (Zugriff am 12.05.25)

²⁴ McKenzie Himelein-Wachowiak et al (2021). Bots and Misinformation Spread on Social Media: Implications for COVID-19. JOURNAL OF MEDICAL INTERNET RESEARCH. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8139392/> (Zugriff am 04.02.23)

²⁵ Shao, C., Ciampaglia, G.L. et al. (Nat. Commun. 2018): “The spread of low-credibility content by social bots” – Nachweis, dass Social Bots unglaubwürdige Nachrichten überproportional verbreiten. nature.com. (Zugriff am 04.02.23)

²⁶ Lalor, J., Berente, N. & Safadi, H. (MIS Quarterly 2025): “The Effect of Bots on Human Interaction in Online Communities” – Studie zeigt: Bots steigern Engagement (Kommentare, Likes), verringern aber tiefere Mensch-zu-Mensch-Interaktion. techxplore.com. (Zugriff am 04.02.25)

²⁷ Li, L., Vászrhelyi, O. & Vedres, B. (Sci. Rep. 2024): “Social bots spoil activist sentiment” – Experimentell: Begegnung mit Bots führt zu negativerem Nutzer-Feedback (Klimastreik-Twitter; *negativer Sentiment-Effekt*). nature.com. (Zugriff am 12.12.24)

²⁸ Serrano, J. C. M. (2019). The Rise of Germany’s AfD: A Social Media Analysis.

https://www.researchgate.net/publication/334155685_The_Rise_of_Germany's_AfD_A_Social_Media_Analysis (Zugriff am 02.02.20)

Anmerkungen zum Abschnitt

Filterblasen

Die **Personalisierung von Inhalten durch Algorithmen** erzeugt bei den Nutzern oft eine trügerische Empfindung der Kontrolle. Obwohl es scheint, als würden die Nutzer aktiv die Inhalte auswählen, die sie interessieren, werden ihre Entscheidungen tatsächlich auf subtile Weise durch die algorithmische Kuration gelenkt. Diese Illusion der Autonomie kann dazu führen, dass sich Nutzer der algorithmischen Beeinflussung ihrer Informationsaufnahme und Meinungsbildung nicht bewusst sind, was sie anfälliger für Manipulationen macht. Wenn Nutzer davon ausgehen, dass ihre Informationsauswahl rein selbstbestimmt ist, neigen sie möglicherweise weniger dazu, die präsentierten Inhalte kritisch zu hinterfragen. Die Algorithmen lernen kontinuierlich aus den Interaktionen der Nutzer und verstärken so die Exposition gegenüber ähnlichen Inhalten, was die Entstehung von Filterblasen begünstigt, ohne dass die Nutzer dies aktiv steuern oder überhaupt bemerken.²⁹

Anmerkungen zum Abschnitt

Fake News

Fake News sind gezielt irreführende Nachrichten im Stil journalistischer Beiträge. Die Literatur zeigt, dass sich solche Falschmeldungen häufig schneller und weiterverbreiten als korrekte Inhalte. Vosoughi et al. (Science 2018) fanden, dass **Fake News deutlich schneller viral** gehen: Es dauert etwa **6-mal länger, bis wahre Informationen die gleiche Reichweite** (z.B. 1.500 Personen) wie falsche erreichen. **Falschmeldungen werden 70 Prozent häufiger retweetet** als wahre Geschichten. In Twitters „Kaskaden“, den ununterbrochenen Retweet-Ketten, erreichen Unwahrheiten eine Kaskadentiefe von zehn – also **etwa 20-mal schneller als Fakten**. Und Unwahrheiten werden von einzelnen Nutzern in jeder Kaskadentiefe häufiger retweetet als wahre Aussagen³⁰. In einer US-Studie berichteten Allcott & Gentzkow (JEP 2017), dass 14 % der Amerikaner soziale Medien als Hauptnachrichtenquelle nutzen und Fake-News-Seiten im Wahlkampf 2016 stark rezipiert wurden (z.B. 30 Mio. Facebook-Shares für Trump-freundliche Fake-Artikel vs. 8 Mio. für Clinton-freundliche)³¹. Umfragen (z.B. Pew Research 2019) zeigen, dass viele Nutzer häufig auf irreführende Meldungen stoßen und dies das Vertrauen in Medien untergräbt: über 50 % sehen Fake News vor allem im anderen Lager und journalistischen Medien³².

Verbreitung von Videos zu Fake News, Verschwörungserzählungen und Radikalisierung: Der deutsche Journalist und Autor Mats Schönauer hat laut eigenen Angaben ein «ein riesiges Netzwerk» von YouTube-Kanälen entdeckt, deren Video-Inhalte komplett KI-generiert sind. Die „KI-Horror“-Videos auf YouTube, die oft gefälschte Szenarien und Clickbait nutzen, ergaben, werden millionenfach angesehen, was auf ein beträchtliches Publikum für bestimmte Arten von KI-fabrizierten Inhalten hindeutet (watson 2024³³). Angesichts der enormen Größe von YouTube, mit über einer Milliarde täglich gestreamter Stunden im Jahr 2024 und Milliarden monatlich aktiver Nutzer:innen³⁴, könnte selbst wenn Fake News nur einen kleinen Bruchteil des gesamten Inhalts ausmachen, die absolute Anzahl der täglichen Aufrufe immer noch beträchtlich sein. Angesichts des exponentiellen Wachstums von YouTube³⁵ und der damit erreichten Reichweite ist von einer **täglich millionenfachen, massiven Verbreitung von Videos zu Fake News, Verschwörungserzählungen und Radikalisierung** auszugehen.

