



RAHEL ESTERMANN

46

MIT DATEN SPRECHEN

50

PRAKTIKEN, EXPERTISEN
UND VISUALISIERUNGSMODI
IM DATENJOURNALISMUS

54

55

56

57

58

59

60

61

62

63

64

65



[transcript] DIGITALE SOZIOLOGIE

Rahel Estermann
Mit Daten sprechen

Digitale Soziologie | Band 2

Editorial

Die Digitalisierung der Gesellschaft eröffnet neue soziologische Forschungsfelder, beeinflusst aber auch klassische Phänomene wie Infrastrukturen, Vergemeinschaftungen und soziale Praktiken grundlegend. Die Reihe **Digitale Soziologie** bietet sowohl theoretisch-konzeptionellen als auch empirischen Untersuchungen hierzu eine gemeinsame editorische Plattform. Insbesondere versammelt sie Arbeiten, die neue, auch unorthodoxe Zugriffe auf substanzelle Fragestellungen der Soziologie anbieten, die aktuelle methodische und konzeptionelle Entwicklungen aufgreifen oder innovative Methodenkombinationen nutzen. Dabei legt sie besondere Wert auf die Diskussion der Rolle der Soziologie angesichts der digitalisierten Gesellschaft und auf die Reflexion der daraus folgenden Implikationen bspw. für Infrastrukturen und die wissenschaftliche Praxis. Die Reihe schließt deutsch- und englischsprachige Monografien und Sammelände ebenso ein wie herausragende Qualifikationsarbeiten.

Die Reihe wird herausgegeben von Roger Häußling, Katharina Kinder-Kurlanda, Sophie Mützel, Jan-Hendrik Passoth und Andreas Schmitz.

Bei Interesse an der Publikation in dieser Reihe können die Herausgeber*innen kontaktiert werden.

Rahel Estermann, geb. 1987, ist im Luzerner Kantonsrat, für die Grünen Schweiz und in der Zivilgesellschaft politisch aktiv. Die Beziehung von Digitalisierung und Gesellschaft ist für sie dabei genauso relevant wie in ihrer soziologischen Forschung zu Datenjournalismus und Datenvisualisierungen.

Rahel Estermann

Mit Daten sprechen

Praktiken, Expertisen und Visualisierungsmodi
im Datenjournalismus

[transcript]

Die vorliegende Arbeit wurde 2021 vom Soziologischen Seminar der Kultur- und Sozialwissenschaftlichen Fakultät der Universität Luzern als Dissertation angenommen.

Gutachter*innen: Prof. Dr. Sophie Mütsel (Universität Luzern) und Prof. Dr. Jan-Hendrik Passoth (Europa-Universität Viadrina)

Die Open-Access-Ausgabe wird publiziert mit Unterstützung des Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnetet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.



Dieses Werk ist lizenziert unter der Creative Commons Attribution-ShareAlike 4.0 Lizenz (BY-SA). Diese Lizenz erlaubt unter Voraussetzung der Namensnennung des Urhebers die Bearbeitung, Vervielfältigung und Verbreitung des Materials in jedem Format oder Medium für beliebige Zwecke, auch kommerziell, sofern der neu entstandene Text unter derselben Lizenz wie das Original verbreitet wird.

Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z.B. Schaubilder, Abbildungen, Fotos und Textauszüge erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

Erschienen 2023 im transcript Verlag, Bielefeld

© Rahel Estermann

Umschlaggestaltung: Maria Arndt, Bielefeld

Umschlagabbildung: NZZ / mjk., awi.

Korrektorat: Anette Nagel, contexta

Druck: Majuskel Medienproduktion GmbH, Wetzlar

<https://doi.org/10.14361/9783839465905>

Print-ISBN 978-3-8376-6590-1

PDF-ISBN 978-3-8394-6590-5

Buchreihen-ISSN: 2751-3149

Buchreihen-eISSN: 2751-3157

Gedruckt auf alterungsbeständigem Papier mit chlorfrei gebleichtem Zellstoff.

Besuchen Sie uns im Internet: <https://www.transcript-verlag.de>

Unsere aktuelle Vorschau finden Sie unter www.transcript-verlag.de/vorschau-download

*Für Erwin,
dessen Lebensbaum als Sprössling schon zu akzeptieren lernte, dass manche Jahrringe
breit und andere dünn ausfallen, und dass das eigene Blühen und Früchtetragen keine
Selbstverständlichkeiten sind;
den wahre Grösse auszeichnete, indem er mich vertrauensvoll und frei blühen liess,
hoch hinaus und in fremden Farben, mir gönnte was ihm vergönnt war, stets auf meine
starken Wurzeln vertrauend, stets mit stilem Stolz.*

Inhalt

Dank	11
Abbildungsverzeichnis	13
1 Einleitung	17
1.1 Beitrag der Arbeit	20
1.2 Beautiful Data, Big Data	23
1.3 Abgrenzungen	25
1.4 Aufbau der Arbeit	26
2 Forschungsdesign	31
2.1 Erkundung des Feldes	31
2.2 Die Vorgehensweise anleitende Konzepte	34
2.2.1 Der abduktive Ansatz	34
2.2.2 Praktiken und Relationen: <i>zooming in and out</i>	36
2.2.3 Relationale Ethnografie	37
2.3 Teamarbeit untersuchen	39
2.3.1 Datenerhebung	40
2.3.2 Datenauswertung	47
2.4 Das Feld analysieren	49
2.4.1 Datenerhebung	49
2.4.2 Datenauswertung	53
2.5 Überblick	54
3 Datenjournalismus: theoretische und empirische Einordnung	57
3.1 Die datenjournalistische Community	58
3.1.1 Hacks und Hackers: »bridge the two worlds«	58
3.1.2 Die ›Ankommenshalbstunde‹: informell verbinden und kooperieren	60

3.1.3	Bedeutungsvolle Visualisierung: die Rolle eines Design-Studios	62
3.1.4	Flaggschiffe ennet des Atlantiks: über die Grenzen hinweg	63
3.1.5	Datenjournalismus als interpretative Community	65
3.2	Datenjournalismus als versammelte Expertisen	68
3.2.1	Kulturelle Rahmung: Ethos, Mindset und <i>Open</i> -Kultur	68
3.2.2	Werken mit Datensätzen, geometrischen Formen und <i>Tools</i>	77
3.2.3	Quantitativ, qualitativ, iterativ, visualisiert – epistemologische (Ent-)Spannungen	85
3.2.4	Die Expertisen des datenjournalistischen Feldes	91
3.3	Zirkulierende Narrative	95
3.3.1	Narrative über Datenjournalismus	96
3.3.2	Das lokale Editieren von Narrativen	111
3.4	Fazit und aufgeworfene Fragen	113
4	Heuristiken der Aushandlung und Visualisierung	115
4.1	Die »trading zone«	115
4.2	Durch Visualisierungen Daten erkennen und Evidenz befestigen	117
4.2.1	Das Labor als Fabrik der Erkenntnis	118
4.2.2	Visualisierung als Erkenntnismaschinerie des Sehens	120
4.2.3	Visualisierung als Befestigung von Evidenz	127
4.3	Visualisierungen in der <i>trading zone</i> : Kombination der Heuristiken	133
5	Daten erkennen: Visualisierungen als Arbeitsoberfläche	137
5.1	Arbeitspraktiken des Erkennens von Daten	137
5.1.1	Datafizierung: Geschichten in Daten übersetzen	138
5.1.2	Saubere Daten: Granularität, Ordnung, Interoperabilität	140
5.1.3	Geschichten drehen	145
5.1.4	Auswertungen plausibilisieren	158
5.2	Erkenntnisproduktion als Aushandlung	171
5.2.1	Koordinationsmuster in den Arbeitspraktiken	173
5.2.2	<i>Traders I</i>	181
5.2.3	Fazit: koordinierte Expertisen für das Erkennen von Daten	183
5.3	Visuelles Vokabular	185
5.4	Tools	190
5.4.1	Formalisierte Visualisierungssysteme	191
5.4.2	Computational notebooks	195
5.4.3	Papier	196
5.4.4	Fazit	198

5.5 Ideen imitieren	199
5.6 Fazit	202
6 Evidenz befestigen: Visualisierungen als Montage	205
6.1 Arbeitspraktiken der Evidenzbefestigung	205
6.1.1 Hauptaussage hervorheben	206
6.1.2 ›Humanizing data‹: Personalisierung und Konkretisierung	216
6.1.3 Datenquellen und Methoden transparent machen	220
6.1.4 Ausdifferenzierte Expertisen ausweisen	223
6.2 Evidenzbefestigung als Aushandlung	227
6.2.1 Koordinationsmuster in den Arbeitspraktiken	227
6.2.2 <i>Traders II</i>	237
6.2.3 Fazit	238
6.3 Leitlinien und <i>Tools</i>	239
6.3.1 <i>Styleguide</i>	240
6.3.2 Q und Datawrapper	242
6.4 Ausweichen auf neuartige Formate	251
6.4.1 Der <i>Explainer</i> : Ereignisse verknüpfen	251
6.4.2 Der <i>Longread</i> : ein Format für Datenvisualisierungen	256
6.5 Fazit	258
7 Schlussbetrachtung: Mit den Daten sprechen	261
8 Literaturverzeichnis	273
9 Anhang	289

Dank

Ganz viele Menschen haben meine Forschung unterstützt – ganz direkt mit ihren Beiträgen, oder still und im Hintergrund.

Für meine Feldforschung war ich angewiesen auf die Zeit und Offenheit vieler Datenjournalist*innen. David Bauer und das Team von NZZ Storytelling gewährten mir einen profunden Einblick in ihre Arbeit und ihre Überlegungen, auch wenn der Redaktionsschluss drängte. Genauso das Team von BR Data, wo ich insbesondere auf die Unterstützung von Robert Schoeffel und Vérona Nierle zählen durfte. Viele weitere Datenjournalist*innen aus Europa und den USA nahmen sich Zeit für meine Fragen, gewährten mir einen Einblick in ihre Kurse und Treffen oder liessen ihre Kontakte für mich spielen. Ich danke der Datenjournalismus-Community für ihre Kooperationsbereitschaft, ihre grosse Offenheit und die interessanten und auch unterhaltsamen Stunden.

Das Projekt »Facing Big Data« machte meine Dissertation erst möglich. Sophie Mützel förderte mich durch die Anstellung im Forschungsprojekt und begleitete mich als Erstbetreuerin durch meine Promotionszeit – anregend und stets hilfsbereit, aber auch viel Freiheit gewährend. In gemeinsamen Stunden im Büro, an Workshops oder bei Kaffee und Bier unterstützten mich die weiteren Mitglieder des Lehrstuhls, insbesondere Lisa Kressin und Philippe Saner, mit ihren wissenschaftlichen und persönlichen Anregungen oder einem offenen Ohr. Viele weitere Angehörige des Soziologischen Seminars und der Universität Luzern insgesamt leisteten materielle, intellektuelle und moralische Unterstützung, speziell möchte ich Christina Cavedon erwähnen. Jan-Hendrik Passoth danke ich für sein Engagement als Zweitbetreuer, Reto Estermann für das Lektorieren meiner Arbeit. Der Dank für die Finanzierung meiner Dissertation geht an den Schweizerischen Nationalfonds mit dem Nationalen Forschungsprogramm 75.

Tagtäglich begleitete mich Oliver Odermatt durch die Achterbahn eines Dissertationsprojekts, mittragend als ruhender Pol und mit liebevollem Ver-

ständnis. Nadia Nicolet lektorierte nicht nur meine englischsprachigen Texte, sondern erörterte mit mir Freud und Leid im Gefühlshaushalt. Viele weitere Freund*innen haben mich moralisch unterstützt. Das Traselinger Nest liess mich stets Geborgenheit spüren.

Ich danke ganz besonders meinen Eltern Ruth und Erwin Estermann-Schmid für ihr grosses Vertrauen: Sie ermutigten mich stets, meine Wege in ihnen unbekanntes Gelände zu beschreiten. Ohne das Rampenlicht zu suchen, liessen sie mich trotzdem ihre Wertschätzung und Unterstützung spüren.

Abbildungsverzeichnis

Abbildung 1: Verbindung von ›W-Fragen‹, dem ›klassischen‹ Mittel der journalistischen Erkenntnisproduktion, mit grundlegenden geometrischen Abbildungsformen. Folie aus einem datenjournalistischen Weiterbildungskurs. (MAZ_Ausbildung_2018_Folie104) S. 79

Abbildung 2: Die drei Expertisen und ihre Materialität, ihre epistemologische Kultur und ihre Professionskultur. (Eigene Darstellung) S. 94

Abbildung 3: Visualisierung der Anzahl an Teilnehmenden an der NICAR, der bekanntesten Datenjournalismus-Konferenz, 1999 bis 2018 (2001 fand keine Konferenz statt). Die Zahlen der ersten Austragungen (1993, 1994, 1995, 1997, 1998) sowie von 2019 und 2020 bleiben unbekannt. <https://twitter.com/christinezhang/status/971765540439064578/photo/1> (09.01.2021). S. 98

Abbildung 4: John Snows Cholera-Karte von London, abgebildet im Data Journalism Handbook. (Gray et al. 2012:202) S. 103

Abbildung 5: Ausschnitt aus der Untersuchung von Amann und Knorr-Cetina 1988:161: Ein Autoradiogramm als Arbeitsoberfläche. S. 121

Abbildung 6: Datenvisualisierungen als Arbeitsoberfläche strukturieren den Aushandlungsraum des Erkennens von Daten. (Eigene Darstellung) S. 135

Abbildung 7: Datenvisualisierungen als Montage strukturieren den Aushandlungsraum der Befestigung von Evidenz. (Eigene Darstellung) S. 136

Abbildung 8: Choroplethenkarte des Meinungsbeitrags »Europas verdrängte Revolution wirkt nach«. NZZ vom 2. August 2018, Seite 10. S. 156

Abbildung 9: Visualisierung des Verständlichkeitsniveaus von Datenschutzerklärungen aus dem BR-Data-Artikel »Komplizierte Datenschutzerklärungen: Der Haken mit dem Häkchen«. <https://web.br.de/interaktiv/datenschutzerklärungen/> (25.02.2020). S. 169

Abbildung 10: Koordinierte Expertisen im Modus des Erkennens von Daten. (Eigene Darstellung) S. 184

Abbildung 11: Aufstellung des »Visual Vocabulary« der Financial Times in den Redaktionsräumen von NZZ Storytelling. S. 187

Abbildung 12: Grafik aus dem Artikel »Die meisten Frauen nehmen noch immer den Namen ihres Mannes an«. <https://www.nzz.ch/gesellschaft/im-namen-m-eines-mannes-ld.1371279> (17.02.2020). S. 190

Abbildung 13: Von NZZ Storytelling erstelltes Schema der informatik-technischen Infrastruktur des Teams: Zusammenspiel von *Tools*, Dateiformaten und Arbeitsprozessen in der Produktion von Visualisierungen. S. 193

Abbildung 14: Screenshot der Visualisierungsoptionen im Menü der NZZ-Software Sue. S. 194

Abbildung 15: Pinnwand in den Redaktionsräumen von BR Data. Bei den meisten der aufgehängten Visualisierungen handelt es sich um die Datenvisualisungsentwürfe, welche den finalen Versionen in den Online-Artikeln sehr nahe kommen. S. 208

Abbildung 16: Pinnwand in den Redaktionsräumen von NZZ Storytelling, an der eine Auswahl von Visualisierungen aus der Printausgabe hängt. S. 209

Abbildung 17: Visualisierung von Fussballtransfer-Daten sowie begleitende Illustrationen von Einzelfällen. NZZ vom 31. März 2018, Seite 54. S. 212

Abbildung 18: Datenvisualisierung aus der Analyse von Datenschutzerklärungen durch BR Data. <https://web.br.de/interaktiv/datenschutzerklärungen/> (20.04.2020). S. 214

Abbildung 19: Ausschnitt aus dem Artikel »No City for Old Men«. NZZ vom 27. März 2018, Seite 21. S. 215

Abbildung 20: Foto des ausgedruckten Visualisierungsentwurfs zum Projekt Lauf-/Scoring-Analyse. Illustriert herausgehoben ist Lionel Messi; in den beiden Datenvisualisierungen plant David die Extremwerte (darunter Messi) sowie besonders prominente Fussballer farblich abzuheben. S. 218

Abbildung 21: Ausschnitt aus dem BR-Data-Projekt zu den Datenschutzerklärungen: Illustration anhand des Beispiels WhatsApp – der weisse Kasten bewegt sich durch Scrolling nach oben. <https://web.br.de/interaktiv/datenschutzerklaerungen/> (20.04.2020). S. 219

Abbildung 22: Koordinierte Expertisen im Modus der Evidenzbefestigung. (Eigene Darstellung) S. 239

Abbildung 23: Aus textlicher eine visuelle Evidenz machen – Folie aus der Schulungspräsentation von NZZ Storytelling für die Toolbox Q. S. 244

Abbildung 24: Eingabemaske von Q. Titel und Quellenangabe sind Pflichtfelder. S. 245

Abbildung 25: Foto des Konzept-Boards eines NZZ-*Explainers* zum Thema Waldbrände. Die abgebildete Konzeptionierung zeigt, wie der *Explainer* versuchen soll, das weltweite Phänomen der Waldbrände zuerst generell und dann die jeweiligen Aspekte aus der Schweizer Perspektive zu beleuchten. (Foto: NZZ-Storytelling-Teammitglied, veröffentlicht auf Twitter) S. 253

Abbildung 26: Codes der ethnografischen Auswertung der Beobachtungen bei NZZ Storytelling, platziert im Prozessablauf von Daten zu Geschichten. S. 296

Abbildung 27: Codes der ethnografischen Auswertung der Beobachtungen bei BR Data, platziert im Prozessablauf von Daten zu Geschichten. S. 296

1 Einleitung

Etwas zurückhaltend, fast brav hält Charlie Smart den Award in den Händen. Erst als er sieht, dass die weiteren Ausgezeichneten neben ihm auf der Bühne ihren Preis enthusiastisch und mit in die Höhe gereckter Trophäe feiern, drückt sich seine Freude mit einem Lachen aus.¹

Smart befindet sich an jenem Abend des 14. Juni 2019 für seinen damaligen Arbeitgeber, The Pudding, auf der Bühne.² Im Rahmen des Jahrestreffens des Global Editors Network (GEN Summit) vereinigen sich die Chefredakteur*innen wichtiger Medienorganisationen und verleihen (unter anderem) die jährlichen Data Journalism Awards, die wichtigste Auszeichnung des Feldes. Sie werden in einem Dutzend Kategorien vergeben, unter anderem gehen die Preise 2019 an Reuters, die New York Times, La Nación aus Argentinien, den deutschen Tagesspiegel, Vox, AP, Quartz und die bekannte Datenjournalistin und Aktivistin Eva Belmonte. The Pudding gewinnt die Kategorie »Data Journalism website of the year«³ und setzt sich dabei unter anderem gegen Reuters und das Wall Street Journal durch.⁴ Die Jury begründet ihre Entscheidung für The Pudding folgendermassen:

-
- 1 Video-Kurz-Zusammenfassung der GEN Data Journalism Awards 2019: <https://www.youtube.com/watch?v=yjykUEOvoE8> (30.01.2021), Sekunden 0:50 bis 0:56. Smart ist als zweiter von rechts zu sehen.
 - 2 Seit Ende 2019 arbeitet Smart für die New York Times: <https://www.nytimes.com/by/charlie-smart> (30.01.2021)
 - 3 Die Kategorie wird folgendermassen beschrieben: »Best journalism website or section of a website that is focused on data journalism – the collection, presentation, analysis of data and the creation of content around that data and its presentation. [...]« <https://datajournalismawards.org/category/> (14.11.2019, inzwischen nur noch über das Web Archive verfügbar)
 - 4 <https://datajournalismawards.org/2019-shortlist/> (14.11.2019, inzwischen nur noch über das Web Archive verfügbar)

»The site is a remarkable assortment of data visualisations, typified by extraordinary data collection (as in ›Women's pocket sizes‹), sophisticated data processing (as in ›Vogue covers‹) and a design that is both sumptuous (like ›Population mountains‹) and straightforward (like ›American anxieties‹). Taken together, The Pudding sets the high-water mark of the medium.«⁵

Es ist ein in mehrfacher Hinsicht denkwürdiger Abend in Athen. Einerseits ist es, wie sich ein halbes Jahr später herausstellt, der letzte GEN Summit und die achte bleibt somit die letzte Verleihung der Data Journalism Awards: Das GEN hat Finanzierungsprobleme und geht ein.⁶ Andererseits: Charlie Smart mit Award auf der Bühne. Auf den ersten Blick erscheint es nicht erstaunlich, dass The Pudding einen Award holt – 2019 ist The Pudding in der gesamten datenjournalistischen Community bekannt für seine ausgeklügelten Arbeiten. Überraschend ist jedoch, dass es Charlie Smart ist, welcher aus den USA nach Europa an den GEN Summit fliegt. Erst einige Wochen vorher – und befristet auf ein halbes Jahr – hat Smart bei The Pudding zu arbeiten begonnen.⁷ An den in der Jury-Begründung erwähnten Projekten war er nicht beteiligt. Die prägenden Köpfe des Pudding-Teams bleiben zuhause. Stattdessen Smart, der an den Verdiensten wenig beteiligte Neuling und Kurzzeit-Mitarbeiter, inmitten des ›Who is Who‹ des Datenjournalismus.

Auf den zweiten Blick jedoch ist auch schon die Tatsache bemerkenswert, dass The Pudding eine datenjournalistische Auszeichnung erhält. Erst zwei Jahre zuvor, Anfang 2017, hatten einige Mitarbeitende der US-amerikanischen Design-Agentur Polygraph The Pudding ins Leben gerufen. Während sie mit den bezahlten Aufträgen von Polygraph ihr Auskommen finanzieren, sollte The Pudding eine Spielwiese sein, wo sie ihren eigenen Interessen nachgehen, ihre »passion project[s]«⁸ verwirklichen und frei veröffentlichen können. Mit ihren »visual essays«, mit welchen »The Pudding explains ideas debated in culture«⁹, hat sich das Team innerhalb kurzer Zeit zu einer der meistbeachteten Publikationsplattformen der Community gemausert.

5 <https://datajournalismawards.org/2019-winners/> (14.11.2019, inzwischen nur noch über das Web Archive verfügbar)

6 <https://www.globaleditorsnetwork.org/about-us/farewell/> (14.11.2019, inzwischen nur noch über das Web Archive verfügbar)

7 <https://www.linkedin.com/in/charliesmart/> (30.01.2021) Es scheint, dass Smart sein Profil seit Ende 2019 nicht mehr aktualisiert hat.

8 Interview D17

9 <https://pudding.cool/about/#team> (30.01.2021)

Und nun, an jenem 14. Juni 2019 – vor den Augen der globalen Medienelite, welche nach neuem Vertrauen und Geschäftsmodellen in der digitalen Welt sucht –, erhält das Team die Auszeichnung mit der Begründung, die Exzellenz in Sachen Datensammlung, -verarbeitung und -visualisierung zu verkörpern. Ein Team, das im Kern aus nur acht Köpfen besteht. Und vor allem gängige journalistische Prozesse und Konventionen auf den Kopf stellt¹⁰: Es gibt keine Redaktionsräume, alle Teammitglieder arbeiten verstreut im englischen Sprachraum; es gibt keine*n Chefredaktor*in, keine formellen Hierarchien, sondern ein »collective«; die Löhne sind öffentlich; es gibt keinen Termindruck durch Publikationsfristen; auf der Webseite gibt es weder eine Bezahlsschranke noch Werbung – stattdessen finanziert sich das Team durch bezahlte Aufträge von Unternehmen an eine eng verflochtene Schwester-Firma¹¹; sämtliche Datensätze und Auswertungsmethoden der Projekte sind auf GitHub transparent und reproduzierbar, teilweise sogar in langen Youtube-Videos aufgezeichnet; das Team ist offen für Themenvorschläge und Zusammenarbeit mit Interessierten, welche eigene Ideen und Datensätze mitbringen; Geschichten im Ideenstadium finden sich in einem öffentlichen »Backlog« aufgeführt – in der Hoffnung, interessierte Mitstreiter*innen zu finden; Preisverleihungen verfügen über keinen grossen Stellenwert.

In der Episode rund um Charlie Smart und den Award für The Pudding kondensieren sich viele interessante Charakteristika eines Phänomens, das sich seit einem guten Jahrzehnt etabliert hat: Datenjournalismus. The Pudding steht nicht stellvertretend für das gesamte Feld oder für alle Datenjournalismus-Teams, die aufgezählten Eigenschaften treffen auf die meisten anderen Organisationen nur sporadisch oder überhaupt nicht zu. Trotzdem verdeutlicht sich: Datenjournalismus meint zwar auch neue Technologien; Datenjournalismus ist aber vor allem ein soziales und kulturelles Phänomen. Diese Aspekte zu erforschen, insbesondere ihre Interaktion mit neuen Werkzeugen, hat mich in rund dreieinhalb Jahren Forschung angetrieben.

-
- ¹⁰ Die Quellen für die folgenden Beschreibungen sind die Webseite von The Pudding (<https://padding.cool/about/>, 30.01.2021), ihr Github-Repositorium (<https://github.com/the-pudding>, 30.01.2021), ihr Youtube-Kanal (<https://www.youtube.com/channel/UCFu9vMFVluW9CAOdS4Oluw/videos>, 30.01.2021) sowie das Interview D17 mit einem Kern-Mitglied des Teams.
- ¹¹ Wie erwähnt arbeiten alle Teammitglieder auch für die Schwester-Agentur Polygraph, welche bezahlte Aufträge ausführt. Mehr zur Verflechtung von Polygraph und The Pudding: <https://padding.cool/faq/#polygraph> (30.01.2021)

1.1 Beitrag der Arbeit

Das initiale Anliegen des Forschungsprojektes bestand darin, das im Laufe der 2010er-Jahre entstandene organisatorische und professionelle Feld des Datenjournalismus zu untersuchen. Der Soziologe Schudson (2000) postuliert drei verschiedene Ansätze, um journalistische Felder und Praktiken zu untersuchen: erstens aus einer politisch-ökonomischen Perspektive; zweitens aus einer (organisations-)soziologischen Perspektive (Routinen, Redaktionsorganisation, Weltbilder der Journalist*innen); drittens aus einer (soziologisch-)kulturellen Perspektive. Die vorliegende Arbeit verbindet soziologische Perspektiven aus der Journalismusforschung und den Science and Technology Studies (STS) und untersucht damit Datenjournalismus in einer Kombination von Schudsons zweitem und drittem Ansatz.

Damit leistet die Dissertation einen Beitrag zu aktuellen Diskussionen in verschiedenen wissenschaftlichen Disziplinen. Sie geht auf soziologische Fragen nach der Datafizierung und Quantifizierung der Gesellschaft ein, indem sie das Zusammenspiel von Technologie und sozialen Strukturen untersucht. Für die STS ist die Diskussion der Wissensbildung mit digitalen Daten als neuartiger Materialität ein Anschluss an die Tradition der Laborstudien und die Überlegungen zur Rolle von Visualisierungen in der Produktion wissenschaftlichen Wissens. Diese Arbeit erweitert die Journalismusforschung um einen Beitrag zu sozialen Prozessen in den Redaktionen, welche die bestehende journalistische Kultur im Zusammenhang mit der Materialität digitaler Daten neu verhandeln: Datengetriebene journalistische Berichterstattung über die Welt entwickelt sich hin zu einem Projekt mit zahlreichen beteiligten Expertisen, einer anspruchsvollen Koordination ebenjener, organisiert durch das Adaptieren wissenschaftlich geprägter Konventionen wie Peer-Review und Reproduzierbarkeit. Die Dissertation bietet durch die Untersuchung der epistemologischen Hintergründe und der Überzeugungskraft von Bildlichkeit aber auch Anschlusspunkte an die generelleren Diskurse der Medien- und Kommunikationswissenschaften, welche die zunehmende Nutzung von Datenvisualisierungen diskutieren.

Diese Beiträge kondensieren im Schema eines datenjournalistischen Dreiecks der Expertisen, deren Aushandlung strukturiert ist durch bestimmte Modi von Datenvisualisierungen. Die vorliegende Arbeit hat mit dem Datenjournalismus ein im digitalen 21. Jahrhundert entstandenes Feld als Fall erforscht, deren Erkenntnisse aber für die genannten wissenschaftlichen Diskurse einen über den Fall hinausweisenden Beitrag leisten.

Die Arbeit untersucht insbesondere zwei Fragestellungen:

- a) Warum gibt es den Datenjournalismus?

Beim Vorhaben, das Feld zu definieren, zeigte sich schnell, wie vieldeutig der Begriff *Datenjournalismus* Verwendung findet – sowohl in der Theorie wie in der Empirie. Es stellt sich folgerichtig die Frage: Warum ist von *dem* Datenjournalismus die Rede, warum erscheint dies als relevante soziale Kategorie, obwohl damit eine Vielzahl an Phänomenen und Praktiken verbunden wird? Oder andersherum: Es gilt zu ergründen, was dieses Feld über den Begriff hinaus verbindet – sodass sich eine Community formiert, welche sozial eng verflochten ist und über geteilte epistemologische Vorgehensweisen und kulturelle Vorstellungen verfügt. Und sodass die Diskussion darüber und die Auszeichnung derer, welche *the high-water mark of the medium* setzen, erst möglich ist.

Bisher sind die Deutungsangebote aus der Forschungsliteratur dazu lückenhaft geblieben. Dies hat damit zu tun, dass die Kategorie *des* Datenjournalismus dabei meist vorausgesetzt und definiert, aber selten hinterfragt und im Zusammenspiel seiner Teilelemente untersucht wurde – auch wenn schon mehrfach Autor*innen darauf verwiesen, dass die verschiedenartigen epistemologischen Zugänge und professionellen Kulturen im Datenjournalismus einer ethnografischen Untersuchung bedürfen (siehe beispielsweise Parasie und Dagiral 2012). Die Dissertation nimmt diesen Faden auf und folgt einer ethnografischen Vorgehensweise, die sich über mehrere Ebenen des Sozialen aufspannt – und ihre Interpretationen darauf abstützt, die verschiedenartigen Expertisen der Akteur*innen – Programmierung, Visualisierungen, Journalismus – zu differenzieren und Datenjournalismus als Manifestation ihrer soziotechnischen Aushandlung zu betrachten.

- b) Welche Rolle spielen Datenvisualisierungen?

Die datenjournalistische Produktion von Erkenntnissen ist ein oft untersuchter Prozess in der Journalismusforschung. Ein Element blieb dabei aber bisher unterbeleuchtet: Datenvisualisierungen. Ihre Rolle als Darstellungsform ist zwar anerkannt und die vorliegende Arbeit wird diesen Aspekt vertiefen. Gleichzeitig wird sie aufzeigen, dass Datenvisualisierungen auch bereits den Produktionsprozess von Berichterstattung prägen. Denn Datenvisualisierungen spielen an der aus kultureller Sicht entscheidenden Schnittstelle eine Rolle: »[...], the cultural view finds symbolic determinants of news in the relations

between ‚facts‘ and symbols‘ (Schudson 2000:189). Die symbolische Form bestimmt, welche Art von Erkenntnissen der Datenjournalismus überhaupt über die Welt produzieren kann – und wann diese als evident anerkannt sind. Wenn The Pudding von *visual essays* als ihren Produkten spricht, so impliziert dies nicht blass ein Visualisierungen inkludierendes Präsentationsformat, sondern eine andersartige Produktion der Inhalte.

Für die Untersuchung der Rolle von Datenvisualisierungen kann dabei auf einen grossen Wissensbestand der STS zurückgegriffen werden. Obwohl in den Forschungskontexten der in dieser Arbeit referenzierten STS-Autor*innen Daten in der Regel nicht in digitaler Form vorlagen, bieten deren grundlegende theoretische Überlegungen eine gute Basis für die Einordnung auch von computergestützt produzierten Datenvisualisierungen. Denn die Frage nach dem Verhältnis von Daten, Visualisierungen und Evidenz ist die gleiche, und die STS-Forschung stellte sie in den Achtzigerjahren in den Naturwissenschaften: »But are sense data problematic? Consider that in the natural sciences *evidence appears to be embodied in visibility; in a literal sense, it is embodied in what we can see on a data display*. Thus understood, the notion of evidence is built upon the difference between what one can see and what one may think, or have heard, or believe« [Hervorhebung RE] (Amann und Knorr-Cetina 1988:134).

Es kann kaum überschätzt werden, wie weitreichend die hervorgehobene Überlegung ist: Sie postuliert, dass die Evidenz von Daten gebunden ist an eine Abbildung. Datenvisualisierungen nehmen also nicht nur *eine* Rolle ein in der Produktion von Erkenntnis und Evidenz aus Daten, sondern *die* zentrale. Diese Dissertation wird zeigen, dass Datenvisualisierungen die strukturierende Materialität der Sprache des Datenjournalismus sind. Die Sprachmetapher ist dabei mit einer doppelten Bedeutung versehen: Datenvisualisierungen dienen einerseits dazu, *mit* den Daten zu sprechen – in einem Sinne der Produktion von Erkenntnissen durch die Mathematisierung und Geometrisierung von Datensätzen, welche in der Überblendung mit einem visuellen Vokabular resultiert.¹² Andererseits ist es dank Visualisierungen möglich, *mittels*

¹² Die Sprachmetapher wird auch im Diskurs des datenjournalistischen Feldes oft verwendet, wenn von *conversations with data* die Rede ist. So nennt sich der Newsletter von DataJournalism.com, der gemeinsamen Datenjournalismus-Plattform des European Journalism Centre und der Google News Initiative, »conversations with data« (siehe <https://datajournalism.com/read/newsletters, 19.07.2020>). Ebenso wurde die Metapher genutzt, um Konferenzpanels über Datenjournalismus zu betiteln (siehe <https://www.journalismfestival.com/programme/2018/conversations-with-data, 19.07.2020>). Al-

(beziehungsweise anhand von) Daten zu sprechen, um eine Aussage zu treffen – dabei kommt diesen Aussagen aufgrund ihrer visuellen Montage eine besondere Überzeugungskraft als Evidenz zu.

1.2 Beautiful Data, Big Data

Die vorliegende Dissertation betrachtet die zwei Fragestellungen nicht getrennt – im Gegenteil zeigt sich, wie eng die Antworten auf die beiden zusammenhängen. Die Entstehung immer neuer Schnittstellen und Überlappungen durch gesellschaftliche Ausdifferenzierung und Querbezüge (Passoth und Rammert 2016) spiegelt sich im Datenjournalismus in der Verflechtung von Expertisen, unterschiedlichen Narrativen und ausdifferenzierten Übersetzungsskripten ebenjener. Dass diese Diversität trotzdem durch *einen* Begriff gefasst wird, erscheint auf den ersten Blick erstaunlich. Die Arbeit erschliesst die Erklärung dafür, gerade indem sie die zweite Fragestellung zu Datenvisualisierungen miteinbezieht. Denn die Verflechtung der Expertisen basiert auf Koordinationsprozessen zwischen ihnen, wofür das Konzept der *trading zone* (Galison 1997) eine analytische Fassung bietet. Es verweist in letzter Konsequenz darauf: Erfolgt die Aushandlung nicht nur zwischen zwei, sondern drei Expertisen, so kann eine neue Kultur mit einer eigenen Sprache entstehen und sich festigen. Diese Arbeit postuliert, dass es an diesem Punkt Datenvisualisierungen sind, welche als Materialität institutionalisierend und strukturierend wirken.

Die Rolle von Datenvisualisierungen als Materialitäten in gesellschaftlichen Prozessen mitzubetrachten erscheint nicht erst, aber gerade im Zeitalter der »Data Revolution« (Kitchin 2014b) angezeigt. Die Referenz auf die visuelle Dimension von Daten ist in den letzten Jahren zweideutig erfolgt. Einerseits in einem direkten, praktischen Sinne die Abbildung von Daten meinend, die als zugehörige Dimension der zunehmenden Quantifizierung aufscheinen (Espeland und Stevens 2008). Andererseits erfolgt die Evaluation des Werts von Daten mittels Verweis auf deren Ästhetik (Mützel et al. 2018). Halpern erklärt:

Ierding verwendet das Feld die Metapher, ohne spezifisch auf die Rolle von Datenvisualisierungen hinzuweisen.

»These narratives come with labels such as ›Beautiful Data‹ and ›Beautiful Evidence‹. Opening with the premise that the web today is above all about the collection of personal data, many data visualization sites and textbooks urge the designers, engineers, and programmers of our future to address the important aesthetic component of making this data useful, which is to say, ›beautiful‹. But data is not always beautiful. It must be crafted and mined to make it valuable and beautiful.« (Halpern 2014:5)

Auch wenn in dieser Arbeit eher die praktische und weniger die metaphorische Ebene im Zentrum steht, verweist die zweitere darauf: In der ästhetischen Komponente liegt eine grosse Macht, welche auf das Publikum wirkt, indem sie Evidenz erzeugt und die Arbeit zu deren Erzeugung gleichzeitig verbirgt. Diese Überlegung wird sich auch auf einer praktischen Ebene reflektieren, wie diese Arbeit später zeigt.

Über die beiden ausgelegten Forschungsfragen hinaus leistet die Arbeit einen Betrag zum Verständnis, wie die digitalen Technologien des 21. Jahrhunderts mit gesellschaftlichen Prozessen, insbesondere denjenigen der Erkenntnisproduktion, interagieren. Sie hat dabei den Anspruch, aufzuzeigen, dass die Werkzeuge einhergehen mit neuen sozialen und kulturellen Verflechtungen. Dies zeigt sich im Phänomen des Datenjournalismus, welcher als Teil des Journalismus entstand und diesen gleichzeitig erweiterte.¹³

»I think it's generally good?«, lautet die Antwort, als ich ein Teammitglied von The Pudding frage, weshalb sie alles Mögliche teilen und transparent machen.¹⁴ Das Verpacken der Antwort in eine rhetorische Frage deutet an, dass diese kulturellen Vorstellungen über ›guten Datenjournalismus‹ keine schon seit langem institutionalisierten Handlungsanleitungen, sondern umstrittene Aushandlungsprozesse jüngeren Datums sind. Sie gehen einher mit der Frage, wie die gesellschaftliche Selbstbeobachtung sich in einer zunehmend digitalisierten Welt weiterentwickelt – indem sie auch im Journalismus neue Akteur*innen mit ihren Expertisen versammelt und sich Sinngebungs- und Kommunikationsmuster verändern.

13 Auch wenn das Verhältnis von Datenjournalismus und Journalismus, Ersteres verstanden als Teil des Zweiteren, sowohl in diesen einleitenden Absätzen wie auch im Folgenden immer wieder mitschwingt, so ist dieses Verhältnis doch nicht das Analyseobjekt der Forschung (es wurde schon andernorts eingehend beleuchtet, siehe beispielsweise Hermida und Young 2019).

14 Interview D17

»Journalism is the socially valued and structured work of crafting and distributing socially significant news and discussion«, definiert Vos (2018:4) Journalismus, rudimentär und gekürzt. Als Berichterstatter – und somit Produzent von Erkenntnissen – über die Geschehnisse in der Welt ist der Journalismus von den Digitalisierungsprozessen der letzten Jahrzehnte geprägt: Digitale Daten und computergestützte Auswertungsmethoden, gemeinhin mit dem Schlagwort Big Data versehen, bergen das Potenzial, die journalistischen Erkenntnisprozesse zu verändern. Lewis und Westlund (2015:453) verweisen auf die beiden von Vos erwähnten Aspekte der journalistischen Tätigkeit – »crafting news« und »distributing news« – und folgern mit Bezug auf Big Data: »Big data, as a set of technological processes as well as a key source of power, thus opens new paths for imagining how journalistic investigations develop epistemologically relevant revelations. More broadly, big data has implications for rethinking the epistemologies of news production and news distribution.« Die Dissertation greift durch ihre Forschungsfokusse folgerichtig auch die Fragestellung nach dem epistemologischen Potenzial von Big Data auf: Datenvisualisierungen nehmen genau in den beiden genannten Bereichen – »news production and news distribution« – eine signifikante Rolle ein, welche diese Dissertation untersucht und einordnet. Dabei ermöglicht es der Fokus auf das Phänomen des Datenjournalismus, aufzuzeigen, wie ein Feld von sozialen und kulturellen Strukturen – wie der Community und dem visuellen Vokabular – erst stabilisiert und neue Standards – wie beispielsweise das Teilen oder Peer-Review-Prozesse – institutionalisiert werden. Und diese dann erklären, wieso ein Team wie The Pudding die Erfolgsgeschichten eines Feldes erzählt, das nach ›gutem Datenjournalismus‹ strebt.

1.3 Abgrenzungen

Visualisierungen im Journalismus können sehr viele verschiedene Formate annehmen. Die Überlegungen dieser Arbeit beziehen sich auf »wissenschaftlich-technische Bilder« (Burri 2008:350), das heisst die (interaktive oder nicht-interaktive) Visualisierung von Daten, gespeichert in Datenbanken, durch kalkulierende Instrumente – seien dies Programmiersprachen oder spezifische Visualisierungssoftwares.

Die Dissertation hat nicht weiter die Rolle von anderen Formen von Visualisierungen im Journalismus untersucht – zum Beispiel Infografiken, also gezeichnete visuelle Elemente, welche bereits seit Jahrzehnten im Journalis-

mus verbreitet sind, aber keine Datenbanken als Grundlage und Kalkulationen in der Herstellung benötigen. Weitere visuelle Formate, die ich ausgeklammert habe, sind Fotos, Illustrationen oder Karikaturen. Auch der genuin visuelle Journalismus des Fernsehens und der Einsatz von Videos gehören nicht zum Untersuchungsfeld. Auch wenn diese Formate ab und zu auch im Rahmen von datenjournalistischen Praktiken zum Einsatz kommen – immerhin gehören auch einige der bekanntesten datenjournalistischen Teams zu öffentlich-rechtlichen (Fernseh-)Anstalten –, so sind doch ihre technischen und epistemologischen Dimensionen sehr verschieden (Wintsch 2006). Dies spiegelt sich auch darin, dass Videos im Rahmen datenjournalistischer Projekte die digitalen Daten meist doch wieder in Formen von Diagrammen oder Karten integrieren.¹⁵

1.4 Aufbau der Arbeit

Im Aufbau der Arbeit widerspiegelt sich das induktive, felderkundende und theoriegenerierende Vorgehen dieser Dissertation. Das Phänomen des Datenjournalismus erforschte ich theoretisch imprägniert, aber stark von der Empirie geprägt. Deshalb folgt gleich nach dieser Einleitung das Forschungsdesign (Kapitel 2), um den iterativen Prozess auszulegen, der die Arbeit rahmt. Das Design zeigt auf, dass ich die Forschung auf verschiedenen Ebenen und mit verschiedenen Fokusseiten gleichzeitig betrieb. Dabei leiteten mich drei theoretisch-methodische Konzepte an: der abduktive Ansatz, das Changieren der Fokusse zwischen Praktiken und Relationen (*zooming in and out*) sowie die relationale Ethnografie als methodisches Konzept, um diese beiden Ebenen zu integrieren. Kapitel 2 erläutert detailliert mein Vorgehen sowie die Quellen, aus welchen ich meine empirischen Daten gewann. Dabei rückt es die ethnografische Untersuchung von zwei datenjournalistischen Teams – NZZ Storytelling und BR Data – ins Zentrum.

Kapitel 3 nimmt eine Einordnung des datenjournalistischen Feldes anhand dreier soziologischer Konzepte vor: Community, Expertisen und Narrative. Das Skizzieren der lokalen Community (3.1) verweist auf eine Gruppierung, welche sich unter dem Dach von Hacks/Hackers regelmässig trifft und dabei

¹⁵ Beispielhaft dieses Video des BR, welches auf der Datenauswertung des Datenjournalismus-Teams aufbaut: <https://www.br.de/mediathek/video/das-verbaute-land-alternativen-zum-neubaugebiet-av:5abc23cb7659070018e784b0> (25.11.2020)

enge informelle Beziehungen in lockerer Atmosphäre aufbaut. Dabei spielen Visualisierungsaspekte, insbesondere ein Design-Studio, eine grosse Rolle – genauso wie die internationale Vernetzung und Orientierung, häufig über Twitter, und die omnipräsenten Flaggschiff-Medienorganisationen aus dem englischen Sprachraum. Der Abschnitt beschreibt, wie diese *interpretative community* Datenjournalismus trägt, indem sie die Sinngebung teilt und über gemeinsam ausgehandelte Bewertungsstandards verfügt. Und dass die Community dafür gemeinsame Kanäle und Plattformen des Austauschs rege pflegt.

Dann zeige ich beschreibend die Diversität und Vielschichtigkeit des Feldes auf (3.2): Im Datenjournalismus versammeln sich drei Expertisen – Visualisierung, Journalismus, Programmierung – rund um drei verschiedene Dimensionen. Erstens die professionelle Kultur: ein Ethos des gemeinsamen Lernens, den die Datenjournalist*innen teilen; eine Ausbildung, die zu einem *Mindset* befähigen soll, um in der Community mitsprechen zu können; und der Spirit der *Open*-Kultur mit Werten wie Partizipation, Bastelei, Teilen und Transparenz. Zweitens das Werken mit verschiedenen Materialitäten: mit strukturierten und interoperablen Datensätzen; mit einem visuellen Vokabular aus einer Palette von geometrischen Formen, von einfachen Diagrammen bis hin zu *Data Art*; mit computergestützten *Tools*, welche die Auswertungen möglichst reproduzierbar und transparent machen. Und drittens Spannungen, welche sich zwischen den epistemologischen Grundsätzen des konventionellen Journalismus und den quantitativen Verfahren des Datenjournalismus ergeben – und die mittels iterativer Vorgehensweisen (*data-driven* vs. *story-driven*) sowie durch Datenvisualisierungen (teilweise) entspannt werden. Diese drei Dimensionen, eine Triade, nuancieren das Expertisen-Dreieck von Journalismus, Visualisierung und Programmierung. Das Dreieck der Triaden ordnet die empirischen Beobachtungen meiner Forschung, insbesondere der Teams, ein.

Danach zeichne ich die Narrative des datenjournalistischen Feldes nach (3.3). Ihre Unterschiedlichkeit ist sehr auffällig, sie schliessen sich aber weder auf der Feld- noch auf der Teamebene aus: Es kann beiderorts sehr verschiedenenartige und gleichzeitig zirkulierende Narrative und Erfolgsgeschichten geben. Das Konzept des Narratifs baut dabei auf den theoretischen Überlegungen aus der Organisationssoziologie auf (Sahlin-Andersson 1996), welche darauf hinweisen, dass auf lokaler Ebene die Narrative editiert (und nicht kopiert) werden. Das heisst, die strukturellen und institutionellen Gegebenheiten von Datenjournalismus-Teams prägen, welche und wie sie die Narrative in eige-

ne datenjournalistische Praktiken umsetzen. Zum Schluss des dritten Kapitels (3.4) rolle ich aus der Forschung auf Feld-Ebene relevante Fragen für die beiden zentralen und ausführlichen Kapitel mit den Team-Ethnografien aus – es sind die beiden bereits hier in der Einleitung erwähnten Fragestellungen nach der Persistenz der vielfältigen sozialen Kategorie Datenjournalismus sowie der Rolle von Datenvisualisierungen.

Das Kapitel 4 greift die aufgeworfenen Fragen auf und entwickelt sie theoretisch weiter – mit dem Ziel, Gerüste zu entwickeln, mit welchen die Team-Ethnografien die Fragen zu beantworten vermögen. Dabei sind theoretische Überlegungen aus den STS, der Wissenschaftsforschung und ihrer Laborstudien, der Anthropologie und den Medienwissenschaften zentral: Die Heuristiken der *trading zone*, von *seeing data* und *becoming evident* kombiniere ich schliesslich zu zwei Modi der Visualisierung, welche Aushandlungen von Expertisen strukturieren.

Die beiden Kapitel 5 und 6 bilden das Herzstück dieser Arbeit. Sie untersuchen in den zwei datenjournalistischen Teams die zwei Modi: die Produktion von Erkenntnis auf Datenvisualisierungen als Arbeitsoberfläche (5) respektive die Befestigung von Evidenz auf Datenvisualisierungen als Montage (6). Dabei lassen sich die Aushandlungen von Expertisen in Diskussionen, aber auch Werkzeugen und Formaten beobachten. Je nach Visualisierungsmodus stehen dabei andere Konstellationen und Verschränkungen im Vordergrund. Der ethnografische Ansatz ermöglicht mir, mit einer reichen Beschreibung des datenjournalistischen Alltags die Strukturen der Praktiken und die subtilen Verbindungen zur Community, ihrer Kultur, ihren Werkzeugen und den Narrativen der Feld-Ebene zu erschliessen.

In Abschnitt 5.1 beschreibe ich zuerst vier zentrale Praktiken zur Produktion von Erkenntnis im Datenjournalismus: die Datafizierung, welche Geschichten in Datensätze übersetzt; die Herstellung sauberer Daten; das Drehen von Geschichten rund um die durch Visualisierungen produzierten Muster; das Plausibilisieren der Auswertungen durch Samples, Einzelfälle und Vergleiche, welche die qualitativen und quantitativen Welten passend machen. Die folgenden Abschnitte analysieren diese Praktiken dann als Aushandlungen zwischen den Expertisen – zuerst in Diskussionen (5.2), in welchen die *traders* eine koordinierende Rolle spielen; dann in Relationierungen durch das visuelle Vokabular (5.3). Abschnitt 5.4 beleuchtet dann, wie sich diese Koordinationsprozesse in Werkzeugen reflektieren: in formalisierten Visualisierungssystemen wie beispielsweise der Programmiersprache R, in *computational notebooks*, aber auch auf Papier. Schliesslich weist 5.5 darauf

hin, wie das Imitieren von Ideen aus der Community – Erfolgsgeschichten, Projekte – direkt datenjournalistische Praktiken in den Teams anleitet.

Kapitel 6 folgt dann demselben Aufbau, bezieht ihn aber auf den Modus der Evidenzbefestigung. Abschnitt 6.1 beschreibt die Praktiken der Her- vorhebung der Hauptaussage durch Leseführung sowie den Vergleich; die Praktik des *humanizing data* durch Personalisierung und Konkretisierung; Transparenzpraktiken rund um Datenquellen sowie Methoden und Visuali- sierungen; das differenzierte Ausweisen von Expertisen in den Autor*innen- Angaben. Wie im ersten Modus koordinieren die Teams diese Praktiken in Aushandlungen durch Diskussionen, wieder vermögen die *traders* eine wichtige Stellung einzunehmen (siehe Abschnitt 6.2). Die Aushandlungen spiegeln sich in Leitlinien und Werkzeugen (Abschnitt 6.3), welche die Teams nutzen – und institutionalisieren sich insbesondere in datenjournalistischer Visualisierungssoftware wie Q und Datawrapper, welche sich an ein Publikum ausserhalb des eigenen Feldes richtet. Das Ausweichen auf neuartige Formate (Abschnitt 6.4) durch die beiden Teams weist dann nochmals auf die ma- terielle, epistemologische und kulturelle Abgrenzung datenjournalistischer Praktiken in den Redaktionen hin.

Die Schlussbetrachtung in Kapitel 7 webt zuerst den roten Faden des For- schungsvorhabens nochmals durch die Kapitel hindurch, um dann darauf auf- bauend Datenjournalismus als Kreol-Sprache zu umreissen. Sie rahmt, wie Datenjournalist*innen anhand der zwei Visualisierungsmodi *mit* den und *mit- tels* der Daten sprechen. Zum Schluss ordne ich die Erkenntnisse dieser Arbeit in einen weiteren Forschungskontext ein und weise auf offene Fragen hin, welche als Desiderat meiner Dissertation verbleiben.

2 Forschungsdesign

Das folgende Kapitel erläutert das Forschungsdesign meiner Arbeit. Es schliesst dabei an die in der Einleitung aufgeworfene Forschungsfrage über das Phänomen Datenjournalismus an. Es transformiert sie in ein definiertes Forschungsfeld und Methoden als Instrumente zur Erkenntnisgewinnung über das Phänomen.

Anschliessend an die offenen Fragen bestand meine erste methodische Entscheidung darin, nicht selbst zu definieren, was *der* Datenjournalismus ist – sondern den Fokus zuerst so weit zu öffnen, dass alles in den Blick rückt, was mit dem Begriff in Verbindung gebracht wird. Mit dieser Ausgangslage begann ich die Erkundung des Feldes, die ich im ersten Abschnitt 2.1 erläutere. Der folgende Abschnitt (2.2) beschreibt drei methodologische Konzepte, welche meine Vorgehensweise angeleitet und dadurch meine Forschungsperspektive mitgeprägt haben. Die Abschnitte 2.3 und 2.4 beschreiben dann die Methoden der Erhebung und Auswertung meiner Forschungsdaten – bezüglich der Untersuchung der datenjournalistischen Praktiken in den Organisationen (Teams) sowie bezüglich der Relationen, welche diese mit den Dynamiken und Diskursen des Feldes verbinden.

2.1 Erkundung des Feldes

Um eine komplexe soziale Realität zu erforschen, muss ein*e Forscher*in ein (zu untersuchendes) Feld definieren. Diese Aufgabe ist nicht zu unterschätzen: Es gilt, die Abgrenzungen in mehrdimensionalen, durch viele Verbindungen durchsetzten »fuzzy fields« (Nadai und Maeder 2005) vorzunehmen. Ein ethnografischer Ansatz nimmt diese Herausforderung an, indem er den Prozess des »constructing the field« (Amit 2000) thematisiert und reflektiert (Amit 2000b).

Von besonderer Bedeutung ist es dabei für ein*e Forscher*in, vorerst offen zu bleiben für »different places and processes that are interconnected rather than subjecting herself to boundaries that are set beforehand« (Naess 2016:5) – es geht also darum, die Bezüge zwischen verschiedenen Orten genauso zu betrachten wie die Beziehungen an einem Ort selber, »the fields are not some mere collection of local units. One must establish the translocal linkages, and the interconnections between those and whatever local bundles of relationships which are also part of the study« (Hannerz 2003:206).

Es ist auch der Ansatz dieser Arbeit, das Feld von seinen Beziehungen her zu erschliessen und abzugrenzen, um es dann als ethnografisch zu untersuchendes Objekt zu verstehen und Zugänge zu finden. Dafür braucht es einerseits anleitende Konzepte, welche Abschnitt 2.2 erörtert. Vorgelagert ist aber eine Erkundung des Gegenstandes und seiner Relationen, um überhaupt einen Überblick zu erhalten.

Die ethnografische Erkundung mittels des Verfolgens von Objekten, Phänomenen und ihren Verbindungen über mehrere Orte hinweg hat sich in der soziologischen Forschung mehrfach bewährt. So hat beispielsweise die Organisationssoziologie diesen Ansatz genutzt, um organisationale Felder zu erforschen und zu konstruieren: Saxenian (1996) hat mittels ethnografischer Methoden zwei regionale Ökonomien rekonstruiert und deren Kultur analysiert; Stark (2009) spannte seine Untersuchung heterarchischer Organisationsformen ebenfalls über mehrere ethnografische *field sites* – die Forschung dieser beiden Personen wird später nochmals Thema sein in dieser Arbeit. Mische (2008) hat ihr Feld, welches sie schliesslich mit einer Netzwerk-Analyse untersuchte, zuerst ethnografisch erkundet: »As I delved into this world, I explicitly sought out participants from a wide range of groups and attempted to understand the accounts of contending factions. While this helped me to understand the range of perspectives in play, it occasionally led to tensions and difficulties« (Mische 2008:9). In der Journalismusforschung skizzierte Anderson (2013b) zuerst ein Netzwerk des Medienplatzes Philadelphia, um es als »ecosystem« zu verstehen und es abzugrenzen – um dann anschliessend verschiedene wichtige Knotenpunkte vertieft zu untersuchen.

In meiner Arbeit führte ich die Erkundung mit verschiedenen Instrumenten durch: formelle und informelle Gespräche mit Akteur*innen; Teilnahme an Kursen und Netzwerktreffen zum Thema Datenjournalismus; Online-Beob-

achtung der datenjournalistischen Aktivitäten durch Literatur über das Feld¹, diverse Newsletter² sowie Twitter³.

Nach den ersten Erkundungen verfügte ich über ein Abbild des datenjournalistischen Feldes, seiner Akteur*innen und Verbindungen, welche ich als Netzwerk aufzeichnete. Dies erlaubte es mir, die internationale Vernetzung des Feldes, aber auch die lokal enge Verflechtung und Community-Bildung zu identifizieren. Auch liess die Abbildung zentrale Knotenpunkte erkennen, welche wichtige Stimmen in der Community verkörpern. Diese drei Charakteristika aus der Erkundung leiteten mich in der weiteren Erforschung der Feld-Ebene an.

Bereits die ersten Gespräche mit zentralen Akteur*innen sowie die Teilnahmen an Treffen zeigten eine In-Bezug-Setzung verschiedener Expertisen

-
- 1 Zum Beispiel: Data Journalism Handbook 1.0 (Gray et al. 2012); Data Journalism in 2017 (Google News Lab 2017); Jonathan Strays »The Curious Journalist's Guide to Data« (<https://towcenter.gitbooks.io/curious-journalist-s-guide-to-data/content/>, 10.12.2020)
 - 2 Ausgehend von den Empfehlungen von Datenjournalismus-Professor und Blogger Paul Bradshaw (<https://onlinejournalismblog.com/2017/09/19/email-newsletters-about-data-journalism/>, 10.12.2020) wurden ab Ende September 2017 folgende Newsletter über datenjournalistische Berichterstattung abonniert: in other news (Giuseppe Sollazzo) > <https://puntofisso.net/>, 10.12.2020); Fair Warning (Sophie Warnes) > <https://www.getrevue.co/profile/FairWarning>, 10. 12.2020); Real World Data Dispatch (Heather Krause/Dataassist) > <https://idatassist.com/>, 10. 12.2020); Data Is Plural (Jeremy Singer-Vine) > <https://tinyletter.com/data-is-plural>, 10.12.2020); Data & Eggs (The DataFace) > <http://thedataface.com/>, 10.12.2020); 1801 newsletter (Peter Yeung) > <https://medium.com/@peteryeung>, 10.12.2020). Weiter kamen dazu Newsletter von Journalismus-Vereinigungen (Global Investigative Journalism Network, European Data Journalism Network, Netzwerk-Recherche, European Journalism Centre) und -Produzenten (Correctiv, The Upshot/New York Times, Bellingcat). Da es sich um Hintergrund-Materialien und wenig systematische Feld-Beobachtungen handelt, wird auf eine komplette Auflistung von Newslettern und deren Details verzichtet.
 - 3 Einerseits durch »Folgen« von im Feld aktiven Personen und Institutionen, andererseits durch den wöchentlichen Service »Top Ten Data Journalism Links« des Global Investigative Journalism Network. Er basiert auf einer Analyse, welche jede Woche die zehn in der Twitter-Community meistdiskutiertesten Tweets mit dem Hashtag #ddj (»data-driven journalism«) herausrechnet. Dabei handelt es sich um datenjournalistische Projekte, Daten-Ressourcen, Handbücher, Ausbildungsinformationen, et cetera. Die Tweet-Liste inklusive Links wird dann über einen Blog-Post und wiederum über Twitter in die Community zurückgespiegelt. Leider ist die Methodik, welche die Tweets rangiert, äußerst komplex und kaum nachvollziehbar. Dies schmälert die Belastbarkeit dieser Quelle für die eigene Forschung. <https://gijn.org/series/top-10-data-journalism-links/> (18.11.2020)

aus bisher eigenständigen Professionen. Um Datenjournalismus als heterogenes und trotzdem eng verflochtenes Feld verstehen zu können, erachtete ich es als unumgänglich, datenjournalistische Praktiken auch auf der organisationalen Ebene zu untersuchen. Um die Relationen zwischen den Ebenen analysieren zu können, sollten die gewählten Schauplätze zudem durch möglichst viele Verbindungen im datenjournalistischen Netzwerk verflochten sein.

Aus dem Untersuchungsinteresse folgte in der Konsequenz, dass es sich um eine empirisch ausgerichtete, der Exploration durch Feldforschung verpflichtete Arbeit handelt. Die Prüfung bisheriger Theorien steht nicht im Vordergrund – stattdessen hat die Forschung den Anspruch, induktiv, aus einer reichen Empirie schöpfend, neue theoretische Einsichten zu generieren.

2.2 Die Vorgehensweise anleitende Konzepte

Nach der ersten Erkundung leiteten mich drei methodologische Konzepte in diesem Prozess an, um in der Datenerhebung schliesslich methodische Kohärenz und damit eine wissenschaftliche Qualität der Schlüsse zu gewährleisten.

2.2.1 Der abduktive Ansatz

Meine Forschung folgte keiner klaren chronologischen Trennung von theoretischer und empirischer Arbeit. Die bereits beschriebene explorative Felderkundung führte ich ab Sommer 2017 bis zum Frühling 2020 weiter. In diesen drei Jahren entwickelte ich meine Erkenntnisse über das Feld, Relationen und Praktiken iterativ weiter, ständig zwischen empirischer Datensammlung und deren Auswertung, theoretischen Annahmen und Einordnungen sowie weiterer empirischer Forschung changierend.

Dabei handelt es sich um einen abduktiven Ansatz (Reichert 2013; Tavory und Timmermans 2014), zurückgehend auf den Philosophen und Semiotiker Charles Sanders Peirce. Das theoretische Einordnen und Generieren von neuen Hypothesen ist dabei eine den Forschungsprozess begleitende Praktik und stützt sich auf bereits gewonnene Daten ab. Aufgrund dieser Hypothesen wiederum werden weitere Daten gesammelt. Ausgehend von der bereits erwähnten Frage, weshalb es trotz einer Diversität an Expertisen und Narrativen *den* Datenjournalismus gibt, präsentiert sich der abduktive Ansatz als passend: »Die Abduktion ist gefordert, wenn man mit dem Wissen, das man hat, nicht weiter kommt, weil man für etwas Problematisches *keine*

entsprechende Erklärung oder Regel findet. Deshalb muss man etwas Neues erfinden, welches das Unverständliche verständlich macht [Hervorhebung im Original]« (Reichert 2013:18). Insofern bietet die Abduktion eine passende Vorgehensweise, das Phänomen eines entstehenden Feldes, welches sich durch seine Verschiedenartigkeit auszeichnet und trotzdem in einer eng verflochtenen Community versammelt ist, zu ergründen.

Der abduktive Ansatz führte dazu, dass ich bereits nach wenigen Wochen erste Memos über die gesammelte Empirie schrieb und sie dabei mit verschiedenen theoretischen Perspektiven einordnete. Woraus sich wiederum neue Fragen und Annahmen entwickelten. Der Ansatz rahmte meine Vorgehensweise auf mehreren Ebenen. In der Beobachtung des Feldes führte die sich im Laufe der Forschung herauskristallisierende zentrale Rolle von Datavisualisierungen dazu, dass die Überschneidungen mit der Datavis-Community vermehrt in den Blick rückten.⁴ Generell transkribierte ich qualitative Interviews, welche ich mit Feld-Akteur*innen führte, umgehend, ordnete sie in theoretische Überlegungen ein und leitete daraus neue Fragen ab, welche wiederum meinen Forschungsprozess mitprägten. Aus diesen Gründen erfolgte im Laufe der Zeit in den Interviews eine gezieltere Thematisierung von Aushandlungsmechanismen und Visualisierungspraktiken.

Der abduktive Ansatz prägte insbesondere meine ethnografischen Untersuchungen in den beiden Teams. Vor der ersten Feldforschungsphase beschrieb ich mehrere theoretische Fokusse, mit denen ich die Feldarbeit begann. Nach der Forschungsphase im ersten Team erfolgte eine induktive Kodierung der Feldnotizen und deren Reflexion mittels Memos. Die vorläufigen Schlüsse aus der ersten Feldforschung schärften die Fokusse für Aushandlungen, Projekte als Analyseeinheiten und die Bedeutung der Materialität (Werkzeuge) in der Untersuchung des zweiten Teams und beeinflussten dadurch auch dessen Auswahl. Die ersten Kodierungsdurchgänge der

4 Beobachtet beispielsweise durch die Anfang 2019 gegründete Data Visualization Society. Die Vereinigung wuchs im ersten Jahr explosionsartig und verzeichnete innert Kürze mehrere tausend Mitglieder, schuf eine eigene Publikation (namens Nightingale) und eine eigene Konferenz (namens Outlier). <https://www.datavisualizationsociety.com/> (11.12.2020). Weitere wichtige Einblicke in das Visualisierungsfeld erfolgten über den Podcast datastori.es (<https://datastori.es/>, 11.12.2020) sowie über Preisverleihungen wie die Malofiej- (<https://www.malofiejgraphics.com/>, 11.12.2020) und Information-is-Beautiful-Awards (<https://www.informationisbeautifulawards.com/>, 11.12.2020), die immer wieder Datenjournalist*innen und ihre Projekte diskutierten.

Feldnotizen habe ich später verworfen und stattdessen mit neuen Fokusse erneut kodiert.

Die hier angedeuteten Methoden werde ich in den Abschnitten 2.3 und 2.4 ausführlicher beschreiben.

2.2.2 Praktiken und Relationen: *zooming in and out*

Praktiken und Relationen bilden die ontologischen Grundeinheiten meiner Untersuchungen. Praktiken konstituieren sowohl »Doings« als auch »Sayings« (Schatzki 2000:103), das heisst, diese Analyseeinheit fasst sowohl Handlungen als auch Aussagen. Praktiken finden sich einerseits auf einer organisationalen Ebene, beispielsweise in der Arbeit auf den Redaktionen, wenn mit soziomaterialien Praktiken aus Daten Erkenntnisse in Form von Geschichten produziert werden: »Focusing on practices is thus taking the social and material doing (of something: doing is never objectless) as the main focus of inquiry« (Nicolini 2013:221). Andererseits steckt in Praktiken immer auch ein Teil des Feldes: Ansichten darüber, was richtig und falsch ist, gehören nicht (nur) zu Individuen, sondern sind als Wissenstand Teil von Praktiken (Reckwitz 2003) und verdichten sich zu Handlungsanleitungen und Bewertungskriterien auf der Feld-Ebene, welche verschiedene Akteur*innen zueinander in Beziehung setzt.

Praktiken sind durch Relationen verbunden: Praktiken sind die (temporäre) Verfestigung von sozialen Strukturen und Verbindungen und verstetigen diese gleichzeitig weiter – durch die Performanz von Praktiken werden soziale Institutionen (beispielsweise Communitys oder Organisationen und ihre Prozesse) erst aufrechterhalten (Nicolini 2013:3). Um Datenjournalismus zu erforschen, sind folgerichtig Praktiken und Relationen als Grundeinheiten relevant.

Um mit Praktiken und Relationen zu analysieren, bietet sich das Konzept von »Zooming In and Out« an (Nicolini 2009, 2013) an. Es besteht einerseits darin, auf verschiedene Fokusse heranzuzoomen – je nach Forschungsperspektive(n) können dies einige der folgenden Aspekte sein: »Sayings and Doings; Interactional order; Timing and Tempo; Bodily choreography; Tools, artefacts, and mediation work; Practical concerns; Tension between creativity and normativity; Processes of legitimization and stabilization« (Nicolini 2013:220). Durch das Heranzoomen wird sichtbar, welche Elemente durch die Praktiken gebündelt werden und wie diese Elemente interagieren (Nicolini 2009:140off). Mit dieser Perspektive will ich aufzeigen, wie datenjournalis-

tische Praktiken in Redaktionen Expertisen versammeln und dabei durch Aushandlung Erkenntnisse und Evidenz entstehen.

Das Herauszoomen macht sichtbar, ob und wie menschliche und nicht-menschliche »mediators« im Feld Relationen zwischen lokalen Praktiken, beispielsweise in Organisationen, knüpfen und aufrechterhalten (Nicolini 2013:228ff). Diese Relationen können verschiedene Formen annehmen: Protokolle, Konventionen, Software, Rezepte, Ideologien, Kategorisierungen, Geschichten usw. (Nicolini 2009:1408ff, 2013:228ff), denen es dann in der Forschung zu folgen gilt (Czarniawska 2007). Praktiken sind das Resultat eines »complex interplay between local and global« (Nicolini 2009:1411).

zooming in and out erweist sich mit seiner Offenheit für verschiedene Aspekte von Praktiken und mit der Untersuchung von Relationen zwischen verschiedenen Orten als geeignete Vorgehensweise für meine Forschung. Das Konzept integriert den abduktiven Ansatz durch sein Changieren zwischen verschiedenen Perspektiven, welche neue theoretische Fragen und empirische Einsichten mit sich bringen – genauso wie es anschlussfähig ist für das Gerüst der relationalen Ethnografie, welches der folgende Abschnitt einführt.

2.2.3 Relationale Ethnografie

Um das Phänomen Datenjournalismus als Feld mit seinen Relationen wie auch in der Aushandlung der Praktiken untersuchen zu können, benötigte ich ein Gerüst, welches diese Aspekte zu integrieren vermag. Praxisstudien basieren oft auf ethnografischer Feldforschung, weil diese soziale Gegebenheiten als Resultat alltäglicher Interaktionen von Struktur und Agency (O'Reilly 2005:6) beleuchtet. Verschiedene ethnografische Konzepte versuchen, den traditionellen ethnografischen Fokus auf einen Ort oder eine einzelne Gruppe als Forschungsobjekt zu weiten.

Marcus (1995) entwickelte die »Multi-Sited Ethnography«, ein Ansatz, der den Forschungsgegenstand aus seinen Verbindungen und Bewegungen heraus konstruiert – und sich damit insbesondere für gegenwärtige Gesellschaften, geprägt von Mobilität, Vernetzung und Kommunikation, anbietet (siehe Hine 2007:655). Burrell (2009) entwickelte Marcus' Ansatz weiter und spricht von einer »Network ethnography«. Diese schlägt vor, damit zu beginnen, ein Phänomen an einem Ort zu untersuchen und dann spezifischen Verbindungen von diesem Startpunkt aus zu folgen und weitere Orte ethnografisch zu untersuchen – unabhängig davon, ob diese virtuell oder reell sind, womit das Konzept die Integration von Online-Welten leisten soll.

Für meine Forschung stütze ich mich auf eine neuere Weiterentwicklung dieser methodologischen Linie: Desmonds »relational ethnography« (2014). Desmond (2014:554) hält fest, dass der Ansatz der »Multi-Sited Ethnography« angetrieben sei von der Logik des Vergleichs – und schlägt sein Konzept als Alternative vor, welche die Verbindungen ins Zentrum stellt: Die lokalen Praktiken dürfen nicht von ihrer Einbettung ins Feld separiert werden. Die Erklärung von Phänomenen liegt nicht in strukturellen Zwängen oder individuellen Agenten, sondern in kontingenten Verbindungen (Desmond 2014:553f) – Ideen und Konventionen, welche durch eine eng verflochtene Community geprägt sind, wie die Arbeit zeigen wird.

Durch das relationale Konzept soll eine Ethnografie ihre »substantivist perspective«, den Fokus auf Gruppen und Orte, überwinden (Desmond 2014:548). Der Autor schlägt vor, ein Feld aus verschiedenen Positionen (mindestens zwei) mit unterschiedlichen Perspektiven zu untersuchen:

»Most basically, a relational approach incorporates fully into the ethnographic sample at least two types of actors or agencies occupying different positions within the social space and bound together in a relationship of mutual dependence or struggle« (Desmond 2014:554).

Dadurch können die Relationen eines Feldes und der Praktiken in den Fokus rücken, indem Forschende nach den Verbindungen der Elemente fragen. Diese Verbindungen können verschiedenartig sein, je nach Akzent der Forschungsfrage(n) und dem daraus konstruierten ethnografischen Untersuchungsobjekt. Gemäss Desmond (2014:555ff) kann die Forscherin dafür auf Punkte des Kontakts und Konflikts, auf die Produktion koordinierter Handlungen zwischen Organisationen und die Ökologie eines Feldes fokussieren. Es gilt, Felder statt Orte zu studieren; Abgrenzungen in den Blick zu nehmen; Prozesse statt Personen zu erforschen; die Koordination verschiedener Kulturen zu untersuchen.

Der Rahmen einer relationalen Ethnografie ist so angelegt, die Praktiken in den Organisationen als mit den Dynamiken im Feld verbunden zu betrachten – und schliesst damit an die Idee von *zooming in and out* an. Wie ich diese Konzepte für meine Forschung angewandt habe, erläutern die nächsten Abschnitte.

2.3 Teamarbeit untersuchen

Nach den ersten Felderkundungen, beschrieben im Abschnitt 2.1, entschied ich mich, die datenjournalistischen Praktiken innerhalb von Organisationen – konkret Redaktionen – eingehender zu erforschen. Datenjournalismus-Teams und ihre Projekte hatten sich im Feld als wichtige Organisationsform präsentiert, sodass ich mir von einem Einblick in diese Organisationen einen grossen Erkenntnisgewinn versprach.

In Redaktionen ist journalistische Kultur unter dem Brennglas sichtbar. Sie haben deshalb in der Journalismusforschung schon immer eine grosse Rolle als Kontext für die Untersuchung von journalistischer Erkenntnisproduktion und deren soziomateriellen Wechselwirkungen gespielt. Diese Untersuchungen erfolgten sehr oft mittels der ethnografischen Methode, welche es erlaubt, die angesprochenen Prozesse sehr umfassend und feinkörnig festzuhalten. Zu den bedeutendsten ethnografischen Untersuchungen in journalistischen Redaktionen gehören die Bewertungskriterien für »Deciding what's news« in der Arbeit von Journalist*innen (Gans 1980) und die Routinen der Nachrichtenproduktion (Tuchman 1980). Seit den Neunzigerjahren ist die Adoption von digitalen Technologien in den Redaktionspraktiken in den Fokus gerückt (siehe beispielsweise Boczkowski 2004; Anderson 2013a; sowie für eine Übersicht Paterson und Domingo 2008).

Viele Journalismusforschende schlossen dabei für ihre Feldforschung in Redaktionen an die ethnografischen Stränge der STS und der Organisationsforschung an. Die STS haben mit den Laborstudien die Bedeutung von Instrumenten und Dingen in der Wissensproduktion ethnografisch erforscht, Abschnitt 4.2.1 wird diese noch genauer vorstellen. In der Organisationsforschung wiederum haben bekannte Werke mittels ethnografischer Beobachtung auf die strukturierende Rolle von Technologien in der Arbeitsorganisation hingewiesen (Barley 1986, 1996; Orr 1996, 2006). Neuere Ansätze haben die Organisationsforschung explizit mit medienwissenschaftlichen Perspektiven beleuchtet (Conrad 2017).

Auch einige Analysen datenjournalistischer Praktiken nutzten ethnografische Methoden (beispielsweise Royal 2012; Usher 2016), wobei sich die Forschungsdauer sehr oft auf wenige Tage oder Wochen der Beobachtung beschränkte. Ausserhofer et al. (2017), welche eine Meta-Studie über Datenjournalismusforschung erarbeiteten, identifizieren denn auch ethnografische Untersuchungen als Forschungslücke im Feld.

2.3.1 Datenerhebung

Mit dieser Arbeit möchte ich diese Lücke etwas verkleinern. Ich habe dafür zwei Datenjournalismus-Teams und deren Praktiken ethnografisch untersucht.

Die Teams: NZZ Storytelling und BR Data

Durch die Erkundung der Deutschschweizer Community verfügte ich über Kontakte in fast alle institutionalisierten Datenjournalismus-Teams in diesem Landesteil, wobei diese aufgrund der Kleinräumigkeit und der wenigen grossen Medienorganisationen an einer Hand abzuzählen sind. Durch persönliche Kontaktaufnahme mit dem Teamleiter gelang ein Feldzugang für eine ethnografische Beobachtungsstudie in der Organisation der Neuen Zürcher Zeitung (NZZ) im März 2018.⁵ Ich beobachtete während dreier Wochen das Team von NZZ Storytelling, das datenjournalistische Team innerhalb der NZZ-Redaktion⁶ – dies natürlich unter Zusicherung der Anonymität der Personen sowie der Verschwiegenheit über beobachtete Aspekte, die für meine Forschung keine Relevanz aufweisen.⁷ In Beobachtungen und direkt wiedergegebenen

5 Bei NZZ Storytelling, genauso wie später auch bei BR Data, gehörte es zur Abmachung für den Feldzugang, dass ich den beiden Teams in den Wochen danach je eine ›Reflexion‹ über aus meiner Sicht interessante Beobachtungen aus dem Team-Alltag zusammenstellte. Die jeweils mehrseitige Reflexion überreichte ich dem jeweiligen Teamleiter. Der Inhalt oblag allein mir und wurde vom Team nicht mitbestimmt, weder im Voraus noch durch Wünsche im Nachhinein. Insofern hatte dieses Dokument keinen Einfluss auf meine Beobachtungsfokusse. Ebenfalls deklarierte ich dem Team, dass es sich bei der Reflexion nicht um eine wissenschaftliche Einordnung handelt.

6 Im Folgenden ist in dieser Arbeit oft von ›Redaktor‹ oder ›Redaktorin‹ die Rede, um Mitarbeitende der Redaktion ausserhalb der untersuchten Teams zu bezeichnen. Die Bezeichnung verweist allgemein auf eine Mitgliedschaft in der Redaktion. Ich nehme aber keine Abgrenzung zu anderen Redaktionsrollen – beispielsweise ›Reporter‹ – vor. In beiden Redaktionen war eine solche Rollenabgrenzung bestenfalls unscharf und für die Untersuchung der Praktiken wenig relevant. Deshalb habe ich eine einheitliche Bezeichnung für alle Redaktionsmitglieder gewählt.

Der kontinentaleuropäische Kontext meiner Untersuchung unterscheidet sich in diesem Punkt vom englischen Sprachraum, in dem die Rollen von ›editor‹ und ›reporter‹ klar abgegrenzt sind, siehe Schudson 2000:186.

7 Da es sich bei Beobachtungen am Arbeitsplatz um einen Eingriff in die Persönlichkeitsrechte handelt, habe ich mit sämtlichen Teammitgliedern von NZZ Storytelling sowie BR Data eine persönliche Einverständniserklärung besprochen, welche sie über ihre

Feldnotiz-Zitaten, die Personen direkt referenzieren und Teil dieser Arbeit sind, habe ich die Namen der Teammitglieder mit einem Pseudonym ersetzt.⁸

NZZ Storytelling⁹ entstand Ende 2015 aus dem Team von NZZ Data heraus. Der Namenswechsel ging einher mit einer Neuausrichtung der Rolle: NZZ Data war sehr frei in seinen Methoden und der Geschichtenwahl, es resultierten aufwendige, technisch hochstehende Geschichten. NZZ Storytelling produziert immer noch eigene Geschichten, ein wichtiger Teil der Rolle des Teams besteht gemäss eigener Beschreibung aber auch darin, die anderen Ressorts »mit Beratung, Formatvorlagen, Tools und spezifischen grafischen Elementen« zu unterstützen.¹⁰ So entwickelte das Team in der Folge Q, eine Toolbox für simple Datenvisualisierungen, welche der gesamten Redaktion zur Verfügung steht. Dies führt zu einer Parallelität verschiedener Rhythmen: Fast alle NZZ-Storytelling-Teammitglieder sind tagtäglich in die Produktion tagesaktueller und schneller genauso wie längerfristiger und aufwendiger Geschichten involviert.

Im Rahmen der Neuausrichtung auf eine (auch) transversale Rolle in der gesamten Redaktion integrierte die NZZ das Infografik-Team, welches seit Jahrzehnten als Dienstleisterin für die Redaktion funktioniert hatte, in das dann dreizehnköpfige Storytelling-Team. Ebenso beinhaltet es mehrere Programmierer, deren Auftrag es primär ist, die technische Infrastruktur des Teams zu unterhalten sowie diese weiterzuentwickeln. Dazu kommen Mitglieder, die einen (daten-)journalistischen Hintergrund mitbringen, sowie *Designer*, welche neben ihren gestalterischen Kompetenzen auch programmieren können.

Rechte sowie meine Pflichten aufklärt und ihnen Anonymität zusichert. Alle Teammitglieder haben diese Einverständniserklärung unterschrieben und als Kopie erhalten.

- 8 Das Pseudonym habe ich dem Geschlecht angepasst gewählt. Die einzige Person ohne Pseudonym ist der NZZ-Storytelling-Teamleiter David Bauer, der als Kopf des Teams einer interessierten (Journalismus-)Teil-Öffentlichkeit bekannt und in seiner Rolle (als Teamleiter) in den Aufzeichnungen so oder so zu identifizieren ist. Ich danke ihm, dass er sich mit der Nennung seines realen Namens einverstanden erklärt hat.
- 9 Seit September 2018 nennt sich das Team »NZZ Visuals«. Es wurde seit der Forschungsphase vergrössert durch eine Motion Designerin, drei zusätzliche Programmierer sowie den Zusammenschluss mit dem Video-Team (fünf Personen). Seit Ende 2018 bilden die Programmierer eine eigene Abteilung namens Editorial Tech, arbeiten aber immer noch sehr eng mit NZZ Visuals zusammen. [https://www.nzz.ch/visuals/ \(12.12.2020\)](https://www.nzz.ch/visuals/ (12.12.2020))
- 10 NZZ Schulungspräsentation_Q (Seite 4)

Sie alle versammeln sich unter dem Dach des traditionsreichsten und wohl auch renommieritesten Schweizer Medienhauses (für eine Einordnung siehe (Maissen 2005). Obwohl auch die NZZ, wie alle traditionell im Printbereich starken Medienorganisationen, von den Zäsuren eines erodierenden Finanzierungsmodells betroffen ist, sieht sich die NZZ immer noch höchsten Qualitätsansprüchen verpflichtet. Aus diesen historischen Gegebenheiten heraus erscheint es konsequent, dass für NZZ Storytelling die hohe visuelle Qualität in der Präsentation ihrer Erkenntnisse sehr wichtig ist.

Nach einer ersten Auswertung und theoretischen Einordnung der Feldforschung bei NZZ Storytelling galt es, ein zweites Team im datenjournalistischen Feld zu untersuchen, das eine etwas andere Position und damit Perspektive mitbringt. Um die Transnationalität des Feldes zu würdigen, aber trotzdem aus forschungspraktischen Gründen bei derselben Sprache zu bleiben, fragte ich drei ausgewählte Datenjournalismus-Teams aus Deutschland für eine zweite Praxisstudie an.

Zugang gewährte mir das datenjournalistische Team des öffentlich-rechtlichen Bayerischen Rundfunks in München, BR Data. Auch in diesem Fall erfolgte eine persönliche Kontaktaufnahme mit dem Teamleiter. Die Erforschung der Praktiken bei BR Data fand während dreier Wochen im Dezember 2018 und Januar 2019 statt. BR Data existiert seit 2015 und war gemäss Angaben der Teammitglieder, die bereits seit Beginn dabei sind, bereits an verschiedenen Stellen im Organigramm >angehängt<. Im Spätherbst 2018, also kurz vor meiner Forschungsphase, wurde BR Data als Partner der investigativen Einheit BR Recherche deren Leiterin Verena Nierle unterstellt, blieb aber als Team bestehen.¹¹ Von dessen zehn Mitgliedern sieht sich die Hälfte als (Daten-)Journalist*innen, welche Geschichten recherchieren, konzipieren und schreiben – die aber alle auch programmieren können. Ebenso gibt es Teammitglieder, welche keinen oder sehr wenig journalistischen Hintergrund mitbringen und sich stattdessen fast ausschliesslich auf Programmierpraktiken beschränken. Einige wenige sehen ihre Expertise beiderorts.

¹¹ Da im beobachteten Arbeitsalltag immer wieder investigativ konzipierte Geschichten diskutiert wurden, welche später als >Primeur< die Öffentlichkeit erreichen sollten, war es dem BR ein grosses Anliegen, dass ich eine zusätzliche, BR-eigene Einverständniserklärung unterschrieb. Diese hält fest, dass meine Arbeit nicht die Investigationen der Journalist*innen in den Fokus rückt. Aus diesem Grund habe ich die Themen der Projekte des Teams erst dann in meinen Vorträgen erwähnt, wenn sie abgeschlossen waren. Inzwischen sind alle damals diskutierten und bearbeiteten Projekte entweder veröffentlicht oder wurden fallen gelassen.

BR Data funktioniert innerhalb der Medienorganisation als eigenständiges Team mit grossen Freiheiten, das für seine Projekte punktuell mit anderen Teams oder Ressorts zusammenarbeitet. Die geplante enge Zusammenarbeit mit BR Recherche befindet sich während der Forschungsphase im Aufbau. Mehrfach kooperiert BR Data auch innerhalb der deutschen Datenjournalismus-Community mit Teams anderer Medienorganisationen und kann dabei auf einen guten Ruf als eines der grössten und renommierertesten Teams zählen. Innerhalb der eigenen Organisation kann BR Data nur online selbst publizieren und ist darüber hinaus auf die Zusammenarbeit mit Radio- oder Fernseh-sendungs-Redaktionen angewiesen. Die schon seit 2011 laufende Umstrukturierung hin zur »Trimediaität«¹² – der BR will sich nicht mehr nach Ausspiel-kanälen organisieren, sondern nach Themenfeldern, welche dann alle Kanäle bedienen sollen – zieht sich noch einige Zeit hin und war in der Forschungsphase im Team kaum ein Thema. Stattdessen nahm die Frage Raum ein, wie datenjournalistische Geschichten auch für Radio- und Fernsehsendungen attraktiv produziert werden können.

Die Charakteristika der beiden untersuchten Teams zusammengefasst:

Tabelle: Eckwerte zu den beiden Praxisstudien.

	NZZ Storytelling	BR Data
Institutioneller Hintergrund	traditionsreicher Print-Verlag	öffentlich-rechtliches Radio und Fernsehen
Organisationale Verortung des Teams	querliegend, Schnittstellen zu allen Ressorts	Programmbereich Politik und Wirtschaft, zusammengeschlossen mit Recherche/Investigativ-Team
Anzahl Mitarbeitende im Team	13	10

12 <https://www.br.de/unternehmen/inhalt/organisation/reform-br-intern-organisation-100.html> (15.12.2020)

Ausspielkanäle Geschichten	Texte und Visualisierungen online und gedruckt	Texte und Visualisierungen online (eigenes Web- Template), Radio- und Fernseh-Sendungen in Kooperation mit Sendungsredaktionen
Werkzeuge Datenanalyse	eigenes <i>Tool</i> Sue, Programmiersprachen (R, Python)	Programmiersprachen (R, Python)
Werkzeuge Datenvisualisierung	eigene grafische Zeichnungen (Adobe Illustrator/Sketch), ei- gene <i>Tools</i> (Q, Sue) für Dia- gramme	Programmiersprachen (via Notebooks), Datawrapper
Feldforschungsphase	März 2018, 3 Wochen	Dezember 2018/Januar 2019, 3 Wochen
Material	Feldnotizen, Kurz-Interviews (13), Fotos, Arbeitsprozess- Dokumentation, Überblick IT-Infrastruktur, Strategie-Dokumente, Aufgabenbeschriebe, Projekt-Pitches, <i>Tool</i> -Schulungsunterlagen, <i>Tool</i> -Zugänge (Editor- Interfaces), <i>Styleguide</i> , Veröffentlichungen ¹³	Feldnotizen, Kurz-Interviews (10), Abbild Projekt-Board, Abbild Visualisierungsskizzen, Strategie-Dokumente, Organigramme, Veröffentlichungen

Die je mehrwöchige Dauer der Feldforschung ermöglichte einerseits ei-
nen vertieften Einblick und das Festhalten eines redaktionellen Alltags (O'Reil-
ly 2005:1), liess aber andererseits genügend zeitliche Ressourcen meines For-
schungsprojekts unverbraucht, um nach dem anschliessenden Herauszoomen

13 In der Regel referenziere ich veröffentlichte Geschichten, auf die meine Arbeit Bezug
nimmt, direkt mit einem Link. An einigen Stellen fehlt ein solcher, um die Anonymität
der einzelnen Teammitglieder zu wahren.

sich weitere Fälle anzuschauen, die Verbindungen zu studieren und weitere Facetten von datenjournalistischen Praktiken zu vertiefen.

Vorgehensweise in den Teams

Die ethnografische Methode hat es mir ermöglicht, nicht verbalisierte soziale und technologische Dimensionen datenjournalistischer Praktiken in den Teams zu erfassen und zu beschreiben. Für die Beobachtungen nutzte ich das Konzept des »Shadowing« (Czarniawska 2007, 2014) und folgte einzelnen Projekten oder Teammitgliedern. Aufgrund der offenen Fragestellung meiner Forschung zeichnete ich die Praktiken und Abläufe im Team breit auf – insbesondere während der ersten Praxisstudie. Die Beobachtungen und informellen Gespräche im zweiten Team trugen dem in der Zwischenzeit theoretisch etwas geschärften Fokus auf Expertisen, Aushandlungen, Legitimierungsnarrative und Datenvisualisierungen Rechnung, ohne aber den Team-Alltag nicht trotzdem umfangreich und offen aufzuzeichnen.

Meine Beobachtungen und Gesprächsaufzeichnungen hielt ich in handschriftlichen und später auf dem Computer vertexteten Feldnotizen fest (Emerson et al. 2011), jeweils eine Feldnotiz für jeden Tag (siehe Anhang; A Vorlage Feldnotiz). Die Feldnotizen beinhalteten eine Spalte, in der ich – direkt nach der Niederschrift, während eines vollständigen Durchlesens der Notizen – erste analytisch imprägnierte Fragen und Feststellungen festhielt. Zudem notierte ich in jeder Feldnotiz erstens die Rahmenbedingungen des und Erwartungen an das Tagesgeschehen und reflektierte zweitens meine persönliche Rolle in den Geschehnissen sowie relevante Empfindungen den Teammitgliedern gegenüber. Dies dient der Reflexion des eigenen Einflusses auf die soziale Situation und deren Interpretation (O'Reilly 2005:211ff).

