

Datenwissenschaften und Gesellschaft: Die Genese eines transversalen Wissensfeldes

Saner, Philippe

Veröffentlichungsversion / Published Version

Dissertation / phd thesis

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
transcript Verlag

Empfohlene Zitierung / Suggested Citation:

Saner, P. (2022). *Datenwissenschaften und Gesellschaft: Die Genese eines transversalen Wissensfeldes*. (Digitale Soziologie, 1). Bielefeld: transcript Verlag. <https://doi.org/10.14361/9783839462591>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

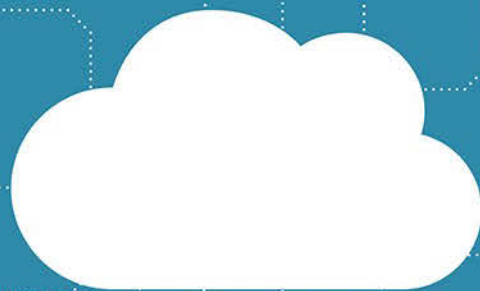
Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:
<https://creativecommons.org/licenses/by/4.0>

PHILIPPE SANER

DATENWISSEN- SCHAFTEN UND GESELLSCHAFT

DIE GENESE EINES
TRANSVERSALEN
WISSENSFELDES



[transcript] DIGITALE SOZIOLOGIE

Philippe Saner
Datenwissenschaften und Gesellschaft

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 substantielle Fragestellungen der Soziologie anbieten, die aktuelle methodische und konzeptionelle Entwicklungen aufgreifen oder innovative Methodenkombinationen nutzen. Dabei legt sie besonderen 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 Sammelbände 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.

Philippe Saner lehrt und forscht in den Bereichen Wissenschafts- und Techniksoziologie, Wirtschaftssoziologie und Hochschulforschung. Er promovierte an der Universität Luzern und erhielt für seine Dissertation den Ulrich-Teichler-Preis der Gesellschaft für Hochschulforschung e.V. Zu seinen Themenschwerpunkten gehören Digitalisierung, Bildung, Arbeit sowie Professionen.

Philippe Saner

Datenwissenschaften und Gesellschaft

Die Genese eines transversalen Wissensfeldes

[transcript]

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

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

Gutachter*innen: Prof. Dr. Sophie Mützel (Universität Luzern) und Prof. Dr. Andrea Glauser (mdw – Universität für Musik und darstellende Kunst Wien)

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet 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 4.0 Lizenz (BY). 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.

(Lizenztext: <https://creativecommons.org/licenses/by/4.0/deed.de>)

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 2022 im transcript Verlag, Bielefeld

© **Philippe Saner**

Umschlaggestaltung: Maria Arndt, Bielefeld

Korrektur: Anette Nagel, Osnabrück

Satz: Francisco Bragança, Bielefeld

Druck: Majuskel Medienproduktion GmbH, Wetzlar

Print-ISBN 978-3-8376-6259-7

PDF-ISBN 978-3-8394-6259-1

<https://doi.org/10.14361/9783839462591>

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

Inhalt

Vorwort	11
Abbildungsverzeichnis	13
Tabellenverzeichnis	15
Danksagung	17
Kapitel 1 – Einleitung	21
1.1 Wie entstehen transversale Wissensfelder?	22
1.2 Die Datenwissenschaften als transversales Phänomen	25
1.3 Beiträge der Arbeit	29
1.4 Aufbau der Arbeit	30
Teil I – Grundlagen	
Kapitel 2 – Transversale Wissensgebiete als Räume zwischen Feldern	35
2.1 Einleitung	35
2.2 Die Praxistheorie Pierre Bourdieus als Ausgangspunkt	36
2.3 Die Entstehung neuer Wissensfelder	40
2.4 Erweiterungen der Bourdieu'schen Feldtheorie zur Erklärung neuer Wissensfelder	42
2.5 Zwischenräume als analytische Erweiterung der Feldtheorie	48
2.6 Das trennende und verbindende Potenzial von Grenzarbeit	52
2.7 Die Konstruktion von Vielstimmigkeit neuer Wissensgebiete durch Begriffsarbeit	57
2.8 Analytisches Modell: Transversale Wissensgebiete als Räume zwischen Feldern	61

Kapitel 3 –

»Data Science« als soziales Phänomen:

Genese und multiple Perspektiven 67

- 3.1 Begriffliche, disziplinäre und theoretische Perspektiven
auf die Datenwissenschaften 68
- 3.2 Konstruktionen der Datenwissenschaften im Arbeitsmarkt 75
- 3.3 Konstruktionen der Datenwissenschaften in der Bildungs-
und Forschungspolitik 81
- 3.4 Konstruktionen der Datenwissenschaften im akademischen Feld 84
- 3.5 Zwischenstand 89

Kapitel 4 –

Forschungsdesign 93

- 4.1 Die verteilte Analyse von Zwischenräumen 93
- 4.2 Stellenanzeigen im Arbeitsmarkt 97
- 4.3 Strategiedokumente im Feld der Politik 105
- 4.4 Curricula und Interviews im akademischen Feld 108
- 4.5 Zusammenfassung 115

Teil II –

Repräsentationen und Imaginationen

von Datenwissenschaften in Arbeitsmarkt und Politik

Kapitel 5 –

Repräsentationen der Datenwissenschaften

im schweizerischen Arbeitsmarkt 119

- 5.1 Einleitung 119
- 5.2 Entwicklung und Struktur
des Arbeitsmarktsegments Datenwissenschaften 122
- 5.3 Repräsentationen von »Data Science«
in den englischsprachigen Stellenanzeigen 132
- 5.4 Repräsentationen der Datenwissenschaften
in deutschsprachigen Stellenanzeigen 139
- 5.5 Diskussion 144

Kapitel 6 –

Zukunftsentwürfe der Datenwissenschaften

in Diskursen der Bildungs- und Forschungspolitik 147

6.1	Einleitung	147
6.2	Geordnete Arbitrarität im politischen Diskurs zur Digitalisierung	149
6.3	Zukunftsentwürfe von Bildung und Forschung in der Strategie »Digitale Schweiz«	158
6.4	»Data turn«: Auf dem Weg zur Datengesellschaft?	167
6.5	Diskussion	176

Teil III –

Konstruktionen der Datenwissenschaften im akademischen Feld

Kapitel 7 –

Die Konstruktion der Datenwissenschaften

im akademischen Feld durch Begriffsarbeit und *boundary work* 181

7.1	Einleitung	181
7.2	Paradigmenwechsel? »Data Science« zwischen Universalisierung und Disziplinierung	183
7.3	Die Suche nach den richtigen Kategorien	185
7.4	Innerwissenschaftliche Konfliktlinien	188
7.5	Die Logik der Domänen: Interdisziplinarität im Kernbereich und <i>domain knowledge</i>	191
7.6	Spielarten von Anwendungsorientierung: »Echte Probleme« und »operative Exzellenz«	193
7.7	Diskussion	195

Kapitel 8 –

Die Verhandlung der Datenwissenschaften

in Universitäten und Hochschulen 197

8.1	Einleitung	197
8.2	»Nützlichkeitserwägungen«: Die Nachfrage auf dem Arbeitsmarkt	199
8.3	Das Zusammenspiel politischer Bewilligungsverfahren und organisationaler Planungsprozesse	201
8.4	Die Mechanik hinter den Kulissen: Zur Finanzierung der Studiengänge	203
8.5	Datenwissenschaften als Organisationsproblem – zur Verortung und Implementierung der Studiengänge in den Hochschulen	206
8.6	Die Synchronizität von Kooperations- und Konkurrenzverhältnissen im akademischen Feld	208
8.7	Diskussion	217

Kapitel 9 –	
Die Strukturlogik datenwissenschaftlicher Curricula	219
9.1 Einleitung	219
9.2 Zulassungsvoraussetzungen	221
9.3 Kanonisierungsprozesse im Kernbereich	223
9.4 Zur Bedeutung der Statistik im Curriculum	225
9.5 Wahlfächer: Die Übersetzung der Logik der Domänen in die Curricula	227
9.6 Praxisorientierte Kurse zur Überschreitung feldspezifischer Differenzierungen	228
9.7 Das Prinzip der Komplementarität	231
9.8 Das Subfeld der Weiterbildungsstudiengänge	233
9.9 Diskussion	236
Kapitel 10 –	
Die Suche nach den richtigen Kompetenzen	239
10.1 Einleitung	239
10.2 »Jack of all trades and a master of some«: Kompetenzprofile in den Curricula	241
10.3 Datenkompetenzen: Die Suche nach neuen Begrifflichkeiten	242
10.4 Problemlösungskompetenzen: Die Übersetzung der Anwendungsorientierung in die Praxis	244
10.5 Kommunikation und Teamarbeit: Zur »Moderationsrolle« von Data Scientists an organisationalen Schnittstellen	246
10.6 Denkweisen und persönliche Einstellungen: Das »Data Mindset« als Grundlage für Datafizierungsprozesse	248
10.7 Vielversprechende Zukünfte: Zur »Sexyness« von Data Scientists	252
10.8 Diskussion	254

Teil IV – Schlussbetrachtungen

Kapitel 11 – Synthese	259
11.1 Transversale Konstruktionen der Datenwissenschaften durch Praktiken von Grenz- und Begriffsarbeit	259
11.2 Widersprüche in der Herausbildung der Datenwissenschaften zwischen Innovation und Reproduktion	262
11.3 Konkurrenz- und Kooperationsverhältnisse strukturieren die Datenwissenschaften als Raum zwischen Feldern	263
11.4 Die beschleunigte Verarbeitung neuer sozialer Kategorien	266
11.5 Die Durchsetzung neuer Denk- und Wahrnehmungsweisen fundiert die Mächtigkeit datenwissenschaftlicher Expertise	267
11.6 Zur Generalisierbarkeit des analytischen Modells	269
Bibliografie	273
Anhang	305

Vorwort

Mit diesem Buch liegt ein erster Band der neuen Reihe Digitale Soziologie beim transcript Verlag vor, weitere Publikationen sind in Arbeit. Die neue Reihe soll sowohl theoretisch-konzeptionellen als auch empirischen Untersuchungen eine gemeinsame editorische Plattform bieten. Ihr Gegenstand sind neue Forschungsfelder des Digitalen, aber auch klassische soziologische Phänomene wie Infrastrukturen, Vergemeinschaftungen und Praktiken, die sich durch die bzw. mit der Digitalisierung wandeln. Insbesondere werden in der Reihe Digitale Soziologie Arbeiten veröffentlicht, die neue, auch unorthodoxe Zugriffe auf substantielle Fragestellungen der Soziologie anbieten, die aktuelle methodische und konzeptionelle Entwicklungen aufgreifen oder innovative Methodenkombinationen nutzen. Dabei legt die Reihe besonderen Wert auf Arbeiten, die die Rolle der Soziologie angesichts der digitalisierten Gesellschaft diskutieren und beispielsweise Implikationen von Infrastrukturen und wissenschaftlicher Praxis reflektieren.

Wir Herausgeber*innen der Reihe Digitale Soziologie sehen einen wachsenden Bedarf an einem Publikationsformat, welches den disziplinären und interdisziplinären Verständigungsprozess darüber, was alles Digitale Soziologie ist, anstößt und begleitet. Gerade wenn man das Verhältnis zwischen Soziologie und Computational Social Sciences in den Blick nimmt, und danach fragt, was das Gemeinsame ist und worin die Unterschiede bestehen, wird deutlich, dass in den verschiedenen Communities hierüber keine Einigkeit besteht. Gleichzeitig wird mit der fortschreitenden Digitalisierung moderner Gesellschaften aber auch deutlich, dass damit nicht nur der Gegenstand der Soziologie digitaler wird, sondern dass auch ihre Methodik und ihr Selbstverständnis zumindest eine Erweiterung, wenn nicht gar eine Transformation erfahren.

Über diese Aspekte laden wir Herausgeber*innen zu einem offenen und konstruktiven Diskurs ein, der nicht zuletzt eine Vielfalt von methodischen Zugängen einschließt. So sind sowohl quantitative als auch qualitative Ansätze wie auch inventive, Digital-Methods- und Mixed-Methods-Zugänge willkommen. Neben empirisch ausgerichteten Studien soll die Reihe auch ein Ort des Austauschs über Theoriebildung einer Digitalen Soziologie sein. Dabei soll der Breite möglicher theoretischer Ansätze gegenüber der Fixierung auf ein theoretisches Paradigma der Vortritt gewährt werden.

Gerade weil wir davon überzeugt sind, dass die Digitale Soziologie das Selbstverständnis und das forschersische Tun der Soziologie verändern wird, halten wir eine möglichst breit und inklusiv angelegte Diskussion über die mögliche Ausgestaltung dieses Wandels für notwendig.

Philippe Saners Dissertationsschrift *Datenwissenschaften und Gesellschaft. Die Genese eines transversalen Wissensfeldes* eröffnet diese Diskussion passenderweise mit einem soziologischen Blick auf das entstehende Feld der Datenwissenschaften. Auf Grundlage zeitgenössischer Formen der relationalen Sozial- und Gesellschaftstheorie legt Saner eine tiefenscharfe Optik auf das Zusammenspiel wissenschaftlicher, technologischer und kultureller Felder an und zeigt, wie diese die modernen Datenwissenschaften in ihrer mehrdimensionalen, multipolaren und umkämpften gegenwärtigen Gestalt hervorbringen und prägen.

Die historischen Verflechtungen zwischen den Sphären der Wissenschaften, der Religion, der Ökonomie, der Politik und des westlichen Nationalstaates, die ihren frühen Ausdruck etwa als ›Statistik als die Lehre von den Daten über den Staat‹ fanden, erfahren heute ihren Fortgang in der multipolaren Konstitution eines Feldes der Datenwissenschaften. Und was zunächst mit der in den verschiedenen Wissenschaften separierten und dann zunehmend akademisch vereinheitlichten institutionalisierten Statistik ihren Anfang nahm, schreibt sich dabei in diesem Feld unaufhaltsam fort. Dieses Feld im Werden kann nicht einer einzelnen Disziplin oder auch nur alleinig der wissenschaftlichen Sphäre zugeordnet werden. Unter Rückgriff auf qualitative Auswertungen von Strategiepapieren, Interviews und Curricula sowie auf die computerunterstützte Textanalyse eines großen Korpus aus Stellenanzeigen kann Philippe Saner zeigen, dass sich die Datenwissenschaften als transversales Feld vielmehr dadurch auszeichnen, das in ihm Akteur*innen, Wissen und Bedeutungen aus unterschiedlichen wissenschaftlichen wie außerwissenschaftlichen Feldern aufeinandertreffen und dieses – teils in Konkurrenz, teils in Kooperation – prägen. Das analytische Modell der Arbeit integriert Praktiken von Grenz- und Begriffsarbeit mit einer auf Zwischenräumen basierten Konzeption von hybriden Wissensfeldern. So leistet die Arbeit auch einen wichtigen theoretischen Beitrag zu aktuellen Diskussionen der Feldtheorie, insbesondere zur Genese neuer Felder. Wie Saners Arbeit aufzeigt, entsteht gleichsam vor den Augen der wissenschaftlichen Beobachtung mit den Datenwissenschaften ein neues Feld, dessen weitreichende Bedeutung für Wissenschaft und Gesellschaft schon heute kaum überschätzt werden kann.

Mit seiner Arbeit, die 2021 mit dem Ulrich-Teichler-Preis der Gesellschaft für Hochschulforschung prämiert wurde, liefert Philippe Saner einen wichtigen soziologischen Beitrag zur Reflexion der Form und Rolle der Datenwissenschaften. *Datenwissenschaften und Gesellschaft. Die Genese eines transversalen Wissensfeldes* erweist sich so als vorzüglicher Auftakt dieser neuen transcript Reihe Digitale Soziologie, deren Gegenstände neue Forschungsfelder des Digitalen und ihre Geschichte ebenso umfassen wie klassische Fragestellungen der Soziologie. Saners Arbeit steht daher sowohl in ihrem Beitrag zur soziologischen Diskussion für sich als auch prototypisch für die Art und Form von Arbeiten, die wir in dieser Reihe versammeln wollen.

*Die Herausgeber*innen der Reihe Digitale Soziologie
im Mai des Jahres 2022.
Roger Häußling, Katharina Kinder-Kurlanda,
Sophie Mützel, Jan-Hendrik Passoth
und Andreas Schmitz*

Abbildungsverzeichnis

- Abbildung 1: Analytisches Modell – Räume zwischen Feldern | 61
- Abbildung 2: Methodische Vorgehensweise | 115
- Abbildung 3: Entwicklung der Stellenanzeigen im Korpus | 123
- Abbildung 4: Entwicklung der Stellenanzeigen nach Suchbegriffen | 124
- Abbildung 5: Wordclouds der häufigsten Begriffe nach ausgewählten Berufsgruppen | 126
- Abbildung 6: Vergleich der relativen Häufigkeiten von technologieindizierenden Themenbereichen in den Strategiedokumenten 1998–2020 | 151
- Abbildung 7: Relative Häufigkeiten ausgewählter Themenbereiche in den Strategiedokumenten 1998–2020 | 153
- Abbildung 8: Relative Häufigkeiten der Themenbereiche »Herausforderungen, Risiken« und »Chancen, Möglichkeiten« in den Strategiedokumenten 1998–2020 | 155
- Abbildung 9: Relative Häufigkeiten der Themenbereiche »Verhältnis Schweiz-Welt« sowie »Bund, Kantone, Verwaltung« in den Strategiedokumenten 1998–2020 | 165
- Abbildung 10: Das Data Science Venn Diagramm | 192
- Abbildung 11: Die Datenwissenschaften als transversaler Raum zwischen Feldern | 266
- Abbildung 12: Häufigste Begriffe der Stellentitel | 309
- Abbildung 13: Geschätzte relative Häufigkeit von Stellenanzeigen nach Topics | 310

Tabellenverzeichnis

- Tabelle 1: Übersicht der Studiengänge und Interviews nach institutionellem Kontext | 110
- Tabelle 2: Merkmale und Ausprägungen der Studiengänge im Sample | 111
- Tabelle 3: Die zehn am häufigsten inserierenden Organisationen im Sample | 128
- Tabelle 4: Die zehn häufigsten Standorte im Sample | 130
- Tabelle 5: Topic-Modell der englischsprachigen Stellenanzeigen über 20 Topics mit den 20 häufigsten Begriffen | 134
- Tabelle 6: Topic-Modell der deutschsprachigen Stellenanzeigen über 11 Topics mit den 20 häufigsten Begriffen | 141
- Tabelle 7: Curricula der Masterstudiengänge in Data Science der ETHZ und EPFL | 224
- Tabelle 8: Übersicht über Dimensionen, Codes und Zitate | 307
- Tabelle 9: Die zehn häufigsten Berufsgruppen im Sample | 308
- Tabelle 10: Sprachen der drei häufigsten Berufsbezeichnungen | 309
- Tabelle 11: Untersuchte Dokumente | 312
- Tabelle 12: Die Strategien »Informationsgesellschaft Schweiz« (1998, 2006, 2012) und »Digitale Schweiz« (2016, 2018) im Vergleich | 315
- Tabelle 13: Studiengänge und Vertiefungen in Datenwissenschaften an Schweizer Universitäten und Hochschulen | 316

Danksagung

An dieser Stelle möchte ich allen Menschen herzlich danken, die mich bei der Entstehung dieses Buchs inspiriert, unterstützt und begleitet haben. Zuallererst gilt mein Dank all jenen Personen, die sich für Interviews bereit erklärt und mit mir ihre Sichtweisen auf das vielschichtige Phänomen der Datenwissenschaften geteilt haben.

Sophie Mützel hat mir durch die Anstellung als wissenschaftlicher Mitarbeiter im Forschungsprojekt »Facing Big Data. Methods and skills needed for a 21st century sociology«, gefördert durch den Schweizerischen Nationalfonds, am Soziologischen Seminar der Universität Luzern die Möglichkeit geboten, ein Dissertationsprojekt auszuarbeiten und durchzuführen. Sie hat mir den nötigen wissenschaftlichen Freiraum gewährt und gleichzeitig durch viele Anregungen den Entstehungsprozess der Arbeit begleitet. Andrea Glauser danke ich für ihre Denkanstösse, die stets eine wertvolle Unterstützung waren, und die kritische Lektüre des Manuskripts. Beide haben mit ihren Fragen, Hinweisen und Anregungen massgeblich zur Fokussierung und zum Abschluss der Arbeit beigetragen.

Lisa Kressin und Rahel Estermann danke für ich den wissenschaftlichen und persönlichen Austausch sowie ihre hilfreichen Kommentare bei ersten Forschungsideen und Textentwürfen in der Erarbeitung der Dissertation. Désirée Waibel verdanke ich den Hinweis auf das theoretische Konzept von Räumen zwischen Feldern bei Gil Eyal. Methodische Unterstützung beim Web Scraping und Topic Modeling in R erhielt ich insbesondere von Tobias Philipp, Pablo Barberá und Alex Flückiger. Mein Dank gilt ferner den Kolleginnen und Kollegen am Soziologischen Seminar der Universität Luzern, die Teile der Arbeit mit mir diskutiert und wertvolle Anregungen gegeben haben: Markus Unternährer, Roman Gibel, Judith Nyfeler und Bernd Wurpts. Pius Zihlmann und Fabian Saner haben durch ihre äusserst genaue Lektüre und das Korrektorat des Manuskripts entscheidend zum Gelingen der Arbeit beigetragen. Schliesslich danke ich dem transcript Verlag, insbesondere Katharina Wierichs sowie Johanna Tönsing, für die Begleitung und Unterstützung bei der Umsetzung dieses Buchprojekts.

Meine Eltern haben meinen Werdegang bedingungslos unterstützt und meine Mutter hat insbesondere in der letzten Phase der Dissertation einen grossen Teil meiner väterlichen Betreuungsarbeit übernommen. Franziska Zihlmann danke ich von ganzem Herzen für ihre Liebe und die wiederholten Aufmunterungen bei der Durchführung und Fertigstellung dieses Projekts.

Gewidmet ist diese Arbeit meinem Sohn Navin Juri, der mich stets mit seinem herzlichen Lachen begrüsst und der durch seine eigenen Relevanzsetzungen die Herausforderungen wissenschaftlichen Arbeitens vergessen lässt.

Bern, im Februar 2022

Philippe Saner

Für Navin Juri

Kapitel 1 - Einleitung

Daten strukturieren das Soziale. Sie präformieren und repräsentieren uns und die soziale Wirklichkeit. Als Individuen ›lesen‹ wir die Welt durch die ›Brille‹ von Daten und werden umgekehrt in verschiedenen gesellschaftlichen Handlungssphären, in Wissenschaft, Medizin, Wirtschaft oder Politik, durch und mit Daten repräsentiert und ›gelesen‹ (Beer 2019). Die Corona-Pandemie macht exemplarisch deutlich, wie weitgehend die soziale Welt durch soziotechnische Systeme und Praktiken der Datenanalyse durchdrungen ist: Infektions- und Mortalitätszahlen, Reproduktionsraten, Contact Tracing, Visualisierungen, Dashboards und weitere soziotechnische Devices sind Elemente jenes datengestützten epidemiologischen Dispositivs, das zur Eindämmung der Corona-Pandemie etabliert wurde. Die gesellschaftliche Wahrnehmung und Bearbeitung der Pandemie vollziehen sich – jenseits der medizinischen Praxis – demnach primär durch Daten.

Ein zentrales Wissensgebiet für die Wahrnehmung bzw. Konstruktion sozialer Wirklichkeit durch und mit Daten sind die Datenwissenschaften, die den Untersuchungsgegenstand meiner Arbeit bilden. Die Datenwissenschaften sind überall: Als Wissensgebiet erlebten sie im letzten Jahrzehnt einen beschleunigten Aufstieg und die Diffusion in verschiedene gesellschaftliche Sphären wie kaum ein anderes Feld zuvor. Heutzutage beschäftigen sich so heterogene kollektive Akteure wie technische Hochschulen, die schweizerische Bundesverwaltung, global tätige Pharmaunternehmen oder Nichtregierungsorganisationen mit der Frage, was die Datenwissenschaften sind, wie sie sich zusammensetzen, worauf sie Anwendung finden können und welche Kompetenzen sie erfordern. In der vorliegenden Arbeit untersuche ich den Entstehungs- und Diffusionsprozess der Datenwissenschaften als neues Wissensfeld. Im Zentrum des Erkenntnisinteresses stehen dabei die sozialen Felder, Akteur*innen und deren Praktiken, die zur Herausbildung der Datenwissenschaften über verschiedene gesellschaftliche Sphären hinweg beitragen.

Die Wissenschaftsforschung verwendet seit mehreren Jahrzehnten den Begriff der Interdisziplinarität, um die Überschreitung wissenschaftlicher Teilgebiete zu beschreiben (Gibbons et al. 1994; Klein 1996; Nowotny et al. 2003). Der Begriff bezieht sich primär auf die Wissenschaften selbst, während Bezüge zu anderen sozialen Feldern seltener sind. Die Referenz auf unterschiedliche wissenschaftliche Disziplinen ist für die vorliegende Arbeit zwar ebenfalls bedeutend, allerdings geht es mir in der Arbeit gerade um das feldübergreifende, d. h. transversale Moment des Untersuchungsgegenstandes: So wird den Datenwissenschaften von vielen verschiedenen Akteur*in-

nen transformatives Potenzial zugeschrieben, das sie über die Wissenschaft hinaus entfalten würden.

Unter sozialen Feldern verstehe ich in Anlehnung an die soziologische Feldtheorie (Bourdieu & Wacquant 1996; Fligstein & McAdam 2012) gesellschaftliche Handlungsräume, in denen alle relevanten Akteur*innen und Institutionen in objektiven Relationen zueinander stehen (dazu Kapitel 2.2). Die Akteur*innen eines jeweiligen Feldes nehmen sich wechselseitig wahr, orientieren ihr Handeln aneinander und bilden gemeinsame Sinnsysteme, Regeln und Normen aus, die die jeweiligen Möglichkeiten und Praktiken rahmen. Die daraus resultierenden feldspezifischen »Spielregeln« (Bourdieu & Wacquant 1996: 126ff.) und Handlungslogiken prägen die weitere Entwicklung des Feldes. Zu den relevanten Feldern, die die Entwicklung der Datenwissenschaften massgeblich prägen, zähle ich in dieser Arbeit die Wissenschaft, die Hochschulbildung, die Ökonomie sowie die Bildungs- und Forschungspolitik.

Bevor ich auf einige elementare Charakteristika der Datenwissenschaften eingehe, geht es im Folgenden zunächst darum, das Erkenntnisinteresse der Arbeit zu umreißen und daraus abgeleitet die Fragestellungen einzuführen.

1.1 Wie entstehen transversale Wissensfelder?

Sowohl die Soziologie im Allgemeinen als auch die soziologische Wissenschafts- und Technikforschung im Speziellen beschäftigen sich mit der Frage nach der Herausbildung neuer Wissensgebiete. Die Wissenschaftsforschung hat dazu das Konzept der Disziplin geprägt. Das System wissenschaftlicher Disziplinen bildete sich im Laufe des 19. Jahrhunderts heraus (Abbott 2001; Stichweh 2013). Disziplinen sind stets durch Relationen zu wissenschaftsexternen Feldern geprägt: Beispielsweise sind Berufsrollen in der Ökonomie durch ihre Nachfrageorientierung mit disziplinären Kategorien gekoppelt, oder es bestehen Verknüpfungen zwischen Disziplinen und Fächereinteilungen im Bildungssystem (insbesondere auf Sekundarschulstufe) (Stichweh 1993: 244ff.). In einer historischen Perspektive werden die Beziehungen zwischen Institutionen der Wissenschaft und solchen anderer Felder als lose Kopplungen bezeichnet (Stichweh 2005; Weick 1976), was auf die relativ hohe Autonomie der Wissenschaft verweist.

In den letzten Jahrzehnten stiegen allerdings ökonomische und politische Anforderungen an die Organisationen im wissenschaftlichen Feld, durch Innovation, Wissenstransfer oder Wettbewerb mit industriellen oder regionalen Anspruchsgruppen zu kooperieren (Jongbloed et al. 2008). Diese als »third mission« benannte – neben Forschung und Lehre – dritte Aufgabe von Universitäten und Hochschulen gewann durch marktorientierte Steuerungsmechanismen, Medialisierung und Internationalisierung des Wissenschaftsfeldes an Bedeutung (Grande et al. 2013; Münch 2014; Zomer & Benneworth 2011). Solche wissenschaftsexternen Anforderungen wirken auch auf Disziplinen als epistemische Kerneinheiten zurück.

Die Emergenz neuer Forschungs- und Wissensgebiete ist durch vermeintlich konträre Prozesse von Differenzierung und Spezialisierung einerseits sowie Rekombination und Aggregation andererseits gekennzeichnet (Biniok 2013: 49ff.; Stichweh 2013): Neue Subdisziplinen und Spezialgebiete (*specialties*) differenzieren sich durch inhaltliche oder konzeptuelle Schwerpunktsetzungen sowie methodische Innovationen aus,

sodass sie kaum mehr inhaltliche Bezüge untereinander aufweisen. Disziplinen sind folglich die Summe solcher ausdifferenzierter Spezialgebiete (Lemaine et al. 1976: 2). Die Binnendifferenzierung von Wissenschaft stellt zudem einen Mechanismus dar, mit dem das Feld auf starkes Wachstum und zunehmende Konkurrenz in stark beforschten Gebieten reagiert (Weingart 2001). Andererseits werden solche Spezialgebiete immer wieder zu umfassenden thematischen und disziplinären Zusammenhängen kombiniert und aggregiert. Neue Wissensgebiete kombinieren demnach ähnliche Forschungsobjekte und Problemstellungen, aber auch Wissensbestände und Methoden unterschiedlicher Disziplinen, was ihre Hybridität ausmacht (Bösch 2016; Klein 1996).¹ Demgegenüber bezeichnet Aggregation die Kopplung mehrerer Spezialgebiete, wobei »trotz gemeinsamer Interaktion der Akteure [...] überdauernde Heterogenität« bestehen bleibt (Biniok 2013: 53).

Während sich die frühe Wissenschaftsforschung am Wachstumsparadigma stetig steigender Publikationen, personeller und finanzieller Ressourcen orientierte (Lemaine et al. 1976), fand allmählich ein Übergang hin zu fach- bzw. disziplinspezifischen Differenzen in den Expansionsraten sowie Fragen der Ressourcenverteilung statt (Merz & Sormani 2016). So rückten auch die Kooperations- und Konkurrenzverhältnisse von existierenden und neuen Wissensfeldern in den Blick. Damit einhergehend wandelte sich das primäre Erkenntnisinteresse von der Frage nach dem *Wann* und *Warum* der Genese neuer Wissensfelder hin zur Frage nach dem *Wie* (ebd.: 9) – ein Primat, das ich in dieser Arbeit teile.

Die jüngere Wissenschafts- und Technikforschung verweist darauf, dass neue Wissensgebiete bzw. Expertisefelder als Grenzgebiete *zwischen* verschiedenen Disziplinen bzw. Wissensfeldern entstehen (Eyal 2013; Eyal & Pok 2015; Marcovich & Shinn 2011; Méndez Fernández & Passoth 2019; Shinn & Joerges 2002). Interdisziplinäre Wissensgebiete wie die Materialwissenschaften (Bensaude-Vincent 2001; Bensaude-Vincent & Hessenbruch 2004), Nanowissenschaften (Biniok 2013) oder Klimawissenschaften (Allan 2017; Edwards 2001) differenzierten sich aus, indem einerseits neue Themen erschlossen und Instrumente oder Methoden entwickelt wurden und andererseits neue Gruppen von Wissenschaftler*innen hinzukamen und grenzüberschreitende Kooperationen dauerhaft institutionalisiert wurden. Trotzdem behalten die involvierten Forschenden ihre disziplinären Identitäten als – im Falle der Materialwissenschaften – Physikerinnen, Chemiker oder Ingenieurinnen (Bensaude-Vincent & Hessenbruch 2004: 347). Lettkemann (2016: 38) spricht in diesem Zusammenhang von »stabiler Interdisziplinarität«, wenn neue Wissensgebiete ihre hybride Zusammensetzung über disziplinäre Grenzen nicht nur behalten, sondern auch über Generationen von Forschenden hinweg reproduzieren können.

Durch den Fokus auf feld- bzw. grenzüberschreitende Interaktionen distanzieren sich Arbeiten, die die Transversalität in der Herausbildung von Wissensfeldern betonen, von postmodernen Entdifferenzierungstheorien (Gibbons et al. 1994; Nowotny et al. 2003), wonach die Wissensproduktion von disziplinär strukturierten Einheiten und Grundlagenforschung (Mode 1) in einen von disziplinären Schranken entbundenen, an ausserwissenschaftlichen Fragestellungen und Anwendungen orientier-

¹ Burren (2010: 12) spricht am Beispiel der Betriebswirtschaftslehre von einer »hybriden Disziplin«, die in einem »Spannungsfeld von Wissenschaftlichkeit und Praxisorientierung« entstanden ist und bis heute operiert.

ten Modus (Mode 2) übergehe. Empirische Studien legen dabei nahe, dass die starke Zunahme inter- und transdisziplinärer Projekte die disziplinären Referenzsysteme keineswegs obsolet macht (Lettkemann 2016: 32): Vielmehr werden Disziplinen über Grenzen hinweg miteinander in Beziehung gesetzt, ohne dass sie dadurch ihre kulturellen Identitäten verlieren würden. Wissensgebiete bzw. wissenschaftliche Disziplinen sind als heterogene, dynamische Einheiten zu denken, die regelmässigen Wandlungen, Grenzverschiebungen und Neuformierungen unterliegen. Disziplinen verstanden als disziplinäre Felder im Sinne Bourdieus bieten die Elastizität und Wandelbarkeit, können gleichzeitig aber auch als Gefässe und Mechanismen der Reproduktion der Strukturen des wissenschaftlichen Feldes konzipiert werden. In diesem Sinne differenzieren sich Wissensgebiete als disziplinäre Felder aus, die allerdings stets nur relativ autonom sind (Schultheis 2005) – gegenüber anderen disziplinären Feldern der Wissenschaft, aber auch den Feldern der Politik, Ökonomie, Hochschulbildung und weiteren.

Die Anzahl interdisziplinärer Wissensgebiete (Van Noorden 2015) ist in den letzten Jahrzehnten rapide gewachsen und ihre epistemischen Praktiken wurden in eine Vielzahl anderer gesellschaftlicher Felder übertragen. Daraus haben sich veränderte Formen des Austausches und der Kollaboration zwischen verschiedenen Praxissphären etabliert. Die entstehenden Wissensgebiete werden als verteilt, hybrid und vernetzt beschrieben und lassen sich kaum noch einzelnen Disziplinen oder gesellschaftlichen Sphären zuordnen (Passoth & Rammert 2016; Witte & Schmitz 2019). Veränderte Expertisen transformieren zudem organisationale Praktiken, die allerdings in den einzelnen Feldern unterschiedlich implementiert werden. Dabei werden sie neben disziplinären Differenzierungen und veränderten Bedingungen der Wissensproduktion auch grundlegend von Motiven und Strategien ökonomischer und (wissenschafts-) politischer Akteur*innen gerahmt (Merz & Sormani 2016).

An diese Arbeiten der Wissenschafts- und Technikforschung anknüpfend, untersuche ich in der vorliegenden Arbeit die Herausbildung und Etablierung der Datenwissenschaften als transversales Wissensfeld, das Elemente verschiedener sozialer Felder umfasst. Das Erkenntnisinteresse richtet sich dabei auf die Austauschbeziehungen und Verknüpfungen zwischen den sozialen Feldern der Wissenschaft, Hochschulbildung, Ökonomie und Politik mit den jeweiligen – kollektiven wie individuellen – Akteur*innen, die alle ihre eigenen Definitionen, Bedeutungszuschreibungen oder Zukunftsszenarien der Datenwissenschaften entwerfen. Dabei stellt sich unter anderem die Frage, wie die Akteur*innen in den unterschiedlichen Untersuchungsfeldern zur Etablierung der Datenwissenschaften beitragen: Welche neuen Begriffe führen sie zur Konzeptualisierung des Wissensgebietes ein? Inwiefern schliessen sie an bestehende disziplinäre und feldspezifische Wissensinhalte an? Öffnen sie sich gegenüber anderen Akteur*innen und Feldern oder versuchen sie die Deutungshoheit über die neue Form von Expertise für sich zu beanspruchen und andere davon auszuschliessen? Zusammengefasst leiten folgende Forschungsfragen meine empirische Untersuchung:

Wie entstehen neue Wissensgebiete über verschiedene soziale Felder hinweg?

Welche Akteur*innen sind wichtig bei der Genese von entstehenden Wissensfeldern?

Wie, d. h. durch welche Praktiken tragen sie zur Etablierung von Wissensgebieten bei?

Obwohl die Datenwissenschaften ein globales Wissensfeld darstellen, zeichnen sich gleichzeitig nationalstaatliche, regionale und lokale Differenzierungen und Besonderheiten ab (Biniok 2013; Merz & Sormani 2016; Saxenian 1996). Ich fokussiere deshalb primär die Herausbildung und Bedeutungskonstruktionen der Datenwissenschaften in der Schweiz: Die Struktur und Repräsentationen der Datenwissenschaften im schweizerischen Arbeitsmarkt, die Zukunftsvisionen über Digitalisierung und Datenwissenschaften im Feld der schweizerischen Bildungs- und Hochschulpolitik sowie die Etablierung von Forschungseinrichtungen und Studiengängen an Schweizer Hochschulen dienen als Untersuchungsfelder. Die nachfolgend skizzierten Prozesse bei der Herausbildung der Datenwissenschaften legen dar, dass diese einen geeigneten empirischen Fall darstellen, um die Genese einer interdisziplinären, feldübergreifenden Wissensformation in actu zu untersuchen.

1.2 Die Datenwissenschaften als transversales Phänomen

Die Datenwissenschaften werden je nach Verortung in den genannten Feldern sehr unterschiedlich beschrieben. Trotzdem können einige gemeinsame Charakteristika festgehalten werden, auf die ich in dieser Arbeit eingehen werde (dazu Kapitel 3): Datenwissenschaften sind hybrid, sie setzen sich aus verschiedenen wissenschaftlichen Disziplinen und Wissensgebieten zusammen. Sie finden Anwendung auf diverse wissenschaftliche und nichtwissenschaftliche Felder. Sie stützen sich elementar auf unterschiedliche Technologien und Methoden zur Prozessierung und Analyse von Daten. Ihnen wird das Potenzial zugeschrieben, grosse Herausforderungen der Menschheit bewältigen zu können. Und schliesslich erfordern sie neue Qualifikationen, Fähigkeiten und persönliche Eigenschaften, die verschiedenen Wissensbereichen entstammen.

Die kurze Darstellung illustriert zweierlei: Zum einen konstruieren sowohl wissenschaftliche als auch nichtwissenschaftliche Akteur*innen die Datenwissenschaften. Trotz der Referenz auf die Wissenschaft lässt sich das Wissensgebiet nicht auf eine feldspezifische Logik, etwa der Wissenschaft oder der Industrie zurückführen, sondern muss als Mischform, als ein hybrides Feld betrachtet werden. Zum anderen werden die Datenwissenschaften als multidimensionales und vielstimmiges Wissensgebiet repräsentiert.

Zum Begriff: Ich verwende in der vorliegenden Arbeit den (selten verwendeten) deutschsprachigen Begriff der »Datenwissenschaften« im Plural, der gegenüber der im Untersuchungsfeld dominierenden englischsprachigen Bezeichnung »Data Science« diskursiv die Vielstimmigkeit und Mehrfachzugehörigkeit des Wissensgebiets andeutet. In diesem Sinne verstehe ich Datenwissenschaften nicht – wie manche Definitionen – als singuläre Entität mit klaren disziplinären Grenzen, sondern als ein multidimensionales Phänomen, das bestehende Grenzen überschreitet und sich auf epistemische Praktiken, Expertisen, Methoden, Instrumente und weitere kognitive Elemente verschiedener sozialer Felder abstützt.

Um die Genese von transversalen, interdisziplinären Wissensfeldern in soziologischen Begrifflichkeiten fassbar zu machen, stütze ich mich auf das Konzept von Räumen zwischen Feldern bei Gil Eyal (2013; Eyal & Pok 2015). Zwischenräume sind offene, unterbestimmte Sphären, in denen Akteur*innen aus nahen Feldern aufein-

andertreffen und in Interaktion treten können, d. h. sowohl in Konkurrenz- als auch in Kooperationsverhältnissen stehen können. Es handelt sich um Räume ausserhalb etablierter Handlungslogiken, die sich durch poröse Grenzen, Vielstimmigkeit sowie eine permanente Instabilität auszeichnen. Eyals Konzept von Räumen zwischen Feldern bietet eine Möglichkeit, die Entstehung neuer Wissensgebiete ausserhalb der etablierten »Spielregeln« sozialer Felder zu situieren. Dabei nehme ich eine praxisorientierte Perspektive ein, um deren Herausbildung und Permanenz von Zwischenräumen zu erklären: Erstens halten Praktiken von Grenzarbeit (*boundary work*) Räume zwischen Feldern offen und durchlässig. Sie ermöglichen es, dass Akteur*innen mit unterschiedlichen Motiven und Strategien darin agieren können. Zweitens befördern diskursive Praktiken von Begriffsarbeit multiple Verständnisse von Gegenständen und tragen so zur Dauerhaftigkeit hybrider Zwischenräume bei.

Die Datenwissenschaften bilden sich synchron in verschiedenen sozialen Feldern heraus und werden von und in diesen geprägt. Dadurch sind sie in ihrer relativen Autonomie von multiplen Effekten anderer Felder durchzogen (Passoth & Rammert 2016; Witte & Schmitz 2019). Der bisherige Forschungsstand indiziert, dass insbesondere die sozialen Felder der Wissenschaft, Hochschulbildung, Ökonomie sowie der Bildungs- und Forschungspolitik die Herausbildung der Datenwissenschaften massgeblich befördert und geprägt haben. Sie dienen deshalb in dieser Arbeit als empirische Untersuchungsfelder: Auf der Makroebene ist dies zunächst der Arbeitsmarkt. Unternehmen und andere Organisationen publizieren Stelleninserate, legen Tätigkeits- und Kompetenzprofile fest und formulieren organisations- und feldspezifische Perspektiven, worin die Datenwissenschaften bestehen, welche Disziplinen sie umfassen und welche Qualifikation und Fertigkeiten sie erfordern. Zweitens ist es das Feld der Bildungs- und Forschungspolitik: Staatliche Institutionen, Interessensgruppen und weitere Akteur*innen entwerfen in Strategiepapieren und Aktionsplänen kollektive Visionen und Szenarien einer vielversprechenden Zukunft, die durch die Datenwissenschaften möglich werden soll. Sekundiert werden solche diskursiven Prioritätensetzungen durch substanzielle Investitionen, die die Bedeutung der geförderten Wissens- und Forschungsfelder auch in materieller Hinsicht unterstreichen. Schliesslich ist es drittens das akademische Feld: Hier entwerfen Forschende als individuelle sowie Universitäten und Hochschulen als kollektive Akteur*innen disziplinäre bzw. epistemische Perspektiven auf die Datenwissenschaften und adressieren sowohl wissenschaftliche Praktiken als auch deren Vermittlung in der Hochschulbildung.

Durch solche kollektiven Praktiken von Grenz- und Begriffsarbeit streben die genannten Akteur*innen danach, bestimmte Sichtweisen auf die Datenwissenschaften gegenüber anderen möglichen Bedeutungen hervorzuheben und zu etablieren: So delegitimieren sie auf der einen Seite konkurrierende Sichtweisen, grenzen sich gegenüber anderen ab und schliessen somit potenzielle Kollaborationen aus. Andererseits überschreiten sie durch offene Begriffe aber auch disziplinäre und soziale Grenzen. Sie schaffen damit Anschlussfähigkeit für Akteur*innen anderer Felder, was deren Möglichkeitsräume erweitert und Potenzial für feldübergreifende Kollaborationen und Kooperationen bietet. Durch solche symbolischen Deutungen tragen sie zur Konstruktion der Datenwissenschaften als heterogenes, vielstimmiges Wissensfeld bei.

Aufgrund des transversalen Moments des Untersuchungsgegenstandes verfolge ich in der vorliegenden Arbeit den Ansatz einer verteilten Analyse von Zwischenräumen:

Dabei untersuche ich ausgewählte Ausschnitte der drei genannten Untersuchungsfelder im Hinblick auf das interessierende Phänomen, die Genese der Datenwissenschaften. Es geht mitunter also darum, den jeweiligen Beitrag eines sozialen Feldes bei der Herausbildung der Datenwissenschaften zu identifizieren. Die Datenwissenschaften sind kein ausschliesslich wissenschaftliches, ökonomisches, bildungsbezogenes oder politisches Phänomen, sondern vielmehr über die verschiedenen Felder verteilt und fragmentiert zugleich: Es manifestieren sich darin Effekte der Praxisformen verschiedener Felder, indem sie gewissermassen zu deren – feldanalytisch formuliert – »Spielball« werden. Das Interesse der Arbeit gilt somit neben den Akteur*innen den Wirkungen der skizzierten Praktiken von Begriffs- und Grenzarbeit, und zwar nicht nur innerhalb der Felder der Wissenschaft, Ökonomie sowie Bildungs- und Forschungspolitik, sondern auch in den Räumen dazwischen.

Charakteristisch für die Vielstimmigkeit und Mehrfachverortung entstehender Wissensgebiete ist die zu beobachtende Suche nach den »richtigen« Begriffen und Kategorien. Empirisch zeigt sich dies insbesondere im Changieren zwischen verschiedenen Disziplinen, Wissensbeständen, Werkzeugen und Methoden. Das heisst, es geht in der Herausbildung des neuen Wissensgebiets fundamental um Such- und Rekrutierungsprozesse, Taxonomien und Unsicherheiten über die zirkulierenden Begrifflichkeiten. Damit markieren widersprüchliche Spannungsverhältnisse die Emergenz der Datenwissenschaften als neues Wissensfeld: Auf der einen Seite zeigen sich kollaborative Prozesse und disziplinäre Grenzüberschreitungen in der Suche nach Neuheit und nach den »adäquaten« Begriffen. Diese werden von der Absicht begleitet, bestehende organisationale Gefässe und Strukturen zu transformieren. Andererseits werden darin Prozesse von Institutionalisierung und Kanonisierung manifest, d. h., es geht um die beschleunigte Verarbeitung der neuen Begriffe, Kooperationsarrangements sowie Kompetenzprofile in bestehenden disziplinären und organisationalen Strukturen durch Profilbildungen und Grenzziehungen.

Diese Synchronizität und Widersprüchlichkeit multipler Such- und Bestimmungsprozesse einerseits und von Institutionalisierungs- und Kanonisierungsprozessen andererseits in der Emergenz transversaler Wissensfelder ist ein zentrales Ergebnis der Studie. Sie äussert sich in den einzelnen Untersuchungsfeldern auf je unterschiedliche Weise: Im Arbeitsmarkt stehen einer erheblichen Konzentration der Stelleninserate auf einige wenige ökonomische Felder und geographische Standorte feldübergreifende Suchprozesse gegenüber. Die Stelleninserate sind gekennzeichnet durch Aufzählungen von Qualifikationen, Wissensbeständen, Methoden und Tools. Obwohl die Listen unterschiedliche Praktiken, Instrumente oder Konzepte zusammenfassen und ein breites Spektrum abdecken, deutet die Analyse an, dass sich in Stelleninseraten gewisse Begriffe etabliert haben, die feldübergreifend und sprachunabhängig Verwendung finden. Die Offenheit und Ambiguität der verwendeten Begriffe erweisen sich als konstitutiv für das untersuchte Arbeitsmarktsegment.

Ähnlich manifestiert sich dies im Feld der Bildungs- und Forschungspolitik: In Strategiedokumenten werden bestimmte Begriffe (wie Digitalisierung, Innovation oder Kompetenzen) eingeführt, die bestehende Diskurse zum Verhältnis von Technologie und Gesellschaft überschreiben. Trotz der neuen Begrifflichkeiten sowie einer permanenten Transformationsrhetorik verweisen die Strategiedokumente und die damit einhergehenden politischen Massnahmen auf Permanenz und die Orientierung an etablierten politischen Konzepten (wie »Wohlstand«, »Standortwettbewerb« oder

»Fachkräftemangel«). Die Offenheit und die Mehrdeutigkeit der gewählten neuen Begriffe stellen dabei ein verbindendes Element dar: Sie erlauben Akteur*innen anderer Felder, sich strategisch darauf zu beziehen und feldübergreifende Kollaborationen einzugehen.

Im akademischen Feld werden die Synchronizität und Widersprüchlichkeit der genannten Prozesse auf unterschiedlichen Ebenen deutlich: Zunächst lösen die Datenwissenschaften eine intensive Suche nach neuen Begriffen aus, um die datengetriebene Transformation von Wissenschaft adäquat zu beschreiben. Dazu werden einerseits gemeinsame Kategorien und Gefässe geschaffen, d. h. etablierte Grenzen zwischen wissenschaftlichen Disziplinen überschritten. Gleichzeitig erweisen sich sowohl disziplinäre als auch feldspezifische Grenzziehungen (zwischen Wissenschaft und Industrie oder zwischen akademischer Grundlagen- und anwendungsorientierter Forschung) in den Stellungnahmen von Lehrenden und Studiengangleitenden als äusserst persistent.

Auf der Ebene der Organisationen im akademischen Feld werden ausser Studiengängen kaum neue Gefässe geschaffen. Die Einführung der Datenwissenschaften als neue Kategorie im Schweizer Hochschulsystem vollzieht sich innerhalb der bestehenden organisationalen und disziplinären Strukturen an Universitäten und Hochschulen. Trotzdem beeinflussen politische, organisationale und ökonomische Überlegungen die Positionierungen der Akteur*innen im Feld. Zudem deutet die Orientierung an anderen Akteur*innen die Synchronizität vielfältiger Konkurrenz- und Kooperationsverhältnisse, d. h. von feldspezifischen Grenzziehungen und Grenzüberschreitungen, an.

Auf der Ebene der Curricula sind trotz unterschiedlicher Hochschultraditionen und disziplinärer Verankerungen ausgeprägte Kanonisierungsprozesse zu einem frühen Zeitpunkt in der Entstehung des Feldes zu beobachten. Diese prägen sowohl die Inhalte als auch die Struktur der Studiengänge: So lehnt sich die curriculare Differenzierung in einen interdisziplinären Kernbereich, in Wahlbereiche (*domains*), Praxismodule sowie komplementäre Inhalte eng an bestehende technikwissenschaftliche Ausbildungstraditionen an und schreibt diese fort.

Schliesslich machen die Curricula und Interviews auch feldübergreifende Suchprozesse nach den »richtigen« Kompetenz- sowie Berufsprofilen deutlich: Die untersuchten Studiengänge schaffen neue Kompetenzbereiche und kombinieren sie im Baukastenprinzip miteinander. Die neuen Skillsets, sogenannte »Datenkompetenzen«, gehen mit Narrativen von Offenheit und Demokratisierung einher. Trotz solcher Versprechungen bleiben sie an existierenden Ausbildungsstrukturen orientiert und reproduzieren damit soziale Ungleichheiten, die in diese eingeschrieben sind. Die Analyse zeigt ferner, dass sich die Kompetenzprofile auf ein mentales Dispositiv abstützen, das im Feld als »Data Mindset« gefasst wird: Es handelt sich dabei um Denkweisen und persönliche Überzeugungen, wissenschaftliche und andere Problemstellungen datengestützt bearbeiten zu wollen. Diese bilden die kulturellen und kognitiven Grundlagen dafür, dass sich die Selbstwahrnehmung von Gesellschaft durch und mit Daten etablieren kann.

1.3 Beiträge der Arbeit

Die Untersuchung unternimmt eine umfassende Studie der Herausbildung der Datenwissenschaften als soziales Phänomen, die über Einzelfallstudien hinausgeht. Sie macht deutlich, dass es sich bei den Datenwissenschaften – trotz ihrer einschlägigen Bezeichnung – nicht in erster Linie um eine neue wissenschaftliche Disziplin handelt: Vielmehr sind eine Vielzahl von Akteur*innen in unterschiedlichen Feldern in die Emergenz der Datenwissenschaften involviert. Somit richte ich den Fokus der Arbeit über die üblichen Perspektiven disziplinärer oder feldspezifischer Differenzierung hinaus auf die Akteur*innen, Praktiken sowie Aushandlungs- und Koordinationsprozesse, die die Herausbildung der Datenwissenschaften fundieren und befördern.

Sozialtheoretisch integriert das vorgeschlagene analytische Modell Praktiken von Grenz- und Begriffsarbeit in eine zwischenräumliche Konzeption von hybriden Wissensfeldern. Somit folgt die Arbeit der Tradition der konstruktivistisch argumentierenden soziologischen Wissenschafts- und Technikforschung. Eine stärker akteur- und praxistheoretische Betrachtung von Räumen zwischen Feldern erweitert das analytische Erklärungspotenzial auf kollektive Bedeutungs- und Sinnkonstruktionen, indem sie neben den feldspezifischen Strategien und Interessen auch die beobachtbaren Praktiken von Begriffs- und Grenzarbeit berücksichtigt, die die Datenwissenschaften in den involvierten Feldern rahmen und strukturieren. Damit pluralisiert das Modell das Begriffsinstrumentarium existierender soziologischer Feldtheorien und stärkt insbesondere deren Potenzial zur Erklärung der Genese neuer Wissensgebiete.

Empirisch rekonstruiere ich in der vorliegenden Arbeit die Herausbildung der Datenwissenschaften am Beispiel ihrer spezifischen Konfiguration in der Schweiz. Eine verteilte Analyse der Datenwissenschaften als Zwischenraum erlaubt es, den Effekten von Begriffs- und Grenzarbeit durch die einzelnen Untersuchungsfelder »zu folgen«. Damit richte ich den Blick auf das transversale Moment in der Herausbildung der Wissensformation: Die Analyse erschliesst die multiplen Bedeutungskonstruktionen der Datenwissenschaften auf unterschiedlichen gesellschaftlichen Ebenen, von der feldübergreifenden Ebene des Arbeitsmarktes über feldspezifische Verortungs- und Organisationsprozesse in der Bildungs- und Forschungspolitik, Wissenschaft und Hochschulbildung bis hin zu subjektorientierten Kompetenzkonstruktionen in Curricula und Interviews.

Methodisch verknüpfe ich sozialwissenschaftliche Methoden der qualitativen Inhalts-, Curricula- und Interviewanalyse mit dem Topic Modeling von Stellenanzeigen, einem neueren Verfahren der computerlinguistischen Textanalyse. Die Triangulation verschiedener Datentypen, Erhebungs- und Auswertungsmethoden ermöglicht es, die feldübergreifenden Bezüge und Darstellungen zu identifizieren sowie Repräsentationen der Datenwissenschaften zu analysieren. Indem ich den Untersuchungsgegenstand durch Stellenanzeigen – eine bisher in der Soziologie kaum genutzte Datenquelle – erschliesse und mit einem computergestützten Verfahrens auswerte, erweitere ich zudem das soziologische Methodenrepertoire und leiste einen innovativen Beitrag zur Arbeitsmarkt- und Wirtschaftssoziologie.

Schliesslich trägt die Arbeit zu einem vertieften soziologischen Verständnis der sozialen Welt als einer Datengesellschaft (Houben & Prietl 2018; Süssenguth 2015) bei, für die die Datenwissenschaften das zentrale kognitive Element bilden. Durch die

Rekonstruktion feldspezifischer Prozesse der Entstehung und Diffusion der Datenwissenschaften in den einzelnen Untersuchungsfeldern und der Analyse von deren Wechselwirkungen leiste ich einen wichtigen Beitrag zum Verständnis der Funktionsweisen von zeitgenössischen Datengesellschaften.

1.4 Aufbau der Arbeit

Die Arbeit ist in vier Teile gegliedert. Im ersten Teil lege ich die theoretischen und methodischen Grundlagen, um die beschriebenen Fragestellungen der Arbeit unter Rückgriff auf das analytische Modell und empirisch fundiert beantworten zu können.

Im Kapitel 2 erarbeite ich ein theoretisches Modell, um durch die Verknüpfung verschiedener theoretischer Perspektiven die feldübergreifende Entstehung neuer Wissensgebiete zu konzeptualisieren. Den Ausgangspunkt bildet die Praxistheorie Bourdieus, die ich durch ergänzende Theorieelemente erweitere, um die transversale Etablierung neuer Wissensfelder analytisch fassen zu können. Das resultierende analytische Modell lokalisiert die Entstehung neuer Wissensgebiete in einem Raum zwischen Feldern, der jenseits der hierarchischen Logiken sozialer Felder verfasst ist. Durch Praktiken von Grenz- und Begriffsarbeit strukturieren die Akteur*innen die Herausbildung neuer Wissensgebiete in solchen Zwischenräumen.

Im Kapitel 3 rekapituliere ich wichtige Arbeiten zum epistemischen Feld der Datenwissenschaften. In einem ersten Schritt skizziere ich die verschiedenen theoretischen Perspektiven, mit denen die Etablierung der Datenwissenschaften interpretiert wird. In einem zweiten Schritt untersuche ich, wie Akteur*innen in den verschiedenen Untersuchungsfeldern selbst durch die Konstruktion von Kompetenzen und Tätigkeitsprofilen zur Genese der Datenwissenschaften beitragen. Abschliessend fasse ich in einem Zwischenstand die wichtigsten Punkte des dargestellten Forschungsstandes zusammen.

Im Kapitel 4 führe ich das Forschungsdesign und die methodischen Vorgehensweisen aus, um die Fragestellungen empirisch zu untersuchen. Ich verorte die Arbeit im allgemeinen Forschungsprogramm der Feldanalyse und erläutere den Ansatz einer verteilten Analyse von Zwischenräumen. Darauf folgt die Beschreibung der Datengrundlage und des methodischen Vorgehens der drei empirischen Kapitel: erstens die Erhebung und Auswertung der Stellenanzeigen im Arbeitsmarkt, zweitens die Strategiedokumente im Feld der Politik und das inhaltsanalytische Vorgehen und drittens die Inhaltsanalyse von Curricula sowie Interviews mit Lehrenden der Datenwissenschaften im akademischen Feld. Abschliessend fasse ich das empirisch-analytische Potenzial des gewählten Forschungsdesigns zusammen.

Der zweite und der dritte Teil umfassen die empirischen Kapitel der Arbeit. Im zweiten Teil gehe ich zunächst den unterschiedlichen Bedeutungskonstruktionen und Zukunftsvorstellungen nach, durch welche den Datenwissenschaften im Arbeitsmarkt sowie in der Hochschul- und Forschungspolitik Sinn zugeschrieben wird.

Das Kapitel 5 widmet sich einerseits der Struktur und Entwicklung des Arbeitsmarktsegments Datenwissenschaften, andererseits den zentralen Berufsprofilen, Begrifflichkeiten und Deutungen, die in unterschiedlichen Organisationen und Feldern entworfen werden. Die Analyse von Stellenanzeigen für Data Scientists und verwandte Berufsbezeichnungen im schweizerischen Arbeitsmarkt ermöglicht es, die Konst-

ruktionsleistungen unterschiedlicher Akteur*innen zu untersuchen. Es geht weniger um eine umfassende soziologische Analyse des Arbeitsmarktsegments, sondern um die zentralen Organisationen, Felder und Profile, die die Datenwissenschaften als neues Berufsfeld charakterisieren. Das computergestützte Topic-Modeling-Verfahren identifiziert dabei eine grosse Vielfalt unterschiedlicher Konnotationen, durch die die Datenwissenschaften in den englisch- sowie deutschsprachigen Stellenanzeigen repräsentiert werden.

Im Kapitel 6 untersuche ich, wie bildungs- und forschungspolitische Akteur*innen durch gesellschaftliche Zukunftsentwürfe die Genese und Sinnkonstruktion der Datenwissenschaften als Wissensfeld rahmen und beeinflussen. Allgemein stellt sich die Frage, wie der politische Diskurs zur Digitalisierung operiert und wie Bildung und Forschung darin gerahmt werden. Spezifisch frage ich danach, welche Zukünfte über Daten bzw. Datenwissenschaften Akteur*innen der Bildungs- und Forschungspolitik entwerfen. Dabei skizziere ich zunächst zentrale Elemente und Operationsmodi des politischen Diskurses zur Digitalisierung, analysiere anschliessend unterschiedliche Zukunftsentwürfe von Bildung und Forschung in den Strategiedokumenten zur Bildungs- und Forschungspolitik. Schliesslich untersuche ich die Rolle von Daten bzw. Datenwissenschaften in den soziotechnischen Visionen und welche Bedeutungen und Zukunftsszenarien diesen zugesprochen werden.

Im dritten Teil der Arbeit gehe ich vertieft auf die Rolle des akademischen Feldes bei der Genese der Datenwissenschaften ein. Im Zentrum stehen jene Praktiken, die die Datenwissenschaften im akademischen Feld konstruieren.

Im Kapitel 7 frage ich danach, wie die Datenwissenschaften im akademischen Feld eingeführt werden und in welchem Verhältnis sie zu verwandten Disziplinen stehen. Die Analyse identifiziert die Bedeutung der beiden zentralen Praxismodi von Begriffs- und Grenzarbeit: Die Einführung der Datenwissenschaften vollzieht sich durch multiple Definitionen, Kategorisierungen und Narrative, in denen sich sowohl Unklarheiten über zukünftige Entwicklungen als auch Suchprozesse darüber manifestieren, wie die beobachtbaren Veränderungen im akademischen Feld »adäquat« beschrieben werden können. Praktiken von *boundary work* artikulieren sich im empirischen Material in intensiven Konfliktlinien zwischen den involvierten Disziplinen. Zugleich transzendiert die dominante Konzeption der Datenwissenschaften, d. h. die Differenzierung in eine interdisziplinäre methodisch-technische Expertise und eine domänenspezifische Fachexpertise, etablierte Grenzen zwischen verschiedenen Disziplinen sowie wissenschaftlichen und nichtwissenschaftlichen Anwendungsbereichen.

Im Kapitel 8 untersuche ich die Verhandlung der Datenwissenschaften in Universitäten und Hochschulen als kollektiven Akteur*innen vor dem Hintergrund hochschulpolitischer Regelwerke, organisationaler Abläufe und ökonomischer Anreizstrukturen. Dabei stellt sich die Frage, welche wissenschaftlichen und ausserwissenschaftlichen Faktoren die Planung und Ausgestaltung der Curricula beeinflussen. Welche Akteur*innen und Disziplinen sind an der Ausarbeitung beteiligt? Es zeigt sich, dass komplexe Wechselwirkungen organisationaler, feldspezifischer und anderer Einflussfaktoren die Möglichkeits- und Handlungsräume der untersuchten Organisationen im akademischen Feld beim Aufbau neuer Studiengänge strukturieren und prägen. Neben der Analyse der ökonomischen, hochschulpolitischen und organisationalen Rahmenbedingungen liegt ein weiterer Fokus auf den feldspezifischen Strategien und Positionierungen von Universitäten und Organisationen, in denen sich

die synchronen Konkurrenz- und Kooperationsverhältnisse des akademischen Feldes insgesamt manifestieren.

Das Kapitel 9 fragt danach, wie die Datenwissenschaften an Schweizer Universitäten und Hochschulen konzeptualisiert sind, d. h. wie die Inhalte des interdisziplinären Wissensfeldes in Curricula implementiert und gelehrt werden. Dazu untersuche ich die zentralen Bestandteile zeitgenössischer Curricula in den Datenwissenschaften wie Zulassungsvoraussetzungen, die Definition eines Kernbereichs, Wahlbereiche und Wahlfächer, praxisorientierte Kurse, komplementäre Inhalte und schliesslich das Subfeld der Weiterbildung, das sich von den Regelstudiengängen erheblich unterscheidet. Der identifizierte Aufbau der untersuchten Curricula ist charakteristisch für technikwissenschaftliche Wissensfelder, die nicht primär die Ausbildung methodisch und technisch versierter Wissenschaftler*innen anstreben, sondern durch »ergänzende Massnahmen« entscheidungsfähige Praktiker*innen formen möchten.

