

Neue Zürcher Zeitung

«Das Bundesamt für Gesundheit muss die Daten gar nicht so Die bleiben schön bei ihm»

Georges-Simon Ulrich, Direktor des Bundesamts für Statistik (BFS), und der n...
Walliser Datenwissenschaftler Diego Kuonen sprechen über das Corona-Dat...
Durcheinander beim Bund und darüber, warum jetzt alles besser werden soll.

Barnaby Skinner

Kommentare →

28.05.2020, 19.30 Uhr



Viele Bundesämter haben Mühe damit, Daten über die Ämtergrenzen hinweg miteinander zu teilen.
Das neue Datenkompetenzzentrum des BfS will für das Problem eine Lösung bieten.

Martial Trezzini / Keystone

Herr Ulrich und Herr Kuonen, wie haben Sie sich während der Corona-Krise informiert?

Ulrich: Über alle möglichen Kanäle. Soziale Netzwerke waren für mich wichtig. Das Berufsnetzwerk LinkedIn etwa. Dort war die Diskussion eine andere als bei Facebook, viel sachlicher. Und natürlich habe ich mich auch über die klassischen Medien informiert: Fernsehen, Radio, Zeitung – logischerweise habe ich auch die NZZ gelesen. Ich bin ein News-Junkie.

Kuonen: Ich las vor allem meine Hauszeitung, den «Walliser Boten». Und ich bin auf dem sozialen Netzwerk Twitter aktiv. Dort habe ich viel Zeit verbracht.

Wurden Sie gut informiert?

Ulrich: Ich habe in den Medien eine Lernkurve beobachtet. Zuerst ging es...
78% Of desktop users scrolled to here...ten zu Corona-Infizierungen zu haben.
Hauptsache, man hatte irgendetwas. Auch das BfS spürte diese grosse

45

1:05
Recirculation
13%

☐ Desktop
☐ Mobile
☐ Tablet

View in:
 Real-Time
 Historical

Nachfrage. Wir zählten im Monat März 1,5 Millionen Seitenaufrufe unserer Website. Normalerweise sind es etwa 1 Million, also zwei davon. Dann setzte die zweite Phase ein. Plötzlich begannen sich Medien zu fragen: Wie repräsentativ sind diese Daten? Wie vergleicht man sie? Wer hat die Daten wie erhoben?

Kuonen: Ich habe davon nicht genug gesehen und hätte mehr Aufklärungsarbeit von den Medien erwartet. Ich sah viele sinnlose Datenvisualisierungen. Ich finde, die Zeitungen und das Fernsehen verbringen mehr Zeit damit verbringen müssen, den Leuten zu erklären, was die dargestellten Kurven eben nicht zu lesen war.

Ulrich: Im Grossen und Ganzen war die Pandemie doch ein Paradebeispiel dafür, wie der gesellschaftliche Meinungsbildungsprozess funktioniert, welche Rolle Daten dabei spielen. Die letzten Monate haben mir gezeigt, dass das Zusammenspiel der Medien, der Politik und der Verwaltung in unserer Demokratie gut funktioniert. Und die genannte Lernkurve sehe ich überall – nicht nur bei Medienkonsumenten und den Journalisten selber. Auch in der Verwaltung hat man viel gelernt.

Hat das BAG als Hüter der Daten zu Corona-Infizierten einen guten Job gemacht?

Ulrich: Es trägt die Verantwortung für die erhobenen Daten. Und im Gesundheitswesen sind Daten eben besonders heikel. Näher beim Menschen kann man ja kaum sein. Der Datenschutz spielt hier eine grosse Rolle. Das Amt muss deshalb sehr genau prüfen, welche Daten es freigibt. Und welche es zurückhält. Ich glaube, was die Publikation von Daten angeht, in welchem Format, zu welchem Zeitpunkt, da hat wie gesagt auch das BAG in den letzten Wochen viel gelernt. Und eben nicht nur das BAG, die ganze Verwaltung.



BfS-Direktor Georges-Simon Ulrich.
PD

Mitten in der Corona-Krise haben Sie kommuniziert: Ab 2021 wird das Bundesamt für Statistik ein Kompetenzzentrum für Datenwissenschaften schaffen. Warum?

Ulrich: Wir arbeiten schon länger daran. Das Ganze wurde schon im Jahr 2017 vom Gesamtbundesrat angestossen. Vielleicht wird jetzt in der Krise und dem Daten-Durcheinander, das teilweise geherrscht hat, die Dringlichkeit davon der Öffentlichkeit irgendwie klarer. Der Bundesrat hat seine eigenen Daten erhebt und analysiert,

78% Of desktop users scrolled to here

dann entstehen ganz viele Doppelspurigkeiten. Und das neue Kompetenzzentrum kann zur Lösung dieses Problems beitragen.

