

Keine Angst vor Zahlen

Datenjournalismus ist seit einigen Jahren schwer in Mode. Kleinere Redaktionen aber scheitern oft an dem sehr speziellen Know-how, das man dafür braucht. Dabei ist der Einstieg gar nicht so schwer, wenn man die richtigen Tools kennt.

Vieles rund um den Datenjournalismus ist in den vergangenen Jahren hochgejubelt worden. Hochkomplexe, schwer lesbare Online-Infografiken werden gefeiert, wie etwa der Jahresrückblick der „NZZ“, eine Visualisierung aller Artikel aus dem Jahr 2014, die online und in der Zeitung erschienen sind. Hübsch anzusehen, aber ohne erkennbaren Nutzen. Gleichzeitig werden simpelste Rechenspiele als das nächste grosse Ding ausgerufen, etwa während der Fussball-WM der Blog FiveThirtyEight des US-Statistikers Nate Silver, herausgegeben vom US-Sportnetzwerk ESPN. Die Wahr-

scheinlichkeit übrigens, dass Deutschland im Halbfinale gegen Brasilien sieben oder mehr Tore erzielen würde, betrug laut Zahlenanalytiker Nate Silver nur 1 zu 4.500. Das Ergebnis ist bekannt.

Solche Übungen sind nett. Doch sie lenken davon ab, was die Arbeit mit Daten für den Beruf des Journalisten tatsächlich bedeutet. Es geht dabei nicht um irgendwelche glorifizierte Infografiken oder Online-Spiele, sondern um den Einsatz von Werkzeugen zur schnelleren Informationsbeschaffung; darum, neue Zusammenhänge zu erkennen – und diese einfacher zu vermitteln. Im

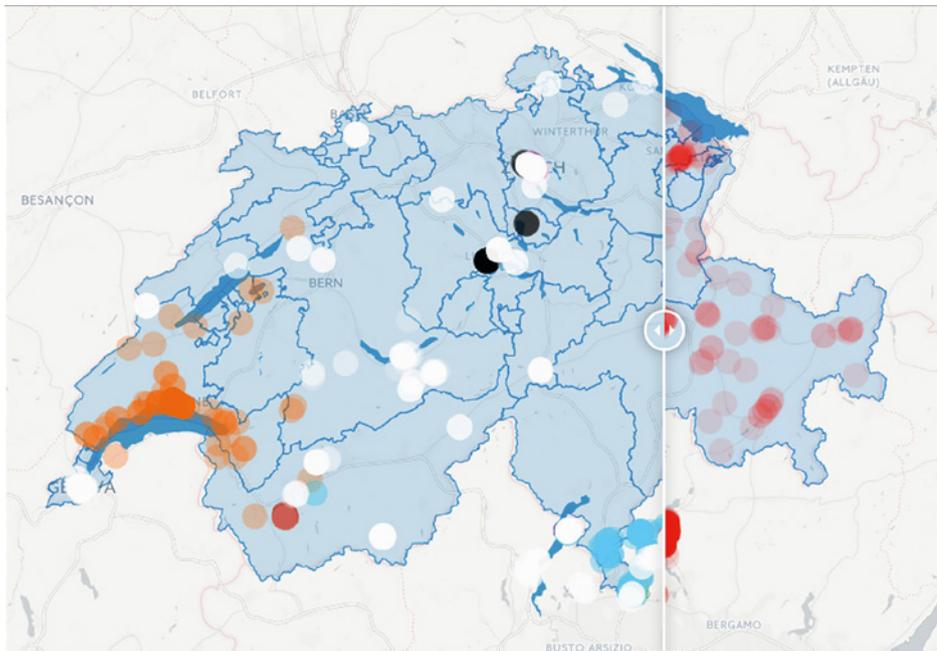
Grunde also darum, bessere Geschichten zu erzählen. Nicht mehr und nicht weniger.

In der digitalen Welt ist nämlich nicht nur der Strauss an Medienformen bunter geworden, sondern eben auch die Mittel, um Journalismus zu machen. Grundsätzlich hat sich am Beruf nichts geändert: Am Anfang jeder journalistischen Geschichte steht immer eine Frage. Wie zum Beispiel kürzlich diejenige des Wirtschaftsressorts des „Tages-Anzeigers“. Es wollte wissen, ob sich die Preise auf der Übernachtungsplattform Airbnb in der Schweiz ähnlich steil nach oben entwickelten, wie dies in New York der Fall ist.

