



# Data Journalism

Indira Dupuis



# Data Journalism

**Indira Dupuis**

## Impressum

© 2016 DFJV Deutsches Journalistenkolleg GmbH, Berlin  
Alle Rechte vorbehalten.

Der gesamte Inhalt des vorliegenden Studienbriefs (Texte, Bilder, Grafiken, Design usw.) und jede Auswahl davon unterliegt dem Urheberrecht und anderen Gesetzen zum Schutze geistigen Eigentums der DFJV Deutsches Journalistenkolleg GmbH oder anderer Eigentümer. Jede Verwertung außerhalb der engen Grenzen des Urheberrechts ist ohne Zustimmung des Eigentümers unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Zuwiderhandlungen werden zivil- und strafrechtlich verfolgt.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Text berechtigt auch ohne besondere Kennzeichnung nicht zur Benutzung solcher Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung. Sämtliche verwendete Marken sind Eigentum der jeweiligen Rechteinhaber.

Die DFJV Deutsches Journalistenkolleg GmbH und ihre Dozenten und Autoren haben höchste Sorgfalt bei der Erstellung des vorliegenden Studienbriefs angewandt. Dennoch übernehmen sie keinerlei Verantwortung oder Haftung für Richtigkeit oder Vollständigkeit sowie eventuelle Fehler oder Versäumnisse innerhalb des Studienbriefs. Die Inhalte und Materialien werden unter Ausschluss jeglicher Gewährleistung zur Verfügung gestellt. Insbesondere erfolgt die Anwendung von im Studienbrief dargestellten Erkenntnissen auf Gefahr des Teilnehmers.

Printed in Germany.

[www.journalistenkolleg.de](http://www.journalistenkolleg.de)

# Data Journalism

## Allgemeine Lernziele

Wenn Sie dieses Paper durchgearbeitet haben, können Sie

- das Genre und seine Ableitungen definieren;
- einen Überblick der historischen Entwicklung des Genres wiedergeben;
- Data Journalism kritisch reflektieren.

Datenjournalismus hat sich historisch punktuell in verschiedenen Redaktionen und an einzelnen Journalistenschulen entwickelt, abhängig vom Engagement einzelner computeraffiner Redakteure oder Professoren. Das Genre ist thematisch nicht spezifisch, in den letzten Jahren immer bedeutender geworden und inzwischen im journalistischen Mainstream angekommen.<sup>1</sup>

Datenjournalismus unterscheidet sich von anderen Genres durch den Ansatz, die Arbeitsmethoden und die Darstellungsweisen, insbesondere durch die Fokussierung auf die übersichtliche Aufbereitung großer Datenmengen („Big Data“)<sup>2</sup> und die Nachvollziehbarkeit der darauf basierenden Aussagen in der journalistischen Berichterstattung. Datenjournalismus kann insbesondere komplexe Sachverhalte veranschaulichen, interessante Zusammenhänge aufzeigen und die Berichterstattung analytisch vertiefen.

Der Begriff Datenjournalismus entspricht dem englischen Begriff des Data Journalism oder auch des Data Driven Journalism.<sup>3</sup> Weitere Begriffe lassen sich zuordnen: Computational Journalism, Scientific Journalism, Database Journalism oder Structured Journalism und die älteren Formen des Datenjournalismus, Precision Journalism und Computer Assisted Reporting. Sie werden hier unter Datenjournalismus zusammengefasst und nach der summarischen Beschreibung dieses journalistischen Genres kurz einzeln erläutert. Nach dieser Erläuterung und Abgrenzung erfolgt zuletzt eine historische und kritische Einordnung des Datenjournalismus im journalistischen Feld.

1 Vgl. Howard (2014).

2 Siehe Beitrag „Computational Journalism“ in diesem Band.

3 Vgl. Mair (2013).

## ■ Definition und Geschichte

Um einen konsistenten Begriff zu verwenden, wird hier Datenjournalismus ausgehend von den ihn kennzeichnenden Strukturen und nicht ausgehend von bestimmten Themen oder einer spezifischen Präsentationsweise von Inhalten definiert.<sup>4</sup> Unter diesen Strukturen werden die an den gesamtgesellschaftlichen Funktionen des Journalismus orientierten Organisationsstrukturen verstanden, die sich unter den jeweiligen gesellschaftlichen Rahmenbedingungen ausprägen. Journalisten agieren in diesen Strukturen in ihrem journalistischen Umfeld und eignen sich die dafür notwendigen Kenntnisse, Arbeitsweisen und Haltungen an. Bei ihrer Arbeit erhalten sie selbst diese Strukturen aufrecht und entwickeln sie zielorientiert weiter.

In dieser Definition ist der Marktwert von journalistischen Angeboten als zentrale Größe in den weitgehend privatwirtschaftlich organisierten Medienmärkten inbegriffen. Es wird berücksichtigt, dass Journalismus als gehandeltes Gut nicht unabhängig von den wirtschaftlichen Organisationsformen zu denken ist und dabei auch von der Nachfrage nach den verschiedenen journalistischen Angeboten abhängt.