Die Feldnotizen habe ich ergänzt mit Fotos, Raumskizzen und Dokumenten, die mit den Praktiken verknüpft (Nicolini 2013) und deshalb für deren Interpretation von Bedeutung sind (O'Reilly 2005:2). Im Fall von NZZ Storytelling handelt es sich bei diesen Artefakten um Materialien wie online zugängliche Werkzeuge, Präsentationen und Handbücher. Dazu kommen produzierte und veröffentlichte Arbeiten (Artikel). Bei BR Data ergänzten neben Fotos und Raumskizzen ebenfalls Dokumente die Datensammlung, wobei aufgrund der weniger standardisierten Prozesse und fehlender Tool-Eigenentwicklungen weniger Zusatzmaterial anfiel (siehe Tabelle).

Die Beobachtungen und informellen Gespräche ergänzte ich mit je einem semistandardisierten Kurz-Interview, das ich mit allen Team-Mitgliedern durchführte. Im Kurz-Interview fragte ich nach den jeweiligen dem Mitglied

zugeteilten Aufgaben, der Einschätzung der eigenen Expertise, Arbeitsbeziehungen, den Ansichten über den Forschungsgegenstand, den eigenen Lernstrategien und Weiterbildungswünschen. Den für den ersten Fall konzipierten Fragebogen passte ich moderat den Unterschieden im Organisationskontext des zweiten Teams an und fügte eine zusätzliche Frage zum Entwicklungsprozess von Geschichten hinzu (siehe Anhang; B und C Fragebogen Kurz-Interviews). Das Kurz-Interview deckte in diesem Sinne viele Bereiche ab, welche im Alltag nur punktuell beobachtbar waren, und es ergänzte die Aufzeichnungen aus den Feldnotizen. Durch die Standardisierung der Fragen stellte ich zudem sicher, über die abgefragten Aspekte ein vollerhobenes und vergleichbares Bild der Akteur*innen und der Unterschiede und Gemeinsamkeiten im jeweiligen Team zu erhalten. Gleichwohl galt es für die Auswertung zu berücksichtigen, dass es sich bei den Interview-Antworten um diskursive und nicht beobachtete Praktiken handelt.

In beiden Teams führte ich einige Monate nach der Forschungsphase mit je einem Teammitglied ein Nachgespräch durch. Dabei ging es darum, zu erfragen, was mit verschiedenen längerfristigen Projekten, die während der Forschungsphase in Bearbeitung waren, passierte. Diese Informationen ergänzten die weitergehende stetige, wenn auch distanzierte Beobachtung der beiden Teams und ihrer veröffentlichten Geschichten auch nach den jeweiligen Forschungsphasen.

In beiden Teams führte ich zudem je ein rund halbstündiges Interview mit einer Leitungsperson der Medienorganisation¹⁴, welche mir über die strategische Rolle des jeweiligen Datenjournalismus-Teams für die gesamte Redaktion Auskunft erteilte. Diese Interviews lieferten mir wichtige Hintergrundinformationen, um die Abläufe in den Teams besser interpretieren zu können. Sie waren auch aufschlussreich bezüglich der Identität der jeweiligen Medienorganisation sowie der damit zusammenhängenden Narrative, welche die Leitungsgremien von aussen an ihre datenjournalistischen Teams herantragen. Der Einbezug dieser Aspekte sprengt aber den Rahmen dieser Arbeit und bietet wenig analytischen Mehrwert für die im Fokus stehenden Fragen. Deshalb habe ich darauf verzichtet, diese beiden Interviews in die Auswertung mit Kodes miteinzubeziehen.

¹⁴ Bei der NZZ handelte es sich um den Chefredaktor Eric Gujer, beim BR um den Informationsdirektor Thomas Hinrichs.

2.3.2 Datenauswertung

Um dem Anspruch von Praxisstudien als »ethnographies as narrative tales« (Emerson et al. 2011:202, siehe auch van Maanen 2000) gerecht zu werden, gilt es, sich durch das komplette Durchlesen aller Feldnotizen erneut in die Atmosphäre, die ›Welt‹ des beobachteten ethnografischen Objekts zu versetzen. Gleich anschliessend habe ich innerhalb weniger Tage mithilfe der Software atlas.ti die Feldnotizen der jeweiligen Praxisstudie vollständig durchkodiert. Dabei bin ich in beiden Fällen gleich vorgegangen, habe sie aber in separaten atlas.ti-Projekten ausgewertet, um dem unterschiedlichen Kontext der Organisationen gerecht zu werden. In einem ersten Durchgang habe ich in beiden Fällen induktiv sämtliche Praktiken mit Kodes versehen und habe auf Abstrahierungen verzichtet (Kategorisierung im Sinne des »Open Coding«, siehe Emerson et al. 2011:175ff). Mit dem Ziel der Komplexitätsreduktion erfolgte dann eine Vereinigung von ähnlichen Kodes. Später erfolgte eine komplette Re-Kodierung.

Kennzeichnend für die Auswertung qualitativer Untersuchungen sind die wenig rigiden Regeln: Die Komplexität der Einzelfall-Realität, welche im aufgezeichneten empirischen Material verblieben ist, kann nicht mit einem Patentrezept in abstrahierte wissenschaftliche Erkenntnisse transformiert werden – wobei eine »high-wire abstraction« (van Maanen 2006) so oder so nicht zum Potenzial einer Ethnografie gehört. Für die ethnografischen Auswertungen habe ich mich an den Vorschlägen aus Standardwerken der ethnografischen Methode (Emerson et al. 2011; O'Reilly 2005) orientiert, diese aber durch eigene Sortierungs- und Auswertungsformate ergänzt. Über den Auswertungsprozess habe ich ein Log-Buch geführt (siehe Anhang; D Log-Buch Feld-Material).

Durch das Schreiben von Memos über die in der Auswertung beobachteten Phänomene und Zusammenhänge (»code memos«, siehe Emerson et al. 2011) sowie die gleichzeitig weitergeführte Beobachtung des Feldes und der vertieften Theorie (Abschnitt 2.2 hat diese Konzepte erläutert) schälten sich die für meine Forschung relevanten Kategorien und Relationen heraus. Auf dieser Basis habe ich die Kodes in 14 (NZZ Storytelling) beziehungsweise 10 (BR Data) grobe Kode-Gruppen eingeteilt. Über diese Kode-Gruppen habe ich wiederum Memos geschrieben, die ihren Inhalt (Praktiken, Relationen) einordneten und sie wo sinnvoll auch mit den weiteren gesammelten Dokumenten verknüpften. Um die Auswertung um eine zusätzliche Perspektive zu erweitern, habe ich zwei Phänomene mit der Netzwerk-Funktion aus atlas.ti, welche die Ver-

bindungen von Kodes und Zitaten anzeigt, abgebildet und dies wiederum mit Memos reflektiert.

Der Prozess der Produktion von (Daten-)Geschichten hat sich im Laufe der Forschung als ein nützlicher Analyse-Faden herausgestellt, auf dem sich die beobachteten Praktiken und Relationen in den Teams aufreihen liessen. In diesem Zusammenhang habe ich einerseits je das meistdiskutierte Projekt aus den beiden Teams chronologisch ausgewertet und beschrieben. Andererseits habe ich als relevant betrachtete Kodes auf diesem Prozessablauf eingeordnet und für beide Fälle je ein physisches Papierplakat daraus erstellt (siehe Anhang; E Prozessabläufe, Abbildung 26 und Abbildung 27). Ab diesem Zeitpunkt habe ich die theoretischen Einordnungen in Memos fallübergreifend geschrieben.

Parallel zu den Feldnotizen habe ich auch die Kurz-Interviews mit den Mitgliedern aus beiden Teams ausgewertet. Wiederum habe ich dafür je ein neues atlas.ti-Projekt eröffnet, um dem jeweiligen Organisationskontext Rechnung zu tragen. Die Antworten aus den Interviews habe ich induktiv kodiert und diese den jeweiligen Interviewfragen zugeordnet. Sofern sich eine Frage (beziehungsweise die Antworten) in meinem fortlaufenden Forschungsprozess als relevant herausstellten¹⁵, habe ich sie vertieft ausgewertet, in Memos beschrieben und in die Interpretation der Team-Ethnografie miteinbezogen. Da ich die Kurz-Interviews vollständig erhob – alle Teammitglieder haben alle Fragen beantwortet –, konnte ich die Antworten zudem nach Expertisen-Rolle im Team¹⁶ aufschlüsseln und die Antworten zwischen diesen Gruppen vergleichen.

¹⁵ Bei NZZ Storytelling betraf das die drei Fragen: Welche Fähigkeiten, Methoden oder *Tools* würdest du dir gerne aneignen? Auf welchem Weg? Wie eignest du dir (neues)

Wissen und (neue) Kompetenzen an? Was ist aus deiner Sicht der (Mehr-)Wert von Visualisierungen und vermehrter Nutzung von Daten für die Medienberichterstattung? Bei BR Data betraf dies die drei Fragen: Welche Fähigkeiten, Methoden oder *Tools* würdest du dir gerne aneignen? Wo liegt der Nutzen von Datenjournalismus? Wie macht man aus Daten Geschichten?

¹⁶ Bei NZZ Storytelling habe ich nach den folgenden Expertisen-Rollen aufgeschlüsselt: Infografiker*in (5 Personen), Datenjournalist*in (3); IT-Mitarbeiter*in (2); Designer*in (2); Teamleiter (1). Bei BR Data habe ich nach den folgenden Expertisen-Rollen aufgeschlüsselt: Datenjournalist*in (7); Programmierer*in (3). Da eine solche klare Expertisenzuweisung sich in den beobachteten Arbeitspraktiken immer wieder als problematisch erwies, habe ich sie einzig für die Auswertung der Kurz-Interviews als fixe Kategorie genutzt.

Die Auswertungsschritte der Feldnotizen und Kurz-Interviews habe ich wie erwähnt mit Memos begleitet, welche ich durch die fortdauernde Forschung stets etwas stärker mit Kontext und analytischen Kategorien und Referenzen *in-formiert* habe (siehe Kapitel 4). Sobald sich dieser Schreibprozess hin zu Texten entwickelte, welche als Teil der Ethnografie in Frage kamen, testete ich meine gezogenen Schlüsse mittels Vorträgen an Konferenzen sowie im Austausch mit Forschungskolleg*innen. Die Rückmeldungen flossen wiederum in den weiteren Erkenntnisprozess ein.

2.4 Das Feld analysieren

Das Feld des Datenjournalismus habe ich während knapp drei Jahren (April 2017 bis Frühling 2020) auf Online-Kanälen stets beobachtet – Abschnitt 2.1 hat diese Kanäle bereits ausgeführt. Neben dieser distanzierten Beobachtung sowie dem Heranzoomen in den beiden Teams habe ich die Relationierungen im Feld an bestimmten Orten vertieft erforscht. Auf Plattformen, wo sich die Community trifft; wo Datenjournalist*innen ausgebildet werden; anhand von Werkzeugen, die alle kennen; und vor allem durch Gespräche mit zentralen Akteur*innen des Feldes. Mein Ziel war, zu untersuchen, wie sich Datenjournalist*innen und ihre Praktiken verbinden und die Verbindungen sich verstetigen.

2.4.1 Datenerhebung

Die Analyse des Feldes und seiner Relationen lässt sich in zwei verschiedene Bereiche aufteilen: Interviews mit zentralen Akteur*innen sowie Beobachtungen an ausgewählten Verbindungsplattformen.

Interviews mit Akteur*innen

Aufgrund der Erkundung und Beobachtung des Feldes konnte ich zentrale Akteur*innen darin identifizieren. Als Knotenpunkte weisen sie besonders viele Verbindungen zu anderen Community-Mitgliedern, Medien- und Ausbildungsorganisationen auf. Über diese Personen will diese Arbeit die Relationierungen im Feld nachvollziehen. Dafür habe ich mit ihnen semistrukturierte, narrative Interviews durchgeführt (O'Reilly 2005:112ff; Czarniawska 2014:28ff; Silverman 2001). Diese Art Interview erlaubte mir, Ansichten zu bestimmen,

die immer gleichen Themen abzufragen und gleichzeitig Räume offenzulassen für die Überlegungen und Relevanzsetzungen der Gesprächspartner*innen.

Da die Erkenntnisse aus den Interviews in meiner Dissertation eine eher rahmende, das Feld skizzierende und auf Verbindungen hindeutende Rolle haben, erlaubte dies mir, die Gesprächsführung offen zu gestalten und eine Vielfalt an institutionellen Hintergründen in das Sample zu integrieren. Das Ziel ist dabei nicht, Variablen von bestimmten Personen oder Organisationen vollständig zu erfassen und sie vergleichbar zu machen, sondern ein reichhaltiges Bild des Feldes zu erhalten, welches auf bestimmte Schwerpunkte und gemeinsame Kategorisierungen hinweist.

Mein Fokus in der Datenerhebung lag in der deutschsprachigen Datenjournalismus-Community, insbesondere in der Schweiz. Während der Feld-Beobachtung war aber auch schnell klar: Die grossen Vorbilder waren Teams und Ausbildungsorganisationen in den USA. Eine Forschungsreise erlaubte mir 2019, mit zentralen Akteur*innen in New York und Washington D.C. sprechen zu können. Die Interviews fanden folgerichtig entweder in deutscher oder englischer Sprache statt.

Das Kriterium für die Auswahl und die Adressierung der Interviewpartner*innen war insofern eine Mitgliedschaft und gewisse Bekanntheit in der Datenjournalismus-Community. Die meisten angefragten Personen zeigten sich sehr offen für ein Interview.¹⁷ Zwischen November 2017 und Mai 2019 habe ich 18 Interviews durchgeführt, die zwischen 30 und 110 Minuten dauerten. Die Interviews fanden, wenn immer möglich, in einer persönlichen Begegnung statt, was schliesslich in 13 Interviews gelang. Fünf Interviews führte ich per Video durch. Zwei Interviews in amerikanischen Redaktionen fanden in je einer mehrköpfigen Runde statt: einmal mit drei, einmal mit vier Teammitgliedern aus derselben Redaktion. Von den insgesamt 23 Personen waren 17 männlich und 6 weiblich – dies widerspiegelt die auch in anderen Studien erwähnte Unterrepräsentation von Frauen im datenjournalistischen Feld (Usher 2016; Hermida und Young 2017; Heravi 2018).

Die Mehrzahl der interviewten Personen bezeichne ich in dieser Arbeit als Datenjournalist*innen. Obwohl sie selbst ihre Tätigkeit teilweise mit nuancierteren Bezeichnungen beschreiben und/oder auch Teamleitungsaufgaben

¹⁷ Ein (sehr subjektiver und unsystematischer) Vergleich mit anderen Forschungsprojekten lässt mich vermuten, dass sich das datenjournalistische Feld durch grosse Offenheit und Unterstützung für Forschende wie mich auszeichnet. Ich bin allen Akteur*innen für ihre Zeit und Gesprächsbereitschaft sehr dankbar.

wahrnehmen, so will ich mit dieser Bezeichnung darauf hinweisen, dass diese Person in einer Redaktion journalistisch tätig ist (in einem Fall handelte es sich um einen freischaffenden Datenjournalisten).¹⁸ Drei Interviews führte ich mit Verantwortlichen von Ausbildungsinstitutionen in der Schweiz (2) beziehungsweise den USA (1). Ein weiterer Interviewpartner war sowohl als Ausbildner (USA) als auch *Tool*-Entwickler tätig. Zwei Interviews führte ich mit Personen durch, welche in der Entwicklung von Datawrapper tätig sind, dem bekanntesten standardisierten Visualisierungstool im Datenjournalismus. Ein Interviewpartner war zum Zeitpunkt des Gesprächs der Betreiber einer internationalen datenjournalistischen Vernetzungsplattform (die inzwischen aufgelöst wurde). Ein Merkmal dieses Interview-Samples ist, dass sich die Personen in den beiden Community-Clustern (deutschsprachig beziehungsweise US-Ostküste) meist kennen, sehr oft sogar persönlich.¹⁹

Die Interviews habe ich zu fast voller Länge als Sprachaufnahme aufgezeichnet, jeweils mit dem Einverständnis der interviewten Person(en), und danach transkribiert. Falls relevante Gesprächsteile vor oder nach der offiziellen Aufzeichnung anfielen, habe ich diese danach aus dem Gedächtnis niedergeschrieben und als Hintergrundinformation bewertet. Ebenso habe ich, falls das Interview mit einem Besuch in einer Redaktion verbunden war, Beobachtungen über die Umgebung zusätzlich aufgezeichnet. Die Namen der interviewten wie auch ihre Arbeitgebenden habe ich in dieser Arbeit anonymisiert.

Die Interviews bestanden aus einerseits vorgegebenen Themenfeldern, die ich bei allen Personen abfragte (falls sie nicht von sich aus auf die Themen zu sprechen kamen), andererseits aus spontanen Nachfragen sowie zusätzlichen Fragen je nach institutionellem Hintergrund und Expertise der Person. Den

18 Die Personen arbeiteten in folgenden Redaktionen: Tamedia, Republik (beide Schweiz), Der Spiegel, Die Zeit (beide Deutschland), New York Times, Washington Post, Pro Publica, The Pudding (alle USA).

19 Während meiner Forschungsreise in den USA führte dies sehr oft zu folgender Situation: Zur Einleitung erzählte ich grob über mein Forschungsprojekt und dass ich aus der Schweiz angereist war. Zum Abschluss des Interviews, längst *off-the-record*, fragten mich die Gesprächspartner*innen, wie lange ich noch in den USA bliebe, wen ich noch treffen würde oder bereits getroffen hatte. Auf meine Antwort hin folgte dann die Aufforderung, den noch kommenden Interviewpartner*innen Grüsse mitzubringen beziehungsweise – im Falle von bereits vergangenen Interviews – die Bemerkung, dass man diese Personen von da und dort kenne und mit ihnen gemeinsam schon dieses und jenes Projekt durchgeführt habe.

Ideen aus den Konzepten des *zooming in and out* sowie der relationalen Ethnografie folgend, standen für die Feld-Ebene die relationierenden Elemente besonders im Fokus. Zu den fixen Themenfeldern gehören

- die persönliche Motivation;
- der eigene Weg ins datenjournalistische Feld;
- Werkzeuge, die man selbst verwendet;
- eigene Aktivitäten, um auf dem aktuellen Stand des Handwerks zu bleiben;
- die im Datenjournalismus benötigte(n) Expertise(n);
- die Bedeutung der datenjournalistischen Community;
- die Sicht auf das Feld und eine Einschätzung seiner Entwicklung;
- die Aufgaben und der Wert datenjournalistischer Praktiken;
- Vorbilder und beispielhafte Projekte für ›guten Datenjournalismus‹.

Im Falle von Datenjournalist*innen kamen insbesondere Fragen nach den eigenen Praktiken und Expertisen hinzu, bei Ausbildungspersonen Fragen nach dem datenjournalistischen Kanon. Personen mit Rollen in Verbindungsplattformen wie Verbänden oder Hacks/Hackers habe ich entsprechende Fragen dazu gestellt. Ein Musterfragebogen (für Datenjournalist*innen) findet sich im Anhang (F Interview-Leitfaden Expert*innen).

Beobachtung von Verbindungsplattformen

Im Rahmen der stetigen Feld-Beobachtung haben sich verschiedene Plattformen als Verbindungsorte herausgestellt: dort, wo sich Datenjournalist*innen persönlich treffen; wo sie Projekte diskutieren; wo Datenjournalismus gelehrt wird; wo sich Datenjournalist*innen fortbilden; wo datenjournalistische Geschichten Auszeichnungen erhalten. Zwar liessen sich diese Plattformen auch anhand der Interview-Aussagen charakterisieren und analysieren. Um jedoch auch nicht verbalisierbare Dynamiken zu erfassen, entschied ich mich, mit der direkten Beobachtung dieser Plattformen die Analyse um eine zusätzliche Perspektive zu erweitern.

In Übereinstimmung mit dem Konzept der relationalen Ethnografie (Abschnitt 2.2.3) habe ich wieder das ethnografische Beobachten als Methode gewählt – die vorherigen Abschnitte haben es bereits beschrieben. Aus praktischen Gründen habe ich die Verbindungsplattformen hauptsächlich in der Schweiz und im angrenzenden Ausland beobachtet. In der Forschungsphase beobachtete ich folgende Plattformen:

- Hacks/Hackers-Treffen der Sektion Zürich, 5 Abendtreffen (eines davon online) zwischen April 2017 und April 2020;
- Datenjournalismus-Seminar an der Universität Zürich (Herbstsemester 2017), 6 Seminar-Teilnahmen sowie Präsentationsfolien;
- International Journalism Festival 2018 in Bologna, diverse Sessions während zwei Tagen;
- European Investigative Journalism Conference (»DataHarvest«) 2018 in Mechelen, diverse Sessions während drei Tagen;
- Einführungskurs Datenjournalismus 2018 am MAZ Luzern (1 Tag), inklusive Folien;
- Besuch und Mittagessen im Unternehmen Datawrapper in Berlin, 2019;
- Netzwerk Recherche Jahrestreffen 2019, eine ausgewählte Session zu Datenjournalismus-Teams (Online-Aufzeichnung).

Von diesen Anlässen erstellte ich jeweils während oder direkt danach je eine ausführliche Feldnotiz, inklusive der bereits anlässlich der Team-Ethnografien erläuterten Randbedingungen (Erwartungen, Empfindungen, erste Einschätzungen). Im Sommer 2019 zeigte sich, dass nach zwei Praxisstudien, der Forschungsreise in die USA sowie der andauernden Feldbeobachtung in der Schweiz sich der Zustand der empirischen Sättigung (Czarniawska 2014:143ff) einstellte: Der abduktive Ansatz generierte nur noch wenige neue Hypothesen und stattdessen verfestigten sich die mit den theoretischen Überlegungen verzahnten Interpretationen. Mit einer Ausnahme²⁰ führte ich nach dem Sommer 2019 keine ethnografischen Beobachtungen mehr durch.

2.4.2 Datenauswertung

Als erster Schritt in der Auswertung der Relationen folgte die Selektion des Materials. Ausgehend von der Auswertung der ethnografischen Beobachtung in den Teams zeichneten sich relevante Analysekategorien – beispielsweise die Community und das Zusammenspiel der Expertisen – ab. Zusätzlich zog ich die bereits in ersten Memos vorgenommenen Feld-Einordnungen, die mit

20 Bei dieser Ausnahme handelt es sich um das im April 2020 (online) durchgeführte Hacks/Hackers-Treffen der Sektion Zürich, bei dem sich diverse für meine Forschung relevante Akteur*innen versammelten. Es ist dies die einzige ethnografische Beobachtung, die nach dem Ausbruch der Corona-Pandemie erfolgte und aufgrund der Community-Relevanz Eingang in diese Arbeit findet.

theoretischen Kategorien arbeiteten, sowie erste analytische Notizen über die Interview-Inhalte heran.

Im Sinne der relationalen Ethnografie sollte das Material verschiedene Perspektiven auf die Prozesse des Feldes aufweisen. Um es mit den ethnografischen Beobachtungen in den Teams sinnvoll verknüpfen zu können, war die gegenseitige Relevanzzuschreibung wichtig. Die meisten geführten Interviews und besuchten Anlässe des Feldes habe ich bereits nach diesem Kriterium für ein Gespräch beziehungsweise eine Beobachtung selektiert. Für die Analyse fielen dann die Feldnotizen der Sessionen an den beiden besuchten Konferenzen weg. Schliesslich verblieben 18 Interviews, die Feldnotizen der Hacks/Hackers-Treffen, die Feldnotizen und Folien aus den Ausbildungsinstitutionen, eine Feldnotiz über das Unternehmen Datawrapper sowie jene über die Session zu Datenjournalismus-Teams.

Dieses Feldmaterial habe ich in einem gemeinsamen atlas.ti-Projekt ausgewertet. Obwohl es sich um Material unterschiedlichen methodischen Zuschnitts – narrative Interviews, eigene Feldnotizen aus Beobachtungen, Dokumente – handelte, ist die gemeinsame Auswertung insofern opportun, als alles Material die Verbindungen und Wechselwirkungen zwischen Teams und Feld beleuchtet. Die Kodierung erfolgte, ganz im Sinne des *zooming in and out*, anhand der aus den Team-Untersuchungen bestimmten relevanten Kategorien: Expertise, Ideen über ›guten Datenjournalismus‹, Konventionen des Handwerks, Werkzeuge, Vernetzung. Die Kodes und Kategorien habe ich dann wiederum in Memos reflektiert und schliesslich das Feld aus theoretischen Perspektiven eingeordnet. Diese Einordnungen ergänzen das Kernstück dieser Dissertation, welches auf den umfassenden ethnografischen Beobachtungen in den Teams aufbaut.

2.5 Überblick

Die ganze Untersuchung ist gerahmt durch die Beobachtung des Feldes, eine offene Erkundung seiner Personen, Institutionen und ihrer Verbindungsplattformen (Ausbildungen, Newsletter, Handbücher, Twitter). Das Feld zeichnete ich in seinen engen Verflechtungen auf, was mir aus Ausgangslage für das Heranzoomen auf Personen und Organisationen diente. Wenn die Darstellung des Materials und der Selektionsentscheidungen in diesem Kapitel äusserst umfangreich und penibel erfolgt, so dient dies dazu, aufzuzeigen, in welchem

Kontext schliesslich aufgrund der Relationen zwischen den Ebenen generalisierbare Aussagen über den Datenjournalismus getroffen werden können.

Das Herzstück meiner Arbeit bilden die beiden Praxisstudien aus den Datenjournalismus-Teams NZZ Storytelling und BR Data. Dadurch habe ich auf die datenjournalistischen Praktiken von Organisationen herangezoomt, um sie vertieft zu erforschen: was sie bündeln, verstetigen und in die eigene Organisation übersetzen – und dies aus dem Blickwinkel zweier unterschiedlicher Positionen im Feld. Dafür habe ich einen ethnografischen Ansatz gewählt, der soziale Gegebenheiten im Alltäglichen, verbalisierten und nicht verbalisierten, untersucht.

Gleichzeitig habe ich herausgezoomt auf das gesamte Feld, um die Praktiken der Teams zu relationieren mit den Dynamiken und Diskussionen des Feldes. Diese stellen Verbindungen dar, Intermediäre, welche schliesslich das zu erklärende Phänomen ausmachen, indem sie organisationale Praktiken mit dem Feld verbinden. Um diese Relationen zu erforschen, habe ich zentrale, gut verflochtene Personen des Feldes mittels narrativer Interviews befragt und Verbindungsorte (Hacks/Hackers, Ausbildungskurse, Konferenzen, Werkzeuge) ethnografisch erforscht.

Da ich sowohl Organisationen wie auch das Feld durch den iterativen Ansatz gleichzeitig erforschte, dient die für diese Arbeit gewählte Chronologie – zuerst offeriere ich das Eintauchen und Einordnen in das Feld – rein der Vermittlung meiner Ergebnisse und der dafür aufgebauten *ethnographic tales*.

3 Datenjournalismus: theoretische und empirische Einordnung

Im Alltag Antworten auf die Frage, was Datenjournalismus ist, zu finden, verspricht vorerst nicht allzu schwierig zu werden. Schliesslich gibt es beispielsweise einen deutschen Wikipedia-Eintrag zu genau diesem Begriff.¹ Besucht man diesen, verweisen dann aber Indizien auf eine kompliziertere Definitionssituation. Abgesehen davon, dass der deutsche Eintrag seit mehreren Jahren kaum mehr aktualisiert wurde, gelangt man bei der automatisch verlinkten englischen Version auf »Data-driven journalism«.² Der erste Warnhinweis: »Not to be confused with Database journalism«. Der zweite: »This page is currently being merged. After a discussion, consensus to merge this page into Data Journalism was found.« Verschiedene Begriffe, die vermeintlich Ähnliches beschreiben – so klar scheint es plötzlich nicht mehr, was Datenjournalismus denn nun ist.

Konsultiert man die wissenschaftliche Literatur, so könnte man schnell die Übersicht verlieren. Obwohl der Begriff Datenjournalismus beziehungsweise data journalism erst seit rund zehn Jahren in Gebrauch ist, hat sich schon eine beachtliche Menge an Forschungsliteratur angesammelt. Und trotzdem scheint man sich nicht auf eine gemeinsame Definition geeinigt zu haben: In einer Meta-Studie zählen Ausserhofer et al. (2017:4f) in demjenigen Absatz, der sich mit der Definition befasst, 30 Studien auf, welche Datenjournalismus sehr verschieden fassen, das heisst verschiedene Aspekte betonen. Die Diversität an Definitionen ist gross und ist auch mehrere Jahre später nicht kleiner geworden. So schreibt Coddington (2019:230) 2019, nach einem knappen Jahrzehnt Forschung über Datenjournalismus:

1 <https://de.wikipedia.org/wiki/Datenjournalismus> (23.11.2020)

2 https://en.wikipedia.org/wiki/Data-driven_journalism (23.11.2020)

»Scholars have had a difficult time arriving at an agreed-upon definition for data journalism, which reflects the amorphous nature of the practice itself. Howard's (2014) is probably the most robust and straightforward, but even this is presented in a few parts: He defines data journalism as, ›gathering, cleaning, organizing, analyzing, visualizing, and publishing data to support the creation of acts of journalism‹ and also as the application of data science, or the extraction of knowledge from data, to journalism, as well the combination of treating data as a source, applying statistics to interrogate it, and using visualizations to present it (ibid.:4-5).«

Einem induktiven Ansatz folgend, verflieht diese Arbeit ethnografische Beobachtungen und Interview-Material mit theoretischen Rahmungen, um Datenjournalismus einzuordnen. Der erste Abschnitt (3.1) nutzt dafür den Begriff der Community, da er sich fast ausschliesslich auf Aussagen von und Beobachtungen über eine kleine Gruppe von Personen stützt, die sich kennen. Später erfolgt eine Öffnung des Blicks über den Begriff des Feldes, um Expertisen (3.2) und Narrative (3.3) über eine kleinere soziale Gruppe hinaus und auf vielfältigerer empirischer Grundlage zu fassen – gleichwohl anerkennend, dass die Community ein wichtiger Teil des Feldes ist und dass die Grenzen zwischen diesen Konzepten analytisch schwierig zu ziehen sind.

3.1 Die datenjournalistische Community

Die induktive Skizzierung des Feldes beginnt vor Ort. Dort, wo sich Datenjournalistinnen und Datenjournalisten treffen, wo sie teilen und diskutieren. Mehrere Abschnitte skizzieren die sozialen Beziehungen und ordnen sie im letzten Abschnitt theoretisch ein.

3.1.1 Hacks und Hackers: »bridge the two worlds«

Dosenbier und Pommes-Chips. Die günstigen aus dem Grossverteiler, in der Familienpackung. Die Community gibt sich betont unprätentiös: Hier soll Feierabend-Apéro-Stimmung sein, Formalitäten und sonst im beruflichen Umfeld gewohnte Statussymbole zählen wenig, und schliesslich ist das hier ein Low-Budget-Anlass – obwohl man sich in den offiziellen Räumlichkeiten grosser Medienhäuser trifft.

Mit dieser Rahmung beginnen die Treffen der Hacks/Hackers-Sektion Zürich. Bei Hacks/Hackers handelt es sich um ein weltweites Netzwerk von Schreiberlingen/Journalist*innen (»Hacks«) und Technikbegeisterten (»Hackers«), organisiert in lokalen Sektionen, deren Ziel nichts weniger als die »media revolution« und »rebooting journalism« ist.³ Das globale Netzwerk wurde 2009 gegründet⁴, die erste Veranstaltung der Sektion Zürich fand Ende 2013 statt.⁵ Seither hat sich die Gruppe 26 Mal getroffen (Stand Oktober 2020), über 1500 Personen zählen sich auf der Meetup-Plattform zum Netzwerk. Dort lässt sich folgender Beschreibungstext lesen:

»The worlds of journalists, designers and developers are coming together as reporting goes digital and Internet companies become media empires. Journalists call themselves ›hacks‹, someone who can churn out words in any situation. Hackers use the digital equivalent of duct tape to whip out code. Designers and hacker-journalists work to bridge the two worlds.«⁶

Das Dach der Hacks/Hackers-Anlässe wird breit gespannt und die offenen, kostenlosen und informellen Treffen senken die Zutrittsschwellen zum Netzwerk. So ist die Durchlässigkeit der Teilnehmenden aus Journalismus, Grafik/Design sowie Technologie sehr gross. Es gibt verschiedene weitere Plattformen, welche genauso versuchen, diese Expertisen zusammenzubringen und zu koordinieren. Dass Hacks/Hackers Zürich es schafft, sich über Jahre zu institutionalisieren, verweist auf einen stabilen, engagierten Kern an Mitgliedern, der an den meisten Anlässen teilnimmt.

-
- 3 Die Zitate sind der Startseite des Netzwerkes entnommen: <https://www.hackshacker.s.com/> (03.10.2020)
- 4 Für weitere Details zur Geschichte des weltweiten Hacks/Hackers-Netzwerkes, siehe Lewis und Usher 2014 sowie direkt <https://www.hackshackers.com/about/history/> (03.10.2020)
- 5 Sämtliche Aktivitäten der Zürcher Sektion von Hacks/Hackers lassen sich auf der Meetup-Plattform verfolgen, über die sie sich organisiert: <https://www.meetup.com/de-DE/Hacks-Hackers-Zurich> (03.10.2020)
- 6 Einleitungstext der Hacks/Hackers Zürich Meetup-Seite: <https://www.meetup.com/de-DE/Hacks-Hackers-Zurich> (05.10.2020)

3.1.2 Die ›Ankommenshalbstunde‹: informell verbinden und kooperieren

Während der Forschungsphase (April 2017 bis April 2020) besteht dieser Kern aus Organisator*innen und regelmässigen Teilnehmenden und grenzt sich von unregelmässig Teilnehmenden und nur punktuell Interessierten ab. Der Kern besteht hauptsächlich aus Personen, die in den datenjournalistischen Teams in den grossen Deutschschweizer Medienhäusern SRF, NZZ und Tamedia arbeiten. Man kennt sich in dieser Gruppe persönlich und weit über die Arbeit hinaus. So beginnen die Treffen meist ähnlich: Wer sich kennt, ist in der ›Ankommenshalbstunde‹ – an jedem Treffen öffnen die Türen eine halbe Stunde vor dem Beginn der Vorträge und dieses Zeitfenster ist gemäss Zeitplan reserviert für »Beers&chats⁷ – verwickelt in Smalltalk mit den Kolleginnen und Kollegen aus dem Kern. Dazu ganz oft ein Dosenbier in der Hand, das man sich selber aus einem Kühlschrank, der irgendwo im Raum steht, holte. Und in dessen Nähe steht eine Familienpackung Pommes-Chips. Die bereitliegenden Plastikteller warten dabei oft vergebens darauf, ihrer Aufgabe – befüllt mit Chips den Weg an einen anderen Ort des Raumes zu finden – nachzukommen; man isst direkt aus der Packung.

Die Betonung von Lockerheit und Informalität zelebriert einerseits die Abgrenzung von Althergebrachtem und Traditionellem – schliesslich pflegt man die eigene Rolle am Rande des journalistischen Feldes: zwar meist in den Sälen und Seminarräumen der grossen Medienhäuser, aber deutlich abgegrenzt von den ›steifen Chefetagen‹.⁸ Andererseits soll die Informalität dazu beitragen,

7 Ein Beispiel für einen typischen Zeitplan: <https://www.meetup.com/de-DE/Hacks-Hackers-Zurich/events/264695029/> (05.10.2020)

8 Dies zeigt sich besonders offensichtlich, als sich der Rahmen für eine einzige Hacks/Hackers-Veranstaltung ändert: Als Alberto Cairo Ende Juni 2018 in Zürich gastiert, scheint dies für den Tamedia-Chef Pietro Supino und sein Unternehmen eine wichtige Sache zu sein. So nimmt er selbst teil und tauscht sich beim Smalltalk persönlich mit Cairo aus. Statt Dosenbier vor und nach dem Vortrag gibt es nur nach dem Vortrag Drinks (Bier und Wein), dafür dazu Sandwiches, Stehtische mit Tischtüchern sowie Personal, welches bedient. Trotzdem ist die Teilnahme auch bei diesem Anlass kostenlos, das Unternehmen Tamedia offeriert. Die formelleren Rahmenbedingungen sowie das Fehlen einiger Hacks/Hackers->Stammgäste< bewirken eine veränderte soziale Rahmung: Der Apéro erinnert nun an konventionelles Business-Netzwerken: Die professionelle Status-Präsentation der Teilnehmenden überlagert die sonst gut sichtbaren informellen Verbindungen der Kern-Mitglieder. –46_20180629_Hacks-Hackers-Cairo

gen, Statusunterschiede zu überdecken und einen Austausch auf Augenhöhe zu etablieren. Jede*r bedient sich schliesslich selbst aus dem Kühlschrank mit einem Dosenbier. Und man erlaubt sich auch mal einen Witz über die Macken der Kolleg*innen.⁹

Dahinter steht die Idee der Kooperation, welche tief in der Community verankert ist und den Grundgedanken hinter dem Netzwerk bildet: Man unterstützt sich bei offenen Fragen, man teilt Daten und Methoden und die Bemühungen, auf der eigenen Redaktion immer wieder zu erklären, was es nun mit *diesen Daten* auf sich hat. Ein Datenjournalist erklärt exemplarisch:

»Also, also man kennt sich untereinander, weil es so wenige sind, man hat dieselben Probleme gehabt, irgendwie, ehm, überhaupt irgendwie einen Anfang finden, man hat dieselben technischen Herausforderungen, ehm, viele berichten offen darüber auf Twitter, man schreibt sich hin und her, und eh hilft sich im Zweifelsfall oder kann einen Kollegen fragen wie der es eigentlich damals gemacht hat, was für Daten stecken dahinter, und wie habt ihr die Geschichte verkauft und so. Also das ist super angenehm, ehm, ich, ich hab kein Problem damit eigentlich jeden, jeden aus dem Bereich jederzeit einfach anzuhauen und zu fragen wenn ich, ehm, wenn irgendwas offen ist. Ehm, das ist... nen sehr kollegialer Austausch, klar ist man irgendwie auch im Wettstreit, so, weil man für konkurrierende Medien arbeitet, ehm, aber, ich glaube, das wäre jetzt meine Einschätzung, gegenüber dem restlichen Journalismus, noch offener seinen Kollegen aus anderen Medienhäusern zu zeigen, wie man arbeitet.«¹⁰

Im (deutschsprachigen) Datenjournalismus sind die Hacks/Hackers-Treffen wichtige Austauschplattformen. Sie bieten viel Raum für das persönliche Kennenlernen und damit den Aufbau von zwischenmenschlichem Vertrauen, welches für die kooperative Kultur von grossem Wert ist.

⁹ beispielsweise: 11_20171220_SeminarUZH-6

¹⁰ Interview D12.

3.1.3 Bedeutungsvolle Visualisierung: die Rolle eines Design-Studios

Auch wenn das personelle Einzugsgebiet mit über 1500 Personen sehr gross ist, so fällt auf, dass sich der Kern der Community in der Forschungsphase nicht allzu stark verändert hat.¹¹ Im Organisationsteam der Hacks/Hackers-Sektion befindet sich seit den Anfangstagen der Mitgründer eines grossen Design-Studios aus Zürich. Seit Jahren sind das Studio und einige Mitarbeitende persönlich und beruflich eng verflochten mit der datenjournalistischen Community und sie haben in vielen grossen Projekten mit den datenjournalistischen Teams der grossen Medienhäuser zusammengearbeitet. So erzählt eine Datenjournalistin, welche sich seit Jahren in der Community engagiert:

Da haben wir weiter mit ihnen [Design-Studio] zusammengearbeitet, ich bin dann von [grosser Schweizer Medienorganisation] zu [andere grosse Schweizer Medienorganisation] und hab da wieder 'n Team aufgebaut, und auch da haben wir mit [Design Studio] zusammengearbeitet. Ehm... [Design-Studio] hat auch schon hier im Haus mitgearbeitet, wir sind was, wir wollen was Gemeinsames wieder entwickeln, ehm... Das sind, das sind auch einfach Weggefährten, ne. Die... die mich, von denen ich ganz viel gelernt hab. Nur ich glaub dass die *massgeblich* dran beteiligt sind, wie sich das hier entwickelt hat, in der Schweiz.¹²

Viel grafische und technische Expertise des datenjournalistischen Feldes sammelt sich in jenem Design Studio und macht es zu einem Akteur, der immer wieder für grosse datenjournalistische Projekte hinzugezogen wird und den visuellen Aspekten darin viel Gewicht gibt. Die Verankerung des Studios in der Hacks/Hackers-Veranstaltungsreihe geht einher mit einigen Anlässen, die Datenvisualisierungen und die Konventionen, welche sie anleiten, thematisieren. Die Mitarbeitenden des Studios nehmen denn auch regelmässig an der Veranstaltungsreihe teil. Obwohl weder *Hack* noch *Hacker*, geniesst grafische Expertise ein grosses Gewicht in der datenjournalistischen Community, wie sich aus der Hacks/Hackers-Beschreibung herauslesen lässt – welche explizit auf die »Brückenfunktion« von Designern verweist (siehe Zitat Fussnote 6, Kapitel 3).

¹¹ Es wäre interessant zu beobachten, ob das auch so bleibt – schliesslich sind während des Jahres 2020 bei einem grossen Teil des Organisationsteams der Sektion grössere Veränderungen in der Arbeitssituation eingetreten (Stellenwechsel, insbesondere aus den Redaktionen heraus; Wegzug aus der Schweiz; Selbstständigkeit).

¹² Interview D19

Neben der Bedeutung der grafischen Expertise verweist das Zitat darauf, dass man für die Zusammenarbeit auf persönliche Bekanntschaften zurückgreift – so wie man in verschiedenen Redaktionen immer wieder mit demselben Design Studio zusammenarbeitete. Im datenjournalistischen Feld erfolgen Rekrutierungen nicht immer, aber doch immer mal wieder über persönliche Kontakte. Dies betrifft Rekrutierungen für Arbeitsstellen, wo das Portfolio an Projekten sowie die teilweise öffentliche Reflexion darüber – beispielsweise auf Blogs – eine Rolle spielt. Aber auch Dozierenden-Tätigkeiten und Gast-Seminare, welche in der Schweiz hauptsächlich Praktiker*innen leiten, werden über persönliche Bekanntschaften aus der Community vergeben¹³ (mehr zu den Ausbildungen in Abschnitt 3.2.1). Da viele Fäden des datenjournalistischen Feldes in der Schweiz bei dieser Kern-Community zusammenlaufen, liegt es auf der Hand, dass deren kulturelle und praktische Hintergründe, Vorstellungen und Ideen aktuell praktisch alle Facetten des datenjournalistischen Feldes in der Deutschschweiz prägen.

3.1.4 Flaggschiffe ennet des Atlantiks: über die Grenzen hinweg

Während die Beobachtungen aus den Hacks/Hackers-Treffen sowie den Ausbildungen hauptsächlich die Deutschschweizer datenjournalistische Community beleuchten, so lässt sich dieser relativ kleine geografische Ausschnitt doch gut mitten in einer räumlich weitläufigen, internationalen Community verorten. Nicht nur aus den Aussagen der Schweizer, sondern auch denjenigen der deutschen Protagonist*innen wird deutlich: Die Community sieht die grossen Redaktionen in den USA, hauptsächlich die New York Times (NYT), als Taktgeberinnen des Feldes. Immer wieder referenziert die Community neuartige Werkzeuge der *Times* und sie diskutiert intensiv deren Methoden der Analyse und Visualisierung von Daten.¹⁴ Diese Diskussionen folgen nicht nationalen

13 So begegnet man dem Mitgründer des Design Studios auch wieder als Dozent im Programm des Datenjournalismus-Kurses an der Universität Zürich.

14 Die NYT ist die meistreferenzierte Medienorganisation in der Community: So hat sie »brilliant people« (Interview D5); eine vorbildhafte Organisationsstruktur durch Vereinigung von Grafik- und Interaktiv-Team (Interview D6); liefert regelmässig den datenjournalistischen-Höhepunkt des Jahres (Interview D11); hat funktionierende Darstellungsmethoden (Interview D11); hat eine gewichtige Frontseite (Interview D14); hat ein innovatives Tech-Team (Interview D14); ist diejenige Organisation, wo man als Datenjournalist*in arbeiten will (Interview D15); hat ein fantastisches Grafik-Team (Interview D17), das auf Augenhöhe mit anderen Ressorts mitreden kann (Interview D3), un-

oder sprachlichen Grenzen, sondern spannen sich im amerikanisch-englisch-westeuropäischen Raum auf. Eine grosse Rolle dabei spielt das soziale Netzwerk Twitter:

»Tatsächlich such ich auf Twitter, ehm... und dort hab ich... auch schon, schon jahrelang nen sehr... sehr... eh handverlesene, eh oder sehr ausgewählte Zahl von Accounts, denen ich einfach folge, ja. Also ich hab da das Prinzip, dass ich eigentlich nur Leuten folge, wo ich... das Gefühl hab da kommen regelmässig Sachen, die mich irgendwie inspirieren. Ehm, und das sind am Ende... schon auch viele Kollegen, ehm, aus dem internationalen Umfeld dann so, ja, die zu denen halt alle irgendwie aufschauen, also über die New York Times, über die Washington Post, über FiveThirtyEight. Dann Pudding Viz, die coole Sachen machen, also dass, die, ja, die halt auch überall die, die Preise gewinnen, die fetten Projekte machen und die grossen Teams haben und so.«¹⁵

Egal ob man Community-Mitglieder direkt fragt oder sich die Verweise auf ihren Ausbildungsskripten anschaut: Twitter ist omnipräsent und die erste Wahl als Quelle von Inspiration, von Debatten über Praktiken oder Projekte, und als selbst genutzte Plattform zur Präsentation von Projekten oder Information über diejenigen anderer. Zhang (2018) untersuchte die Twittersphäre des Datenjournalismus, indem er rund 7000 Tweets einer Textanalyse unterzog. Dabei zeigt sich eine multiperspektivische Sphäre, in der sich Themen-Kategorien nur schwer abgrenzen lassen und die verschiedenartige Nutzer*innen umfasst: Praktiker*innen, Akademiker*innen, Blogger*innen, publizistische und zivilgesellschaftliche Organisationen. Der Inhalt der Tweets bewegt sich rund um die Themen Journalismus, Datenvisualisierung und Datenanalyse-Techniken. Als Netzwerk abgebildet, verweist die Hashtag-Kookkurrenz-Analyse auf organisierende Kategorien,

»[...] such as ›#ddj‹ (data-driven journalism), ›#dja‹ (the annual held Data Journalism Award), ›#dataviz‹ (an abbreviation and homophonic of ›data visualization‹), ›#dd‹ (data-driven), and ›#ogp‹ (Open Government Partnership). These terms are mostly located in the central positions in the hashtag

ter anderem weil der Grafikleiter in der Chefredaktion sitzt (Interview D18); produziert »compelling stories«, weil sie oft Personalisierung als narrative Strategie wählt (Interview D17); ist ein Vorbild mit ihrer Datenvisualisierungsexpertise (6_20171108_SeminarUZH-3) und ihrer Toolbox (11_20171220_SeminarUZH-6).

15 Interview D12

co-occurrence network, indicated by their high degree of centrality (highly connected to other hashtags) and high betweenness [sic!] centrality (located in the bridging positions linking other hashtags). They act as important topic organizers and markers facilitating the discussion of data-driven journalism in the Twittersphere.« (Zhang 2018:753)

Unabhängig von der Plattform: Das Beobachten der anderen Community-Mitglieder dient dazu, sich stets über die neuesten methodischen und erzähltechnischen Ideen und Werkzeuge zu informieren. Zur Verständigung über die Sprachgrenzen hinweg dient Englisch – und es kommt gar nicht selten vor, dass nicht-englischsprachige Datenjournalismus-Teams ihre Projekte auf Englisch übersetzen, um die Verbreitung in der Community, insbesondere über Twitter, zu ermöglichen.¹⁶ Die internationale Vernetzung führt dazu, dass die ressourcenstarken datenjournalistischen Teams der amerikanischen Flaggenschiff-Medienhäuser mit ihren »fetten Projekten« (siehe Zitat oben) stets als Vorbild und Vergleichsfolie mitschwingen – und mit ihren methodischen und narrativen Konventionen das datenjournalistische Feld weltweit mitprägen.

3.1.5 Datenjournalismus als interpretative Community

Die US-amerikanische Journalismusforscherin Barbie Zelizer leitete 1993 eine ihrer Arbeiten folgendermassen ein:

»This article suggests that the notion of ›profession‹ may not offer the most fruitful way of examining community among American journalists. It proposes viewing journalists as members of an interpretive community instead, one united by its shared discourse and collective interpretations of key public events.« (Zelizer 1993:219)

¹⁶ So übersetzten zwei deutsche Medienhäuser ihre (später) preisgekrönten Projekte: BR Data dasjenige zum Schufa-Algorithmus (<https://web.br.de/interaktiv/erhoehtes-risiko/english/index.html>, 07.10.2020) sowie Die Zeit ihres über die Abwanderung aus der ehemaligen DDR seit dem Mauerfall (<https://www.zeit.de/politik/deutschland/2019-05/east-west-exodus-migration-east-germany-demography>, 07.10.2020). In beiden Fällen lag das Interesse der internationalen Community nicht primär beim Thema, sondern beim ›Handwerk‹ der Sammlung/Auswertung beziehungsweise Visualisierung der Daten. Für diese Aspekte wurden die beiden Projekte auch ausgezeichnet.

Zelizer illustriert ihre theoretischen Überlegungen mit Beobachtungen, die schon länger her sind, nämlich aus den USA der Sechziger-, Siebziger- und Achtzigerjahre. Dieser Abschnitt zeigt auf, dass die theoretische Rahmung – von Zelizer aus der Anthropologie, der Volkskunde und der Literaturwissenschaft abgekupfert – jedoch für die datenjournalistische Community viele Jahrzehnte später anschlussfähig ist.

Das in den vorangegangenen Abschnitten beschriebene Netzwerk des Datenjournalismus, die darin stattfindenden Diskussionen und gegenseitigen Vergleiche, das Übernehmen von Ideen anderer – dies sind gemäss Zelizer zentrale Charakteristika interpretativer Communitys, die verbunden sind durch eine geteilte Sinngebung der Wirklichkeit und sich daraus entwickelnde Bewertungsstandards:

»They establish conventions that are largely tacit and negotiable as to how community members can ›recognize, create, experience, and talk about texts‹ [...] These views suggest that communities arise less through rigid indicators of training and education – as indicated by the frame of the profession – and more through the informal associations that build up around shared interpretations.« (Zelizer 1993:223)

Diese geteilte Sinngebung bedingt Plattformen und Kanäle, um den Diskurs aufzuspannen. Zelizer (1993:223) schreibt: »Through channels like informal talks, professional and trade reviews, professional meetings, autobiographies and memoirs, interviews on talk shows and media retrospectives, they [journalists, Anm. RE] create a community through discourse.« In der datenjournalistischen Community des 21. Jahrhunderts haben sich diese Kanäle ergänzend ins Internet verlegt und internationalisiert – auf Twitter wurde bereits verwiesen, daneben existieren viele weitere Online-Plattformen, siehe Fussnote 2 in Kapitel 2. Persönliche Begegnungen spielen für den Aufbau informeller Netzwerke aber immer noch eine entscheidende Rolle, wie bereits anhand der Hacks/Hackers-Sektion Zürich aufgezeigt wurde. Eine weitere Plattform, die online und offline in der internationalen Community verbindet, sind Berufsverbände und ihre Konferenzen. Die grösste, die NICAR¹⁷, versammelt jährlich Datenjournalist*innen aus der ganzen Welt und ermöglicht so

¹⁷ NICAR ist die Abkürzung für ›National Institute for Computer-Assisted Reporting‹, ein Programm des amerikanischen Berufsverbandes Investigative Reporters and Editors (IRE). <https://www.ire.org/nicar/about> (03.11.2020)

einen geteilten Diskurs über und geteilte Praktiken zur Produktion datenjournalistischer Erkenntnisse (Hermida und Young 2019:49ff). Bradshaw (2018:7) verweist zudem auf die NICAR *mailing list*, welche Datenjournalismus exemplarisch als »community of practice« (Wenger 2008) charakterisiere: Mehrere hundert Personen – wer eingetragenes Mitglied ist, bleibt unbekannt – unterhalten sich über diese Liste über datenjournalistische Praktiken und fragen nach methodischen Tipps. Weitere Verweise auf die zentrale Stellung von NICAR für die Entwicklung des Datenjournalismus finden sich bei Anderson (2015), Coddington (2015) und Usher (2016); ein kurzer historischer Rückblick bei Hermida und Young (2019:32f).

Abschnitt 3.1 hat die datenjournalistische Community eingeführt und gezeigt, wie in der Forschungsphase eine eng verflochtene, auf persönlicher Bekanntschaft sowie informeller Kooperation beruhende und gleichzeitig international orientierte Community einen Raum der geteilten Sinngebung schuf. Insbesondere in der Deutschschweiz hat ein Kern von einem guten Dutzend Personen das Feld geprägt. Dieser Kern der Community formte mit seinen Expertisen und Narrativen die geteilte Sinngebung – genau deshalb entwickelte Zelizer ihre theoretische Rahmung der interpretativen Community: »We need a frame that might explain journalism by focusing on how journalists shape meaning about themselves« (Zelizer 1993:222).

Was in den interpretativen Communitys zirkuliert, sind Narrative, Erzählungen, welche die Mitglieder versammeln: »They [journalists, Anm. RE] come together not only through training sessions, university curricula, or formal meetings, but through stories that are informally repeated and altered as circumstances facing the community change« (Zelizer 1993:234). Diese Narrative über das eigene Feld leiten die Sinngebung an – Abschnitt 3.3 zeigt auf, welche Erzählungen und Ideen im datenjournalistischen Feld zirkulieren.

Um gemeinsam interpretieren und erzählen zu können, bedingt es einer gemeinsamen Sprache. In dieser Arbeit gilt es deshalb, Zelizers Überlegungen um diesen Aspekt zu erweitern. Denn Datenjournalismus »bridges the two worlds«, wie es Hacks/Hackers beschreibt, bringt also verschiedene Welten und ihre Expertisen zusammen – die dann den gemeinsamen Aushandlungsraum für Praktiken und Narrative bilden. Vor dem Aufzeigen der Narrative geht deshalb der Abschnitt 3.2 der Frage nach, welche kulturellen, epistemologischen und materiellen Aspekte von verschiedenen Expertisen das datenjournalistische Feld versammelt.

3.2 Datenjournalismus als versammelte Expertisen

Datenjournalismus – genauso wie Journalismus generell – hat den Anspruch, valide Erkenntnisse über die Welt zu generieren und als Evidenz zu befestigen. Dies ist nicht nur ein technischer oder intellektueller Akt, sondern ist abhängig von den von Communitys geschaffenen Versammlungs- und Sprachräumen und den Narrativen, die darin zirkulieren – Zelizer hat diese sozialen Gebilde als »interpretive communities« beschrieben, wie der vorherige Abschnitt ausführte. Rund 15 Jahre nach dem Original-Artikel setzte Zelizer in einem Buchkapitel ihr theoretisches Konzept folgendermassen in Bezug zur Produktion von Erkenntnissen (»knowledge«):

»The idea of interpretive communities [...] helps situate the strategies that go into the sharing of knowledge as integral to the knowledge that results. Recognizing that groups with shared ways of interpreting evidence shed light on the way that questions of value are settled and resettled, the persons, organizations, institutions and fields of inquiry engaged in journalism's analysis become central to understanding what journalism is. [...] Inquiry, then, is not just an intellectual act but a social one too.« (Zelizer 2009:30f)

Es stellt sich also die Frage, welche Bewertungskriterien, kulturelle und epistemologische Vorstellungen und professionelle Standards die Feldakteur*innen in ihren »inquiries« versammeln, wie diese mit den technologischen Werkzeugen interagieren und welche Basis so für die Aushandlung von Narrativen und Praktiken entsteht. Die folgenden Abschnitte gehen dieser Frage nach, indem sie von empirischen Beobachtungen und Aussagen angeleitet verschiedene Aspekte des Feldes ausleuchten und schliesslich in einer zusammenfassenden Systematisierung münden (Abschnitt 3.2.4). Viele ethnografische Beobachtungen basieren im folgenden Abschnitt – wie schon in 3.1 – auf der datenjournalistischen Community in der Deutschschweiz. Die Interviews mit Akteur*innen aus der internationalen Community zeigen aber, dass die Charakteristika aus dem Lokalen sich auf sehr viele weitere Orte übertragen lassen.

3.2.1 Kulturelle Rahmung: Ethos, Mindset und Open-Kultur

Die folgenden Abschnitte skizzieren die Kultur des datenjournalistischen Feldes mit den Einflüssen der verschiedenen Expertisen. Sie sind strukturiert nach drei Aspekten: erstens die Grundhaltung auf persönlicher Ebene, das Ethos des gemeinsamen Lernens; zweitens darauf aufbauend die Aus-

bildungsphilosophie; und drittens der kulturelle Kontext, konkret die *Open-Kultur*.

Das Ethos des gemeinsamen Lernens

Zur Grundhaltung in der Community gehört die persönliche Überzeugung, dass die verschiedenen Expertisen, welche die Teamarbeit oder der Community-Austausch zusammenführen, und eine multiperspektivische Aushandlung die datenjournalistische Praxis verbessern. Exemplarisch zeigt sich dies in der Institutionalisierung von Ideen-*Pitches* innerhalb der datenjournalistischen Teams – so auch bei The Pudding, wie ein Teammitglied beschreibt. Das Team versammelt sich regelmässig über GoogleHangout und einzelne Mitglieder präsentieren ihr »proposal« für mögliche neue Geschichten. Dem folgt ein kollaborativer und gleichzeitig kompetitiver Prozess des Konzipierens:

»And so the, the group that goes one by one and asks questions about the proposal. And once that's done, eh, again one by one people give the reactions to the proposal. And, eh, in the end of that... the idea is to give in three ideas per meeting. And people pick their favorite idea out of those three. Ehm, that doesn't mean that this is the idea you're gonna choose, because again, you have... freedom to explore what you wanna explore. But, it's good to engage people's enthusiasm and to engage that perspectives and reactions and potential, [...]. So, so that's, that's kind of a communal brain storming almost.«¹⁸

Kooperation und gemeinsames Lernen verbessern aber nicht nur datenjournalistische Praktiken, sondern sind der Weg zur Weiterentwicklung der eigenen Fähigkeiten. Zugespitzt findet sich dieses Ethos auf einer Folie der Datenjournalismus-Ausbildung der Universität Zürich, welche das Selbstbild der Datenjournalist*innen von den Journalist*innen abgrenzt und es unter andrem beschreibt mit den Begriffen »Teamplayer« (gegenüber »Einzelkämpfer«) und »mit anderen lernen« (gegenüber »weiss alles«).¹⁹ Im Datenjournalismus tätig zu sein bedeutet in der Selbstbeschreibung des Feldes, dass man die Motivation mitbringt, stetig hinzuzulernen – einerseits weil sich die Techniken schnell verändern, andererseits aber auch aus persönlicher Bescheidenheit gegenüber der Expertise anderer, auf die man angewiesen ist. Man lernt *on the job*, Arbeit bedeutet im Datenjournalismus ein kontinuierlicher Lernprozess.

18 Interview D17

19 UZH_1004_Folie60

Das Lernen soll dabei betont *projektorientiert* erfolgen: Das bedeutet, dass jedes Projekt eine Gelegenheit ist, dazuzulernen. Deshalb streben viele Community-Mitglieder danach, stets Neues auszuprobieren, wie es dieser Datenjournalist beschreibt:

»You have to say like: You have to, you have to teach yourself. But you just have to make sure that you are finding projects where you continue to push your own technical boundary.«²⁰

Neben dieser Grundhaltung der ›Entdeckung‹ neuer Erkenntnisse durch die Verschiebung der technologischen Grenzen ist es für die meisten Praktiker*innen attraktiv, von Projekten anderer zu lernen, weil die Resultate der Bemühungen sichtbar sind: An den Konferenzen und über Twitter präsentieren die Teams die Erfolgsgeschichten ihrer Arbeit. Als wichtigste Referenz für diese Best Practices erscheinen wiederum überdurchschnittlich oft amerikanische Projekte, insbesondere diejenigen der New York Times (siehe Zitat Fussnote 15 in Kapitel 3).

Im datenjournalistischen Feld sind verschiedene Formate für die Präsentation der eigenen Projekte und deren Peer-Review institutionalisiert. Bei grösseren Teams im deutschsprachigen Raum besteht beispielsweise die Erwartung, Daten, Methoden und Hintergründe auf der Sharing-Plattform Github zu veröffentlichen.²¹ Weiter gehört es zu den üblichen Programmfpunkten der Hacks/Hackers-Treffen wie auch von Konferenzen, aufwendige, vom Feld als exemplarisch bewertete Projekte – in der Regel als *Leuchtturm*-Projekte bezeichnet – zu diskutieren. Der Fokus liegt dabei auf den technischen Details, den Methoden. Dies wiederum schafft Anreize, insbesondere für grosse Datenjournalismus-Teams, neue Technologien und Vorgehensweisen zu verwenden und so neben den individuellen *technical boundaries* auch gleich diejenigen des Feldes zu *pushen*. Eine weitere Plattform zur Sichtbarmachung und Diskussion von datenjournalistischen Projekten sind Preisverleihungen wie die Data Journalism Awards. Diese geniessen im Feld eine grosse Beachtung und rücken immer wieder methodisch innovative Pro-

20 Interview D15

21 Interview D2; Interview D16; 6_20171101_SeminarUZH-2 – der Gesprächspartner im Interview D12 lobt das Schweizer SRF-Datenjournalismus-Team als »vorbildlich«, weil es konsequent die komplette Methodik auf Github stelle.

ekte in den Fokus, welche als Goldstandard datenjournalistischer Praktiken gelten können (Loosen et al. 2017; Ojo und Heravi 2018; Young et al. 2018).²²

Die Präsentation eigener Projekte darf aber nicht zur Nabelschau ohne Reflexion der gemachten und gemäss dem Ethos nie versiegenden Lernprozesse werden. Insbesondere wenn ein Projekt in einem offenen Diskussionsraum wie Hacks/Hackers vorgezeigt wird, soll es Gelegenheit geben, um Schwächen oder alternative Vorgehensweisen zu thematisieren.²³

Ausbildung zur »Community-Fähigkeit«

Aufgrund der beschriebenen Community und ihres Ethos gilt für Datenjournalismus noch ausgeprägter, was Carlson (2018:1758) für den Journalismus insgesamt konstatierte: »Journalism has never attained the clear-cut boundaries associated with professional status. Journalists, in the United States and many other nations, lack the mandatory educational requirements, credentialing practices, interpersonal encounters, legal protections, and esoteric knowledge that mark off the classic professions.« Entsprechend vielfältig sind die beruflichen Wege und Bildungshintergründe, welche in das datenjournalistische Feld führen.

Gerade in der Deutschschweiz, welche grosses Gewicht auf praxisorientierte Berufsbildung legt, sind es hauptsächlich Praktiker*innen des Feldes,

22 Das Gewicht von Preisverleihungen im Datenjournalismus wird unterschiedlich hoch eingeschätzt. Während US-amerikanische Interviewpartner*innen – auch preisgekrönte – die Bedeutung eher herunterspielen, scheint die Beachtung in Europa grösser. Vereinzelte Personen weisen darauf hin, dass in Datenjournalismus, Datenvisualisierung und Journalismus insgesamt (zu) viele Preise vergeben werden. In den letzten Jahren galten die Data Journalism Awards des Berufsverbands Global Editors Network (GEN) als wichtigster Preis im Feld. Allerdings stellte das GEN seinen Betrieb im Herbst 2019 ein, inklusive der Data Journalism Awards. Kurzfristig riefen verschiedene Akteur*innen des Feldes Anfang 2020 die Sigma Awards ins Leben (<https://datajournalism.com/awards>, 03.11.2020). Es wird interessant zu beobachten sein, ob die Sigma Awards und Preisverleihungen insgesamt auch in Zukunft das beziehungsweise ein datenjournalistisches Feld zu versammeln und zu prägen vermögen.

23 Dass ›Shows‹ mit einem hohen Inszenierungsgrad nicht der Erwartung entsprechen, liess sich wiederum an der bereits in Fussnote 8 (Kapitel 3) beschriebenen Hacks/Hackers-Veranstaltung beobachten, die durch den Besuch der Szenegrosse Alberto Cairo eine andere Rahmung erfuhr. Als ich eines der Kern-Mitglieder der Community nach dem Vortrag frage, was sie davon mitnehme, meint sie mit kritischem Unterton, dass Cairo »halt schon sehr amerikanisch« und »ein guter Verkäufer« sei. – 46_20180629_Hacks-Hackers-Cairo

welche die wenigen datenjournalistischen Ausbildungen nebenberuflich konzipieren und lehren. Auch die Verantwortlichen in den Ausbildungsinstitutionen verfügen über Verflechtungen mit der Community und rekrutierten entsprechend ihre Bekanntschaften für Dozierenden-Aufgaben.²⁴ Das Ziel einer praxisorientierten datenjournalistischen Ausbildung ist denn auch nicht, nach dem Abschluss datenjournalistische Werkzeuge perfekt anwenden zu können, sondern die »Community-Fähigkeit«, wie es ein Bereichsleiter der Schweizer Ausbildungsinstitution MAZ nennt:

»Es ist ganz entscheidend, und das hat jetzt [Ausbildungsleiter*in] auch mit dieser Ausbildung probiert, dass man Teil ist dieser Community, die international ausgerichtet ist, die auch im Internet stattfindet, auf diesen Plattformen, wo man sich trifft an Hack-Days oder an internationalen Konferenzen. Wo man Leute trifft, die gleich ticken, vielleicht auch mal ein gemeinsames Projekt macht, über die Grenzen hinweg zusammenarbeitet. Ja, ich glaube das muss dort stattfinden, in der Community. Und nur bedingt an einer klassischen Ausbildungsstätte. Weil sich das so schnell bewegt und innerhalb der Community so schnell weiterentwickelt. Man kann nur über die Ausbildung die Leute dorthin führen, den Zugang schaffen, dass man weiß, wie man sich bewegt, wie man Code-Probleme löst in der Community, wie man Fragen stellt, wie man es formuliert, was man sucht, so dass es jemand aus der Community versteht, was man meint, und einem dann auch hilft, was man wieder einbauen kann. So wie ›community-fähig‹ werden, das kann eine Ausbildung leisten.«²⁵

Dies widerspiegelt sich dann im berufsbegleitenden Ausbildungsprogramm der Institution insofern, dass die periodischen gemeinsamen Ausbildungsblöcke stets mit der gleichen Aufgabe beginnen:

Wir stellen den Stand unserer Projekte vor (Vormittag), helfen einander bei der Problemlösung und allenfalls beim Weiterentwickeln. Wir üben dabei, wie wir Probleme mit Stackoverflow, Github und weiteren Debugging-Strategien selbst lösen können.²⁶

Das projektorientierte Lernen und Anwenden von in den Redaktionen verankerten Auswertungs- und Visualisierungstechnologien sowie das Diskutieren

²⁴ Interview D2; Interview D6; Interview D9

²⁵ Interview D2

²⁶ MAZ_CAS_Datenjournalismus_18-19

dieser in kollaborativ-kompetitiver Struktur widerspiegelt sich also in der sehr praxisorientierten Ausbildung.

Sie lehnt sich dabei an die im Feld bekannteste und beliebteste Ausbildung an, das Lede Program der Columbia University New York (das auch viele europäische Datenjournalist*innen absolvieren²⁷). Wie Splendore et al. (2016) gezeigt haben, sind Datenjournalismus-Ausbildungen international sehr unterschiedlich konzipiert und geprägt von länderspezifischen Charakteristika wie der Grösse des Medienmarkts, der journalistischen Ausbildung sowie der Bildungspolitik generell. Insbesondere in Deutschland gibt es inzwischen diverse universitäre Ausbildungsgänge zu Datenjournalismus. Es liegt aber ausserhalb des Fokus dieser Arbeit, diese Unterschiede und Hintergründe detailliert auszurollen. Für die in das empirische Material eingeflossene Schweizer Datenjournalismus-Ausbildung lässt sich festhalten, dass sie stark von den bisher im Feld vertretenen Expertisen geprägt ist und auf das Ethos des gemeinsamen Lernens aufbaut: Wer befähigt worden ist, die Sprache der Community zu sprechen, wird von der Arbeit und den Best Practices anderer lernen.

Das beschriebene Ethos des gemeinsamen Lernens in der Praxis wie auch das Ausbilden zur »Community-Fähigkeit« haben Brown und Duguid (1991:48) als »becoming an ›insider‹« beschrieben, als »learn to function in a community«: »They acquire that particular community's subjective viewpoint and learn to speak its language. In short, they are enculturated (Brown, Collins, and Duguid 1989).« Die Autoren verweisen so auf die inhärent kollaborative und situative Form von Lernprozessen. Und obwohl sich diese Mechanismen grundsätzlich in allen Arbeits- und Lernpraktiken finden lassen, so fällt mit Bezug auf das datenjournalistische Feld auf, dass diese sozialen Prozesse nur selten von den abstrakten Dimensionen – »formal descriptions of work« (Arbeitsprozesse) und »subject matters« (Ausbildungen), wie sie Brown und Duguid beschrieben haben (1991:40) – verdeckt sind.

27 So schickt das grösste Schweizer Medienunternehmen Tamedia (inzwischen TX Group), welches mehrere der reichweitenstärksten Print- und Online-Nachrichtenportale betreibt, jährlich mehrere Redaktionsmitglieder in das dreimonatige Lede Program. Dabei handelt es sich nicht nur um Datenjournalist*innen, sondern auch Personen mit leitenden Positionen in der Redaktion. (Interview D8)

Der Spirit der Open-Kultur

Im datenjournalistischen Feld finden sich verschiedene Verweise auf und Überschneidungen mit dem *open movement*, welche der folgende Abschnitt ausleuchtet. Diese Bewegung umfasst ein Netzwerk mit verschiedenen Verästelungen, wovon die bekanntesten *Open Source*, *Open Data* und *Open Knowledge* sind. Zugrunde liegt allen Ausformungen sowohl eine materielle wie eine kulturelle Abgrenzung: Es geht darum, Wissen nicht markt- oder vertragsgetrieben zu teilen, sondern mit einer breiten und offenen Palette an Nutzenden. Und das, ohne sofort einen Gegenwert dafür zu erhalten – sondern mit der Intention, zu einer gemeinsam geteilten, partizipativen Entwicklung beizutragen (Lewis und Usher 2013:606). Konkret bedeutet dies, dass der Zugang zu Daten und Wissen – zum Beispiel Software-Quellcodes – möglichst für alle offen sein soll (Coddington 2015:333). Lewis und Usher (2013:603) betonen mit spezifischem Bezug auf *Open Source* die Relevanz dieser Elemente für den Journalismus: »Open source – as both a practical approach to transparent and participatory coding, as well as a philosophy of sharing – is central to the hacker ethic and provides common ground for groups and activities that bring together journalists and technologists.«

Die in der Community gepflegte Kooperation (siehe Abschnitt 3.1.2) ist eng verknüpft mit den kulturellen Vorstellungen des *open movement* (Borges-Rey 2016:838; Hermida und Young 2019:54). Sie manifestiert sich entsprechend in (weitgehender, aber nicht absoluter) Offenheit und Transparenz über die eigenen datenjournalistischen Praktiken, genauso wie im Zugänglichmachen von Daten und Methoden sowie deren Reflexion, womit das gemeinsame Lernen akzeleriert wird, welches die Community besser machen soll. Eine schon viele Jahre im datenjournalistischen Feld verankerte Journalistin bringt den Spirit der Community sofort mit der *Open*-Kultur in Verbindung:

»Ich glaube der kommt total stark aus dieser Open-Data-Community, zum Beispiel, mit der wir ja sehr viele Werkzeuge und Herangehensweisen teilen, und die auch viele Daten bereitstellen, die für Journalisten wichtig sind, und Daten überhaupt auch ausgraben und verfügbar machen. Wenn ich zum Beispiel an die OKFN [Open Knowledge Foundation, Anm. RE] denke, das sind Kollegen, mit denen teilen wir ganz viel, und zugleich sind die auch viel viel transparenter, und viel viel offener, als wir es manchmal sein können. Das ist ein grosser Einfluss, der glaube ich die Szene auch geprägt hat.«²⁸

28 Interview D4.

Wenn es um die Entwicklung von Werkzeugen und Auswertungsmethoden geht, besteht die Absicht, diese *open source* der Öffentlichkeit, und damit vor allem dem datenjournalistischen Feld, zugänglich zu machen. Ein Datenjournalist, der früher für konventionelle Redaktionen Texte schrieb und nun in einem *Tech*-orientierten Datenjournalismus-Team arbeitet, erzählt:

»And the idea of open source is something... is really important to what we did, right, because... I'm having seen the way that... journalistic culture and media culture operates. It was crazy coming into the more tech world. And seeing that people are... very open to share their code and visualizations, and processes. But I think it's something that is much more, eh... much more common in tech.«²⁹

Dass die *Open*-Kultur seit dem Entstehen des datenjournalistischen Feldes eng mit dessen Erfolgsgeschichten verknüpft ist, zeigt sich bereits in einer der frühesten Studien über ein datenjournalistisches Team: Royal untersuchte 2009 das Interactive News Technology Department der New York Times und konstatiert in ihrer Publikation:

»Part of the success of this team comes from the culture that has been developed in the department, the roots of which are in creativity and innovation, driven in many ways by the open-source or hacker culture. This mindset is something that is infiltrating the entire Times organization and is recognized as important to future success.« (Royal 2012:16f)

Es stellt sich die Frage, durch welche Mechanismen sich dieses »Mindset« im datenjournalistischen Feld und seinen Praktiken einbettet.³⁰ Baack (2018:678f) weist darauf hin, dass die *Open*- und die Datenjournalismus-Community über geteilte Sinngebungsprozesse verfügen: (1) weil sich ihre Fähigkeiten überlappen; (2) weil sie beide über ein Bekenntnis zur Lern- und *Open*-Kultur verfügen; (3) weil sich nicht nur der Journalismus, sondern auch die *Open*-Kultur als öffentliche Dienstleistung sieht (»facilitating«). Aus diesen drei Gründen würden

29 Interview D17

30 Einige Forschungsstudien über Datenjournalismus haben die Verschmelzungen und/oder Grenzziehungen zwischen der Technologie- und der Journalismus-Kultur aus einer Identitätsperspektive beleuchtet, siehe beispielsweise Royal 2012 und Hanaford 2015. Diese Perspektive ist für diese Arbeit wenig relevant, weil Datenjournalismus als verteilte und nicht personengebundene Expertise aufgefasst wird.

die beiden Communitys eine gemeinsame »community of practice« (Wenger 2008) bilden, so Baack.

Lewis und Usher fokussieren darauf, dass aus Sicht der *Open-Source*-Kultur verschiedene Werthaltungen in Technologien eingebettet sind. Sie nennen dabei vier Aspekte:

»Like the architecture of open source, this open-source culture carries with it certain assumptions about technology's normative values: namely, transparency (coding in the open, with bug-tracking); iteration (continuously releasing unfinished code for beta-testing); tinkering (privileging play and experimentation, focusing on the process of work more than its outcome); and participation (encouraging input from the widest possible group of collaborators).« (Lewis und Usher 2013:607)

Einige Jahre später ordnen Lewis und Zamith (2017) das Zusammenspiel von Technologie und Journalismus in verschiedene »worlds« ein, die unterschiedliche datenjournalistische Ausprägungen hervorbringen. Ihnen gemeinsam ist, dass die »nerds« an Status in den Redaktionen gewonnen haben und sich nicht nur ihre Rollen wandeln, sondern sie auch die »conventions« in den Redaktionen verändert haben – eben genau in die Richtung von Experimentierfreude, Iteration und Bastelei.

Saxenian (1996) hat die Arbeitskultur im Silicon Valley der Siebziger- und Achtzigerjahre porträtiert – sehr viele der von ihr beschriebenen Charakteristika finden sich im datenjournalistischen Feld wieder: eine eng verflochtene, netzwerkartige Struktur von Unternehmen; projektbasierte Arbeitsstrukturen und Teams ohne steile Hierarchien, stattdessen auf persönlicher Motivation aufbauend; eine Kultur, die Experimente und »Entrepreneurship« begünstigt; eine Community mit einem Entdeckungs-*Mindset* bezüglich technologischer Möglichkeiten; viel informelle und offene Kommunikation zwischen den Angestellten verschiedener Unternehmen; gleichzeitig vertrauensvolle wie kompetitive Beziehungen und Lernprozesse zwischen ihnen; institutionalisierte Strukturen im Feld, welche dieses gegenseitige Lernen und kollaborative Praktiken fördern; sogar die Ausbildung einer eigenen Sprache zwischen den Community-Mitgliedern. Das datenjournalistische Feld verfügt über eine Durchlässigkeit bezüglich Bildungshintergrund, solange man die eigenen Praktiken stets verbessern will und die »Sprache« der Community spricht, und über eine starke Prägung durch Ideen der *Open*-Kultur wie Transparenz, Experimentierfreude, Bastelei, Iteration und Partizipation. Wie bereits im letzten Zitat von

Lewis und Usher angedeutet stehen diese in einem engen Bezug mit den technologischen Werkzeugen – darauf geht der folgende Abschnitt ein.

3.2.2 Werken mit Datensätzen, geometrischen Formen und Tools

Der folgende Abschnitt beleuchtet, wie das datenjournalistische Feld die Materialität der Datensätze, der geometrischen Formen und der Software für deren Bearbeitung (*Tools*) diskutiert und bewertet.

Datensätze

Datensätzen, oder kurz ›Daten‹, schreibt das Feld ein grosses Potenzial zu, ein Potenzial für neue Geschichtenoptionen. Für journalistische Erkenntnisproduktion Daten zu nutzen gehört für das Feld zur Zukunft des Journalismus in einem datafizierten Zeitalter. Daten sind eine zusätzliche Materialität, welche die journalistische Erkenntnisproduktion ergänzt und mit anderen Quellen in ein Zusammenspiel tritt, wie eine Interviewpartnerin es in der Antwort auf die Frage, ob Datenjournalismus ›die Zukunft‹ sein wird, formuliert:

»Ja ein Teil der Zukunft, ja. Ich glaube das andere braucht es auch, aber jetzt haben wir die Möglichkeit, grosse Datenmengen zu verarbeiten, also, und es gibt die grossen Datenmengen, also brauchen wir sie doch auch. Aber das ist ein Teil von... Es braucht auch immer noch das direkte Gespräch, und es wird immer noch Leute geben die irgendwo hingehen müssen um etwas anzuschauen, also, ja, ich glaube es erweitert einfach unseren Fächer an Möglichkeiten um zu recherchieren.«³¹

Meistens werden für datenjournalistische Praktiken öffentlich zugängliche Datensätze verwendet, beispielsweise des Statistikamtes, von Forschungsinstitutionen oder NGOs – dies gilt für »everyday data journalism« (Zamith 2019) wie auch *Leuchtturm-Projekte* (Loosen et al. 2017; Young et al. 2018). Trotzdem idealisiert das datenjournalistische Feld die eigene Datenerhebung: Sie gilt als erstrebenswert, insbesondere bei auf Investigation ausgerichteten Redaktionen, welche journalistische Kriterien hoch gewichten: Die Daten sind im Falle einer eigenen Erhebung ›unverbraucht‹ und der Neuigkeitswert einer daraus entwickelten Erkenntnis ist wahrscheinlich. Ebenso ist der Kontext der Erhebung und somit die Parameter des Datensatzes besser kontrollierbar. Die soziomaterielle Formung durch die Erhebungsinstrumente wird im

³¹ Interview D9

Feld selten explizit, aber meist implizit bestätigt – wie beispielsweise im folgenden Zitat: Ein Datenjournalist beschreibt, welche »banale[n] Sachen« Journalist*innen prüfen sollten, wenn sie von anderen erhobene Datensätze nutzen:

»So ganz, ganz banale Sachen wie gro- wie, wie viel Umfrage-Teilnehmer gibt es überhaupt? Wie, wie ist die Umfrage durchgeführt worden? Also, so, so ganz methodische Grundlagen. Die sollte glaub ich jeder Journalist können in seiner Alltagsarbeit, dass er, so wie er normalerweise auch jede Quelle und jede, eh, jeden Text einschätzen kann, muss er halt auch Zahlen ein bisschen grundlegend einschätzen können.«³²

Auch die in datenjournalistischen Weiterbildungen vermittelten Best Practices stellen Datensätze in eine Reihe mit anderen journalistischen Quellen – deshalb gilt es sie ebenso zu »verifizieren« wie eine menschliche Quelle³³ sowie ihren Erhebungszweck zu hinterfragen und sie auf Fehler und Manipulationen zu prüfen.³⁴

Das wichtigste Kriterium für die Bewertung eines Datensatzes ist aber dessen Strukturierung und Interoperabilität, welche eine schnelle und multioptionale Verarbeitung von Daten ermöglicht. Am auffälligsten sind dabei die Verweise auf das *Tidy-Data*-Prinzip (Wickham 2014), welche das Feld immer wieder vornimmt. Die Ein-Ordnung und gleichzeitige Reduktion von Ereignissen in ein Tabellenformat mit Spalten und Zeilen bezeichnen die Datenjournalist*innen als »eines der wichtigsten Prinzipien«³⁵ beziehungsweise als Gebet des »Gott[s] von R«, Hadley Wickham.³⁶ Für das datenjournalistische Feld ist die digital angeschlossfähige Datenbank die Grundlage zur Erkenntnisproduktion über die Welt (Manovich 2001; Burkhardt 2015).

Geometrische Formen

Die Sinngebung in der Verarbeitung wiederum verläuft über die Materialität der geometrischen Formung. Eine Interviewpartnerin beschreibt das zweiseitige Zusammenspiel zwischen Formung und Sinngebung folgendermassen:

32 Interview D12

33 MAZ_Ausbildung_2018_Folie058

34 UZH_0927_Folie29

35 MAZ_Ausbildung_2018_Folie060

36 UZH_1101_Folie09, 6_20171101_SeminarUZH-2

»Was will ich eigentlich zeigen? Ich hab keine Ahnung, lass mich nachdenken, ja, stimmt, ich will ja eigentlich das zeigen was Darstellungstyp A am besten zeigen kann. Und dann war das immer... so die, so die, die Lösung. Und das hat es auch, also das hat funktioniert, also, das funktioniert im Journalismus und mit diesen explanatory halt am besten, und das funktioniert sehr schön. Und, eh, schlüssig, dass man einfach so fragt: Okay, was ist denn das Statement, das ich machen will? Hier ist die Darstellungsform. [...] Also wenn du noch nicht weisst, was du zeigen willst, dann ja, nutze alle möglichen Chart-Typen, und nutze Histogramme, und Scatterplots, also, dann sind halt solche... exploratory Chart-Typen wie Histogramme und Scatterplots hervorragend. In Scatterplots kann man Korrelationen sehen, in Histogrammen kann man Verteilungen sehen, ehm... um erstmal die Daten zu verstehen und zu kucken: Okay, was gibt es für Zusammenhänge, was ist hier eigentlich spannend.«³⁷

Abbildung 1: Verbindung von ›W-Fragen‹, dem ›klassischen‹ Mittel der journalistischen Erkenntnisproduktion, mit grundlegenden geometrischen Abbildungsformen. Folie aus einem datenjournalistischen Weiterbildungskurs.

Wer hat am meisten? Balkengrafik

Wie unterscheidet sich das? Kuchen

Wann? Zeitverlauf: Linien-Grafik

Wie ist das passiert? Flow-Chart

Wie hängt X mit Y zusammen? Scatterplot

Tool: [DataVizCatalogue](#)

(MAZ_Ausbildung_2018_Folie104)

Im datenjournalistischen Feld hat sich ein visuelles Vokabular etabliert, dessen Basis einige Grundformen wie Balken-, Linien- und Streudiagramme

37 Interview D11

sind. Dabei verweist das Feld insbesondere auf Kataloge wie den DataVizCatalogue des Designers Severino Ribecca³⁸ sowie das davon inspirierte Visual Vocabulary der englischen Medienorganisation Financial Times³⁹, die als wichtiger Orientierungspunkt für Best Practices mehrmals im Feld erwähnt wird.

Am anderen Ende des Vokabular-Kontinuums finden sich experimentelle Darstellungsformen der »Data Art«, denn es gebe »eigentlich so 'nen fliessen-den Übergang zwischen Datavisualisierung und [...] Data Art, also Sachen, die einfach nur, gemacht werden weil sie schön sind«⁴⁰, wie ein Datenjournalist Visualisierungsformen einordnet.

Geometrische Formen sind Repräsentationen, die – über die zwischenge-schalteten Daten – Sachverhalte aus der ›echten Welt‹ in bestimmter Weise relationieren. Das Abbilden dieser Formen an einer bestimmten Position auf einem Ordnungssystem wie dem Gitternetz aktiviert wiederum bestimmte Muster und neue Relationierungen (Manovich 2011). Die Medienwissen-schaftler Bauer und Ernst haben diesen Prozess mit Bezug auf Diagramme folgendermassen beschrieben:

»Diagramme sind visuelle Darstellungen, die Beziehungen respektive Verhältnisse aufzeigen: die Stimmverteilung bei einer Wahl, die Zimmeraufteilung in einem Gebäude oder die Entwicklung von Aktienkursen in einem bestimmten Zeitraum. Dabei sind die Relationen und Proportionen, die Strukturen und Funktionen, die Diagramme vor Augen führen, prinzipiell veränderbar – auf dem Papier, am Computer oder kurzerhand im Kopf. Mittels Diagrammen können verschiedene Konfigurationen eines Sachverhaltes oder einer Ereignisfolge durchgespielt und miteinander ver-glichen werden, um zu praktischen oder theoretischen Schlussfolgerungen zu gelangen; welche Parteien eine Mehrheitskoalition bilden können, wie sich ein Gebäude einrichten oder umbauen lässt und welche Risiken eine Börsenspekulation birgt.« (Bauer und Ernst 2010:9)

Die datenjournalistischen Relationierungen vollziehen sich durch das visuelle Vokabular, das sich im Spannungsfeld zwischen den journalistischen Aussageimperativen der W-Fragen und ästhetischen Kriterien folgenden,

38 <https://datavizcatalogue.com/> (24.10.2020)

39 <https://github.com/ft-interactive/chart-doctor/blob/master/visual-vocabulary/Visual-vocabulary.pdf> (24.10.2020)

40 Interview D12; auch in Interview D18 wird auf »data visualization artists« als Inspiration verwiesen.

künstlerischen Datenverkörperungen bewegt. Auch wenn raffinierte, mehrschichtige Datenvisualisierungen als Referenzobjekte im Feld dienen⁴¹, so zeigen Analysen über den alltäglichen Datenjournalismus, dass ein Grossteil der Datenvisualisierungen die erwähnten Grundformen sind (Stalp 2017:9). Zamith (2019:15) verweist auf die im Feld wenig artikulierte, aber in den Praktiken grosser Medienhäuser zum Ausdruck kommende Ansicht von »semiotic simplicity as a strength rather than a weakness«. In dieser unterschiedlichen Bewertung widerspiegeln sich die verschiedenen Publika der Feld-Diskussionen beziehungsweise der Praktiken. In Zweiterem geht es schliesslich darum, möglichst sicherzustellen, dass Datenvisualisierungen von den Lesenden verstanden werden. Ethnografische Beobachtungen aus Datenjournalismus-Teams werden diese Aspekte im Verlaufe der Arbeit vertiefen: Abschnitt 6.1.1 wird das Zusammenspiel geometrischer Formen als Materialität und journalistische Expertise in datenjournalistischen Praktiken beleuchten. Im gleich folgenden zweiten Teil dieses Abschnitts wird deutlich, dass *Tools* massgeblich mitbestimmen, welches visuelle Vokabular, ergo welche Relationierungsoptionen als Palette dem Feld zur Verfügung stehen.

Tools

Obwohl einige Interviewpartner*innen erwähnen, dass ihre Redaktionen Daten schon *vor dem Datenjournalismus* nutzten – besonders oft wird dabei auf das Wirtschaftsressort verwiesen –, so fällt die Entstehung des Feldes mit der Entwicklung der Werkzeuge zur quantitativen Auswertung der grossen Datenmengen zusammen. Das zu Beginn dieses Kapitels erwähnte Potenzial von Daten als journalistische Quelle ist für das Feld erst durch die *Tools* und ihre Methoden der Auswertung und Visualisierung zu bestimmen. Dementsprechend sind sie oft und intensiv diskutierte Themen im Feld. Und die *Tool*-Entwickler*innen sind nicht nur Teil des Feldes, sondern auch wichtige Stimmen darin – auf den Entwickler der Programmiersprache R, Hadley Wickham, und

⁴¹ Die grosse Bedeutung aufsehenerregender Datenvisualisierungen für die Community zeigt sich beispielsweise daran, dass es mindestens zwei Dutzend Personen oder Organisationen gibt, welche Jahres-Rückblicke über die oder mithilfe der ›beste(n)‹ Datenvisualisierungen publizieren. Der international bekannte Datenjournalist Maarten Lambrechts stellt jährlich ›The list of [year] visualization lists‹ zusammen – hier für 2019: <https://www.maartenlambrechts.com/2019/12/23/the-list-of-2019-visualization-lists.html> (24.10.2020)

sein *Tidy-Data*-Prinzip wurde bereits verwiesen. Mehrere Personen des Datenvisualisierungstools Datawrapper⁴² sind in der datenjournalistischen Community zentrale Personen und haben auch selbst schon in Redaktionen gearbeitet. Dasselbe gilt für den in Interviews mehrfach erwähnten Entwickler der JavaScript-Datenvisualisierungspalette D3.js, Mike Bostock, der mehrere Jahre auf der Redaktion der New York Times arbeitete. Die persönlichen und fachlichen Verflechtungen zwischen den Redaktionen und den *Tool*-Entwickler*innen sind gross.