Tun Sie da dem Bundesrat nicht zu viel Ehre an? Er reagierte damals politischen Druck. Aus Sicht des Steuerzahlers ergibt es tatsächlich Sinn, wenn Daten vom Bund zwei- oder gar dreimal erhoben werden, weil die Ämter keine Daten miteinander teilen wollen.



Diego Kuonen, Professor für Datenwissenschaft an der Universität Genf und CEO Statoo Consulting, berät das BfS.
PD

Ulrich: Ja, am Anfang stand eine FDP-Fraktion: «Keine Doppelspurigkeiten in der Verwaltung». Für uns war dies eine Steilvorlage. Er forderte den Bundesrat im Jahr 2016 auf, sicherzustellen, dass Unternehmen nicht die gleichen Daten durch unterschiedlichen Verfahren an den Bund übermitteln müssen. Und die Idee des Kompetenzzentrums ist es nun, solche Lösungen zu unterstützen, indem innerhalb der Verwaltung Daten-Brücken gebaut werden.

Kuonen: Im Grunde geht es darum, einen Kulturwandel herbeizuführen. Die Bundesverwaltung soll offener mit ihren Datenschätzen umgehen und das Potenzial von innovativen datenwissenschaftlichen

Methoden ausschöpfen.

Der vielzitierte Kulturwandel also. Das hört sich immer so wunderbar an. Aber warum sollen Bundesämter, die seit Jahrzehnten Datenhoheiten geniessen, diese plötzlich aufgeben?

Kuonen: Als ich 2016 als Statistiker das Mandat beim BfS im Bereich Datenwissenschaft begann, traf ich auf die Haltung: Ah, da kommt jetzt der von der Privatwirtschaft, um mit seinen neuen Ideen aus dem Bund ein Startup zu machen, um agiler zu werden. All diese nichtssagenden Schlagwörter eben. Aber ich habe ein inneres Feuer für das Statistik-System Schweiz. Und dieses Feuer versuche ich allen in der Verwaltung zu vermitteln.

Aber was heisst das? Reden wir wieder über den Fall des Bundesamts für Gesundheit. Seit Wochen weigert sich das BAG, Daten zu Corona-Infektionen mit der Öffentlichkeit oder mit der Forschung zu teilen. Dabei könnte man anhand dieser Daten verstehen, wo aktuell die grösste Gefahr besteht, sich mit dem Virus anzustecken.

Ulrich: Der Bundesrat hat letztes Jahr beschlossen, dass nicht sofort per Default mit der Öffentlichkeit geteilt werden müssen. Wir sind die operative Geschäftsstelle dieser Strategie.

78% Of desktop users scrolled to here

45

1:05

Recirculation

13%

Desktop

Mobile

Tablet

View in:

Real-Time

Historical

Und wir müssen dafür sorgen, dass sie umgesetzt wird. Wenn das aufgrund rechtlicher Abwägungen zu dem Schluss kommt, dass die Infizierungen etwa auf Gemeindeebene als besonders schützenswert eingestuft werden, weil damit in kleinen Gemeinden Personen identifiziert werden können, dann gilt das.

Aber grundsätzlich kann dieses Argument doch für alle Daten gelten?

Ulrich: Unser neues Kompetenzzentrum für Datenwissenschaften nutzt nun Möglichkeiten, einen neuen Weg einzuschlagen. Wir werden Daten nutzen und datenwissenschaftlich auswerten können, ohne den Datenschutz zu verletzen und ohne dass das BAG die Rohdaten abzurufen muss. Im Grunde schaffen wir hier eine geschützte Umgebung, um datenwissenschaftliche Methoden testen und dann die Testergebnisse publizieren zu können. Die Rohdaten bleiben bei den Fachverantwortlichen – und trotzdem kann das Potenzial der Daten ausgeschöpft werden. Zudem können die Fachverantwortlichen die Resultate ja auch selber nutzen und weiterbearbeiten.

Das ist interessant. Könnten beispielsweise auch Medien und damit die Öffentlichkeit Teil einer solchen geschützten Datenplattform werden?

Ulrich: Warum nicht? Vielleicht nicht mit Zugriff auf alle Rohdaten, aber auf einen Teil der Daten und auf die datenwissenschaftlichen Methoden, damit für die Öffentlichkeit nachvollziehbar bleibt, wie eine gewisse datenbasierte Aussage zustande kommt. Um beim BAG-Beispiel zu bleiben: Es könnte der Eindruck entstehen, dass in der Gemeinde X das Risiko einer Covid-19-Ansteckung grösser ist als in der Gemeinde Y. Dann könnte dieser Verdacht in einem nächsten, unabhängig geführten Schritt überprüft werden, bevor Schlussfolgerungen gezogen und allfällige Massnahmen ergriffen werden.