Im Kapitel 10 schliesslich stehen individuelle Zuschreibungen an (künftige) Data Scientists im Zentrum des Erkenntnisinteresses. Wie werden die Lehrinhalte in individuelle Zuschreibungen übersetzt und an welche bestehenden Kategorien und Wissensbestände lehnen sich diese an? Die Diskussionen über die jeweiligen Qualifikationen und Fähigkeiten, die für eine Tätigkeit in Datenwissenschaften als notwendig erachtet werden, machen Suchprozesse über die Bedeutung und Sinnhaftigkeit von Datenwissenschaften als Wissensfeld manifest. Auf der curricularen Ebene wird aus der Schnittstellenfunktion von Data Scientists zwischen verschiedenen disziplinären Wissensbeständen und organisationalen Einheiten eine Komplementierung ihres Profils durch soziale, kommunikative und mentale Eigenschaften abgeleitet.

Im abschliessenden vierten Teil nehme ich zentrale Ergebnisse der empirischen Kapitel auf und diskutiere sie vor dem Hintergrund des Erkenntnisinteresses und der Fragestellungen.

Kapitel 11 bildet eine Synthese der theoretischen Konzepte und der methodischen Vorgehensweise und wendet diese auf die Resultate der empirischen Analyse an. Dazu rekapituliere ich zentrale Ergebnisse der Arbeit und bette diese in das analytische Modell ein. Ich erörtere das Erklärungspotenzial des analytischen Modells sowie dessen Generalisierbarkeit über den konkreten Untersuchungsfall hinaus. Abschliessend formuliere ich Perspektiven für weitere Arbeiten, die bisher unterbeleuchtete Aspekte des Untersuchungsgegenstandes adressieren und vertiefen können.

Teil I - Grundlagen

Kapitel 2 - Transversale Wissensgebiete als Räume zwischen Feldern

2.1 Einleitung

Die Charakterisierung der Datenwissenschaften impliziert verschiedene Annahmen, die für die theoretische Grundierung der Arbeit leitend sind: Erstens untersucht die Arbeit empirisch die Entstehung eines sozialen Phänomens, das synchron in unterschiedlichen gesellschaftlichen Sphären verortet ist. Die Einteilung von Gesellschaft in soziale Sphären schliesst an differenzierungstheoretische Annahmen soziologischer Klassiker wie Durkheim (Arbeitsteilung), Weber (Wertsphären) oder Simmel (soziale Differenzierung) an, die verschiedene Mechanismen als Triebkräfte für die Ausbildung relativ autonomer gesellschaftlicher (Handlungs-)Räume identifiziert haben. Ich greife daher zunächst auf das Feldkonzept Bourdieus zurück, welches das begriffliche und analytische Instrumentarium bietet, um soziale Phänomene wie emergente Wissensgebiete theoretisch-analytisch zu verorten und zu untersuchen. Zweitens wird erkennbar, dass eine Vielzahl individueller wie kollektiver Akteur*innen aus unterschiedlichen Feldern in die Konstruktion und Entwicklung der Datenwissenschaften involviert sind. Auch hier beziehe ich mich auf Bourdieus Praxistheorie, wobei ich deren Akteurskonzeption und somit die Handlungsfähigkeit auf kollektive Formationen wie Universitäten oder Unternehmen erweitere. Drittens hat das praxistheoretische Feldkonzept kaum Verständnis für soziale Sphären, die nicht umfassend organisiert sind und kollektiven Regeln unterliegen. Ich ergänze Bourdieus Feldmodell daher mit dem Konzept von Räumen zwischen Feldern, das solche unterbestimmten Sphären in die differenzierungstheoretischen Annahmen der Feldtheorie integriert, diese aber auch in zentralen Punkten ergänzt. Schliesslich stellt sich viertens die Frage nach den Praktiken, durch die neue Wissensgebiete konstruiert werden. Während die Feldtheorie akteuriale Handlungen primär in den Möglichkeitsräumen verortet, die in den Habitus inkorporiert sind, werden im empirischen Fallbeispiel organisationale Praktiken von Grenz- und Begriffsarbeit erkennbar, die die Herausbildung neuer Wissensgebiete strukturieren. Hier beziehe ich mich auf Konzepte der *Science and Technology Studies*, die die Koexistenz multipler Verständnisse von sozialen Phänomenen reflektieren und damit der Multiperspektivität neuer Wissensgebiete gerecht werden.

Das Kapitel macht demnach einen Vorschlag, wie durch die Verknüpfung verschiedener theoretischer Perspektiven die feldübergreifende Entstehung neuer Wissensgebiete konzeptualisiert und erklärt werden kann. Es ist wie folgt strukturiert: Zunächst beschreibe ich die Praxistheorie Bourdieus als Ausgangspunkt und führe

die grundlegenden theoretischen Konzepte der Arbeit wie Feld, Raum, Praktiken und Akteur*innen ein (Kap. 2.2). Ich frage nach dem Potenzial von Bourdieus Feldtheorie zur Erklärung der Emergenz neuer Wissensgebiete (Kap. 2.3). Anschliessend diskutiere ich einige Schwachstellen von Bourdieus Feldtheorie, die einer Erweiterung durch ergänzende Theorieelemente bedürfen, um die transversale Entstehung von Wissensfeldern fassen zu können (Kap. 2.4). Das Konzept von Räumen zwischen Feldern bietet eine Möglichkeit, die Entstehung neuer hybrider Wissensgebiete in einem Raum zu lokalisieren, der ausserhalb etablierter Logiken sozialer Felder situiert ist (Kap. 2.5). Anschliessend fokussiere ich zwei zentrale Praxismodi (Grenz- und Begriffsarbeit), durch die Akteur*innen in solchen Zwischenräumen die Herausbildung neuer Wissensgebiete strukturieren: Erstens bespreche ich Praktiken von *boundary work*, d. h. Grenzziehungen und Grenzüberschreitungen, die ein widersprüchliches, produktives Potenzial entfalten können. Zweitens zeige ich auf, wie diskursive Praktiken der Repräsentation und Imagination multiple Verständnisse von solchen Räumen befördern, die Handlungen unterschiedlicher Akteur*innen koordinieren und so zu deren Permanenz beitragen (Kap. 2.7). Schliesslich beschreibe ich mein analytisches Modell, das die verschiedenen Theorieperspektiven integriert und somit einen Beitrag zur empirischen Untersuchung der Entstehung hybrider Wissensgebiete zwischen sozialen Feldern leistet (Kap. 2.8).

2.2 Die Praxistheorie Pierre Bourdieus als Ausgangspunkt

Die Praxistheorie Pierre Bourdieus stellt ein integriertes Theoriekonzept dar, das klare Vorstellungen von Begriffen wie Feld, Raum, Akteur*innen sowie den Formen der Praxis hat, die für diese Arbeit zentral sind. Ich beginne die Ausführungen mit dem Begriff des sozialen Feldes und der damit einhergehenden Akteurskonzeption, die ich in einem zweiten Schritt durch die Habitus- und Kapitaltheorie ergänze und dabei auch die Begriffe des sozialen Raumes und der Praxis erläutere.

Allgemein beschreiben Felder gesellschaftliche Handlungsräume, in denen soziale Beziehungen weder auf gesellschaftliche Makrostrukturen (Kapitalismus, soziale Klassen etc.) noch auf die Handlungen individueller Akteur*innen reduziert werden. Felder umfassen alle für einen gesellschaftlichen Teilbereich relevanten Akteur*innen und Institutionen, die in objektiven Relationen zueinander stehen (Bourdieu & Wacquant 1996: 126ff.). Sie nehmen sich wechselseitig wahr, orientieren ihre Praktiken aneinander und prägen jeweils die Entwicklung eines Feldes, wobei auf jedem Feld andere Handlungslogiken, andere »Spielregeln« gelten. Felder strukturieren somit die Interessen und Erwartungen der Akteur*innen, die gemeinsame Sinnsysteme, Regeln und Normen ausbilden und teilen, die wiederum die jeweiligen Opportunitäten und Aktivitäten rahmen.

Bourdieu konzipiert in seiner umfassenden Praxistheorie Felder in erster Linie als »Kampffelder« (Bourdieu 1993, 1997b; Bourdieu & Wacquant 1996). Akteur*innen (*agents*) – ob Individuen oder Kollektive – konkurrieren miteinander um Ressourcen, Status und vor allem um die Definition der »Spielregeln«, die die Beziehungen auf dem Feld bestimmen. Indem der feldanalytische Akteurbegriff soziale Entitäten beschreibt, die »Eigenschaften besitzen, die erforderlich sind, um im Feld Wirkungen zu entfalten, Effekte zu produzieren« (Bourdieu & Wacquant 1996: 139), ist er expli-

zeit nicht auf Individuen beschränkt, sondern umfasst auch Kollektive wie Gruppen, Organisationen oder Netzwerke (Baier & Schmitz 2012: 195). Akteur*innen sind selbst kulturelle Konstrukte (Meyer & Jepperson 2000), deren Herausbildung es feldtheoretisch in Wechselwirkung mit dem Gegenstand des Feldes zu untersuchen gilt (Bernhard & Schmidt-Wellenburg 2012: 40).

Die Relationen *innerhalb* von Feldern sind hierarchisch, wobei dominante Akteur*innen ihre Macht über dominierte Akteur*innen aufgrund ihrer Fähigkeit durchsetzen, das Feld zu kontrollieren, was auf dem Spiel steht und welche Kapitalformen zählen. Die Kräfteverhältnisse innerhalb von sozialen Feldern basieren stets auf der Struktur der Kapitalverteilung:

»Die Struktur des Feldes gibt den Stand der Machtverhältnisse zwischen den am Kampf beteiligten Akteuren oder Institutionen wieder bzw., wenn man so will, den Stand der Verteilung des spezifischen Kapitals, das im Verlauf früherer Kämpfe akkumuliert wurde und den Verlauf späterer Kämpfe bestimmt« (Bourdieu 1993: 108).

Felder sind also dynamische Entitäten, da sie ständigen Wandlungen unterliegen und stets nur eine Momentaufnahme der aktuellen Machtverhältnisse abbilden. Voraussetzung für eine Teilhabe am Feld ist der von allen Beteiligten geteilte Glaube (*croyance*) an die Existenz des Feldes sowie an die Relevanz und den Wert ihrer Spieleinsätze in Form eines unausgesprochenen Konsenses (*illusio*). Dazu müssen die jeweils geltenden Regeln akzeptiert und ein Interesse für das Spiel entwickelt werden. Die involvierten Akteur*innen, sowohl diejenigen auf den herrschenden wie diejenigen auf den beherrschten Positionen, tragen dadurch zur Reproduktion des jeweiligen Praxisfeldes bei (Bourdieu 1993: 109).

Die Relationen *zwischen* Feldern sind ebenfalls hierarchisch: Das Feld der Macht strukturiert aufgrund der Kapitalvolumina und Zusammensetzung die Relationen – und damit Machtverhältnisse – zwischen den jeweiligen Feldern (Bourdieu 2004; Schmitz et al. 2017; Wacquant 2004; Witte & Schmitz 2017). In den Feldern dominante Akteur*innen und Institutionen ringen hier um die Kräfteverhältnisse zwischen den jeweiligen Feldern. Das Machtfeld bezeichnet somit eine Kampfzone zwischen den nach eigenen Logiken ausdifferenzierten, relativ autonomen Feldern. Bourdieu beschreibt es auch als das dominante, kapitalstarke ›obere‹ Spektrum des sozialen Raumes (vgl. unten) bzw. als Feld der herrschenden Klasse (Bourdieu 1987).

Bourdieu's Feldtheorie ist ohne Berücksichtigung seiner Kapital- und Habitus-theorie nicht vollständig zu erfassen: Die Struktur eines Feldes, d. h. die Positionen und Positionierungen der Akteur*innen, ergibt sich in Homologie zu derjenigen des sozialen Raumes durch die unterschiedliche Struktur und Volumen an Kapital, über das die Akteur*innen verfügen können. In welchem Verhältnis welche Kapitalsorten¹ jeweils Gültigkeit haben, unterscheidet sich je nach Feld und bildet einen zentralen Konfliktgegenstand der Genese, aber auch der Reproduktion von Feldern. So zeigt seine Analyse des französischen Universitätssystems in *Homo Academicus* (Bourdieu 1988), inwiefern die bipolare Struktur des universitären Feldes die Struktur des sozialen Raumes abbildet: »Weltlich« dominierte Fakultäten stehen dabei sozial dominanten

1 Zur Differenzierung und zum Verhältnis der allgemeinen Kapitalsorten verweise ich auf Bourdieu (1983).

Fakultäten gegenüber, d. h., kapitaltheoretisch formuliert befinden sich Inhaber*innen von Positionen mit viel kulturellem Kapital in Konkurrenz mit solchen, die über viel ökonomisches Kapital verfügen (ebd.: 90). Daraus resultieren zwei gegensätzliche Hierarchisierungs- und Legitimationsprinzipien der Positionen bzw. Positionierungen der Akteur*innen des Feldes: Einer sozialen Hierarchie, die auf ökonomischem und geerbtem kulturellem Kapital beruht, steht eine kulturelle Hierarchie gegenüber, die auf objektiviertem kulturellem Kapital und wissenschaftlicher Autorität (gilt vor allem für die Naturwissenschaften) bzw. intellektueller Prominenz (gilt primär für die Geisteswissenschaften) basiert (ebd.: 100ff.). Durch die Spieleinsätze, d. h. Investitionen des jeweiligen verfügbaren Kapitals, reproduzieren die Inhaber*innen somit nicht nur ihre Positionen, sondern auch die Struktur des universitären Feldes insgesamt.

Das Konzept des sozialen Raumes, das Bourdieu in seiner Auseinandersetzung mit der französischen Sozialstruktur der 1960er- und 1970er-Jahre entwickelt, entwirft Gesellschaft als dreidimensional: Die vertikale Achse wird bestimmt durch die verfügbaren Kapitalvolumina, auf der horizontalen Achse differenziert er Positionen aufgrund der Kapitalstruktur, des Verhältnisses von ökonomischem und kulturellem Kapital. Eine dritte Dimension eröffnet die zeitliche Entwicklung der beiden, welche den Raum der Positionen konstituiert. Bourdieu zeigt dies durch die Veränderung sozialer Kategorien – der Berufsgruppen – über die Zeit an (Bourdieu 1987: 212f.). Die beobachtbaren Transformationen geschehen jedoch nicht zufällig: Einem bestimmten Punkt im sozialen Raum mit spezifischem Kapitalvolumen und -struktur entspricht jeweils ein Bündel gleich wahrscheinlicher, zu ähnlichen Positionen führender Lebensverläufe (*trajectoires*) – nach Bourdieu das »objektiv gegebene *Möglichkeitsfeld*« (Bourdieu 1987: 188; Hervorhebung im Original). Die einzelnen Positionen und Verläufe sind nicht unabhängig voneinander: Start- und Endpositionen können durch individuelle Laufbahneffekte, d. h. die Einflüsse der Entwicklung von Umfang und Struktur des verfügbaren Kapitals auf sozialen Auf- oder Abstieg, voneinander divergieren (Bourdieu 1987: 189).

Im Habitus werden Volumen, Zusammensetzung und die zeitliche Entwicklung der einzelnen Kapitalformen inkorporiert, welche die Formen der Praxis, der Klassifikationen und der Distinktionen in einem Feld massgeblich definieren und diesen zugrunde liegen (Bourdieu 1987: 195). Der Habitus klassifiziert somit die verschiedenen Praxisformen und bildet synchron das vermittelnde Scharnier zwischen dem sozialen Raum und dem Raum der Lebensstile, d. h. der Formen der Praxis, dar. Durch ihn weisen alle Praxisformen von Akteur*innen einen systematischen Charakter auf und unterscheiden sich gleichzeitig systematisch von denjenigen anderer Akteur*innen. In diese fundamentalen Denk-, Wahrnehmungs- und Beurteilungsschemata fließen die früher gemachten Erfahrungen mit ein: »Die fundamentalen Gegensatzpaare der Struktur der Existenzbedingungen (oben/unten, reich/arm etc.) setzen sich tendenziell als grundlegende Strukturierungsprinzipien der Praxisformen wie deren Wahrnehmung durch« (Bourdieu 1987: 279). Im Habitus sind die gesellschaftlichen Strukturen in den Akteur inkorporiert, deshalb beschreibt Bourdieu ihn auch als »Leib gewordene Geschichte« (Bourdieu 1976: 200). Die inkorporierte Ausübung von Denken, Wahrnehmen und Handeln strukturiert somit das zukünftige Denken, Wahrnehmen und Handeln, was er als *opus operatum*, als strukturierte Struktur, bezeichnet. Demgegenüber steht der Habitus als *modus operandi*, als das »Erzeugungsprinzip objektiv

klassifizierbarer Formen von Praxis« (Bourdieu 1987: 277). Der Habitus bildet sich stets in Relation zu bestimmten Feldern aus (Bourdieu 1987: 164), was impliziert, dass die dort verinnerlichten Fähigkeiten zwar potenziell übertragbar, aber in anderen Feldern nicht zwingend wertvoll und praktikabel sind. Der Begriff der Doxa bezeichnet die symbolische Anerkennung der verinnerlichten Strukturen im Habitus, die somit den objektiven Strukturen des sozialen Handlungskontextes bzw. Feldes entsprechen.

Der Habitus wird nicht nur durch die sozialen Positionen im Klassengefüge, sondern auch durch die vergeschlechtlichten (Bourdieu 1997a) und ethnifizierten bzw. rassifizierten Strukturen (Kastner 2002, 2011) von Gesellschaft geprägt. Gesellschaftliche Machtverhältnisse in solchen binären Kategorisierungen (Mann – Frau, Staatsbürgerin – Ausländerin oder die Konstruktion in weiss – nicht-weiss) strukturieren ebenso die Denk-, Wahrnehmungs- und Beurteilungsschemata von Akteur*innen, die in diesem Sinne auch als vergeschlechtlicht und vergeschlechtlichend bzw. als ethnifiziert und ethnifizierend zu verstehen sind.

Im Konzept des Habitus zeigt sich der Anspruch Bourdieus, strukturdeterministische und subjektivistische Handlungs- und Akteurmodelle zu überwinden (vgl. Diaz-Bone 2010: 34). Damit werden Akteur*innen weder als vollständig den sozialen Macht- und Herrschaftsverhältnissen unterworfen (wie im marxistischen Strukturalismus), noch als inhärent autonom sowie unabhängig von gesellschaftlichen Strukturen konzipiert (wie in Rational-Choice-Modellen). Vielmehr bilden sich durch die verinnerlichten Dispositionen Muster aus, die zukünftige Handlungen prägen. Das Handeln bewegt sich so innerhalb der vom *opus operatum* vorgegebenen Grenzen des Möglichen – dem »*Möglichkeitsfeld*« (vgl. oben) –, die an die materiellen Existenzbedingungen gekoppelt sind.

Für die vorliegende Arbeit von Bedeutung ist schliesslich das Verhältnis von Feld und sozialem Raum, das im umfangreichen Werk Bourdieus unterschiedlichen Schwerpunktsetzungen unterliegt (Reed-Danahay 2019): Während in früheren Arbeiten zur französischen Sozialstruktur die Klassendimension und Lebensstile, d. h. der soziale Raum, im Zentrum stehen (Bourdieu 1987; Bourdieu et al. 1981), entwickelt er in seinem späteren Werk eine stärker differenzierungstheoretische Konzeption von Gesellschaft, indem er spezifische Sphären von Gesellschaft als Kräftefelder denkt, die unabhängig von weltlichen oder überirdischen Instanzen eigene Handlungs- und Legitimationsprinzipien ausbilden.² Der soziale Raum wird somit nicht nur durch die Lebensstile individueller Akteur*innen reproduziert, sondern auch in und durch Macht- und Ungleichheitsverhältnisse in den einzelnen Feldern. Damit kommt Bourdieu das Verdienst zu, dass er in seiner Praxistheorie differenzierungstheoretische Fragen mit solchen der sozialen Ungleichheit integriert. Allerdings bleibt das exakte Verhältnis von Feldern und dem sozialen Raum, ob Felder im sozialen Raum enthalten sind (Reed-Danahay 2019: 5, 27ff.), ob es sich um Mikrokosmen des sozialen Raumes handelt (Bourdieu 2001a: 41) oder ob Feld und Raum letztlich identisch sind, aufgrund uneindeutiger Aussagen ungelöst.³ Mit Feld und sozialem Raum koexistieren vielmehr

2 Zu nennen sind etwa die Felder der Wissenschaft (Bourdieu 1988, 2004), Ökonomie (Bourdieu 1998, 2002), Politik (Bourdieu 2001a), Religion (Bourdieu 2000) oder Kunst (Bourdieu 2001b), die sich als relativ autonome Sphären mit eigener Kapitalstruktur ausgebildet und differenziert haben.

3 Schmitz et al. (2017: 62) argumentieren, dass das Feld der Macht und der soziale Raum zwei unterschiedliche Konstruktionen derselben ontologischen Entität – der Gesellschaft – darstellen.

zwei unterschiedliche topologische Modelle des Sozialen (Rehbein 2003), deren Verhältnis sich nicht im Vorhinein festlegen lässt, sondern das stets empirisch zu analysieren ist.

2.3 Die Entstehung neuer Wissensfelder

Mit seiner umfassenden Praxistheorie hat Bourdieu ein sozialtheoretisches Modell entworfen, das er selbst sowie daran anschließende Forschende auf sehr unterschiedliche empirische Gegenstände angewandt haben. Im Folgenden frage ich nun nach dem spezifischen Beitrag seines Theoriekonstrukts zur Erklärung der Entstehung neuer Wissensgebiete, die das zentrale Erkenntnisinteresse der Arbeit bildet.

Prozesse der Feldgenese sind in Bourdieus konflikttheoretischer Lesart primär durch Abgrenzung, Definitionsmacht und Ressourcenmonopolisierung charakterisiert (Bourdieu 1988; Fröhlich 2009): Neue Felder entstehen in heteronomen Grenzbereichen bzw. an der Schnittstelle bestehender Felder. In seiner kapitaltheoretischen Lesart investieren Akteur*innen Kapital in neue Gebiete, wenn sie sich davon einen materiellen (oder symbolischen) Vorteil bzw. Profit versprechen (Bourdieu 1993). Da die dominanten Gruppen in etablierten Feldern zur Orthodoxie streben, sind es primär die marginalisierten, d. h. beherrschten Akteur*innen, die versuchen, die Regeln des Feldes zu ihren Gunsten neu zu bestimmen (für das Beispiel Data Science: Grommé et al. 2018). Sie werden von den herrschenden Positionen des Feldes als häretisch wahrgenommen (Bourdieu 1993: 109). Die Praktiken, d. h. Positionierungen und Investitionen jener Akteur*innen in neuen bzw. entstehenden Feldern müssen sich gegenüber jenen in existierenden, etablierten Feldern behaupten. Letztere verfügen nicht nur über eigene Geschichten und Logiken, sondern auch über ausdifferenzierte Formen der Kapitalverteilung. Im Zuge seiner relativen Autonomisierung bildet ein entstehendes Feld eigene Anerkennungsstrukturen aus, nach denen feldspezifisch legitimes Kapital zugeteilt wird. Der Bruch mit den etablierten Traditionen und Einteilungen des Sozialen in relativ autonome Felder ist stets ein stark risikobehafteter, da sich die Neulinge gegenüber der konzentrierten Macht der herrschenden sozialen Verhältnisse durchsetzen müssen. Umso lohnender ist das Unterfangen dann jedoch bei positivem Ausgang.

Im Anschluss an Bourdieus Feldmodell und Webers Religionssoziologie skizziert Schultheis (2005: 71ff.) sechs Schritte der Feldgenese, die er auf die »Disziplinierung« neuer Praktiken im Feld der Wissenschaft und der Kunst anwendet: Zunächst wird durch Kritik und Anspruch (1) der Status quo problematisiert und vernachlässigte Praktiken, Forschungslücken sowie die Bedeutung eines Perspektivenwechsels betont. Darauf folgen Kooptation und Selbstorganisation (2), d. h. ein »Prozess der Vergemeinschaftung« von Akteur*innen unterschiedlicher Felder und Disziplinen, die durch ihre »gemeinsame soziale Lage«, in der Regel dominierte Positionen in etablierten Feldern, ein Interesse an Grenzüberschreitung entwickeln. In der Phase der »Repräsentationsarbeit« müssen die beteiligten Akteur*innen Sozialkapital und Legitimationsressourcen (3) mobilisieren, beispielsweise durch Einladung etablierter Vertreter*innen der neuen Disziplin, um die neuen Ansprüche gegenüber Widerständen aus bestehenden Disziplinen zu legitimieren. Durch Selbstreferenz und Kodifizierung (4) können sie postulierte Forschungslücken plausibel machen und nach Regeln wis-

senschaftlicher Argumentation erläutern, da diese sonst nicht als legitim gelten. Dies erfolgt einerseits durch Verweise auf ein eigenständiges Set an epistemologischen, theoretischen und methodologischen Koordinaten, andererseits durch die Kanonisierung von »Klassikern«, Standardwerken oder »Gründungsvätern«. Besonders identitätsstiftend können dazu Mythen, Legenden und »Heldengeschichten« sein (ebd.: 72f.).

In der Phase von Institutionalisierung und Akademisierung (5) muss der prekäre Status im stark umkämpften Feld der Wissenschaft stabilisiert werden: Neben dem Ringen um die stets knappen Mittel legitimieren ›weiche‹ Formen wie die Nachfrage auf dem Arbeitsmarkt oder unter Studienaspirant*innen das neue Feld. Es muss ein Hochschultyp als ›richtiger‹ Ort der Institutionalisierung identifiziert und ausgewählt werden; zentral ist auch der Aufbau von Curricula, Studienprofilen und Lehrangeboten, die die Anschlussfähigkeit für Berufskarrieren gewährleisten. Schliesslich gilt es, sich durch Corporate Identity und Monopolisierung (6) öffentliches Gehör und Sichtbarkeit durch Fachzeitschriften, Verbände, Kongresse etc. zu verschaffen. Selbstorganisation ist notwendig, um das Monopol über die legitime Definition des Gegenstandsbereiches beanspruchen zu können, wobei die Kriterien der Monopolisierung universalistisch sein müssen. Dabei muss der Anspruch auf Autonomie gegenüber konkurrierenden wissenschaftlichen Disziplinen ohne Referenz auf ausserwissenschaftliche Grössen (Staat, Markt, Kirche etc.) begründbar sein.

Schultheis verweist auf die widersprüchlichen Effekte solcher Disziplinierungsprozesse: So können am Ende zwar Autonomie und Anerkennung stehen, sie führen aber auch zu einer Disziplinierung »im Sinne einer Unterwerfung unter Regeln«, was »mit einer deutlichen Einschränkung der individuellen Gestaltungsfreiheit zu Gunsten kollektiver Autonomie« einhergeht (ebd.: 67f.). Die einzelnen Akteur*innen unterwerfen sich demnach kollektiven Regeln, die ihre Möglichkeitsräume massiv einschränken. Die Disziplinwerdung wird insofern als ein »Herrschaftsprozess« gedeutet (ebd.: 74), wodurch die Grundstruktur des wissenschaftlichen Feldes anerkannt und reproduziert wird.

Der beschriebene Prozess, der in sich sehr schlüssig ist, macht deutlich, dass die Genese wissenschaftlicher Disziplinen zum einen langfristig angelegt, zum anderen stark konfliktbehaftet ist. Insgesamt erscheint die Entstehung neuer Wissensgebiete in einem Modell, das primär auf die Reproduktion etablierter Strukturen ausgerichtet ist, als höchst unwahrscheinlich. Entweder autonomisieren sich solche Wissensgebiete als neue Felder oder sie verfallen und verschwinden. Es gibt gewissermassen kein Dazwischen, keinen Platz für unorganisierte, nicht hierarchisch strukturierte Räume (Eyal 2013b) und kaum Möglichkeiten für nicht-konfliktuelle, kooperative Praktiken. Eine Ausnahme bilden die literarischen Salons des 19. Jahrhunderts, in denen sich vor allem marginalisierte Positionen aus Politik und Künsten vermischten und austauschten. Bourdieu bezeichnet sie als »Zwitterinstitutionen« (Bourdieu 2001b: 89; Hervorhebung im Original), was die Idee einer Verschmutzung der ›reinen‹, abgeschlossenen Sphäre des Feldes indiziert. Andererseits bilden die Salons aber auch »regelrechte Mittlerinstanzen zwischen den Feldern« (ebd.: 88): Künstler*innen und Politiker*innen kommen zusammen und versuchen, durch strategische Allianzen materielle oder symbolische Vorteile für sich zu erwirken. Die Salons sind ein Beispiel für soziale Räume – heute würde man sie als *Innovationsinkubatoren* bezeichnen –, in denen feldübergreifende Interaktionen koexistieren und zur Herausbildung neu-

er Expertisen und Praktiken, letztlich zur relativen Autonomie des Kunstfeldes (Eyal 2013b: 177), beitragen.

Trotz dieser Ausnahme erklärt das Feldmodell Bourdieus jene sozialen Sphären nur unzureichend, die marginal und instabil sind. Feldtheoretisch formuliert handelt es sich um heteronome Felder bzw. heteronome Pole, die dominierte Positionen innerhalb relativ autonomer Felder markieren. Diesen Positionen bleiben letztlich nur Strategien von Heterodoxie bzw. Häresie, indem sie sich bewusst den legitimen Stellungnahmen im Feld widersetzen (Bourdieu 1976: 325, 1987: 668). Panofsky (2011) beschreibt am Beispiel der menschlichen Verhaltensgenetik ein Wissensfeld, das am Rande der Wissenschaft liegt. Trotz seines liminalen Status erweist sich das Wissensgebiet aufgrund hohen Medieninteresses und finanzieller Alimentationen durch staatliche Akteur*innen als dauerhaft – und wurde so zu einem Schauplatz von Kämpfen verschiedener Akteur*innen umliegender Felder. Es hat sich demnach weder autonomisiert noch ist es wieder verschwunden, sondern hält sich beständig in einem wissenschaftlichen Grenzbereich. Das Beispiel deutet das Potenzial einer adäquateren Berücksichtigung solcher Grenzräume sowie der Relationen zwischen Feldern an.

Ähnlich zeigt sich dies in der Beschreibung des empirischen Fallbeispiels (vgl. Kap. 3): Erstens ist ein signifikanter Einfluss ausserwissenschaftlicher Referenzfelder (Industrie, Staat) auf die Herausbildung erkennbar, was im Sinne der Feldtheorie ihrer Autonomisierung entgegensteht. Zweitens ist die Herausbildung der Datenwissenschaften weniger durch Langfristigkeit, sondern vielmehr durch eine forcierte Beschleunigung gekennzeichnet, die für technologieintensive Wissensgebiete charakteristisch ist (Beer 2017; Wajcman 2019). Ein dritter Unterschied liegt darin, dass es nicht nur marginale Akteur*innen sind, die in neue Gebiete investieren. Auch zentrale Positionen in den jeweiligen Feldern prägen durch umfangreiche Investitionen das neue Wissensgebiet. Schliesslich streben viertens die beteiligten Akteur*innen gar nicht nach Abschliessung und Monopolisierung der Ressourcen, sondern das Teilen von Inhalten, Daten und Methoden, d. h. Kooperation und kooperatives Verhalten über Feldgrenzen hinweg, ist zentraler Bestandteil der Entstehung und Funktionsweise des Wissensgebiets.

2.4 Erweiterungen der Bourdieu'schen Feldtheorie zur Erklärung neuer Wissensfelder

Die Ausführungen zum Erklärungspotenzial des Bourdieu'schen Feldmodells für entstehende Wissensgebiete haben deutlich gemacht, dass es ergänzender Theorieperspektiven bedarf, um dem empirischen Gegenstand dieser Arbeit gerecht zu werden. Bezogen auf mein Erkenntnisinteresse sehe ich dabei folgende drei Punkte: Erstens erhält die Gleichzeitigkeit von konflikthaften und kooperativen Praktiken zu wenig Bedeutung innerhalb von Bourdieus Feldtheorie (Kap. 2.4.1). Zweitens fehlt eine angemessene Berücksichtigung der Relationen zwischen Feldern, d. h., die Erklärung von Neuheit bleibt mehr oder weniger auf den jeweiligen Untersuchungsbereich, die Relationen innerhalb von Feldern, limitiert (Kap. 2.4.2). Drittens braucht es eine Erweiterung des Akteurmodells, um die kollektiven Akteur*innen des Untersuchungsfeldes adäquater in das Theoriegerüst zu integrieren (Kap. 2.4.3).

2.4.1 Die Parallelität und Synchronizität von Konflikt und Kooperation

Bourdieu's Feldmodell erklärt primär die Reproduktion und Stabilität von Feldern und fokussiert dabei die konflikthafter Relationen zwischen den jeweiligen Positionen im Feld. Obwohl kooperative Beziehungen in bzw. über Felder hinaus in Bourdieus Konzeption nicht ausgeschlossen sind (für das ökonomische Feld: Bourdieu & Wacquant 1996: 144; für das Kunstfeld: Bourdieu 2001b), werden sie letztlich immer machttheoretisch als strategische Positionierungen im Konkurrenzkampf um die zentralen Kapitalia gewendet. Beispielhaft äussert sich dies wiederum in Bourdieus Beschreibung der literarischen Salons:

»Die Salons bilden auch, über die vielfältigen Formen des darin sich vollziehenden Austauschs, regelrechte Mittlerinstanzen zwischen den Feldern: Die Inhaber der politischen Macht wollen ihre Sicht den Künstlern aufzwingen und sich deren Konsekrations- und Legitimationsmacht, die diese insbesondere über das innehaben, was Sainte-Beuve die »literarische Presse« nennt, zu eigen machen; die Schriftsteller und Künstler wiederum, die als Bittsteller und Fürsprecher, zuweilen sogar als regelrechte *pressure group* auftreten, sind darauf aus, eine mittelbare Kontrolle über die verschiedenen vom Staat verteilten materiellen und symbolischen Gratifikationen zu gewinnen« (Bourdieu 2001b: 88f.; Hervorhebung im Original).

Obwohl Bourdieu beiden Seiten, den »Inhaber[n] der politischen Macht« als auch den »Künstler[n]«, ein Interesse an Austausch und somit strategischen feldübergreifenden Koalitionen zuschreibt, sind sie damit letztlich auf ihren eigenen Vorteil bedacht und wollen durch konfliktbehaftete Formen der Praxis (»aufzwingen«, »zu eigen machen«, »gewinnen«) ihre Positionierungen im Feld verbessern. Durch eine solche konflikttheoretische Wendung von kooperativen Situationen lässt Bourdieus Feldmodell sowohl die Potenziale von Institutionen zur koordinierten, kooperativen Problemlösung als auch jene Situationen ausser Acht, die nicht durch Macht- und Herrschaftsverhältnisse strukturiert sind (Maurer 2006).

Die Theorie strategischer Handlungsfelder von Fligstein und McAdam bietet hier einen Anknüpfungspunkt: Sie setzt Felder nicht immer mit Konkurrenzkämpfen und Hierarchien gleich, sondern schreibt den kooperativen, konsensbetonenden Elementen grössere Bedeutung zu (Fligstein & McAdam 2012: 89f.; Kluttz & Fligstein 2016: 191). Ähnlich wie bei Bourdieu sind strategische Handlungsfelder durch etablierte Positionen (*incumbents*) und Herausforder*innen (*challengers*) strukturiert, die über die zentralen Ressourcen des Feldes konkurrieren. Dennoch sind Kooperation und kollektives Handeln in ihrer Theoriesynthese ebenso strategisch angelegt – weshalb die Theorie strategischer Handlungsfelder als eine interaktionistische Feldvariante gedeutet wird (Florian 2008: 134).⁴

4 Trotz der Unterschiede der beiden Spielarten der Feldtheorie manifestiert sich der relationale Charakter von Feldern: Die Emergenz neuer Gegenstände wird in den Beziehungen der involvierten Akteur*innen verortet. Diese werden nicht substantialistisch gedacht, sondern unterliegen kollektiven Aushandlungen. Neuheit wird folglich nicht reduktionistisch durch die Handlungen einzelner Akteur*innen (Erfinder*innen, Genies etc.) erklärt.

In strategischen Handlungsfeldern ringen individuelle wie kollektive Akteur*innen um Ressourcen zu ihrem Vorteil; um sich durchzusetzen, müssen sie die Handlungen anderer berücksichtigen und strategische Beziehungen und Koalitionen eingehen (Fligstein & McAdam 2012: 11). Die Zugehörigkeit basiert – analog zu Bourdieu – auf einem kollektiven Verständnis darüber, um was es im Feld geht, was auf dem ›Spiel‹ steht. So entwickeln die Akteur*innen ein Gespür für die anderen Positionen im Feld. Auch verfügen sie über eine gemeinsame Auffassung der als legitim erachteten Regeln. Schliesslich teilen die Akteur*innen in bestimmten Positionen des Feldes einen Interpretationsrahmen, der innerhalb des Feldes variiert, aber von Akteur*innen an ähnlichen Orten geteilt wird (Fligstein & McAdam 2012: 12f.).

In der Erklärung von Wandel und Neuheit sind strategische Handlungsfelder durch Einsichten der sozialen Bewegungsforschung geprägt (Kluttz & Fligstein 2016: 198). Entstehende Felder werden entlang des Kontinuums von Kooperation und Koalition einerseits sowie Hierarchie und Machtunterschiede andererseits organisiert (ebd.: 199). Die feldinternen Verhältnisse zwischen Konsens und Konflikt sind ständigem Wandel unterworfen und hängen insbesondere vom Grad der Etabliertheit ab: Neue Felder sind konflikthafter, während ältere, etablierte Felder stärker Kooperation betonen. Stabilisierte Felder haben gegenüber entstehenden Feldern höhere Wahrscheinlichkeiten, dass sie reproduziert werden.

Fligstein und McAdam (2012: 86ff.) unterscheiden zwei Arten von Wandel in strategischen Handlungsfeldern: zum einen kontinuierlicher Wandel in kleinen Schritten innerhalb des Feldes, zum anderen exogen induzierter revolutionärer Wandel. Im ersten Fall können Akteur*innen durch den Bruch etablierter Bündnisse und durch neue strategische Koalitionen relative Verbesserungen ihrer Feldposition erreichen. Über die Zeit zeitigen solche inkrementellen Veränderungen aggregierte Effekte und führen schliesslich zu neuen Machtverhältnissen in Feldern.

Wahrscheinlicher ist jedoch exogen induzierter Wandel in nahen Feldern (ebd.: 99ff.): Erstens sind es »Outsider«, die bislang keine Akteur*innen des Feldes waren, allerdings über entsprechende Ressourcen verfügen, sodass sie in das Feld eintreten und neue Spielregeln etablieren können. Aktuelle Beispiele sind Technologiekonzerne, die dank ihren enormen finanziellen und technologischen Möglichkeiten ganze Felder (wie Journalismus, Werbung, Tourismus, Automobilität etc.) nachhaltig prägen und verändern können. Zweitens sind es gesellschaftliche Makroereignisse (wie Kriege, ökonomische Krisen oder Pandemien), die zu allgemeinen Krisen führen, in deren Folge die Beziehungen zwischen Akteur*innen sowie jene zwischen Feldern dauerhaft transformiert werden. Schliesslich können drittens die Relationen zwischen Feldern, auf denen Ressourcenabhängigkeiten aufbauen, zu Wandel und Neubildung von strategischen Handlungsfeldern führen (vgl. Furnari 2016), etwa wenn aufgrund feldübergreifender Lieferketten Krisen in einem Feld in andere ›übertragen‹ werden, was die Auflösung und Neustrukturierung der existierenden Feldbeziehungen begünstigt.

Eine Perspektive auf Konflikte in und zwischen Feldern ist wichtig, da sie das zentrale Strukturprinzip von Feldern als »Kampffelder« bilden. Ebenso bedeutsam für Prozesse der Feldgenese und Wandel sind jedoch auch kooperative Logiken des Austauschs, da ansonsten die Herausbildung neuer Felder und Wissensgebiete äusserst unwahrscheinlich ist. Die Theorie strategischer Handlungsfelder von Fligstein und McAdam ermöglicht es, konfliktbehaftete und kooperative Praktiken als räum-

lich und zeitlich koexistierend zu lesen denn als sich gegenseitig ausschliessend.⁵ Die Synchronizität und Parallelität von Konflikt und Kooperation muss in diesem Sinne stärker für die Erklärung von neuen Wissensgebieten berücksichtigt werden.

2.4.2 Die Berücksichtigung von Relationen zwischen Feldern

Eine zweite Limitation von Bourdieus Feldtheorie liegt darin, dass er den Relationen zwischen Feldern zu wenig Bedeutung zumisst. Felder werden als relativ abgeschlossene Sphären konzipiert, die räumlich gewissermassen parallel nebeneinander existieren. Die Feldtheorie Bourdieus interessiert sich jedoch erstaunlich selten für die Relationen *zwischen* einzelnen Feldern; oft wird dieser Bereich einfach als ein eigenes Feld ausdifferenziert (Eyal 2013b; Liu 2021). Ein grundlegendes Problem liegt gemäss Eyal darin, dass Bourdieu und daran anschliessende Feldkonzeptionen Relationalität nur *innerhalb* von Feldern, jedoch nicht *zwischen* sozialen Feldern situieren. Begrifflichkeiten wie »politisches Feld«, »ökonomisches Feld« oder »künstlerisches Feld« grenzen demnach einen fixen Gegenstand gegenüber anderen Gegenständen ab, sind also gerade nicht relational.⁶

»It follows, therefore, that there is no easy, clear cut, self-evident criteria by which the various fields could be distinguished in terms of the content of the activity that takes place within them. Nothing is ›economic‹ or ›scientific‹ or ›artistic‹ by itself. These are also shorthand for bundles of relations.« (Eyal 2013b: 159)

Bei Bourdieu gibt es zwei Ausnahmen, in denen er explizit feldübergreifende Phänomene thematisiert: Zum einen konzipiert er das Feld der Macht als feldübergreifendes »Meta-Feld« (Bourdieu & Wacquant 1996: 142; Schmitz et al. 2017: 56; Swartz 1997: 136), in dem die verschiedenen Felder – vermittelt über ihre jeweiligen Akteur*innen bzw. Positionen – um die Durchsetzung spezifischer Interessen gegenüber denjenigen anderer Felder konkurrieren. Dieser feldübergreifende Raum steht in Homologie zum sozialen Raum an sich und ist elitesoziologisch als Netzwerk an Kapital reicher Akteur*innen, sowohl in ökonomischer, kultureller als auch sozialer Hinsicht, angelegt (Bourdieu 2014: 348f.; Schmitz et al. 2017).⁷

5 Lahire (2014) schlägt dafür den Begriff der Welten (*worlds*) vor, um solche gesellschaftlichen Sphären von Feldern als Spiel- bzw. Kampffelder im Bourdieu'schen Sinne zu differenzieren.

6 Exemplarisch dafür: »Es geht jeweils darum zu definieren, *was* eine feldspezifische Praxis (Kunst, Wissenschaft, Religion, Politik usw.) ist und *was* nicht, und es geht dabei primär darum zu definieren, *wer* Akteur eines Feldes ist und *wer* nicht (Künstler, Wissenschaftler, professioneller religiöser Dienstleister, Politiker usw.). Die Art und Weise, in der dies geschieht, legitimiert die Praxis und definiert mithin die *Grenzsicherung* der relativ autonomen Felder« (Bongaerts 2011: 117; Hervorhebung im Original).

7 Neben dem Feld der Macht ist es der Staat, ebenfalls als »Meta-Feld« bezeichnet und über ein »Metakapital« verfügend, der die grundlegenden Verhältnisse, wie die ›Wechselkurse‹ zwischen den einzelnen Kapitalsorten, zwischen verschiedenen Feldern beeinflussen kann (Bourdieu & Wacquant 1996: 142). So ist es im Fall des französischen Eigenheimmarktes der Staat, der durch Neuausrichtung der Wohnbauförderung (Angebotsseite) zur Herausbildung eines neuen Produktes, des Eigenheims, und damit eines neuen Feldes beiträgt (Bourdieu 2002). Ähnlich argumentieren auch Fligstein und McAdam (2012: 67ff.), wenn sie den Einfluss des Staates auf die Herausbildung neuer Handlungsfelder theoretisieren.

Das zweite Beispiel sind die bereits erwähnten literarischen Salons, in denen feldübergreifende Interaktionen und Kooperationen parallel zu konflikthafter Beziehungen existieren und die Herausbildung neuer Praktiken und Expertisen in den Feldern der involvierten Akteur*innen begünstigen. Als »Mittlerinstanzen zwischen den Feldern« (Bourdieu 2001b: 88) koppeln sie gewissermaßen die beteiligten Felder, ohne – wie im Fall des Machtfeldes – ein hierarchisches Verhältnis zwischen diesen zu definieren. Damit ist eine zwischenräumliche Idee angelegt, die dazu beitragen kann, die Genese neuer Wissensgebiete adäquater zu erklären (dazu Kapitel 2.5).

Fligstein und McAdam verfügen über eine elaboriertere Vorstellung der wechselseitigen Bezugnahmen zwischen unterschiedlichen Feldern. Sie thematisieren nicht nur die internen Feldbeziehungen, sondern auch die Beziehungen dazwischen. Felder sind in diesem Sinne stets in komplexe Beziehungen anderer Felder eingebunden (Fligstein & McAdam 2012: 18, 59, 100f.): Sie sind hierarchisch in breiteren, umfassenderen Feldern eingebettet oder – einer russischen Puppe gleich – ineinander verschachtelt.⁸ Felder stehen in Abhängigkeitsbeziehungen zu übergeordneten Feldern aufgrund unterschiedlicher Ressourcenverhältnisse (Fligstein & McAdam 2012: 16). Daneben werden »Zwischen-Feld-Beziehungen« (*inter-field relations*) auch durch legale und hierarchische Autoritäten definiert. In der Frage nach dem konkreten Verhältnis zwischen Feldern bleiben Fligstein und McAdam etwas vage, wenn sie am Beispiel der Relationen der Felder der Produzent*innen und Zuliefer*innen schreiben:

»Producers in a market frequently orient their actions to their competitors [...]. Producers are obviously dependent for success on their suppliers, but suppliers generally do not command all that much of the producer's attention. Instead, the suppliers comprise a field of their own. In our analysis, at the boundary between the fields, there *may be another field that exists* that defines the relationship between the two fields« (Fligstein & McAdam 2012: 168; eigene Hervorhebung).

Das Beispiel macht die wechselseitige Orientierung und Abhängigkeit der beiden Felder aufgrund unterschiedlicher Ressourcen deutlich. Zudem deuten sie bei der Analyse der Grenzen an, dass es sich weniger um eine lineare Grenzziehung, sondern vielmehr um ein eigenes Feld *handeln könnte* (vgl. Hervorhebung im Zitat), das die Beziehungen zwischen den zwei Feldern definiert – mit der Andeutung schlagen sie eine Brücke, um solche zwischenräumlichen Grenzphänomene in die Feldtheorie zu integrieren.

In dieser Arbeit verzichte ich auf eine primär machttheoretische Deutung wie bei Bourdieu, sondern konzipiere und untersuche die Beziehungen zwischen Feldern analytisch als eigene zwischenräumliche Phänomene. Sie können sowohl zu eigenständigen Feldern werden, wodurch die Relationen wiederum hierarchisch organisiert sind (wie bei Fligstein & McAdam), oder aber wenig institutionalisierte soziale Sphären bleiben und so die Beziehungen vage und uneindeutig halten. Ich werde das Konzept von Räumen zwischen Feldern, das solche instabilen sozialen Sphären beschreibt, die durch wechselseitige Konflikte und Kooperationen strukturiert sind, in Kapitel 2.5 vertiefen.

8 Fligstein und McAdam (2012: 59) führen das Bild des Feldes als einer russischen Puppe (*russian doll*) an, in der immer kleinere Felder enthalten sind.

2.4.3 Die Erweiterung des Akteurskonzepts auf Organisationen und Kollektivformationen

Ein dritter Punkt besteht in der Erweiterung des Akteurskonzepts von singulären auf kollektive Akteur*innen. In der vorliegenden Arbeit sind es primär Unternehmen, staatliche Institutionen, universitäre Departemente sowie Hochschulen, die als handelnde Akteur*innen auftreten. Obwohl in Bourdieus Feldmodell bereits angelegt, braucht es eine stärkere Berücksichtigung kollektiver Entitäten. Es geht dabei nicht um ein organisationales Feld wie im Neo-Institutionalismus (DiMaggio & Powell 1983; Scott 1994; Wooten & Hoffman 2017), sondern um Organisationen und andere kollektive Entitäten als soziale Felder einerseits und als Akteur*innen in sozialen Feldern andererseits (Dederichs & Florian 2004: 91ff.).

Bourdieu verweist an verschiedenen Stellen seines Werks auf kollektive Akteur*innen: So besteht etwa das ökonomische Feld aus einer »Gesamtheit von Teilfeldern«, die dem entsprechen, was üblicherweise unter wirtschaftlichen »Sektoren«, Branchen oder »Industriezweigen« verstanden wird (Bourdieu 1998: 174). In einer »zweifach relationale[n] Perspektive auf Organisationen« (Dederichs & Florian 2004: 91) sind Unternehmen Akteur*innen im Feld der Ökonomie, während sie gleichzeitig selbst jeweils ein Kräftefeld unterschiedlicher Akteur*innen (Produktionseinheiten, F&E, Marketing, Management etc.) darstellen. Äquivalent dazu sind Hochschulen und Universitäten Akteur*innen im Feld der Wissenschaft (Meier 2009), bilden aber gleichzeitig ein eigenes soziales Feld unterschiedlicher Fakultäten, Institute und anderer organisationaler Einheiten (Rektorat, Verwaltung etc.) (Bourdieu 2004; Baier & Schmitz 2012). Schliesslich bilden staatliche Institutionen als Akteur*innen das bürokratische Feld, können aber auch selbst als relativ autonome Felder konzipiert und untersucht werden (Bourdieu 2014).

Trotz der vielfältigen Verwendung des Akteurbegriffs bleibt Bourdieus Habituskonzept ein Knackpunkt für eine Übertragung auf kollektive Akteur*innen (Dederichs & Florian 2004: 91; Florian 2008: 144), da der Habitus aufgrund seiner Verinnerlichung mit der Konzeption von individuellen Akteur*innen verknüpft ist. Die zweifach relationale Perspektive eröffnet nun die Möglichkeit, einen solchen *organisationalen Habitus* (Baier & Schmitz 2012: 200; Emirbayer & Johnson 2008: 19) als analytisches Konstrukt zu fassen: Die einzelnen organisationalen Einheiten setzen sich aus unterschiedlichen Mitgliedern zusammen, deren soziale Praktiken in je eigenen beruflichen, fachlichen oder disziplinären Habitus gründen. Die Praxis von organisationalen Akteur*innen in sozialen Feldern kann nicht alleine aufgrund der Positionierung im Feld und des verfügbaren Kapitals erklärt werden. Sie wird durch ein System von Dispositionen geprägt, das auf der Entwicklung und Geschichte der Organisation im Feld basiert (Baier & Schmitz 2012: 200).

»One might speak here of an organizational habitus [...]. [E]mergent position-takings on the part of an organization must always be understood, not as the self-expressions of a singular actor, but rather, as compromise products of a whole complex of negotiations and contestations unfolding over time within that organization *understood as itself a field*« (Emirbayer & Johnson 2008: 19; Hervorhebung im Original).

Die jeweiligen Stellungnahmen – »ein strukturiertes System der Praktiken und Äusserungen der Akteure« (Bourdieu & Wacquant 1996: 136) – basieren auf der Korrespondenz zwischen den »objektiven Positionen«, die Organisationen (bzw. Organisationseinheiten) in einem bestimmten sozialen Feld (bzw. einer Organisation als soziales Feld) einnehmen, und dem durch Institutionalisierung ökonomischer und sozialer Verhältnisse erworbenen System von Dispositionen, dem organisationalen Habitus. Dieser prägt somit und wird gleichzeitig geprägt durch das Konvolut aller formellen und informellen organisationalen Strukturen, Kulturen, Identitäten und Traditionen. Der organisationale Habitus ist mit anderen Worten weniger Leib, sondern vielmehr »Ding« gewordene Geschichte« (Florian 2008: 144).

Die Erweiterung des Akteurskonzepts auf kollektive Akteur*innen (wie Unternehmen, Universitäten oder staatliche Organe) ermöglicht es nun, kollektive Praktiken und Äusserungen (wie Strategien, Forschungsschwerpunkte oder Investitionen) als »Kompromissprodukte« organisationaler Konflikte und Kollaborationen zu untersuchen. Durch kollektive Stellungnahmen formulieren die Akteur*innen ihre Sicht auf soziale Phänomene, versuchen aber gleichzeitig, ihre Sichtweisen als die jeweils gültigen durchzusetzen. Konflikttheoretisch betrachtet streben sie damit zum einen nach der Dominanz innerhalb des noch wenig konturierten Raumes zwischen ihren Feldern, andererseits aber auch nach der Verstetigung, d. h. Reproduktion ihrer Organisationen als soziale Felder.

2.5 Zwischenräume als analytische Erweiterung der Feldtheorie

Wo entstehen neue Wissensgebiete? Wie können diese in aktuellen soziologischen Feldtheorien verortet werden? Verschiedene relationale Ansätze lokalisieren die Entstehung von Neuheit in einem Bereich, der ausserhalb der Routinen und etablierten Logiken in sozialen Praxisfeldern situiert ist (Passoth & Rammert 2016; Powell & Sandholtz 2012; Stark 2009). Solche Orte sind nicht als abgrenzbare Entitäten mit fixen Identitäten und Sinnhaftigkeit zu betrachten, sondern ihre Bedeutung wird kollektiv durch die Beziehungen der involvierten Akteur*innen ausgehandelt. Aus theoretischer Perspektive stellt sich deshalb die Frage, wie ein analytisches Modell beschaffen sein müsste, um dieses Dazwischen und gleichzeitig die vielfältigen Verbindungen, die zu etablierten Feldern existieren, analytisch greifbar machen zu können. Zwischenräume werden in verschiedenen Theorietraditionen behandelt, ohne allerdings bis dato einen gebührenden Platz in der Sozialtheorie gefunden zu haben.⁹

9 Konzepte von Zwischenräumlichkeit existieren in verschiedenen Theorietraditionen: In der an Bourdieu anschliessenden Feldtheorie (Witte & Schmitz 2019; Herberg 2018a) werden zwischenräumliche Phänomene als »transversale Felder« bzw. »Felder zwischen Feldern« konzipiert und netzwerkanalytisch analysiert. Ebenfalls netzwerkanalytisch untersuchen Arbeiten der neo-institutionalistischen, organisationalen Feldtheorie die Relationen und Austauschbeziehungen von Akteur*innen und Organisationen in Zwischenräumen (»interstices«) (Korff et al. 2015, 2017; Powell et al. 2017). Rammert und Passoth (2016) sprechen von reflexiven »Innovationsfeldern«, in denen die Wissensproduktion disziplinäre und institutionelle Grenzen, die funktional differenzierte Subsysteme charakterisieren, überschreitet und in Netzwerken heterogener Elemente an deren Rändern oder im Bereich dazwischen organisiert ist. Im Symbolischen Interaktionismus wiederum wird dafür der Begriff der Arena verwendet, der die Schnittflächen sozialer Welten bezeichnet (Abbott 2005: 249; Clarke 1991). Shinn

Das Feldmodell Bourdieus deutet, wie bereits ausgeführt, die feldübergreifenden Relationen solcher Räume bzw. zwischen Räumen und etablierten Feldern als Dominanz- und Abhängigkeitsverhältnisse innerhalb des Machtfeldes (Schmitz et al. 2017). Ähnlich verhält es sich bei der Theorie strategischer Handlungsfelder von Fligstein und McAdam. Sie differenzieren zwar Felder und Räume als unterschiedliche Sphären des Sozialen. Dabei definieren sie strategische Handlungsfelder als sozial konstruierte Sphären, die organisiert sind und bestimmten Regeln unterliegen, während der soziale Raum »unorganisiert« sei.

»We call the terrain of action within which all of these collective actors operate a strategic action field when it is well defined and unorganized social space when it is not« (Fligstein & McAdam 2012: 5).

Trotz dieser konzeptuellen Unterscheidung äussern sie sich aber kaum weiter zur Ausgestaltung und Verortung von sozialen Räumen bzw. wann solche sozialen Räume zu strategischen Handlungsfeldern werden und wann nicht (ebd.: 171).¹⁰

Gil Eyal's Konzeption von Räumen zwischen Feldern (*spaces between fields*) bietet hier einen Anknüpfungspunkt: Es handelt sich um eine zwischenräumliche Differenzierung des Feldmodells von Bourdieu, das geeignet ist, solche »unorganisierten« Phänomene des Sozialen analytisch zu fassen. Zwischenräume bzw. Räume zwischen Feldern, die ich in dieser Arbeit synonym verwende, beschreiben offene und unbestimmte Sphären, in denen Akteur*innen unterschiedlicher etablierter Felder aufeinandertreffen und spezifische Austauschbeziehungen knüpfen können (Eyal 2013b, 2013a).¹¹ Eyal und Pok (2015: 45) charakterisieren Zwischenräume wie folgt, um diese von temporären Phänomenen, die gerade nicht dauerhaft sind (Furnari 2014: 12ff.), und entstehenden Feldern (»fields-in-the-making«), die sich im Prozess der Ausdifferenzierung und Autonomisierung befinden, zu differenzieren:

1. *Durchlässigkeit*: Durchlässige Grenzen erlauben das Eintreten unterschiedlicher Akteur*innen aus umliegenden Feldern und das Knüpfen von Beziehungen zwischen diesen.
2. *Unter-Regulierung*: Es existieren keine oder nur schwach ausformulierte Regeln dazu, was legitimerweise in diesem Raum getan werden darf – eine Eigenschaft, die an die Abgrenzung von unorganisiertem sozialem Raum gegenüber organi-

und Joerges (2002, 2004) haben daran anknüpfend in der pragmatistischen Wissenschafts- und Technikforschung den Begriff der interstitiellen Arena eingeführt. Schliesslich beschreibt das Konzept der *trading zone* (Galison 1997) die Koexistenz und den sprachlichen Austausch verschiedener Kulturen (vgl. Kap. 2.6). Liu (2021) schliesslich leistet in Anlehnung an Simmels Formalsoziologie einen Versuch der Systematisierung von Theorien des Zwischenräumlichen, indem er die Relationen zwischen sozialen Räumen in einem Kontinuum zwischen Heterogenität und sozialer Distanz verortet.

10 So schreiben Fligstein und McAdam (2012: 171) bezeichnenderweise an dieser Stelle: »A field that is not yet organized might have multiple groups with quite different conceptions of what is at stake. These groups could differ on means, ends, and/or methods«.

11 Die Konzeption von Räumen zwischen Feldern unterscheidet sich deutlicher vom Bourdieu'schen Begriff des sozialen Raumes (*social space*) (vgl. Kap. 2.2), den Fligstein und McAdam etwas irreführend zur Abgrenzung gegenüber organisierten, hierarchisch strukturierten Handlungsfeldern verwenden.

- sierten, hierarchisch strukturierten Feldern bei Fligstein und McAdam (2012: 5) anknüpft (vgl. Zitat oben).
3. *Hohe Einsätze*: Auf dem Spiel stehen hohe und heterogene (Kapital-)Einsätze und Gewinne, die sich auf unterschiedliche Ursprungsfelder beziehen.
 4. *Schwache Institutionalisierung*: Multiple Expertise-Netzwerke können koexistieren, wobei sich weder eine klare Arbeitsteilung noch eine spezifische Wertigkeit der verschiedenen Beiträge einstellt. Eyal und Pok (2015: 45) führen dies auf drei unterschiedliche Dynamiken zurück: Erstens können ungelöste Konflikte bei der Feldgenese und Institutionalisierung zu *Pattsituationen (stalemate)* zwischen verschiedenen Akteur*innen *führen*. Zweitens machen es durchlässige Grenzen, Unter-Regulierung sowie die hohen Einsätze vor allem für marginale Akteur*innen – die Häretiker*innen bei Bourdieu – attraktiv, sich für *Beutezüge (raids)* kurzfristig in solche Räume zu begeben (ebd.). Die Beute kann dann in feldspezifisch legitimes Kapital transferiert werden, was die eigene, marginale Position im Herkunftsfeld verbessert. Drittens sind das Dazwischen-Sein und die Unterbestimmtheit nicht Folge gescheiterter Professionalisierungsbemühungen bzw. schwacher Institutionalisierung, sondern es handelt sich vielmehr um strategische Ambiguität (Eisenberg 1984; Leitch & Davenport 2007), einen Raum offen zu halten, damit gerade keine Institutionalisierung erfolgt, die mit staatlichen und anderen Regulierungen einhergeht.

Die Charakteristika erlauben es, Räume zwischen Feldern als empirisch überprüfbare Erweiterung der Feldtheorie zu fassen: Durchlässigkeit, Unter-Regulierung und schwache Institutionalisierung sind keine temporären Phänomene, sondern Grundlage für die Hybridität und die permanente Instabilität solcher Räume. Die Konzeption von Zwischenräumen als Möglichkeitsräumen (*spaces of opportunity*) (Eyal 2013b; Eyal & Pok 2015) verweist auf Orte ohne fixe Regeln, in denen Gegenstände, Praktiken oder Expertisen kombiniert werden können, die ansonsten separiert bleiben müssen. Das »objektiv gegebene Möglichkeitsfeld« (Bourdieu 1987: 188; Hervorhebung im Original) der Akteur*innen erweitert sich somit über Feldgrenzen hinaus, da in Räumen zwischen Feldern ohne »Gesichtsverlust« experimentiert und Beziehungen eingegangen werden können, die in hierarchisch organisierten Feldern nicht möglich sind. Solche zwischenräumlichen Settings schaffen gute Bedingungen für die Genese von Expertisen, da sich diese in Netzwerke etablieren können, die vielfältige Anschlussmöglichkeiten sowie Kooperationsmöglichkeiten für umgebende Felder eröffnen (Callon & Rabeharisoa 2003: 199). Trotz des produktiven Potenzials von Räumen zwischen Feldern existieren konflikthafte Beziehungen fort; zudem lauern auch Gefahren, nicht zuletzt da sich hier eine Vielzahl an Akteur*innen – Hybride, Bastarde und andere »Monster« (Bowker & Star 1999: 300ff.; Law 1991) – mit (vermeintlich) illegitimen Motiven und Interessen tummeln.¹²

Eyal verwendet das Konzept des Zwischenraumes zur Verortung neuer, hybrider Wissensbereiche, was auch Teil des Erkenntnisinteresses dieser Arbeit ist. In Anlehnung an die *Actor-Network-Theory* konzipiert Eyal Autismus-Therapien als »Expertise-Netzwerke« (2013a: 863; Eyal et al. 2010: 250ff.), die neben kollektiven und menschlichen Akteur*innen, Kompetenzen und Fähigkeiten auch Regulatorien, Handbücher,

¹² Ich danke Markus Unternährer für den Hinweis.