An die Materialität von Werkzeugen hat das datenjournalistische Feld klare Anforderungen: Ein *Tool* ist gut, wenn es Nachvollziehbarkeit, Transparenz und Reproduktion (der Analyse) ermöglicht – hier spiegeln sich besonders die kulturellen Vorstellungen der *Open*-Kultur. Besonders konsequent eingeschrieben sind diese Anforderungen in den Programmiersprachen, welche entsprechend als Werkzeug-Goldstandard im Feld gelten: Sie integrieren verschiedene Teilpraktiken des datenjournalistischen Prozesses mit derselben Logik und gemeinsamen Datenstrukturen⁴³ in eine »eigene Programmier-Welt«⁴⁴. Dadurch erfüllen sie die Anforderung, Datensätze schnell und flexibel auswerten zu können: Die interoperable Schnittstelle zwischen den Datensätzen und Werkzeugen ermöglicht ein schnelles Durchprobieren verschiedener Auswertungsmethoden. So antwortet ein Datenjournalist auf die Frage, was ein Werkzeug zu einem guten Werkzeug mache:

»It needs to be flexible, so you will be able to code beyond your normal charts. Ehm, so both of those tools, ggplot [Datenvisualisierungspaket in R, Anm. RE] and d3, they allow you to be quite creative. But it also needs to be fast. So I used to do a lot more of my work in d3, but now I'm moving more to ggplot and R because – yeah, it integrates with the data preparation and the data cleaning, and then it is much faster to come up with the results. So ehm... Speed and flexibility.«⁴⁵

Die Nutzung von Programmiersprachen bedeutet innerhalb des Feldes auch ein Abgrenzungskriterium: Es ist »richtig sauberes Handwerk«⁴⁶, weil das Feld

42 <https://www.datawrapper.de/> (28.10.2020)

43 Interview D5

44 Interview D19

45 Interview D5

46 Interview D12

datenjournalistischen Praktiken mit Verweis auf transparente und reproduzierbare Skripte Überprüfbarkeit zuschreibt.

Die zwei meistgenutzten Programmiersprachen im Feld sind R und Python. Dabei gehört es zwar dazu, eine gewisse »Show-Rivalität« zwischen der Nutzung von jeweils einer der beiden zu pflegen, indem man sie als Religionen darstellt (siehe auch Zitat Fussnote 36 in Kapitel 3). Immer wieder wird dann aber auch betont, dass in den meisten Teams mehrere Programmiersprachen im Einsatz sind und dass dies problemlos funktioniert. Ein Datenjournalist, der Python nutzt, fasst zusammen:

»Ich hab so, letzte Woche hab ich so ein bisschen mich in ein R-Skript rein-gelesen, und es ist, es ist einem sehr vieles vertraut wenn man so Python macht, inzwischen. Ich glaube diese Auseinandersetzung R/Python ist gar nicht so... eh... ist, ist, ist lustig, aber ist nicht so wichtig.«⁴⁷

Obwohl in der Regel das eigene Programmieren in den Vordergrund gestellt wird, anerkennt das Feld den Wert von *Tools*, welche Datenauswertungen und -visualisierungen ermöglichen, ohne dass die Nutzenden programmieren müssen. Excel gilt als eine Art Startpunkt⁴⁸ beziehungsweise »wichtiges Instrument«⁴⁹, das aber keine Nachvollziehbarkeit gewährleistet⁵⁰ und über untaugliche Standardeinstellungen⁵¹ verfügt. Die Datenvisualisierungssoftware Datawrapper geniesst im Feld eine grosse Reputation, weil es ihr zuschreibt, dass sie einerseits eine gewisse Interoperabilität mit Programmiersprachen (und damit dem Handwerk der »Cracks«) bietet, aber andererseits auch für »normale« Journalist*innen, die Daten selbst visualisieren möchten, beherrschbar ist. Ein Datenjournalist, in dessen renommierter Redaktion Datawrapper im Einsatz ist, erzählt, wieso man sich für dieses *Tool* entschieden hat:

»Ehm, natürlich gibt es viele andere Tools, nur da haben wir einfach... festgestellt dass sie dann irgendwann schon wieder eigentlich zu gross sind. Ne, also, Datawrapper hat grade noch so... schafft es grad noch so beherrschbar zu sein ohne dass es jetzt wieder die eierlegende Wollmilchsau ist, die wieder alles kann, und wenn man hier nochmals einstellt, also, es ist... und es

47 Interview D6

48 Interview D5

49 MAZ_Ausbildung_2018_Folie082

50 Interview D6

51 Interview D3

ist halt für Journalisten gemacht, ne, also für Sachen wie, da gibt's 'n Feld wo drin steht Quelle. Das hast du bei anderen Tools nicht selbstverständlich. Also das ist irgendwie, es ist journalistisch gedacht und deswegen ist es eigentlich so unser, unser Tool was wir verwenden.«⁵²

Aus Sicht des Feldes kondensiert in Datawrapper eine Art Grundstock an visuellem Vokabular, das *Storytelling* mit Daten auch ohne Programmieren ermöglicht, und verbindet es mit spezifischen Anforderungen der journalistischen Standards. Ein Datenjournalist erklärt dieses Zusammenspiel folgendermassen:

»Yeah, well, I think Datawrapper ehm... hits a very sweet spot. Because you can make, yeah, you can make very fancy graphics with the tool, but... all the basic chart types are there. And it allows you for example to annotate your charts. So you can start telling a story with your charts. Or you can tell your readers well you should look here, this is the interesting part of the visualization. Ehm, and it's also very fast. And in the newsroom, yeah, you don't find many coders. So the tools I use are code-based, while Datawrapper is aimed at non-coders. Ehm, so – I'm a fan of Datawrapper. And they... The tools have good defaults, so you have quality visualisations, and it allows you to tell a story, and it's fast for non-coders as well.«⁵³

Kürzer fasst sich ein Dozent in einem Datenjournalismus-Kurs: Er lobt Datawrapper, weil das Produzieren von Datenvisualisierungen dank den guten Grundeinstellungen »tubelischer« [Schweizer Mundart für idiotensicher] sei.⁵⁴ Das Feld schreibt Datawrapper zu, den datenjournalistischen »state of the art« abzubilden, den man in dieser Form als Handwerk aller Journalist*innen erwartet, und diesen gleichzeitig mit den gesetzten Standardeinstellungen festzuschreiben.

Datawrapper ist das meistdiskutierte *Tool* des Feldes, sicher auch dank der bereits erwähnten starken Verflechtung der *Tool*-Entwickler*innen mit der Community. Im Laufe der Zeit rückte dieses *Tool* deshalb auch in den Fokus der Feldbeobachtung, weshalb überdurchschnittlich viel Forschungsmaterial darüber vorhanden ist (siehe Kapitel 2 zu Methoden). Neben den bereits erwähnten Programmiersprachen und ihrer Visualisierungspackages (ggplot2

⁵² Interview D10

⁵³ Interview D5

⁵⁴ MAZ_Ausbildung_2018_Folie106

für R, matplotlib für Python) sowie Excel referenzieren die Akteur*innen auch Visualisierungstools wie Tableau, Flourish, QGIS, RAWgraphs, Infogram oder die redaktionsspezifischen Mr. Chartmaker (New York Times) oder Q (NZZ). Im Unterschied zu Datawrapper erscheinen diese aber nur wenig diskutiert und insofern wenig formend für das Feld – oder sie verfügen zwar über eine sehr lebendige eigene Community, wie im Falle von Tableau, die aber nicht allzu grosse Überschneidungen mit der datenjournalistischen Community aufweist.⁵⁵

Daten, geometrische Formen und *Tools* sind als Materialitäten der datenjournalistischen Praktiken eng verflochten mit epistemologischen und kulturellen Aspekten, wie sich dies in der hier abgebildeten Empirie beispielsweise durch die enge Community-Verbindung zeigt. Anderson und De Maeyer (2015:4) plädieren deshalb für »a relational understanding of technology rather than a deterministic one, an understanding that sees the material aspects of objects as inevitably imbricated in a web of human and non-human relations.«

Die Entwicklung des datenjournalistischen Feldes ist also eng verknüpft mit den verwendeten Werkzeugen und deren sozialem Kontext, mit deren »affordances« und »constraints« (Hermida und Young 2019:80ff). Dieser Abschnitt hat zwar aufgezeigt, dass einige Werkzeuge wie die Programmiersprachen R und Python sowie das Online-Tool Datawrapper wichtige Institutionen des Feldes sind. Allerdings ist nicht zu unterschätzen, dass innerhalb des Feldes eine grosse Menge an verschiedenen Werkzeugen zum Einsatz kommt, was wiederum darauf verweist, dass es sich um ein sehr durchlässiges, sich im ständigen Fluss befindliches Feld handelt (Hermida und Young 2019: 74ff).

3.2.3 Quantitativ, qualitativ, iterativ, visualisiert – epistemologische (Ent-)Spannungen

Sprechen Datenjournalist*innen über sich selbst, so taucht ab und zu der Begriff *Nerd* oder *nerdig* auf. Ein Schweizer Datenjournalist bezeichnet seine Community sogar als »Weirdo Nerds« in den Redaktionen.⁵⁶ Dies ist nur eine von zahlreichen Abgrenzungspraktiken, der Abschnitt zum Ethos hat bereits auf weitere hingewiesen. Diese Grenzziehungen erfolgen innerhalb

55 Interview D11

56 11_20171220_SeminarUZH-6

des journalistischen Feldes. Das heisst, dass sich die Datenjournalist*innen grösstenteils durchaus als Teil des Journalismus sehen, oder wie es eine deutsche Datenjournalistin formuliert: »Wir sind ja mit denselben Relevanzkriterien unterwegs.«⁵⁷ Auch geniesst die Zusammenarbeit mit thematischen Ressorts, über welche die meisten Redaktionen verfügen, im Diskurs des Feldes einen hohen Stellenwert. Gleichzeitig reiben sich datenjournalistische Praktiken aus Sicht des Feldes immer wieder an den althergebrachten Redaktionsstrukturen und Bewertungskriterien, die man als verschlossen und die Zusammenarbeit erschwerend wahrnimmt. So verweisen einige Interview-Partner*innen darauf, dass Datenjournalist*innen »Erklärbär[en]«⁵⁸ seien, weil sie anderen aufzeigen, was mit Programmieren (nicht) möglich ist oder weshalb Visualisierungen mehr als eine Dienstleistung am Ende des journalistischen Prozesses sind.⁵⁹ Dass viele datenjournalistische Teams über eigene Publikationsrubriken verfügen oder gar ein eigenes Ressort bilden, empfinden viele Datenjournalist*innen als Privileg: Die als Hürden empfundenen Bewertungskriterien der anderen Ressorts sind damit umgangen. Ein Datenjournalist, der in einer grossen amerikanischen Medienorganisation in einem eigenständigen Team arbeitet, vergleicht diese Hürden mit der Mindestgrössen-Massgabe vor der Achterbahnfahrt in einem Vergnügungspark:

»When you go to an amusement park in the United States, they have a little like... stick. And it says if you're not as tall as this stick, you can't get on this, right. [...] Like a little kid goes on a dangerous rollercoaster. So like the, before you get on like, are you this tall? [...] you have to be this tall to be able to like understand, there's a floor you understand and you have to have, eh... Having someone at that editorial level, of like being a department head, who does have that expertise, like you need that to be able to edit conceive and execute lots of these... ideas. [...] Yeah, eh, not just data, and coding, for sure, but like... a certain kind of visual literacy, a certain value for information graphic [...]«⁶⁰

57 Interview D4

58 Interview D10

59 Diese Perspektive auf andere Redaktionsmitglieder wird dadurch verstärkt, dass fast alle Interview-Partner*innen aus Europa erwähnen, dass sie in ihren Redaktionen Weiterbildungsaufgaben wahrnehmen. Diesem interessanten Befund nachzugehen, liegt aber ausserhalb des Fokus dieser Arbeit.

60 Interview D15

Neben dem Umgehen dieser Hürden versuchen verschiedene Medienorganisationen, das Verstehen von datenjournalistischen »ideas« auch auf den Leitungspositionen zu fördern.⁶¹ Besonders relevant ist das Zitat des Datenjournalisten aber für die Frage: In welchen Aspekten unterscheiden sich datenjournalistische »ideas« und worin besteht »that expertise«, die es braucht, um die »ideas« des Datenjournalismus angemessen bewerten zu können? Der Schluss des Zitats verweist auf die Perspektive unterschiedlicher Expertisen – Programmieren und Visualisieren –, deren Verstehen als notwendig erachtet wird, um datenjournalistische Erkenntnisse einschätzen zu können. Viele der genannten Abgrenzungspraktiken verweisen im Grunde darauf, dass die epistemologischen Grundlagen der datenjournalistischen Praktiken sich von den »journalist's knowledge claims about the empirical world« (Parasie und Dagiral 2012:854) unterscheiden. Der folgende Abschnitt skizziert zuerst kurz einführende Grundlagen der quantitativen Epistemologie und Forschungsperspektiven darauf, widmet sich dann aber vor allem der Frage, wie das untersuchte Feld darüber spricht.

Einordnung der quantitativen Epistemologie

Datenbasierte Analysen sind empiristisch (Abbott 2004), das heisst, sie abstrahieren ein Phänomen und unterstellen gleichzeitig positivistisch, dass Natürliches oder Soziales messbar ist. Die meisten anderen journalistischen Methoden sammeln Informationen qualitativ und wenig abstrahiert, basierend auf bisherigen Annahmen über die Wirklichkeit. Sie teilen aber ein realistisches und in der Empirie verankertes Weltbild (Ward 2018): Konventioneller Journalismus bewegt sich auf einer Ebene des gesunden Menschenverstands (»common sense realism«), Fakten sind real und stabil – es gilt, die Geschehnisse der Welt zu *entdecken* und Geschichten darin zu *finden*.⁶²

C.W. Anderson untersucht in seinem Buch »Apostles of Certainty« (2018), wie sich die »knowledge work«, also die Erkenntnisproduktion, im Journalismus seit dem frühen zwanzigsten Jahrhundert entwickelt hat, insbesondere mit einem Fokus auf die Nutzung quantitativer Verfahren.

Anderson beschreibt die Entwicklungen im (Daten-)Journalismus in Abgrenzung zu jenen in den Sozialwissenschaften: Beide Tätigkeiten haben den

61 Fussnote 27 des Kapitels 3 (Tamedia/Lede) hat bereits auf ein solches Beispiel hingewiesen.

62 Dies im Gegensatz zu einem konstruktivistischen Ansatz, in welchem sowohl Fakten als auch Geschichten einer Konstruktionsleistung bedürfen.

Anspruch, Evidenz über die Geschehnisse in der Welt zu produzieren. Sie unterscheiden sich aber in ihren Vorgehensweisen und Konventionen, diese zu generieren. Der Autor beschreibt, wie die neuartigen »evidentiary objects« des Datenjournalismus neue epistemologische Ansprüche und Vorgehensweisen mit sich bringen:

»[...] these new evidentiary objects [statistics, databases, data visualizations] had a particular materiality (particularly documents and databases) and made new epistemological claims (they were quantitative) in ways that challenged some fundamental reporting practices, practices which were still largely based on oral and qualitative evidence.« (Anderson 2018:186)

Eine quantitative Epistemologie bedeutet, Erkenntnisse aus etwas Zählbarem zu schliessen. Präziser: Quantifizierung und die Formalisierung in Datenbanken transponieren qualitative Evidenzobjekte in quantitative, Informationen (»Daten«) sind so mathematisch behandelbar, vergleichbar und rekombinierbar (Heßler und Mersch 2009:36, für generelle Überlegungen siehe Heintz 2010; Krämer 2009; Passoth und Wehner 2013).

Quantitative Erkenntnisproduktion ist aber nicht nur eine Frage der Datengrundlage und -auswertung, sondern ist mit bestimmten Praktiken und kulturellen Vorstellungen und dem daraus resultierenden Verständnis dieser Datengrundlage verbunden (Anderson 2015:354) – beispielsweise mit einer vom investigativen Journalismus geprägten Ansicht, dass Daten in sich selbst noch wenig Erkenntnisse bergen, sondern zuerst kontextualisiert und visualisiert werden müssen (Parasie und Dagiral 2012:859f); oder mit dem radikalen Transparenzideal der programmiertechnischen Expertise, dass Datensätze *dann* gar keine Information verstecken können, *falls* sie genügend granular, komplett und für das Publikum zugänglich sind (Parasie und Dagiral 2012:862ff). Daraus entwickeln sich verschiedene Ausprägungen von Datenjournalismus – für eine weitere Übersicht siehe insbesondere Coddington (2015) –, die Abschnitt 3.3.1 in verschiedene Narrative des Feldes einbettet. Zuerst beleuchtet aber der folgende Abschnitt, ob und wie die hier beschriebenen epistemologischen Zugänge im datenjournalistischen Feld sich tatsächlich herauskristallisieren und verhandelt werden.

Erkenntnisse iterativ produzieren und visualisiert aufzeigen

Im Feld zeigt sich, dass die Unterscheidung verschiedener Ausprägungen von Datenjournalismus für die Akteur*innen kaum Relevanz besitzt. Stattdessen verhandeln die epistemologischen Reflexionen im Forschungsmaterial das

Verhältnis zwischen quantitativ und qualitativ produzierten Erkenntnissen und den für deren Verflechtung optimalen organisatorischen Strukturen. Die althergebrachten journalistischen Methoden (»shoe-leather epistemologies«, Lewis und Waters 2018) bauen bezüglich Empirie auf Erfahrungswissen, das Lesen von Dokumenten, Beobachtung und qualitative Befragungen auf (Anderson 2018:188ff, für eine konzeptionelle Systematisierung siehe Hanitzsch 2007).

In den methodischen Diskussionen des datenjournalistischen Feldes finden sich die quantitativen Auswertungen verbunden mit der Heuristik, dass die Vorgehensweise zur Auswertung des Datensatzes entweder explorativ beziehungsweise *data-driven* oder erläuternd beziehungsweise *story-driven* war.⁶³ Der Konsens im Feld besteht darin, dass gute datenjournalistische Erkenntnisse in einem iterativen Prozess von *data-driven* und *story-driven* produziert sind. Das heisst, mit ‚traditionellen‘ journalistischen Methoden aufgebautes thematisches Wissen verflechten ‚gute‘ Datenjournalist*innen mit ihren quantitativen Auswertungen. Die Bedeutung dieses thematischen Wissens bewegt sich dabei auf einem Kontinuum zwischen den Endpunkten ‚unterstützt dabei, die richtigen Fragen zu stellen, um Muster und Trends in den Daten zu finden‘ einerseits und ‚bisherige Erkenntnisse zusätzlich mit Daten zu illustrieren‘ andererseits. Abschnitt 5.1.3 vertieft diese Heuristik auf einer praktischen Ebene.

Zur Iteration gehört es, dass Datenauswertungen aus Sicht des Feldes es in einem weiteren Schritt ermöglichen, gemäss journalistischen Kriterien (Nachrichtenwerten) relevante Einzelfälle zu identifizieren und diese dann für eine illustrierende qualitative Fokussierung auszuwählen. Diese Iteration ist oft verbunden mit dem Überschreiten einer organisatorischen Abgrenzung – sprich der Zusammenarbeit mit anderen Ressorts oder Projektteams. Eine Datenjournalistin erläutert diese Zusammenarbeit folgendermassen:

»Die Ideen aus den Ressorts sind natürlich eher inhaltlich getrieben, was toll ist, weil man da schon näher an der Story ist. Und wir kommen dann mit Ideen, zum Beispiel dass wir eine neue Datenquelle gesehen haben, oder eine Möglichkeit entdeckt haben, etwas auszuwerten, und dann versuchen da inhaltlich anzudocken. Das passiert beides, zu ungefähr gleichen Teilen.«⁶⁴

63 UZH_0927_Folie08f; Interview D4; Interview D5; Interview D14; Interview D15; Interview D18

64 Interview D4

Zwei Schweizer Datenjournalismus-Ausbildungen beschreiben den epistemologischen Ansatz – in Abgrenzung zum »gängigen Journalismus«, der auf »einzelne Handlungen« fokussiere – als »auf das grosse Ganze« blickend.⁶⁵ Aus dieser »Heli-Ansicht«⁶⁶ liessen sich dann Trends und Muster sehen, welche anderen journalistischen Vorgehensweisen verborgen blieben – *sichtbar* gemacht mittels Datenauswertungen. Gefragt danach, was eine ›gute‹ Datengeschichte ausmache, kommt ein Interviewpartner denn auch schnell auf Datenvisualisierungen zu sprechen:

»Ich glaub erstmal dass sie [die ›gute‹ Datengeschichte, Anm. RE], dass es ihr nicht um den Effekt geht, sondern wirklich der Einsatz von Daten richtig ist, und die Geschichte weiter bringt. Und dass... alles was visuell gezeigt wird auch echte starke Muster sind [Pause] wo es auch nicht um den Effekt geht, sondern wo ich einfach, also, wo die Essenz der Geschichte eh mir klarmacht: Okay, jetzt, jetzt versteh ich um was es geht, jetzt versteh ich dass es ein Thema ist.«⁶⁷

Genauso wie im Zitat ist es auch im ganzen Feld so, dass Datenvisualisierungen eine hohe Bedeutung für den Datenjournalismus zugeschrieben wird. Nur vereinzelt verweisen Interviewpartner*innen darauf, dass sie Datenjournalismus eher als Recherche- denn als Darstellungsform sehen – während für die meisten die datenjournalistische Epistemologie eng mit Datenvisualisierungen verbunden scheint. Sie zeigen etwas, was man sonst nicht sehen würde; sie vermitteln die »Essenz« auf einen Blick; machen etwas sofort verständlich; sie können und sollen »wirkungsvoll« sein.⁶⁸ Datenvisualisierungen schreibt das Feld eine höhere Autorität und Überzeugungskraft zu als Text.⁶⁹

Verschiedene Akteur*innen des Feldes verweisen darauf, dass eine Datenvisualisierung eine Aussage haben und kommunizieren müsse. So antwortet ein Datenjournalist auf die Frage, was eine gute Datenvisualisierung ausmache: »It communicates the message, that you are trying to communicate. Above all. You know, so... It... might not be exciting, but if it communicates

65 MAZ_Ausbildung_2018_Folie018ff; UZH_1004_Folie62ff

66 UZH_1004_Folie65 – »Heli« dient als Abkürzung für »Helikopter«.

67 Interview D10

68 eigene Übersetzung der in Interview D14 spezifischen Datenvisualisierungen zugeschriebenen Eigenschaft »impactful« und »powerful«.

69 Interview D12; Interview D14; Interview D15; 8_20171115_SeminarUZH-4; MAZ_Ausbildung_2018_Folie010; 46_20180629_Hacks-Hackers-Cairo; 66_20200423_Hacks-Hackers-Covid19

some sort of central idea, that you want to communicate as part of your story structure.«⁷⁰ Diese gelte es mit Farben, Titel und Annotationen zu unterstützen, wie verschiedene Akteur*innen erwähnen. Eine Datenvisualisierung ist schliesslich aus Sicht des Feldes dann gut, wenn sie das Publikum auch darin unterstützt, die erklärte Erkenntnis zu verstehen. Zudem spielen Datenvisualisierungen eine Rolle im Zusammenspiel von quantitativer und qualitativer Erkenntnisproduktion, wie dieser Abschnitt andeutet – aber ohne diese Rolle genau zu definieren.

Zusammenfassend lässt sich mit Blick auf die datenjournalistische Epistemologie nicht nur festhalten, dass es sich – sozialwissenschaftlichen Jargon nutzend – um eine Art *Mixed Methods* handelt, sondern auch, dass die Visualisierung der Daten eine entscheidende, wenn auch durch das Feld selbst nicht vollständig spezifizierte Bedeutung einnimmt, um datenjournalistische Erkenntnisse zu produzieren und zu befestigen. Abschnitt 4.2 wird diesen Aspekt deshalb theoretisch ausleuchten und eine Heuristik entwickeln, um ihn dann in den Kapiteln 5 und 6 in der empirischen Analyse von datenjournalistischen Praktiken fassen zu können.

3.2.4 Die Expertisen des datenjournalistischen Feldes

Zu Beginn des Abschnitts 3.2 hat das Zitat von Zelizer darauf hingewiesen, dass Erkenntnisse von Communitys davon bestimmt sind, welche Bewertungskriterien, Standards und Erzählungen aus dem Feld selbst darin einfließen. Die Untersuchung des datenjournalistischen Feldes benötigt ergänzende Theorien und Heuristiken, welche erstens aufzeigen, wie ein gemeinsamer Interpretations- und Erkenntnisraum dank einer gemeinsamen Sprache entstehen kann, und zweitens die Bedeutung von Werkzeugen im Feld fassen können. Anderson nutzt dafür in seinem Werk »Apostles of Certainty« (2018:186ff) folgendes Modell: Die Produktion von Erkenntnis und die Befestigung von Evidenz geschieht in der Schnittmenge von Materialität, epistemologischer und professioneller Kultur. Mit dieser Triade untersucht er verschiedene Formen und Zeitalter im Journalismus.

Die Abschnitte 3.2.1, 3.2.2 und 3.2.3 haben das Feld des Datenjournalismus und einige theoretische Überlegungen aus der Forschung ausgeleuchtet. Die Strukturierung in die Abschnitte erfolgte aufgrund von Andersons Modell und es wurden jeweils professionskulturelle, materielle und epistemologische

Facetten beschrieben und eingeordnet – auch wenn diese Trennung eine rein analytische ist und sich in der Empirie nicht wiederfindet.

Gemäss Andersons theoretischen Überlegungen definiert die Triade mit den drei genannten Aspekten – Materialität, epistemologische und professionelle Kultur – die Bestandteile von (journalistischer) Expertise (Anderson 2013a). Sowohl die kulturellen Aspekte, die Werkzeuge (verstanden als Materialitäten) wie auch die epistemologischen Grundlagen des datenjournalistischen Feldes sind aber von vielfältigen Einflüssen – Expertisen – geprägt, wie der Streifzug durch das Feld gezeigt hat.

Für das Feld ist datenjournalistische Berichterstattung über die Welt das Resultat von Teamarbeit, welche Elemente aus den Expertisen Programmierung, Visualisierung und Journalismus kombiniert. Dieses Expertisen-Dreieck ist kein rein analytisches Gebilde, sondern eine gängige Heuristik für das ganze Feld; man begegnet dem Dreieck auch in Selbstbeschreibungen von Personen und Organisationen. Fast alle Gesprächspartner*innen betonen das Zusammenspiel der drei Expertisen, die sich ergänzen und in dem so aus allen Bereichen vertieftes Fachwissen zusammenkommt. Dabei wird auch darauf verwiesen, dass das Team nicht umhinkommt, dass einige Personen die verschiedenen Expertisen »verbinden«, also eine »trader«-Rolle (siehe Abschnitt 5.2.2) einnehmen können:

»We on the news apps team, on the news apps side, it's, we are all kind of generalists, and we all kind of fall into those three categories. I mean, it's kind of like, you know, our strengths are varied on each point. Some of us are better designers, some of us are better developers, some are better reporters. But you'll have some combination of those three, that you are in our team. [lacht]«⁷¹

Diese Perspektive auf Datenjournalismus als Team-Expertise ergibt sich auch daraus, dass die Organisation als Team die dominante Form im datenjournalistischen Feld ist: Fast alle bekannten Praktiker*innen des Feldes arbeiten in Teams. Hermida und Young (2019:10) streichen die Zusammenarbeit als zentrales Charakteristikum des Datenjournalismus heraus: »Collaboration emerges from the interdisciplinary nature of data journalism.«

71 Interview D14

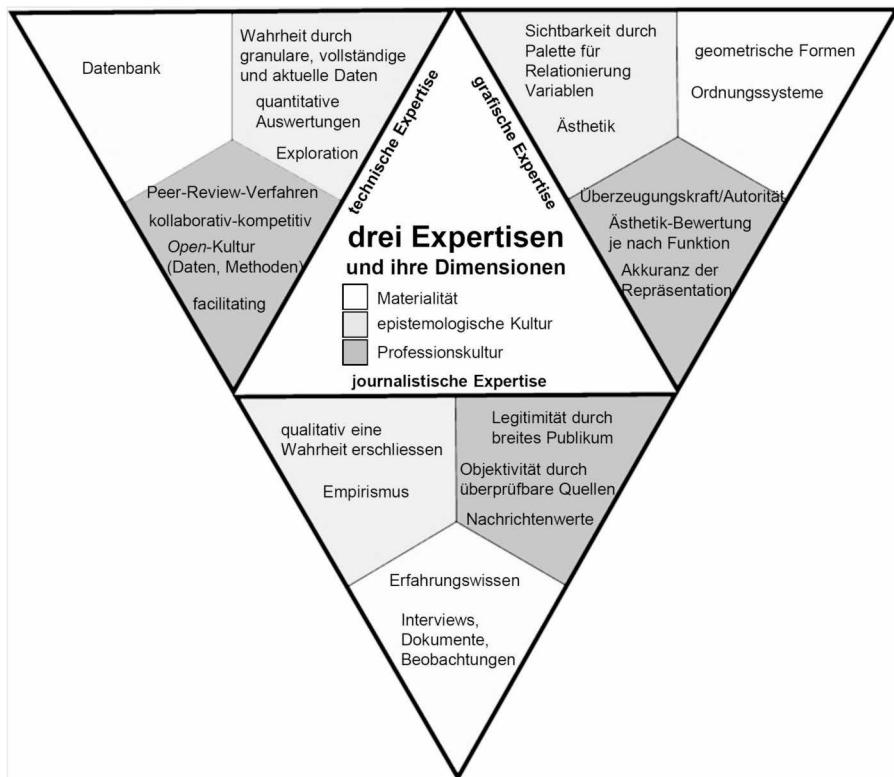
Systematisierung der Expertisen: das Dreieck der Triaden

In der Beschreibung des datenjournalistischen Feldes in den vorangehenden Abschnitten scheinen die verschiedenen Expertisen immer wieder durch. In der Zusammensetzung der Community, den kulturellen Hintergründen, in den verwendeten Werkzeugen und ihren eingravierten Affordanzen, in den epistemologischen Ansprüchen – überall zeigten sich Einflüsse und Elemente der drei Expertisen Programmierung, Visualisierung und Journalismus.

- Die *programmiertechnische Expertise* ist geprägt von der *Open-Kultur*, welche offen mit Daten und Methoden umgeht und davon ausgeht, dass Peer-Review einzelne Akteur*innen und das Feld insgesamt besser macht durch kollaborativ-kompetitives Lernen. Diese Expertise basiert auf der Materialität der Datenbank und geht davon aus, dass vollständige und granulare Datensätze valide Erkenntnisse über die in den Daten repräsentierten Sachverhalte ermöglichen. Nutzer*innen soll es möglich sein, selbst Datensätze zu explorieren und eigene Schlüsse zu ziehen.
- Die *journalistische Expertise* orientiert sich daran, dass aus Erfahrung, Dokumenten, Beobachtungen und Gesprächen gewonnene Erkenntnisse neben Datensätzen wichtige Grundlagen für die Sinngebung von Geschehnissen sind. Die journalistische Kultur basiert auf der Rolle, einem breiten Publikum gemäss Nachrichtenwerten relevante Ereignisse zu vermitteln und zu erklären, die sich über »objektiv« feststellbare Auffälligkeiten definieren. Die empirischen Quellen dafür gilt es kritisch zu überprüfen, unabhängig von ihrer Form, und wenn möglich zu nennen.
- Die *grafische Expertise* baut auf die Materialität der geometrischen Formen und eines Ordnungssystems auf und hat daraus ein visuelles Vokabular entwickelt. Visualisierungen können etwas sichtbar machen, was sonst umständlich zu erklären ist oder gar ganz verborgen bleibt. Mit den Abbildungen sind bestimmte Konventionen verbunden, von denen die Expertise ausgeht, dass sie zu einem gewünschten Erfolg der Kommunikation beitragen.

Abbildung 2 ordnet die beschriebenen Elemente aus der technischen, der journalistischen und grafischen Expertise jeweils in eine Triade aus Materialität, epistemologischer und professioneller Kultur ein. Das Dreieck der Triaden wird als Schema auch im weiteren Fortlauf der Arbeit als Gerüst dienen, um die ethnografischen Beobachtungen einzusortieren.

Abbildung 2: Die drei Expertisen und ihre Materialität, ihre epistemologische Kultur und ihre Professionskultur.



(eigene Darstellung)

Die vielfältigen versammelten Expertisen stehen in einer Wechselwirkung mit den Narrativen im Feld: Erstere bestimmen die zweiteren, indem sie einen gewissen Rahmen dafür setzen, welche Erfolgsgeschichten möglich scheinen. Gleichzeitig können Erfolgsgeschichten zur weiteren Vergrösserung oder Verkleinerung eines Feldes führen. Der folgende Abschnitt 3.3 wird diese verschiedenen Narrative näher beleuchten.

3.3 Zirkulierende Narrative

Die vorangehenden Abschnitte haben Datenjournalismus als Community sowie Datenjournalismus als Triade der drei Expertisen Grafik, Programmierung und Journalismus eingeordnet. Verschiedene empirische Ausschnitte haben bereits angedeutet, dass Teams und Exponent*innen mit dem Datenjournalismus unterschiedliche Schwerpunkte verbinden: Für manche ist es eher die Recherche mit Daten, welche Datenjournalismus ausmacht; für andere gehört das Mitdenken der visuellen Darstellung von Anfang an dazu – und die Teams sind in ihren Redaktionen sehr unterschiedlich verortet.

Am Ende von Abschnitt 3.1 hat ein Zitat von Zelizer darauf hingewiesen: Die Community erzählt sich Erfolgsgeschichten, welche informell zirkulieren, sich wiederholen und auch wieder ändern. Diese Narrative bestimmen den Sinngebungshorizont in einem Feld. Mit Bezug auf das Feld als analytische Einheit bietet sich das Konzept des »editing success« (Sahlin-Andersson 1996) an. Es vermag Zelizers theoretische Überlegungen zu präzisieren, indem es Mechanismen aufzeigt, wie und warum bestimmte Narrative in Praktiken umgesetzt werden. Dies ist insbesondere wertvoll, wenn das Feld – wie im Falle des Datenjournalismus – über verschiedene, heterogene Narrative verfügt, wie dieser Abschnitt zeigen wird.

Kerstin Sahlin-Andersson beschreibt das Konzept in ihrem Aufsatz »Imitating by Editing Success: The Construction of Organizational Fields« (Sahlin-Andersson 1996). Aus der Organisationssoziologie kommend geht sie dabei von der Frage aus, ob und wieso Organisationen ähnliche Praktiken entwickeln. Der Schlüssel zur Antwort liegt in den Narrativen, die in einem Feld zirkulieren und dann von den Organisationen nicht imitiert, sondern editiert werden.

Was meint die Autorin mit »editing«? Als erster Schritt des Prozesses konstatieren Organisationen eine eigene problematische Situation: »One common definition of a problem is the difference between a desired state and a present state. The perception of a problem is thus a result of a comparison« (Sahlin-Andersson 1996:71). Diese Prämisse erscheint folgerichtig: Organisationen in einem Feld befinden sich in einem ständigen Austausch und damit auch impliziten Vergleich ihrer Arbeit. Auch im Datenjournalismus: Vorherige Abschnitte haben auf Peer-Review-Prozesse an Treffen und auf Twitter, Preisverleihungen, Konferenzen – um hier nur die wichtigsten zu

nennen – hingewiesen. Diese Plattformen⁷² bedienen Datenjournalist*innen stetig mit Projekten und Ideen darüber, was (guter) Datenjournalismus ist⁷³ – Czarniawska und Joerges (1996) haben es »Travels of Ideas« genannt. Die folgenden Abschnitte ordnen diese Projekte und Ideen in Narrative ein. Dabei geht es auch darum, zu fragen, ob sich die von Sahlin-Andersson beschriebenen Formungen der Narrative hin zu Prototypen oder Rezepten beobachten lassen, sowie die Begründungen und Legitimationsstrategien zu erfassen.

In einem darauffolgenden Abschnitt 3.3.2 werde ich erläutern, weshalb Organisationen gewisse Narrative und deren Projekte zu Erfolgsgeschichten machen (und andere nicht). Und wie die theoretischen Überlegungen darlegen, dass Organisationen die Narrative dann in datenjournalistische Praktiken übersetzen.

3.3.1 Narrative über Datenjournalismus

Für die Einordnung und Beschreibung der Narrative greife ich auf dasjenige empirische Material über das Feld und seine Community zurück, welches bereits die vorherigen Abschnitte angeleitet hat, und werde es dementsprechend referenzieren. Gleichzeitig kommen ergänzende Schlüsse aus dem Material dazu, die spezifisch auf die Narrative Bezug nehmen.

Die folgenden sechs Narrative sind nicht als sich ausschliessend zu betrachten, sondern zirkulieren im Feld parallel, sich überkreuzend und zeitlich versetzt. Akteur*innen können sich durchaus zu mehreren Narrativen hingezogen fühlen. Allen gemein ist, dass sie zum Begriff Datenjournalismus, nach dem sich mein Material gerichtet hat, Bezug nehmen. Sie sind rudimentär geordnet nach Präsenz in den Diskussionen – ohne den Anspruch zu haben, dies genau auszählen zu können.

72 Natürlich kreieren auch Beratungs-, Forschungs- und Branchenorganisationen von ausserhalb des Feldes Narrative mit Erfolgsgeschichten über Datenjournalismus – und teilweise ist die Feld-Grenze auch schwer zu ziehen aufgrund der engen Verflechtungen. Der Fokus dieser Arbeit liegt aber auf Narrativen, die innerhalb des Feldes über das Feld selbst erzählt werden.

73 Exemplarisch dafür ist die »Top 10 in Data Journalism«-Zusammenstellung des Global Investigative Journalism Network. Fussnote 3 (Kapitel 2) hat diese bereits vorgestellt – und gleichzeitig erläutert, dass sie aufgrund der intransparenten Methodik für die eigene Erforschung des Feldes nicht genügend belastbar war.
<https://gijn.org/series/top-10-data-journalism-links/> (18.11.2020)

Neuartige, weil datengetriebene Geschichten

Die meistgenutzte Erzählung im Feld verweist darauf, dass man dank datenjournalistischen Praktiken neuartige Geschichten *findet*. Geschichten, die es ohne die Nutzung digitaler Daten und neuer Werkzeuge nicht geben würde. Die Aussage eines Interviewpartners illustriert das Narrativ:

»Like a lot of... my favorite pieces of data journalism don't have any graphics at all. They're just like... really good reporting that makes more use of numbers. And I think that this is gonna a specially great moment for that, because... there are just... techniques available... that people wouldn't have considered before. Like the one available to traditional reporters. Like- I forget who did it. But my all time favorite... data reporting piece... was... some representative was like flying around the country with private jets with [unverständlich] money, and AP busted him because his like Insta post location tags. And that's like... that's data journalism. [...] So there's like a whole types of investigation... that didn't use to exist.«⁷⁴

Das Narrativ verweist darauf, dass digitale Daten als neue Quelle und somit neues Evidenzobjekt neue Schlüsse zulassen – Abschnitt 3.2.3 hat diese epistemologischen Hintergründe bereits ausgeleuchtet. Die Bedeutung von Datenvisualisierungen als Präsentationsmodus wird dabei in den Hintergrund gerückt, wie der erste Satz des Zitats verdeutlicht. Eine andere Interviewpartnerin möchte den Begriff Datenjournalismus für Projekte vorbehalten, welche eher *data-driven* und weniger *story-driven* sind:

»Also wenn du... irgendwie die Arbeitslosenquote zeigst in den letzten zwölf Monaten und dass, und dir schnell die Daten dafür raussuchst von irgendwie dem statistischen Bundesamt, und... und die schnell plottest, innerhalb von einer halben Stunde, und das in einen Artikel packst, dann find ich nicht, dass das Datenjournalismus genannt werden sollte. Ich weiss nicht wie es genannt werden sollte, ich finde es verdient einen Begriff, aber ich glaube der Begriff Datenjournalismus ist... eher reserviert für... für die Leute die wirklich mit Daten Journalismus machen, also die aus den Daten etwas Neues... und journalistisch Wertvolles herausbekommen.«⁷⁵

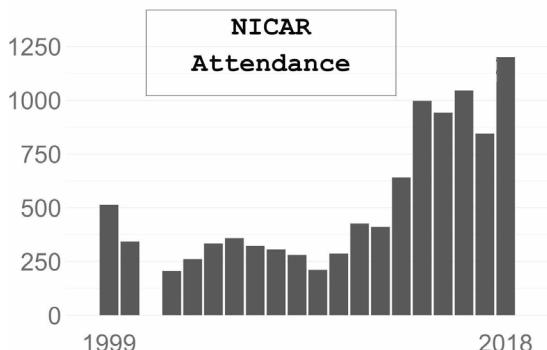
Dieses Narrativ bewertet Datenjournalismus als etwas, das dank seinen Werkzeugen die Welt aus einer neuen Perspektive beleuchten soll, ein neuar-

74 Interview D18

75 Interview D11

tiger Sinngebungsprozess auf Basis der Daten als Quellen steht im Zentrum. Das Narrativ schliesst dabei an Grays epistemologische Metapher der »Data Worlds« an. Diese beruht darauf, dass Datenpraktiken neue Möglichkeiten für »making things up« aufbringen, indem sie spezifische Dinge sagbar machen und neue Formen des Argumentierens nach sich ziehen (Gray 2018:7). Dies bringt veränderte Nachrichtengegenstände, Technologien, Methoden, Software-Programme und ästhetische Zugänge mit sich.

Abbildung 3: Visualisierung der Anzahl an Teilnehmenden an der NICAR, der bekanntesten Datenjournalismus-Konferenz, 1999 bis 2018 (2001 fand keine Konferenz statt). Die Zahlen der ersten Austragungen (1993, 1994, 1995, 1997, 1998) sowie von 2019 und 2020 bleiben unbekannt.



<https://twitter.com/christinezhang/status/971765540439064578/photo/1> (09.01.2021)

Diese Erzählungen des Narrativs wie auch diese Einordnung bleiben sehr allgemein und schwer eingrenzbar: eine neue Perspektive, die vorher nicht möglich war – und das, obwohl Philip Meyer mit seinem »precision journalism« (2002, im Original 1973 erschienen) bereits vor fünfzig Jahren einen Journalismus skizziert hat, welcher computergestützt Muster aufzudecken versucht und dafür auf quantitative positivistische Verfahren baut (in den folgenden Jahrzehnten wurde diese Journalismusform dann meist mit *computer-assisted reporting* beziehungsweise CAR adressiert). Bis in die späten Nuller-

jahre blieb dies aber ein wenig beachtetes Nischenphänomen. Stellvertretend sei hier auf die NICAR-Konferenz (siehe Abschnitt 3.1.5) verwiesen, welche ihr Volumen an Teilnehmenden in den letzten Jahren mehr als verdoppelte (Abbildung 3).

Das Narrativ der neuartigen, datengetriebenen Geschichten verbreitete sich also erst in den 2010er-Jahren wie ein Lauffeuer – es scheint, als hätte es die Kombination mit weiteren Narrativen und deren Erfolgsgeschichten gebraucht, um die journalistische Arbeit mit Daten breiter zu verankern und als eigenes Feld entstehen zu lassen. Es folgt die Beschreibung weiterer Narrative, die folglich sehr oft mit dem soeben beschriebenen ersten kombiniert sind.

Watchdogging Big Data

Im Feld oft erzählt wird ein ausgeprägtes *Watchdog*-Narrativ über Datenjournalismus. Dieses erzählt, dass der Datenjournalismus in einer Welt, die immer stärker datafiziert ist (»Big Data«) und »immer komplexer«⁷⁶ wird, die vom Journalismus selbst zugeschriebene Aufsichts-, eben Wachhund-Aufgabe über die digitale Welt wahrnimmt (Hamilton und Turner 2009; Lesage und Hackett 2014:48). »Datenkompetenz« wird zu einem wichtigen, sogar essenziellen Bestandteil der investigativen Recherche überhaupt.⁷⁷ In besonders enger Verbindung mit diesem Narrativ steht die Idee, dass man Datensätze möglichst selbst erheben soll, weil sie dann als Quelle »unverbraucht« und »besser kontrollierbar« sind (siehe Ausführungen zu Datensätzen im Abschnitt 3.2.2).

Insbesondere die Investigation der Konstruktion und der Konsequenzen von algorithmengestützten Entscheidungen rückt das Narrativ in den Fokus:

»[...], es [das Befassen mit den Technologien an sich, Anm. RE] muss ein Teil vom Journalismus sein den wir besser beackern. Zu verstehen wie funktioniert Google, zu verstehen wie funktioniert Facebook mit der Information die da vorhanden ist, das ist ja nicht wahnsinnig viel. Ehm, und bis hin zu Algorithmen die von Unternehmen privatwirtschaftlich und- privatwirtschaftlicher Natur verkauft werden an die öffentliche Hand.«⁷⁸

Die Aussage der Datenjournalistin stellt die Investigation von Technologien zusätzlich in einen Kontext eines Machtgefälles: private Konzerne, welche die

76 Interview D9

77 Interview D10

78 Interview D19

weitverbreiteten Informationsplattformen betreiben und so einen »new digital divide« mit asymmetrisch verteilter Macht provozieren (boyd und Crawford 2012:673ff, siehe auch Diakopoulos 2015). Neben diesen verschlossenen Mediennplattformen gibt es weitere digitale Datensätze, welche vom (Daten-)Journalismus an die Öffentlichkeit gebracht werden sollen. Das Stichwort lautet hier gemeinhin *Leaks*, wie ein Ausbildungsleiter ausführt:

»Mich faszinieren Bewegungen im Journalismus, die es erlauben, Know-how zu den Journalisten zu bringen, die dann wiederum auch wirtschaftlich-technologische Prozesse kritisch überwachen können. Wenn man schaut, was jetzt läuft mit den globalen Steueroptimierungsmechanismen, dann hat man erst jetzt geschafft, mit international zusammenarbeitenden Journalismus-Konsortien, mit Programmierung grosser Datenbanken, mit den *Leaks*, diese Abläufe der mächtigen und reichen Leute, die damit Geld sparen, das dann in Krippen und Schulen und sonstwo fehlt, diese überhaupt zu begreifen, kritisch zur Diskussion zu stellen, der breiten Menge zu erzählen was hier läuft. Und damit einem politischen Prozess zuzuführen. Und das gleich beim Datenjournalismus: Langsam rekonstruieren, was Facebook, Google, was algorithmisch gesteuerte Technologie kann. Und Journalisten bekommen damit ein Wissen, um ihre eigene Rolle, nämlich das kritische Hinterfragen und der Öffentlichkeit verständlich machen, was eigentlich läuft, so dass politische Antworten der Bevölkerung möglich werden, die dann sagen kann, das wollen wir oder das wollen wir nicht. Das ist total lässig, wenn ich merke, wir holen wieder etwas auf!«⁷⁹

Journalismus insgesamt rahmt der Interviewpartner als den »Abläufe[n] der mächtigen und reichen Leute« ständig hinterherhinkend – und *Leaks* deshalb insofern legitim, weil sie diesen Rückstand etwas verkleinern. Das bekannteste *Leak*, die Panama Papers (veröffentlicht 2016)⁸⁰, gehört zu den meistreferenzierten Projekten des datenjournalistischen Feldes⁸¹, sie sind zweifellos eine von dessen prägenden Erfolgsgeschichten.⁸² Das damit verbundene ›Watch-

79 Interview D2

80 <https://www.icij.org/investigations/panama-papers/> (12.11.2020)

81 Interview D2; Interview D6; Interview D7; Interview D11; Interview D19

82 Zweifellos ist es so, dass es bereits vor den Panama Papers aufsehenerregende Datenlecks gab. So sammelt und enthüllt beispielsweise die Plattform WikiLeaks seit 2006 regelmässig als geheim klassifizierte Datensätze. Jedoch scheinen diese Lecks für die datenjournalistische Community von weniger Bedeutung zu sein, tauchen sie doch in den Community-Diskussionen kaum auf.

dog-Narrativ prägte das Feld stark – und dies betraf nicht nur die Außenwahrnehmung, wie man aufgrund der grossen Aufmerksamkeit für die Panama Papers über das datenjournalistische Feld hinaus vermuten könnte. Auch innerhalb des Feldes gelten derart ausgerichtete Projekte besonders oft als Erfolgsgeschichten: Loosen et al. (2017:15) haben in ihrer Untersuchung der Data Journalism Awards gezeigt, »that awarded projects are slightly more sophisticated and oriented towards fulfilling journalism's watchdog role through investigation and scrutinising those in power.«

Das *Watchdog*-Narrativ wird von der datenjournalistischen Literatur vielfach aufgegriffen und beschrieben – in verschiedenen Nuancierungen: siehe beispielsweise Gynnilds (2013) »newsroom approach«; Stavelins (2013) »watchdogging in code«; Borges-Reys (2017) Systematisierung als »newshound paradigm«. Sie sehen diesen Ansatz als Fortsetzung der journalistischen Tradition.

Datenvisualisierungen als effektives Storytelling-Element

Ein weiteres Narrativ erzählt, dass Datenjournalismus die Grundlage ist, um mittels Datenvisualisierungen ein überzeugendes Abbildungsformat von Information zu nutzen, die Erkenntnisse ermöglicht, die ohne Visualisierungen nicht möglich wären. Sie zeigen etwas auf, was man sonst nicht sieht, und/oder machen ein Phänomen verständlich. Ein Datenjournalist erklärt:

»I think, ehm, the best projects are the projects are, ehm, where you can point at the problem, or look at the problem from data itself. And in a lot of cases it hasn't been done before. So you might have a lot of ehm, journalists writing about it, but without a solid data base behind it. Or you might have some editorial people writing about what they think is going on. But then, when you come up with real data, and you have good visuals, ehm, explaining what your problem is, I think that's the most valuable, the most valuable part of data journalism. So... showing things with visuals, with data visualization, that, ehm, was not visible before. Because people were not using data, or not using data visualization in a proper way to point things out.«⁸³

Neben der starken Evidenzbasis, welche das Feld dem Datenjournalismus und damit auch den Datenvisualisierungen generell zuschreibt (siehe nachfolgendes Narrativ), werden sie mit einer starken Überzeugungskraft und Autorität begründet. Ein Interviewpartner, der bereits mehrere Visualisierungswerkzeuge im Journalismus mitentwickelt hat, verbindet die starke Wirkung

83 Interview D5

von Datenvisualisierungen mit Kognitionsprozessen – sie sind besonders überzeugend und wirkungsvoll, weil man etwas *sieht*:

»Visualisierung hat halt einfach den Vorteil, dass man sie schneller wahrnehmen kann als Texte. [lange Pause] Ich könnte jetzt noch tiefer reingehen, wieso das so ist. Wir speichern ja Information als mentale Bilder in unserem Kopf. Und Visualisierungen sind dann quasi externe Bilder, die wir aber genauso schnell lesen können, und wir merken uns die räumliche Verteilung einfacher. Ich hab mal gelesen, dass es eine Hochgeschwindigkeits-Datenleitung ins Gehirn ist. Man kann halt viel mehr Sachen gleichzeitig, man kann hervorgehobene Sachen viel schneller gegen Text vergleichen, man kann hundert Jahre in Zahlen darstellen und ich sehe sie sofort in einer Linie, und einen Trend darüber. Wir sind Muster-Erkennungs-Maschinen, und da sind wir richtig gut drin, dafür sind unsere Augen und unser Gehirn gemacht. Und wenn man eine Visualisierung benutzt, kann man diese Fähigkeiten ausnutzen, die wir sowieso schon haben.«⁸⁴

Auch für dieses Narrativ gibt es bestimmte Erfolgsgeschichten, welche als besonders gelungene Beispiele referenziert werden. Auf die unzähligen Jahresbestenlisten für Visualisierungen hat bereits Fussnote 41 (Kapitel 3) hingewiesen. Die Grafik- und Design-Kultur sieht sich als die Fortsetzung einer weit zurückgehenden genealogischen Linie – und bringt Projekte aus einer Zeit, bevor man computergestützt visualisierte, als Erfolgsgeschichten in den Datenjournalismus ein. So verweist der vom datenjournalistischen Feld oft gelesene Tufte auf Minards Visualisierung von Napoleons Russlandfeldzug⁸⁵, gezeichnet 1869 (Tufte 1983:40). Das Data Journalism Handbook wiederum erzählt: »In 1854 John Snow created his now famous ›Cholera Map of London‹ by adding a small black bar over each address where an incident was reported. Over time, an obvious density of the outbreak could be seen and action taken to curb the problem.« (Gray et al. 2012:201) – siehe Abbildung 4.

Kennedy et al. (2016) haben gezeigt, wie Redaktionen spezifische Konventionen⁸⁶ aus der grafischen Expertise für ihre datenjournalistischen Projekte nutzen, um den journalistischen Ansprüchen nach Objektivität, Transparenz

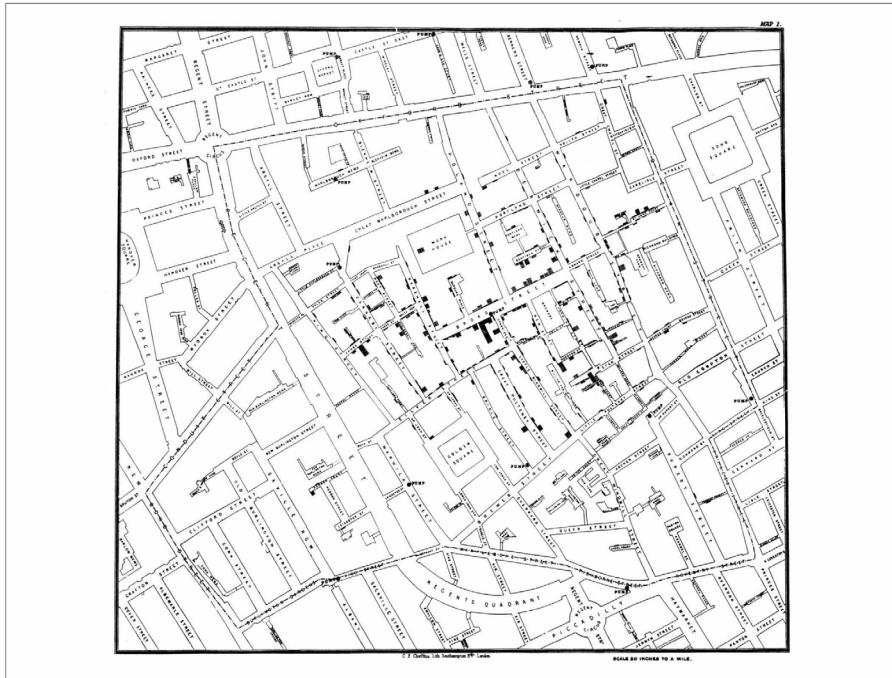
84 Interview D3

85 <https://commons.wikimedia.org/wiki/File:Minard.png> (14.11.2020)

86 Es handelt sich dabei um die teilweise bereits in vorherigen Kapiteln erwähnten Konventionen »(a) two-dimensional viewpoints; (b) clean layouts; (c) geometric shapes and lines; (d) the inclusion of data sources.« Kennedy et al. 2016:716.

und Faktizität gerecht zu werden: Datenvisualisierungen erscheinen dadurch ›neutral‹ und weisen eine starke Überzeugungskraft auf, so die Autor*innen.

Abbildung 4: John Snows Cholera-Karte von London, abgebildet im Data Journalism Handbook.



(Gray et al. 2012:202)

Ein grosser Teil des Feldes ist überzeugt, dass es für die effektive Kommunikation durch Datenvisualisierungen weiterer Präparationen bedarf. Ein Interviewpartner verweist auf den »roten Faden«, dem eine Geschichte – und darin eingebettet auch eine Datenvisualisierung – folgen soll.⁸⁷ Das Zeigen von Ambiguitäten und Unsicherheiten oder interaktive Elemente haben sich

87 Interview D10

dieser journalistischen Logik des *Storytelling* unterzuordnen.⁸⁸ Eine grosse Bedeutung nehmen dabei Instrumente wie Annotationen und Lesehilfen ein; darauf verweist das Feld immer wieder. Eine Interviewpartnerin erklärt, weshalb sie dieses Rezept bei Datenvisualisierungen immer wieder einsetzt:

»Kuck hier, das ist hier spannend, ich erkläre dir die Daten direkt im Diagramm, ich gebe dem Ganzen eine gute Headline, ne gute Description, und mehrere Annotations irgendwie wo ich sage: Hier dieser Peak in diesem Line- Liniendiagramm erklärt sich dadurch, dass. [...] Also in fast allen Daten-, eh, Chart-Typen kann man, ehm, kann man bestimmte Elemente hervorheben. Und, und das ist wieder dieses, also dieses: Kuck dort hin, das ist hier das Wichtigste, das ist die wichtigste Aussage des Charts. Was für den Journalismus so wichtig ist.«⁸⁹

Sowohl die im Zitat beschriebenen Hervorhebungstechniken – diese Arbeit wird sie später als Evidenzbefestigung einordnen – wie auch die Produktion von Erkenntnissen aus Datenvisualisierungen, beispielsweise dass sich die Cholera-Fälle rund um eine verseuchte Wasserstelle gruppieren, benötigen ein visuelles Vokabular. Abschnitt 3.2.2 hat die Bedeutung dieses Vokabulars bereits angedeutet. Abschnitt 4.2 nimmt diesen *roten Faden* wieder auf und wird theoretische Grundlagen ausbreiten.

Der Schluss dieses Abschnitts verweist nochmals auf das grosse Gewicht, das Datenvisualisierungen im datenjournalistischen Feld einnehmen. So erzählt ein Ausbildungsleiter, der das renommierte Lede Program (siehe Abschnitt Ausbildung in 3.2.1) weiterentwickelt hat, wieso Visualisierung inzwischen ein wichtiger Teil des Curriculums ist:

»And... when we first started, the first year we did, or the first year I was in charge, we didn't know visualization, zero visualization. There was all data analysis, all number crunching. All trying to go from... ehm, we wanna answer this question, and this dataset to how do we get the information we need out of that dataset. It turns out that it's so hard to get a job as a data

⁸⁸ Ob sich die Datenjournalist*innen in diesem Grundsatz auf ihr ›Bauchgefühl‹ oder auf Untersuchungen der Rezeption von Datenvisualisierungen stützen, bleibt unklar. Auf jeden Fall liesse sich dieser Grundsatz durchaus mit wissenschaftlichen Studien stützen, siehe beispielsweise de Haan et al. 2017:1293, die beschreiben, dass Lesende Infografiken vor allem dann begrüssen, wenn »they are coherently integrated into a news story and thus fulfill a function that can be easily understood.«

⁸⁹ Interview D11

journalist if you can't make pretty pictures! [wir lachen] In the end of the day, even no... you can do a ton of work with data, no one believes that happened unless they see a chart, unless they see a map, unless they see a graph. [...] It's like: Everyone wants a map and everything, even if a map is not the right way to show the data. Maybe they should just see a chart or something. Ehm... or not even a chart, like a table, a spreadsheet, a tiny spreadsheet of information, everyone just loves to actually *see* that, cause it feels like something happened, if you see it.«⁹⁰

Es hat sich also im Feld die Idee verankert, dass sich ohne Visualisierungen datenjournalistische Projekte in Medienorganisationen kaum lohnen. Wie im Zitat angedeutet, ist dieses Narrativ mit der Etablierung des Feldes über die letzten Jahre tendenziell weiter verbreitet und gestärkt worden – auch die journalistische Forschungsliteratur liefert dazu Indizien (für eine Übersicht siehe Engebretsen et al. 2018:13).

Valide und faktenbasierte Erkenntnisse dank Datenjournalismus

Datenjournalismus wird in einem weiteren Narrativ als eine der Wissenschaft angelehnte Art, journalistische Erkenntnisse zu produzieren, gerahmt. Diese Rahmung erfolgt in zwei Aspekten.

Erstens greift das Feld in seiner Selbstbeschreibung der Kultur und seiner Methoden auf viele Begriffe und Verfahren aus der Wissenschaft zurück. Diverse Abschnitte in 3.2 haben diese bereits erwähnt: die Verknüpfung von Legitimität und Validität eines Schlusses mit einem transparenten und überprüfbaren Verfahren; die Betonung von Peer-Review als Qualitätsverbesserungsmerkmal; die Nutzung von in den Wissenschaften entwickelten Analyse-Methoden. Ein Datenjournalist beschreibt dies:

»Ehm, und deswegen ist es wichtig, dass man das von Anfang bis Ende irgendwie, ehm, möglichst programmiert, bearbeiten kann, dass man, ja, verlässliche, reproduzier- reproduzierbare und auch exakte Ergebnisse hat. Wenn man die Zeit dazu hat, ehm... kann man das natürlich am Ende auch noch publizieren, also wie der, der SRF das vorbild- vorbildlich macht, auf Github. Also die komplette Methodik, alle Rohdaten, und dann sind wir wieder bei dem Punkt, wo man auch sagen kann bei gesellschaftlich wichtigen Themen, oder bei, bei kritischen Recherchen kann ich auf die Art und Weise die, die Transparenz und das Vertrauen erhöhen, in dem

ich sag: Hier liegt wirklich alles an Rohdaten, jeder einzelne Schritt. Es ist vielleicht nicht für jeden, nicht für jeden wieder durchführbar, aber es ist transparent dokumentiert, und dann bin ich so auf nem Standard, der so in der, der Open-Science-Bewegung auch da ist. Dass ich sagen kann: Meine Wissenschaft, bei uns jetzt meine kleinen Studien, meine kleine... Arbeit die ich jetzt hier publiziere, ist eh nachprüfbar. Und hat dann auch so 'nen, so 'nen Qualitätsstandard, wo ich sagen kann: Das was ich hier mache, ist... dann richtig sauberes Handwerk. Und dafür brauch ich diese, diese handwerklichen Grundlagen. Ehm und das find ich wichtig dass wir glaub ich im Datenjournalismus inzwischen da sind, dass die, die es ein paar Jahre machen, und die ein bisschen mehr Erfahrung haben, dass die auf dem Level angekommen sind. Das betone ich deswegen, weil es vor ein paar Jahren nicht selbstverständlich war. Ehm... Und da sind dann glaub ich schon, ja, hoffentlich einige Parallelen mit der, mit der Data Science, die hoffentlich so ähnlich arbeitet.«⁹¹

Das datenjournalistische Feld – und dies ist der zweite Aspekt – pflegt das Narrativ der Wissenschaftlichkeit, indem es das eigene Evidenzobjekt mit höherer Validität bewertet als andere Grundlagen für journalistische Erkenntnisse: »Validitäts-Kultur, das ist genau eben dieses Fakten-basierte, Evidenz-basierte Denken aus der Wissenschaft und auch im Datenjournalismus, das wir versuchen voranzutreiben«, meint eine Datenjournalistin.⁹² Folgerichtig werden Datensätze als Grundlage mehrmals mit dem Begriff der Fakten verbunden. Interviewpartner*innen erwähnen Vermutungen, Vorwissen, Meinungen oder Bauchgefühl, die man mit Datenanalysen validieren könne.⁹³ Einer der untersuchten Ausbildungskurse expliziert diese Abgrenzung mittels der Unterscheidung von »datengetriebene[m] Journalismus« und »anekdotengetriebene[m] Journalismus« und beschreibt Datenjournalismus mit der im Abschnitt zur Epistemologie (3.2.3) bereits erwähnten »Heli-Ansicht«.⁹⁴

Dieser ›Blick von oben‹ ist im Zusammenhang zu sehen mit der häufigen Nutzung von Datenvisualisierungen. Ihnen wird zugeschrieben, dass sie eine starke Evidenzbasis verkörpern und kommunizieren (Weber 2018) und dass sie mit ihren Konventionen der zweidimensionalen Perspektive (Kennedy et

91 Interview D12

92 Interview D4

93 Interview D8; Interview D11; Interview D12; Interview D14; Interview D18

94 UZH_1004_Folie64f

al. 2016:723f) und ihrer »god's eye view« (Rettberg 2020:39) Objektivität kreieren. Konsequenterweise argumentieren gewisse Akteur*innen, Datenjournalismus sei eine Strategie von Medienorganisationen »to regain public trust« (Stalph und Borges-Rey 2018:1078).

Anderson (2018) spricht bereits mit dem Titel seines Buches – »Apostles of Certainty« – das starke Narrativ des Feldes an, welches datenjournalistische Erkenntnisse mit Tatsachen und Fakten quasi gleichsetzt. Der Autor beleuchtet das Verhältnis zwischen (Daten-)Journalismus und Sozialwissenschaften über mehr als ein Jahrhundert und mahnt zum Schluss an, dass Datenjournalist*innen mehr Zweifel und Unsicherheit zulassen sollten, so wie es auch die Wissenschaft tue.

Datenjournalismus als Experimentierfeld und Ort der Innovation

Datenjournalismus-Teams gelten in Redaktionen als Orte, an denen die Organisation neue, bisher ungewohnte Expertisen und damit verbundene neue *Mindsets* und Kulturen aufnehmen und eingliedern, also *innovativ* sein kann. Ein Datenjournalist, der bereits seit mehreren Jahren in einem solchen Team arbeitet, erzählt:

»Und gleichzeitig ist Datenjournalismus ja nicht so scharf abgegrenzt und auch immer viel was, visueller Journalismus ist, ehm... und Innovation, ehm, passiert bei uns auch mit. Also Sachen wie Personalisierung von Geschichten, oder interaktive Elemente, manchmal auch einfach nur ein Video, oder, oder Animation, oder neue Formate ausprobieren, also bei vielem was wir machen fällt am Ende auch fürs Haus was bei ab im Sinne von: wir probieren eh neue Techniken, Tools, Artikelformate und so weiter aus, in unseren Stücken, die wiederum für, für ganz, für ganz, ganz normale, für alle Formen des Journalismus, die wir machen, eh auch wieder, auch wieder nützlich sein können, also es ist auch so ein Experimentierfeld, wo, wo das Haus auch was lernen kann. Ich glaub das ist auch für viele noch ein Argument, dass man in dem Bereich stärkt, wenn man, wer sich Innovation, und eh... und mal das eh Outside-of-the-Box-Denken so, so ein bisschen davon verspricht.«⁹⁵

Aufgrund der in Abschnitt 3.2.1 beschriebenen Kultur ist es folgerichtig, dass neue methodische Ideen im datenjournalistischen Feld eine hohe Anerkennung geniessen.⁹⁶ Insbesondere grosse Projekte seien eine Möglichkeit für

95 Interview D12

96 UZH_0927_Folie07

Individuen und Teams, Neues auszuprobieren und sich weiterzuentwickeln, erzählt ein Dozent in einer Ausbildungslektion.⁹⁷ Dies ist denn auch ein wichtiges Element, um sie zu den bereits in den vorherigen Abschnitten erwähnten *Leuchtturm*-Projekten zu machen.

Was mit der Kombination von programmiertechnischer und grafischer Expertise hin zum Datenjournalismus begann, soll auch künftig offen sein für die Kombination mit weiteren Expertisen, so das Feld – datenjournalistische Teams sollen ein »Experimentierfeld« bleiben.⁹⁸ In einem Team zu arbeiten, das mehr Zeit und Ressourcen zum Ausprobieren erhält, erscheint für die Community ein Privileg, das man gerne behalten würde. Gleichzeitig scheint sich dies zu reiben mit dem Wunsch, (noch) mehr mit anderen Ressorts zusammenzuarbeiten (Abschnitt 3.2.3). So kommen denn im Feld verschiedene Stimmen auf, welche kritisch anmerken, dass die Zeit der *Leuchtturm*-Projekte langsam vorbei sei, Datenjournalismus müsse in den redaktionellen Alltag einfließen.⁹⁹

Datenjournalismus mit dem gesamten Feld des Journalismus in Bezug setzend, verknüpfen mehrere Interviewpartner*innen ihn mit dem Begriff der *Innovation*. Wie das Zitat zur Einleitung des Abschnitts illustriert, erscheint *Innovation* als Begriff für Redaktionen erstens mit sehr unterschiedlichen Aspekten befüllbar und wird zweitens dennoch hauptsächlich anschlussfähig gemacht an die Expertisen der datenjournalistischen Teams. Gynnild (2013:13ff) hat Datenjournalismus denn auch als einen »door opener to innovation at large« beschrieben.

Aufgrund der technologischen Entwicklungen müssten sich Geschäftsmodelle, Organisationsstrukturen wie auch Inhalte verändern, »change« und »innovation« seien unerlässlich – so lautet seit über einem Jahrzehnt der Imperativ im Journalismus (siehe beispielsweise Kueng 2017). Ein Koordinator einer datenjournalistischen Austauschplattform nimmt das Narrativ explizit auf:

»And I was interested in data journalism specifically, because I think it's one of the, ehm, most interesting and most innovating and most, eh, promising, if you can say that, field, eh areas, in journalism. Ehm. It combines a lot of

97 8_20171115_SeminarUZH-4

98 Der Begriff des Experimentierens fällt in vielen Interviews, teilweise auch mehrfach: Interview D5; Interview D10; Interview D11; Interview D12; Interview D14; Interview D15; Interview D19

99 Interview D11; Interview D18; Interview D19

interesting things, it has a lot of, eh, well, the technical side is interesting as well, eh, something that's new to a lot of journalists I think.«¹⁰⁰

Einige Community-Mitglieder sehen Datenjournalismus durchaus als Möglichkeit, Innovationen in die Redaktionen einzubringen – erwähnen aber auch, dass Medienorganisationen datenjournalistische Teams als »Branding«¹⁰¹ brauchen, um sich als innovativ darzustellen.¹⁰² Tatsächlich ist es so, dass ›Datenjournalismus als Innovation im Journalismus‹ ein Narrativ ist, welches hauptsächlich von Personen gepflegt wird, welche nicht selber datenjournalistische Praktiker*innen sind – beispielsweise der Koordinator, von welchem das obige Zitat stammt. Oder Personen aus den Redaktions- und Verlagsleitungen: »Er glaubt daran, dass das [Datenjournalismus, Anm. RE] EIN wichtiger Bestandteil ist eines erfolgreichen Medienhauses der Zukunft«, sagt ein Datenjournalist über den Tamedia-Chef Pietro Supino.¹⁰³

Gerade das Beispiel Tamedia mit seinem Textroboter Tobi¹⁰⁴ führt zu einem weiteren Narrativ, das mit Datenjournalismus verbunden wird: die Möglichkeit, Prozesse zu automatisieren.

Datenjournalismus als automatisierte Berichterstattung

Während der Corona-Pandemie traf sich die Zürcher Hacks/Hackers-Community online, um verschiedene datenjournalistisch interessante Aspekte der Pandemie zu beleuchten. Einer der Präsentierenden meinte, dass Datenjournalismus »the art of automation« sei – dieser Meinung sei er schon lange gewesen, aber die datenjournalistischen Projekte während der Pandemie hätten dies nun besonders deutlich gezeigt.¹⁰⁵

Automatisierung meint im Feld, dass Redaktionen durch datenjournalistische Praktiken Prozesse der Datensammlung, -auswertung und -visualisierung automatisieren – also die Erkenntnisproduktion standardisieren und

100 Interview D7

101 Interview D5

102 Interview D9; UZH_0927_Folie07

103 Interview D6 – die Fussnoten 8 und 27 des Kapitels 3 haben bereits auf verschiedene Aktivitäten von Tamedia hingewiesen, welche Datenjournalismus fördern. Supino selber hat in einem Gastkommentar in der NZZ betont, dass er Datenjournalismus als eine der grossen Chancen der technologischen Entwicklung im Journalismus sieht (<https://www.nzz.ch/meinung/kommentare/medien-die-chancen-sind-mindestens-so-gross-wie-die-herausforderungen-ld.1316577, 15.11.2020>).

104 <https://interaktiv.tagesanzeiger.ch/2018/tobi/> (15.11.2020)

105 66_20200423_Hacks-Hackers-Covid19

skalierbar machen. Ein Datenjournalist spricht (unter anderem) Automatisierung an, als die Interviewfrage lautet, weshalb sein Team trotz schwieriger Finanzierungslage im Journalismus viele Ressourcen erhält:

»Dann gibt es auch, halt auch, sicher den Gedanken, dass man, eh, gewisse Sachen automatisieren kann. Ehm... Jeder einzelne Journalist kann das. Und Leute die, eh... ein bisschen etwas von Algorithmen verstehen, programmieren können, die können das dann auch skalieren, sie können auch Sachen entwickeln, von denen dann andere profitieren können, dass es automatisiert ist, dass es da eine Automatisierung gibt.«¹⁰⁶

Gerade aus datenjournalistischen Teams, die nicht isoliert, sondern oft mit anderen Ressorts arbeiten, scheint das Narrativ ab und zu auf: Automatisierung durch Datenjournalismus bedeutet für die gesamte Redaktion einen Gewinn des raren Gutes Zeit. An der Jahrestagung 2019 von netzwerk recherche, einem der wichtigsten Treffen des deutschsprachigen Datenjournalismus-Feldes, steht eine Veranstaltung zum Thema »schneller« Datenjournalismus im Programm. Der Einleitungstext lautet:

»Nur für die Galerie, nur für Preisverleihungen, viel zu langsam – das war und ist teilweise immer noch der Vorwurf an den Datenjournalismus. Es kann aber auch anders gehen. Drei Teams von Datenjournalisten präsentieren in einem kleinen Wettbewerb, wie sie unter Zeitdruck datengetriebene, visuelle Geschichten realisieren.«¹⁰⁷

Drei Datenjournalismus-Teams präsentieren anschliessend ihre *Tools*, mit denen sie automatisiert produzieren. Es ist die Rede von »Skripten« und von »Templates«, die »rumliegen« und schnell und einfach einsetzbar sind.¹⁰⁸ Diese seien teilweise auch von anderen Redaktionsmitgliedern autonom nutzbar – aber oftmals gebe es schon viele »edge cases«, die ein vertieftes Wissen über Datensätze und *Tools* benötigten, so einer der präsentierenden Datenjournalisten.

Das nur sporadische Auftauchen des Automatisierungs-Narrativs deutet darauf hin, dass es sich aus Sicht des Feldes nur um eine sekundär anzustrebende Weiterentwicklung handelt. Es gibt im Automatisierungsbereich kaum

¹⁰⁶ Interview D6

¹⁰⁷ <https://nr19.sched.com/event/Lx7A> (15.11.2020)

¹⁰⁸ 63_20190614_nr19-bestes-schnelles-dj-team

vieldiskutierte Erfolgsgeschichten¹⁰⁹, auch gibt es keine Award-Kategorien dafür. Stattdessen scheint es sich eher um eine von aussen aufgedrängte Organisationsweise zu handeln, der die ›Faszination‹ der Exploration unbekannter Datensätze oder der Produktion erinnerungswürdiger Datenvisualisierungen fehlt. Automatisierung vermindert die journalistische Legitimität der selbstbestimmten Konstruktion von Erkenntnissen (Carlson 2018) aus Sicht des datenjournalistischen Feldes zu stark.

3.3.2 Das lokale Editieren von Narrativen

Datenjournalistische Teams als Organisationen betrachtend stellt sich die Frage: Welche der Projekte und Ideen aus welchen Narrativen werden von ihnen als Erfolgsgeschichten bewertet? Sahlin-Andersson (1996:72f) argumentiert, dass diese Bewertung mit der Identität der Organisation zusammenhängt. Die Identität bildet sich relational zu anderen Akteur*innen des Feldes aus, indem sie mit den (geteilten) Narrativen des Feldes interagiert. Die Identität ist gleichzeitig verankert in der Institution und soll Kontinuität gewährleisten. So ist der Vergleich und die Bewertung der Narrative und ihrer Erfolgsgeschichten gerahmt durch die Identität:

»Thus, identity focuses attention both in the sense of directing attention towards certain criteria in the organizational operations, and of helping to determine what the organizational actors will attend to in the environment, i.e. what parts of the environment are seen as relevant. The organizational identity both follows from and is the basis for that with which the organization interacts and is compared.« (Sahlin-Andersson 1996:73)

Die dadurch in den Fokus rückenden Erfolgsgeschichten werden durch die Organisationen aber nicht einfach imitiert, sondern übersetzt.¹¹⁰ Dieser Prozess dekontextualisiert und abstrahiert »success stories« bezüglich Zeiten und Situation und resultiert in »not exact copies but instead something which dif-

¹⁰⁹ Ob die Corona-Pandemie dies ändert, wie die zu Beginn des Abschnitts geschilderte Situation des Hacks/Hackers-Treffens es andeutet, wird sich weisen.

¹¹⁰ Eine wichtige Feststellung in diesem Zusammenhang ist, dass sich das theoretische Modell an dieser Stelle immer noch auf einer diskursiven und nicht auf einer praktischen Ebene befindet. Anders gesagt: Es geht nicht darum, Praktiken oder konkrete Handlungsanleitungen *per se* zu imitieren – sondern Ideen in eigene Praktiken zu übersetzen. Siehe Sahlin-Andersson 1996:78.

fers considerably from the imitated model« (Sahlin-Andersson 1996:80). Die Autorin beschreibt dies als Übersetzungsprozess in einem latourschen Sinne:

»The follower interprets and translates experiences presented by others. Such translations, in turn, enrich the source of the knowledge with new ways of interpreting their experiences. It is a process in which the story of what has happened is constantly rewritten as it spreads. The imitating organization and the imitated prototype are continuously reinterpreted and reformulated. A success story cannot be explained by its origin, as the success is produced in the process of translation.« (Sahlin-Andersson 1996:82)

Die Autorin nutzt dafür das aktive Verb editieren, weil durch die kontinuierliche Reformulierung stets kleine Änderungen auftreten. Diese Variantenproduktion ist geprägt von »social control«: »Unlike Latour's concept of translation, editing is also a process of social control, conformism and traditionalism. Just as numerous rules – explicit and implicit – are applied to various publications, so are various rules applied to success stories.« (Sahlin-Andersson 1996:82) Auf die editierende Organisation bezogen bedeutet das, dass ihre Rahmenbedingungen formen, wie ein Narrativ übersetzt wird. Sie geben die Regeln vor: »They [die Editoren, Anm. RE] also translate their stories in terms of those prescriptions given by the greater institutional context. Editing rules restrict and direct the translation in each phase of the process, and are re-created in each such phase« (Sahlin-Andersson 1996:85).

Es handelt sich bei diesen Regeln eher um die lokalen Strukturen einer organisationalen Umgebung als um eine fixe Handlungsanleitung. Das Zitat spricht zudem von Narrativen in der Mehrzahl, diese können sich im Übersetzungsprozess also auch gegenseitig beeinflussen. Ebenso ist festzuhalten, dass eine Organisation nicht ein homogenes Gebilde ist und die strukturellen Voraussetzungen von Akteur*innen innerhalb der Organisation differieren können.

Mit Bezug auf die Untersuchung des datenjournalistischen Feldes lassen sich, wie bereits erwähnt, die Datenjournalismus-Teams analytisch als Organisationen fassen. Die in ihnen sich versammelnden Expertisen verkörpern als lokale Strukturen die Identität dieser Organisationen. Es stellt sich die Frage, welche Narrative die Teams aufgreifen und in eigene Praktiken übersetzen. Abschnitt 3.3.1 hat heterogene Narrative über Datenjournalismus beschrieben. Diese sind nicht einzelnen Expertisen zuordenbar – auch wenn bei einigen eine dominante Expertise durchscheint: Das *Watchdog*-Narrativ ist angetrie-

ben von einer starken (investigativ-)journalistischen Professionskultur, während das Narrativ der Datenvisualisierungen als *Storytelling*-Element fundiert ist durch die epistemologische Kultur der grafischen Expertise. Die Narrative bieten Erfolgsgeschichten an, welche Anschlussfähigkeit für sehr verschiedene Arten von Teams und deren Expertisen erlauben.

Die Diagnose eines heterogenen, breit gefassten datenjournalistischen Feldes haben Journalismusforscher*innen über die letzten zehn Jahre sehr oft gestellt und sie wird hiermit durch diesen ersten Teil der Arbeit bestätigt. Andere Forscher*innen haben die Diversität mittels Typologien zu fassen versucht, abgegrenzt innerhalb des Feldes mittels verschiedenartiger Kriterien: newsroom vs. academics vs. entrepreneurial approach (Gynnild 2013); computer-assisted reporting, data journalism, computational journalism (Coddington 2015; Hermida und Young 2019); newshound vs. techie paradigm (Borges-Rey 2017); hacker journalist, programmer journalist, data journalist (Usher 2016); ambient journalism, data journalism, algorithmic journalism (Lewis und Zamith 2017) – um hier nur einige Feldbeschreibungen zu nennen. Keine der Typisierungen hat sich in der Forschung als dominant durchgesetzt, genauso wenig wie eine einheitliche Definition des Begriffs. Stattdessen bestätigt sich auch in dieser Literatur die Vielfältigkeit und Vieldeutigkeit des Feldes. Die Antwort auf die Frage, was Datenjournalismus ist, bestimmt folgerichtig jedes Team in seinen eigenen datenjournalistischen Praktiken.

3.4 Fazit und aufgeworfene Fragen

Die Kapitel 5 und 6 werden zwei solche Teams und deren Praktiken ethnografisch erforschen. Die Einordnungen des Kapitels 3 bilden dafür die Grundlagen.

- (1) Abschnitt 3.1 hat Datenjournalismus als Community diskutiert – seine Plattformen sowie sozialen und professionellen Hintergründe; seine formellen Vernetzungen und informellen Verflechtungen; seine Internationalität und deren Konsequenzen. Die Mitglieder der beiden Teams sehen sich selbst als Teil dieser Community und engagieren sich teilweise aktiv in ihren lokalen Hacks/Hackers-Sektionen. Sie sind also Teil des skizzierten geteilten Raumes der Sinngebung und werden gleichzeitig davon geprägt.

- (2) Die Expertisen-Triaden, welche Abschnitt 3.2 aus der Empirie herausgeschält hat, bilden eine wichtige strukturelle Rahmung im Feld: Sie ordnen die beobachteten Elemente in eine materielle, epistemologische und kulturelle Dimension der drei Expertisen ein. Damit lässt sich feinkörnig nachvollziehen, wie die Expertisen prägen, welches die Fokusse und Perspektiven eines Teams sind und wie es seine Praktiken aushandelt – das gilt auch für die beiden untersuchten Teams.
- (3) Die im Abschnitt 3.3 beschriebenen Narrative verweisen auf jeweilige Erfolgsgeschichten, die im Feld zirkulieren. Sie werden lokal in den Medienorganisationen in eigene Praktiken übersetzt. Als Teil des Feldes sind auch die beiden untersuchten Teams sowohl an der Pflege der Narrative und ihrer Erfolgsgeschichten wie auch an deren Übersetzungen beteiligt.

Die verschiedenartigen Einordnungen des Kapitels 3 haben zwei zentrale Aspekte immer wieder durchscheinen lassen:

1. Erstens die Diversität von Expertisen und Narrativen, welche sich im Feld versammeln und für die eine internationale Community wie ein Klebstoff wirkt. Es stellt sich daran anschliessend für den folgenden empirischen Teil die Frage, wie zwischen diesen Expertisen datenjournalistische Praktiken ausgehandelt werden und weshalb dieser Raum trotz der Heterogenität ein stabiles Feld hervorzubringen vermag.
2. Zweitens die koordinierende Rolle von Datenvisualisierungen beziehungsweise grafischer Expertise, wie sie in den Einordnungen immer wieder durchschien: im Aufbau der Hacks/Hackers-Community der Deutschschweiz (3.1.3); geometrische Formen als relationierendes Vokabular (3.2.2); Datenvisualisierungen als Koordination von quantitativ und qualitativ gewonnenen Erkenntnissen (3.2.3); als ein weitverbreitetes Narrativ im Feld (3.3.1). Mit Bezug auf diesen Aspekt stellt sich die Frage, wie diese koordinierende Rolle der Visualisierungen im Datenjournalismus sich genau ausgestaltet. Diese Frage ist besonders wichtig, weil es bisher an Forschungsliteratur fehlt, welche die Rolle von Datenvisualisierungen in den Praktiken des Datenjournalismus explizit und ethnografisch untersucht hat.

Es gilt also, theoretisch konsolidierte Heuristiken einzuführen, welche diese zwei Fragen im ethnografischen Material ausreichend herauszuschälen vermögen. Dieser Aufgabe widmet sich das nächste Kapitel.

4 Heuristiken der Aushandlung und Visualisierung

Die folgenden beiden Abschnitte führen zwei theoretisch von den Science and Technology Studies beziehungsweise der Wissenschaftsforschung inspirierte Konzepte ein, welche in der folgenden ethnografischen Arbeit als Heuristiken dienen, um auf die aufgeworfenen Fragen Antworten zu finden. Es handelt sich dabei um das Konzept der »trading zone« (4.1) sowie die zwei Visualisierungsmodi Daten erkennen und Evidenz befestigen (4.2).

4.1 Die »trading zone«

Wo – wie im Datenjournalismus – verschiedene Expertisen zusammenkommen, um Erkenntnisse zu produzieren, unterliegt dies einem Prozess der Aushandlung und Koordination. Diese Arbeit nutzt im Folgenden die Heuristik der »trading zone«. Sie ist Peter Galisons (1997) theoretisches Destillat aus seiner monumentalen historischen Nachzeichnung des Feldes der Physik, in welchem sich über Jahrzehnte Gruppierungen beziehungsweise Subkulturen der theoretischen und der experimentellen Physik (und ihre jeweiligen Erkenntnisinstrumente) parallel und schliesslich zueinander hin entwickelten. Durch das Aufeinandertreffen dieser verschiedenen Subkulturen in der Physik entsteht ein Raum, in dem sie ihre verschiedenartigen Datengrundlagen, Annahmen, Instrumente und Evidenzansprüche austauschen und koordinieren. Für ihre Kommensurabilität entwickeln die Subkulturen vermittelnde Konzepte, Galison spricht von Austauschregeln sowie »interlanguages«. Er nutzt dafür auch den anthropologischen Begriff der Kreolsprache: eine Mischform, mit Hilfe derer man sich mindestens in gewissen Bereichen versteht und woraus wiederum Erkenntnisse und allenfalls eine neue Subkultur entstehen kann.

Natürlich eignen sich weder das Feld des Datenjournalismus noch die beiden Teams direkt für einen Vergleich, zu unterschiedlich sind die Ausgangs-

gen. Die kurze Beschreibung der *trading zone* zeigt aber, dass sie einen analytischen Rahmen auch für die Aushandlung von Expertisen auf einer Mikro-Ebene bieten kann. Galison (1997:783) beschreibt, dass Subkulturen verschiedene Vorstellungen davon haben, »what entities there were, how they were classified, and how one demonstrated their existence«. Das entwickelte Expertisen-Dreieck (Abbildung 2) skizziert ähnliche Aspekte: Es geht um die Koordination von verschiedenartiger Materialität, epistemologischen Ansprüchen und professioneller Kultur sowie die Aushandlung darauf basierender Praktiken. Zusätzlich zur ursprünglich diskursfokussierten Heuristik der *trading zone* erweitert Anderson mit seinem Expertisen-Begriff die Verhandlungsmasse um eine explizit materielle Perspektive – es geht um die Vernetzung von Welten und Dingen (in seinem Artikel mit Bezug zum Journalismus):

»This article, then, is an attempt to trace the operation of journalistic expertise through an analysis of worlds and things. It concludes that the social and material aspects of journalistic knowledge do not congeal into solid domains of expertise battled over by well-bounded social groups. Rather, it argues that journalistic expertise is a socio-material property of messy networks. To the degree that these structures of expertise momentarily cohere, they do so on the basis of temporary network assemblages that are partly structured by power games, but also structured by different cultural understandings of what constitutes valid digital evidence, and thus what >belongs< in the journalistic network.« (Anderson 2013a:1010)

Auch an Galison anschliessende Arbeiten nutzten den Expertisen-Begriff – zum Beispiel Collins et al. (2007), welche Aushandlungsräume typisieren anhand der Merkmale Homo-/Heterogenität des Feldes sowie ob die Parteien (freiwillig) kollaborieren oder dazu gezwungen sind. Die Autoren weisen darauf hin, dass Galisons Konzept sehr sprachlastig ist und Materialitäten nicht inkludiert, ausser sie sind Teil des Diskurses (Collins et al. 2007:661).

Aufgrund der festgestellten Diversität der Narrative und ihrer Erfolgs geschichten geht meine Arbeit nicht von einer dominanten Expertise und einer erzwungenen Aushandlung aus, sondern von einem »mutual agreement to trade« (Collins et al. 2007:658) innerhalb des datenjournalistischen Feldes. Es stellt sich die Frage, ob es sich bei Datenjournalismus um ein noch heterogenes Feld (mit gemäss Collins entstehender »interactional expertise«) handelt oder ob sich dieses bereits homogenisiert hat und daraus eine neue »interlanguage« entstanden ist.

Ein Anhaltspunkt zur Beantwortung dieser Frage besteht darin, dass »inter-language trading zones operate by developing new cultural tools« (Collins et al. 2007:660) – die Entwicklung neuer kultureller Werkzeuge weist also auf eine stabile ›Sprache‹ hin. Allerdings verbleibt die Aufgabe, die diskurslastige Galison-Konzeption um eine materielle Perspektive zu erweitern – dieser Aufgabe wird sich der nächste Abschnitt widmen.

Zuerst verbleibt festzuhalten: Obwohl also Galisons Konzept von verhandelnden Subkulturen ausgeht, während diese Arbeit eher von Expertisen im Sinne Andersons spricht, so sind die damit referenzierten Prozesse sich genügend ähnlich, um in beiden Fällen die *trading zone* als Heuristik einzusetzen. So nutzen denn auch verschiedene weitere Arbeiten aus der Journalismusforschung den Rahmen der *trading zone*, um das Aufeinandertreffen verschiedener Kulturen beziehungsweise Expertisen als Aushandlung zu betrachten: Journalist*innen und Programmierer*innen (Lewis und Usher 2014, 2016; Hannaford 2015); Foto- und Textjournalist*innen (Lowrey 2002); Journalist*innen und Software-Entwickler*innen (Ananny 2012); Informationsdesigner*innen und ihre Auftraggebenden aus verschiedenen Bereichen (Smit et al. 2014).