Journalismus hat in einer allgemein normativen, demokratietheoretisch fundierten Definition, wie im Presserecht oder in der Selbstbeschreibung des Deutschen Journalisten Verbands<sup>5</sup>, die Funktion der Herstellung von politischer Öffentlichkeit durch die Selektion und verständliche Aufarbeitung von gesellschaftlich relevanten Themen für die politische Willensbildung. Der hier verwendete funktional-strukturelle Begriff des Journalismus ist breiter und schließt auch andere Informations- und Orientierungsfunktionen ein, zum Beispiel den Kultur-, Wissenschafts- oder Servicejournalismus.

Der normative Teil der Definition betrifft vor allem die Qualitätsanforderungen, die an journalistische Produkte gestellt werden.<sup>6</sup> Sie ist von zentraler ideologischer Bedeutung in einigen journalistischen Genres, betrachtet man beispielsweise die politisch hochmotivierten investigativen Journalisten, die politisch brisante Themen bearbeiten, oder die Berichterstattung über geheime staatliche

4 Mit diesem systemtheoretisch-institutionellen Ansatz im Sinne u. a. von Ulrich Saxer (vgl. Künzler 2012) unterscheidet sich die Begriffsdefinition im vorliegenden Text von anderen, z. B. der des französischen Datenjournalisten Nicolas Kayser-Bril, der 2011 über den Begriff eine Fachdiskussion angestoßen und diese zusammengefasst hat. Er bezeichnet den Begriff Datenjournalismus als „Mantel-“ oder „Container-Begriff“, „für einen Trend im Journalismus und des Informationsmanagements“ (nach Matzat 2011a).

5 Vgl. DJV (2009), S. 4.

6 Vgl. McQuail (1992).

Dokumente, die einzelnen datenjournalistisch versierten Redaktionen durch die Organisation Wikileaks zur Verfügung gestellt wurde und die einen datenjournalistisch bedeutenden Wendepunkt darstellt.

Die hier verwendete breite Definition kann dagegen alle datenjournalistischen Ausprägungen, nicht nur die investigativen Projekte, in einem Begriff zusammenfassen. Datenjournalismus wird aus dieser Perspektive über seine spezielle Funktion, die exponentiell wachsenden Datenmengen in der Informationsgesellschaft zu bewältigen, begriffen.

Historisch ist Datenjournalismus im Zuge der Digitalisierung und Vernetzung entstanden, die große gesellschaftliche Potenziale auch für die Formierung der politischen Öffentlichkeit und für Innovationen in der Wirtschaft haben. Gleichzeitig bergen diese Entwicklungen jedoch auch große politische und wirtschaftliche Herausforderungen. Gesellschaften müssen heute eine deutlich komplexere Informationslage bewältigen als früher. Dies führt sowohl zu neuen Anforderungen an die Fähigkeiten der einzelnen Menschen im Umgang mit Medien in Alltag und Beruf als auch zu einer neuen gesellschaftlichen Ungleichheit (digital divide). Datenjournalismus gehört dabei in eine Entwicklungslinie mit den journalistischen Ansätzen, die einen methodisch-wissenschaftlichem Anspruch verfolgten. Diese Tradition reicht bis ins frühe 19. Jahrhundert zurück und wurde von Emile Zola mit seinen Sozialreportagen oder Arthur Young mit seinen Reiserreportagen begonnen. In dieser Zeit entstanden auch statistische Bureaus, die amtliche Statistiken für verschiedene Publika aufbereiteten.<sup>7</sup>

## Computer Assisted Reporting

Die computergestützte Berichterstattung ist ein Vorläufer und immer noch Bestandteil des heutigen Datenjournalismus.<sup>8</sup> Der Begriff Computer Assisted Reporting bezieht sich auf die Recherchetechnik, die sich technischer und methodischer Ansätze aus der sozialwissenschaftlichen Forschung bedient.

Mit dem Boom quantitativer und deskriptiver empirischer Forschungsprojekte in allen Bereichen der Sozialwissenschaften und Psychologie in den 1960er-Jahren fanden insbesondere die quantitativen und statistischen Methoden der Datenerhebung und -auswertung Eingang in den Journalismus. Sie wurden in einzelnen aufsehenerregenden investigativen journalistischen Projekten effektiv eingesetzt und konnten zur sachlichen Diskussion von die Öffentlichkeit polarisierenden politischen Konflikten beitragen. Beispiele sind die Berichterstattungen von Cla-

<sup>7</sup> Vgl. Haas (1999), S. 145 ff.; Weaver und McCombs (1980).