Auch auf der **TikTok-Plattform** ist die Verbreitung von Desinformation und Manipulation massiv: Eine Studie aus dem Jahr 2022 ergab, dass **fast 20 % der bei Suchanfragen nach Top-Nachrichtmeldungen zurückgegebenen Videos**

²⁹ Elearncollege (2025). Social Media Algorithms and the Illusion of Choice. <https://elearncollege.com/technology/social-media-algorithms-and-the-illusion-of-choice/> (Zugriff am 12.03.25)

³⁰ Vosoughi, S., Roy, D. & Aral, S. (Science 2018, 359, 1146–1151): “The spread of true and false news online”. [Study: On Twitter, false news travels faster than true stories | MIT News | Massachusetts Institute of Technology](https://www.mit.edu/news/2018/08/07/fake-news-travels-faster-than-true-stories) (Zugriff am 07.08.24)

³¹ Allcott, H. & Gentzkow, M. (J. Econ. Persp. 2017, 31(2): 211–36): Social Media and Fake News in the 2016 Election. [aeaweb.org](https://www.aeaweb.org/conference/papers/2017/Allcott-Gentzkow-2017-02). (Zugriff am 07.08.24)

³² Pew Research Center (2019): Many Americans Say Made-Up News Is a Critical Problem”. https://www.pewresearch.org/wp-content/uploads/sites/20/2019/06/PJ_2019.06.05_Misinformation_FINAL-1.pdf (Zugriff am 07.08.24)

³³ Watson (2012.24). KI-Horror bei YouTube – das steckt hinter dem Phänomen. <https://www.watson.ch/digital/kuenstliche-intelligenz/835898352-ki-fake-news-und-clickbait-bei-youtube-das-steckt-hinter-dem-phaenomen> (Zugriff am 03.01.25)

³⁴ 49 YouTube Stats 2025: Growth, Views & Engagement in 2025 – Descript. <https://www.descript.com/blog/article/49-youtube-stats-2023-engagement-views-revenue-and-more> (Zugriff am 2.05.25)

³⁵ Schwab, P.-N. (2024). Research reveals YouTube’s most secret stats. <https://www.intotheminds.com/blog/en/research-youtube-stats/> (Zugriff am 25.08.24)

Fehlinformationen enthielten³⁶. Die Rolle der Plattform als Suchmaschine für jüngere Nutzer verstärkt das Potenzial für die Verbreitung gegenüber Fake News zusätzlich. Eine Google-Studie ergab, dass 40 % der Gen-Z in den USA TikTok mittlerweile für die Internetrecherche nutzen, wobei die Top-Ergebnisse oft Fehlinformationen enthalten³⁷. Auch KI-generierte Inhalte sind ein bedeutender Verbreitungsweg für Fake News auf TikTok. BBC-Ergebnisse aus dem Jahr 2024 zeigen, dass jungen Wählern irreführende KI-generierte Videos über Politiker empfohlen werden³⁸. KI-generierte Verschwörungstheorien erreichen Millionen von Nutzer:innen erreichen, wobei einige Konten Hunderte von Millionen Aufrufen pro Video erzielen³⁹. Eine Studie schätzte, dass ungefähr 0,1 % aller Videos Verschwörungstheorien fördern.⁴⁰ Angesichts der Milliarden jährlich hochgeladener Videos und weltweit über eine Milliarde Nutzer:innen⁴¹ stellt dieser Prozentsatz immer noch ein beträchtliches Volumen potenziell angesehener Inhalte dar.

Manipulation von Algorithmen und KI durch Plattformbetreiber für politische Zwecke: Die Betreiber sozialer Medien verfügen über erhebliche Möglichkeiten, ihre Algorithmen und die eingesetzte künstliche Intelligenz potenziell für politische und/oder kommerzielle Zwecke zu manipulieren, um politische Zwecke zu verfolgen. Da die genauen Details der Algorithmen oft als Geschäftsgeheimnisse behandelt werden, ist eine externe Überprüfung solcher Manipulationen erheblich erschwert. Plattformen haben die technische Fähigkeit, ihre Algorithmen so zu verändern, dass bestimmte Inhalte oder Meinungen bevorzugt oder benachteiligt werden. Dies kann auf subtile Weise geschehen, beispielsweise durch die **Anpassung der Gewichtung bestimmter Engagement-Metriken** wie Likes, Kommentare oder Shares⁴². Weniger subtile Einflussnahmen hat kürzlich Elon Musks Chatbot Grok offenbart, der selbst angab dass seine "Erschaffer" ihn angewiesen hätten, auf X über den vermeintlichen Massenmord an weißen Farmern in Südafrika zu sprechen und so Verschwörungstheorien über den "White Genocide" zu verbreiten⁴³.

Die Manipulation von Algorithmen und KI kann erhebliche Auswirkungen auf das politische Agenda-Setting haben. Algorithmen beeinflussen, welche Themen in der öffentlichen Debatte an Bedeutung gewinnen (Themensetting). Durch die algorithmische Selektion von Inhalten kann auch die Vermittlung von Faktenwissen und des Meinungsspektrums zu einem bestimmten Thema beeinflusst werden. Personalisierte Empfehlungssysteme können dazu führen, dass Nutzer einem weniger vielfältigen Spektrum an Nachrichten und Informationen ausgesetzt sind. Studien deuten darauf hin, dass **Algorithmen möglicherweise Inhalte von rechten politischen Akteuren und Medien stärker verstärken** als von linken. Eine Analyse von politischen Nachrichten-URLs, die während der US-Präsidentenwahlen 2020 auf Facebook gepostet wurden, ergab, dass deutlich mehr Websites, die sich an Konservative richten, in diesen Beiträgen enthalten waren als solche, die sich an Liberale richten⁴⁴.

Es gibt bereits zahlreiche Beispiele und Studien, die auf eine **politische Einflussnahme durch algorithmische Manipulation** hinweisen. Sie deuten darauf hin, dass die Algorithmen von Plattformen wie **TikTok und X in Deutschland unverhältnismäßig viele rechte, pro-AfD-Inhalte an nicht-parteiische Nutzer** ausspielen. Eine Studie von Global Witness aus dem Jahr 2025 ergab, dass nicht-parteiische Nutzer in Deutschland auf TikTok 74 % rechte politische Inhalte sahen, von denen der größte Teil pro-AfD war. Auf X waren es 72 % rechte und 64 % pro-AfD-

³⁶ U.S. PIRG Education Fund (2023). How misinformation on social media has changed news

<https://pirg.org/edfund/articles/misinformation-on-social-media/> (Zugriff am 25.08.24)

³⁷ New York Times (2022). For Gen Z TikTok ist the new Search Engine. <https://www.nytimes.com/2022/09/16/technology/gen-z-tiktok-search-engine.html> (Zugriff am 25.11.12)

³⁸ BBC (02.06.2024). TikTok users being fed misleading election news, BBC finds.

