

DATEN RÜCKEN IN DEN FOKUS

WIE FACHJOURNALISTEN DAVON PROFITIEREN KÖNNEN

von Ulrike Langer

Fachjournalisten müssen nicht unbedingt programmieren können. Aber sie sollten wissen, wie interaktive Datenaufbereitung den Sportjournalismus, den Wirtschaftsjournalismus oder andere Fachjournalismen bereichern kann. Noch wird Datenjournalismus nicht systematisch gelehrt. Das verschafft Pionieren auf diesem Gebiet Wettbewerbsvorteile.

Ein Journalist, der ohne Computer- und Internetkenntnisse arbeitet? Unvorstellbar im Jahr 2011. Genauso vorgestrig werden in einigen Jahren Journalisten wirken, die sich dem Umgang mit Excel-Tabellen verweigern, glaubt SIMON ROGERS, Datenchef des GUARDIAN. Die britische Tageszeitung gehört zusammen mit der NEW YORK TIMES zur Speerspitze des Datenjournalismus. Dazu gehört mehr, als bloß Infografiken zu drucken oder ins Netz zu stellen: Datenjournalisten werten Datensätze systematisch aus, visualisieren sie in interaktiven Grafiken¹, rücken Daten in den Mittelpunkt einer Geschichte, machen Originärquellen transparent und stellen Datensätze der Allgemeinheit zur Verfügung.

Der GUARDIAN machte schon vor zwei Jahren von sich reden, als er fast eine halbe Million Dokumente zum Spesenskandal der Unterhausabgeordneten über eine offene Schnittstelle ins Netz stellte und seine Nutzer bat: „Können Sie uns mit der Überprüfung dieser Daten helfen?“. Die Resonanz war überwältigend und machte die weitere umfassende Berichterstattung der Zeitung zum Thema erst möglich.² Seitdem haben die Londoner einen Datablog und einen Datastore eingerichtet. Letzteres ist eine Serviceabteilung, die sich um das Aggregieren, Gewichten und Auswerten von Daten für redaktionelle Beiträge kümmert. Mehrfach wurde der GUARDIAN für seine Aufbereitung der WIKILEAKS-Dokumente gelobt, die nicht nur in herkömmliche und multimedial aufbereitete Berichte mündete, sondern auch in eine von Nutzern recherchierbare Datenbank.

teilweise stehen sie kostenlos im Netz zur Verfügung. Das Informationsfreiheitsgesetz, die „Open Data“-Bewegung, die sich für frei zugängliche, öffentliche Daten einsetzt, und Whistleblower-Plattformen wie WIKILEAKS haben riesige Datenberge erzeugt. Gleichzeitig hat die Berichterstattung über WIKILEAKS weitere Aufmerksamkeit auf das Thema gelenkt. Auch SASCHA VENOHR, Entwicklungsredakteur bei ZEIT ONLINE sieht im Datenjournalismus einen wichtigen Trend: „Gerade im angelsächsischen Raum haben Gesetzesinitiativen, die vom Gedanken des Open Government geprägt sind, neue spannende Quellen mit maschinenlesbaren Daten geöffnet.“³

Für VENOHR und ZEIT ONLINE ist Datenjournalismus „ein wichtiger Baustein in den Darstellungsformen, die uns zur Verfügung stehen – keine Modeerscheinung.“ Im September 2010 veröffentlichte ZEIT ONLINE ein Dossier über die Todesopfer rechter Gewalt.⁴ Die Beiträge sind in erster Linie das Ergebnis einer aufwendigen investigativen Recherche – den 47 offiziell dokumentierten Fällen von Mord oder Totschlag von 1990 bis 2010 stehen frappierende 137 von ZEIT ONLINE belegte Fälle gegenüber. Eine interaktive Karte mit einer Zeitachse, in der sich die markierten Orte der Übergriffe sukzessive anhäufen, zeigt, dass rechtsradikale Gewalt weder zeitlich noch räumlich ein punktuell, sondern ein manifestiertes gesamtdeutsches Phänomen ist. (→ **Abb. 1**) Alle recherchierten Detailinformationen zu den 137