<sup>8</sup> Vgl. Hewett (2013), S. 3.

rence Jones über das Justizsystem für den Miami Herald im Jahr 1969 und von Philipp Meyers über die Rassenunruhen in Detroit 1967. Während sich Jones auf der Basis des gerade in Kraft getretenen Informationsfreiheitsgesetzes Zugang zu amtlichen Dokumenten und Statistiken verschaffen konnte, die er mithilfe von Computerprogrammen auswertete, tat sich Meyer mit Nathan Caplan vom Institut für Sozialforschung an der Universität von Michigan zusammen, um mithilfe eigener Umfragen valide Daten für die Berichterstattung zu erheben. Meyer veröffentlichte später ein grundlegendes Lehrbuch zu seiner Vorgehensweise, die er selbst – ganz im Stil der Zeit – als „Precision Journalism“, Präzisionsjournalismus, bezeichnete.

## Präzisionsjournalismus

Der Begriff Precision Journalism wurde als Titel des Lehrbuchs von Philip Meyer aus dem Jahr 1973 bekannt. In diesem Grundlagenwerk wird der Einsatz der quantitativ-statistischen Methoden der empirischen Sozialwissenschaft und Verhaltensforschung im Journalismus beschrieben. In Deutschland wurde Präzisionsjournalismus auch unter dem Namen empirischer Journalismus gefasst.<sup>9</sup> Mit dem Begriff Precision Journalism wurde die journalistische Richtung auch vom gleichzeitig entstehenden New Journalism abgegrenzt.<sup>10</sup> Beide Formen wandten sich vom traditionellen Rollenbild des amerikanischen Journalismus, dem des distanziert beobachtenden, „professionellen Amateurs“<sup>11</sup>, ab, zugunsten eines am Geschehen Beteiligten.

Während sich die Vertreter des New Journalism für ihre persönliche und authentische Darstellung von Themen literarischer Stilmittel bedienten und damit die Grundlage der heute als Storytelling bezeichneten Technik im Journalismus legten, wurde dieser subjektive Ansatz im Präzisionsjournalismus abgelehnt. Vielmehr wurde das Recherchieren und Auswerten von umfangreichen Datensätzen nach aus der Sozialwissenschaft stammenden methodischen Gütekriterien als neue Form eines Journalismus propagiert. Dessen Vertreter nahmen die Rolle von informierten Experten ein, die selbst aktiv an der Themensetzung beteiligt sein wollten.

Problematisch ist dieses journalistische Selbstverständnis im Hinblick auf die gesellschaftliche Aufgabenteilung und eine vor allem kontrollierende Funktion des Journalismus, gefährdet es doch genauso wie der subjektive New Journalism die Distanziertheit im Journalismus von den Inhalten der Berichterstattung. Durch eigene Erhebungen sind Journalisten in einer Auftraggeber- oder sogar forschungsleitenden Rolle und nicht mehr unabhängig. So sind sie nicht mehr

9 Vgl. Haas (1999), S. 108 ff.

10 Mit Tom Wolfe als wohl bekanntestem Vertreter dieses New Journalism.

11 Der Begriff wurde 1967 vom damaligen Herausgeber des Wallstreet Journals, Vermont Royster, geprägt (zitiert nach Meyer 1991).

in der Lage, das Thema und seine Relevanz für die Öffentlichkeit so kritisch zu hinterfragen, wie es bei der Berichterstattung über andere wissenschaftliche Forschungsergebnisse üblich ist.<sup>12</sup>

Eine der Begründungen im Präzisionsjournalismus, Datengewinnung und Datenanalyse zur Aussagengewinnung zu nutzen, war, dass die Welt heute zu komplex wäre, um lediglich mit den bisherigen journalistischen Mitteln zu arbeiten. Es war also nicht lediglich ein aktivistischer Ansatz, sondern es ging den Vertretern um die Etablierung eines höheren Qualitätsanspruchs im Journalismus beim Umgang mit Daten.

Ein wesentliches Ziel war dementsprechend auch die Qualitätssicherung im Journalismus bei der Recherche für die Berichterstattung über datenintensive Themen. Philipp Meyer vertrat daher die nachvollziehbare Meinung, dass es in allen Redaktionen wenigsten einige Journalisten geben müsse, die statistische und sozialwissenschaftliche Erhebungsmethoden beherrschen.<sup>13</sup> Die Kenntnis der statistischen Analyse- und Auswertungsmethoden sollte verhindern, nur die Interpretationen von Forschungsdaten aus Presstexten von Organisationen oder Unternehmen sowie Zusammenfassungen von wissenschaftlichen Autoren, Marktforschungsinstituten, Behörden oder (lobbyistischen) Institutionen unkritisch zu übernehmen. So sollte dafür gesorgt werden, dass Daten differenziert erläutert und nicht als Feigenblatt für eine sensationalistische, meinungslastige Präsentation missbraucht werden.<sup>14</sup>

## Scientific Journalism

Der Begriff Scientific Journalism wurde von Julian Assange geprägt, dem Gründer und Chefredakteur von Wikileaks.<sup>15</sup> Er ist abzugrenzen vom „journalisme scientifique“, der französischen Übersetzung des Begriffs Precision Journalism, die als Bezeichnung der journalistischen Berichterstattung über wissenschaftliche Themen gebräuchlich ist. Im Deutschen spricht man bei diesem eigenständigen Genre von „Wissenschaftsjournalismus“, das im englischen Sprachgebrauch wiederum mit „science journalism“ bezeichnet wird.