Methoden, Checklisten und weitere Devices umfassen.¹³ Solche Netzwerke erlauben es den verschiedenen Akteur*innen, an den materiellen und symbolischen Ressourcen, die in Zwischenräumen verteilt werden, teilzuhaben. Expertise wird in diesem Sinne nicht in erster Linie durch Expert*innen, eine bestimmte Profession oder soziale Gruppe, die über ein Monopol an Wissen verfügen, von dem sie andere ausschließen können, getragen, sondern durch ein Dispositiv bzw. ein institutionelles, politisch-ökonomisches Arrangement, das solche Netzwerke bzw. Expert*innen-Statements überhaupt erst möglich macht. In einer solchen Konzeption von Expertise äussert sich ein anderer Machtbegriff als im traditionellen Professionsverständnis, bei dem es stark um Autonomie, d. h. die Kontrolle von Jurisdiktion (Zuständigkeit) geht: Professionen setzen bestimmtes Wissen gegenüber anderen Gruppen als Standard durch und schliessen diese damit davon aus (Abbott 1988). Eyal verwendet in Anlehnung an Foucault einen verteilten Machtbegriff, d. h., die Mächtigkeit einer Expertise liegt darin, dass sie offen und anschlussfähig für verschiedene Professionen, Disziplinen und Felder ist.¹⁴

Neben der Anwendung des Konzepts von Räumen zwischen Feldern auf die Genese von Wissensfeldern (Eyal 2002, 2013a; Eyal & Pok 2015; Navon & Eyal 2014; Stampnitzky 2013) greifen Arbeiten in unterschiedlichen Theorietraditionen (Feldtheorie, Neo-Institutionalismus, Netzwerktheorie, Ecology-Ansatz) darauf zurück, um die Etablierung hybrider Praktiken (Lee 2005; Furnari 2014; Heimstädt & Reischauer 2019), institutionellen Wandel (Furnari 2016; Heimstädt & Ziewitz 2019; Morril 2017) sowie die Herausbildung von »Hybridorganisationen« (Laux 2016; Meier & Meyer 2020) wie Stiftungen (Korff et al. 2017; Powell et al. 2017), Think Tanks (McLevey 2015; Medvetz 2012) oder wissenschaftlichen Grenzorganisationen (Åm 2013; Guston 1999, 2000) in mehrdeutigen Governancekontexten zu erklären. Die Beispiele beschreiben hybride Wissensbereiche, Praktiken und Organisationen, die synchron in verschiedenen Feldern verortet sind und sich nicht auf feldspezifische Handlungslogiken reduzieren lassen.¹⁵ Sie verknüpfen die involvierten Felder, begünstigen den Fluss von Informationen, Wissen sowie Ressourcen und halten gleichzeitig multiple Logiken aufrecht. Damit überschreiten und vermischen solche zwischenräumlichen Entitäten im Sinne der »fragmentalen Differenzierung« sowohl disziplinäre und institutionelle Grenzen als auch »funktional fein säuberlich [getrennte] Leitreferenzen«, die funktional differenzierte Subsysteme oder soziale Praxisfelder kennzeichnen (Passoth & Rammert 2016: 41).

13 Expertise wird dabei in Anlehnung an Nikolas Rose (1992) nicht als Zuschreibung bzw. Verfügen eines spezifischen Skills-Set (≠ Professionelle als Expert*innen), sondern verteilt, als Netzwerk von Akteur*innen, Instrumenten, Wissensbeständen und institutionellen Arrangements verstanden.

14 Rose spricht davon, dass die Wirkmächtigkeit psychologischer Expertise, d. h. die Kalkulierbarkeit von Intersubjektivität, darin liegt, dass sie in unterschiedliche gesellschaftliche Bereiche und Organisationen (Spitäler, Schulen, Fabriken, Betriebe etc.) diffundieren konnte (*generosity*). Allianzen mit verschiedenen Akteur*innen hätten dann zur Re-Definition von Normalität und Krankheit geführt (Rose 1992: 356).

15 Bösch (2016) spricht für wissenschaftlich-technologische Felder in diesem Zusammenhang von »hybriden Wissensregimen«, die »regulierte soziale Räume zur Artikulation, Definition und Lösung sozio-technischer Probleme und damit der problemzentrierten Wissensproduktion dar[stellen]« (ebd.: 63).

Während manche der von Eyal und Pok identifizierten Charakteristika wie hohe Einsätze, Pattsituationen oder Beutezüge eher an den konfliktbetonenden Kern der Bourdieuschen Feldtheorie anschliessen, betonen Durchlässigkeit, Unter-Regulierung sowie strategische Ambiguität eher das grenzüberschreitende, kooperative Moment von Zwischenräumen, in denen dynamische Potenziale mobilisierbar sind.¹⁶ Dabei sind es vor allem zwei Praktiken, die zur Ausbildung und Instabilität von Räumen zwischen Feldern beitragen: Zum einen halten Praktiken von Grenzziehung und Grenzüberschreitung die Räume offen und durchlässig, was neben konflikthaften Beziehungen auch kooperatives Verhalten zwischen Akteur*innen ermöglicht. Zum anderen führen diskursive Praktiken zu multiplen, koexistierenden Verständnissen: Akteur*innen formulieren unterschiedliche Perspektiven darauf, was legitimerweise in solchen Räumen getan werden darf. Wie ich noch zeigen werde, tragen diskursive wie nicht-diskursive Praktiken von Repräsentation und Imagination zur Permanenz und Kontinuität solcher Arrangements bei. Zunächst erörtere ich das widersprüchliche Potenzial von Grenzarbeit.

2.6 Das trennende und verbindende Potenzial von Grenzarbeit

Durchlässige Grenzen sind konstitutiv in der Herausbildung neuer Wissensgebiete als zwischenräumlichen Phänomenen. Indem Grenzen soziale Entitäten synchron trennen und verbinden, schaffen sie einen Raum, in dem widersprüchliche, aber auch produktive Dynamiken in Gang gesetzt werden, die sowohl zu Kollaborationen und dem Knüpfen neuer als auch zur Auflösung existierender Beziehungen sowie zum Verfall gemeinsamer Institutionen führen können. Durch Praktiken von Grenzarbeit stecken Akteur*innen nicht nur bestimmte Territorien ab und schliessen andere davon aus, sondern etablieren synchron dazu Verbindungen über etablierte Grenzlinien hinweg und verknüpfen getrennte Bereiche.

Die Feldtheorie Bourdieus berücksichtigt die konstitutive Rolle von Grenzen und Grenzziehungen in der Herausbildung neuer Felder zu wenig: Obwohl die Frage nach dem Gegenstandsbereich und seinen Grenzen für Feldtheorien analytisch zentral ist – es geht eigentlich immer um die Frage, was dazugehört und was nicht (Bongaerts 2011: 120f.) –, werden die Grenzen selbst folglich nur selten thematisiert. Die Theoretisierung der Grenzen von Feldern ist jedoch notwendig, um feldübergreifende Beziehungen fassen zu können. Da es sich bei Feldern um analytische Konstrukte handelt, stellt sich stets die Frage danach, was diese umfassen und was nicht: Wo liegen die Grenzen eines Feldes? Welche Akteur*innen sind Teil eines Feldes, welche nicht?

Bourdieu antwortet auf die Frage nach den Grenzen eines Feldes damit, dass diese »[dort] liegen, wo die Feldeffekte aufhören« (Bourdieu & Wacquant 1996: 131). In seiner konstruktivistischen Lesart handelt es sich um eine empirisch zu klärende Aufgabe, die »keine Antwort a priori zu[lässt]« (ebd.). So stellt sich die Frage nach dem »Zuschnitt und der Begrenzung von Feldern« primär hinsichtlich ihrer (relativen) Autonomisierung, den Logiken der Praxis sowie den auf dem Spiel stehenden Einsätzen bzw.

16 Die Eigenschaften beschreiben auch Bourdieus literarische Salons (vgl. Kap. 2.3).

Gewinnen (Witte & Schmitz 2019: 33; Bongaerts 2011). Es obliegt also dem Forschenden zu zeigen, wo die »Feldeffekte« beginnen und wo sie enden.¹⁷

In der Theorie strategischer Handlungsfelder wiederum besteht ein dynamisches Verständnis: Die Grenzen eines Feldes sind je nach Situation und Gegenstand variabel (Fligstein & McAdam 2012: 10). Veränderte Bedingungen in übergeordneten Feldern, solchen mit mehr Ressourcen oder legaler Autorität können die Grenzen von Feldern verschieben, zur Auflösung oder Neubildung führen. Gleichzeitig stellen Fligstein und McAdam (2012: 166) die Frage nach den Grenzen explizit nur für stabile Felder. Das widersprüchliche und zugleich produktive Potenzial von Grenzen bzw. Grenzziehungen zeigt sich allerdings gerade im Falle von zwischenräumlichen Phänomenen, deren Entwicklung und Status noch unklar ist. Demgegenüber haben sich Arbeiten der Wissenschaftssoziologie der Theoretisierung von Grenzen angenommen und zeigen das produktive Potenzial synchroner Grenzziehungen und Grenzüberschreitungen auf.

2.6.1 Die Synchronizität von Praktiken der Grenzziehung und Grenzüberschreitung

Grenzen und Grenzziehungen sind zentrale soziologische Konzepte zur Erforschung gesellschaftlicher Sphären (Bongaerts 2011; Lamont & Molnár 2002; Wimmer 2013). Dabei kann zwischen symbolischen und sozialen Grenzen unterschieden werden: Symbolische Grenzen sind konzeptuelle Unterscheidungen, die Akteur*innen treffen, um Objekte, Personen und Praktiken zu kategorisieren, während soziale Grenzen objektivierte Formen gesellschaftlicher Differenzen darstellen, die sich im ungleichen Zugang und Verteilung von Ressourcen und Möglichkeiten manifestieren. Sie werden damit auch zur Grundlage der Entstehung von Kollektiven wie Gruppen, Organisationen oder Disziplinen (Lamont & Molnár 2002: 168f.). Symbolische und soziale Grenzziehungen sind inhärent miteinander verknüpft: Symbolische Grenzziehungen bilden oft das Fundament, auf dem soziale Grenzen ihre Wirkungen entfalten können. Umgekehrt konkurrieren soziale Kollektive über die Etablierung und Ausgestaltung symbolischer Klassifikationssysteme (Bourdieu 1987).

Die Wissenschaftssoziologie hat sich intensiv mit den Grenzen und Grenzziehungsprozessen ihres Gegenstandsbereichs zu anderen Feldern beschäftigt: *Boundary work* beschreibt symbolische wie soziale Grenzziehungen der Wissenschaft gegenüber nichtwissenschaftlichen Bereichen als auch innerhalb wissenschaftlicher Disziplinen und Professionen (Gieryn 1983, 1999). Komplementär dazu beschreibt *boundary crossing* das Überschreiten solcher Grenzen und das Verknüpfen ehemals getrennter Bereiche (Akkerman & Bakker 2011; Klein 1996). Akteur*innen ziehen und überschreiten Grenzen im Sinne einer strategischen Praxis, die ihren Zielen und Interessen entspricht (Gieryn 1999: 23), d. h. der Korrespondenz ihrer sozialen Positionen im jeweiligen Feld und ihren inkorporierten Dispositionen. Grenzen als soziales Phänomen sind demnach hybrid, sie trennen und verbinden zugleich und überschreiben die Räume,

17 Andererseits argumentiert Rehbein (2003: 88f.), dass die Grenzen von Feldern immer unscharf sind, denn einzelne »Spiele« liessen sich selten auf einen Bereich begrenzen.

die sie entwerfen.¹⁸ Interdisziplinäre Felder oder heterarchische Organisationen mit horizontalen Strukturen basieren geradezu auf der Durchlässigkeit, dem *boundary crossing* von Ideen, Methoden oder epistemischen Praktiken (Beunza & Stark 2004; Kellogg et al. 2006; Powell & Sandholtz 2012; Stark 2009).

Der in Kapitel 2.3 skizzierte idealtypische Prozess der Herausbildung neuer Wissensfelder kann als Wechselspiel von symbolischen und sozialen Grenzziehungen einerseits und Grenzüberschreitungen andererseits betrachtet werden: Akteur*innen aus verschiedenen Feldern überschreiten die – symbolischen wie sozialen – Grenzen, die zwischen ihren Bereichen existieren. Sie gehen strategische transversale Kooperationen ein, auch wenn sie möglicherweise nicht in allen Ansichten über den zentralen Gegenstand des Feldes übereinstimmen. Trotzdem bekräftigen die Akteur*innen die Notwendigkeit eines »Perspektivenwechsels« oder der Anwendung neuer Methoden auf alte Fragestellungen, sie formulieren eigene epistemologische, theoretische und methodologische Prämissen, auf deren Grundlage sie andere Akteur*innen einladen oder ablehnen. Anschließend formieren sie sich in kollektiven Formationen wie Interessensgruppen, Fachgesellschaften, d. h., disziplinäre Gemeinschaften nutzen epistemologische Grenzziehungen dazu, um eigene Identitäten auszubilden bzw. zu bekräftigen, Territorien abzustecken und gegenüber Aussenstehenden zu verteidigen (Mäkinen 2018). Werden diese von den legitimen, d. h. herrschenden Instanzen des Feldes (wie Förderorganisationen, Akademien, Universitäten etc.) anerkannt, können symbolische, diskursive Praktiken somit in der Zuteilung von finanziellen Mittelzuschüssen oder der Strukturierung organisationaler Einheiten wie Instituten und Fakultäten münden. Gerade die Herausbildung neuer kollektiver Formationen zeigt, dass die Wechselwirkungen von symbolischen und sozialen Grenzziehungen räumliche Effekte (Gieryn 1999: 10ff.) haben: Sie binden manche Akteur*innen in die Konstruktion neuer Wissensgebiete ein, während andere aussen vor bleiben. Sie schaffen somit für die Inkludierten Möglichkeitsräume und erweitern deren Handlungsoptionen über die Ursprungsfelder hinaus.

Die Beispiele zeigen, dass Praktiken der Grenzziehung und Grenzauflösung koexistieren. Darin liegt das widersprüchliche, aber auch produktive Potenzial von Grenzen: Indem sie trennend und verbindend zugleich wirken, schaffen sie ein spannungsgeladenes Setting, in dem Dynamiken in Gang gesetzt werden, die sowohl zu Kollaborationen und dem Knüpfen neuer Beziehungen führen können, was die Herausbildung neuer Gegenstände begünstigt, als auch zur Auflösung existierender Beziehungen und zum Verfall kollektiver Institutionen.

Ein weiteres, damit verknüpft Charakteristikum liegt darin, dass Grenzziehungen oft latent bleiben (Mäkinen 2018): Kollektive Akteur*innen beteiligen sich am Aufbau neuer Entitäten, ohne zwingend dieselben Ziele und Strategien zu teilen; sie sind vielmehr durch die Kopräsenz unterschiedlicher Interessen, Strategien und Machtverhältnisse charakterisiert. Würden diese permanent aktiv mobilisiert, stellt sich die Frage, wie und weshalb es überhaupt zur Bildung stabiler kollektiver Formen kommen kann. Indem sie Grenzziehungen latent halten, d. h. im Hintergrund vor sich hin schwelen lassen, stellen sie das verbindende Moment in den Vordergrund und si-

18 Die Idee geht auf Simmel (1995[1902]) zurück, der sie in seinem Essay zum Bilderrahmen entwickelte. Ich danke Andrea Glauser für diesen Hinweis.

gnalisieren zugleich Einheit nach aussen, womit die Kohäsion des Kollektivs erhalten bleibt.¹⁹

2.6.2 Die zwischenräumlichen, materialen und akteurialen Dimensionen von Grenzarbeit

Die Herausbildung neuer Wissensgebiete durch Praktiken der Grenzziehung und Grenzüberschreitung vollzieht sich in drei unterschiedlichen Dimensionen: Erstens verweist die Synchronizität und Parallelität solcher Praktiken bei der Genese von Neuheit auf die zwischenräumliche Dimension von Grenzen (Abbott 1988, 1995; Eyal 2013b; Gieryn 1999; Liu 2018): Wo finden Prozesse der Grenzziehung und Überschreitung statt? Erfolgen diese innerhalb, ausserhalb oder zwischen den involvierten Feldern? An diese Fragen anknüpfend haben sich verschiedene Arbeiten mit der Hybridität solcher Räume beschäftigt: Konzepte wie *trading zone* (Galison 1997), *boundary spaces* (Beauchamp & Thomas 2011; Mäkinen 2018; Williams 2013), *hybrid space* (Gilmour 2006) oder *relational spaces* (Kellogg 2009) bezeichnen weniger fixe, stabile Entitäten und trennende Grenzen im Sinne von geometrischen Linien, sondern hybride, wandelbare Räume als Zonen von Verknüpfungen und Verhandlungen dazwischen – Grenzen können deshalb als »fuzzy zones of separation and connection« (Eyal 2013b: 175) verstanden werden. Allerdings bieten die Begriffe kaum konzeptuelle Heuristiken oder Werkzeuge, um solche Grenz- oder Zwischenräume analytisch fassbar zu machen. Vielmehr operieren sie als Chiffren, um Überlappungen zwischen Feldern oder institutionellen Settings zu beschreiben.

Eine zwischenräumliche Konzeption von Grenzen bzw. Grenzarbeit bietet gegenüber einem geometrischen Verständnis den Vorteil, dass sie in existierende Feldtheorien integriert werden kann. Anstatt das produktive Potenzial von Praktiken der Grenzziehung und -überschreitung in neue Felder zu verlagern und damit in gewisser Weise zu neutralisieren, ist vielmehr eine Berücksichtigung der Hybridität von Grenz- bzw. Zwischenräumen in der Feldtheorie notwendig (Eyal 2013b).

Eine zweite Dimension umfasst jene materiellen Elemente und Objekte (*boundary objects*), die grenzüberschreitende Kommunikation und Kooperation zwischen verschiedenen Feldern erlauben (Bowker & Star 1999; Star & Griesemer 1989). Die materielle und die zwischenräumliche Dimension von Grenzen sind miteinander kompatibel bzw. komplementär, denn *boundary objects* haben eine verbindende Funktion:

»Die Begriffe ›Grenze‹ und ›Objekt‹ müssen aber vielleicht auch ein wenig näher erläutert werden. Oft impliziert Grenze so etwas wie Rand oder Peripherie, wie im Falle der Grenze eines Staates oder eines Tumors. Hier jedoch soll Grenze einen gemeinsamen Raum bedeuten, in dem genau diese Wahrnehmung von Hier und Dort durcheinandergerät. Diese gemeinsamen Objekte bilden die Grenzen zwischen Gruppen durch Flexibilität und gemeinsame Struktur – sie sind das Material des Handelns. Ursprünglich zog ich in Erwägung, sie ›marginale Objekte‹ zu nennen, was aber noch verwirrender gewesen wäre« (Star 2017: 214).

19 Gleichzeitig können sie aber, wenn sie manifest werden, zur Spaltung von Entitäten führen (Mäkinen 2018).

Solche *boundary objects* sind in verschiedenen epistemischen Gemeinschaften enthalten und erfüllen die unterschiedlichen Anforderungen der Praxisphären. Sie können den jeweiligen lokalen Bedingungen angepasst werden, behalten aber auch gemeinsame Identitäten und Stabilität über Feldgrenzen hinweg (Bowker & Star 1999: 15f.). (Meta-)Daten (Acker & Donovan 2019; Reichert 2018), Methoden (Ruppert et al. 2013; Savage 2013), Tools (Coté 2014; Mackenzie 2013), Algorithmen (Lowrie 2017) und soziotechnische Systeme (Beer 2019; Chrisman 1999) begünstigen so den Austausch von Informationen und Wissen und ermöglichen Kollaborationen über Feldgrenzen und Handlungslogiken hinweg, was zur Herausbildung neuer Wissensgebiete beiträgt.

Während die Logik der Multi- oder Interdisziplinarität (Klein 1996), wie sie dem Konzept der *trading zone*²⁰ unterliegt, die Grenzen zwischen Bereichen zu lockern oder aufzulösen versucht, werden Felder durch *boundary objects* miteinander verlinkt und in Beziehung gesetzt. Daraus resultiert keine neue Sprache, sondern die Einsicht, dass die andere Seite dieselben Gegenstände anders konzipiert, operationalisiert und misst, was dann zur Grundlage von Übersetzung wird (Chrisman 1999). Dabei bleiben die Grenzen bestehen und die Logik der Aufteilung der sozialen Welt in Felder wird aufrechterhalten (Ribes 2019; Ribes et al. 2019).²¹

Schliesslich betrifft eine dritte Dimension von Grenzarbeit die Ebene der Akteur*innen: Grenzüberschreitende Rollen (*boundary-spanning figures*) haben eine lange Kontinuität in High-Tech-Industrien, in denen Forschende zwischen den involvierten Feldern hin und her wechseln können (Powell & Sandholtz 2012; Ribes 2019).²² Grenzorganisationen (Guston 1999, 2000), Intermediäre (Åm 2013; Bessy & Chauvin 2013; Meyer & Kearnes 2013) und *knowledge broker* (Herberg 2018b; Mäkinen et al. 2020) bilden bzw. schaffen Schnittstellen (*interfaces*) (Hookway 2014; Lipp 2017) zwischen verschiedenen Feldern wie Wissenschaft, Industrie und Politik, was die Übertragung von Ressourcen und Informationen unterstützt. Sie sind dabei allerdings nicht nur passive Vermittler, sondern prägen und mobilisieren bestimmte Expertisen oder politische Imperative und strukturieren somit die involvierten Felder (Meyer & Kearnes 2013). Darin liegt gewissermassen ihr performatives Potenzial, indem sie nicht nur übersetzen, sondern auch neue, hybride Räume schaffen und koordinieren (ebd.: 424; Kearnes 2013). Instabile, feldübergreifende Arrangements werden also gewissermassen in solchen Grenzorganisationen dauerhaft gehalten.

Stärker konflikthafte soziale und symbolische Grenzbeziehungen und Praktiken der Grenzüberschreitung können also koexistieren. Indem deren Wechselwirkungen räumliche Effekte zeitigen, schaffen sie für die inkludierten Akteur*innen neue Mög-

20 Die Metapher der »trading zone« (Galison 1997) geht über *boundary objects* hinaus, indem sie einen linguistischen Raum, das *Pidgin*, eröffnet, in dem verschiedene Gruppen in unterschiedlichen »Sprachen« miteinander kommunizieren und agieren können. Dies setzt oft eine gemeinsam geteilte Architektur bzw. Infrastruktur voraus wie physikalische Labore und Arbeitsräume. In zeitgenössischen wissenschaftlichen Kollaborationsarrangements nehmen diese Rolle etwa Programmiersprachen, Software oder Daten ein.

21 Ribes (2019) und Ribes et al. (2019) sprechen nicht von Feldern, sondern von Domänen (*domains*), die aber kompatibel sind mit der räumlichen Differenzierung des Sozialen in unterschiedliche Felder.

22 Powell und Sandholtz (2012: 95) bezeichnen Unternehmer*innen und Wissenschaftler*innen, die zwischen den verschiedenen Sphären des entstehenden Biotech-Sektors in den USA (Universitäten und Biotech-Start-ups) hin und her wechseln, als »boundary crossers« bzw. »amphibious entrepreneurs«.

lichkeitsräume und erweitern deren Handlungsoptionen. Praktiken von Grenzarbeit wohnt insofern ein widersprüchliches, aber auch produktives Potenzial inne: Sie wirken trennend und verbindend zugleich und schaffen damit einen spannungsgeladenen, dynamischen Raum, der sowohl zur Herausbildung neuer als auch zum Niedergang existierender kollektiver Formationen führen kann. Ein weiteres zentrales Moment liegt in ihrer Latenz: Indem soziale Kollektive ihre internen Konflikte zurückstellen, bleibt die Kohäsion erhalten, was in Prozessen der Genese neuer Felder oder Institutionen ein strategisches Mittel darstellt und Einheit nach aussen signalisiert.

2.7 Die Konstruktion von Vielstimmigkeit neuer Wissensgebiete durch Begriffsarbeit

Diskursive Praktiken von Begriffsarbeit stellen einen zweiten zentralen Praxismodus dar. Sie etablieren multiple Verständnisse von neuen Wissensfeldern und tragen so zur Hervorbringung und Permanenz hybrider Räume zwischen Feldern bei. Dabei unterscheidet sich zwischen Repräsentationen, womit Akteur*innen durch sprachliche Begriffe eine bestimmte Deutung des Phänomens fixieren und gegenüber anderen möglichen Bedeutungen durchsetzen (Kap. 2.7.1), und Imaginationen, d. h. kollektiven, öffentlichen Zukunftsentwürfen, mit denen politische und andere Akteur*innen Vorstellungen sozialen Zusammenlebens und sozialer Ordnung artikulieren, die durch technologisch-wissenschaftlichen Fortschritt erreicht werden können und sollen (Kap. 2.7.2). Zunächst bedarf es allerdings einleitend einer Klärung des Verhältnisses von Begriffsarbeit und Grenzarbeit.

Praktiken von Begriffsarbeit und Grenzarbeit sind einander konzeptuell nahe, befinden sich jedoch auf unterschiedlichen Analyseebenen: Begriffsarbeit dient zunächst dem Aufspannen neuer Wissensgebiete. Diese werden diskursiviert, d. h. in bestimmten Begriffen und Kategorien gefasst und in Form gebracht (Vilsmaier 2018: 125). Heterogene Perspektiven bestehen parallel nebeneinander, noch bevor eine bewertende Deutung und Privilegierung vorgenommen wird. Feldanalytisch formuliert eröffnet dies neue Möglichkeitsräume (Rip & Voss 2013: 43), was die Handlungsoptionen für die Akteur*innen erweitert. Im Untersuchungsfeld zeigt sich dies daran, dass multiple Begriffe (wie Big Data, Data Science, Data Analytics etc.), die inhaltlich miteinander verwandt sind, koexistieren können. Dies erlaubt es dann, Akteur*innen aus anderen Wissensgebieten zu inkludieren.

Daran anknüpfend dienen Praktiken von Grenzarbeit dann eher dazu, die jeweiligen Begriffe voneinander zu differenzieren, d. h. durch symbolische Grenzziehungen unterscheidbar zu machen (beispielsweise »Data Science scheint nicht schlecht als Begriff, besser als Big Data, weil damit noch mehr Themen assoziiert werden«), und anschliessend soziale Grenzen (wie die Etablierung von Forschungskollaborationen oder Studiengänge) zu markieren, durch die wiederum bestimmte Begriffsdeutungen reproduziert werden. Somit wirkt Begriffsarbeit eher fundierend, während Grenzarbeit darauf aufbaut.²³ Selbstverständlich kann das Verhältnis auch umgekehrt sein, indem

23 Konträr dazu konzipiert Vilsmaier (2018: 125) Begriffsarbeit neben dem Erkunden professioneller Profile und Identitäten sowie der quantitativen Arbeit mit Variablen, Indikatoren und Schwellenwerten als eine »Methode der Grenzarbeit«; eine Festlegung, die ich für meine Arbeit nicht übernehme.

zunächst Differenzen artikuliert werden, die dann in der Suche nach neuen Begriffen münden.

2.7.1 Repräsentationen als symbolische Deutungen neuer Wissensgebiete

Die Emergenz neuer Wissensgebiete wird durch eine Vielzahl unterschiedlicher – technologischer, politischer, ökonomischer, medialer etc. – Diskurse begleitet (Jasanoff & Kim 2015). Die involvierten Akteur*innen entwerfen unterschiedliche Perspektiven darauf, worin neue Wissensgebiete bestehen, welche Entwicklungen möglich sind und schliesslich wo die Grenzen zu verwandten Gebieten liegen. Es handelt sich um Stellungnahmen kollektiver Akteur*innen, die auf der Korrespondenz ihrer Positionierung im jeweiligen Feld und dem organisationalen Habitus entsprechen. Als kollektive Äusserungen sind diese stets als »Kompromissprodukte« (Emirbayer & Johnson 2008: 19) von feld- bzw. organisationsinternen Kämpfen und Kollaborationen zu betrachten. Sind solche Kompromisse einmal erzielt, repräsentieren die Akteur*innen die Wahrnehmung neuer Wissensgebiete nach aussen (gegenüber anderen kollektiven Akteur*innen) genauso wie nach innen (in den jeweiligen Organisationen als Feldern), d. h., sie fixieren durch sprachliche Begriffe eine bestimmte Deutung des Phänomens und privilegieren diese fortan gegenüber anderen möglichen Bedeutungen.²⁴

Bourdieu konzipiert Repräsentationen als die »symbolische Darstellung der sozialen Verhältnisse« (Fröhlich & Rehbein 2009: 414). Dabei ergibt sich eine Nähe zum Repräsentationsbegriff bei Stuart Hall (Marquardt 2016: 161), der Repräsentation wie folgt fasst:

»Representation is the production of meaning of the concepts in our minds through language. It is the link between concepts and language which enables us to refer to either the »real« world of objects, people or events, or indeed to imaginary worlds of fictional objects, people and events« (Hall 1997: 17).

Sowohl Bourdieu als auch Hall betonen also, dass der Sinn den Phänomenen nicht inhärent ist, sondern sozial konstruiert wird (Marquardt 2016: 161) – durch die gegebenen sozialen und ökonomischen Beziehungen, d. h. die Machtverhältnisse, bei Bourdieu, während bei Hall die Verknüpfungen zwischen Konzepten und Kategorien einerseits und Sprache andererseits im Zentrum stehen.

Am Beispiel des Untersuchungsgegenstandes kann dies wie folgt exemplifiziert werden: Neben der Suche nach den adäquaten Begriffen (vgl. oben) entwerfen soziale Akteur*innen wie Unternehmen, Universitäten oder politische Entitäten, indem sie Stellenanzeigen publizieren, Kompetenzprofile festlegen, Curricula aufbauen oder politische Fördermassnahmen lancieren, organisations- und feldspezifische Perspektiven darauf, worin die Datenwissenschaften bestehen (und worin nicht), welche Disziplinen sie umfassen (und welche nicht) sowie welche epistemologischen, theoretischen und methodologischen Prämissen das Wissensgebiet charakterisieren (und welche nicht). Diese multiplen symbolischen Deutungen tragen dadurch zur Konstruktion

24 Die Kompromisse sind dabei stets nur temporärer Art, sie verändern sich je nach Zustand der sozialen, ökonomischen etc. Beziehungen in den kollektiven Akteur*innen (sprich: Universitäten, Unternehmen, politischen Institutionen etc.) als sozialen Feldern.

und Wahrnehmung der Datenwissenschaften als heterogenes, transversales Wissensfeld bei. Gleichzeitig streben die genannten Akteur*innen danach, eine bestimmte Sichtweise des Phänomens gegenüber anderen möglichen Bedeutungen zu etablieren: Einerseits delegitimieren sie andere, konkurrierende Sichtweisen, grenzen deren Verfasser*innen aus und schliessen somit die sich eröffnenden Möglichkeitsräume ab. Andererseits überschreiten sie durch offene Begrifflichkeiten aber auch disziplinäre und soziale Grenzen: Damit schaffen sie Anschlussfähigkeit für andere Akteur*innen, was wiederum feldübergreifende Kollaborationen und Kooperationen ermöglicht.

2.7.2 Die koordinierende Rolle von Zukunftsentwürfen bei der Etablierung neuer Felder

Neben Repräsentationen spielen soziotechnische Imaginationen eine zentrale Rolle bei der gesellschaftlichen Konstruktion, Adaption und Verwendung von neuen Wissensgebieten oder technologischen Innovationen.²⁵ Nach Sheila Jasanoff sind soziotechnische Imaginationen kollektive, öffentliche Zukunftsentwürfe, die Vorstellungen sozialen Zusammenlebens und sozialer Ordnung artikulieren, die durch wissenschaftlich-technologischen Fortschritt erreicht werden können und sollen (Jasanoff 2015: 4). Soziotechnische Imaginationen können eine Vielzahl materieller und immaterieller Praktiken und Diskurse umfassen: Wissenschaftliche Berichte, politische Strategien, ökonomische Analysen, (massen-)mediale Inhalte und andere kulturelle Produkte formulieren und inszenieren Zukunftsvisionen für soziale Kollektivformationen. Soziotechnische Imaginationen stützen sich auf Metaphern, Narrative und andere kulturelle Formate, die ihrem jeweiligen Publikum vertraut und damit anschlussfähig sind (Dourish & Gómez Cruz 2018; Mische 2009). Sie erhalten ihre Sinnhaftigkeit demnach durch ihre Einbettung in kulturelle, zeitliche und räumliche Kontexte (Jasanoff 2015: 23ff.). Jasanoff betont, dass es primär kollektive Akteur*innen (wie Nationalstaaten, Organisationen oder soziale Bewegungen) sind, die soziotechnische Wissensgebiete konstituieren. Diese ihrerseits wiederum rahmen die Herausbildung individueller Identitäten (beispielsweise durch Prozesse der Subjektivierung von Individuen durch Self-Tracking), wodurch gesellschaftliche Zukunfts- und Technikdiskurse in den Habitus der Individuen inkorporiert werden (Jasanoff 2015: 23). Sie prägen damit auch die möglichen Lebens- bzw. Entwicklungsverläufe von Akteur*innen.

Soziotechnische Imaginationen sind nie neutral: Sie enthalten neben Beschreibungen des Status quo und prospektiven Szenarien, wie die Welt aussehen *könnte*, immer auch normative Vorstellungen darüber, wie die Welt aussehen *sollte*. Wenn beispielsweise konstatiert wird, dass die Erhebung und Analyse von Daten in einem bestimmten gesellschaftlichen Bereich wichtiger werden, ist darin meist auch die Forderung impliziert, dass sich die entsprechenden Akteur*innen an dieser Entwicklung ausrichten sollten. Kollektive Äusserungen über das zukünftige Verhältnis von sozialer Ordnung, Wissenschaft und Technologie prägen deren weitere Entwicklung immer auch mit; sie haben insofern performative Effekte auf das, was sie beschreiben (Beckert 2016; Esposito 2016). ‚Zukunft‘ erweist sich in diesem Sinn als ein umkämpf-

25 Neue soziotechnische Systeme sind stets instabil und kontingent, weil viele Entwicklungsverläufe offenstehen; diese müssen stabilisiert, d. h. viele Entscheidungen getroffen werden, die wiederum den weiteren Verlauf beeinflussen (Gugerli 2018).

ter Raum an Äusserungen und Praktiken von Akteur*innen aus unterschiedlichen sozialen Feldern (Brown et al. 2000: 5; Mische 2014).

Zukunftsentwürfe spielen in unterschiedlichen Feldern eine wichtige Rolle, indem sie die Genese neuer Wissensgebiete, Technologien oder Märkte rahmen (Beckert 2016; Jasanoff 2015; Mische 2014; Mützel 2010). In Wissenschaft und Forschung kommt Zukunftsprognosen und Erwartungen eine strukturierende Rolle zu (Jasanoff 2015; Jasanoff & Kim 2015). So verknüpft etwa die Forschungsförderung die Produktion von Zukunftspotenzialen mit der Zuweisung von ökonomischem Kapital und anderen Ressourcen (wie Infrastruktur) (Blümel 2018). Projektanträge und Forschungsprogramme werben damit, dass ihre allfällig geförderte Praxis in der Lage sein wird, die Entwicklung eines Wissensfeldes, von Methoden oder Technologien in der Zukunft mitzugestalten. Dies gilt insbesondere für Wissensbereiche, deren feldspezifische Zugehörigkeit (noch) offen ist oder verhandelt wird. Akteur*innen unterschiedlicher Felder formulieren divergierende Perspektiven darauf, worin für sie diese Expertisen bestehen, und eröffnen dadurch einen Raum zwischen den eigenen, etablierten Feldern (Eyal 2013b; Furnari 2014), der durch Vielstimmigkeit und Multiperspektivität gekennzeichnet ist. Durch solche nach »ausser« gerichteten diskursiven Praktiken, Formate und Materialisierungen tragen sie zur Konstitution, Entwicklung und Permanenz solcher zwischenräumlichen Phänomene bei.

In der Ökonomie werden Theorien, Prognoseinstrumente und andere Devices (Mackenzie & Millo 2003; MacKenzie et al. 2007) eingesetzt, um zukünftiges Marktgeschehen formal und berechenbar zu machen; aufgrund der hohen Unsicherheiten über zukünftige Entwicklungen dienen sie in erster Linie der Abstimmung der Aktivitäten zwischen wirtschaftlichen und anderen Akteur*innen (Beckert 2013, 2016; Beckert & Bronk 2018). Dadurch begründen solche Prognosen nicht nur neue Zukunftsperspektiven, sondern schreiben diese auch in organisationale Praktiken und politische Massnahmen ein (Holmes 2018). Durch die wechselseitige Beobachtung beginnen die Akteur*innen in verschiedenen Feldern, ihre Aktionen zu koordinieren und zur Erreichung der Ziele, die implizit oder explizit in solchen Strategien und Visionen artikuliert werden, beizutragen (Reichmann 2018). Die verteilten Aktivitäten tragen im Ergebnis dazu bei, die weitere Entwicklung eines entstehenden oder sich verändernden Feldes zu strukturieren und zu stabilisieren.

Im Gegensatz zu Wirtschaftsprognosen basieren Zukunftsszenarien im Politikfeld meist nicht auf komplexen Berechnungsschemata und mathematischen Formeln. Vielmehr kanalisieren sie die politische Aufmerksamkeit auf bestimmte Entwicklungen, während andere ausser Blick geraten (Brown et al. 2000). Durch Strategien und Aktionspläne nehmen politische Akteur*innen eine Priorisierung bestimmter Politikfelder vor, die in der Regel mit der Zuweisung von Ressourcen verbunden ist, die wiederum zur (relativen) Marginalisierung anderer Bereiche führt (Beckert 2016: 185). In diesem Sinne sind politische Prognosen und Zukunftsszenarien elementarer Bestandteil politischen Handelns (Koselleck 2006).

Die Akteur*innen des Politikfeldes sind dabei von besonderer Bedeutung: Ihre Absichtserklärungen und Investitionen signalisieren anderen Akteur*innen, dass sich ihrerseits Investitionen in neue Wissensgebiete oder Technologien auszahlen. Dies zieht wiederum weitere Akteur*innen an, die sich davon rasche und hohe Profite versprechen (Eyal & Pok 2015: 45). Die Beteiligten bemühen sich deshalb trotz möglicherweise divergierender Interessen darum, die neu entstehenden Räume zwischen

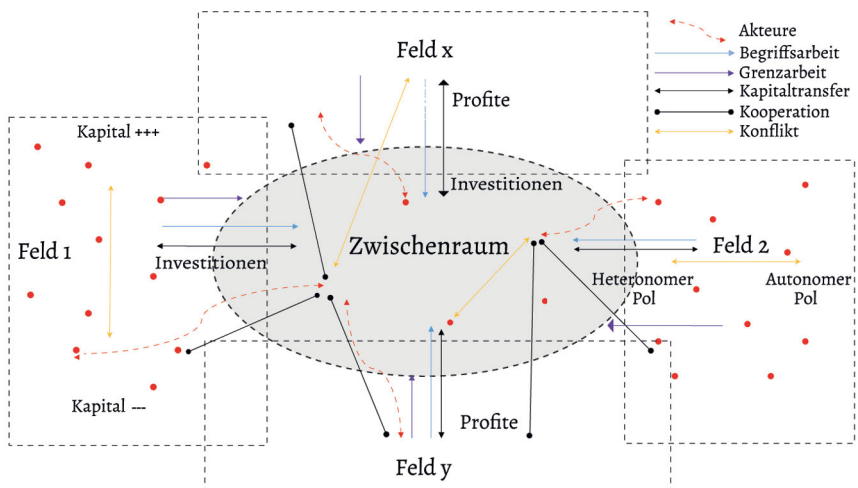
Feldern möglichst offen und anschlussfähig zu halten, ohne dass es zu einer Institutionalisierung, beispielsweise durch politische Regulierung, kommt.

Praktiken von Begriffsarbeit bilden somit ein weiteres Element im Theoriegerüst der Arbeit: Repräsentationen neuer Wissensgebiete sind als kollektive Äußerungen Produkte von feld- bzw. organisationsinternen Kämpfen und Kollaborationen. Die involvierten Akteur*innen fixieren dadurch eine bestimmte Deutung des Gegenstandes und privilegieren diese fortan gegenüber anderen möglichen Bedeutungen. Sozio-technische Imaginationen sind kollektive Zukunftsentwürfe neuer Wissensgebiete oder Technologien, die Vorstellungen sozialer Ordnung artikulieren. Hier hat sich gezeigt, dass Akteur*innen in verschiedenen Feldern mit Zukunftserwartungen und Prognosen operieren, um die Entwicklung ihrer Felder zu antizipieren, berechenbar zu machen oder bestimmte Bewertungen und Ressourcenzuweisungen vorzunehmen.

2.8 Analytisches Modell: Transversale Wissensgebiete als Räume zwischen Feldern

Im letzten Teilkapitel skizziere ich nun, wie die verschiedenen Theorieperspektiven zu einem analytischen Modell zusammengeführt werden können, das einen Beitrag zur Erklärung der Entstehung neuer Wissensgebiete zwischen etablierten sozialen Feldern leistet. Das analytische Modell baut auf Eyals Konzeption von Räumen zwischen Feldern (Eyal 2013b) auf, erweitert dieses aber auch in zentralen Punkten. Ich führe die zentralen Konzepte Zwischenraum, Feld, Akteur*innen, Grenzarbeit und Begriffsarbeit (Repräsentationen und Imaginationen) am Beispiel des empirischen Untersuchungsgegenstandes Datenwissenschaften aus. Zwischenräume sind offene, unterbestimmte Sphären, in denen Akteur*innen umliegender Felder aufeinander treffen und Austauschbeziehungen knüpfen können (vgl. Abbildung 1).

Abbildung 1: Analytisches Modell – Räume zwischen Feldern



Es handelt sich um Räume ausserhalb des Etablierten und Alltäglichen, die sich durch poröse Grenzen, multiple Logiken sowie eine permanente Instabilität auszeichnen. Zwischenräume sind von hierarchisch strukturierten Feldern umgeben, in denen Konflikte zwischen verschiedenen Akteur*innen die Regeln der Felder und deren Grenzen bestimmen. Die umgebenden Felder (Feld 1, 2, x, y) sind im Sinne des Bourdieuschen Feldmodells strukturiert, wobei die Kapitalvolumina die vertikale Achse und die Differenzierung in Autonomie und Heteronomie (bzw. legitimes und illegitimes Kapital) die horizontale Achse prägen.

Hervorzuheben sind einige Adaptionen gegenüber der Konzeption von Eyal: So ist es zum einen nicht einsichtig, weshalb nur marginale Akteur*innen in solchen Zwischenbereichen tätig sein sollen.²⁶ Als »Space Travellers« (Liu 2021: 133) bewegen sie sich zwischen sozialen Sphären, ohne allerdings in die Grenzpraktiken eines Feldes involviert zu sein. Vielmehr können sie multiple Identitäten ausbilden, was sie gegenüber anderen Akteur*innen privilegiert. Somit lohnt es sich gerade auch für kapitalmässig gut situierte Akteur*innen oder den Staat, sich in diese Bereiche zu begeben: Da es sich um Möglichkeitsräume handelt, werden hier Kooperationen und Praktiken opportun, die in hierarchisch organisierten und regulierten Feldern nicht legitim sind (Panofsky 2011). Beispielsweise können staatliche Institutionen prototypisch mit neuen Formen oder Expertisen experimentieren, die innerhalb bürokratischer Systeme mit vorgegebenen Abläufen und Regeln nicht vorgesehen sind.

Ein zweiter Punkt sind eindimensionale Kapitaltransfers: Eyal und Pok (2015: 45) argumentieren, dass Akteur*innen dank durchlässiger Grenzen Beutezüge durchführen und die erbeuteten Kapitalia, d. h. die Investitionen anderer Akteur*innen im Zwischenraum, in den Herkunftsfeldern in legitimes Kapital transferieren können. Da es sich um unterregulierte Räume handelt, ist es durchaus schlüssig, dass sich Akteur*innen diesen Umstand ohne eigene Investitionen zunutze machen wollen. Allerdings stellt sich die Frage, weshalb solche Kapitaltransfers bloss unidirektional ausgestaltet sein sollten: Einseitige Investitionen koexistieren vielmehr mit bidirektionalen Kapitaltransfers, beispielsweise in Form von staatlichen oder unternehmerischen Investitionen, die in zwischenräumliche Kooperationen oder Institutionen getätigt werden. Sie stehen für andere Akteur*innen zur Verfügung, eröffnen dem Staat dafür im Gegenzug den Zugriff und Transferoptionen auf darin zirkulierende Expertisen. Für Unternehmen ergeben sich aus den Kooperationen neue Profitmöglichkeiten, indem sie Zugang zu personalen Ressourcen oder methodisch-technischen Expertisen erhalten. Eine weitere Option sind verlorene Investitionen, die in neuen, »unsicheren Terrains« nicht unüblich sind: Gewisse Akteur*innen eines Feldes investieren grosse Summen in die Entwicklung neuer Technologien, die sich letztlich aber nicht auszahlen (Feld 1). Davon profitieren andere Akteur*innen (Feld y), die solche Investitionen ausnutzen (*pirating*), um eigene Ideen weiterzuentwickeln und im eigenen Feld profitabel zu machen (Poor 2012).

Ich konzipiere die Felder der Wissenschaft, der Hochschulbildung, der Ökonomie und der Politik als hierarchisch strukturierte Felder, die den Zwischenraum Daten-

26 Dies gilt besonders für ein komplexes Phänomen wie Innovation, das nicht auf die Praktiken einzelner marginaler Akteur*innen zurückzuführen, sondern vielmehr zu einem gesellschaftlichen Leitmotiv mutiert ist (Blümel 2018).

wissenschaften umgeben.²⁷ Deren Logiken sind primär auf die Reproduktion etablierter Strukturen ausgerichtet, was die Herausbildung eines neuen Wissensgebiets innerhalb der Felder unwahrscheinlich macht. Die Datenwissenschaften als neues Wissensgebiet entstehen demnach nicht innerhalb eines der genannten Felder, sondern in einem Raum dazwischen.

Aus der Literatur ergeben sich folgende Charakteristika für Zwischenräume: Die Grenzen neuer Wissensgebiete sind porös. Auf individueller Ebene können Akteur*innen zwischen den verschiedenen Feldern der Wissenschaft, der Ökonomie und den staatlichen Institutionen hin und her wechseln. Auf kollektiver Ebene eröffnen sich Möglichkeiten zur Kooperation von Akteur*innen wie Unternehmen, Hochschulen, Forschungseinrichtungen und staatlichen Institutionen. Dabei verfolgen alle ihre eigenen, feldspezifischen Ziele, wodurch sich die Koexistenz multipler (Handlungs-) Logiken (Passoth & Rammert 2016) ergibt: Unternehmen können ihre technologischen, sozialen und kulturellen Kapitalien erweitern, indem sie einerseits vom Wissens- und Methodentransfer in solchen Kollaborationen profitieren und andererseits Absolvent*innen und Forschende anwerben. Universitäten und Hochschulen können ihre Beziehungsnetzwerke in die ökonomischen Felder erweitern und erhalten Zugang zu »echten Daten«. Zudem können sie Anwendungsorientierung und Innovationsfähigkeit unter Beweis stellen, die zu dominanten politischen Imperativen geworden sind. Bildungs- und forschungspolitische Akteur*innen wiederum tätigen hohe Investitionen in neue Wissensgebiete, profitieren aber umgekehrt auch von Expertisen, die vermehrt zur Grundlage staatlicher Politiken werden. Dies eröffnet staatlichen Einheiten wiederum neue Handlungsoptionen.

Für die involvierten Akteur*innen lohnen sich solche Einsätze und Kooperationen in Zwischenräumen insbesondere dann, wenn aufgrund wechselseitiger Orientierungen feldübergreifend grosse Summen in diese investiert werden. Dies erlaubt es auch, heterogene Kapitalerträge zu erzielen: Während manche Akteur*innen primär ökonomische Interessen verfolgen, streben andere kulturelle, soziale oder feldspezifische Kapitalien bzw. Mischformen davon an.²⁸ Als Effekt solcher Kapitalakkumulation resultieren schliesslich auch Reputationsgewinne (symbolisches Kapital), beispielsweise indem Akteur*innen ihre Erneuerungsfähigkeit demonstrieren können.

Neben diesen Charakteristika zielt das analytische Modell in seiner Erklärungsleistung auf die fundierenden Praktiken von Grenz- und Begriffsarbeit, die zur Herausbildung neuer Wissensgebiete beitragen. Durch Praktiken von Begriffsarbeit schreiben die involvierten Akteur*innen dem Zwischenraum multiple Bedeutungen zu: Im Feld der Wirtschaft haben Berichte und Prognosen hohe Erwartungen an Data Scientists als Schlüsselfiguren datengetriebener Produktionsmodelle geweckt. Die

27 Eine vertiefte Analyse der umgebenden Felder der Wissenschaft (bzw. der Technikwissenschaften), Ökonomie (bzw. der datengetriebenen Ökonomie) und Politik (bzw. der Bildungs- und Forschungspolitik) kann in dieser Arbeit nicht geleistet werden. Ich verweise auf Arbeiten zum Feld der Wissenschaft (Bourdieu 1988, 2004; Fröhlich 2009), der Hochschulbildung (Bourdieu & Passeron 1971, 1973, 2007), der Ökonomie (Bourdieu 1998, 2002; Dederichs & Florian 2004; Florian & Hillebrandt 2006) und der Politik (Bourdieu 2001a, 2014; Janning 1998).

28 Im Sinne eines Möglichkeitsraumes können bestimmte Kapitalsorten konvertiert werden, ohne von einer übergeordneten Instanz, wie dem Feld der Macht oder dem Staat, abgesegnet werden zu müssen.

daraus resultierenden individuellen Zuschreibungen an Individuen (»Kompetenzprofile«) repräsentieren die Datenwissenschaften als einen ökonomischen Möglichkeitsraum, in dem eine äusserst grosse »Nachfrage« einem sehr knappen »Angebot« gegenübersteht. Damit strukturieren die Stellungnahmen ökonomischer Akteur*innen zudem die Erwartungen der Akteur*innen anderer Felder, namentlich in Politik und Wissenschaft hinsichtlich eines zu schaffenden »Angebots« an hochqualifizierten, kompetenten Data Scientists.

Akteur*innen im Feld der Politik rahmen die Datenwissenschaften als »Basiswissenschaft« oder »Zukunftstechnologie« und schreiben ihnen ein fundamentales Transformationspotenzial hinsichtlich der Lösung globaler Menschheitsprobleme (wie Klimawandel, Migration, Pandemien etc.) zu. Neben Technologien wie künstlicher Intelligenz und Robotik bilden sie ein zentrales Element einer neuen Datenpolitik, die sich eng an bestehende Narrative des Standortwettbewerbs anlehnt. Die bildungs- und forschungspolitische Förderung des Wissensgebiets artikuliert sich insofern primär vor dem Hintergrund wirtschaftspolitischer Interessenlagen.

Im akademischen Feld schliesslich werden die Verschränkungen und Wechselwirkungen von Begriffs- und Grenzarbeit deutlich: Wissenschaftliche Akteur*innen entwerfen multiple Bedeutungen des Gegenstandes Datenwissenschaften, die von einer Einbindung in eine primär disziplinär strukturierte Landschaft bis hin zu einem neuen wissenschaftlichen Paradigma reichen. Dabei werden epistemologische und disziplinäre Grenzziehungen artikuliert, die den Gegenstand von anderen Wissensgebieten differenzieren sollen. Parallel dazu werden allerdings auch disziplinäre Grenzen in Forschungskollaborationen und -netzwerken als »Grenzorganisationen« transzendiert. Solche synchron ablaufenden Grenzpraktiken finden in der Hochschulbildung im Aufbau von Curricula oder Kompetenzprofilen eine organisationale Form, wodurch wiederum bestimmte Begriffe und Wissensbestände reproduziert werden. Repräsentationen und Imaginationen sind Produkte kollektiver Aushandlungen der involvierten Akteur*innen innerhalb der jeweiligen Felder. Ihre Effekte eröffnen und konturieren für die Akteur*innen jedoch Möglichkeitsräume, die über etablierte Feldgrenzen hinausgehen. Demgegenüber haben Grenzpraktiken, indem sie sich zentral um den Gegenstand eines jeweiligen Feldes drehen (was gehört dazu, was nicht? Welche Akteur*innen sind legitimiert zu sprechen, welche nicht? etc.), primär Effekte auf die etablierten Felder selbst.

In Zwischenräumen zirkulieren und koexistieren somit verschiedene Begriffe, die als *boundary objects* flexibel und adaptiv sind, um mit den jeweiligen »Sprachen« der verschiedenen Felder kompatibel zu sein. Grenzobjekte haben in diesem Sinne eine zwischenräumliche Funktion (Staley 2002), indem sie verschiedene Felder über Grenzen hinweg verknüpfen und so den Austausch von Informationen und Wissen ermöglichen. Wie das Beispiel übergreifender Forschungskollaborationen zeigt, werden auch Grenzorganisationen geschaffen, die Logiken verschiedener Felder integrieren und somit zur Dauerhaftigkeit zwischenräumlicher Phänomene beitragen.

Das analytische Modell versucht einen Brückenschlag zwischen der Feldtheorie Bourdieus, daran anschliessenden Erweiterungen sowie ausgewählten Theorieperspektiven der Science and Technology Studies. Es bietet den Vorteil, dass das widersprüchliche, produktive Potenzial von Grenzziehungen und Grenzüberschreitungen sowie die Dynamiken von Kooperation und Konkurrenz, die Räume zwischen Feldern charakterisieren, in eine räumliche Konzeption von Gesellschaft integriert werden

können. Anstatt solche hybriden Räume zu neuen Feldern zu erklären, die klaren Grenzen und Regeln unterliegen, was deren Spezifika in gewisser Weise neutralisiert, berücksichtigt das analytische Modell die Vermischungen und multiplen Logiken. Dadurch ist das Modell schliesslich in der Lage, die Genese und den Wandel sozialer Kollektivformationen adäquater zu erklären.

Kapitel 3 - »Data Science« als soziales Phänomen: Genese und multiple Perspektiven

In diesem Kapitel diskutiere ich wichtige Forschungsarbeiten, die das Phänomen »Data Science« aus unterschiedlichen Perspektiven untersuchen. Die Auswahl der existierenden Literatur schliesst dabei an das in der Einleitung dargelegte zentrale Erkenntnisinteresse, die Fragestellungen und die theoretisch-analytische Perspektive an. Die Diskussion der Studien dient dazu, den Blick auf den Untersuchungsgegenstand zu schärfen und in der weiteren Forschungslandschaft zu verorten. In einem ersten Teil widme ich mich Arbeiten, die durch disziplinäre und theoretische Perspektiven die Etablierung der Datenwissenschaften interpretieren (Kap. 3.1). Dabei skizziere ich die Genese und Etablierung des Begriffs »Data Science« an der Schnittstelle von Wissenschaft, Technologieindustrie sowie Forschungs- und Wissenschaftspolitik. Danach untersuche ich Positionsbezüge innerhalb der Wissenschaft, die den Gegenstand einerseits in einer disziplinären Differenzierungs-, andererseits in einer universalen Entdifferenzierungsperspektive betrachten. Schliesslich adressiere ich Arbeiten, die die Entstehung der Datenwissenschaften primär als Professionalisierung deuten.

Parallel zu den begriffshistorischen, disziplinären und sozialtheoretischen Erklärungs- und Deutungsversuchen beschäftigt sich ein weiteres Spektrum der Forschungsliteratur intensiv mit der Frage nach den relevanten Kompetenzen, Qualifikationen oder Tätigkeitsprofilen. Kompetenzen stellen – wie ich empirisch noch zeigen werde – einen zentralen Begriff des Untersuchungsfeldes dar. Einer konstruktivistischen Perspektive folgend verstehe ich unter Kompetenzen »sozial zugeschriebene Qualitäten, die sich über vielgestaltige Kommunikationen und Interaktionen manifestieren bzw. als sich manifestierend dem Subjekt attestiert werden« (Kurtz 2010: 8). Die Suche nach den »richtigen« Kompetenzen stellt ein zentrales Erkenntnisinteresse für Akteur*innen verschiedener Felder dar, die in die Etablierung der Datenwissenschaften involviert sind. Obwohl viele Arbeiten eher instrumentelle Ziele verfolgen, indem sie darauf ausgerichtet sind, Unternehmen und anderen Akteur*innen Wissen über die Praxis von Data Scientists zu vermitteln und somit deren Rekrutierung zu erleichtern, liefern solche Kompetenzkonstruktionen auch instruktive Rückschlüsse über die transversale Konstruktion der Datenwissenschaften als soziales Phänomen. Entsprechend sind sie von Interesse für die vorliegende Arbeit. Ich fasse deshalb im zweiten Teil zusammen, wie Akteur*innen durch Kompetenzkonstruktionen und Tätigkeitsprofile als feld- und organisationspezifische Perspektivierungen zur Ge-

nese der Datenwissenschaften im Arbeitsmarkt (Kap. 3.2), in der Bildungs- und Forschungspolitik (Kap. 3.3) sowie im akademischen Feld (Kap. 3.4) beitragen. Mit der Anordnung nehme ich eine Strukturierung vor, die ich in den beiden empirischen Teilen der Arbeit wieder aufgreifen werde. Abschliessend rekapituliere ich in einem Zwischenstand die wichtigsten Punkte des dargestellten Forschungsstandes (Kap. 3.5).

3.1 Begriffliche, disziplinäre und theoretische Perspektiven auf die Datenwissenschaften

3.1.1 »Data Science« als Label: Eine kurze Begriffsgeschichte

»The absence of clear boundaries defining data science, and the many people co-opting the term for their own, is a good thing for the burgeoning function. It creates more interest in data science, both to support organizational decision-making, as well as to attract more talent into the field of data science« (EMC 2011: 4).

Der Begriff »Data Science« ist keineswegs neu, sondern findet seit den 1960er-Jahren Verwendung in bestimmten akademischen Disziplinen (Chatfield et al. 2014: 3). Dem dänischen Informatiker Peter Naur wird die Verwendung der Begriffe »Data Science« bzw. »Datalogy« als Substitute für Computerwissenschaften zugeschrieben (Irizarry 2020; Donoho 2017: 763). Ebenfalls aus jener Zeit stammen einige hinsichtlich der Emergenz der Datenwissenschaften als Wissensfeld zentrale Texte: Der Artikel »The Future of Data Analysis« des Statistikers und Bell-Labs-Ingenieurs John Tukey (1962) sowie das spätere Lehrbuch »Exploratory Data Analysis« gelten in der Statistik und den Computerwissenschaften als wichtige Beiträge für jene methodisch-technische Expertise, die in den 1990er-Jahren als »Data Mining« Bedeutung erlangt (Dasu & Johnson 2003; Fisher 2001; Han et al. 2012).

Im Laufe der 1990er-Jahre erhält das Label »Data Science« breitere Akzeptanz an der Schnittstelle von Computerwissenschaften, Statistik und weiteren Disziplinen: Verschiedene Repräsentanten im wissenschaftlichen Feld äussern sich dazu (Hayashi 1998; Wu 1997) und erste Konferenzen unter dem Label finden statt (Hayashi et al. 1998; Ueno 2017). Parallel dazu werden wissenschaftliche Zeitschriften gegründet. Als eigentliches Manifest in der jüngeren Geschichte kann ein Artikel des Statistikers William S. Cleveland in der *International Statistical Review* 2001 gelten. Während Friedman (2001) in derselben Ausgabe dazu aufruft, das Feld der »Data Analysis« nicht den Computerwissenschaften alleine zu überlassen, legt Cleveland (2001: 21f.) mit direktem Bezug auf Tukey einen »Action Plan« für »Data Science« als eigenständige wissenschaftliche Disziplin inklusive eines Curriculums mit eigenen Schwerpunkten und Gewichtungen vor. Obwohl zu jener Zeit in den USA erste Studienprogramme unter dem Label lanciert werden, stiess der Weckruf Clevelands zumindest im Feld der Statistik lediglich auf geringen Nachhall (Kane 2014).

Auch aus wissenschaftspolitischer Sicht häufen sich in jenen Jahren Studien und Bezugnahmen auf das Label »Data Scientist« zur Beschreibung jener Professionen, die innerhalb wissenschaftlicher Forschungsgruppen Aufgaben des Managements, der Kuratierung und Aufbereitung von Daten besorgen (NSB 2005; Swan & Brown 2008; IWGDD 2009). Das US-amerikanische National Science Board (NSB) definierte »Data

Scientists« als jene »information and computer scientists, database and software and programmers, disciplinary experts, curators and expert annotators, librarians, archivists, and others, who are crucial to the successful management of a digital data collection« (NSB 2005: 27). Neben der Durchführung wissenschaftlicher Analysen werden die Beratung und Ausbildung in statistischen Analysemethoden, die Visualisierung und Exploration digitaler Daten sowie die Entwicklung neuer Datenbanktechnologien genannt. Die Definition des NSB zeugt von einem breiten, inklusiven Rollenverständnis, das disziplinäre und professionelle Grenzen überschreitet.

Eine breitere Anerkennung über das wissenschaftliche Feld hinaus setzt jedoch erst gegen Ende der 2000er-Jahre ein,¹ als Repräsentanten grosser Social-Media-Plattformen sich selbst als »Data Scientists« zu bezeichnen beginnen bzw. in Anspruch nehmen, die neue Berufsrolle überhaupt erst erschaffen zu haben:

»At Facebook, we felt that traditional titles such as Business Analyst, Statistician, Engineer, and Research Scientist didn't quite capture what we were after for our team. The workload for the role was diverse: on any given day, a team member could author a multistage processing pipeline in Python, design a hypothesis test, perform a regression analysis over data samples with R, design and implement an algorithm for some data-intensive product or service in Hadoop, or communicate the results of our analyses to other members of the organization in a clear and concise fashion. To capture the skill set required to perform this multitude of tasks, *we created the role of »Data Scientist«* (Hammerbacher 2009: 84; eigene Hervorhebung).

In der Folge wird die Bezeichnung medial weit verbreitet, was sich in unzähligen Selbst- und Fremdzuschreibungen äussert (Varian 2009; Conway 2010; Davenport & Patil 2012; Voulgaris 2014). Zur selben Zeit häufen sich Bedarfsanalysen von Beratungsunternehmen wie McKinsey und anderen ökonomischen Akteur*innen, die im Einklang mit dem »Hype« um »Big Data« (Carter & Sholler 2016; Donoho 2017; Elish & boyd 2018) eine stetig steigende Nachfrage nach Datenwissenschaftler*innen für die kommenden Jahre prognostizieren, womit das gegenwärtige Angebot der Universitäten nicht mithalten könne (Manyika et al. 2011). Um den wachsenden Bedarf decken zu können, entstanden in den Folgejahren Studiengänge in Datenwissenschaften an zahlreichen Universitäten und Hochschulen rund um den Globus. Trotz der Bemühungen von Wissenschaftler*innen und wissenschaftspolitischen Institutionen waren es Akteur*innen in der Industrie, die das Wissensgebiet über die Professionsbezeichnung »Data Scientist« popularisierten und personifizierten (Davenport & Patil 2012) und diesem somit zum Durchbruch verhelfen (Gehl 2015).

Ein Vergleich verschiedener Definitionen zeigt, dass die meisten Texte Elemente aus Statistik, Mathematik, Informatik und (Software) Engineering, bisweilen auch Ökonomie, Natur- und Sozialwissenschaften umfassen (Chatfield et al. 2014). Insbesondere Definitionen von industriellen Akteur*innen legen den Schwerpunkt auf die

1 Ein Vergleich der Begriffe »Big Data«, »Social Media« und »Data Science« im Web of Science (wissenschaftliche Artikel) sowie bei Google Trends (Google-Suchbegriffe) indiziert, dass die drei Begriffe in den Jahren 2006–2008 wissenschaftlich breit auftauchen und teilweise erst einige Jahre später in Suchanfragen auf Google Prominenz erlangen (»Social Media« 2007/8; »Big Data« 2011/12; »Data Science« ab 2014) (vgl. auch Kane 2014).

Verknüpfung der computerwissenschaftlichen Komponente mit der Lösung konkreter Business-Probleme (Provost & Fawcett 2013). Sowohl wissenschaftliche Definitionen² als auch professionelle Selbstbeschreibungen³ verweisen auf ein grundsätzlich weites Verständnis, unter dem verschiedene wissenschaftliche Disziplinen gefasst werden können. Nur wenige wissenschaftliche Disziplinen arbeiten nicht auf die eine oder andere Weise mit Daten oder computergestützten Auswertungsinstrumenten. Solche inklusiven Definitionen eröffneten vielfältige Möglichkeiten für verschiedene wissenschaftliche Disziplinen, sich fortan unter dem Label »Data Science« zu präsentieren. Als »umbrella term« (Giabbanelli & Mago 2016: 1970; Irizarry 2020) erlaubt es das Label anderen Akteur*innen und Disziplinen, sich affirmativ (oder auch distanzierend) darauf zu beziehen. »Umbrella terms« operieren als Mediatoren zwischen wissenschaftlicher Forschung, Wissenschaftspolitik und industriellen Anwendungsfeldern (Rip & Voss 2013): Durch einheitliche, inklusive Bezeichnungen werden entstehende Wissensfelder gefestigt und mit ökonomischen sowie wissenschaftspolitischen Ansprüchen verknüpft, was deren Identität sowohl nach innen wie nach aussen prägt. Etabliert sich dann ein Wissensfeld, werden die transversalen Verknüpfungen durch Forschungsfinanzierung und Infrastrukturen stabilisiert und erhalten einen dauerhaften Charakter (ebd.: 11). Eine solche konsensbefördernde »Politik von Buzzwords« (Bensaude-Vincent 2014) ist auch im Falle der Datenwissenschaften an der Schnittstelle von Technikwissenschaften, Industrie und Wissenschaftspolitik zu beobachten.