<https://www.bbc.com/news/articles/c1ww6vz1l81o> (Zugriff am 25.08.24)

³⁹ New York Public Radio (09.02.2024). AI-Generated Conspiracy Theories Are Flooding TikTok.

<https://www.wnycstudios.org/podcasts/otm/segments/ai-generated-conspiracy-theories-are-flooding-tiktok-on-the-media> (Zugriff am 25.08.24)

⁴⁰ TikTok Statistics: Revenue & Usage (Updated Apr. 2025) – SendShort. <https://sendshort.ai/statistics/tiktok/> (Zugriff am 25.08.24)

⁴² Morgan, C. J. (2018). The Silencing Power of Algorithms: How the Facebook News Feed Algorithm Manipulates Users' Perceptions of Opinion Climates. Portland State University. [The Silencing Power of Algorithms: How the Facebook News Feed Algorithm Manipulates Users' Perceptions of Opinion Climates](https://www.psu.edu/research/center-for-social-media-research/the-silencing-power-of-algorithms-how-the-facebook-news-feed-algorithm-manipulates-users-perceptions-of-opinion-climates) (Zugriff am 13.03.20)

⁴³ Der Standard (15.05.2025). Musks Chatbot Grok verbreitet Verschwörungstheorien über den "White Genocide".

<https://www.derstandard.at/story/3000000269902/musks-chatbot-grok-verbreit-verschwoerungstheorien-ueber-den-white-genocide> (Zugriff am 15.05.25)

⁴⁴ Gogarty, K. (2021). Facebook tweaks its news Feed algorithm, and right-leaning pages are reaping the benefits.

<https://www.congress.gov/118/meeting/house/115561/documents/HHRG-118-IF16-20230328-SD038.pdf> (Zugriff am 28.03.22)

Inhalte⁴⁵. Im Gegensatz dazu zeigte Instagram mit 59 % eine geringere rechte Tendenz, wobei 96 % der Inhalte von den gefolgten Konten stammten.

Plattformbetreiber verfügen über erhebliche Macht über die Informationsumgebung ihrer Nutzer durch die Gestaltung ihrer Algorithmen. Diese Macht birgt das **Risiko einer unbeabsichtigten oder sogar beabsichtigten Manipulation des politischen Diskurses**, insbesondere im Hinblick auf Wahlen. Die genauen Funktionsweisen der Algorithmen sind oft Geschäftsgeheimnisse, was es externen Beobachtern erschwert, potenzielle Manipulationen aufzudecken oder deren Auswirkungen vollständig zu verstehen. Darüber hinaus können die Geschäftsmodelle sozialer Medien, die auf Nutzerengagement und Verweildauer basieren, unbeabsichtigt Anreize für die Verbreitung von polarisierenden oder sensationalistischen Inhalten schaffen, da diese oft höhere Interaktionsraten aufweisen und somit die Einnahmen der Plattform steigern können. Die Optimierung von Algorithmen auf maximale Interaktion kann dazu führen, dass qualitativ hochwertige oder ausgewogene Informationen gegenüber aufmerksamkeitsstarken, aber möglicherweise irreführenden Inhalten benachteiligt werden⁴⁶.

Anmerkungen zum Abschnitt

Deepfakes

Politische Akteure erkennen zunehmend das immense Potenzial von Algorithmen und künstlicher Intelligenz, um die öffentliche Meinung zu beeinflussen und ihre politischen Ziele zu erreichen. Sie nutzen soziale Medien aktiv, um ihre Botschaften zu verbreiten und die öffentliche Meinung in ihrem Sinne zu beeinflussen. Dabei greifen sie auf verschiedene Strategien zurück, darunter der Einsatz von **KI-gestützten Bots und gefälschten Konten**⁴⁷ (s.o.), um Konversationen zu manipulieren, Fehlinformationen zu verbreiten und schädliche Inhalte zu verstärken. Eine besonders besorgniserregende Entwicklung ist die **Verwendung von Deepfakes, also KI-generierten Videos oder Audioaufnahmen**, die täuschend echt wirken und dazu dienen können, Nutzer zu manipulieren und politische Narrative zu verfälschen.

Im Bereich der **kognitiven Kriegsführung** werden Algorithmen und KI gezielt eingesetzt, um die Meinungen und Einstellungen der Bevölkerung zu beeinflussen. So fluten beispielsweise russische Netzwerke das Internet mit Propaganda, mit dem Ziel, KI-Chatbots zu "groomen" und sie dazu zu bringen, russische Fehlinformationen zu reproduzieren. KI kann auch dazu verwendet werden, realistische, aber gefälschte Online-Personas zu erstellen, die in sozialen Medien agieren und subtil oder offen politische Meinungen beeinflussen können. Ein konkretes Beispiel hierfür ist der Einsatz einer russischen Bot-Farm, die KI nutzte, um Tausende von gefälschten amerikanischen Profilen in sozialen Medien zu erstellen und über diese Anti-Ukraine- und Pro-Russland-Narrative in den Vereinigten Staaten zu verbreiten. Es gibt auch Hinweise darauf, dass pro-Kreml-Akteure versuchen, Suchmaschinenalgorithmen zu manipulieren, um die Verbreitung pro-Kreml-freundlicher Propaganda zu verstärken. Ähnliche Strategien der KI-gestützten Manipulation werden ebenfalls in anderen autokratisch regierten Ländern eingesetzt.⁴⁸

Auch in **Deutschland** gibt es – wie oben bereits ausgeführt - deutliche Beispiele für den Einsatz von Algorithmen und KI durch politische Akteure. Die **AfD** nutzt soziale Medien, insbesondere TikTok, sehr effektiv, um ihre Botschaften zu verbreiten und junge Wähler zu erreichen, wobei sie die Funktionsweise der Algorithmen zu ihrem Vorteil nutzt. Studien zeigen, dass die Algorithmen von TikTok und X (ehemals Twitter) in Deutschland unverhältnismäßig viele rechte, pro-AfD-Inhalte **an nicht-parteiische Nutzer ausspielen**. Darüber hinaus gibt es Berichte über den Einsatz von KI-generierten Inhalten, einschließlich KI-Influencern, durch die AfD im Rahmen ihrer Wahlkampagnen.⁴⁹ In den Vereinigten Staaten nutzen politische Akteure soziale Medien ebenfalls intensiv, und es gibt Hinweise darauf, dass Algorithmen dazu beitragen können, Wut und Zynismus in der politischen Auseinandersetzung zu schüren. Elon Musk, der Eigentümer von X, wird vorgeworfen, den Algorithmus der Plattform so angepasst zu haben, dass pro-Trump-