4.2 Durch Visualisierungen Daten erkennen und Evidenz befestigen

In Kapitel 3 schien immer wieder die koordinierende Rolle von Datenvisualisierungen beziehungsweise grafischer Expertise durch. Die abgeleitete Hypothese daraus besteht darin, dass Elemente der Materialität (beispielsweise digitale Daten), epistemologische Ansprüche (beispielsweise Objektivität) sowie die professionelle Kultur im Datenjournalismus nur mit einem Fokus auf Datenvisualisierungen vollständig gefasst werden können. Um in der ethnografischen Untersuchung der datenjournalistischen Praktiken diese Vermutung zu untersuchen, benötigt es eine entsprechende Heuristik. Der Blick fällt dabei auf die Science and Technology Studies (STS), welche sich seit Jahrzehnten damit beschäftigen, wie im Zusammenspiel von Wissen, Kultur und Technologien Erkenntnisse entstehen – und welche zur Rolle von Visualisierungen in diesem Prozess entsprechende Konzepte entwickelt haben.¹ Die STS haben ge-

¹ In den folgenden Abschnitten ist meist von ›Visualisierungen‹ (und nicht Datenvisualisierungen) die Rede. Dies deshalb, um zu differenzieren, dass die Laborstudien teilweise wissenschaftliche Visualisierungen analysierten, die nicht auf Zahlen als Da-

zeigt, dass Datenvisualisierungen nicht ignorieren kann, wer die Produktion von Evidenz unter dem epistemologischen Anspruch nach Objektivität untersucht (Daston und Galison 2007). Visualisierungen sind Teil der »Erkenntnismaschinerie« in den Wissenschaften, indem sie »collaborative seeing« organisieren (Amann und Knorr-Cetina 1988:164).

4.2.1 Das Labor als Fabrik der Erkenntnis

»Wir schauen mit den eigenen Augen, wir sehen mit den Augen des Kollektivs«, schrieb der Wissenschaftssoziologe Fleck 1947 (Fleck 1983:154). Rund vier Jahrzehnte danach griffen innerhalb der STS die Laborstudien das Thema der visuellen Erkenntnisproduktion und deren soziale Hintergründe auf, indem sie den Konstruktionsprozess wissenschaftlicher Fakten in den Labors ethnografisch untersuchten. Was mit Latours und Woolgars »Laboratory Life« (1979) begann, führten Knorr-Cetina (1981), Lynch (1988, 2012) und wiederum Latour (2012) weiter, dabei die Rolle von Datenvisualisierungen vertiefter beleuchtend.

Insbesondere die Arbeiten Karin Knorr-Cetinas – mit ihrem Fokus auf epistemische Prozesse und damit verbundene Kommunikationsmodi – erweisen sich für die vorliegenden Fragestellungen als besonders aufschlussreich. Sie charakterisiert Labore als »Orte der Weltkonstruktion« (Knorr-Cetina 1988:320), verstanden (auch) in einem metaphorischen Sinne: Das Labor lässt sich in verschiedensten Disziplinen und in unterschiedlicher Form wiederfinden – es ist überall dort, wo Wirklichkeit erzeugt wird: »Laboratorien sind Instrumente, die in meist lokaler und eingegrenzter Weise auf bestimmte Erzeugungsvorgänge spezialisiert sind. Die Vorstellungen von Laboratorien als Erzeugungsquellen der verschiedenen Wissenschaften ist [sic!] also durchaus wörtlich zu verstehen« (Knorr-Cetina 1988:334f). Weiter charakterisiert Knorr-Cetina das Labor als Ort, zu dem weit mehr als Maschinen und Apparaturen gehören: »Verkörperte Tätigkeit, mündliche Interaktion und Vertextung sind Aspekte allen Forschungshandelns« (Knorr-Cetina 1988:335). Im Folgenden fokussiert sie auf die Mündlichkeit. Sie argumentiert, dass »Mündlichkeit im

ten und auch nicht auf digitalen Datensätzen basieren. Zwar könnte man auch andere Grundlagen als ›Forschungsdaten‹ bezeichnen. Da die Digitalität als Charakteristik der Datenvisualisierungen im Datenjournalismus eine wichtige Rolle spielt, wie später Kapitel 5 nochmals aufzeigen wird, verwende ich hier für andere wissenschaftlich-technische Bilder den Begriff »Visualisierungen«.

Arbeitsbanklabor zum *technischen Instrument gesteigert* wird: in Form von ›shop talk‹, von technischen Gesprächen, ist sie *Instrument der Erzeugung von Wissen aus dem Händischen und Apparativen des Labors*. [alle Hervorhebungen im Original]« (Knorr-Cetina 1988:336). Bereits an dieser Stelle verweist die Autorin darauf, dass Gespräche über Visualisierungen die »unklaren Spuren« darauf erst erkennbar machen, sie interpretiert und in lesbare »Daten« transformiert werden (Knorr-Cetina 1988:336f). Das Gespräch dient »als *Denkprothese*« genauso wie als ›Sinnesprothese‹« (Knorr-Cetina 1988:340).

Erst in ihrer Studie über die Wissenskulturen in den Disziplinen der Molekularbiologie und der Teilchenphysik (Knorr-Cetina 2002, im englischen Original 1999 erschienen) führt Knorr-Cetina den Begriff der Erkenntnismaschinerie ein. Der Begriff wird nicht einheitlich verwendet, sondern in Varianten wie Wissensmaschinerie oder -technologie oder empirische Maschinerie. Knorr-Cetina (2002:13) ordnet ein: »In der vorliegenden Studie bin ich nicht an der Konstruktion von Erkenntnis interessiert, sondern an der Konstruktion der Maschinerien, durch die Erkenntnis konstruiert wird.« Erkenntnismaschinerien sind Vergleichsmethoden, Kriterien, Konventionen, Gewohnheitsregeln, Selbstverständlichkeiten – kulturelle Strukturen, welche die Produktion und Evaluation von Erkenntnissen rahmen.

Man findet den Begriff der Maschinerie bereits in Werken aus den Achtzigerjahren (allerdings ohne genauere Beschreibung) – nämlich als »machinery of seeing« in Klaus Amann und Knorr-Cetinas (1988:138) grundlegendem Papier über die Produktion visueller Evidenz in Laboren. Die beiden verweisen gleich zu Beginn des Papiers auf die Autorität von visueller Information und die damit einhergehende starke Wahrheitszuschreibung:

»Consider that in the natural sciences evidence appears to be embodied in visibility; in a literal sense, it is embodied in what we can see on a data display. Thus understood, the notion of evidence is built upon the difference between what one can see and what one may think, or have heard, or believe. Among these modes of relating to an object, only seeing bestows on objects an accent of truth.« (Amann und Knorr-Cetina 1988:134)

Autor und Autorin nehmen diese Ausgangslage zum Anlass, um das, was (Natur-)Wissenschaftler*innen *sehen*, als Endprodukt eines langen Prozesses des Erkennens und Befestigens von visueller Evidenz zu betrachten. In diesem Prozess nehmen Visualisierungen zwei Funktionen ein: Sie dienen einerseits als Arbeitsoberfläche der Erkenntnismaschinerie, andererseits als Montage, auf welcher die Erkenntnisse in spezifischer Weise befestigt werden, sodass

sie als Evidenz erscheinen. Beide Visualisierungsmodi sind mit einem Set an Praktiken verbunden.

Die zwei Modi und ihre Praktiken bilden als Heuristik das Gerüst für die folgenden Abschnitte und dienen später dazu, die datenjournalistischen Praktiken zu untersuchen.² Die folgenden beiden Abschnitte führen dieses Gerüst ein, indem sie auf den Ideen Amann und Knorr-Cetinas aufbauen und sie mit theoretischen Überlegungen weiterer Autor*innen – meist aus den STS – ergänzen.

4.2.2 Visualisierung als Erkenntnismaschinerie des Sehens

Amann und Knorr-Cetina bezeichnen das Sehen als Erkenntnismaschinerie und subsumieren dazu unter dem Begriff *seeing data*³ ein Set an Praktiken, bei denen das Ziel darin besteht, »to achieve the work of seeing what the data consist of [Hervorhebung im Original]« – die Visualisierung wird zur Arbeitsoberfläche (»workplace«), das Sehen zur Maschinerie, so Amann und Knorr-Cetina (1988:138). Das Sehen ist für die beiden eine wechselseitig ausgeführte Praxis, welche die »sense data« mit der externen Welt verbindet – hauptsächlich durch Konversationen (Amann und Knorr-Cetina 1988:138ff).

Die Überlegungen verschiedener anderer STS-Forscher*innen sind daran anschlussfähig. Sie beschreiben und präzisieren die Erkenntnismaschinerie des Sehens durch verschiedene Elemente, die sich ergänzen und ineinander greifen: Die Metapher der externalisierten Netzhaut verweist auf Praktiken der Selektion und Sichtbarmachung durch Visualisierungen; die Sinngebung erfolgt durch die Abstimmung von Theorie, Empirie und Instrumenten; Mathematisierung, Geometrisierung und das visuelle Vokabular bilden eine visuelle Kultur. Diese Elemente sind eng miteinander verknüpft, sich gegenseitig beeinflussend – sie werden hier in Abschnitte gebündelt, um die theoretischen Bezüge besser verständlich zu machen.

2 Wie dem Forschungsdesign (Kapitel 2) zu entnehmen ist, passiert(e) dieser Prozess iterativ mit der Beobachtung und Analyse des Feldes – auch wenn die theoretischen Überlegungen durch dieses Kapitel separiert und vorangestellt sind.

3 Im Folgenden nutze ich als Übersetzung den Begriff ›Daten erkennen‹. Dies trägt insbesondere dem Prozess, den die Praktiken des *seeing data* beschreiben (siehe nachfolgender Satz), Rechnung: Zu ›sehen‹, woraus Daten bestehen, deutet bereits auf den dafür nötigen Erkenntnisprozess hin.

Abbildung 5: Ausschnitt aus der Untersuchung von Amann und Knorr-Cetina 1988:161: Ein Autoradiogramm als Arbeitsoberfläche.

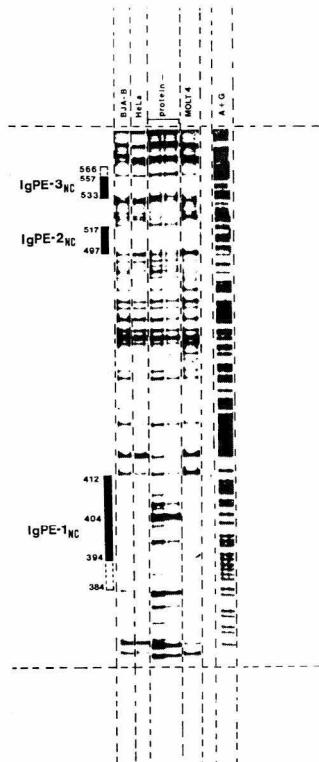


Exhibit 4. Montaged autoradiograph display assembled from various original films. The lines indicate the cuts participants made in piecing the display together. Note also the pointers (bars) and labels at the edges of the image. Compare this autoradiograph with the unedited film reproduced in Exhibit 1

Die externalisierte Netzhaut: Sichtbarmachung durch Selektion

Verschiedene Forscher*innen beschreiben Visualisierungen als Heuristik der Sichtbarmachung: Sie machen Zusammenhänge und Muster sichtbar, die sich sonst dem Wahrnehmungsvermögen entziehen (Heintz und Huber 2001:17); sie visualisieren, was zuvor ein unsichtbares Phänomen war (Burri 2008:349). Dabei wird explizit die körperliche Dimension der Praktik des Erkennens,

das Auge in seiner äusseren wie inneren Dimension, miteinbezogen (Johnson 2007).

Es ist die Netzhaut, welche im menschlichen Auge die Projektionsfläche darstellt, auf welche die Umwelt abgebildet wird:

»Die Netzhaut, Retina (von lateinisch *rete* >Netz<) oder innere Augenhaut ist das mehrschichtige, spezialisierte Nervengewebe, das die Innenseite des Auges der Wirbeltiere, einiger Tintenfische und Schnecken auskleidet. In der Netzhaut wird das einfallende Licht, nachdem es die Hornhaut, die Linse und den Glaskörper durchquert hat, in Nervenimpulse umgewandelt. Dem abdunkelnden retinalen Pigmentepithel liegt von innen die Schicht der lichtempfindlichen Sinneszellen (Fotorezeptoren) an. Deren Impulse werden in weiteren Schichten von Nervenzellen, die nach innen folgen, verarbeitet und weitergeleitet. Daneben enthält die Retina verschiedene unterstützende und versorgende Strukturen.

Vereinfacht ausgedrückt wird die Netzhaut damit zu einer Art Projektionsfläche für die Abbildung unserer Umgebung, ähnlich einer Leinwand oder einem lichtempfindlichen Film, und leitet die durch Lichtreize hervorgerufenen Erregungen weiter an Hirnregionen.«⁴

Michael Lynch nutzte in den Achtzigerjahren die Netzhaut als Metapher, um die Funktion von Visualisierungen zu verorten. Im Aufsatz »Discipline and the Material Form of Images« (im Original 1985 veröffentlicht, hier nutze ich eine Abschrift von 2012) führt Lynch die Metapher folgendermassen aus:

»Graphic formats, instrumental fields and preparatory techniques in histology penetrate both the field of what is visible and the means for perceiving it. It is as though they operate as elements of an externalized retina, activating the perceptible and schematically processing it. Analogous to the specialized retinal cells which are said to constitute our visual world by detecting lines, edges and coded ranges of colour, the externalized retina acts flexibly to manipulate graphic and instrumental edges and codes in order to constitute the sensible, palpable, tangible and appreciable properties of data. [...] An instrument, and the accompanying project of >domesticating< or routinizing space and time in accordance with the instrument's use, prepares the

4 Einführung in den Wikipedia-Eintrag zum Begriff »Netzhaut«: <https://de.wikipedia.org/wiki/Netzhaut> (6. Dezember 2019)

way for perception by pre-coding, geometrizing and normalizing the properties of what comes to be perceived. An active reconstruction of the world is achieved.« (Lynch 2012:402)

Visualisierungen als externalisierte Netzhaut lassen Objekte und Relationen sichtbar und greifbar werden als Resultat von technischen Fähigkeiten und komplexen Instrumenten (Lynch 2012:384). Dieses Zusammenspiel gilt es im Folgenden zu untersuchen.

Sinngebung durch die Abstimmung von Theorie, Empirie, Instrumenten

Daston und Galison (2007:440) haben das »Bild-als-Werkzeug« bezeichnet, als eine neue Form, »eine Form, die das Ideal der Naturneue zugunsten der richtigen Machart aufgegeben hat« – oder wie es Heintz und Huber (2001:30) formuliert haben: Die Visualisierungen verweisen nicht auf eine Realität, sondern auf Messdaten und Regeln der Bildgenerierung. Die Überlegungen verweisen auf das doppelte Repräsentationsproblem: Erstens, wie verkörpern Daten die beobachtete Welt? Und zweitens, was ist eine adäquate Abbildung dieser Daten und damit Sichtbarmachung der dadurch beobachteten Welt?

Die Autorin und der Autor falten in ihrem Beitrag ein Panorama an Visualisierungsformen und -strategien auf, um zu zeigen, welcher Konstruktionsleistung von Wissenschaftler*innen es bedarf, um im Labor etwas sichtbar zu machen. Als Basis dient die Idee, Wissenschaft als einen offenen Prozess zu skizzieren, während dessen folgende Elemente variiert und aktiv aufeinander abgestimmt werden müssen (Heintz und Huber 2001:25):

- theoretische Erwartungen
- beobachtete Ereignisse
- Annahmen über die experimentellen Bedingungen und die Funktionsweise der Aufzeichnungs- und Auswertungsgeräte

Erst wenn es gelungen ist, zwischen diesen drei Elementen eine Übereinstimmung herzustellen, macht ein Phänomen »Sinn« – es wird dann zu einem »Faktum«, das als objektives Merkmal der Außenwelt zugerechnet wird und dessen Herstellungsprozess in Vergessenheit gerät (Heintz und Huber 2001:25). Dieser erfolgreiche Sinngebungsprozess bedingt eine mindestens vorübergehende konstruktive Lösung des doppelten Repräsentationsproblems.

Diese Rekapitulation über die Konstruktionsprozesse von wissenschaftlichen Erkenntnissen verweist darauf, dass es nicht nur darum geht, technische Instrumente zu analysieren. Um soziologisch die Prozesse zu untersuchen, wie Erkenntnisse produziert werden, sind die Visualisierungen als Ort der »rendering practices« (Lynch 2012:384) von Bedeutung. Der Begriff verweist auf die Ethnomethodologie Garfinkels (Garfinkel und Rawls 2002) und dient Lynch dazu, Praktiken zu beschreiben, mithilfe derer Labor-Proben als sichtbare und analysierbare Daten wiedergegeben werden. Diese Praktiken, die konstitutiv sind für die Produktion von Wissen, basieren auf ästhetischen Elementen aus dem grafischen Bereich wie Formen, Farben, Linien und Kontrasten (Hefpler und Mersch 2009:44).

Lynch erarbeitet die Praktiken anhand zweier Beispiele, einer Karte und eines Projekt-Reports. Die zentrale Frage lautet dabei nicht, wie objektive Eigenschaften der Proben den grafischen Elementen entsprechen, sondern »How do graphic properties merge with and come to embody the ›natural object?« (Lynch 2012:388f). Der Autor nennt drei zentrale Praktiken:

- (1) hervorheben – durch indexieren, bezeichnen, die Visibilität erhöhen
- (2) einen grafischen Raum konstituieren – und dafür in geometrischen Formen aufbereiten und ausstellen
- (3) Beobachtungen normieren – und sie dadurch auf einem zweidimensionalen Raum abbilden

Drei Jahre später, im Aufsatz über die externalisierte Netzhaut, beschreibt Lynch die Visualisierungen als »revelatory objects, and, moreover, objects which simultaneously analyze what they reveal« (Lynch 1988:202). Lynch identifiziert nun nicht mehr drei, sondern zwei zentrale Praktiken, die durch Visualisierungen passieren: Selektion und Mathematisierung. Beides sind kollektive methodische Praktiken, »accomplished by researchers working together in groups, which transform previously hidden phenomena into visual displays for consensual ›seeing‹ and ›knowing‹« (Lynch 1988:203).

Praktiken der Selektion und Hervorhebung reduzieren ein Phänomen auf eine ganz bestimmte Weise: »Instruments, graphic inscriptions, and interactional processes take the place of ›mind‹ as the filter, serving to reduce phenomena of study into manageable data.« (Lynch 1988:204) Der folgende Abschnitt zeigt auf, dass sich dabei Vorgänge der Mathematisierung und Geometrisierung abspielen.

Visuelle Kultur: Mathematisierung, Geometrisierung, visuelles Vokabular

Mathematisierung als Methodik ist ein weiteres Element der Erkenntnismaschinerie des kollaborativen Sehens. Mathematisierung bedeutet, dass visualisierte Labor-Proben durch Diagramme zu verrechenbaren Darstellungen werden: »Mathematization is embodied in the graph. The graph has become an emblem of science which even popular advertisements exploit. Graphing a phenomenon identifies the thing or relationship with the analytic resources of mathematics« (Lynch 1988:218). Diagramme pressen die Phänomene in die Formalsprache der Geometrie – als Beispiele nennt Lynch Punkte, Linien und das zweidimensionale Gitternetz (kartesisches Koordinatensystem) als einige der meistgenutzten geometrischen Formen. Die beschriebene Formalisierung setzt die im vorherigen Kapitel erläuterten Praktiken der Sichtbarmachung und Selektion voraus: Erst dadurch werden die Daten definiert und rechenbar (Lynch 1988:229f), die Möglichkeit zur Verkörperung in kalkulierbaren Zeichen wird zum Existenzkriterium (Krämer 2001). Die geometrische Formgebung ist in Datenvisualisierungen das bestimmende Ordnungskriterium der Zeichen (Burri 2008:346ff).

Das zweidimensionale Gitternetz bringt dabei zwei Eigenschaften mit sich, welche die Relationierungsprozesse organisieren: Erstens ermöglicht seine Zweidimensionalität das Übereinanderlegen und Verbinden von Objekten mit verschiedener Herkunft – »Most of what we call ›structure‹, ›pattern‹, ›theory‹, and ›abstraction‹ are consequences of these superimpositions (Bertin 1973)« (Latour 2012:23ff). Der Autor verweist auf Bertins Werk zur »Graphischen Semiologie« (1974, im französischen Original 1973 erschienen) aus gutem Grund: Bertin hat sich eingehend mit dem semiotischen System in Visualisierungen als Forschungsinstrument befasst (Bertin 1974:11). Er unterscheidet zwischen den Komponenten der Information, des grafischen Systems (Grenzen, Formate, Farben und Muster) sowie den Regeln für dieses grafische System, die das ›Sprechen‹ und ›Verstehen‹ der visuellen Sprache bestimmen. Heßler und Mersch (2009:38) haben die *superimposition* – oder deutsch »Überblendung« – konziser gefasst: »Dabei werden Tabellen, Diagramme, Graphen etc. zu multiplen ›ikonischen Modellen‹ zusammengeführt, was nicht nur zu einer Verdichtung von Informationen unterschiedlicher Quellen führt, sondern gleichzeitig auch zu deren Relationierung.«

Für diese Relationierung spielt die *superimposition* mit den geometrischen Formen (als kulturelles Ordnungskriterium der Datenvisualisierungen) zusammen – für deren Nutzung das zweidimensionale Gitternetz die Voraussetzung ist (Latour 2012:23ff). Neben der Formalisierung und Abbildung auf

ein mathematisiertes Ordnungssystem zeichnen sich Datenvisualisierungen als visuelle Sprache durch einen spezifischen Wortschatz, das visuelle Vokabular, aus. Im Falle von Datenvisualisierungen sind dies geometrische Formen als kulturelle Sinnstrukturen, die auf der beschriebenen Mathematisierung und den Selektions- und Hervorhebungspraktiken aufbauen (Burri 2008:346ff). Allerdings ist diese Sprachform nicht in der Lage, etwas über die Qualität und Modalität der Verbindung auszusagen: Während die gesprochene und die geschriebene Sprache die Relationen beispielsweise mittels Verben, deren Modi und Personalpronomen einordnet, ist das visuelle Vokabular entpersonalisiert und dessen Beziehungen objektiviert (Vaara 2019:105ff).

Ein visuelles Vokabular ist die Voraussetzung, um Daten sichtbar zu machen und die entstehenden Zahlen und visuellen Elemente mit Bedeutung zu versehen und zu deuten: »Scientists start seeing something once they stop looking at nature and look exclusively and obsessively at prints and flat inscriptions« (Latour 2012:19). Latour erläutert die Bedeutung des visuellen Vokabulars anhand der Disziplin der Chemie und des Mendelejewschen Periodensystems, welches die Grundlage für die Artikulierung chemischer Prozesse bildet: »Chemistry becomes powerful only when a visual vocabulary is invented that replaces the manipulations by calculation of formulas. Chemical structure can be drawn, composed, broken apart on paper, like music or arithmetic, all the way to Mendeleiev's table« (Latour 2012:17). Der Grund, warum eine solche visuelle Sprache derart wirkungsvoll ist, liegt darin, dass sie die optische Konsistenz sichert (Latour 2012:18). Das bedeutet, dass sie zu einer Vereinheitlichung der Perspektive führt. Die Konsequenzen daraus sind nicht zu unterschätzen. So erläutert Latour, wie das Zeichnen aus der Zentralperspektive eine ganze visuelle Kultur verändert hat – der Begriff »worldview« ist dabei wörtlich zu verstehen: »how a culture *sees the world*, and makes it visible. A new visual culture redefines both what it is to see, and what there is to see« [Hervorhebungen im Original] (Latour 2012:11f). Kosminsky et al. (2019) haben diese genealogische Linie weitergezogen und beleuchtet, wie Datenvisualisierungen als visuelle Kultur genauso an die enge Verflechtung von Wahrheit, Perspektive und Objektivität anschliessen wie früher das Zeichnen aus der Zentralperspektive und die Fotografie.

Zur Sprache in einer Erkenntnismaschinerie gehört aber nicht nur ein Vokabular, sondern auch eine Sprachlehre, die Grammatik: Jedes Medium verfügt über Buchstaben und eine Grammatik, welche die validen Kombinationsmöglichkeiten bestimmen – was wiederum den Variationsraum und die Sinngebung einschränkt (Heintz 2010:171f). Geht es um die Konstruktion von

Datenvisualisierungen, mangelt es denn auch nicht an Referenzen auf die Analogie der Sprache und der Grammatik (beispielsweise Bauer und Ernst 2010; Wickham 2010). Wilkinson (2005:634) schreibt: »Grammar gives language rules. Graphics are generated by a language. The syntax of graphics lies in their specification. The semantics of graphics lies in their data.« Dieses Zusammenspiel, die Regeln der Grammatik und die Semantik des Vokabulars sind sozial ausgehandelte Konventionen. Wie diese Aushandlung im Visualisierungsmodus des Erkennens von Daten vor sich geht, wird Gegenstand des empirischen Kapitels 5 sein.

4.2.3 Visualisierung als Befestigung von Evidenz

Im zweiten Teil des Gerüsts geht es darum, wie auf Visualisierungen Evidenz befestigt wird – also derjenige Zustand erreicht ist, in dem etwas als ›objektives Faktum‹ anerkannt wird, weil die Betrachterin den Konstruktionsprozess nicht mehr hinterfragt: »The result of the construction of a fact is that it appears to be unconstructed by anyone« (Latour und Woolgar 1979:240). Amann und Knorr-Cetina haben diesen zweiten Visualisierungsmodus als Montage-Prozess des *becoming evident*⁵ beschrieben: »As indicated before, data become evidence, i.e., the data included in scientific texts, only after they have undergone an elaborate process of transformation« (Amann und Knorr-Cetina 1988:159f). Die folgenden Abschnitte zeigen, dass die Montage durch Selektionspraktiken mittels verschiedener analytischer Ordnungen erfolgt: Sie nutzen ästhetische Wertungen und geben die Rahmung vor, um die interpretative Flexibilität von Visualisierungen zu reduzieren.

Das Zeigen von Visualisierungen geniesst ein bemerkenswertes rhetorisches Gewicht, wie Latour betont: Wir könnten uns kaum vorstellen, etwas zu wissen (»to know«), ohne uns dafür nicht auf Bibliographien, Indizes, Referenzen, Tabellen, Fotografien und so weiter zu berufen (Latour 2012:16). Die folgenden Abschnitte verweisen auf verschiedene Überlegungen aus der Medientheorie, welche die starke Überzeugungskraft von Datenvisualisierungen

5 Ich übersetze den Begriff des *becoming evident* mit ›Evidenz befestigen‹. Wörtlich übersetzt wäre von ›evident werden‹ die Rede. Dieser Ausdruck impliziert jedoch eine Passivität, die den Beobachtungen nicht gerecht werden kann. Im Zusammenhang mit dem von Amann und Knorr-Cetina (1988:135) im Original genutzten Begriff der ›Montage‹, auf welcher die Akteur*innen »visually flexible phenomena« fixieren, habe ich mich für das Verb ›befestigen‹ entschieden.

begründen. Dabei verweisen sie auf die Aura der Wissenschaftlichkeit und Technizität sowie auf Eigenheiten des Mediums und seine Form des Zeigens. Die soziologische Perspektive relativiert dies wiederum, indem sie auf die verschiedenen Publika hinweist, welchen andere Anforderungen zugeschrieben werden – insbesondere entstehende Communitys wie im Datenjournalismus mit ihren Expertisen verändern möglicherweise die Evidenzbefestigung, wie der letzte Abschnitt vermutet.

Wie für die Elemente der Erkenntnismaschinerie gilt auch für diejenigen aus diesem Teil: Sie sind nicht trennscharf und hängen eng zusammen. Sie werden hier in Abschnitte gebündelt, um die theoretischen Bezüge besser verständlich zu machen.

Selektions- und Transformationspraktiken

Waren es im Modus des Erkennens von Daten Praktiken der Selektion zur Sichtbarmachung, so finden sich im zweiten Modus genauso Praktiken der Selektion, die nun aus dem sichtbar Gemachten Evidenz befestigen. Diese Transformationsprozesse lassen sich drei analytischen Ordnungen zuordnen (Amann und Knorr-Cetina 1988:159ff):

- (1) Markierungen und ästhetische Wertungen: Es werden »signals« in den Vordergrund, »noise« in den Hintergrund gerückt (Amann und Knorr-Cetina 1988:160). Es werden also spezifische Elemente ausgewählt, markiert und hervorgehoben – und zwar dadurch, dass sie »relatively ›clean‹, ›pure‹ or ›beautiful‹ signals [sind] according to aesthetic criteria which specify, in an area of research, what counts as a ›good‹ figure. Note that such judgments are not extrinsic to scientists' judgment on what the figure reveals« (Amann und Knorr-Cetina 1988:160ff).
- (2) Sicherstellung von Analysierbarkeit und Vergleichbarkeit: Es wird eine Matrix übergestülpt, welche die Analysierbarkeit und Vergleichbarkeit mit anderen Visualisierungen sichert.
- (3) Hinweise auf Lesart und Rahmung: Die Präsentation von Daten als Evidenz basiert auf Hinweisen (»pointers«). Sie indizieren eine bestimmte Lesart der Visualisierung, indem sie gewisse Elemente als wichtig herausheben und andere ignorieren. Neben Anhaltspunkten innerhalb des Bildes gehören auch der Titel und geschriebene Lesehilfen dazu.

Amann und Knorr-Cetina (1988:162) schliessen, dass die Praktiken des zweiten Modus den Sinngebungsprozess des ersten Modus auf ein Ergebnis zuspitzen:

»Now to avoid misunderstanding, let us stress that we are not suggesting here that the evidence thus created is purely fictional – however fabricated it may be. But neither does it correspond to the ›data‹ or signals obtained in the laboratory. Rather, this montage is a members' way of visually reproducing the sense of ›what was seen‹ which is an *upshot* of participants' shop talk negotiations; an accomplishment of – not a precondition for – their work. Talk attached to (in the sense of Section 4) visual materials was crucial insofar as it provided participants with candidate formulations of the reality they ›saw. [Hervorhebungen im Original]« (Amann und Knorr-Cetina 1988:162)

Visualisierungen werden also zu Evidenz durch einen Montage-Prozess, der diese Evidenz mittels verschiedener Selektionspraktiken erarbeitet. Obwohl auch im ersten Modus Selektionspraktiken passieren, unterscheiden sie sich fundamental in ihrer Funktion: In der Erkenntnismaschinerie des ersten Modus richten sie sich an die Forschenden selbst, um sie in der Wiedergabe von unsichtbaren Phänomenen und ihrer Transformation in eine visuelle Sprache zu unterstützen. Im Montage-Prozess richten sie sich an ein äusseres Publikum, welches von der Evidenz überzeugt werden soll (im Falle von Amann und Knorr-Cetina an die Wissenschaftsgemeinde). Aufgrund des unterschiedlichen Motivs und des unterschiedlichen Publikums folgt die Aushandlung in den Modi jeweils anderen Dynamiken.

Die drei analytischen Ordnungen lassen sich verschiedenen Ansprüchen zuordnen, auf welche die weiteren Abschnitte eingehen.

Reduktion der interpretativen Flexibilität durch ästhetische Wertungen und Rahmung

Während es sich im naturwissenschaftlichen Labor von Amann und Knorr-Cetina (1988:135) noch um »visually flexible phenomena« handelte, sind computergestützte Datenvisualisierungen »provisional and interactive. The researchers constantly tweak them, altering parameters changing color scales, substituting different algorithms or statistical analyses« (Burri und Dumit 2008:303). Das Autor*innen-Duo beschreibt, dass aus diesen Datenvisualisierungen Evidenz mit einem Anspruch auf Objektivität wird, indem die Selektions- und Hervorhebungspraktiken die Unsicherheit der Beobachtungen und damit die interpretative Flexibilität reduzieren (Burri und Dumit 2008:303). Eine Datenvisualisierung erhält dadurch die Überzeugungskraft, die Phänomene ›wie sie sind‹ – und damit definitionsgemäss mit dem Anspruch auf Objektivität – abzubilden.

Um Evidenz mit einem Anspruch auf Objektivität zu befestigen, ist es nötig, eine »meaningful visualization of the *data as knowledge* [Hervorhebung im Original]« zu konstruieren (Burri und Dumit 2008:303). Die Konstruktion muss sich deshalb am Ziel eines erfolgreichen Kommunikationsprozesses der Erkenntnisse orientieren. Dieser ist zweiseitig: einerseits die bereits beleuchtete Formalisierung mittels geometrischer Formen und einer visuellen Grammatik in der Erkenntnismaschinerie (*encoding*⁶). Andererseits verlangt er auch Lesekompetenz vom Publikum (*decoding*) – diese bleibt aber im Konstruktionsprozess unbekannt und muss geschätzt werden. Für Datenvisualisierungen im Journalismus gilt das aufgrund seines breiten Publikums sogar noch ausgeprägter als für die Wissenschaft, wo die Wissenschaftsgemeinde enger eingegrenzt ist.

Um ein möglichst kohärentes Verstehen von Datenvisualisierungen sicherzustellen, stützt man sich auf ästhetische und wissenschaftliche Konventionen ab (Tufte 1983; Burri und Dumit 2008:303). Dabei geht es darum, die Konstruktionsleistung und arbiträren Entscheide, welche die Produktion einer Datenvisualisierung mit sich bringt, möglichst vollständig zu verbergen.⁷ Wiederum basieren die Praktiken dabei auf ästhetischen Ordnungen, welche so auszulegen sind, dass sie der Idee gehorchen, was ein ›gutes Bild‹ ausmacht – und diese Idee meist eher nach »Klarheit statt nach Wahrheit« verlangt (Heßler und Mersch 2009:45), auch in den Naturwissenschaften. Es sind dies genau die analytischen Ordnungen (1) und (3) aus dem vorherigen Abschnitt, welche diese Ideen umsetzen. Der folgende Abschnitt ordnet diese aus verschiedenen Perspektiven ein.

-
- 6 Die Begriffe *encoding* und *decoding* stammen aus der Kommunikationstheorie. Folgendes Zitat von Few (2016:6) erklärt sie: »Data visualization is only successful to the degree that it encodes information in a manner that our eyes can discern and our brains can understand. Getting this right is much more a science than an art, which we can only achieve by studying human perception. The goal is to translate abstract information into visual representations that can be easily, efficiently accurately, and meaningfully decoded.«
- 7 Espeland und Stevens (2008:425) haben den interessanten Hinweis gemacht, dass diese Verbergungspraktik gleichwohl durch ein Medium (Datenvisualisierungen) erfolgt, das oftmals vorgibt, höchsten Ansprüchen der Transparenz zu genügen.

Überzeugungskraft und Autorität

Die Praktiken der Reduktion der interpretativen Flexibilität dienen dazu, die Überzeugungskraft einer Datenvisualisierung zu steigern. Verschiedene Forschende aus der Kultur- und Wissenstheorie verweisen aber auch auf die generell starke Überzeugungskraft visueller Medien. Merz und Hinterwaldner ergänzen im Handbuch Wissenschaftssoziologie den Aspekt des wissenschaftlich-technischen Kontextes:

»Jenseits ihrer Entstehungskontexte beziehen wissenschaftliche, technische und medizinische Bilder ihre Macht aus der Annahme, dass sie auf eindeutige und direkte Weise das Dargestellte repräsentieren. Die Anlehnung an etablierte Darstellungskonventionen verstärkt den Eindruck eines im Bild objektiv Gegebenen. Zugleich bezieht das Bild seine Überzeugungskraft aus der Autorität des wissenschaftlich-technischen Kontextes, dem es entstammt.« (Merz und Hinterwaldner 2012:309)

Burri (2008:349) ergänzt, dass die Eindringlichkeit visueller Medien oft anthropologisch-universalistisch mit dem Menschen als »visuelles Wesen« und einem privilegierten Sehsinn begründet wird. Sie spricht dabei von einer Verführung und Inszenierung, welche auch auf einer emotionalen Ebene spielt:

»Andererseits ist es die Verführung durch das Bild selbst, auf welche die visuelle Wirkung zurückzuführen ist. Die Ästhetik von Visualisierungen und ihre Inszenierung in einem bestimmten diskursiven und lokalen Kontext sprechen Emotionen an und beeinflussen die Wahrnehmung intuitiv. [...] Die Wirksamkeit der *visual persuasiveness* ist dabei von individuellen Dispositionen und Erfahrungen, aber auch von sozialen Sinnstrukturen, lokalen Denk- und Wahrnehmungskonventionen und kulturellen Sehtraditionen abhängig.« (Burri 2008:350)

Die Kultur- und Medienwissenschaftler*innen Heßler und Mersch verweisen weiter auf spezifische Eigenschaften des Mediums und seine »Form des Zeigens«, aus welcher sich ein starker »Evidenzeffekt« ergibt:

»Doch folgen die Evidenzeffekte aus der ›affirmativen Kraft‹ des Bildes, ihrer bereits erwähnten eigentümlichen Intensität und Suggestibilität, die unmittelbar mit seiner medialen Form des *Zeigens* verquickt ist, die von sich her mit der Schwierigkeit der Unverneinbarkeit, der Nichthypothetizität und Nichtkonjunktivität behaftet ist und den Scheincharakter unterstreicht, [...]« (Heßler und Mersch 2009:29)

Weber greift diese Überlegungen auf und münzt sie auf die Evidenzkraft von Datenvisualisierungen aus dem Journalismus um:

»Im Gegensatz zur Sprache können Visualisierungen nicht ›nichts zeigen‹, sie können (erstmal) nicht negieren, sie können kein ›Vielleicht‹ darstellen, keine Unsicherheit, keine Wahrscheinlichkeit, kein Futur und keinen Konjunktiv [...]. Zum anderen umgibt Graphen und andere diagrammatische Formen eine Aura von Technizität, Exaktheit und Wissenschaftlichkeit [...]. Trotz ihrer wissenschaftlichen Aura und Nüchternheit darf aber nicht übersehen werden, dass Infografiken und Datenvisualisierungen Artefakte sind und daher immer ästhetischen Interpretationen, stilistischen Vorlieben und redaktionellen Machbarkeiten unterliegen.« (Weber 2018:7)

Die Annahme besteht also darin, dass sich die theoretischen Überlegungen aus Kultur- und Medienwissenschaften wie auch der Wissenschaftstheorie auf Datenvisualisierungen im Journalismus übertragen lassen – aber gleichzeitig die spezifischen Dynamiken des Feldes, der Redaktion und ihrer Expertisen die Praktiken der Evidenzbefestigung anpassen.

Incomplete models

Zudem relativiert eine soziologische Perspektive diese Ansichten insofern, als sie darauf hinweist, dass Visualisierungen je nach Publikum andersartige Wirkungen und Konsequenzen verursachen (Burri und Dumit 2008:304ff). Gerade mit Blick auf die zwei sehr verschiedenen Publika des Datenjournalismus – die Öffentlichkeit als Medienkonsument*innen sowie die eigene Community – ist eine Differenzierung angezeigt. So besteht die Vermutung, dass Latours (2012) Charakterisierung von Visualisierungen als »immutable mobiles« ergänzt werden muss. Dieser Begriff verweist auf die uneingeschränkte und unhinterfragte Autorität und optische Konsistenz von Visualisierungen.

Burri und Dumit (2008:307f) bringen jedoch mit den »incomplete models« eine alternative theoretische Perspektive auf Visualisierungen ein. Als Modelle sind Datenvisualisierungen verschieden interpretierbar und teilweise unmittelbar veränderbar. Sie sind offen für Anpassungen und Basteleien durch Akteur*innen mit der nötigen programmiertechnischen Expertise und/oder einer interoperablen Software. Diese Offenheit verweist auf eine junge Bewegung: »Another response to the openness of visualizations to manipulation is a movement among scientists to ›open-source‹ their data, making the raw data available online for other experimental groups to download and analyze themselves« (Burri und Dumit 2008:308). Dies beeinflusste insbeson-

dere die Arbeitsorganisation und -kultur zwischen verschiedenen beteiligten Gruppierungen, fügen die beiden hinzu.

Insgesamt erscheint es wichtig, diese Perspektive mit dem Blick auf die datenjournalistische Community in die Heuristik zu integrieren. Die Montage-Praktiken richten sich nach der Vorstellung, wie Erkenntnisse mittels einer Datenvisualisierung möglichst effektiv und überzeugend transportiert werden können und als Evidenz erscheinen. Diese Vorstellung ergibt sich dabei aus den Aushandlungsdynamiken zwischen den Expertisen der Produzent*innen. Die journalistische Expertise legt den Schwerpunkt auf die Evidenzbefestigung gegenüber einer breiten Öffentlichkeit als Publikum. In den Dynamiken mit den weiteren Expertisen und der insgesamt starken Orientierung an der eigenen datenjournalistischen Community könnte aber auch eine Grundlage sein, dass Datenvisualisierungen eher als *incomplete models* evident werden. Die Befestigung erfolgt dann über Praktiken der Offenheit der Datenquellen und Methoden.

4.3 Visualisierungen in der *trading zone*: Kombination der Heuristiken

Abschnitt 4.1 hat die Heuristik der *trading zone* eingeführt, um Sinngebung als Aushandlung zwischen verschiedenen Subkulturen beziehungsweise Expertisen zu fassen. Dabei hat sich gezeigt, dass Galisons ursprüngliche Konzeption einen starken Fokus auf die Verständigung im sprachlichen Diskurs legt und Materialität ausklammert. Dabei hat gerade diese Dimension das Potenzial, zu zeigen, wie (materielle) Werkzeuge die Verständigung stabilisieren.

Abschnitt 4.2 hat eine zweite Heuristik hinzugezogen, welche dies leisten kann: Datenvisualisierungen als Werkzeug mit zwei verschiedenen Modi zu sehen: Daten erkennen und Evidenz befestigen. Die Ausführungen haben sich dabei vor allem auf die Materialität und ihre Praktiken konzentriert. Dass die beiden Gerüste zusammenpassen, zeigt sich insbesondere dadurch, dass praktisch sämtliche STS-Forschende, welche mit ihren Überlegungen zur Heuristik beigetragen haben, auf die soziale Praktik der Aushandlung hinweisen. Die folgenden Absätze erläutern die Bedeutung von Gesprächen über Datenvisualisierungen für den Aushandlungsprozess.

So sind Konversationen für Amann und Knorr-Cetina die zentrale Praktik in der Erkenntnismaschinerie des Sehens. Ihr kollaboratives Sehen ist eine wechselseitig ausgeführte Praxis, welche die ›sense data‹ durch Gespräche mit

der externen Welt verbindet (Amann und Knorr-Cetina 1988:138ff). Weitere Autor*innen stützen sich auf ähnliche Überlegungen:

- Die Konversationen der Labor-Mitarbeitenden widerspiegeln, was als adäquate Abbildung eines Objekts gilt (Lynch 2012:384);
- Die Urteile über die Validität und Ästhetik einer Visualisierung hängen von Konventionen beziehungsweise konsensuellen Aushandlungen ab: Muster und ihre Erkennung haben mit der Wirklichkeit zu tun, sind aber keine absoluten Wahrheiten – »Die Bilder sind arbiträr, aber nicht willkürlich« (Heintz und Huber 2001:31);
- Die sozialen Aushandlungen über Visualisierungen sind als Sinngebungsprozess zu deuten: Darstellung und Wahrnehmung, das heißt herstellende und interpretierende Praktiken, sind eng miteinander verknüpft und konstituieren sich gegenseitig (Burri 2008:349).

Amann und Knorr-Cetina (1988:138ff) spezifizieren dann, welche Charakteristika solche Konversationen über Visualisierungen ausmachen: Sie sind an das Objekt (in ihrem Fall: die Visualisierung der Laborproben) gebunden und durch dieses strukturiert; sie sind eingebettet in eine Serie von Konversationen, die durch das Objekt verbunden sind; und die Sprechenden sind zueinander orientiert, das Resultat (die Erkenntnis) wird im gemeinsamen Gespräch konstruiert. Die in ihrem Praxisbeispiel beobachteten Konversationen, als Fachsimpeln (»shop talk«) bezeichnet, sind eher chaotisch, weil sie keiner logischen Struktur folgen (Amann und Knorr-Cetina 1988:145). Amann und Knorr-Cetina haben zudem auf die hohen Hürden verwiesen, die zu überwinden sind, damit derartige sprachliche Konversationen überhaupt zustande kommen – denn das zu Referierende nicht nur zu sehen, sondern auch sprachlich zu fassen, ist eine sehr anspruchsvolle Übersetzungsarbeit (Amann und Knorr-Cetina 1988:135).

Ein erster Aspekt in der Kombination von Datenvisualisierungen als Materialität und Aushandlungen ist also, dass Erstere strukturieren und Konversationen organisieren. Ein zweiter Aspekt besteht darin, dass Datenvisualisierungen aus Aushandlungssituationen bestimmte Akteur*innen versammeln – und somit auch die Frage nach deren Expertisen aufwerfen (Burri und Dumit 2008:301f): Wer produziert, wer weiß was – und wem ist es erlaubt, was zu wissen? Durch die Expertisen der Beteiligten fliessen epistemologische Grundlagen und professionskulturelle Standards von »guten Bildern« und Evidenz in die herstellenden und interpretierenden Praktiken

der Erkenntnismaschinerie ein (Burri 2001). Entwickeln sich diese Expertisen weiter, allenfalls auch durch neue Kombinationen und Aushandlungsdynamiken, bilden sich neue Wissensbestände und Subdisziplinen, die wiederum neue Standards und Aushandlungssituationen mit sich bringen (Burri und Dumit 2008:301f).

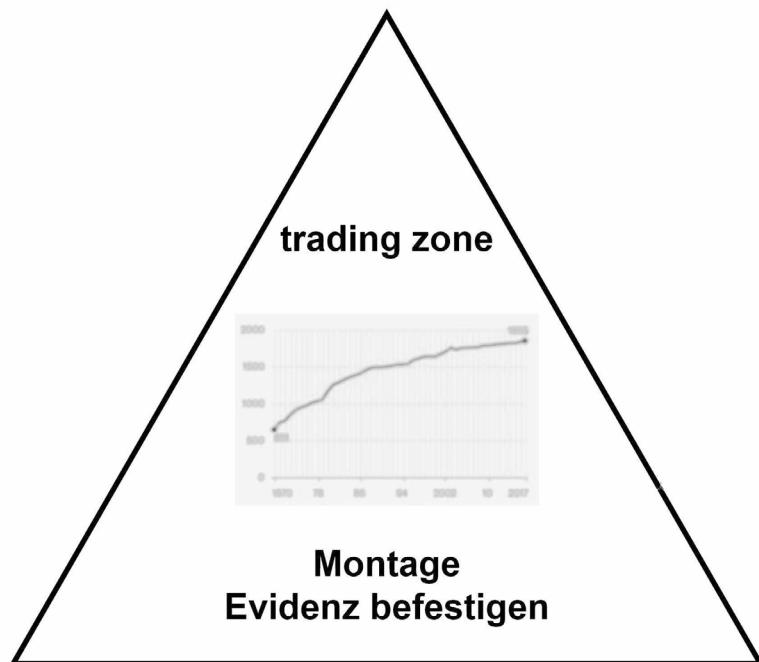
So bildet die Kombination der beiden im Kapitel 4 vorgestellten Heuristiken, die Aushandlungen in den Modi der Datenvisualisierungen, den von den Expertisen-Triaden ›eingeschlossenen‹ Raum im bereits vorgestellten Schema (siehe Abbildung 2). Die Untersuchung der aufgeworfenen Fragen (siehe Abschnitt 3.4) mittels des ethnografischen Materials aus den Teams erfolgt anhand der folgenden Heuristiken:

Abbildung 6: Datenvisualisierungen als Arbeitsoberfläche strukturieren den Aushandlungsraum des Erkennens von Daten.



(eigene Darstellung)

Abbildung 7: Datenvisualisierungen als Montage strukturieren den Aushandlungsraum der Befestigung von Evidenz.



(eigene Darstellung)

Die beiden Abbildungen vervollständigen das Schema der Expertisen mit einem Aushandlungsraum, in dem je nach Modus auf einer Arbeitsoberfläche Daten erkannt werden (Abbildung 6) beziehungsweise auf einer Montage Evidenz (Abbildung 7) befestigt wird.

Diese theoretisch fundierten Gerüste dienen in den beiden folgenden Kapiteln dazu, die erforschten datenjournalistischen Praktiken einzuordnen und daraus zu verstehen, ob und wie Datenjournalismus sich in den Aushandlungen als eigener Wissensbestand stabilisiert und Aussagen ermöglicht.

5 Daten erkennen: Visualisierungen als Arbeitsoberfläche

Der Abschnitt 4.2 hat zwei verschiedene Modi von Visualisierungen dargelegt: das Erkennen von Daten (*seeing data*) und das Befestigen von Evidenz (*becoming evident*). Abschnitt 4.3 hat die beiden Modi mit der Heuristik der Aushandlung verbunden. Kapitel 5 untersucht mit diesem Gerüst den erstgenannten Modus in den zwei datenjournalistischen Teams: das Erkennen von Daten auf Visualisierungen¹ – die als Arbeitsoberfläche dienen, um aus den Daten Sinn zu generieren und Erkenntnisse zu produzieren.

Der Abschnitt 5.1 erörtert die vier aus dem empirischen Material resultierenden Arbeitspraktiken: Datafizierung, saubere Daten, Geschichten drehen und Auswertungen plausibilisieren. Durch diese machen sich die beiden Teams ihre drei Expertisen mit ihren materiellen und kulturellen Hintergründen nutzbar. Die folgenden Abschnitte beleuchten die soziale Verwobenheit dieser Expertisen – in den Konversationen (5.2), dem visuellen Vokabular als semiotische Ressource (5.3), den soziotechnischen Werkzeugen (5.4) sowie den Verweisen auf andere Akteur*innen des Feldes.

5.1 Arbeitspraktiken des Erkennens von Daten

Im folgenden, ausführlichen ersten Teil des Kapitels zum ersten Visualisierungsmodus geht es darum, die Arbeitspraktiken zu beschreiben, mit welchen

1 Ich nutze konsequent den Begriff Datenvisualisierungen, auch wenn im referenzierten empirischen Material dafür Begriffe wie »Grafik«, »Infografik« oder »Chart« verwendet werden. Zwischen diesen Begriffen gibt es theoretische Differenzierungen, die aber im beobachteten Arbeitsalltag der beiden Teams nicht relevant sind.

die beiden Teams das Erkennen von Daten organisieren. Es erfolgt eine Gliederung in vier Praktiken beziehungsweise Sets an Praktiken, die einen analytischen Mehrwert bietet, aber in der Empirie als solche nicht klar abgegrenzt wird. Im ersten Abschnitt folgt ein Schlaglicht auf Datafizierung und einige grundlegende ontologische Überlegungen. Danach folgen Sets an Praktiken zur Datengrundlage, Geschichtenproduktion und -plausibilisierung. Die Gliederung insinuiert auch eine gewisse prozessorientierte Chronologie der Praktiken, die zwar empirisch tatsächlich sinnhaft ist, aber keineswegs immer eingehalten wird. Stattdessen sind die Praktiken und Schritte eng verwoben und der Prozess ist iterativer und abduktiver Natur.²

5.1.1 Datafizierung: Geschichten in Daten übersetzen

Waren es in den Laborstudien natürliche Stoffe, die mathematisiert und geometrisiert wurden, so stellt sich im Datenjournalismus die Frage: Wie werden Daten konstruiert, die auf ›reale Begebenheiten‹ referenzieren? Es handelt sich dabei um einen Prozess der Datafizierung, das heißt, die soziale Welt wird in quantifizierter Form wiedergegeben. Die Sammlung von Daten ist keine neue Praktik, wie die Historie der Statistik aufzeigt. Die Welt in Zahlen auszudrücken ermöglicht es, über Grenzen (beispielsweise von Ländern oder Feldern) hinaus Verbindungen herzustellen (Desrosières 1998), zu vergleichen (Lamont 2012; Heintz 2010) und sich auf Objektivität zu berufen (Porter 1995). Die Datafizierung einer Geschichte ermöglicht eine auf der Logik der Datenbank (Manovich 2001) rationalisierte Sichtweise.

Bei BR Data suchen die Teammitglieder, inspiriert von einer Idee, einfach mal Daten und Indikatoren, um sich ein Thema, eine *Geschichte* mit Daten vorstellen zu können. Während der Recherche nach Daten geht es darum, zu explorieren, welche Variablen man kombinieren könnte; welche Sachverhalte überhaupt quantifiziert existieren, das heißt in Zahlenform vorhanden und mit Skalen erfassbar sind; welche Daten Grundlage kalkulatorischer Praktiken werden können. Für dieses Set an Praktiken ist das Internet die wichtigste Quelle: Fast alle Recherchen passieren online. Es gibt Themenbereiche, in denen BR Data eine stärkere Quantifizierung erwartet. Während Treffen, an denen das Team Geschichten-Ideen diskutiert, erfolgen immer wieder ad hoc

² Wiederum war es das Team von The Pudding, welches Transparenz auch über diesen iterativen Prozess anstrebt und ihn in einem eigenen Artikel aufzuzeichnen versucht: <https://pudding.cool/process/pivot-continue-down/> (27.12.2020)

derartige Abschätzungen: Welche Daten sind zu jenem Thema verfügbar und wie aufwendig wäre die Sammlung und Aufbereitung?

Das ›Finden von Geschichten in Daten‹ ist eigentlich ein ›Geschichten in Daten übersetzen‹. Anderson (2018:138ff) und Coddington (2018) haben in verschiedenen Organisationen geforscht, welche Informationseinheiten (Einzelereignisse beziehungsweise einzelne Nachrichtenmeldungen) sie in Datenbanken versammeln und mittels Geschichten verbinden. Beide interpretieren dies als »computational thinking« (Wing 2006) im Journalismus. Dies ist dadurch charakterisiert, dass die Datenbanken nicht aufgrund von Geschichten und dazugehörigen Fakten (Dokumente, Zitate) aufgebaut werden, sondern dass die journalistische Erkenntnisproduktion so geordnet ist, dass man sich die Realität im Datenbankformat vorstellt – »A fact is true, in other words, if it exists in the database« (Anderson 2018:14) – und die Geschichte auf dieser ontologischen Grundlage imaginiert.

Somit erscheint es logisch, dass das BR-Data-Team keine Geschichte produziert, falls keine digitalen Daten vorhanden sind.³ Basiert die Erkenntnisproduktion nicht auf Daten, so ordnet das Team dies ausserhalb der eigenen Zuständigkeit und Kompetenz ein. BR Data beruft sich damit implizit auf ein »data *a priori*« (Hervorhebung im Original), das Gray (2018:7) beschreibt als »not only designates but also provides the conditions of possibility for seeing and engaging with different aspects of collective life – making possible particular styles of reasoning and particular forms of knowledge and experience.« Diese »Data Worlds« von Gray sind charakteristisch für das Narrativ der neuartigen, datengetriebenen Geschichten (Abschnitt 3.3.1) – die *Datenwelten* zu beherrschen wird in den Teams gerne zelebriert und NZZ-Datenjournalist Yannick kategorisiert Personen mit entsprechender Expertise als »data-savvy«⁴.

Der Imaginationsprozess, eine Geschichte in datafizierte Form zu übersetzen, passiert bei NZZ Storytelling häufig über die Antizipierung geometrischer Verkörperungen der Daten. So sind händische Skizzen, die nicht aus Daten errechnet werden, ein wichtiger Aspekt der Datafizierungspraktik (siehe Walny 2016). Die Teammitglieder beginnen die Datenauswertung sehr oft mit Skizzen auf Papier, mithilfe derer sie eine Visualisierung als Geschichte modellieren (Abschnitt 5.4.3 beleuchtet die Rolle von Papierskizzen als Werkzeug detailliert). Visualisierungen sind insbesondere dann ein Instrument, um die

3 58_20190110_BR-Data-12 (1)

4 22_20180323_NZZStorytelling-10 (11)

Geschichte in Daten zu übersetzen, wenn es darum geht, andere Kolleg*innen mit weniger datenjournalistischer Erfahrung für ein Projekt zu gewinnen. So stossen NZZ Storytelling und das Auslandressort in einer Sitzung viele gemeinsame Geschichten-Ideen an – nicht indem sie überlegen, welche Daten es zu einem Thema gibt, sondern wie die Datenvizualisierung schliesslich aussehen könnte.⁵ Durch die Visualisierung – ob als Vorstellung oder Skizze – werden Daten, die sonst als Zahlenwerte oder in grossen Textmengen stark abstrahiert sind, auf eine Ebene re-konkretisiert, auf der ein Anschluss an die Sinngebung für die Journalist*innen mit spezifischen Themenspezialisierungen möglich ist. Die folgenden Abschnitte werden zeigen, welche datenjournalistischen Praktiken dies in den Teams gewährleisten sollen. Die Produktion von journalistischer Erkenntnis aus Daten basiert damit nicht nur auf journalistischer und programmiertechnischer Expertise, sondern insbesondere die grafische nimmt durch visuelle Verkörperung eine zentrale Rolle ein.

Dieser Abschnitt hat die ontologischen und epistemologischen Grundlagen der datenjournalistischen Arbeitspraktiken in den beiden Teams aufgezeigt: Sie basieren auf einer durch das Datenbankformat rationalisierten Sicht auf reale Ereignisse, also der Datafizierung der Informationen über die Welt.⁶ Gleichzeitig gilt es, die epistemologische Form der Geschichte in Daten zu übersetzen. Dies passiert insbesondere über Imaginationen von visuellen Verkörperungen der Daten. Datafizierung selbst basiert in den beiden Teams demnach oft auf einer visuellen Rationalität, welche »die Strukturlogik sozialer Praxis um den Visualitätsaspekt erweitert« (Burri 2008:354).

Die folgenden Abschnitte vertiefen die erwähnten verschiedenen Elemente der datenjournalistischen Arbeitspraktiken. Der nächste fokussiert auf die Datensätze.

5.1.2 Saubere Daten: Granularität, Ordnung, Interoperabilität

Die zuweilen geäusserte Binsenweisheit, saubere Daten seien die Grundlage für eine saubere Arbeit, reibt sich mit der Feststellung, dass in fast jedem Datensatz Verschmutzungen aller Art vorkommen. Die folgenden Abschnitte beleuchten, anhand welcher Leitlinien die beiden Teams die datenjournalistische

5 21_20180322_NZZStorytelling-09 (4)

6 Die Einordnung des Datafizierungsphänomens bleibt an dieser Stelle knapp. Für weiterführende Überlegungen dazu sei auf die Sammelwerke von Prietl und Houben 2018 und Mämecke et al. 2018 sowie Mayer-Schönberger und Cukier 2013 verwiesen.

Praktik der Herstellung einer analysierbaren Datengrundlage, »sauberer Daten« eben, strukturieren: Granularität, Ordnung und Kategorisierung, Interoperabilität.

Granularität

Ein wichtiges Kriterium in der Herstellung analysierbarer Datensätze ist die Sicherstellung von Granularität. So zum Beispiel in einem Projekt über Strafverfahren zu spezifischen Delikten, für das BR Data einen Datensatz selbst erhebt. Die zwei Teammitglieder Lea und Lukas konzipieren eine Excel-Tabelle, die danach die Grundlage für die Datensammlung und -analyse bilden soll. Insbesondere Lukas betont immer wieder, dass die Tabelle möglichst viele Differenzierungen zulassen muss.⁷ Dies ermöglicht mehr Optionen für Variablen und potenziert entsprechend die Anzahl möglicher Kalkulationen. Und es verringert die Chance, dass in den Daten Verzerrungen und Aggregate unentdeckt bleiben, so eine weitere Argumentationslinie von BR Data.

Das NZZ-Storytelling-Team zieht in einer Situation den Umkehrschluss: Ist ein Phänomen nicht erklärbare, sind die Daten noch zu wenig granular. So lassen sich die Verschiebungen in der politischen Landschaft bei den kurz vor der Forschungsphase stattgefundenen Wahlen auf der Ebene der Wahlkreise nicht anhand bestimmter soziodemografischer Variablen statistisch erklären. Deshalb versucht Datenjournalistin Annalena zusammen mit einem Redaktor aus dem Zürich-Ressort die Wahlstatistiken eines Urnengangs auf einer noch granulareren Ebene als den stadtzürcherischen Wahlkreisen auszuwerten.⁸

Datenjournalistischen Praktiken nutzen grosse Datenmengen, »Big Data«, als Methode, um soziale Phänomene in höherer temporaler und geografischer Granularität zu fassen (Lazer und Radford 2017). Hadley Wickham, Statistiker und Chefentwickler der Programmiersprache R, skizzierte in einem Artikel die Prinzipien hinter den Visualisierungen, die mit dem von ihm entwickelten R-Paket ggplot2 möglich sind (Wickham 2010). Ein besonders wichtiges Prinzip ist dabei die Granularität der Daten. Denn Wickhams Paket basiert wie die meisten Visualisierungswerzeuge darauf, dass verschiedene Schichtungen (»layers«) zusammenkommen, die beispielsweise je eine statistische Transformation oder die Ausgestaltung mittels einer geometrischen Figur umfassen. Und dass Skalen und Aspekte (»facets«) differenzierbar sind, um die Einordnung grafisch darstellen zu können. Die Granularität eines

7 58_20190110_BR-Data-12 (6-8)

8 20_20180321_NZZStorytelling-08 (2)

Datensatzes ist die Voraussetzung der »geschichteten« (»layered«) Visualisierungen, die verschiedene Aspekte abbilden. Dabei wird jede Komponente auf eine unterschiedliche Art ästhetisch umgesetzt: teilweise in geometrischen Formen in einer Skala, oder kategoriale Variablen als Farb-, Größen- oder Kontur-Attribute (siehe ausführlicher in Abschnitt 5.3). Das Befolgen dieser Prinzipien ermöglicht es, einerseits schnell Visualisierungen zu entwickeln und anzupassen; es macht sie replizierbar; es ermöglicht aber vor allem auch die Exploration: Einzelne Komponenten können einfach ausgetauscht und so verschiedene Varianten ausprobiert werden (Healy und Moody 2014:111ff; Tukey 1977). Die Granularität von Daten ermöglicht eine grosse Varianz an explorativen Visualisierungen und stellt damit in den Teams die Vielfältigkeit und Flexibilität der Interpretation sicher.

Daten ordnen und kategorisieren

Für eine Mehrzahl an Projekten nutzen sowohl BR Data wie auch NZZ Storytelling Daten, die bereits in digitaler Form existieren. Sehr oft verwenden die beiden Teams offizielle Statistiken, die in tabellarischer Form vorhanden sind. Durch die vorgegebene Formatierung sind die Daten mit gewissen Kategorisierungen vorstrukturiert. Insbesondere die Teammitglieder von BR Data äussern explizit den Anspruch, diese Kategorisierungen zu erfragen und zu verstehen sowie die Verzerrungen zu antizipieren. BR Data erfragt die Herkunft und die Erhebungsumstände der Datensätze; die Teammitglieder vergleichen dabei Daten mit menschlichen Quellen, die Informationen zur Verfügung stellen und die kritisch hinterfragt werden müssen (Parasie 2014; Ettema und Glasser 1985). Staatlichen Daten begegnen sie oft mit der Attitüde, dass sie lückenhaft und unsystematisch erhoben sind – dies verweist einerseits darauf, dass die Teammitglieder die Vollständigkeit des Datensatzes als sehr wichtiges Kriterium erachten, weil lückenhafte Datensätze die Interpretation stark erschweren. Andererseits wird damit auf den teameigenen Anspruch verwiesen, mächtigen Akteur*innen wie dem Staat kritisch gegenüberzustehen und eine Wachhund-Funktion einzunehmen (siehe Watchdog-Narrativ, Abschnitt 3.3.1). Ein mögliches Spannungsfeld zwischen der enthusiastischen Nutzung von Datensätzen als Quellen sowie dem kritischen Hinterfragen ebendieser, wie es Lesage und Hackett (2014:48) erwähnen, war in den Teams nicht feststellbar.

Neben der Evaluation der Datenquelle verläuft die Einordnung der Qualität eines Datensatzes vor allem über das Kriterium der systematischen Ordnung – bei BR Data wird die positive Bewertung mit dem Adjektiv »sau-

ber« ausgedrückt.⁹ Der Ordnungsaspekt wird in zwei Eigenschaften bewertet: erstens seine Nachvollziehbarkeit und zweitens eine durchgehend konsistente Umsetzung der Ordnung. Nachvollziehbar ist eine Ordnung dann, wenn die abgebildeten Relationen zwischen Beobachtungen und Eigenschaften bzw. kategorisierenden Variablen mit den Vorstellungen möglicher Relationen der Teammitglieder stimmig sind. Der Einordnungsprozess ist ein essenzieller Teil der Sinngebung von Daten – eine Position, die verschiedene Autor*innen aus den Science and Technology Studies untermauert haben: Beispielsweise haben Bowker und Star (1999) Kategorisierungspraktiken als Setzung von ontologischen Grundlagen beschrieben.

Die konsequente Umsetzung des Ordnungsprinzips wurde in der neueren Zeit wiederum von Hadley Wickham und seinem Prinzip der »tidy data« vorangetrieben (Wickham 2014) – Abschnitt 3.2.2 hat es bereits als Teil der programmiertechnischen Expertise im Feld angetippt. Wickham schlägt eine Standardisierung der Verbindung zwischen der Struktur eines Datensatzes und seiner Sinngebung vor – und zwar im folgenden Format: »In tidy data: 1. Each variable forms a column. 2. Each observation forms a row. 3. Each type of observational unit forms a table.« (Wickham 2014:4) Die von Wickham skizzierte Standard-Strukturierung ist heute ein wichtiges Prinzip in der Anwendung der Programmiersprache R, welche sehr oft in datenjournalistischen Praktiken zur Anwendung kommt. Auch in den beiden Teams basieren einige Praktiken auf R, wie die weiteren empirischen Ausführungen zeigen.

Bauen die Datenjournalist*innen einen Datensatz selbst auf oder strukturieren sie ihn wieder oder neu, erfolgt eine zeitaufwendige Phase der Konzeption, während derer die Aspekte Zeitaufwand, Verfügbarkeit, vor allem aber die Einordnung mittels Variablen abgestimmt werden. Dieser Abstimmungsprozess erfolgt vor dem Horizont, dass die angewandte Einordnungsweise der Beobachtungen mittels Variablen oder Metadaten zu spezifischen Sinngebungen und Formen der Visualisierung der Daten einlädt und andere ausschliesst (Flyverbom und Madsen 2015:135).

9 beispielsweise 57_20190109_BR-Data-11 (4)

Interoperabilität herstellen

NZZ Storytelling bewertet Datensätze besonders gut, wenn die Daten ohne grossen Bearbeitungsaufwand in ihre Visualisierungstools einfügbar sind. So lobt ein Teammitglied das lokale statistische Amt, welches Daten als Excel-Datei liefert und über eine *Open-Data*-Strategie verfügt, was das Einspeisen der Daten in die Tools vereinfache.¹⁰ In einem anderen Projekt, für das Wahlresultate aus dem Internet gescraped werden, die weniger »sauber« sind, unternehmen die Programmierer des Teams einen zusätzlichen Aufwand zur Standardisierung mittels eines Skripts. Dadurch waren die Daten ohne weitere Bearbeitung mit dem Datenvisualisierungstool Q kompatibel und eine Live-Update-Funktion führte dazu, dass die Datenvisualisierungen sich jeweils selbst aktualisierten, sobald neue Daten online gestellt wurden. Ein NZZ-Storytelling-Teammitglied präsentierte den Prozess stolz auf Twitter.

NZZ-Datenjournalist Yannick erzählt, dass er für das Team Prozesse entwickeln soll, um Terroranschläge in ihrem Ausmass schnell grafisch einordnen zu können. Auf die Frage hin, wie der Prozess funktioniert, antwortet er, dass zwei Webseiten die Datengrundlage für die Opferzahlen bieten.

Er habe dann ein Python-Script gebaut, das diese beiden Seiten mittels eines Protokolls schnell absuchen kann. Wir schauen uns das Skript direkt im Python-Interface Jupyter an und lassen es einmal laufen. Gerade heute Mittag hat es in Frankreich einen Terroranschlag gegeben und wir nutzen die dazu bekannten Zahlen und Daten. Diese werden gleich beim Script eingespeist und dann als Vergleichswerte genutzt. Dann, erklärt [Yannick], würden standardmässig sechs bis acht Charts ausgeworfen.¹¹

Die Praktik der Herstellung von Interoperabilität zwischen Datensatz und Tools sichert erst die Einlesbarkeit der Daten in die Tools und ermöglicht Analysen. Das störunglose Einlesen gehört deshalb zur Praktik, über *saubere Daten* zu verfügen. Abschnitt 5.4 geht auf die Organisationsformen ein, welche die Teams anwenden, um die Interoperabilität herzustellen und Dateninfrastrukturen zu ihrer Sicherstellung zu etablieren.

Insgesamt verfolgen die Teammitglieder mit dem Grundsatz der *sauberen Daten* ähnliche Ziele wie beispielsweise Datenanalysten mit »schönen Daten«, nämlich die Sicherung möglichst vieler künftiger Analysemöglichkeiten:

¹⁰ 22_20180323_NZZStorytelling-10 (16)

¹¹ 22_20180323_NZZStorytelling-10 (9)

»Die Schönheit von Daten liegt in ihrer Potentialität. Schöne Daten bzw. das Potential von schönen Daten erlauben es, die spezifischen Ansprüche, Hoffnungen und Unsicherheiten verschiedener AkteurInnen momentan zu absorbieren. Schöne Daten sind verheißungsvoll. Im Falle des untersuchten Unternehmens versprechen schöne Daten eine erfolgreiche Suche nach noch unbekannten Fakten und Erkenntnissen.« (Mützel et al. 2018)

Diese Potenzialität versuchen die Teammitglieder mit einem Set von Praktiken auszuschöpfen, indem die Datengrundlage mithilfe computergestützter Verfahren visualisiert und die Variablen dabei relationiert werden – darum dreht sich der folgende Abschnitt.

5.1.3 Geschichten drehen

Werden Geschichten bei BR Data oder NZZ Storytelling produziert, so nutzen die Teammitglieder häufig die Metapher des »Drehens« von Geschichten.¹² In diesem Set an Praktiken »verdrehen« bzw. »verwickeln« sich besonders ausgeprägt soziale und technologische Elemente: »..., capacities for action are seen to be enacted in practice and the focus is on constitutive entanglements (e.g., configurations, networks, associations, mangles, assemblages etc.) of humans and technologies.« (Orlikowski 2010:135) Die folgenden Abschnitte beleuchteten Praktiken des »Drehens« von Geschichten, wie sie empirisch in den Redaktionen beobachtbar waren, um aus Daten Erkenntnisse und Sinn zu generieren. Die Abschnitte gliedern sich in verschiedene Visualisierungs- und damit Relativerierungstypen sowie die daran angehefteten Diskussionen.

Zuerst aber eingeschoben ein Abschnitt zum aufgespannten methodologischen Raum, in dem diese Praktiken stattfinden: Sie sind verwickelte, iterative und abduktive Vorgehensweisen zwischen explorativer Datenanalyse und Thesen aus Themenwissen – im Argot *data-driven* beziehungsweise *story-driven* genannt. Abschnitt 3.2.3 hat diese epistemologischen Aspekte bereits im Feld thematisiert, auf einer diskursiven Ebene. Die folgenden Ausführungen zeigen auf, was dies in den beiden Teams praktisch bedeutet.

¹² 48_20181211_BR-Data-02 (3); 48_20181211_BR-Data-02 (3); 51_20181214_BR-Data-05 (5); 14_20180313_NZZStorytelling-02 (27).

Die Iteration zwischen *data-driven* und *story-driven*

Als einen der ersten Schritte in einem Projekt explorieren NZZ Storytelling und BR Data die Daten. Sie schliessen dabei an die explorative Datenanalyse (EDA) an, die John Tukey (1977) in den Siebzigerjahren begründete. Tukey erkannte, dass die statistischen Analysen und der Erkenntnisgewinn daraus in engem Zusammenhang mit Möglichkeiten der grafischen Abbildung und Kommunikation durch Visualisierungen standen, obwohl sie sich in Teilen in ihrer Zielsetzung unterscheiden (Gelman und Unwin 2013). Später schloss Edward Tufte an Tukeys Überlegungen an und schrieb: »At their best, graphics are instruments for reasoning about quantitative information« (Tufte 1983:9).

Die EDA provozierte im Feld der Statistik über lange Zeit viel Skepsis (und tut dies immer noch), weil die Exploration ohne oder nur mit schwachen Hypothesen manchen opportunistisch und unstrukturiert erscheint (Healy und Moody 2014:113). Trotzdem hat die EDA einen Aufschwung erlebt durch die Verfügbarkeit von computerbasierten grafischen und statistischen Verfahren, insbesondere *Tools* für interaktive Visualisierung – was auf der Hand liegt, denn »[e]xploratory graphics are all about speed and flexibility and alternative views« (Gelman und Unwin 2013:4). Die Geschwindigkeit und die Flexibilität haben sich durch die technologische Entwicklung der letzten Jahrzehnte zweifellos vergrössert.

Im Zeitalter von Big Data erfolgt eine Annäherung zwischen der explorativen Datenanalyse und konfirmatorischen Ansätzen: Wer mit grossen Datenmengen arbeitet, wird immer – auch wenn sie oder er über eine Hypothese verfügt – explorativ mittels verschiedener Visualisierungsmethoden einen Datensatz erkunden, um ihn zu säubern und seine Validität zu überprüfen (Healy und Moody 2014:113). Verschiedene neue Werkzeuge ermöglichen Interaktivität und damit das Ausprobieren verschiedener Beziehungen, die mit schnell wechselnden Interpretationen einhergehen (Healy und Moody 2014:116). Über ein solches *Tool* verfügt NZZ Storytelling mit Sue. Sue ermöglicht es, Daten explorativ zu visualisieren mittels sehr vieler verschiedener geometrischer Formen der Relationierung. Zum Beispiel zu Beginn des Projekts der Fussballtransfers:

[Kaspar] schiebt und probelt dann auf Excel, wie er die Ligen einteilen und kategorisieren kann. [Kaspar] sagt [zu mir], während er mit den Grafiken rumspielt und seinen Notizblock studiert, er überlege, was man aus der Visualisierung Spannendes herausziehen könnte – in welchen Ländern die Schweizer Liga Ge-

winn mache? Wie liesse sich das herausheben? [Kaspar] probiert einen Area-Chart aus, aber mit nur zwei Punkten: Auf der einen Seite die Einnahmen aus einer Liga, auf der anderen Seite die Ausgaben. Er meint, das sehe spannend aus. Er dreht sich zu Teamleiter David und fragt, ob er kurz vorbeischauen möge. Dieser kommt und fragt gleich, ob [Kaspar] das nun mit Sue gemacht habe. [Kaspar] bejaht. David schaut den Chart an und meint, dieser sei schwierig zu interpretieren. [Kaspar] fragt den Teamleiter, was einzelne Elemente in der Grafik genau bedeuten (zum Beispiel warum die Fläche der Schweizer Liga konstant ist), David erklärt (hierbei handelt es sich um die Transfers liga-intern). [Kaspar] meint, dass er nochmals daran weiterarbeite.¹³

Designer Kaspar tritt ohne starke Hypothesen und ohne viel thematisches Wissen zum Thema Fussballtransfers an die Analyse heran. Er probiert aus und evaluiert die entstehende Datenvisualisierung mit einem Mix aus ästhetischen und journalistischen Kriterien. Einige Stunden später hat Kaspar das Repertoire an Variablen in der Tabelle erweitert, um zusätzliche Visualisierungs- und Geschichtenoptionen zu eröffnen. Er sitzt am Pult und diskutiert mit Yannick, gleich neben ihm sitzend. Ich begebe mich zu ihnen:

Ich blicke auf die Tabelle, die [Kaspar] inzwischen um einige Spalten erweitert hat. Er hat Berechnungen und Sortierungen vorgenommen, die zeigen, ob es Ligen gibt, in die nur verkauft oder aus denen nur gekauft wird. So liessen sich alle Ligen in vier Muster einteilen. Auf dem Notizpapier hat er die vier Muster in Small Multiples aufgezeichnet. [Kaspar] sagt, Muster seien das Spannende. [Yannick] stimmt zu und sagt, dass sei spannend. [Kaspar] hat auch die Differenz berechnet, die aus Ein- und Verkäufen aus verschiedenen Ligen ausmachen, sowieso die Prozentzahl der Steigerungen. [Yannick] zeigt sich kritisch, das könne man nicht so darstellen. Die beiden diskutieren dann während rund zwanzig Minuten, wie und was man visualisieren könne. Sie haben einfach die Daten vorliegen und versuchen, daraus eine Visualisierung zu erreichen. [Yannick] versucht immer wieder, sich zu fragen, was man denn damit aussagen könnte, was die »Geschichte« sei. Er verfügt über etwas Wissen des Fussballgeschäfts und bringt Beispiele. [Kaspar] bringt kaum Fussballwissen-Inputs. Er zeichnet immer wieder Vorschläge auf seinen Notizblock. Auf meine Frage hin,

13 14_20180313_NZZStorytelling-02 (10)

was denn die Sportredaktion für einen Text dazu schreibe, ob denn da noch thematische Inputs oder Vorgaben kämen, verneinen beide. Sie denken nicht. Außerdem lief der Kontakt mit der Sportredaktion über David. [Yannick] sagt, dass sei »manchmal das Problem«, dass man hier selber vor den Daten sitze und eine Geschichte daraus entwickeln müsse. [Kaspar] vermutet, dass NZZ Storytelling die Visualisierung liefert und dann der Sportredaktor noch einen Text/Erklärung dazu liefert. [Yannick] findet, dass das Betrachten einzelner Fussballer-Karrieren spannend wäre. [Kaspar] ist mehr an den Mustern interessiert. Ihre Vorschläge legen immer wieder einen anderen Fokus: Geldflüsse zwischen den Ligen, den Status aus Ausbildungsliga, Management-Strategien, Scouting-Länder, ...¹⁴

In vielen anderen, im Falle von NZZ Storytelling gar den meisten Situationen, ist die datenjournalistische Praktik angeleitet von einer starken Hypothese beziehungsweise Geschichte – wie in einer Situation, als ein Wirtschaftsredaktor nach Vorankündigung beim Teamleiter vorbeikommt und erklärt:

Es gehe um die Firmenbeteiligungen von Pierin Vincenz [ehemaliger CEO der Schweizer Raiffeisenbank]. Der Redaktor legt eine handgezeichnete Skizze einer Visualisierung auf einem gehäuselten Papier auf den Tisch und erklärt seine Vorstellung: Eine Zeitachse von 2005 bis 2015, er erklärt auch die Achsenbeschriftungen. Er erläutert zudem kurz die Sachlage, was Vincenz für Firmenbeteiligungen hatte. Zudem kämen noch »Sub-Firmen« dazu, die visualisiert werden müssten. Es seien noch nicht alle Daten vorhanden, einige Kaufpreise fehlten noch. Der Teamleiter meint, dass einzelne Werte noch nicht entscheidend sind für einen Entwurf, aber der »range« muss klar sein. Der Redaktor schlägt eine logarithmische Skala vor, der vermutliche teuerste Wert ist ihm schon bekannt.¹⁵

Fragt man die Mitglieder der beiden Teams nach dem Kernstück des datenjournalistischen Prozesses, so verweisen sie sehr oft auf die Methodik der Produktion von Geschichten aus Daten, das Zusammenspiel von These und thematischem Wissen. Insgesamt oszilliert die Herangehensweise an die Datenanalyse zwischen den beiden Endpunkten (1) *data-driven* und (2) *story-driven*:

14 14_20180313_NZZStorytelling-02 (20)

15 16_20180315_NZZStorytelling-04 (10)

- (1) *data-driven*: Die Geschichte wird erst *aus den Daten* und deren Analyse heraus entwickelt (NZZ Storytelling) bzw. es gibt noch keine Vorstellung, *was in den Daten steckt* (BR Data). Obwohl in den beiden Redaktionen etwas andere Begriffe verwendet werden, referenzieren sie dieselbe Herangehensweise.
- (2) *story-driven*: Es gibt eine Geschichte oder eine klare Idee dafür, die mit Daten als zusätzlicher Perspektive *ergänzt* wird (NZZ Storytelling), bzw. es gibt eine Fragestellung oder Hypothese, welche an einem Datensatz *überprüft* wird (BR Data). Auch hier unterscheiden sich die gewählten Begriffe, aber es wird dieselbe Herangehensweise referenziert.

Im redaktionellen Alltag sind diese beiden Herangehensweisen selten in reiner Form zu finden. Viel eher ist es plausibel, sich die beiden Typen als Endpunkte eines Spannungsfeldes vorzustellen. Die Erkenntnisproduktion bewegt sich spiralförmig in diesem Spannungsfeld und ist folglich abduktiv (Kitchin 2014a; Halford und Savage 2017): In einem iterativen Vorgehen wechseln sich das Erheben und Analysieren von quantitativen Daten und das Sammeln von Kontextwissen mit qualitativen Methoden ab, sich stetig zu einer Geschichte hin entwickelnd – Abschnitt 5.1.4 geht darauf ein, wie diese unterschiedlichen Perspektiven passend gemacht werden. Zuerst folgt ein Blick in die Erkenntnismaschinerie: Praktiken des *Geschichten-Drehens* aus den Redaktionsräumen von BR Data und NZZ Storytelling. Diese Relationierungspraktiken sind implizit strukturiert durch das Iterieren zwischen *data-driven* und *story-driven* und operieren im dadurch aufgespannten Raum.

Suchen von Mustern und Auffälligkeiten

Man schaut sich die Datensätze an, sucht nach Auffälligkeiten, Peaks, Verteilungen; mit den statistischen Methoden kann man dann Geschichten herausholen – so beschreibt BR-Data-Teammitglied Johannes, wie man aus Daten Geschichten »macht«.¹⁶ Die Suche nach Auffälligkeiten beschrieb bereits John Dewey (1927) als Charakteristikum von *news*, dessen Überlegungen Wihbey (2019:151) wie folgt auf den Punkt bringt: »The search for discontinuity is what makes news news.« Für diese Suche nach Diskontinuitäten in Datensätzen, die einem *auffallen*, nutzen die beiden Teams Datenvisualisierungen. Das Überstülpen eines visuellen Vokabulars produziert visuell erkennbare Muster und Auffälligkeiten:

16 BR_Fragebogen_Mitarbeitende_10

[Brigitte] hat nun auf ihrem Computer html-Dokumente mit Säulengrafiken der Anzahl Berichte nach Quartal, dabei ist in der Säule der Anteil der Berichte mit Herkunftsennnung ausgewiesen. [Brigitte] und [Lea] vergleichen verschiedene Städte und Regionen (jeweils eigenes html-File nach Polizeistation – teilweise sind dies Städte, Landkreise, ganze Bundesländer, die Einteilung ist sehr unterschiedlich). Die Zahl der Berichte aus den einzelnen Stationen ist sehr unterschiedlich, auch der Anteil der Herkunftsennnung. [Brigitte] und [Lea] finden den Vergleich »spannend«, vor allem zwischen Städten ähnlicher Grösse, die aber einen sehr unterschiedlichen Anteil haben. Dieser schwankt teilweise auch über die Zeit.¹⁷

Der Ausschnitt aus der Feldnotiz beleuchtet ein Projekt von BR Data, von dem die Daten – kurze Polizeimitteilungen (»Berichte«) in Textform, verteilt auf fünf Jahre, insgesamt rund eine Million – gesammelt sind. Seitens des Teams besteht die Vermutung, dass sich die Herkunftsennnungen der Personen in diesen Berichten verändert haben – über die Zeit, möglich wären aber auch Differenzen zwischen verschiedenen Polizeistationen. Mittels Textanalyse hat die Programmiererin Brigitte deshalb extrahiert, wie oft die Berichte die Herkunft nennen. Der Anteilsquotient wird Variablen wie dem Jahr oder der Absenderin (Polizeistation einer Stadt, Landkreis, Bundesland) zugeordnet und dann diese Relation über die Zeit oder die Verteilung nach Absenderin verglichen. Die Produktion der Geschichte folgt dem Kriterium der Differenz der Anteile und eine hohe Differenz bewerten Brigitte und Lea mit »spannend«.

Der Vergleich erfolgt mittels einer Datenvisualisierung in Form eines Säulendiagramms. Grosse Differenzen widerspiegeln sich in sehr unterschiedlich hohen Säulen und sind für Brigitte und Lea schnell visuell erkennbar. In dieser Phase des Projekts sind noch viele *Geschichten-Drehs* offen, das heisst, der Möglichkeitsraum der Erkenntnisproduktion wurde durch die granulare Datenerhebung und diverse Metadaten der Berichte – die nun zu Variablen für die Auswertung werden – offen gehalten. Aus der hohen Anzahl an Vergleichsmöglichkeiten selektieren die beiden Teammitglieder visuell anhand der erkennbaren Höhendifferenzen der Säulengrafiken, welche Muster und Auffälligkeiten relevant sind für die Erkenntnisproduktion.

Ausschlagende Kurven

Die Kurve ist bei BR Data keine eng definierte mathematische Gleichung, sondern wird ganz generell als Beschreibung des Graphen eines Liniendiagramms genutzt. Der Verlauf der Kurve ist während der Phase der Erkenntnisproduktion ein wichtiger visueller Indikator für den *Geschichten-Dreh*:

Ein Programmierer hat dann mittels eines R-Scripts diverse Auswertungen vorgenommen, er stand dafür in Austausch mit [Martin]. Diese Auswertungen hat er in ein [html-]Dokument eingefügt, in dem der Datensatz mittels R ausgewertet wird, sichtbar in diversen grafischen Darstellungen. Der Code dazu ist ein-/ausklappbar. Es wurden die Anzahl der Meldungen, deren Verteilung auf die Bundesländer und Polizeistellen, deren zeitliche Entwicklung, die Aufschlüsselung auf einzelne Delikte sowie mittels bestimmter Wörter auch auf Nationalitäts- und ethnische Zuschreibungen vorgenommen. Diese Explorations wurden mittels Linien- und Balken-Diagrammen (produziert mit R) abgebildet. [Lea] erwähnt, dass es darum ging, Auffälligkeiten, »Peaks« in Kurven und ungewöhnliche Entwicklungen über die Zeit zu finden.¹⁸

Das Finden von Abweichungen bedingt eine Vorstellung der Normalität. Beides wird in diesem Projekt aus den Datenvisualisierungen heraus konstruiert. Denn was ein normaler Anteil an Herkunftsennungen sei, davon hat BR Data keine klare Vorstellung: Man geht davon aus, dass es keine kollektiven Leitlinien für Polizeistationen gibt, wann die Herkunft zu nennen ist – und deshalb ist auch nicht klar, welche Verteilung ›normal‹ wäre.¹⁹ Wenn die Kurve ausschlägt, die Richtung ändert, dann ist das ›nicht normal‹. Dann ist es ›spannend‹. Dann begibt man sich auf die Suche nach der Variablen, welche die Veränderung verursacht hat.

Adrian untersucht für ein gemeinsames Projekt mit einer Redaktorin aus einem anderen Ressort Verschreibungszahlen von verschiedenen Medikamenten derselben Sorte. Zusammen mit Martin versucht er, Auffälligkeiten in den Kurven zu finden.

[Martin] und [Adrian] wechseln zu Tabelle und Linien-Diagramm, die den allgemeinen Verbrauch (Verschreibungszahl) anzeigen. Die Kurven gehen in der

18 48_20181211_BR-Data-02 (5)

19 48_20181211_BR-Data-02 (5)

Tendenz zurück. [Martin] meint, man könne [der Redaktorin aus einem anderen Ressort] Daten anbieten, die zeigten, dass der Verbrauch eher zurückgehe. Auf einem weiteren Liniendiagramm zeigt sich, dass die Verschreibungszahl eines Medikaments entgegen dem sinkenden Trend stark ansteigt. [Martin] fragt [Adrian], was das für ein Medikament sei. [Adrian] erklärt kurz, [Martin] findet diese Entwicklung spannend.²⁰

Der mehrheitlich sinkende Trend als Normalfall dient als Abgrenzungsfolie für das eine Medikament mit steigendem Trend, welcher das Interesse weckt: »spannend« – die Spannung baut sich auf zwischen der konstruierten Normalität und der mindestens für den Moment unerklärbaren Abweichung (wobei sich diese Spannung wenig später auflöst, weil man sich den gegenläufigen Trend des einen Medikaments mit dem abgelaufenen Patentschutz erklärt).

Der um die Kurve tanzende *Geschichten-Dreh* findet seine Fortsetzung im nächsten Excel-Tabellenblatt. Dort gibt es zu einem verwandten Aspekt des Themas (Resistenzen) Daten in Tabellenform, Balken- und Liniendiagramme. Adrian und Martin betrachten das Tabellenblatt:

[Adrian] zeigt auf die Diagramme und meint, dass die Resistenzen etwas zunehmen, aber nicht allzu stark. [Martin] meint, dass es aus Datensicht »nichts gebe, dem man nachgehen sollte«. [Martin] und [Adrian] interpretieren zusammen die Kurve und wie sie sich verändert – [Martin] schlägt auch vor, sich die Kurve vorzustellen, würde man die ausschlagenden Werte einiger Jahre weglassen. [Martin] meint, man könne dies [der Redaktorin aus dem anderen Ressort] mitteilen, dass man in den Daten nichts sehe – was aber nicht heisse, dass das Problem nicht existiere.²¹

Adrian und Martin erscheint die Entwicklung der Kurve aufgrund des Hintergrundwissens plausibel. Auch das gedankliche Experiment, sich einige (offenbar plausibel begründet) ausschlagende Werte »wegzudenken«, führt nicht zu Auffälligkeiten, denen man nachgehen sollte. Der etwas erfahrene Martin grenzt die eigene Perspektive als »Datensicht« ab, in der man »nichts sieht« – oder von der anderen Seite betrachtet: *seeing data* heisst für die beiden Datenjournalisten, mittels Auffälligkeiten sich potenzielle Geschichten imaginieren zu können.