Eine begriffshistorische Perspektive auf den Gegenstand macht also deutlich, dass sowohl Akteur*innen in der Wissenschaft, in der Wissenschaftspolitik und der IT-Industrie in die Genese der Datenwissenschaften involviert sind. Die Durchsetzung des breiten, inklusiven Labels »Data Science« markiert dabei ein integrierendes Moment zwischen den unterschiedlichen Akteur*innen der genannten gesellschaftlichen Sphären.

3.1.2 Die Datenwissenschaften zwischen disziplinärer Differenzierung und Entdifferenzierung

Eine zweite Perspektive interpretiert das Phänomen – vor allem in der Wissenschaft selbst – in einer disziplinären Lesart. Zahlreich sind die Bezüge, die »Data Science« als neue Disziplin rahmen (Cao 2017; NASEM 2017; Song & Zhu 2017). Die disziplinäre Perspektive wird besonders virulent in der Auseinandersetzung über die Ursprünge, die verschiedene »Kerndisziplinen« der Datenwissenschaften für sich in Anspruch nehmen. Womöglich die grösste epistemologische, aber auch politische Herausforderung stellt das als »Data Science« bezeichnete Wissensfeld für die Statistik dar, indem dieses gewissermassen Anspruch auf eine traditionelle Domäne des Feldes, nämlich die Entwicklung von Methoden zur Analyse von (grossen) Datensätzen, erhebt (Grommé et al. 2018). In Reaktion darauf kam es zu einer Flut an Publikationen, die sich mit

2 »Data science is the discipline of drawing conclusions from data using computation« (Adhikari & DeNero 2016).

3 »»Data Science« means the scientific study of the creation, validation and transformation of data to create meaning«, vgl. Data Science Association (2020): Code of Conduct. Online: www.datascience-assn.org/code-of-conduct.html (Zugriff: 03.02.2022).

dem Verhältnis von Statistik und Datenwissenschaften beschäftigen.⁴ Einige Kommentator*innen argumentieren, dass »Data Science« lediglich eine Art Re-Branding von Statistik sei (Donoho 2017; Yu 2014; Wu 1997). So hielt der Statistiker Jeff Wu bereits 1997 einen Vortrag mit dem Titel »Statistics = Data Science?«, in dem er Statistik als eine Trilogie von Datenerhebung, Datenmodellierung und Analyse sowie *decision making* beschreibt. In seiner Schlussfolgerung fordert er – wie Cleveland (2001) – dazu auf, fortan die Begriffe »Data Science« anstelle von Statistik bzw. »Data Scientists« für Statistiker*innen zu verwenden (Wu 1997).

Ähnlich argumentiert David Donoho, wonach der »Hype« um Big Data, die erforderlichen Kompetenzen und die vielen neuen Jobs für Data Scientists die disziplinären Grundlagen der Statistik, auf denen die Datenwissenschaften aufbauten, verdecken würden (Donoho 2017: 745–49). Er bezieht sich dabei auf den Statistiker Leo Breiman (2001), der die Auffassung zweier unterschiedlicher Statistik-Kulturen (*Data Modeling Culture vs. Algorithmic Modeling Culture*) vertrat.⁵ Unter »Generative Modeling«, das der »Data Modeling Culture« entspreche, kann nach Breiman ein Grossteil der mathematischen Statistik, Ökonometrie und quantitativen Sozialforschung subsumiert werden. In dieser epistemischen Tradition wird für einen gegebenen Datensatz ein stochastisches Modell entwickelt, zu welchem die vorhandenen Daten »passen«. Auf dieser Grundlage werden dann Schlussfolgerungen über die zugrundeliegende Struktur des Modells gezogen (Inferenzstatistik). »Predictive Modeling«, das Breiman als Teil einer »Algorithmic Modeling Culture« sieht, umfasst hingegen Felder wie angewandte bzw. industrielle Statistik oder Computerwissenschaften und verschreibt sich primär der Entwicklung von Algorithmen. Das Modell testet die Vorhersagekraft (*prediction*) für einen gegebenen Datensatz, ohne Annahmen über dessen Entstehung zu machen (Donoho 2017: 751; Hofmann & VanderPlas 2017: 776).⁶

Trotz dieser divergierenden Positionsbezüge wird den Datenwissenschaften von vielen Autor*innen das Potenzial zugeschrieben, die unterschiedlichen Methoden, Datenverständnisse und Denkweisen zu integrieren und disziplinäre Grenzziehungen zu überschreiten.⁷ Konträr dazu wird in einer konflikttheoretischen Perspektive die Machtverschiebung über soziale und kulturelle Fragestellungen von den Geistes- und Sozialwissenschaften zu den Computerwissenschaften und Engineering beklagt (Roberge & Seyfert 2016: 10). Manche warnen gar vor einer Kolonisierung sozialwissenschaftlicher Fragestellungen durch ingenieur- bzw. computerwissenschaftliche Disziplinen (Dagiral & Parasio 2017; McFarland et al. 2016). Barocas und boyd (2017: 25)

4 Vgl. die Beiträge im *Journal of Computational and Graphical Statistics* 26(4), 2017, oder das Special Issue on Statistics and the Undergraduate Curriculum in *The American Statistician* 69, 2015.

5 Die Diagnose zweier »Statistik-Kulturen« weist darauf hin, dass es nicht nur um bestimmte Modelle, Messinstrumente oder Analysetools geht, sondern auch um divergierende Denktraditionen und Weltverständnisse im Sinne epistemischer Kulturen (Knorr-Cetina 2002).

6 DiMaggio (2015) und Manovich (2015: 22) entwickeln ähnliche Argumente zum Verhältnis von sozialwissenschaftlicher (Inferenz-)Statistik und Computer- bzw. Datenwissenschaften.

7 Auch sozialwissenschaftliche Autoren argumentieren, dass bestehende Grenzziehungen zwischen wissenschaftlichen Disziplinen und Denktraditionen aufgelöst (DiMaggio 2015: 1) und fruchtbare Kollaborationen zwischen diesen entstehen würden (Salganik 2017). Kauermann (2018: 88) bezeichnet Data Scientists als »Zwitter«, die beide von Breiman skizzierten statistischen Kulturen beherrschen und somit verknüpfen könnten.

verweisen auf die Unterschiedlichkeit von Computer- und Datenwissenschaften auf der einen Seite sowie Geistes- und Sozialwissenschaften auf der anderen Seite bezüglich der Diskussion von ethischen Fragestellungen und der gesellschaftlichen Implikationen, die sich durch zeitgenössische Verwendungsweisen und Methoden in »Data Science« oder »Machine Learning« ergeben würden (ähnlich Wallach 2018).

Gegenüber dieser Perspektive disziplinärer Differenzierung sind jene Beobachter*innen zahlreich, die in »Data Science« weniger die Entstehung einer neuen wissenschaftlichen Disziplin als vielmehr eine grundlegende Neuordnung des Verhältnisses von Wissenschaft, Technologieentwicklung und Ökonomie diagnostizieren: In Verknüpfung mit Big Data als eigentlicher »Revolution« (Mayer-Schoenberger & Cukier 2013; McAfee & Brynjolfsson 2012) entstehe damit eine fundamental neue Wahrnehmung von Gesellschaft, eine neue »Kultur« (Barlow 2013).⁸ Hey et al. (2009) erkennen in den veränderten Bedingungen wissenschaftlicher Erkenntnisproduktion mit Bezug auf den Informatiker und Turing-Preisträger Jim Gray ein »viertes Paradigma«: Es bezeichnet in Anlehnung an den Wissenschaftshistoriker Thomas S. Kuhn (1996) einen epochalen Wandel der Wissensproduktion, die auf Experimenten (experimentelle Wissenschaft), Modellen und Generalisierungen (theoretische Wissenschaft) oder Simulationen (computergestützte Wissenschaft) basiert, hin zu einer »explorativen Wissenschaft« (Kitchin 2014: 3), die sich massgeblich auf grosse Datenmengen, immense Rechenkapazitäten und algorithmische Verfahren zu deren Auswertung abstützt (Gray 2009: xviii; Kitchin 2014).⁹ Das Verständnis der Datenwissenschaften als neues »Forschungsparadigma« hat zudem Eingang in die Wissenschaftspolitik gefunden (ETH-Rat 2016a: 2).

Ebenfalls im Sinn dieser Entdifferenzierungsperspektive präsentiert Donoho (2017: 758f.) »Data Science« als eine Art »Meta-Wissenschaft«, die Sekundäranalysen für jegliche Formen wissenschaftlicher Daten und Forschungsprobleme erlaube. Gemäss Ribes (2019: 516; Ribes et al. 2019) positionieren sich die Datenwissenschaften als »Universalwissenschaft«, die nicht nur für die Wissenschaften selbst, sondern auch für die Wirtschaft, den Staat und andere soziale Felder fundamental sei. Daten spielten in jeder wissenschaftlichen Disziplin eine zentrale Rolle, mehr noch als komplexe Algorithmen und hohe Rechenkapazitäten. Durch diese Positionierung, die durch Akteur*innen in Industrie und Wissenschaftspolitik gestützt werde, würden die Datenwissenschaften gewissermassen von Beginn weg die Gefahr überwinden, zwischen formaler Mathematik und angewandtem Engineering auf den Status einer »Hilfswissenschaft« reduziert zu werden, wie es den Computerwissenschaften über Jahrzehnte hinweg drohte (Ribes et al. 2019: 296).

8 In ihrer »populärwissenschaftlichen Variante« taucht das Argument in Form globaler Heilsversprechungen und technikutopischer Rhetoriken auf: Jede denkbare Herausforderung (wie die Rettung der Menschheit vor dem Klimawandel, globaler Armut, Hunger etc.) gilt als technisch lösbar. Morozov (2013) spricht in diesem Zusammenhang von einer »solutionistischen« Perspektive, die nicht nur geographische, kulturelle und ökonomische Kontexte ausblendet, sondern auch politische und soziale Machtverhältnisse negiert.

9 Ein Paradigma besteht laut Kuhn (1996) in einer akzeptierten Art, die Welt zu untersuchen und das Wissen einer bestimmten Disziplin zu einem gegebenen Zeitpunkt zusammenzufassen. Periodisch erscheint eine neue Art zu denken, die bestehende Theorien und Ansätze herausfordert, wie beispielsweise die Darwin'sche Evolutionstheorie. Es kommt zum Paradigmenwechsel, wenn die dominante wissenschaftliche Vorgehensweise nicht mehr in der Lage ist, ein bestimmtes Phänomen oder Schlüsselfragen zu beantworten und zu erklären.

Die kurze Rekapitulation einiger wichtiger Positionen der disziplinären Differenzierungsperspektive einerseits und universaler Entdifferenzierung andererseits macht deutlich, dass unter Beobachter*innen kaum Konsens über die wissenschaftliche bzw. gesellschaftliche Bedeutung der Datenwissenschaften auszumachen ist. Dennoch können die intensiv geführten Debatten als Beleg dafür gelesen werden, dass sich Akteur*innen in unterschiedlichen Disziplinen und Feldern in vielfältiger Weise mit dem Gegenstand auseinandersetzen. »Data Science« als Wissensgebiet prägt und strukturiert demnach über seinen »Buzzword«-Charakter hinaus die Interessen und Aufmerksamkeiten heterogener Akteur*innen in multiplen sozialen Feldern.

3.1.3 Die Genese der Datenwissenschaften in einer Professionalisierungsperspektive

Ein dritter Strang der Literatur diskutiert die Genese der Datenwissenschaften in einer Professionalisierungsperspektive (Avnoon 2021; Carter & Sholler 2016; Demchenko et al. 2016; Dorschel & Brandt 2021; H. D. Harris et al. 2013; Walker 2015).¹⁰ Die Arbeiten referieren mehrheitlich auf das Professionsverständnis bei Abbott (1988), wonach Professionen ein interdependentes System bilden, das die Kontrolle über Wissen, Kompetenzen und Arbeitsinhalte ausübt. Kennzeichnend ist, dass Professionen generalistisches, abstraktes Wissen auf konkrete Anwendungsfälle applizieren (ebd.: 8).¹¹

Die Entwicklung der Datenwissenschaften in den letzten zwei Jahrzehnten entspricht durchaus jener Professionalisierung mit eigenen Konferenzen, Zeitschriften, Lehrbüchern, Fachgesellschaften oder Grundsätzen der Selbstregulierung (Walker 2015), wie sie auch bei anderen, etablierten Professionen zu beobachten ist. Brandt (2016) zeigt in seiner Dissertation auf, wie Data Scientists allgemeine Wissensbestände in Form von Methoden und Technologien auf konkrete – wissenschaftliche, ökonomische, politische – Fragestellungen anwenden, und stellt dies in Beziehung zu etablierten Professionen wie Jurisprudenz oder Systembiologie. Er schlägt den Begriff einer »Denkgemeinschaft« (*thought community*) (ebd.: 3) vor: Daten-Nerds würden formale Ideen mit informellen Interpretationen kombinieren. Daraus resultiere eine Form von Geheimwissen bzw. Expertise (*arcane knowledge*). Durch diese Form der Improvisation gelingt es ihnen, inhaltliche Probleme durch Darstellungen miteinander zu verknüpfen und so bürokratische Kontrolle in den jeweiligen Arbeitskontexten zu unterbinden. Dorschel und Brandt (2021) sprechen von »Professionalisierung mittels Ambiguität«: Sie identifizieren »eine soziale Logik der Ambiguität«, die die Konstruktion von Data Scientists in Wissenschaft und Wirtschaft strukturiere und auf einer »Grammatik aus Differenz und basalem Konsens« aufbaue (Dorschel & Brandt 2021: 21; Hervorhebung im Original). Die Rahmungen als »multipolare Akademikerinnen«

10 Gleichzeitig wird aber bisweilen auch bestritten, dass es sich bei Data Scientists überhaupt um eine eigenständige Profession handle, sondern im Rahmen des »Hypes« um Big Data das Label Science für bereits existierende Praktiken der Datenerhebung, -auswertung und -analyse in Feldern wie Business Intelligence angeeignet und Letztere dadurch aufgewertet worden seien (Watson 2014).

11 Professionen befinden sich in einer Art konstanten Auseinandersetzung mit anderen Professionen um jurisdiktische Monopole, d. h. die Legitimität der jeweiligen Expertise auf einem bestimmten Gebiet. Der Begriff der Jurisdiktion bezeichnet die Verknüpfung zwischen einer Profession und ihren Arbeitsinhalten (Abbott 1988: 20).

(Wissenschaft) bzw. »Schnittstellenprofession« (Wirtschaft) einerseits sowie als »Grenzgängerin«, »die sich von Nerds abgrenzt, aber für Weltverbesserung einsetzt« (ebd.), andererseits markieren Mehrdeutigkeit als strategisches Mittel (Eisenberg 1984; Leitch & Davenport 2007) in der Genese und Etablierung der Profession.

Grommé et al. (2018) konzipieren das Verhältnis zwischen traditioneller staatlicher Statistik und den neuen Datenwissenschaften als einen professionellen Konflikt um die Produktion legitimen Wissens über den Staat. Die Genese eines neuen Gegenstandes (Big Data) als epistemisches Objekt intensivierte die Auseinandersetzungen über die legitimen Formen der Expertise. Grommé et al. argumentieren feldanalytisch, dass Data Scientists eine neue Fraktion und somit Differenzierung des Statistikfeldes darstellen, die die bestehende Hierarchie herausfordern und zu einer Rekonzeption der Rollenverständnisse und epistemischen Praktiken nationalstaatlicher Statistikinstitute bzw. Statistiker*innen geführt haben.

Die vielfältigen Definitions- und Bestimmungsversuche in akademischen Publikationen genauso wie in Selbstbeschreibungen (Bowne-Anderson 2018; Strachnyi 2017; Voulgaris 2014) indizieren demnach ein professionelles Feld, dessen Grenzen offen sind, was Anschlussmöglichkeiten an verschiedene wissenschaftliche Disziplinen und Wissensgebiete ermöglicht. Zu den verbindenden Elementen, die Kommunikation über disziplinäre Grenzen hinweg ermöglichen, zählen neben gemeinsamen Karrierewegen und epistemischen Objekten vor allem methodische Herangehensweisen: Fragestellungen aus wissenschaftlichen, ökonomischen oder politischen Zusammenhängen werden als Datenprobleme gerahmt, die mit einem ähnlichen Set an Methoden, Instrumenten und Tools bearbeitet werden können (Schutt & O’Neil 2013: 44). Eine zeitgenössische Sozialform hierfür sind Diskussionen, wie sie an »Data Science« Hackdays oder Meetups¹² geführt werden, in denen Problemstellungen feld- und themenübergreifend vor allem in Bezug auf die verwendeten Methoden diskutiert werden. Brandt (2016) hat dies ethnographisch im Hinblick auf Rollenbeschreibungen von Data Scientists bei der Genese der Profession untersucht. Dahl (2020) untersucht die performative Aushandlung beruflicher Identitäten von Data Scientists auf digitalen Plattformen anhand von Podcasts. In Anlehnung an Goffmans Konzept der Selbstdarstellung identifiziert Dahl drei Typen eines Data-Scientist-Selbst (*Coach*, *Business* und *Academic*), die auf die interne Differenzierung der Berufsgruppe sowie unterschiedliche Betätigungsfelder hinweisen.

Neue Studienangebote und die stark steigende Nachfrage transformieren sowohl die Verfügbarkeit von Data Scientists als auch deren Ausbildungsniveau. Die Lancierung neuer Studiengänge löst dabei etablierte Rekrutierungswege ab: So verfügt die »zweite Generation« von Data Scientists häufiger über Masterabschlüsse in »Data Science«, während die »erste Generation« breiter zusammengesetzt war und Professionelle technischer, natur- und sozialwissenschaftlicher Disziplinen umfasste. Befragungen von Data Scientists zeigen, dass die zunehmende Verfügbarkeit von Absolvent*innen erste Effekte auf das Ausbildungsniveau und die Saläre der US-amerikanischen »Data Science«-Community zeitigt (Burtch 2016, 2018).

Somit stehen professionssoziologische Diagnosen eines »Upgradings« (Dröge 2019: 23) bzw. »Upskillings« datenwissenschaftlicher Praxis, wonach Data Scientists legiti-

12 Dabei handelt es sich um informelle Treffen zu bestimmten Themen mit starkem Networking-Charakter.

me Expertise für das Feld der Datenarbeit beanspruchen können, einem »Downgrading« (ebd.: 24) bzw. »Deskilling« entgegen: Gehl (2015) vermutet, dass der ökonomische »Wert« der »zweiten Generation« von Data Scientists rasch sinken werde: Sobald die Hochschulen genügend Absolvent*innen in hoher Zahl produzierten, seien Data Scientists nicht mehr rar, würden billiger und entsprechend bald auch nicht mehr so »sexy« (ebd.: 422). Strategisches Ziel insbesondere von Technologieunternehmen sei es, das Wissen von Data Scientists zu entpersonalisieren und in Applikationen oder Algorithmen einzubauen, damit es unabhängig von der jeweiligen Person zur Verfügung steht. Solche Anwendungen könnten dann auch von Nicht-Expert*innen benutzt werden und seien entsprechend günstiger (ebd.: 423).

Ob »Buzzword«, »Umbrella Term«, Disziplin, Feld, Paradigma oder Profession: So vielfältig die begrifflichen, disziplinären und theoretischen Perspektiven auf die Emergenz der Datenwissenschaften sind, so konträr wirken die resultierenden Interpretationen. Was aus diesen divergierenden Deutungsangeboten übrig bleibt, scheint mir die Erkenntnis zu sein, dass die Datenwissenschaften ein multidimensionales Phänomen darstellen, das synchron in verschiedenen sozialen Sphären in je eigenen Ausprägungen konzeptualisiert, verhandelt, praktiziert und implementiert wird. Dabei bleiben das Label »Data Science«, die damit assoziierten Wissensbestände, Praktiken, Tools und Methoden sowie die daraus resultierenden Auswirkungen auf die Wissensproduktion die zentralen Bezugspunkte der unterschiedlichen Deutungen. Insofern sind die Interpretationen weniger als konkurrierend zu lesen, sondern vielmehr als Ausprägungen jener strukturellen Eigenschaft des Gegenstandes, nämlich selbst vielstimmig und mehrdeutig zu sein.

3.2 Konstruktionen der Datenwissenschaften im Arbeitsmarkt

»Despite the recent sensational declaration of a data scientist as ›the sexiest job of the 21st century‹, however, there is a lack of published rigorous studies of what a data scientist is, and what job skills this hottest job title may require« (Chatfield et al. 2014: 1).

3.2.1 Qualifikations- und Kompetenzanforderungen in Stellenausschreibungen

Arbeiten, die die Entstehung der Datenwissenschaften im Arbeitsmarkt untersuchen, wenden im Wesentlichen zwei unterschiedliche Strategien an, um die relevanten Kompetenzen, Qualifikationen oder Tätigkeitsprofile von Data Scientists zu eruieren. Eine erste Strategie in der Forschungsliteratur besteht darin, die neuen Qualifikations- und Kompetenzanforderungen von Data Scientists durch Analysen von Stellenausschreibungen empirisch herzuleiten (Dadzie et al. 2018; Debortoli et al. 2014; Djumaliev et al. 2018; Gardiner et al. 2018; Wowczko 2015). Im Fokus sind entstehende sowie sich verändernde Berufsfelder, vor allem im Bereich der Informations- und Kommunikationstechnologien sowie der digitalen Transformation von Wissensarbeit (dazu Boes et al. 2018; Dröge & Glauser 2019):¹³ Bildungs- und Qualifikationsanforderungen

13 Die veränderten Anforderungen im schweizerischen Arbeitsmarkt (Sacchi et al. 2005; Salvisberg 2010) und spezifisch in Bezug auf die Digitalisierung (SECO 2017a; Sheldon 2020) sind seit längerem Gegenstand der sozialwissenschaftlichen Arbeitsmarktforschung.

sowie die nachgefragten Tools in diesen Berufsfeldern entwickeln sich derart dynamisch, dass sie kaum mehr zur Festlegung von Tätigkeitsprofilen verwendet werden können. Berufsbezeichnungen sind in diesem Sinne »Container« (Wowczko 2015: 39) für spezifische Kompetenzprofile, die sich kontinuierlich verändern. Die Analyse von digital verfügbaren Stelleninseraten bietet insofern eine einfache, gegenüber anderen Erhebungsmethoden (wie Befragungen) kostengünstige Möglichkeit, um die kurzfristige, stark volatile Nachfrage nach bestimmten Kompetenzen innerhalb sich transformierender Berufsfelder zu eruieren (Boselli et al. 2018). Im Folgenden fasse ich einige wichtige Ergebnisse dieser Forschungsliteratur im Hinblick darauf zusammen, wie Stellenanzeigen und Kompetenzdefinitionen die Datenwissenschaften beschreiben und konstruieren.

Chatfield et al. definieren sechs Kompetenzbereiche von Data Scientists als Schnittmenge von zwei Dutzend akademischen und industriellen Definitionen: »(1) Entrepreneurship and business domain knowledge, (2) Computer scientist, (3) Effective Communication skills, (4) Create valuable and actionable insights, (5) Inquisitive and curious, and (6) Statistics and modeling« (Chatfield et al. 2014: 7). Die Verknüpfung computerwissenschaftlicher und statistischer Fertigkeiten mit unternehmerischem Wissen sowie den dazugehörigen Softskills fördert sehr ähnliche Ergebnisse zutage wie die Analyse der Kompetenzanforderungen der untersuchten Curricula. Schumann et al. (2016) konstruieren Data Scientists als »Allrounder«, die vielfältige Fachkompetenzen, soziale Fähigkeiten und Selbstkompetenzen wie beispielsweise Kreativität besitzen. Je nach Anforderungen bezüglich Tätigkeit bzw. der Rolle innerhalb der stellenausschreibenden Organisation empfehlen sie die Bildung von unterschiedlichen »Kompetenzprofile[n] (Data-Scientist-Typen)«, wobei keine der drei Kompetenzdimensionen (Fach-, Sozial-, Selbstkompetenz) vernachlässigt werden soll (ebd.: 465).

Dadzie et al. (2018) eruieren durch die Analyse zehntausender Stellenanzeigen aus diversen europäischen Ländern Datenbanken, Statistik, Tools (vor allem NoSQL, Hadoop) sowie Erfahrungen bezüglich Datenvisualisierung (vor allem Tableau, D3.js) als die am häufigsten genannten Kompetenzanforderungen. Debortoli et al. (2014) zeigen, dass bei Inseraten mit dem Suchbegriff »Big Data« zwischen technischen und Business-Kompetenzen differenziert werden kann, wobei erstere in rund 70 % aller Anzeigen nachgefragt werden. Die am häufigsten genannten technischen Kompetenzen sind »NoSQL Databases«, »Software Engineering« und »Programming«, während »Machine Learning« und »Quantitative Analysis« weniger bedeutend sind (Debortoli et al. 2014). Gardiner et al. (2018) identifizieren in der Analyse von Stellenanzeigen mit »Big Data« im Titel ebenfalls starke technische Orientierungen, insbesondere im Management von Daten sowie bei analytischen Informationssystemen, und folgern daraus, dass die Anforderungen an Fertigkeiten und Wissen von Datenwissenschaftler*innen hauptsächlich aus der Informatik stammen (ebd.: 9). Eine Industriestudie von IBM und Burning Glass Technologies erschliesst durch Data Mining von 130 Millionen Stellenanzeigen 300 analytische Kompetenzen für datenbezogene Professionen, die anschliessend aufgrund von Ähnlichkeiten in den Kompetenzprofilen in sechs Gruppen eingeteilt werden (Markow et al. 2017). Für jeden Beruf wird sodann ein analytischer Wert berechnet, wobei Data Scientists die höchsten Werte erzielen würden:

»Data Scientists are the most analytical roles in the market. They require proficiency with a large range of specialized analytical skills and tools, such as Machine Learning, Apache Hadoop, and Data Mining, in addition to generalized [Data Science and Analytics] skills like SQL, R, and Data Analysis« (Markow et al. 2017: 9).

Das Beispiel zeigt, dass viele der Kompetenzen, über welche Datenwissenschaftler*innen verfügen sollten, trotz solcher Kategorisierungen unscharf sind: So bezeichnet »Machine Learning« ein Kompendium statistischer Modelle und Algorithmen, mit deren Hilfe Computersysteme Aufgaben ohne spezifische Instruktionen ausführen können, sondern sich auf die Analyse von Mustern und Inferenzen aus bestehenden Daten stützen. In der thematischen Anwendung und Komplexität reichen diese von linearen Regressionsmodellen bis zu neuronalen Netzwerken, die divergierende Anforderungen an mathematische Wissensbestände, technisch-methodische sowie analytisch-interpretative Fähigkeiten stellen. »Machine Learning« kann folglich in verschiedenen Anwendungskontexten Unterschiedliches bedeuten (Engemann & Sudmann 2017). Der Umstand, dass notwendige methodisch-technische Kenntnisse oft summarisch (beispielsweise »Erfahrung in Machine Learning«, »Statistische Analysen in R«) repräsentiert werden (Ismail & Abidin 2016), indiziert demnach eine disziplinen- bzw. feldübergreifende Offenheit, die eine Vielzahl von Praktiker*innen ansprechen und potenziell inkludieren soll.¹⁴ Solche Listen mehrdeutiger Begriffe stellen insofern strategische Positionierungen der rekrutierenden Organisationen dar.

Data Scientists werden in vielen der zitierten Studien als multidisziplinär orientierte Praktiker*innen imaginiert, die als sogenannte »data unicorns« (Hermida & Young 2017) über eine ganze Reihe verschiedener Kompetenzen verfügen sollen. In einer kultursoziologischen Perspektive werden Data Scientists sodann als »Allesfresser« (*omnivores*) bezeichnet, die ihre professionelle Identität durch einen generalistischen, inklusiven Zugang zum Kompetenzerwerb konstruieren und aufrechterhalten würden (Avnoon 2021). Die grundlegende Distinktion gegenüber anderen Berufsgruppen erfolge gerade nicht durch Spezialisierung, sondern durch eine generalistische Haltung, die Expertise für unterschiedlichste Gegenstände beansprucht (ebd.: 345).

Da solche vielseitig ausgebildeten Generalist*innen jedoch auf dem Arbeitsmarkt nicht verfügbar waren, fand allmählich ein Übergang zur Wahrnehmung als Teamplayer statt und rückte die Zusammensetzung von »Data Science Teams« in den Vordergrund (Davenport 2020; H. D. Harris et al. 2013; Kim et al. 2018; Patil 2011; Schumann et al. 2016). Diese werden zudem als Lösung für den andauernden »Fachkräftemangel« in den Datenwissenschaften präsentiert (J. G. Harris et al. 2013). Aufgrund der anhaltenden Unsicherheit bezüglich der verfügbaren Kompetenzprofile streben vor allem Akteur*innen in der Geschäftswelt formale Klassifikationsschemata sowie Zertifizierungen der nachgewiesenen Kompetenzen von Data Scientists an (Davenport 2020), um diese gegenüber anderen Praktiker*innen, die ebenfalls Daten verarbeiten, abgrenzen zu können.

Unabhängig von der Frage, wie sich ein Profil von Data Scientists zusammensetzt und welche Kompetenzen dazu vonnöten sind, zeigt die Rekapitulation dieses ersten

14 Umgekehrt vermuten manche Autor*innen eine »Aufpolierung« der Stellenanzeigen durch Unternehmen, weil konkrete Definitionen für das entstehende Berufsfeld noch weitgehend fehlen würden (Debortoli et al. 2014).

Teils von Arbeitsmarktstudien zu Datenwissenschaften, dass unterschiedliche Kombinationen von Skills charakteristisch sind. Qualifikations- und Kompetenzanforderungen werden durch offene Aufzählungen von zugeschriebenen Fähigkeiten, Bildungstiteln, Methoden oder Tools formuliert. In dieser Offenheit manifestieren sich einerseits Unsicherheiten über die zentralen Begrifflichkeiten des noch jungen Feldes. Auf der anderen Seite kennzeichnen die Listen bestimmte Kategorien, die sich bereits in einer frühen Phase etabliert haben. Die empirische Analyse der Stellenanzeigen wird zeigen, inwiefern und welche Begriffe sich als flexibel genug erweisen, um die multiplen Bedeutungen und Anforderungen an datenwissenschaftliche Praktiker*innen zu vereinen und zu repräsentieren.

3.2.2 Arbeitspraktiken von Data Scientists

Neben der Analyse von Stellenanzeigen besteht eine zweite Strategie der Forschungsliteratur darin, praktizierende Data Scientists und verwandte Professionen nach den erforderlichen Kompetenzen, Ausbildungs- und Karrierewegen sowie den epistemischen Praktiken im Arbeitsalltag zu befragen (EMC 2011; Feldman et al. 2017; H. D. Harris et al. 2013; Ismail & Abidin 2016; Kandel et al. 2012; Kim et al. 2018; Roberts & Roberts 2013; Swan & Brown 2008). Aus den Erkenntnissen werden oft Forderungen an die Strukturen und Inhalte von Ausbildungsprogrammen abgeleitet.

Eine Studie des Hard- und Softwareherstellers EMC (2011) untersucht Datenpraktiken, Tools und Ausbildungswege der »data science community« sowie die Frage, wie die Organisationen mit datengetriebener Problemlösung umgehen. Die Analyse von Surveydaten von knapp 500 Data Scientists und Business-Intelligence (BI)-Professionellen zeigt, dass sich Data Scientists von den BI-Professionellen insbesondere dadurch unterscheiden, dass sie öfter Programmierkenntnisse aufweisen und in der Lage sind, Experimente mit grossen Datenmengen durchzuführen.¹⁵ Bezüglich der Ausbildung verfügen sie häufiger über Master- und PhD-Abschlüsse als die Vergleichsgruppe, bei der der Bachelor der häufigste Abschluss ist. Als zentral für die Praxis von Data Scientists werden zudem Interdisziplinarität und Arbeitsteilung mit anderen Professionellen in Unternehmen betrachtet (ebd.: 4).

Kandel et al. (2012) untersuchen, wie die Arbeitssituation und die verfügbaren Tools die Praktiken dreier Gruppen von Analyst*innen (»Business Analysts«, »Data Analysts« und »Data Scientists«) beeinflussen. Mittels qualitativer Interviews mit Praktiker*innen aus verschiedenen ökonomischen Feldern identifizieren sie drei »Archetypen«: »Hackers« sind erfahrene Programmierer*innen, verwenden die meisten Tools und operieren meist unabhängig von der IT. Sie führen die Schritte der Datenerhebung und -analyse selbstständig durch, erschliessen Datenquellen ausserhalb der Organisation und verknüpfen sie mit internen Daten. »Scripters« sind erfahren in Programmier- und Statistiksoftware wie R, erheben aber die Daten nicht selbst. Die Gruppe verwendet viel Arbeitszeit für ihre Modelle, die sie meist innerhalb einer Programmierumgebung erarbeitet. Die »Application Users« schliesslich verwenden hauptsächlich Excel-Tabellen und Software wie SAS oder SPSS. Sie benötigen Hilfe

15 Zu ähnlichen Schlussfolgerungen gelangen Debortoli et al. (2014: 325), die vergleichend Stellenanzeigen von BI- und Big-Data-Professionellen analysieren.

bei der Datenaufbereitung, arbeiten meist mit kleineren Datenmengen und selten ausserhalb der vertrauten Software (Kandel et al. 2012: 2918f.).

Roberts und Roberts (2013) unterscheiden anhand von Surveydaten (N = 300) vier Gruppen von Analytics- und Data-Science-Professionellen, die sie unter anderem aufgrund der aufgewendeten Arbeitszeit für verschiedene Analysetätigkeiten sowie zugeschriebenen Fähigkeiten zuordnen. Daraus leiten sie detaillierte Beschreibungen ab, wie die vier Gruppen von Analytics- und Data-Science-Professionellen rekrutiert, gemanagt und in Organisationen gehalten werden können (ebd.: 8ff.).

Harlan D. Harris et al. (2013) untersuchen, weshalb für unterschiedliche berufliche Tätigkeiten die gemeinsame Bezeichnung »Data Scientists« verwendet wird. Sie argumentieren, dass professionelle Bezeichnungen aus Statistik, Software- und Datenbank-Engineering sowie Natur- und Sozialwissenschaften wie im »Fleischwolf« zusammengeführt und durch relativ unklare Profile, Fähigkeiten und Kompetenzen abgelöst worden sind (ebd.: 3). Dies habe einerseits zusammen mit dem »Hype« um »Data Science« zu nicht einlösbaren Erwartungen geführt, andererseits zu einem unnötigen Aufwand bei der Rekrutierung von Mitarbeitenden.

Die Crowdsourcing-Plattform CrowdFlower (2015) (heute Figure Eight Inc.) befragte 153 Data Scientists, um deren Einbettung in Organisationen sowie den Beitrag zu deren Operationsweise zu verstehen. Interessant sind die Ergebnisse der zeitlichen Aufwendungen für die unterschiedlichen Tätigkeiten: Die zeitintensivste Tätigkeit besteht in der Aufbereitung und Bereinigung von sogenannten »messy data«, d. h. das Vorhandensein von Daten in schmutzigem und unordentlichem Zustand (Mützel et al. 2018: 115). Daraus resultieren sehr eingeschränkte zeitliche Kapazitäten für die beliebtesten Tätigkeiten, nämlich die »predictive analysis« sowie das »mining data for patterns« (CrowdFlower 2015: 7f.).

Die Feststellung, dass die ungeliebte Tätigkeit der Datenerhebung und -aufbereitung rund 80 % der Arbeitszeit einnimmt, wird in vielen Selbstbeschreibungen der *Community* geteilt (Schutt & O'Neil 2013; Wickham 2014; Wilson 2017). Wie Mützel et al. (2018) zeigen, bleiben solche Praktiken in der Diskussion oftmals auf der »Hinterbühne« verborgen, obwohl die darin implizierten Entscheidungen (der Datenerhebung, -bereinigung etc.) spätere Analyseergebnisse und somit organisationale Entscheidungen vorstrukturieren.

Carter und Sholler (2016) diagnostizieren anhand von semistrukturierten Interviews mit achtzehn Datenanalyst*innen in unterschiedlichen ökonomischen Feldern eine Diskrepanz zwischen der medial konstruierten Wahrnehmung von »Data Science« und den Alltagspraktiken von Data Scientists »on the ground«. Dies betrifft sowohl die Rolle von Theorie und traditionellen statistischen Methoden, die Bedeutung von Objektivität sowie den Zugang zu – öffentlichen oder proprietären – Datenquellen. Dabei zeigt sich, dass die Datenanalyst*innen ein reflexiveres und differenzierteres Verständnis der Herausforderungen von Datenpraktiken in zeitgenössischen Organisationen entwickeln, als dies die medialen Verlockungen und Verheissungen nahelegen (ebd.: 2317). Obwohl sie meist aus der Wissenschaft in industrielle Felder gewechselt sind mit der Motivation, rasch und unbürokratisch Zugang zu grossen Datenbeständen zu erhalten und diese bearbeiten zu können, werden neue Limitationen (insbesondere im Umgang und der Kommunikation mit Kund*innen) deutlich, die die Handlungsspielräume der Interviewten in verschiedener Hinsicht einschränken.

Feldman et al. (2017) untersuchen am Beispiel von Data Analytics die Personalrekrutierung auf Freelancing-Plattformen. Durch qualitative Interviews eruieren die Autor*innen, welche Tätigkeiten festangestellte Data Scientists an Freelancer*innen auslagern (würden) und welche Kompetenzen dazu notwendig sind.¹⁶ Neben Programmierfähigkeiten sowie statistischen und mathematischen Skills erachten die Interviewten vor allem Fachexpertise (*domain knowledge*), Interdisziplinarität und Kommunikation als wichtig, da mit unterschiedlichen Professionen und Funktionen kollaboriert wird. Auch die Freelancer*innen werden vor allem für jene Tätigkeiten eingesetzt und hoch bewertet, die supplementär zu den Aufgaben der Data Scientists selbst sind, nämlich Dokumentation, Visualisierung oder Datenbank-Engineering (ebd.: 13f.).

Viele existierende Studien fokussieren primär Arbeitspraktiken und Kompetenzerfordernisse in industriellen Anwendungsfeldern der Datenwissenschaften. Allerdings setzte sich eine frühe empirische Arbeit, die im Auftrag des britischen *Joint Information Systems Committee* entstand (Swan & Brown 2008), mit den Berufsrollen und der Karriereentwicklung von Datenprofessionen in der Wissenschaft auseinander. Auch sollte der Bedarf entsprechender Kompetenzen und Fähigkeiten für die Forschungsgemeinschaft evaluiert werden. Mittels Interviews und Fokusgruppen mit verschiedenen Berufsgruppen, die im wissenschaftlichen Feld mit grossen Datensätzen arbeiten, entwickeln die Autorinnen verschiedene Professionsbezeichnungen und unterscheiden zwischen »data creators or data authors«, »data scientists«, »data managers« sowie »data librarians« (ebd.: 8). Bei den Data Scientists sind oft keine klaren Karrierewege erkennbar; zudem ziehen sie informelle berufliche Weiterbildung gegenüber formellen Ausbildungsprogrammen vor. Die Interviewten sind auch der Ansicht, dass »computational skills« alleine nicht ausreichen, um in interdisziplinär zusammengesetzten Forschungsgruppen mitarbeiten zu können; vielmehr sei dazu auch eine spezifische Fachexpertise (*domain knowledge*), z. B. in den Naturwissenschaften, notwendig (ebd.: 15).

Ein Jahrzehnt später untersuchen Geiger et al. (2018) Karrierewege und Zukunftsaussichten von Data Scientists an renommierten US-amerikanischen Forschungsuniversitäten. Sie zeigen, dass sich die Karrierewege und Zukunftsvorstellungen von akademischen Data Scientists teilweise stark von den angestammten disziplinären Werdegängen unterscheiden, indem etwa gewisse Praktiken (wie der Unterhalt von Forschungsinfrastruktur oder das Schreiben statistischer Pakete) hinzugekommen sind, die zwar von den Forschenden geschätzt werden, die allerdings von Berufungsgremien kaum honoriert würden (ebd.: 27ff.).¹⁷

Die Ergebnisse des zweiten Teils der diskutierten Forschungsliteratur zum Arbeitsmarkt der Datenwissenschaften deuten intensive Suchprozesse und Bestrebungen nach einer Taxonomie im Bereich der Datenprofessionen an, die Aufschluss, Überblick und Eingrenzung bieten soll über das breite Feld datenbezogener Praktiken, Berufsrollen und dazugehöriger Kompetenzen in zeitgenössischen Organisationen. Die

16 Bei den Freelancer*innen handelt sich oftmals um junge, hochqualifizierte Studierende mit noch wenig Berufserfahrung.

17 Metzler et al. (2016) legen eine ähnliche Analyse zu den Arbeitspraktiken und Herausforderungen von Sozialwissenschaftler*innen (Computational Social Scientists) vor, die mit grossen Datenmengen arbeiten.

Vielfalt der identifizierten Bezeichnungen indiziert, dass trotz des neuen Labels »Data Science« bzw. der entstehenden Profession der »Data Scientists« kaum Klarheit, geschweige denn Konsens darüber besteht, wie sich eine solche Gruppe zusammensetzt, was sie charakterisiert und was sie von anderen abgrenzt. Die empirische Analyse von Stellenanzeigen und Curricula wird zeigen, ob und inwiefern solche Taxonomien die Such- und Rekrutierungsstrategien von Organisationen im schweizerischen Arbeitsmarkt bzw. in der Hochschulbildung prägen.

3.3 Konstruktionen der Datenwissenschaften in der Bildungs- und Forschungspolitik

Bereits in den 2000er-Jahren beschäftigten sich forschungs- und wissenschaftspolitische Akteur*innen vermehrt mit der Transformation der Wissensproduktion infolge der exponentiellen Zunahme digital verfügbarer Daten (NSB 2005; IWGDD 2009; Swan & Brown 2008). Kennzeichnend für die Studien ist ein tendenziell »pessimistischer« Grundton (Swan & Brown 2008: 18f.; IWGDD 2009: 18; Brown 2009), da Probleme bei der Anerkennung, mangelnde Ausbildungsmöglichkeiten sowie Karriereoptionen für Data Scientists innerhalb von wissenschaftlichen Forschungsgruppen diagnostiziert werden. Diese eher negativen Einschätzungen weichen innert weniger Jahre fast komplett einer (sehr) positiven Beurteilung: Fortan dominiert die Zuschreibung eines transformativen Potenzials und die Betonung der grossen Nachfrage nach ausgebildeten Fachkräften die Bewertung der Datenwissenschaften in der Wissenschafts- und Forschungspolitik (National Research Council 2014; ETH-Rat 2014; Geetoor et al. 2016; Horvitz & Mitchell 2010; Labrinidis & Jagadish 2012). Damit einher geht nicht nur eine Perspektiven-, sondern auch eine Deutungsverschiebung, indem sich die dominanten Positionen im Diskurs zu »Data Science« vom wissenschaftlichen und wissenschaftspolitischen Feld hin zum technologischen und ökonomischen Feld – und dabei insbesondere den grossen Internetfirmen – verschieben.

Fortan richtet sich die wissenschaftspolitische Aufmerksamkeit auf die zukünftigen Potenziale der Datenwissenschaften zur Lösung zentraler wissenschaftlicher, politischer oder ökonomischer Herausforderungen. Zugleich bleibt ein negativer Gegenhorizont in den Publikationen bestehen, indem sie unter Verweis auf einen McKinsey-Bericht von 2011 einen gravierenden »Fachkräftemangel« diagnostizieren (Manyika et al. 2011) – alleine in den USA würden bis ins Jahr 2018 zwischen 140'000 und 190'000 Data Scientists fehlen.¹⁸ Das fehlende Angebot an datenwissenschaftlichen Praktiker*innen würde demnach die Erkenntnisgewinne und Handlungsmöglichkeiten in verschiedenen sozialen Feldern gefährden. Diese Rahmung der Datenwissenschaften erwies sich als äusserst einflussreich und schrieb sich sowohl in die Aktivitäten ökonomischer Akteur*innen als auch in bildungs- und forschungspolitische Massnahmen ein (Saner 2019).

In der Folge widmeten sich zahlreiche Förderprogramme als auch universitäre Initiativen der Etablierung, Vernetzung und Priorisierung von disziplinen- und hoch-

18 Vorhersagen für die EU gingen sogar von 500'000 fehlenden Data Scientists aus, vgl. Offerman, Adrian (2016): »500'000 data scientists needed in European open research data«. Online: <https://joinup.ec.europa.eu/collection/open-government/news/500000-data-scientists-need> (Zugriff: 03.02.2022).

schulübergreifenden Gefässen der Datenwissenschaften in Forschung und Lehre. Die US-amerikanische *National Science Foundation* (NSF) lancierte verschiedene Instrumente zur Förderung der Datenwissenschaften: Dazu zählen unter anderem geographische Innovationssysteme, sogenannte Big Data Hubs (NSF 2015), oder das Schwerpunktprogramm »Harnessing the Data Revolution« (NSF 2019), das als langfristiger, strategischer Forschungsbereich an der Schnittstelle von Wissenschaft und Industrie etabliert wurde.¹⁹ In der Schweiz wird unter Verweis auf das Primat »liberaler Innovationspolitik« (Merz & Sormani 2016: 9) in der Regel keine Förderung bestimmter Technologien vorgenommen. Im Sinne einer kompetitiven Förderpolitik engagierten sich trotzdem eine Reihe wissenschaftspolitischer Akteur*innen in der Förderung der Datenwissenschaften: Der Schweizerische Nationalfonds (SNF) schrieb 2015 das Nationale Forschungsprogramm (NFP) Big Data (SNF 2015) aus. Parallel dazu lancierte der bundesstaatlich geförderte ETH-Bereich im Jahr 2016 die »Initiative for Data Science in Switzerland« (ETH-Rat 2016a), wozu die Gründung des Swiss Data Science Centers (SDSC) sowie neue Masterstudiengänge in Data Science an den ETH Lausanne und Zürich gehören (vgl. ausführlich zu den Massnahmen Kap. 6.4.2). Die bildungs- und forschungspolitischen Massnahmen erläutert und konkretisiert der Bericht »Herausforderungen der Digitalisierung für Bildung und Forschung in der Schweiz« (SBFI 2017). Die Datenwissenschaften werden darin als eine neue »Basiswissenschaft« präsentiert, »auf welcher andere Wissenschaften und Anwendungen aufbauen können« (ebd.: 70). Die bildungs- und forschungspolitischen Akteur*innen erhoffen sich zudem nicht nur technologisch-wissenschaftliche Innovationen, sondern auch ein gänzlich neues Verhältnis zwischen Hochschulen, Forschung und Industrie.

Des Weiteren beteiligen sich auch zahlreiche nichtstaatliche Akteur*innen an der Forschungsförderung im Bereich Datenwissenschaften: Die Gordon und Betty Moore Foundation sowie die Alfred P. Sloan Foundation schufen mit den »Moore-Sloan Data Science Environments« eine gemeinsame disziplinen- und organisationsübergreifende Struktur zur Förderung von Forschung und Ausbildung in Datenwissenschaften an drei renommierten US-amerikanischen Forschungsuniversitäten (University of California Berkeley, New York University und University of Washington-Seattle) (Moore-Sloan Data Science Environments 2018). Damit förderten die beiden Stiftungen nicht nur Forschung, Lehre und Beratung in datenintensiver Wissenschaft an den beteiligten Institutionen, sondern schufen eigentlich Rollenmodelle, die durch Veröffentlichung und Diffusion der erarbeiteten Inhalte und Prozesse an zahlreichen anderen Universitäten und Hochschulen aufgenommen und implementiert wurden (Katz 2019; Geiger et al. 2019). Ein weiteres Beispiel bildet die feldübergreifende Initiative »Social Science One« (King & Persily 2020): Damit lancierten Akteur*innen im akademischen Feld in Kollaboration mit Facebook, dem *Social Science Research Council* sowie mehreren gemeinnützigen Stiftungen ein neues Förder- und Begutachtungsmodell, das Daten von Facebook-Nutzer*innen unter Berücksichtigung von Datenschutzrichtlinien für sozialwissenschaftliche Forschung zugänglich macht. Die geförderten Projekte werden durch Stiftungsgelder finanziert, um die Unabhängigkeit gegenüber den Unternehmensinteressen von Facebook zu wahren.

Ein zweites Aktionsfeld ist die Förderung und Weiterentwicklung von Lehrprogrammen und Studiengängen im Bereich Datenwissenschaften. Die US-amerikani-

19 Für eine Übersicht zu staatlichen Förderaktivitäten in den Datenwissenschaften vgl. SBFI 2017: 94ff.

schen *National Academies of Sciences, Engineering and Medicine* (NASEM) beauftragten mehrere Arbeitsgruppen, die sich mit der Implementierung der Datenwissenschaften auf den unterschiedlichen Stufen der Hochschulbildung beschäftigten (NASEM 2017, 2018). Parallel dazu führten mehrere Divisionen der NASEM gemeinsam eine Serie von Roundtables durch (Kloefkorn et al. 2020), die Vertreter*innen von Universitäten, Förderinstitutionen, Stiftungen und der Industrie zusammenbrachten, um Ausbildungs- und Praxismodelle sowie die Bedürfnisse der »Data Science«-Community und von Unternehmen als Arbeitgeber*innen zu diskutieren (NASEM 2020).

Auch in der europäischen Forschungspolitik wurden zahlreiche Initiativen lanciert, um Studiengänge in Datenwissenschaften zu fördern und die Eigenschaften des entstehenden Feldes schärfer zu konturieren. Das im Rahmen von Horizon 2020 geförderte Projekt EDISON²⁰ setzt sich zum Ziel, durch eine koordinierte Zusammenarbeit von Forschungspolitik, Wissenschaft und Industrie die Datenwissenschaften als Profession zu etablieren, die Bedürfnisse des Arbeitsmarktes mit den verfügbaren Skills in Übereinstimmung zu bringen und die Curricula an Hochschulen anzupassen. Damit sollen die Zahl ausgebildeter Data Scientists in Europa signifikant erhöht und deren Kompetenzen sowie die Ausbildungsqualität verbessert werden (Demchenko et al. 2016). Ähnliche Ziele verfolgt die ebenfalls durch Horizon 2020 geförderte *European Data Science Academy* (Mikroyannidis et al. 2018).

Sowohl die wissenschaftspolitischen Fördermassnahmen als auch die Vorbildfunktion etablierter Forschungs- und Ausbildungszentren unterstützten die Diffusion der Datenwissenschaften als Wissensfeld und führten zur Etablierung unzähliger Studienprogramme an Universitäten und Hochschulen weltweit (Anderson et al. 2014; Asamoah et al. 2015; Buckingham Shum et al. 2013; Giabbanelli & Mago 2016). Dies trug einerseits zur Professionalisierung bestehender Methodenausbildungen hinsichtlich der gelehrten Kompetenzen bei und andererseits zu einer klareren Grenzziehung gegenüber Nachbardisziplinen wie Computerwissenschaften, Software Engineering, Mathematik oder Statistik.

Der Literaturüberblick macht deutlich, dass Akteur*innen der Bildungs- und Forschungspolitik die Diagnosen eines Mangels an Data Scientists sowie damit einhergehende Forderungen – sowohl aus dem Feld der Ökonomie als auch der Wissenschaft – nach einem verstärkten Aufbau von Kompetenzen in den Datenwissenschaften aufgenommen und in ihre Förderpolitiken integriert haben. Sie schreiben den Datenwissenschaften hohe Potenziale zu und artikulieren kollektive Visionen und Entwicklungsverläufe des neuen Wissensgebiets. In der qualitativen Inhaltsanalyse von bildungs- und forschungspolitischen Strategiedokumenten wird zu überprüfen sein, welche Zukunftsszenarien der Datenwissenschaften entworfen werden und inwiefern sie die Entwicklung des entstehenden Wissensfeldes strukturieren. Ferner sind die begleitenden Massnahmen und Investitionen von Interesse, durch welche den Datenwissenschaften gegenüber anderen Wissensgebieten prioritäre Förderung zukommt.

20 Das Akronym steht für »Education for Data Intensive Science to Open New science frontiers«. Für eine Projektbeschreibung vgl. online: <https://edison-project.eu/edison/edison-project/> (Zugriff: 03.02.2022).

3.4 Konstruktionen der Datenwissenschaften im akademischen Feld

3.4.1 Curricula als Forschungsgegenstand

Curricula bilden ein zentrales Forschungsgebiet der Bildungs- und Hochschulforschung (Pinar et al. 1995; Brüsemeister 2008). Neue bzw. veränderte Inhalte und Lehrformate in Curricula gehen oft mit intensiven Debatten über die Positionierung, Rolle und Ziele der verantwortlichen Bildungsinstitutionen einher. Insofern artikulieren sich in einer konflikttheoretischen Perspektive in den Auseinandersetzungen um Curricula auch Konzeptionen sozialer Ordnung und damit gesellschaftliche Wertvorstellungen (Bernstein 1975; Binder 2002).

Im Feld der Hochschulbildung sind Curricula ein zentraler Mechanismus, mit der sich wissenschaftliche Disziplinen reproduzieren (Holley 2009). Sie repräsentieren für die disziplinäre Gemeinschaft zentrale Werte und Normen und schaffen damit die Bedingungen für Forschungs- und Lehrpraktiken. Curricula sind viel mehr als nur die Summe der disziplinären Wissensbestände; sie umfassen auch soziale Prozesse und Relationen zwischen verschiedenen Akteur*innen wie Dozierenden, Studierenden, den Universitäten und Hochschulen als Organisationen sowie verschiedenen externen Entitäten wie professionellen Fachgesellschaften, politischen Behörden, Akkreditierungsagenturen und Akteur*innen im Feld der Ökonomie; Letztere motivieren insbesondere die Nachfrage nach Absolvent*innen mit bestimmten Kompetenzprofilen. Entsprechend existiert eine Vielzahl wissenschaftsinterner und -externer Faktoren, die die Planung, Etablierung oder Anpassung von Curricula rahmen und beeinflussen (Knight et al. 2013; Lattuca & Stark 2009).

Die Herausbildung eines neuen Wissensfeldes wie der Datenwissenschaften bedingt insofern neben der Festlegung von Vermittlungsformaten, Lehrmethoden und Wissensinhalten auch die Positionierung gegenüber verwandten disziplinären Feldern, die über etablierte Curricula verfügen. Darüber hinaus formulieren sie stets auch Konzeptionen über die Rolle und Bedeutung des Wissensgebiets in der Gesellschaft: Indem bestimmte Protagonisten die Datenwissenschaften als unverzichtbar für die Bewältigung globaler Menschheitsprobleme deuten, setzen sie einen Definitionsanspruch weit über das Feld der Wissenschaft hinaus. Manche Curricula nehmen diese Positionierung auf und vermitteln sie an Studierende und prägen somit nicht nur deren Verständnis auf die Datenwissenschaften, sondern auch ihr Weltbild an sich. Dies zeigen insbesondere auch Debatten über die sozialen und epistemologischen Bedingungen der Wissensproduktion in den Datenwissenschaften (Bates et al. 2020; Kross et al. 2020).

Curricula bilden insofern einen interessanten Forschungsgegenstand, um einerseits zeitgenössische Vorstellungen über die gesellschaftliche Verortung eines Wissensfeldes und andererseits die Austauschbeziehungen zwischen verschiedenen Akteur*innen im Feld der Hochschulbildung selbst und jenen anderer Felder zu untersuchen.

3.4.2 Charakteristika und Herausforderungen interdisziplinärer Curricula

In den letzten Jahrzehnten ist ein signifikanter Anstieg an interdisziplinären Curricula auf unterschiedlichen Stufen des tertiären Bildungswesens zu beobachten (Brint et al. 2009; Holley 2009; Knight et al. 2013). Interdisziplinäre Curricula erheben den Anspruch, den limitierten Blick einer einzelnen Disziplin auf einen Gegenstand oder ein Forschungsgebiet zu überwinden und stattdessen multiple Perspektiven darauf zu eröffnen. Beispiele aus jüngerer Zeit sind Neuro-, Klima- oder Umweltwissenschaften, aber auch die diversen Ausprägungen der Cultural Studies in den Geistes- und Sozialwissenschaften (Brint et al. 2009). Insbesondere bei den Studiengängen mit natur- und technikkundenschaftlicher Ausrichtung handelt es sich oft um technologieintensive, forschungsorientierte Wissensgebiete, die meist auf Masterstufe angeboten werden, jedoch keine eigene departementale Entsprechung haben und sich eng an ausseruniversitäre Forschungsbereiche, insbesondere in der Industrie, anlehnen (Abbott 2005: 265; Brint et al. 2009). Daraus resultiert eine zunehmende Entkopplung von interdisziplinären Forschungsgebieten und organisationalen Einheiten des akademischen Feldes (wie Instituten und Fakultäten). Dies lässt sich empirisch auch am Beispiel der Datenwissenschaften beobachten (vgl. Kap. 8.5).

Interdisziplinäre Curricula entstehen in Räumen zwischen etablierten disziplinären Strukturen, in denen die existierenden institutionellen Einheiten kaum Kontrolle ausüben können (Lindvig et al. 2019). Dies schafft Opportunitäten für Lehrende und Lernende, interdisziplinäre Aktivitäten zu kreieren, die sich nicht in die existierenden disziplinären Strukturen einfügen müssen – und die sich, sofern sich eine entsprechende Nachfrage einstellen sollte, auch in hohe Profite feldspezifischen Kapitals transformieren lassen. Eine schwache Positionierung interdisziplinärer Aktivitäten kann insofern auch strategische Absicht sein, Räume zu schaffen bzw. offenzuhalten, damit gerade keine Institutionalisierung erfolgt (Eyal & Pok 2015).

Im organisationalen Alltag führen interdisziplinäre Curricula allerdings zu diversen Herausforderungen und Konflikten (Holley 2009: 242): Zum einen sind interdisziplinäre Curricula Teil von Universitäten, die grossmehrheitlich disziplinär, in Instituten und Fakultäten, organisiert sind. Verschiedene Disziplinen können demnach Expertise auf einem solchen Gebiet beanspruchen. Dabei droht gewissermassen die Übertragung existierender disziplinärer Grenzziehungen und Konflikte auf interdisziplinäre Gegenstände und Gefässe. Zum anderen sind die Lehrenden eines interdisziplinären Curriculums in der Regel in einem disziplinären Feld trainiert, d. h., es fehlen gemeinsame Werte und Normen, die eine spezifische Fachkultur mit entsprechend sozialisierten Praktiker*innen begründen.

Neben der institutionellen Verankerung, Grösse und Struktur der lokalen organisationalen Einheiten (Brint et al. 2009; Merz & Sormani 2016; Small 1999) beeinflussen auch historische Besonderheiten, beispielsweise bei der Lancierung des Programms, die Ausgestaltung und das Ausmass an Interdisziplinarität massgeblich (Holley 2009; Lindvig et al. 2019). Augsburg und Henry (2009) charakterisieren interdisziplinäre Curricula als ein Kontinuum zwischen »starken« und »schwachen« Programmen: Erstere sind hochgradig strukturiert und weisen einen hohen Anteil verpflichtender Studienleistungen im interdisziplinären Bereich eines Programms auf. Dies geht oft mit einer höheren Anzahl Lehrender einher, deren Denomination im Kernbereich verortet ist. Schwache Programme hingegen bieten mehr Wahlmöglichkeiten ausserhalb des

definierten Kernbereichs und übertragen somit die Integration der involvierten Wissensbereiche und Forschungsperspektiven auf die Studierenden selbst.

3.4.3 Curricula in den Datenwissenschaften

Zeitgenössische Curricula in den Datenwissenschaften sind meist nach dem »Baukastenprinzip« organisiert und gemäss aktuellen berufs- und wirtschaftspädagogischen Gesichtspunkten erarbeitet werden (Schumann et al. 2016). Gemeinhin sind Curricula in Datenwissenschaften derart strukturiert, dass ein interdisziplinärer »Kern« in den Computerwissenschaften, Statistik, Mathematik und Engineering verortet wird, der sich auf Methoden, Technologien und andere Formalisierungen abstützt. Bisweilen werden zusätzliche Disziplinen wie Ökonomie, Semantik, Linguistik, Business, Design oder Visualisierung zum erweiterten Kern gezählt (Chatfield et al. 2014). Auch kommunikative und soziale Kompetenzen werden in empirischen Studien wiederkehrend als bedeutend identifiziert (Feldman et al. 2017; Ismail & Abidin 2016; Schumann et al. 2016).

Der nachfolgende Forschungsstand zu Curricula in den Datenwissenschaften umfasst äusserst heterogene Perspektiven unterschiedlicher Akteur*innen und ist in drei Teilen strukturiert. Erstens werden in einer disziplinenorientierten Perspektive – wie bereits gezeigt – normative Anforderungen an solche Programme formuliert (Cleveland 2001; Donoho 2017; Song & Zhu 2017; De Veaux et al. 2017). So versuchen primär hochschulpolitische Akteur*innen, datenwissenschaftliche Kompetenzprofile für unterschiedliche Bildungsstufen zu identifizieren und zu definieren (BHEF 2016; ETH-Rat 2016a; NASEM 2018). Parallel dazu skizzieren Praktiker*innen aufgrund eigener Forschungs- und Lehrerfahrungen sowie bestehender Ausbildungsgänge normativ die Anforderungen an solche Curricula (Cleveland 2001; Donoho 2017; Gupta et al. 2015; Kane 2014; De Veaux et al. 2017). In den USA haben sich verschiedene wissenschaftliche Institutionen intensiv mit den Herausforderungen der epistemologischen Transformation von Wissenschaft und der Rolle der Datenwissenschaften auseinandergesetzt (Berman et al. 2016; Kloefkorn et al. 2020; NASEM 2017, 2018; NASEM & The Royal Society 2018). Eine Expert*innenkommission der NASEM fasst zehn verschiedene Kompetenzbereiche als »Data Acumen«²¹ (was mit »Datenscharfsinn« oder »Datenverstand« übersetzt werden kann) für die Konzeption datenwissenschaftlicher Studiengänge auf *Undergraduate*-Stufe zusammen (NASEM 2018: 2–8). Das bereits erwähnte Projekt EDISON entwickelte ein Kompetenzprofil sowie ein Modell-Curriculum für die Datenwissenschaften. Damit wurde somit eine Art »Landkarte« der unterschiedlichen Berufsprofile der entstehenden Data Economy formuliert, die zu einer europaweiten Vereinheitlichung und Zusammenführung von universitärem Angebot und industrieller Nachfrage nach datenwissenschaftlichen Kompetenzen beitragen sollten (Demchenko et al. 2016).²²

21 Es handelt sich um die folgenden zehn Kompetenzbereiche: »Mathematical foundations, Computational foundations, Statistical foundations, Data management and curation, Data description and visualization, Data modeling and assessment, Workflow and reproducibility, Communication and teamwork, Domain-specific considerations, and Ethical problem solving« (NASEM 2018: 2–8ff.).

22 EDISON entwirft die folgenden Kompetenzen als Bestandteil einer »Data Science Literacy«: »Statistical techniques [...], Computational thinking and programming with data [...], Programming languages

Unter solche normativen Anforderungen fallen auch Bemühungen, Curricula in den Datenwissenschaften zu »demokratisieren« (Cornelissen 2018; Kross et al. 2020), d. h. ihre soziale Verortung und die gesellschaftlichen Implikationen datenwissenschaftlicher Praxis aktiv zum Thema zu machen (Green 2018). Dabei sind insbesondere zwei Themen virulent: Einerseits ist das Teilen von Inhalten, Daten, Methoden, Algorithmen etc. unter Losungen wie Open Science und Open Data von Beginn weg in die Praxis des neuen Wissensgebiets eingeschrieben. Dies wird besonders durch Kurse zur Reproduzierbarkeit datenwissenschaftlicher Forschung, beispielsweise als Teil des erwähnten »Data Acumen« (NASEM 2018: 2–8), gefördert. Andererseits sollen Gender, Race, Diversity und weitere Dimensionen von *Social Justice* in den Curricula und Praktiken der Datenwissenschaften Berücksichtigung finden (Berman et al. 2016, 2018; Berman & Bourne 2015; Durantón et al. 2020; Geiger et al. 2019; Rawlings-Goss 2018).

