Hannah Metzler

Mythos Fehlinformation

Warum soziale Medien weniger gefährlich sind, als wir denken

Sorgen über die Folgen von Desinformation für die Demokratie sind weit verbreitet. Doch die öffentliche Wahrnehmung und die wissenschaftlichen Daten klaffen bei diesem Thema weit auseinander. Aufklärung ist wichtig, damit sich öffentliche Aufmerksamkeit und gesetzliche Regulierungen auf Lösungen für die tatsächlichen Probleme konzentrieren.

Viele weit verbreitete Annahmen über Social Media lassen katastrophale Folgen für unsere Gesellschaft befürchten. Dazu gehören Sorgen vor Fake News und anderen Formen der Fehlinformation, wie etwa absichtlich verbreiteter Desinformation. Wir glauben zum Beispiel, dass Social Media uns – und vor allem andere – mit einer Unmenge an sensationsheischenden, polarisierenden und faktisch fragwürdigen Inhalten überfluten. Ebenso hält der Glaube, dass Donald Trump die Präsidentschaftswahl mit Hilfe russischer Propaganda gewann, die dank Cambridge Analytica genau auf einzelne Nutzer:innen zugeschnitten werden konnte. Wir befürchten auch, dass wir alle in digitalen Echokammern gefangen sind, die permanent unsere eigenen Sichtweisen bestätigen. Es gibt Geschichten darüber, wie Menschen auf YouTube durch immer extremere Video-Empfehlungen in eine Radikalisierungsspirale hineingezogen werden. Das Ganze scheint von Algorithmen angetrieben zu werden, die nur ein Ziel haben: uns möglichst lange auf einer Plattform zu halten und Klicks zu generieren. Extreme, schockierende und hasserfüllte Inhalte können das am besten. So spaltet sich die Gesellschaft zunehmend, befürchten wir, und immer mehr Menschen wählen rechtsextreme Parteien. Sachliche Debatten zu dringenden Fragen sind kaum noch möglich. Kurz: Die Demokratie ist in Gefahr.

Die ernstzunehmende Sorge, dass Social Media – durch Algorithmen und künstliche Intelligenz (KI) noch weiter angefeuert – im großen Stil Einstellungen manipulieren, steht in starkem Kontrast zu unserer Erfahrung mit Meinungsänderungen im Alltag. Wenn wir versuchen, jemanden von etwas zu überzeugen – oder jemand versucht, uns zu überzeugen –, gelingt das in der Regel nicht. In Diskussionen rund um Politik, die COVID-Pandemie, den Klimawandel und Co. entsteht leicht das Gefühl, gegen eine Wand anzureden: Jeder glaubt, Recht zu haben, keiner hört dem anderen geduldig zu. Auch wir selbst ändern unsere Meinung nur selten.

Es gibt eine wissenschaftlich belegte Erklärung für diesen scheinbaren Starrsinn: Wir sammeln ein Leben lang Erfahrungen und Wissen, die sich zu unserem Weltbild zusammensetzen. Alles interpretieren wir durch diese Brille – und auch Menschen, die uns nahestehen, sehen die Welt so. Gegensätzliche Argumente in einzelnen Gesprächen oder Medienberichten wiegen da wenig. Es wäre auch ziemlich töricht, einfach etwas zu glauben, das mit unserem restlichen Wissen nicht zusammenpasst.

Meinungsänderungen im Alltag sind also selten, weil es so schwer ist, andere zu überzeugen. Wenn Menschen ihre Meinung aber nicht leicht ändern, wie können dann Social Media die Einstellung ganzer Gesellschaftsgruppen manipulieren? Und warum halten die vom Weltwirtschaftsforum befragten Expert:innen Desinformation für die größte globale Gefahr der nächsten zwei Jahre? Irgendetwas scheint hier nicht zusammenzupassen.



Fehlinformation betrifft nur eine kleine Minderheit

Die öffentliche, mediale und politische Meinung zu Social Media und Algorithmen ist weit von dem entfernt, was wissenschaftliche Studien nahelegen.