TOOLS

TOOLS WERDEN PREISWERT, DATENBERGE WACHSEN

Mehrere Faktoren haben den Datenjournalismus jüngst aus seinem „nerdigen“ Nischendasein geholt. Die Tools, um Daten aufzubereiten, sind preiswert geworden,

Data-Scraping: Liste mit Werkzeugen zum Auslesen ("Scrapen") von Daten

→ <http://www.opendataday.org/wiki/Tools>

International Open Data Hackathon: Linkliste zu frei zugänglichen Tools

→ <http://www.opendataday.org/wiki/Tools>

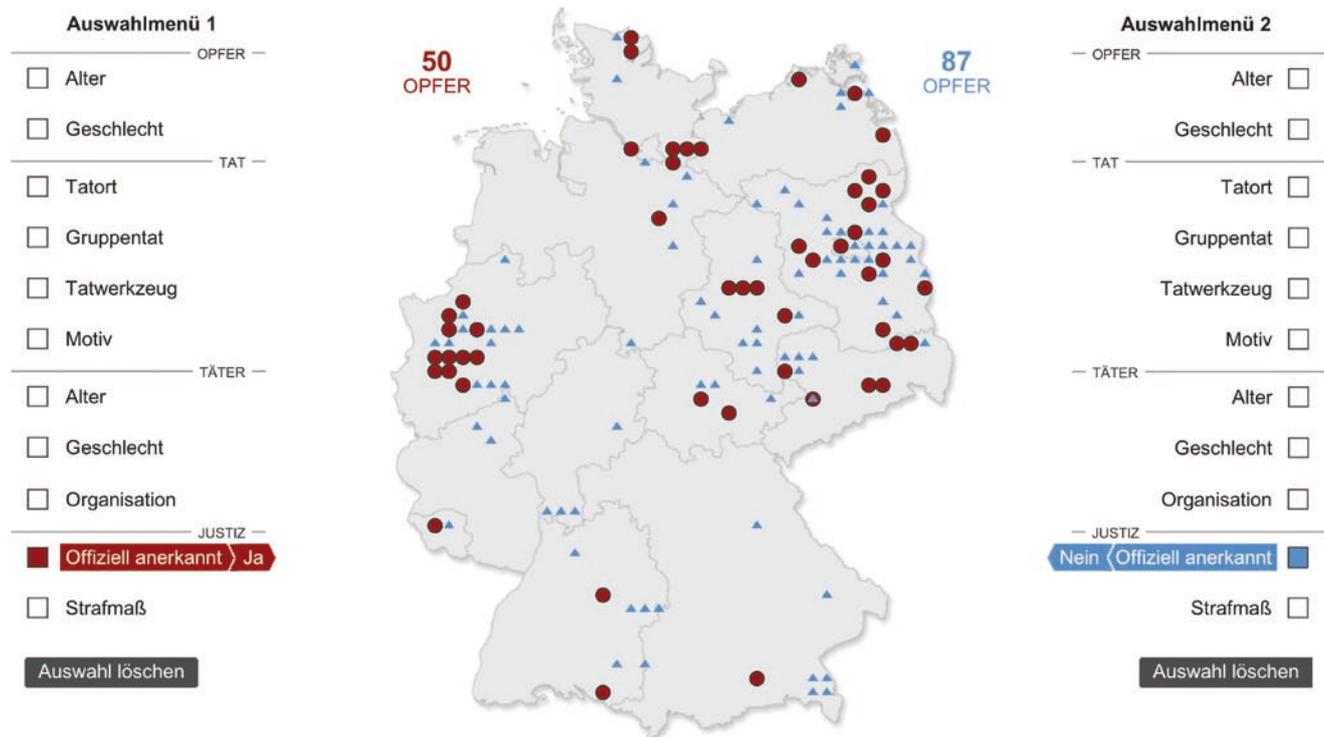
Periodensystem der Datenvisualisierungs-Methoden:

→ http://www.visual-literacy.org/periodic_table/periodic_table.html

Zehn hervorragende freie Tools zum Erstellen von Infografiken:

→ <http://www.makeuseof.com/tag/awesome-free-tools-infographics/>

ABB. 1: „TODESOPFER RECHTER GEWALT 1990 – 2010“



Quelle: ZEIT ONLINE (<http://www.zeit.de/gesellschaft/zeitgeschehen/todesopfer-rechter-gewalt>)