Mit anderen Worten: Das neue Kompetenzzentrum für Datenwissenschaften des BfS funktioniert als Schnittstelle nicht nur für den Datenaustausch innerhalb der Verwaltung, sondern auch zur Öffentlichkeit oder zur Privatwirtschaft?

Kuonen: Ich möchte hier die Daten-Diskussion abbrechen. Es geht um den Einsatz innovativer datenwissenschaftlicher Methoden. Daten sind das Mittel zum Zweck. Um nochmals Ihr Beispiel zu nennen: Das BAG muss die Daten gar nicht schicken, die bleiben schön bei ihm. Es geht darum, das Amt mithilfe datenwissenschaftlicher Ansätze darin zu befähigen, die entwickelten Algorithmen auf seine Daten in seinem Haus anzuwenden.

Auch Sie brauchen also keinen direkten Zugriff auf die Rohdaten?

Kuonen: Genau, wir selber brauchen keinen Dateneinblick, vorausgesetzt, das einzige, was zurückkommt, ist das Resultat der Rechnung des Algorithmus. Das ist die Zukunft der Datenwissenschaft. Es

78% Of desktop users scrolled to here

geht nicht darum, mit den Daten herumzuspielen – hier etwas C bauen, da etwas Cooles. Im Fokus ist immer die ganz konkrete ge Leistung vom Zentrum. Das könnte zurzeit also sein: die Häufung Corona-Infizierungen räumlich darstellen, ohne den Datenschutz verletzen.

Wie viele Leute werden künftig in diesem neuen Kompetenzzentrum arbeiten?

Ulrich: Wir beginnen mit zehn Vollzeitstellen.

Finden Sie diese Leute mit dem nötigen Know-how in der Schweiz? müssen Sie international aktiv werden?

Kuonen: Datenwissenschaft ist Team-Sport. In der Privatwirtschaft sehe ich eine grosse Nachfrage nach «Data-Scientists». Das ist «Guguus». Das Kernteam muss aus Leuten bestehen, die etwas von maschinellem Lernen verstehen. Also davon, wie man mit Daten einen Algorithmus trainiert. Dann braucht es Leute, die Datenbanken pflegen können und wissen, was Daten-Management bedeutet; dann braucht es ein paar Leute mit sehr guten Statistik-Kenntnissen oder Leute, die wissen, wie man Daten verknüpft. Das sind ganz unterschiedliche Skills, die da gefragt sind. Vor allem müssen die Leute kreativ, kommunikationsfähig und anpassungsfähig sein und die Probleme erkennen und lösen wollen.

Haben Sie eine Vision für das Kompetenzzentrum?

Kuonen: Ich bin nur der Berater des Bundes. Langfristig stelle ich mir vor, dass die gesamte öffentliche Hand eine interne Suchmaschine hat, in der die Datensätze aller Ämter beschrieben sind; genauso die Methoden, mit denen die Daten analysiert wurden und werden; und eine nachvollziehbare Dokumentation der einzelnen Schritte im datenwissenschaftlichen Prozess. Davon träume ich.

Das tut man auch als Journalist oder als interessierter Bürger.

Ulrich: Es geht uns grundsätzlich darum, ein System von Wissen aufzubauen. Covid-19 zeigt dessen Dringlichkeit. In der Schweiz ist der Datenschutz sehr wichtig. Unsere Idee ermöglicht es uns als Land, alle verfügbaren Daten zur Bekämpfung des Coronavirus zu nutzen, ohne dabei den Datenschutz auszuhebeln. So etwas haben wir heute nicht.

Haben Sie denn konkrete Anwendungen, die Ihnen bereits vorschweben?

Ulrich: Das Zentrum startet offiziell erst im Januar 2021. Aber unter der Internetadresse www.experimental.bfs.admin.ch finden Sie erste kleinere Vorprojekte: etwa ein Mobilitäts-Monitoring von 2500 Personen. Es zeigt, wie das Bewegungsverhalten der

78% Of desktop users scrolled to here

Schweizerinnen und Schweizer verändert hat. Oder eine Schätzung, wie sich die Erwerbsbeteiligung von Personen auf Gemeindeebene im Jahr 2020 verändert.

Viel Arbeit des Kompetenzzentrums dient wohl internen Zwecken. Künftig in der Verwaltung Algorithmen zum Einsatz kommen: Werden diese immer transparent machen, wie diese Algorithmen funktionieren?

Kuonen: Wir bauen hier ein Dienstleistungszentrum. Die genutzten datenwissenschaftlichen Methoden werden grundsätzlich – unter Wahrung des Datenschutzes – immer dokumentiert und öffentlich gemacht.

Unter Wahrung des Datenschutzes – das heisst also: Nicht alle Algorithmen, die das Zentrum künftig entwickelt, werden transparent sein?