Zweitens schildern Fallstudien ihre Erfahrungen und *Best Practices* einzelner Studiengänge (Anderson et al. 2014; Borne et al. 2009; Kreuter et al. 2018; McNamara et al. 2017; Stockinger et al. 2016) oder Veranstaltungen (Asamoah et al. 2015; Kross et al. 2020; Schuff 2018). Eine der ersten Publikationen stammt von Borne et al. (2009), die die Etablierung des Undergraduate-Programms »Data and Computational Sciences« an der George Mason University beschreiben. Das Programm vereint datengetriebene und rechenintensive Praktiken mit Anwendungsfeldern der Naturwissenschaften. Die Studiengangbezeichnung ist Ausdruck davon, dass zum Zeitpunkt der Etablierung multiple Bezeichnungen für die epistemische Transformation der Wissensproduktion koexistierten. Entsprechend fokussiert das Programm stärker auf das Management und die Analyse wissenschaftlicher Daten (ebd.: 78f.), als dies spätere Angebote unter dem Label »Data Science« tun.

Der erste Studiengang, der explizit »Data Science« im Namen trägt, ist ein Programm am College of Charleston in den USA, das seit Mitte der 2000er-Jahre existiert (Anderson et al. 2014). Neben dem Umstand, dass es sich um ein Provinzcollege handelt, ist auch die Verortung auf Undergraduate-Stufe aussergewöhnlich. Solche wurden an den meisten Universitäten erst nach Erfahrungen mit (professionellen) Masterprogrammen implementiert. Inhaltlich fällt auf, dass von den fünfzehn Kursen, die den »Kern« des Studienprogramms bilden, lediglich deren drei in »Data Science« sind, gegenüber sieben Kursen in Informatik und acht in Mathematik (inkl. Statistik) (ebd.: 147). Dennoch skizziert das Programm eine curriculare Struktur, die später von diversen Hochschulen in ähnlicher Form aufgenommen wurde und sich auch in den Empfehlungen der National Academies (2018) wiederfindet.

Stockinger et al. (2016) schildern ihre Erfahrungen anhand eines Weiterbildungsprogramms in Datenwissenschaften (Diploma of Advanced Studies in Data Science) an der Zürcher Hochschule für angewandte Wissenschaften (ZHAW) in Winterthur. Das Curriculum legt den Schwerpunkt auf die Verknüpfung von theoretischen Wissensbeständen in Datenanalyse und Information Engineering mit praktischen Umsetzungen (Data Science Applications), um damit Berufstätige zu adressieren. Die Ergebnisse eines Surveys unter Absolvent*innen deuten die Heterogenität der beruflichen und feldspezifischen Verortung der Teilnehmenden an (ebd.: 73f.). Die Befrag-

and tools for data analysis [...], Data visualization languages and tools [...], Data Management« (Demchenko et al. 2016: 30).

ten erwarten positive Effekte der Weiterbildung auf ihre berufliche Stellung, was sich etwa darin manifestiert, dass sie »mehr Verantwortung bei analytischen, quantitativen und technischen Aufgaben« erwarten. Auch verwenden sie fortan häufiger die Bezeichnung »Data Scientist« anstelle anderer Stellentitel (ebd.: 77f.).

Kross et al. (2020) beschreiben die Einführung des MOOCS (Massive Open Online Course) »Data Science Specialization« der John Hopkins University. Die Ubiquität grosser Datenmengen und die postulierte »Demokratisierung« des Datenzugangs und der Verfügbarkeit entsprächen allerdings nicht der Verfügbarkeit der akademischen Ausbildungsmöglichkeiten. Aufgrund der grossen Nachfrage unter Studierenden habe das neue Programm, das auf existierenden Onlineangeboten aus dem Bereich Biostatistik basierte, zu einer Angleichung anderer Studienangebote (online und offline) an die curricularen Inhalte des MOOC geführt.

Schliesslich entwerfen Kreuter et al. (2018) ein Online-Mastercurriculum für Professionelle der Sozial- und Marktforschung, das Surveyforschung mit Methoden und Technologien der Datenwissenschaften verbinden soll. Zu den fünf Schwerpunkten des Programms zählen die Neudefinition von Forschungsfragen, Datenerhebung, Datenaufbereitung und -speicherung, Datenanalyse sowie die Produktion von Output und Zugänglichkeit der Daten (ebd.: 2). Das Programm beabsichtigt so eine datenwissenschaftliche Ergänzung existierender Curricula.

Drittens existieren schliesslich einige Studien, die die Strukturen und Inhalte von Curricula in den Datenwissenschaften, primär von Universitäten in den USA, empirisch untersuchen (Aasheim et al. 2014, 2015; Bukhari 2020; Ortiz-Repiso et al. 2018; Tang & Sae-Lim 2016). Aasheim et al. (2014) untersuchen die Implementierung von Business-Analytics-, Data-Analytics- und Data-Science-Programmen auf Undergraduate-Stufe. Zum damaligen Zeitpunkt boten lediglich 21 Universitäten im Sample Majors in den genannten Feldern an, davon lediglich fünf in Datenwissenschaften. Ein Vergleich der curricularen Strukturen und Inhalte der Studiengänge verweist auf signifikant ähnliche Verteilungen von Kursen und *Credit Hours* in den Bereichen Programmieren, Statistik, Mathematik, Data Analytics/Modeling sowie Data Mining. Differenzen existieren insbesondere in den Bereichen Datenbanken, Big Data und Visualisierung (ebd.: 17). Die Programme in Business und Data Analytics werden primär von Business Schools angeboten, während diejenigen in »Data Science« mehrheitlich in computerwissenschaftlichen Departements und Schools verortet sind.

Tang und Sae-Lim (2016) untersuchen Strukturen, Inhalte und Beschreibungen von dreissig Data-Science-Studiengängen an US-amerikanischen iSchools (vgl. auch Ortiz-Repiso et al. 2018). Trotz einer gewissen Heterogenität der Studiengänge hinsichtlich der involvierten Disziplinen zeigt die Inhaltsanalyse der Materialien, dass die Studiengänge mathematische Kursinhalte insbesondere in den Kernbestandteilen der Curricula gegenüber informationswissenschaftlichen sowie Visualisierungskompetenzen vernachlässigen würden.

Schliesslich analysiert auch Bukhari (2020) Strukturen und Inhalte von dreissig Studiengängen der Datenwissenschaften an US-amerikanischen Universitäten. Es zeigen sich signifikante Überschneidungen im curricularen Aufbau, in der Anzahl erforderlicher *Credit Hours* sowie der Bedeutung, die Praktika und Capstone-Kurse zukommt. Bezüglich der institutionellen Verortung in den Hochschulen sind Management und Business Schools führend, während Schools of Engineering, Sciences sowie Computer Science nachgelagert sind (ebd.: 6). Es dominieren Kursinhalte in Daten-

analyse und -management, (Business) Analytics, Statistik und Mathematik gegenüber solchen in Informatik, Programmierung, künstlicher Intelligenz und Machine Learning sowie Datenschutz und Ethik (ebd.: 8).

Für das deutsche Hochschulfeld geben Lübcke und Wannemacher (2018) sowie Heidrich et al. (2018) einen Überblick über existierende Studiengänge, Kursangebote und deren institutionelle Verortung. Die ab 2014 entwickelten Studiengänge sind überwiegend generalistisch angelegt und weisen nur selten eine fachliche Konzentration auf. Mehrheitlich handelt es sich um Masterstudiengänge an staatlichen Universitäten und Fachhochschulen, die auf vorangehenden Informatik- oder Mathematikstudiengängen auf Bachelorstufe aufbauen. Viele Studienangebote sind eher forschungsorientiert, was aufgrund der steigenden Nachfrage im Arbeitsmarkt als Nachteil interpretiert wird (Lübcke & Wannemacher 2018: 3). Bezüglich der in den Studiengängen und weiteren Ausbildungsformaten vermittelten Data Literacy orten Heidrich et al. (2018) eine hohe Diversität und unklare Definitionen der verwendeten Begrifflichkeiten. Obwohl in vielen Implementierungen ähnliche Ansätze verfolgt würden, resultiere durch die Vielfalt der Begriffe eine gewisse Ambiguität. Deshalb brauche es standardisierte Kompetenz-Frameworks für »Data Literacy« (wie EDISON, vgl. oben) auf den unterschiedlichen Stufen des Bildungssystems.

Die Diskussion existierender Studien indiziert eine Breite disziplinärer und institutioneller Implementierungen und Positionierungen von Curricula in den Datenwissenschaften im Feld der Hochschulbildung. Trotz der identifizierten Diversität, die im Kontext der Herausbildung des Wissensfeldes zu betrachten ist – der Grossteil der Studiengänge existiert seit weniger als fünf Jahren –, zeigen sich jedoch bereits Prozesse von Angleichung bzw. mimetischer Isomorphie (DiMaggio & Powell 1983: 151f.), die in strukturellen und inhaltlichen Ähnlichkeiten der Curricula resultieren. Die Forschung zu interdisziplinären Studiengängen verweist auf Herausforderungen in der Etablierung und Implementierung solcher Curricula. Hierbei ist insbesondere auch der Einfluss von curricularen Empfehlungen und anderer normativer Anforderungen an solche Programme zu prüfen. Die Etablierung von akademischen Studiengängen in Datenwissenschaften an Schweizer Universitäten und Hochschulen stellt insofern einen geeigneten empirischen Fall dar, um solche Prozesse von Diffusion und Kanonisierung, aber auch organisationaler Aushandlung und Koordination sowie die Herausforderungen, die die Genese interdisziplinärer Curricula markieren, exemplarisch zu untersuchen.

3.5 Zwischenstand

Der Forschungsstand hat die Bedeutung multipler Begriffsverständnisse, disziplinärer und theoretischer Perspektiven, von Qualifikations- und Kompetenzzuschreibungen sowie kollektiven Visionen und Entwicklungsszenarien bei der Konstruktion der Datenwissenschaften als entstehendes transversales Wissensfeld aufgezeigt. Nachfolgend fasse ich einige zentrale Erkenntnisse zusammen, um die Analyse der empirischen Untersuchungsfälle im bestehenden Forschungs- und Erkenntnisstand zu verorten.

Bei den Datenwissenschaften handelt es sich um einen empirischen Gegenstand, dessen Genese in verschiedenen sozialen Sphären synchron verortet werden kann. Im

Zentrum steht weniger die Frage, welche Disziplin oder welches soziale Feld das Primat bei der Genese der Datenwissenschaften für sich beanspruchen kann, sondern der Umstand, dass Akteur*innen aus verschiedenen sozialen Feldern darin involviert sind. Insofern sind die Datenwissenschaften ein transversales Phänomen: Als epistemisches Feld basieren sie auf einer Reihe von technologischen und wissenschaftlichen Innovationen (vor allem in Methoden, algorithmischen Systemen und leistungsfähiger Hardware), die in andere soziale Felder diffundiert sind. Sie stützen sich auf technologische Infrastrukturen, die wiederum durch Kollaborationen und Wissenstransfer von industriellen Akteur*innen und wissenschaftlicher Forschung entstanden sind. Trotz der wissenschaftlichen Konnotation des Begriffs »Data Science« lässt sich das entstehende Wissensgebiet nicht auf eine feldspezifische Handlungslogik, beispielsweise der Wissenschaft oder der Industrie, zurückführen, sondern muss als Mischform, als Hybrid betrachtet werden.

Die Auseinandersetzung mit verschiedenen Definitionen legt nahe, dass Interdisziplinarität und Multiperspektivität konstitutive Merkmale der Datenwissenschaften darstellen: Als Kernbereich werden gemeinhin Statistik, Mathematik und Informatik genannt, die sich je nach Ausrichtung mit weiteren Wissensfeldern bzw. Disziplinen (Engineering, Wirtschaftswissenschaften, Linguistik etc.) verknüpfen. Anwendungsfelder (*domains*) liegen in der Wissenschaft, in der Medizin, in ökonomischen Feldern, in der staatlichen Verwaltung, im Non-Profit-Bereich und in anderen sozialen Feldern. Analog zu anderen technischen Wissensformationen zirkulieren nicht nur Methoden, Tools und Wissensbestände zwischen den involvierten Feldern, sondern auch individuelle Akteur*innen können zwischen technischen Hochschulen, Forschungslabors und der Tech-Industrie, aber auch zwischen der staatlichen Verwaltung, internationalen Organisationen oder Think Tanks hin- und herwechseln (»*revolving doors*«) (Safavi et al. 2018; Beckert et al. 2008). Somit profitieren auch Organisationen als kollektive Akteur*innen davon, denn es eröffnen sich vielfältige Möglichkeiten zur Kooperation: Während Unternehmen ihre humanen und methodischen Ressourcen erweitern können, bietet sich Universitäten und Hochschulen die Gelegenheit, Innovationsfähigkeit und Praxisnähe unter Beweis zu stellen. Auch bildungs- und forschungspolitische Institutionen tätigen hohe Investitionen in die Datenwissenschaften. Schliesslich profitiert auch die Verwaltung von der Expertise, indem sie zur Grundlage staatlicher Politiken oder direkt in staatliche Strukturen inkorporiert wird (Bundesrat 2020).

Über die Grenzen, d. h. was noch dazugezählt werden kann und was nicht, herrscht tendenziell Unklarheit. Sie sind Gegenstand von Aushandlungsprozessen, die unterschiedliche Formen annehmen. In der Verhandlung der Grenzen der Datenwissenschaften, was noch dazugehört und was nicht, wird die fundierende Rolle von Kompetenzen und Begriffen erkennbar: So beschäftigt sich eine Vielzahl unterschiedlicher Akteur*innen mit der Frage nach den »richtigen« Kompetenzen und Tätigkeitsprofilen. Charakteristisch für solche Diskussionen ist, dass umfangreiche Listen unterschiedlicher Methoden, Tools, Bildungsqualifikationen und individueller Eigenschaften angeführt und miteinander verglichen werden. Solche Aufzählungen verweisen einerseits auf manifeste Unsicherheiten über die relevanten Begrifflichkeiten, andererseits aber auch darauf, dass sich bestimmte Kategorien in den Datenwissenschaften etabliert haben, die in Stellenausschreibungen, hochschulpolitischen Strategiedokumenten oder Curricula Verwendung finden.

Schliesslich entwerfen Akteur*innen aus Wissenschaft, Ökonomie und Politik durch multiple Bedeutungszuschreibungen divergierende Perspektiven auf die Datenwissenschaften als einen Raum zwischen Feldern. Bedeutsam sind insbesondere Zuschreibungen von zukünftigen Potenzialen und Entwicklungsverläufen, durch die kollektive Visionen der gesellschaftlichen Nutzung des Wissensgebiets artikuliert werden (Jasanoff 2015; Mische 2014). So strukturiert etwa die Identifikation einer grossen »Nachfrage« nach Data Scientists in der Wirtschaft die Aktivitäten eines zu schaffenden »Angebots« und so die Erwartungen anderer Akteur*innen, namentlich von Universitäten sowie in der Bildungs- und Forschungspolitik.

Kapitel 4 – Forschungsdesign

Ich verfolge den Ansatz einer verteilten Analyse von Zwischenräumen, in dem ich die Emergenz und die Bedeutungszuschreibungen der Datenwissenschaften in verschiedenen sozialen Feldern untersuche. In seinem Aufbau folgt das Kapitel der Grundstruktur der Arbeit: Zunächst ordne ich die Arbeit in das allgemeine Forschungsprogramm der Feldanalyse ein und erläutere das Forschungsdesign, um die Fragestellungen adäquat untersuchen zu können (Kap. 4.1). Danach beschreibe ich die Datengrundlage und das methodische Vorgehen für die drei empirischen Kapitel: erstens die Erhebung und Auswertung der Stellenanzeigen im Arbeitsmarkt (Kap. 4.2), zweitens die Strategiedokumente im Feld der Politik sowie das inhaltsanalytische Vorgehen (Kap. 4.3) und drittens Curricula sowie Interviews mit Lehrenden der Datenwissenschaften im akademischen Feld, die ich ebenfalls inhaltsanalytisch auswerte (Kap. 4.4). Abschliessend fasse ich das empirisch-analytische Potenzial des gewählten Forschungsdesigns hinsichtlich der verschiedenen Untersuchungsebenen zusammen (Kap. 4.5).

4.1 Die verteilte Analyse von Zwischenräumen

Die Arbeit verortet sich im »Forschungsprogramm« der Feldanalyse (Bernhard & Schmidt-Wellenburg 2012), das über den heterogenen Begriff des Feldes ein breites Spektrum empirisch-methodischer Vorgehensweisen abdeckt. Bernhard und Schmidt-Wellenburg (2012: 28) schlagen in Anlehnung an Lakatos (1978) vor, die Methodologie der Feldanalyse in einen programmatischen »Kern« von grundlegenden, stillschweigend vorausgesetzten Annahmen und einen flexibleren »Schutzgürtel«, der regelmässig empirischer Kritik und Auseinandersetzung unterzogen wird, zu differenzieren: Zum Kern des feldanalytischen Forschungsprogramms zählen die Autoren etwa die Einsicht, dass die soziale Realität als Ansammlung einander überlagernder und konkurrierender sozialer Felder zu verstehen und zu untersuchen ist. Dagegen sei im Schutzgürtel die Frage offen, »wie diese Felder zugeschnitten sind, welche Auseinandersetzungen sie im innersten bewegen und wie sie sich über die Zeit entwickeln« (Bernhard & Schmidt-Wellenburg 2012: 30f.).

Ich folge dieser forschungsprogrammatischen Unterscheidung, da das vorgeschlagene analytische Modell damit kompatibel ist: Räume zwischen Feldern als analytisches Konstrukt stützen sich elementar auf das Feldkonzept, erweitern es jedoch in einem zentralen Punkt, nämlich in der Frage der Lokalisierung neuer Wissensgebiete,

die nicht relativ autonome Felder, sondern unorganisierte, durchlässige und hybride Sphären sind. Daraus resultiert allerdings eine empirisch-methodische Herausforderung: Da die Feldtheorie Bourdieus, in geringerem Masse auch jene von Fligstein und McAdam, keinen ›Platz‹ für Zwischenräume als soziale Phänomene hat, finden sich in der feldanalytischen Methodendiskussion entsprechend keine Hinweise für deren Analyse (Bernhard & Schmidt-Wellenburg 2012; Blasius et al. 2019; Fligstein & McAdam 2012; Schmidt-Wellenburg & Bernhard 2020b).¹ Wie also können Zwischenräume als analytische Erweiterungen des Feldansatzes empirisch erforscht werden?

Eyal arbeitet in seinen Analysen zur Genese von Räumen zwischen Feldern bzw. Netzwerken der Expertise primär historisch, d. h., er rekonstruiert die Entwicklung solcher Wissenskonfigurationen durch eine umfassende Genealogie verschiedener Phasen, darin erscheinender, konkurrierender Expertisen und veränderter Akteurskonstellationen (Eyal 2002, 2013a, 2013b; Eyal & Pok 2015). Eine historische Rekonstruktion *post hoc* ist allerdings für die vorliegende Arbeit aus zwei Gründen keine Option. Erstens befindet sich das Wissensgebiet der Datenwissenschaften nach wie vor im Prozess seiner Etablierung – es ist noch unklar, wohin die Entwicklung führen wird: Etabliert es sich als ein Feld mit eigenen Spielregeln und Logiken, nach denen feldspezifisches Kapital zuteilwird? Existiert es auf Dauer als ein wenig regulierter, durchlässiger Raum zwischen Feldern? Oder verfällt es gar als eigener Bereich und existiert in anderen sozialräumlichen Phänomenen weiter? Zweitens ist das Phänomen in seiner Breite schlicht zu umfassend, als dass diese Arbeit eine vollständige Analyse der involvierten Felder wie Wissenschaft, Wirtschaft oder Politik leisten kann – vielmehr müssen dies spezifische Feldanalysen tun (Brandt 2016; Grommé et al. 2018).

Ich lege in dieser Arbeit deshalb einen anderen Fokus: Wie bereits ausgeführt, interessiert mich vor allem das feldübergreifende Moment des Phänomens, also der Umstand, dass Akteur*innen in verschiedenen Feldern an der Konstruktion der Datenwissenschaften beteiligt sind. Ich verfolge deshalb den Ansatz einer verteilten Analyse von Zwischenräumen²: Dabei untersuche ich ausgewählte Ausschnitte von Feldern im Hinblick auf das interessierende transversale Phänomen, die Emergenz der Datenwissenschaften. Dies ist möglich, da sich die Wirkungen der feldspezifischen Praktiken – d. h. die empirisch beobachtbaren »Feldeffekte« (Bourdieu & Wacquant 1996: 131) – nicht nur innerhalb der jeweiligen Felder der Wissenschaft, Ökonomie oder Politik, sondern ebenso im dazwischenliegenden Raum verorten lassen.

Mit der empirischen Vorgehensweise einer verteilten Analyse von Räumen zwischen Feldern schliesse ich an einen Kerngedanken der *multi-sited ethnography* bei George Marcus (1995; Nadai & Maeder 2005) an.³ Marcus plädiert dafür, von einer

1 Bourdieu hat ein klares Vorgehen in der Analyse sozialer Felder formuliert (Bourdieu & Wacquant 1996: 136): Erstens gilt es, die Position eines Feldes im Verhältnis zum Feld der Macht zu untersuchen; zweitens um die Ermittlung der objektiven Struktur der Relationen zwischen den Positionen der in diesem Feld miteinander konkurrierenden Akteur*innen; drittens um die Analyse der Habitus, in denen die Dispositionen der Akteur*innen verinnerlicht sind.

2 Die Feldanalyse im Anschluss an Bourdieu schlägt eine ähnliche Vorgehensweise für transnationale Felder (Schmidt-Wellenburg & Bernhard 2020b: 2) bzw. transversale Felder (Witte & Schmitz 2019) vor.

3 Scheel et al. (2020) sprechen in ihrem ähnlich gelagerten methodischen Vorgehen von einer »transversalen Ethnographie«, was allerdings etwas irreführend ist in Bezug auf den Untersuchungsgegenstand, nämlich transversale Wissensfelder.

Fokussierung auf stationäre lokale Subkulturen und deren Einbettung in ein übergeordnetes Makrosystem wegzukommen und überkommene Dichotomien wie global vs. lokal, System vs. Lebenswelt etc. zu verwerfen. Vielmehr gelte es, die verteilten Phänomene in ihrer wechselseitigen Durchdringung zu betrachten, um fragmentierte, komplexe soziale Beziehungen erschliessen zu können (Marcus 1995: 98).

Eine solche Konzeption bietet sich analog für die Analyse von Zwischenräumen an: Anstelle einer monolithischen Analyse einzelner Felder, die binäre Strukturlogiken (dominant vs. dominiert, legitimes vs. illegitimes Kapital, autonom vs. heteronom etc.) reproduziert, strebt die Arbeit danach, den jeweiligen Beitrag der involvierten sozialen Felder bei der Herausbildung eines neuen Wissensgebietes zu betrachten, das eben gerade nicht ausschliesslich wissenschaftlich, ökonomisch, kulturell etc. ist. Vielmehr ist der Gegenstand in Zwischenräumen verteilt und fragmentiert zugleich: Es manifestieren sich in ihm Effekte verschiedener Felder, da er zum Objekt von konflikthafter wie kooperativer Praktiken – feldanalytisch formuliert quasi zum »Spielball« – dieser Felder wird. Dennoch hat kein Feld das Primat oder die Deutungshoheit inne, sondern die Bedeutung ergibt sich erst durch die Gesamtheit der multiplen Perspektiven und Praktiken, die im Zwischenraum koexistieren. Es geht deshalb auch nicht darum, einen Vergleich der einzelnen Felder im Hinblick auf das interessierende Phänomen anzustellen, sondern darum, die Analyseteile des Puzzles zusammenzufügen (Nadai & Maeder 2005: 8). Dies erfolgt letztlich in der Absicht, mehr Erkenntnisse über das Phänomen zu gewinnen als die Summe der Analysen seiner einzelnen Teile – ein Postulat, das auch die Feldanalyse teilt (Bernhard & Schmidt-Wellenburg 2012: 35).

Von den unterschiedlichen Strategien der Feldkonstruktion, die Marcus in seinem Aufsatz vorschlägt, schliesst sich die vorliegende Arbeit jener der »Follow the Metaphor« (Marcus 1995: 108) an: Dabei leitet die Produktion und Zirkulation von Zeichen, Symbolen und anderen Repräsentationen in verschiedenen Bereichen das Forschungsdesign und empirische Vorgehen an. Marcus bezieht sich dabei prominent auf die Arbeit von Emily Martin (1993), die in ihrer Studie den Metaphern bzw. Konstruktionen folgt, wie in verschiedenen gesellschaftlichen Sphären über das Immunsystem gesprochen wird, von den Massenmedien über die »Strasse«, AIDS-Therapien, Alternativmedizin bis hin zur Wissenschaft.⁴

Ich gehe in der Arbeit ähnlich vor, um die Genese neuer Wissensgebiete zu erklären: Dazu folge ich gewissermassen den Effekten von Begriffs- und Grenzarbeit, durch die in unterschiedlichen Feldern multiple Bedeutungen des Gegenstandes Datenwissenschaften entworfen werden. Somit bietet es sich an, den Untersuchungsgegenstand in den relevanten Feldern verteilt zu untersuchen. Dabei verfolge ich einen stärker akteursorientierten Ansatz: Repräsentationen eines (neuen) Gegenstandes zirkulieren nicht einfach so in sozialen Feldern. Sie werden primär von Organisationen, d. h. kollektiven Akteur*innen, mobilisiert, die damit spezifische Ziele und Interessen in ihren jeweiligen Feldern verfolgen. Sie versuchen eine bestimmte Deutung des Gegenstandes zu etablieren, befördern allerdings synchron dazu auch die Entstehung neuer Möglichkeitsräume, um dadurch ihre Handlungsoptionen zu erweitern.

Um die Herausbildung von Zwischenräumen empirisch-analytisch untersuchen zu können, schlage ich ein dreistufiges Vorgehen vor: Erstens gilt es die relevanten umgebenden Felder zu identifizieren. In Anknüpfung an die (unscharfen) Grenzen

4 Eine ähnliche Strategie schlägt Abbott (2005: 249) für das vielschichtige Phänomen Alkoholismus vor.

von Feldern stellt sich die Frage: Die Effekte welcher Felder sind im Zwischenraum zu beobachten? An die Auswahl der umgebenden Felder anknüpfend ergibt sich zweitens die Frage nach den Akteur*innen: Welche Akteur*innen dieser Felder sind wie an der Konstruktion solcher Räume beteiligt? Inwiefern tragen sie zur Herausbildung neuer Räume bei? Sind die Relationen zwischen den Akteur*innen eher konflikthaft oder eher kooperativ? In einem dritten Schritt stellen sich Fragen nach den fundierenden Praktiken, die zur Etablierung von Zwischenräumen beitragen: Welche Praxismodi, d. h. welche kollektiven Stellungnahmen sind zu beobachten? Auf welchen Ebenen, d. h. wo finden diese statt bzw. von wo werden sie geäußert?

Die Auswahl der relevanten Felder leitet sich primär aus dem Forschungsstand ab, in dem ich die Bedeutung der Wissenschaft, der Ökonomie sowie der Politik als relevante soziale Sphären in der Konstruktion der Datenwissenschaften herausgearbeitet habe. Im Feld der Wirtschaft werden einerseits hohe Erwartungen an Data Scientists als Schlüsselfiguren datengetriebener Produktionsweisen formuliert, die insbesondere in der Hochschulbildung stark rezipiert werden (Saner 2019), andererseits zeigen sich manifeste Unsicherheiten und divergierende Deutungen über die relevanten Kategorien und Inhalte. In der Wissenschaft beschäftigen sich verschiedene Disziplinen seit Jahrzehnten mit den konkreten Praktiken und Wissensinhalten der Datenwissenschaften. Die Transformation hin zu einer datengetriebenen Wissensproduktion reartikuliert dabei existierende Konflikte und Grenzziehungen. Gleichzeitig eröffnen sich Opportunitäten, solche Konfliktlinien durch grenzüberschreitende Kooperationen in Forschung und Lehre zu überwinden und heterogene Perspektiven auf die Datenwissenschaften zu etablieren. Schliesslich prägen Akteur*innen der Hochschul- und Forschungspolitik Szenarien einer vielversprechenden Zukunft, die durch Investitionen und Fördermassnahmen begleitet werden. Somit koexistieren in den drei zentralen Feldern multiple Deutungen und konturieren über ihre Effekte den Raum dazwischen.

Geographisch beschränke ich mich dabei auf die Herausbildung der Datenwissenschaften und die Konfiguration der relevanten umgebenden Felder in der Schweiz. Obwohl neue Wissensfelder oft global strukturiert sind, bleiben sie durch lokale, regionale und nationalstaatliche Spezifika gekennzeichnet (Biniok 2013; Merz & Sorman 2016). Dennoch versuche ich, die transnationalen Relationen und wechselseitigen Abhängigkeiten mit zu berücksichtigen, um die Räume zwischen Feldern nicht als nationalstaatlich verfasst zu reifizieren (Schmidt-Wellenburg & Bernhard 2020b).

Um die in der Einleitung genannten Forschungsfragen empirisch-methodisch umsetzen zu können, verknüpfe ich das für die Soziologie relativ »neue« Verfahren des Topic Modeling (DiMaggio et al. 2013; Mützel 2015a; Papilloud & Hinneburg 2018) mit »traditionellen« sozialwissenschaftlichen Methoden der qualitativen Inhalts-, Curricula- und Interviewanalyse. Zunächst ermöglicht es das Topic Modeling von Stellenanzeigen, transversale Repräsentationen der Datenwissenschaften auf der gesellschaftlichen Makroebene zu analysieren. Obwohl hier das Material feldspezifisch nicht begrenzt ist, zeigt sich bei der Datenvorbereitung, dass Akteur*innen im ökonomischen Feld zentral sind. Das gewählte methodische Vorgehen ist jedoch in der Lage, sowohl feldspezifische als auch feldübergreifende Darstellungen zu identifizieren und letztere in allgemeinen Topics zu modellieren.

Im Fall des Politikfeldes bilden Strategiedokumente unterschiedlicher Akteur*innen der Hochschul- und Forschungspolitik das Datenmaterial. Durch eine qualitative Inhaltsanalyse können die spezifischen Zukunftsvisionen, die im Rahmen des Diskur-

ses über die Digitalisierung auf die Datenwissenschaften entworfen werden, rekonstruiert werden. Effekte dieser kollektiven Praktiken sind nicht nur in der Politik selbst, sondern auch in anderen Feldern, insbesondere der Wissenschaft und der Ökonomie zu beobachten.

Schliesslich bilden Curricula aller verfügbaren Studienangebote in Datenwissenschaften an Schweizer Universitäten und Hochschulen sowie Interviews mit Lehrenden das empirische Untersuchungsmaterial im akademischen Feld, die ich ebenfalls mittels qualitativer Inhaltsanalysen analysiere. Dies erlaubt es zum einen, die fundierende Rolle von Praktiken der Begriffsarbeit in der Herausbildung des Zwischenraumes als eines Möglichkeitsraums zu bestimmen. Zum anderen bietet die Kombination von kollektiven (Curricula) und subjektiven Stellungnahmen (Interviews) die Möglichkeit, die Widersprüche und Dynamiken synchroner Praktiken der Grenzziehung und Grenzüberschreitung zu adressieren.

4.2 Stellenanzeigen im Arbeitsmarkt

Der Forschungsstand hat deutlich gemacht, dass in der Suche und Diskussion über die relevanten Kompetenzen ein zentrales Moment in der Konstruktion der Datenwissenschaften liegt. Durch die Erhebung von Stellenanzeigen für Data Scientists können divergierende Repräsentationen der Datenwissenschaften in Stellenanzeigen feldübergreifend analysiert werden. Um ein Sample von Stellenausschreibungen für Data Scientists und verwandten Bezeichnungen zusammenstellen, aufbereiten und auswerten zu können, war ein mehrstufiges Verfahren notwendig, das ich hier ausführlich darlegen werde. Zunächst lege ich im ersten Abschnitt einige allgemeine Charakteristika von Stelleninseraten als Daten zugrunde (Kap. 4.2.1), bevor ich im zweiten Teil auf die Prozesse der Datenerhebung (Kap. 4.2.2) sowie der Kodierung und Datenbereinigung (Kap. 4.2.3) eingehe. Im dritten Teil erläutere ich schliesslich die Vorbereitung und Validierung der Texte (Kap. 4.2.4) sowie das methodisch-analytische Vorgehen mittels Topic-Modeling-Verfahren (Kap. 4.2.5).

4.2.1 Stellenanzeigen als Forschungsgegenstand

Stelleninserate sind in der Soziologie ein vergleichsweise wenig beachtetes empirisches Material (Geser 1983; Kriesi et al. 2010; Salvisberg 2010; Schreiber 1995). So enthält ein Standardwerk zur (deutschsprachigen) Arbeitsmarktsoziologie keinerlei Bezüge zu Stelleninseraten als Daten (Abraham & Hinz 2018). Dies ist umso erstaunlicher, da es sich um eine interessante Datenquelle handelt: Stelleninserate formulieren die unterschiedlichen Repräsentationen des Untersuchungsgegenstandes, etwa indem sie divergierende Tätigkeitsfelder, Kompetenzzuschreibungen oder Qualifikationsanforderungen nennen und einfordern. In solchen Konstruktionsleistungen verdichten sich berufsspezifische Rollenerwartungen (Geser 1983): Sie artikulieren feld- und organisationsspezifische Perspektiven darauf, worin bestimmte Praktiken und Expertisen bestehen und wo die Berührungspunkte sowie Differenzen zu verwandten Gebieten liegen.

Bei Stellenanzeigen handelt es sich um prozessgenerierte Daten, da sie im ordentlichen Handlungsfluss von Akteur*innen entstehen und entsprechend auf einem

»nicht-reaktiven Erhebungsprozess« basieren (Weischer 2015: 73). Sie folgen kulturellen und sprachlichen Konventionen, beispielsweise was die erwarteten Qualifikationsanforderungen oder Aussagen zur finanziellen Entlohnung betrifft (Geser 1983: 478), aber auch rechtlichen Regulierungen (Aratnam 2012). In der Regel weisen Stelleninserate einen stark standardisierten Aufbau auf: Sie enthalten Informationen über die inserierende Organisation (*wer sucht*), die Stelle und die zu erledigenden Aufgaben (*für was wird gesucht*), Anforderungen an Ausbildung, Erfahrung und Sachkenntnisse, zugeschriebene Eigenschaften und Fähigkeiten (»Kompetenzen«) oder Kriterien wie Alter und Geschlecht (*wer wird gesucht*) (Salvisberg 2008: 2567).⁵ Abschliessend wird neben Bewerbungsfristen und Kontaktmöglichkeiten bisweilen auf weitere Faktoren wie das Arbeitsumfeld und Chancengleichheits- bzw. Diversity-Statements (vor allem in englischsprachigen Stelleninseraten) verwiesen.

Die Bestandteile von Stelleninseraten sind oft mit unterschiedlich strukturierten Textformaten verknüpft: Während die organisationalen Selbstbeschreibungen sowie die Tätigkeitsprofile in der Regel prosaisch in integralen Sätzen verfasst sind (vgl. die Abschnitte [1]–[5] im Beispielinserat im Anhang), werden die Anforderungen und Qualifikationen oft in Listen unvollständiger Satzkonstruktionen oder nur stichwortartig formuliert (vgl. den Abschnitt [6] »Ihr Profil« im Beispielinserat im Anhang). Auf der Ebene des Vokabulars äussert sich dies in spezifischen Kombinationen von organisationsbezogenen, feld- sowie stelleninseratetypischen Begriffen.

Als Spezifika von Stelleninseraten in technologiebezogenen Wissensgebieten wie Informatik oder Datenwissenschaften gelten in der fach- und berufspädagogischen Diskussion divergierende Annahmen über die Qualifikationsanforderungen von Bewerber*innen (Bott et al. 2000; Schumann et al. 2016; Wowczko 2015). Was für die Akteur*innen im Feld ein Problem darstellen mag, weil immer unklar bleibt, über welche Kompetenzen nun eine Datenwissenschaftlerin verfügen muss, ist für die empirische Analyse hingegen bereits eine erste konstitutive Erkenntnis: Die Debatten über die notwendigen Skillsets oder Methoden verweisen auf die Suche nach den »richtigen Kompetenzen«. Sie sind gewissermassen fundierend für die Herausbildung des Phänomens an sich. Weiter manifestiert sich in Stellenanzeigen die eminente Bedeutung unterschiedlicher Werkzeuge wie der verwendeten Betriebssysteme, Software, Programmiersprachen und weiterer Instrumente: So ist im empirischen Material zu beobachten, dass Methoden, Tools oder Qualifikationsanforderungen oft als Synonyme verwendet werden (»Mehrjährige Erfahrung mit statistischer Software, R, Python, SAS oder SPSS«, vgl. Beispielinserat 1). Solche nicht-hierarchisierenden Aufzählungen deuten eine gewisse Beliebigkeit im Sinne »unverbindliche[r] Wunschlisten« an (Geser 1983: 479).

Die empirische Aussagekraft der Datenquelle Stellenanzeigen wird insofern eingeschränkt, dass sie nur die formale, nachfrageorientierte Seite des Arbeitsmarktes abbilden, während informelle Wege (wie Praktika, Abschlussarbeiten, persönliche

5 Salvisberg (2010: 116f.) unterscheidet folgende acht Analyseeinheiten: Selbstbeschreibung des inserierenden Unternehmens; Angaben zum Grund der Stellenausschreibung; Informationen zur administrativen Abwicklung der Bewerbung; ggf. Angaben zur Stellenvermittlerin; Hinweise auf materielle Gegenleistungen des Arbeitgebers; Stellen- und Aufgabenbeschreibung; Anforderungen an Ausbildung, Erfahrung und Sachkenntnisse; notwendige Charaktereigenschaften, Sozialkompetenzen oder Zuschreibungskriterien wie Alter oder Geschlecht.

Netzwerke etc.) oder Spontanbewerbungen auf der Angebotsseite nicht berücksichtigt werden.⁶ Auch organisationsinterne Arbeitsmärkte (Köhler & Krause 2010; Struck 2018) werden durch das gewählte Verfahren nicht erfasst. Nichtsdestotrotz bilden Stelleninserate aus den genannten Gründen einen interessanten Forschungsgegenstand, weil sie es ermöglichen, die multiplen Bedeutungen eines sozialen Phänomens über die unterschiedlichen sozialen Zuschreibungen, Qualifikationsanforderungen oder Tätigkeitsfelder zu rekonstruieren. Dies stellt insbesondere im Falle neuer Wissensgebiete bzw. solcher in Transformation eine Möglichkeit dar, prozessinduzierte Dynamiken jenseits subjektiver Stellungnahmen (wie Kommentare, Blogs, Tweets etc.) als auch reaktiver Erhebungsmethoden (wie Umfragen) zu untersuchen.

4.2.2 Datenerhebung

Die Stellenausschreibungen für Data Scientists und verwandte Bezeichnungen (wie Data Analyst, Data Engineer etc.) bilden die Datengrundlage für die Analyse im Kapitel 5. Um das Korpus zu bilden, habe ich auf der Schweizer Stellenplattform jobs.ch veröffentlichte Stellenbeschreibungen, die die Suchbegriffe »Data AND Scientist«, »Data AND Science« oder »Big AND Data« enthalten, mittels Webscraping-Verfahren (Munzert et al. 2015) in der Statistiksoftware R erfasst und als Textdateien gespeichert. Zwischen Januar 2017 und Juni 2019 konnte ich ein Sample von $N = 4341$ Stellenanzeigen zusammenstellen.⁷ Die Suchbegriffe wurden bewusst kombiniert, um einerseits die Breite des Phänomens zu adressieren und andererseits keine Engführung auf bestimmte mit den Begriffen verknüpfte Orientierungen (beispielsweise »Big Data« als informatisches gegenüber »Data Science« als statistisches Konzept) zu riskieren.

Während des Erhebungsprozesses fiel auf, dass diverse inserierte Stellen über Monate, manchmal sogar über mehrere Jahre hinweg verfügbar blieben. Oft wurden die Inserate mehrmals leicht adaptiert und unter verändertem Berufstitel oder mit einem veränderten Anforderungsprofil erneut publiziert.⁸ Dies verweist auf andauernde Schwierigkeiten in der Rekrutierung von Datenwissenschaftler*innen (Burch 2016; H. D. Harris et al. 2013; Markow et al. 2017) sowie deren Konstruktion als stark nachgefragte Berufsgruppe.

Das Sample habe ich fortlaufend sowohl manuell als auch mittels automatisierter Texterkennung auf Duplikate überprüft, um Mehrfachzählungen derselben Inserate zu vermeiden. Dabei habe ich folgende Regel festgelegt: Verwenden zwei oder meh-

6 Aufgrund der sehr hohen Nachfrage und der Vorbereitung der Studierenden auf informelle Kontaktkanäle kommt diesen für das untersuchte Arbeitsmarktsegment eine gewisse Bedeutung zu (vgl. Gerlach 2014: 145 für das ähnlich gelagerte Beispiel von Automobilingenieur*innen).

7 Das Sample stellt keine Vollerhebung dar; es handelt sich um einen selektiven Ausschnitt aus dem Total aller Stellenanzeigen. Nimmt man den »Schweizer Jobradar« der Firma x28 AG, der in Anspruch nimmt, ein »vollständiges und repräsentatives Bild der Nachfrage nach Arbeitskräften« in der Schweiz zu liefern, als Vergleichsmass, lässt sich näherungsweise festhalten, dass das Sample durchschnittlich rund 2.4 % der pro Quartal erfassten Stellenanzeigen entspricht (eigene Berechnungen). Vgl. online: <https://www.x28.ch/jobradar/> (Zugriff: 03.02.2022).

8 Zudem ergab sich die Schwierigkeit, dass der HTML-Code der Webseite mehrmals geändert wurde. Dies führte dazu, dass unterschiedliche Stellenanzeigen in die Suchanfrage inkludiert wurden. Eine Änderung des HTML-Codes ist beispielsweise für Oktober 2017 dokumentiert, was sich in einem deutlich erkennbaren Anstieg der Anzahl erfasster Anzeigen niederschlägt (vgl. Abbildung 3).

rere Anzeigen einen identischen Stellentitel und Ausschreibungstext, aber ein unterschiedliches Datum, so werden diese aus dem Sample entfernt. Enthalten zwei Anzeigen einen identischen Titel, aber unterschiedliche Stellentexte (beispielsweise indem das Anforderungsprofil adaptiert wurde), oder umgekehrt variierende Stellenbezeichnungen, aber identische Stellentexte, so sind beide im Sample enthalten.⁹

4.2.3 Kodierung und Datenbereinigung

Eine zentrale Herausforderung von Daten, die durch soziotechnische Systeme (wie Social-Media-Plattformen und andere internetbasierte Dienste) prozessual generiert werden, ist, dass sie inhärent »messy« sind.¹⁰ Für den ersten Erhebungs- und Auswertungsschritt der Arbeit ergab sich insofern das Problem, dass parallel zur Datenerhebung mehrere Teilschritte der Kodierung und Datenbereinigung notwendig wurden, um die Daten zu vereinheitlichen und so für die Analyse verfügbar zu machen.¹¹ Aufgrund der hohen Anzahl Stellenanzeigen im Sample sollte dies automatisiert erfolgen. So liessen sich aus den erfassten Textdateien automatisiert diverse Informationen extrahieren, die für die deskriptive Analyse relevant sind: Stellenbezeichnung, Datum, inserierende Organisation (Arbeitgeber), lokale Standorte, die URL sowie der Text der Anzeige. Die Attribute wurden als einzelne Variablen kodiert.¹² Deren Häufigkeiten und Verteilungen bilden die Grundlage für die deskriptive Analyse des Samples.

Um Informationen über die Akteur*innen der Datenwissenschaften in den einzelnen Feldern zu erhalten, mussten die Namen der Organisationen, die Stellenanzeigen publizieren, vereinheitlicht werden (Gross- und Kleinschreibung, mit oder ohne rechtsformindizierende Angaben wie AG, GmbH etc.). Anschliessend wurden diese nach ökonomischen bzw. anderen sozialen Feldern kodiert. Als Grundlage dazu dienten die Selbstbeschreibungen der Firmen auf ihren Webseiten. Die Kodierung der Unternehmen und Organisationen in ökonomische und andere Felder entspricht damit nicht wirtschaftsstatistischen Klassifikationen des Bundesamts für Statistik (BFS), sondern hat sich während des Kodierprozesses induktiv aus dem Material gebildet.¹³ Die inserierenden Organisationen im Datensatz entstammen achtzehn unterschiedlichen Feldern.

9 Ein Beispiel: Die Firma Helbling AG publizierte ein Inserat für eine/n »Data Scientist (m/w) – Schwerpunkt Signalverarbeitung/Algorithmen für optische Messsysteme«, datiert auf den 11. Februar 2019. Dieses wurde rund einen Monat später, am 15. März 2019, unter der Bezeichnung »Entwicklungsingenieur (m/w) – Schwerpunkt Signalverarbeitung/Optik« über dieselbe URL erneut erfasst. Beide Inserate bleiben im Sample enthalten.

10 Der Umgang damit ist ein zentraler Aspekt datenwissenschaftlicher Praxis (Mützel et al. 2018: 114ff.).

11 Die bereinigten und kodierten Daten sind auf dem Datenrepositorium FORSbase hinterlegt (Saner & Mützel 2021a).

12 Zusätzlich wurde jeder Anzeige eine eigene Identifikationsnummer zugewiesen.

13 Bei der ersten Kodierung hatten Unternehmen im Feld der Personaldienstleistungen den höchsten Anteil. Eine detaillierte Betrachtung der Inserate zeigte, dass diese nur in der Minderheit der Fälle (<5 %) für sich selbst rekrutieren, sondern dies überwiegend im Auftrag anderer Firmen tun (indiziert durch Begriffe wie *client*, *customer*, *Kunde/Kundin* oder *Auftraggeber*in* im Inseratetext). Bei diesen handelte es sich mehrheitlich um die bedeutenden Felder der Pharmaindustrie sowie der Banken/FinTech (erschieden in je rund 34 % der Inserate). Auch Felder wie IT/Software (19 %), Industrie (13 %) oder Versicherungen (7 %) wurden oft genannt. In der Folge wurden alle Inserate von Personalrek-

Die Stelleninserate sind in vier verschiedenen Sprachen (Englisch, Deutsch, Französisch und Italienisch) publiziert. Computerlinguistische Verfahren der automatisierten Spracherkennung ermöglichen es hier, den Anzeigentext eindeutig einer der vier Sprachen zuzuordnen, was Analysen über die sprachräumliche Verteilung der Datenwissenschaften in der Schweiz erlaubt.¹⁴ Zusätzlich stellt die Sprache einer Stellenanzeige einen Indikator für die Internationalität einer Organisation bzw. eines Feldes dar.

Auch die Standorte der inserierenden Unternehmen und Organisationen wurden vereinheitlicht (Entfernung von Doppelnamen wie Zürich-Oerlikon oder englischen Städtenamen [Geneva] in eindeutige Standortangaben), um diese eindeutig geographisch lokalisieren zu können. Anschliessend erfolgte die Einteilung der Standorte der Unternehmen und Organisationen in die sieben geographischen Grossregionen der Schweiz (Genferseeregion, Espace Mittelland, Nordwestschweiz, Zürich, Ostschweiz, Zentralschweiz sowie das Tessin) gemäss BFS.¹⁵ Waren zwei oder mehr Standorte genannt, die sich über mehrere der genannten Grossregionen hinweg verteilen, wurde die Region als »Mehrfachnennung« kodiert. Rund ein Prozent der Inserate liess sich keinem eindeutigen Standort zuweisen oder dieser liegt ausserhalb der Schweiz.

4.2.4 Vorbereitung und Validierung der Daten für Text Mining

Die beschriebenen Informationen über die feldspezifische, geographische und sprachliche Verortung der Stelleninserate geben Hinweise über die Struktur und Entwicklung der Datenwissenschaften im schweizerischen Arbeitsmarkt. Das Hauptkenntnisinteresse der Untersuchung gilt allerdings den Ausschreibungstexten der Stelleninserate: In den Texten werden sehr unterschiedliche Tätigkeitsfelder, Kompetenzzuschreibungen oder Qualifikationsanforderungen entworfen, d. h. multiple Repräsentationen der Datenwissenschaften formuliert. Aufgrund ihrer »messyness« infolge des Webscraping habe ich die Texte zunächst manuell und anschliessend automatisiert überprüft und bereinigt, um eine möglichst hohe Datenqualität zu garantieren und ihre Interpretierbarkeit in textanalytischen Verfahren zu gewährleisten (Maier et al. 2018: 100f.).¹⁶ Dies beinhaltete mehrere Schritte der Datenvorbereitung mittels des *Quanteda Package* in R (Benoit et al. 2018).¹⁷

rutierungsfirmen, die sich über den Text eindeutig einem bestimmten ökonomischen Feld zuordnen liessen, manuell rekodiert.

- 14 Das *textcat Package* in R (Hornik et al. 2013) verwendet dazu eine n-gram-basierte Textkategorisierung (Jurafsky & Martin 2018): Texte werden in einzelne Bestandteile zerlegt und sodann mittels Häufigkeitsanalysen der jeweiligen Sprache zugeordnet. Die Ergebnisse dieses Verfahrens wurden regelmässig auf ihre Validität geprüft.
- 15 Vgl. Bundesamt für Statistik (2020): *Analyseregionen*: <https://www.bfs.admin.ch/bfs/de/home/statistiken/querschnittsthemen/raeumliche-analysen/raeumliche-gliederungen/analyseregionen.html> (Zugriff: 03.02.2022).
- 16 Die Länge der in der Analyse verwendeten Dokumente weist einen Durchschnitt von 530 und einen Median von 488 Wörtern auf, wobei die englischsprachigen Dokumente (Durchschnitt = 581 Wörter, Median = 547 Wörter) etwas länger sind als die deutschsprachigen (Durchschnitt = 414 Wörter, Median = 381 Wörter).
- 17 Zu den Vorbereitungsschritten gehörten die Transformation aller Token in Kleinbuchstaben, das Entfernen von HTML-Tags, Satzzeichen sowie sprachenspezifischer Stoppwörter. Zusätzlich entfer-

Nach diesen Vorbereitungsschritten erstellte ich je eine Dokument-Feature-Matrix (DFM) für die englisch- sowie die deutschsprachigen Stelleninserate, wobei jedes Dokument einen eigenen Vektor darstellt. Die Texte werden in ihre einzelnen Bestandteile (Tokenisierung) zerlegt, d. h., die tatsächliche Reihenfolge der Wörter und damit auch deren syntaktische Struktur werden ignoriert. Dadurch wird auch der Umstand neutralisiert, dass es sich bei Stellenanzeigen um ein Textgenre handelt, das integrale Satzbestandteile mit Wortauflistungen kombiniert. Das Entfernen seltener Begriffe, die in weniger als fünf unterschiedlichen Dokumenten im Korpus erscheinen, reduziert die Dimensionalität der Matrizen um rund drei Viertel (-75.6 % im englischsprachigen bzw. -78.8 % im deutschsprachigen Korpus). Die resultierende englischsprachige DFM enthält 2941 Dokumente mit 5771 Features, die deutschsprachige DFM 1282 Dokumente mit 4671 Features.

Solche Schritte der Textvorbereitung machen den Datensatz schlanker und somit für die Analyse leichter handhabbar, reduzieren jedoch gleichzeitig auch die Menge an Informationen, die im Datensatz enthalten sind. Entsprechend muss jeder einzelne Schritt der Textvorbereitung sorgfältig überprüft und validiert werden (Denny & Spirling 2018; Maier et al. 2018). Insbesondere das Stemming (Stammformreduktion) von Wörtern, d. h. das regelbasierte Kürzen der Endungen von Wörtern auf ihren Stamm, kann zu sehr unterschiedlichen Ergebnissen beim Topic Modeling führen (Blei 2012; Denny & Spirling 2018; Maier et al. 2018). Denny und Spirling (2018) haben mit preText ein Tool entwickelt, mit dem die Effekte der einzelnen Textvorbereitungsschritte insbesondere für nicht-überwachte Formen maschinellen Lernens (*unsupervised learning*) evaluiert werden können. Je tiefer der resultierende Wert, desto weniger beeinflusst die Kombination der gewählten Textvorbereitungsschritte das Korpus. Trotz der Information, inwiefern ein bestimmtes Modell das Korpus beeinflusst, gibt der Wert in diesem Sinne keine finale Antwort für ein zu wählendes Modell, sondern schafft die Möglichkeit, die gewählte Variante im Vergleich zu anderen Modellen zu evaluieren. Die gewählten Varianten¹⁸ erwiesen sich im Vergleich mit anderen evaluierten Modellen als robust und in der Analyse mittels Topic Modeling auch als tatsächlich interpretierbar.¹⁹

te ich Begriffe, die spezifisch für die erhobenen Stelleninserate sind (*jetzt, bewerben, Originalinserat, anzeigen*) oder solche, die aufgrund des Aufbaus und der Struktur der Webseite jobs.ch in die Datenerhebung einfließen (*bitte, beziehe, dich, bei, deiner, Bewerbung, auf, jobs.ch*), da sie für die inhaltliche Analyse keinen Informationswert besitzen, umgekehrt allerdings die Ergebnisse des Topic-Modeling-Verfahrens verzerren können (vgl. unten).

- 18 Die gewählte Variante mit dem Akronym P(unctuation)-N(umbers)-L(owercase)-S(tem)-W(Stopword Removal) weist für das englischsprachige Korpus mit preText = 0.132 den neuntniedrigsten Wert aller 128 möglichen Kombinationen auf (die Spannweite ist bei $r = 0.225$). Der Wert für das deutschsprachige Sample ist mit preText = 0.127 etwas niedriger. Dieser liegt zwar nur im Mittelfeld aller möglichen Kombinationen, wobei die Differenz zum ›besten‹ Wert lediglich 0.027, gegenüber dem ›schlechtesten‹ Wert jedoch 0.179 beträgt (die Verteilung der Werte ist stark linksschief).
- 19 Dies habe ich exemplarisch anhand der drei Topic-Modelle mit den niedrigsten preText-Werten für das englisch- und deutschsprachige Sample überprüft: Die Ergebnisse waren unter anderem aufgrund der Kombination von Bi- und Trigrammen kaum interpretierbar, da keine Stopwörter entfernt wurden.

4.2.5 Analytisches Vorgehen mittels Topic Modeling

Nun liegen die Textdateien als bereinigte DFM vor und können mittels textanalytischer Methoden (Evans & Aceves 2016; Ignatow & Radev 2016) untersucht werden. Die Analyse grosser Mengen an Textdaten erfährt gegenwärtig unter der Losung *Text as Data* (Grimmer & Stewart 2013) eine erneuerte Konjunktur in den Sozialwissenschaften. In einer kultursoziologischen Perspektive erweist sich insbesondere Topic Modeling als vielversprechendes Verfahren, um grosse Textdaten auf latente Muster und Strukturen hin zu untersuchen und mit der Analyse sozialer bzw. kultureller Felder zu verknüpfen (Bail 2014; DiMaggio et al. 2013; Mohr & Bogdanov 2013; Mützel 2015b; Papilloud & Hinneburg 2018). Topic Modeling umfasst computerlinguistische probabilistische Verfahren, um Textdokumente statistisch zu annotieren und somit ein Korpus dimensionenreduzierend zu strukturieren.

Bei Topic Modeling werden »aufgrund einer bestimmten Verteilungsannahme Themen (topics) aus Texten generiert. Topic Modeling Algorithmen analysieren Wörter aus einem Textkorpus, um darin enthaltene Themen zu identifizieren. Dies ermöglicht dann, Themen zueinander in Beziehung zu setzen und im zeitlichen Verlauf abzubilden« (Mützel 2015b: 411). Ein prominenter Algorithmus ist die »Latent Dirichlet Allocation« (LDA) (Blei et al. 2003): Das Topic-Modell verwendet Ko-Okkurrenzen von Wörtern, um latente Themen in den Dokumenten zu eruieren. Die Dokumente werden nicht als einem Thema alleine zugehörig betrachtet, sondern können verschiedenen Themen angehören. Gleichzeitig variieren die Verhältnisse der Themen über die Dokumente. So setzt sich ein bestimmtes Dokument beispielsweise zu 62 % aus Topic 1, zu 23 % aus Topic 3 und zu 15 % aus einer Mischung der restlichen Topics zusammen. Die Zahl der Topics muss zu Beginn manuell definiert werden. Ausgehend von dieser Definition berechnet LDA die Topics, die aus unterschiedlichen Themenvokabularen bestehen. Während also den Dokumenten Topics zugeordnet werden, zeichnen sich Topics durch eine charakteristische Begriffssammlung aus.

Der Nutzen einer Modellierung des Textkorpus mittels LDA liegt darin, dass die Multidimensionalität von Begriffen einer grossen Anzahl von Dokumenten automatisiert untersucht werden kann (DiMaggio et al. 2013): Begriffe wie »Daten«, »Methoden« oder »Wissenschaft« verfügen je nach Kontext über multiple Bedeutungen. Ein bestimmter Begriff kann in Kombination mit weiteren Begriffen eines Topics eine andere inhaltliche Bedeutung aufweisen wie derselbe Begriff in einem anderen Topic. Topic-Modelle sind in der Lage, die vielfältigen Konnotationen in unterschiedlichen Topics zu identifizieren und zu rekonstruieren. Dies erlaubt es, die multiplen Dimensionen und Bedeutungsebenen des Untersuchungsgegenstandes Datenwissenschaften zu analysieren.

Die latente Struktur der Topics eines Korpus berechne ich mittels Gibbs-Sampling (Griffiths & Steyvers 2004: 5229f.) und 2000 Iterationen: Dabei wird nach einer zufälligen Initialisierung die Topic-Zuordnung jedes einzelnen Wortes in allen Dokumenten in einem iterativen Verfahren wiederholt vorgenommen. Die Neuordnung hängt einerseits von der Dominanz eines Topics innerhalb des Dokuments ab und andererseits von der Wahrscheinlichkeit eines Wortes, zu einem bestimmten Topic zu gehören (Steyvers & Griffiths 2007: 7f.). Der Prozess wird 2000-mal repetiert und mit jeder Iteration verändert sich die Berechnung der Dokumentaufteilung bzw. der Topics, damit die Fehler der Rekonstruktion von Worthäufigkeiten in den Original-

dokumenten sinken. Ziel ist eine Balance der Homogenität zwischen Dokumenten – alle Wörter eines Dokuments werden dem gleichen Topic zugeordnet – und Topics – einige Wörter eines Topics haben eine hohe Wahrscheinlichkeit. Die Auflösung der widersprüchlichen Kriterien produziert für jedes Topic Schlüsselbegriffe, die häufig kookkurrent auftreten und damit eine semantische Interpretierbarkeit aufweisen. Aus der Anwendung von LDA geht schliesslich eine Topic-Struktur hervor, die gemäss den Modellannahmen das vorliegende Textkorpus statistisch am besten beschreibt. Topic Modeling mit LDA kombiniert somit eine induktive, *struktursuchende* Vorgehensweise (Baumer et al. 2017) mit statistischen Messverfahren und begünstigt insbesondere explorative Analysen. Das Verfahren eignet sich folglich gut, um ein Wissensgebiet, dessen Konturen noch nicht endgültig umrissen sind, in seiner Heterogenität und Vielschichtigkeit zu erforschen.

Ich berechne die Topic-Modelle sprachgetrennt, weil gemeinsame Modelle aufgrund der Ungleichverteilung (~68 % englisch, ~30 % deutsch) dazu tendieren, die deutschsprachigen Begriffe in einzelnen separaten Topics zu repräsentieren. Dies steht allerdings der Absicht entgegen, die semantischen Strukturen der Dokumente sprachunabhängig zu analysieren. Eine separate Analyse bietet zudem die Möglichkeit, den Einfluss des ökonomischen Feldes auf das Modell zu berücksichtigen, da im englischsprachigen Sample Pharmaunternehmen sehr häufig sind, während sie im deutschsprachigen Sample kaum auftreten.

Die Modellierung der Topic-Struktur mithilfe der LDA erfolgt mit dem Softwarepaket *Topicmodels* in R (Grün & Hornik 2011). Um ein geeignetes Topic-Modell zu bestimmen, habe ich Dutzende von Durchläufen mit jeweils unterschiedlicher Anzahl Topics berechnet und dabei auch quantitative Metriken berücksichtigt. Ein datenbasiertes Vorgehen nimmt meist die *Perplexity*- und *Log-Likelihood*-Werte eines Topic-Modells als Grundlage für die Festlegung der Anzahl Topics (Grün & Hornik 2011: 7; Maier et al. 2018: 99).²⁰ Die *Perplexity* ist eine Metrik, um die statistische Güte (*goodness-of-fit*) eines Modells zu bestimmen. Dabei wird geschätzt, wie gut ein Modell, das für einen grösseren Teil des Korpus berechnet wurde, einen kleineren Teil der Dokumente vorhersagt (Maier et al. 2018: 99). Die logarithmierte *Likelihood* bewertet, »wie gut ein Set von Parametern (Koeffizienten) geeignet ist, um ein Modell (z. B. eine logistische Regressionsfunktion) an vorhandene Daten anzupassen« (Diaz-Bone & Weischer 2015: 245). Ein höherer *Likelihood*-Wert korrespondiert dabei mit tieferen *Perplexity*-Werten (Maier et al. 2018: 102).

Für das englischsprachige Korpus ergibt die *Perplexity* eine optimale Anzahl bei $K = 100$ Topics, für die *Log-Likelihood* bei $K = 75$ Topics. Beim deutschsprachigen Sample liegen die entsprechenden Werte bei $K = 90$ sowie $K = 40$ Topics. Da die Struktur von Modellen mit dieser Granularität nicht sinnvoll interpretierbar ist, reduziere ich in nachfolgenden Berechnungen die Anzahl Topics, bis die Modelle inhaltlich valide beschrieben werden können. Dazu dürfen sie nicht zu weit gefasst sein, da sich sonst viele Überschneidungen von Topics ergeben, aber auch nicht zu eng, da sonst viele inhalt-

20 Das R Package *LDAtuning* umfasst vier unterschiedliche Metriken bzw. Informationskriterien (»Griffiths2004«, »CaoJuan2009«, »Arun2010«, »Deveaud2014«), die für die Festlegung der Anzahl Topics herangezogen werden können. Für das englischsprachige Korpus ergaben die Metriken eine optimale Anzahl Topics zwischen $K = 40$ und $K = 100$, für das deutschsprachige Sample lagen die entsprechenden Werte zwischen $K = 10$ und $K = 40$.

liche Aspekte zusammengefasst werden (Maier et al. 2018: 98). Für die semantische Interpretation der Topics ziehe ich sowohl die statistischen Auftretenswahrscheinlichkeiten der häufigsten Begriffe als auch Dokumente, die durch die jeweiligen Topics gut beschrieben werden, heran. Die resultierenden Modelle, bei denen die häufigsten Begriffe der Topics inhaltlich valide beschrieben werden können, umfassen $K = 20$ Topics für das englischsprachige und $K = 11$ Topics für das deutschsprachige Korpus (vgl. Tabellen 5 und 6).

Wie die Festlegung der Anzahl Topics zeigt, spielen qualitative Interpretationsleistungen trotz der Quantifizierung der Textdokumente bzw. ihrer einzelnen Bestandteile eine zentrale Rolle. Die Interpretation der Topics basiert zwar teilweise auf quantitativen Kennzahlen, lässt sich jedoch als solche nicht berechnen (Roberts et al. 2019: 9ff.). Im Sinne eines relationalen Verfahrens erhalten die Topics ihre Sinnhaftigkeit nur im Verhältnis zu anderen Topics und werden erst dadurch interpretierbar. Insofern äussert sich darin eine Anschlussfähigkeit von Topic Modeling an kultursoziologische Ansätze, die die Relationalität von Sinnstrukturen untersuchen (Kirchner & Mohr 2010; Mohr 2000, 2013; Mohr & Bogdanov 2013).