Das Wirtschaftsressort trug diese Frage an mich. Ich hätte mich nun an einen Branchenexperten wenden können – einen solchen hätten wir mit Bestimmtheit gefunden. Stattdessen veranstalteten wir eine eigene kleine Untersuchung. Etwas, das mir vor 20 Jahren nicht möglich gewesen wäre. Natürlich hat es Airbnb damals nicht gegeben. Aber nehmen wir einmal an, es hätte etwas Ähnliches auf Papier existiert. Vielleicht ein Prospekt mit Übernachtungsangeboten in Privatwohnungen, jährlich flächendeckend in allen Briefkästen der Schweiz verteilt.

Für einen Preisvergleich hätte ich jedes dieser Prospekte zur Hand nehmen müssen. Ich hätte die Angebote auszählen, die Preise der Angebote aufschreiben, vergleichbare Anbieter herausuchen müssen. In der ungeduldigen Branche des Journalismus eine besonders undankbare Aufgabe. Heute kann ich dafür einen Computer und ein sogenanntes Scraping-Programm einsetzen, eine Software, die mir diese mühsame Auszählarbeit abnimmt.

Das in diesem Fall eingesetzte Programm namens Outwit Hub (siehe Kasten) brauchte eine Weile, bis es verstand, was es machen sollte. Denn Computerprogramme



Die Airbnb-Könige in der Schweiz: „Tages-Anzeiger“-Geschichte vom 2. März 2015 von Angela Barandun, Barnaby Skinner und Marc Fehr. Im Bild eine Slidergrafik, die zeigt, wie viele Angebote es in der Schweiz auf dem Übernachtungsangebot Airbnb gibt und wie viele Nutzer mehr als fünf Angebote machen. Barnaby Skinner brauchte etwa eineinhalb Tage, um sich in die Software Outwit Hub einzuarbeiten. Die eigentliche Analyse dauerte nur rund drei bis vier Stunden. Softwaregestützt untersuchte er alle Angebote des Übernachtungsportals und fand zum Beispiel heraus, dass der durchschnittliche Preis für eine Nacht stattliche 135 Franken betrug. Das teuerste Angebot im Wallis sollte gar 2.741 Franken pro Nacht kosten.

sind, ganz entgegen der Behauptungen der IT-Branche, alles andere als smart. Sie benötigen ganz präzise Befehle, damit sie korrekt arbeiten. Von Outwit verlangten wir, täglich alle Airbnb-Angebote von Schweizern auszuführen und die Preise in einer Excel-Tabelle abzuspeichern, damit wir sie vergleichen konnten. Eine unglaublich langweilige Arbeit also.

Das Ergebnis hingegen war äusserst interessant. Tatsächlich stellten wir bei den Preisen einen Aufwärtstrend fest. Bei der Datensammlung hatte Outwit aber neben den Kosten weitere Angaben gesammelt. Zum Beispiel die geografischen Koordinaten der Angebote. Somit waren wir in der Lage, eine interaktive Karte der Schweiz mit allen Airbnb-Angeboten zu erstellen – gewissermassen als Abfallprodukt der Recherche.

Am spannendsten waren aber die Informationen zu den Anbietern der Wohnungen und Zimmer auf Airbnb. Mit ein paar wenigen simplen Filtereinstellungen in der Excel-Tabelle stellten wir fest, dass sich längst nicht mehr nur Privatanbieter auf Airbnb tummelten. Das tun auch immer mehr professionelle Anbieter. Wir entdeckten gar einige Anbieter mit Hunderten Zimmern und Wohnungen. Das ist problematisch, weil sie damit zum Beispiel die Wohnungspreise in den Stadtzentren in die Höhe treiben oder die Sicherheitsbestimmungen, die Hotels zum Beispiel anbieten müssen, umgehen.

Das Resultat der Recherche war neben der interaktiven Online-Karte auch ein Porträt der Frau, die am meisten Angebote bei Airbnb betreut (am 2. März 2015 im „Tages-Anzeiger“ erschienen, siehe auch Grafik). Knapp 200 Angebote hat sie insgesamt. Airbnb ist für sie längst keine Sharing-Plattform mehr, wie der Webdienst immer angepriesen wird, sondern ein knallhartes Geschäft.