⁴⁵ Global Witness (20.02.2025). X and TikTok algorithms push pro-AfD content to non-partisan German users: New analysis. <https://globalwitness.org/en/press-releases/x-and-tiktok-algorithms-push-pro-afd-content-to-non-partisan-german-users-new-analysis/> (Zugriff am 15.03.25)

⁴⁶ Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag ((o.J.) Mit zunehmender Bedeutung für die Meinungsbildung geraten Algorithmen bei Onlinemedien in den Fokus von Forschung und Regulierung. https://www.tab-beim-bundestag.de/projekte_algorithmen-in-digitalen-medien-und-ihr-einfluss-auf-die-meinungsbildung.php (Zugriff am 02.10.22)

⁴⁷ University of Oxford (2021). Social media manipulation by political actors an industrial scale problem - Oxford report. <https://www.ox.ac.uk/news/2021-01-13-social-media-manipulation-political-actors-industrial-scale-problem-oxford-report> (Zugriff am 02.10.22)

⁴⁸ Rand (2024). Social Media Manipulation in the Era of AI. <https://www.rand.org/pubs/articles/2024/social-media-manipulation-in-the-era-of-ai.html> (Zugriff am 23.10.24)

⁴⁹ European Center for Digital Action (18.03.25). The social media domination of the German far right. <https://www.centerfordigitalaction.eu/post/report-the-social-media-domination-of-the-german-far-right>. (Zugriff am 23.04.25)

Inhalte verstärkt werden. Eine Studie ergab, dass neue Nutzer auf X, die kein Interesse an Politik bekundeten, mit politischen Inhalten überschwemmt wurden, die oft Donald Trump unterstützten.⁵⁰

Die **Kombination aus der Leistungsfähigkeit von KI zur Erstellung realistischer, schwer zu erkennender Fake-Inhalte und der algorithmischen Verstärkung in sozialen Medien** ermöglicht politischen Akteuren eine beispiellose Skalierbarkeit und Effektivität bei der Verbreitung von Propaganda und Desinformation. Diese Entwicklung untergräbt das Vertrauen in Online-Informationen und kann zu einer weiteren Polarisierung der Gesellschaft führen, indem sie die Verbreitung von Fakten und evidenzbasierten Argumenten erschwert. Die beobachteten Muster der algorithmischen Verstärkung rechter politischer Inhalte in Deutschland und der USA deuten auf eine mögliche systemische Voreingenommenheit in den Algorithmen oder deren Manipulation hin, was die politische Landschaft in diesen Ländern beeinflussen könnte. Diese Tendenz könnte zu einer Verschiebung des politischen Diskurses nach rechts führen und die Sichtbarkeit und Reichweite anderer politischer Perspektiven einschränken.

Anmerkungen zum Abschnitt

Emotionale Manipulation

Social-Media-Plattformen basieren auf komplexen Algorithmen, die Inhalte kuratieren, vorschlagen und priorisieren, um das Nutzerengagement (Likes, Kommentare, Shares) zu maximieren und Nutzer möglichst lange auf der Plattform zu halten⁵¹. Diese Algorithmen berücksichtigen verschiedene Faktoren wie Nutzerinteraktionen, Videoinformationen und Geräteeinstellungen⁵².

Die Ausrichtung der Algorithmen auf Engagement schafft Schwachstellen, die von Akteuren mit hoher "algorithmischer Kompetenz" gezielt ausgenutzt werden. Durch koordinierte Aktionen kann künstlich Engagement erzeugt werden, das von den Algorithmen als authentisch interpretiert und verstärkt wird. Dies ermöglicht es, bestimmte **Narrative oder Beiträge künstlich sichtbarer zu machen und ihre Reichweite zu erhöhen**⁵³. Spammer und Scammer nutzen beispielsweise KI-generierte Bilder, die oft visuell sensationell sind und Engagement anziehen, um ihre Reichweite auf Plattformen wie Facebook zu erhöhen⁵⁴.

Insbesondere bei visuellen Inhalten wie Bildern und Videos spielt das Engagement eine grosse Rolle für die algorithmische Verbreitung: Eine Studie der Stanford Universität beispielsweise hat 120 Facebook-Seiten untersucht, die jeweils mindestens 50 KI-generierte Bilder gepostet hatten, und klassifizierten die Seiten in die Kategorien Spam, Betrug und „andere Ersteller“. Einige waren koordinierte Gruppen von Seiten, die von denselben Administratoren betrieben wurden. Diese **Bilder erhielten insgesamt Hunderte von Millionen von Zugriffen**. Ein Beitrag, der ein KI-generiertes Bild enthielt, gehörte im 3. Quartal 2023 zu den 20 meistgesehenen Inhalten auf Facebook (mit 40 Millionen Aufrufen). Spam-Seiten verwendeten Clickbait-Taktiken und versuchten, Nutzer zu plattformfremden Inhaltsfarmen und minderwertigen Domains zu leiten⁵⁵.

Wir haben gesehen, wie stark TikTok von Jungen genutzt wird und wie verbreitet Falschinformationen auf dieser Plattform sind. Ika Idris weist in einer Analyse auf die psychologischen Wirkungsmechanismen hin, die von TikTok genutzt werden⁵⁶: Die multimodale Natur visueller Inhalte fördere eine schnelle Aufnahme von Inhalten. Auf TikTok verbindet sich die leicht verdauliche visuelle Darstellung mit dem einfachen Wechseln von Videos per Daumenwisch, was das Betrachten mühelos und passiv macht. Visuelle Botschaften, insbesondere „Bildausschnitte“, beeinflussen die

⁵⁰ Benton, J. (2025). In Germany, social media algorithms are pumping out huge amounts of far-right, pro-AfD content. <https://www.niemanlab.org/2025/02/in-germany-social-media-algorithms-are-pumping-out-huge-amounts-of-far-right-pro-afd-content/> (Zugriff am 18.03.25)

⁵¹ Bundeszentrale für politische Bildung (2025). Systematische Manipulation sozialer Medien im Zeitalter der KI. <https://www.bpb.de/shop/zeitschriften/apuz/wahlkampf-2025/558872/systematische-manipulation-sozialer-medien-im-zeitalter-der-ki/> (Zugriff am 2.02.25)

⁵² Jourdan Aldredge (2024). How Does the TikTok Algorithm Work? Tips and Insights to Boost Your Content Strategy, <https://www.soundstripe.com/blogs/how-does-the-tiktok-algorithm-work> (Zugriff am 09.02.25)

⁵³ Katja Muñoz, Influencers and their Ability to Engineer Collective Online Behavior. A Boon and Challenge for Politics, Deutsche Gesellschaft für Auswärtige Politik, DGAP Policy Brief 23/2024.