4.3 Strategiedokumente im Feld der Politik

Akteur*innen im Feld der Politik prägen die Herausbildung des Gegenstandes Datenwissenschaften primär durch unterschiedliche Strategiepapiere: Sie entwerfen im Rahmen des Diskurses über die Digitalisierung Szenarien einer vielversprechenden Zukunft und verknüpfen diese mit bildungs- und forschungspolitischen Fördermassnahmen. Politische Strategiepapiere können als kollektive Stellungnahmen und Kompromisse verschiedener, koexistierender »Flügel«, Überzeugungen und Weltanschauungen von organisationalen Akteur*innen als sozialen Feldern betrachtet werden. Einerseits sind sie »Resultat einer sozialen Praxis, die von positionsabhängigen Akteursstrategien geprägt ist und die sich wiederum als Teil eines Ensembles von Strategien verstehen lässt, die als Ganzes von einem feldspezifischen Machtkonflikt um Positionen, Ressourcen und Anerkennung getrieben sind, der sich um die Definition von politischen Kategorien dreht« (Bernhard & Schmidt-Wellenburg 2012: 48f.). Andererseits eröffnen die Stellungnahmen, die durch Vielstimmigkeit und Multiperspektivität gekennzeichnet sind, auch Möglichkeiten zur Kooperation mit anderen Akteur*innen. Sowohl durch stärker konflikthafte als auch durch kooperative Praktiken tragen sie zur Konstitution und Permanenz der Datenwissenschaften als neues Wissensgebiet bei.

Ich rekonstruiere die Zukunftsentwürfe durch eine qualitative Inhaltsanalyse und stelle die Verknüpfung der hochschulpolitischen Stellungnahmen zu anderen Feldern, insbesondere der Wissenschaft und der Ökonomie her. Zunächst lege ich dar, wie ich das Korpus der politischen Strategiepapiere zusammengestellt habe (Kap. 4.3.1). Anschliessend erläutere ich das inhaltsanalytische Vorgehen und den Prozess der Kodierung des Materials (Kap. 4.3.2).

4.3.1 Das Korpus der Dokumente

Das Sample umfasst Dokumente und Strategiepapiere von Akteur*innen im Feld der schweizerischen Hochschul- und Forschungspolitik, die sich mit der Digitalisierung in Bildung und Forschung im weiteren Sinne und mit der Herausbildung der Datenwissenschaften als neuem Wissensgebiet im engeren Sinne beschäftigen. Ausgangspunkt für die Zusammenstellung des Samples waren die Dokumente zur Strategie »Digitale Schweiz« des Bundesrates. Über diese konnte ich 34 Strategiepapiere und weitere Dokumente erschliessen, die sich ebenfalls ganz oder teilweise dem genannten Themenfeld widmen. Zu den Akteur*innen gehören sowohl politische Institutionen wie der Bundesrat, das Staatssekretariat für Bildung, Forschung und Innovation (SBFI), das Staatssekretariat für Wirtschaft (SECO), das Bundesamt für Kommunikation (BAKOM), der ETH-Rat, die Konferenz der kantonalen Erziehungsdirektorinnen und -direktoren (EDK) sowie wissenschaftliche Kommissionen, Unternehmensverbände und Think Tanks (vgl. die Liste der untersuchten Dokumente in Tabelle 11 im Anhang).²¹

Der Zeitraum der untersuchten Dokumente reicht von 1998 bis 2020, wobei die überwiegende Mehrheit davon nach 2014 publiziert wurde. Da die Dokumente ab 2014 vermehrt »Big Data« und »Data Science« sowie deren rechtliche, ökonomische und edukative Aspekte in der Bildungs- und Forschungspolitik adressieren, bildet dieser Zeitraum den Schwerpunkt der Analyse. Die älteren Strategiedokumente zur Informationsgesellschaft Schweiz wurden berücksichtigt, um Kontinuitäten und Brüche in den bundesrätlichen Strategien zu untersuchen.

Einen zentralen Bestandteil des politischen Diskurses zur Digitalisierung in der Schweiz bilden die bundesrätlichen Strategiedokumente, Aktionspläne und Berichte zur »Informationsgesellschaft Schweiz« bzw. zur »Digitale[n] Schweiz« (vgl. Tabelle 12 im Anhang): Ende der 1990er-Jahre gab sich der Bundesrat zum ersten Mal eine Strategie für den Umgang mit den »neuen Informations- und Kommunikationstechniken«. Die »Strategie Informationsgesellschaft« wurde im Februar 1998 (im Folgenden: SIG 1998) verabschiedet und 2006 (im Folgenden: SIG 2006) sowie 2012 (im Folgenden: SIG 2012) überarbeitet, wobei Ziele, Grundsätze und Politikbereiche fortlaufend adaptiert wurden (Abun-Nasr 2009; Ciarla 2018). Die im April 2016 lancierte Strategie »Digitale Schweiz« (im Folgenden: SDS 2016) des Bundesrates löst die vorherige »Informationsgesellschaft Schweiz« ab, die im September 2018 (im Folgenden: SDS 2018) revidiert wurde.

Die fünf Strategiedokumente der »Informationsgesellschaft Schweiz« und »Digitale Schweiz« bilden ein Subsample innerhalb des Korpus. Neben ihrer gemeinsamen Autorschaft (Bundesrat) formulieren sie politische Strategien über das Verhältnis von Technologie und Gesellschaft und eignen sich daher besonders für eine vertiefende inhaltliche Analyse über den betrachteten Zeitraum hinweg. Während andere Dokumente im Korpus in erster Linie bildungs- und forschungspolitische Akteur*innen adressieren, richten sich die bundesrätlichen Strategien an ein breiteres, hybrides Publikum.

21 Die untersuchten Dokumente, das Kodierschema sowie das Kodebuch sind auf dem Datenrepositorium FORSbase hinterlegt (Saner & Mützel 2021b).

Formal zu beobachten ist neben den kürzer werdenden Erscheinungszyklen die kontinuierliche Ausweitung des inhaltlichen Umfangs der bundesrätlichen Strategie-papiere. Die Strategie »Digitale Schweiz« 2018 ist fünfmal umfangreicher als die erste Strategie »Informationsgesellschaft Schweiz« von 1998. Der Vergleich bestimmter Wortfrequenzen muss deshalb in Relation zum stark anwachsenden Textumfang der Strategiedokumente gesetzt werden (vgl. Tabelle 12 im Anhang).

4.3.2 Das inhaltsanalytische Vorgehen

Alle Dokumente lagen als Textdateien vor und wurden in der Analysesoftware ATLAS.ti erfasst. Ich habe ausschliesslich jene Textstellen kodiert, die einen expliziten Bezug zu den Bereichen Bildung und Forschung haben. Daraus resultierte ein reduziertes Textsample von rund 350 A4-Seiten. Ich habe die für die Fragestellung relevanten Textstellen in einem offenen, induktiven Kodierungsprozess (Flick 2016: 388ff.) mit zusammenfassenden bzw. erklärenden Codes versehen (Friese 2012: 92ff.). Dabei entstand in einem ersten Durchgang ein Kategoriensystem, das ich in insgesamt drei Durchgängen überprüft und überarbeitet habe. Durch das mehrmalige Lesen und Kodieren der Textstellen konnte ich das Kategoriensystem erweitern und ausdifferenzieren (vgl. das Kodierschema in Saner & Mützel 2021b). Ziel war dabei neben der inhaltlichen Strukturierung des Materials das Eruiieren der relevanten Themen und Dimensionen im Korpus der politischen Dokumente.

Die Kodierung der Texte in drei Durchläufen ergab 75 unterschiedliche Codes, die insgesamt 1144 Zitate umfassen.²² Bei Passagen, die mit mehr als einem Code versehen sind, handelt es sich um inhaltlich dichte Textstellen mit unterschiedlichen thematischen Bezügen. Ein Beispiel für ein inhaltlich äusserst dichtes Zitat lautet wie folgt:

»Um die mit dem Strukturwandel verbundenen Chancen zu nutzen und Herausforderungen erfolgreich bewältigen zu können, müssen diese bereichsübergreifend sowohl national als auch international vernetzt angegangen werden.« (SDS 2018: 4)

Der erste Teilsatz bezieht sich auf den »Strukturwandel«, der auf die Digitalisierung als gesellschaftlicher Transformationsprozess (1. Kode) verweist, wodurch sich sowohl »Chancen« (2. Kode) als auch »Herausforderungen« (3. Kode) eröffnen – dies bildet, wie ich noch zeigen werde, ein zentrales Element im politischen Diskurs über Digitalisierung von Bildung und Forschung. Der zweite Teilsatz bezieht sich auf die Strategien, um die »Chancen zu nutzen und Herausforderungen erfolgreich zu bewältigen«, wobei hier Interdisziplinarität (4. Kode) sowie Kooperation und Vernetzung (5. Kode) erwähnt werden.

Die Kodierung der Zitate erfolgte in erster Linie über bestimmte Begriffe. Beispielsweise umfassen die Textstellen, die inhaltlich mit »Digitalisierung als gesellschaftlicher Transformationsprozess« kodiert werden, sehr oft prozessindizierende Begriffe wie Automatisierung, Modernisierung, (Struktur-)Wandel, Veränderung oder Entwicklung. Mittels einfacher Häufigkeitszählung können diese Begriffe in den einzelnen Dokumenten aggregiert ausgewertet werden. Dies erlaubt es, Entwick-

22 Zwei Drittel der zitierten Stellen sind einmalig kodiert (66.7 %), die restlichen Zitate sind zweimalig (25.7 %), dreimalig (6.6 %) oder häufiger (1 %) kodiert.

lungen bestimmter zentraler Kategorien im politischen Diskurs über den Zeitverlauf hinweg zu analysieren. Dies ist insbesondere für das Subkorpus der fünf Strategiedokumente interessant, da diese über zwei Jahrzehnte hinweg soziotechnische Zukunftsszenarien für die Schweiz formulieren.

Zudem lassen sich die identifizierten Codes in neun übergreifenden Dimensionen zusammenfassen (vgl. Tabelle 8 im Anhang). Am häufigsten sind Zitate in der politischen Dimension zu verorten, was auf die Verortung der Akteur*innen im politischen Feld verweist, gefolgt von der wissenschaftlichen Dimension, was die thematische Orientierung der untersuchten Dokumente indiziert. An dritter Stelle folgt die Datendimension, die das Erhebungskriterium für die Auswahl des Korpus darstellt. Ferner sind auch die organisationale, technische, ökonomische sowie edukative Dimension in den Dokumenten bedeutend. Etwas weniger prävalent sind die internationale Dimension sowie die Kompetenzen, die aber dennoch in der grossen Mehrheit der Dokumente enthalten sind.

Knapp die Hälfte der Codes (46 %) bildet eine Dimension ab. Die übrigen Codes sind in zwei (39 %) oder mehr (15 %) Dimensionen enthalten. Die Multidimensionalität vieler Codes bildet das relationale Moment der Analyse, indem Codes zwei oder mehrere Dimensionen miteinander verknüpfen. So kann der Code »Standortwettbewerb« sowohl in der ökonomischen, wissenschaftlichen als auch internationalen Dimension massgebend sein. Um dies anhand eines Zitats aus der aktuellen Strategie »Digitale Schweiz« zu veranschaulichen:

»Um den Spitzenplatz der Schweiz als Innovations- und Forschungsstandort zu halten, müssen die Forschungskompetenzen bezüglich digitaler Technologien in ihrer ganzen Breite gestärkt und der Wissenstransfer in die Wirtschaft beschleunigt werden.« (SDS 2018: 5)

Über die internationale Dimension, den »Spitzenplatz der Schweiz« in einem nicht weiter bezeichneten Vergleich, wird die wissenschaftliche mit der ökonomischen Dimension verknüpft (»Innovations- und Forschungsstandort«, »Wissenstransfer in die Wirtschaft beschleunigen«). Durch die Einbettung und den Kontext der identifizierten Textstellen sowie die Beziehungen zwischen den verschiedenen Dimensionen ergibt sich damit die Bedeutung bzw. Sinnhaftigkeit eines Konzeptes wie »Standortwettbewerb«.

4.4 Curricula und Interviews im akademischen Feld

Die Curricula aller verfügbaren Studienangebote in Data Science an Schweizer Universitäten und Hochschulen sowie Interviews mit Lehrenden bilden das empirische Material für das akademische Feld, das ich mittels qualitativer Inhaltsanalysen untersuche. Während die Curricula im Falle der Datenwissenschaften meist Produkte grenzüberschreitender disziplinärer Kooperation darstellen, treten in den Interviews die epistemologischen und disziplinären Perspektiven auf den Gegenstand deutlicher zutage. Die Kombination der Materialien erlaubt es, die Widersprüche und Dynamiken synchroner Praktiken von Grenzziehung und Grenzüberschreitung

in der Herausbildung der Datenwissenschaften als einem Möglichkeitsraum zu adressieren.

Zunächst beschreibe ich das Sample, das die Curricula aller Studiengänge bzw. Vertiefungen in Datenwissenschaften an Schweizer Universitäten und Hochschulen umfasst. Dieses kodiere und untersuche ich mithilfe der Inhaltsanalyse (Kap. 4.4.1). Anschliessend skizziere ich den Prozess der Erhebung der qualitativen Interviews mit den Studiengangleitenden ausgewählter Studiengänge (Kap. 4.4.2), die ich ebenfalls inhaltsanalytisch auswerte (Kap. 4.4.3).

4.4.1 Curricula und Curricula-Analyse

Wie bereits im Forschungsstand dargelegt, sind Curricula mehr als die Summe der relevanten Wissensbestände eines Feldes: In ihnen artikulieren sich stets auch Konzeptionen über das Verhältnis eines Wissensgebiets zu seinen ökonomischen, technologischen, politischen und weiteren Bezugsfeldern. Curricula der Datenwissenschaften können als Kompromissprodukte verschiedener Disziplinen verstanden werden, die Aushandlungen der involvierten Disziplinen innerhalb von Universitäten oder Hochschulen als organisationalen Feldern abbilden (Knight et al. 2013). Insofern repräsentieren Curricula – ähnlich wie Stellenanzeigen – durch die Kombination von Strukturen und Inhalten eine bestimmte Deutung des Wissensgebiets Datenwissenschaften. Durch die Vielzahl konkurrierender Curricula resultieren und koexistieren demnach multiple, vielstimmige Perspektiven.

Um die verschiedenen curricularen Deutungen der Datenwissenschaften analysieren zu können, habe ich ein Sample aller Studienangebote in Data Science an Schweizer Universitäten, ETH und Fachhochschulen zusammengestellt. Bezeichnend für das Material ist, dass das Sample kontinuierlich angewachsen ist: Gab es zum Zeitpunkt der ersten Erhebungsphase (Sommer 2017) lediglich neun Studiengänge oder Vertiefungen, so ist die Zahl bis zum jetzigen Zeitpunkt (Februar 2022) auf 42 Angebote in unterschiedlichen Hochschultypen und Studienniveaus angewachsen (vgl. Tabelle 1). Die Entwicklung meines Sample bildet insofern den Diffusionsprozess einer Grenzkategorie ab, die von einer wissenschaftlich-industriellen Schnittstelle zu einem disziplinen- und feldübergreifenden Wissensgebiet mutierte, dem hohe Legitimität zugesprochen wird.

Tabelle 1: Übersicht der Studiengänge und Interviews nach institutionellem Kontext²³

Hochschultyp	Studienstufe	Anzahl Studiengänge (davon Vertiefungen)	Interviews durchgeführt
ETH	Master	2	3
Universität	Master	11 (3)	3
	Bachelor	1 (1)*	1
ETH & Universität	Weiterbildung (CAS/DAS/MAS)	7***	1
Universität & Fachhochschule	PhD-Programm ²⁴	1	0
Fachhochschule	Master	3 (1)	1
	Bachelor	6 (3)	0**
	Weiterbildung (CAS/DAS/MAS)	12	2
Forschungsinstitute (SDSC, Datalab)	Weiterbildung (CAS/DAS/MAS)	1***	2
Total		42	13

* es handelt sich um ein Lehrprogramm mit Zertifikat (ohne Abschluss BA/MA)

** die Studiengänge wurden erst nach der qualitativen Erhebungsphase etabliert

*** ein Programm wird in Kooperation von Universität und ETH Lausanne (SDSC) durchgeführt

Mittels Websuche erstellte ich zunächst eine Liste aller verfügbaren Studienangebote in »Data Science«. Da einige Studienangebote lediglich als Nebenfach (Minor), Vertiefungen oder Profile innerhalb anderer Studiengänge (vor allem Informatik) angeboten werden, schloss ich fortan auch solche Programme ein. Parallel öffnete ich den Fokus und inkludierte nun auch Studiengänge, die Data Analytics oder Data Management in der Bezeichnung tragen. Zudem wurden ab 2017 erste Studienangebote in Artificial Intelligence oder Machine Learning etabliert, die ebenfalls die Verarbeitung und Analyse grosser Datenmengen umfassen (vgl. Tabelle 13 im Anhang).

Nachdem die Suchstrategie nun differenziert und verfeinert war, sammelte ich die für die Studienangebote verfügbaren öffentlichen Dokumente und Webseiten wie Werbe- und Informationsmaterialien, Vorlesungsverzeichnisse und Kursbeschreibungen. Ich begann hochschulübergreifende Themen zu eruieren, die die verschiede-

²³ Stand: Februar 2022.

²⁴ Das PhD Network in Data Science der Universitäten Zürich und Neuenburg in Kooperation mit den Fachhochschulen ZHAW und SUPSI wird von *Swissuniversities* im Rahmen des Teilprojekts 2 »Koope-ration zwischen Schweizer FH/PH und UH« gefördert. Vgl. <https://www.swissuniversities.ch/themen/nachwuchsfoerderung/p-1-doktoratsprogramme/tp2-kooperation-zwischen-fh/ph-und-uh> (Zugriff: 03.02.2022).

nen Studiengänge und Vertiefungen charakterisieren. Anschliessend kodierte ich die Materialien inhaltsanalytisch nach unterschiedlichen Merkmalen (Tabelle 2).

Tabelle 2: Merkmale und Ausprägungen der Studiengänge im Sample

Merkmal	Ausprägungen
Hochschultyp	Universität, ETH, Fachhochschule
Name	Name Programm/Vertiefung
Studienstufe	Bachelor, Master, PhD, Weiterbildung
Organisationale Verortung	Fakultäten bzw. Departemente
Kooperation intern	Fakultäten bzw. Departemente
Kooperation extern	Universität bzw. Hochschule
Fachliche Verortung der Lehrenden	Denomination der Professuren*
Zeitpunkt der Etablierung	Jahr
Zulassungsvoraussetzungen	Unterschiedliche Kriterien (Bachelorabschlüsse, Lehrveranstaltungen etc.)
Umfang	Anzahl ECTS
Pflichtbereich	ja/nein, Anzahl ECTS
Wahlbereich	ja/nein, Anzahl ECTS
Abschlussarbeit	ja/nein, Anzahl ECTS
Praxisorientierte Kurse	ja/nein, Anzahl ECTS
Weitere Inhalte	ja/nein, Anzahl ECTS
Publikum	Zielgruppen, Rekrutierungsbemühungen

* Datenlage unvollständig

Die erhobenen Merkmale erfassen demnach sowohl die institutionelle Verankerung, die Struktur und den Aufbau sowie inhaltliche Aspekte der Studiengänge im Sample. Der Hochschultyp markiert die Verortung im akademischen Feld der Schweiz, die nicht nur mit der Grösse und historischen Traditionen, sondern auch mit politischen Verantwortlichkeiten, Vorgaben und finanziellen Ressourcen korrespondiert. Dazu gehört, wie ich noch zeigen werde, auch der Etablierungszeitpunkt, der auf unterschiedliche Strategien und Positionierungen, aber auch organisationale Möglichkeiten zur Einrichtung neuer Studienangebote verweist. Merkmale wie die Programmbezeichnung, Studienstufe, Verortung sowie interne bzw. externe Kooperationen verweisen einerseits auf die organisationale Eingliederung, andererseits auch auf die inhaltliche Ausrichtung des Studiengangs. Die Analyse des Aufbaus der Studiengänge fördert eine spezifische Struktur zutage, die sich einerseits an ingenieurwissenschaftliche Traditionen anlehnt, diese gleichzeitig aber auch erweitert (vgl. Kap. 9): So können die curricularen Elemente – in der Regel Module genannt – von Kernbereich,

Wahlbereich, Abschlussarbeit, praxisorientierte Kurse sowie weitere (»komplementäre«) Inhalte differenziert werden, die für die vorliegende Analyse als Merkmale festgelegt werden. Über die fachliche Verortung der Lehrenden sowie die jeweilige Anzahl Credits in den einzelnen Bereichen erschliessen sich die disziplinären Verankerungen bzw. die eigentliche Interdisziplinarität des Lehrangebots der Studiengänge.

In Anlehnung an Knight et al. (2013: 147) unterscheide ich die Regelstudiengänge (auf Bachelor- und Masterstufe) im Sample in ein Kontinuum zwischen starken und schwachen interdisziplinären Programmen anhand der Zusammensetzung sowie der Anzahl verpflichtender Credits im Kernbereich: Setzt sich der Kernbereich aus mehreren disziplinären Veranstaltungen zusammen und liegt der Anteil verpflichtender Kreditpunkte über 50 %, handelt es sich in diesem Sinne um starke interdisziplinäre Programme.²⁵

4.4.2 Interviews mit Studiengangleitenden und Professor*innen

In einem zweiten Erhebungsschritt habe ich qualitative Interviews mit Leitungspersonen und Lehrenden der Studiengänge durchgeführt.²⁶ Die Interviews dienen dazu, die strukturellen und inhaltlichen Merkmale der Curricula-Analyse mit den Verantwortlichen der Studiengänge zu diskutieren. Einerseits trägt dies dazu bei, die Erkenntnisse des ersten Analyseschritts besser zu verstehen und mit Expert*innen aus dem Untersuchungsfeld zu verifizieren. Andererseits konnte ich so weitere Erkenntnisse, etwa über strategische Überlegungen und organisationale Herausforderungen, gewinnen, die durch die institutionelle Verortung und Interdisziplinarität der Curricula induziert sind und sich nicht aus den öffentlich zugänglichen Dokumenten eruieren lassen. Die Kombination von Curricula- und Interviewanalysen erschliesst somit jene Konstruktionsleistungen, durch welche Akteur*innen im akademischen Feld die Datenwissenschaften objektiv wie subjektiv rahmen.

Insgesamt habe ich dreizehn semistandardisierte Interviews mit vierzehn Studiengangleitenden, Professor*innen und Dozierenden durchgeführt, die ich als Expert*innen für »Data Science« adressiert habe. Die Interviews, die zwischen dreissig Minuten und zwei Stunden dauerten, habe ich auf Tonband aufgezeichnet.²⁷ Unmittelbar im Nachgang an die Interviews habe ich jeweils Memos verfasst, in denen ich neben Notizen und ersten wichtigen Erkenntnissen des Gesprächs auch meine subjektiven Eindrücke über die Interviewsituation festhielt. Zudem sind auch Reflexionen über meine Rolle als Interviewer enthalten. Memos unterstützen nicht nur den

25 Knight et al. schlagen als zweites, organisationales Merkmal den prozentualen Anteil jener Lehrenden vor, deren Denomination innerhalb des Programms liegt, und ob die Studiengangleitung ihre Professur innerhalb oder ausserhalb des Programms hat. Da allerdings im akademischen Feld der Schweiz bis dato kaum Professuren in Data Science existieren, liefert die Erhebung dieses Merkmals keinen zusätzlichen Erkenntnisgewinn.

26 Das ursprüngliche Vorhaben, eine Erhebung bei allen Studiengängen im Feld durchzuführen, liess sich aufgrund mangelnder Teilnahmereitschaft leider nicht realisieren. Deshalb habe ich das Ziel verfolgt, mit unterschiedlichen Positionen im Bereich »Data Science« im universitären Feld der Schweiz zu sprechen.

27 Ein Interview habe ich via E-Mail geführt, wodurch sich die Möglichkeit, Nach- und Vertiefungsfragen zu stellen, erheblich reduziert hat.

späteren Analyseprozess, indem sie Anleitungen geben für die Relevanzsetzungen in den Interviews, sondern sind auch ein wichtiges Hilfsinstrument einer reflexiven Wissensproduktion (Kühner et al. 2013). Die Interviews habe ich in vollständiger Länge nach vorgängig definierten Regeln transkribiert (Langer 2010).²⁸

Eingeleitet habe ich die Interviews jeweils mit der Frage nach den wissenschaftlichen und berufsbiographischen Werdegängen der Befragten (vgl. den Interviewleitfaden im Anhang). Nimmt man die Antworten sowie die gegenwärtigen Stellenbezeichnungen als Ausgangspunkt für eine fachliche Verortung der Interviewpartner*innen, so präsentiert sich angesichts der postulierten Heterogenität des Wissensfeldes ein eher einseitiges Bild: Gesprächspartner*innen waren neun Informatiker*innen sowie je eine Person mit disziplinärem Hintergrund in Chemie, Mathematik, Ökonomie, Informationssysteme sowie Marketing und Kommunikation. Obwohl diverse Befragte teilweise lange Berufsbiographien in Datenwissenschaften aufweisen, verwendet niemand die Selbstbeschreibung als »Data Scientist« für das eigene Tätigkeitsgebiet. Zum einen sind die Lehrenden in bestimmten Disziplinen fachkulturell sozialisiert (Holley 2009), sodass die Neuheit des interdisziplinären Gegenstandes nachgelagert ist: Die Offenheit des Raumes zwischen etablierten Disziplinen bietet die Möglichkeit, über Forschung, Lehre und anderweitig darin tätig zu sein, ohne sich selbst damit identifizieren zu müssen. Zum anderen manifestieren sich darin auch innerakademische Distinktionsmechanismen (Bourdieu 1988), die »Data Science« als professionelle Praxis ausserhalb des akademischen Feldes situieren und primär kommerziellen Tätigkeitsbereichen in der »Industrie« zuschreiben. Schliesslich existieren, obwohl »Data Science« als Wissens- und Forschungsgebiet anerkannt und explizit in Stelleninseraten gesucht wird, trotz der zahlreichen Studiengänge und Vertiefungen an den untersuchten akademischen Einrichtungen bis dato nur einzelne Professuren mit Denomination »Data Science«.²⁹

Der Interviewleitfaden beinhaltet Fragen zu den folgenden Themengebieten: Aufbau und Curriculum des Studiengangs, Kursinhalte, Studienplätze und Zusammensetzung der Studierenden, Relationen zu anderen Hochschulen, Arbeitsmarkt, Definitionen von »Data Science« sowie Hinweise für die weitere Forschung (vgl. den Interviewleitfaden im Anhang). Den Leitfaden habe ich je nach Interview flexibel und situativ eingesetzt, wobei ich den Relevanzsetzungen der Interviewten, die sich im Gespräch ergaben, und den sich daraus ergebenden Anschlussfragen hohes Gewicht beimessen habe.

Zur Anonymisierung der Aussagen der Befragten unterscheide ich nicht zwischen Studiengangleitenden, Professor*innen und Dozierenden, sondern verwende eine Abkürzung für alle Interviewten in Kombination mit dem Hochschultypus (ETH, UH, FH), einem individuellen Buchstaben sowie der Interview- und Zitatnummer, um eine Textstelle eindeutig einem Interview zuzuordnen. So steht die Abkürzung »Prof_ETH_A« für Professor*innen an den beiden ETH, »Prof_UH_B« für Interviewte

28 Die Interviewdateien, Transkripte und der Abkürzungsschlüssel zur Identifizierung der Befragten sind in einem persönlichen, passwortgeschützten Ordner auf dem akademischen Server SWITCHdrive gespeichert. Dieses Vorgehen schützt die Persönlichkeitsrechte der Interviewten und garantiert den Datenschutz.

29 Zudem werden an akademischen Einrichtungen kaum Positionen für »Data Scientists« eingerichtet (Geiger et al. 2018; Moore-Sloan Data Science Environments 2018).

an Universitäten und »Prof_FH_C« für Gesprächspartner*innen an Fachhochschulen. Ist eine Identifizierung einer Universität oder Hochschule notwendig für den Kontext einer Aussage, wird die Nennung der interviewten Person sowie der Interview- und Zitatnummer weggelassen.

4.4.3 Qualitative Analyse der Interviews

Zur Unterstützung der qualitativen Analyse habe ich die Transkripte in der Analysesoftware ATLAS.ti erfasst. In einem ersten Schritt habe ich die für das Erkenntnisinteresse und die Fragestellungen relevanten Textstellen mit zusammenfassenden bzw. erklärenden Codes versehen (Friese 2012: 92ff.). Bei der Auswertung gemäss der strukturierenden qualitativen Inhaltsanalyse wird eine inhaltlich-typisierende Strukturierung vorgenommen und der Umfang des Materials reduziert (Mayring 2010: 92). In die Erarbeitung des Kategoriensystems flossen neben dem Forschungsstand zu interdisziplinären Curricula (und insbesondere jene der Datenwissenschaften) auch die Ergebnisse über die strukturelle und inhaltliche Verfasstheit der untersuchten Curricula mit ein. Ferner habe ich im Sinne eines offenen, induktiven Kodierungsprozesses (Flick 2016: 388ff.) auch Codes aus den Interviews selbst gewonnen. Dabei entstand ein integrales Kategoriensystem, das ich nach dreimaligem Kodieren derselben Dokumente kontinuierlich erweitern und ausdifferenzieren konnte.

Grundlegendes Ziel dabei war das Eruiieren der relevanten Aussagen, Einschätzungen und Reflexionen, die Lehrende über das Feld der Datenwissenschaften anstellen. Die Stellungnahmen der Interviewten erweitern und vertiefen insbesondere die strukturellen und inhaltlichen Merkmale, die ich aus der Analyse der Curricula gewinnen konnte. Sie geben mit anderen Worten den beobachtbaren organisationalen Implementierungen und curricularen Eigenheiten einen subjektiven Sinn. Zusätzlich ermöglicht die Bezugnahme auf die disziplinäre Verankerung der Befragten sowie deren Positionen im akademischen Feld eine Kontrastierung verschiedener Aussagen, wodurch sich sowohl disziplinäre als auch feldspezifische Grenzziehungen und -überschreitungen rekonstruieren lassen.

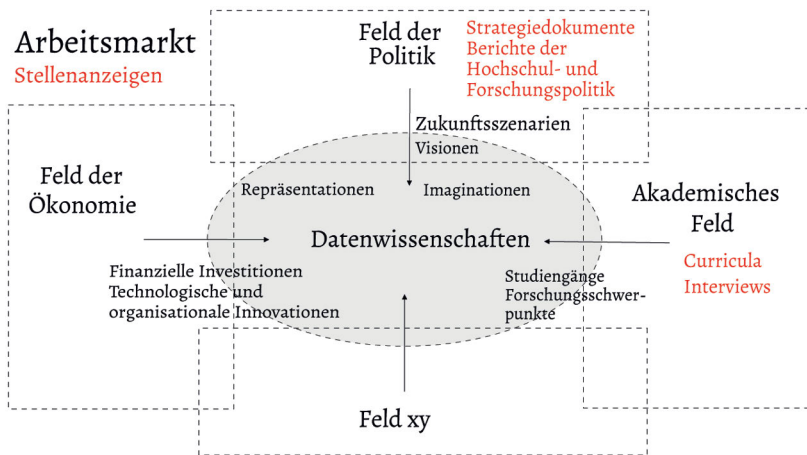
Die qualitative Inhaltsanalyse identifiziert vier zentrale Themenbereiche in den Interviews, die auf unterschiedlichen gesellschaftlichen Ebenen verortet sind und zwischen denen multiple Wechselwirkungen existieren. Zunächst beschreiben und verorten die Befragten die Genese der Datenwissenschaften im Kontext eines fundamentalen Wandels wissenschaftlicher Erkenntnisproduktion. Sie äussern je nach disziplinärer Sozialisation und institutioneller Verankerung unterschiedliche Perspektiven auf den Gegenstand. Die multiplen Verständnisse markieren nicht nur die Heterogenität der Datenwissenschaften, sondern strukturieren wiederum die Positionierung, Ausgestaltung und Profilbildung von Studienprogrammen. Ein zweites Themenfeld umfasst politische und ökonomische Faktoren sowie hochschulstrategische Überlegungen, die den Aufbau konkreter Studienprogramme beeinflussen. Drittens findet das Zusammenspiel hochschulexterner und -interner Einflussgrössen eine Umsetzung in konkreten Implementierungsprozessen der Curricula in Hochschulen und Universitäten. Schliesslich skizzieren die Befragten viertens umfangreiche Kataloge jener als Kompetenzen bezeichneten individuellen Eigenschaften, die sie Praktiker*innen der Datenwissenschaften zuschreiben. Indem die für Datenwissenschaften relevanten Kompetenzen als Zukunftsressourcen gerahmt werden, werden

sie wiederum mit den durch makrostrukturelle Transformationen induzierten neuen Modi der Wissensproduktion verknüpft.

4.5 Zusammenfassung

Die verteilte Analyse von Zwischenräumen bildet ein flexibles, adaptierbares empirisches Modell, das mit Mehrebenenanalysen kompatibel ist: In der vorliegenden Arbeit fokussiere ich primär auf die feldübergreifende Makroebene von Stellenanzeigen sowie die Feldebene der Hochschul- und Forschungspolitik einerseits und auf die verschiedenen Organisationsformen im akademischen Feld andererseits (vgl. Abbildung 2).

Abbildung 2: Methodische Vorgehensweise



Auf der Makroebene ermöglicht die Erhebung von Stellenanzeigen, die multiplen Bedeutungskonstruktionen der Datenwissenschaften durch organisationale Akteur*innen im Arbeitsmarkt zu analysieren. Die deskriptive Analyse erschliesst die zentralen Strukturen des Arbeitsmarktes und zeigt Häufigkeiten und Verteilungen an. Die Auswertung mittels Topic-Modeling-Verfahren identifiziert sodann latente Themen in den Dokumenten. Diese werden nicht ausschliesslich einem Thema alleine zugeordnet, sondern sind in verschiedenen Themen enthalten. Über deren Distribution und Zusammensetzung lässt sich die Vielstimmigkeit sowohl feldtypischer als auch feldübergreifender Themen und Inhalte der Datenwissenschaften in den Stellenanzeigen zugleich rekonstruieren.

Danach wechsele ich die Untersuchungsebene und untersuche Ausprägungen des Gegenstands in zwei zentralen Feldern: Im Feld der Politik entwerfen Akteur*innen durch Strategiepapiere im Rahmen des Diskurses über die Digitalisierung Szenarien einer vielversprechenden Zukunft durch die breite Anwendung der Datenwissenschaften. Durch die inhaltsanalytische Kodierung der Dokumente erfasse ich die artikulierte Multidimensionalität der verwendeten Begriffe und Konzepte und setze sie

in Beziehung zu den anderen Feldern, insbesondere der Wissenschaft sowie der Ökonomie. Den Gegenstand analytisch als verteilt zu konzipieren, erlaubt es somit, die multiplen Bedeutungsdimensionen wechselseitig aufeinander zu beziehen.³⁰

Während im Arbeitsmarkt und im Feld der Politik der Fokus der Betrachtung auf Praktiken von Begriffsarbeit liegt, deren feldübergreifende Effekte Zwischenräume als Möglichkeitsräume eröffnen, stehen im akademischen Feld die Wechselwirkungen von Begriffs- und Grenzarbeit im Zentrum: In Curricula und Interviews, d. h. in kollektiven und individuellen Stellungnahmen, formulieren Akteur*innen der Wissenschaft disziplinäre bzw. epistemische Perspektiven auf die Datenwissenschaften. Durch das inhaltsanalytische Vorgehen erfasse ich nicht nur die verbindenden Momente der verwendeten Begriffe, sondern auch die widersprüchlichen Dynamiken und Potenziale, die sich in den synchronen Praktiken von Grenzziehung und Grenzüberschreitung artikulieren.

Eine verteilte Analyse von Räumen zwischen Feldern erschliesst somit die multiplen, mitunter divergierenden Bedeutungskonstruktionen des sozialen Phänomens Datenwissenschaften auf unterschiedlichen Ebenen. Indem der Schwerpunkt der Analyse weniger auf die Strukturen der einzelnen Felder, sondern auf die feldübergreifenden Effekte von kollektiven Praktiken gelegt wird, können die Synchronizitäten und Parallelitäten von Praktiken der Begriffs- und Grenzarbeit in der Fundierung und Konturierung der Datenwissenschaften als verteiltes, zwischenräumliches Phänomen herausgearbeitet werden.

30 Daran kann auch eine mikrosoziologisch fundierte Studie anschliessen, die sich die Sinnkonstruktion in Datenpraktiken in einzelnen Organisationen oder Settings anschaut (Mützel et al. 2018).

Teil II – Repräsentationen und Imaginationen von Datenwissenschaften in Arbeitsmarkt und Politik

Im ersten empirischen Teil der Arbeit gehe ich den unterschiedlichen Bedeutungen und Zukunftsvorstellungen nach, durch welche den Datenwissenschaften Sinn zugeschrieben wird. Die Datenwissenschaften sind – anders als die Begriffsbezeichnung vermuten lässt – nicht primär ein wissenschaftliches, sondern auch ein ökonomisches, ein politisches, ein kulturelles, letztlich also ein gesellschaftliches Phänomen. Wissen und Wissenschaft werden nicht nur endogen im wissenschaftlichen Feld geprägt, sondern auch durch ihre Aussenbeziehungen mit anderen Feldern wie Wirtschaft, Politik, Medien etc. Umgekehrt rahmen diese auch die Entwicklung neuer Wissensgebiete. Bevor ich im dritten Teil der Arbeit auf das akademische Feld eingehe, untersuche ich im Folgenden die Bedeutungskonstruktionen der Datenwissenschaften im Arbeitsmarkt sowie in Dokumenten der Bildungs- und Forschungspolitik.

Inhaltlich untersuche ich in den beiden Kapiteln des zweiten Teils, wie durch Stelleninsetrate für Data Scientists im Arbeitsmarkt (Kap. 5) einerseits, durch Strategiedokumente der Hochschul- und Forschungspolitik (Kap. 6) andererseits den Datenwissenschaften als Gegenstand multiple Bedeutungen zugeschrieben, wie sie *ausserhalb* des wissenschaftlichen Feldes konturiert und gerahmt werden. Der Fokus des Erkenntnisinteresses liegt hierbei auf Praktiken von Begriffsarbeit, die die Herausbildung der Datenwissenschaften fundieren: In beiden empirischen Untersuchungsfällen verwenden die identifizierten Akteur*innen offene, inklusive Begriffe zur Kategorisierung des Wissensgebiets. So charakterisieren die Stelleninsetrate publizierenden Organisationen die Datenwissenschaften durch feldspezifische Kombinationen epistemischer Praktiken, Wissensobjekte und Inhalte. Dies ermöglicht es ihnen, ihre Such- und Rekrutierungsprozesse einerseits in dem Wissensgebiet zu positionieren, andererseits aber auch flexibel und offen zu halten, was Möglichkeitsräume schafft und damit Handlungsoptionen erweitert. Akteur*innen im Feld der Politik rahmen die Datenwissenschaften zunächst innerhalb existierender, verwandter Diskurse und setzen sie mit ambiguen Begriffen wie Digitalisierung, Kompetenzen oder Innovation in Beziehung. Damit erhalten sie nicht nur die Vielstimmigkeit des Gegenstandes selbst, sondern schaffen gleichzeitig weitreichende Erwartungen und Hoffnungen auf gesellschaftliche Transformation durch die Entstehung der Datenwissenschaften als neues Wissensgebiet.

In beiden Fällen ermöglichen es die identifizierten Begriffspraktiken den Akteur*innen unterschiedlicher Felder, strategisch daran anzuknüpfen und feldübergreifende Kollaborationen zu etablieren. Insofern tragen die Multidimensionalität und Ambiguität der verwendeten Begriffe zur Koordination der Akteur*innen über die Untersuchungsfelder hinweg bei.

Kapitel 5 – Repräsentationen der Datenwissenschaften im schweizerischen Arbeitsmarkt

5.1 Einleitung

»Sherlock Holmes wäre von [Ihrem] Data-Spürsinn begeistert? Ist für [Sie] Data Science der rote Faden[,] der zum Erfolg führt? Sie haben einen Hochschulabschluss in Wirtschaftsinformatik, Informatik, Statistik, Mathematik oder Physik oder eine vergleichbare Ausbildung. Sie haben in ein oder mehreren Technologien und Analysetools ihre Erfahrungen gesammelt. Interessant sind zum Beispiel Technologien wie Python, R, Scala, Hadoop, Spark, SAP HANA ... Oder Analysetools wie Hue, SAS Customer Intelligence, SPSS, SAP BO, Matlab, Shiny, Splunk, TensorFlow, Tableau, Spark und Qlikview? ... Zögern [Sie] nicht, mir Ihre Fähigkeiten mitzuteilen! Neben Ihrer Data-Kompetenz liegen Ihre Stärken in geschicktem, initiativem, zielorientiertem, verantwortungsbewusste[m] und unternehmerischem Denken, Handeln und Verhandeln. Sie können gesprächspartnerorientiert komplexe Sachverhalte kompetent, einfach und verständlich kommunizieren. Auch fungieren Sie gerne selbstständig und sind zugleich ein intuitiver Teamplayer. Sie schätzen es[,] Ihre Umgebung für Ihre innovativen Ideen und Projekte zu begeistern!«¹

Das einleitende Zitat beschreibt das Profil einer »Persönlichkeit« für eine Position als Data Scientist im Schweizer Arbeitsmarkt. Auf der inhaltlichen Ebene wird deutlich, dass eine Person mit quantitativ-technischem Fachhintergrund gesucht wird, die mit verschiedenen Programmiersprachen und Datenanalysetools versiert ist und zudem unternehmerische, soziale und kommunikative Kompetenzen aufweist – ein »Skillset«, das so oder ähnlich für verschiedene Berufsbezeichnungen artikuliert wird, gleichzeitig allerdings äusserst selten vorzufinden ist. Trotzdem steht es exemplarisch für zeitgenössische technikwissenschaftliche Wissensgebiete (Paulitz & Prietl 2017: 158ff.).

¹ Stelleninserat »Data Scientist: Customer Journey« des Unternehmens Addexpert GmbH: [https://www.addexpert.ch/jobs/job-detail.aspx?Bereich=Informatik&\]jobID=6244](https://www.addexpert.ch/jobs/job-detail.aspx?Bereich=Informatik&]jobID=6244) (Original nicht mehr verfügbar; letzter Zugriff: 04.07.2017, Ergänzungen zur besseren Lesbarkeit in eckigen Klammern durch den Autor).

Interessant an diesem Ausschnitt ist, dass sich das Aufzählen äquivalenter Alternativen durch die verschiedenen Ebenen des skizzierten Persönlichkeitsprofils zieht. Das Zitat beginnt gleich mit zwei Metaphern, die unterschiedliche Deutungsangebote für die Persönlichkeit als Data Scientist bereitstellen: Zum einen wird das Bild einer detektivisch-analytischen Persönlichkeit (»Sherlock Holmes«) bemüht, die dank ihrem »Data-Spürsinn« komplexe Fälle lösen kann. Es wird also eine persönliche Eigenschaft der möglichen Bewerber*innen angesprochen, die eine führende (literarische) Referenzfigur der kriminalistischen Forensik »begeister[n]« würde, sofern sie denn existierte. Ein zweites Bild wird durch das Grundmotiv des »rote[n] Faden[s]« evoziert, wonach in »Data Science« die leitende Strukturierung besteht, »[die] zum Erfolg führt«. Die zweite Metapher setzt »Data Science« also als Mittel zum Erfolg, wobei unklar im Inserat insgesamt bleibt, um welche Art von Erfolg (kommerziell, wissenschaftlich, in der Auflösung von Kriminalfällen etc.) es sich handelt.

Auf die zwei unterschiedlichen Imaginationen zur Persönlichkeit von Data Scientists zum Einstieg folgt eine Aufzählung von disziplinären (»Wirtschaftsinformatik, Informatik, Statistik, Mathematik oder Physik«) und qualifikatorischen (»Hochschulabschluss [...] oder vergleichbare Ausbildung«) Hintergründen. Diese werden in der Stellenanzeige als Grundlage des Persönlichkeitsprofils präsentiert, ohne dies allerdings explizit zu machen. Kryptisch wird es anschliessend für Nicht-Insider*innen, wenn die gesuchten »Erfahrungen« und »Fähigkeiten« in knapp zwei Dutzend unterschiedlichen »Technologien« und »Analysetools« ausbuchstabiert werden. Schliesslich nennt der Ausschnitt eine ganze Reihe von Kompetenzen, d. h. individuell zugeschriebenen Eigenschaften, die sehr unterschiedliche Betätigungsfelder und Praktiken indizieren, jedoch divergente Entstehungshintergründe aufweisen (Salvisberg 2010).

Strukturell, so lässt sich der zitierte Ausschnitt des Stelleninserats interpretieren, bleibt also unklar, aus welchen Imaginationen, Disziplinen, Qualifikationen, Technologien, Analysetools und Kompetenzen sich die gesuchte Persönlichkeit von Data Scientists zusammensetzt. Sie präsentiert sich als eine Aufzählung vermeintlich äquivalenter Zuschreibungen und Eigenschaften. Sinngemäss sind solche Aufzählungen im Sinne »unverbindliche[r] Wunschlisten« (Geser 1983: 479) ein konstitutives Merkmal von Stellenanzeigen, um ein möglichst breites Spektrum an Bewerber*innen anzusprechen. Der Forschungsstand hat aber deutlich gemacht, dass sich eine Vielzahl unterschiedlicher Akteur*innen mit der Frage nach den Begriffen, Kompetenzen und Persönlichkeitsprofilen der Datenwissenschaften beschäftigt. Die Suche nach und das Zusammenstellen der »richtigen« »Skillsets«, die sich im zitierten Ausschnitt des Stelleninserats exemplarisch manifestieren, kulminierten in den letzten Jahren zu einem zentralen Erkenntnisinteresse im Untersuchungsfeld selbst. Ein Merkmal der Diskussion besteht darin, umfangreiche Listen unterschiedlicher Methoden, Tools, Bildungsqualifikationen oder individueller Fähigkeiten und Eigenschaften zu artikulieren. Das Formulieren solcher äquivalenter Alternativen signalisiert insofern sowohl Inklusivität bezüglich disziplinärer Wissensbestände, Qualifikationen oder methodisch-technischer Expertisen als auch Unklarheit darüber, was noch dazugehört und was nicht.

Das Kapitel fragt danach, wie solche Bedeutungszuschreibungen über die Datenwissenschaften durch unterschiedliche Listen und Aufzählungen multipler Begrifflichkeiten konstruiert werden. Allgemein erfuhren die Datenwissenschaften im Arbeitsmarkt in den letzten Jahren einen signifikanten Bedeutungsgewinn, der sich

in einer steigenden Anzahl von Stellenangeboten für Data Scientists äussert. Synchron dazu werden die Datenwissenschaften in der gesellschaftlichen Diskussion vornehmlich mittels ökonomischer Deutungen, wie der Diagnose eines gravierenden Fachkräftemangels (Manyika et al. 2011; Markow et al. 2017), gerahmt (Saner 2019). Wie ich noch ausführen werde, hat dies auch Effekte auf andere Felder: So werden die finanzielle Förderung von Forschungsinitiativen durch die Politik sowie die rasche Implementation neuer Studiengänge an Hochschulen und Universitäten primär mit der hohen Nachfrage nach entsprechenden Abschlüssen und Fähigkeiten im Arbeitsmarkt begründet. Stellenanzeigen für Data Scientists stellen insofern eine wichtige Konstruktionsleistung der Datenwissenschaften durch Akteur*innen in verschiedenen Feldern dar. Die Analyse ihrer strukturellen und inhaltlichen Charakteristika ist deshalb elementar für das Verständnis der Datenwissenschaften als Wissensgebiet.

Wie die Analyse zeigen wird, sind es neben wissenschaftlichen Institutionen vor allem Unternehmen in ökonomischen Feldern, aber auch öffentliche Einrichtungen, Non-Profit- sowie internationale Organisationen, die Kompetenzprofile von und für Data Scientists konstruieren. Sie alle bauen »Analytics Teams« auf (J. G. Harris et al. 2013; Kim et al. 2016; Patil 2011), formulieren Tätigkeitsbeschreibungen, publizieren Stelleninserate und rekrutieren schliesslich Datenwissenschaftler*innen. In den Anforderungen an Stellenbewerber*innen verdichten sich einerseits intraorganisatorische Rollenerwartungen (Geser 1983), andererseits werden feld- und organisationspezifische Perspektiven darauf formuliert, worin datenwissenschaftliche Praktiken und Expertisen bestehen, wo die Berührungspunkte sowie Grenzziehungen zu verwandten Gebieten liegen. Stellenanzeigen als Stellungnahmen kollektiver Akteur*innen repräsentieren die Wahrnehmung des Wissensfeldes nach aussen genauso wie innerhalb der jeweiligen Organisation, d. h., sie fixieren durch sprachliche Kategorien eine bestimmte Deutung der Datenwissenschaften und privilegieren diese gegenüber anderen möglichen Bedeutungen. Das Changieren zwischen und das Ringen um die relevanten Kategorien, das sich in Stellenanzeigen in umfangreichen Listen kontingenter Werkzeuge, Methoden oder Bildungsanforderungen manifestiert, strukturiert umgekehrt die Erwartungshaltungen von Akteur*innen in unterschiedlichen Feldern. Indem deren multiple, divergierende Praktiken dadurch koordiniert werden, fundiert dies die Datenwissenschaften als zwischenräumliches Phänomen und trägt zu dessen Stabilisierung bei.

Stellenanzeigen für Data Scientists und verwandte Berufsbezeichnungen ermöglichen es nun, von einem makroskopischen Standpunkt aus die Konstruktionsleistungen unterschiedlicher Akteur*innen zu analysieren, aber auch Gemeinsamkeiten und Unterschiede der involvierten Felder herauszuarbeiten. Das Kapitel erhebt keinen Anspruch auf eine umfassende soziologische Analyse des schweizerischen Arbeitsmarktes (dazu Aratnam 2012; Kriesi et al. 2010; Sacchi et al. 2005; Salvisberg 2010), sondern es fragt nach den für das Segment der Datenwissenschaften zentralen Organisationen, Feldern, Profilen und Bedeutungszuschreibungen. Folgende Fragestellungen leiten die Analyse an:

- Wie ist das Segment der Datenwissenschaften im Arbeitsmarkt strukturiert? Welche Organisationen in welchen Feldern publizieren Stelleninserate für Data Scientists?
- Welche Profile von Data Scientists werden in den Inseraten konstruiert?

- Wie werden die Datenwissenschaften im Arbeitsmarkt repräsentiert, d. h., welche Deutungen werden in den Stelleninseraten entworfen?

Die Datengrundlage zur Untersuchung der Fragestellungen bilden Stelleninserate für Data Scientists und verwandte Professionen, die ich mittels Webscraping-Verfahren auf einer Schweizer Online-Jobplattform erhoben habe (dazu Kap. 4.2). Das resultierende umfangreiche Korpus an Textdokumenten habe ich mithilfe von Topic Modeling, einem computerlinguistischen Verfahren, ausgewertet.

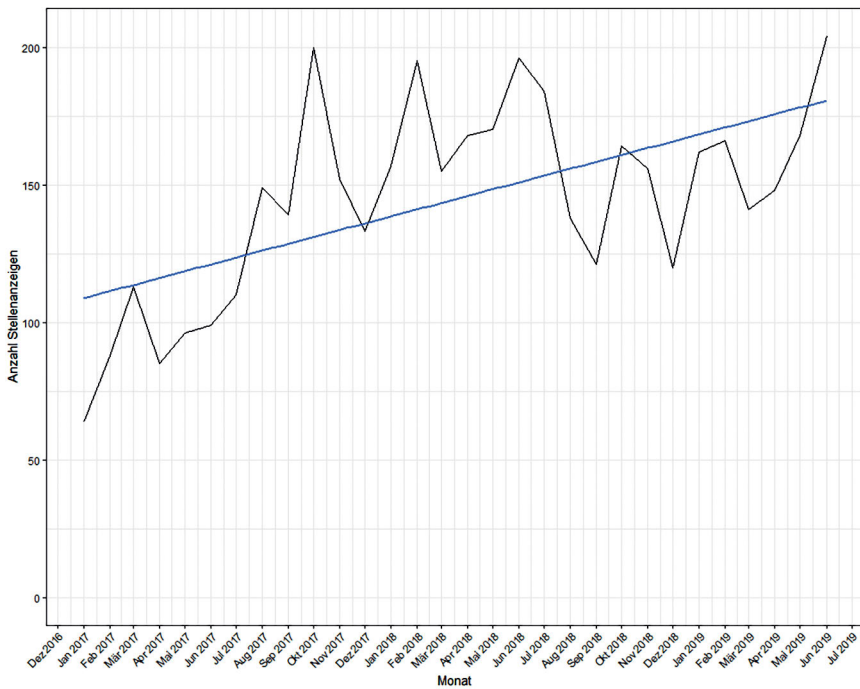
Das Kapitel ist wie folgt strukturiert: Zuerst erläutere ich anhand der deskriptiven Auswertung der Daten die Entwicklung, untersuche die zentralen Begriffe sowie die feldspezifische, geographische und sprachliche Verteilung der Datenwissenschaften im schweizerischen Arbeitsmarkt (Kap. 5.2). Den Hauptteil des Kapitels bilden die Analysen der Repräsentationen der Datenwissenschaften in den englischsprachigen (Kap. 5.3) sowie den deutschsprachigen Stellenanzeigen (Kap. 5.4). Schliesslich diskutiere ich die zentralen Ergebnisse im Hinblick auf die übergeordneten Fragestellungen der Arbeit (Kap. 5.5).

5.2 Entwicklung und Struktur des Arbeitsmarktsegments Datenwissenschaften

5.2.1 Entwicklung

Die Anzahl Stellenanzeigen nimmt über den Erhebungszeitraum hinweg betrachtet leicht, aber kontinuierlich zu. Sind die Stellenanzeigen nach Monaten gruppiert, wird die Entwicklung visuell nachvollziehbar (vgl. Abbildung 3). Durchschnittlich sind pro betrachteten Monat rund 145 Anzeigen erfasst. Die meisten Anzeigen wurden im Juni 2019 publiziert ($n = 204$), die wenigsten im Januar 2017 ($n = 66$). Zudem sind gewisse Schwankungen nach Jahreszeiten zu erkennen, beispielsweise die rückläufigen Zahlen zum Jahresende (Dezember 2017 und 2018), während in den darauffolgenden Monaten ein Anstieg zu beobachten ist. Dies ist möglicherweise die Folge regulatorischer Beschränkungen: Arbeitsbewilligungen für sogenannte »Drittstaatenangehörige«, d. h. Arbeitskräfte ausserhalb der EU/EFTA-Staaten, die primär für hochspezialisierte Berufstätigkeiten vergeben werden, wie es die untersuchten Professionen darstellen, sind stark kontingentiert (Aratnam 2012). Dafür spricht der hohe Anteil englischsprachiger Stellenanzeigen, der als ein Indikator für deren internationale Reichweite konzipiert werden kann, um ein breites Feld an Kandidat*innen anzusprechen, das nicht auf nationalstaatliche bzw. sprachräumliche Grenzen limitiert ist.

Abbildung 3: Entwicklung der Stellenanzeigen im Korpus

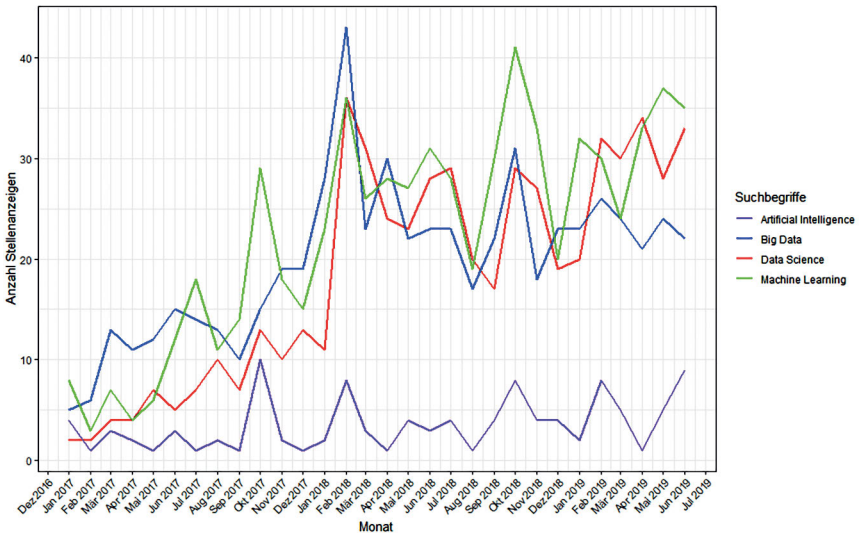


Bei der Unterscheidung der Stelleninserate nach den Begriffen Data Science, Big Data, Machine Learning und Artificial Intelligence offenbaren sich nur leicht divergierende Entwicklungsverläufe (vgl. Abbildung 4): Trotz inhaltlicher Differenzen weisen die drei Kategorien »Data Science«, »Big Data« und »Machine Learning« ähnliche Entwicklungen wie das Gesamttotal der Stellenanzeigen auf.² Obwohl »Machine Learning« nicht zur Erhebung des Samples verwendet wurde, bildet der Begriff die Entwicklung sogar leicht besser ab als »Data Science« und erfährt im Oktober 2018 einen Peak. Dies weist darauf, dass »Machine Learning« zwar als Synonym bzw. Äquivalent für »Data Science« verwendet wird, während umgekehrt »Data Science« einen umfassenderen Begriff darstellt. Die Anzeigen für den Suchbegriff »Big Data« erreichen nach starkem Wachstum im ersten Erhebungsjahr im Februar 2018 einen Höchststand und sind danach rückläufig. Trotz seiner gegenwärtigen Konjunktur ist der Begriff »Artificial Intelligence« nur selten in den Stellenanzeigen vertreten.³

2 Die Korrelation der drei Datenreihen im Verhältnis zur Gesamtentwicklung ist mit Pearson R-Werten von 0.68 (»Big Data«), 0.75 (»Data Science«) und 0.81 (»Machine Learning«) hoch. Der Korrelationskoeffizient Pearson R ist ein Maß für die Stärke und Richtung eines linearen Zusammenhangs zwischen zwei intervallskalierten Variablen X und Y (Diaz-Bone & Weischer 2015: 230).

3 »Artificial Intelligence« bildet zudem die Gesamtentwicklung am schlechtesten ab (Pearson R = .53).

Abbildung 4: Entwicklung der Stellenanzeigen nach Suchbegriffen



Die temporale Entwicklung verläuft ähnlich, wenn die Stellenanzeigen nach Berufstiteln, Feldern oder geographischen Regionen differenziert betrachtet werden. Insgesamt erfährt das Segment der Datenwissenschaften demnach eine kontinuierlich zunehmende Relevanz im schweizerischen Arbeitsmarkt.

5.2.2 Häufige Begriffe in den Inseratetexten

Die Texte der Stelleninserate bilden das Hauptuntersuchungsobjekt des Kapitels. Mittels einfacher Wortfrequenzen und Vergleichen lassen sich deskriptiv Muster erkennen, noch bevor komplexere Verfahren wie Topic Modeling zur Anwendung kommen.

Im englischsprachigen Korpus wird – wenig erstaunlich – *data* ($n = 21'479$) am meisten verwendet. Weitere zentrale Terme sind *experience*, *team*, *work*, *skills*, *knowledge*, *working*, *job*, *support* oder *strong*, die charakteristisch für die Textgattung Stellenanzeigen sind. Spezifischer für das Sample sind Begriffe wie *business*, *development*, *management*, *research*, *science*, *analytics*, *analysis* oder *software*, die bestimmte Tätigkeitsfelder indizieren. Des Weiteren sind auch Rollenbezeichnungen (*scientist*, *analyst*, *engineer* etc.) sowie Standorte (*switzerland*, *basel*, *zürich*) unter den hundert häufigsten Termen zu finden.

Auch im deutschsprachigen Sample ist *data* ($n = 3'582$) der wichtigste Begriff, der immerhin in 92.6 % der Anzeigen erscheint, oft in Kombination mit weiteren englischsprachigen Termen *analytics*, *big* oder *science*. Dies zeigt, dass die Verwendung kategorialer Komposita wie »Data Science« oder »Big Data« sich auch im deutschen Sprachraum durchgesetzt hat. Begriffe wie Datenwissenschaften oder das gelegentlich verwendete Massendaten (Kolany-Raiser et al. 2018: 76ff.) finden sich nicht in den deutschsprachigen Inseraten. *daten* ($n = 781$) als deutschsprachiges Pendant zum Suchbegriff ist jedoch immerhin in 60.9 % der deutschsprachigen Dokumente vertreten. Weiter zeigen sich beim allgemeinen Vokabular diverse Überschneidungen (*erfahr-*

zung, bereich, kenntnisse, team, arbeiten, suchen, aufgaben etc.) zum englischen Korpus, aber auch einige sprachspezifische Unterschiede wie *gute, kunden, m* (= »männlich«), *w* (= »weiblich«), *unsere** oder *ag*.

Inhaltlich interessant ist, dass sprachübergreifend verschiedene Analysetools bzw. Programmierumgebungen (*sql, r, python, excel, sap*) unter den hundert häufigsten Begriffen vertreten sind. Dies verweist einerseits auf die fundierende Bedeutung solcher Tools und Devices für die Datenwissenschaften. Andererseits werden Methoden und Analysetools wie erwähnt in den Inseratetexten oft austauschbar und in nicht-hierarchisierenden Aufzählungen angeführt (vgl. Beispielinserat 1). Da methodische Praktiken in einer Vielzahl von Softwareumgebungen programmierbar sind, rückt die Bedeutung des konkreten Tools in den Hintergrund – so können auch in SPSS »neuere Ansätze des [...] Machine Learnings« implementiert werden (Beispielinserat 1). Stellenanzeigen für Data Scientists verwenden demnach ein breites, inklusives Verständnis der erforderlichen methodischen und analytischen Tools.⁴ Als *boundary objects* erfüllen sie die multiplen Anforderungen der Praxisfelder und können den jeweiligen lokalen Bedingungen angepasst werden, behalten dennoch ihre Identität und Stabilität über Feldgrenzen hinweg.

5.2.3 Berufsgruppen

Über die Titel der Stellenanzeigen lassen sich verschiedene Gruppen bilden, die auf unterschiedlichen Rollenbezeichnungen basieren. Die Einteilung in verschiedene Berufsgruppen dient dazu, um spezifische Zuschreibungen von Fertigkeiten, technologischen Kenntnissen oder Bildungsanforderungen im Sample zu untersuchen (Wowczko 2015). Darüber lassen sich auch die verschiedenen Profile von Data Scientists erschliessen, die in der Literatur angeführt werden (vgl. Kap. 3.2). Besonders relevant sind die Gruppen der *Scientists* (23.3 %), der *Analysts* (17.6 %) sowie der *Engineers* (13.7 %). Die drei Gruppen stehen demnach für 54.6 % aller Anzeigen. Ebenfalls von Bedeutung sind die Gruppen *Manager* (7.8 %), *Specialists* (6.4 %), *Interns* (6.1 %), *Consultants* (5.1 %) sowie *Developer* (4.8 %). Die zehn wichtigsten Berufsgruppen im Sample machen rund 90 % des Samples aus (vgl. Tabelle 9 im Anhang). Betrachtet man die exakten Stellenbezeichnungen, so machen die drei zentralen Stellenbezeichnungen im Sample – *Data Scientist* (n = 587), *Data Analyst* (n = 497) und *Data Engineer* (n = 128) – rund 29 % der Anzeigen aus (vgl. Tabelle 10 im Anhang). Für den Berufstitel *Data Scientist* existiert keine deutsche oder französische Entsprechung im Sample – ein Hinweis darauf, dass sich die Rollenbezeichnung auch für anderssprachige Stelleninserate im Arbeitsmarkt etabliert hat.

