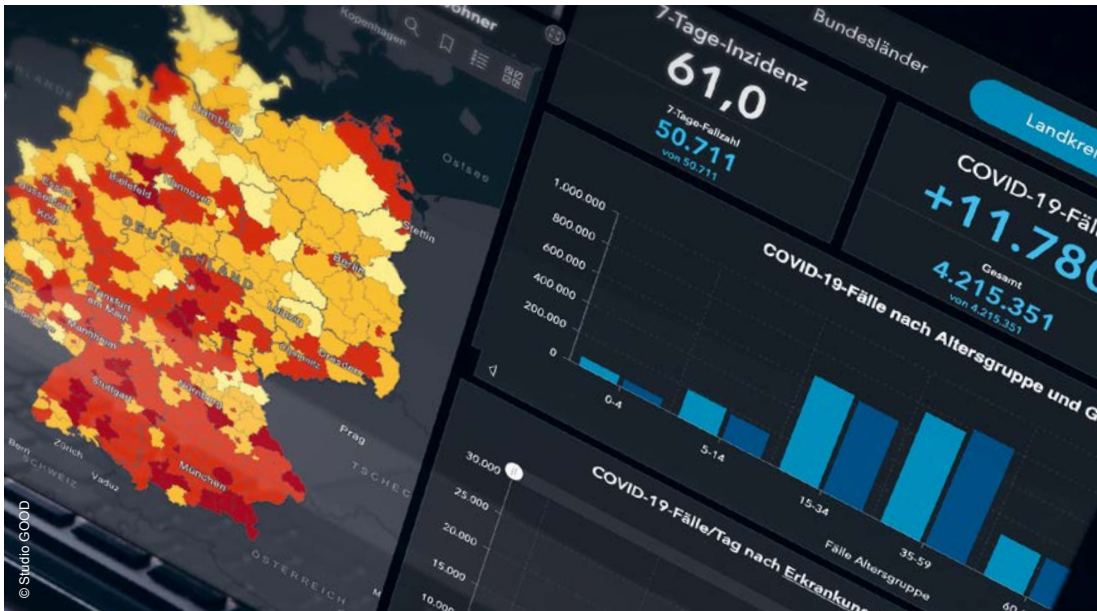


# „Corona zeigt, dass wir Daten besser verstehen müssen“

Die Statistikerin Katharina Schüller gehört zum Team der populärwissenschaftlichen Aktion „Unstatistik des Monats“, die den öffentlichen Umgang mit Zahlen kritisch hinterfragt. Im Interview erklärt sie, warum wir Daten oft unbewusst falsch einordnen.



**Frau Schüller, seit mehr als anderthalb Jahren werden wir mit Hochrechnungen zu Corona regelrecht „bombardiert“. Gerade zu Beginn erwiesen sich Prognosen über die Ausbreitung der Pandemie häufig als unzutreffend. Warum war das so?**

Die Bewertung eines Risikos als Grundlage für weitere Entscheidungen erfolgt stets auf Basis von Daten, die schon vorhanden sind. Doch genau diese Daten sind ausgerechnet zu Beginn einer Krise nicht besonders gut für eine Lagebeurteilung geeignet, da sie nicht dafür erhoben wurden und deshalb beispielsweise nicht repräsentativ sind. Wir können aus ihnen keine präzisen Handlungsempfehlungen ableiten. Sie können allenfalls als Leitplanken unseres Handelns dienen.

**War man sich zu früh zu sicher?**

Die Prognosen erweckten häufig den Anschein. Doch man muss zwei Quellen der Unsicherheit beachten. Die eine steckt in der Natur jeder Datenanalyse: Schätzungen sind stets zu einem gewissen Grad ungenau. Diese Unsicherheiten können als Schwankungsbreiten benannt werden, in sogenannten Konfidenzintervallen. Außerdem kann es immer Einflüsse auf das Geschehen

geben, die nicht absehbar sind wie zum Beispiel Virusmutationen, Wetterschwankungen oder Reaktionen auf die Prognose. Dies war bei der Corona-Pandemie der Fall.

**Welche Lehren ziehen Sie daraus?**

Corona zeigt uns, wie wichtig solide Informationen und eine robuste Daten-Infrastruktur sind. Es ist außerdem entscheidend, welche Daten vorliegen. „You can't manage what you can't measure“, lautet ein Spruch – „Was man nicht messen kann, kann man nicht steuern.“ Das bedeutet: Wo Informationen fehlen, gibt es blinde Flecken beim Management einer Krise sowie Probleme, die man nicht wahrnimmt. Zudem existiert eine kognitive Verzerrung.

**Was ist damit gemeint?**

Unvollständige Daten werden von uns überbewertet. Einfach deshalb, weil sie vorhanden sind.

**Können Sie ein Beispiel nennen?**

Wenn jeden Tag die aktuelle Zahl gemeldeter Corona-Fälle in den Schlagzeilen rauf und runter wandert, dann geht es irgendwann nur noch um diese eine Größe.