⁵⁴ Renee DiResta & Josh A. Goldstein (2024). How Spammers, Scammers and Creators Leverage AI-Generated on Facebook for Audience Growth. Stanford University. <https://cyber.fsi.stanford.edu/news/ai-spam-accounts-build-followers> (Zugriff am 04.05.24)

⁵⁵ Renee DiResta & Josh A. Goldstein (2024). How Spammers, Scammers and Creators Leverage AI-Generated on Facebook for Audience Growth. Stanford University. <https://cyber.fsi.stanford.edu/news/ai-spam-accounts-build-followers> (Zugriff am 04.05.24)

⁵⁶ Idris, I. (o.J.). Deceptive Trends: The Societal Impact of Disinformation on TikTok. <https://www.internationalaffairs.org.au/australianoutlook/deceptive-trends-the-societal-impact-of-disinformation-on-tiktok/> (Zugriff am 04.05.24)

Meinungsbildung von Wählern und verbessern die Verankerung im Gedächtnis effektiver als verbale Botschaften. Da visuelle Informationen auf TikTok direkter und müheloser verarbeitet werden, werden sie als präzise wahrgenommen und besser ins Gedächtnis integriert. Diese Effektivität hat jedoch auch eine Kehrseite: Irreführende visuelle Darstellungen können aufgrund der „Realismus-Heuristik“, bei der Menschen Bilder und Audio als realistischer wahrnehmen als Text, falsche Wahrnehmungen erzeugen. Solche videobasierten Plattformen werden durch einen „Fluency“-Effekt verstärkt: Bekannte visuelle Darstellungen, wie beispielsweise Deepfake-Videos bekannter Persönlichkeiten, werden unabhängig von ihrer Genauigkeit eher als zutreffend akzeptiert, was die Bekämpfung von Fehlinformationen erschwert (Idris o.J.).

Anmerkungen zum Abschnitt

Clickbait

Studien zeigen, dass ein verstärktes Nutzer-Engagement besonders durch angst- oder wutinduzierende multimediale Video-plus-Text-Desinformationen erzeugt wird (Lee et al 2024)⁵⁷. Zu den emotionalisierenden Strategien gehören Clickbaits, reißerische Aufmacher und Überschriften, die Nutzer:innen zum Klick verleiten sollen (oft mit irreführendem Inhalt). Die Wissenschaft zeigt, dass Clickbait-Headlines deutlich mehr Klicks erzeugen als sachliche Titel (Zannettou et al 2019)⁵⁸. Studien betonen, dass Plattformen like Facebook/YouTube zwar gegen Clickbait vorgehen (Abwertung dieser Inhalte), aber trotzdem Millionen Nutzer durch Clickbait-Angebote gelenkt werden (insb. auf alternativen News-Sites). Gesamthaft zeigt die Forschung, dass Clickbait ein gängiges Mittel der Desinformation ist – es lockt Nutzer und verschiebt Aufmerksamkeit, oft ohne sachlichen Mehrwert (Jacobo-Morales et al. 2024⁵⁹).

Anmerkungen zum Abschnitt

Gesellschaftliche Auswirkungen

Die Verstärkung politischer Botschaften über Social Media führt zu einer zunehmenden Fragmentierung der öffentlichen Debatte⁶⁰. „Während traditionelle Medien noch einen gemeinsamen Referenzrahmen für gesellschaftliche Diskussionen boten, entstehen unsere Informationsräume durch die algorithmische Gestaltung zunehmend isoliert voneinander. Die KI-gestützte Hyperpersonalisierung von Inhalten schafft individuelle Informationswelten, die sich von Person zu Person fundamental voneinander unterscheiden können. Menschen, die formal in der gleichen Gesellschaft leben, bewegen sich faktisch in völlig unterschiedlichen Informationsökosystemen. Verstärkt wird diese Entwicklung durch die Engagement-Optimierung der Plattformen, die polarisierende Inhalte bevorzugt und damit zu einer systematischen Verstärkung extremer Positionen führt. Diese technologisch gestützte Polarisierung untergräbt zunehmend die Möglichkeit eines ausgewogenen gesellschaftlichen Diskurses⁶¹.

Desinformation hat weitreichende gesellschaftliche Effekte. Untersuchungen zeigen, dass verbreitete Fake News und manipulierte Inhalte zu Polarisierung, Vertrauensverlust und Radikalisierung beitragen können. Eine Befragung (Pew 2019⁶²) ergab z.B., dass **Misstrauen in traditionelle Medien und Politik zugenommen hat, weil Nutzer:innen kaum zwischen Information und Desinformation unterscheiden** können. Allen et al. (2020, Harvard) fanden, dass Fake-

⁵⁷ Lee, J.; Hameleers, M.; Shin, S.Y. (2024). The emotional effects of multimodal disinformation: How multimodality, issue relevance, and anxiety affect misperceptions about the flu vaccine. University of Amsterdam. [The emotional effects of multimodal disinformation.pdf](#) (Zugriff am 12.05.24)

⁵⁸ Zannettou, S., Caulfield, T. et al. (2019). The web of false information: Rumors, fake news, hoaxes, clickbait, and various shenanigans. In. (Journal of Data and Information Quality 2019, 11(3)) https://www.researchgate.net/publication/324435607_The_Web_of_False_Information_Rumors_Fake_News_Hoaxes_Clickbait_and_Various_Other_Shenanigans (Zugriff am 12.05.24)

⁵⁹ Jacobo-Morales, A. & Marino-Jiménez, R. (2024): Clickbait: Research, challenges and opportunities. Online Journal of Communication and Media Technologies, 2024, 14(4). <https://www.ojcm.net/download/clickbait-research-challenges-and-opportunities-a-systematic-literature-review-15267.pdf#:~:text=,15> (Zugriff am 12.05.24)