Ein erster Mythos betrifft die Annahme, dass Fehlinformationen weit verbreitet sind. Je nach Plattform, Land und Jahr machen Fehlinformationen nur zwischen 0,1 und 6,7 % des Nachrichtenkonsums aus. Dazu kommt, dass nur wenige Menschen überhaupt Nachrichten in Social Media lesen. Auf Facebook in den USA machen politische Nachrichten zum Beispiel nur 3 % aller geteilten Posts aus – und davon stammen wiederum nur 1 bis 6 % aus wenig vertrauenswürdigen Quellen.

Hinzu kommt, dass nur wenige Nutzer:innen für den Großteil der Interaktion mit Fehlinformationen verantwortlich sind. Bei der US-Präsidentschaftswahl 2016 teilten 0,1 % der Twitter-User 80 % der Nachrichten aus unzuverlässigen Quellen – und 1 % der User machte 80 % der Kontakte mit diesen Nachrichten aus. Bei der Wahl 2020 waren es insgesamt 2.107 sogenannte Supersharer; sie erreichten damit 5 % aller registrierten Wähler:innen. Die überwiegende Mehrheit der Nutzer:innen sieht also noch viel weniger Fehlinformationen, als die Durchschnittszahlen vermuten lassen.

Wer sind die Menschen, die den Großteil dieser Fehlinformationen lesen und teilen? Es sind Menschen mit extremen politischen Einstellungen. Sie haben tendenziell ein niedriges Vertrauen in Institutionen wie Politik, Medien, Wissenschaft und Medizin und fühlen sich von ihnen nicht vertreten. Einige Studien zeigen, dass unter den Supersharern in westlichen Demokratien mehr Menschen mit rechter politischer Einstellung sind – vermutlich auch, weil ihre Sichtweisen in traditionellen Medien und Institutionen wie der Wissenschaft tatsächlich seltener abgebildet werden.

Algorithmen und Echokammern im Faktencheck

Auch wenn wir also nicht alle von Desinformation überflutet werden, bleibt die Sorge vor einer Radikalisierung durch Algorithmen und digitale Echokammern. Bis vor Kurzem gab es nur wenige Studien zu den Effekten von Algorithmen, da einfach die Daten fehlten. In den letzten Jahren zeigten jedoch über zehn Studien zu Twitter, Facebook, Instagram, YouTube und Google relativ durchgängig: Die Nachfrage nach extremen Inhalten spielt die größte Rolle. Menschen mit extremen Einstellungen suchen aktiv nach Inhalten, die ihr Misstrauen gegenüber dem System bestätigen. Gruppen auf Facebook, Telegram und Co. sind Orte, an denen Menschen mit extremen, oft rechten Einstellungen diese Bestätigung finden. Algorithmen hingegen schlagen den meisten Nutzer:innen moderate Posts und Videos vor – auf YouTube lässt sich beispielsweise sogar ein linker politischer Bias beobachten. Extreme Inhalte empfehlen sie vor allem Menschen, die bereits viele solcher Inhalte konsumiert haben.

Eine besonders elegante Studie ließ „kontrafaktische Bots“ zunächst echte YouTube-Nutzer:innen imitieren und anschließend ausschließlich den Empfehlungen des Algorithmus folgen: Diese Empfehlungen führten im Schnitt zu *moderateren* Inhalten – und zwar umso stärker, je extremer das Sehverhalten des Ausgangsprofils war. Die Klicks auf radikale YouTube-Videos stammen meist von externen Links oder von Kanälen, denen die Nutzer:innen selbst folgen.

Bei einer der vier groß angelegten „Facebook and Instagram Election Studies“ wurde vor den US-Wahlen 2020 (Biden gegen Trump) der Algorithmus im Feed von 45.000 Nutzer:innen für drei Monate deaktiviert. Der Effekt war groß: Die Teilnehmer:innen verbrachten weniger Zeit online und sahen mehr Inhalte von Gruppen statt von Freunden sowie mehr politische Inhalte. Dabei stammten weniger Posts aus gleichgesinnten oder politisch konträren Quellen, dafür mehr aus moderaten. Leider enthielt der Feed ohne Algorithmus auch etwas mehr Inhalte aus unzuverlässigen Quellen (4,4 % statt 2,6 %).