Wikileaks betreibt eine Internet-Plattform, über die anonym geheimes Informationsmaterial veröffentlicht werden kann. Seit der Gründung 2006 wurden hier zahlreiche Geheimdokumente veröffentlicht, die zu Enthüllungsskandalen führten, unter anderem Militärdokumente zu den Kriegen in Afghanistan und Irak sowie zum Gefangenenlager Guantanamo Bay.

<sup>12</sup> Vgl. Haas (1999), S. 108.

<sup>13</sup> Diese Idee wurde auch in Deutschland u. a. an der Freien Universität Berlin von Lutz Erbring und Hans-Jürgen Weiß vertreten.

<sup>14</sup> Weischenberg und Rakers (2001), S. 42 f.

<sup>15</sup> Vgl. Meyer (1991).



Durch die Bezeichnung dieser Aktivitäten als wissenschaftlicher Journalismus plädieren die beteiligten Personen auf Schutz vor rechtlichem Zugriff durch die US-amerikanische Regierung, unter Berufung auf den 1. Zusatzartikel der Verfassung der Vereinigten Staaten, der die Pressefreiheit garantiert.

Im Unterschied zu Precision Journalism, bei dem es um die Anwendung von wissenschaftlichen Methoden im Journalismus geht, forderte Assange bei seinen Ausführungen zu Scientific Journalism, das eigene Datenmaterial mit zu veröffentlichen, wie dies bei wissenschaftlichen Publikationen üblich ist. Damit könnten Rezipienten, ganz im Sinne (sozial-) wissenschaftlicher Qualitätsstandards in der Forschung, die Auswertung dieser Dokumente selbst nachvollziehen<sup>16</sup> – unter der Voraussetzung, dass die Rezipienten mit den Methoden der statistischen Analyse und Auswertung vertraut sind.

Über den Ansatz des Scientific Journalism wurde nach der Prägung des Begriffs durch Assange diskutiert, ob er den traditionellen Journalismus revolutionieren würde. Insbesondere vor dem Hintergrund, dass Wikileaks die Strategie, Analysen und geleakte Dokumente auf den eigenen Seiten zu veröffentlichen, zugunsten der Kooperation mit investigativen Journalisten von etablierten Medienhäusern änderte.<sup>17</sup> Grundsätzlich kann jedoch gerade der Journalismus, der aus diesen Kooperationen entstand, klar in das Genre des Datenjournalismus eingeordnet werden.

## Database oder Structured Journalism

Mit Database oder auch Structured Journalism wird eine spezifische Organisationsstruktur von im Internet publizierten, klassisch massenmedialen Inhalten bezeichnet.

Der Begriff geht zurück auf den Informatiker und Journalisten Adrian Holovaty, der im Jahr 2006 dieses Konzept in einem Blogpost skizzierte.<sup>18</sup> Er wandte sich darin gegen die „Story-Zentriertheit“ des herkömmlichen Journalismus, das heißt die Veröffentlichung von in sich abgeschlossenen Artikeln. Stattdessen schlägt er eine weniger statische Form der Präsentation von Information vor: Online-Journalismus solle Inhalte beständig mit neuen Daten aktualisieren, fortschreiben, auslesbar machen und so die Möglichkeiten des Internets besser ausnutzen. Jeff Jarvis, Professor für Journalismus an der City University of New York und Betreiber des Blogs Buzzmachine,<sup>19</sup> bezeichnete vor diesem Hintergrund den journalistischen Artikel in der Online-Welt als Luxus: „When digital comes first and print

<sup>16</sup> Vgl. Khatchadourian (2010).

<sup>17</sup> Vgl. Lynch (2012).

<sup>18</sup> Vgl. Holovaty (2006).

<sup>19</sup> <http://www.buzzmachine.com/about-me/>

last, then the article is something you need to put together to fill the paper; it's not the goal of the entire process. The process is the goal of the process: keeping the public constantly informed."<sup>20</sup>

Die Idee der journalistischen strukturierten Datenaufbereitung wurde inzwischen weiterentwickelt und in bedeutenden Datenblogs der Qualitätsmedien integriert.<sup>21</sup> Allerdings wird sie nicht in der geforderten Breite auf alle Bereiche des Online-Journalismus angewendet.

## Computational Journalism

Computational Journalism ist der neueste Begriff innerhalb des Genres Datenjournalismus. Unter diesem wird insbesondere das Problem fokussiert, dass interessante Daten nicht für den Journalismus, sondern für verschiedene andere Zwecke mit verschiedenen Computerprogrammen und mit unterschiedlichen wissenschaftlichen Methoden erstellt wurden. Dadurch sind sie oft in einer nicht leicht zugänglichen Form strukturiert. Dies führt dazu, dass sie ohne spezialisierte Vorkenntnisse einfach ausgewertet werden können und so der Allgemeinheit und insbesondere dem Journalismus daher kaum zur Verfügung stehen. Da die Daten außerdem in völlig unterschiedlichen Bereichen und Datensets vorliegen, muss zwischen diesen erst ein Zusammenhang hergestellt werden.