Solche Software ist doch viel zu schwierig zu bedienen, wird sich der eine oder andere Kollege nun denken. Tatsächlich braucht es dazu gewisse IT-Grundkenntnisse. Es gilt zu verstehen, wie eine Website funktioniert. Wo man den Quellcode einer Site abrufen kann. Und wie diese Informationen dazu genutzt werden können, Software wie Outwit für die eigene Datensammlung einzusetzen.

Doch all das kann man sich heute vollkommen umsonst mit ein wenig Geduld

im Internet aneignen. Es finden sich im Web mittlerweile Dutzende Kurse, in denen man sich selbst das Programmieren beibringen kann, bei www.jsforcats.com oder www.codeschool.com zum Beispiel. Oft reicht es dabei, nur die ersten paar Übungen dieser Online-Kurse durchzuexerzieren. Selbst programmieren braucht der Journalist von heute nicht. Ihm oder ihr ist schon geholfen, wenn er versteht, wie Code aufgebaut ist und wie er funktioniert.

Wie nützlich IT-Kenntnisse im journalistischen Alltag sind, zeigte kürzlich eine Geschichte in der „Basler Zeitung“ über den „gelungenen Wahlkampf des Gewerbeverbandes gegen die neue RTVG-Revision“. Die Kampagne sei ganz allein auf dem Mist des Gewerbepräsidenten Erich Bigler gewachsen, schrieb der Journalist.

Hätte der Schreiber etwas von IT verstanden, hätte er mit einem Mausklick bei der Online-Einblendung der Anti-Kampagne den Code prüfen können. Dort steht, dass hinter der Kampagne das Werbebüro Goal steckt. Landesweit ist diese Agentur etwa für die schwarzen Schafe der Ausschaf-

fungsinitiative bekannt oder für die Morscheen, die aussahen wie Raketen, auf den Plakaten der Minarettverbotsinitiative.

Solche peinliche Fehler sind mit ein wenig IT-Kenntnissen und Daten-Know-how zu vermeiden. Die Digitalisierung und all die nützlichen Onlinewerkzeuge bedeuten nicht das Ende des Journalismus, wie viele immer wieder herausposaunen. Im Gegenteil. Richtig eingesetzt, machen sie den Journalismus nur besser.

BARNABY SKINNER

ist Datenjournalist der „Sonntagszeitung“ und beim „Tages-Anzeiger“ in Zürich.

barnaby.skinner@sonntagszeitung.ch



INFO

Fünf Einsteigertools

Google Docs (Alternativen: Open Office, Excel, Numbers)

Das Pendant zu Excel. Mit dem Online-Werkzeug von Google lässt sich allerdings einfacher arbeiten. Die Formeln sind logischer aufgebaut. Und vor allem lassen sich bearbeitete Tabellen einfach online teilen. Gratis-Software.

Open Refine

Ab 10.000 Zeilen werden Tabellen unübersichtlich. Mit Open Refine lassen sich grosse Datenmengen analysieren und reinigen. Zum Beispiel dann, wenn die Städtenamen einer Datenbank unterschiedlich geschrieben worden sind. Das Programm kann locker mit Abermillionen Datensätzen umgehen. Gratis-Software.

Outwit (Alternativen: Wiki Scraper, Web Scraper)

Um dieses Programm anzuwenden, muss der Anwender in die Tasche greifen. Aber

nicht tief. Die komplette Software kostet umgerechnet 90 Franken. Sie kann auch für umgerechnet 60 Franken jährlich lizenziert werden. Outwit erlaubt auf Knopfdruck Millionen Websites in wenigen Stunden abzugrasen.

PDF Extractor (Abbyy)

Die spannendsten Zahlen werden oft in PDFs vergraben, die nicht von Computerprogrammen gelesen werden können. Mit dem PDF Extractor können sie – nomen est omen – extrahiert und dann mit Spreadsheets wie Excel oder Google Docs verwendet werden. 25 Franken jährlich.

Tableau

Die ersten Schritte mit Tableau sind schwierig. Und die Vollversion kostet auch einiges: Umgerechnet 1.000 Franken im Jahr. Doch wer erst einmal damit zu arbeiten begonnen hat, kann nicht mehr aufhören. Selbst riesige Datensätze lassen sich innert Minuten mit dem Programm einlesen und visualisieren.