Trotz all dieser Veränderungen im Feed änderten sich weder die politische Einstellung noch das Wahlverhalten oder die Gefühle gegenüber der Gegenseite (die sogenannte affektive Polarisierung) – entgegen den Erwartungen der Forschung. Das Abschalten des Algorithmus verringerte die Polarisierung nicht.

Diese Nullbefunde wurden mittlerweile jedoch relativiert. Zum einen hatte Meta genau im Studienzeitraum wegen der US-Wahlen kurzfristig Maßnahmen aktiviert, die Inhalte aus unzuverlässigen Quellen herabstuften und die Aufrufe von Fehlinformationen um mindestens 24 % senkten – gemessen wurde also teilweise ein bereits entschärfter Algorithmus. Zum anderen zeigen zwei neuere Feldexperimente, die per Browser-Erweiterung ganz ohne Kooperation mit den Plattformen auskamen, sehr wohl Effekte.



In einem Experiment mit knapp 5.000 X-Nutzer:innen verschob der Wechsel vom chronologischen zum algorithmischen Feed über sieben Wochen politische Positionen messbar nach rechts: Die Nutzer:innen priorisierten um 4,7 Prozentpunkte häufiger Themen wie Inflation, Immigration und Kriminalität. Der Effekt blieb selbst dann bestehen, nachdem der Algorithmus wieder abgeschaltet wurde – und der umgekehrte Wechsel zeigte keine vergleichbare Wirkung. Das könnte erklären, warum das bloße *Abschalten* des Algorithmus bei Facebook und Instagram nichts bewirkte. Die gefühlte Polarisierung veränderte sich allerdings auch hier nicht.

Für die gefühlte Polarisierung könnte es dagegen entscheidend sein, welche Inhalte ein Algorithmus verstärkt. In einem zehntägigen Experiment auf X ließen Forschende ein Sprachmodell den Feed neu sortieren und feindselige, antidemokratische Beiträge entweder hoch- oder herabstufen. Das veränderte die Gefühle gegenüber der politischen Gegenseite um gut zwei Punkte auf einer 100-Punkte-Skala – in beide Richtungen, bei Linken wie Rechten gleichermaßen.

Zwei Punkte klingen zwar wenig, entsprechen aber ungefähr dem, was sich in der US-Bevölkerung über Jahre hinweg aufgebaut hat. Was die Forschenden hier künstlich herstellten, geschieht im Alltag von selbst: Beiträge über die politische Gegenseite werden etwa doppelt so oft geteilt wie Beiträge über die eigene. Algorithmen, die auf Interaktion optimieren, stufen sie deshalb ohnehin bevorzugt hoch.

Echokammern sind offline tendenziell stärker ausgeprägt als online. Wie stark sie im Netz wirken, variiert jedoch stark: Während 23,1 % der Facebook-Nutzer:innen weniger als ein Viertel ihrer Inhalte aus Quellen sehen, die mit ihrer eigenen Sichtweise übereinstimmen, stammen die Beiträge bei weiteren 20,6 % der Nutzer:innen zu mehr als 75 % aus der eigenen Blase. Das Auflösen von Echokammern ist zudem nicht unbedingt hilfreich: Ein Experiment auf Twitter zeigte, dass Nutzer:innen noch stärker polarisiert werden, wenn sie einen Monat lang Posts von Meinungsführer:innen der anderen politischen Seite ausgesetzt sind. Der Kontakt mit extremen Meinungen in Social Media führt also eher zu Konflikten als zu gegenseitigem Verständnis.