²⁰ 51_20181214_BR-Data-05 (2)

²¹ 51_20181214_BR-Data-05 (2)

Bemerkenswert an der Situation ist der reflexive Moment, der zum Schluss auftaucht (aber schliesslich nicht weiter ausgeführt wird). Martin grenzt die »Datensicht« nicht nur ab, sondern relativiert deren Aussagekraft: Ein von der Redaktorin mit ihren journalistischen Werkzeugen postulierter Sachverhalt widerspiegelt sich nicht zwingend »in den Daten«. Damit anerkennt er nicht nur die Grenzen der eigenen Perspektive, sondern beugt vor allem einer konfliktbeladenen Situation im darauf geplanten Gespräch mit der Redaktorin vor: Es braucht kein Streit darüber ausgetragen werden, welches Werkzeug nun die validere Erkenntnis hervorbringt.

Ausreisser

Neben sich »ungewöhnlich« verhaltenden Kurven wecken auch Ausreisser das Interesse. Als Ausreisser wird gemeinhin eine Beobachtung bezeichnet, die nicht in einen erwarteten Bereich fällt. Diese Definition teilen Datenjournalist*innen mit Statistiker*innen, welche sich seit Bernoulli 1777 mit dem Phänomen beschäftigen (Beckman und Cook 1983). Treffen die Praktiker*innen aus den jeweiligen Bereichen auf einen Ausreisser, so unterscheiden sich jedoch ihre Schlussfolgerungen: Während die Statistik nach Wegen sucht, trotz Ausreissern ihre Resultate sinnvoll interpretieren zu können (Zimek und Filzmoser 2018), so bezeichnet das Datenjournalismus-Handbuch Ausreisser entweder als Fehler – oder als potenzielle »good stories« (Gray et al. 2012:186). Ausreisser und starke Abweichungen sind wichtige Attribute, welche die Aufmerksamkeit für eine Datenvisualisierung erhöhen (Segel und Heer 2010:1140).²² Murray Dick bestätigte in seiner Forschung über die Infografik-Produktion in englischen Medienorganisationen die vorhandene, wenn auch nicht unhinterfragte Attraktivität von Ausreissern (Dick 2014:500).

Lea schlägt für das Polizeibericht-Projekt vor, Ausschau nach solchen zu halten:

[Lea] schlägt vor, die Ausreisser in den Polizeistationen zu finden mit [Brigitte] zusammen. [Martin] meint auf jeden Fall, das sei ja auch spannend, sich dann

²² Welche Attraktivität die Figur des Ausreissers auf die Datenvisualisierungs-Community ausübt, zeigt sich daran, dass die Data Visualization Society – Fussnote 4 (Kapitel 2) hat bereits auf sie hingewiesen – ihr Konferenzformat »Outlier« nennt. <https://www.outlierconf.com/> (21.10.2020) <https://medium.com/nightingale/introducing-outlier-a-conference-hosted-by-the-data-visualization-society-801ca8ff87ff> (21.10.2020)

genau diese Stationen anzuschauen. Dann könne man auch schauen, wie die Mitteilungen sich aus diesen Stationen verhalten.²³

Im Gegensatz zu den für statistische Analysen entwickelten Tests (Beckman und Cook 1983), welche Ausreisser mathematisch definieren, dient der Ausreisser in der datenjournalistischen Praktik als vage Zuschreibung für einen gegenläufigen Trend oder Extremwert, hier in einem Projekt von NZZ Storytelling:

[Susanne] erzählt [Annalena], dass sie gerade die Reportagen für das Gemeinde-Projekt im Wallis vorbereitet und eine Reportage aus Naters eine Option sei, sie sei gerade am Abklären mit der Gemeinde. Naters sei genau das Gegenteil der Tendenz und gerade darum spannend. Es gebe diverse Gründe, wieso die Leute aus den Seitentälern nach Naters zögen. Solche, die aus ähnlichen Fällen bestens bekannt seien – aber es werde auch ein neues Altersheim gebaut. [Annalena] meint, sie fände das sehr spannend, super.²⁴

Die Sinngebung der Datenanalyse und Produktion von Erkenntnis aus der Visualisierung fällt Susanne und Annalena in diesem Beispiel besonders leicht, weil sie sowohl die Grundtendenz (Abwanderung aus den Alpentälern des Wallis) wie auch der Ausreisser mit der gegenläufigen Tendenz (Naters) mithilfe des sonstigen Wissens (Altersheim) plausibilisieren können.

Die Beispiele verweisen auf drei Eigenschaften, welche die Attraktivität des Ausreissers ausmachen: Erstens ist er visuell auffällig und weckt damit Aufmerksamkeit. Zweitens sind Ausreisser eng verknüpft mit der Idee, Neuheiten zu ›entdecken‹ – einerseits durch die statistischen Verfahren, welche für die Feststellung von Ausreissern sowie für Neuheit die gleichen sind (Pimentel et al. 2014), andererseits bergen sie, gespiegelt auf den Journalismus, das Versprechen, dass Neuheiten *Neuigkeiten* und damit *news* sind. Drittens bieten Ausreisser, die per Definition auf einzelne Fälle referenzieren, eine ideale Ausgangslage für die Ergänzung von qualitativen und quantitativen Erkenntnissen und deren Passendmachung – Abschnitt 5.1.4 geht auf diese Praktiken der Plausibilisierung näher ein.

23 56_20190108_BR-Data-10 (4)

24 24_20180327_NZZStorytelling-12 (4)

Daten geografisch verteilen

Das Verteilen von Daten nach geografischen Variablen, das *Mappen*²⁵ bzw. *Kartieren*, ist eine oft gewählte Vorgehensweise zur Erkenntnisgewinnung. Sie schafft Übersicht durch ein vertrautes Ordnungssystem (Nohr 2004). Beide Achsen sind mit vertrauten Variablen besetzt: Breiten- und Längengrad. Bei diesem Wiedererkennen handelt es sich um wichtiges methodisch-technisches Wissen, das dem Publikum erst erlaubt, eine Datenvisualisierung zu verstehen (Passmann 2013). Deshalb schreiben Datenjournalist*innen ihm zu, dass es Visualisierungen auf Karten versteht – und gleichzeitig ist die Karte auch für Redaktionsmitglieder ausserhalb der datenjournalistischen Teams aus den gleichen Gründen die zugänglichste Datenvisualisierungsform. Sie wird bei der NZZ mit am häufigsten gebraucht, sei es als ›Auftrag‹ für das Infografik-Team oder produziert mittels Q.

Jedes dritte NZZ-Storytelling-Mitglied wünscht, sich Fähigkeiten in QGIS anzueignen oder diese zu verbessern – wobei es sich dabei um diejenige Software handelt, mit welcher im Team sogenannte Choroplethenkarten produziert werden.²⁶ Einige Monate nach meiner Feldforschung bei der NZZ ist es dann auch eine Choroplethenkarte, welche es als (vermutlich) erste auf die Meinungsseite der NZZ-Printausgabe schafft²⁷:

Die Geschichte *dreht* sich hier um die Emigration aus osteuropäischen Staaten, was durch die Einfärbung *augenfällig* wird. Begonnen hat das Projekt an einer Sitzung während meiner Forschungsphase, wobei damals erst von Daten zu (demografisch) ›schrumpfenden‹ osteuropäischen Ländern die Rede war²⁸ – woraus schliesslich in Kooperation mit einem Auslandredaktor, ›mit dem wir immer gut zusammenarbeiten‹²⁹, ein Beitrag über die Emigration und deren Gründe wird.

-
- 25 Ist im deutschsprachigen Datenjournalismus von »Mapping« die Rede, so meist im Sinne des Kartografierens auf einem geografischen Koordinatensystem – auch wenn der englische Begriff »mapping« ganz generell das Abbilden auf einem Koordinatensystem (welcher Art auch immer) meinen kann.
- 26 13_20180312_NZZStorytelling-01 (28) – Choroplethen sind Karten, auf denen die Einheiten (beispielsweise Länder) auf der Ausprägung einer Variablen eingefärbt oder konturiert sind.
- 27 Twitter-Austausch mit David Bauer vom 2. August 2018.
- 28 21_20180322_NZZStorytelling-09 (4)
- 29 Twitter-Austausch mit David Bauer vom 2. August 2018.

Abbildung 8: Choroplethenkarte des Meinungsbeitrags »Europas verdrängte Revolution wirkt nach«.



NZZ vom 2. August 2018, Seite 10

Mittels Karten wird das geografische Koordinatensystem überlagert mit einer ordinalen oder metrischen Variable (in Abbildung 8 mit Zu- bzw. Abnahme der Bevölkerung). Das Koordinatensystem wird dafür in der Regel in politischen Gebietskörperschaften kategorisiert: Kommunen, Regionen, Bezirke, im Beispiel hier Länder, oder Kontinente. Diese Relationierung der Daten mit politischen Entitäten ermöglicht einerseits einen schnell augenschein-

lichen Vergleich. Andererseits korrelieren mit politischen Entitäten und deren Grenzziehungen weitere strukturelle Merkmale. So lassen sich an politische und gesellschaftliche Diskussionen schnell anschlussfähige Erkenntnisse treffen und die politischen Verantwortlichkeiten klar benennen – wie im Meinungsbeitrag der NZZ. Dies macht Choroplethenkarten im Journalismus besonders attraktiv.

Karten als Visualisierungen weisen ihre Tücken auf (siehe Cairo 2019). Bei BR Data wird ihre Aussagekraft kritisch reflektiert, als Johannes an einem Teammeeting von einem Projekt mit Daten über die Funktionstüchtigkeit der Aufzüge an den deutschen Bahnhöfen erzählt:

[Mario] habe ihm [Johannes] SQL so eingestellt, dass man die defekten Aufzüge auf einer Karte ansehen und verorten könne. Es gebe keine Überraschungen, viele defekte Aufzüge befänden sich in den Ballungszentren. Allenfalls müsse man eine Verbindung zur Bevölkerungsdichte machen, um die Intensität der Nutzung abzuschätzen. [Martin] meint, das könnte schwierig sein. [Moritz] meint, es sei eigentlich egal, ob ein Aufzug vielbenutzt sei. [Thomas] sagt, es sei eigentlich räumlich nur spannend, wenn es irgendwo besonders schlimm sei – sonst erscheine ihm eine Karte nicht sinnvoll, auch nicht der Abgleich mit der Bevölkerungsdichte. Es wäre besser, den Anteil an defekten Aufzügen auszuweisen. [Johannes] meint, auch die Dauer der Störung wäre spannend, er müsse mit [Mario] absprechen ob das gehe. [Martin] meint, es wäre auch gut, Einzelfälle aufzugreifen.³⁰

Trotz der verbreitet kritischen Einschätzung von Karten – auch bei der NZZ³¹ – gehören Karten in den beiden Teams zu den meistgenutzten Datenvisualisierungsformaten, weil die damit produzierten Erkenntnisse schnell auf die politische Ebene spiegelbar sind.

Auf den Abschnitt zurückblickend lässt sich zusammenfassen, dass die Metapher des *Geschichten-Drehens* soziotechnisch verwickelte Praktiken beschreibt, durch welche die Teams Relationierungen von Variablen vornehmen. Die Praktiken verorten sich in einem Spannungsfeld von *data-driven* und *story-driven*, zwischen denen die Teammitglieder iterieren.

Für die Relationierungen setzen die Teams Verrechnungen durch geometrische Verkörperungen ein. Die *Geschichten*, das heißt aus journalistischer

³⁰ 58_20190110_BR-Data-12 (1)

³¹ 13_20180312_NZZStorytelling-01 (29)

Perspektive relevante Erkenntnisse, entstehen durch journalistisch anschlussfähige Zuschreibungen, welche die Praktiken provozieren. Der Abschnitt hat beschrieben, wie beispielsweise Säulen-, Balken- und Liniendiagramme Auffälligkeiten und (Dis-)Kontinuitäten sichtbar machen, wie Normalitäts- und Abnormalitäts-Einteilungen aktiviert werden, wie Ausreisser Neuheit und Neuigkeit versprechen, wie Karten die Verbindung von Mustern mit politischen Entitäten und ihren Strukturen ermöglichen.

Die Datenvisualisierungen sind Objekte, welche die analysierenden Aushandlungen in Konversationsform strukturieren und so eine Wissensordnung zugrunde legen (Burri 2008). Gleichzeitig erfolgt eine ständige Verarbeitung und Anpassung der Objekte, um sie zu Geschichten zu drehen.

5.1.4 Auswertungen plausibilisieren

Der folgende Abschnitt beleuchtet datenjournalistische Praktiken, welche die Datenauswertungen und die Sinngebung daraus zu Erkenntnissen plausibilisieren. Zuerst werden die Voraussetzungen eingeführt, welche für die Plausibilisierung erst erforderlich sind.

Die Differenzierung von ›Daten‹ und ›Realität‹

Die beiden Teams von NZZ Storytelling und – akzentuierter – BR Data differenzieren zwischen zwei Welten: die »echte Welt« und die »Datenwelt«. Roman von BR Data erzählt begeistert über ein vergangenes Projekt, das eine Recherche vor Ort beinhaltete: Es sei toll gewesen, vom Bildschirm wegzukommen und auch mal Verbindungen zur »realen Welt« hinter den Daten zu machen.³² Und sein Teamkollege Thomas blickt folgendermassen auf eine vergangene investigative Recherche zurück:

Bei der Datenanalyse selber hätten sich dann die Hypothesen teilweise bestätigt, teilweise nicht. Er habe eher die journalistische Seite abgedeckt, und bei den Analysen hätte er sich schon stark auf [Lukas] und die [andere Medienorganisation]-Leute verlassen, die seien da wirklich »tief drin« gewesen. Den Code habe er selten angeschaut, manchmal habe er die Notebooks herausgefordert, kritische Fragen gestellt. Er habe viel Wissen über den Bereich gesammelt, mit Experten gesprochen, Ergebnisse »in die echte Welt« übertragen und geprüft.

³² 47_20181210_BR-Data-01 (12)

Auch habe er geschaut, dass die Nachrichtenwerte drin seien – er habe fast schon eine Redakteurs-Rolle gehabt.³³

Thomas zählt sich innerhalb von BR Data eher zu den ›Autoren‹, also denjenigen, welche aus dem journalistischen Bereich kommen und entsprechende Expertise mitbringen. Er kann zwar programmieren, aber meistens ist er es, der die Geschichten *schreibt*. Er sieht seine Aufgabe darin, die »Verbindung [der Daten] zur echten Welt« sicherzustellen.³⁴

Besonders für Personen mit grosser journalistischer Expertise ist die Differenzierung der Welten ein Reibungspunkt in der datenjournalistischen Praktik: Daten können ein schlechtes Abbild der Welt sein und man kann sich nicht nur darauf verlassen³⁵; man muss die Daten mit der Realität abgleichen³⁶; man blickt für eine Geschichte auch darauf, was hinter den Daten steckt³⁷ – so widerspiegelt sich die Differenzierung in einigen Aussagen der Datenjournalist*innen. Sie hängt eng zusammen mit der bereits erläuterten epistemologischen Unterschiedlichkeit von digitalen Daten und Methoden der Erkenntnisproduktion (Abschnitt 3.2.3) und manifestiert sich besonders häufig im iterativen Prozess von *data-driven* und *story-driven* – nun mit einem Bezug zur Sinngebung der hergestellten Relationierung: Was bilden die Daten *wirklich* ab?

Dass sich diese Frage stellt, insbesondere bei Teammitgliedern mit grosser journalistischer Expertise, verweist auf die Unterschiedlichkeit des datengetriebenen epistemologischen Zugangs im Vergleich zum ›klassischen‹ – Abschnitt 3.2.3 hat diese Unterschiedlichkeit bereits erläutert und insbesondere zwei Punkte herausgehoben: den abstrahierten Zugang zu einem Phänomen sowie die Annahme, dass Fakten real und stabil sind. Während der letztere Punkt eine von datenjournalistischen Analysen geteilte Annahme ist, gilt es den ersten Punkt, den abstrahierten Zugang zu einem Phänomen, mit den anderen Zugängen abzulegen, *passend* zu machen.

33 57_20190109_BR-Data-11 (5)

34 BR_Fragebogen_Mitarbeitende_08

35 BR_Fragebogen_Mitarbeitende_08

36 BR_Fragebogen_Mitarbeitende_01

37 BR_Fragebogen_Mitarbeitende_02

Die Passendmachung von qualitativen und quantitativen Zugängen

Konflikte auf epistemologischer Ebene bei investigativen, datengetriebenen Projekten hat Parasie empirisch ausführlich erforscht. Während die Studie aus Chicago (Parasie und Dagiral 2012) in dieser Arbeit zur Einordnung der epistemologischen Grundlagen im Feld diente, bietet Parasies Folgeprojekt (2014) einen geeigneten Anschlusspunkt zur Untersuchung in einem Team auf Projektebene. Parasie untersuchte die Produktion einer investigativen Recherche (»On Shaky Ground«) 2011/12 in Kalifornien mithilfe der Perspektive der »epistemischen Kulturen« (Knorr-Cetina 2002). Er beschreibt, wie die Spannung zwischen der datengetriebenen und der ›klassischen‹ Epistemologie des investigativen Journalismus – es gibt *eine* aufzudeckende Wahrheit; dafür werden verschiedene Darstellungen gesammelt und abgestimmt; die ›Fakten‹ und die Geschichte entwickeln sich simultan – nur unter hohem Ressourcen-Aufwand aufgelöst werden kann (Parasie 2014:373): »Adjusting the artifacts to the established epistemologies was a demanding process for the organization. [...] But it ended up producing collective confidence in the data.« Das Vertrauen basiert auf einer geteilten epistemischen Kultur, die dazu führt, dass Akteur*innen, die eigentlich verschiedene epistemologische Grundannahmen vertreten, das-selbe in den Daten *sehen*.

»Adjusting the artifacts« meint in Parasies (2014:372f) Beitrag das Hinzufügen von »real-life evidence« – womit genau der oben beschriebene qualitative Zugang durch Methoden wie Interviews gemeint ist – zu Datenbanken und -visualisierungen. Im Beispiel wird auf einer im Projekt entstehenden Karte eine zusätzliche Ebene hinzugefügt. Diese Ebene soll einen bestehenden Unsicherheitsbereich anzeigen, weil das Recherche-Team herausfand, dass sie die Präzision der Kartendaten nicht garantieren konnten. Gleichzeitig verpflichteten sie die Standards der investigativen Recherche dazu, wahrheitsgemäß über die Verortung der Institutionen auf der Karte zu berichten. Schliesslich war die interaktive Karte (Parasie 2014:369) nicht nur ein zentrales Resultat der Recherche, sondern ermöglichte es, die quantitativ und qualitativ gesammelten Daten zusammenzubringen und abzustimmen.

Je grösser ein Projekt, desto umfangreicher sammeln die beiden untersuchten Teams weiteres Material mittels ›klassischer‹ Methoden: Man spricht mit mindestens einer*m Expert*in, sammelt wissenschaftliche Literatur zum Thema, recherchiert bereits erschienene Berichterstattung.³⁸ Aus diesen

³⁸ In einem Projekt von BR benötigt es gar eine *readme*-Datei, um die Übersicht in der Recherche-Dokumentation zu behalten, siehe 52_20181217_BR-Data-06 (5).

qualitativen Recherchen gewinnen die Teams Wissen über das Thema, treffen Annahmen und evaluieren mögliche Betrachtungsperspektiven (»angles«) auf die Daten – wie Lea anhand des Polizeiberichte-Projekts erläutert:

Dort habe sie [Lea] ein Dossier angelegt mit einem Datenordner (alle Daten in CSV-Format), einem Literatur-Ordner (sie habe viel gelesen, auch Studien über psychologische Hintergründe), ein Daten-Exposé, wo die Ergebnisse der Datenanalyse niedergeschrieben seien, die vertraulich geführten Hintergrundgespräche unter anderem mit Polizei-Angehörigen – diese hätten viel erzählt, sobald sie wüssten, dass es nicht veröffentlicht würde, wobei es einen dann wieder vor Schwierigkeiten stelle, wenn man Zitate brauche –, dann ein sehr detailliertes Recherche-Protokoll, das alle Schritte auflistet, die im Projekt gemacht wurden, und ein Dokument namens »Background« ([Lea]: »eine kleine Master-Arbeit«), in dem sie die Hintergründe, Hypothesen, die möglichen Recherche-Richtungen (»angles«) festhalte – Möglichkeiten seien hier, die Polizei als Publizisten zu sehen und sie an entsprechenden Vorgaben zu messen, oder die Geschichte mit der »gefühlten Kriminalität« zu rahmen.³⁹

Mit welchen Praktiken versuchen die beiden Teams, die Zugänge passend zu machen, »gemeinsam zu sehen« und das »Vertrauen in die Daten« zu sichern? Es handelt sich um Praktiken der Plausibilisierung: Sample ziehen, Einzelfälle untersuchen, vergleichen.

Sample ziehen

Im Projekt über die Polizeimeldungen besteht eine der Annahmen darin, dass sich diejenigen Meldungen mit Nationalitätsnennung viel öfter auf fremdenfeindlichen Plattformen und Blogs im Netz wiederfinden als diejenigen ohne. Um dies zu prüfen, arbeitet BR Data mit dem externen Experten Gabriel zusammen, der ein *Tool* entwickelt hat, welches die Weiterverbreitung von Text im Internet nachzeichnen kann. Die Projektverantwortliche Lea versuchte bereits vor der Zusammenarbeit mit Gabriel, mittels eines kleinen Samples aus der gesamten Menge an Meldungen herauszufinden, ob sich diese im Internet wiederfinden lassen – und sie diskreditiert später Gabriels (intransparente) Methode, weil sich ihr Anteil an manuell wiedergefundenen Meldungen als

weit höher herausstellt als derjenige Gabriels. Sie macht Gabriels Resultate damit unplausibel – zuerst implizit⁴⁰, etwas später dann noch explizit, indem sie ihre manuellen Resultate höher bewertet als die (fehlenden) Resultate Gabriels: Ihre manuelle Sample-Analyse habe ergeben, dass »da etwas ist«, und man müsse nach weiteren Methoden suchen, um dies zu automatisieren.⁴¹

Kurz darauf beraten Lea und Martin über weitere quantitative Methoden. Um diese zu explorieren, soll die Programmiererin Brigitte Ähnlichkeitssuchen mit einem Sample von Meldungen vornehmen.⁴² Einige Stunden später, das Büro hat sich bereits geleert, sitzen Lea und Martin am frühen Abend gemeinsam vor dem Bildschirm und probieren eine weitere Methode aus. Tatsächlich ergibt die Suche von einigen Textbausteinen über Google Resultate – die beiden schalten über Skype Brigitte dazu:

[Martin] erklärt [Brigitte], er habe bei den Polizei-Meldungen jeweils das Kürzel am Schluss und die Ortsbezeichnung am Anfang weggenommen und diese dann in eine Google-Suche eingefügt – Google finde dann verschiedene Portale mit derselben Meldung. Es gebe aber auch Treffer bei Google, die aber nur in einigen (fett markierten) Begriffen übereinstimmten, aber eigentlich nicht die gleiche Meldung seien – dies wären gemäss [Martin] dann vermutlich False Positives, die man also anhand der fett markierten Schlagwörter identifizieren könnte. Er würde dann nur die wortgleichen Meldungen anschauen, die auf der ersten Seite der Google-Suche erscheinen. [Brigitte] hat in der Zeit ebenfalls die gleiche Meldung – [Martin] hat sie ihr über Slack zugestellt – über Google gesucht und teilweise die gleichen Ergebnisse erhalten, teilweise nicht. [Lea] meint, das sei ja auch schon nur zu verdauen, dass das so unterschiedlich sei... [Brigitte] meint, normalerweise suche sie Artikel über einzelne Schlagwörter, bei ganzem Text gebe es jeweils doch recht viele False Positives. [Martin] meint, wenn er in der Struktur der Google-Webseite der Ergebnisse schaue, dann sehe er in den Kurztexten, dass diese in einem »<div>« steckten, und dort könne man die fettgedruckten Wörter mittels »« automatisiert identifizieren (und so die False Positives aussortieren). Bei »schlechten Meldungen«, also eher »Allerwelts-Meldungen«, so [Martin], würde Google mehr Treffer ausspucken, habe er nun bei einigen Versuchen gemerkt. Aber man könne ja mal die html-Seite der Google-Seite wegspeichern und dann die <div>s anschauen. [Brigitte]

40 57_20190109_BR-Data-11 (3)

41 57_20190109_BR-Data-11 (4)

42 57_20190109_BR-Data-11 (3)

meint, sie könnte das mal versuchen zu automatisieren. [Lea] meint, so müssten sie halt die Idee, auch die umgeschriebenen Meldungen erfassen zu können, verabschieden.⁴³

Kurz darauf wird Martin beim Gedanken an die Transparenz und Nachvollziehbarkeit etwas bange⁴⁴ und am Tag darauf ergeben sich bei der Automatisierung weitere Probleme mit falschen Treffern.⁴⁵ Obwohl erwogen wird, nur ein Sample an Polizeistationen zu untersuchen oder nur ein Sample an Medienplattformen, welche die Berichte weiterverbreiten, wird schliesslich keine Geschichte zu den Polizeimeldungen veröffentlicht. Die möglichen Perspektiven auf die Daten aus den qualitativen Recherchen liessen sich nie mit der Sinngebung aus den automatisierten Datenanalysen passend machen – auch weil die Untersuchung von Samples die Nachvollziehbarkeit immer wieder in Frage stellte.⁴⁶ Die genutzten Werkzeuge (Gabriels Tool beziehungsweise Google) waren extern und die Transparenz über deren Funktionsweise nicht gewährleistet. Für die beiden Teams ist die Legitimität einer datenjournalistischen Erkenntnis eng mit Transparenz und Nachvollziehbarkeit der Vorgehensweise verbunden. Transparenz und Nachvollziehbarkeit im Datenjournalismus wurden in der Forschung bisher vor allem als Glaubwürdigkeitszuschreibung von aussen diskutiert (siehe Abschnitt 6.1.3) – hier zeigt sich, dass es bereits für die Sinngebung in der Erkenntnisproduktion unabdingbar ist, dass die Beteiligten die automatisierten Schritte nachvollziehen können. Das Changieren zwischen manuell untersuchten Stichproben und automatisierten Analysen grosser Datensätze gehört in beiden Redaktionen zur datenjournalistischen Praktik der Erkenntnisproduktion. Ein Sample zu untersuchen dient dazu, die automatisierten Analyseschritte, welche sonst vom Computer übernommen werden, selbst zu absolvieren. Manuelles Nachvollziehen der Vorgehensweisen ermöglicht, die Validität der Automatisierung zu überprüfen und so die Datenauswertungen zu plausibilisieren – oder wie im Falle von Gabriels Methode zu diskreditieren.

43 57_20190109_BR-Data-11 (6)

44 57_20190109_BR-Data-11 (6)

45 58_20190110_BR-Data-12 (9)

46 Im Follow-Up-Interview rund ein halbes Jahr später meint Martin zum Polizeimeldungen-Projekt, dass sich die Thesen in den Daten nicht bestätigt und die Datenquellen für die Nachverfolgung der Meldungen im Internet sich immer wieder verändert hätten – schliesslich sei es »zu kompliziert« geworden. (Follow-Up-Interview mit Martin von BR Data vom 10. Juli 2019)

Einzelfälle untersuchen

Eine weitere Praktik zur Plausibilisierung ist die Untersuchung von Einzelfällen. Abschnitt 5.1.3 hat bereits erläutert, wie die Suche nach Ausreisern und Extremfällen wichtige Ziele der explorativen Datenanalyse sind. Beides sind Klassifizierungen, die per Definition nur auf wenige Fälle zutreffen. Die Praktik verweist darauf, dass beide Teams neben der Analyse von grossen Datenmengen ergänzend einzelne Fälle beleuchten. NZZ-Datenjournalistin Susanne erklärt die Vorgehensweise in ihrem Projekt, welches die Bevölkerungsstruktur der Schweizer Gemeinden untersucht:

Auf einem Scatterplot würde man sehen, welche Gemeinden wachsen würden. Es seien unglaublich viele Daten und es sei nun die Frage, was genau sie zeigen würden. Die Standard-Annahme des Bundesamts für Raumentwicklung, wo von sie die Daten hätte, sei halt, dass wenn eine Gemeinde wachse, sie weniger stark altäre. Es gebe diese Korrelation schon, aber sie sei nicht so stark. [Emma] habe nun die gute Idee gehabt, einen Scatterplot aufzustellen, der auf der x-Achse aufzeige, wie viel eine Gemeinde gewachsen sei, auf der y-Achse dann wie viel älter eine Gemeinde geworden sei. Diese Grafik sei relativ komplex im Verständnis, aber wenn man sie mehrmals bringe und gut einbette, würde es funktionieren. Ich frage [Susanne], welche Aspekte der Geschichte sie denn nun betonen werde und wieso. [Susanne] meint, dass man Sachen wie die »jüngste« und die »älteste« Gemeinde schon kenne, es gebe aber noch viele Aspekte und spezielle Fälle, die noch nicht gezeigt wurden. Die Geschichte werde zudem ergänzt mit verschiedenen Reportagen, das sei gut, so könne man die »Menschen hinter den Daten« zeigen.⁴⁷

Die Datenvisualisierung dient in diesem Projekt dazu, in der Phase der explorativen Analyse *spannende* Einzelfälle zu identifizieren. Im Feldnotiz-Ausschnitt oben wie in demjenigen in Fussnote 24, Kapitel 5, (aus demselben Projekt) verläuft die Sinngebung über *spannende* Einzelfälle. Das heisst, für die Beteiligten geht es darum, dass die sinnhafte Erschliessung der Relationierung der Variablen – in diesem Falle Durchschnittsalter und Bevölkerungswachstum – über deren Untersuchung am Einzelfall passiert.

Auch in der Fortsetzung von Adrians Datenanalyse zu den Medikamenten (siehe Fussnoten 20 und 21 des Kapitel 5) erfolgt die Erkenntnisproduktion über die »Extremfälle«. Er hat die in die Recherche bisher nicht involvierte Lea

47 25_20180328_NZZStorytelling-13 (7)

für eine Einschätzung angefragt. Diese schlägt sofort vor, über die beiden *Extremfälle* die Beziehungen zwischen den Variablen Land und Anteil Antibiotika qualitativ zu erschliessen, mittels eines Interviews mit einem Experten:

[Lea] und [Adrian] sitzen gemeinsam vor dem Bildschirm, vor sich eine Excel-Datei, in der die sehr simple Tabelle und ein gestapeltes Säulendiagramm angezeigt sind. Die Tabelle zeigt den Anteil der Antibiotika, die auf der »Watchlist« sind (also die Reserve), und zwei weitere Kategorien. [Lea] stellt fest, es sei ja in den Ländern sehr unterschiedlich. Italien sei der Extremfall auf der einen Seite. Schweden und Island dagegen auf der anderen Seite seien »super«. [Adrian] solle doch die »Extremfälle rauspicken« und den Experten (ein Uni-Forscher zu Allgemeinmedizin, der sich mit der Verschreibung von Antibiotika befasst hat), den er noch befragen wird, darauf ansprechen, was das heisse und was die Auswirkung sei.⁴⁸

Diverse weitere Situationen würden die Praktik illustrieren können, quantitative Auswertungen mittels einer Vertiefung von Einzelfällen zu plausibilisieren⁴⁹ – Datenjournalist Roman bezeichnet diese Vorgehensweise denn auch als »klassisches« datenjournalistisches Konzept.⁵⁰ In den Sozialwissenschaften würde man von einem Mixed-Methods-Design sprechen, wobei diese Be grifflichkeit im datenjournalistischen Feld wenig gebräuchlich ist.

Das Ziel der Teams ist es, durch die Untersuchung eines Einzelfalls die durch Visualisierung aktivierte Relationierung nachzuvollziehen und zu verstehen. Geometrische Verkörperungen bilden als »visuelles Vokabular« die Syntax (siehe Abschnitt 5.3), sind aber nicht in der Lage, die Qualität oder Modalität der Relationierung zu spezifizieren (siehe Abschnitt 4.2.2). Um mittels einer Datenvisualisierung Erkenntnisse zu produzieren, untersuchen die Teammitglieder deshalb Einzelfälle mit qualitativen Methoden – beispielsweise wie in der erwähnten Situation mit der Befragung eines Experten –, um die Art der Relationierung nachzuvollziehen.

48 53_20181218_BR-Data-07 (5)

49 siehe bei BR Data weitere Praktiken und Aussagen: 47_20181210_BR-Data-01 (12); 47_20181210_BR-Data-01 (8)

50 47_20181210_BR-Data-01 (12)

Vergleichen

Um Datenanalysen zu plausibilisieren, nutzen die Teammitglieder von BR Data und NZZ Storytelling sehr oft die Praktik des Vergleichs. Für NZZ-Storytelling-Teammitglied Christoph gehört der Vergleich ganz essenziell zur Arbeit mit Daten im Journalismus und er definiert die Praktik mittels des Inhalts: Es [Datenjournalismus] ist der Versuch, Daten zu vergleichen, um eine Aussage zu machen, Trends und Strukturen zu erkennen.⁵¹

Bei BR Data wird fast in jedem Projekt während der Datenanalyse verglichen, um aus den Daten Sinn zu generieren. Beliebt sind insbesondere Vergleiche unter geografischen Entitäten wie Bundesländern oder Ländern⁵² – bereits die Ausführungen zur geografischen Verteilung in Abschnitt 5.1.3 haben gezeigt, dass die Überlappung von Datenkategorisierungen mit politischen Entitäten das Drehen von Geschichten vereinfacht, weil die Sinngebung anhand polit-struktureller Gegebenheiten erfolgen kann. Die schnörkellosen Balken- oder auch Liniendiagramme der explorativen Phase laden das Auge zu einem sofortigen visuellen Vergleich ein – die Feldnotiz (Fussnote 17, Kapitel 5) zu den Polizeimeldungen hat bereits eine solche Praktik beschrieben. Für Barlow (2015) ist es eine der Hauptfunktionen von Datenvisualisierungen, schnell Vergleiche zu ermöglichen – Brooker et al. (2018) haben dies bereits empirisch für Datenvisualisierungen in der Astrophysik aufzeigt, Burri (2001) für Visualisierungen in der Medizin. Lowrey und Hou (2018:7) stellen den (visuellen, quantitativen) Vergleich gar ins Zentrum ihrer inhaltsorientierten Definition von Datenjournalismus: »informational, graphical accounts of current public affairs for which data sets offering quantitative comparison are central to the information provided.«

Ob mittels Visualisierungen oder nicht: Vergleiche können einerseits ad hoc und simpel passieren – wie in der Besprechung zu einem Projekt der Nutzungsbedingungen und Datenschutzerklärungen, als Lea energisch darauf hinweist, dass Microsofts Nutzungsbedingungen 49'000 Wörter umfassten, ihre BA-Arbeit aber nur 10'000.⁵³ Der Vergleich kann aber auch ein sehr aufwendiger Schritt sein. So schlägt NZZ-Datenjournalist Yannick in einer Gruppendiskussion zum Fussballtransfer-Projekt vor, dass die Aussage der Datenvisualisierung sein könnte, dass die Schweizer Fussballliga eine

51 NZZ_Fragebogen_Mitarbeitende_08

52 47_20181210_BR-Data-01 (13); 50_20181213_BR-Data-04 (3); 51_20181214_BR-Data-05 (2); 59_20190111_BR-Data-13 (5)

53 48_20181211_BR-Data-02 (3)

Ausbildungsliga sei – man müsse dazu die Daten aus der Schweizer Liga mit einer anderen *vergleichen* und dafür deren Daten scrapen.⁵⁴ An diesem Beispiel zeigt sich, wie sich die Teammitglieder die Operationsweise von Vergleichen nutzbar machen: die Kombination von Gleichheitsunterstellung und Differenzbeobachtung (Heintz 2010:164). Das heisst, durch den *Vergleich* der Ligen soll die *Differenz* bezüglich der Kategorie Ausbildungsliga verifiziert werden. So verläuft die Sinngebung durch eine Vergleichspraktik.

Neben dem Vergleich über verschiedene Kategorien wie Länder hinweg ist die Relationierung verschiedenartiger Datensätze eine weitere Form der Vergleichspraktik. Diese bringt meistens die Notwendigkeit mit sich, Datensätze auf ihre Kategorienbildung und Variablensetzung hin zu untersuchen, sie allenfalls zu transformieren, um dann einen Vergleich anstellen zu können. Sie müssen *kommensurabel* gemacht werden: »Commensuration is characterized by the transformation of qualities into quantities that share a metric, a process that is fundamental to measurement« (Espeland und Sauder 2007:16). Lea und Martin diskutieren, ob sie die Herkunftsennennung in den Polizeimeldungen mit einer Variable (Anteil an ausländischen Straftäter*innen) aus der Kriminal-Statistik korrelieren wollen – oder, anders ausgedrückt, ob die beiden Statistiken überhaupt *kommensurabel* sind:

[Lea] meint, es sei ja auch spannend, der Standard [österreichische Zeitung] habe untersucht, ob sich die Anweisung des Innenministers ausgewirkt habe. Und das dann auch mit der Polizei-Kriminal-Statistik verglichen. [Lea] zeigt [Martin] den Online-Artikel und die Grafik mit Über-/Unterrepräsentation. [Lea] sagt, der Vergleich mit der Statistik sei »völlig ohne Hemmungen« passiert, vielleicht müsste BR Data auch nochmals darüber diskutieren. [Martin] meint, ein Vergleich sei halt schwierig, die Datenerhebungen seien sehr verschieden. Ein Ereignis könne mehrere Meldungen [von Straftaten] auslösen. Deshalb würde er es nicht machen. [Lea] meint, es habe sie verunsichert, dass der Standard das einfach so mache.⁵⁵

Es trifft hier Leas Perspektive, sich durch eine Verbindung mit einem anderen Datensatz und einem damit ermöglichten Vergleich eine zusätzliche Interpretationsoption zu verschaffen, auf diejenige Martins, der auf die Fallstri-

54 14_20180313_NZZStorytelling-02 (27)

55 55_20190107_BR-Data-09 (8)

cke verschiedenartiger Datenerhebungskriterien hinweist und argumentiert, dies verunmögliche einen sinnhaften Vergleich.

Der Vergleich durch das Übereinanderlagern von Variablen aus anderen Datensätzen stellt stets eine reizvolle Möglichkeit dar, Korrelationen zu produzieren und sie als sinnhafte Kausalitäten zu plausibilisieren. Martin erzählt an einem Teammeeting von einer Projektidee zum Thema Bildung, zu der Johannes bereits einige Recherchen vorgenommen hat:

BR Data selber habe ja schon mal das Thema Schulsubventionen angedacht und die These gehabt, die Verteilung dieser sei in Bayern sehr unklar. [Johannes] sagt, er habe ein Pilotprojekt aus München gefunden, das gute Erfahrungen mit gezielten Subventionen gebracht habe. [Martin] sagt, das Thema würde auch in anderen Städten diskutiert. [Johannes] meint, er werde weitere Städte recherchieren. [Lukas] sagt, man müsste eine Datenbank aufbauen, die über Einzelfälle hinausgehe, und dann das Verhältnis zum Sozialindex anschauen.⁵⁶

»For statisticians, there always have to be comparisons; numbers on their own are not enough« (Gelman und Unwin 2013:5). Vergleiche sind also seit langem ein bewährtes Mittel, um aus Zahlen Sinn zu generieren und so Auswertungen zu plausibilisieren. Durch die Quantifizierung und Datafizierung unserer Lebenswelt – auch im Datenjournalismus – sind Vergleichspraktiken ubiquitär geworden. Visualisierungen unterstützen sie durch spezifische Formen wie Balkendiagramme oder Small Multiples⁵⁷ (Healy und Moody 2014:110).

Wenn Daten nicht abbilden

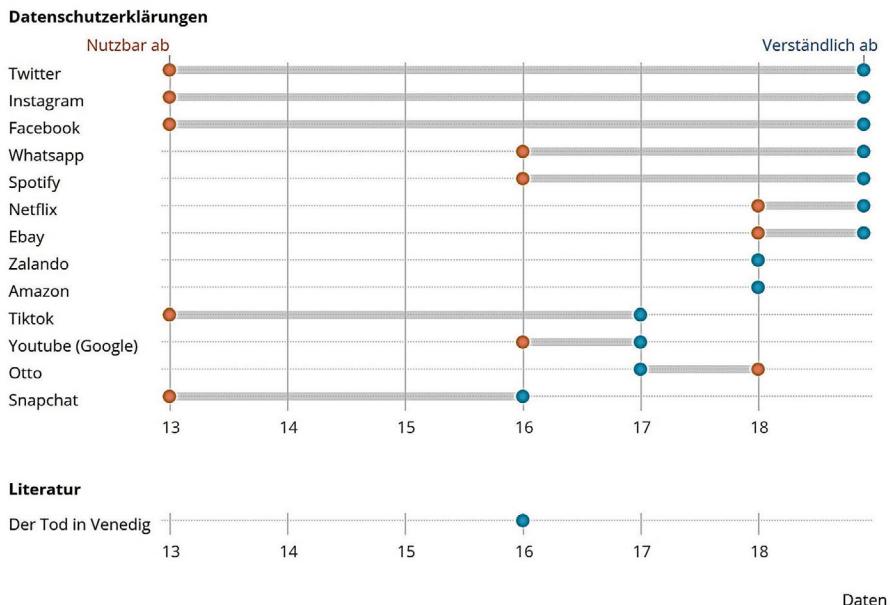
Wie bereits der erste Abschnitt dieses Kapitels gezeigt hat, drehen sich viele Praktiken in beiden Teams um die Beziehung zwischen ›Daten‹ und ›Realität‹ und deren Passendmachung. Die darauffolgenden Abschnitte zeigten ein Set an Praktiken der Passendmachung auf, welche Auswertungen plausibilisieren und so die Produktion von sinnhaften Erkenntnissen ermöglichen. Den Abschluss dieses Kapitels bilden folgend Ausführungen darüber, was in den beiden Teams passiert, wenn Daten nicht abbilden beziehungsweise die Passendmachung nicht erfolgreich ist.

56 56_20190108_BR-Data-10 (2)

57 Small Multiples referenziert auf die Visualisierung in Form vieler kleiner, simpler Diagramme, in welche die Visualisierung aufgeteilt wird (faceting). Small Multiples sollen einen schnellen Vergleich verschiedener Variablen ermöglichen (Tufte 1990:67).

Die Frage, was Daten abbilden können, ist für BR Data eine Abgrenzungspraktik gegenüber anderen Redaktionsteilen: Das Team sieht sich zuständig für Fragen (*Thesen*), welche man mittels Daten untersuchen (*prüfen*) kann. Die Mitglieder sehen ihre Expertise (auch) darin, dass sie die »Einschränkungen« der Abbildungen durch Zahlen reflektierten⁵⁸ und sie den »nicht daten-affinen« Kolleg*innen aus anderen Redaktionen erklären müssten, dass sich vieles halt nicht in den Daten zeige.⁵⁹

Abbildung 9: Visualisierung des Verständlichkeitsniveaus von Datenschutzerklärungen aus dem BR-Data-Artikel »Komplizierte Datenschutzerklärungen: Der Haken mit dem Häkchen«.



<https://web.br.de/interaktiv/datenschutzerklaerungen/> (25.02.2020)

58 47_20181210_BR-Data-01 (5)

59 49_20181212_BR-Data-03 (9)

In einem eigenen Projekt, in dem BR Data die Komplexität von Nutzungsbedingungen und Datenschutzerklärungen aufzeigen und insbesondere auf die Diskrepanz zwischen dem für das Verstehen eigentlich erforderlichen Bildungsniveau und dem Mindestalter von Apps und Medienplattformen hinweisen will⁶⁰, kann nur ein Teil der Annahmen mit der Datenauswertung passend gemacht werden: Man beschränkt sich auf Datenschutzerklärungen und lässt die Nutzungsbedingungen weg, weil es sonst »zu kompliziert« geworden wäre, wie Martin im Rückblick erklärt.⁶¹ Durch die Datenvisualisierungen ergab sich nur bei den Datenschutzerklärungen ein plausibles Abbild einer (auch) qualitativ bestätigten Annahme, eine Geschichte, die sich *sehen* lässt (Abbildung 9). Die Nutzungsbedingungen kommen schliesslich im Artikel gar nicht mehr vor, weil die Auswertungen unplausibel (»zu kompliziert«) blieben.

Auch in einem gemeinsamen Projekt zwischen NZZ Storytelling und der Lokalredaktion über Wahlergebnisse erfolgt eine Abgrenzung der Expertise darüber, was die Zahlen »sagen«.⁶² Die im Raum stehende These der Lokalredaktion, dass die Mobilisierung bestimmter politischer Lager und junger Frauen insgesamt in einem Zusammenhang mit gleichzeitig stattfindenden Abstimmungen sowie mit der Quartierentwicklung stehe, sorgt bei Annalena für Stirnrunzeln:

Die These des [Lokal-]Redaktors ist es, dass neuere Wohnangebote auch eher links-liberale Leute anzögen. [Annalena] mahnt zur Vorsicht. Der Redaktor entgegnet, es gebe aber schon wissenschaftliche Studien, die eindeutig sagten, dass Junge generell links-liberal, Alte eher konservativ und Studierte nicht SVP wählten – das seien signifikante Ergebnisse. [Annalena] meint, man sollte trotzdem »ultra-vorsichtig sein«. Der Redaktor meint, er fange es [die These] qualitativ auf, er werde die Reportage mit qualitativen Elementen aus einem eben-solchen Quartier begleiten.⁶³

Die These, dass die Mobilisierung mit dem Wohnangebot zusammenhängt, lässt sich mittels Datenauswertungen und -visualisierungen nicht prüfen und passend machen. Annalena besteht deshalb auf einer Abgrenzung von einer

60 48_20181211_BR-Data-02 (3)

61 Follow-Up-Interview mit Martin von BR Data vom 10. Juli 2019.

62 20_20180321_NZZStorytelling-08 (2)

63 22_20180323_NZZStorytelling-10 (1)

quantitativen Auswertung über die Mobilisierung. Im Artikel bleiben schliesslich der datengestützte Teil mit der Datenvisualisierung (siehe Abbildung 19) und ein Reportage-Teil mit qualitativ gestützten Annahmen getrennt.⁶⁴

So lässt sich das Fazit ziehen, dass die Plausibilisierung von Datenvisualisierungen ein Prozess der Passendmachung ist, der nicht zwingend erfolgreich endet. Insbesondere gegenüber von ›Common-Sense‹-gestützten Annahmen, welche andere Redaktionsteile einbringen und überzeugt vertreten, grenzen sich die beiden Teams ab.

Die datenjournalistischen Arbeitspraktiken des Erkennens von Daten – Geschichten in Daten übersetzen (Datafizierung), die Herstellung sauberer Daten, das Drehen von Geschichten und die Plausibilisierung von Auswertungen – hat dieser Abschnitt beschrieben und mit theoretischen Überlegungen ergänzt. Durch diese Arbeitspraktiken verfügen die beiden Teams über ein in Datensätzen eingepresstes Abbild der Welt, welches verrechenbar und visualisierbar ist. Die visuelle Verkörperung der Daten durch geometrische Formen auf Ordnungssystemen produziert Anschlusspunkte für Geschichten wie beispielsweise Diskontinuitäten, Abnormalitäten, Neuheit und die Verbindung mit politischen Strukturen. Die hervorgerufenen Geschichten-Ideen der quantitativen Auswertungen plausibilisieren die Teams in einem iterativen Prozess mit qualitativen Zugängen zum Thema, indem sie Samples manuell nachprüfen, Einzelfälle beleuchten und vergleichen.

Diese Arbeitspraktiken sind angeheftet an materielle Objekte wie die Datenbank und Datenvisualisierungen. Um mithilfe dieser Objekte Daten zu erkennen und Sinn zu generieren, braucht es ein gemeinschaftliches Sehen. Diesem Aspekt widmet sich der nächste Abschnitt.

5.2 Erkenntnisproduktion als Aushandlung

Seeing data, das Erkennen von Daten, ist eine geteilte, kollaborative Praktik. Sie erfolgt mittels Aushandlungen in Diskussionen, anhand eines visuellen Vokabulars, auf formalisierter Ebene mittels Werkzeugen und durch das Imitieren von Ideen. Die folgenden Abschnitte zeigen auf, wie die Teams ihre Expertisen koordinieren, um aus den erläuterten Arbeitspraktiken sinnhafte datenjournalistische Erkenntnisse zu produzieren. Dies geschieht auf der Basis der in den datenjournalistischen Teams vorhandenen Expertisen, welche sich in

64 Zitat aus dem Artikel »No City for Old Men«, NZZ vom 27. März 2018, Seite 21.

einem Dreieck der Triaden aufspannen lassen. Abschnitt 3.2 hat die Expertisen als Schema eingeführt und anhand der Feld-Untersuchungen definiert, in jeweils den drei Aspekten Materialität, epistemologische Kultur und Professionskultur (Abbildung 2).

In den im Abschnitt 5.1 beschriebenen datenjournalistischen Arbeitspraktiken des Erkennens von Daten scheinen die verschiedenen Expertisen immer wieder durch. Vorerst lassen sie sich folgendermassen skizzieren:

- Die technische Expertise orientiert sich daran, welche Variablen überhaupt datafizierbar und kalkulierbar sind, indem sie in Datenbanken aufbewahrt und verschiedenseitig anschliessbar für Verrechnungen gemacht werden.
- Die journalistische Expertise orientiert sich daran, welche Relationen interessant als Geschichte (*story*) zu erzählen sind, indem sie mit thematischem Wissen und weiteren qualitativ gewonnenen Erkenntnissen verbunden und damit plausibilisiert werden können.
- Die grafische Expertise orientiert sich daran, welche Relationen durch verschiedene geometrische Formen und auf welchem Ordnungssystem sinnhaft verkörpert und damit sichtbar gemacht werden.

Das allgemein gehaltene erste Schema lässt sich damit auf den Visualisierungsmodus des Erkennens von Daten beziehen und mit den spezifischen Eigenschaften auffüllen, welche die Aushandlungen ausmachen, wenn Datenvisualisierungen als Arbeitsoberfläche dienen. Zuerst gilt es aber, diese Aushandlungen und Koordinationsmuster zwischen den Expertisen zu beschreiben und einzuordnen. Dafür hat der Abschnitt 4.1 die Heuristik der *trading zone* eingeführt.

Wichtig ist die Prämisse, dass sich die einzelnen Teammitglieder nicht eindeutig und ausschliesslich jeweils einer Expertise zuordnen. Im Gegenteil: Die meisten sind aufgrund ihrer Ausbildungen und/oder langjähriger Arbeitserfahrungen mit mehreren Expertisen in Berührung gekommen und haben sich diese angeeignet. Die hier vorgenommene analytische Trennung der Expertisen ist für sie im Arbeitsalltag kaum je relevant. Stattdessen sind sie eng verwoben nicht nur innerhalb des Teams, sondern oft auch in den persönlichen Rollen. Diejenigen, welche persönlich fähig sind, verschiedene Expertisen besonders gut zu koordinieren, übernehmen gar die wichtige Funktion als »traders, strategically coordinating parts of interpreted systems against parts of others«

(Galison 1997:816), siehe Abschnitt 5.2.2. Nun zeichnen aber zuerst einige empirische Situationen die Aushandlungen und Koordinationsmuster nach.

5.2.1 Koordinationsmuster in den Arbeitspraktiken

Der folgende Abschnitt beleuchtet Situationen, in denen die verschiedenen Expertisen und ihre materiellen, epistemologischen und professionellen Dimensionen für das Erkennen von Daten in den Teams koordiniert werden.

Wie Abschnitt 5.1.2 gezeigt hat, bildet die Herstellung *sauberer Daten*, das heißt granularer, geordneter und interoperabler Datensätze, die Grundlage für die Produktion von Erkenntnissen. Ein Ausschnitt aus einer Feldnotiz von BR Data lässt die dabei ablaufenden Koordinationsprozesse nachvollziehen. Im darin verhandelten Projekt wollen die Verantwortlichen Lea und Lukas Daten zur Bearbeitung von bestimmten Strafverfahren bei Staatsanwaltschaften untersuchen. Die journalistische Potenzial-Abschätzung für eine Geschichte rund um das Thema Befangenheit der Staatsanwaltschaft wird koordiniert mit den Datafizierungsprozessen: Welche Daten können erhoben werden und wie können sie geordnet werden, um sie danach verrechnen und das Phänomen mit Daten erkennbar machen zu können?

In einem ersten Schritt geht es darum, überhaupt an die Staatsanwaltschaften zu gelangen und bei ihnen die Daten für die Analyse abzufragen – für dieses Schreiben sind die beiden in telefonischem Kontakt mit einer BR-eigenen Juristin. Es ist unklar, in welcher Form die Staatsanwaltschaften die Daten bereits erheben. Bevor BR Data die Stellen anschreibt, diskutieren die beiden Teammitglieder über die Tabellen-Strukturierung und damit darüber, mit welchen Variablen und mit welcher Granularität sie die Daten erheben wollen. Sie sitzen dafür gemeinsam am Pult und blicken auf einen Tabellenentwurf auf dem Bildschirm vor ihnen:

[Lukas] fragt, ob die Tabelle als Vorlage alle nötigen Optionen der Differenzierung offen lasse für die Analyse. [Lea] fragt, ob die Zuordnung zur Einstellungs-Instanz wohl möglich sei. Sie will die Tabelle noch gemeinsam durchgehen, ihr sei die Systematik nicht immer ganz klar. [Lukas] sagt, dass sie sich ja auf die Staatsanwaltschaften konzentrierten, und ob diese als Einstellungs-Instanz immer klar zuordenbar seien? [Lea] und [Lukas] lesen einen der Artikel der Strafprozessordnung zu Einstellungen im Internet nach. Sie diskutieren über die Deutung der Aussagen der Juristin. [Lukas] meint, sie sollten in der Tabelle eine zusätzliche Spalte einfügen, um differenzieren zu können. [Lea] meint,

sie werde die Tabelle nochmals durchgehen. [...] [Lukas] fragt, ob man die Differenzierung zwischen Einstellungen der Staatsanwaltschaften und der Gerichte aus den Zahlen lesen könne oder ob man sie abfragen müsste im Schreiben? Falls man sie abfrage, bestehe die Gefahr, dass es Widersprüche gebe zu anderen Zahlen. [Lea] meint, sie werde die Tabelle durchgehen und versuchen, diese zwischen den Instanzen auszudifferenzieren, um dies herauszufinden. [...] [Lukas] fragt, ob man eigentlich nicht gleich die vorhandene Tabelle des Ministeriums nutzen könne, um Auswertungen zu machen – [Lea] erklärt ihm dann aber, dass in dieser noch viele weitere Straftatbestände zu finden seien, weshalb man die genauen Daten erfragen müsse. Man wolle sich auf einen [Straf-]Artikel beschränken. [Lukas] fragt, wie viele Staatsanwaltschaften es eigentlich seien, die sie anfragten. [Lea] antwortet: 117. [Lukas] ist dann entschlossen, dass sie eine Vorlage mitschicken müssten, sonst würden sie wahnsinnig, wenn es darum gehe, die Daten zu sammeln und zu analysieren, das sei zu viel Arbeit. [Lea] meint, dass sie davon ausgehe, dass andere Bundesländer eine ähnliche Struktur zur Datensammlung hätten wie Bayern. [Lukas] betont nochmals, dass sie gut überlegen müssten, wie die einzelnen Zeilen zu interpretieren und den Instanzen zugeordnet seien. [Lea] meint, sie fände es auch bemerkenswert, wie unterschiedlich von den Gerichten teilweise bestraft werde und Massnahmen gesprochen würden. [Lukas] meint, das sollten sie sich auch anschauen, wenn sie das Gefühl hätten, »da steckt was drin«. [Lea] sagt, sie habe die These, dass die Staatsanwaltschaften schon viel Ermessensspielraum hätten. [Lukas] meint, sie sollten die Tabelle nochmals mit dem Bundesjustizministerium besprechen und nachfragen, ob es ähnliche Schemata in allen Bundesländern so gebe – und falls ja, ob sie die eigenen Punkte separat abfragen müssten. [Lea] schaut sich die Tabelle vom Bundesjustizministerium an. [Lukas] gleicht die Zahlen für Bayern darin mit den Zahlen aus dem bayerischen Ministerium ab – es gibt einen grossen Unterschied. [Lukas] meint, sie sollten die Juristin dann beim Telefonat auch noch um eine Einschätzung bitten, wie auf Bundesebene die Daten erhoben würden und wie diese vergleichbar seien.⁶⁵

Dieses Gespräch zeigt, was mit *sauberen Daten* gemeint ist: durchgehend konsistent strukturierte Daten, deren Eigenschaften eindeutig einzelnen Variablen zuordenbar sind. Diese Variablen sollen dann wiederum konsistent mit Kategorien aus dem Thema übereinstimmen – hier beispielsweise Art

65 58_20100110_BR-Data-12 (8)

der Straftat (operationalisiert mit dem Strafartikel) oder Instanz, welche das Verfahren einstellt. Diese Abstimmung zwischen Kategorien aus der thematischen Recherche, einer daraus generierten vagen Hypothese (in Bayern ist der Anteil der Verfahrenseinstellungen besonders hoch) und der Strukturierung in Tabellenform erfolgt zu zweit am Pult, auf Basis des Tabellenentwurfs auf dem Bildschirm sowie weiterer Materialien auf dem Tisch oder auf dem Bildschirm. Die so vorhandenen Informationen stimmen Lea und Lukas mittels Diskussion ab. Lukas verweist immer wieder auf die programmiertechnische Konvention einer hohen Granularität der Datenbank. Aufgrund der noch sehr offenen Fragestellung tippt die Konversation viele verschiedene mögliche Thematiken im Bereich der Staatsanwaltschaften an, wobei die beiden diese immer auf ihre mögliche *saubere* Datafizierung hin abklopfen – die tabellarische Ordnung muss sinnhaft sein, die Interoperabilität ist durch die Excel-Tabelle gegeben – und dann im Raum stehen lassen. Die Praktik der Herstellung *sauberer Daten* ist stark geprägt durch die programmiertechnische Expertise, ihre Vorstellung von granularen, vollständigen, systematischen und interoperablen Datensätzen – und sie wird dergestalt mit der journalistischen Expertise koordiniert, dass die Ordnung in der Datenbank journalistische Sinngebung ermöglicht. Und zwar zu Beginn, wenn die explorative Auswertung eine grosse Varianz an Fragestellungen ermöglicht.

Die Existenz verschiedener Ansatzpunkte und offener thematischer Fragen scheint insofern wünschenswert, als damit verschiedene *Geschichten-Drehs* offen sind – die Potenzialität wirkt als Versprechen: Im Hintergrund schwingt die Vorstellung mit, dass sich bei einer guten, eben *sauberen* Datengrundlage mittels explorativer Auswertung Muster finden lassen – ein Paradigma aus dem Big-Data-Zeitalter (Kitchin 2014a; Mützel et al. 2018). Mögliche Schwierigkeiten in der Interpretation von Daten antizipieren die beiden und lassen diese stehen mit dem Verweis auf zukünftige Abklärungen mit weiterem Material, der Juristin oder einem Ministerium.

Bei BR Data liess sich das Erkennen von Daten häufig dann beobachten, wenn zwei Teammitglieder gemeinsam vor einem Bildschirm am selben Pult sassen. Gemeinsam betrachtet man Datenanalysen und -visualisierungen, um aus den Daten Erkenntnisse zu schliessen, Annahmen zu bilden und zu hinterfragen. Die Tandem-Konstellation strebt BR Data explizit für alle grösseren Projekte an.⁶⁶ Die Legitimation von Erkenntnisansprüchen bedingt

66 49_20181212_BR-Data-03 (1); 50_20181213_BR-Data-04 (3); 53_20181218_BR-Da-
ta-07 (1)

für das Team einen interpersonalen Abgleich in Form einer Diskussion der Instrumente und der produzierten Erkenntnisse. Damit verbunden ist stets die Koordination der verschiedenen Expertisen: Bei BR Data fliesst die programmiertechnische Expertise durch viele Personen mit entsprechendem Hintergrund ein und die Idee, Datenbanken als »wahre« Perspektive auf »die Realität« als neue Quellen zu erschliessen, bietet einen geteilten Koordinationspunkt mit der (investigativ-)journalistischen Kultur. Evident wird dies auch in den geführten Kurz-Interviews: Auf die Frage nach dem Nutzen von Datenjournalismus antworten viele Teammitglieder von BR Data mit einer Kombination von Verweisen auf dadurch neu ermöglichte Geschichten wie auch die unbedingt nötige technische Kompetenz, um Praktiken von mächtigen Digital-Konzernen und datengestützte Prozesse in der Gesellschaft insgesamt als *Watchdog* hinterfragen zu können – sie schliessen damit an die entsprechenden Narrative des Feldes an.

Nicht immer werden dabei alle Details verbalisiert wie exemplarisch in der Situation der Datensammlung über die Staatsanwaltschaften. In der im Abschnitt zum Drehen von Geschichten geschilderten Situation über die Analyse von Herkunftsennungen in Polizeimeldungen (Fussnote 17, Kapitel 5), in der Brigitte und Lea Säulenografiken am Bildschirm betrachten, werden wenige Schritte der Erkenntnisproduktion expliziert. Durch die gewählte geometrische Verkörperung (Säulen) erscheint beiden implizit klar, dass die Analyse auf einen Vergleich hinausläuft. Treten Unterschiede im Anteil der Herkunftsennungen zutage, sichtbar durch die unterschiedliche Höhe der Säulen, so ist dies »spannend« – eine weitere Explikation der Annahme, dass auffällige Unterschiede auf ein journalistisch interessantes Muster deuten, bleibt aus. Die grafische Expertise ist bei beiden Teammitgliedern verinnerlicht und bleibt implizit. Zur Exploration werden einfache grafische Formen, deren Variablen-Relationierung wenig komplex ist, genutzt – denn die grafische Expertise bei BR Data überlappt mit der technischen Expertise. Oder anders gesagt: Erstens stützt man sich auf Auswertungen mittels Programmiersprachen, welche ein bestimmtes, eingeschränktes Set an visuellem Vokabular für explorative Auswertungen zur Verfügung stellen. Und zweitens gibt es keine ausgebildeten Designer*innen im Team und es sind meist die Programmierer*innen, welche die Datenvisualisierungen produzieren.

So sind die Expertisen in den weiteren Situationen und Besprechungen des Polizeimeldungen-Projekts klar zugewiesen: Brigitte programmiert und rechnet und visualisiert, Lea verfügt über die Hauptrolle in der Sinngebung der Daten. So entscheidet sie mit journalistischer Expertise, dass die visuellen

Muster noch keine sinnhaften Auffälligkeiten ergeben und dass deshalb noch mehr Daten gesammelt werden sollen, um weitere Muster zu produzieren – die Situation endet wie folgt:

[Lea] fragt [Brigitte], ob es ein Problem sei, nochmals nachzuscrepen bis Ende 2018 – [Brigitte] meint nein, der Scraper würde wohl noch »irgendwo rumliegen«. Der Scraper ist gespeichert in Github, der Code und beschreibender Text dazu sind in einem html-File sauber abgelegt und abrufbar.⁶⁷

In den Situationen, in denen Adrian und Martin gemeinsam am Bildschirm die Excel-Liniendiagramme über Medikamentenverschreibungen interpretieren (siehe Feldnotizen mit Fussnoten 20 und 21 des Kapitels 5), verfügt Adrian über das grössere thematische Wissen, weil er in den Tagen zuvor die Daten online gesammelt und sich ins Thema eingelesen hat. Die Daten sind in Balken- und Liniendiagrammen abgebildet, womit Relationen zwischen Variablen oder verschiedenen Zeitpunkten sichtbar vorliegen. Dem journalistisch routinierteren Martin obliegt es dann, nach thematischer Erläuterung durch Adrian die vorliegenden Datenvisualisierungen zu koordinieren und mit der journalistischen Expertise zu bewerten – das heisst, ob eine der sichtbar gemachten Relationen als Geschichte erzählenswert ist. Martin schliesst aus den Kurvenverläufen, dass es »nichts gebe, dem man nachgehen sollte«, und dass man in den Daten nichts sehe.⁶⁸ Er schlägt während der Diskussion vor, sich den Kurvenverlauf visuell vorzustellen ohne einige ausschlagende Werte – aber auch diese imaginierte Kurve verbleibt ohne Auffälligkeiten, die Martin als relevant erachtet (siehe Ausschnitt Fussnote 21, Kapitel 5).

Die grafische Expertise koordiniert in diesem Sinne mittels ihrer Materialität, in der Regel geometrische Formen und ein Koordinatensystem, die Sinngebung und journalistische Bewertung – wie dies in den verschiedenen Aspekten des *Geschichten-Drehens* (Abschnitt 5.1.3) deutlich wird. Das vorhandene visuelle Vokabular bestimmt, mit welchen Relationierungen Daten überhaupt verkörpert und in diesen Formen von den Datenjournalist*innen *erkannt* werden können. Abschnitt 5.3 erläutert das visuelle Vokabular als koordinierende Organisationsform.

67 56_20190108_BR-Data-10 (8)

68 51_20181214_BR-Data-05 (2)

Im Team von NZZ Storytelling, einerseits auf die Kollaboration mit anderen Ressorts ausgerichtet und andererseits eigene grössere Projekte verfolgend, fungieren andere Koordinationspunkte für die verschiedenen Expertisen. Die bei vielen Teammitgliedern ausgeprägte grafische Expertise führt dazu, dass Datenvisualisierungen oft hinterfragt und diskutiert werden – begründet darin, dass die Konvention der Prüfung der akkuraten Abbildung von Informationen nicht nur im Journalismus verankert (Parasie 2014), sondern auch eine in der grafischen Professionskultur verbreitete Vorstellung ist (Kosminsky et al. 2019:50).

Die Teammitglieder holen dabei Perspektiven nicht nur von denjenigen ein, die am Projekt mitarbeiten, sondern auch von mehreren anderen Teammitgliedern. Diese Feedbacks passieren ad hoc, manchmal sogar ungefragt. Zum Beispiel über die Fussballtransfers-Datenvisualisierung: Kaspar verfügt zum Diskussionszeitpunkt über einen Entwurf in Form einer Sankey-Grafik (auch die schliesslich veröffentlichte Datenvisualisierung wird eine Sankey-Grafik sein, siehe Abbildung 17). Aufgrund der Idee aus der Sportredaktion ist es zwar klar, dass es um die Ein- und Verkäufe der Schweizer Fussballklubs gehen soll, über präzise journalistische Hypothesen für die Datenauswertung verfügt Kaspar aber nicht. Er praktiziert einen datengetriebenen Ansatz und interessiert sich primär für Muster, um überhaupt eine interessante Relatiorierung konstruieren zu können.⁶⁹ Nun entspinnt sich an seinem Pult, vor dem Bildschirm, eine Diskussion zwischen diversen Teammitgliedern, die über eine halbe Stunde dauert – hier ein Ausschnitt aus dem abschliessenden Teil der Diskussion:

[Yannick] versucht allen zu erklären, wieso die jetzige Darstellung problematisch ist: Nicht nur wegen der Durchschnittswerte, sondern auch, weil der Chart Punkte verbinde, die aber in der Realität nicht die gleichen seien (nicht die gleichen Spieler werden in der einen Liga gekauft und wieder in dieselbe zurückverkauft). Der Chart suggeriere aber genau das, dabei seien doch nicht die einzelnen Spieler-Flüsse, sondern die Geldflüsse zwischen Ligen visualisiert. [Annalena] betont, sie fände das nicht problematisch. [Emma] gibt sich zurückhaltend. [Yannick] sagt, Spielerflüsse müsste man mit einer Grafik visualisieren, die man auch für Wählerflüsse braucht. [Annalena] belehrt ihn, dass auch diese nicht auf tatsächlichen Wanderungen, sondern Regressionen beruhen. [Yannick] ist überrascht, das wusste er nicht. Bisher ging er davon aus, es handelte

69 14_20180313_NZZStorytelling-02 (20)

sich bei solchen Grafiken immer um tatsächliche Wanderungen. [...] [Yannick] bringt nochmals seine Idee vom frühen Nachmittag ein (ein Bar-Chart mit einer Null-Linie). Die Diskutanten fragen nach meiner Meinung, ich halte mich aber zurück. Mehrmals steht die Frage im Raum, ob sich denn nun die Geschichte um die einzelnen Spieler, das ländervielfältige Scouting oder einen Geldfluss-Vergleich drehen soll. Was für eine Aussage kann getroffen werden? Die Diskussion landet am Schluss aber bei ziemlich genau jener Grafik, die bereits David am Morgen vorgeschlagen hat. [Yannick] betont, dass die eine Aussage, die man treffen könne, wohl die sei, dass die Schweizer Liga eine Ausbildungsliga sei. Von [Annalena] und [Kaspar] kommt das Argument, dass man dies nur sagen könne, wenn man einen Vergleich anstelle. Am besten mit einem der Top-5-Ligen. Alle finden das eine spannende Idee. [Yannick] schlägt England vor. Er sagt, dass sein Scraper einfach umgebaut werden kann. Man müsse halt einfach bei David nochmals »einen halben Tag rausreden«. Die Diskussion kommt zu einem Ende. [Kaspar] wird morgen mit David sprechen und sagt, die anderen könnten sicher auch dazustossen. [Kaspar] muss gehen. [Yannick] sagt ihm noch halb-fragend, ob er denn nun schon mal den Scraper umbauen solle – [Kaspar] meint ja.⁷⁰

Yannick versucht, die in Flussform (Sankey) relationierten Daten mit journalistischer Expertise über Fussball zu interpretieren (er ist kein Sportredaktor, sondern aktiviert sein Allgemeinwissen darüber). Die von ihm vorgeschlagenen *Geschichten* sind aber nicht mit der vorliegenden Darstellung sinnhaft, beispielsweise die Verbindung von Punkten durch die Verkörperung als Fluss. Sein Vorschlag für die Nutzung eines Balkendiagramms (»bar chart«) ist der Versuch, eine andere Art der Relationierung der Daten (Mengen- statt Flussdarstellung) zu visualisieren. Es gelingt in der Diskussion aber lange Zeit nicht, eine journalistisch relevante Erkenntnis in Form einer Hypothese mit einer akkurate, sinnhaften geometrischen Form und ihrem impliziten Relationierungstyp übereinstimmend zu machen. Solche Aushandlungsprozesse, in denen sowohl aus journalistischer wie auch aus grafischer Perspektive eine grosse Palette an Optionen und somit grosse Kontingenz besteht, erfordern einen grösseren Abstimmungsaufwand für das Erkennen von Daten.

Schliesslich erfolgt eine Koordination anhand des Vorschlags, eine Aussage aus einem Vergleich mit einer anderen Fussballliga heraus zu treffen. Yannick

70 14_20180313_NZZStorytelling-02 (27)

proklamiert aufgrund seines thematischen Wissens, dass dies die englische Liga sein sollte. Dieser Vorschlag muss abgestimmt werden mit den technischen Möglichkeiten: Die Frage, ob Informationen über die englische Liga datafiziert verfügbar und genauso kalkulierbar sind wie die Informationen über die Schweizer Liga, beantwortet Yannick positiv, weil dasselbe methodische Vorgehen (Scraping) anwendbar ist. Die Koordination haben die Akteure dadurch erreicht, dass sie auf die programmiertechnische Perspektive zurückgreifen und den Datensatz als nicht vollständig einstufen – dies impliziert, dass sie nach der Vervollständigung eine sinnhafte Erkenntnis durch einen Vergleich produzieren können.

Solche Einigungen sind stets temporär und können im abduktiven Prozess lange folgenlos revidiert werden. So verfolgt der Verantwortliche Kaspar am nächsten Vormittag eine Visualisierungsidee, die Emma ihm auf ein Blatt Papier gezeichnet hat.⁷¹ (Abschnitt 5.4.3 vertieft den Aspekt von Handzeichnungen für die Konstruktion von datenjournalistischen Erkenntnissen.) Auch für das Projekt zur Alterung der Bevölkerung in den Gemeinden hat Emma eine Visualisierungsidee eingebracht, welche die Daten auf ein Koordinatensystem spiegelt und welche die Verantwortliche Susanne deshalb überzeugt, weil sie anschlussfähig an die journalistische Konvention ist, eine Erkenntnis sowohl als generellen Trend in Form eines Musters sichtbar zu machen als auch dieses mit Einzelfällen zu plausibilisieren (siehe Ausschnitt der Feldnotiz mit Fussnote 47, Kapitel 5).

Die Praktiken der Plausibilisierung (Abschnitt 5.1.4) koordinieren die mit technischer und grafischer Expertise hergestellten Datenvisualisierungen mit journalistischen Konventionen wie der Zuspritzung auf eine ›Wahrheit‹, der Prüfung der Nachvollziehbarkeit durch Ziehen eines Samples, der Sinngebung und Narration durch die Beleuchtung von Einzelfällen und dem Herstellen von Vergleichen. In beiden Teams prägt die eingebrachte journalistische Expertise diese Praktiken, denn für die Plausibilisierung der Erkenntnisse verlässt man sich primär auf qualitativ gewonnenes Wissen aus der Recherche, Allgemeinwissen oder Themenwissen aus Fachbereichen. Dabei adressiert die journalistische Expertise die Differenz zwischen der ›Zahlenwelt‹ und der ›echten Welt‹, die sie mittels Plausibilisierungen zu überbrücken versucht. Diese Passendmachung soll mittels grafischer Expertise und ihres Ideals der

⁷¹ 15_20180314_NZZStorytelling-03 (3) – leider lässt sich aus den Feldnotizen nicht nachvollziehen, in welcher Rahmung Emmas Idee zu Kaspar gelangte und welche Aushandlungen dabei stattfanden.

akkuraten Abbildung von realen Phänomenen koordiniert werden. Scheitert diese Passendmachung, so bleibt die Erkenntnisproduktion stehen – das Projekt zu den Polizeiberichten hat dies gezeigt.

5.2.2 *Traders I*

Die beschriebene Passendmachung kann mittels einer verbalen Einigung wie in der Diskussion über die Fussballtransfers-Visualisierung passieren. Aber auch die Vorschläge einzelner Teammitglieder, welche die verschiedenen Expertisen überbrücken, ermöglichen Koordination. Bei NZZ Storytelling wie auch etwas weniger ausgeprägt bei BR Data lässt sich feststellen, welche Personen mit ihren Hintergründen besonders oft als Vermittler*innen in Koordinationssituationen fungieren. Es sind Personen, welche sich in der Selbsteinschätzung im Expertisen-Dreieck (siehe Tabelle) nicht in einer der drei Ecken verorten, sondern deren Expertisen mehrdimensional sind. Es handelt sich um Teammitglieder wie Martin (bei BR Data) oder Emma (NZZ Storytelling), die in den Situationen eine koordinierende Stellung einnehmen, indem sie unterschiedliche Expertisen an bestimmten Punkten beidseitig anschlussfähig machen. In besonderem Masse gilt das für David, von welchem sehr oft eine Bewertung eingeholt wird, um eine Datavisualisierung zu evaluieren. Sowohl gemäss eigener Verortung wie auch aufgrund seiner Teamleiterfunktion verbinden sich bei ihm Expertisen aus dem journalistischen, grafischen und dem technischen Bereich. Die Wertschätzung seiner Bewertungen hat auch mit seiner institutionellen Position als Teamleiter zu tun, geht aber über die hierarchische Komponente hinaus: Seine Vorschläge werden meist als sinnhaft aufgefasst und ihm attestiert ein langjähriges Teammitglied, die meisten guten Ideen zu haben.⁷²

Galison (1997:816) hat Personen, welche in den Zwischenräumen verschiedener Subkulturen durch die Entwicklung von vereinigenden Sprachen (»interlanguages«) koordinieren, »traders« genannt. Diese Rolle wird je nach Situation neu verteilt, fällt aber in der Empirie gehäuft denjenigen Teammitgliedern zu, welche mindestens zwei Expertisen sinnhaft koordinieren – sei es explizit oder durch ein analytisch schwer fassbares, implizites Fingerspitzengefühl oder »tacit knowledge«. Dabei zeigt ein Blick in die jeweiligen persönlichen Hintergründe, dass es keine Rolle spielt, ob die Multiperspektivität auf verschiedenartige Berufs- oder Ausbildungserfahrung oder eine

72 20_20180321_NZZStorytelling-08 (11)

hohe Routine in datenjournalistischen Praktiken zurückzuführen ist. Dabei geht es nicht um reine Übersetzungsleistungen – Galison weist den Begriff »translation« explizit zurück (Galison 1997:47) –, sondern um die Performance lokaler Koordinationssprachen, verkörpert durch geteilte Praktiken und Interpretationen. Während Galison die Rolle nicht weiter beschreibt, hat Stark (2009:186f) die Vermittlung als einen konstruktivistischen Prozess von »actively making a new type of association« beschrieben, welche das Erkennen von Daten ermöglicht. Das deckt sich mit den in den beiden Teams beobachteten Situationen. Die koordinierenden Relationierungen entstehen insbesondere im Zusammenhang mit der Produktion oder Interpretation von Datenvisualisierungen. Stark weist darauf hin, dass der Vermittlungsprozess immer eines interaktionalen, geteilten Raumes der Sinngebung bedingt (Stark 2009:187). Abschnitt 5.5 wird zeigen, dass diese koordinierenden Relationierungen auch über Referenzierungen und Imitation anderer datenjournalistischer Geschichten verlaufen – wobei es wiederum die *traders* sind, welche diese in die Teams einbringen.

In beiden Teams befinden sich am Rande dieses Raumes diejenigen, welche Expertise nur aus einem Bereich mitbringen. Diese Mitglieder sind unterdurchschnittlich oft Teil von Aushandlungssituationen beziehungsweise befinden sich darin in einer peripheren Rolle. Bei BR Data trifft das auf die meisten Programmierer*innen zu: Sie erfüllen oft klar abgegrenzte Aufgaben in ihrem spezifischen Zuständigkeitsbereich. Dies steht in einem sich gegenseitig verstärkenden Zusammenhang damit, dass diese Teammitglieder selbst ihre journalistische oder grafische Expertise als schwach einschätzen. Bei der NZZ ist es vor allem das »Subteam«⁷³ der Infografik – ihre periphere Rolle in datenjournalistischen Praktiken gründet auf einer mehrfachen Bruchlinie: Neben der Selbsteinschätzung von fast ausschliesslich grafischer Expertise sowie der traditionell verankerten Zuschreibung einer Service-Rolle in der Redaktion unterscheiden sie sich auch durch eine eigene Organisationsstruktur, welche nicht derselben Zeittaktung folgt wie die Projekte der restlichen Teammitglieder. Die Infografiker*innen folgen einem printorientierten Tages-Nachrichten-Zyklus, während die meisten anderen mehrtägige oder -wöchige Projekte verfolgen. NZZ Storytelling unternimmt zwar vielfältige und aufwendige

73 Bei diesem Begriff handelt es sich um eine Kategorisierung aus dem empirischen Material – besonders oft manifestiert in der Abgrenzung des Infografik-Teams, beispielsweise 15_20180314_NZZStorytelling-03 (10–14).

Versuche, den »Brückenschlag«⁷⁴ zum Infografik-Subteam so oft als möglich zu vollziehen und die Organisationsprozesse und Gestaltungsstandards aller Datenvisualisierungen zu vereinheitlichen.⁷⁵ Trotzdem brechen die professionalkulturellen wie organisationalen Bruchlinien in Situationen der Erkenntnisproduktion wie auch in weiteren Aspekten immer wieder auf, wie sich in der weiteren Arbeit zeigen wird.

Neben der konzeptionellen Fassung als koordinierende Vermittler*innen oder Randfiguren liessen sich die verschiedenen Rollen auch typisieren, wie dies Baack (2018) vorschlägt. Der Autor fasst die datenjournalistische Community in Berlin als Überlappung von Journalismus und »civic tech« und identifiziert vier Typen darin, welche die beiden Konventionen »gatekeeping« (traditionelle journalistische Filterung von Information vor der Präsentation) und »facilitating« (civic tech, andere ermächtigen, selbst die Daten zu nutzen) auf einem geteilten Kontinuum ausbalancieren. Gemünzt auf die vorliegende Empirie liesse sich mit der Orientierung an Ästhetik durch die grafische Expertise eine weitere Dimension hinzufügen und ermöglichte eine Typen-Verortung. Dafür ist die Anzahl der in den beiden Teams vorliegenden Personen allerdings zu klein und dies muss deshalb der Gegenstand weiterer Forschung sein.

5.2.3 Fazit: koordinierte Expertisen für das Erkennen von Daten

Der Abschnitt hat gezeigt, wie die beiden Teams den datenjournalistischen Erkenntnisprozess innerhalb der drei Expertisen – technisch, journalistisch, grafisch – aushandeln und dadurch koordinieren. Sinnhaftes Erkennen von Daten entsteht durch die Abstimmung der grundlegenden Materialität der Datenbank und ihrer quantitativen Auswertung, deren Visualisierung mittels spezifischer geometrischer Formen und dadurch verkörperter Relationierungen, deren Passendmachung mit qualitativen Recherchen durch Hypothesen und Plausibilisierungen.

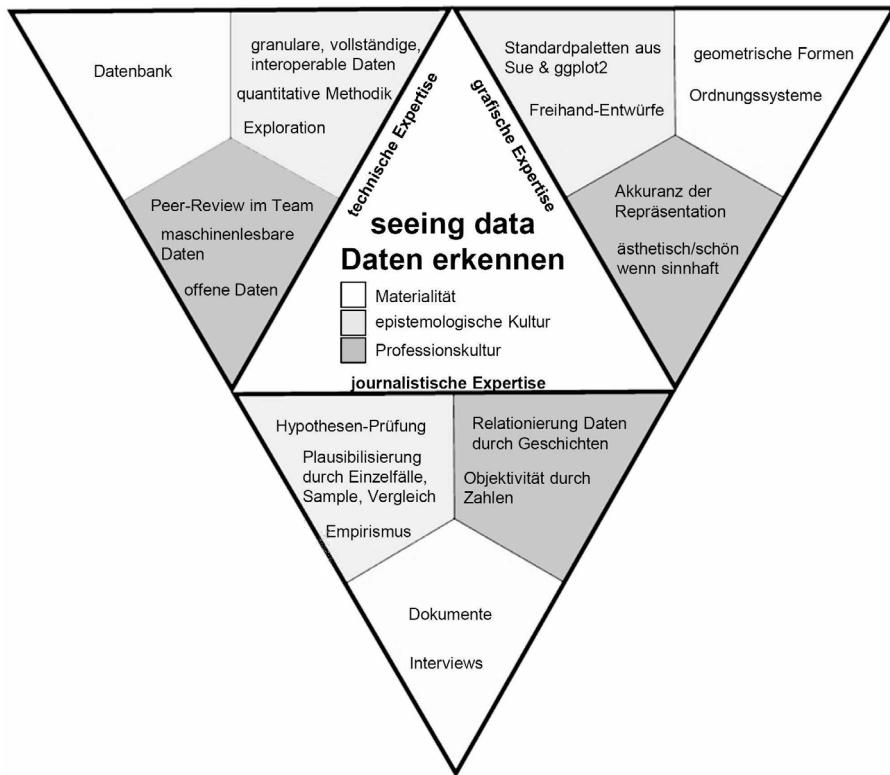
Die Abstimmung geschieht durch Aushandlungen, oft mittels mündlicher Diskussionen direkt vor einem Bildschirm, gebunden und strukturiert durch vorliegende Datenvisualisierungen – es handelt sich um kollaboratives, interaktives »Erkennen«, wie es Amann und Knorr-Cetina (1988:138ff) in der Beschreibung der »machinery of seeing« skizziert haben. Diese konversatio-

74 13_20180312_NZZStorytelling-01 (4)

75 15_20180314_NZZStorytelling-03 (10–14)

nale Praktik verbindet mittels der Visualisierung die Daten mit der ›realen‹ Welt, indem koordiniert wird, was als akkurate und objektive Erkenntnis gilt.

Abbildung 10: Koordinierte Expertisen im Modus des Erkennens von Daten.



(eigene Darstellung)

In die Aushandlungssituationen fliessen die epistemologischen Vorgehensweisen, Standards und professionellen Konventionen dreier Expertisen ein. Um die Koordination zu ermöglichen, sind vermittelnde Diskussionen oder Interpretationsleistungen von *traders*, welche mehrere Perspektiven sinnhaft zu verbinden vermögen, nötig. In den abgefragten Weiterbildungswünschen aller Teammitglieder spiegelt sich dieses Bedürfnis nach Multi-

perspektivität darin, dass die meisten sich mehr Expertise aus den jeweils anderen Bereichen wünschen.

Die unterschiedlichen Ausrichtungen – BR Data als eher investigativ-journalistische Einheit, NZZ Storytelling mit der dominierenden Aufgabe der Entwicklung von visuellem Journalismus auf der Basis von Daten – und entsprechenden Teamkompositionen widerspiegeln sich in den Gewichtungen der jeweiligen Expertisen. Und doch finden sich dieselben datenjournalistischen Praktiken zur Erkenntnisproduktion (siehe Abschnitt 5.1) schliesslich in beiden Teams. Bezuglich des Erkennens von Daten lässt sich das aufgespannte allgemeine Schema über die drei Expertisen (Abbildung 2) folgendermassen anpassen:

Das andauernde Auftreten von Aushandlungssituationen – aufgrund der Teamkomposition und der Arbeitsprozesse – hat zwei Konsequenzen: Einerseits verwenden die Teams Formalisierungen in Form von *Tools*, um den ständigen Koordinationsaufwand durch internalisierte Regeln und Standards zu senken – darauf gehen die folgenden Abschnitte ein. Andererseits bilden sich neue Wissensbestände und Subdisziplinen (Burri und Dumit 2008:301ff) – das Feld des Datenjournalismus, auf das Kapitel 3 einging.

5.3 Visuelles Vokabular

Die datenjournalistischen Arbeitspraktiken, insbesondere das *Geschichten-Drehen* (Abschnitt 5.1.3), haben die soziotechnische Verwicklung von (sozialer) Sinngebung und (technischer) Visualisierung gezeigt. Die Grundoperation der Sinngebung besteht dabei darin, die Variablen aus den Datensätzen zu relationieren – durch spezifische geometrische Verkörperungen, beispielsweise durch Diagramme oder durch Karten. Bei den geometrischen Verkörperungen handelt es sich um ein Symbol-System, das in einer Reihe mit verbalen und anderen visuellen Sprachen (zum Beispiel Gesten, Fotografie) steht (Kosminsky et al. 2019:45). Abschnitt 4.2.2 hat gezeigt, dass Datenvisualisierungen und ihre Elemente ihre eigenen semiotischen Eigenschaften, Ausformungen und Restriktionen mitbringen – sie bestimmen also durch ihre »jeweiligen Leistungen und Grenzen [...] als auch die verschiedenen Operationsarten der ›Übersetzung‹ und ›Transposition‹ selber, wobei die Frage entscheidend ist, wie diese das in ihnen aufgezeichnete oder dargestellte Wissen jeweils verändern [Hervorhebungen im Original]« (Heßler und Mersch 2009:34). Um diese Veränderungsprozesse zu untersuchen, steht hier nicht

eine medienwissenschaftliche Perspektive, sondern die Kontexte, in denen die Interpretationsleistungen passieren, im Zentrum – insbesondere die Frage nach den aktivierten Expertisen und kulturellen Perspektiven (siehe auch Aiello 2020).

Die Relationierung durch eine geometrische Verkörperung zielt auf den Kern der grafischen Expertise. *Seeing data* bedeutet, eine Visualisierungsform der Daten zu finden, welcher Akkuranz in der Abbildung eines Phänomens zugeschrieben wird – insbesondere die Diskussion rund um die Fussballtransfers (siehe Ausschnitt mit Fussnote 70 in Kapitel 5) hat dies illustriert. Das Abbilden bedingt, zuerst eine Relationierung zu konstruieren. Basierend auf der Begrifflichkeit der ›Aussage‹, die es aus den Daten zu produzieren gelte, erscheint hierfür die Bezeichnung des ›visuellen Vokabulars‹ treffend. Es erzeugt als Formalisierung eine Syntax, die Semantiken der Daten reduziert, sie aber andererseits in logische Strukturen einbettet (Heßler und Mersch 2009:35). Am Ende von Abschnitt 4.2.2 wurde bereits darauf verwiesen, dass theoretische Überlegungen das ›visuelle Vokabular‹ einbetten in eine Sprachlehre (Grammatik), welche die fundamentalen Prinzipien und Regeln für eine spezifische Fertigkeit vorgibt.

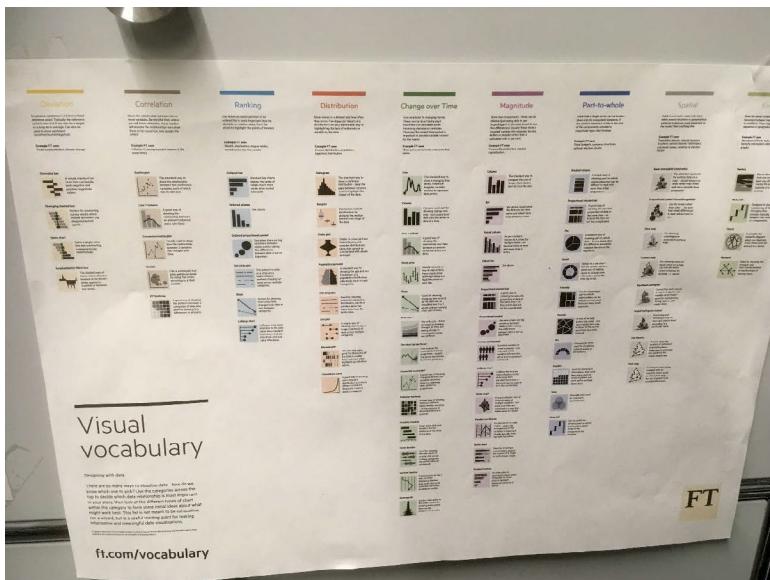
Das visuelle Vokabular referenziert dabei auf eine im Grundsatz unbegrenzte Palette an möglichen Relationierungen von Variablen durch geometrische Verkörperungen. Als grundlegende Materialität der grafischen Expertise wird sie durch alle Infografiker*innen und Designer*innen bzw. Personen mit solchem Bildungswissen in die beiden Teams gebracht. Das visuelle Vokabular fließt so implizit durch die grafische Expertise in die Aushandlungssituationen ein. Es hat aber auch bereits mehrere Stufen der Formalisierung und Institutionalisierung durchlaufen – einerseits durch verschiedene Zusammensetzungen aus dem Feld der Datenvisualisierung, andererseits dient es in verschiedenen *Tools* als Grundlage der Visualisierungsoptionen. Während der folgende Abschnitt 5.4 sich den *Tools* widmet, zuerst im Folgenden ein Blick auf die zugrundeliegende Palette.