99

## Unvollständige Daten werden von uns überbewertet. Einfach, weil sie vorhanden sind.



**Katharina Schüller** regt einen bewussten Umgang mit Statistiken an. Sie ist im Vorstand der Deutschen Statistischen Gesellschaft sowie Leiterin der Münchner Beratungsfirma „Stat-up“.

Gleichzeitig ignorieren wir die Wissenslücken. Denn, was wäre, wenn in der „Tagesschau“ ebenfalls täglich berichtet würde, wie viele Menschen wegen der Pandemie ihren Job verloren haben? Oder wie viele Schulstunden ausgefallen sind? Oder wie häufig Depressionen festgestellt wurden? Dann hätten wir eine ganz andere Vorstellung von den Folgen der Pandemie.

### Es ist also falsch, sich nur auf die Infektionszahlen zu konzentrieren?

Ja, absolut. Nur der Zugang zu den unterschiedlichsten Datenquellen zeichnet ein ganzheitliches Bild von einer Krise. Diese müssen wir auswerten und in Relation zueinander setzen. Von der Corona-Krise sind doch Millionen Menschen mit sehr unterschiedlichen Lebensumständen betroffen. Etwa alleinerziehende Mütter mit Schulkindern, die es häufig schwer haben. Diese Vielfalt der Perspektiven müssen wir ernst nehmen. Sonst erkennen wir nicht, dass es Zielkonflikte geben kann und Maßnahmen unerwünschte Nebenwirkungen haben.

### Wie können wir diese Vielfalt der Perspektiven besser berücksichtigen?

Es geht um die Frage: Welche Aspekte der Realität will ich einbeziehen? Wir sollten eine Pandemie wie Corona als ein komplexes System begreifen: Welche Gesichtspunkte sind relevant, um dieses System zu steuern und um gut aus der Pandemie rauszukommen? Welche Ziele sind wichtig? Daten allein helfen nicht. Doch ihre kompetente Analyse ist ein wesentlicher Faktor, um Konsequenzen aus einer solchen Situation zu ziehen und zu steuern – und vieles von dem, was wir im vergangenen Jahr erlebt haben, besser beherrschbar zu machen.

### Ein Schwerpunkt Ihrer Arbeit ist „Data Literacy“ – der kompetente Umgang mit Daten. Kann sich durch Data Literacy unsere Fähigkeit, Risiken zu beurteilen, verbessern?

Für mich ist der Umgang mit Daten eine Voraussetzung für Risikokompetenz. Wir müssen lernen, Daten und Informationen kritisch zu hinterfragen. Wir müssen verstehen, was in den Daten steckt und was erst durch unsere Interpretation hinzugefügt wird. Die Bewertung erfolgt nie nur objektiv, sondern abhängig davon, welche Ziele man verfolgt. Bei Corona etwa stellt sich die Frage: Geht es bei den Maßnahmen nur darum, eine unmittelbare Bedrohung abzuwenden oder interessieren uns mittel- und langfristige Folgen und Fragen der Lebensqualität? Je nachdem, welche Ziele wir verfolgen, müssen wir Daten nach spezifischen Kriterien einordnen und sie entsprechend bewerten.

### Welche praktischen Konsequenzen sollte man daraus ziehen?

Die Entscheider, zum Beispiel Politiker, brauchen ein besseres Verständnis für Daten: Was sind die Stärken von Daten, wo liegen ihre Grenzen, wo ihre Möglichkeiten? Sie sollten wissen, wie man Daten kommuniziert – einschließlich deren Unwägbarkeit, die wir immer mitdenken müssen. Zudem brauchen wir hochwertige öffentliche Daten und Statistiken, denen eine belastbare Infrastruktur zugrunde liegt. Diese wird häufig vergessen, wenn wir über Kompetenz in der Krise sprechen. Der Aufbau einer Dateninfrastruktur klingt nicht so hip und sexy wie die Schlagworte Big Data, Künstliche Intelligenz oder Dashboards – und doch brauchen wir ein qualitätsgesichertes professionelles System für die Datenbereitstellung und -analyse, auf das Politik und Verwaltung zuverlässig zugreifen können.

### Wie moralisch ist das Sammeln von Informationen? Stichwort Datenethik.

Wenn es ums Sammeln und Verarbeiten von Informationen geht, wird die Moral-Frage sehr oft nur dahingehend gestellt, was man nicht darf. Als ob es bei der Datenethik einzig darum geht, Daten nicht zu missbrauchen. Zur Datenethik gehört ebenso, was man soll: Daten für einen guten Zweck nutzen und zum Wohl der Gesellschaft einsetzen. Denn genauso unethisch wie der Missbrauch ist es, Daten nicht zu nutzen, obwohl sie helfen könnten, Probleme wie die jetzige Pandemie besser und schneller zu lösen. ■