

## Doktor Google und der Kampf gegen Corona

Augenbrennen, Tinnitus, Hautausschlag, Schulterschmerzen... Wer sich nicht gut fühlt, geht heutzutage nicht direkt zum Arzt, sondern erst einmal ins Internet. Umfragen zufolge nutzen zwei Drittel der Deutschen das Netz, um sich über mögliche Krankheiten zu informieren. Nun wollen Forscherinnen und Forscher herausfinden, ob sich bei "Doktor Google" Hinweise auf die Verbreitung des Coronavirus finden lassen – oder sogar bislang unbekannte Symptome. Wissenschaftler in Harvard und London haben Google-Suchanfragen in acht Ländern untersucht, die mit bestätigten Covid-19-Symptomen einhergehen, etwa Fieber, Husten oder der Verlust des Geruchssinns. Das Ergebnis: überall stiegen in den vergangenen Wochen die Suchanfragen nach Covid-19-Symptomen. Diese Erkenntnis ist zunächst nicht überraschend, denn schließlich ist die Angst, infiziert zu sein, in der Bevölkerung insgesamt gestiegen. Und damit auch das Bedürfnis, Symptome zu googeln. Interessant aber ist, was die Modelle der Forschenden über die Verbreitung und Entwicklung der Pandemie aussagen könnten. In einem Beitrag in der *New York Times* nahm der Datenforscher Seth Stephens-Davidowitz die Studie zum Anlass, um die Suche nach Symptomen in einzelnen US-Bundesstaaten mit den dort bestätigten Covid-19-Fällen zu korrelieren. Er fand heraus, dass die Anfragen nach Riechverlust in den USA vor allem in den Bundesstaaten häufiger waren, in denen es auch die meisten Infizierten gab, nämlich New York, New Jersey und Louisiana.

Übertrage man diese Korrelation nun auf andere Länder, schreibt Stephens-Davidowitz, könne man theoretisch zwei Dinge feststellen: Erstens, ob die Anzahl der Erkrankten in einem Land oder einer Region höher ist als bislang bekannt. Als Beispiel erwähnt er Ecuador: Dort sei die Zahl der bestätigten Covid-19-Fälle zehnmal geringer als in Spanien, es werde aber zehn Mal so oft nach Riechverlust ("Ich kann nicht riechen") auf Google gesucht – ein möglicher Hinweis, dass in Ecuador viele Menschen erkrankt sind, aber noch nicht getestet wurden. "Die Suchdaten lassen vermuten, dass Ecuador weit stärker von Covid-19 betroffen sein könnte, als die offiziellen Zahlen sagen", schreibt Stephens-Davidowitz. Zweitens könnten die Suchanfragen Symptome aufzeigen, die bislang noch nicht mit der Lungenkrankheit in Verbindung stehen. In Italien etwa stiegen die Suchanfragen nach Riechverlust bereits an, bevor es als bestätigtes Symptom galt. Stephens-Davidowitz fand heraus, dass in den Corona-Hotspots der USA in den vergangenen Wochen vermehrt nach Augenschmerzen gesucht wurde. In Bundesstaaten, die weniger stark betroffen sind, war das nicht der Fall. Tatsächlich gibt es erste Berichte, wonach auch eine Bindehautentzündung ein Symptom für Covid-19 sein könnte. Um das mit Sicherheit zu sagen, ist die Datengrundlage aber noch zu dünn. Überhaupt warnen sowohl Seth Stephens-Davidowitz als auch die Forschenden aus Harvard und London davor, voreilige Schlüsse aus Google-Suchanfragen zu ziehen. Aus gutem Grund.