Fällen stehen als frei zugängliche Onlinetabelle zur Verfügung – auch als Basis für künftige Beiträge anderer Medien. Bei dieser aufwendigen Recherche eignete sich das Team Fähigkeiten im Umgang mit Datensätzen an, die künftig immer wieder, wenn es sinnvoll erscheint, eingesetzt werden sollen. Interessante Entwicklungen beim Trendthema Datenjournalismus begleiten sowohl ZEIT ONLINE als auch die Berliner TAGESZEITUNG in eigenen Blogs zum Thema.⁵

NOCH KEINE SYSTEMATISCHE AUSBILDUNG

Noch wird Datenjournalismus in Deutschland nicht systematisch gelehrt. Doch auch in den USA gibt es bisher nur punktuelle Ausbildungsschwerpunkte. In New York hat im Herbst 2010 die COLUMBIA JOURNALISM SCHOOL damit begonnen, erstmals 15 Journalistikstudenten in einem neuen Studiengang zu Datenjournalisten an der Schnittstelle zwischen Journalismus und Informatik auszubilden. In Deutschland beschäftigen sich bisher nur rund ein Dutzend Journalisten und Programmierer intensiv mit dem Thema. Worauf es dabei ankommt, sind sich die Pioniere allerdings weitgehend einig. „Datenjournalismus ist, sich nicht mit selektiv vorgefertigten Daten zufriedenzugeben, sondern alle Rohdaten auszuwerten und diese auch zur Verfügung zu stellen“, betont SEBASTIAN MONDIAL (Langer, U. 2011, S. 41). Er ist Redakteur beim STERN im Team „Investigative

Recherche“ sowie Trainer an der HENRI-NANNEN-SCHULE, an welcher er in diesem Jahr den neuen Lehrbereich Datenjournalismus mit Seminaren für Schüler sowie Schulungen für Redakteure der Verlage GRUNER + JAHR, ZEIT und SPIEGEL aufbaut. Für LORENZ MATZAT, freischaffender Berliner Journalist und Medientrainer,⁶ liegt Datenjournalismus im Schnittpunkt zwischen Infografiken, interaktivem Storytelling und investigativer Recherche. Das sehen auch andere Experten so. Und alle sind sich einig, dass spezielle Programmierkenntnisse weitaus unwichtiger sind als der richtige Umgang mit dem Tabellenkalkulationsprogramm EXCEL oder Grundkenntnisse beim Arbeiten mit GOOGLE Tools oder dem Programm MANY EYES.

Die meisten datenjournalistischen Beispiele, die zumeist mit frei zugänglichen Tools aufbereitet werden, entdeckt man im Lokalen. Wenn sich in Datensätzen lokale Auffälligkeiten finden, springen Regionalzeitungen gerne auf ein Thema an. Oft lassen sich Geodaten auf einer interaktiven Karte verorten. So stellte die TAZ die wahrscheinlichen lokalen Lärmauswirkungen des geplanten Großflughafens BERLIN BRANDENBURG INTERNATIONAL in einer interaktiven Anwendung dar. Doch auch in einigen Fachressorts spielt Datenjournalismus eine immer größere Rolle. Überall dort, wo immer schon Daten gesammelt und in Form von Tabellen und Kurven dargestellt wurden – im Wirtschafts- und Politikjour-

nalismus, in der Börsenberichterstattung oder im Sportjournalismus – entwickeln Experten innerhalb und außerhalb der Medien neue Methoden, um Daten auszuwerten und interaktiv darzustellen.