⁶⁰ Tabela Timse, Kyle Findlay & Aldu Cornelissen (2024). Influence-for-hire trend is distorting public discourse, poses threat to foundations of democracy. Daily Maverick. <https://www.dailymaverick.co.za/article/2024-05-26-influence-for-hire-distorts-public-discourse-threatens-democracy/> (Zugriff am 26.06.24)

⁶¹ Bundeszentrale für politische Bildung (2025). Systematische Manipulation sozialer Medien im Zeitalter der KI. <https://www.bpb.de/shop/zeitschriften/apuz/wahlkampf-2025/558872/systematische-manipulation-sozialer-medien-im-zeitalter-der-ki/> (Zugriff am 2.02.25)

⁶² Pew Research Center (2019): Many Americans Say Made-Up News Is a Critical Problem“. https://www.pewresearch.org/wp-content/uploads/sites/20/2019/06/PJ_2019.06.05_Misinformation_FINAL-1.pdf (Zugriff am 07.08.24)

News-Konsum die allgemeine Medienvertrauensbereitschaft sinken lässt⁶³. Der Edelman Trust Barometer 2021 konstatiert „eine Epidemie von Fehlinformation und weit verbreitetem Misstrauen in soziale Institutionen“ und stellt in seinem Jahresbericht fest, dass die traditionellen Medien (53 Prozent) **mit acht Prozentpunkten weltweit den größten Vertrauensverlust verzeichnet haben**⁶⁴. Desinformation hat das Vertrauen in Medien, Regierung und Wissenschaft deutlich untergraben. Eine Harvard-Studie⁶⁵ zeigte, dass **Konsum von Fake News mit signifikant geringerem Vertrauen in klassische Nachrichtenmedien einhergeht**, und zwar parteiübergreifend. In dieser Langzeitstudie berichteten Personen mit häufigem Desinformationskonsum deutlich weniger Vertrauen in Medien. Interessanterweise stellte die gleiche Studie fest, dass konservative Amerikaner durch Fake News manchmal *mehr* Vertrauen in ihre Regierung gewannen (unter Trump), während progressive Befragte eher misstrauischer wurden. Generell zeigen viele Umfragen sinkendes Vertrauen: Eine Reuters-Institute-Studie (2020⁶⁶) ergab, dass **60 % der Briten glaubten, ihre Regierung täusche sie über die Herkunft von COVID-19, und 40 % vermuteten eine absichtliche Verbreitung durch mächtige Eliten**. Solche Zahlen zeigen ein stark verbreitetes Misstrauen gegenüber Institutionen – befeuert durch Falschinformationen.

Inzwischen deuten Studien auf messbare Dynamiken hin: In Ländern mit hoher Desinformationsbelastung wächst die **Zustimmung zu extremen Positionen (Verstärker-Effekt von Filterblasen)**. Auf individueller Ebene zeigen Befunde, dass Exposition gegenüber Falschnachrichten insbesondere bei Menschen mit niedriger Medienkompetenz Ängste und Vorurteile schürt. Dadurch können konstruktive Debatten erschwert werden: Die Rutgers-Studie (2020) zeigte, dass **generelle Exposition Fake-News das Vertrauen in Mainstream-Medien über Parteigrenzen hinweg senkt**. Gerade in Zeiten gesellschaftlicher Krisen (Kriege, Pandemie) kann sich Misstrauen deutlich verstärken. Medienwissenschaftler sprechen von „Trust Crisis“. Die demografische Analyse offenbart: Ältere Menschen (60+) waren besonders anfällig für Corona-Fake-News (Pew 2019). Insgesamt betonen Fachleute, dass Desinformation das Vertrauen in Demokratie, Medien und Wissenschaft untergräbt und so langfristig politischen Extremismus fördern kann⁶⁷.

Die Verbreitung von Desinformation in sozialen Medien trägt nachweislich zu einer **Zunahme extremer Positionen** bei. Aktuelle Studien zeigen einen klaren Zusammenhang zwischen Desinformation in sozialen Medien und der Zunahme extremistischer Ansichten. So ergab eine Befragung des Pew Research Centers von 2022⁶⁸, dass **14 % der Erwachsenen angaben, bewusst falsche politische Informationen online geteilt zu haben**. Diese Personen neigten deutlich häufiger zur Befürwortung politischer Gewalt und unterstützten extremistische Gruppen wie QAnon, die Proud Boys oder weiße Nationalisten. Obwohl ein direkter Kausalzusammenhang schwer nachzuweisen ist, deuten diese Korrelationen darauf hin, dass **Falschinformationen Polarisierung begünstigen**. 69 % der US-Amerikaner sagen außerdem, das Internet habe politische Gespräche „weniger zivilisiert“ gemacht. In einer weiteren Studie erklären im Schnitt **64 %, soziale Medien hätten negative Effekte auf das Zusammenleben** (PEW 2020⁶⁹). Diese Wahrnehmung stützt die Annahme, dass soziale Medien extreme Positionen verstärken. Dass die algorithmische Kuratierung von Inhalten auf Social-Media-Plattformen die politische Polarisierung verschärft, ist ein Phänomen, das sich weltweit zeigt. So stellt eine kürzlich durchgeführte Studie aus Pakistan (2023) fest, dass der zunehmende Einsatz von Social-

⁶³ Allen, W., de Melo, G., Shilling, F., Rayo, C. & Rahwan, I. (2020): Exposure to misinformation is linked to lower trust in media. (Harvard Misinformation Review 2020): misinfoview.hks.harvard.edu. (Zugriff am 02.03.24)

⁶⁴ Edelman (2021): Trust Barometer 2021. [Edelman Trust Barometer 2021 | Edelman](https://www.edelman.com/trust-barometer-2021). (Zugriff am 04.03.24)

⁶⁵ Ognyanova, K. et al. (2020). Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. Harvard Kennedy School – Misinformation Review 2020. <https://misinfoview.hks.harvard.edu/article/misinformation-in-action-fake-news-exposure-is-linked-to-lower-trust-in-media-higher-trust-in-government-when-your-side-is-in-power/> (Zugriff am 04.03.24)