Schon einmal galt die Suchmaschine als vielversprechendes Tool, um die Verbreitung einer Krankheit zu analysieren – bis man schließlich herausfand, dass die Daten doch nicht immer ganz so genau waren. Die Rede ist von *Google Flu Trends*, die 2008 gestartet wurde. Mithilfe von Millionen Suchanfragen wollten Google-Ingenieure vorhersagen, wann und wo die nächste Grippewelle ausbrechen würde. Die Grundlage bildeten Suchbegriffe, die mit dem Auftreten einer Grippe korrelierten. Die Idee: Wenn viele Menschen in einer Region plötzlich vermehrt diese Begriffe googeln, könnte dort demnächst eine Grippewelle auftreten. Die Bevölkerung, Mediziner und Behörden könnten sich dann schneller und besser darauf einstellen. Tatsächlich funktionierte *Google Flu Trends* zunächst vergleichsweise gut; die Vorhersagen deckten sich in vielen Fällen mit den späteren Analysen der US-Seuchenschutzbehörde CDC (Center for Disease Control and Prevention). Schon bald aber lagen die Vorhersagen immer öfter daneben: Manche Ausbrüche wurden gar nicht erkannt, andere dagegen, etwa in der Grippesaison 2012-2013, überschätzte GFT maßlos. 2015 wurde das Projekt nach anhaltender Kritik beendet.

Wie Wissenschaftler herausfanden, hatten die Entwickler von *Google Flu Trends* vor allem zwei Dinge nicht beachtet: Zum einen den eigenen Empfehlungsalgorithmus der Suchmaschine. Wer etwa nach "Grippe-symptomen" suchte, bekam auch die Suche nach "Grippeimpfung" empfohlen – beide Anfragen wurden in die Analyse aufgenommen; das System verstärkte sich quasi selbst. Zum anderen gab es schlicht falsche Korrelationen zwischen der Grippe und anderen, saisonalen Suchbegriffen. "*Google Flu Trends* war teils Grippedetektor, teils Winterdetektor", schrieben Kritiker wie der Informatiker und Politikwissenschaftler David Lazer damals. Handelt es sich folglich bei den aktuellen Studien ebenfalls bloß um einen Blick in die Glaskugel? Nicht unbedingt. Inzwischen haben Forscher und Forscherinnen mehr Erfahrung mit Big-Data-Analysen und passen ihre Studien entsprechend an. In der eingangs erwähnten Untersuchung nutzen die Wissenschaftler nicht nur die absoluten Suchanfragen, sondern harmonisieren diese mit Medienberichten aus den jeweiligen Ländern am Tag der Suche. Denn je präsenter ein Thema in den Medien ist, desto stärker wird danach gesucht. Somit können aus Korrelationen falsche Kausalitäten entstehen. Das lässt sich an einem Beispiel aus Deutschland illustrieren: Am 16. März erschien in der *Frankfurter Allgemeinen Zeitung* ein Interview mit dem Virologen Hendrik Streeck, in dem dieser sagte, er habe neue Covid-19-Symptome entdeckt, nämlich den Verlust des Geruchs- und Geschmackssinns. Vergleicht man die Suchanfragen nach "Geruchsverlust" und "Geschmacksverlust" in Google Trends, sieht man einen deutlichen Anstieg ab dem 18. März – zwei Tage nach dem Interview, als weitere Medien die Meldung aufgriffen. In diesem Fall zeigen die Daten folglich nicht, dass plötzlich mehr Menschen in Deutschland an Covid-19 erkrankt sind und ihre Symptome googelten. Sondern bloß, dass es eine neue Entwicklung gab, die viele Menschen interessierte.

Es ist unbestritten, dass Big-Data-Analysen helfen können, die Ausbreitung des Coronavirus besser zu verstehen. Social-Media-Beiträge und Suchmaschinen können auf kommende Ausbrüche hinweisen, wie erste Studien aus China zeigen. Mediziner verwenden bereits große Datenmengen in Kombination mit künstlicher Intelligenz, um Arzneimittel und Impfstoffe zu finden. Und Start-ups wie BlueDot aus Kanada durchforsten täglich Tausende verschiedene Quellen, Nachrichtenseiten, Foren, und Blogs um herauszufinden, ob es irgendwo auffällige Entwicklungen gibt. So konnten sie bereits Ende Dezember einen möglichen Ausbruch in Wuhan entdecken. Auch die Kritiker von *Google Flu Trends* schrieben in ihrer Studie aus dem Jahr 2014, dass die Algorithmen des Projekts für sich genommen zwar nicht sehr genau waren. Kombiniere man sie aber "mit anderen Gesundheitsdaten in Echtzeit", seien die Modelle sogar besser als die der Seuchenschutzbehörde damals. Anders gesagt: Die Suche nach Symptomen im Heuhaufen namens Big Data ist nicht völlig nutzlos. Auf die richtige Technik – und die richtigen Fragen – kommt es an.