Aus dieser Problematik wird die Forderung abgeleitet, dass heute (sozial-) wissenschaftliche Methoden und Informatik im Journalismus organisational integriert werden müssen, um die journalistische Auswertung von Big Data, also großen, komplexen und auch unstrukturierten Datensätzen, zu ermöglichen. Dabei sollen neue, an der journalistischen Funktion orientierte Arbeitstechniken für den Journalismus entwickelt werden, insbesondere im Bereich Informatik und Webdesign. Journalisten sollen so in die Lage versetzt werden, vorhandene Datenmengen zu erschließen, um darin die gesellschaftlich relevanten Themen erkennen und qualifiziert interpretieren zu können, die in der Öffentlichkeit diskutiert werden müssen: „For decades, the computing community has empowered individuals to seek information, improving their lives in the process. Few fields have done more to give citizens the tools they need to govern themselves. Few fields today need computer scientists more than public-interest journalism.“<sup>22</sup>

Während Computerspezialisten auf organisatorischer Ebene bis in die 2010er-Jahre von Journalisten im Datenjournalismus nur als Dienstleister angesehen wurden, steht Computational Journalism für eine ebenbürtige Partnerschaft der

<sup>20</sup> Vgl. Jarvis (2011).

<sup>21</sup> Vgl. u. a. die Open Platform des Guardian unter <http://www.theguardian.com/open-platform> [17.10.2014].

<sup>22</sup> Vgl. Cohen, Hamilton und Turner (2011), S. 71.

Berufsgruppen im Journalismus bzw. für die Kombination beider Qualifikationen in einer Person, dem sogenannten Journo-Coder. Diese Doppelqualifikation ist aktuell unter Journalisten gefragt und wird auf Konferenzen und in Workshops diskutiert und vermittelt. In Großbritannien wurde 2014 sogar ein einjähriges Masterstudienprogramm für Computational Journalism an der Universität Cardiff eingerichtet.<sup>23</sup>

Auch wenn am Markt bereits selbstständig arbeitende Journo-Coders agieren, erscheint es auf organisationaler Ebene als nicht ausreichend, lediglich die Doppelqualifikation der einzelnen Journalisten zu fördern. Vielmehr ist zu vermuten, dass sich Kooperationen zwischen Journalisten, Wissenschaftlern und Programmierern in Zukunft bezahlt machen.

## ■ Techniken, Organisation und Bedeutung

Die Bedeutung des Datenjournalismus nimmt zu, da er die durch Digitalisierung und Vernetzung entstandene, unüberschaubare Datenmenge, auch als „Big Data“ bezeichnet, strukturieren kann. Soziale Netzwerke, Unternehmen, Behörden, Hochschulen und andere Institutionen und Organisationen produzieren große Datenmengen, die in unzähligen Sets unterschiedlicher Rohdaten vorliegen. Aus diesen kann der Datenjournalismus relevante und interessante Erkenntnisse extrahieren und der Öffentlichkeit sowohl zugänglich als auch verständlich machen. Er ist Teil der veränderten Möglichkeiten im Journalismus und steht auch für neue Anforderungen an diesen Beruf.

Datenjournalismus zeichnet sich dadurch aus, dass die der Berichterstattung zugrunde liegenden Daten erkennbar sind und in manchen Fällen auch kontinuierlich aktualisiert werden. Für die Auswertung von Informationen werden vorhandene Datensätze bzw. Daten aus unterschiedlichen Quellen über Computerprogramme zusammengeführt und maschinell ausgewertet – oder auch manuell von Journalisten gesucht, bewertet und zusammengeführt. Datenjournalistische Artikel enthalten Abbildungen und Tabellen, die die Berichterstattung teilweise aus ganz unterschiedlichen Perspektiven illustrieren. Die Angebote werden vielfach separat in Infoblogs der Medienoutlets und oft auch als Applikationen publiziert; damit sind sie interaktiv nutzbar und können von den Rezipienten nach unterschiedlichen Kriterien ausgelesen und – in selteneren Fällen – sogar weiter bearbeitet werden.

<sup>23</sup> Die Internetseite zum Studiengang: <http://courses.cardiff.ac.uk/postgraduate/course/detail/p411.html>.