Die reale Wirkung von Medien und Propaganda

Auch Propaganda, Werbung und politische Kampagnen scheinen Einstellungen kaum zu verändern. Zwei Studien zeigen etwa, dass die russischen Desinformationsattacken während der US-Wahlen 2016 weder das Wahlverhalten noch die Polarisierung messbar beeinflusst haben. Warum? Die Nutzer:innen sahen viel mehr Mainstream-News und Posts von Politiker:innen als Inhalte russischer Accounts. Zudem waren die wenigen, die den Großteil dieser Desinformation sahen, ohnehin schon überzeugte Trump-Wähler:innen und bereits stark polarisiert – auch hier handelte es sich also um eine kleine, politisch motivierte und extreme Minderheit.



Im Gegensatz dazu können groß angelegte politische Kampagnen in den Mainstream-Medien durchaus Effekte haben. Aber auch ihre Wirkung ist klein – unabhängig von der werbenden Partei, den Einstellungen der Wähler:innen und vielen weiteren Faktoren. Politische TV-Werbung steigert oder senkt die Beliebtheit eines Kandidaten um lediglich 1 % und verändert Wahlabsichten um ca. 0,7 %. In den USA hatte der regelmäßige Konsum des konservativen Senders Fox News vor allem Auswirkungen auf Republikaner:innen. So könnte Fox News die Wahlergebnisse der Partei um etwa 0,5 % verbessert haben.

Auch Facebook zeigt manche kleinen Effekte – obwohl es sich auf viele andere Bereiche nicht so negativ auswirkt wie befürchtet. Das Deaktivieren von Facebook für sechs Wochen verringerte in einer Studie mit 20.000 Usern vor den US-Wahlen 2020 das allgemeine News-Wissen und verbesserte das Erkennen von Fehlinformationen. Wenn man das mit dem Effekt eines Universitätsstudiums auf das Erkennen von Fehlinformation vergleicht (0,55 Standardabweichungen), war der Effekt der Facebook-Deaktivierung etwa zehnmal kleiner (0,04 Standardabweichungen). Geschätzte 1,16 % weniger potenzielle Trump-Wähler:innen stimmten statt Trump für Biden. Obwohl dieser Effekt statistisch nicht signifikant war – in einer sehr knappen Wahl kann er den Unterschied ausmachen. Das politische Wissen zur Wahl, die gefühlte Polarisierung, die Einstellungen und die politische Teilhabe offline änderten sich allerdings nicht. Und Instagram zeigte überhaupt keine vergleichbaren Effekte auf Fehlinformationen, Wahlentscheidungen oder News-Wissen.

Eine Facebook-Studie zu COVID-News fand ähnlich große Effekte auf Einstellungen: Geteilte Artikel aus Mainstream-Medien, die Gefahren suggerierten, ohne falsch zu sein ("Gesunder Arzt stirbt zwei Wochen nach Impfung"), reduzierten die Impfbereitschaft viel stärker (im Durchschnitt minus 2 % pro User) als Fake News (minus 0,05 %). Das lag auch daran, dass Fehlinformationen nur 0,03 % der geteilten Links zu COVID ausmachten.

Zusammenfassend lässt sich also sagen, dass Mainstream-Medien und Aussagen von Politiker:innen eine viel wichtigere Rolle spielen als Desinformation. Sie können insgesamt kleine, aber potenziell bedeutende Effekte auf Einstellungen haben. Fehlinformationen und polarisierende Inhalte erreichen hingegen nur wenige Leser:innen, vor allem jene, die bereits extreme Meinungen vertreten. Weil diese Menschen aktiv nach solchen Inhalten suchen und sie in Gruppen verbreiten, bedienen Algorithmen diese Nachfrage, sind aber nicht ihr Ursprung.

Dass diese Effekte klein sind, zeigt: Menschen stehen neuen Informationen grundsätzlich skeptisch gegenüber und sind nicht leichtgläubig. Überzeugungsarbeit ist schwer und braucht gute Argumente aus Quellen, denen die Menschen vertrauen. Deswegen haben wir oft das Gefühl, gegen eine Wand anzureden, wenn wir versuchen, die Einstellung anderer zu verändern.