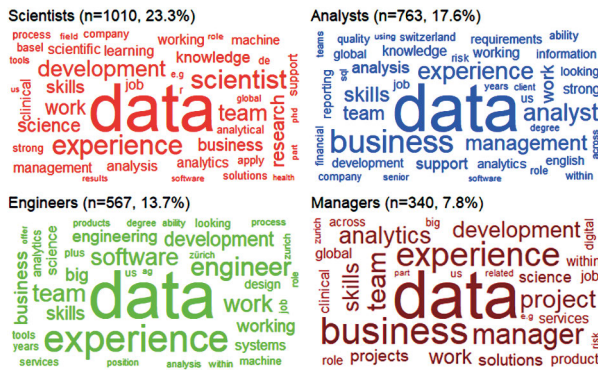
Über die Titel lassen sich zudem oft auftauchende Attribute eruieren (vgl. Abbildung 12 im Anhang): In rund 24 % der Anzeigen werden Pensenangaben verwendet. Am häufigsten sind 100 % (12.6 %), gefolgt von 80–100 % (5.8 %) sowie 80 % (3.9 %). Rund 18 % der Anzeigen enthalten geschlechtsindizierende Angaben (wie *male/female* oder *männlich/weiblich* bzw. deren Abkürzungen). Terme wie *Senior* (12.8 %), *Junior* (3.4 %), *Lead* (1.9 %) oder *Associate* (1.8 %) können der beruflichen Hierarchie der ausgeschriebenen Stellen zugeordnet werden. Auch andere Begriffe sind inhaltlich identifizierbar:

4 Hier offenbart sich ein Widerspruch zu curricularen Konzeptionen an Universitäten und Hochschulen, die datenwissenschaftliche Expertise explizit in der »Fähigkeit zu skripten und programmieren« verorten und »nicht bloss in SPSS das richtige Menü auszuwählen« (Prof_UH_A).

Diese sind einerseits in konkreten Berufsbezeichnungen enthalten (wie *Business Analyst*), indizieren andererseits aber auch ein bestimmtes Tätigkeitsfeld, wie beispielsweise *Business* (7.6 %), *Software* (5 %), *Science* (4.6 %), *Research* (3.5 %) oder *Management* (3.2 %). Attribute wie *digital* (2.9 %), *clinical* (2 %) oder *technical* (1.4 %) charakterisieren gewisse Tätigkeitsfelder (*digital marketing* oder *clinical research*) oder Berufsbezeichnungen (*technical consultant*). In einzelnen Fällen werden auch Programmiersprachen oder Softwarelösungen (wie *SAP*, *R* oder *Java*) erwähnt.

Die Analyse häufiger Begriffe nach aggregierten Berufsgruppen fördert einige Unterschiede zutage (vgl. Abbildung 5): In der Gruppe der *Scientists* indizieren Terme wie *development*, *science*, *research*, *clinical*, *learning* oder *scientific* wissenschaftliche Tätigkeitsfelder. Das pharmazeutische Feld (*pharmaceutical* und *basel*) ist prominent in dieser Berufsgruppe vertreten. Bei den *Tools* figuriert die Statistiksprache *R* unter den dreissig häufigsten Begriffen. Bei den *Analysts* überwiegen demgegenüber für das ökonomische Feld generische Begriffe wie *business*, *management*, *reporting*, *quality* oder *services*. Trotzdem verweisen *risk*, *zürich* oder *financial* auf die primäre Verortung im Banken- bzw. im Versicherungsfeld. Unter den *Tools* findet *sql* am meisten Erwähnung. Bei den *Engineers* wiederum sind es vor allem *software*, *big*, *systems*, *design* oder *technology*. Tokens wie *zürich*, *swisscom*, *computer* und *platforms* geben Hinweise auf entsprechende Organisationen des informations- bzw. kommunikationstechnologischen Feldes. Hier ist *python* als einziges Tool unter den fünfzig häufigsten Begriffen vertreten.

Abbildung 5: Wordclouds der häufigsten Begriffe nach ausgewählten Berufsgruppen



Die Differenzierung der untersuchten Inserate über die Stellentitel legt unterschiedliche Profile von datenbezogenen Rollenbezeichnungen wie *Data Scientists*, *Data Analysts*, *Data Engineers* und weiteren offen. Wie bereits gezeigt prägte die Suche und Identifizierung entsprechender Profile von Beginn weg die Genese der Datenwissenschaften als soziales Phänomen. Die Ergebnisse der Analyse deuten signifikante Unterschiede hinsichtlich feldspezifischer Verortung, Standorten, Arbeitsinhalten, Qualifikationsanforderungen oder methodischen Vorgehensweisen an. Die identifizierten Begriffskonvolute repräsentieren das untersuchte Arbeitsmarktsegment insofern als ein heterogenes Wissensgebiet, das multiple epistemische Praktiken, disziplinäre Wissensbestände sowie technologische Objekte umfasst, die sich – wie ich im folgenden Unterkapitel zeigen werde – nicht auf einzelne soziale Felder begrenzen lassen.

5.2.4 Organisationen und Felder

Über die Organisationen, welche die Stellen ausschreiben, können die ökonomischen und sozialen Felder erschlossen werden, in denen Datenwissenschaftler*innen gesucht werden und tätig sind.⁵ Das Sample enthält 908 Nennungen unterschiedlicher Organisationen, die achtzehn unterschiedlichen Feldern entstammen. Dazu zählen vierzehn ökonomische Felder, das universitäre Feld, das Feld der staatlichen Verwaltung, das Feld der internationalen und Nichtregierungsorganisationen (NGO) sowie das hybride Feld der Gesundheit (Brown & Barnett 2004).⁶ Bei der Analyse der Häufigkeiten fällt sofort die schiefe Verteilung der Inserate ins Auge: So verantworten zehn Organisationen einen Viertel aller Anzeigen (vgl. Tabelle 3), 65 Organisation stehen bereits für rund 50 % aller Inserate. Demgegenüber gibt es 161 Organisationen, die zweimal im Sample enthalten sind, und 408 Organisationen mit nur einem Inserat. Das Pharmaunternehmen F. Hoffmann-La Roche ist mit 331 Inseraten (7.6 %) am häufigsten im Sample vertreten. Es folgen die Grossbanken Credit Suisse mit 125 Inseraten (2.9 %) und UBS mit 118 Anzeigen (2.7 %), das Telekommunikationsunternehmen Swisscom (2.5 %) sowie das Pharmaunternehmen Novartis (2.4 %).

Die Aufschlüsselung der Inserate nach einzelnen Feldern verdeutlicht, dass die fünf Organisationen mit den meisten Stellenanzeigen im Pharma-, Telekom/IKT- sowie im Bankenfeld zu verorten sind. Es handelt sich um Unternehmen, die alle ihren Hauptsitz in der Schweiz haben. Des Weiteren zählen der weltweit grösste Hersteller von Tabakprodukten (Philip Morris), die ETH Zürich, zwei international tätige Firmen der Personalrekrutierung (Harvey Nash und Experis) sowie Nestlé zu den zehn am häufigsten inserierenden Organisationen. Pharma und Banken sind auch über alle Stellenanzeigen betrachtet jene Felder, die am meisten Stelleninserate für Data Scientists publizieren: Das pharmazeutische Feld alleine verantwortet 19 % aller Anzeigen, während das Feld der Banken, Finanzdienstleister sowie FinTech-Unternehmen kombiniert 14.3 % des Samples ausmachen. Damit vereinen die zwei Felder bereits einen Drittel aller Anzeigen. Wichtige Felder sind auch das IT-/Software-Feld, das akademische Feld, das Feld der Consultingunternehmen, die Industrie, Versicherungen, Analytics-Firmen und Start-ups, Transport und Logistik, Unternehmen der Informations- und Kommunikationstechnologie (IKT) sowie der Detailhandel (mit jeweils über 4 %). Trotz der Dominanz einiger weniger Grossunternehmen indiziert die Verteilung insgesamt dennoch eine breite Nachfrage nach datenwissenschaftlicher Expertise über die achtzehn unterschiedlichen Felder hinweg. Dies deutet an, dass sich das Phänomen Datenwissenschaften keineswegs nur auf technologieintensive, datengetriebene Felder konzentriert, sondern im Sinne einer Datengesellschaft (Houben & Prietl 2018) auch soziale Bereiche affiziert, die bis dato als nur geringfügig datafiziert wahrgenommen werden, d. h. deren Funktionsweisen und Entscheidungsverfahren sich elementar auf die Erhebung, Prozessierung und Auswertung von Daten stützen.

5 Die folgenden Angaben beziehen sich auf die jeweiligen Organisationen inkl. Tochterunternehmen.

6 Das Feld der Gesundheit umfasst private, halbstaatliche und staatliche Akteur*innen, die sich zwischen Gewinnorientierung, Gemeinwohlorientierung und Mischformen differenzieren lassen (Brown & Barnett 2004).

Tabelle 3: Die zehn am häufigsten inserierenden Organisationen im Sample

Organisation	Anzahl	Anteil	Feld
F. Hoffmann-La Roche AG	331	7.6 %	Pharma
Credit Suisse AG	125	2.9 %	Banken/FinTech
UBS AG	118	2.7 %	Banken/FinTech
Swisscom AG	108	2.5 %	Telekom/IKT
Novartis AG	105	2.4 %	Pharma
Philip Morris	77	1.8 %	Tabak
ETH Zürich	69	1.6 %	Wissenschaft/Bildung
Harvey Nash AG	62	1.4 %	Personalrekrutierung
Experis	59	1.4 %	Personalrekrutierung
Nestlé	41	0.9 %	Lebensmittel
Total Sample	4341	100 %	

Interessant ist zudem die Aufschlüsselung der identifizierten Berufsgruppen nach einzelnen Feldern: Das pharmazeutische Feld schreibt mit deutlichem Abstand die meisten Stellen in der Berufsgruppe *Scientist* aus (41.3 %), gefolgt von der Tabak- (31.4 %) und der Lebensmittelindustrie (28.4 %). Die Stellenbezeichnung *Data Scientist* wird am meisten von Organisationen in den Feldern Versicherungen (24.8 %), Detailhandel (24.4 %) sowie Transport und Logistik (23 %) verwendet. Es handelt sich oft um Organisationen, die hochgradig datafiziert sind (Muster & Büchner 2018). *Data Science* taucht am häufigsten in den Stellenausschreibungen der öffentlichen Verwaltung (17.4 %), Universitäten und Hochschulen (11.3 %) auf, während sie in den anderen Feldern kaum vorkommt. Die staatliche Verwaltung sowie das akademische Feld greifen demnach auf die Kategorie im Sinne eines epistemischen Feldes zurück, während in den ökonomischen Feldern in erster Linie die Rollenbezeichnung *Data Scientist* dominiert.

Die Gruppe der *Analysts* macht im Banken/FinTech-Feld (34.7 %) über einen Drittel aller Anzeigen aus, gefolgt von den Versicherungen (31.6 %), Medien und Kommunikation (31.34 %) sowie dem Detailhandel (29.7 %). *Engineers* werden am häufigsten im Telekom/IKT-Feld (34.1 %), bei Analytics-Firmen und Start-ups (25.3 %), der Industrie (19.2 %) sowie in Software-Unternehmen (12.4 %) gesucht. Bei den übrigen Berufsgruppen fällt insbesondere der relativ hohe Anteil an *Consultants* bei Beratungsfirmen

(31.2 %) sowie im IT-Feld (17.1 %) auf. *Internships* und *Praktika* in Datenwissenschaften werden oft in der Industrie (14.4 %), im Detailhandel (7.7 %) sowie in der Pharmabranche (6.3 %) vergeben. Schliesslich sind die hohen Anteile anderer Berufsbezeichnungen an Universitäten und Hochschulen (47.2 %) sowie in der Pharmabranche (17.3 %) bemerkenswert. Eine genauere Inspektion zeigt, dass an Universitäten und Hochschulen feldspezifische Stellenbezeichnungen wie *Postdocs* (20.2 %), *Assistent*innen* (10.7 %), *Mitarbeiter*innen* (5.2 %), *Researcher* (5 %), *Dozent*innen* (4.5 %) sowie *Professor*innen* (3.9 %) oft vorkommen. Pharmaunternehmen suchen *Lead/er* (5.6 %), *Directors* (3.8 %), *Heads* (3.4 %) sowie ebenfalls *Postdocs* (3 %), die in den anderen Feldern selten sind.

5.2.5 Geographische Verteilung

Das untersuchte Arbeitsmarktsegment ist grossmehrheitlich auf wenige urbane Zentren in der Deutschschweiz konzentriert: Die Stadt Zürich ist (31.2 %) der häufigste Standort von Organisationen, die Stelleninserate für Data Scientists publizieren. An zweiter Stelle folgt Basel (16.1 %), was wie gezeigt am Pharmafeld liegt, in dem besonders viele Stellen ausgeschrieben werden. Die Städte Zürich und Basel vereinen also bereits knapp die Hälfte (47 %) aller Inserate und sind somit in Relation zur Bevölkerung signifikant übervertreten. Es folgen mit deutlichem Abstand die Städte Genf (6.3 %), Lausanne (5.2 %) sowie Bern (4.9 %). Zu den zehn häufigsten Standorten zählen ferner Winterthur und Luzern (je 1.6 %), Zug (1.5 %), Neuchâtel (1.4 %) sowie Dübendorf (1.1 %). Die zehn wichtigsten Standorte stehen für rund 71 % der Inserate im Sample.

Die Standorte illustrieren – wie die Sprachverteilung – die Fokussierung auf die deutschsprachige Schweiz gegenüber den anderen Sprachregionen. Betrachtet man die Grossregionen gemäss BFS, so verstärkt sich die Konzentration auf die Agglomerationen Zürich und Basel noch zusätzlich: 40.3 % aller Inserate werden im Kanton Zürich ausgeschrieben, der mit der Grossregion Zürich identisch ist.⁷ Weitere 21.1 % der Inserate finden sich in der Grossregion Nordwestschweiz, zu der die Kantone Basel-Stadt, Baselland und Aargau zählen. Die beiden Grossregionen bilden zusammen über 61 % aller Anzeigen im Sample ab. Es folgt die Genferseeregion, zu der die Kantone Genf, Waadt und Wallis gezählt werden, mit 15.6 %. Die zweisprachige und einwohnerstärkste Grossregion Espace Mittelland (mit den Kantonen Bern, Solothurn, Freiburg, Neuchâtel und Jura) umfasst lediglich 10.6 %. Ebenfalls weniger bedeutend sind die Grossregionen Zentralschweiz (5.5 %), Ostschweiz (3.1 %) sowie das Tessin (0.7 %).⁸ Es manifestiert sich in der Analyse eine ausgeprägte Konzentration auf wenige urbane Zentren, die zudem über ökonomische, politische sowie kulturelle Organisationen in ein globales Feld (Buchholz 2016) von Grossstädten eingebunden sind (Sassen 2013, 2016). Daraus lässt sich schliessen, dass Organisationen und Unternehmen in peripheren Regionen noch weniger datafiziert sind und auf grössere Schwierigkeiten bei der Besetzung von offenen Stellen stossen (Stadelmann 2017).

7 Auch ohne die Stadt Zürich weist die Grossregion Zürich beinahe gleich viele Inserate (9.1 %) auf wie die drei Grossregionen Zentralschweiz, Ostschweiz und das Tessin zusammen.

8 2 % der Anzeigen enthalten Mehrfachnennungen (wie Zürich oder Basel), 0.3 % der Anzeigen sind im Ausland angesiedelt und 0.9 % liessen sich keinem eindeutigen Standort zuordnen. Für diese Inserate konnte folglich keine Kodierung nach Grossregion vorgenommen werden.

Tabelle 4: Die zehn häufigsten Standorte im Sample

Standort	Anzahl	Anteil	Feld mit höchstem Anteil
Zürich	1355	31.2 %	Banken/FinTech
Basel	700	16.1 %	Pharma
Genève	272	6.3 %	Transport/Logistik
Lausanne	225	5.2 %	Tabak
Bern	214	4.9 %	Transport/Logistik
Winterthur	71	1.6 %	Wissenschaft/Bildung
Luzern	68	1.6 %	Versicherungen
Zug	67	1.5 %	Pharma
Neuchâtel	61	1.4 %	Tabak
Dübendorf	47	1.1 %	Wissenschaft/Bildung
Total Sample	4341	100 %	

In der geographischen Verteilung spiegelt sich auch die Konzentration der jeweiligen Felder an bestimmten Standorten: In Zürich weist das Bankenfeld mit 32.8 % den höchsten Anteil⁹ auf, wobei diverse weitere Felder ebenfalls häufig vertreten sind (wie IT/Software, Consulting, Analytics/Start-ups, Wissenschaft oder Versicherungen, je 5–11 %). In Basel hingegen vereint das pharmazeutische Feld fast drei Viertel aller Inserate (74 %), während andere Felder nur selten auftauchen. Dahingegen weisen Städte wie Genf, Lausanne und Bern eine heterogenere Verteilung auf mit verschiedenen Feldern, die mit Anteilen zwischen 10 und 20 % im Sample vertreten sind. In Luzern (Versicherungen), Zug (Pharma) oder Neuchâtel (Tabak) wiederum dominieren ebenfalls einzelne ökonomische Felder.¹⁰

9 Die folgenden Werte umfassen keine Inserate von Personalrekrutierungsfirmen.

10 In der Betrachtung nach Grossregionen schwächt sich diese Verteilung ab: Lediglich in der Nordwestschweiz dominiert das Feld der Pharma mit einem Anteil von knapp zwei Dritteln (65.5 %). Zu erwähnen ist ferner die Ostschweiz, wo der Anteil an Industriefirmen (34.6 %) deutlich höher liegt als derjenige anderer Felder. In den übrigen Grossregionen findet sich eine gleichmässigerere Verteilung über die verschiedenen Felder hinweg.

5.2.6 Sprachverteilung

Sprachlich dominieren Stelleninserate auf Englisch: Insgesamt sind zwei Drittel aller Anzeigen ($n = 2941$, 67.7 %) auf Englisch verfasst, rund 30 % ($n = 1282$, 29.5 %) auf Deutsch. Obwohl auch französisch- ($n = 117$, 2.7 %) und italienischsprachige ($n = 1$, 0.02 %) Anzeigen enthalten sind, bildet das Sample wie erwähnt primär die deutschsprachige Schweiz ab.¹¹ Dabei ist zu beobachten, dass der Titel bzw. die Stellenbezeichnung und der Text der Stellenanzeige oft nicht dieselbe Sprache aufweisen. So hat ein signifikanter Teil der deutschsprachigen Anzeigen einen englischen Titel. Von den drei wichtigsten Rollenbezeichnungen sind jeweils rund zwei Drittel auf Englisch verfasst (vgl. Tabelle 10 im Anhang). Beispielsweise enthält das Sample 497 Anzeigen für *Data Analysts*. Davon sind wiederum 59 % englisch, 34.4 % auf Deutsch und 6.4 % auf Französisch verfasst. Demgegenüber gibt es lediglich 38 Anzeigen für *Datenanalyst*innen* auf Deutsch sowie 11 für *Analyste de données* auf Französisch. Für *Data Scientist* finden sich gar keine deutsch- bzw. französischsprachigen Äquivalente im Sample.¹² Die Ergebnisse deuten darauf hin, dass sich die englischsprachige Professionsbezeichnung demnach bereits vollständig etabliert hat. Trotzdem sind ein Drittel der Anzeigen unter dieser Professionsrolle auf Deutsch oder Französisch verfasst.¹³

Die Sprachverteilung gibt schliesslich auch Hinweise auf die Internationalisierung der untersuchten Felder: So verfassen Unternehmen der Tabakindustrie (98.8 %) und jene der Pharmabranche (97 %) Stelleninserate fast ausschliesslich auf Englisch. Ebenfalls überdurchschnittlich hohe Anteile weisen internationale Organisationen und NGOs (83.6 %), das Lebensmittelfeld (83.2 %) sowie das akademische Feld (79.3 %) auf. Die englischsprachigen Anteile sind also – neben dem akademischen Feld, das in der Schweiz stark internationalisiert ist – primär in jenen ökonomischen Feldern überdurchschnittlich, die von der Präsenz einiger weniger internationaler Grosskonzerne geprägt sind. Es ist insofern naheliegend, dass diese ihre Rekrutierungsaktivitäten stärker als andere Organisationen auf ein international mobiles Segment des Arbeitsmarktes ausrichten. Demgegenüber ist der Anteil englischsprachiger Inserate im Detailhandel (54.4 %), im Gesundheitswesen (48.6 %), im Telekom/IKT-Feld (46.5 %), im Transport- und Logistikfeld (46.1 %), im Energie- und Baufeld (31.9 %) sowie bei Versicherungen (27.4 %) unterdurchschnittlich.¹⁴ Den Gegenpol markiert schliesslich das Feld der öffentlichen Verwaltung, wo keine Stelle auf Englisch verfasst ist. Auch hier

11 Dies dürfte in erster Linie auf den Umstand zurückzuführen sein, dass eine Deutschschweizer Stellenplattform die Grundlage für das Webscraping darstellt (vgl. Kap. 4.2.2).

12 Überdies verwendet lediglich ein Inserat die deutschsprachige Bezeichnung »Datenwissenschaften« im Text.

13 Besonders hohe Anteile englischsprachiger Anzeigen sind bei den Berufsgruppen der Experts (86 %), der Scientists (77.9 %) und bei den Interns/Praktikant*innen (76 %) zu beobachten (vgl. Tabelle 9). Umgekehrt ist der Anteil englischsprachiger Anzeigen bei den Gruppen der Consultants (39 %), der Specialists (54.9 %) sowie der Analysts (60.5 %) unterdurchschnittlich. Bei Letzteren deuten die hohen Anteile deutsch- und französischsprachiger Inserate an, dass Arbeitskräfte gesucht werden, die eine der Landessprachen beherrschen.

14 Insofern ist es erstaunlich, dass die nur wenig internationalisierten Felder der Versicherungen und des Detailhandels am häufigsten den Berufstitel »Data Scientist« für ihre Inserate verwenden (je 25 %).

zeigt sich, dass Organisationen dieser Felder durch sprachliche Festlegungen stärker auf ein inländisches Arbeitskräftepotenzial fokussieren.

Insgesamt präsentiert sich das untersuchte Arbeitsmarktsegment demnach zwar als feldübergreifend verteilt, weist allerdings eine erhebliche Konzentration auf Grossunternehmen in ausgewählten ökonomischen Feldern auf, die stark internationalisiert und an wenigen urbanen Standorten wie Zürich, Basel oder in der Genferseeregion lokalisiert sind. Die Differenzierung in unterschiedliche Profile erfolgt einerseits durch die jeweiligen Felder und Standorte, andererseits durch die Kombination multipler epistemischer Praktiken, disziplinärer Wissensbestände, methodischer Vorgehensweisen sowie technologischer Objekte. Wenn die Anzahl publizierter Stellenausschreibungen für Data Scientists und andere datenbezogene Rollenbezeichnungen einen Gradmesser der Datafizierung darstellt, weisen die Analysen in diesem Teilkapitel auf signifikante Differenzen zwischen Feldern, Organisationen, geographischen und sprachlichen Räumen hin.

5.3 Repräsentationen von »Data Science« in den englischsprachigen Stellenanzeigen

Nachdem ich im vorangehenden Abschnitt die Entwicklung und Struktur der Datenwissenschaften im schweizerischen Arbeitsmarkt untersucht habe, gehe ich im Folgenden auf die Ergebnisse des Topic-Modeling-Verfahrens ein. Zunächst charakterisiere ich die Topics der englischsprachigen Stellenanzeigen allgemein und verdeutliche deren Breite, die sich trotz der Dominanz der ökonomischen Felder erkennen lässt (Kap. 5.3.1). Anschliessend gehe ich spezifisch auf drei Methoden-Topics ein, denen feldübergreifend eine zentrale Bedeutung in der Herausbildung der Datenwissenschaften zugesprochen wird (Kap. 5.3.2).

5.3.1 Die Vielfalt der Topics

Tabelle 5 gibt einen Überblick über die thematischen Makrostrukturen, die die englischsprachigen Stellenanzeigen prägen. Darin sind die zwanzig Topics des berechneten Modells mit den jeweils zwanzig wahrscheinlichsten Begriffen aufgeführt.¹⁵ Während die Reihenfolge der Topics zufällig ist, sind die Begriffe der Topics nach ihrer Wahrscheinlichkeit geordnet. Zur Benennung der Topics wurden sowohl die Begriffswahrscheinlichkeiten als auch Dokumente, die durch die jeweiligen Topics gut beschrieben werden, herangezogen.

Dies wird exemplarisch anhand von Topic 6 nachvollziehbar: Die Begriffe *tobacco*, *philip* und *morris*, die hier zur inhaltlichen Charakterisierung des Topics verwendet werden, sind nicht unter den zwanzig häufigsten Begriffen des Topics (allerdings steht das Akronym *pmi* für »Philip Morris International«). Dennoch bilden sie das Topic inhaltlich kohärent ab, was sich manuell überprüfen lässt: Topic 6 beschreibt Stelleninserate, die zu rund 88 % von Philip Morris stammen. Während zunächst vor allem

¹⁵ Aufgrund der statistischen Funktionsweise von LDA weisen nicht alle Topics dieselbe semantische Kohärenz auf und sind wegen der asymmetrischen Schätzung der Parameter verschieden gross (Kap. 4.2.5).

nicht-tabakbezogene Begriffe wahrscheinlich sind (wie *product, intern, work, develop, team, compani, futur, present, market, world*), finden sich des Weiteren auch *tobacco, cigarett, smoke-fre* oder *smoke* unter den wichtigsten Begriffen. Die inverse Verteilung der Auftretenswahrscheinlichkeiten deutet an, dass Philip Morris tabakbezogenes, d. h. feldspezifisches Vokabular zugunsten der organisationalen Selbstbeschreibung in seinen Stelleninseraten in den Hintergrund rückt.

Insgesamt fördert das Modell vier Gruppen von Topics zutage: Eine erste Gruppe von Topics umfasst Begriffe, die Datenwissenschaften primär durch Methoden bzw. epistemische Praktiken repräsentieren. Diese finden transversal Anwendung und können nicht einzelnen Feldern zugeordnet werden. Dazu zählen Project Management (T1), Internships/Science (T5), Business Reporting (T8), Data Analysis (T10), Machine Learning (T12), Digital Marketing (T14), Software Engineering (T16) sowie Business Analytics (T19). Die Topics weisen eine relativ ausgeglichene Verteilung auf, d. h., kein Feld hat einen Anteil über 50 %. Die niedrigsten Werte weisen die Topics Machine Learning (16 %), Business Analytics (19.3 %), Data Analysis (21.7 %) auf.¹⁶ Die geringe Konzentration der drei methodischen Topics auf einzelne Felder deutet an, dass methodenbezogene Inhalte in Stellenanzeigen die am wenigsten feldspezifischen sind (vgl. unten).

Weiter existieren Konvolute epistemischer Praktiken in Verbindung mit einzelnen Feldern. Dabei kann zwischen Topics des Pharmafeldes und solchen mit Beziehungen zu anderen Feldern differenziert werden. Die Gruppe der pharmabezogenen Topics bilden Clinical Development (T3), Research/Biology (T7), Basel/Roche (T9) sowie Development/Laboratory (T18). T3 fokussiert vor allem klinische Studien und laborbezogene Forschungspraktiken, während T18 stärker Begriffe von Entwicklung und Produktion repräsentiert. T7 indiziert in erster Linie die Konvergenz von datengetriebener Forschung in Disziplinen wie Biologie oder Medizin (durch Begriffe wie *biolog, research, translat, medicin, bioinforma*) (Leonelli 2014, 2016; Trajanoski 2012). Schliesslich umfasst T9 Begriffe, die für Stelleninserate der Firma Roche (die häufigste Arbeitgeberin im Sample, vgl. Tabelle 3) typisch sind. Der Anteil des pharmazeutischen Feldes an diesen Inseraten ist mit Werten zwischen 72 % (T18) und 97 % (T9) sehr hoch. In geringerem Masse beschreiben die Topics ferner das akademische Feld sowie die Felder Gesundheit, Industrie und Lebensmittel, was auf eine gewisse Ähnlichkeit hinsichtlich epistemischer Praktiken und Praxisarrangements (wie Forschungslaboratorien) hinweist.

Eine dritte Gruppe umfasst feldspezifische Praktiken wie Consulting Services (T4), Research/Science (T13), Tobacco/Philip Morris (T6), Recruitment (T17), Risk Management/Banking (T20), wobei die Namensgebung der Topics bereits das jeweilige Feld indiziert. Charakteristisch für diese Topics ist, dass epistemische Praktiken, Begriffe, Organisationsnamen und Standorte gemeinsam figurieren. Exemplarisch verdeutlicht dies T20: Für das Bankenfeld typische Begriffe (wie *risk, finance, bank, invest, asset, credit, wealth*) treten hier zusammen mit Standorten (*zurich, switzerland, suisse*) und Organisationsbezeichnungen (*credit, suisse, ubs*) auf. Die feldspezifischen Auftretenswahrscheinlichkeiten liegen bei diesen Topics zwischen 61.8 % (T4) und 96 % (T13).

¹⁶ Es folgen Business Reporting (27.8 %), Software Engineering (28.9 %) sowie Digital Marketing (30.9 %).

Tabelle 5: Topic-Modell der englischsprachigen Stellenanzeigen über 20 Topics mit den 20 häufigsten Begriffen (T = Topic) (n = Anzahl Dokumente)

Project Management (T1)	Diversity/ Inclusion (T2)	Clinical Development (T3)	Consulting Services (T4)	Internships/Science (T5)	Tobacco/ Philip Morris (T6)
n=96	n=88	n=114	n=123	n=129	n=77
manag	work	clinic	servic	work	product
project	applic	studi	client	intern	intern
lead	inform	develop	consult	internship	work
team	googl	medic	team	project	pmi
strategi	experi	safeti	work	opportun	develop
across	skill	report	busi	start	team
global	opportun	review	manag	motiv	compani
ensur	employ	health	career	applic	futur
partner	status	includ	profession	program	present
abil	employe	regulatori	challeng	develop	market
drive	zurich	trial	develop	offer	world
strateg	nation	scientif	solut	industri	employ
plan	orient	statist	industri	letter	lead
develop	requir	plan	offer	date	continu
organ	recruit	activ	experi	month	requir
execut	protect	respons	opportun	environ	appli
leadership	develop	intern	project	person	global
implement	qualif	document	audit	master	potenti
respons	compani	scienc	technolog	lausann	switzerland
effect	locat	de	intern	interest	build

Research/ Biology (T7)	Business Reporting (T8)	Basel/Roche (T9)	Data Analysis (T10)
n=171	n=158	n=221	n=83
biolog	manag	basel	data
job	report	job	experi
research	requir	develop	analysi
develop	process	roch	tool
scientist	busi	switzerland	databas
drug	qualiti	research	process
scientif	support	function	sourc
experi	system	clinic	model
discoveri	master	digit	use
team	oper	patient	work
translat	analyst	opportun	knowledg
medicin	ensur	headquart	analys
high	skill	healthcar	sql
novarti	global	fact	visual
molecular	excel	basel-town	strong
analysi	improv	biomark	understand
bioinformat	provid	appli	collect
appli	respons	pharmaceut	analyst
includ	suppli	scientist	set
candid	inform	innov	integr

Fortsetzung Tabelle 5

Working Environment (T11)	Machine Learning (T12)	Research/Science (T13)	Digital Marketing (T14)	Working Conditions (T15)
n=173	n=150	n=227	n=139	n=29
team	learn	research	market	work
work	machin	applic	custom	benefit
peopl	scienc	scienc	digit	flexibl
product	statist	posit	compani	job
world	model	univers	sale	swisscom
creat	scientist	inform	product	offer
technolog	comput	institut	innov	data
develop	python	pleas	global	continuu
build	experi	candid	develop	care
offic	r	comput	manag	addit
passion	algorithm	includ	base	system
grow	mathemat	appli	group	team
custom	appli	phd	job	compani
environ	program	group	industri	salari
compani	languag	contact	perform	programm
join	field	zurich	opportun	share
part	problem	swiss	strategi	famili
travel	deep	onlin	strong	product
us	engin	project	year	achiev
user	predict	eth	support	employ

Software Engineering (T16)	Recruitment (T17)	Development/ Lab (T18)	Business Analytics (T19)	Risk Management/ Banking (T20)
n=298	n=162	n=175	n=88	n=240
experi	experi	develop	analyt	manag
softwar	look	process	busi	risk
engin	skill	experi	data	team
develop	ag	product	solut	financi
technolog	pleas	technic	insight	bank
big	project	manufactur	advanc	work
solut	english	scientist	technolog	skill
design	posit	support	skill	support
platform	zürich	laboratori	intellig	invest
system	work	analyt	abil	asset
architectur	interest	knowledg	scienc	busi
technic	contact	work	tool	us
applic	send	perform	understand	within
servic	profession	respons	experi	challeng
cloud	german	lab	inform	offer
knowledg	switzerland	valid	drive	client
team	us	method	bi	financ
integr	year	activ	transform	function
secur	cv	evalu	design	corpor
comput	client	life	identifi	credit

Schliesslich umfasst eine vierte Gruppe von Topics stelleninsetratebezogene Begriffe, die für die Analyse nicht relevant sind.¹⁷ Dazu zählen Diversity/Inclusion (T2), Working Environment (T11) sowie Working Conditions (T15). Sie sind zu einem geringen Anteil in den meisten Dokumenten vorhanden, haben inhaltlich jedoch keine Bedeutung. Die Distribution über die einzelnen Felder ist deshalb eher breit (bei T2 und T11 liegen die Anteile jeweils unter 30 %).¹⁸ T2 beschreibt im Wesentlichen Diversity-Statements am Ende von Stellenanzeigen, was eine sprachlich-kulturelle Konvention im englischsprachigen Raum darstellt (Dobbin et al. 2015), die Organisationen in deutschsprachigen Inseraten kaum anwenden. Die Stelleninsetrate, die durch dieses Topic gut beschrieben werden, stammen denn auch überwiegend von Grosskonzernen (Google, Zurich, eBay, Wipro, Biogen oder Siemens) und internationalen Organisationen (UNO, WHO), die ihre Rekrutierungsaktivitäten primär auf internationale Arbeitsmärkte fokussieren.

5.3.2 Die feldübergreifende Zentralität der Methoden-Topics

Nach der Beschreibung der Topics gilt es die Ergebnisse zu interpretieren und zu plausibilisieren (DiMaggio et al. 2013: 583ff.). Die Entwicklung der relativen Häufigkeiten der Topics ist nicht wirklich sinnvoll interpretierbar, da sich visuell keine signifikanten Verschiebungen bestimmter Topics erkennen lassen (vgl. Abbildung 13 im Anhang). Anstelle der Gesamtheit der Topics untersuche ich deshalb vertieft die drei zentralen Methoden-Topics Data Analysis (T10), Machine Learning (T12) sowie Software Engineering (T16). Die durchschnittliche relative Häufigkeit – und damit die Bedeutung über alle Topics hinweg – steigt bei T16 sowie T12 über den Beobachtungszeitraum leicht an, während sie bei T10 konstant bleibt.¹⁹

Inhaltlich macht ein Vergleich der drei Topics-Kombinationen von Praktiken, Tools, Disziplinen und weiteren Kategorien erkennbar: T10 umfasst Praktiken wie *use*, *model*, *analys*, *understand*, *collect*, *integrate*, *analyz*, *integrat* oder *handl* sowie Instrumente wie *tool*, *datas*, *sql*, *warehous*, *access* oder *excel*, während explizite Nennungen bestimmter Felder oder Disziplinen fehlen. T12 hingegen sind oft Terme zugeordnet, die für unterschiedliche soziale (*scienc*, *scientist*, *industri*) bzw. disziplinäre Felder (wie *statist*, *comput*, *mathemat*, *engin*, *physic*) stehen. Enthalten sind auch Terme wie *learn*, *model*, *comput*, *program*, *predict*, *use* oder *develop*, die epistemische Praktiken andeuten, die als zentral für die Datenwissenschaften gesetzt werden. Bezogen auf Tools stechen *python*, *r* oder *c* [für C++] ins Auge. T16 schliesslich wird durch hohe Wahrscheinlichkeiten von informatik- und ingenieurwissenschaftlichen Begrifflichkeiten wie *softwar*, *engin*, *develop*, *technolog*, *solut*, *platform*, *system*, *architectur*, *technic*, *applic*,

17 Maier et al. (2018: 108) und DiMaggio et al. (2013: 586) sprechen von »boilerplate topics« bzw. »boilerplate terms«, die keine inhaltliche Bedeutung haben, jedoch andere, inhaltlich relevante Topics schärfen.

18 T15 tritt nur selten auf und umfasst Stellenanzeigen von zwei Unternehmen (Swisscom sowie Siemens).

19 T10 weist (Spannweite $R = 0.02$) die grösste Konstanz auf, während T12 ($R = 0.037$) sowie T16 ($R = 0.041$) stärkeren Schwankungen unterliegen. Die Spannweite $R = x_{\max} - x_{\min}$ ist ein Streuungsmass, das sich aus der Differenz zwischen dem höchsten und dem niedrigsten Wert einer Verteilung ergibt (Diaz-Bone & Weischer 2015: 389).

servic oder *cloud* charakterisiert. Nachgelagert sind auch Praktiken wie *develop*, *design*, *integr*, *secur*, *comput*, *implement*, *test*, *deploy* und *distribut* sowie Tools wie *hadoop*, *java*, *spark*, *python* und *linux* enthalten.

Das Modell segmentiert demnach für Datenwissenschaften zentrale Praktiken, Disziplinen, Tools und weitere Begriffe in den drei Topics. Es umreißt damit drei methodische Praktiken, die feldunabhängig durch die Berufsgruppen der Data Scientists (T12), Data Analysts (T10) und Data Engineers (T16) repräsentiert werden. Die Benennung des Topics als »Machine Learning« deutet an, dass jenes Begriffskonvolut als zentral innerhalb der Datenwissenschaften identifiziert wird. Dadurch wird auch klar, weshalb der Begriff »Machine Learning« das Sample adäquater abbildet als die eigentlich angeführten Suchbegriffe »Data Science« oder »Big Data« (vgl. Kap. 5.2.1). Sowohl die geringe Konzentration auf einzelne Felder als auch die überdurchschnittlichen Auftretenswahrscheinlichkeiten der drei Topics verweisen darauf, dass die in-sierierenden Unternehmen und Organisationen Methoden über die Felder hinweg als relevant konstruieren.²⁰ Damit wird eine bestimmte Deutung der Datenwissenschaften anderen möglichen Alternativen vorgezogen und fixiert. Allerdings bleibt trotz der Bedeutung der Methoden-Topics ihr kumulierter Anteil im Vergleich zu den anderen Topic-Gruppen minoritär. Methoden sind für die Datenwissenschaften zwar wichtig, allerdings sind darüber hinaus epistemische Praktiken mit Bezug zu bestimmten, vor allem ökonomischen Feldern ebenso zentral.²¹

5.4 Repräsentationen der Datenwissenschaften in deutschsprachigen Stellenanzeigen

Nach den englischsprachigen untersuche ich nun die deutschsprachigen Stellenanzeigen. Zunächst beschreibe ich das identifizierte Topic-Modell und die Ausprägungen der einzelnen Topics, wobei englischsprachige Begriffe vor allem in den methodischen und ökonomischen Topics dominieren (Kap. 5.4.1). Danach gehe ich anhand eines Beispielinstruments auf die Offenheit der verwendeten Begriffe ein (Kap. 5.4.2).

5.4.1 Die Dominanz englischsprachiger Begriffe in deutschsprachigen Stellenanzeigen

Tabelle 6 gibt einen Gesamtüberblick über die prägenden thematischen Makrostrukturen. Darin enthalten sind die elf Topics des berechneten Modells der deutschsprachigen Stellenanzeigen mit den jeweils zwanzig wahrscheinlichsten Begriffen. Das

20 Das Ergebnis deckt sich insofern mit der Curricula-Analyse, wonach Methoden transdisziplinär unterrichtet werden, während die inhaltliche Vertiefung (*domain knowledge*) in den einzelnen Disziplinen erfolgt (vgl. Kap. 9).

21 Bezüglich berufsunspezifischer Anforderungen ist zu beobachten, dass kommunikative und soziale Kompetenzen sowie persönliche Eigenschaften wie Kreativität oder Neugier sowohl im englisch- als auch im deutschsprachigen Modell erscheinen. Sie lassen sich jedoch keinem Topic ausschliesslich zuordnen, sondern sind jeweils über mehrere verteilt. Obwohl dies auf ihre allgemeine Bedeutung über datenbezogene Berufsfelder hinaus verweist (Salvisberg 2010), haben Soft Skills im untersuchten Korpus somit nur eine geringe Relevanz.

Modell ist breiter verteilt, auch weil kein (ökonomisches) Feld dominiert wie im Falle des englischsprachigen Samples: Lediglich zwei Topics weisen einen feldspezifischen Anteil über 50 % auf, nämlich T8 Arbeitsumfeld/Swisscom²² und T4 Datenwissenschaften/Forschung. T8, T2 Verantwortung/Entwicklung und T11 Stelleninsetrate zählen zu jenen Topics, die für Stelleninsetrate typisches Vokabular umfassen und nur wenige Dokumente umfassend beschreiben. T5 Schweiz – International enthält vor allem Begriffe, die das Verhältnis der Schweiz zur Welt charakterisieren.²³ Hier liegt die Annahme nahe, dass es sich um Insetrate von Organisationen handelt, die Arbeitskräfte für Praktiken des Austausch mit internationalen Handelspartner*innen und anderen Akteur*innen suchen – was bei englischsprachigen Anzeigen möglicherweise zwar enthalten ist, aber nicht explizit angegeben wird. Demgegenüber beschreibt T4 Inhalte, die für die Datenwissenschaften im akademischen Feld stehen, ohne allerdings Methoden oder epistemische Praktiken zu nennen. Neben Wissenschaft und Bildung (51 % der Dokumente in T4) sind darin ferner der öffentliche Dienst (18 %) sowie das Gesundheitswesen (14 %) vertreten.

Neben T4 segmentiert das deutschsprachige Modell jene Topics sehr ähnlich wie im Falle des englischsprachigen Samples, die epistemische Praktiken in den Stelleninsetraten beschreiben: Dazu gehören T1 Big Data/Software Engineering, T3 Data Science/Machine Learning, T6 Business Analytics, T7 Datenanalyse/Reporting sowie T9 Digital Marketing. Der Umstand, dass die Topics sprachübergreifend ähnlich differenziert werden, deutet darauf hin, dass sich datenbezogene Praktiken in verschiedenen Sprachräumen etabliert haben und Anwendung finden. T1 und T3 zeichnen sich vor allem durch viele englischsprachige Terme, die Nennung spezifischer Disziplinen und Tools unter den wahrscheinlichen Begriffen aus. Analog zum englischsprachigen Modell repräsentieren die zwei Topics epistemische Praktiken, die feldübergreifend Anwendung finden (die Anteile liegen jeweils unter 30 %) und nach T7 die höchste Auftretenswahrscheinlichkeit haben. Die zwei Methoden-Topics sind demnach etwas öfter als im statistischen Durchschnitt zu erwarten, erklären aber auch im deutschsprachigen Modell lediglich 20 % der Inhalte.

T7 und T9 beinhalten – anders als beim englischsprachigen Modell – nur wenig methodisch-technisches Vokabular oder bestimmte Tools, dafür hohe Anteile an Begriffen, die in ökonomischen Feldern verwendet werden. T9 umfasst ausserdem Terme, die mit der digitalen Transformation assoziiert werden, während T10 Innovation/Produktentwicklung vor allem die Entwicklung neuer Produkte und Innovation beschreibt. T7, T9 und T10 weisen alle niedrige feldspezifische Anteile auf ($\leq 30\%$), während T6 vergleichsweise stark im Consulting-Feld (Muzio et al. 2011; Schmidt-Wellenburg 2013) verankert ist. T6, T7, T9 und T10 bilden zusammen die businessbezogenen Topics im Modell der deutschsprachigen Insetrate. Insgesamt übertrifft die kumulierte durchschnittliche Auftretenswahrscheinlichkeit dieser für die ökonomischen Felder relevanten Topics – analog zum englischsprachigen Modell – diejenige der methodenbezogenen Topics deutlich.

22 T8 wird als einziges Topic im deutschsprachigen Modell massgeblich von einem Unternehmen geprägt, da 97 % der Dokumente von der Swisscom stammen.

23 Ein ähnliches Themenfeld fördert auch die Analyse der politischen Dokumente zutage (vgl. Kap. 6.3.4).

Tabelle 6: Topic-Modell der deutschsprachigen Stellenanzeigen über 11 Topics mit den 20 häufigsten Begriffen ($T = \text{Topic}$) ($n = \text{Anzahl Dokumente}$)

Big Data/Software Engineering (T1)	Verantwortung/Entwicklung (T2)	Data Science/Machine Learning (T3)	Data Science/Forschung (T4)	Schweiz - International (T5)	Business Analytics (T6)
n=184	n=55	n=167	n=125	n=125	n=65
data	uns	data	bereich	gut	uns
big	schweiz	statist	data	kenntnis	kund
erfahr	mitarbeit	scienc	verfug	international	business
softwar	arbeit	scientist	scienc	ag	analytics
losung	verschied	analytics	informat	deutsch	consultant
engine	bereich	dat	weit	such	intelligenc
kenntnis	verantwort	r	information	bereich	sap
entwickl	sbb	neu	per	prozess	bereich
sql	vielfalt	python	gern	data	data
technologi	sich	analys	forschung	schweiz	person
kund	entwickl	mathemat	wissenschaft	aufgab	arbeit
technisch	stell	erfahr	vereinbar	fuhrend	projekt
design	moglich	learning	team	resourc	mitarbeit
informat	wichtig	machin	erfahr	m	hast
bern	gut	method	hochschul	durchfuhr	mehr
microsoft	beweg	kenntnis	mitarbeit	erfahr	bi
datenbank	biet	z.b	entwickl	global	technisch
team	dat	analyt	auskunft	team	senior
zurich	unternehmen	modell	wirtschaft	human	losung
w	intern	w	engineering	manag	mocht

Fortsetzung Tabelle 6

Datenanalyse/ Reporting (T7)	Arbeitsumfeld/ Swisscom (T8)	Digital Marketing (T9)	Innovation/ Produkte (T10)	Stelleninserate (T11)
n=203	n=32	n=112	n=102	n=112
gut	arbeit	digital	uns	zurich
erstell	swisscom	kund	neu	ag
analys	gut	business	team	such
analyst	woch	erfahr	arbeit	aufgab
hoh	kann	data	produkt	m
analyt	beruf	analytics	bring	partn
kenntnis	uns	schweiz	zukunft	per
weiterentwickl	grosszug	management	gross	spannend
vorteil	unterstutz	m	innovativ	tatig
fahig	zusatz	neu	biet	profil
intern	famili	zurich	entwickeln	unterstutz
dat	fair	w	zusamm	bereich
business	weit	marketing	unternehmen	uns
ausgepragt	moglich	digitalisier	entwickl	deutsch
management	vorteil	strateg	hast	w
reporting	wichtig	projekt	ag	freu
erfahr	profitierst	manag	herausforder	interess
person	leistung	berat	ide	unternehmen
arbeitsweis	arbeitsbeding	umsetz	gemeinsam	send
auswert	gewahr	innovation	erst	bewerbungsunterlag

5.4.2 Die Offenheit der verwendeten Begriffe

Im Folgenden untersuche ich am Beispiel des bereits eingeführten Beispielinserats (vgl. Anhang), wie durch offene, vieldeutige Begriffe Relationen zwischen verschiedenen Profilen und Tätigkeitsfeldern hergestellt werden, die im Sinne ambiguer Kategorien fundierend für die Datenwissenschaften sind.

Das Stelleninserat mit dem Titel »Data Scientist/Analyst (w/m) BMW Financial Services« stammt von der BMW Group Switzerland. Im Rahmen der Datenerhebung habe ich es am 1. September 2017 auf der Online-Stellenplattform jobs.ch erfasst. Bereits der Stellentitel markiert eine gewisse Unentschiedenheit, ob ein/e Data Scientist, ein/e Data Analyst oder beides gesucht wird. Dies kann eine Strategie der inserierenden Organisation sein, erst aufgrund der eingehenden Bewerbungen für eine bestimmte Berufsbezeichnung zu optieren. Da der Inseratetext weder eine Präferenz noch eine Differenz (im Sinne einer oder-Formulierung) ausdrückt, sondern an mehreren Stellen von einer »Position« bzw. einer »Persönlichkeit« spricht, artikuliert sich darin eine Offenheit über die Abgrenzung der zwei Profile bzw. Berufsgruppen, die sowohl mit unterschiedlichen epistemischen Praktiken, Objekten sowie ökonomischen Feldern verbunden werden.

Aufgrund der inserierenden Organisation habe ich das Inserat dem Transport- und Logistikfeld zugeordnet. Die Angabe »Financial Services« im Titel indiziert allerdings ein inhaltliches Praxisfeld, das eher mit Finanzdienstleistungen und damit dem Banken-/FinTech-Feld assoziiert ist. Hier bietet sich ein Rückgriff auf die identifizierten Topics an: Das Topic-Modell der deutschsprachigen Inserate beschreibt das Transport- und Logistikfeld vor allem durch die Topics 2 und 5, während jene des Banken- und Finanzdienstleistungsfeldes hohe Anteile der Topics 7 und 9 aufweisen (vgl. Tabelle 6). Das gewählte Verfahren beschreibt aber nicht nur auf der aggregierten Feldebene, sondern auch einzelne Inserate als Zusammenstellung unterschiedlicher Themen.

So setzt sich das Beispielinserat gemäss dem berechneten Topic-Modell als Kombination aus 23 % T9 Digital Marketing, 16 % T7 Datenanalyse/Reporting, 15 % T3 Data Science/Machine Learning, 12 % T11 Stelleninserate zusammen. Die restlichen sieben Topics (≤ 8 %) machen 34 % der Inhalte aus. Der bedeutende Anteil von T9 wird insbesondere in den Abschnitten [3] und [4] deutlich: Die Position ist im »Vertriebs- und Marketingteam« angesiedelt und zählt unter anderem die Analyse und Bewertung der »Vertriebs- und Marketingperformance«, die Erstellung »komplexe[r] Vertriebs-, Markt- und Wettbewerbsanalysen« oder die Verantwortung der »Pricingstrategie« als »Aufgabengebiet« auf. T7 und T3 finden sich primär in den Abschnitten [5] und [6]: Disziplinäre und methodische Terme (»systematische Muster«, »cluster«, »Information Retrieval«, »Machine Learning«, »Naturwissenschaft«, »Statistik«, »statistische[r] Software, R, Python, SAS oder SPSS«) sind T3 zuzurechnen, während T7 eher das betriebswirtschaftliche Vokabular (»Kundensegmente [...], Trends, Kundenwerte und Handlungsbedarf (next best action)«) sowie Qualifikationsanforderungen (»betriebswirtschaftliche Ausbildung (HF, FH, Uni) mit vertieften Kenntnissen in [...] Analysen und Controlling«) erfasst. T11 verweist schliesslich auf einen vergleichsweise hohen Anteil stelleninseratetypischen Vokabulars (Abschnitte [3], [6] und [7]). Über den gesamten Inseratetext dominieren also betriebswirtschaftliche, marketingbezogene Inhalte, während das explizit methodenbezogene Vokabular dieses komplementiert.

Das Topic-Modell gibt keine Antwort darauf, ob in diesem Inserat eher ein/e Data Scientist oder ein/e Data Analyst gesucht wird, sondern identifiziert die Inhalte, die durch diese Topics beschrieben werden, als quasi gleichwertig. Damit löst die Analyse des Inseratetextes auch die Unklarheit bezüglich der Kategorisierung der Stellenbezeichnung nicht auf. Obwohl die Differenz *Data Scientist* oder *Data Analyst* offengelassen wird, weisen Stellentitel und Text dennoch eine inhaltliche Kohärenz auf, die durch das deutschsprachige Topic-Modell und die Verteilung der Inhalte nachvoll-

ziehbar wird. Die Offenheit bezüglich der in diesem Inserat verwendeten Kategorien ist jedoch konstitutiv, da sie multiple Handlungs-, d. h. in diesem Fall Rekrutierungsoptionen bereithält. Ambiguität (Eisenberg 1984; Leitch & Davenport 2007) in Stellenanzeigen als organisationalen Praktiken wird so zu einem zentralen Merkmal in der Herausbildung der Datenwissenschaften.

5.5 Diskussion

Das Kapitel hat untersucht, wie Akteur*innen in unterschiedlichen Feldern durch die Publikation von Stellenanzeigen multiple Bedeutungen der Datenwissenschaften konstruieren. Ich habe zunächst gefragt, wie das Segment der Datenwissenschaften im Arbeitsmarkt strukturiert ist und welche Organisationen in welchen Feldern Stelleninserate für Data Scientists publizieren. Danach erschloss die Analyse die unterschiedlichen Profile von Data Scientists, die in den Inseraten konstruiert werden. Schliesslich interessierte sich das Kapitel dafür, welche übergeordneten Deutungen der Datenwissenschaften in den Stelleninseraten entworfen und fixiert werden.

Die Ergebnisse der deskriptiven Analyse verweisen einerseits auf die zunehmende Bedeutung der Datenwissenschaften während des Untersuchungszeitraums. Andererseits indizieren sie eine ausgeprägte Konzentration der Stelleninserate auf bestimmte Grossunternehmen in ökonomischen Feldern, die in urbanen Standorten wie Zürich und Basel lokalisiert sind. Der hohe Anteil englischsprachiger Inserate im Sample kann als ein Indikator für die starke Internationalisierung des Arbeitsmarktsegments interpretiert werden. Insbesondere Organisationen in den ökonomischen Feldern Arzneimittelherstellung, Banken, IT/Software und Consulting sowie im akademischen Feld publizieren besonders viele Stelleninserate für Data Scientists. Dies kann als Indiz dafür gelesen werden, dass sie stärker als andere datafiziert sind. Es zeigen sich zudem signifikante Unterschiede zu Organisationen in anderen Feldern, Landesteilen und Sprachregionen, die noch weniger durch Datenpraktiken durchdrungen sind. Insofern manifestieren sich erhebliche Differenzen in der Diffusion und Adoption datenwissenschaftlicher Expertise und Praktiken in den einzelnen Feldern.

Über die Titel der Stelleninserate können verschiedene Gruppen von datenbezogenen Berufsbezeichnungen differenziert werden, die im Wesentlichen jenen in der Literatur geforderten und identifizierten Profilen von Data Scientists entsprechen. Über die Rollenbezeichnungen hinaus lassen sich diese einerseits den genannten (ökonomischen) Feldern und Standorten zuweisen. Andererseits sind die Profile durch distinkte Begriffskonvolute charakterisiert, die epistemische Praktiken, disziplinäre Wissensbestände, methodische Vorgehensweisen und Objekte miteinander verknüpfen. »Data Science« und »Big Data« als der Analyse zugrundeliegende Suchterme erweisen sich damit als inklusive Begriffe, durch die multiple Alternativen formuliert werden. Solche Praktiken von Begriffsbearbeitung eröffnen den inserierenden Organisationen neue Möglichkeitsräume und erweitern somit deren Handlungs- bzw. Rekrutierungsoptionen.

Gegenüber der strukturellen Konzentration der Stelleninserate fördert das Topic-Modeling-Verfahren auf der inhaltlichen Ebene sowohl für das englisch- als auch das deutschsprachige Sample die Multidimensionalität der Datenwissenschaften zutage. Die computergestützte Textanalyse der Stellenanzeigen indiziert eine grosse Vielfalt

unterschiedlicher Konnotationen, durch die die Datenwissenschaften in den Topics repräsentiert werden. Dies sind zum einen Topics, die feldübergreifend wissenschaftliche Methoden und epistemische Praktiken beschreiben. Zum anderen lassen sich Topics differenzieren, die für Wissensinhalte, Begriffe und Objekte in bestimmten Feldern, namentlich im pharmazeutischen, im Banken- sowie im akademischen Feld stehen. Die Heterogenität der identifizierten Topics bestätigt insofern die theoretische Annahme, dass sich – trotz der Dominanz der ökonomischen Felder hinsichtlich der Struktur des Arbeitsmarktsegments – die Datenwissenschaften weder auf ein soziales Feld noch auf eine Disziplin reduzieren lassen.

Insofern ist die Prävalenz pharmabezogener Topics, die im englischsprachigen Modell einen gewichtigen Anteil der Inhalte prägen, erklärungsbedürftig: Deren Auftretenswahrscheinlichkeiten schwanken in der Untersuchungsperiode zwar relativ stark, weisen jedoch eine ähnliche Spannweite auf wie die Methoden-Totics. Die Relevanz der Topics bemisst sich in der hohen Anzahl von Stellenanzeigen des Pharmafeldes, das über ein Viertel des englischsprachigen Samples verantwortet.²⁴ Zwar werden die Naturwissenschaften bisweilen als wichtige Bezugsdisziplinen der Datenwissenschaften genannt, allerdings deutlich seltener als Statistik, Mathematik und Informatik. Gerade laborbezogene epistemische Praktiken finden in der Diskussion um die notwendigen Kompetenzen für Datenwissenschaftler*innen kaum je Erwähnung. Möglicherweise spiegelt das Ergebnis die datengetriebene Transformation der für die Pharmaindustrie relevanten epistemischen Felder wider (Leonelli 2016; Trajanoski 2012), was sowohl hinsichtlich der ökonomischen Wertschöpfung als auch der kollektiven Selbstbeschreibung der Schweiz von elementarer Bedeutung ist.

In den zentralen methodischen Topics werden Praktiken, Tools und Disziplinen, die in der Diskussion gemeinhin mit Datenwissenschaften assoziiert werden, in den sprachgetrennt berechneten Modellen separat segmentiert. Die Kombination disziplinen- und feldspezifischer Begriffe, Methoden und Tools in diesen Topics verweist darauf, dass sich in der Diskussion bestimmte Kategorien etabliert haben, die in Stelleninseraten feldübergreifend und sprachunabhängig Verwendung finden. Sowohl die Präsenz englischsprachiger Rollenbezeichnungen als auch die Frequenzen englischsprachiger Terme in den Topics des deutschsprachigen Modells machen klar, dass es sich dabei um englischsprachige Kategorien handelt, die von ihrem Popularisierungskontext im Feld der US-amerikanischen Technologiekonzerne in diverse ökonomische und soziale Felder diffundiert sind. Die Analyse der Stelleninserate verdeutlicht auch, dass Methoden, Analysetools oder Qualifikationsanforderungen oft austauschbar sind und in nicht-hierarchisierenden Aufzählungen angeführt werden. Solche Listen rücken die unklare Verortung der Datenwissenschaften in den Vordergrund: Da für viele Organisationen nach wie vor Unsicherheit über die zentralen Begriffe besteht, signalisiert das Changieren zwischen verschiedenen Vorgehensweisen, Werkzeugen oder Bildungszertifikaten Inklusivität methodischer Expertisen sowie disziplinärer Wissensbestände. Die in den Stelleninseraten zu beobachtenden Aufzählungen äquivalenter Alternativen operieren in diesem Sinne als Praktiken der Grenzüberschreitung, die darauf abzielen, möglichst viele Wissensbereiche und Expertisen zu inkludieren.

24 Das Ergebnis könnte auch eine Folge der Suchstrategie mit den Begriffen »data« und »scientist« sein, indem der Suchalgorithmus der Jobplattform die zwei Begriffe mit Inseraten der Pharmaindustrie assoziierte und insofern zu einer Überrepräsentation dieses Feldes führte.

dieren und den Gegenstand flexibel, adaptiv und offen zu halten. Damit markiert sich eine erste Differenz zum akademischen Feld, in dem die Synchronizität divergierender Praktiken der Grenzziehung und Grenzüberschreitung fundamental in die Herausbildung der universitären Curricula eingeschrieben ist (vgl. Kap. 9).

Trotz der Bedeutung jener Topics, die Datenwissenschaften als Kombinationen methodenbezogener Begriffe und Praktiken repräsentieren, bleibt deren kumulierter Anteil in beiden Modellen in Relation zu anderen Topics minoritär. In der Verknüpfung von methodischen Praktiken und ökonomischen Wissensinhalten manifestiert sich die Priorität der oft eingeforderten Anwendungskomponente. Dies zeigt sich ferner in der Unabhängigkeit der Methoden von jenen Topics, die für das institutionelle akademische Feld stehen, ohne bestimmte Disziplinen oder Methoden zu beschreiben. Somit ergibt sich eine weitere Differenz zum akademischen Feld, wo in den Curricula die methodische Ausbildung im Kernbereich durch eine *domain knowledge* in Wahlfächern ergänzt wird. Im Korpus der Stelleninserate bilden hingegen die feldspezifischen Inhalte der *domains* den Kern, während die methodischen Kompetenzen diesen komplementieren.

Insgesamt setzen Stellenanzeigen eine Vielzahl unterschiedlicher Organisationen, Felder, Praktiken, Objekte und Begriffe miteinander in Beziehung, ohne die jeweiligen Relationen zwischen diesen zu definieren. Vielmehr sind die verwendeten Begriffe feld- und sprachübergreifend durch Anpassungsfähigkeit und Mehrdeutigkeit charakterisiert, was den Akteur*innen affirmative Bezugnahmen erlaubt, wobei die Bedeutung der jeweiligen Begriffe von Inserat zu Inserat, von Feld zu Feld differieren kann. Die Offenheit und Ambiguität der verwendeten Begriffe ist konstitutiv für das untersuchte Segment des schweizerischen Arbeitsmarktes und fundiert insofern die Emergenz der Datenwissenschaften als ein Wissensgebiet, das sprachliche Grenzziehungen transzendieren und verschiedene etablierte Felder miteinander verknüpfen kann.

Kapitel 6 – Zukunftsentwürfe der Datenwissenschaften in Diskursen der Bildungs- und Forschungspolitik

»Die Schweiz muss ihr Know-how und ihre Kompetenzen in Datenwissenschaften stärken. Darunter versteht man die wissenschaftliche Nutzung und den sicheren Umgang mit riesigen Datenmengen (Big Data), um damit zu neuen Erkenntnissen zu gelangen. Intensivierte Forschung und Lehre im Gebiet der Datenwissenschaften trägt dazu bei, globale Herausforderungen der Gesellschaft wie Gesundheit, Energie- und Nahrungsmittelversorgung sowie den Umgang mit Ressourcen zu bewältigen« (ETH-Rat 2016b: 2).

6.1 Einleitung

In den letzten Jahren ist der bildungs- und forschungspolitische Diskurs der Schweiz eminent durch die Digitalisierung geprägt. Universitäten, wissenschaftliche Akademien, Wirtschaftsverbände und staatliche Behörden formulieren »digitale Strategien« und Aktionspläne, um sich der »Chancen und Herausforderungen der Digitalisierung für Bildung und Forschung« (SBFI 2017) anzunehmen. Die Diskussion über die Digitalisierung ist Teil soziotechnischer Imaginationen: Politische, ökonomische und wissenschaftliche Akteur*innen entwerfen Zukunftsvisionen, in denen soziale Relationen von und zu digitalen Technologien beschrieben und gerahmt werden (Jasanoff 2015; Jasanoff & Kim 2015; Meyer 2020). Das Formulieren von politischen Strategien und Zielen sowie das Beschliessen von Massnahmen umfasst dabei sowohl diskursive wie nicht-diskursive Praktiken: Indem die politischen Akteur*innen die zukünftige Entwicklung gesellschaftlicher Bereiche skizzieren, nehmen sie auch eine Wertung und Zuweisung von Aufmerksamkeit, finanziellen und weiteren Ressourcen vor (Beckert 2016; Jasanoff 2015). Ihre Absichtserklärungen und Investitionen signalisieren anderen Akteur*innen, dass sich ihrerseits Aktivitäten in diesen Bereichen auszahlen. Dies zieht wiederum Akteur*innen aus angrenzenden Feldern an, die sich rasche und hohe Profite versprechen. Die Beteiligten bemühen sich deshalb trotz möglicherweise divergierender Interessen darum, den Raum offen und anschlussfähig zu halten, ohne dass es zu einer Institutionalisierung, etwa durch politische Regulierung, kommt.