⁶⁶ Nielsen, R. K. et al. (2020). Navigating the 'Infodemic': How People in Six Countries Access and Rate News and Information about Coronavirus. Oxford Reuters Institut. <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-04/Navigating%20the%20Coronavirus%20Infodemic%20FINAL.pdf> (Zugriff am 04.03.24)

⁶⁷ World Economic Forum (2023): Report "Global Risks". <https://www.weforum.org/publications/global-risks-report-2023/> (Zugriff am 21.04.24)

⁶⁸ Pew Research Center (2022). Partisan Divides in Views of Political Violence, False Information. Washington, D.C. <https://www.pewresearch.org/politics/2022/08/11/partisan-divides-in-views-of-political-violence-false-information/> (Zugriff am 18.04.24)

⁶⁹ Pew Research Center (2020). 64% of Americans say social media have a mostly negative effect on the way things are going in the U.S. today. Pew Research Center, Washington, D.C. <https://www.pewresearch.org/short-reads/2020/10/15/64-of-americans-say-social-media-have-a-mostly-negative-effect-on-the-way-things-are-going-in-the-u-s-today/> (Zugriff am 18.04.24)

Media-Algorithmen die politische Polarisierung im Land weiter verschärfen dürfte.⁷⁰ Internationale Analysen zeigen, dass die weltweite Polarisierung (Aussagen „die Gesellschaft ist stark gespalten“) seit den 2000er Jahren deutlich zugenommen hat – besonders ausgeprägt in angloamerikanischen Ländern⁷¹. In den letzten 5 Jahren verstärkte sich dies noch weiter: Umfragen etwa in Großbritannien ergaben, dass über 80 % der Bevölkerung Gesellschaftsspaltungen wahrnehmen; viele machen online kursierende Hasskommentare und Verschwörungspropaganda mitverantwortlich.

Weitere Studien zeigen, dass Online-Plattformen, insbesondere solche mit audiovisuellen Inhalten, von rechtsextremen Akteuren intensiv für strategische Kommunikation genutzt werden. Dabei ist festzustellen, dass extremistische Inhalte auf allen Plattformen präsent sind, einige Angebote jedoch besonders gute Voraussetzungen für deren Verbreitung bieten. So hat das Vorherrschen **extremistischer Inhalte auf Plattformen wie X** (vormals Twitter) in jüngster Zeit zugenommen, während TikTok in der Vergangenheit vergleichsweise weniger intensiv gegen solche Inhalte vorgegangen ist⁷².

Eine Studie des Institute for Strategic Dialogue (ISD) belegt, dass extremistische Inhalte, insbesondere **neo-nazistische und faschistische Inhalte, auf TikTok millionenfach abgerufen** werden konnten. Dies ermöglichte es Propagandisten, ein Netzwerk von pro-nazistischen Konten aufzubauen. In einer Stichprobe von 50 Konten, die dem Unternehmen gemeldet wurden und zusammen über 6,2 Millionen Aufrufe generierten, wurde keines sofort gesperrt. Innerhalb eines Monats wurden lediglich 23 dieser Konten gesperrt, nachdem sie zuvor mindestens zwei Millionen Aufrufe gesammelt hatten. Es wurde beobachtet, dass der Algorithmus der "For You Page" auf TikTok nach dem Ansehen von nur zehn Videos von pro-nazistischen Konten sofort begann, ähnliche Inhalte von zuvor unentdeckten Konten zu empfehlen⁷³.

Die Studie "Hass in den sozialen Medien hat stark zugenommen" aus dem Jahr 2024 zeigt, dass **89 Prozent der Internetnutzer in Deutschland der Aussage zustimmen, dass Hass im Netz in den letzten Jahren zugenommen hat**. Dabei bezieht sich Hass am häufigsten auf politische Ansichten (41 Prozent) oder das Aussehen (37 Prozent). Diese Zunahme von Hass und extremen Positionen **gefährdet den demokratischen Diskurs und die Meinungsvielfalt im öffentlichen Raum**.⁷⁴

Die veränderten Kommunikationsbedingungen und Gruppennormen in sozialen Online-Medien begünstigen ein negatives Kommunikationsverhalten und tragen somit zum **Rückgang konstruktiver Dialoge** bei. Die erhöhte Sichtbarkeit und öffentliche Zugänglichkeit von Debatten online führen dazu, dass Nutzer häufiger mit unzivilisierter Kommunikation in Kontakt kommen. Die vereinfachten Möglichkeiten zur schnellen und weitreichenden Verbreitung von Inhalten verstärken diesen Effekt, da emotional aufgeladene und polarisierende Inhalte tendenziell eine große Reichweite erzielen (vgl. Kumpel & Riege 2019)⁷⁵. Die (relative) Anonymität vieler Online-Plattformen kann zu einer Enthemmung und einem verringerten Verantwortungsbewusstsein führen, was Nutzer dazu verleitet, sich online anders zu verhalten als in persönlichen.

Viele Studien berichten, dass politische Diskussionen in sozialen Medien deutlich an Qualität verloren haben. Besonders in den USA herrscht das Gefühl, Online-Diskurse seien polarisiert und feindselig geworden. Fachliteratur erklärt dies u. a. mit Empfehlungsalgorithmen, die Polarisierung und Provokation belohnen, während differenzierte Argumentation untergeht⁷⁶.

Verschwörungsideologien haben in den letzten Jahren massiv zugenommen, häufig durch algorithmisch verbreitete Inhalte. Laut Pew Research **stieg die Bekanntheit der QAnon-Bewegung in den USA von 23 % (Anfang 2020) auf 47 %**

⁷⁰ Javed, U. & Javed, U. (2023). The Influence of Social Media Algorithms on Political Polarization and Public Opinion. https://www.researchgate.net/publication/387068259_The_Influence_of_Social_Media_Algorithms_on_Political_Polarization_and_Public_Opinion (Zugriff am 18.04.24)

⁷¹ Blattner, A. (2025). It's time to unite against polarization. [Allianz | It's time to unite against polarization](#). (Zugriff am 18.04.25)

⁷² Sophia Rothut, Darian Harff & Cornelius Puschman (2024). Rechtsextreme Narrative und Social-Media-Influencende. BzKJAKTUELL 2/2024. [Rechtsextreme Narrative und Social-Media-Influencende](#) (Zugriff am 12.12.24)