Wird im Feld auf das visuelle Vokabular verwiesen, so führt diese Referenz sehr oft auf das »Visual Vocabulary«⁷⁶ der englischen Medienorganisation Financial Times (FT) – so hängt beispielsweise ein kleines Plakat mit dieser Palette in den Redaktionsräumen von NZZ Storytelling (siehe Abbildung 11). Die Palette der FT hat sich als Standard in der Community etabliert –

⁷⁶ <https://github.com/ft-interactive/chart-doctor/tree/master/visual-vocabulary>
(20.06.2020)

auch die Teammitglieder der NZZ verweisen immer wieder auf die hohe grafische Qualität der FT-Publikationen, und einige sind gar in direktem Kontakt mit FT-Datenjournalist*innen. Die Palette gruppiert verschiedene Visualisierungsformen nach ihrem Relationierungstyp: »Deviation, Correlation, Ranking, Distribution, Change over Time, Magnitude, Part-to-whole, Spatial, Flow« – also Abweichung, Korrelation, Rangordnung, Verteilung, Veränderungen über die Zeit, Größenordnung, Teil-zum-Ganzen-Verhältnis, räumliche Kartierungen, Flüsse (Übersetzungen RE).

Abbildung 11: Aufstellung des »Visual Vocabulary« der Financial Times in den Redaktionsräumen von NZZ Storytelling.



Das Plakat der FT verweist auf die Inspiration aus dem Feld der Datenvisualisierung: Das »Graphic Continuum by Jon Schwabish and Severino Ricca«⁷⁷, bei beiden handelt es sich um Datenvisualisierungsexperten. Vergleicht

77 <https://policyviz.com/2014/09/09/graphic-continuum/> (20.06.2020) Die beiden Autoren berufen sich auf verschiedene Inspirationen für ihre Klassifikation – darunter

man die beiden Paletten, so fallen die zahlreichen Gemeinsamkeiten auf; einzig die gruppierenden Relationierungstypen sind etwas breiter gefasst und die Visualisierungsformen spärlicher beschrieben. So lässt sich festhalten, dass grafische Expertise fast ohne weitere Formung durch journalistische Konventionen in Form des »Visual Vocabulary« in die Redaktionen einfließt.

Das »Visual Vocabulary« bietet einen Ansatz für eine optische Konsistenz und eine vereinheitlichte Perspektive, wie sie Latour theoretisierte, oder eine kulturelle Sinnstruktur, wie sie Burri bezeichnete – Abschnitt 4.2.2 hat dies ausgeführt. In den beiden Teams prägt sie, wie Daten verkörpert und dadurch gesehen werden können. Das Vokabular bestimmt in den datenjournalistischen Praktiken die validen Kombinationsmöglichkeiten und den Variationsraum der Sinngebung – die Kapitel zu den Praktiken des Geschichtendrehens und der Plausibilisierung von Auswertungen (insbesondere durch Vergleiche) haben die Rolle des visuellen Vokabulars bereits implizit beschrieben. Es entspricht damit dem, was Knorr-Cetina (2002) die Technologien beziehungsweise Maschinerien zur Produktion von Erkenntnis in bestimmten Wissenskulturen genannt hat – »[they] refer to the social and cultural structures that channel, constrain, define, and enable the production and evaluation of knowledge—indeed, such structures are both preconditions and constraints for the latter«, schreibt Lamont (2012:211) über die Bedeutung dieser Maschinerien (Abschnitt 4.2.1 hat dieses Konzept bereits ausführlich theoretisch erläutert).

Je vertiefter dabei die grafische Expertise, desto vielfältiger das Vokabular, das aus der Palette heraus auch zur Anwendung gebracht wird. Die datenjournalistischen Arbeitspraktiken in beiden Teams zeigen, dass sie im Modus der Erkenntnisproduktion hauptsächlich grundlegende Formen wie Balken-/Säulen-Vergleiche, Liniendiagramme und Kartierungen nutzen. Weil diese bereits seit längerem eingeübt sind und deshalb auch für Personen mit wenig grafischer Expertise eine schnelle Sinngebung ermöglichen, sind sie stimmig mit den schnellen und flexiblen explorativen Datenanalysen. Visuell weniger gewohnte Visualisierungsformen erfordern einen erhöhten Aufwand, um Daten sinnhaft zu erkennen. Insbesondere in der Koordination mit der journalistischen Expertise und deren Relevanzkriterien zur Produktion von Geschichten verläuft die Sinngebung anhand von visuell unmittelbar erkennbaren Auffälligkeiten: ausschlagende Kurven (Link 2001:88), Ausreisser,

Bertins »Grafische Semiole« von 1974, siehe https://policyviz.com/2014/10/06/grafic_continuum_inspiration/ (20.06.2020).

sich farblich abhebende Entitäten auf Karten. Die Sinngebung mit diesem Standardvokabular weckt gleichzeitig das Bedürfnis, dieses zu erweitern: In beiden Teams äussern viele Mitglieder den Wunsch, sich zusätzliche Visualisierungsexpertise anzueignen – sei es durch spezifische *Tools* oder das vertiefte Beherrschung von Programmiersprachen.

Diese Sinngebung passiert sehr oft im Stillen, sich nur in einer reibungslosen Produktion von Datenvisualisierungen manifestierend. Sinngebungsprozesse brechen in beiden Teams meist erst dann hervor und werden explizit, wenn sie mit Reibung verbunden sind – beispielsweise im Fussballtransfer-Projekt, als die Koordination zwischen grafischer Formgebung und journalistischer Geschichte Diskussionen provoziert (siehe Feldnotiz mit Fussnote 70 in Kapitel 5). Oder dann, wenn die *story* eben nicht mit dem gängigen Muster einer ausschlagenden Kurve verbunden werden kann, sondern NZZ Storytelling spezifische Massnahmen ergreift, um eine »Nicht-Kurve« zu einer Geschichte zu machen:

[Susanne] geht zu [Yannick] und [Annalena]: Sie habe Daten zur Namenswahl der Frauen bei Heirat – es verändere sich nichts, die Frauen nähmen immer noch genau gleich oft den Namen des Mannes an! Das sei doch echt erstaunlich. [Yannick] und [Annalena] stimmen zu – und meinen das wäre es doch wert, gerade deshalb eine Grafik zum Artikel zu machen. [Susanne] meint, sie müsste mit den Verantwortlichen auf der Redaktion schauen, ob eine Grafik zum Artikel dazukommen soll. Es wäre aber schon viel Aufwand, und sie sei wirklich nicht sicher, ob sich das lohne. [Annalena] und [Yannick] machen Vorschläge, wie man diese »Nicht-Kurve« umsetzen und anschreiben könnte. [Susanne] meint ja, dann würde sie nochmals insistieren. Es sei wohl ihre erste Grafik ohne wirkliche Kurve. [Annalena] meint, sie solle doch einen Vorschlag machen, auf Q mal entwerfen, dann sei es oft einfacher, dass sie darauf einsteigen würden. [Susanne] sagt, dann schaue sie nochmals.⁷⁸

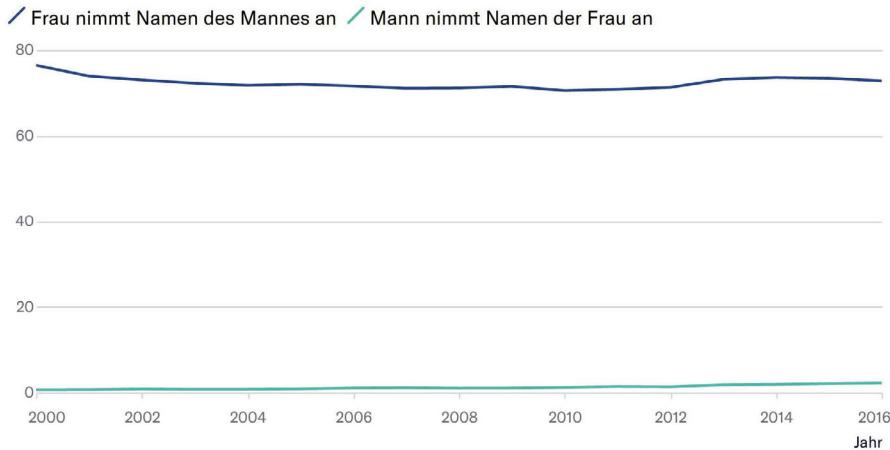
Obwohl die Kurve aus der Perspektive der drei Teammitglieder abnormal verläuft – dass sie nämlich den vermeintlichen Trend nicht widerspiegelt, Frauen würden bei der Heirat ihren Nachnamen behalten –, es sich also um eine Auffälligkeit handelt, zweifelt vor allem die federführende Susanne daran, ob sich hier eine Grafik »lohnt«. Ist plötzlich das konstant Verlaufende das Auffällige, verweigert die Datenvisualisierung als »Auflösungsmaschine«, welche

78 25_20180328_NZZStorytelling-13 (10)

die Daten mittels kognitiv wahrnehmbarer Differenz (Abweichung) zu einer Geschichte transformiert, ihren Dienst. Schliesslich erscheint der Artikel in der Print-Ausgabe ohne Datenvisualisierung⁷⁹, während sich im Online-Artikel die folgende Grafik wiederfindet:

Abbildung 12: Grafik aus dem Artikel »Die meisten Frauen nehmen noch immer den Namen ihres Mannes an«.

**Nach wie vor geben meistens die Frauen bei der Heirat ihren Namen ab
Namensführung, in Prozent**



Seit dem 1. Januar 2013 gilt das neue Namensrecht: Ehepartner können bei der Heirat ihre Namen behalten, die Wahl eines gemeinsamen Familiennamens ist freiwillig. Den Doppelnamen ohne Bindestrich gibt es nicht mehr.

<https://www.nzz.ch/gesellschaft/im-namen-meines-mannes-ld.1371279> (17.02.2020)

5.4 Tools

Die Datenvisualisierungswerzeuge der beiden Teams formalisieren bestimmte Paletten des visuellen Vokabulars und betten sie dadurch in die Arbeitsprozesse ein. Der erste Abschnitt zeichnet die Traditionslinien dieser

79 NZZ vom 6. April 2018, Seite 51.

Paletten nach und erläutert, wie sich diese mit den technischen Infrastrukturen verwickeln. Darauf folgen zwei Abschnitte, welche sich mit spezifischen Werkzeugen auseinandersetzen – *computational notebooks* und Papier – und wie diese die Koordinationssituationen zwischen den Expertisen formen und institutionalisieren.

5.4.1 Formalisierte Visualisierungssysteme

BR Data nutzt für seine visuellen Datenauswertungen die Programmiersprache R, insbesondere das *Package* ggplot2. Wie der Hauptentwickler (Wickham 2010:4) festhält, beruft sich dieses hauptsächlich auf Wilkinsons »The Grammar of Graphics« (Wilkinson 2005) sowie auf Bertins »Graphische Semiologie« (Bertin 1974). Bertin entwickelte seine grafischen Grundsätze lange bevor computergestützte Visualisierungen grosse Verbreitung erfuhren. Wilkinson schloss an Bertins Konzeptionen an, betrachtete dann Visualisierungen aber als datengestütztes System mit verschiedenen Komponenten, die variabel kombinierbar sind.⁸⁰ Das macht sein System anschlussfähig für die Entwicklung einer programmiersprache-gestützten Grammatik der Datenvisualisierungen – eben das ggplot2-*Package* in R. Es entwickelt Wilkinsons System weiter, lehnt sich aber immer noch eng an dessen Komponenten für die Produktion von Datenvisualisierungen an (Wickham 2010): *data*, *mapping*, *statistics*, *scales*, *geometries*, *facets*, *coordinates*, *theme*. Die Möglichkeit der Zergliederung in diese Komponenten ergänzt sich dabei mit dem *tidy-data*-Ordnungsprinzip in R (Wickham 2014, siehe Abschnitt 5.1.2).

Auch NZZ Storytelling nutzt R, bedient sich aber zusätzlicher Software zur explorativen Datenauswertung: das eigenentwickelte *Tool* Sue. Es orientiert sich am *Visual Vocabulary* der FT⁸¹ und basiert für die technische Umsetzung auf »Vega – A Visualization Grammar«⁸², welches sich auf verschiedene Wurzeln aus dem Visualisierungsfeld beruft, darunter wiederum Bertin (siehe Satyanarayan et al. 2017). Vega offeriert einen Baukasten für die Gestaltung einer

80 Bei diesen Komponenten handelt es sich um: *algebra*, *scales*, *statistics*, *geometry*, *coordinates*, *aesthetics*. Sie alle bauen auf *data* und *data sets* sowie *variables* und *variable sets* – siehe Schema von Wilkinson 2005:24.

81 14_20180313_NZZStorytelling-02 (7)

82 <https://vega.github.io/vega/> (27.06.2020)

Palette von Visualisierungen, welchen Designer Kaspar während zweieinhalb Monaten an die Bildsprache der NZZ angepasst hat.⁸³

Ohne hier tiefer in die Details einzutauchen, lässt sich festhalten, dass in beiden Teams die jeweilige Software für die Datenauswertung sich auf bestimmte Visualisierungspaletten und -systeme aus dem Feld der Visualisierung stützt. Es handelt sich um eine Traditionslinie von Bertin über Wilkinson bis hin zu Umsetzungen in R oder Vega – und damit um jene Systeme, welche sich durch ihre Zergliederung in Komponenten besonders gut mit der Expertise der Programmierer*innen und ihren Konventionen der Granularität und der Materialität der Datenbanken koordiniert. Um die Transponierung der Daten in die Visualisierungswerzeuge sicherzustellen, ist diese Sicherung von Interoperabilität essenziell – insbesondere für NZZ Storytelling, welches mit eigener Visualisierungssoftware und -standards arbeitet. Programmierer Jonas erklärt, wieso sich das Team für Vega als »Motor«⁸⁴, welcher die verschiedenen *Tools* integriert, entschieden hat:

Es nehme die »Grammar of Charts, quasi die ›Bibel der Visualisierungen«, auf und habe somit viele Visualisierungskonzepte schon drin. Man brauche nur noch die Parameter zu setzen, beispielsweise die Grösse und die Achsenbeschriftungen, und Vega rechne und gruppere dann vieles selber. Zudem könne Vega-Code in verschiedenen Programmiersprachen geschrieben werden (bzw. wird mit Packages übersetzt). Derzeit machten viele NZZ-Storytelling-Mitglieder mit Code ihre Grafiken – aber die Sprachen seien nicht einheitlich, es kommen verschiedene zur Anwendung. Man hätte auch sagen können, man konzentriert sich nun auf eine Sprache, das hätte sicher einiges einfacher gemacht. Aber man würde so die Fähigkeiten einiger einfach aussen vor lassen, und einige müssten neue Sprachen lernen. So habe man als Kompromiss Vega gewählt, das alle integriere.⁸⁵

Die Ziele der Infrastruktur-Wartung durch die Programmierer des Teams besteht also darin, Interoperabilität mit den gängigen Datenbankformaten und Programmiersprachen zu sichern und so eine Vielfalt an Daten für die Visualisierungssoftware maschinenlesbar zu machen. In dieser Software wiederum bilden die in der Visualisierungscommunity entwickelten Paletten

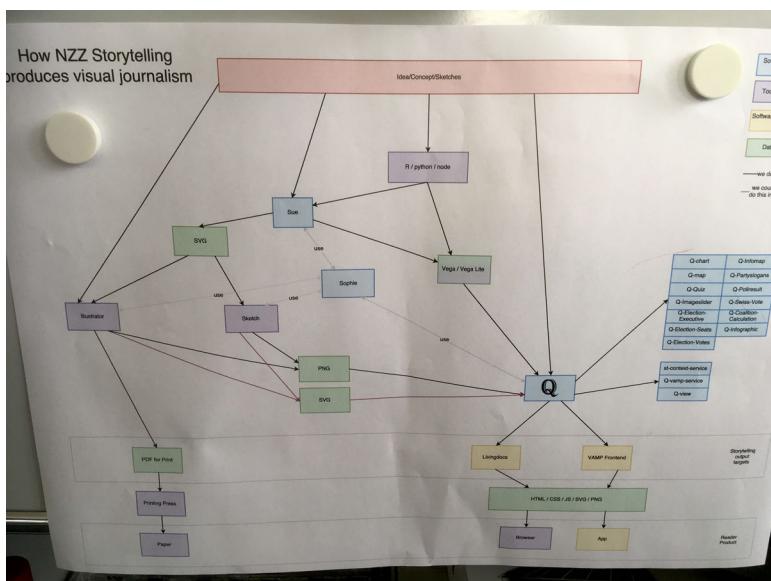
83 14_20180313_NZZStorytelling-02 (7)

84 14_20180313_NZZStorytelling-02 (7)

85 19_20180320_NZZStorytelling-07 (12)

an *grammars* oder *vocabularies* die Formung, um Daten zu erkennen und zu relationieren. Es lässt sich eine Institutionalisierung bestimmter Paletten – wie das »Visual Vocabulary« der FT sowie Wilkinsons »Grammar of Graphics« – beobachten, welche die Formalisierung durch Software prägt und gleichzeitig von ihr geprägt wird (Star 1999).

Abbildung 13: Von NZZ Storytelling erstelltes Schema der informatik-technischen Infrastruktur des Teams: Zusammenspiel von Tools, Dateiformaten und Arbeitsprozessen in der Produktion von Visualisierungen.



Die Formalisierung in Software beeinflusst die Koordination mit der journalistischen Expertise in dem Sinne, dass sie die Exploration mittels verschiedener Visualisierungstypen beschleunigt: Mit Sue kann NZZ Storytelling verschiedene Visualisierungsformen innerhalb von kurzer Zeit ausprobieren (siehe Projekt zu den Fussballtransfers, Feldnotiz mit Fussnote 13 in Kapitel 5). Designer Kaspar spricht dabei von einem »intuitiveren Zugang« zu den Daten.⁸⁶

Abbildung 14: Screenshot der Visualisierungsoptionen im Menü der NZZ-Software Sue.



NZZ Storytelling eröffnen sich noch weitere Automatisierungs- und damit Beschleunigungsmöglichkeiten der explorativen Datenanalyse, die Programmierer Fabian an einem Teammeeting vorstellt:

[Fabian] zeigt Vega Voyager, ein Interface, wo man einfach Daten einfügen kann, und das dann direkt Vorschläge für Visualisierungsarten liefert. Wenn man eine Grafik anwählt, werden direkt »related views« vorgeschlagen, also Abwandlungen der gewählten Grafik, aber mit noch mehr in die Visualisierung integrierten Daten (zum Beispiel die gewählte Darstellung des Geschlechts mittels Balken nun noch gestapelt nach Gebietseinteilung). Von der angewählten Grafik kann man auch direkt den Code rauskopieren.⁸⁷

Die Präsentation schliesst mit wohlwollendem Applaus – und obwohl ich in meiner Forschungsphase den Einsatz von Voyager nicht beobachten konnte, so zeigt sich, wie hoch der Stellenwert einer flexiblen und schnellen Datenvizualisierung für die Exploration von potenziellen Geschichten ist.⁸⁸ Durch die

87 15_20180314_NZZStorytelling-03 (6)

88 Einer der Hauptentwickler von Voyager, Jeffrey Heer, betont diesen Punkt in einem Podcast: »So the idea with Voyager is: How do you get this more broad exposure

verschiedenen möglichen Visualisierungsformen bleiben – innerhalb der angebotenen Palette an Relationierungstypen – möglichst viele Optionen offen, um sie mit den Geschichtenideen und Hypothesen aus den begleitenden Recherchen abzugleichen.

5.4.2 Computational notebooks

Vega Voyager ist ein *computational notebook*⁸⁹ und leitet über zu jenem *Tool*, auf welchem die journalistische Expertise mit der technischen und der grafischen koordiniert wird, um datenjournalistische Erkenntnisse zu produzieren. Bei Notebooks handelt es sich um eine Programmierumgebung für »literate programming«, also Programmierabläufe, die auch für Nicht-Programmierer*innen lesbar sind. Notebooks bieten ein ›Interface‹⁹⁰ für die Passendmachung von qualitativen und quantitativen Zugängen. Es existieren für weitverbreitete Programmiersprachen wie R, Python oder JavaScript Notebook-Programmierumgebungen: R Markdown, Pythons Jupyter Notebook oder ObservableHQ für JavaScript.

Bei BR Data werden Notebooks im html-Format ausgegeben, das alle Teammitglieder auf ihrem Computer problemlos anzeigen können. Die Feldnotiz mit Fussnote 18 (Kapitel 5) beschreibt eine Situation, in der ein solches Notebook zum Einsatz kommt: In einem html-Dokument sind diverse Datenvisualisierungen abgebildet. Das Notebook vereinigt dabei zwei semiotische Ebenen: eine Code-Ebene, die direkt das Schreiben in und die Ausführung

to your dataset early on?« Dies solle nicht nur mögliche Qualitätsprobleme in einzelnen Variablen aufzeigen und deren Verständnis verbessern, sondern auch verhindern, dass man sich zu schnell auf *einen Geschichten-Dreh* fokussiert. »What are the different variables I might combine?« – solche Varianten solle Voyager durch automatisierte Visualisierungsproduktion und dadurch weniger Aufwand aufzeigen. Voyager soll sogar Visualisierungsvarianten empfehlen. Dafür baut Voyager auf einem fein austarierten soziotechnischen Zusammenspiel auf, »a guided tour«. Für die Empfehlung bewertet das Programm die Datenvisualisierungen nach »perceptual effectiveness principles«, welche Variante »present[s] the data in the most effective way«. Dabei sollen auch verschiedene »visual encodings« vergleichbar sein. Siehe: <https://datastori.es/72-jeff-heer>, (30.08.2020), Minuten 19 bis 28.

- 89 Die vollständige und präzise englische Bezeichnung lautet *computational notebook*. In der beobachteten Empirie war jedoch immer von *Notebooks* die Rede.
- 90 Im Englischen wird es sogar explizit ›notebook interface‹ genannt, um es von anderen Notebook- bzw. Notizbuch-Formen abzugrenzen, siehe auch https://en.wikipedia.org/wiki/Notebook_interface (04.10.2019).

von Code erlaubt. Auf dieser Ebene lassen sich Kalkulationen direkt ausführen und Datenvizualisierungen abbilden. Die Code-Ebene kann man zu- und ausklappen. Man kann sie damit der anderen Ebene unterordnen. Diese andere Ebene ist die Text-Ebene dank *markdown*-Sprache (die sich sinnigerweise von der *markup language*, der maschinenlesbaren Sprache, abgrenzt⁹¹). Sie erlaubt es, Wörter und einfache Formatierungen in gewohnter Weise zu schreiben – also *rich text* zu produzieren, das für Journalist*innen bisher gewohnte Werkzeug, um Erkenntnisse zu rationalisieren. Da ein Notebook eine Computer-Software ist, basiert auch diese Text-Ebene in letzter Konsequenz auf Code. Dieser bleibt aber verborgen. Schliesslich lässt sich noch auf die zellartige Anordnung der Elemente in Notebooks verweisen: »This new generation of notebooks is based on a linear collection of cells, each of which contains rich text or code that can be executed to compute results or generate visualizations. These cells are linearly arranged, but can be reorganized, reshuffled, and executed in any order.« (Rule 2018:14)

Notebooks bieten mit ihrer auf Code basierenden Struktur ein interoperables, flexibles Instrument, um die auf verschiedene Weise generierten Erkenntnisse und Recherchen eines Projekts zusammenzubringen. Die Notebooks lösen so die Spannung zwischen den verschiedenen epistemologischen Kulturen auf beziehungsweise liefern eine Grundlage, um diese in einer mündlichen Aushandlung zu koordinieren, um Daten zu erkennen.

Die anspruchsvolle Koordination der drei Expertisen gelingt aber nicht immer. Die Teams greifen deshalb auf ein weiteres Werkzeug zurück, wie der folgende Abschnitt zeigt.

5.4.3 Papier

Auf allen Pulten der Teammitglieder von NZZ Storytelling liegt Notizzettel immer bereit. Ab und zu setzen sie es ein, um der anwesenden Forscherin einen komplizierten Sachverhalt zu erklären. In den meisten Fällen nutzen sie es aber, um Visualisierungen zu entwerfen. Obwohl das Team über diverse Werkzeuge verfügt, die schnell computergestützt visualisieren, ist Notizzettel immer noch in fast jeder Geschichte irgendwo involviert.

91 Die Zentralität des *Markup*-Begriffs für den Datenjournalismus zeigt sich darin, dass eines der meistdiskutierten Projekte des Feldes, das investigativen Journalismus in und über die digitale Welt betreiben will, sich »The Markup« nennt: <https://themarup.org/> (04.10.2019)

Die Teammitglieder setzen dabei schon ganz zu Beginn handgezeichnete Notizen ein – zu einem Zeitpunkt, an dem ihnen die vorliegenden Daten noch nicht im Detail bekannt sind. Es dient dazu, von Anfang an *das* Hauptargument ins Zentrum zu stellen, es entsprechend visuell herauszustellen.

[Caroline] zeichnet auf vor ihr liegendes Notizzettel erste Skizzen. [Die Auslandkorrespondentin] meint, der rote Faden der Geschichte sei ungefähr, dass das Wassermanagement der Stadt einfach unfähig sei. Die Stadt sinke ab, das Problem sei multifaktoriell. [Caroline] schlägt vor, eine grosse Darstellung mit einem Querschnitt zu machen. Und dazu »Small Multiples« mit verschiedenen Zukunftsszenarien, dargestellt als Skizzen.⁹²

Handgezeichnete Entwürfe begleiten auch die explorative Datenanalyse im Fussballtransferprojekt: Als David den Auftrag für die Datenanalyse und -visualisierung im Rahmen eines kleinen Meetings an Kaspar und Caroline überträgt, liefert er gleich einen Visualisierungsentwurf mit:

Es ist eine »Sankey«-Grafik. Es sei ja auch spannend, woher die Spieler kämen [meint David]. Außerdem schlage er vor, dann einzelne Spieler herauszupicken, um »es« greifbar zu machen. David präsentiert auf einem Excel-Sheet direkt auf seinem Laptop gleich eine Auswahl von fünf Spielern, die teuersten Transfers aus der Schweiz. Die Datensätze hat er bereits vorsortiert und präsentiert sie ebenfalls auf dem Laptop. [Caroline] nimmt dann einen Notizzettel, den sie mitgebracht hat, und zeichnet einen eigenen Visualisierungsvorschlag. Sie findet, dass eine Bar-Chart-Variante bessere Vergleichbarkeit biete.⁹³

Etwas später am selben Tag bespricht sich Kaspar mit seinem Kollegen Yannick. Die beiden diskutieren die möglichen Geschichten sowie Datenvisualisierungen zum Fussballtransferprojekt intensiv. Das iterative Datenerkennen geschieht dabei nicht nur anhand der explorativen Visualisierungen mit Sue und dem thematischen Wissen der beiden, sondern bezieht auch die handgezeichneten Visualisierungsentwürfe von Kaspar mit ein. Die Produktion der Erkenntnis richtet sich danach aus, wie sie sich *griffig*, das heisst schnell erfassbar, visualisieren lässt. Das heisst, die Daten müssen sofort *als Geschichte* erkennbar sein. Diese Evidenzmontage (siehe Kapitel 6) wirkt in den beiden

92 24_20180327_NZZStorytelling-12 (9)

93 14_20180313_NZZStorytelling-02 (3)

Teams bereits in die Erkenntnisproduktion hinein. Die Zeichnung auf Papier während der Datenanalyse sorgt für eine Selektion derjenigen Analysemöglichkeiten, die erlauben, dass am Schluss die »Visualisierung als Geschichte« taugt (Segel und Heer 2010).

Traders mit viel grafischer Expertise nutzen besonders häufig Papier, um die journalistische mit der grafischen Expertise zu koordinieren. So liegt einen Tag nach der Diskussion mit Yannick auf dem Pult von Kaspar, neben seiner Tastatur, ein Entwurf auf einem Notizblock. Er stammt von Emma, die bisher nicht in das Projekt involviert war und auch wenig Wissen über das Fussballgeschäft mitbringt. Der Entwurf ähnelt stark demjenigen, welcher bereits am initialen Meeting von David vorgelegt wurde, und Kaspar versucht nun, diesen mit der explorativen Grafiksoftware Sue nachzustellen.⁹⁴ Indem also die technischen Formalisierungen mit ihren Begrenzungen (temporär) ausgeschaltet werden, vereinfacht sich die Koordination einer Visualisierung mit einer Geschichte. Ist demnach die Abstimmung zwischen den drei Expertisen besonders komplex, entstehen oftmals datenjournalistische Erkenntnisse ohne formalisierte Datengrundlage.

5.4.4 Fazit

Der vorangegangene Abschnitt über *Tools* hat gezeigt, wie Werkzeuge zur explorativen Datenauswertung in den beiden Teams die Koordinationssituationen zwischen den Expertisen formen. Software zur Datenvisualisierung wie das R-Package *ggplot2* oder NZZ Storytellings Sue und Vega bauen auf Paletten von Visualisierungsformaten auf, welche sich in Komponenten zergliedern lassen. Sie ermöglichen so im Zusammenspiel mit der Datenbank-Materialität und der technischen Expertise eine hohe Varianz an Visualisierungsoptionen und eine teilweise Automatisierung von explorativen Methoden. Die Interoperabilität zwischen den Programmen ist dabei eine Voraussetzung für eine vernetzte technische Infrastruktur, wie sie NZZ Storytelling aufgebaut hat. Diese Grundlagen ermöglichen ein flexibles, schnelles und variables Erkennen von Daten mit computergestützten *Tools*. Das Werkzeug des *computational notebooks* bietet dann ein *interface* zwischen den grafisch-technischen Datenauswertungen und der journalistischen Expertise: In der Verwicklung von Code und Text werden unterschiedliche Relevanzbewertungen und epistemologische Zugänge abgeglichen, um aus Daten Geschichten zu erkennen.

Für dasselbe Ziel kommt dabei häufig Papier als Werkzeug zum Zuge. Dabei werden die computertechnischen Restriktionen temporär ausgeschaltet und die Expertisen ausserhalb von Software-Struktur koordiniert. Dadurch fokussiert die Koordination stärker auf die visuelle Verkörperung als Geschichte, indem die Visualisierung bereits in der explorativen Phase der Konstruktion einer Hauptaussage untergeordnet wird.

5.5 Ideen imitieren

[Lea] scrollt durch den BBC-Artikel und meint, so eine griffige Grafik wie die erste aus dem Artikel zu haben wäre schon gut.⁹⁵

Als Lea mit Roman und Adrian das Projekt zu den Nutzungsbedingungen und Datenschutzerklärungen diskutiert, referenziert sie auf einen Artikel der BBC⁹⁶, welcher ein halbes Jahr vorher erschien. Er basiert auf den (englischen) Nutzungsbedingungen verschiedener Internet-Plattformen und ordnet ihre Verständlichkeit im Vergleich mit klassischer englischer Literatur ein. Nach einem Austausch mit Personen der BBC war BR Data gewillt, die Geschichte auf Deutschland zu adaptieren. Lea referenziert primär auf die Grafik (so im Ausschnitt aus der Feldnotiz) sowie die Illustration mittels Kindern, welche die für sie viel zu komplizierten Texte vorlesen.

Dies ist ein Beispiel für die Praktik, dass BR-Data-Teammitglieder andere Medien dann referenzieren, wenn sie eine Geschichte adaptieren wollen. Referenziert werden dabei sehr oft Visualisierungen – so erfolgt während eines Ideen-Brainstormings bei BR Data die Bewertung eines anderen Projekts⁹⁷ anhand des Kriteriums, etwas *sichtbar* gemacht zu haben, wie eine Teilnehmerin ausführt:

95 48_20181211_BR-Data-02 (3)

96 <https://www.bbc.com/news/business-44599968> (10.03.2020)

97 <https://interaktiv.tagesspiegel.de/radmesser/> (10.03.2020)

[...], die Radmesser-Recherche habe schon gewirkt, auch wenn man vorher schon wusste, dass Radfahrer eng überholt werden. Es habe einen schon beeindruckt, das nochmals zu sehen.⁹⁸

Im Projekt zu den Polizeiberichten wiederum führt eine mitten in der Datenanalyse erscheinende Grafik des österreichischen Der Standard dazu, dass die eigene Entscheidung, einen Vergleich mit einer anderen Statistik nicht vorzunehmen, nochmals evaluiert wird (siehe Feldnotiz Fussnote 55 in Kapitel 5): Liesse sich die Erkenntnis, welche Der Standard durch den Vergleich ermöglicht hat, auch für das eigene Projekt übernehmen?

Das Referenzieren auf Visualisierungen als Sichtbarmachungen oder auf die Auswertungsmethoden – wie es in anderen Situationen in beiden Teams vorkam – ist ein Referenzieren auf Sinngebungen von Daten: Die Datengrundlage ändert sich in der Regel bei der Adaption auf den lokalen Kontext, aber es wird die Erkenntnis, die Aussage aus den Daten (und damit auch das methodische Vorgehen, mit dem diese produziert wurde) übernommen.

NZZ Storytelling verfügt über einen »Inspirations-Channel« auf Slack, wo einige Teammitglieder pro Woche rund fünf bis zehn Einträge platzieren und darauf hinweisen, wieso sie jene Geschichte »inspirierend« finden. Designer Kaspar hat die Einträge gesammelt und sie auf die Organisationsplattform trello hochgeladen. Er zeigt die Sammlung:

Er hat selbst Kategorien erstellt (ca. 20), differenziert nach Formen und Formaten von Visualisierungen. Es gibt zudem eine Kategorie, die sich ›Einschweizer‹ bzw. ›Einzürchern‹ nennt: Diese beinhaltet Ideen, die man lokal adaptieren könnte. Die Einteilung funktioniere manuell, erklärt [Kaspar], und er habe die Plattform schon länger nicht mehr nachgeführt. Es brauche sie derzeit auch niemand. Der Inspirationschannel werde schon gepflegt, aber die Dinge dann wieder zu finden sei natürlich mühsam. Ich frage [Kaspar], welche Medien oder Plattformen besonders oft vertreten seien im Channel – er meint, das seien so ›die Üblichen‹, wie Financial Times und New York Times. Er geht etwas durch die Quellen und meint dann, es seien aber auch viele kleine dabei. Ich frage ihn, wie sie zu den Quellen kämen – [Kaspar] meint, vieles laufe wohl über Twitter. Als er die Einträge der Kategorie »Scrollytelling« anschaut, meint er, das sei auch etwas ›vorbei‹. Ich frage ihn, ob es denn Trends gebe – das bejaht er.⁹⁹

Die Sammlung ist kategorisiert nach Visualisierungsformaten, was den Relationierungstyp als auch das Präsentationsformat – beispielsweise »Scrollytelling«¹⁰⁰ – als für die Teammitglieder relevantes Ordnungskriterium ausweist. Ein weiteres Beispiel ist die Grafik-Sammlung der Designerin Emma auf Pinterest. Als sie überlegt, wie NZZ Storytelling die vielen Wechsel in der Trump-Administration beleuchten könnte, erklärt sie mir ihr Vorgehen am Pult:

Ausserdem hätten die New York Times und die Washington Post schon Artikel dazu gemacht, die schaue sie sich an, was halt andere schon gemacht haben. Sie sammle nun Ideen. Ich frage sie, wie sie diese Ideen nun sucht. Sie meint, das sei eine gute Frage – sie schaue sich halt viele Beispiele und Grafiken an, in diesem Falle Ideen für eine Timeline. Sie habe dazu auf Pinterest Sammlungen angelegt – geordnet je nach Typen. Sie zeigt mir die Alben gleich auf dem Bildschirm. Es sind sehr viele für alle möglichen Sorten von Grafiken. Sie sammle diese selber zusammen, schaue sich Beispiele im Internet an.¹⁰¹

In die datenjournalistischen Praktiken beider Teams fliessen die Ideen der datenjournalistischen Geschichten anderer ein. Die Imitation verläuft über Datenvisualisierungen und deren Sinnzusammenhänge. Damit übernehmen die Teams ein bereits ausgehandeltes Schema für die Produktion von datenjournalistischen Erkenntnissen. Das Imitieren nimmt insofern die gleiche Rolle ein wie die neuen Relationierungstypen, welche die *traders* einbringen: Sie schlagen ein aus Sicht der Teams sinnhaftes Expertisen-Koordinationsmuster (in Form einer geometrischen Verkörperung) vor. So überrascht es nicht, dass in beiden Teams Personen, welche überdurchschnittlich oft die *trader*-Rolle einnehmen, genauso überdurchschnittlich oft datenjournalistische Geschichten aus anderen Redaktionen referenzieren, diskutieren und Sammlungen darüber anlegen.

Dieser Abschnitt zeigt auf einer zusätzlichen Ebene, wie sich die datenjournalistische Community gegenseitig beobachtet und welche Prozesse dies lokal, in den Teams, antreibt. Imitation bedeutet also nicht nur ein Editieren von Narrativen über die Potenziale von Datenjournalismus, wie es Abschnitt

99 26_20180329_NZZStorytelling-14 (11)

100 Unter »Scrollytelling« versteht man gemeinhin eine Erzählweise mit sich abwechselnden Text- und Visualisierungselementen, durch welche man sich einfach »durchscrollen« kann. Detaillierter: <https://medium.com/nightingale/from-storytelling-to-scrollytelling-a-short-introduction-and-beyond-fbda32066964> (31.12.2020)

101 20_20180321_NZZStorytelling-08 (5)

3.3 illustriert hat. Imitation bezieht sich *auch* auf konkrete Projekte, gar auf einzelne Datenvisualisierungen, die besonders *griffig* erscheinen in der Erkenntnisproduktion. Der Imitationsprozess ist geprägt durch die Expertisen und Kulturen in den Teams: So re-evaluerte BR Data zwar aufgrund des österreichischen Berichts zu den Polizeimeldungen die eigene Auswertungsmethode, verzichtete aber aufgrund der eigenen Ansichten über statistische Qualität auf die gleiche Vorgehensweise.¹⁰²

5.6 Fazit

Das Kapitel 5 hat empirisch nachgezeichnet, welche datenjournalistischen Arbeitspraktiken den ersten Modus der Datenvisualisierungen, das Erkennen von Daten (*seeing data*), ausmachen, welche Aushandlungs- und Koordinationsprozesse sich in diesem Zusammenhang in den beiden untersuchten Teams abspielen und welche Rolle dabei Visualisierungswerkzeuge einnehmen.

»Other researchers present in the laboratory are attracted by such events, gather around the visual materials, finger the documents and gaze about their surfaces. As they examine the film, scientists begin a series of verbal exchanges. [...] When embedded in talk, ›seeing‹ is interactively accomplished« (Amann und Knorr-Cetina 1988:138).

Genauso wie sich die Forscher*innen um ihre molekulargenetischen Laborproben versammeln, sie betrachten und diskutieren, um kollektiv Erkenntnisse zu produzieren, nutzen die beiden datenjournalistischen Teams Arbeitspraktiken, um Daten zu erkennen: Visualisierungen dienen als Arbeitsoberfläche für die Praktiken der Datafizierung, der Herstellung sauberer Daten, des Drehens von Geschichten und der Plausibilisierung der Auswertungen. Der Abschnitt 5.1 hat diese Arbeitspraktiken nachgezeichnet:

- (1) Den Praktiken zugrunde liegen die ontologischen und epistemologischen Annahmen, dass sich in datafizierter Information reale Ereignisse widerspiegeln und sich das Geschichten-Format in Daten übersetzen lässt. Dies passiert insbesondere über die Vorstellungen von visuellen Verkörperungen der Daten.

¹⁰² 56_20190108_BR-Data-10 (4)

- (2) Die Herstellung »guter« datenjournalistischer Geschichten bedingt *saubere Daten* als deren Grundlage: Die Datensätze sollen granular, systematisch geordnet und kategorisiert sowie interoperabel sein. Diese Konventionen bringt vor allem die programmiertechnische Expertise ein. Die Eigenschaften sind für die journalistische Expertise besonders anschlussfähig, weil sie ein Maximum an Optionen für mögliche Geschichten offenhalten.
- (3) Die Relationierungen der Variablen in den Datensätzen erfolgen unter dem Begriff des *Geschichten-Drehens*. Die Teams bewegen sich dabei iterativ in einem Spannungsfeld von datengetriebener (*data-driven*) und hypothesengetriebener (*story-driven*) Vorgehensweise. Geometrische Verkörperungen wie Säulen-, Balken und Liniendiagramme, Ausreisser oder Karten sind das visuelle Vokabular, um Daten zu erkennen, und diejenigen Objekte, welche die Aushandlungen strukturieren. Die dabei produzierten Auffälligkeiten, (Dis-)Kontinuitäten, Normalitätseinteilungen und Neuheiten bieten sinnhafte Anschlusspunkte für die Relevanzzuschreibungen der journalistischen Expertise.
- (4) Mittels Praktiken der Plausibilisierung machen die Teams dann die Datenvisualisierungen, deren Grundlage in der Regel eine quantitative Auswertung ist, mit den für den »klassischen« Journalismus typischen Common-Sense-Annahmen passend. Für die Plausibilisierung werden Samples und Einzelfälle vertieft qualitativ untersucht oder Sinnzusammenhänge durch Vergleiche konstruiert.

Diese Praktiken des Erkennens von Daten handeln die Teams in einem Spannungsfeld dreier Expertisen aus: der journalistischen, der programmiertechnischen sowie der grafischen. Viele Teammitglieder sind dabei in der Lage, als *traders* mehrere Expertisen zu überbrücken, und sie nehmen dadurch Vermittlungspositionen ein, welche die Koordination ermöglichen.

Die Sinngebung steht dabei in einem engen Zusammenhang mit dem visuellen Vokabular, das über die grafische Expertise in die Teams einfließt. Die Koordination kann einerseits situativ, das heißt für jede Datenvisualisierung neu erfolgen, was meist über mündliche Diskussionen passiert. Oder die entsprechenden Arbeitspraktiken sind bereits als Visualisierungssysteme formalisiert in Programmiersprachen und Werkzeugen zur explorativen Datenauswertung. Als besonders geeignete Arbeitsoberflächen zur Sinngebung haben sich das *computational notebook* und Skizzen auf Papier herausgestellt: Während Ersteres den Abgleich verschiedener Expertisen und ihrer epistemologischen Zugänge erlaubt, indem es sie in einem Werkzeug kombiniert, ermöglicht Pa-

pier das Ausschalten der technischen Beschränkungen und dadurch eine erleichterte Koordination von journalistischer und grafischer Bewertung.

6 Evidenz befestigen: Visualisierungen als Montage

Der Abschnitt 4.2.3 hat gezeigt, dass bei wissenschaftlich-technischen Visualisierungen im zweiten Modus der Evidenzbefestigung (*becoming evident*) das Ziel der Praktiken die Herstellung von Autorität und Überzeugungskraft sowie der Adaptierbarkeit (*incomplete models*) ist.

Die beiden untersuchten Teams transformieren Daten zu Evidenz, indem sie die produzierten Erkenntnisse mittels Datenvisualisierungen dem Publikum präsentieren. Es geht dabei darum, »the sense of what was seen« (Amann und Knorr-Cetina 1988:162) aufzuzeigen. Beim zweiten Visualisierungsmodus handelt sich um eine Art Fazit, eine ästhetisch aufgewertete, kuratierte Montage der im ersten Visualisierungsmodus »erkannten« Realität.

Die folgenden Abschnitte unter 6.1 identifizieren aus den empirischen Untersuchungen vier datenjournalistische Arbeitspraktiken, welche Daten in Evidenz transformieren. Wie bereits für die Arbeitspraktiken für das Erkennen von Daten koordinieren die Teams für die Montage von Evidenz ihre verschiedenen Expertisen: journalistische, programmiertechnische, grafische. Abschnitt 6.2 beleuchtet die Koordinationsmuster in Konversationen; Abschnitt 6.3 erläutert die Kodifizierung der ausgehandelten datenjournalistischen Praktiken und Konventionen in Leitlinien und *Tools*; Abschnitt 6.4 beleuchtet das Ausweichen der Teams auf neuartige Präsentationsformate, um ihre Praktiken der Evidenzbefestigung zu manifestieren.

6.1 Arbeitspraktiken der Evidenzbefestigung

Im folgenden, ausführlichen ersten Teil des Kapitels geht es darum, die Arbeitspraktiken zu beschreiben, mit welchen die beiden Teams die Befestigung von Evidenz organisieren. Sie werden innerhalb der drei Expertisen ausgehan-

delt und koordiniert. Aus analytischen Gründen erfolgt eine Gliederung in vier Praktiken beziehungsweise Sets an Praktiken, die aber in der Empirie als solche nicht klar abgegrenzt sind. Es sind dies das Hervorheben der Hauptaussage durch Leseführung und ausgewählten Vergleich (6.1.1), ›humanizing data‹ durch Personalisierung und Konkretisierung (6.1.2), die Transparentmachung von Datenquellen und Methoden (6.1.3) sowie das Ausweisen differenzierter Expertisen in der *Byline* (6.1.4). Die hier gewählte Reihenfolge soll dabei keine Chronologie der Praktiken abbilden, im Gegenteil, die Praktiken können simultan, nachgelagert oder iterativ ausgeführt werden.

6.1.1 Hauptaussage hervorheben

Die erste Praktik ist das Hervorheben einer Hauptaussage, welche die interpretative Flexibilität der Datenvisualisierung reduziert. Der Visualisierungsexperte Alberto Cairo (2016:79) beschreibt in seinem für das datenjournalistische Feld programmatischen Lehrbuch: »What you design is never exactly what your audience ends up interpreting, so reducing the chances for misinterpretation becomes crucial.« Oder wie Kaspar den Anspruch im Kurz-Interview formuliert: Visualisierungen sollen ein effizientes Stilmittel sein, um »die Essenz herauszuschälen«.¹ Um dies zu erreichen, steht den beiden Teams ein Set an Praktiken zur Verfügung, die einzeln oder kombiniert eingesetzt werden: Lesart, Leseführung und Lesebeispiele sowie montierte Vergleiche. Vor dem Beschrieb dieser Praktiken leiten die folgenden Absätze den Abschnitt ein, indem sie das Hervorheben einer Hauptaussage mit dem Grundsatz der Visualisierung als *Geschichte* verbinden.

Das Team von NZZ Storytelling hat viele dieser Methoden schriftlich festgehalten: in einem *Styleguide*². Abschnitt 6.3 wird diesen Leitfaden kontextualisieren und seine Rolle als Werkzeug ausleuchten – der hier unmittelbar folgende Teil begnügt sich vorerst damit, auszurollen, dass NZZ Storytelling einige Gestaltungsregeln expliziert hat, die sich als Praxis in den Datenvisualisierungen des Teams wiederfinden.

Das erste Kapitel des *Styleguide* stellt vier »visual principles«³ auf:

1 NZZ_Fragebogen_Mitarbeitende_04

2 Der *Styleguide* ist im Internet zugänglich: <https://nzzdev.github.io/Storytelling-Styleguide/#/> (18.04.2020).

3 <https://nzzdev.github.io/Storytelling-Styleguide/#/principles> (18.04.2020)

Determined – Entschlossen

- Every graphic has a purpose and a message that becomes apparent immediately
- Clarity and readability is never sacrificed for beauty

Accentuated – Akzentuiert

- We highlight what's most important
- We guide the viewer

Precise – Präzis

- We use the best available data and take no shortcuts
- We use clear and simple language

Authentic – Authentisch

- Each graphic is reflective of our collective voice and vision as the NZZ Visuals team⁴

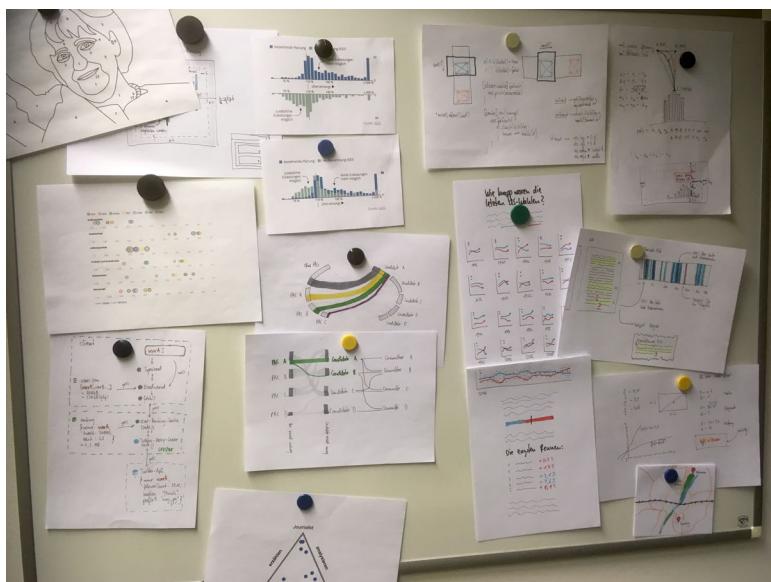
Die ersten beiden der Prinzipien weisen das Hervorheben einer Aussage, die Klarheit und die Akzentuierung dieser sowie die Leser*innen-Führung als Richtlinien für Visualisierungen aus. Diese kodifizierten »sayings« lassen sich als »doings« nicht nur an diversen Beispielen, die im *Styleguide* integriert sind, beobachten, sondern auch in laufenden Projekten während der Forschungsphase – beispielsweise als Susanne mit David über die Visualisierung der Namensstatistik (siehe Abbildung 12) spricht:

[Susanne] und David diskutieren die Namensstatistik-Grafik – David empfiehlt, eine Linie wegzulassen, die auf geht – diese sei inhaltlich nicht relevant, absorbiere aber viel Aufmerksamkeit.⁵

⁴ Zur Zeit des Schreibens dieser Ethnografie nennt sich das Team NZZ *Visuals*, dies wurde im Forschungsdesign (2.3.1) bereits ausgeführt. Was den *Styleguide* besonders interessant macht, ist, dass das Team während der Forschungsphase mit der Entwicklung begonnen hat. Der »Changelog« des *Styleguide* weist aus, dass er als Online-Plattform erstmals aufgesetzt wurde im Juni 2018, rund drei Monate nach der Forschungsphase.

Es werden einzelne Komponenten der Datenvisualisierung herausgehoben, indem andere weggelassen werden – in diesem Beispiel mit der Legitimation, dass sie für das Aufzeigen der produzierten Erkenntnis (weiterhin nimmt eine gleichbleibend grosse, überwiegende Mehrheit der Frauen bei einer Heirat den Namen des Ehemannes an) als nicht relevant beziehungsweise störend eingestuft wird. Als Resultat dieser Selektion erwartet man die Aufmerksamkeitslenkung auf diejenige Aussage, welche das Team als die *Geschichte* identifiziert hat. Für die Montage gilt es, durch Hervorhebungen eine Datenvisualisierung zu einer möglichst selbsterklärenden *Geschichte* zu machen – die Datenvisualisierung dient als Storytelling-Element, welches die Daten in eine Geschichte einordnet (Segel und Heer 2010).

Abbildung 15: Pinnwand in den Redaktionsräumen von BR Data. Bei den meisten der aufgehängten Visualisierungen handelt es sich um die Datenvisualisierungsentwürfe, welche den finalen Versionen in den Online-Artikeln sehr nahekommen.



So hängen in den Redaktionsräumen beider Teams jeweils Datenvisualisierungen von besonders erinnerungswürdigen *Geschichten* an den Wänden. Während es sich bei NZZ Storytelling um herausgerissene Seiten aus der Printausgabe handelt (Abbildung 16), sind es bei BR Data die finalen Papier-Entwürfe von wichtigen Visualisierungen aus mehreren Projekten (Abbildung 15). Als Darstellung der gewonnenen Erkenntnisse sind sie die Referenzpunkte zu den jeweiligen Projekten, denen meist lange Recherchen vorangingen.

Abbildung 16: Pinnwand in den Redaktionsräumen von NZZ Storytelling, an der eine Auswahl von Visualisierungen aus der Printausgabe hängt.



Lesart, Leseführung, Lesebeispiele

In der Fussballtransfers-Visualisierung (Abbildung 17) sollen Annotationen die Komplexität des Lesevorgangs vermindern:

[Kaspar] hat ausserdem die einzelnen Ströme eingefärbt und dann eine Farblegende gemacht. Der Platz für eine direkte Beschriftung sei knapp. David sagt, dass es den Leser jedes Mal aus der Grafik herausreisse. Er finde, es sollte sofort

sichtbar sein: »Bundesliga, 200 Millionen«. [Kaspar] meint: »Ach so, auch noch direkt die Zahlen...?« [bisher sind keine Zahlen in der Grafik], was David bejaht. [Kaspar] blickt dann nochmals auf die Grafik und spricht die Legenden unten und oben an, die er noch eingefügt hat: Aus welchen Ligen wurden nur Spieler ein-, in welchen nur Spieler verkauft. Das sei vielleicht nicht so wichtig und könnte man weglassen, meint [Kaspar]. David nickt.⁶

Vorgaben zu Leseführung und Lesart sind bei NZZ Storytelling in zwei Leitfäden festgehalten: Das Infografik-Subteam verfügt über eine kleine Palette an festgehaltenen Standards bezüglich Farben, Grafikformen und Platzbedarf auf einem A4-Blatt, das in den Redaktionsräumen hängt. Es fokussiert auf gute Lesbarkeit – vor allem für ein älteres Publikum, wie Christoph festhält: Dieses sei ja »60+, oder sogar eher 70+«.⁷ Es habe einige Leserbriefe gegeben, die sich über zu kleine Schrift in den Datenvisualisierungen beschwert hätten. Die Standards des Infografik-Subteams sind einer der wenigen Orte, wo Publikumsrückmeldungen direkt zu einer Anpassung der Praxis geführt haben.

Andererseits hält der *Styleguide* in seinen Prinzipien⁸ fest, mit welchen Methoden die bereits genannten Vorgaben (*determined, accentuated, precise, authentic*) erreicht werden sollen:

Introducing the reader to the data

The information we show through our graphics have a range of complexity. Our work is largely centered around breaking down complex themes or dense datasets, which means we need to find ways to display this information in an understandable format for our readers. Often we can count on basic chart types like bar and line charts to show our data but occasionally we use more complex visualizations that might require a little more introduction for our regular readers. No matter the complexity of the chart, we strive to create a clear introduction to each chart with a main message in the title. The following subtitle and legend provide a straight forward description of exactly what the chart is showing.

6 15_20180314_NZZStorytelling-03 (17)

7 18_20180319_NZZStorytelling-06 (4)

8 <https://nzzdev.github.io/Storytelling-Styleguide/#/principles> (18.04.2020)

Guiding the reader through the data

For more complex graphics or abstract topics it is sometimes helpful to provide the readers with an exact example of how they should read the information shown in the chart. When creating these snippets, take care to formulate them in a colloquial fashion. Read them out loud to see if they reflect the way you would explain the information to someone else. The reader should be able to use it as a template to interpret all other data points in the chart. Make sure to place the reading example in a prominent place, preferably to be read right after the introduction information (title, subtitle and legend).

Putting things into context

Annotations are a great tool for giving additional context or highlighting interesting bits of information that would otherwise not be apparent to the reader.

Communication over decoration

Use color, illustrations and icons to communicate and explain the data that you're displaying. When the illustration helps the reader get a better understanding of what they're looking at on the first look, then it's a good use of illustration. Illustrations and icons should add information and provide context to the data you're showing in the chart. The illustration should also not compete with the main message of the chart.

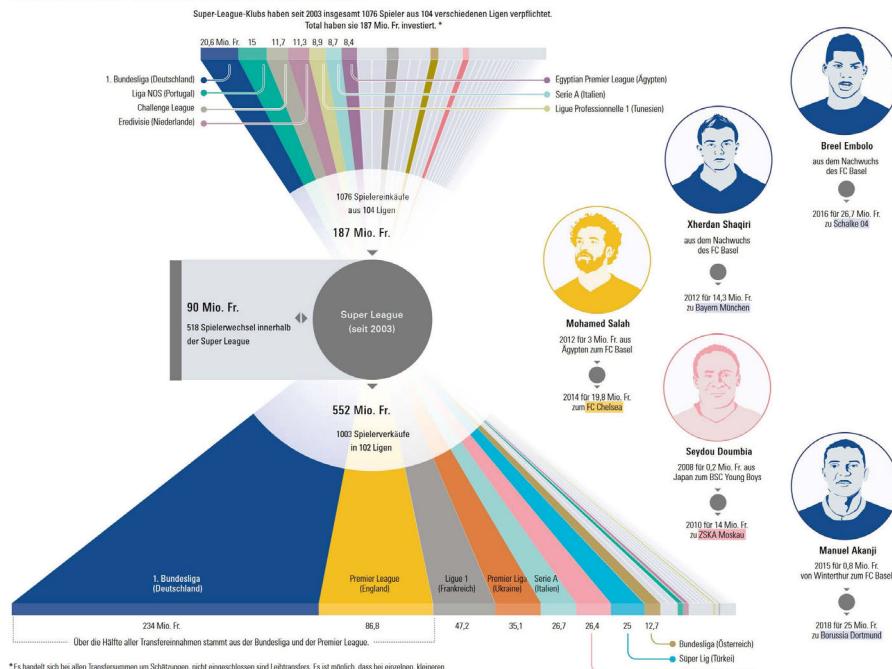
Der *Styleguide* schlägt also verschiedene flankierende Massnahmen vor, um die interpretative Flexibilität von Datenvisualisierungen zu reduzieren, indem es die Einengung auf eine Hauptaussage sicherstellt: Der Titel soll die »main message« bestimmen – eine Vorgabe, die überhaupt nur Datenvisualisierungen mit Hauptaussage zulässt. Nach der Einführung mittels Titel und Untertitel, welche für jede Datenvisualisierung eingefordert wird, muss die Leseführung sicherstellen, dass die Interpretation eines Falles auf alle Datenpunkte übertragbar ist – die möglichen Bedeutungsunterschiede zwischen abgebildeten Fällen werden dadurch eingebettet. Die Beschriftungen sollen auf weiteren Kontext aufmerksam machen, wo es der oder die Designer*in für angemessen hält. Der letzte Absatz verweist dann auf die Wichtigkeit des ersten Blicks: Diesem wird untergeordnet, wie weitere Gestaltungs- und Illustrationsmöglichkeiten zu erfolgen haben.

Lesebeispiele finden sich in diversen Datenvisualisierungen von NZZ Storytelling, beispielsweise in derjenigen mit den Wahlstatistiken (siehe Abbildung 19). Auch für die Fussballtransfers-Visualisierung erwägt Kaspar ein

solches, schliesslich übernehmen aber die Illustrationen der Einzelfälle diese Funktion (Abbildung 17).⁹

Abbildung 17: Visualisierung von Fussballtransfer-Daten sowie begleitende Illustrationen von Einzelfällen.

Wo Schweizer Klubs Talente finden und wer sie ihnen abkauft



NZZ vom 31. März 2018, Seite 54

Vergleich

Der explizite Vergleich ist eine weitere Methode zur Hervorhebung einer Hauptaussage in einer Datenvisualisierung. Bei BR Data sind die Methoden zur Hervorhebung weniger stark kodifiziert als bei der NZZ und werden situativ abgewogen. So lassen sich weniger Methoden der expliziten Leseführung

beobachten, ebenso gibt es keine Leitfäden wie den *Styleguide*. Stattdessen erwägt BR Data in einem Projekt das Montieren eines ausgewählten Vergleichs, um auf subtile Weise eine Hauptaussage einer Datenvisualisierung hervorzuheben:

[Lea] stellt dann noch ihre Überlegungen bezüglich Ausspielwegen vor, die sie gestern mit [Roman] besprochen hat. Sie fände es ein gutes Thema, um viele bunte Grafiken zu haben. Sie stellt sich vor, die gleichen Aspekte zu beleuchten wie im BBC-Bericht. [Martin] findet, die BBC-Idee mit dem Roman als Benchmark gut, er würde auch denselben Benchmark über alle Altersstufen hinweg nehmen. [Lea] meint, sie fände es toll, auch noch weitere Vergleiche einzubauen, beispielsweise mit der Verfassung. Und ihre dritte Idee sei, die Verständnis-Schwierigkeiten mit dem Bildungsniveau in Relation zu setzen, dies sei ja auch eine Berechnung.¹⁰

Der Vergleich dient hier nicht mehr dazu, die Erkenntnis zu produzieren (wie in Abschnitt 5.1.4 erläutert), dass die Dokumente lang und kompliziert geschrieben sind – dies wurde bereits ohne entsprechenden ›Benchmark‹ erkannt. Stattdessen soll der Vergleich mit einem Referenzwert, der einem breiten Publikum bekannt ist, die Hauptaussage *augenscheinlich* machen. So verwirft BR Data den ursprünglichen Plan, Max Frischs ›Homo Faber‹ als literarisches Referenzwerk einzusetzen; weil es viele kurze Sätze habe und daher nicht als Vergleich tauge, so Martin rückblickend.¹¹ Die Wahl des Vergleichswerks erfolgt so, dass damit die Hauptaussage der Geschichte durch den Vergleich so augenscheinlich wie möglich wird (Abbildung 18).

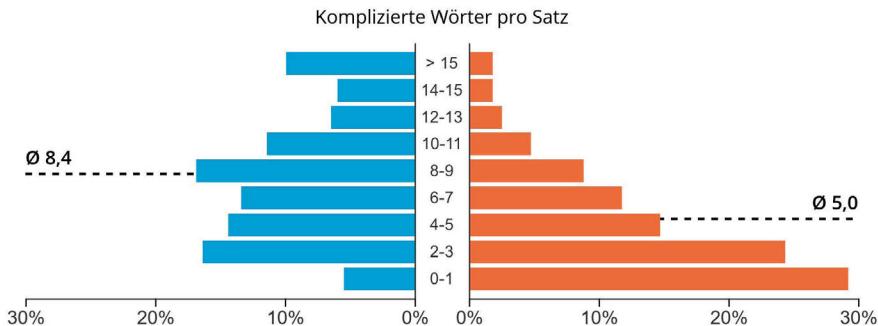
Im Modus der Visualisierung zur Evidenzbefestigung wendet BR Data Vergleiche an, um Abweichungen sichtbar zu machen und als Missstände zu werten – hier spiegelt sich die vom Team selbst proklamierte investigativ-journalistische Ausrichtung und die Orientierung am entsprechenden Narrativ wider. Dies gilt besonders für die oft angewandten Vergleiche zwischen politischen Entitäten wie Städten, Bundesländern oder Ländern.

10 49_20181212_BR-Data-03 (5)

11 65_20190710_BR-Data-14

Abbildung 18: Datenvisualisierung aus der Analyse von Datenschutzerklärungen durch BR Data.

Datenschutzerklärung und Der Tod in Venedig im Vergleich



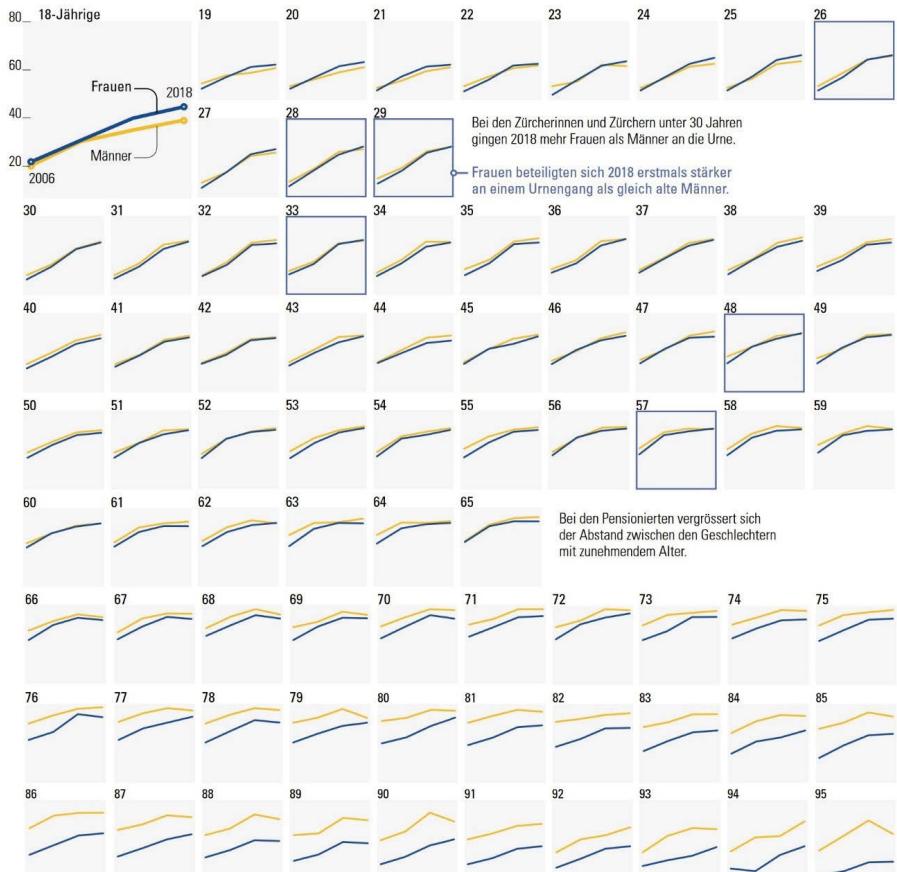
<https://web.br.de/interaktiv/datenschutzerklaerungen/> (20.04.2020)

Im Beispiel der Wahlbeteiligungsstatistik (siehe Abbildung 19) montiert NZZ Storytelling die Geschichte als Vergleich auf möglichst granularer Ebene. Klassischerweise hätte man für die Aussage, dass junge Frauen anteilmäßig zahlreicher an die Urne gehen, Altersgruppen gebildet. Dies schlägt die Statistikerin Annalena denn auch vor, aber Designerin Emma votiert für Small Multiples jeden Altersjahres.¹² Die beiden sowie Infografikerin Caroline und der zuständige Lokalredaktor besprechen dann die Finalisierung der Datenvisualisierung:

Der Redaktor, [Caroline] und [Annalena] gehen danach zu [Emma]s Pult, die gerade an den Small Multiples der Wahlstatistik arbeitet. [Emma] erklärt, sie habe einige der Grafiken hervorgehoben. [Annalena] ergänzt, das seien diese Grafiken, wegen »denen es ja diese Geschichte gibt«. Der Redaktor schaut sich die Grafik an und stellt einige Fragen zu den Details. Er macht darauf aufmerksam, dass es sehr wichtig ist, einen präzisen Titel zu setzen. Und macht auf die noch fehlende y-Achsen-Beschriftung aufmerksam. [Annalena] erklärt, diese werde man bei der ersten der vielen Small Multiples einsetzen, und danach sei es ja klar.¹³

Abbildung 19: Ausschnitt aus dem Artikel »No City for Old Men«.

Bei älteren Frauen ist die Beteiligung an Urnengängen viel tiefer als bei Männern. Junge Frauen aber haben aufgeholt.
Anteil der Stimmberechtigten, die bei den Wahlen 2006, 2010, 2014 und 2018 an die Urnen gingen, in %



NZZ vom 27. März 2018, Seite 21

Während der Redaktor es als wichtig erachtet, dass die Aussage schnell erfassbar ist mit einem Titel und Beschriftungen, so entscheiden sich Annalea und Emma für eine umfassende Vergleichsmöglichkeit für das Publikum durch 78 Small Multiples bei gleichzeitig möglichst starker Simplifizierung des einzelnen Liniendiagramms.

Die Geschichte aus dem Lokal-Ressort ist schliesslich die Gelegenheit, dass die Printausgabe der NZZ Small Multiples als Visualisierungen abdruckt – was von NZZ Storytelling offenbar schon länger herbeigesehnt wurde.¹⁴ Auf dem eigenen Twitter-Account feiert das Team die umfassende Vergleichsleistung der Daten, indem es den Artikel mit einem GIF der hintereinandergereihten 78 Kurven und dem Ausruf »Small mu[...]lultiples.« anpreist.¹⁵

So erfolgt also die Hervorhebung der Hauptaussage mittels der expliziten Kennzeichnung (Lesart, Leseführung, Lesebeispiele) und/oder einer expliziten Kontrastierung durch Vergleiche.

6.1.2 ›Humanizing data‹: Personalisierung und Konkretisierung

Als David anlässlich einer Besprechung das Projekt über die Fussballtransfers initialisiert, präsentiert er den Teammitgliedern Kaspar und Caroline einen Vorschlag und legt dafür eine Papierskizze auf den Tisch (siehe auch erweiterten Feldausschnitt, Fussnote 93 in Kapitel 5):

Es ist eine »Sankey«-Grafik. Es sei ja auch spannend, woher die Spieler kämen [meint David]. Ausserdem schlage er vor, dann einzelne Spieler herauszupicken, um »es« greifbar zu machen. David präsentiert auf einem Excel-Sheet direkt auf seinem Laptop gleich eine Auswahl von fünf Spielern, die teuersten Transfers aus der Schweiz.¹⁶

Es ist eine von vielen Situationen in beiden Teams, in der es darum geht, »es« greifbar zu machen oder eine *griffige Grafik* zu konstruieren. ›Greifbarkeit‹ verweist eigentlich darauf, dass man etwas anfassen, berühren kann – und es damit direkt wahrnehmen, als wenn man es ›bloss‹ liest. In dieser Bewertung

¹⁴ So fragt David explizit nochmals nach, ob denn nun die Small Multiples in die Zeitung kämen (23_20180326_NZZStorytelling-11 (7)).

¹⁵ <https://twitter.com/nzzvisuals/status/978256285291483139> (26.02.2020)

¹⁶ 14_20180313_NZZStorytelling-02 (3)

schwingt auch die Zuschreibung mit, dass damit den Daten ein höherer Evidenz-Status zukommt. *Griffig* und *greifbar* sind deshalb Eigenschaften, welche eine »gute« Datenvisualisierung ausmachen und denen die Teammitglieder zuschreiben, Evidenz zu produzieren. Die abgebildete Situation mit den Fussballspielern (tatsächlich zieren schliesslich fünf illustrierte Fussballer-Gesichter die finale Version der Datenvisualisierung, siehe Abbildung 17) zeichnet die Praktik nach: Es werden Einzelfälle »herausgepickt«, oftmals Personen, um den Sinnzusammenhang zu konkretisieren – es gilt, »herausstechende Fälle«¹⁷ zu erzählen. Die erste Version des »Data Journalism Handbook« bezeichnet diese Praktik als »Making Data Human« und begründet: »Data are tethered to real lives of real people, and when we engage with the numbers, we must consider the real-world systems from which they came.« (Gray et al. 2012:181f) – oder konzis ausgedrückt: Datenvisualisierungen sollen dadurch »approachable«, »accessible« und »relatable« sein (Alamalhodaei et al. 2020). Die Praktik schliesst zudem an die Plausibilisierungsstrategien (Abschnitt 5.1.4) an – die sich aber an die Auswertenden selbst richten, während es im zweiten Modus darum geht, Evidenz gegenüber Aussenstehenden zu befestigen.

So erzählt BR Data die über die Schufa¹⁸ gewonnenen Erkenntnisse nicht nur anhand eines fiktiven Falles, welcher vom Algorithmus diskriminiert wird (und schneidet die Lesart der Datenvisualisierungen auf den Fall zu [jung, männlich]¹⁹), sondern der *Longread* wird ergänzt mit Illustrationen von (fiktiven) jüngeren Männern. Das im Jahr vor der Forschungsphase von BR Data veröffentlichte Projekt »Hanna und Ismail« war sogar, wie der Titel bereits deutlich macht, grundsätzlich rund um zwei (ebenfalls fiktive) Personen aufgebaut, um strukturelle Verzerrungen im Mietmarkt aufzuzeigen.²⁰ Auch für das während der Forschungsphase initiierte Projekt zur Untersuchung der Staatsanwaltschaften taucht bereits in einem sehr frühen Stadium während eines Teammeetings die Frage nach der Personalisierung auf – man müsse bereits während der Datenanalyse nach »konkreten Fällen« Ausschau halten, empfiehlt ein Teammitglied.²¹

17 BR_Fragebogen_Mitarbeitende_07

18 Die Schufa ist ein Unternehmen, welches die Kreditwürdigkeit deutscher Personen und Unternehmen bewertet. Dafür hat sie unter anderem einen Score entwickelt, der nicht einsehbar ist. <https://www.schufa.de/> (31.01.2021)

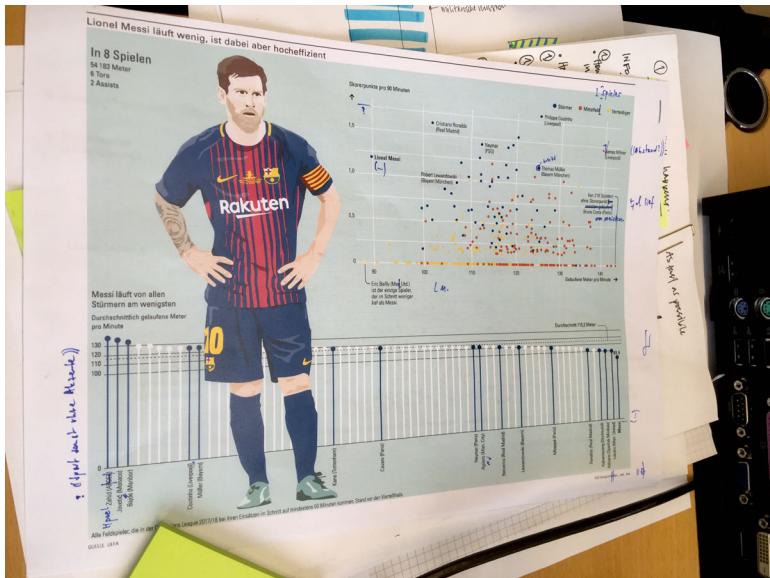
19 <https://web.br.de/interaktiv/erhoehtes-risiko/> (21.04.2020)

20 <https://www.hanna-und-ismail.de/> (21.04.2020)

21 50_20181213_BR-Data-04 (3) – für BR Data hat dieser Aspekt ein besonderes Gewicht, weil die Verbreitung der Geschichten über den Fernsehkanal nach bewegten Bildern

Eine besondere Attraktivität üben Extremwerte aus – dies hat sich bereits bei der Erkenntnisproduktion gezeigt. Prominente Einzelfälle bieten sich nicht nur als Plausibilisierung, sondern auch als Illustration an. In einem kleineren Projekt mit einem Sportredaktor untersucht David die Champions-League-Fussballer hinsichtlich ihres Verhältnisses von Laufdistanz und Scoring-Punkten, mit einem besonderen Fokus auf die Effizienz von Lionel Messi (Abbildung 20).

Abbildung 20: Foto des ausgedruckten Visualisierungsentwurfs zum Projekt Lauf-/Scoring-Analyse. Illustriert herausgehoben ist Lionel Messi; in den beiden Datenvisualisierungen plant David die Extremwerte (darunter Messi) sowie besonders prominente Fussballer farblich abzuheben.



verlangt, wofür insbesondere die Leitung von BR Data die Konkretisierung mit Fällen für nötig hält (52_20181217_BR-Data-06 (7)).

Konkretisierungen von Einzelfällen gehen über Personalisierungen hinaus. So baut BR Data die Geschichte über die Komplexität von Datenschutzerklärungen rund um das Beispiel der WhatsApp-Datenschutzerklärung auf. Mittels eines *Scrollytelling*-Formats – die Lesenden scrollen nach unten, wodurch nacheinander erklärende Annotationen im Vordergrund vorüberziehen, während sich im Hintergrund die Hervorhebung ändert – erklärt der *Longread* die verschiedenen Aspekte, welche das Lesen von Datenschutzerklärungen besonders kompliziert machen (Abbildung 21).

Falls nicht direkt in der Datenvisualisierung, so findet sich in praktisch allen Geschichten von NZZ Storytelling und BR Data die Praktik des ›humanizing data‹ durch Personalisierungen und Geschichten von Einzelfällen im erläuternden Text.

Abbildung 21: Ausschnitt aus dem BR-Data-Projekt zu den Datenschutzerklärungen: Illustration anhand des Beispiels WhatsApp – der weisse Kasten bewegt sich durch Scrolling nach oben.

auch der **Verantwortliche** bezüglich deiner **Informationen** ist, wenn du **unsere** Dienste nutzt.]
wir gehören zu den Facebook-Unternehmen.] Mithilfe unserer Datenschutzrichtlinie („**Datenschutzrichtlinie**“) erläutern wir unsere Praktiken in Bezug auf Informationen (einschließlich Nachrichten), einschließlich der Informationen, die wir zur Unterstützung unserer Dienste verarbeiten. So erklären wir beispielsweise, welche **Informationen** wir sammeln

➔ Kaum zu verstehen

Der hervorgehobene Satz ist nicht nur lang, sondern hat besonders viele komplizierte Wörter: nämlich 15.

zierst. |

<https://web.br.de/interaktiv/datenschutzerklaerungen/> (20.04.2020)

Bereits Abschnitt 5.1.4 hat gezeigt, wie auf der Basis von Datenvisualisierungen quantitative und qualitative Vorgehensweisen für die Erkenntnisproduktion koordiniert werden können. Im zweiten Modus machen Datenvisualisierungen primär quantitative Ergebnisse evident, durch Personalisierung und Konkretisierung werden aber bereits auf den Visualisierungen selbst Anknüpfungspunkte für eine Beleuchtung von Einzelfällen begleitend im Text montiert – Messi und die Fussballer-Köpfe illustrieren dies.

6.1.3 Datenquellen und Methoden transparent machen

Transparenz kann viele verschiedene Aspekte datenjournalistischer Praktiken betreffen, der Abschnitt 3.2.1 hat dies im Zusammenhang mit der *Open-Kultur* erläutert. Der Transparenz-Begriff ist zudem anschlussfähig an die journalistische Kultur – durch das Ideal, so viel als möglich über Datenquellen und Methoden offenzulegen (Kennedy et al. 2020:170; Allen 2008), also Faktoren aufzudecken, welche die Nachrichtenproduktion beeinflussen (Zamith 2019:5). Auch in der Praxis schwingt dieses Ideal ständig mit, ebenfalls mit einem Fokus auf Datenquellen und Methoden – wie der folgende Ausschnitt zeigt, in dem Teammitglieder von BR Data einen Artikel über Kriminalität von Migrant*innen eines anderen Mediums betrachten und kritisieren:

Die Grafik mit den gestapelten Balkendiagrammen zu den Delikten und Nationalitäten wird kritisiert, da sie sehr kompliziert und relativ schwer verstehbar sei – das sei halt das Problem mit den gestapelten Balkendiagrammen, meinen [Lukas] und [Moritz], dass man nicht genau die Proportionen abschätzen könne, und »was ist 1?«. Geschrieben sei die Geschichte aber gut, meint [Lea], sie habe sie schon online gelesen. Intransparent sei aber, wie genau die abgebildeten Nationalitäten und Delikte ausgewählt wurden.²²

Die Diskussion verweist einerseits auf den Anspruch auf einen schnellen und eindeutigen Sinngebungsprozess aus der Datenvisualisierung, der für die Teammitglieder die statistische Nachvollziehbarkeit bedingt (»was ist 1?«), und andererseits die Transparenz über die Methodik der Datengrundlage. Die hohen Ansprüche an das Ideal der Transparenz finden sich in den Praktiken der beiden Teams situativ umgesetzt. Die Q-Grafiken von NZZ Storytelling geben meist die Quelle an, woher der Datensatz stammt, stellen ihn teilweise

bei grossen Projekten auch zur Verfügung (siehe Abschnitt 6.3). Das Tool Q selbst ist Open-Source-Software²³, man kann es sogar in einer Demoversion selbst nutzen.²⁴ In vielen anderen Beispielen bleiben aber die Datensätze unzugänglich und die methodischen Entscheidungen nicht öffentlich reflektiert.

Die situativ gehandhabte Transparenzpraktik deckt sich mit inhaltsanalytischen Untersuchungen von datenjournalistischen Projekten, welche die vorher aufgezeigten Idealvorstellungen kontrastieren: Loosen et al. (2017) finden bei preisgekrönten datenjournalistischen Projekten nur in einer Minderheit Angaben darüber, wie auf die Datenquellen zugegriffen wurde. Dies deckt sich mit Stalphs (2017:16) Resultaten, der nur in rund der Hälfte von datenjournalistischen Artikeln direkt Daten fand – auch Zamith (2019) stellte eine niedrige Quote an direktem Zur-Verfügung-Stellen von Daten bei grossen amerikanischen Medienorganisationen fest. Breda (2019) schaute sich nur die Datenvizualisierungen selber an und konstatiert, dass Angaben über die Datenerhebungsart und Einschätzungen der Datenqualität nur bei einer Minderheit zu finden sind.

Bei den grossen Recherche-Projekten von BR Data machen die Teammitglieder die Daten und die Methodik – bis hin zum Code – so weit als möglich transparent, wägen dies aber mit dem Schutz von Datenquellen und der Personen dahinter ab, einem altbekannten Dilemma aus dem investigativen Journalismus.²⁵ Die Skepsis bezüglich der Erklärbarkeit und Nachvollziehbarkeit der Datenanalyse-Methoden im Projekt zu den Polizeimeldungen dominiert die vielen Diskussionen – so beispielsweise in einer Situation, in der Martin und Lea eine neue Methode ausprobieren:

[Martin] blickt zu [Lea] und meint, er finde das gerade auch eine spannende Idee – aber habe jetzt schon Angst, wenn sie erklären müssten, wie sie das gemacht hätten.²⁶

-
- 23 David hat darüber gebloggt: <https://medium.com/nzz-open/lets-build-together-nzzs-storytelling-toolbox-q-is-now-open-source-39decb92aeca> (25.04.2020)
- 24 <https://editor.q.tools/> (25.04.2020)
- 25 Die öffentlich gemachte Dokumentation zu einer Recherche zu Pestizid-Zulassungen beispielsweise ist sehr umfangreich: <https://github.com/br-data/2018-pestizidzulassungen-analyse> (25.04.2020). Bei der Analyse des Schufa-Score-Algorithmus hingegen wurde auf die Offenlegung von Quellen und detaillierter Methodik verzichtet, weil es sich um sensible, personenbezogene Daten handelt: <https://web.br.de/interaktiv/erhoeftes-risiko/> (25.04.2020)

Martin antizipiert, dass das methodische Vorgehen in der Präsentation einer Rechtfertigungspflicht untersteht und nachvollziehbar sein muss – und zweifelt, ob die angedachte Methodik diesen Kriterien gerecht werden kann.

Das nachgezeichnete ambivalente Bild der Transparenzpraktiken bezüglich Datenquellen und Methoden deckt aber längst nicht alle Ebenen ab. So thematisieren Kennedy et al. (2020:174ff) als zusätzliche Ebene das Offenlegen der Konstruktionsprozesse und dessen Unsicherheiten in der Datenvisualisierung selbst. Die Autor*innen konstatieren, dass diese Transparenzpraxis fehle: Unsicherheiten würden nicht visuell kommuniziert. Dies deckt sich mit den Beobachtungen aus den zwei untersuchten Teams: Ihre Datenvisualisierungen sind stets eindeutig, Unsicherheiten und Konstruktionsprozesse visuell nicht erkennbar, ausser der temporale Bezug betrifft die Zukunft (Prognosen²⁷). Die Bewertungspraktiken und alternative visuelle Verkörperungen werden nicht ausgewiesen, der fragile Konstruktionsprozess der Referenzen auf die *echte Welt*, welche Abschnitt 5.1.4 beleuchtet hat, bleibt verborgen. In den beiden Teams erfolgt also eine Grenzziehung zwischen Transparenz über Datenquellen und Methoden – diesbezüglich folgt man situativ abgewogen dem journalistischen Ideal – und dem visuellen Ausweisen von Konstruktionsprozessen und Unsicherheiten – diese Praxis wird kaum angewendet. Eine mögliche Erklärung liefern Kosminsky et al. (2019) in ihrem Aufsatz »Belief at first sight«: Datenvisualisierungen rationalisieren das Erkennen im Sehprozess und verwickeln es dabei mit der Idee der Objektivität. So mag der Verzicht auf die visuelle Kennzeichnung von Konstruktionsprozessen darin begründet sein, dass man sonst befürchtet, der Datenvisualisierung würde nicht mehr die erhoffte Objektivität zugeschrieben. Transparenz im Produktionsprozess bedeutet potenziell eine Unterminierung für die Autorität zahlenbasierter Kommunikation (Heintz 2010:171f).

Schreibt man Transparenzpraktiken eine doppelte Funktion zu – verbesserte Zugänglichkeit für Akteur*innen aus dem eigenen Feld (*peers*) sowie bessere Legitimation der eigenen Schlüsse gegenüber dem Publikum, siehe

26 57_20190109_BR-Data-11 (6)

27 Zu Prognosen hat NZZ Storytelling zweimal methodische Überlegungen transparent gemacht – bezüglich Wahlen (<https://www.nzz.ch/international/deutschland-im-wahljahr/methodik-statistik-wie-die-nzz-umfragen-zur-bundestagswahl-darstellt-und-warum-ld.148951, 25.04.2020>) wie auch Wetterphänomenen (<https://www.nzz.ch/p-anorama/warum-die-gaengigste-hurrikan-visualisierung-problematisch-ist-und-was-die-nzz-an-irher-stelle-macht-ld.1394033, 25.04.2020>).

die Narrative des datenjournalistischen Feldes (3.3.1) –, so pflegen die beiden Teams einen unterschiedlichen Umgang mit den jeweiligen Typen: Intransparenz über die Visualisierungskonstruktion gegenüber dem Publikum sowie teilweise Transparenz über Daten und Methoden, ein Angebot, das sich vor allem an die *peers* richtet. Einerseits sichern sie durch die semiologische Eindeutigkeit die Produktion von Evidenz mit hohen Objektivitätsanspruch, andererseits sichert das Transparenzangebot über Datensätze und Analysemethoden die Möglichkeit, dass die eigenen Schlüsse von Sachverständigen aus der eigenen Community durch eine Replikation der Analyse geprüft werden könnten (Peer-Review) – was wiederum ein Vertrauen signalisiert, dass die eigenen Analysen einer solchen Prüfung standhalten würden, und somit die Evidenz der Schlüsse untermauert. Es ist gar nicht relevant, dass das Replikationsangebot äusserst selten wahrgenommen wird:

Ich frage [Moritz], ob mal schon jemand eine BR-Data-Analyse repliziert habe – [Moritz] meint nein.²⁸

6.1.4 Ausdifferenzierte Expertisen ausweisen

Die epistemologische Idee von Datenvisualisierungen *als* Geschichten spiegelt sich in der (veränderten) Verwendungspraktik der *Byline* (Verfasser*innenzeile – in den Gesprächen wird fast ausschliesslich der englische Begriff genutzt). Es ist insbesondere David, welcher eine Änderung der Praktik anstrebt: Bei grösseren Datenvisualisierungen wie auch Illustrationen sollen die Infografiker*innen und Designer*innen des Teams Teil der *Byline* werden (bisher ›unterzeichneten‹ sie mit ihrem Kürzel die Grafiken). So beispielsweise beim Fussballtransfers-Projekt:

Er [David] sagt zu [Kaspar], dass er schauen müsse, dass er in die Byline komme – schliesslich sei die Grafik die »halbe Story«. [Kaspar] sagt dann, [Caroline], David selber und [Yannick] hätten ja auch mitgearbeitet. David zuckt mit den Schultern und meint, es sei jetzt mal wichtig, dass er [Kaspar] aufgeführt sei. Für alles Weitere müsse er jetzt zuerst mal das neue »Byline-Konzept« aufstellen.²⁹

28 49_20181212_BR-Data-03 (9)

29 18_20180319_NZZStorytelling-06 (22)

Kaspar bewertet die Arbeit von Caroline (Illustrationen), David (Mitarbeit Datenanalyse) sowie Yannick (Daten-Scraping) als Beiträge zur Geschichte, die für das Publikum in der *Byline* erkennbar sein und deshalb ausgewiesen werden sollen. Einige Monate später findet sich im *Styleguide* das von David angekündigte »Konzept«:

Bylines, both in print and online, show our audience who contributed substantially to the story. Traditionally, bylines were reserved only for those who wrote text. The graphic and methodic work that our team creates is equally as significant to the final text as it carries the visual weight of the main message.

Whenever we contribute work essential to a story, it is accepted practice to reflect that in the byline. In that case, we remove the initials below the graphic. **When multiple team members worked on a graphic, we use »NZZ Visuals« in the byline or initials.**³⁰

Der Ausschnitt verweist auf die traditionell grosse Bedeutung der *Byline* für Text-Autor*innen, hebt nun aber mit dem Verweis auf das »visual weight of the main message« hervor, dass die grafische und technische Arbeit des Storytelling-Teams »equally as significant« sei. Folgerichtig beansprucht das Team einen Teil der *Byline*, was eine Emanzipation aus der Service-Rolle und eine Infragestellung der traditionellen Hierarchien in der Redaktion bedeutet, wie dies auch Hermida und Young (2017:167f) bereits in anderen datenjournalistischen Teams festgestellt haben.

Die *Byline* des Fussballtransfers-Artikel wird Kaspar schliesslich an zweiter Stelle nennen, nach dem Hauptautor aus der Sportredaktion. Carolines Kürzel zierte sowohl in der Print- wie in der Online-Version unterhalb der Illustrationen. Die Online-Version erwähnt zudem David in der *Byline* und ganz am Ende des Textes Yannicks Mitarbeit. Für den Print-Bereich sind die Hürden also offensichtlich höher, als Mitautor*in auftreten zu können.