KI und die Grenzen der Überzeugung

Könnte KI so überzeugend werden, dass sie die Meinungen der Massen manipuliert? Auch hier erwarten Fehlinformationsforscher mit guten Argumenten keine Katastrophe. In westlichen Demokratien wird der Nachrichtenkonsum vor allem über die Nachfrage gesteuert. Verschiedene Medien kämpfen um Aufmerksamkeit und publizieren deshalb das, was die Menschen lesen wollen. Nur eine Minderheit interessiert sich wirklich für Nachrichten und selbst diese liest oft nur die Headlines.

KI wird kaum dabei helfen, mehr von dieser Aufmerksamkeit zu generieren. Es gibt schon jetzt eine Menge Unsinn im Netz, aber niemand hat die Zeit, all das zu konsumieren. Viele Inhalte – inklusive Social Media – werden vor allem zur Unterhaltung konsumiert und nicht blind geglaubt.

Die bisher größte Untersuchung zur Überzeugungskraft von KI – mit fast 77.000 Teilnehmenden, 19 Sprachmodellen und 707 politischen Themen – zeigt: KI kann durchaus überzeugen, vor allem wenn sie ihre Argumente mit vielen überprüfbaren Fakten spickt – die allerdings umso häufiger falsch sind, je überzeugender das Modell auftritt.

Ob Menschen einer Information letztendlich vertrauen, hängt stark von der Quelle ab. Traditionelle Zeitungen und Fernsehsender genießen nach wie vor ein viel höheres Vertrauen als dubiose Internetseiten. Solange etablierte Medien sorgfältig recherchieren und sich nicht blind auf KI verlassen, lässt sich die breite Masse nicht leicht hinters Licht führen.

Nicht fehlinformiert, sondern uninformiert

Vielleicht wirft dieser Beitrag bei Ihnen mehr Fragen auf, als er beantwortet. Wenn Fehlinformation, Algorithmen und KI nicht das Problem sind, warum glauben dann so viele

Menschen so viel Unsinn, der nichts mit Wissenschaft oder der Realität zu tun hat? Warum gibt es so starke Tendenzen zu populistischen und rechtsextremen Parteien? Das sind gute und wichtige Fragen.

Viele Menschen sind tatsächlich sehr schlecht informiert und glauben an unseriöse Dinge. Nur sind Fehlinformationen nicht die Ursache dafür. Die richtige Frage lautet eigentlich: Warum glauben Menschen überhaupt an wahre Dinge über die Welt? Noch vor ein bis zwei Jahrhunderten hatte die Menschheit keine Ahnung, wie Impfungen, Antibiotika, Klimawandel, Atome oder Gene funktionieren. Diese Phänomene sind hochkomplex und nicht intuitiv zu erfassen. CO_2 und Bakterien kann man nicht sehen, und sich einen Krankheitserreger unter die Haut zu spritzen, um sich vor Krankheit zu schützen, ist alles andere als intuitiv. Unser menschlicher Grundzustand ist einer des Nichtwissens. Dank Bildung, Medien und Informationstechnologien haben bereits erstaunlich viele Menschen „richtige" Annahmen über solche Phänomene.

Umso tragischer ist es, dass das Interesse vieler Menschen an Medien, Politik und Wissenschaft sinkt. Die Interaktion mit Nachrichten, auch auf Social Media, nahm seit 2015 in 46 Ländern um 25 % ab. Die meisten Menschen sind *nicht fehlinformiert, sondern uninformiert*. Sie glauben nicht an viele Fehlinformationen, sondern vertrauen der Wissenschaft und den Medien zu wenig. Im mangelnden Vertrauen in gesellschaftliche Institutionen liegt das eigentliche Problem – und auch der Ansatz für eine nachhaltige Lösung.

Fehlinformation ist das Symptom, nicht die Krankheit

Was bei Diskussionen rund um Social Media, Algorithmen und KI leider übersehen wird, sind die eigentlichen Ursachen der Fehlinformation und die entscheidende Rolle von Eliten und Mainstream-Medien für ihre Verbreitung. Fehlinformationen sind vor allem ein Symptom gesellschaftlicher Polarisierung und mangelnden Vertrauens in Institutionen wie Politik, Medien, Wissenschaft und Medizin.