Die automatisierteste Form des Datenjournalismus ist das wobbing, die Anwendung programmierter Recherchertools für Echtzeitdaten. Eingesetzt wurden sie beispielsweise bei der Auswertung der Kommunikation über die Twitter-Accounts der Kandidaten zu den US-Kongresswahlen von 2010 durch die New York Times.<sup>24</sup> Weitere Recherchetechniken des Datenjournalismus sind Data Scraping, bei dem Informationen aus Websites im Internet automatisch ausgelesen werden, und Crowdsourcing, bei dem die Rezipienten aufgerufen werden, die ihnen vorliegenden Informationen, zum Beispiel zu regional differierenden Benzinpreisen, der Redaktion zuzuschicken.<sup>25</sup>

Auch Angebote wie der Zugmonitor auf sueddeutsche.de<sup>26</sup> oder der Kriminalitätsmonitor für Chicago von Adrian Holovaty<sup>27</sup> werden unter dem Begriff Datenjournalismus zusammengefasst, auch wenn es sich hier um automatisierte Informationsservices ohne große journalistische Zusatzleistungen handelt.<sup>28</sup>

Datenjournalismus ist einerseits als eine Spezialisierung zu verstehen und andererseits als eine Zusatzqualifikation. Datenjournalisten recherchieren mithilfe von Algorithmen, käuflichen und selbst programmierten Computerprogrammen und analysieren die Ergebnisse auf der Basis ihrer Kenntnisse von Statistik und wissenschaftlichen Methoden.

Organisational steht Datenjournalismus jedoch auch für die Integration von Grafikern und Programmierern in das journalistische Team.<sup>29</sup> In Deutschland werden viele datenjournalistische Projekte nicht in den Redaktionen selbst, sondern von Agenturen wie OpenDataCity für die Online-Seiten traditionsreicher Printmedien wie Die Zeit und Der Spiegel realisiert.

Datenjournalismus steht in einer langen Tradition journalistischer Ansätze, bei der Recherche und Argumentation hinsichtlich der methodischen Exaktheit die Nähe zur Wissenschaft gesucht haben. Dabei wird von den Vertretern dieses jour-

24 [http://www.nytimes.com/interactive/us/politics/2010-twitter-candidates.html?\\_r=0](http://www.nytimes.com/interactive/us/politics/2010-twitter-candidates.html?_r=0)

25 Vgl. Matzat (2011).

26 <http://www.sueddeutsche.de/thema/Bahn-Versp%C3%A4tungen>

27 <http://chicago.everyblock.com/crime/>

28 In den USA sind Auskunftsportale auf der Basis von Behördendaten verbreiteter als in Deutschland. Diese Daten stehen dort aufgrund des bereits 1967 etablierten Informationsfreiheitsgesetzes und einer anderen Informationskultur der Behörden umfangreicher zur Verfügung. Die Bedeutung des Informationsfreiheitsgesetzes (Freedom of Information Act) für den Journalismus und insbesondere im Datenjournalismus hat sich auch in Großbritannien nach seiner Einführung 2000 gezeigt (vgl. Hewett 2013). In Deutschland ist diese Informationsfreiheit gesetzlich noch nicht in allen Bundesländern etabliert. Ein Paradigmenwechsel im Umgang mit der Öffentlichkeit hat noch nicht stattgefunden, sodass in Behörden oft die Haltung, das Amtsgeheimnis wahren zu wollen, überwiegt (vgl. ROG 2013).

29 Vgl. Matzat (2011).

nalistischen Genres keineswegs eine Objektivität in der Darstellung angestrebt und durchaus eine subjektive, in der Sache engagierte Rolle eingenommen. Lediglich Fakten sollen die Argumentation leiten, die in der Berichterstattung mit den Quellen offengelegt werden und dadurch nachvollziehbar sind. Darüber hinaus ist ein zentrales Anliegen im Datenjournalismus, dass allgemein Daten möglichst gut strukturiert zugänglich sind. So sollen die Rezipienten in die Lage versetzt werden, sie nach eigenen Interessen zu sortieren oder zu analysieren und sich damit nicht nur umfassend zu informieren, sondern unter Umständen sogar zur Berichterstattung beizutragen zu können.

Die Darstellungstechniken des Datenjournalismus setzen sich von den konventionellen journalistischen Techniken ab, die durch Personalisierung komplexe Themen zu veranschaulichen suchen. Es soll gerade nicht vom Einzelfall ausgehend eine abgeschlossene Geschichte erzählt werden; die Daten sollen möglichst für sich sprechen und, falls machbar, auch fortlaufend aktualisiert und zu öffentlichen thematischen Datenbanken ausgebaut werden. Daher sind Navigationshilfen und Visualisierungstools für die datenjournalistische Darstellung essenziell. Vertreter des Datenjournalismus wenden sich aber nicht nur aus diesen pragmatischen Gründen gegen die Personalisierung von Inhalten in der Berichterstattung, sondern auch wegen des journalistischen Selbstverständnisses. Während sich der konventionelle Journalismus häufig an Eliten orientiert und beispielsweise Interviews mit Unternehmern ganze Seiten im Wirtschaftsteil füllen oder parteipolitische Diskussionen über Inhalte als persönliche Auseinandersetzungen zwischen einzelnen Funktionsträgern dargestellt werden, entspricht dies aus der datenjournalistischen Perspektive nicht der Vorstellung vom journalistischen Auftrag. Dieser wird hier als detaillierte Beschreibung der Sachverhalte in ihren verschiedenen Dimensionen sowie deren Veranschaulichung verstanden, um gerade komplexe Themen möglichst kenntnisreich öffentlich diskutieren zu können.