Am Ende desselben Jahres folgt ein symbolischer Schritt: Fast die Hälfte derjenigen, denen Chefredaktor Eric Gujer das sogenannte Zeichnungsrecht³¹

³⁰ <https://nzzdev.github.io/Storytelling-Styleguide/#/charts-guidelines?a=bylines> (20.04.2020)

³¹ Ursprünglich durfte nur, wer das Zeichnungsrecht hatte, seinen Namen unter einen Artikel in der NZZ setzen. Während der Forschungsphase wurde dies in der Redaktion aber bereits nicht mehr strikt gehandhabt und generell schien dem Zeichnungsrecht keine grosse funktionale Wirkung mehr zuzukommen – mindestens nicht im Team

für die NZZ verleiht, sind Teammitglieder von NZZ Storytelling³²: der Teamleiter, drei Infografiker*innen, ein Designer sowie ein Programmierer – was David als »wichtiges Signal« bezeichnet, weil das Zeichnungsrecht vorher nur an Journalist*innen verliehen wurde.³³

In dieser Arbeit stehen aber nicht entsprechende organisationssoziologische Dynamiken zwischen verschiedenen Gruppierungen innerhalb der Redaktion im Zentrum, sondern die Praktiken der Teammitglieder, das Konzept der Autor*innenschaft breiter zu fassen – wovon der Anspruch, in der *Byline* vertreten zu sein und das Zeichnungsrecht zu erhalten, nur eine Ausprägung ist. Eine andere ist der Verweis auf die Konstrukteur*innen der Werkzeuge, mithilfe derer Datenanalysen und -visualisierungen vorgenommen wurden – zum Beispiel im *Explainer* zu Schweizer Waffenexporten in kriegsführende Länder.³⁴ Bereits die *Byline* nennt sieben Personen (wovon zwei von NZZ Storytelling sind), was für das *Explainer*-Format, ein längerfristig aktuelles und in der Regel mehrfach aktualisiertes *Erklärstück*, nicht erstaunt. Schliesslich wird zum Ende des *Explainers* die Mitarbeit von David sowie von drei Programmierern unter dem Vermerk »Tools« erwähnt. Die Autor*innenschaft für die produzierten journalistischen Erkenntnisse wird also nicht nur für die verschiedenen journalistischen Expertisen reklamiert, sondern auch für diejenigen aus dem Bereich Design und Programmierung ausgewiesen. Die Grenze der berücksichtigten Expertisen wird rund um das eigene Team gezogen (auf die verwendete Analysesoftware R verweisen erst die Methodik-Ausführungen auf Github³⁵).

Bei BR Data stellt sich die Frage um die Autor*innenschaft weniger oft, weil die datenjournalistischen Publikationen nicht in einem täglichen Rhythmus erfolgen. Es widerspiegelt sich eine ähnliche Praxis wie bei NZZ Storytelling, wenn sie auch nicht derart formell geregelt und explizit thematisiert wird. Publikationen von BR Data, die im eigenen Webseiten-Template erscheinen, warten mit einer ausdifferenzierten Aufführung der beteiligten Team-

von NZZ Storytelling. Trotzdem scheint der symbolische Wert innerhalb der Redaktion immer noch von gewisser Bedeutung für die informelle Hierarchie.

32 NZZ vom 29. Dezember 2018, Seite 11.

33 <https://twitter.com/davidbauer/status/1078989280456962050> (20.04.2020)

34 <https://www.nzz.ch/schweiz/waffenexporte-schweiz-liefert-vermehrt-in-laender-mit-konflikten-ld.1422907> (19.02.2019)

35 <https://github.com/nzzdev/st-methods/tree/master/1834-waffenexporte-konfliktarten> (20.04.2020)

mitglieder auf. Der Schluss des *Longreads* über die Schufa-Algorithmen³⁶ differenziert zwischen Autorschaft (2), Mitarbeit (7), Redaktion (3), Illustration (1) und Grafiken (2) aus – insgesamt 13 Personen (zwei Personen werden in je zwei Kategorien genannt). Andere Projekte nennen als weitere Kategorien der Mitarbeit Programmierung, Beratung, Infografiken oder Digitales Design. Das Team weist also für ein datenjournalistisches Projekt diverse Expertisen bezüglich Fachwissen, Programmierung wie auch Design als signifikante Beiträge zur Geschichte aus und kategorisiert sie situativ mit verschiedenen Funktionen. Erscheint die Berichterstattung von BR Data jedoch im gängigen Webseiten-Template des BR, so sind die Autor*innen (maximal fünf) in der rechten Spalte aufgeführt und ihre Funktion nicht weiter ausdifferenziert.

Abschnitt 6.4 wird die Nutzung von eigenen Templates durch BR Data näher beleuchten – es kann aber jetzt schon festgehalten werden: Verfügt das Team selbst über die Kompetenz, zu benennen, wer hinter einer Geschichte steckt, so strebt es an, eine Komplexität und Diversität der Expertisen aufzuzeigen. Loosen et al. (2017:8) setzen die für preisgekrönte Datenjournalismus-Projekte festgestellte hohe Durchschnittszahl an Autor*innen und Mitarbeitenden (zirka fünf) in Verbindung mit den vielen Kollaborationen innerhalb und ausserhalb der eigenen Redaktion. Dies ist ein relevanter Faktor auch in den Praktiken von BR Data und NZZ Storytelling. Die Beobachtungen in beiden Teams zeigen aber auch, dass die Gründe über die organisatorischen hinausgehen. Die Praktik der Ausweisung möglichst vieler Personen und ihrer Expertisen, die in eine Geschichte einfließen, kontrastiert mit früheren Vorstellungen von *Bylines*: Die Nennung einer*^s Verfasser*^s suggerierte eine gewisse ›Kontamination‹ der Objektivität durch die menschliche Subjektivität – und einige Medienorganisationen verzichteten früher deshalb komplett auf *Bylines* (Ward 2018:71). Die neuartige Auszeichnung von Expertisen und Technologien verweist darauf, dass die Teams reflexiv ihren angewandten Technologien eine gewisse Agentschaft in der Erkenntnisproduktion zuschreiben – dies wurde bisher vor allem mit Bezug auf den Einsatz von Algorithmen und Bots zur Nachrichtenproduktion diskutiert (Lewis et al. 2019; Diakopoulos und Kolika 2016). Mit der direkten Nebeneinanderstellung von journalistischen, grafischen und technischen Expertisen in den *Bylines* verwirklicht sich aber vor allem der Evidenzanspruch, der den datenjournalistischen Praktiken der bei-

³⁶ <https://web.br.de/interaktiv/erhoehtes-risiko/> (20.04.2020) – diese Geschichte veröffentlichte BR Data nicht während, sondern kurz vor der Forschungsphase.

den Teams inhärent ist: Die mit ihren *Tools* produzierten Datenvisualisierungen sind genauso *Geschichten* wie textbasierte Abschnitte und Artikel.

Die vier ausgeführten Arbeitspraktiken der Evidenzbefestigung sind – genauso wie diejenigen der Erkenntnisproduktion – ein Resultat der Aushandlung zwischen den Expertisen sowie deren Kodifizierung und Formalisierung. Darum geht es in den folgenden Abschnitten.

6.2 Evidenzbefestigung als Aushandlung

Abschnitt 5.2 hat jene Aushandlungssituationen nachgezeichnet, in denen die programmiertechnische, die journalistische und die grafische Expertise koordiniert werden, um Erkenntnisse zu produzieren. Zur Rekapitulation: Dabei treffen verschiedene Materialitäten, epistemologische und professionelle Kulturen aufeinander, welche unterschiedliche Entitäten als Ausgangspunkte für datenjournalistische Praktiken nutzen:

- Programmiertechnische Expertise: Was ist datafizierbar und, in Datenbanken gespeichert, für Verrechnungen anschlussfähig?
- Journalistische Expertise: Welche Relationen sind interessant als Geschichte zu erzählen?
- Grafische Expertise: Welche geometrischen Formen können auf welchem Ordnungssystem Relationen zwischen Variablen sinnhaft verkörpern?

Dieser Abschnitt untersucht Aushandlungssituationen, in denen die Expertisen koordiniert werden, um die Evidenz dieser Erkenntnisse zu befestigen. Wiederum werden die Aushandlungen in den Teams als »trading zones« (Garrison 1997) analytisch gefasst, innerhalb derer sich die verschiedenen Expertisen verständigen und ihre unterschiedlichen materiellen, epistemologischen und professionellen Perspektiven koordinieren. Abschnitt 4.3 hat die Heuristik eingeführt, mit welcher im folgenden Abschnitt der Blick auf die Koordinationsmuster in den Arbeitspraktiken erfolgt.

6.2.1 Koordinationsmuster in den Arbeitspraktiken

Die grafischen Anpassungen sollen eine Datenvisualisierung übereinstimmend machen mit »the story«, wie es Datenjournalist Yannick in einer Diskussion mit Designerin Emma ausdrückt. Die beiden diskutieren darüber,

welche Assoziationen und Sinnzusammenhänge sie durch die Farbgebung betont haben wollen und wie diese nach ästhetischen Kriterien zu beurteilen sind.³⁷ Und Designer Kaspar legitimiert die eigene Gestaltungsvariante der Fussballtransfers-Visualisierung (Abbildung 17) damit, dass sie die Hauptaussage unterstreicht:

[Rudolf] schlägt vor, man könnte die Grafik noch besser machen, indem man die Ströme anders anordnen sollte, sodass die Farben besser gruppiert sind. [Kaspar] beruft sich auf die »Hauptaussage«, dass es halt um die grössten Umsatzländer insgesamt geht.³⁸

Bezüglich der ersten Arbeitspraktik (Abschnitt 6.1.1) lässt sich festhalten: Grafische Expertise macht NZZ Storytelling für die Evidenzbefestigung nutzbar mit dem Ziel, dass das Publikum eine Datenvisualisierung und ihre Hauptaussage möglichst schnell und ohne interpretativen Streuverlust versteht und zielsicher in eine Geschichte einbettet. Dabei gilt es, Abwägungen vorzunehmen, wie eine kurze Diskussion über die Small Multiples der Wahlstatistik (siehe Abbildung 19) und das daraus erstellte GIF (siehe auch Fussnote 15, Kapitel 6) zeigen:

[Annalena] und [Emma] diskutieren danach nochmals die produzierten Small Multiples. Die vom Redaktor gewünschten Beschriftungen »seien visuell schon viel«, aber ja, sie ([Emma]) habe das jetzt so angepasst. [Annalena] meint, es sei gut zu wissen, was ihre visuellen Überlegungen seien. Sie diskutieren das GIF für die mobile Version hinsichtlich der Legenden: Sind diese noch nötig, werden diese überhaupt wahrgenommen? Wird das GIF noch an anderen Orten verwendet? Braucht es einen Titel? Sie gehen gemeinsam zu David, um dies mit ihm zu diskutieren. [Emma] plädiert dafür, aus ästhetischen Gründen möglichst wenig Beschriftungen vorzunehmen.³⁹

Legenden grenzen die Lesart ein – und gleichzeitig konfigurieren sie in dieser Situation mit dem Mittel des Vergleichs, den die Small Multiples implizieren, wofür Emma allzu zahlreiche Beschriftungen als irritierend und von der Hauptaussage ablenkend betrachtet. Die GIF-Datenvisualisierung strebt als

37 17_20180316_NZZStorytelling-05 (9)

38 16_20180315_NZZStorytelling-04 (18)

39 23_20180326_NZZStorytelling-11 (3)

Montage an, die Hauptaussage durch wenig zusätzliche Informationen möglichst schnell und überzeugend *augenscheinlich* zu machen.

Die journalistische Kultur präsentiert Erkenntnisse über die Welt als Geschichte mit hervorgehobener Hauptaussage. Diese Reduktion der interpretativen Flexibilität, gleichbedeutend mit einer Reduktion der Komplexität, verweist auf das Ziel, Erkenntnisse möglichst eindeutig und zielsicher zu vermitteln. Carlson (2015:9) macht darauf aufmerksam, dass jede Profession, die Erkenntnisse produziert, vor der Herausforderung steht, diese so zu übersetzen, dass sie für das Publikum verständlich werden. Spezifisch für den Journalismus gelte, dass er sich an ein sehr breites Publikum richte, das nicht aus dem eigenen professionellen Feld stamme. Es entsteht ein Spannungsfeld zwischen einerseits dem Anspruch des Journalismus, von einem breiten Publikum verstanden zu werden und es rund um die Geschehnisse der Welt zu versammeln, und andererseits kaum etwas über die Rezeptionsprozesse zu wissen (Carlson 2018:1759).

Die offene Frage der Rezeption akzentuiert sich im Hinblick auf Datenvisualisierungen, da die diesbezügliche Lesekompetenz weit weniger instituionalisiert ist als diejenige von Texten. Im Montage-Modus treffen die Teammitglieder Annahmen über ihr Publikum, die in der Regel nicht oder nur anekdotisch überprüft sind.⁴⁰ Zwar befassen sich mit der Publizistik und der Rezeptionsforschung wissenschaftliche Disziplinen mit Fragen nach der Rezeption von journalistischen Produkten, diese sind aber selten auf ein einzelnes Medium oder gar einzelne Artikel heruntergebrochen. Die Rezeptionsanalyse von technisch ausgefeilten Datenvisualisierungen steckt erst in den Kinderschuhen (Wihbey 2019:168). Ob und wie viel die NZZ oder der BR innerhalb ihrer eigenen Organisation die Rezeption ihrer Geschichten erforschen, bleibt unbekannt – zu konstatieren gilt es, dass, wenn solches Wissen vorhanden gewesen wäre, es in den datenjournalistischen Teams nicht verfügbar war. So erzählt David, dass man wenig konkrete Ahnung davon habe, wie die Datenvisualisierungen verstanden würden, man müsste den Umgang damit im Publikum eigentlich viel genauer erforschen und das eigene Bild des Lesers hinterfragen.⁴¹ Die Rückmeldungen auf einzelne Geschichten und ihre Datenvisualisierungen

40 Trotz der unbekannten Rezeptionsprozesse ist insbesondere bei NZZ Storytelling die Annahme verbreitet, dass die Menschen grundsätzlich »visuelle Wesen« seien, wie in zwei Kurz-Interviews explizit erwähnt und in diversen weiteren (mittels verwandter Formulierungen) angedeutet wird.

41 13_20180312_NZZStorytelling-01 (44)

bleiben anekdotisch, beispielsweise mittels Kommentaren auf Twitter oder Leser*innenbriefen. Neuere Vermessungspraktiken des Publikums auf den Online-Auftritten der Medienorganisationen nehmen zwar auf Redaktionen viel physischen (Stichwort: Chartbeat) und diskursiven Raum ein, verbleiben aber in ihrer Interpretation sehr offen (Christin 2018) und erlauben kaum den Bezug auf einzelne Datenvisualisierungen.

So stützen sich die datenjournalistischen Praktiken auf die grafische Expertise, um die Rezeptionsprozesse bei einem breiten Publikum zu strukturieren – salopp ausgedrückt: »komplexe Themen anhand von Visualisierungen verständlich machen«⁴². Diese Expertise kann verschieden geformt sein, von formellen Regeln bis zu einem Fingerspitzengefühl. Dieses verbleibt oft implizit – mit Ausnahme beispielsweise der folgenden Situation: Ihr ästhetisches Fingerspitzengefühl lässt Infografikerin Caroline auf technische Details verzichten, um stattdessen der Formgebung mehr Priorität einzuräumen, um die ›richtige‹ Assoziation bei den Lesenden hervorzurufen:

[Zora] diskutiert mit [Caroline], wie die Rakete genau illustriert werden soll. [Caroline] ist der Meinung, dass das Schema technisch nicht allzu ausgefeilt sein muss, lieber mit ›Verkleidung‹. Formen würden von den Lesern immer mit bekannten Gegenständen (wie dem Wallholz) in Verbindung gebracht.⁴³

Caroline bezieht sich wenige Tage später ein weiteres Mal auf das implizite Wissen und beschreibt es als »Gefühl«, das bei ihrer grafischen Arbeit eine grosse Rolle einnehme:

Als ich frage, was das Kriterium sei, Illustrationen und nicht Fotos zu verwenden, meint sie, dafür gebe es keine Richtlinien. Für sie wirke eine Infografik mit Fotos drin halt nicht »aus einem Guss«. Es gehe um das »Gefühl«, ästhetische Kriterien, die man auch nicht immer explizieren könne. Sowas könne man nicht in Richtlinien packen [...]⁴⁴

⁴² BR_Fragebogen_Mitarbeitende_08 – wobei hier interessanterweise die Vergleichsfolie nicht andere Journalismus-Formate in der eigenen Organisation sind, sondern das Teammitglied betont, dass andere Datenjournalismus-Teams Visualisierungen besser und öfter einsetzen.

⁴³ 13_20180312_NZZStorytelling-01 (25)

⁴⁴ 17_20180316_NZZStorytelling-05 (5)

Die Feldnotiz deutet bereits an, dass die Formalisierung der grafischen Expertise durch Richtlinien Reibungspunkte bietet – Abschnitt 6.3 wird diese vertiefen.

Differenzierung der Service-Rolle

Die Koordination von journalistischer und grafischer Expertise in den beiden Teams fällt aber nicht ins traditionelle Schema der Infografik als »Service-Abteilung« zurück. Dies zeigt sich in einem Vergleich in einer Situation, in der die rein journalistische Expertise die grafische Umsetzung bestimmt, ja geradezu diktiert. NZZ-Infografikerin Zora finalisiert in dieser Situation eine Datenvisualisierung für die Print-Ausgabe. Obwohl die Datengrundlage wie auch die Visualisierungsform längst festgelegt sind, kommt es vor dem Bildschirm von Zora zu einer weiteren Aushandlungssituation mit dem Wirtschaftsressort über die Annotationen.⁴⁵ Dieses ist mit dem Umfang der gewählten Beschriftungen nicht einverstanden. Dabei übernimmt eine Redaktorin aus dem Ressort sogar selbst den Stuhl von Zora und fügt direkt im Grafikprogramm Beschriftungen ein. Zora bemüht sich, ästhetische Kriterien einzubringen, damit die Beschriftungen nicht überhandnehmen. Die mittels Zwang durchgesetzten Ansprüche des Wirtschaftsressorts stehen stellvertretend für die Mechanismen, wie Datenvisualisierungen in traditionellen Service-Konstellationen montiert werden – nämlich so, wie es die journalistische Expertise für richtig hält. Davon heben sich die fein zisierten Koordinationsmuster innerhalb von NZZ Storytelling deutlich ab, indem ästhetische Kriterien eine stärkere Position einnehmen und mehr Berücksichtigung finden.

Die Situation weist aber darauf hin, woraus der Arbeitsalltag für das Team – neben der Koordination innerhalb – eben auch sehr oft bestand: aus der Verteidigung eines Zuständigkeitsbereiches mit fragilen und stets neu auszuhandelnden Grenzen an der organisationalen Bruchlinie mit anderen Ressorts. Dabei tritt immer wieder ein Machtgefälle hervor, wie es Smit et al. in ihrer Untersuchung von »trading zones« in der Informationsvisualisierung festgestellt haben: »Secondly, organizations with a dominant culture rooted in verbal thinking seem to be inclined to coercive forms of collaboration as far as the production of information visualisations is concerned. Our research shows that this led to underestimation of the skills and analytical competences of the designers« (Smit et al. 2014:49). Für die Aushandlungen innerhalb der beiden Teams trifft diese Schlussfolgerung sicher nicht mehr zu, wie die sonstigen

45 17_20180318_NZZStorytelling-05 (15)

Beispiele aufgezeigt haben: Die grafische Expertise hat die Evidenz in vielen Beispielen nach ästhetischen Montage-Kriterien befestigt, wie die Diskussion bezüglich der Wahlstatistiken (Fussnote 39, Kapitel 6) sowie die Überlegungen Carolines (Fussnoten 43 und 44 des Kapitels 6) zeigen.

Die Praktiken der Personalisierung und Konkretisierung (‘humanizing data’, siehe Abschnitt 6.1.2) sind in der journalistischen Epistemologie tief verankert und viel erforscht (beispielsweise Schudson 2000, Galtung und Ruge 1965). Von den von Schudson (2000:191) vermuteten Gründen – Anschluss an das westliche Bild der Individualisierung; Steigerung der persönlichen Identifikation durch ein Narrativ; bessere Einpassung der Handlungen von Einzelpersonen in den täglichen Nachrichtenzyklus – erscheint besonders der zweite für datenjournalistische Praktiken relevant. Durch ästhetische Markierungen oder Illustrationen von einzelnen Personen und Fällen stimmt die grafische Expertise eine Datavisualisierung mit der journalistischen Konvention der Personalisierung ab, welche Aufmerksamkeit erregen und das Sich-Einlassen auf eine Geschichte erleichtern will.

Diese Vorstellung gründet darauf, dass sich Einzelfälle als Erzählform auf einer tieferen Abstraktionsebene befinden als quantitative Erkenntnisse. Die Komplexitätsreduktion induziert in der journalistischen Expertise eine bessere Zugänglichkeit der Sachverhalte. Espeland (2015:71) hat die Funktion von Narrativen⁴⁶ anhand der quantitativen Ordnungsform der Schulrankings beschrieben: »But if we look closely at the structure of reactions to quantification we can see that many take on something like a narrative, with many of the same components, and these narratives often put back the agency, context, identity, sequence, reasoning and emotions that quantification removes.« Auch wenn Espeland in ihrem Artikel die Nutzung von Narrativen aus einer anderen Perspektive – derjenigen der durch Quantifizierung reduzierten Subjekte – beschreibt, so ist die Legitimation zur Nutzung von Narrativen dieselbe: um »agency, context, identity, sequence, reasoning and emotions« in ihre Erkenntnisse zurückzubringen, denn diese Elemente sind aus Sicht der journalistischen Expertise nötig, um sie als Geschichten einem breiten Publikum zugänglich zu machen. Narrativisierung wie auch Annotations machen Datavisualisierungen für die Rezipierenden bewegend und

46 Es ist deshalb kein Zufall, dass die Autorin für die Definition des Begriffs Narrativ ihn in Verbindung setzt mit »to recount« (erzählen, nacherzählen) sowie »story« (Geschichte), siehe Espeland 2015:60f.

resonierend und somit anschlussfähig – so auch die Vorstellung im Feld der Datenvisualisierung (Meyer und Dykes 2019).

Die in den Abschnitten 6.1.1 und 6.1.2 beschriebenen Praktiken der Hervorhebung der Hauptaussage sowie der Personalisierung und Konkretisierung durch ‚humanizing data‘ zeigen, dass grafische Expertise in den datenjournalistischen Praktiken in die journalistische Kultur der Komplexitätsreduktion eingepasst wird: Datenvisualisierungen sollen *eine* Geschichte unmittelbar sichtbar und zugänglich machen.

Koordination verschiedener Transparenzpraktiken durch Entflechtung

Während die Teams mit den Praktiken der Hervorhebung der Hauptaussage und der Personalisierung und Konkretisierung bekannte journalistische Vorgehensweisen der Komplexitätsreduktion und Zugänglichkeit auf Datenvisualisierungen münzen, scheint die (teilweise) Transparentmachung von Daten und Methoden in eine andere Richtung zu zielen: Das Sichtbar- und sogar Verfügarmachen von Datenquellen und Methoden erhöht die Informationsdichte in der Präsentation journalistischer Erkenntnisse. Ausgewählte Transparenzpraktiken über die Quellen und die Methodik gehörten explizit oder implizit schon immer zur journalistischen Kultur (Allen 2008), Detaillierungsgrad und die Veröffentlichung der Datengrundlage in datenjournalistischen Praktiken gehen aber weit darüber hinaus. Abschnitt 6.1.3 hat nachgezeichnet, welche Transparenzpraktiken die beiden Teams für notwendig erachteten, um mittels Datenvisualisierungen Evidenz zu befestigen. Welche Expertisen werden dafür wie koordiniert?

Die Veröffentlichung von Datensätzen folgt einer fein austarierten Abstimmung von technischer und journalistischer Expertise, wobei gerade im investigativen Bereich, in dem sich BR Data oft bewegt, auch rechtliche Aspekte abgewogen werden. Die Feld-Übersicht hat bereits auf den Einfluss der programmiertechnischen Expertise und ihrer (Open-)Kultur verwiesen (Abschnitt 3.2.1). Dies zeigt sich auch in den beiden untersuchten Teams: Ansprüche an offene Verwaltungsdaten sind immer wieder zu hören, und genau so will man eigene Werkzeuge und Methoden offenlegen (siehe beispielsweise Abschnitt 6.3.2 zu Q). Eine wichtige Rolle spielen dabei die Überschneidungen der »civic tech« und der datenjournalistischen Community, die sich in einigen Biografien von Mitgliedern von beiden Teams zeigt (siehe Baack 2018 sowie Erläuterungen dazu im Abschnitt 3.2.1). Baack schliesst, dass es zwischen den beiden Communitys keine konfigurerenden Logiken gibt, sondern die beiden Ansätze sich ergänzen, indem neue Typen auf einem Kontinuum entstehen,

die jeweils situativ die beiden Kulturen koordinieren. Genauso situativ wägen jeweils beide untersuchten Teams die Veröffentlichung von Datensätzen ab, wie Abschnitt 6.1.3 gezeigt hat.

Ähnliche Abwägungen machen die Teams in Bezug auf die Veröffentlichung der Methodik. Fragen der Transparentmachung der benutzten Methoden und Werkzeuge wirken bereits in den Prozess der Erkenntnisgewinnung hinein: Die Projektverantwortliche Lea diskutiert sehr oft mit Martin, aber auch Lukas und Moritz darüber, wie sie damit umgehen wollen, dass das für einige Datenanalysen der Polizeimeldungen genutzte externe *Tool* für sie eine Blackbox darstellt. Alle verfügen über programmiertechnische Expertise zur Datenauswertung und antizipieren in ihren Diskussionen den Konflikt zwischen der verschleierten Methodik des externen *Tools* (die sie als komplex antizipieren, aber in das auch das Team nicht einsehen kann) und der Legitimität der präsentierten Erkenntnisse. Die Evidenz der Ergebnisse ist für sie mit einer transparenten Methodik verknüpft. Gleichzeitig äussern BR-Data-Teammitglieder in den Kurz-Interviews, dass Datenjournalismus umstrittene Diskussionen auf eine »Faktenbasis« zurückholen kann⁴⁷ oder mindestens eine zusätzliche Perspektive darstellt, um andere Quellen zu kontrollieren, »eine Art Fact-Checking«⁴⁸. Immer wieder schmiegt das Team sich dabei an »quasi wissenschaftliche Standards«⁴⁹ an, welche ebenfalls Transparenz verlangen und die so die Glaubwürdigkeit des eigenen Vorgehens nicht nur sich selbst gegenüber, sondern auch gegenüber dem Publikum sichern sollen. Vielen Teammitgliedern ist das wissenschaftliche Arbeiten aufgrund eines abgeschlossenen sozialwissenschaftlichen oder technischen Studiums bekannt. Insgesamt widerspiegelt sich in diesen Aussagen und Werthaltungen das bereits im Abschnitt 3.3.1 beschriebene Narrativ, welches datenjournalistische Praktiken in die Nähe der wissenschaftlichen rückt.

Im Hintergrund schwingt dabei auch immer die Idee eines Peer-Review-Prozesses mit, der auf die vielfältigen Austauschplattformen der datenjournalistischen Community referenziert (siehe Abschnitt 3.1). Während die Teams mit einer Simplifizierung (Hervorhebung Hauptaussage, Personalisierung und Konkretisierung) das breite Publikum ihrer Medienorganisationen ansprechen, adressieren sie das eigene Feld durch zusätzliche Informationen und Transparenzpraktiken. Der Bewertung durch die Community wird viel

47 BR_Fragebogen_Mitarbeitende_04

48 BR_Fragebogen_Mitarbeitende_03

49 47_20181219_BR-Data-01 (12)

Relevanz zugeschrieben und die Teams investieren entsprechend viel Zeit, den Konventionen wie Transparenz über Daten und Methoden gerecht zu werden.

Die verschiedenartigen Ansprüche bezüglich der Sichtbarmachung von Komplexität koordinieren die Teams durch Entflechtung. Das heisst, es werden in der Evidenzbefestigung an verschiedenen Orten unterschiedliche Ansprüche bedient. Während die Datenvisualisierungen eindeutig und konkret auf eine Hauptaussage abzielen, *eine* Geschichte evident machen, werden die methodischen Überlegungen in Kästchen, am Ende eines *Longreads*, in einem zusätzlichen Artikel oder – sehr oft – auf Github veröffentlicht.⁵⁰ Beide Teams verfügen über gut genutzte Github-Repositories.

Die Vorstellung der Entflechtung zeigt sich, als ich Moritz von BR Data frage, wie viel und welchen Platz das Ausweisen der eigenen Methoden und der Konstruktion der Datenvisualisierung einnehmen soll:

[Moritz] meint dann, man müsse sich immer die Frage stellen, wie stark die eigene Recherche und das methodische Vorgehen Teil der Geschichte werden sollen. Er konzentriere sich lieber auf die Erkenntnisse. Es sei ja in den USA inzwischen auch so, dass Datenjournalismus selten mehr als Begriff gebraucht werde, im Gegensatz von vor 4, 5 Jahren. Es sei schon wichtig die Methodik aufzuzeigen, aber nicht mehr zwingend in der Geschichte selbst.⁵¹

Moritz verbindet die Entflechtung von Geschichte und Methodik mit einer Fortentwicklung⁵² des Datenjournalismus beziehungsweise mit der Distinktion zwischen der Präsentation, in welcher datenjournalistische Erkenntnisse als Geschichte evident werden, und dem Aufzeigen der Methodik, das neben- oder nachgelagert werden kann. Die Ambiguitäten und Kontingenzen des Konstruktionsprozesses erreichen so einen kleinen, aber in der datenjournalistischen Perspektive relevanten Kreis von Interessierten – die Community –, während sich die Evidenz für das breite Publikum in Form von Datenvisualisierungen an konventionellen journalistischen Praktiken orientiert.

50 Die Aufzählungen betreffen vor allem Veröffentlichungen online. Auf Papier (NZZ) oder in Fernseh- und Radiobeiträgen (BR) bleiben methodische Überlegungen praktisch komplett unsichtbar.

51 49_20181212_BR-Data-03 (9)

52 Das datenjournalistische Feld teilt im Grossen und Ganzen die Vorstellung, dass US-amerikanische Medienorganisationen mit ihren datenjournalistischen Praktiken stets einige Jahre voraus, also ›Trendsetter‹ sind. Gemeinhin wird dies damit begründet, dass das Feld dort auf die längste Historie zurückblickt.

Interne Sichtbarkeit von differenzierten Expertisen

Abschnitt 6.1.4 hat beschrieben, dass die beiden Teams veränderte Praktiken der Autor*innenschaft anwenden, indem sie die beteiligte technische und grafische Expertise differenziert ausweisen. Wie die kurze Diskussion zwischen Kaspar und David (siehe Fussnote 29, Kapitel 6) zeigt, ist die Reflexion über die Beteiligung von technischer, journalistischer und grafischer Expertise an der Produktion datenjournalistischer Erkenntnisse breit verankert. Aufgrund dieser Einigkeit über alle Teammitglieder hinweg entstehen kaum je Situationen, in denen ein ausgeweitetes *Byline*-Konzept innerhalb des Teams kontrovers diskutiert wird. Es scheint sowohl aus der grafischen wie der technischen Expertise heraus wünschenswert, dass die eigenen Beiträge gleichberechtigt mit dem ›klassischen‹ textbasierten Journalismus genannt sind. Mitglieder der beiden Teams mit starker journalistischer Expertise wiederum reflektieren eine enge Verwicklung ihrer Praktiken mit den grafischen und technischen Expertisen, weshalb auch ihnen eine differenzierte *Byline* sinnhaft erscheint.

Dies verweist auf den grossen Einfluss der *Open*-Kultur, der sich nicht nur in den Feld-Diskursen (Abschnitt 3.2.1) zeigt. Mit dem Ausweisen von technischer oder grafischer Expertise in der *Byline* öffnen die Teams einen Aspekt des Konzepts der journalistischen Glaubwürdigkeit, die Rechenschaftspflicht durch Autor*innenschaft, für Programmierer*innen und Designer*innen. Diese Praktik zur Evidenzbefestigung richtet sich dabei weniger an das breite Publikum, wie die Beobachtungen in den beiden Teams zeigen. Eher geht es darum, die mittels datenjournalistischer Praktiken neu definierte Materialität, Methodik und Kultur der Produktion von Evidenz – montiert auf Datavisualisierungen – innerhalb der eigenen Redaktion und Medienorganisation einzubringen. Durch die Rekonzeptualisierung der *Byline* werden diese Vorstellungen mit dem traditionellen journalistischen Konzept der Autor*innenschaft verbunden. Gleichzeitig reartikulieren die Teams dabei die Idee der *Open*-Kultur, dass Kollaborationen und vielfältige, ausdifferenzierte Expertisen zur Produktion beitragen. Dies impliziert die Akzeptanz, dass die Erkenntnisproduktion nicht mehr auf das individuelle Subjekt zurückzuführen, sondern eine kollektive epistemische Arbeit ist (Knorr-Cetina 2002) – und dass dieses Modell auch gegenüber der eigenen Organisation und der Öffentlichkeit zu kommunizieren ist.

6.2.2 *Traders II*

Die Erläuterungen über *trader*-Rollen und deren Verteilung in den Teams aus Abschnitt 5.2.2 treffen in ähnlicher Weise auf das Befestigen von Evidenz zu: Mitglieder, welche mehrere Expertisen und ihre spezifischen Perspektiven überbrücken, vermögen überdurchschnittlich oft sinnhaft zu koordinieren. Die grafische Expertise gewinnt an Bedeutung, wenn es nun darum geht, datenjournalistische Erkenntnisse evident zu machen – schliesslich bietet die grafische Expertise Werkzeuge an, von welchen sich die Teams versprechen, dass Datenvisualisierungen für das Publikum überzeugend und verständlich montiert werden.

In beiden Teams werden Datenvisualisierungen sehr oft gegenseitig evaluiert, indem man andere Teammitglieder nach einer Bewertung fragt. Eine ›gute Datenvisualisierung‹ ist eine, welche innerhalb der Teams durch mehrere Evaluationen geprüft wurde. Durch diese Multiperspektivität hoffen die Teammitglieder, die Rezeptionssituation zu simulieren und entsprechende Defizite in der Evidenzbefestigung aufzuspüren. Gleichzeitig verbleibt die Vorstellung, wie das Publikum eine Datenvisualisierung interpretiert, sehr von den eigenen *peers* geformt – dies hat Dick (2014) auch bereits in englischen Redaktionen festgestellt. Die *trader*-Rolle übernehmen aus diesen Gründen insbesondere Teammitglieder, welche die journalistische mit der grafischen Expertise zu verbinden vermögen.

Insbesondere bei noch weniger oft genutztem visuellem Vokabular wie der Sankey-Grafik der Fussballtransfers (Abbildung 17) oder den ›Barcode‹-Visualisierungen der Pestizid-Geschichte⁵³ bedingt das Befestigen von Evidenz in den Teams multiperspektivische und mehrfache Bewertungs- und Koordinationssituationen, damit die Datenvisualisierung zur Geschichte wird. Im Gegensatz dazu ist die Nutzung von Balken- oder Liniendiagrammen zur Montage von Evidenz gut eingeübt – was die Einigung auf Leitlinien und Standards sowie deren Kodifizierung und die Entwicklung von Tools ermöglicht. Die nach dem Zwischenfazit folgenden Abschnitte beleuchten diese Leitlinien und Kodifizierungen sowie die Aushandlungsprozesse, welche damit einhergehen.

53 <http://web.br.de/interaktiv/pestizide/> (11.06.2020)

6.2.3 Fazit

Es lässt sich bezüglich der vier beschriebenen Praktiken festhalten, dass für die Hervorhebung der Hauptaussage und die Personalisierung und Konkretisierung grafische Expertise so eingepasst wird, dass ihre ästhetischen Kriterien die epistemologischen und professionskulturellen Vorstellungen der journalistischen Expertise unterstützen. Die Befestigung von Evidenz bedeutet, dass dem breiten Publikum eine Datenvisualisierung als *Geschichte* augenscheinlich wird. Das Ideal der Transparenz, bezogen auf die Veröffentlichung von Datensätzen und des methodischen Vorgehens, geniesst in beiden Teams einen hohen Stellenwert; die Programmierer*innen-Expertise belebt mit ihren diesbezüglichen Vorstellungen ein ›klassisches‹ journalistisches Ideal neu. Beide Teams setzen mehr Transparenzpraktiken um, wenn Geschichten innerhalb des Teams produziert und eigene Formate genutzt wurden (siehe Abschnitt 6.4). Beide Teams lagern aber die Transparenzpraktiken der Datenvisualisierung nach, das heisst, Datensätze und Erläuterungen über das methodische Vorgehen finden sich am Rand der entsprechenden Artikel oder durch einen Link zugänglich auf anderen Seiten. Mit den zusätzlichen Informationen wird primär die datenjournalistische Community referenziert, um im Feld peer-review-ähnliche Diskussionen über die Methodik zu ermöglichen.

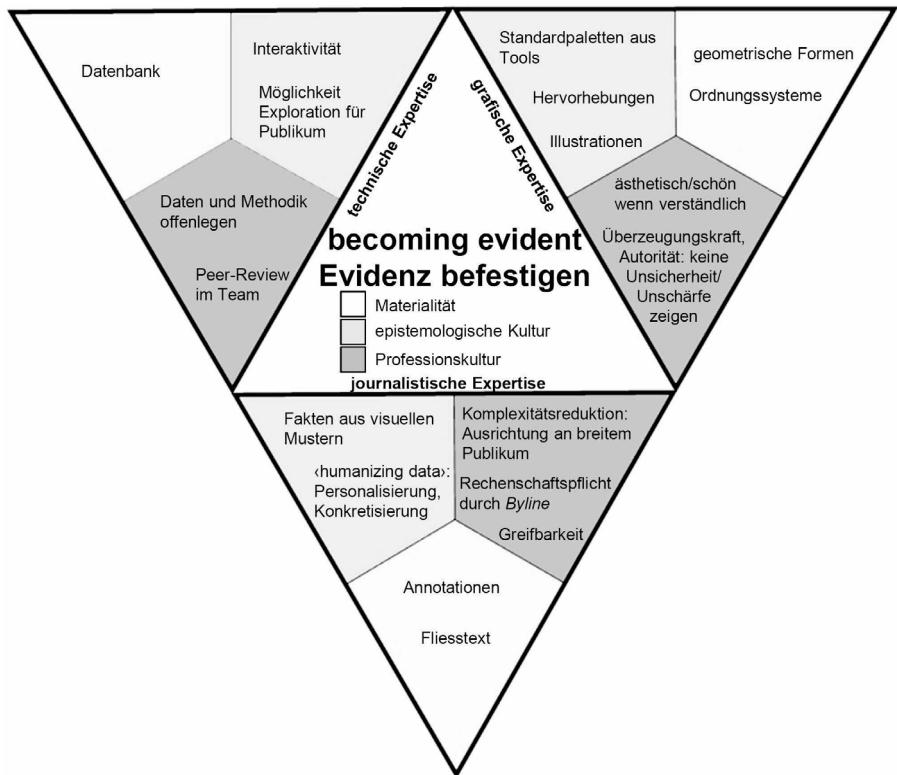
Datenvisualisierungen selbst verbergen ihre Konstruktionsmethodik. Spuren von Unsicherheiten, Annahmen und Hinweise auf die Kontingenz von Entscheidungen fehlen. Grafische Expertise könnte dafür eingesetzt werden, Hinweise darauf in die Präsentation einzufügen – allerdings sind solche Praktiken in der Datenvisualisierungscommunity selbst umstritten und bisher in kaum einem Anwendungsbereich üblich: Man fürchtet um die Autorität und Überzeugungskraft der Evidenz (siehe dazu Hullman 2020).

Durch die Ausweitung des *Byline*-Konzepts auf technische und grafische Expertise nutzen beide Teams eine journalistische Institution – mit der Autor*innenschaft wird Glaubwürdigkeit und Rechenschaftspflicht verbunden, schliesslich hat man ›skin in the game‹ –, um die technischen und grafischen Beiträge als Teil der Erkenntnisproduktion und Evidenzmontage einzuordnen, indem sie in die *Byline* gehoben werden. Diese Praktik soll für ein breites Publikum sichtbar sein, richtet sich mit dem hohen symbolischen Wert aber vor allem an die Kolleginnen und Kollegen in der eigenen Redaktion und Medienorganisation.

Die im Modus der Befestigung von Evidenz koordinierten Expertisen lassen sich – die Ausgangslage bezüglich Expertisen (Abbildung 2) weiterentwi-

ckeln und mit dem Modus des Erkennens von Daten (Abbildung 10) vergleichend – mit folgendem Dreieck aufzeigen:

Abbildung 22: Koordinierte Expertisen im Modus der Evidenzbefestigung



(eigene Darstellung)

6.3 Leitlinien und Tools

Leitlinien und Datenvisualisierungswerkzeuge der beiden Teams kodifizieren bestimmte Prinzipien beziehungsweise schreiben sie in Software fest und betten sie so in die Arbeitspraktiken ein. Der erste Abschnitt erläutert die

Entstehung des *Styleguide* von NZZ Storytelling, welcher sowohl gegenüber der Redaktion wie auch innerhalb des Teams bestimmte Koordinationsmuster der Expertisen kodifiziert und damit institutionalisieren soll. Der zweite Abschnitt fokussiert auf die zwei in den Teams meistgenutzten Software-Tools, um Evidenz mittels Datenvizualisierungen zu befestigen: Q und Datawrapper. Sie formalisieren spezifische Koordinationsmuster durch die Festschreibung von Montage-Optionen. Der Abschnitt beleuchtet aber auch die Grenzen der Institutionalisierung in *Tools* und Leitlinien.

6.3.1 *Styleguide*

Abschnitt 6.1 hat bereits einen Einblick in den NZZ Storytelling *Styleguide*⁵⁴ erlaubt, welcher verschiedene Praktiken kodifiziert, wie das Team eine Datenvizualisierung evident montieren will: dass eine Hauptaussage hervorgehoben wird, dass dies mittels Lesart, Leseführung und Lesebeispiele geschieht; und die Konzeptionierung der *Byline*. Die Entwicklung dieses Leitfadens begann zur Zeit der Forschungsphase bei NZZ Storytelling, während derer insbesondere Emma das Projekt vorantrieb – sie war noch nicht allzu lange im Team und bezeichnete die Vereinheitlichung der Design-Prinzipien (eines der vom Teamleiter vorgegebenen Jahresziele) als eine ihrer Hauptaufgaben.⁵⁵

Emma wie auch Kaspar betrachten sich als *Designer*, zwar dieselben Wurzeln aufweisend, aber abgegrenzt von der seit Jahrzehnten in Redaktionsräumen traditionellen Infografik-Profession und dem entsprechenden *Subteam* bei NZZ Storytelling. Die Unterschiedlichkeit der Rollenverständnisse und Arbeitsweisen der beiden *Subteams* manifestiert sich in jenem Moment, in dem sich beide auf Initiative von Emma zusammensetzen und sie vorschlägt, gemeinsam einen *Styleguide* zu entwickeln.⁵⁶ Dabei trifft der Wunsch nach einer Festlegung von Design-Prinzipien, die sowohl für handgezeichnete wie auch für Software-generierte Visualisierungen gelten, auf einen teilweise impliziten, teilweise explizierten Bestand an Leitlinien der Infografiker*innen (siehe Abschnitt 6.1.1). Während die Festlegung einheitlicher Design-Prinzipien für das eine *Subteam* eine Voraussetzung ist, um überhaupt die Software

⁵⁴ Der *Styleguide* ist im Internet zugänglich: <https://nzzdev.github.io/Storytelling-Styleguide/#/> (18.04.2020).

⁵⁵ NZZ_Fragebogen_Mitarbeitende_05

⁵⁶ 15_20180314_NZZStorytelling-03 (10–14)

für computergenerierte Datenvisualisierungen konfigurieren zu können, interpretieren einige Infografiker*innen dies als Einschränkung der kreativen Gestaltungsfreiheit und somit als Misstrauensvotum gegen die Kompetenz, auf Basis der eigenen Expertise eine ästhetisch und kommunikativ stimmige Visualisierung zu montieren. Später wird die Infografikerin Caroline erklären, dass es bei ihrer Arbeit ganz oft um das »Gefühl« gehe, um ästhetische Kriterien, die man nicht immer explizieren und in Richtlinien packen könne (siehe Zitat Fussnote 44, Kapitel 6).⁵⁷

Als schon seit langem und relativ autonom vom Rest von NZZ Storytelling funktionierendes *Subteam* verfügen die Infografiker*innen über einen eigenen Arbeitsrhythmus und eigene Organisationsprozesse. Infografiken werden gemäss mehr oder weniger detaillierter Auftrag durch ein Redaktionsmitglied ausgeführt, allenfalls im Entwurfsstadium iterativ mit dessen Erwartungen abgeglichen. Die Koordination von grafischer und journalistischer Expertise erfolgt je Auftrag, aber innerhalb eines engen Spielraums einer über Jahre gewachsenen (Service-)Rollenzuschreibung. Durch den Aufbau des Storytelling-Teams, inklusive des Infografik-Subteams, erweitert sich die Querschnittsfunktion der Befestigung der visuellen Evidenz, indem sie auf weitere Bereiche der journalistischen Erkenntnisproduktion, nämlich auf die Zusammenarbeit in einzelnen grösseren Projekten wie bei den Fussballtransfers oder die Visualisierungssoftware Q, ausgebaut wird. Deshalb erscheint es den Teammitgliedern, insbesondere den Designer*innen, sinnvoll, die Visualisierungsprinzipien zu vereinheitlichen und zu kodifizieren.

Die situative Zusammenarbeit mit den Ressorts und die Nutzung von Q haben Räume geschaffen, in denen die Koordination der Expertisen noch nicht gefestigt ist – und dafür soll der *Styleguide* für das Team eine Rückzugsposition darstellen, um auch aus einer situativ vorkommenden Aussenseiterposition die Montage-Prinzipien der grafischen Expertise zu sichern. Als es um die Entwicklung des *Styleguide* geht, merkt Emma an, dass dieser auch helfen würde, in der Kritik von Q-Visualisierungen auf eine Basis mit gemeinsamen Regeln zurückgreifen zu können.⁵⁸ Explizierte und gemeinsam getragene Regeln und deren Kodifizierung sollen die grafische Expertise über das Befestigen von Evidenz aus Daten gegenüber dem Rest der Redaktion absichern und die Aushandlungsposition stärken.

57 17_20180316_NZZStorytelling-05 (5)

58 15_20180314_NZZStorytelling-03 (12)

In dem halben Jahr nach der Forschungsphase entsteht der NZZ Storytelling *Styleguide*. Man habe schliesslich kleine gemeinsame Nenner zwischen den verschiedenen Design- bzw. Infografik-Rollen gesucht und gefunden, blickt ein Teammitglied in einem Nachgespräch zurück.⁵⁹ Der ursprüngliche Plan sei gewesen, dass das gesamte Team den *Styleguide* gemeinsam pflegt und weiterentwickelt – durch den inzwischen erfolgten Weggang von Emma laufe man aber Gefahr, dass sich niemand mehr verantwortlich fühle, weil es schon sie gewesen sei, welche den Leitfaden initiiert und gepflegt habe.

Ob es mit der fehlenden Verantwortungszuweisung zu tun hat, liess sich nicht eruieren – aber in der derzeit online zugänglichen Version fehlt ein Me-nü-Punkt des *Styleguide*, welcher zeitweise aufgeschaltet war. Dieser war mit »TBD« (»to be determined«, auf Deutsch »noch zu bestimmen«) bezeichnet und sollte Themen und offene Fragen in der Gestaltung der NZZ-Datenvisualisierungen sammeln, welche das Team gemeinsam ausdiskutieren und Leitlinien dafür festlegen wollte. Jeden Monat sollten dann zwei bis drei dieser offenen Punkte herausgepickt und an eine Person aus dem Team delegiert werden, welche dann Lösungsvorschläge erarbeitet – so erläuterte es die Beschreibung. Diese Lösungsvorschläge sollten dann diskutiert, die bestimmte Leitlinie dokumentiert und implementiert werden. So waren im Januar 2019 im *Styleguide* die beiden Themen »colors to represent gender« sowie »how to visualize uncertainty« vermerkt.⁶⁰ Es handelte sich dabei um das Bestreben, die Koordination der Expertisen für die Montage in bestimmten Punkten zu explizieren und zu kodifizieren. Damit einhergegangen wäre nicht nur die bereits angesprochene Absicherung gegenüber der Redaktion, sondern auch Kohäsion innerhalb des Teams, indem Leitlinien gemeinsam und unter Einbezug der verschiedenen Expertisen erarbeitet und koordiniert werden. Wieso die Idee der explizierten Aushandlung von weiteren Montage-Leitlinien nicht weiterverfolgt wurde, bleibt unklar.

6.3.2 Q und Datawrapper

Beide untersuchten Teams arbeiten für einen Teil ihrer Datenvisualisierungen mit standardisierten Ausgabe-Tools: Q und Datawrapper. Beide Tools kodifizieren die Montage-Praktiken von Datenvisualisierungen durch die Festbeschreibung in Technologien. Sie schränken den Koordinationsspielraum der

59 64_20200129_NZZStorytelling-15

60 Memo Styleguide (23.01.2019)

Expertisen ein auf eine Bandbreite, die von den Entwickler*innen als angemessen eingeschätzt wird: Welche Arten von Daten in welchen Formaten sind verrechenbar? Welches visuelle Vokabular (Ordnungssysteme und geometrische Formen) ist wählbar? Welche Annotationen und Hervorhebungen sind möglich? Die Tools engen die Kontingenz der lokalen Koordination ein, indem sie gewisse Konventionen aus den Expertisenbereichen festigen, allenfalls unumgänglich machen. Die nötige Ein- und Anpassung der eigenen Praktiken in diese Affordanzen verursacht einen steten Diskussionsbedarf – insbesondere bei Q, dem Tool der NZZ, welches das Storytelling-Team selbst entwickelt hat.⁶¹ Es ist benannt nach der Figur in den James-Bond-Filmen: Dort steht »Q« für den Leiter der Forschungs- und Entwicklungsabteilung des Geheimdienstes respektive die Abteilung selbst. Seine bzw. ihre Aufgabe ist es, den Hauptakteur mit neu erfundenen und teilweise extravaganten Instrumenten auszustatten, mit welchen er dem Feind überlegen ist – und das jederzeit. Mit der Referenz zu Q steckt NZZ Storytelling die Aufgabe der Toolbox ab: »It provides tools for others to do their job.«⁶²

Q ermöglicht es allen Mitgliedern der NZZ-Redaktion, selbst einfache visuelle Elemente in eine Geschichte einzufügen. Q bietet als Formate Karten, Tabellen, Balken- und Liniendiagramme, Wahlauswertungen und einige mehr.⁶³ Q ist ein Bestandteil der 2017 implementierten Strategie »Digitale Entwicklung der NZZ« und soll den Inhalt der NZZ verbessern: »So wie wir bei Technologie und optischem Markenerlebnis Marktführer werden wollen, müssen wir auch bei digitalen Inhalten, ihrer Aufbereitung und ihrer Präsentation Vorreiter sein.«⁶⁴ Man habe mit Q »etwas Nachhaltiges« schaffen wollen, anstatt Datenvisualisierungen nur bei multimedialen Grossprojekten zu nutzen – verschiedene Personen aus der NZZ hätten zu derselben Zeit ähnliche Ideen gehabt, darunter er und David Bauer, so Chefredaktor Eric Gujer über die Entstehung von Q.⁶⁵

61 Tool-Informationsseite: <https://q.tools/> (02.02.2019)

62 David Bauer: Why newsrooms need storytelling tools and what we've learned building them. Blog-Post auf medium.com vom 18.01.2017: <https://medium.com/nzz-open/why-newsrooms-need-storytelling-tools-and-what-weve-learned-building-the-m-87df4802b737> (01.02.2019)

63 Q ist in einer Demo-Version zugänglich: <https://editor.q.tools/> (02.02.2019)

64 Präsentation: »Digitale Entwicklung der NZZ – Die nächsten Schritte für die Publizistik« (Eric Gujer, Carola Ettenreich, 3. Juli 2017), Seite 29.

65 Interview mit Eric Gujer vom 18. April 2018.

Q ermöglicht es für alle Redaktionsmitglieder, einzelne Aspekte der journalistischen Berichterstattung visuell evident zu machen – ohne Inanspruchnahme eines Infografik-Auftrags. Visuelle Evidenz soll dafür in einer grösseren Anzahl Geschichten und innert kürzerer Fristen bereitgestellt werden – so wie es die NZZ-Strategie vorsieht. NZZ Storytellings Auftrag sei es »to improve the average story«, so schreibt ein Blogbeitrag des Teams über die Anfänge von Q.⁶⁶ Es steht in dieser Arbeit aber weniger die organisationssoziologische Beziehung zwischen NZZ Storytelling und anderen Ressorts im Zentrum, sondern die Frage, welche Praktiken NZZ Storytelling wie in Q formalisiert.

Abbildung 23: Aus textlicher eine visuelle Evidenz machen – Folie aus der Schulungspräsentation von NZZ Storytelling für die Toolbox Q.

Diagramme

Das mag auf den ersten Blick erstaunen, gilt das Unternehmen doch eigentlich als äusserst erfolgreich. Netflix hat in 190 Ländern 104 Mio. Nutzer, die sich täglich 125 Mio. Stunden Inhalte aus der Bibliothek der Firma ansehen. Damit erwirtschaftete der Konzern im ersten Semester 2017 einen Umsatz von 5,4 Mrd. \$ und einen Gewinn von 244 Mio. \$. Das Wachstum der Abonnentenzahl übertraf im zweiten Quartal die Erwartungen der Analytiker deutlich und sorgte für einen starken Anstieg des Aktienkurses. Andere Technologiefirmen in einem ähnlichen Alter – wie Twitter und Uber – schreiben noch immer Hunderte Millionen Dollar Verlust.

Die Probleme von Netflix liegen denn auch nicht in der Erfolgsrechnung, sondern in der Bilanz. Dorthin fließt nämlich das Geld, das der Streaming-Anbieter in den Ausbau seiner Bibliothek steckt. Netflix stellt drei verschiedene Typen von Filmen und Serien zur Verfügung. Erstens sind das colbat produzierte Inhalte wie

Das mag auf den ersten Blick erstaunen, gilt das Unternehmen doch eigentlich als äusserst erfolgreich. Netflix hat in 190 Ländern 104 Mio. Nutzer, die sich täglich 125 Mio. Stunden Inhalte aus der Bibliothek der Firma ansehen. Damit erwirtschaftete der Konzern im ersten Semester 2017 einen Umsatz von 5,4 Mrd. \$ und einen Gewinn von 244 Mio. \$. Das Wachstum der Abonnentenzahl übertraf im zweiten Quartal die Erwartungen der Analytiker deutlich und sorgte für einen starken Anstieg des Aktienkurses. Andere Technologiefirmen in einem ähnlichen Alter – wie Twitter und Uber – schreiben noch immer Hunderte Millionen Dollar Verlust.

Netflix schreibt schwarze Zahlen

Gewinn 2016, in Mio. \$



Quellen: Geschäftsberichte, Bloomberg News

NZZ

Datenvisualisierungen mit Q sind wenig komplex und ihnen gehen kaum computergestützte Datenanalysen voraus, das legt die beschränkte Auswertungs- und Visualisierungspalette des Tools fest. Q-Visualisierungen sind eine

⁶⁶ David Bauer: Why newsrooms need storytelling tools and what we've learned building them. Blog-Post auf medium.com vom 18.01.2017: <https://medium.com/nzz-open/why-newsrooms-need-storytelling-tools-and-what-weve-learned-building-the-m-87df4802b737> (01.02.2019)

Errungenschaft klassischer journalistischer Berichterstattung, aber in anderem Format dargestellt (Abbildung 23). Q stülpt also visuelle Evidenzformen des Storytelling-Teams über die Erkenntnisse der Kolleg*innen aus den anderen Ressorts. Q schreibt dafür Montage-Praktiken in einer Software-Toolbox fest, indem es vordefinierte Formate und verschiedene Optionen festlegt.

Abbildung 24: Eingabemaske von Q. Titel und Quellenangabe sind Pflichtfelder.

Q

Speichern

Titel - erforderlich

Der Titel bringt die Kernaussage der Grafik auf den Punkt.

Untertitel

Der Untertitel benennt präzise, was in der Grafik zu sehen ist, inklusive Einheiten.

Pflichtfeld

Dieses Feld muss zwingend ausgefüllt sein.

Daten

	A	B	C	D	E
1					
2					
3					
4					
5					
6					
7					
8					

Zeilen und Spalten tauschen

Daten zum Herunterladen anbieten

Quelle(n)

Quelle löschen

Quelle - erforderlich

Pflichtfeld

Dieses Feld muss zwingend ausgefüllt sein.

Url

Direklink zur Quelle, http://...

Quelle hinzufügen

So finden sich in der Eingabemaske von Q die in den vorherigen Abschnitten beschriebenen Praktiken als Pflichtfelder wieder (siehe Abbildung 24): Eine Datenvisualisierung muss zwingend einen Titel und eine Quellenangabe aufweisen, eine Quellen-Verlinkung ist genauso möglich wie der Daten-Download. In einem weiteren Schritt offeriert Q einige Optionen für die Konfiguration der Visualisierung: Diagrammtyp (Säulen-, Linien-, Flächen-, Punkt- oder Pfeildiagramm); Hervorhebungsarten für Zeilen und Datenreihen; Einzelbeschriftungen für Höchst- und Tiefstwerte.

NZZ Storytelling hat Q selbst konzeptioniert und programmiert, ist für den Unterhalt und die Weiterentwicklung zuständig. Q schreibt viele der Koordinationsmuster der Expertisen fest – so finden sich die Grundsatzüberlegungen des *Styleguide* in Q als Vorgaben mit wenigen Freiheitsgraden wieder. Die Hervorhebungsoptionen erfordern, dass sich die Datenvisualisierung auf einen Aspekt, auf einen Sinnzusammenhang beschränkt. Bei der engen Palette an Diagrammtypen handelt es sich um die gängigsten, die sich an das *Visual Vocabulary* (siehe Abschnitt 5.3) anlehnen und die vor allem auch die bereits seit langem existierenden Infografik-Formate widerspiegeln. Die ausgewählte Palette schliesst daran an, was NZZ Storytelling den anderen Redaktionsteilen an Vorstellungen von visueller Evidenz aus Daten zuspricht und in der täglichen Zusammenarbeit auch bestätigt erhält: Eine Karte erstellt einen geografischen Zusammenhang, ein Balkendiagramm verkörpert einen Vergleich, eine Zeitreihe ordnet temporal. Dieses Vokabular ist einem Grossteil der Redaktion bereits aus der Zusammenarbeit mit der Infografik-Abteilung bekannt und die Verkörperung von Erkenntnis mit diesem visuellen Vokabular eingeübt. So kann von einer gewissen Grund-*Literacy* ausgegangen werden – Heflser und Mersch (2009:48) haben dies als »eingeübte Codierungen« visueller Referenzen bezeichnet. Das bestätigt sich beispielsweise im Austausch mit dem Auslandressort über mögliche gemeinsame Projekte: Die Ressort-Mitarbeitenden schlagen hauptsächlich bekanntes Vokabular, vor allem Karten, vor.⁶⁷

Grenzen der Festschreibung in *Tools*

Trotz der Formalisierung mit wenig Spielraum in Q verbleibt für NZZ Storytelling viel Arbeit, um die Visualisierungspraktiken in Q zu begleiten. Die Vermittlung von Best Practices in der Nutzung des Werkzeugs und ganz generell Konventionen aus dem Datenvisualisierungsbereich erscheinen dem Team wichtig, weshalb es sie über verschiedene Gefässe vermittelt: Schulungen, den »Storytelling-Dienst« sowie die Ausgestaltung von Q.

Für alle neu eintretenden Redaktionsmitglieder bietet NZZ Storytelling periodisch eine Schulung für die Nutzung von Q an. So weist Annalena an einer Schulung darauf hin, dass der für jede Q-Datenvisualisierung erforderliche Titel die wichtigste Aussage aufnehmen soll – und dass man Y-Achsen in Säulendiagrammen überhaupt nicht und in Liniendiagrammen nur sehr vor-

⁶⁷ siehe beispielsweise in einem Austausch über gemeinsame Projekte mit dem Auslandressort: 21_20180322_NZZStorytelling-09 (4)

sichtig beschneiden solle; dies sei heikel und man solle bei Achsenkürzungen immer Vorsicht walten lassen.⁶⁸

Den »Storytelling-Dienst« nimmt täglich ein*e Datenjournalist*in oder ein*e Designer*in wahr. Die oder der Tagesverantwortliche überwacht einen spezifischen Kanal der Chat-Anwendung Slack, in der ein Bot darauf hinweist, wann immer eine neue Q-Datenvisualisierung erstellt wurde. Diese soll dann von der Dienstperson kontrolliert und allenfalls »mit einem Handgriff« selbst optimiert und der oder die ursprüngliche Ersteller*in darüber informiert werden – so lautet die Anweisung im kurzen Leitfaden zum Storytelling-Dienst.⁶⁹ Der Leitfaden empfiehlt weiter, man solle auf »den Redaktor zugehen«, falls »eine Grafik überarbeitet werden muss«⁷⁰ – die Kompetenz, zu beurteilen, was als gute Datenvisualisierung gilt, beansprucht NZZ Storytelling im Leitfaden für sich. Die Bewertung erfolgt anhand der teamintern koordinierten Konventionen bezüglich Datenvisualisierung – der Leitfaden expliziert diese nicht, sondern nennt als »typische Fehler in Q« eine Reihe von Abweichungen von diesen Best Practices.⁷¹ Es wird impliziert, dass die diensthabenden Datenjournalist*innen und Designer*innen über diese Expertise verfügen und sie auch bereits internalisiert haben. Und dass sie die Best Practices weitergeben mit direkten Überarbeitungen von Datenvisualisierungen oder Diskussionen darüber. Die »Qualitätssicherung« – so ein formuliertes Ziel des Storytelling-Dienstes – zeigt: Obwohl es Q allen Redaktionsmitgliedern möglich macht, Datenvisualisierungen selbst zu erstellen, liegt die Bewertung und Qualitätssicherung gemäss dem Leitfaden bei NZZ Storytelling. Was eine *gute* Datenvisualisierung ist, welche eine Geschichte visuell evident befestigt, sollen schliesslich Personen beurteilen, welche nicht nur über journalistische Expertise verfügen, sondern die mit technischer und grafischer Expertise mindestens teilweise vertraut sind und mit entsprechenden Perspektiven auf die Datenvisualisierungen zu blicken vermögen.

Dass die Formalisierung der Evidenzbefestigung journalistischer Erkenntnisse durch Q mit aufwendigen Qualitätssicherungsmassnahmen begleitet werden muss, deutet darauf hin, dass die von Healy und Moody (Healy und

68 20_20180321_NZZStorytelling-08 (14)

69 NZZ_Beschreibung_Storytelling-Dienst, Seite 1.

70 NZZ_Beschreibung_Storytelling-Dienst, Seite 2.

71 Beispielsweise: »Datumsformat nicht richtig; Achsen nicht korrekt angeschrieben; Falscher Charttyp gewählt; Y-Achse zu krass beschnitten, so dass Aussage dramatisiert wird; [...]« (siehe NZZ_Beschreibung_Storytelling-Dienst, Seite 1).

Moody 2014:110) beschriebene kritische Lücke zwischen abstrakten Vorgaben und spezifischen Beispielen durch *Tools* doch nicht allzu trivial zu füllen ist:

»The tendency for the codification of data visualization to vacillate between overly abstract maxims and overly specific examples is characteristic of any craft where a practical sense of how to proceed—a taste or feeling for the right choice—matters for successful execution. A long-standing and plausible response to the problem is to have the designer make many of the judicious choices in advance and then embed them for users in the default settings of graphics applications. Given that graphical software aimed at regular users has been around for several decades now, however, these efforts have proven less successful than initially hoped.« (Healy und Moody 2014:110)

In der Erklärung, weshalb die »graphical software« die hohen Erwartungen nicht erfüllen konnte, verweisen die Autoren darauf, dass »[i]n the end, high-quality graphical presentation requires crafting a deliberately designed message rather than accepting the pre-established setting. Recent theoretical work explicitly recognizes the limits of relying on defaults« (Healy und Moody 2014:110). Auch Wickham verweist in seiner Umsetzung der »Grammar of Graphics« in das Datenvisualisierungs-Package *ggplot2* (bereits angesprochen in Abschnitt 5.4.1) darauf, dass Standardeinstellungen und Optionen nicht automatisch zu guten Visualisierungen führen: »A grammar may also help guide us on what a well-formed or correct graphic looks like, but there will still be many grammatically correct but nonsensical graphics. This is easy to see by analogy to the English language: good grammar is just the first step in creating a good sentence« (Wickham 2010:3). Konkrete Versuche, die Umwandlung von Text in Visualisierungen – wie es Abbildung 23 exemplifiziert – zu automatisieren, verdeutlichen die komplexen Prozesse der Produktion von visueller Evidenz (Cui et al. 2019).

Bezogen auf den Einsatz von Q bei der NZZ heisst das, dass das *Tool* den Aspekt der visuell adäquaten Wiedergabe einer Aussage nicht automatisieren kann. Es handelt sich um schwer kodifizierbare Elemente der datenjournalistischen Expertise. So bleibt die visuelle Evidenzbefestigung durch Q stets kritisch und in hoher Abhängigkeit von der einzelnen Nutzerin und dem einzelnen Nutzer. Die Bestrebungen des Storytelling-Teams, die ›User Experience‹ zu verbessern und mehr Anleitung innerhalb des *Tools* einzubringen, können als Antwort auf diese Abhängigkeit gesehen werden.

Healy und Moody (2014:110) weisen auf Strategien hin, wie »good taste« gesichert werden kann, auch wenn Standardeinstellungen dies nicht leisten können: »expose the mechanics of good practice«. Das heisst, dass sowohl Daten wie auch der Code der Datenvisualisierung offengelegt sind (*Open Source*). Dies trifft für Q als Software zu, während die Offenlegung der Daten den jeweiligen Nutzenden aus der Redaktion überlassen ist. NZZ Storytelling begründet den Entscheid für *Open Source* damit, dass es Software besser mache – und dass man selber sehr oft von *Open-Source*-Software für die eigene Arbeit profitiere.⁷² Das Team stützt sich dabei auf die von den Programmierer*innen geprägte *Open*-Kultur (siehe Abschnitt 3.2.1), welche auch Healy und Moody (2014:122) als Komponente guter Datenanalyse- und Visualisierungspraktiken sehen:

»It seems clear, though, that information visualization tools will become ever more widespread. In keeping with our general argument that good visualization is a component of broader good practice around data analysis, a key issue is the openness of standards and tools for data analysis on the web.«

Bei BR Data ist die Varianz an genutzten *Tools* für Datenvisualisierungen grösser. Es kommen verschiedene Programmiersprachen (bereits angesprochen in Abschnitt 5.4.1), sehr oft aber auch Datawrapper zum Einsatz – die browserbasierte und im Feld weitverbreitete Datenvisualisierungssoftware, die Abschnitt 3.2.2 bereits vorstellt. Auch die NZZ nutzte Datawrapper, bevor sie Q entwickelte.⁷³ Datawrapper selbst ist nur zu einem kleinen Teil *Open Source*. NZZ Storytelling referenziert das *Tool* aber mehrmals aufgrund seiner Deliberation über Entscheidungskriterien in der Konstruktion von Datenvisualisierungen, welche auf dem eigenen Blog veröffentlicht werden.⁷⁴ In der Bewertung eigener Datenvisualisierungen und der Diskussion der Darstellungspraktiken referenziert das Team die Datawrapper-Umsetzungen⁷⁵ und verweist im *Styleguide* auf einen Datawrapper-Blogbeitrag über Kuchendiagramme, um aufzuzeigen, wann und wie diese einzusetzen sind.⁷⁶ Datawrapper gilt in beiden Teams als Referenzpunkt für Best Practices der journalistischen Datenvisualisierung zur Evidenz-Montage – einerseits durch

72 <https://medium.com/nzz-open/lets-build-together-nzz-s-storytelling-toolbox-q-is-now-open-source-39dec92aeca> (10.05.2020)

73 19_20180320_NZZStorytelling-07 (16)

74 <https://blog.datawrapper.de/> (10.05.2020)

75 19_20180320_NZZStorytelling-07 (16); 23_20180326_NZZStorytelling-11 (4)

76 <https://nzzdev.github.io/Storytelling-Styleguide/#/charts-guidelines> (10.05.2020)

die gesetzten Standardeinstellungen und Gestaltungsoptionen, welche die Teams offensichtlich als sinnhaft beurteilen, und andererseits durch seine Erörterungen über das Abwegen und Entscheidungskriterien im Rahmen der Konstruktion von Datenvisualisierungen.

Im Gegensatz zu NZZ Storytelling und dessen umfangreicher Präsentation und Dokumentation von Q erscheint es für auf investigative datenjournalistische Praktiken spezialisierte Teams wie BR Data eigentlich nicht angemessen, die Evidenz der oft aufwendigen und ambitionierten Datenanalysen schliesslich innerhalb von voreingestellten Formatvorlagen zu montieren. So erzählt Martin, dass eine der Datenvisualisierungen in der Geschichte über die Datenschutzerklärungen mittels Datawrapper erstellt wurde – aber dass man das kaum sehe, da das *Tool* inzwischen doch einige Konfigurationsoptionen habe.⁷⁷ Aus dem Methodik-Teil des Artikels ist die Nutzung von Datawrapper nicht ersichtlich. Dies spiegelt die Ansprüche der Teams an die Evidenzbefestigung der eigenen Erkenntnisproduktion durch Daten: Ihr würde eine fein abgestimmte und austarierte Datenvisualisierung mehr gerecht, die eigenen Geschichten aus Daten und den aufwendigen Prozess der Erkenntnisproduktion hält man eigentlich für schlecht montierbar in formalisierten Vorlagen. Die Nutzung von Datawrapper erscheint der Community gegenüber deshalb wenig opportun.

Der Abschnitt über Leitlinien und *Tools* hat gezeigt, wie diese in beiden Teams – vor allem aber bei NZZ Storytelling, welches die Kodifizierung bereits stärker vorangetrieben hat – die Koordinationsmuster zwischen den Expertisen festschreiben. Der *Styleguide* der NZZ dient einerseits dazu, die eigene grafische Expertise dauerhaft in die visuelle Evidenz-Montage einzuschreiben. Andererseits verfolgt das Team damit auch das Ziel, die interne Kohäsion in den Koordinationsmustern zu erhöhen und die immer noch mit dem *Info-grafik-Subteam* verknüpfte Service-Rolle abzustreifen.

Als Eigenentwicklung hat die Koordinationsmuster zwischen grafischer, technischer und journalistischer Expertise in einem *Tool* für die gesamte Redaktion festgeschrieben. BR Data nutzt mit Datawrapper ein ähnliches Werkzeug, welches im datenjournalistischen Feld ein hohes Renommée geniesst. Es zeigen sich in der Anwendung der *Tools* schnell die Grenzen der Festschreibung von Expertisen: Das Montieren einer prägnanten journalistischen Aussage mittels visueller Evidenz lässt sich nicht in Werkzeugen institutionalisieren. Es lässt sich deshalb insbesondere bei der NZZ beobachten, wie die Nut-

77 65_20190710_NZZStorytelling-14

zung von Q mit Schulungen und einem Qualitätssicherungsdienst begleitet wird, um die im Team ausgehandelten datenjournalistischen Qualitätsstandards auf sämtliche Datavisualisierungen in der Redaktion zu übertragen.

6.4 Ausweichen auf neuartige Formate

Beide Teams nutzen für ihre datenjournalistischen Erkenntnisse zum Teil Präsentationsformate, die sie selbst aufgebaut beziehungsweise auf ihre Bedürfnisse zugeschnitten haben: *Explainer* beziehungsweise *Longreads*. Durch diese Nutzung neuartiger, eigener Formate erfolgt eine Koordination der Expertisen ohne die Einpassung in bisherige journalistische Ausspielwege mit ihren spezifischen medialen Evidenzordnungen. Der folgende Abschnitt beleuchtet die beiden Formate, auf welche die beiden Teams ausweichen, weil sich organisationale Bruchlinien mit den verschiedenen Ausgangslagen bezüglich Materialität und epistemologischer Kultur überschneiden.

6.4.1 Der *Explainer*: Ereignisse verknüpfen

Bei NZZ Storytelling ist es der *Explainer* (auf Deutsch Erklärstück, wird aber fast ausschliesslich englisch bezeichnet). Ein *Explainer* beleuchtet ein Thema, das über eine längere Zeit und mit verschiedenen Facetten immer wieder aktuell ist.⁷⁸ Er soll eine Einordnung über längere Zeit vornehmen, indem er verschiedene Ereignisse verknüpft und in Beziehung setzt. David betont, dass es in *Explainern* darum gehe, komplexe Zusammenhänge *anschaulich* zu erklären, wofür sich visuelle Elemente *natürlich* besonders gut eigneten – in *Explainern* kondensierten sich die Vorgaben und Aufgaben von NZZ Storytelling, erklärt er.⁷⁹

78 Drei *Explainer* als Beispiel:

Syrien: Antworten zur Lage im Syrien-Konflikt: <https://www.nzz.ch/international/syrien-antworten-zur-lage-im-syrien-konflikt-ld.1377102> (07.02.2019)

Temperaturanstieg, Treibhauseffekt und die Rolle der Schweiz: Das sind die wichtigsten Fakten zum Klimawandel: <https://www.nzz.ch/wissenschaft/klimawandel-die-wichtigsten-fakten-im-ueberblickld.1420749> (07.02.2019)

Nordkoreas Raketen und ihre Reichweiten im Überblick: <https://www.nzz.ch/international/nordko-re/a/atommaecht-nordkorea-nordkoreas-raketen-und-seine-reichweite-ld.151690> (07.02.2019)

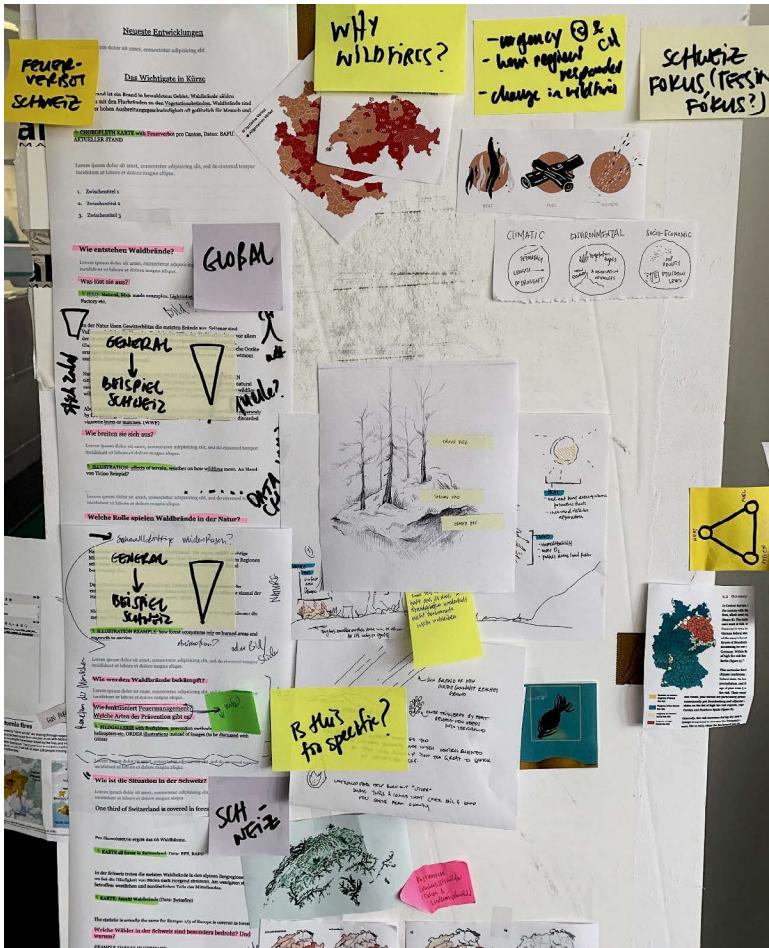
79 24_20180327_NZZStorytelling-12 (18)

Das Verbinden von einzelnen Ereignissen, als Datum materialisiert, und die daraus folgende Ableitung längerfristiger Trends beschrieb Adrian Holovaty in seinem als eine Art Manifest des datenjournalistischen Feldes zu werten- den Beitrag »A fundamental way newspaper sites need to change« von 2006. Das Manifest geniesst deshalb seinen Status, weil Holovaty mit der bisherigen epistemologischen Kultur des Journalismus brach und sie durch eine Daten- bank-Strukturlogik ersetzte. Entsprechend sollte die Auswertung dieser Da- tenbank nicht einzelne Ereignisse, sondern ihre Zusammenhänge aufzeigen:

»For example, say a newspaper has written a story about a local fire. Being able to read that story on a cell phone is fine and dandy. Hooray, technology! But what I *really* want to be able to do is explore the raw facts of that story, one by one, with layers of attribution, and an infrastructure for comparing the details of the fire — date, time, place, victims, fire station number, dis- tance from fire department, names and years experience of firemen on the scene, time it took for firemen to arrive — with the details of previous fires. And subsequent fires, whenever they happen. [Hervorhebung im Original]« (Holovaty 2006)

Die von Holovaty beschriebene Integration unterschiedlicher Ereignisse findet sich in *Explainern* auf verschiedenen Ebenen. Einerseits listet der typische *Explainer* gleich zu Beginn einige Aufzählungspunkte unter dem Titel *Neueste Ent- wicklungen* und schafft damit eine Verbindung zu aktuellen Ereignissen, wel- che andere Artikel beleuchten. Nach weiteren Aufzählungspunkten unter der Überschrift *Das Wichtigste in Kürze* folgt dann ein Inhaltsverzeichnis, wobei die Kapitelüberschriften oftmals Fragen sind. Diese Fragen vernetzen einerseits bestehende Artikel zu einzelnen Ereignissen, andererseits zeigen Datenvisua- lisierungen und ergänzender Beschreibungstext langfristige Entwicklungen. Zum Schluss folgen Abschnitte zu Methodik, Daten und/oder weiterführen- der Literatur. Die Strukturierung zeigt sich beispielhaft in Abbildung 25. NZZ- *Explainer* berufen sich auf eine hohe Glaubwürdigkeit, indem sie sich an wis- senschaftliche Vorgehensweisen und Standards anlehnern, ohne sie aber mit derselben Rigidität zu befolgen.

Abbildung 25: Foto des Konzept-Boards eines NZZ-Explainers zum Thema Waldbrände. Die abgebildete Konzeptionierung zeigt, wie der Explainer versuchen soll, das weltweite Phänomen der Waldbrände zuerst generell und dann die jeweiligen Aspekte aus der Schweizer Perspektive zu beleuchten.



(Foto: NZZ-Storytelling-Teammitglied, veröffentlicht auf Twitter)

Um die in den Ressorts vorhandene journalistische Themenexpertise hoch zu gewichten, ist die inhaltliche Hauptverantwortung für *Explainer* den jeweiligen Ressorts zugeordnet. Dies führt immer wieder zu Diskussionen und Reibungen, beispielsweise an einer Sitzung von NZZ Storytelling, an der die Datenjournalist*innen des Teams das aktuelle Geschehen und mögliche Geschichten darüber diskutieren.⁸⁰ Aufgrund aktueller Ereignisse zur Katalonien-Krise, des Kriegs in Syrien sowie einer politischen Entscheidung zum Brexit drängten sich die Aktualisierung bestehender sowie die Entwicklung eines neuen *Explainers* auf, meint Yannick. David verweist darauf, dass diese in der Verantwortung der Ressorts liegen würden und die Initiative von diesen kommen müsse. Aus der epistemologischen Perspektive der Datenjournalist*innen erscheinen die neu hinzugekommenen Ereignisse erst in Relation mit den bisherigen Ereignissen evident und deshalb eine Aktualisierung bzw. Erstellung eines *Explainers* nötig, während es die zuständigen Ressorts offenbar bei einer Berichterstattung über das einzelne Ereignis belassen. Diese epistemische Unterschiedlichkeit benennt David explizit in einem Tweet, in dem er Datenjournalismus mit dem ›normalen Journalismus‹ als ›Anekdotenjournalismus‹ kontrastiert und so dessen Erkenntnisproduktion in Frage stellt.⁸¹

Die bezüglich Materialität und epistemologischer Kultur abgegrenzte datenjournalistische Evidenz-Montage reibt sich im beschriebenen Team-meeting an der ›klassisch-journalistischen‹ Vorgehensweise der Ressorts – ein ›intra-group struggle‹ innerhalb der journalistischen Profession (Schudson und Anderson 2009:94). »The key aspect of professional struggle, argues Abbott (1988), is the struggle over jurisdiction, or the struggle over the link between knowledge and work« (Schudson und Anderson 2009:94). Bezogen auf das Format des *Explainers* lässt sich festhalten, dass das Nachzeichnen langfristiger Trends mittels Datenvisualisierungen über und das Gruppieren der Berichterstattung rund um ein längerfristig präsenzes Thema eine Sorte der Evidenzbefestigung ist, die den koordinierten Expertisen datenjournalistischer Praktiken folgt. Gleichzeitig gilt in der NZZ-Redaktion, was Schudson und Anderson (2009:95) für die USA festhalten: Journalismus mit Daten bzw. quantitativen Methoden verharrt (bisher) in der Aussenseiterposition, weil der Journalismus insgesamt Erkenntnisse qualitativ produziert: »In Abbott's account, journalism, at least in the United States, has claimed jurisdiction

80 23_20180326_NZZStorytelling-11 (5–8)

81 <https://twitter.com/davidbauer/status/1166229522448080896> (05.02.2020)

over the collection and distribution of qualitative, current information about general events.«

Neben den geschilderten organisationalen und professionskulturellen Reibungen stellt sich die Frage, wie die Veränderungen der Materialität – im Falle des Datenjournalismus digitalisiert verfügbare Daten – und der Epistemologie – quantitative Auswertungen – sich im Zusammenspiel mit der angestrebten Präsentation langfristiger Trends in veränderten Arbeitspraktiken und andersartiger Evidenzbefestigung koordinieren.

»One of the key questions in the use of data for journalistic purposes is whether it is extending the time horizon of reporting, shifting the focus from chronicling the daily or even hourly event in a de-contextualized fashion versus looking at longer trends over time, embedded in a variety of thickly nuanced places and times. Is there, in other words, a relationship between the materiality of the document and the idea of journalism as a record, versus the orality of the interview and the decontextualized notion of journalism as a report?« (Anderson 2015:357).

Andersons Frage, die an viele Überlegungen zur epistemologischen Kultur des Datenjournalismus in Abschnitt 3.2.3 anschliesst, scheint die Empirie dieser Arbeit zu bejahen. Abgesehen vom möglichen Frust darüber, dass der Anstoss dafür aus den Ressorts kommen soll, aber schliesslich doch die Initiative und die meiste Arbeit bei NZZ Storytelling liegt, bewerten mehrere Teammitglieder *Explainer* explizit als äusserst geeignetes Format für ihre Arbeit – weil es sehr eng mit Visualisierungen verbunden ist. *Explainer* sollen »so visuell wie möglich« sein, so David.⁸² Designerin Emma beschreibt das *Explainer*-Format als perfekt, weil es viele Grafiken hat und diese dann mit Text verbindet.⁸³ Datenvisualisierungen wirken darin als relationierendes und koordinierendes Ordnungssystem. NZZ Storytelling hat darauf hingewirkt, dass das Produktionsmanagement-Team für die ganze Redaktion Davids Postulat als Vorgabe institutionalisiert, nämlich dass *Explainer* »so visuell wie möglich« sein sollen, um komplexe Zusammenhänge *anschaulich* zu erklären.⁸⁴

82 23_20180326_NZZStorytelling-11 (8)

83 NZZ_Fragebogen_Mitarbeitende_5

84 24_20180327_NZZStorytelling-12 (18)

6.4.2 Der *Longread*: ein Format für Datenvisualisierungen

Der Bayerische Rundfunk verbreitete seine Informationen wie alle deutschen Landesrundfunkanstalten über Jahrzehnte über die Ausspielkanäle Radio und Fernsehen. Seit 1995 gibt es einen Internetauftritt, seit 2015 versucht BR die Redaktionen nicht mehr nach Sendung beziehungsweise Ausspielweg zu unterscheiden, sondern sie zu verzähnen (siehe Abschnitt zu BR in der Fallbeschreibung unter 2.3.1). Als Folge davon sind Online-Textartikel in der Regel an ein Video oder einen Radio-Ausschnitt gebunden und selten reine Textartikel – und wenn, dann meist kurze. BR Data hingegen schliesst fast alle seine grösseren Projekte mit einem *Longread* ab. Für die *Longreads* weicht das Team auf eine eigene Web-Plattform aus, die es selbst verwaltet.⁸⁵ Martin spricht dabei von »Vertrauen«, das man sich dafür bei der BR-Leitung erarbeitet habe, und dass es einfach technisch notwendig sei, weil das Team mit anderen Werkzeugen arbeite.⁸⁶ Der Verantwortliche im Team, Moritz, erzählt über die eigene Infrastruktur:

Ich spreche [Moritz] darauf an, dass er ja ein eigenes BR-Data-Website-Template aufgebaut habe. Er bestätigt, dies sei, weil es zu wenig Platz gebe und die normale BR-Seite technisch zu wenig biete, zu unruhig sei für lange Geschichten mit Grafiken. Das eigene Template, so die Überlegung, würde lange recherchierte BR-Data-Geschichten auch herausheben, »Premium« aussehen lassen. [Moritz] steht heute aber kritisch zu dieser Strategie: Es mache keinen Sinn, eine zusätzliche Marke aufzubauen. Deshalb versuchten sie sich auch wieder visuell anzunähern, langfristig wieder zu integrieren in das BR-CMS. Derzeit sei dieses aber noch zu unflexibel für die grafischen und interaktiven Formate. Die BR-IT bzw. eher das Management sei derzeit noch zu wenig offen, das CMS entsprechend anzupassen, es sei nicht prioritätär. Grundsätzlich hätten sie etwas eigene IT-Infrastruktur hier aufbauen können, so dass sie Anpassungen in den Daten und Artikeln schnell selber machen könnten. Sie hätten je eigene Server für Applikationen und Ausspielungen.⁸⁷

Die Vorstellung, dass eine eigene datenjournalistische Geschichte vor allem dann »Premium« aussieht, wenn sie als *Longread* und mit funktionierenden,

85 Die URL für den *Longread* beginnt dann mit ›<http://web.br.de/interaktiv>‹

86 52_20181217_BR-Data-06 (7)

87 49_20181212_BR-Data-03 (9)

nichttrivialen Datenvizualisierungen daherkommt, verweist auf deren zentrale Rolle: Die Geschichte muss ästhetischen Kriterien der grafischen Expertise genügen, um Evidenz zu montieren. Zudem ermöglicht die eigene Infrastruktur schnelle Anpassungen – gerade aufgrund der Bestrebungen für Transparenz und Peer-Review (siehe Abschnitt 6.1.3), das zu kritischer Begutachtung einlädt, eine Notwendigkeit.

Gleichzeitig unterliegt BR Data organisationsintern derselben Bewertungsordnung wie andere Teams, die Inhalte für verschiedene Sendungen vorbereiten: Sendezeit am Radio und im Fernsehen zählt.⁸⁸ Das Team muss also einerseits im Laufe jedes Projekts darüber nachdenken, wie und für welche Sendung abstrahierende Datenauswertungen für einen Fernseh- oder Radiobeitrag kurzweilig visualisiert beziehungsweise erzählt werden könnten. Dies bezeichnen die meisten Teammitglieder als grosse Herausforderung, da nur wenige Kompetenzen für Radio- oder Fernseh-Journalismus mitbringen. Deshalb arbeitet man dafür mit Sendungsredaktionen zusammen – wobei man dabei einen beträchtlichen Teil der Kontrolle über die Präsentation, das heisst die Montage der Evidenz, verliert.⁸⁹ Die Nutzung des für BR ungewöhnlichen *Longread*-Formats sowie der Aufbau und die Wartung einer eigenen Infrastruktur dafür verdeutlichen, dass für die Teammitglieder vor allem im *Longread*-Format die Evidenz datenjournalistischer Praktiken befestigt ist.

Somit gilt für beide Teams, dass sie sich nicht nur in ihrer Materialität – digitalen Daten – von der restlichen Redaktion abgrenzen, sondern dass sich daraus folgend epistemologische und professionelle Reibungen ergeben: Etwas evident machen, das bedingt ein Format, welches nichtriviale Datenvizualisierungen mit Text und anderen Illustrationen verbindet. Dafür pflegen beide Teams mit dem *Explainer* bzw. dem *Longread* auf ihre Bedürfnisse der Präsentation angepasste Formate, in welchen sich die koordinierten Expertisen aus der Perspektive der Teams als Evidenz befestigen. Gleichzeitig ist die

-
- 88 So führt die übergeordnete Vorgesetzte des Teams eine Sendezeitstatistik: Eine Tabelle, die minutiös die Zeit in Minuten und Sekunden festhält, während der Projekte von BR Data und BR Recherche in den Sendegefäßen des BR über den Äther liefen (54_20181219_BR-Data-o8 (4))
- 89 So meint Teammitglied Roman einmal, dass man die Social-Media-Video-Redaktion genau über die Aussage einer Analyse briefen und mit mindestens zwei Korrekturschleifen rechnen müsse (48_20181211_BR-Data-o2 (3)).

Arbeit mit diesen Formaten auszubalancieren mit dem Anspruch von organisatorischen Bewertungsordnungen, welche die Präsenz auf klassischen Auspielkanälen der beiden Medienorganisationen bedingt.

6.5 Fazit

Das Kapitel 6 hat empirisch nachgezeichnet, welche datenjournalistischen Arbeitspraktiken den zweiten Modus der Datenvisualisierungen, die Befestigung von Evidenz durch Montage, ausmachen – und welche Aushandlungs- und Koordinationsprozesse sich in diesem Zusammenhang in den beiden untersuchten Teams abspielen.