Dass Institutionen dabei versagen, reale Probleme zu lösen, und es nicht schaffen, große Bevölkerungsgruppen angemessen zu vertreten, führt zu sinkendem Vertrauen und steigender Polarisierung. So ist der Glaube an Verschwörungsmythen in Ländern mit mehr Korruption – wo Politiker:innen also weniger vertraut werden kann – deutlich weiter verbreitet. Politiker:innen selbst tragen auch beträchtlich zur Verbreitung von Fehlinformation bei, weil sie ein großes Publikum erreichen. Über verschiedene Länder hinweg korreliert die gefühlte Polarisierung der Gesellschaft am stärksten mit der Polarisierung der politischen Eliten und mit der Größe des nicht-weißen Bevölkerungsanteils. Dies könnte zum Beispiel ein Hinweis auf Integrationsprobleme sein.

Social Media spiegeln diese gesellschaftlichen Konflikte wider. Sie lassen sie schärfer erscheinen, als sie sind, indem sie bevorzugt extremen Stimmen eine Bühne bieten. Menschen, die sich nicht wirklich für ein Thema interessieren, melden sich nicht zu Wort. Deswegen bleibt die moderate Mehrheit auf Social Media weitgehend unsichtbar (die sogenannte „silent majority"). So entsteht eine scheinbare Polarisierung, die die tatsächliche Polarisierung verstärken kann.

Mögliche Lösungswege

Versuche, die Verbreitung von Fehlinformationen in der allgemeinen Bevölkerung einzudämmen, können nur kleine Effekte erzielen, weil das Hauptproblem von politischen Eliten und einer kleinen, politisch motivierten Minderheit ausgeht. Wirksam ist stattdessen alles, was das Vertrauen in demokratische Institutionen stärkt: Das verringert die Polarisierung und lässt damit auch die Fehlinformationen zurückgehen.

Leider gibt es dafür keine schnellen und einfachen Rezepte. Wir müssen Schritt für Schritt transparentere und effizientere Institutionen schaffen, die reale gesellschaftliche Probleme wirklich lösen. Wir brauchen Medien, die tatsächlich nützlich für ihre Leser:innen sind, statt nur vom Spiel um Aufmerksamkeit zu leben. Wir brauchen Interventionen, die das Vertrauen in die Medien stärken. Beispielsweise, indem Menschen lernen, wie Journalist:innen recherchieren und Fakten überprüfen. Wir brauchen eine verlässliche, transparente Wissenschaft, die sich selbst kritisch hinterfragt. Und wir brauchen vertrauenswürdige Politiker:innen, die offen kommunizieren, den Menschen die Wahrheit zutrauen und nicht nur die Interessen von Eliten bedienen.

Weil diese Lösungswege komplex und schwer umzusetzen sind, ist es verständlich, dass lieber mit dem Finger auf Social Media und auf Ingenieur:innen im Silicon Valley gezeigt wird. Leider lenken übertriebene Ängste vor Fehlinformation und Social Media die Aufmerksamkeit von den echten Ursachen ab. Wir müssen lernen, die Rolle von Social Media realistisch einzuordnen, um sie schlussendlich effektiv zu regulieren: indem wir uns auf Politiker:innen und extreme Akteur:innen konzentrieren, anstatt pauschale Warnungen vor Fake News zu verbreiten, die das Vertrauen in die Medien nur weiter senken. Eventuell könnten wir Social Media und ihre Algorithmen auch nutzen, um gezielt Inhalte zu fördern, die zwar kleine, aber dafür positive Effekte haben könnten.

Zu diesem Thema erschien 2024 in *Spektrum der Wissenschaft* ein Artikel der Autorin. Der vorliegende Text ist eine eigenständige, an die aktuelle Studienlage angepasste Fassung.

Hannah Metzler ist Psychologin und psychologische Wissenschaftlerin. Sie forschte an der Medizinischen Universität Wien und am Complexity Science Hub zu psychologischen Aspekten sozialer Medien, insbesondere zu Emotionen und Desinformation.