Es gibt jedoch auch im Datenjournalismus Angebote, die weniger journalistisch anspruchsvoll und investigativ sind und eher flüchtigen, unterhaltenden Charakter haben. Mit Spielereien, wie beispielsweise der einfachen Auswertung der meistgesuchten Schlagwörter auf Google in Deutschland auf [vice.de](http://vice.de)<sup>30</sup>, oder sehr sinnvollen, aber wenig journalistischen Serviceleistungen wie dem Zugmonitor auf [sueddeutsche.de](http://sueddeutsche.de) wird teilweise sehr erfolgreich der für die Anzeigenerlöse maßgebliche Datenträffic auf den Online-Seiten der anbietenden Informationsmedien stimuliert.

Problematisch ist eine Tendenz im Datenjournalismus, durch die zentrale Präsentation von Daten eine Deutungshoheit erlangen zu wollen.<sup>31</sup> In diesen Fällen wird durch die wissenschaftliche Anmutung der Infografiken Seriosität simuliert oder

<sup>30</sup> Vgl. Tartaglia (2014).

<sup>31</sup> Vgl. Schragger (2014).

Objektivität in der Berichterstattung signalisiert. Dazu gehört ebenfalls die bereits angesprochene fehlende Distanz und dementsprechend fehlende Kritikfähigkeit der Journalisten zu ihren Inhalten, wenn sie nicht über Forschungsergebnisse aus autonomen Quellen berichten, sondern die Erhebung zu einer gesellschaftlichen Fragestellung von der Zielsetzung bis zur Auswertung selbst erstellt haben.

## **Fazit und Ausblick**

In den Medienmärkten haben Digitalisierung und Vernetzung alte Geschäftsmodelle untergraben, zum Beispiel das des „Generalanzeigers“ für Zeitungen, die über einen großen Anzeigenteil einen unabhängigen Journalismus finanzieren. Die Abwanderung von Anzeigen ins Internet hat den Zeitungsverlagen nicht nur wirtschaftliche Einbußen beschert, sondern auch zu einem Bedeutungsverlust der Zeitung als alltägliche Informationsplattform geführt.

Das Phänomen „Big Data“ hat andererseits dazu geführt, dass heute Unternehmen wie Google, die in der Lage sind, Daten qualifiziert auszuwerten, extrem erfolgreich sein können. Diese überall vorhandenen Mengen von noch nicht ausgewerteten Daten werden auch als die neuen Rohstoffe bezeichnet, bei deren Raffinierung wertvolle Produkte geschaffen werden.<sup>32</sup>

Auf dieser Basis wird dem Datenjournalismus von manchen Protagonisten und Marktanalysten eine große wirtschaftliche Bedeutung für die strauchelnde Medienbranche prognostiziert. Erhofft wird, dass der Datenjournalismus nicht nur innovative Ideen für die journalistische Recherche und die Aufbereitung von Themen beisteuert, sondern auf ihm beruhende Angebote sich auch zu ertragreichen neuen Geschäftsmodellen für Medienunternehmen entwickeln lassen. Die Vision ist, dass Medienunternehmen mit datenjournalistischen Informationsservices einen Ersatz für den weggebrochenen Anzeigenmarkt entwickeln und den seriösen, immer schon teuren und tatsächlich sich noch nie selbsttragenden investigativen Journalismus damit weiterhin querfinanzieren können.

Schon der US-amerikanische Marketing-Experte Michael Palmer bezeichnete 2006 Datenjournalismus als „das neue Öl“. Darauf verweisend spricht der deutsche Datenjournalist Lorenz Matzat von den großen, zugänglichen Datenmengen als den „Rohstoffen des Internets“, die dafür neu zu organisierenden Medienunternehmen oder -sparten gute Gewinne sichern könnten (Matzat 2011).

## Quellen

Arthur, C. (2013): Tech giants may be huge, but nothing matches big data, The Guardian 23.8.2013. <http://www.theguardian.com/technology/2013/aug/23/tech-giants-data> [24.7.2014].

DJV (2009): Grundsatzprogramm des Deutschen Journalisten-Verbands, Bonn.

Garrison, B. (1995): Computer Assisted Reporting, Hillsdale, New Jersey.

Haas, H. (1999): Empirischer Journalismus. Verfahren zur Erkundung gesellschaftlicher Wirklichkeit Wien.

Hewett, J. (2013): Learning in Progress: From Computer-Assisted Reporting to Data Journalism, via Freedom of Information Open Data and More, in: Mair, J. et al. (Hrsg.): Data Journalism. Mapping the Future, Bury St Edmunds, S. 3-14.