Kuonen: Die ersten Projekte, die wir seit 2018 durchgeführt haben, hatten dieses Mantra: eine nachvollziehbare Dokumentation, und der gesamte Code musste reproduzierbar sein. Künftig werden wir die erforderliche Transparenz jedes einzelnen Schrittes im datenwissenschaftlichen Prozess in jedem einzelnen Fall prüfen müssen. Neben der Transparenz werden aber auch den Grundwerten Datenschutz, Datensicherheit, Daten-Governance und Reproduzierbarkeit grösste Bedeutung beigemessen.

So definiert der Bund Daten und deren Quellen

Primärdaten: geschaffene (erhobene) Daten, welche vom BfS zu statistischen Zwecken generiert werden, um die Gültigkeit von spezifischen bestehenden Annahmen anhand der Umsetzung theoretischer Konzepte zu erklären und zu überprüfen.

Sekundärdaten: beobachtete (gefundene) Daten, welche aus unterschiedlichen Gründen erhoben werden (oft nicht unter der Aufsicht des BfS; und die oft Anforderungen der öffentlichen Statistik nicht erfüllen) und zum Entwickeln neuer Annahmen bzw. Theorien genutzt werden können.

Identifizierbare Sekundärdaten: Sekundärdaten, welche sinnvoll einer einzelnen Einheit in einer interessierenden Population an einem bestimmten Standort und zu einem gegebenen Zeitpunkt zugeordnet werden können, etwa einer Einzelperson, einer Institution, einem Produkt oder einem Ort.

Nicht identifizierbare Sekundärdaten: Sekundärdaten, welche – im Gegensatz zu den identifizierbaren Sekundärdaten – nicht genau zugeordnet werden können und somit zum Zweck von statistischen Rückschlüssen nur begrenzt genutzt werden können.

78% Of desktop users scrolled to here

45

1:05

Recirculation 13%

Desktop

Mobile

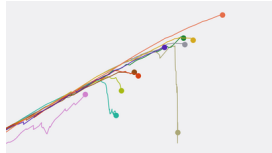
Tablet

View in:

Real-Time

Historical

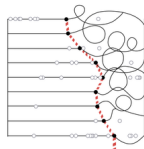
Mehr zum Thema



Die perfekte Corona-Grafik gibt es nicht

In der Corona-Krise sind Datenvisualisierungen so wichtig wie selten zuvor. Die richtige Darstellungsform für die Entwicklung der Pandemie zu finden, ist jedoch gar nicht so einfach.

Nikolai Thelitz, Kaspar Manz 08.04.2020



Von Lockdown zur Lockerung – wie das Coronavirus die Welt auf den Kopf gestellt hat im Zeitverlauf

In nur vier Monaten eroberte das neuartige Coronavirus fast den gesamten Globus und zwang Länder weltweit zur drastischen Einschränkung der Wirtschaft und des öffentlichen Lebens. Ein Rückblick auf die sich überstürzenden Ereignisse.

Jonas Oesch, Manuel Roth, Ulrich von Schwerin, Philip Küng, Nikolai Thelitz, Barnaby Skinner 10.05.2020

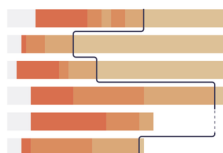
Weitere datengetriebene Artikel



KOMMENTAR

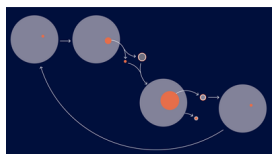
Der Bund muss raus aus den Datensilos

Barnaby Skinner 15.05.2020



Je früher die Massnahmen kamen, desto besser – eine Analyse von sechs stark vom Coronavirus betroffenen Regionen weltweit

Katrin Büchenbacher, Christian Kleeb, Alexandra Kohler, Jonas Oesch 09.04.2020



Tests sind entscheidend, um die Corona-Pandemie zu bekämpfen. Aber nicht immer ist vermehrtes Testen besser

Alan Niederer (Text), Joana Kelén (Infografik) 14.05.2020

Copyright © Neue Zürcher Zeitung AG. Alle Rechte vorbehalten. Eine Weiterverarbeitung, Wiederveröffentlichung oder dauerhafte Speicherung zu gewerblichen oder anderen Zwecken ohne vorherige ausdrückliche Erlaubnis von Neue Zürcher Zeitung ist nicht gestattet.

78% Of desktop users scrolled to here

45

1:05
Recirculation
13%

☐ Desktop
☐ Mobile
☐ Tablet

View in:
 Real-Time
 Historical

45

▲

1:05

Recirculation

13%

Desktop

Mobile

Tablet

View in:

Real-Time

Historical



78% Of desktop users scrolled to here