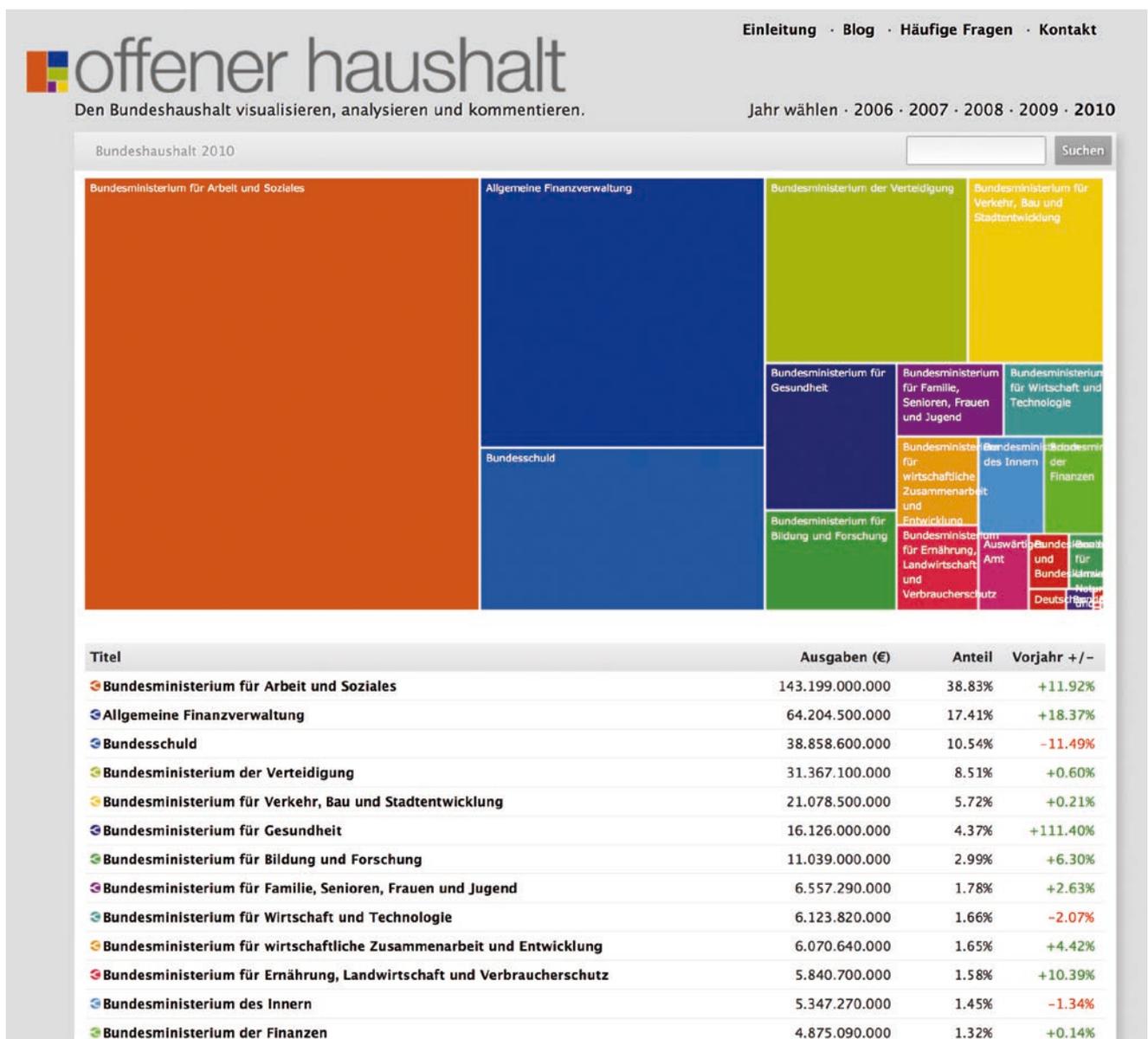
GESCHICHTEN ERGEBEN SICH ERST DURCH DATEN

Die NEW YORK TIMES schuf unmittelbar nach den Olympischen Winterspielen 2010 in Vancouver ein interaktives Klangbild, um darzustellen, wie minimal die Zeitunterschiede zwischen einem siegreichen und einem hinteren Platz im Spitzensport sind.⁷ Der Effekt ist verblüffend und würde mit einer herkömmlichen Tabelle oder Grafik sicher nicht erreicht. Die Berliner Agentur

TACTICAL TOOLS visualisierte die Haushaltsposten des Bundes in einer sogenannten „Tree Map“ (verschachtelte Rechtecke mit maßstabgetreuer Größe). (→ **Abb. 2**) Ihr Projekt „Offener Haushalt“⁸ will mehr Transparenz schaffen, wie der Staat Steuermittel verwendet, und ist das einprägsamere deutsche Pendant zur britischen Webseite „Where does my money go?“. Wer auf einen Posten klickt, gelangt auf Unterbereiche mit feineren Aufschlüsselungen. Insgesamt wurden 7.000 Posten (plus ein Vielfaches an Unterposten) dargestellt. Jeder Haushaltsposten hat eine eigene Internetadresse (URL). So können Onlineberichte illustriert und mit zusätzlichen Informationen ergänzt werden.