Holovaty (2006): A fundamental way newspaper sites need to change. Blogpost vom 6.9.2006, <http://www.holovaty.com/writing/fundamental-change/>[3.3.2015].

Howard, A.B. (2014): The Art and Science of Data-Driven-Journalism. When Journalists combine New Technology with Narrative Skills, They can Deliver Context, Clarity, and a Better Understanding of the World Around Us. Tow Center for Digital Journalism at Columbia Journalism School, New York. A Tow/Knight Report May 2014, <http://towcenter.org/wp-content/uploads/2014/05/Tow-Center-Data-Driven-Journalism.pdf> [24.7.2014].

Jarvis, J. (2011): The article as luxury or byproduct, 28.05.2011, <http://buzzmachine.com/2011/05/28/the-article-as-luxury-or-byproduct/> [17.10.2014]

Katchadourian, R. (2010): No Secrets. Julian Assange's mission for total transparency, in: New Yorker, 07.06.2010, S. 40-51.

Künzler, M. (2012) (Hrsg.): Der Institutionenbegriff in der Publizistik- und Kommunikationswissenschaft, Baden-Baden.

Lynch, L. (2012): "That's not leaking. It's pure editorial": Wikileaks, scientific journalism, and journalistic expertise, in: The Canadian Journal of Media Studies Special issue Fall 2012.

Mair, J. (Hrsg.) (2013): Data journalism. Mapping the future, Bury St Edmunds.

Matzat, L. (2011): Datenjournalismus, in: Dossier „Open Data“ der Bundeszentrale für politische Bildung, <http://www.bpb.de/gesellschaft/medien/ opendata/64069/ datenjournalismus?p=all> [25.7.2014].

Matzat, L (2011a): Neues von der Definition des Datenjournalismus, <http:// datenjournalist.de/neues-von-der-definition-des-datenjournalismus/> [25.07.2014].

McQuail, D. (1992): Media Performance. Mass Communication and the Public Interest, London.

Meyer, P. (1973): Precision Journalism, Blomington.

Meyer, P. (1991): The New Precision Journalism, Blomington, und unter <http://www.unc.edu/~pmeyer/book/> [17.10.2014].

Palmer M. (2006): Data is the new Oil. ANA Marketing Maestros Blogbeitrag vom 03.11.2006, [http://ana.blogs.com/maestros/2006/11/data\\_is\\_the\\_new.html](http://ana.blogs.com/maestros/2006/11/data_is_the_new.html) [24.7.2014].

Rühl, M. (1980): Journalismus und Gesellschaft. Bestandsaufnahme und Theorieentwurf. Kommunikationswissenschaftliche Bibliothek, 9, Mainz.

ROG, Reporter ohne Grenzen (2013): Pressefreiheit in Deutschland. Stand 30.1.2013, [https://www.reporter-ohne-grenzen.de/fileadmin/rte/docs/2013/130130\\_Nahaufnahme-Deutschland\\_layouted.pdf](https://www.reporter-ohne-grenzen.de/fileadmin/rte/docs/2013/130130_Nahaufnahme-Deutschland_layouted.pdf) [17.10.2014].

Cohen, S.; Hamilton, J.T.; Turner, F. (2011): “Computational Journalism: How Computer Scientists Can Empower Democracy’s Watchdogs”, in: Communications of the ACM, Vol. 54, No. 10 (October, 2011), 66-71.

Schrager, A. (2014): it’s not science: The problem with data journalism. 19.03.2014, <http://qz.com/189703/the-problem-with-data-journalism/> [17.10.2014].

Tartaglia, H. (2014): Was man aus den Google-Suchen über Deutschland lernen kann. 1. Juli 2014, <http://www.vice.com/de/read/was-man-aus-den-google-suchen-ueber-deutschland-lernen-kann> [17.10.2014].





Weaver, D.H.; McCombs, M.E. (1980): Journalism and Social Science: A New Relationship?. in: Public Opinion Quarterly 44 (4, Polls and the News Media: A Symposium), S. 477-494.

Weinacht, S.; Spiller, R. (2014): Datenjournalismus in Deutschland, in: Publizistik 59 (4), S. 411-433.

Weischenberg, S.; Rakers, J. (2001): Nachrichten-Journalismus. Anleitungen und Qualitäts-Standards für die Medienpraxis, Opladen.

Williams, F. (1989): Computer-assisted writing instruction in journalism and professional education, New York.

## Über die Autorin



**Indira Dupuis**, Seit Ende 2006 wissenschaftliche Mitarbeiterin am Institut für Medienwissenschaft an der Ruhr-Universität Bochum, im Sommersemester 2015 Vertretung des dortigen Lehrstuhls Mediensysteme im internationalen Vergleich. Aufbau einer Lehr- und Lernredaktion seit September 2009, in der Studierende Grundlagen des audiovisuellen Journalismus lernen und Sendungsformate entwickeln können. Vor ihrer akademischen Karriere war sie in den Medien tätig, unter anderem in den Fernsehnachrichten.

# Data Journalism

Indira Dupuis