⁷³ Joseph König (2024). Report finds extremist content getting millions of views on TikTok as app works to rout it. [Report: Extremist content gets millions of views on TikTok](#) (Zugriff am 12.12.24)

⁷⁴ Rbb (2024). Hass in den sozialen Medien hat stark zugenommen. [Neue Studie: Hass hat in den sozialen Medien stark zugenommen | rbb24](#) (Zugriff am 19.03.24)

⁷⁵ Anna Sophie Kumpel & Diana Riege. Wandel der Sprach- und Debattenkultur in sozialen Online-Medien Ein Literaturüberblick zu Ursachen und Wirkungen von inzivilisierter Kommunikation. Hrsg. Konrad-Adenauer-Stiftung e. V., 2019, Berlin. [6a76553c-7c30-b843-b2c8-449ba18c814e](#) (Zugriff am 19.03.24)

⁷⁶ Pew Research Center (2022). *Most Americans say social media makes political discourse worse*. <https://www.pewresearch.org/short-reads/2020/11/16/5-facts-about-the-qanon-conspiracy-theories/> (Zugriff am 18.04.24)

im September 2020⁷⁷. Viele Studien zeigen zudem, dass Anti-Impfmythen („Microchips im Impfstoff“, „DNA-Veränderung“ usw.) innerhalb kürzester Zeit Millionen Menschen erreichten. Umfragen zeigen, dass ein signifikanter Prozentsatz von Teenagern (81%) in den USA an mindestens eine Verschwörungstheorie glaubt und ihnen häufig in sozialen Medien begegnet⁷⁸.

Soziale Medien spielen eine Rolle bei der Verbreitung von Fehlinformationen und Verschwörungstheorien im Zusammenhang mit COVID-19. Algorithmen begünstigen diese Dynamik: Wer mit extremen Inhalten interagiert, bekommt auf TikTok und YouTube häufig noch radikalere Empfehlungen – ein bekanntes „Echokammer“-Phänomen.

Wie die auf Engagement und lange Nutzerdauern ausgerichteten Designs der Social Media-Plattformen „bei Teilen der Nutzerschaft zu einer verstärkten Polarisierung sowohl im Hinblick auf wahrgenommene Inhalte als auch im Hinblick auf gesellschaftliche und politische Einstellungen“ führen, haben Lischka und Stöcker (2018, S. 8)⁷⁹ in einem Arbeitspapier zur Digitalen Öffentlichkeit ausführlich theoretisch und empirisch beschrieben. „Die Wechselwirkung algorithmisch optimierter "Iss weiter!"-Plattformen mit unseren ungesunden Medienvorlieben bringt im Zweifel unsere schlimmsten Abgründe zum Vorschein. Bei YouTube laufen Videos über [Selbstmordattentäter](#) gut, , [Verschwörungstheorien](#). All das, was man Facebook seit der US-Wahl vorgeworfen hat, trifft auf YouTube mindestens ebenso zu - was daran liegt, dass beide Plattformen von Algorithmen sortiert werden, die auf "Engagement" hin optimieren. Dabei werden verhaltenspsychologische Methoden eingesetzt, die [Interaktion ohne Denken](#) möglichst einfach machen sollen (Stöcker 2018⁸⁰).

Anmerkungen zum Abschnitt

Schulrealität hinkt hinterher

Studien zum Umgang mit Desinformation in Schulen belegen großen Nachholbedarf. PISA 2022 (Auswertung 2025) fand, dass nur 47 % der 15-Jährigen in Deutschland Vertrauen darin haben, die Qualität von Online-Informationen beurteilen zu können⁸¹. Die Befunde zeigen einen erheblichen Mangel an kritischer Medienkompetenz. Gleichzeitig berichten Befragungen, dass Schulen und Lehrkräfte oft schlecht auf digitale Desinformation vorbereitet sind. Eine Studie zur Nachrichtenkompetenz (Universität Halle 2022) ergab, dass Lehrpläne und Schulbücher kaum Fake-News-Themen enthalten, und Lehrkräfte nur in 1/3 der Schulen entsprechende Einheiten anbieten⁸². Deutschlandweiten Untersuchungen (z.B. JIM-Studie 2024⁸³) zufolge gehen bei Jugendlichen geringere Bildung mit höherer Informationsvermittlung bei Social Media einher, was zu einer größeren Gefährdung durch Fake News führt.

⁷⁷ Pew Research Center (2020). 5 facts about the QAnon conspiracy theories. <https://www.pewresearch.org/short-reads/2020/03/30/qanons-conspiracy-theories-have-seeped-into-u-s-politics-but-most-dont-know-what-it-is/> (Zugriff am 18.04.24)

⁷⁸ Government Technology (2024). Teens Frequently See Conspiracy Theories on Social Media. <https://www.govtech.com/public-safety/study-teens-frequently-see-conspiracy-theories-on-social-media> (Zugriff am 18.11.24)

⁷⁹ Lischka, K. & Stöcker, Ch. (2017). Digitale Öffentlichkeit Wie algorithmische Prozesse den gesellschaftlichen Diskurs beeinflussen – Arbeitspapier. Bertelsmann Stiftung. [Digitale Öffentlichkeit - Wie algorithmische Prozesse den gesellschaftlichen Diskurs beeinflussen](#) (Zugriff am 23.01. 2020)

⁸⁰ Stöcker, Ch. (2018). YouTube, Facebook & Co. Sehend ins Verderben. SPIEGEL, 11.02.2018. <https://www.spiegel.de/wissenschaft/mensch/youtube-facebook-co-sehend-ins-verderben-kolumne-a-1192615.html> (Zugriff am 14.02.18)

⁸¹ PISA-Studie 2022 (Auswertung 2025, TUM) (Deutschland): Jugendliche unsicher bei der Beurteilung von Online-Informationen. [PISA-Studie: Jugendliche unsicher bei der Beurteilung von Online-Informationen - TUM](#) (Zugriff am 28.01.25)

⁸² Universität Halle (2022). Nachrichtenkompetenz – Kinder und Jugendliche sicher durch die Informationsflut navigieren <https://deutschdidaktik.germanistik.uni-halle.de/digitale-didaktik-nachrichtenkompetenz/> (Zugriff am 14.10.23)

⁸³ JIM-Studie 2024 (mpfs): Jugend, Information, Medien.