Fortsetzung folgt auf Seite 8 >>

ABB. 2: „OFFENER HAUSHALT“



„WÜNSCHENSWERT SIND GRUNDKENNTNISSE STATISTISCHER ZUSAMMENHÄNGE“

ULRIKE LANGER im Gespräch mit CHRISTIANE SCHULZKI-HADDOUTI über die Bedeutung von Datenjournalismus im Fachjournalismus

Bei vielen datenjournalistischen Anwendungen stehen lokale Themen im Vordergrund. Profitiert auch der Fachjournalismus vom Trend Datenjournalismus?

Ich glaube, dass der Fachjournalismus extrem vom Datenjournalismus profitieren kann, weil viele Statistiken oder Daten aus Behörden sehr spezialisiert sind, wie zum Beispiel Umweltdaten. GREGOR AISCH hat jüngst Energiedaten der EU visualisiert. Das ist ein sehr umfassendes Thema, das man für viele Facetten datenjournalistisch aufbereiten kann: Kernenergie, Fusionsforschung, erneuerbare Energien. Dort lassen sich sicher plakative Vergleiche anstellen, die man wiederum in journalistische Beiträge umsetzen kann.

Sind das auch Ihre eigenen Erfahrungen?

Ich habe selbst den Bundeshaushalt nach Daten zu Energiequellen und Finanzierung durchforstet. Gerade bei der Kernfusionsforschung stößt man ziemlich rasch an Grenzen, wenn es darum geht, wie viele Institute wie viel Geld bekommen. Wenn man das vertieft, erkennt man bald, wo gemauert wird und wo es anfängt politisch zu werden. Dort muss man traditionell weiterrecherchieren und kann zu politischen Fragestellungen gelangen, die man vorher nicht gesehen hat. Datenjournalismus muss gar nicht unbedingt zu Visualisierungen führen. Allein schon der Researchweg über die Zahlen kann zu interessanten und manchmal auch ernüchternden Erkenntnissen führen.

Können interaktive Datenvisualisierungen komplexe wirtschaftliche oder wissenschaftliche Sachverhalte besser illustrieren als herkömmliche Grafiken und Schaubilder?

Ja, auf jeden Fall. Man kann die Grafiken entsprechend den eigenen Interessen individuell erforschen und aufbereiten. Es ist auch sinnvoll, Daten kommentierbar zu machen, weil dann über das Feedback zusätzliche Details und Zusammenhänge erfasst werden können. Das lohnt sich natürlich nur bei sehr komplexen und umfangreichen Datensätzen.

Gehört dazu auch, die Rohdaten zur Verfügung zu stellen?

Grundsätzlich ist das wünschenswert, aber es bietet nur dann einen Mehrwert, wenn die Leser damit arbeiten können. Ich habe die Erfahrung gemacht, dass Rohdaten aufgearbeitet und redigiert werden müssen, weil die Formate oftmals nicht zueinanderpassen und Doppelungen enthalten sind. Die Daten müssen immer bereinigt werden. Aber dann sollten sie durchaus einsehbar sein, auch wenn sie ohne Visualisierung nicht wirklich greifbar sind.

Was sollte man als Fachjournalist beherrschen, um Datenjournalismus sinnvoll einsetzen zu können?

Die Mindestvoraussetzung ist: Keine Angst vor Zahlen haben. Wünschenswert sind Grundkenntnisse statistischer Zusammenhänge und dass man Datenmaterial auf seine Verwertbarkeit hin zuverlässig beurteilen kann. Dafür sollte man mindestens mit einem Tabellenkalkulationsprogramm umgehen können, sonst wird es sehr schwierig, wenn nicht gar unmöglich. Man muss mit Excel wirklich spielen können.