»Selection of the lanes from several runs of gels which best display the features proposed. These lanes are cut out (some scientists insert an additional marker in the middle of the gel in order to have a lane in which to cut) and glued together« (Amann und Knorr-Cetina 1988:160). Genauso wie die Forscher*innen ihre molekulargenetischen Laborproben selektieren und präparieren, um ihre produzierten Erkenntnisse hervorzu stellen und so Evidenz zu befestigen, wenden die Datenjournalist*innen beider Teams Praktiken zur Hervorhebung der Hauptaussage, Personalisierung und Konkretisierung sowie Transparentmachung und differenzierten Expertisenausweisung an. Der Abschnitt 6.1 hat diese Arbeitspraktiken nachgezeichnet und es lassen sich folgende Schlüsse ziehen, wie die beiden Teams datenjournalistische Erkenntnisse evident befestigen:

1. Die Montage soll durch das gezielte Hervorheben einer Hauptaussage verständlich und eindeutig sein. Die tatsächliche Rezeption beim Publikum bleibt unbekannt. Die Praktiken orientieren sich deshalb an den Erfahrungen der grafischen Expertise, was als eine gute Gestaltung gilt.
2. Die Montage soll als *eine* Geschichte sinnhaft gemacht werden. Dafür bedienen sich beide Teams bei den Praktiken der Personalisierung und der Konkretisierung, zusammengefasst mit dem Begriff ‚humanizing data‘. Die Verbindung verschiedener Abstraktionsebenen durch Datenvisualisierungen verspricht eine verbesserte Zugänglichkeit der Erkenntnisse, indem es auf klassische journalistische Methoden zurückgreift.
3. Im Kontext der Montage weisen die beiden Teams oftmals, aber nicht immer, die Datenquellen sowie die Vorgehensweise aus. Die dadurch erfolgende Exponierung für ein Peer-Review durch das eigene Feld soll das

Vertrauen in die Erkenntnisse fördern – der *Open*-Kultur entsprechend. Gleichzeitig verbergen die Datenvisualisierungen Unsicherheiten und Unschärfen und bleiben semiotisch eindeutig, um nicht ihren Objektivitätsanspruch zu vermindern.

4. Das Ausweisen differenzierter Expertisen in der *Byline* sowie das Zeichnungsrecht als Autor*in für Personen mit grafischer und technischer Expertise verweisen auf den Anspruch, dass journalistische Evidenz nicht nur mit dem Textformat verbunden ist. Diese Arbeitspraktik ist ein Signal, das sich vor allem an die Kolleg*innen organisationintern richtet.

Diese datenjournalistischen Praktiken zur Transformation von Daten in Evidenz handeln die Teams im Spannungsfeld der drei Expertisen aus: der journalistischen, der programmiertechnischen sowie der grafischen. Datenjournalistische Evidenz zu montieren heisst in den beiden Teams in der Regel, dass mittels grafischer Expertise eine Hauptaussage einer Datenvisualisierung hervorgehoben und diese zudem personalisiert oder konkretisiert wird – typische journalistische Vorgehensweisen zur Konstruktion einer für das breite Publikum verständlichen Geschichte. Es werden also Vorstellungen aus den einzelnen Expertisen übernommen, aber in der Koordination mit den anderen abgeschliffen und dadurch koordiniert. Viele Teammitglieder sind in der Lage, mehrere Expertisen zu überbrücken, und sie nehmen dadurch Vermittlungspositionen (*trader*-Rollen) ein, welche die Koordination ermöglichen.

Die Koordination kann einerseits situativ erfolgen, das heisst für jede Datenvisualisierung neu. Oder die entsprechenden Arbeitspraktiken werden in Leitlinien wie dem *Styleguide* oder in Software-Tools wie Q und Datawrapper kodifiziert. Diese verbinden die Arbeitspraktiken mit einer Palette an fixiertem visuellem Vokabular, das hauptsächlich auf bereits sehr bekannten und deshalb für die Sinngebung anschlussfähigen Visualisierungstypen beruht. Und sie sind Festschreibungen der Bewertungen aus der grafischen Expertise, was als eine ›gute Datenvisualisierung‹ gilt.

Im Fall von NZZ Storytelling treten in der Aushandlung der Kodifizierungen immer wieder epistemologische und organisationale Bruchstellen hervor, da die Kodifizierungen auch Absicherungen der eigenen Visualisierungsexpertise gegenüber anderen Ressorts darstellen, gegenüber denen das Team eine Querschnittsfunktion einnimmt.

Beide Teams setzen datenjournalistische Praktiken zur Montage von Evidenz bevorzugterweise in neuartigen Präsentationsformaten um. Diese können sie selbst formen und in ihnen sind die ›klassischen‹, institutionalisier-

ten journalistischen Arbeitsweisen weniger eingeprägt. Im Falle der NZZ ermöglichen *Explainer* das Verbinden von einzelnen Ereignissen und das visuelle Aufzeigen längerfristiger Trends. Bei BR Data sind es *Longreads*, ein für Rundfunk-Anstalten ungewohntes Format, welches komplexe Visualisierungen von Daten, gleichzeitige Konkretisierung, die Transparenz über Datenquellen und Methoden wie auch eine differenzierte Ausweisung der beteiligten Expertisen ermöglicht. Die Nutzung eigener Formate schafft eine Nische für die eigenen datenjournalistischen Evidenzansprüche, während deren Umsetzung in den konventionellen Formaten beider Medienorganisationen immer wieder viel Reibung produziert.

7 Schlussbetrachtung: Mit den Daten sprechen

Während seiner Zeit bei The Pudding bezeichnet sich Charlie Smart als »Journalist-Engineer«¹ – genauso wie die meisten seiner Teamkolleg*innen.² Von The Pudding wechselt Smart dann in die Grafikabteilung der New York Times, wo er »maps and charts« produziert.³

Auch wenn dies Denominationen sind, die einer gewissen Beliebigkeit und Laune unterliegen, so sind sie doch weitere Indizien, dass Datenjournalismus eine technische, eine grafische und eine journalistische Expertise versammelt. Der kurze Blick auf Smarts Berufsbezeichnungen widerspiegelt, was diese Arbeit ausführlich dargelegt hat: Die Produktion und Evidenzbefestigung datenjournalistischer Erkenntnisse basiert darauf, sie im Spannungsfeld dieses Expertisen-Dreiecks auszuhandeln.

Koordinierte Expertisen

Die Nachzeichnung des datenjournalistischen Feldes beginnt bei der Community (Abschnitt 3.1) – eine eng verflochtene, auf persönlichen Beziehungen und informeller Kooperation beruhende, aber gleichzeitig international orientierte Community. Sie teilt einen Raum der geteilten Sinngebung und ist folgerichtig, gemäss Zelizer, eine interpretative Community. Sie versammelt im Feld die drei Expertisen der Visualisierung, der Programmierung und des Journalismus mit ihren je drei Dimensionen der Materialität, der epistemologischen und der professionellen Kultur (Abschnitt 3.2). Die Community erzählt sich im

1 <https://www.linkedin.com/in/charliesmart> (30.01.2021)

2 Dies wird einerseits aus den Kurzbiographien der Teammitglieder offensichtlich, andererseits zählt auch die Lohntabelle nur drei verschiedene mögliche Jobtitel auf: Jr. Journalist-Engineer; Journalist-Engineer; Sr. Journalist-Engineer. <https://pudding.cool/about/#working> (30.01.2021)

3 Gemäss Twitterprofil: https://twitter.com/charlie_smart_ (30.01.2021)

Feld verschiedene Narrative über Datenjournalismus, die der Abschnitt 3.3 erläutert hat. Vor allem aber baut sie auf einer gemeinsamen Sprache auf, welche den geteilten Interpretations- und Erkenntnisraum erst hervorruft. Um aufzuzeigen, wie eine gemeinsame Sprache diesen Raum entstehen lässt und welche Rolle die Materialität der Datenvisualisierungen darin spielt, entwickelt die Arbeit im Folgenden zwei Gerüste: erstens das Dreieck der Expertisen mit der jeweiligen Triade von Materialität, epistemologischer und professioneller Kultur (Abschnitt 3.2.4), und zweitens die Modi des Erkennens von Daten und der Befestigung von Evidenz (*seeing data* und *becoming evident*, Abschnitt 4.2), in denen die Expertisen ausgehandelt werden (Konzept der *trading zone*, Abschnitt 4.1).

Auf Basis dieser Gerüste untersuchen die Kapitel 5 und 6 die datenjournalistischen Praktiken in den zwei Teams NZZ Storytelling und BR Data. Die ethnografisch erhobene Empirie beschreibt das Dreieck der Triaden für die beiden Visualisierungsmodi: Datenvisualisierungen als Arbeitsoberfläche für das Erkennen von Daten (Abbildung 10) sowie Datenvisualisierungen als Montage für die Befestigung von Evidenz (Abbildung 22). Die beiden Dreiecke der Triaden skizzieren die jeweiligen Elemente der Expertisen, welche es in den Praktiken zu koordinieren gilt. Dieser feingranulare Einblick in datenjournalistische Praktiken ermöglicht es, dem Phänomen des Datenjournalismus im redaktionellen Alltag auf den Grund zu gehen. Dabei gelingt es, nicht nur diskursive Aushandlungen, sondern auch die materielle Dimension der Praktiken zu untersuchen – ein soziotechnisches Zusammenspiel, welches die Datenjournalismus-Forschung sonst kaum so profund ausleuchtet und das hier kurz rekapituliert wird.

Kapitel 5 zeigt, wie die Teams von Daten als ontologischer Grundlage der Berichterstattung ausgehen und die Datensätze in das (journalistische) Geschichten-Format übersetzen – meistens, indem sie sich die Daten in ihrer visuellen Verkörperung vorstellen. Gleichzeitig greifen die Teams in der Bewertung von Datenquellen auf die programmiertechnische Expertise zurück: *Saubere Daten* sind granular, systematisch, interoperabel – das ermöglicht gleichzeitig ein Maximum an Optionen für journalistische Geschichten. Der Begriff des Drehens von Geschichten referenziert darauf, dass ein visuelles Vokabular die Variablen durch geometrische Verkörperungen in einer bestimmten Weise relationiert – und diese Darstellungen dann mit journalistischen Relevanzzuschreibungen wie Auffälligkeit, Diskontinuität oder Neuheit verbunden werden. Datenjournalistische Geschichten erfordern zudem von den Teams Praktiken der Plausibilisierung: das Ziehen von Samples, das Beleuchten von Ein-

zelfällen oder spezifische Vergleiche. Diese Methoden entspannen die epistemologischen Differenzen zwischen qualitativen und quantitativen journalistischen Methoden. Die beschriebenen Praktiken handeln die Teammitglieder in Diskussionen aus, oder sie sind bereits in formalisierten Visualisierungssystemen (wie Programmiersprachen oder Software) festgeschrieben. Bestimmte Personen nehmen dabei als *traders* eine koordinierende Rolle ein, wofür sie ihre Multiperspektivität nutzen. Die Teams setzen zudem zur Koordination verschiedene Werkzeuge als Datenvisualisierungs-Arbeitsoberflächen ein: Dazu gehören *computational notebooks*, welche verschiedene materielle Ebenen integrieren, aber auch Papierskizzen, welche die programmiertechnische Materialität gezielt ausschalten, um die grafische mit der journalistischen Perspektive abzustimmen.

Kapitel 6 beschreibt die Praktiken im Modus der Evidenzbefestigung. Die Montage der datenjournalistischen Geschichten muss verständlich und eindeutig, das heisst als *eine* Geschichte sofort erkennbar sein. Dafür orientieren sich die Teams an den Erfahrungswerten der grafischen Expertise. Die journalistischen Methoden der Personalisierung und Konkretisierung verbinden die quantitativen Auswertungen mit einer weniger abstrakten Ebene – der Journalismus erhofft sich dabei eine verbesserte Zugänglichkeit und Identifikation mit der Berichterstattung, schliesslich legitimiert er sich in der Regel durch das Erreichen eines breiten Publikums. Die Teams weisen in ihren Geschichten teilweise Datenquellen und Methoden aus, um das Vertrauen in die Ergebnisse durch Peer-Review zu fördern – dies schliesst an die *Open*-Kultur an. Gleichzeitig bleibt diese Offenheit aber aussen vor, wenn es um das Abbilden von Unschärfen und Unsicherheiten geht. Die *Bylines* der Teams sind vielfältig: Beide weisen ausdifferenzierte Expertisen aus, welche (insbesondere organisationsintern) vermitteln sollen, dass (daten-)journalistische Evidenz nicht an ein Textformat gebunden ist, sondern von grafischer und programmier-technischer Expertise mitproduziert wird. Wie im ersten Modus werden die Praktiken zwischen den drei Expertisen ausgehandelt, oftmals durch Diskussionen – und wiederum nehmen die *traders* eine wichtige Funktion der Koordination ein. In beiden Teams erfolgt eine (auszuhandelnde) Kodifizierung gewisser Praktiken durch Leitlinien und Software, die ein bestimmtes visuelles Vokabular formalisieren – die grafische Expertise nimmt dabei eine grosse Bedeutung ein, indem sie auf besonders anschlussfähige Visualisierungstypen verweist. Beide Teams pflegen wiederum neuartige Präsentationsformate, welche als Nische für die eigenen Evidenzansprüche dienen, um die Reibung

derer in der Übersetzung in die konventionellen Formate der jeweiligen Redaktionen zu umgehen.

Der Anspruch dieser Dissertation geht darüber hinaus, nur die datenjournalistischen Praktiken zweier Teams zu analysieren. Sie verknüpft die Praktiken mit der Feld-Ebene und schliesst aus der Verbindung von Praktiken und Feld auf die Erklärung eines Phänomens (der Ansatz der relationalen Ethnografie, vorgestellt in Abschnitt 2.2.3). Immer wieder verweisen die beiden Teams auf die Community, auf die epistemologische und die professionelle Kultur des Feldes, auf Erfolgsgeschichten anderer Teams, und sie nutzen Werkzeuge, welche Teil des datenjournalistischen Feldes sind oder mindestens darin diskutiert und adaptiert werden.

Die Verbindung der Ebenen, das Zusammenspiel von Expertisen, das Zirkulieren von Narrativen, die geteilte Sinngebung einer Community – all diese Formationen bedürfen eines relationierenden Elementes: einer gemeinsamen Sprache. Eine Sprache ist es auch, welche den Schlüssel zur Auflösung der Forschungsfragen dieser Arbeit bildet: Warum ist Datenjournalismus trotz grosser Diversität eine relevante soziale Kategorie, ein Feld, eben *der* Datenjournalismus? Und welche Rolle spielen Datenvisualisierungen?

Datenjournalismus als Kreol-Sprache

Die Koordinationsprozesse zwischen den Expertisen und ihren Materialitäten, epistemologischen und professionellen Kulturen erfolgt über eine sich entwickelnde »interlanguage«, eine Zwischensprache. Galison (2010:42f) unterscheidet mit Rückgriff auf anthropologische Untersuchungen verschiedene Stadien der Entwicklung einer solchen Sprache: Zuerst handelt es sich um ein »jargon«, später um ein »pidgin«, und die am weitesten entwickelte Form ist ein »creole«. Eine solche Kreol-Sprache stabilisiert sich insbesondere, wenn drei statt nur zwei Expertisen (bei Galison: Subkulturen) sich vereinigen (Galison 1997:837). Abgestützt auf die Aushandlungen und herausgebildet aus dem Expertisen-Dreieck lässt sich schliessen, dass sich Datenjournalismus als Kreol-Sprache so weit stabilisiert hat, dass Institutionalisierungen in Teams und als eigenes Feld (Kapitel 3) stattgefunden haben – gerade weil es drei Expertisen sind und nicht nur zwei.

Betrachtet man Datenjournalismus als Kreol-Sprache, wie funktioniert sie? In ihrem Aufbau widerspiegeln sich die Machtverhältnisse. Galison (1997:50) skizziert in seinem Beispiel, in welchem sich zwei Subkulturen koordinieren, dass die dominante den Wortschatz und die weniger mächtige die Syntax, also den Satzbau, beträgt. Im Falle der datenjournalistischen

Koordinationsprozesse lässt sich für beide Teams festhalten, dass die organisatorische Verortung in einer journalistisch tätigen Redaktion eine dominante Stellung der journalistischen Expertise und ihrer Kultur mit sich bringt: Die Geschichte, die es zu drehen gilt, und die Passendmachung der Daten mit der ›Welt dahinter‹ sind die beiden prägendsten Logiken im Modus der Erkenntnisproduktion. Die grafische Expertise ermöglicht es, Aussagen zu treffen, indem ihr visuelles Vokabular die Daten auf spezifische Weise relationiert – die Syntax. Zu den syntaktischen Elementen gehört aber genauso die materielle Grundlage der Datenbank aus der programmiertechnischen Expertise.

Der Modus der Montage von Evidenz wird ebenfalls dominiert von Begriffen aus der journalistischen Professionskultur: Es geht um das Hervorheben der Hauptaussage, um die Einbettung in eine Geschichte mittels Personalisierung und Konkretisierung, um Transparenzpraktiken – die Teams kodifizieren diese Praktiken mittels Leitlinien und in Datenvisualisierungssoftware, und sie weichen auf spezifische Präsentationsformate aus, welche die Umsetzung ihrer Ideen vereinfachen. Wiederum ist es die grafische Expertise, welche die Syntax liefert, um eine Datenvisualisierung als Geschichte zu montieren. Und es sind die Materialität und kulturelle Vorstellungen aus der programmiertechnischen Expertise, welche Transparenzpraktiken forcieren.

Galison schreibt an einer Stelle von »Image and Logic«, seiner Arbeit über verschiedene Gruppierungen in der Physik, über das Ziel des Buches:

»By one reading, then, this book aims to show how two subcultures of instrumentation formed a wordless creole: On one side is the image tradition of photographic practices aiming at noninterventionist objectivity – producing homomorphic representations of nature. On the other side is the logic tradition of electronic practices aiming at manipulative persuasion – producing homologous representations of nature. Out of the two, I will argue historically, sociologically, and philosophically, came a wordless pidgin that evolved into a wordless creole: electronically produced images, homologous in source and homomorphic in presentation, a powerful system in which generations of physicists could then ›grow up‹.« (Galison 1997:51)

Sind Datenvisualisierungen das »wordless creole« des datenjournalistischen Expertisen-Dreiecks? Die Kapitel 5 und 6 haben gezeigt, dass Datenvisualisierungen in den Modi der Erkenntnisproduktion wie auch der Evidenzbefestigung eine strukturierende Rolle in der Sprache einnehmen, indem sie es ermöglichen, mit den Daten zu sprechen – einerseits im Sin-

ne von *conversations with data*, andererseits im Sinne einer Kommunikation von Erkenntnissen *mittels* Daten. Datenvisualisierungen sind die materielle Grundlage der gemeinsamen Sprache der Community von Journalist*innen, Programmierer*innen und Designer*innen in den datenjournalistischen Praktiken, mitgeformt durch die technische Infrastruktur der Datenbanken, Programmiersprachen und *Tools*. Datenvisualisierungen ermöglichen erst das Sprechen mit den Daten, in einem doppelten Sinne.

Das Kreol als wortlos zu apostrophieren ginge aber zu weit: Die zahlreichen Aushandlungs- und Interpretationssituationen mit längeren Diskussionssequenzen, aber auch die Wichtigkeit der Papierskizzen als Werkzeug verweisen auf die heterogenen Konstellationen von Materialität, Erkenntnisinstrumenten und Professionskultur aus den drei Expertisen. Die Konstellationen dieser Elemente gilt es immer wieder aufs Neue zu koordinieren mit Aussagen – gesprochenen und visuellen –, Werkzeugen – computer- und papiergestützt –, mithilfe der Imitation von Ideen anderer beziehungsweise neuartiger Präsentationsformate.

In diesen Aushandlungssituationen nehmen die *traders* eine besonders wichtige Rolle ein, weil sie die Expertisen zu koordinieren vermögen (siehe Abschnitte 5.2.2 und 6.2.2). Sie verbinden durch ihre Kompetenzen verschiedene Expertisen in persona und vermögen durch ihre Multiperspektivität neue Assoziationen zu knüpfen, neue Erkenntnisse zu schliessen. Die *traders* sind Schlüsselfiguren, weil sie die versiertesten Sprecher*innen der datenjournalistischen Sprache sind – sowohl wenn sie Erkenntnisse produzieren, indem sie mit den Daten sprechen, als auch wenn sie mittels der Daten relevante journalistische Evidenz aussprechen. Im Idealtypus heisst dies: Sie verfügen über eine datafizierte Vorstellung der Welt, die sie in Geschichten übersetzen können; sie greifen auf ein breites visuelles Vokabular zurück, das ihnen differenzierte visuelle Aussagen ermöglicht; und sie vermögen die datenbasierten, visuellen Geschichten nach den Erfordernissen der journalistischen Professionskultur und ihren Relevanzkriterien auszurichten. Es ist somit kein Zufall, sondern eine schlüssige Folge, dass diese Personen sowohl in den Teams wie auch im Feld zentrale Rollen einnehmen und ein hohes Renommee geniessen. So lässt sich auch für Charlie Smart und das Pudding-Team festhalten: Wie kaum ein anderes Team vermögen sie »datenjournalistisch zu sprechen«. Sie vermögen die Expertisen zu vereinigen, oftmals sogar in einer Person.

Wie bereits in den untersuchten Teams zeigt sich auch bei The Pudding die strukturierende Rolle von Datenvisualisierungen: Die Teammitglieder stel-

len sich ihre Geschichten bereits im Ideenstadium visuell vor.⁴ Das heisst, in ihren *visual essays* nutzen sie Datenvisualisierungen nicht nur, um errechnete Resultate zu präsentieren, sondern stellen von Anfang an die Frage, wie die Geschichte visuell effektiv kommuniziert werden kann. Dies verweist erneut darauf, wie wirkungsvoll Visualisierungen als Kommunikationsmedium sein können – und dass ihre Macht über die mittels der Daten projizierten Informationen hinausgeht: »Bilder synthetisieren Wissen, verräumlichen zeitliche Entwicklungen, verleihen ihnen eine topologische Struktur, überbrücken Sprünge und nichtstetige Stellen und eröffnen auf diese Weise Plausibilitäten, die eher der Bildlichkeit selbst zuzurechnen sind, als z.B. den durch sie dargestellten Messergebnissen« (Heßler und Mersch 2009:17). Oder kürzer und mit dem Verweis auf die Machtdimension gefasst: »The ›great man‹ is a little man looking at a good map« (Latour 2012:32).

Einordnung der Ergebnisse

Aus ethnografischen Praxisstudien gewonnenen Erkenntnissen haftet stets die Frage nach ihrer Generalisierbarkeit an – das ist in diesem Falle nicht anders. Trotzdem hat der breite Blick auf das gesamte datenjournalistische Feld die in den Teampraktiken beobachteten soziotechnischen Muster wiedergespiegelt. Der Ansatz der relationalen Ethnografie hat sogar explizit nach einer Erklärung aus der Verbindung der Ebenen gefragt, was die Arbeit mit der Figur der Sprache beantwortet hat. Auch wenn gewisse Elemente, beispielsweise die eng verflochtene Community, möglicherweise in anderen Feldern so nicht zu beobachten sind: Die vorliegende Dissertation bietet theoretische Überlegungen und Schlüsse an, welche für die Disziplinen der Soziologie, der STS, der Journalismusforschung wie auch der Medien- und Kommunikationswissenschaft interessant sind.

Mit der Materialität der Datenbanken, das hat diese Arbeit gezeigt, sind immer auch spezifische Aspekte einer epistemologischen und professionellen Kultur verbunden, welche der Koordination mit bestehenden Elementen bedürfen. Die Erkenntnisse dieser Arbeit können dazu beitragen, die weitergehenden Implikationen der soziologischen Diagnose einer zunehmend datafizierten und verrechneten Welt abzuschätzen: die Stärkung einer *Open-Kultur*, welche mit dem Bedürfnis nach Teilen, Bastelei, Transparenz und Experimentieren einhergeht, wie auch das Ethos des kompetitiv-kollaborativen Lernens und das Betonen von Teamarbeit. Gleichzeitig die gewichtige Stellung

von Tool-Entwickler*innen und Designer*innen, welche mit ihren Communitys computergestützte Werkzeuge weiterentwickeln und deren Möglichkeitsraum stetig neu abstecken.

Die Untersuchung von Visualisierungen als Wissensträger interessiert die STS seit jeher. Die Produktion datenjournalistischer Erkenntnisse zeigt, dass in digitalen Welten die Kategorisierung und die Geometrisierung zunehmend wichtige Grundoperationen des Wissens sind. Das visuelle Vokabular bietet bestimmte Relationierungen als Konstruktionsmöglichkeiten von Wissen an und schliesst andere aus. Das Drehen datenjournalistischer Geschichten verdeutlicht, dass je nach Feldlogik gewisse Relationierungen besonders anschlussfähig sind, andere nicht. Visualisierungswerzeuge ergänzen die explorative Datenanalyse und ermöglichen das Ausprobieren verschiedenster Visualisierungsoptionen in kurzer Zeit – sie erweitern damit das empiriegetriebene *data mining* (Anderson 2008) um eine visuelle Deutungsebene.

Für die Journalismusforschung sind die materiellen, epistemologischen und kulturellen Konstanten beziehungsweise Weiterentwicklungen der journalistischen Profession durch das Feld des Datenjournalismus, wie es diese Arbeit beschreibt, abtastbar. Konventionen der Transparenz haben sich ausgeweitet, das Verlangen nach Reproduzierbarkeit der eigenen Erkenntnisse und damit die Anlehnung an wissenschaftliche Standards ertönt aus dem Feld des Datenjournalismus. Dies schliesst auch an akademische und praktische Diskussionen rund um das Vertrauen in journalistische Institutionen an. Datenjournalist*innen grenzen sich explizit ab vom Bild des einsamen journalistischen Wolfes und betonen stattdessen den Wert von Teamarbeit – die *Bylines* werden vielfältiger.

Für die Medien- und Kommunikationswissenschaft sind insbesondere die Ergebnisse zur Evidenzbefestigung interessant. Datenvisualisierungen mit ihrer Grundlogik des Zeigens, ihrem inhärenten Evidenzcharakter (was man sieht, ist da) und ihrem Verbergen von Nuancen und Unsicherheiten sind zu einer mächtigen Kommunikationsform geworden, welche sich nicht nur der Datenjournalismus aneignet. Die Überzeugungskraft der Bildlichkeit, aber genauso das kritische Lesen und Hinterfragen der Konstruktion von Aussagen sind nicht nur für das journalistische Feld wichtige Themen, sondern zu führende Debatten in den Medien- und Kommunikationswissenschaften.

Offene Fragen

Gerade mit Bezug auf den letztgenannten Aspekt ist festzuhalten, dass die visuelle Sprache auf Voraussetzungen und Expertisen basiert, deren Aneignung einigen Aufwands bedarf:

»Der Betrachter muss also die jeweilige ›Sprache‹ kennen; er muss wissen, welches Wissen und welches Verständnis transportiert werden, sodass die epistemische Plausibilität von Visualisierungen ganz wesentlich vom Rezipienten sowie vom Wissen darüber abhängt, dass die Abbildung ein stilisiertes Zeichen darstellt. Anders ausgedrückt: Bildprozesse erweisen sich als ausgesprochen voraussetzungsreich, *ohne dass jedoch das Bild die Möglichkeit bietet, dies zu thematisieren*. Die Lesbarkeit beruht auf der Etablierung einer gemeinsamen Praxis, eines geteilten visuellen Codes, der Übereinstimmung über die Verwendung von Formen, Farben usw. [Hervorhebung im Original]« (Heßler und Mersch 2009:46)

Es liegt auf der Hand, dass das Sprechen der datenjournalistischen Sprache das Eintauchen in die Community und die Expertisen des Feldes bedingt. Eine weitere, nicht zu unterschätzende Schlussfolgerung daraus ist, und darauf weisen Heßler und Mersch hin: Auch das Publikum muss über eine gewisse Sprachfähigkeit verfügen, um mit Datenvisualisierungen und den darin verkörperten Erkenntnissen überhaupt in einen Diskurs treten zu können.

Rezeptionsprozesse auf der Seite des Publikums hat diese Arbeit aussen vor gelassen. Trotzdem erscheint es klar, dass es aufgrund der Weiterentwicklung und Weiterverbreitung der datenjournalistischen Sprache, aber vor allem auch aufgrund der zunehmenden Nutzung von Datenvisualisierungen in vielen anderen Bereichen nötig ist, an die Förderung der »data visualization literacy« (Börner et al. 2019) beziehungsweise der »graphicacy« (Balchin 1972; Cairo 2016) zu appellieren. Eine datafizierte Welt wird immer auch eine (daten-)visualisierte Welt sein, und die entsprechende Lesekompetenz wird zum Kompass darin.

Ein weiterer offener Punkt bleibt die Frage, ob bestimmte Typen von Visualisierungen, welche Daten in bestimmter Weise relationieren, im datenjournalistischen Feld besonders oft vorkommen⁵ – was ein Indiz für eine besondere ausgeprägte Wirkung wäre.

5 Der englische Datenjournalismus-Professor Paul Bradshaw kam aufgrund einer Analyse von 100 Datenjournalismus-Publikationen zum Schluss, dass die folgenden sieben »data story angles« dominieren: »Scale, Change, Ranking, Variation, Explore, Relati-

Ebenso bleibt die Frage nach dem Projektmanagement als weiterer Expertise ein Desiderat dieser Arbeit. Es lassen sich sowohl in der Empirie aus dem Feld wie auch aus den beiden Teams Hinweise dazu finden, dass die Heuristik des Projekts eine organisierende Rolle einnimmt. Die Koordination verschiedener Expertisen, welche datenjournalistische Praktiken erfordert, sowie die benötigte Zeitdauer für die datenjournalistische Berichterstattung – meistens Tage bis Wochen oder Monate – führen zu einer Projekt-Strukturierung in den beiden untersuchten Teams. Es handelt sich um eine in der ganzen Community verbreitete Ontologie. Als Folge davon taucht sporadisch eine vierte Expertise auf, welche für Datenjournalismus benötigt wird: Projektmanagement.⁶

Die Organisationssoziologie hat Projekte als Arbeitseinheiten beschrieben, in denen sich durch mehrere Beteiligte verschiedene Expertisen und Identitäten versammeln – dadurch werden Projekte zu »sites of contention« (Stark 2009:33f) und damit zu wichtigen Elementen der organisierten Dissonanz, welche Kreativität und Innovation ermöglicht. Grabher verbindet die Arbeitseinheit Projekt mit den Eigenschaften einer damit verbundenen Aufgabe: »[...] a project is based on a particular *task* that either might be complex and non routine or rather standardized [Hervorhebung im Original]«, dies insbesondere durch interdependente Expertisen, welche für die Entwicklung einer Lösung keine *ex-ante*-Aufteilung der Aufgabe auf die Beteiligten ermöglichen (Grabher 2002:207f).

So verweist die Projekt-Ontologie im Datenjournalismus darauf, dass das Zusammenspiel der verschiedenartigen Expertisen die heuristische Perspektive des Feldes auf datenjournalistische Praktiken darstellt. Dies illustriert eine Aussage eines Datenjournalisten, der Überlegungen anstellt über den Inhalt einer datenjournalistischen Ausbildung:

»Du hast die investigativen Leute, du hast die eh... Designer, du hast, du hast Datenrecherche. Und oft, ehm... Es wäre eigentlich ideal, wenn du jemanden hättest, der dann... so das Projektmanagement übernimmt, von so einem komplexen Projekt. Und das ist dann oft, also ad hoc wird das bestimmt, obwohl, also, obwohl es da eigentlich rechte, rechte Fähigkeiten braucht, und das Verständnis dafür: Was leisten denn die investigativen genau, was macht

onships, Bad/open«. Sein methodisches Vorgehen wie auch Überlegungen zur Kategorienbildung bleiben aber intransparent, seine Stichprobe wenig repräsentativ.
<https://gijn.org/2020/08/18/from-relationships-to-ranking-angles-for-your-next-data-story/> (30.01.2021)

6 Interview D6; Interview D9; Interview D13

der Web-Designer genau, was, was, und was machen denn die, die dieses Zeugs aufbereiten genau.«⁷

Zwei Ausbildungsleitende aus der Schweiz beziehungsweise den USA betonen, dass Projektmanagement für das datenjournalistische Curriculum eine sinnvolle Ergänzung ist⁸ – wobei es in beiden Ausbildungen trotzdem nur insofern Teil ist, als die Studierenden eigene Projekte ausarbeiten, die aber nicht Teamarbeit umfassen.

Die Projekt-Heuristik ist wertvoll, weil sie auf einer zusätzlichen Ebene, der organisatorischen, darauf verweist, dass Datenjournalismus nur als Zusammenspiel verschiedenartiger Expertisen gesehen werden kann. Trotz der kurzen Einordnung an dieser Stelle verbleibt ein grosses Desiderat: Eine weitere Untersuchung der Bedeutung von Projektmanagement, vielleicht sogar deren Kategorisierung als vierte Expertise, kann diese Arbeit nicht vornehmen. Es hätte einen andersartigen Beobachtungsfokus der ethnografischen Praxisstudien bedingt, welcher nicht im Nachhinein zu implementieren war.

Bier, Chips, Erfolgsgeschichten

Es wird interessant zu beobachten sein, ob sich Datenjournalismus als geteilter Raum der Sinngebung, als Sprache einer Community, als Feld weiter institutionalisiert. Oder ob die zentrifugalen Kräfte verschiedener Materialitäten, epistemologischer und professioneller Kulturen sowie unterschiedlicher Narrative nicht zu einer Ausdifferenzierung oder zu einer Zersplitterung des Feldes und schliesslich zur Auflösung der sozialen Kategorie des Datenjournalismus führen. Ohne Zweifel spielt es in der künftigen Entwicklung eine grosse Rolle, ob sich die Community weiterhin so eng zu verflechten vermag. Ob sie Bier und Chips und Expertisen freimütiig unter dem Hacks/Hackers-Dach teilt – und dabei den gemeinsamen Raum der datenjournalistischen Sprache aufspannt, um sich Erfolgsgeschichten zu erzählen.

7 Interview D6

8 Interview D9; Interview D13

8 Literaturverzeichnis

- Abbott, Andrew (1988): *The System of Professions. An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.
- Abbott, Andrew (2004): *Methods of Discovery. Heuristics for the Social Sciences*. New York: W. W. Norton & Company.
- Aiello, Giorgia (2020): Inventorizing, situating, transforming. In: Martin Engebretsen und Helen Kennedy (Hg.): *Data Visualization in Society*: Amsterdam University Press, S. 49–62.
- Alamalhodaei, Aria; Alberda, Alexandra P.; Feigenbaum, Anna (2020): Humanizing data through >data comics<. In: Martin Engebretsen und Helen Kennedy (Hg.): *Data Visualization in Society*: Amsterdam University Press, S. 347–366.
- Allen, David S. (2008): The Trouble with Transparency. The challenge of doing journalism ethics in a surveillance society. In: *Journalism Studies* 9 (3), S. 323–340.
- Amann, Klaus; Knorr-Cetina, Karin (1988): The Fixation of (Visual) Evidence. In: *Human Studies* 11 (2/3), S. 133–169.
- Amit, Vered (Hg.) (2000): *Constructing the Field. Ethnographic Fieldwork in the Contemporary World*. London: Routledge.
- Ananny, Mike (2012): Press-Public Collaboration as Infrastructure: Tracing News Organizations and Programming Publics in Application Programming Interfaces. In: *American Behavioral Scientist* 57 (5), S. 623–642.
- Anderson, C. W. (2013a): What aggregators do: Towards a networked concept of journalistic expertise in the digital age. In: *Journalism* 14 (8), S. 1008–1023.
- Anderson, C.W. (2013b): *Rebuilding the News: Metropolitan Journalism in the Digital Age*. Philadelphia: Temple University Press.
- Anderson, C.W. (2015): Between the Unique and the Pattern. In: *Digital Journalism* 3 (3), S. 349–363.

- Anderson, C.W. (2018): Apostles of Certainty. Data Journalism and the Politics of Doubt. New York: Oxford University Press.
- Anderson, C.W.; De Maeyer, Juliette (2015): Objects of journalism and the news. In: *Journalism* 16 (1), S. 3–9.
- Anderson, Chris (2008): The end of theory. The data deluge makes the scientific method obsolete. *Wired Magazine*. <https://www.wired.com/2008/06/pb-theory/> (31.01.2021).
- Ausserhofer, Julian; Gutounig, Robert; Oppermann, Michael; Matiasek, Sarah; Goldgruber, Eva (2017): The Datafication of Data Journalism Scholarship: Focal Points, Methods and Research Propositions for the Investigation of Data-intensive Newswork. In: *Journalism*, S. 1–24.
- Baack, Stefan (2018): Practically Engaged. The entanglements between data journalism and civic tech. In: *Digital Journalism* 6 (6), S. 673–692.
- Balchin, W. G. V. (1972): Graphicacy. In: *Geography* 57 (3), S. 185–195.
- Barley, Stephen R. (1986): Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments. In: *Administrative Science Quarterly* 31 (1), S. 78–108.
- Barley, Stephen R. (1996): Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organizational Studies. In: *Administrative Science Quarterly* 41 (3), S. 404–441.
- Barlow, Mike (2015): Data visualization. A new language for storytelling: Sebastopol: O'Reilly Media Inc.
- Bauer, Matthias; Ernst, Christoph (2010): Diagrammatik. Einführung in ein kultur- und medienwissenschaftliches Forschungsfeld. Bielefeld: transcript Verlag (Kultur- und Medientheorie).
- Beckman, R. J.; Cook, R. D. (1983): Outlier.....s. In: *American Society for Quality* 25 (2), S. 119–149.
- Bertin, Jacques (1974): Graphische Semiolegie. Diagramme, Netze, Karten. Berlin: de Gruyter Verlag.
- Boczkowski, Pablo J. (2004): Digitizing the News. Innovation in Online Newspaper. Cambridge: MIT Press.
- Borges-Rey, Eddy (2016): Unravelling Data Journalism. In: *Journalism Practice* 10 (7), S. 833–843.
- Borges-Rey, Eddy (2017): Towards an epistemology of data journalism in the devolved nations of the United Kingdom: Changes and continuities in materiality, performativity and reflexivity. In: *Journalism*, 1–18.
- Börner, Katy; Bueckle, Andreas; Ginda, Michael (2019): Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. In:

- Proceedings of the National Academy of Sciences of the United States of America* 116 (6), S. 1857–1864.
- Bowker, Geoffrey C.; Star, Susan Leigh (1999): *Sorting Things Out. Classification and its Consequences*. Cambridge: MIT Press.
- boyd, danah; Crawford, Kate (2012): Critical questions for Big Data. Provocations for a cultural, technological, and scholarly phenomenon. In: *Information, Communication & Society* 15 (5), S. 662–679.
- Bradshaw, Paul (2018): Data Journalism Teaching, Fast and Slow. In: *Asia Pacific Media Educator* 28 (1), S. 1–12.
- Breda, Adrian (2019): Objektive Diagramme gegen die Vertrauens- und Finanzierungskrise? Inhaltsanalyse datenjournalistischer Visualisierungen. Masterarbeit. Universität Leipzig, Leipzig. Institut für Kommunikations- und Medienwissenschaft. https://www.dropbox.com/s/6kf9xqedj7oewoc/DDJ-MA_Objektivit%C3%A4t_Adrian%20Breda.pdf.
- Brooker, Phillip; Sharrock, Wes; Greiffenhagen, Christian (2018): Programming Visuals, Visualising Programs. In: *Science & Technology Studies* 32 (1), S. 21–42.
- Brown, John Seely; Collins, Allan; Duguid, Paul (1989): Situated Cognition and the Culture of Learning. In: *Educational Researcher* 18 (1), S. 32.
- Brown, John Seely; Duguid, Paul (1991): Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation. In: *Organization Science* 2 (1), S. 40–57.
- Burkhardt, Marcus (2015): Digitale Datenbanken. Eine Medientheorie im Zeitalter von Big Data. s.l.: transcript Verlag (Digitale Gesellschaft, 7).
- Burrell, Jenna (2009): The Field Site as a Network: A Strategy for Locating Ethnographic Research. In: *Field Methods* 21 (2), S. 181–199.
- Burri, Regula Valérie (2001): Doing Images. Zur soziotechnischen Fabrikation visueller Erkenntnis in der Medizin. In: Bettina Heintz und Jörg Huber (Hg.): *Mit dem Auge denken. Strategien der Sichtbarmachung in wissenschaftlichen und virtuellen Welten*. Zürich: Edition Voldemeer, S. 277–303.
- Burri, Regula Valérie (2008): Bilder als soziale Praxis: Grundlegungen einer Soziologie des Visuellen. In: *Zeitschrift für Soziologie* 37 (4), S. 342–358.
- Burri, Regula Valérie; Dumit, Joseph (2008): Social Studies of Scientific Imaging and Visualization. In: Edward J. Hackett, Olga Amsterdamska, Michael Lynch und Judy Wajcman (Hg.): *The Handbook of Science and Technology Studies*. Third Edition. Cambridge, London: MIT Press, S. 297–317.
- Cairo, Alberto (2016): *The Truthful Art. Data, Charts, and Maps for Communication*. Thousand Oaks, CA: New Riders.

- Cairo, Alberto (2019): How Charts Lie. Getting Smarter about Visual Information. New York: W. W. Norton & Company.
- Carlson, Matt (2015): Introduction. The many boundaries of journalism. In: Matt Carlson und Seth C. Lewis (Hg.): *Boundaries of Journalism. Professionalism, Practices and Participation*. New York: Routledge, S. 1–18.
- Carlson, Matt (2018): Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism. In: *new media & society* 20 (5), S. 1755–1772.
- Christin, Angèle (2018): Counting Clicks: Quantification and Variation in Web Journalism in the United States and France. In: *American Journal of Sociology* 123 (5), S. 1382–1415.
- Coddington, Mark (2015): Clarifying Journalism's Quantitative Turn. In: *Digital Journalism* 3 (3), S. 331–348.
- Coddington, Mark (2018): Gathering evidence of evidence: News aggregation as an epistemological practice. In: *Journalism* 54 (4), 146488491881760.
- Coddington, Mark (2019): Defining and mapping data journalism and computational journalism. A review of typologies and themes. In: Scott A. Eldridge II und Bob Franklin (Hg.): *The Routledge Handbook of Developments in Digital Journalism Studies*. London and New York: Routledge, S. 225–236.
- Collins, Harry; Evans, Robert; Gorman, Mike (2007): Trading zones and interactional expertise. In: *Studies in History and Philosophy of Science* 38 (4), S. 657–666.
- Conrad, Lisa (2017): Organisation im soziotechnischen Gemenge. Mediale Um- schichtungen durch die Einführung von SAP. Bielefeld: transcript Verlag.
- Cui, Weiwei; Zhang, Xiaoyu; Wang, Yun; Huang, He; Chen, Bei; Fang, Lei et al. (2019): Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements. Microsoft Research Asia.
- Czarniawska, Barbara (2007): Shadowing and Other Techniques for Doing Fieldwork in Modern Societies. Malmö/Kopenhagen: Liber/CBS Press.
- Czarniawska, Barbara (2014): Social Science Research. From Field to Desk. London: Sage.
- Czarniawska, Barbara; Joerges, Bernward (1996): Travels of Ideas. In: Barbara Czarniawska und Guie Sevón (Hg.): *Translating Organizational Change*. Berlin, New York: Walter de Gruyter, S. 13–48.
- Daston, Lorraine J.; Galison, Peter (2007): Objektivität. Frankfurt a.M.: Suhrkamp Verlag.

- de Haan, Yael; Kruikemeier, Sanne; Lecheler, Sophie; Smit, Gerard; van der Nat, Renee (2017): When Does an Infographic Say More Than a Thousand Words? In: *Journalism Studies* 19 (9), S. 1293–1312.
- Desmond, Matthew (2014): Relational ethnography. In: *Theor Soc* 43 (5), S. 547–579.
- Desrosières, Alain (1998): The Politics of Large Numbers. A History of Statistical Reasoning. Cambridge, MA: Harvard University Press.
- Dewey, John (1927): The Public and Its Problems. New York: Holt.
- Diakopoulos, Nicholas (2015): Algorithmic Accountability. In: *Digital Journalism* 3 (3), S. 398–415.
- Diakopoulos, Nicholas; Koliska, Michael (2016): Algorithmic Transparency in the News Media. In: *Digital Journalism*, S. 1–20.
- Dick, Murray (2014): Interactive Infographics and News Values. In: *Digital Journalism* 2 (4), S. 490–506.
- Emerson, Robert M.; Fretz, Rachel I.; Shaw, Linda L. (2011): Writing Ethnographic Fieldnotes. 2. Aufl. Chicago: University of Chicago Press.
- Engebretsen, Martin; Kennedy, Helen; Weber, Wibke (2018): Data Visualization in Scandinavian Newsrooms. Emerging Trends in Journalistic Visualization Practices. In: *Nordicom Review*.
- Espeland, Wendy Nelson (2015): Narrating numbers. In: Richard Rottenburg (Hg.): The world of indicators. The making of governmental knowledge through quantification. Cambridge: Cambridge University Press (Cambridge studies in law and society), S. 56–75.
- Espeland, Wendy Nelson; Sauder, Michael (2007): Rankings and Reactivity: How Public Measures Recreate Social Worlds. In: *American Journal of Sociology* 113 (1), S. 1–40.
- Espeland, Wendy Nelson; Stevens, Mitchell L. (2008): A Sociology of Quantification. In: *Arch. eur. sociol.* 49 (3), S. 401–436.
- Ettema, James S.; Glasser, Theodore L. (1985): On the epistemology of investigative journalism. In: *Communication* 8, S. 183–206.
- Few, Stephen (2016): Data Visualization for Human Perception. In: The Encyclopedia of Human-Computer Interaction. 2nd Edition.
- Fleck, Ludwik (1983): Schauen, sehen, wissen. In: Ludwik Fleck (Hg.): Erfahrung und Tatsache. Gesammelte Aufsätze. Frankfurt a.M.: Suhrkamp, S. 147–174.
- Flyverbom, Mikkel; Madsen, Anders Koed (2015): Sorting data out. Unpacking big data value chains and algorithmic knowledge production. In: Florian

- Süssenguth (Hg.): Die Gesellschaft der Daten. Über die digitale Transformation der sozialen Ordnung. Bielefeld: transcript Verlag, S. 123–144.
- Galison, Peter (1997): Image and Logic. A Material Culture of Microphysics. Chicago: University of Chicago Press.
- Galison, Peter (2010): Trading with the Enemy. In: Michael E. Gorman (Hg.): Trading Zones and Interactional Expertise. Creating New Kinds of Collaboration. Cambridge, MA: MIT Press (Inside technology), S. 25–52.
- Galtung, Johan; Ruge, Mari Holmboe (1965): The Structure of Foreign News. The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. In: *Journal of Peace Research* 2 (1), S. 64–91.
- Gans, Herbert J. (1980): Deciding what's news. A study of CBS Evening News, NBC Nightly News, Newsweek, and Time. New York: Vintage Books.
- Garfinkel, Harold; Rawls, Anne Warfield (2002): Ethnomethodology's program. Working out Durkeim's aphorism. Lanham, Md: Rowman & Littlefield Publishers (Legacies of social thought).
- Gelman, Andrew; Unwin, Antony (2013): Infovis and Statistical Graphics: Different Goals, Different Looks. In: *Journal of Computational and Graphical Statistics* 22 (1), S. 2–28.
- Google News Lab (2017): Data Journalism in 2017. The current state and challenges facing the field today. Unter Mitarbeit von Simon Rogers, Jonathan Schwabish und Danielle Bowers. Google News Lab.
- Grabher, Gernot (2002): Cool Projects, Boring Institutions: Temporary Collaboration in Social Context. *Regional Studies*, 36(3), 205–214. In: *Regional Studies* 36 (3), S. 205–214.
- Gray, Jonathan (2018): Three Aspects of Data Worlds. In: *Krisis Journal for contemporary philosophy* (1), S. 4–17.
- Gray, Jonathan; Chambers, Lucy; Bounegru, Liliana (2012): The data journalism handbook. How journalists can use data to improve the news: O'Reilly Media, Inc.
- Gynnild, Astrid (2013): Journalism innovation leads to innovation journalism. The impact of computational exploration on changing mindsets. In: *Journalism* 15 (6), S. 713–730.
- Halford, Susan; Savage, Mike (2017): Speaking Sociologically with Big Data: Symphonic Social Science and the Future for Big Data Research. In: *Sociology* 51 (6), S. 1132–1148.
- Halpern, Orit (2014): Beautiful Data. A History of Vision and Reason since 1945. Durham and London: Duke University Press.

- Hamilton, James F.; Turner, Fred (2009): Accountability Through Algorithm: Developing the Field of Computational Journalism. Stanford University. Stanford.
- Hanitzsch, Thomas (2007): Deconstructing Journalism Culture: Toward a Universal Theory. In: *Communication Theory* 17 (4), S. 367–385.
- Hannaford, Liz (2015): Computational Journalism in the UK newsroom: hybrids or specialists? In: *Journalism Education* (4), S. 6–21.
- Hannerz, Ulf (2003): Being there... and there... and there! In: *Ethnography* 4 (2), S. 201–216.
- Healy, Kieran; Moody, James (2014): Data Visualization in Sociology. In: *Annual Review of Sociology* 40, S. 105–128.
- Heintz, Bettina (2010): Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. In: *Zeitschrift für Soziologie* 39 (3), S. 162–181.
- Heintz, Bettina; Huber, Jörg (2001): Der verführerische Blick: Formen und Folgen wissenschaftlicher Visualisierungsstrategien. In: Bettina Heintz und Jörg Huber (Hg.): Mit dem Auge denken. Strategien der Sichtbarmachung in wissenschaftlichen und virtuellen Welten. Zürich: Edition Voldemeer, S. 9–40.
- Heravi, Bahareh R. (2018): Data Journalism in 2017: A Summary of Results from the Global Data Journalism Survey. In: Gobinda Chowdhury, Julie McLeod, Val Gillet und Peter Willett (Hg.): *Transforming Digital Worlds*. iConference 2018. Cham: Springer International Publishing, S. 107–113.
- Hermida, Alfred; Young, Mary Lynn (2017): Finding the Data Unicorn. A hierarchy of hybridity in data and computational journalism. In: *Digital Journalism* 5 (2), S. 159–176.
- Hermida, Alfred; Young, Mary Lynn (2019): Data Journalism and the Regeneration of News. London and New York: Routledge.
- Heßler, Martina; Mersch, Dieter (2009): Einleitung. Bildlogik oder Was heißt visuelles Denken? In: Martina Heßler und Dieter Mersch (Hg.): *Logik des Bildlichen*. Bielefeld: transcript Verlag, S. 8–62.
- Hine, Christine (2007): Multi-sited Ethnography as a Middle Range Methodology for Contemporary STS. In: *Science, Technology, & Human Values* 32 (6), S. 652–671.
- Holovaty, Adrian (2006): A fundamental way newspaper sites need to change. [www.holovaty.com/writing/fundamental-change/\(12.05.2017\).](http://www.holovaty.com/writing/fundamental-change/(12.05.2017).)

- Howard, Alexander Benjamin (2014): The Art and Science of Data-Driven Journalism. Hg. v. Tow Center for Digital Journalism. Columbia Journalism School. New York.
- Hullman, Jessica (2020): Why Authors Don't Visualize Uncertainty. In: *IEEE Transactions on Visualization and Computer Graphics* 26 (1), S. 130–139.
- Johnson, Mark (2007): The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Reason. 8th printing. Chicago and London: The University of Chicago Press.
- Kennedy, Helen; Hill, Rosemary Lucy; Aiello, Giorgia; Allen, William (2016): The work that visualisation conventions do. In: *Information, Communication & Society* 19 (6), S. 715–735.
- Kennedy, Helen; Weber, Wibke; Engebretsen, Martin (2020): Data visualization and transparency in the news. In: Martin Engebretsen und Helen Kennedy (Hg.): *Data Visualization in Society*: Amsterdam University Press, S. 169–186.
- Kitchin, Rob (2014a): Big Data, new epistemologies and paradigm shifts. In: *Big Data & Society* 1 (1), S. 1–12.
- Kitchin, Rob (2014b): The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. London: Sage.
- Knorr-Cetina, Karin (1981): The Manufacture of Knowledge. An Essay on the Constructivist and Contextual Nature of Science. Burlington: Elsevier Science.
- Knorr-Cetina, Karin (1988): Laboratorien: Instrumente der Weltkonstruktion. In: Paul Hoyningen-Huene und Gertrude Hirsch (Hg.): *Wozu Wissenschaftsphilosophie? Positionen und Fragen zur gegenwärtigen Wissenschaftsphilosophie*. Berlin, New York: Walter de Gruyter, S. 315–344.
- Knorr-Cetina, Karin (2002): Wissenskulturen. Ein Vergleich naturwissenschaftlicher Wissensformen. Frankfurt a.M.: Suhrkamp.
- Kosminsky, Doris; Walny, Jagoda; Vermeulen, Jo; Knudsen, Søren; Willett, Wesley; Carpendale, Sheelagh (2019): Belief at first sight. Data visualization and the rationalization of seeing. In: *Information Design Journal* 25 (1), S. 43–55.
- Krämer, Sybille (2001): Kann das »geistige Auge« sehen? Visualisierung und die Konstitution epistemischer Gegenstände. In: Bettina Heintz und Jörg Huber (Hg.): *Mit dem Auge denken. Strategien der Sichtbarmachung in wissenschaftlichen und virtuellen Welten*. Zürich: Edition Voldemeer, S. 347–364.

- Krämer, Sybille (2009): Operative Bildlichkeit. Von der ›Grammatologie‹ zu einer ›Diagrammatologie‹? Reflexionen über erkennendes ›Sehen‹. In: Martina Heßler und Dieter Mersch (Hg.): Logik des Bildlichen. Bielefeld: transcript Verlag.
- Kueng, Lucy (2017): Going Digital. A Roadmap for Organisational Change. Hg. v. Reuters Institute for the Study of Journalism. University of Oxford.
- Lamont, Michèle (2012): Toward a Comparative Sociology of Valuation and Evaluation. In: *Annual Review of Sociology* 38 (1), S. 201–221.
- Latour, Bruno (2012): Visualization and cognition. Thinking with eyes and hands. In: Michael Lynch (Hg.): Science and Technology Studies. Critical Concepts in the Social Sciences, Bd. 2. 4 Bände. New York: Routledge, S. 3–44.
- Latour, Bruno; Woolgar, Steve (1979): Laboratory Life. The Construction of Scientific Facts. Beverly Hills: Sage Publications.
- Lazer, David; Radford, Jason (2017): Data ex Machina: Introduction to Big Data. In: *Annual Review of Sociology* (43), 7.1–7.21.
- Lesage, Frédéric; Hackett, Robert A. (2014): Between Objectivity and Openness – The Mediality of Data for Journalism. In: *MaC* 2 (2), S. 42–54.
- Lewis, Norman P.; Waters, Stephenson (2018): Data Journalism and the Challenge of Shoe-Leather Epistemologies. In: *Digital Journalism* 6 (6), S. 719–736.
- Lewis, Seth C.; Guzman, Andrea L.; Schmidt, Thomas R. (2019): Automation, Journalism, and Human–Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News. In: *Digital Journalism* 7 (4), S. 409–427.
- Lewis, Seth C.; Usher, Nikki (2013): Open source and journalism: toward new frameworks for imagining news innovation. In: *Media, Culture & Society* 35 (5), S. 602–619.
- Lewis, Seth C.; Usher, Nikki (2014): Code, Collaboration, And The Future Of Journalism. A case study of the Hacks/Hackers global network. In: *Digital Journalism* 2 (3), S. 383–393.
- Lewis, Seth C.; Usher, Nikki (2016): Trading zones, boundary objects, and the pursuit of news innovation: A case study of journalists and programmers. In: *Convergence* 22 (5), S. 1–18.
- Lewis, Seth C.; Westlund, Oscar (2015): Big Data and Journalism. Epistemology, expertise, economics, and ethics. In: *Digital Journalism* 3 (3), S. 447–466.
- Lewis, Seth C.; Zamith, Rodrigo (2017): On the Worlds of Journalism. In: Pablo J. Boczkowski und C.W. Anderson (Hg.): Remaking the News. Essays on

- the Future of Journalism Scholarship in the Digital Age. Cambridge: MIT Press, S. 111–128.
- Link, Jürgen (2001): Aspekte der Normalisierung von Subjekten. Kollektivsymbolik, Kurvenlandschaften, Infografiken. In: Ute Gerhard, Jürgen Link und Ernst Schulte-Holtey (Hg.): Infografiken, Medien, Normalisierung. Zur Kartografie politisch-sozialer Landschaften. Heidelberg: Synchron Verlag, S. 77–92.
- Loosen, Wiebke; Reimer, Julius; De Silva-Schmidt, Fenja (2017): Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016. In: *Journalism*, S. 1–18.
- Lowrey, Wilson (2002): Word People vs. Picture People: Normative Differences and Strategies for Control Over Work Among Newsroom Subgroups. In: *Mass Communication and Society* 5 (4), S. 411–432.
- Lowrey, Wilson; Hou, Jue (2018): All forest, no trees? Data journalism and the construction of abstract categories. In: *Journalism* 41 (6), 1–17.
- Lynch, Michael (1988): The Externalized Retina: Selection and Mathematization in the Visual Documentation of Objects in the Life Sciences. In: *Human Studies* 11 (2/3), S. 201–234.
- Lynch, Michael (2012): Discipline and the material form of images: an analysis of scientific visibility. In: Michael Lynch (Hg.): *Science and Technology Studies. Critical Concepts in the Social Sciences*, Bd. 1. 4 Bände. New York: Routledge, S. 383–407.
- Maissen, Thomas (2005): Die Geschichte der NZZ 1780–2005. Zürich: Verlag Neue Zürcher Zeitung.
- Mämecke, Thorben; Passoth, Jan-Hendrik; Wehner, Josef (Hg.) (2018): Bedeutende Daten. Modelle, Verfahren und Praxis der Vermessung und Verdattung im Netz. Wiesbaden: Springer VS (Medien, Kultur, Kommunikation).
- Manovich, Lev (2001): The Language of New Media. Cambridge: MIT Press.
- Manovich, Lev (2011): What is visualisation? In: *Visual Studies* 26 (1), S. 36–49.
- Marcus, George E. (1995): Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography. In: *Annual Review of Anthropology* 24, S. 95–117.
- Mayer-Schönberger, Viktor; Cukier, Kenneth (2013): Big Data: Die Revolution, die unser Leben verändern wird. München: Redline.
- Merz, Martina; Hinterwaldner, Inge (2012): Neue Bilder, Modelle und Simulationen: Zwischen Repräsentativität und Produktivität. In: Sabine Maasen, Mario Kaiser, Martin Reinhart und Barbara Sutter (Hg.): *Handbuch Wissenschaftssoziologie*. Wiesbaden: Springer, S. 303–316.

- Meyer, Miriah; Dykes, Jason (2019): Criteria for Rigor in Visualization Design Study. In: *IEEE Transactions on Visualization and Computer Graphics*, S. 1–13.
- Meyer, Philip (2002): Precision Journalism. A Reporter's Introduction to Social Science Methods. 4. Aufl. Oxford: Rowman & Littlefield Publishers.
- Mische, Ann (2008): Partisan Publics. Communication and contention across Brazilian youth activist networks. Princeton: Princeton University Press (Princeton studies in cultural sociology).
- Mützel, Sophie; Saner, Philippe; Unternährer, Markus (2018): Schöne Daten! Konstruktion und Verarbeitung von digitalen Daten. In: Bianca Prietl und Daniel Houben (Hg.): Datengesellschaft. Einsichten in die Datafizierung des Sozialen. Bielefeld: transcript Verlag, S. 111–132.
- Nadai, Eva; Maeder, Christoph (2005): Fuzzy Fields. Multi-Sited Ethnography in Sociological Research. In: *Forum: Qualitative Social Research* 6 (3), S. 1–10.
- Naess, Hans Erik (2016): Creating »the field«: Glocality, Relationality and Transformativity. In: *Forum: Qualitative Social Research* 17 (3), S. 1–20.
- Nicolini, Davide (2009): Zooming In and Out: Studying Practices by Switching Theoretical Lenses and Trailing Connections. In: *Organization Studies* 30 (12), S. 1391–1418.
- Nicolini, Davide (2013): Practice Theory, Work, and Organization. An Introduction. Oxford: Oxford University Press.
- Nohr, Rolf F. (2004): Medien(a)nomalien. Viren, Schläfer, Infiltrationen. In: Rolf F. Nohr (Hg.): Evidenz – »... das sieht man doch!«. Hamburg: Lit-Verlag, S. 57–89.
- O'Reilly, Karen (2005): Ethnographic Methods. London and New York: Routledge.
- Ojo, Adegboyega; Heravi, Bahareh R. (2018): Patterns in Award Winning Data Storytelling. Story Types, Enabling Tools and Competences. In: *Digital Journalism* 6 (6), S. 693–718.
- Orlikowski, Wanda J. (2010): The sociomateriality of organisational life: Considering technology in management research. In: *Cambridge Journal of Economics* 34 (1), S. 125–141.
- Orr, Julian E. (1996): Talking about Machines. An Ethnography of a Modern Job. Ithaca: Cornell University Press.
- Orr, Julian E. (2006): Ten Years of Talking About Machines. In: *Organization Studies* 27 (12), S. 1805–1820.
- Parasie, Sylvain (2014): Data-Driven Revelation? Epistemological tensions in investigative journalism in the age of »big data«. In: *Digital Journalism* 3 (3), S. 364–380.

- Parasie, Sylvain; Dagiral, Eric (2012): Data-driven journalism and the public good: »Computer-assisted-reporters« and »programmer-journalists« in Chicago. In: *new media & society*, 14(1) 124-135.
- Passmann, Johannes (2013): Forschungsmedien erforschen. Über Praxis mit der Daten-Mapping-Software Gephi. In: *Navigationen – Zeitschrift für Medien- und Kulturwissenschaften* 13 (2), S. 113–129.
- Passoth, Jan-Hendrik; Rammert, Werner (2016): Fragmentale Differenzierung und die Praxis der Innovation. Wie immer mehr Innovationsfelder entstehen. In: Werner Rammert, Arnold Windeler, Hubert Knoblauch und Michael Hutter (Hg.): Innovationsgesellschaft heute. Perspektiven, Felder und Fälle. Wiesbaden: Springer VS, S. 39–67.
- Passoth, Jan-Hendrik; Wehner, Josef (Hg.) (2013): Quoten, Kurven und Profile. Zur Vermessung der sozialen Welt. Wiesbaden: Springer.
- Paterson, Chris A.; Domingo, David (Hg.) (2008): Making online news. The ethnography of new media production. New York: Lang (Digital formations, 49).
- Pimentel, Marco A. F.; Clifton, David A.; Clifton, Lei; Tarassenko, Lionel (2014): A review of novelty detection. In: *Signal Processing* 99, S. 215–249.
- Porter, Theodore M. (1995): Trust in Numbers. The Pursuit of Objectivity in Science and Public Life. Princeton: Princeton University Press.
- Prietl, Bianca; Houben, Daniel (Hg.) (2018): Datengesellschaft. Einsichten in die Datafizierung des Sozialen. Bielefeld: transcript Verlag.
- Reckwitz, Andreas (2003): Grundelemente einer Theorie sozialer Praktiken. Eine sozialtheoretische Perspektive. In: *Zeitschrift für Soziologie* 32 (4), S. 282–301.
- Reichertz, Jo (2013): Die Abduktion in der qualitativen Sozialforschung. 2. Aufl. Wiesbaden: Springer VS.
- Rettberg, Jill Walker (2020): Ways of knowing with data visualizations. In: Martin Engebretsen und Helen Kennedy (Hg.): Data Visualization in Society: Amsterdam University Press, S. 35–48.
- Royal, Cindy (2012): The Journalist as Programmer: A Case Study of The New York Times Interactive News Technology Department. In: #ISOJ: the Official Research Journal of the International Symposium for Online Journalism 2 (1), S. 5–24.
- Rule, Adam Carl (2018): Design and Use of Computational Notebooks. Dissertation. University of California, San Diego. <https://escholarship.org/uc/item/0dc498tf>.

- Sahlin-Andersson, Kerstin (1996): Imitating by Editing Success: The Construction of Organization Fields. In: Barbara Czarniawska und Guie Sevón (Hg.): *Translating Organizational Change*. Berlin, New York: Walter de Gruyter, S. 69–92.
- Satyanarayan, Arvind; Moritz, Dominik; Wongsuphasawat, Kanit; Heer, Jeffrey (2017): Vega-Lite: A Grammar of Interactive Graphics. In: *IEEE Transactions on Visualization and Computer Graphics* 23 (1), S. 341–350.
- Saxenian, AnnaLee (1996): *Regional Advantage. Culture and Competition in Silicon Valley and Route 128*. 4. Aufl. Cambridge, London: Harvard University Press.
- Schatzki, Theodore R. (2000): Wittgenstein and the social context of an individual life. In: *History of the Human Sciences* 13 (1), S. 93–107.
- Schudson, Michael (2000): The sociology of News Production Revisited (Again). In: James Curran und Michael Gurevitch (Hg.): *Mass Media and Society*. 3. Aufl. London: Arnold, S. 175–200.
- Schudson, Michael; Anderson, C.W. (2009): Objectivity, Professionalism, and Truth Seeking in Journalism. In: Wahl-Jorgensen, Karin und Thomas Hanitzsch (Hg.): *The Handbook of Journalism Studies*. New York: Routledge, S. 88–101.
- Segel, Edward; Heer, Jeffrey (2010): Narrative Visualization: Telling Stories with Data. In: *IEEE Transactions on Visualization and Computer Graphics* 16 (6), S. 1139–1148.
- Silverman, David (2001): *Interpreting qualitative data. Methods for analysing talk, text and interaction*. 2. ed., reprinted. London: Sage.
- Smit, Gerard; de Haan, Yael; Buijs, Laura (2014): Working With or Next to Each Other? Boundary Crossing in the Field of Information Visualisation. In: *The Journal of Media Innovations* 1 (2), S. 36–51.
- Splendore, Sergio; Di Salvo, Philip; Eberwein, Tobias; Groenhart, Harmen; Kus, Michal; Porlezza, Colin (2016): Teaching Big. Educational strategies in the field of data journalism. A comparative study in five European countries. In: *Journalism* 17 (1), S. 138–152.
- Stalp, Florian (2017): Classifying Data Journalism. In: *Journalism Practice* 3 (1), S. 1–19.
- Stalp, Florian; Borges-Rey, Eddy (2018): Data Journalism Sustainability. In: *Digital Journalism* 6 (8), S. 1078–1089.
- Star, Susan Leigh (1999): The Ethnography of Infrastructure. In: *American Behavioral Scientist* 43 (3), S. 377–391.

- Stark, David (2009): *The Sense of Dissonance. Accounts of Worth in Economic Life*. Princeton und Oxford: Princeton University Press.
- Stavelin, Eirik (2013): Computational journalism. When journalism meets programming. Dissertation. University of Bergen, Bergen.
<http://bora.uib.no/bitstream/handle/1956/7926/dr-thesis-2013-Eirik-Stavelin.pdf>.
- Tavory, Iddo; Timmermans, Stefan (2014): *Abductive analysis. Theorizing qualitative research*. Chicago: The University of Chicago Press.
- Tuchman, Gaye (1980): *Making news. A study in the construction of reality*. First Free Press paperback ed. New York: Free Press.
- Tufte, Edward R. (1983): *The Visual Display of Quantitative Information*. First Edition. Cheshire: Graphics Press.
- Tufte, Edward R. (1990): *Envisioning Information*. Cheshire: Graphics Press.
- Tukey, John W. (1977): *Exploratory data analysis*. Reading, Massachusetts: Addison-Wesley.
- Usher, Nikki (2016): *Interactive Journalism. Hackers, Data, and Code*. Urbana: University of Illinois Press.
- Vaara, Eero (2019): *The Power of Diagrams*. In: Markus A. Höllerer, Theodor Jacob van Leeuwen, Dennis Jancsary, Renate E. Mayer, Thomas Hestbaek Anderson und Eero Vaara (Hg.): *Visual and multimodal research in organization and management studies*. New York: Routledge, S. 104–128.
- van Maanen, John (2000): *Tales of the field. On writing ethnography*. [Nachdr.]. Chicago: Univ. of Chicago Press (Chicago guides to writing, editing, and publishing).
- van Maanen, John (2006): Ethnography then and now. In: *Qualitative Research in Organizations and Management: An International Journal* 1 (1), S. 13–21.
- Vos, Tim P. (2018): *Journalism*. In: Tim P. Vos (Hg.): *Journalism*. Boston, Berlin: Walter de Gruyter (Handbooks of Communication Science, 19), S. 1–18.
- Walny, Jagoda (2016): *Thinking with Sketches: Leveraging Everyday Use of Visuals for Information Visualization*. Ph.D. Thesis. University of Calgary, Calgary.
- Ward, Stephen J. A. (2018): *Epistemologies of Journalism*. In: Tim P. Vos (Hg.): *Journalism*. Boston, Berlin: Walter de Gruyter (Handbooks of Communication Science, 19), S. 63–82.
- Weber, Wibke (2018): *Multidisziplinäre Forschungsperspektiven auf Infografiken und Datenvisualisierungen*. In: Katharina Lobinger (Hg.): *Handbuch visuelle Kommunikationsforschung*. Wiesbaden: Springer VS, S. 1–25.

- Wenger, Etienne (2008): Communities of practice. Learning, meaning, and identity. Cambridge: Cambridge Univ. Press (Learning in doing).
- Wickham, Hadley (2010): A Layered Grammar of Graphics. In: *Journal of Computational and Graphical Statistics* 19 (1), S. 3–28.
- Wickham, Hadley (2014): Tidy Data. In: *Journal of Statistical Software* 59 (10), S. 1–23.
- Wihbey, John (2019): The Social Fact: News and Knowledge in a Networked World. Cambridge: MIT Press.
- Wilkinson, Leland (2005): The Grammar of Graphics. 2. Aufl. New York: Springer.
- Wing, Jeannette M. (2006): Computational thinking. In: *Communications of the ACM* 49 (3), S. 33–35.
- Wintsch, Dani (2006): Doing News – Die Fabrikation von Fernsehnachrichten. Eine Ethnografie videojournalistischer Arbeit. Zugl.: Zürich, Univ., Diss., 2005–2006. Wiesbaden: VS Verlag für Sozialwissenschaften | GWV Fachverlage GmbH Wiesbaden.
- Young, Mary Lynn; Hermida, Alfred; Fulda, Johanna (2018): What Makes for Great Data Journalism? In: *Journalism Practice* 12 (1), S. 115–135.
- Zamith, Rodrigo (2019): Transparency, Interactivity, Diversity, and Information Provenance in Everyday Data Journalism. In: *Digital Journalism*.
- Zelizer, Barbie (1993): Journalists as interpretive communities. In: *Critical Studies in Mass Communication* 10 (3), S. 219–237.
- Zelizer, Barbie (2009): Journalism and the Academy. In: Wahl-Jorgensen, Karin und Thomas Hanitzsch (Hg.): *The Handbook of Journalism Studies*. New York: Routledge, S. 29–41.
- Zhang, Xinzhi (2018): Visualization, Technologies, or the Public? Exploring the articulation of data-driven journalism in the Twittersphere. In: *Digital Journalism* 6 (6), S. 737–758.
- Zimek, Arthur; Filzmoser, Peter (2018): There and back again: Outlier detection between statistical reasoning and data mining algorithms. In: *WIREs Data Mining Knowledge Discovery* 8 (6), S. 123.

9 Anhang

A Vorlage Feldnotiz

Name: Redaktion BR Data		Nummer der Feldnotiz: xx
Datum: Wochentag, dd.mm.yyyy	Zeit: x.xx bis x.xx Uhr	Ort: BR Data, Studio Freimann, München
Erstellungsdatum: dd.mm.yyyy	Quelle: Notizen, Gespräche, ...	
<u>Rahmen:</u>		
Erwartungen, Empfindungen, eigener Einfluss:		

#	Beobachtung/Situation	Überlegung
1		
2		
3		

#	Material	Überlegung
1		
2		
3		

B Fragebogen Kurz-Interviews NZZ Storytelling

Datum, Ort, Rahmenbedingungen des Gesprächs: .

Mitarbeiterin bzw. Mitarbeiter von NZZ Storytelling:

Funktion (Berufs-/Stellenbezeichnung) bei NZZ Storytelling:

Aufgaben und Kompetenzen:

beruflicher Werdegang, Ausbildungen und Schulen:

bei NZZ Storytelling seit und weshalb:

Mit welchen Tools arbeitest du sehr oft, manchmal, selten?

Wie oft hast du Kontakt mit Mitarbeitenden von ausserhalb der Abteilung? Mit welchen?

Welche Fähigkeiten, Methoden oder Tools würdest du dir gerne aneignen? Auf welchem Weg?

Wie eignest du dir (neues) Wissen und (neue) Kompetenzen an?

Verfolgst du Blogs, Newsletter, Diskussionen, ... aus deinem Arbeitsbereich? Welche?

Was ist aus deiner Sicht der (Mehr-)Wert von Visualisierungen und vermehrter Nutzung von Daten für Medienberichterstattung?

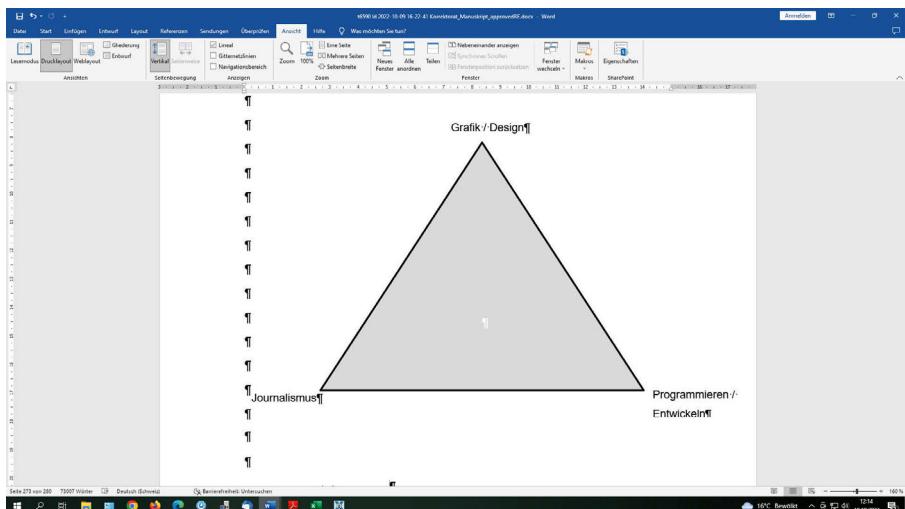
Wie erlebst du die Zusammenarbeit mit Mitarbeitenden innerhalb und ausserhalb von NZZ Storytelling? Wo gibt es Unterschiede?

An was denkst du beim Begriff Datenjournalismus?

Was ich noch sagen wollte...

Wo in diesem Dreieck würdest du dich bezüglich deiner Kompetenzen einordnen?

Abbildung B



C Fragebogen Kurz-Interviews BR Data

Datum, Ort, Rahmenbedingungen des Gesprächs:

Mitarbeiterin bzw. Mitarbeiter von BR Data:

Funktion (Berufs-/Stellenbezeichnung) bei BR Data:

Aufgaben und Kompetenzen:

beruflicher Werdegang, Ausbildungen und Schulen:

bei BR Data seit und weshalb:

Wie erlebst du die Zusammenarbeit mit Mitarbeitenden innerhalb und ausserhalb von BR Data?

Welche Gemeinsamkeiten und Unterschiede gibt es?

Wo liegt der Nutzen von Datenjournalismus?

Wie macht man aus Daten Geschichten?

Und welche?

Mit welchen Tools arbeitest du sehr oft, manchmal, selten?

Wie kam es dazu, dass du gerade diese Tools (und nicht andere) verwendest?

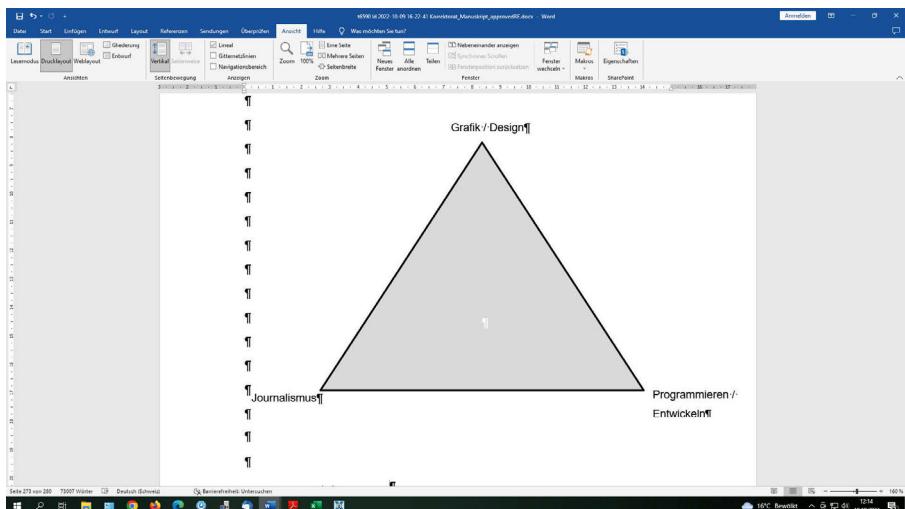
Welche Fähigkeiten, Methoden oder Tools würdest du dir gerne aneignen?

Wie eignest du dir (neues) Wissen und (neue) Kompetenzen an?

Verfolgst du Blogs, Newsletter, Diskussionen, ... aus deinem Arbeitsbereich? Welche?

Wo in diesem Dreieck würdest du dich bezüglich deiner Kompetenzen einordnen?

Abbildung C



D Log-Buch Feld-Material NZZ und BR in atlas.ti

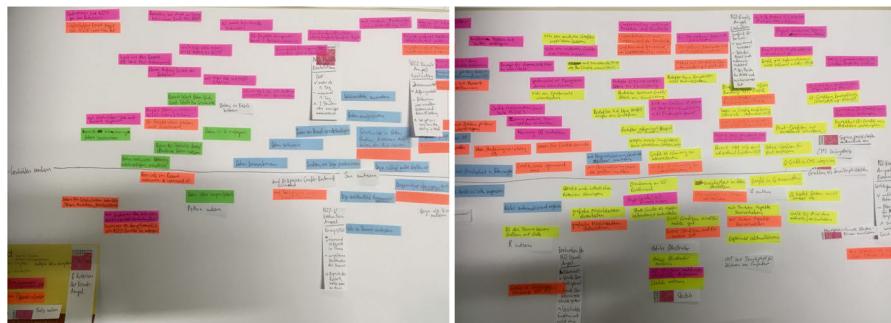
Fall	Schritt	Zeitraum	Vorgehen
NZZ	Erst-Codierung Feldnotizen	Mai 18	Codierung Praktiken
NZZ	Re-Codierung Feldnotizen Bildung Code-Gruppen	Aug 18	Codierung näher an den Praktiken, Bildung Code-Gruppen aufgrund des Editing-Narrativ-Konzepts von Sahlin-Andersson, abgestimmt auf die sich abzeichnenden Bereiche (ohne Anspruch, dass Gruppen so bleiben müssen)
NZZ	Memos aus Feldnotizen Bildung Memo-Gruppen	Aug 2018 - Mar 2019	Memos aufgrund der Überlegungen während des Codierens sowie Betrachtung der Code-Gruppen
NZZ	Codierung Kurz-Interviews Bildung Code-Gruppen nach Fragen	Jul 18	Codierung nahe an den Aussagen, Bildung Code-Gruppen aufgrund der Interview-Fragen
NZZ	Auswertung Code-Gruppen nach Mustern zwischen Sub-Gruppen	Jul 18	Zuordnung Codes/Codegruppen zu Sub-Gruppen DJ, Designer, IGT, IT, Teamleitung mittels Funktion Code-Dokument-Tabelle, Zusammenfassung der Tabellen
NZZ	Memos aus Kurz- Interviews	Jul 18	Memos aufgrund der Überlegungen während des Codierens
NZZ	Auswertung follow Fussball- Projekt	Aug 18	Extrahieren aller Situationen, in denen am Projekt gearbeitet oder darüber geredet wurde, Notizen erster theoretischer Überlegungen dazu
NZZ	Netzwerk Explainer aus Codes/Zitaten Feldnotizen	Aug 18	Vernetzung aller Codes (Praktiken) zu Explainern mit den Zitaten und in einem zweiten Schritten den Codes/Praktiken, die ebenfalls mit diese Zitaten verbunden sind
NZZ	Netzwerk Style-Guide aus Codes/Zitaten Feldnotizen	Jan 19	Vernetzung aller Codes (Praktiken) zu Explainern mit den Zitaten und in einem zweiten Schritten den Codes/Praktiken, die ebenfalls mit diese Zitaten verbunden sind
NZZ	Zusammenführungen Codes Verbindungen zu Memos	Apr 19	Codes ähnlicher Praktiken werden zusammengeführt, aber auch Bildung nächsthöhere Code-Ebene - jeweils Notiz im Code-Kommentar, welche Codes zusammengeführt wurden
NZZ	Prozess der Praktiken	Aug 19	Durchsicht und Auswahl der relevantesten Codes (Anzahl, theoretische Relevanz/Reichhaltigkeit) aus allen Code-Gruppen für ein Schema auf Papier, Einteilung nach Farbe (möglichst analog BR-Übersicht), Positionierung anhand Ablauf datenjournalistischer Prozess der Geschichten-Produktion

BR	Erst-Codierung Feldnotizen	Apr 19	Codierung Praktiken
BR	Re-Codierung Feldnotizen	Jun 19	Codierung Praktiken
BR	Memo Gespräche BRD-Leitung	Jul 19	ausführliches Memo zu den beiden Gesprächen mit Team-Leitungspersonen, die sich in den Feldnotizen finden (keine Codierung, da Interview-Situation)
BR	Auswertung follow PM-Projekt	Jul 19	Extrahieren aller Situationen, in denen am Projekt gearbeitet oder darüber geredet wurde, Notizen (Memo) erster theoretischer Überlegungen dazu
BR	Bildung Code-Gruppen	Jul 19	Bildung Gruppen anhand derer von NZZ-ST, danach iterativ Zuordnung der Codes und Anpassung der Gruppen (Änderung/neue/löschen)
BR	Zusammenführungen Codes	Jul 19	Codes ähnlicher Praktiken werden zusammengeführt, aber auch Bildung nächsthöhere Code-Ebene - jeweils Notiz im Code-Kommentar, welche Codes zusammengeführt wurden
BR	Memos Ideen-Geschichten	Jul 19	alle relevanten Codes aus den jeweiligen Code-Gruppen, die in Verbindung mit der Code-Gruppe "Ideen und Geschichten entwickeln" (IGE) stehen, beschreiben und in einem Memo pro Code-Gruppe festhalten
BR	Prozess der Praktiken	Jul 19	Auswahl der relevantesten Codes (Anzahl, theoretische Relevanz/Reichhaltigkeit) in Verbindung mit IGE für ein Schema auf Papier
BR	Memos ergänzen	Jul 19	Ergänzung der IGE-Code-Gruppen-Memos mit Gedanken zu Codes, die nicht in Verbindung mit IGE stehen - hier selektivere Auswahl, das heisst höherer Anspruch an Relevanz (was nicht in Verbindung mit IGE steht, ist weniger im Fokus).
BR	Prozess der Praktiken IGE ergänzen	Jul 19	Nach Durchsicht der Code ohne IGE-Verbindung Auswahl der wichtigsten Codes und Ergänzung des Prozess-Schemas
BR	weitere Memos Ergänzung Prozess IGE	Jul 19	Memos zu Themen, die während der bisherigen Codierung, Auswertung und Memo-Schreiben als wichtig erachtet wurden
BR	Codierung Kurz-Interviews Bildung Code-Gruppen	Okt 19	Codierung nahe an den Aussagen, Bildung Code-Gruppen aufgrund der Interview-Fragen
BR	Auswertung Code-Gruppen nach Mustern zwischen Sub-Gruppen	Okt 19	Gruppierung Codes zu gewissen Kategorien (Anlehnung an NZZ-ST-Kategorien soweit vorhanden), Zuordnung Codes/Codegruppen zu Sub-Gruppen DJ, IT (Teamleitung weggelassen da kaum wichtig) mittels Funktion Code-Dokument-Tabelle, Zusammenfassung der Tabelle in eine
BR	Memo aus Kurz-Interviews	Okt 19	Memos aufgrund der Überlegungen während des Codierens

E Prozessabläufe von Daten zu Geschichten

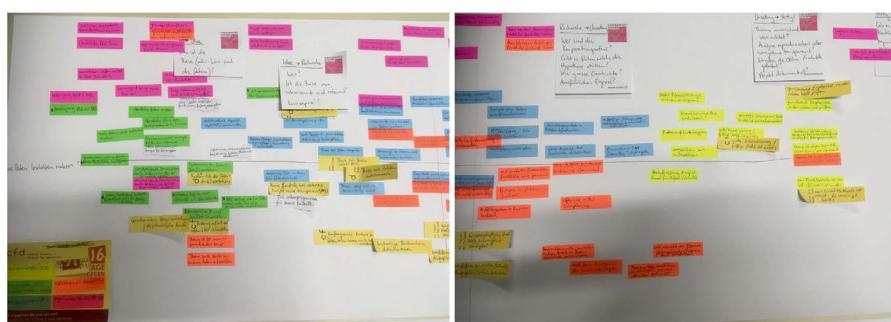
NZZ Storytelling

Abbildung 26: Codes der ethnografischen Auswertung der Beobachtungen bei NZZ Storytelling, platziert im Prozessablauf von Daten zu Geschichten.



BR Data

Abbildung 27: Codes der ethnografischen Auswertung der Beobachtungen bei BR Data, platziert im Prozessablauf von Daten zu Geschichten.



F Interview-Leitfaden Expert*innen

(persönlich)

Was ist/war deine Motivation, Journalismus mit Daten zu betreiben?

Dein Hintergrund?

Wie bleibst du auf dem Laufenden, was im Datenjournalismus gerade der >State of the Art< ist?

(Tools)

Mit welchen Tools arbeitest du?

Hast du Lieblingstools?

Welche Rolle nimmt die Visualisierung ein?

Gibt es ein Tool, das man deiner Meinung nach noch erfinden sollte?

(Feld Datenjournalismus)

Was ist der (Mehr-)Wert von Daten und Visualisierungen im Journalismus?

Was ist aus deiner Sicht eine gute Datenjournalismus-Geschichte? Hast du Beispiele?

Wie schätzt du die Entwicklungen der letzten Jahre ein?

Datenjournalismus hat eine sehr aktive Community. Wie ist sie entstanden, welche Faktoren haben sie begünstigt?

(eigenes Team/Redaktion)

Wie ist Datenjournalismus bei [Redaktion] organisiert? Wie entwickelt ihr Geschichten? Mit welchen Personen und Ressorts arbeitest du zusammen?

Wie erklärst du deine Arbeit den Redaktor*innen, die wenig Ahnung von Datensammlung, -auswertung und -visualisierung haben?

Wenn du für dein Team eine*n Datenjournalist*in einstellt – was muss sie/er können oder mitbringen?

Ihr legt viel Wert auf die Offenlegung von Quellen und Methodik in euren Projekten. Wieso?

Soziologie



Michael Volkmer, Karin Werner (Hg.)
Die Corona-Gesellschaft
Analysen zur Lage und Perspektiven für die Zukunft

2020, 432 S., kart., 2 SW-Abbildungen
24,50 € (DE), 978-3-8376-5432-5
E-Book:
PDF: 21,99 € (DE), ISBN 978-3-8394-5432-9
EPUB: 21,99 € (DE), ISBN 978-3-7328-5432-5



Vera Hofmann, Johannes Euler, Linus Zumühlen, Silke Helfrich
Commoning Art –
Die transformativen Potenziale
von Commons in der Kunst

Juli 2022, 124 S., kart
19,50 € (DE), 978-3-8376-6404-1
E-Book: kostenlos erhältlich als Open-Access-Publikation
PDF: ISBN 978-3-8394-6404-5



Kerstin Jürgens
Mit Soziologie in den Beruf
Eine Handreichung

2021, 160 S., kart.
18,00 € (DE), 978-3-8376-5934-4
E-Book:
PDF: 17,99 € (DE), ISBN 978-3-8394-5934-8

**Leseproben, weitere Informationen und Bestellmöglichkeiten
finden Sie unter www.transcript-verlag.de**

Soziologie



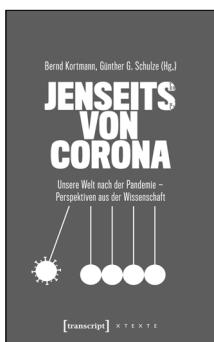
Gabriele Winker
Solidarische Care-Ökonomie
Revolutionäre Realpolitik für Care und Klima

2021, 216 S., kart.
15,00 € (DE), 978-3-8376-5463-9
E-Book:
PDF: 12,99 € (DE), ISBN 978-3-8394-5463-3



Wolfgang Bonß, Oliver Dimbath, Andrea Maurer,
Helga Pelizäus, Michael Schmid
Gesellschaftstheorie
Eine Einführung

2021, 344 S., kart.
25,00 € (DE), 978-3-8376-4028-1
E-Book:
PDF: 24,99 € (DE), ISBN 978-3-8394-4028-5



Bernd Kortmann, Günther G. Schulze (Hg.)
Jenseits von Corona
Unsere Welt nach der Pandemie –
Perspektiven aus der Wissenschaft

2020, 320 S., Klappbroschur, 1 SW-Abbildung
22,50 € (DE), 978-3-8376-5517-9
E-Book:
PDF: 19,99 € (DE), ISBN 978-3-8394-5517-3
EPUB: 19,99 € (DE), ISBN 978-3-7328-5517-9

**Leseproben, weitere Informationen und Bestellmöglichkeiten
finden Sie unter www.transcript-verlag.de**