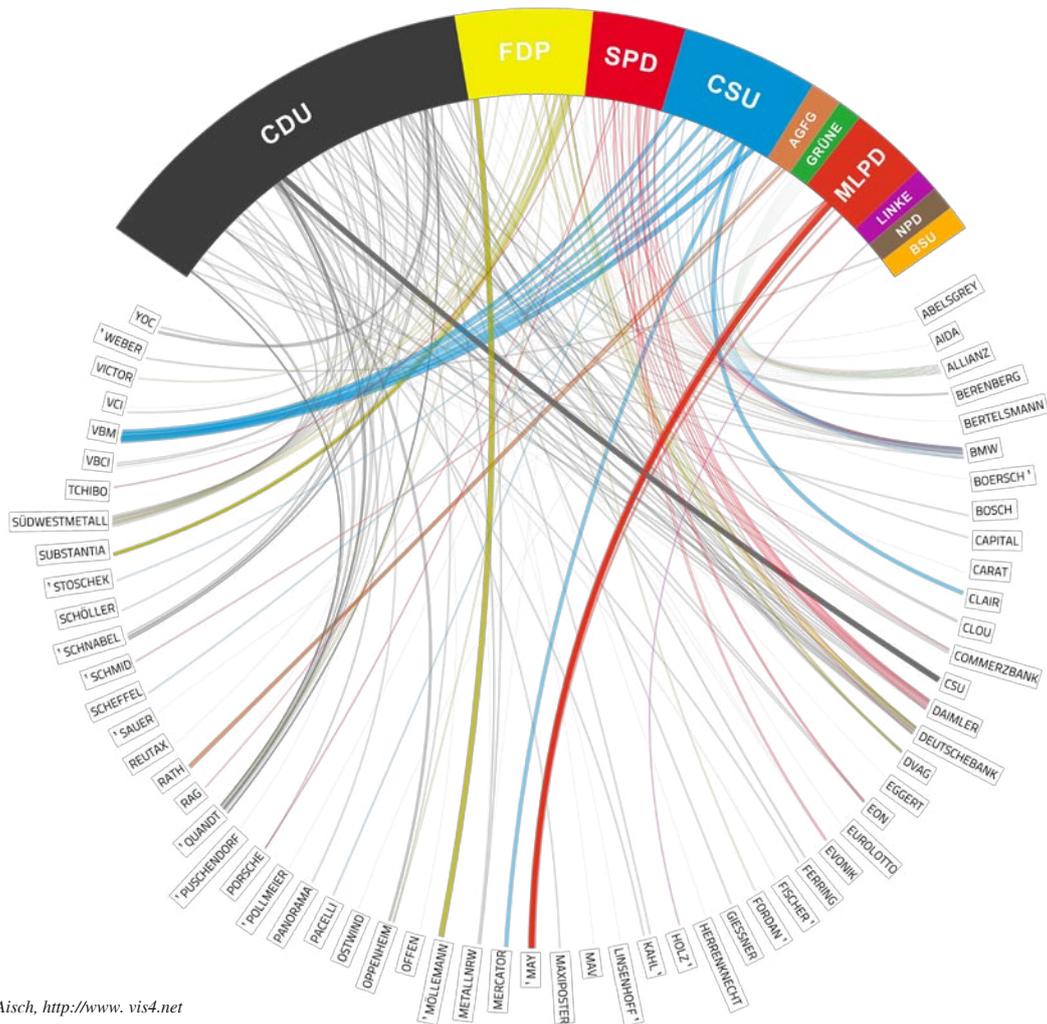
Muss man auch programmieren können?

Nein, aber man muss mit Programmierern einen vernünftigen Dialog führen können und deren Möglichkeiten kennen.

Noch wird Datenjournalismus in Deutschland erst in Ansätzen institutionalisiert gelehrt. Öffnet sich hier gerade ein Zeitfenster für Fachjournalisten, um sich als Pionier mit gefragten Zusatzkenntnissen zu profilieren?

Auf jeden Fall, weil es in Deutschland zu wenig datenjournalistische Angebote und in den Redaktionen zu wenig Ansätze dafür gibt. Es ist zu erwarten, dass Journalisten künftig immer mehr Daten zur Verfügung stehen werden, auch dank der Informationsfreiheitsgesetzte. Wenn man mit diesen Daten umgehen kann, ist das auf jeden Fall ein Wettbewerbsvorteil.

ABB. 3: „PARTEISPENDEN IN DEUTSCHLAND“



Quelle: Gregor Aisch, <http://www.vis4.net>

Auf ganz ähnliche Weise verglich die Münchner Agentur WEBEVANGELISTEN die Haushaltsausgaben eines Durchschnittshaushalts mit einem typischen Hartz-IV-Empfänger-Haushalt.⁹ Datenvisualist GREGOR AISCH,¹⁰ beteiligt an der Umsetzung von „Offener Haushalt“, visualisierte auch Parteispenden in Deutschland (→ Abb. 3), Nebeneinkommen deutscher Bundestagsabgeordneter,¹¹ internationale Hilfen zwischen Staaten¹² oder die Abhängigkeit von Energieimporten im europäischen Vergleich. Er sieht den entscheidenden Unterschied zwischen herkömmlichen Infografiken, die es im Journalismus seit Jahrzehnten gibt, und Datenvisualisierungen neueren Typs in der Interaktivität. Bisweilen ergäben sich Geschichten auch erst durch die Analyse und Visualisierung von Daten, so AISCH im lab-Blog der DEUTSCHEN WELLE: „Wenn ich klassisch Daten analysiere, finde ich Extreme. Ich filtere die Top 10 raus, gucke mir den Besten und den Schlechtesten an, aber die Muster, die sich im uninteressanten Mittelfeld ergeben, die finde ich durch eine Analyse von Tabellen eher selten.“¹³

NICHT JEDE STATISTIK BIETET MEHRWERT

Allerdings sieht der Experte derzeit auch Tendenzen, jede neue Visualisierung zu feiern, nur weil das Thema gerade "en vogue" ist. Das findet auch der Sportjournalist JENS WEINREICH übertrieben. „Im Sport liegt es schon in der Natur der Sache, dass sich alles messen und summieren und analysieren lässt“, so WEINREICH.¹⁴ Nur sei das Sammeln von Daten und Präsentieren von Statistiken noch kein Journalismus, sondern erfülle zumeist nur eine Unterhaltungsfunktion. Beim Thema Korruption im Sport, Schwerpunkt des mehrfach preisgekrönten investigativen Journalisten, seien Fähigkeiten in der Aufbereitung von Daten weniger gefragt als die Kunst, an Dokumente zu kommen, die von interessierter Seite lieber geheim gehalten würden. „Wenn es zum Beispiel darum geht, Bestechungen bei der jüngsten Vergabe der Fußball-Weltmeisterschaften 2018/2022 nachzuweisen, nützt mir erstmal auch keine Excel-Datei“, so WEINREICH.¹⁵ Konkrete Beispiele, wie echte Investigativleistungen im

Sportbusiness mit Tools des Datenjournalismus erreicht worden sind, kennt der Sportjournalist nicht. Zwei Beispiele für Sportjournalismus, der letztlich nur noch aus Daten besteht, stammen aus den USA. Die Smartphone-App GAMECHANGER soll Trainern, Vereinsmitgliedern und Fans erleichtern, Spielverläufe und Ergebnisse im Blick zu behalten. Gleichzeitig sollen lokale Medien auf die von Nutzern erhobenen Daten zugreifen können. Eine weitere Software namens STATMONKEY wurde von Studenten an der NORTHWESTERN UNIVERSITY in der Nähe von Chicago entwickelt. Hier setzt ein Algorithmus ganz ohne Journalisten komplette Spielberichte aus Phrasen und Textbausteinen zusammen. Noch handelt es sich allerdings nur um eine Projektstudie.

CHANCEN FÜR PIONIERE

Doch Projekte wie STATMONKEY oder real existierende Contentfabriken wie DEMANDMEDIA, in denen ein Algorithmus vorgibt, über welche Stichworte geschrieben werden soll, weil sich dort über die GOOGLE-Suchergebnisse lukrative Werbung schalten lässt, sind nur ein Vorgeschmack. Viele Routineaufgaben im Sport- oder Wirtschaftsjournalismus könnten künftig von Software statt Journalisten übernommen werden. Im Gegenzug können sich Fachjournalisten die Datenfülle als zukunftsträchtiges Feld nutzbar machen. Sie können lernen, Informationen und Erkenntnisse optimal aus Datensätzen herauszuholen, und – ob in einer Redaktion oder freischaffend – ihre Expertise zu einem Alleinstellungsmerkmal auszubauen. Pioniere auf diesem Gebiet können über Projektmangel nicht klagen. Denn eines zeichnet sich ab: Große Datensätze, die es zu durchsuchen und aufzubereiten lohnt, wird es künftig immer häufiger geben. //

ENDNOTEN:

- ¹ Zum Beispiel: BP Oil Spill, <http://www.nytimes.com/interactive/2010/05/01/us/20100501-oil-spill-tracker.html>
- ² „MPs' expenses: all the revelations, as a spreadsheet“, <http://www.guardian.co.uk/news/datablog/2009/may/13/mps-expenses-houseofcommons>
- ³ „Thema Datenjournalismus: Das Beispiel ‚Zeit online‘“, Interview von Ulrike Langer mit Sascha Venohr, <http://www.mediummagazin.de/magazin-plus/thema-datenjournalismus-das-beispiel-zeit-online/>
- ⁴ <http://www.zeit.de/gesellschaft/zeitgeschehen/todesopfer-rechter-gewalt>
- ⁵ <http://blog.zeit.de/open-data/>, <http://blogs.taz.de/open-data/>
- ⁶ <http://www.datenjournalist.de/>
- ⁷ „Fractions of a Second: An Olympic Musical“: <http://www.nytimes.com/interactive/2010/02/26/sports/olympics/20100226-olysymphony.html>
- ⁸ <http://bund.offenerhaushalt.de/>
- ⁹ „Durchschnittliche Haushaltsausgaben in Deutschland“, <http://webevangelisten.de/verbraucherpreise/>
- ¹⁰ <http://vis4.net/blog/de>
- ¹¹ <http://www.nebeneinkuenfte-bundestag.de>
- ¹² Zum Beispiel EU-Agrarsubventionen: <http://www.farmsubsidy.org/>
- ¹³ Gregor Aisch im Interview mit Steffen Leidel und Marcus Bösch, <http://training.dw-world.de/ausbildung/blogs/lab/?p=1955>
- ¹⁴ Antworten von Jens Weinreich auf Fragen von Ulrike Langer per E-Mail
- ¹⁵ Antworten von Jens Weinreich auf Fragen von Ulrike Langer per E-Mail



Die Autorin **ULRIKE LANGER** arbeitet seit 1995 als freie Medienjournalistin in Köln und hat sich vor allem auf Themen rund um den digitalen Medien- und Journalismuswandel spezialisiert. Sie schreibt Fachbeiträge, gibt Seminare, hält Vorträge, moderiert und diskutiert auf Kongressen. Ulrike Langer betreibt das Medienfachblog medialdigital.de und ist seit 2011 Mitglied der Nominierungskommission des Grimme Online Award.



CHRISTIANE SCHULZKI-HADDOUTI ist freie Wissenschaftsjournalistin in Bonn. Sie betreibt das Fachblog [Kooptech](http://kooptech.de) und beschäftigt sich mit Datenjournalismus.

LITERATUREMPFEHLUNGEN & QUELLEN FÜR DATENSÄTZE

Eurostat Datenbank: → http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database

Langer, U. (2011): Schaubild statt Klickstrecke, in: *Mediummagazin* 01/02 2011, S. 41.

Open Knowledge Foundation: CKAN, → <http://www.ckan.net/>

Pro Publica: Scraping for Journalism: A Guide for Collecting Data,

→ <http://www.propublica.org/nerds/item/doc-dollars-guides-collecting-the-data>

Stanford University: Journalism in the Age of Data,

→ <http://datajournalism.stanford.edu/>

Statistisches Bundesamt: Genesis Online Datenbank, → <https://www-genesis.destatis.de/genesis/online>

Stray, J.: A computational journalism reading list,

→ <http://jonathanstray.com/a-computational-journalism-reading-list>