Ich argumentiere dafür, die Genese der Datenwissenschaften im umfassenderen politökonomischen Kontext der Digitalisierung zu verorten. Im Zentrum des Kapitels stehen die kollektiven Stellungnahmen von Akteur*innen im Feld der Politik, die stets

als Kompromissprodukte konkurrierender Positionen in den jeweiligen Organisationen zu betrachten sind. Gleichzeitig sind die Stellungnahmen durch Vielstimmigkeit und Multiperspektivität gekennzeichnet und eröffnen Möglichkeiten zur Kooperation mit anderen Akteur*innen. Sowohl durch stärker konflikthafte als auch durch kooperative Praktiken tragen sie zur Konstitution und Permanenz der Datenwissenschaften als neues Wissensgebiet bei. Wie ich noch zeigen werde, geht der politische Diskurs zur Digitalisierung dabei weit über den engeren Bereich der Datenwissenschaften hinaus, markiert diesen allerdings in verschiedener Hinsicht als einen »strategischen Forschungsbereich« (ETH-Rat 2016a) oder eine grundlegende »Basistechnologie« (SBFI 2017). Die im Rahmen des Digitalisierungsdiskurses entworfenen Zukunftsszenarien können als Fallbeispiel eines kollektiven Gesellschaftsentwurfes, der sich auf Aussagen politischer, ökonomischer und wissenschaftlicher Akteur*innen stützt, analysiert werden.

Anders als in den Stelleninseraten, die gleichsam von einer strukturellen Offenheit und Mehrdeutigkeit des Begriffs der Datenwissenschaften geprägt sind, stellt sich im Feld der Bildungs- und Forschungspolitik zumindest die Frage nach der Bedeutung der Datenwissenschaften nicht, wie im einleitenden Zitat des ETH-Rats exemplarisch deutlich wird. Interessant ist hingegen, wie die Datenwissenschaften gerahmt werden: Forschung und Lehre in den Datenwissenschaften sind geradezu ein – je nach Standpunkt – wissenschaftlicher, politischer oder gar moralischer Imperativ, dessen gesellschaftlicher Nutzen ausser Frage zu stehen scheint, tragen sie doch dazu bei, aktuelle »globale Herausforderungen der Gesellschaft [...] zu bewältigen« (ETH-Rat 2016b: 2).

Das Zitat ist instruktiv, denn es verweist auf zwei unterschiedliche Dimensionen, innerhalb derer den Datenwissenschaften im Feld der Hochschul- und Forschungspolitik Bedeutung zugeschrieben wird: Zum einen manifestiert sich ein politischer Handlungsimperativ, etwas tun zu müssen, um durch Investitionen in Bildung und Forschung die Geschehnisse in anderen sozialen Feldern (»Gesundheit, Energie- und Nahrungsmittelversorgung sowie de[r] Umgang mit Ressourcen«) zu beeinflussen. Damit werden unterschiedliche Felder in ein soziales Verhältnis zueinander gestellt. Zum anderen operiert der politische Diskurs in einer temporalen Dimension, indem forschungs- und bildungspolitischen Massnahmen Potenziale hinsichtlich der Strukturierung der Zukunft zugeschrieben werden. Zukunftsentwürfe tragen bei der Genese neuer Wissensgebiete dazu bei, die Erwartungen unterschiedlicher Akteur*innen zu koordinieren und zu strukturieren. Der hochschul- und forschungspolitische Diskurs der Schweiz ist insofern geeignet, um kollektive Visionen und Narrative der Datenwissenschaften hinsichtlich ihrer multidimensionalen Verortungen und Wechselwirkungen zu analysieren.

Ich untersuche in diesem Kapitel durch qualitative Inhaltsanalysen von Strategien und Dokumenten bildungs- und forschungspolitischer Akteur*innen die Frage, welche Rolle gesellschaftliche Zukunftsentwürfe bei der Genese und Sinnkonstruktion der Datenwissenschaften als Wissensfeld spielen. Folgende Fragen leiten die Analyse an:

- Wie operiert der politische Diskurs zur Digitalisierung?
- Wie werden Bildung und Forschung in diesem Diskurs gerahmt?

- Welche Zukünfte über Daten bzw. Datenwissenschaften entwerfen Akteur*innen in der Bildungs- und Forschungspolitik?

Die inhaltsanalytische Auswertung identifiziert Praktiken der Begriffsarbeit als das fundierende Moment: Akteur*innen der Bildungs- und Forschungspolitik verwenden offene, mehrdeutige Begriffe zur Charakterisierung des Wissensgebiets und rahmen die Datenwissenschaften innerhalb existierender, soziotechnischer Diskurse und Zukunftsvisionen. Durch die Verknüpfung mit ambigen Begriffen wie Digitalisierung, Kompetenzen oder Innovation stellen sie somit eine Vielstimmigkeit des Gegenstandes her. Zudem schaffen sie durch positive Rahmungen neue Möglichkeitshorizonte und stimulieren Hoffnungen auf gesellschaftliche Transformation durch die Förderung bestimmter »Zukunftstechnologien« wie »Data Science«, künstliche Intelligenz und Robotik. Die Einbindung zentraler Elemente der neuen »Datenpolitik« in das Narrativ des Standortwettbewerbs und die gelegentliche Gleichsetzung bildungs- und forschungspolitischer mit wirtschaftspolitischen Interessen verweist auf die historische Kontinuität technisch-ökonomistischer Rationalitäten der schweizerischen Bildungs- und Forschungspolitik seit der zweiten Hälfte des 20. Jahrhunderts. Insofern markiert der untersuchte Diskurs gerade in seiner Verwendung offener, mehrdeutiger Begrifflichkeiten sowie einer ubiquitären Transformationsrhetorik eine Orientierung an Stabilität und Kontinuität, die sich nur vermeintlich von früheren Narrativen absetzt.

Das Kapitel ist wie folgt strukturiert: Zunächst skizziere ich zentrale Elemente und Operationsmodi des politischen Diskurses zur Digitalisierung im Hinblick auf die Koordinierung verschiedener Akteur*innen (Kap. 6.2). Danach untersuche ich Zukunftsentwürfe von Bildung und Forschung in der bundesrätlichen Strategie »Digitale Schweiz« (Kap. 6.3). Im Anschluss widme ich mich der Frage, welche Rolle Daten bzw. Datenwissenschaften in den soziotechnischen Zukunftsvisionen einnehmen, wie sie adressiert und gerahmt werden (Kap. 6.4). Abschliessend diskutiere ich die Ergebnisse des Kapitels im Hinblick auf die übergeordneten Fragestellungen der Arbeit (Kap. 6.5).

6.2 Geordnete Arbitrarität im politischen Diskurs zur Digitalisierung

»Das D-Wort ist eine zeitgemässe Art zu sagen: Wandel, Veränderung, Restrukturierung. Das D-Wort hat noch eine weitere Bedeutungsebene, auf der sich Elemente der Techniqueuphorie vergangener Zeiten erhalten haben. Es klingt beim D-Wort die Vorstellung mit, dass Technik eine Urkraft sei, der sich der Mensch nicht in den Weg stellen dürfe« (Betschon 2019: 10).

6.2.1 Die Multidimensionalität und Arbitrarität von Digitalisierung

Keines der untersuchten Dokumente definiert oder erläutert den Begriff der Digitalisierung, obwohl dieser regelmässig Verwendung findet. So bleibt der Begriff im politischen Diskurs unterbestimmt, was ihn offen und anschlussfähig für multiple Perspektiven macht. Dennoch lassen sich verallgemeinert folgende drei divergierende Konzeptionen von Digitalisierung identifizieren (ähnlich Bedenlier & Deimann 2020: 50f.):

- Digitalisierung als die Überführung analoger in digitale Zeichen,
- Digitalisierung als Technologie(feld),
- Digitalisierung als gesellschaftliche Transformation.¹

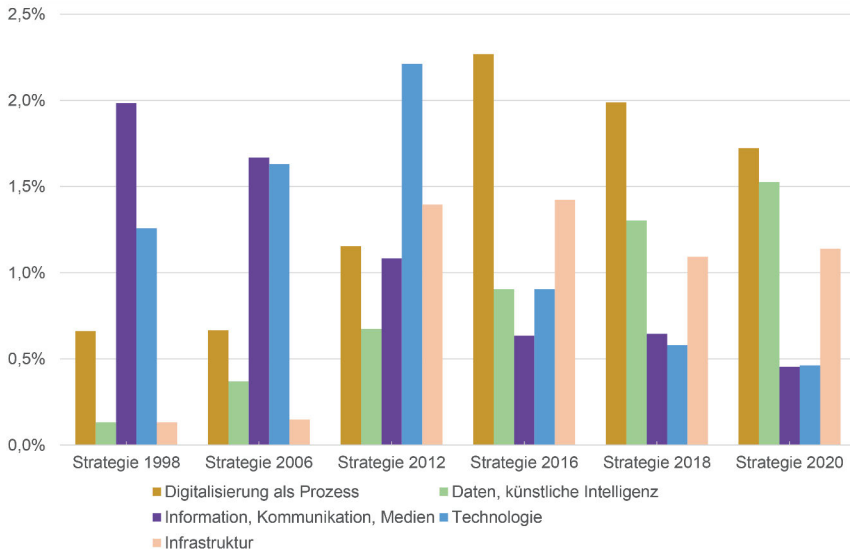
Zunächst bezeichnet Digitalisierung im wörtlichen Sinn die Überführung analoger in digitale, d. h. diskrete, maschinenlesbare Zeichen – ein Prozess, der in den militärischen, akademischen und industriellen Grossrechenmaschinen der 1940er-Jahre einsetzt und bis heute andauert (Gugerli 2018). Die Strategien zur Informationsgesellschaft verwenden das Attribut »digital« lediglich dann, wenn explizit von »digitalen Inhalten«, Zertifikaten oder Signaturen die Rede ist. Ansonsten adressieren sie technologische Aspekte der Informationsgesellschaft unter dem Akronym »IKT«, d. h. Informations- und Kommunikationstechnologien. Auch die Dokumente zur Strategie »Digitale Schweiz« (SDS) schliessen an das ursprüngliche Verständnis an, verknüpfen und erweitern es allerdings mit dem nun ubiquitären Begriff der »Daten«.

In den untersuchten Dokumenten sind die letzteren zwei Konzeptionen vorherrschend. Unter »Digitalisierungstechnologien« werden »neue Technologien aus der Informations- und Kommunikationstechnik (IKT) sowie leistungsfähigere Computer und Netzinfrastrukturen verstanden, welche die technische Grundlage der Digitalisierung darstellen« (SBFI 2017: 3). Gleichzeitig benutzen die Texte den Begriff zunehmend als Synonym für Technologie an sich. Die Bedeutung von Digitalisierung als Technologie(feld) erschliesst sich allerdings erst in Kombination mit dem dritten Verständnis, Digitalisierung als gesellschaftlicher Transformation. Der Begriff indiziert bereits eine Prozesslogik (Grunwald 2019) und zielt auf die unterschiedlichen – politischen, ökonomischen, technologischen, organisationalen und weiteren – Dimensionen der »digitalen Transformation«. Insbesondere der SDS liegt in verstärkter Masse eine solche Prozesssemantik zugrunde: Digitalisierung wird als »fortschreitend« bzw. »zunehmend« charakterisiert (SDS 2016: 3), eng verknüpft mit Begriffen wie »Entwicklung«, »Veränderung«, »Transformation« bzw. »Transformationsprozess« oder »Wandel«.

Diese Bedeutungsverschiebungen im politischen Diskurs lassen sich im Subkorpus der Strategiedokumente empirisch nachzeichnen (vgl. Abbildung 6): Der Themenbereich »Digitalisierung als Prozess« nimmt nach 2006 markant zu. Taucht der Begriff »Strukturwandel« in der SIG 2006 nur einmal beiläufig auf, wird er in der SDS 2016 zu einem zentralen Grundsatz (»den Strukturwandel aktiv angehen«). »Strukturwandel« und »digitale Transformation« sind fortan nicht mehr nur Opportunitäten, die es wahrzunehmen gilt, sondern mutieren zum Hauptorientierungspunkt politischer Aufmerksamkeit.

¹ Vgl. auch die Definition von Hess, Thomas (2019): Digitalisierung. Online: <https://www.enzyklopaedie-der-wirtschaftsinformatik.de/lexikon/technologien-methoden/Informatik--Grundlagen/digitalisierung/index.html> (Zugriff: 03.02.2022).

Abbildung 6: Vergleich der relativen Häufigkeiten von technologieindizierenden Themenbereichen in den Strategiedokumenten 1998–2020



Mit der prononcierten Prozesssemantik geht ein relativer Bedeutungsverlust des Themenbereichs »Technologie« einher: Sind die »IKT« in den Strategien zur Informationsgesellschaft der zentrale Begriff, so verliert das Akronym an Bedeutung (SDS 2016) und taucht später gar nicht mehr auf (SDS 2018). Der damit verwandte Themenbereich »Information, Kommunikation, Medien« erlebt gar einen noch stärkeren Rückgang. Gleichzeitig lässt sich eine Verschiebung von den »IKT« der Informationsgesellschaft hin zu »Daten und künstlicher Intelligenz« sowie »Infrastruktur« der digitalen Gesellschaft identifizieren – zwei Politikfelder, die noch in den ersten beiden Strategiedokumenten irrelevant waren, werden in den letzten Jahren zunehmend wichtiger und nehmen eine bedeutende Rolle in den kollektiven Zukunftsentwürfen des bildungs- und forschungspolitischen Diskurses ein.²

Die Prozesssemantik wird allerdings nicht durchgängig verwendet: Trotz der gehäuftten Verwendung von Begriffen wie »Entwicklung« oder »Transformation« ist neben der »Digitale[n] Schweiz« auch von einer »digitalen Gesellschaft« oder »digitalen Welt« die Rede (SDS 2016: 17). Obwohl Digitalisierung als dynamisch und »fortschreitend« charakterisiert wird, wird zugleich auch die Gegenwart bzw. Existenz einer bereits vorhandenen Digitalität anerkannt. Die Dokumente entwerfen demnach die Vision einer Schweiz, die einerseits bereits digital strukturiert ist, während sie andererseits einem un abgeschlossenen, dynamischen Entwicklungsprozess unterliegt, also immer erst »im Digital werden« begriffen ist. Das komplexe, dialektische Verhältnis von Gegenwart und Zukunft (Jasanoff 2015: 24) verweist auf die Emergenz von Wissen und Technologien innerhalb eingeschlagener Pfade, die wiederum die weitere

2 Trotz dieser Verschiebungen bleiben die technologieindizierenden Themenbereiche in den Strategiedokumenten mit einem kumulierten relativen Anteil von 3,5 bis 3,8 % konstant, mit Ausnahme der SIG 2012.

Entwicklung prägen. Umgekehrt verändern auch zukünftige Implementierungen und Verwendungsweisen die Bewertung historischer Ereignisse und Verläufe.

»Die Digitalpolitik des Bundes stellt den Menschen in den Mittelpunkt einer prosperierenden demokratischen Informations- und Wissensgesellschaft Schweiz« (SDS 2016: 3).

Die Verwendung des Digitalisierungsbegriffs erfolgt auch deshalb uneindeutig und arbiträr, weil verschiedene Gesellschaftsbegriffe (»Informationsgesellschaft«, »Wissensgesellschaft« und »digitale Gesellschaft«) nebeneinander koexistieren, wobei sich weder in früheren noch aktuellen Berichten eine Abgrenzung findet. Die Synchronizität und Parallelität der Begriffe ergibt sich vor allem aus der Fortführung der bundesrätlichen Strategien: Diese weisen – wie in dem obigen Zitat erkennbar – in ihren Einleitungen explizit einen Bezug zu vorangehenden Strategien auf und stellen insofern Kontinuität zwischen den verschiedenen Dokumenten her (Abun-Nasr 2009; Ciarla 2018).

Die bundesrätlichen Strategiedokumente, die sich an eine breite Öffentlichkeit richten, imaginieren die »digitale Transformation« entsprechend als eine kontinuierliche, lineare Entwicklung, die dadurch gewissermassen vorhersehbar und planbar wird – und somit zur Reduktion der mit der Digitalisierung assoziierten Unsicherheiten beiträgt. Sie signalisieren so die Kontinuität gerade durch den eingeschlagenen Transformationsprozess, d. h. »Stabilität durch Veränderung« (Esposito 2016: 428), und übertragen diese mit der SDS in ein politisches Format. Die so organisierte bzw. geordnete Arbitrarität des Digitalisierungsbegriffs erlaubt nicht nur die Anschlussfähigkeit der Strategien für andere Akteur*innen, sondern auch die Bearbeitung von Unsicherheit: Die Kontingenzen und Unsicherheiten der Zukunft, die sich exemplarisch im multidimensionalen Prozessbegriff der Digitalisierung verdichten, werden durch politische Vorstellungen von Linearität und Kontinuität prospektiv extrapoliert.

6.2.2 Die Ziele der Strategie »Digitale Schweiz«: Wohlstand und Wettbewerbsfähigkeit durch Innovation erhalten

Die Kontinuität in den politischen Vorstellungen zur Zukunft der Schweiz äussert sich nicht nur in der Verwendung diskursiver Rahmungen; auch inhaltlich weisen die Dokumente bedeutende Überschneidungen auf. Sie zeigen sich prominent in den grundlegenden Zielen: Die »Nutzung« von »IKT« bzw. »digitalen Technologien« zur Erhaltung des Wohlstandes sowie der Wettbewerbsfähigkeit der Schweiz markiert über den untersuchten Zeitraum hinweg das primäre Ziel der Strategien. Die »Digitale Schweiz« postuliert dies wie folgt:

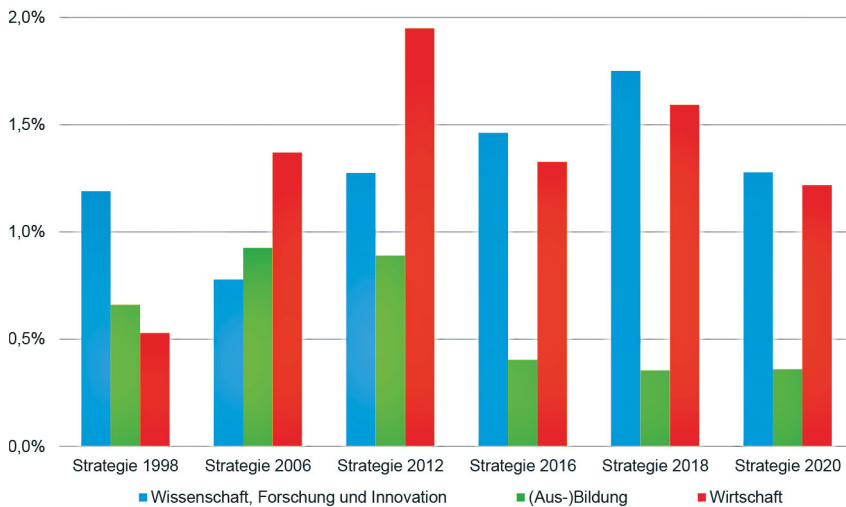
»Im Zentrum der vorliegenden Strategie steht die konsequente Nutzung der Chancen der Digitalisierung, damit sich die Schweiz als attraktiver Lebensraum und innovativer, zukunftsorientierter Wirtschafts- und Forschungsstandort positionieren kann« (SDS 2016: 3).

Das erste Kernziel der SDS 2016 präzisiert den positiven volkswirtschaftlichen Effekt, indem Digitalisierung unmittelbar mit Innovation, Wertschöpfung, wirtschaftlichem Wachstum und Wohlstandssicherung verknüpft wird (SDS 2016: 7). Die weiteren Kernziele betonen die Relevanz digitaler Technologien für die politische Meinungsbildung

und Partizipation, die Transparenz und Sicherheit von digitalen Technologien sowie die nachhaltige Entwicklung (SDS 2016: 6f.). Dadurch stellen sie eine inhaltliche Kohärenz der politischen Ziele über einen längeren Zeitraum hinweg fest und markieren Beständigkeit in einer Phase technologischer, ökonomischer und sozialer Veränderung.

Die Themenbereiche »Wohlstand, Lebensqualität« sowie »(Standort-)Wettbewerb« sind im untersuchten Zeitraum – auf sehr tiefem Niveau – äusserst konstant. Sie bilden neben den technologischen Aspekten das inhaltliche Fundament, indem sie in allen Strategiedokumenten wiederkehrend Erwähnung finden. Umgekehrt wird die Verknüpfung von technologischen Entwicklungen mit Bildungs- und Forschungspolitik die wichtigste Rahmung, um die Prosperität einer wissensbasierten Gesellschaft bzw. Ökonomie zu erhalten: Bildung und Forschung fällt die Aufgabe zu, die Produktion, Distribution und Vermittlung neuer Wissensbestände und technologischer Innovationen zu gewährleisten (Jessop et al. 2008). Politische Massnahmen zur Begünstigung und Übertragung wissenschaftlicher und technologischer Innovationen in die ökonomische Sphäre sind zentrale Elemente der Strategie, um die Ziele zu erreichen.

Abbildung 7: Relative Häufigkeiten ausgewählter Themenbereiche in den Strategiedokumenten 1998–2020



In den Strategiedokumenten lässt sich die zunehmende Bedeutung dieses Narrativs an der Prävalenz der Themenbereiche »Wirtschaft« sowie »Wissenschaft, Forschung, Innovation« ablesen (vgl. Abbildung 7). Obwohl die untersuchten Dokumente Bildung allgemein und die Aus- und Weiterbildung von Fachkräften im Speziellen als einen zentralen Schwerpunkt ausweisen, verliert der Themenbereich »Bildung« nach 2012 an Relevanz. Wie ich noch zeigen werde, hat sich der Kompetenzbegriff zur Bearbeitung bildungsbezogener kollektiver Zukunftsprobleme der Digitalisierung etabliert (vgl. Kap. 6.3.1). »Wissenschaft, Forschung und Innovation« hingegen beschränken sich nicht auf das gleichnamige Themenfeld, sondern rahmen und durchziehen ver-

schiedene wichtige Handlungsfelder (wie »Wirtschaft«, »Infrastruktur«, »Daten und künstliche Intelligenz« etc.), d. h., ihnen wird im politischen Diskurs eine transversale Wirkung zugeschrieben. Zukünftig hohe Lebensqualität und wirtschaftliches Wachstum durch forschungsbasierte Innovation zu erreichen, wurde in den letzten Jahrzehnten zu einem zentralen Bezugspunkt forschungs- und wissenschaftspolitischer Bemühungen (Blümel 2016; Rammert et al. 2016). Dadurch markieren die bundesrätlichen Strategiepapiere eine feldübergreifende Anschlussfähigkeit der diskutierten Ziele anstelle von spezifischen politischen Lösungen und ordnen diese in die angestrebte Kontinuität der »digitalen Transformation« ein.

6.2.3 »Chancen und Herausforderungen«: Positive und negative Horizonte von Digitalisierung

»Für die einen ist die digitale Revolution der perfekte Sturm, der sich zusammenbraut, für die anderen die Chance für den nächsten Entwicklungsschritt der Gesellschaft« (Expertengruppe 2018: 25).

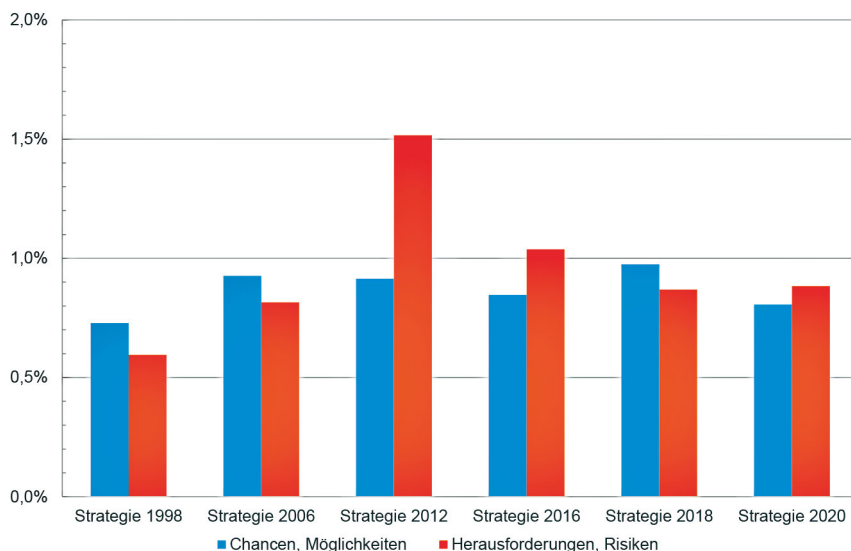
Die Rahmung der Digitalisierung als gesellschaftlicher Transformationsprozess impliziert sozioökonomischen Wandel und Veränderung, die als kontingente Ereignisse in der Zukunft inhärent mit Unsicherheiten behaftet sind. Sie müssen deshalb im politischen Diskurs plausibilisiert und legitimiert werden (Jasanoff 2015; Maasen 2019). Wie Sabine Maasen (2019: 244) argumentiert, müssen neue Technologien »ihre Transzendenz plausibel [...] kommunizieren«, und zwar sowohl durch positive wie negative Bezugnahmen und Horizonte. Eine wichtige Funktion erfüllen dabei Einschätzungen und Bewertungen über die Zukunft von »digitalen Technologien«: Positive und negative Evaluationen von Technologien koexistieren, was es erlaubt, die Unsicherheit von Zukunftsvorstellungen zu adressieren und zu bearbeiten (Esposito 2016). So trägt die Artikulation negativer Möglichkeitshorizonte wie zu erwartende Gefahren oder Risiken, die es zu vermeiden oder minimieren gilt, dazu bei, die weitere Entwicklung durch Massnahmen und Investitionen der Forschungsförderung auf bestimmte Aspekte zu lenken (Beckert 2016: 175).

Die Strategiepapiere rahmen die »digitale Transformation« grundlegend als eine »Chance« zur Wohlstandswahrung bzw. -vermehrung. Indem die untersuchten Dokumente zunächst vor allem die positiven Aspekte von Digitalisierung betonen (SDS 2016: 3), signalisieren sie Kontinuität oder gar eine Optimierung der aktuellen, sozioökonomischen Situation. Die gehäufte Verwendung von »Chancen« und »Potenzialen«, die sich durch die Digitalisierung ergeben, markiert eine »Rhetorik der Potenzialität« (Dickel & Schrape 2015: 442; Hänzi 2015), die weitgehend auf fix konturierte Zukunftsvorstellungen verzichtet; sie operiert vielmehr durch das Eröffnen von Möglichkeitsräumen, in denen »digitale Technologien« über den Horizont des bestehenden Wissens und der gegebenen soziotechnischen Verhältnisse hinaus ihre »Potenziale« freisetzen können.

Da es sich bei Transformationen jedoch um kontingente, entscheidungsoffene Prozesse handelt, ist auch eine Wohlstandsminderung möglich. Eine solche wird indirekt angesprochen, als impliziter Negativhorizont, falls die »Chancen [...] der fortschreitenden Digitalisierung« nicht gepackt würden. Die Strategien betonen demnach nicht nur die »Chancen« und »Potenziale«, sondern legen ebenso einen Fokus auf mög-

liche »Risiken« und »Gefahren« der Digitalisierung, insbesondere im Zusammenhang mit Sicherheit im »Cyberspace«. Synchron lässt sich ein vermehrter Bezug auf Fragen von digitaler Ungleichheit, die es zu verhindern gilt, sowie Datenschutz beobachten.

Abbildung 8: Relative Häufigkeiten der Themenbereiche »Herausforderungen, Risiken« und »Chancen, Möglichkeiten« in den Strategiedokumenten 1998–2020



Insgesamt tauchen im untersuchten politischen Diskurs positive und negative Bewertungen von Digitalisierung bemerkenswert häufig zusammen auf: Auf die Erörterung von Chancen, Möglichkeiten und Potenzialen folgen oft Beschreibungen von möglichen Herausforderungen, Schwierigkeiten und Gefahren.³ Aggregiert betrachtet entwickeln sich in den Strategiedokumenten die Themenbereiche »Chancen, Möglichkeiten« und »Herausforderungen, Risiken« parallel zueinander (mit Ausnahme der SIG 2012), d. h., es finden sich jeweils fast gleich viele positiv und negativ konnotierte Terme (vgl. Abbildung 8).⁴ Die »Rhetorik der Potenzialität« (Dickel & Schrape 2015: 442; Hänzli 2015) verknüpft sich demnach mit dem Risikodiskurs der digitalen Transformation von Gesellschaft, Wirtschaft und Wissenschaft.

Die SDS identifiziert nicht bloss positive und negative Horizonte der nahenden digitalen Zukunft, sondern formuliert auch eine Vision zur Lösung der drohenden Herausforderungen: Zur Bewältigung des sozioökonomischen »Strukturwandels« werden durch die Digitalisierung induzierte Eigenschaften wie »Transversalität«, »Interdisziplinarität« oder »Vernetzung« in Einklang gebracht mit jenen »Trümpfen«, die als für die Schweiz charakteristisch angenommen werden, wie »Multikulturalität, Dialog- und Konsensbereitschaft sowie von Pragmatismus geprägte direktdemokratische Prozesse« (SDS 2016: 5). Die Strategien aktualisieren Grundlagen des schwei-

3 Die diskursive Logik von »Chancen« und »Risiken« digitaler Technologien ist ein wiederkehrendes Element in den untersuchten Dokumenten (Jarchow & Estermann 2015; SBFI 2017; SECO 2017a).

4 Für das gesamte Korpus überwiegt allerdings der Kode »Herausforderungen« um den Faktor 1.3.

zerischen Staatsverständnisses vor dem Hintergrund der Rhetorik einer Netzwerkgesellschaft. Dies suggeriert zwar eine gewisse Kontinuität zwischen gegenwärtigen und zukünftigen Werteeigenschaften, die für die digitale Transformation notwendig sind. Wie diese allerdings in Einklang gebracht werden (können), wird nicht expliziert. Zudem lässt eine technikdeterministische Lesart von Digitalisierung, die in den Dokumenten ebenfalls präsent ist, vermuten, dass der digitale Wandel gerade keine Zeit für langwierige demokratische Aushandlungsprozesse bietet. Durch die Verknüpfung signalisieren die politischen Strategien hingegen eine Ausgewogenheit zwischen positiven und negativen Rahmungen, was wiederum ein Inklusionsangebot für heterogene Anspruchsgruppen darstellt.

6.2.4 Digitalisierung als kollaborativer Prozess: Die Einbindung heterogener Akteur*innen in das Netzwerk »Digitale Schweiz«

Die älteren Strategien zur »Informationsgesellschaft Schweiz« beschränkten sich auf die Bundesebene und verzichteten auf einen »umfassende[n] gesamtgesellschaftliche[n] Regelungsanspruch« (vgl. Tabelle 12 im Anhang). In der SDS werden nun normative Anforderungen an die verschiedenen »Stakeholder« explizit. Die Strategien adressieren fortan auch bundesexterne Akteur*innen wie Kantone, Unternehmen oder die »Zivilgesellschaft«. Seit 2017 umfassen die Aktionspläne zur Digitalen Schweiz auch Massnahmen von nichtstaatlichen Akteur*innen (insbesondere in den Aktionsfeldern »Infrastruktur« sowie »Natürliche Ressourcen und Energie«, vgl. GDS 2017, 2018), womit das Netzwerk über die verantwortlichen Bundesstellen hinaus erweitert wird.⁵ Gleichzeitig fliessen die Rückmeldungen und Interessen von bundesexternen Akteur*innen aus Wirtschaft, Politik und Wissenschaft in die Aktualisierung der Strategie und Aktionspläne mit ein (Staatslabor 2018). Unternehmerische und andere private Initiativen können somit Teil einer nationalstaatlichen Strategie zur Digitalisierung werden. Die Strategie selbst wird als ein dynamischer, sich ständig verändernder Prozess inszeniert, der durch Aktionspläne und Umsetzungsberichte, Umfragen und Workshops, statistische Indikatoren und Messgrössen⁶ sowie weitere Feedbackschlaufen (Ciarla 2018: 51) kontinuierlich aktualisiert wird.

Mit der SDS verfolgt der Bund einen »Multistakeholder-Ansatz« (ISA IG 2016: 3) und lanciert einen »nationale[n] Dialog zur Digitalisierung mit allen Akteur*innen«, den »Dialog Digitale Schweiz« (SDS 2016: 5). Mit den jährlichen »Digitaltage[n] Schweiz«⁷ und den zweijährlich stattfindenden Konferenzen werden Begriffe wie Dialog, Partizipation oder Teilhabe ubiquitär. Im Gegensatz zu jenen Dokumenten im Korpus, die

5 Auch die Lancierung der Strategie »Digitale Schweiz« im November 2017 mit einem Kongress mit 700 Teilnehmenden aus Verwaltung, Politik, Wirtschaft, Wissenschaft, Medien und Zivilgesellschaft expliziert das neue Selbstverständnis und den Regelungsanspruch des Bundes hinsichtlich der »Digitalisierung«.

6 Vgl. ICTSwitzerland (2020): Facts and Figures zur digitalen Wirtschaft aus einer Hand. Online: <https://digitalswitzerland.com/facts-and-figures/> (Zugriff: 03.02.2022).

7 Der privatrechtlich organisierte Verein »digitalswitzerland«, dessen Mitglieder sich aus Grossunternehmen, Kantonen, Universitäten und Hochschulen zusammensetzen, organisiert jährlich die »Digitaltage« im Auftrag und in Zusammenarbeit mit dem Bund. Vgl. <https://www.digitaltage.swiss/> (Zugriff: 03.02.2022).

Digitalisierung als abstrakte Transformation beschreiben, schafft der »Digital Dialog« sinnlich und körperlich erfahrbare Momente: Dezentral werden auf öffentlichen Plätzen technische Innovationen vorgestellt und Diskussionen über das zukünftige Verhältnis von digitalen Technologien und Gesellschaft geführt. Der »Digital Dialog« erprobt Instrumente der Partizipation und Mitbestimmung im politischen Diskurs zur Digitalisierung, wobei sich Parallelen zu Prozessen der wissenschaftlichen Forschung und Technologiegestaltung zeigen, in denen Partizipation, Transparenz und Nachvollziehbarkeit zu Funktionsprinzipien avanciert sind (Maasen & Dickel 2016).

Die SDS kann in der Dringlichkeit ihrer politisch-medialen Inszenierung als Versuch gelesen werden, die Bevölkerung der Schweiz auf die Implikationen der soziotechnischen Transformation vorzubereiten.⁸ Ferner zielen die Partizipationsprozesse auch darauf ab, Zustimmung bzw. Konsens für politische Massnahmen zu erzeugen, die die Einführung bzw. Nutzung bestimmter Technologien (wie aktuell die Debatten über den neuen Mobilfunkstandard »5G« zeigen) oder sozioökonomische Transformationsprozesse (»Strukturwandel«) an sich begleiten, die demokratisch aber kaum legitimierbar sind (Betschon 2019). Die Wirkmächtigkeit soziotechnischer Imaginationen über den politischen Diskurs hinaus liegt in diesem Sinne darin, dass sie Inklusionsangebote für sehr unterschiedliche Akteur*innen, Ziele und Interessen formulieren, um diese in ein heterogenes Netzwerk einzubinden.

Die Analyse macht deutlich, dass die Arbitrarität des zentralen Begriffs der Digitalisierung im politischen Diskurs Anknüpfungspunkte für unterschiedliche Akteur*innen schafft, sich auf die Ziele und Inhalte der Strategien zu beziehen, was diese stabilisiert und ihre Wirkmächtigkeit feldübergreifend erhöht. Offenheit und Anschlussfähigkeit erlauben es Akteur*innen in anderen Feldern, ihre jeweiligen Digitalisierungsstrategien auf die im Kern ökonomischen Ziele zu koordinieren. Die diskursive Rahmung als Chance und Herausforderung vereinigt divergierende Bewertungen digitaler Technologien, wodurch Inklusionsangebote für ein breites Publikum formuliert werden. Die Inszenierung eines kollaborativen Prozesses bindet zudem Akteur*innen über die unmittelbar involvierten politischen, technisch-wissenschaftlichen und ökonomischen Anspruchsgruppen hinaus ein. Das so erweiterte Netzwerk stabilisiert die unterschiedlichen Aktivitäten und trägt durch wechselseitige Bezugnahmen im Ergebnis dazu bei, zukunftsbezogene Unsicherheiten zu reduzieren und die weitere Entwicklung der »digitalen Schweiz« zu strukturieren.

8 Sabine Maasen spricht in Bezug auf soziale Robotik von einem »Vorbereitungs- und Verschaltungsdiskurs«, in dem die Öffentlichkeit auf die künftigen Mensch-Maschinen-Interaktionen und deren Implikationen vorbereitet werden soll (Vortrag Sabine Maasen, Universität Bern, Collegium Generale, 05.12.2018).

6.3 Zukunftsentwürfe von Bildung und Forschung in der Strategie »Digitale Schweiz«

6.3.1 Bildung als die Vielfalt »digitaler Kompetenzen«

Der Bildungsbereich gilt als Schlüsselfeld für den gesellschaftliche Bewältigung der »digitalen Transformation«. Dies äussert sich in den bundesrätlichen Strategiedokumenten etwa darin, dass Bildung als einem der wenigen Politikfelder stets eine prominente Stelle eingeräumt wird und verschiedene Massnahmen den Bildungsbereich fokussieren.⁹ Die »Weiterentwicklung der Wissensgesellschaft« bildet ein konstitutives Ziel einer »Digitalen Schweiz«. Bildung gilt insofern als elementar, um die proklamierten »Chancen der Digitalisierung« zu nutzen sowie Wohlstand und Wettbewerbsfähigkeit des Landes zu steigern. Dies wird an zwei bildungspolitischen Problemstellungen deutlich: Im ersten Fall, der Ausbildung von Arbeitskräften, werden die gegenwärtigen Stärken und Leistungen des Bildungssystems in seiner Bedeutung für den »Forschungs- und Wirtschaftsstandort Schweiz« beschworen und zusätzliche Entfaltungsmöglichkeiten diskutiert, um »Potenziale auszuschöpfen« (SDS 2016: 16) bzw. das »digitale Potenzial« zu aktivieren (Economiesuisse & W.I.R.E. 2017: 46). Ein zweites Handlungsfeld betrifft Anstrengungen, um den »Fachkräftemangel« im MINT-Bereich zu adressieren. Die »Förderung des IKT-Nachwuchses« war bereits Gegenstand früherer Strategien, allerdings noch ohne die Verknüpfung mit drohenden Wettbewerbsverlusten. In der Zwischenzeit setzt sich die bildungsökonomische Rahmung eines akuten »Fachkräftemangels« für die Beziehung zwischen Bildungswesen und Arbeitsmarkt durch (Herberg 2018a), sodass die politischen Appelle zu dessen Behebung immer dringender werden (SBFI 2017). Demgegenüber geraten andere Verständnisse von Bildung im digitalen Zeitalter, wie beispielsweise die erweiterte »Handlungsfähigkeit und [...] Kommunikationsmöglichkeiten [...], [die] Knüpfung und Vertiefung von grenzüberschreitenden, multikulturellen Kontakten« oder die Nutzung der »emanzipatorischen Möglichkeiten« von IKT, wie sie noch in früheren Strategiedokumenten formuliert wurden (SIG 1998), zunehmend aus dem Blick.

Die Beispiele zeigen, dass die politischen Diskussionen über die Digitalisierung keine grundlegende Änderung der Bildungspolitik indizieren, dass aber Zukunftsprognosen die eingeschlagenen Pfade neu situieren. Ein zentrales diskursives Mittel dazu ist das Konzept der »Kompetenzen«¹⁰: »Fit sein für die Digitalisierung« bedeutet demnach, über die durch das Bildungssystem vermittelten IKT-Kompetenzen bzw. »digitalen Kompetenzen« zu verfügen, um »risikobewusst und eigenverantwortlich« in den unterschiedlichen gesellschaftlichen Bereichen an der »digitalen Transformation« teilnehmen zu können (SDS 2016: 16). Um das Strategieziel zu erreichen, wird unter anderem eine bessere Koordinierung der Bemühungen von Bund und Kantonen hinsichtlich der »Integration der IKT im Bildungswesen« angestrebt. Es folgt jedoch

9 Die Benennung des Politikfeldes verändert sich über die zwei Jahrzehnte (Ciarla 2018: 86ff.). In der Strategie »Digitale Schweiz« werden Bildung und Forschung unter »Weiterentwicklung der Wissensgesellschaft« zusammengeführt. In der aktuellen Fassung vom September 2018 bilden sie als »Bildung, Forschung und Innovation« das erste von neun Aktionsfeldern.

10 Kompetenzkonstruktionen können als Bündel heterogener Wissensbestände, Fähigkeiten sowie persönlicher Eigenschaften konzeptualisiert werden (Green 2013).

keine breit angelegte Bildungs- bzw. Weiterbildungskampagne zu deren Vermittlung, wie dies noch Ende der 1990er-Jahre mit einer »Bildungsoffensive« in der SIG 1998 der Fall war.¹¹ Die Aneignung der »digitalen Kompetenzen« bleibt letztlich den Individuen in ihren Bildungsprozessen selbst überlassen.

»Die Kompetenzen der Schweizer Bevölkerung sollen weiter gestärkt werden, damit sie die Chancen der Digitalisierung umfassend nutzen kann. Dank dem lebenslangen Lernen sollen die Menschen stets in der Lage sein, kompetent an digitalisierten politischen, sozialen, kulturellen und wirtschaftlichen Prozessen teilzunehmen und die Konsequenzen des eigenen Tuns möglichst richtig einschätzen zu können« (SDS 2018: 4).

Wie in dem Zitat manifestiert sich allgemein ein Ebenenwechsel bezüglich der Adressat*innen der Strategieziele: Formulieren die Dokumente bis anhin Zukunftsentwürfe, die in erster Linie kollektive Akteur*innen (staatliche Einheiten, Unternehmen, Verbände etc.) adressieren, erfolgt an dieser Stelle für den Bildungsbereich quasi eine Individualisierung. Nun sind es die »Menschen«, die ihrerseits in die Pflicht genommen werden, ihre Bildungsanstrengungen zur Erreichung der kollektiven Ziele zu intensivieren. Der Ebenenwechsel wird begleitet durch ein pädagogisch-bildungswissenschaftliches Verständnis von Kompetenzen als individuellen Fähigkeiten (gegenüber einem juristisch-staatswissenschaftlichen Verständnis von Zuständigkeit im Sinne Max Webers, dazu Traue 2010), das sich als zentrales bildungspolitisches Element zur Bearbeitung der kollektiven Zielerreichung etabliert hat: Wo Lücken in der Qualifikation von Individuen oder der Lösung gesellschaftlicher Problemstellungen adressiert werden können, müssen neue Kompetenzen erlernt und aufgebaut werden.

Die zu beobachtende Konjunktur des Kompetenzbegriffs hängt auch mit dem Wandel des Bildungswesens als zentraler gesellschaftlicher Instanz für den Umgang mit Kompetenz(en) zusammen: Gegenüber der kredentialistischen Organisation von Bildung im Fordismus, in der das staatliche bzw. staatlich beaufsichtigte Bildungswesen die »äusseren« Leistungen von Schüler*innen – Qualifikationen – testete und zertifizierte – und so die Zugänge zu den verschiedenen Bildungswegen regulierte – (Collins 2019 [1979]), beruht das postfordistische Bildungsarrangement der letzten drei Jahrzehnte zusätzlich auf der Feststellung und Evaluation einer Reihe von »inneren« Eigenschaften – Kompetenzen – einer Person mittels Gesprächen, Feedbacks und (Selbst-)Beurteilungen (Traue 2010).¹²

Die untersuchten Dokumente äussern sich uneinheitlich dazu, was »digitale Kompetenzen« genau ausmacht und worin diese bestehen.¹³ Die Analyse fördert nicht weniger als zwölf Kompetenzbereiche zutage, die je nach Dokument und AutorInschaft

11 Lediglich für ältere und gering qualifizierte Arbeitnehmende sollen spezifische Massnahmen zur Weiterbildung im IKT-Bereich ergriffen werden (vgl. SECO 2017a: 99).

12 Arbeitsmarktsoziologische Studien zeigen die Bedeutung von sogenannten Soft Skills in der Transformation von Arbeitsmärkten in den vergangenen Jahrzehnten (Salvisberg 2010). Dies gilt vor allem für Berufe mit hohen Qualifikationsanforderungen, bei denen analytische, soziale und kommunikative Kompetenzen hinzukommen (Salvisberg 2008: 2572).

13 Zur Fokussierung der digitalen Hochschullehre auf Kompetenzen vgl. Getto und Kerres (2018).

unterschiedliche Fertigkeiten umfassen und mit weiteren kombiniert werden.¹⁴ Am häufigsten genannt werden allgemeine Konzeptualisierungen (wie »IKT-Kompetenzen« oder »digitale Kompetenzen«), während detailliertere Formalisierungen seltener Verwendung finden. Dennoch zeichnet sich eine Verschiebung von generischen »IKT-Kompetenzen«, die primär als »Anwendungskompetenzen« verstanden werden, hin zu konkreten methodischen (wie Datenanalyse, statistische Modellierung etc.) und technischen Fertigkeiten (wie Programmieren, Netzwerkverwaltung) ab. Zum anderen identifizieren die Dokumente kognitives Wissen aus Mathematik, Statistik sowie Informatik als eigentlichen »Kernbereich«, den es durch die gezielte Förderung von MINT-Fächern auf allen Stufen des Bildungssystems zu priorisieren gilt (Bundesrat 2016; SBFJ 2017; Economiesuisse & W.I.R.E. 2017). Dies ist nicht erstaunlich, handelt es sich doch bei der Fokussierung auf technisch-naturwissenschaftliche und mathematische Wissensbestände um ein strukturelles Moment schweizerischer Bildungspolitik (Gugerli et al. 2005; Honegger et al. 2007).

Interessanter ist, dass die Prozess- und Netzwerksemantik der »digitalen Transformation« den technisch-mathematischen »Kernbereich« erweitert: Zusätzlich werden »computational thinking« (Wing 2006), d. h. abstraktes Prozessdenken, und kritisch-reflexives Denken (»critical thinking«) als grundlegende »transversale« bzw. »Querschnittskompetenzen« anerkannt, die sowohl technische als auch nicht-technische Domänen durchziehen würden.¹⁵ Sie gelten aufgrund der Unsicherheiten kontingenter technologischer Entwicklungen in der Zukunft als »Schlüssel«, um die »digitalen Potenziale« zu aktivieren und den Herausforderungen zu begegnen. In dieser Konzeption lösen sich »digitale Kompetenzen« von ihrer (ursprünglichen) technischen Fixierung und ermöglichen es anderen Wissensfeldern, ihren Beitrag zur Ausbildung von »digital kompetenten« Subjekten zu leisten.

Trotz der breiten Anerkennung als Zukunftsressourcen existiert bis dato kein »schweizweit gültige[r] Bezugsrahmen für digitale Kompetenz, der in den entsprechenden Lehrplänen verankert ist«, wie es die kantonalen Bildungsdirektionen in ihrer Digitalisierungsstrategie (EDK 2018: 2) formuliert haben. Die Akteur*innen auf den unterschiedlichen Ebenen des Bildungsfeldes arbeiten relativ unabhängig voneinander an pädagogischen Konzepten zur Integration »digitaler Kompetenzen« in den Unterricht. Diese »digitale Disharmonie« (Krummenacher 2019) verweist mitunter auf die beschränkten Handlungs- und Gestaltungsspielräume der Bundespolitik in der Bildung, die ausser dem ETH-Bereich und der Berufsbildung kaum inhaltliche Vorgaben machen kann. Auch für den Hochschulbereich hält das Massnahmenprogramm »Digital Skills«, das zur Stärkung »digitaler (Anwendungs-)Kompetenzen« in Forschung, Lehre und Verwaltung beitragen soll, die Themenfelder explizit offen.¹⁶

14 Es sind dies Datenkompetenzen, digitale Kompetenzen bzw. Computational Thinking, Grundkompetenzen, IKT-Kompetenzen, kommunikative Kompetenzen, Kreativität/Neugier, kritische Reflexion bzw. Critical Thinking, Medien-Kompetenzen, methodisch-analytische Kompetenzen, Querschnittskompetenzen, soziale und interaktive Kompetenzen sowie technische Kompetenzen (inkl. Programmieren).

15 Komplementiert werden diese durch sozial-interaktive, kommunikative Kompetenzen sowie persönliche Einstellungen (wie Neugier, Neugier, Kreativität, eine unternehmerische Haltung etc.).

16 Swissuniversities (2018): Themenfelder & thematische Aspekte. Bern: Swissuniversities. Online: <https://www.swissuniversities.ch/themen/digitalisierung/digital-skills> (Zugriff: 03.02.2022).

Die empirische zu beobachtende Operationalisierung und Umsetzung von »digitalen Kompetenzen« trifft sich insofern mit dem sehr weiten Verständnis von Digitalisierung, wie es den politischen Diskurs insgesamt kennzeichnet. Trotz der identifizierten Vielfalt und Varietät, die sich am Beispiel der Diskussion und Einführung »digitaler Kompetenzen« zeigt, hat sich der Kompetenzbegriff zur Beschreibung und Lösung bildungsbezogener kollektiver Zukunftsprobleme etabliert (Bedenlier & Deimann 2020). Die untersuchten Dokumente operieren mit einem subjektzentrierten, stark auf humankapitalorientierte Sichtweisen angelegten Bildungsverständnis, während andere Konzeptionen von Bildung zunehmend aus dem Blick geraten.¹⁷ Im politischen Diskurs zur Digitalisierung verknüpft sich dieses mit der Potenzialitätsrhetorik der digitalen Transformation und rahmt das Bildungssystem als Ort der Ausbildung von Fachkräften sowie als elementare Stütze im internationalen Standortwettbewerb.

Das Konzept der »Kompetenzen« übersetzt die Multidimensionalität und Arbitrarität von Digitalisierung in konkrete, prozessier- und messbare Grössen. Obwohl die Akteur*innen im Feld der Bildungs- und Forschungspolitik vielfach ihre je eigenen Konzeptionen »digitaler Kompetenzen« entwerfen, entfalten die unkoordinierten Aktivitäten performative Effekte: Sie befördern die politische wie öffentliche Wahrnehmung, dass »digitale Kompetenzen« zur Bearbeitung zukunftsbezogener Kontingenzen implementiert werden (müssen), und indem sie dies tun, verlagern sie diese erneut in die Zukunft, was wiederum neue Unsicherheiten schafft, die bearbeitet werden müssen.

6.3.2 Die Logik der Verwertbarkeit von Forschung und das Primat eines ökonomischen Innovationsbegriffs

Neben der Bildung wird auch die Forschung als zentrales Feld identifiziert, um die Ziele der Strategie zu erreichen. Verweise auf die »Bedürfnisse unserer digitalen Gesellschaft und Wirtschaft« sowie den »Spitzenplatz der Schweiz als Innovations- und Forschungsstandort« markieren die grundlegende Ausrichtung einer Bildungs- und Forschungspolitik, die ökonomische Interessen wie den Bedarf an Fachkräften, Wissens- und Technologietransfer sowie das Bestehen im internationalen (Standort-)Wettbewerb berücksichtigen muss. Die Schwerpunktsetzung äussert sich, wie bereits gezeigt, in der zunehmenden Relevanz der Themenbereiche »Wissenschaft, Forschung und Innovation« sowie »Wirtschaft« in den Strategiedokumenten (vgl. Abbildung 7). Massnahmenseitig manifestiert sie sich in der Förderung von Technologien, die als elementar für die »digitale Transformation« betrachtet werden (vgl. auch Kap. 6.4.3): Dies sind zum einen das NFP 75 »Big Data« (2017–2022) und das NFP 77 »Digitale Transformation« (2020–2024) des SNF. Das Aktionsfeld 7 des Massnahmenpakets des Bundes legt einen besonderen Schwerpunkt auf die »Stärkung der Grundlagenforschung im Bereich Digitalisierung« im Rahmen der 5. Serie der Nationalen Forschungsschwerpunkte sowie auf einen »Kompetenzaufbau« in »Informatik/Compu-

17 Ansatzweise zeichnet sich diese Verschiebung auch im Zeitverlauf ab (vgl. Abbildung 6): Die Bedeutung der Themenbereiche »Information, Kommunikation, Medien« sowie »Gesellschaft, Kultur, Vielfalt«, die inhärent mit anderen Bildungsverständnissen zusammenhängen, nimmt kontinuierlich ab. Zählten diese zwei Bereiche in den Strategien zur Informationsgesellschaft zu den wichtigsten überhaupt, sind sie mittlerweile nur noch sekundär.

ting Science« im ETH-Bereich (SBFI 2017: 70f.). Weitere Investitionen betreffen die Bereiche künstliche Intelligenz, Robotik sowie Internet of Things (SDS 2016; SBFI 2017). Die Massnahmen verdeutlichen nicht nur die forschungspolitische Relevanz und »Zukunftsorientierung« der Wissensgebiete, sondern signalisieren anderen Akteur*innen überdies, dass sich Investitionen in diese Bereiche lohnen.

Das wichtigste Element dieser forschungspolitischen Ausrichtung ist das multidimensionale Konzept der Innovation (Godin 2018; Rammert et al. 2016), das zu einem zentralen diskursiven Instrument für den gesellschaftlichen Umgang mit der Digitalisierung wird. Innovationspraktiken und das Sprechen über Innovation haben ihre Ursprünge in Wissenschaft und Ökonomie verlassen; sie wurden zu einem allgemeinen gesellschaftlichen Imperativ, dem sich kaum noch ein Feld entziehen kann. In den 1970er-Jahren hat das Konzept Eingang in die Wissenschafts- und Forschungspolitik gefunden und definiert diese mittlerweile massgeblich (Blümel 2018). So geht die Vorstellung von – regionalen, nationalen oder gar transnationalen – Innovationssystemen von institutionellen Kooperationen zwischen Wissenschaft, Industrie und Politik aus, die auf unterschiedlichen Ebenen die Entstehung von Innovationen rahmen und beeinflussen (Godin 2009; Nelson 1993).

Um die Inwertsetzung wissenschaftlicher Forschungsergebnisse zu erreichen, werden in der SDS »die Gründung und der Aufbau wissenschaftsbasierter Unternehmen« sowie der »Wissens- und Technologietransfer zwischen Hochschulen, Wirtschaft und Gesellschaft« unterstützt (GDS 2018: 3).¹⁸ Die Dokumente situieren Innovation demnach an der Schnittstelle von wissenschaftlicher Forschung an Universitäten und Hochschulen einerseits sowie anwendungsorientierter Technologieentwicklung und ökonomischer Verwertung in Unternehmen andererseits. Die grundlegende Argumentation, wie Innovation entsteht, lautet dabei wie folgt: Die Digitalisierung bzw. der technologische Fortschritt produzieren in den Forschungsaktivitäten von Universitäten und Hochschulen neues Wissen, das über Wissens- und Technologietransfer in die Wirtschaft transponiert wird und existierende Technologien transformiert. Die Unternehmen nehmen neue Technologien bzw. neues Wissen auf und implementieren sie in neue Produkte und Dienstleistungen. Dies führt zu neuen Kompetenzprofilen und einer veränderten Nachfrage im Arbeitsmarkt. Folglich muss sich das Bildungssystem anpassen und die erforderlichen Qualifikationen zur Verfügung stellen. Obwohl Bildung und Forschung »eine zentrale Rolle in der Entwicklung, Anwendung und Nutzbarmachung der neuen Technologien« (SDS 2016: 17; SBFI 2017: 3) zuerkannt wird, bleibt »Digitalisierung als Basistechnologie« in dieser Logik ein undurchdringbarer, von Zeit und Raum unabhängiger Bereich, der kontinuierlich technische Innovationen und Wissen produziert.¹⁹

Die Strategien imaginieren die »Entwicklung neuer Technologien« bzw. Innovation anhand eines linearen Entwicklungsmodells (Blümel 2016; Godin 2006). In dieser sequentiellen Abfolge, die auch eine konzeptuelle Trennung und Gegenüberstellung von »Technologie«, »Wissenschaft« und »Gesellschaft« beinhaltet, artikuliert sich ein

18 Der Bedeutungszuwachs spiegelt sich auch in der Reorganisation der zuständigen Verwaltungseinheiten wider, die mit finanziellen Mitteln für die Periode 2016–2020 einherging (vgl. Bundesrat 2016: 3206).

19 Vgl. exemplarisch: Bundesrat (2018): Gesamtschau der Innovationspolitik. Bericht des Bundesrates in Erfüllung des Postulates 13.3073, Derder, 13. März 2013, S. 4ff.

instrumentelles Verständnis von Technologieentwicklung und -nutzung. Die untersuchten Dokumente diagnostizieren und lokalisieren in der Entwicklung und Anwendung von »Digitalisierungstechnologien« insofern die Grundlage für gesellschaftlichen und ökonomischen »Strukturwandel«. Darin begründet sich die Logik eines »konsequentiellen Technikdeterminismus«, wonach eine bestimmte technologische Entwicklung »äußere Wirkungen hat, die üblicherweise als »gesellschaftliche Folgen« bezeichnet werden« (Schulz-Schaeffer 2019: 2).

Dieses lineare Verständnis (Godin 2006) reduziert die Multidimensionalität von Innovation, indem es sie auf eine sequentielle Abfolge in sozialen Feldern fest schreibt.²⁰ Dabei zeigt sich eine »Engführung auf ökonomische Innovation« (Passoth & Rammert 2016: 42), wodurch Eigenschaften wie Reziprozität oder Transversalität aus dem Blick geraten, die in den Ausführungen über Innovation ebenfalls enthalten sind (SBFI 2017): Bildung und Forschung werden sowohl als Triebkräfte als auch als »Getriebene« der Digitalisierung imaginiert, indem sie ständig Neuheit produzieren, andererseits aber immer auch die Neuheit anderer gesellschaftlicher Felder verarbeiten und adaptieren müssen. Die Vorstellung beruht auf einem ökonomistischen Fortschrittsdiskurs, wonach Innovation massgeblich die Performanz und das Wachstum von Unternehmen, Regionen oder ganzen Staaten beeinflusst; ohne Innovation wären nicht nur ökonomische Systeme stationär (Beckert 2016: 169; Blümel 2018), sondern – dem linear-deterministischen Modell zufolge, das den Strategiedokumenten zugrunde liegt, auch die Schweiz an sich. Mit Innovation verbunden sind zudem organisationale und epistemologische Aspekte der Wissensproduktion: Das Beispiel der Datenwissenschaften zeigt, dass sich diese bei feld- bzw. hochschulübergreifenden Kooperationen, Forschungsschwerpunkten, der Anpassung von Lehr- und Lernressourcen und der Einrichtung von Studiengängen überkreuzen (dazu Kapitel 8). Es handelt sich gleichfalls um soziale oder organisationale Innovationen (Baur et al. 2016; Godin 2016), die durch »digitale Technologien« induziert werden, diese aber gleichzeitig auch prägen und gestalten. Allerdings gerät dieser erweiterte, offene Innovationsbegriff in den untersuchten Dokumenten aus dem Blick.

Das vielschichtige Konzept der Innovation übersetzt die Kontingenzen, die mit einem offenen, arbiträren Digitalisierungsbegriff einhergehen, und verortet Neuheit sozial-räumlich an der Schnittstelle von wissenschaftlich-technologischer Forschung und ökonomischer Verwertung. Dies stabilisiert gesellschaftliche Erwartungen darüber, wo und wie Lösungen für die soziotechnischen Herausforderungen der Digitalisierung gefunden werden können, etwa durch zusätzliche Investitionen in Forschungsförderung sowie Wissens- und Technologietransfer. Der zugrundeliegende ökonomistische Innovationsbegriff befördert ein Verständnis gesellschaftlicher Problemlösung, das sich eng an technischen Rationalitäten und Machbarkeiten orientiert. Damit wird ein einseitiges, lineares Verständnis von Innovation determiniert, das soziale und organisationale Aspekte solcher Prozesse mehrheitlich ausblendet.

20 Innovation funktioniert gerade nicht linear, sondern ist vielmehr durch iterative Prozesse des Suchens oder das Rekombinieren von Praktiken gekennzeichnet (Padgett & Powell 2012; Powell & Sandholtz 2012). Mehrdeutigkeit, Verwirrung und Unsicherheit sind zentrale Bestandteile von Innovationsprozessen, die sich nicht rational planen und kalkulieren lassen (Esposito 2016; Lester & Piore 2004; Stark 2009).

6.3.3 Temporalität und Beschleunigung: Die kontinuierliche Anpassung an die Anforderungen der Digitalisierung

Die Dynamik und Prozesshaftigkeit verweisen auf die Temporalität als konstitutive Funktion der Digitalisierung. Das eher statische Verständnis von IKT in der Informationsgesellschaft wird durch die Konzeption von Digitalisierung als kontinuierlicher Transformation abgelöst. Parallel dazu verlagert sich die gesellschaftliche Selbstbeschreibung von einer Gegenwarts- hin zu einer stärkeren Zukunftsorientierung: Zukunftsfähigkeit wird Bestandteil der Selbstpositionierung der Schweiz als wettbewerbsfähiger, innovativer Bildungs- und Forschungsstandort (SDS 2016: 3). Die Zeitlichkeit des Diskurses äussert sich in der gehäuften Verwendung zeitbezogener Substantive und Adjektive (wie »zukunftsorientiert«, »zukünftig«, »heute«, »rechtzeitig«, »zeitgemäss«, »rasant« oder »schnell«) in den untersuchten Dokumenten.

Dabei werden insbesondere wissenschaftliche und technologische Innovationen als Triebkräfte der gesellschaftlichen Entwicklung gerahmt. Der Bildungsbereich hingegen wird als statisch beschrieben, der auf diese Anforderungen reagieren muss. So wird dem Bildungssystem wiederholt die kontinuierliche Anpassung an externe, durch die Digitalisierung induzierte »Herausforderungen« und »Entwicklungen« auferlegt:

»In Zukunft werden alle Bildungsangebote in schnellerer Kadenz auf die sich durch die Digitalisierung ergebenden Herausforderungen überprüft werden müssen [...] Das Bildungssystem als Ganzes muss sich insofern anpassen, dass es flexibler und schneller als bisher auf die Entwicklungen reagieren kann« (SBFI 2017: 39).

Die Dokumente konstruieren eine spezifische Rationalität von Geschwindigkeit und Beschleunigung (Beer 2019: 51), bei der Nicht-Handeln oder Abwarten keine Optionen sind, da dies dazu führen würde, den »Anschluss« zu verpassen. In dieser Logik sind Adaptivität und Flexibilität die Strategien, die die Aktualität von Bildungsinhalten an Schulen und Universitäten gewährleisten sollen, wobei dies nicht nur einzelne Stufen oder bestimmte Vertiefungen oder Studiengänge, sondern »alle Bildungsangebote« bzw. das »Bildungssystem als Ganzes« betrifft. Die Berichte unterwerfen dadurch die Akteur*innen des Bildungswesens einem normativen Handlungsdruck, während gleichzeitig aufgrund der Arbitrarität und Heterogenität der verwendeten Konzepte (wie »digitale Kompetenzen«) sowie der Kontingenz möglicher Transformationspfade unklar bleibt, an welche »Entwicklungen« im Bildungsbereich überhaupt angepasst werden soll. Die politischen Forderungen nach Flexibilität und Adaption zielen in diesem Sinne vielmehr auf die Reformfähigkeit an sich, denen Bildungsorganisationen – genauso wie andere zeitgenössische Organisationen – unterworfen sind (Brunsson 2017).

Gerade die Datenwissenschaften können als ein Beispiel dafür gesehen werden, dass zumindest das Hochschulsystem zur Adaption neuer Wissensbereiche innert relativ kurzer Zeit imstande ist: Als eine Massnahme, »[um] den Bedürfnissen des Arbeitsmarkts und der Studierenden gerecht zu werden« (ETH-Rat 2016b: 2), präsentiert der ETH-Rat die Schaffung zweier Masterstudiengänge in Data Science an der ETH Zürich (ETHZ) sowie der EPFL. Parallel dazu etablierten diverse Universitäten und Fachhochschulen ebenfalls in wenigen Jahren ähnliche Studiengänge oder Vertiefungen. Wie die Analyse zur Implementierung der Datenwissenschaften im Schweizer Hochschulfeld (vgl. Kap. 8) zeigt, handelt es sich dabei oft um mehrjährige Prozesse, die aufgrund

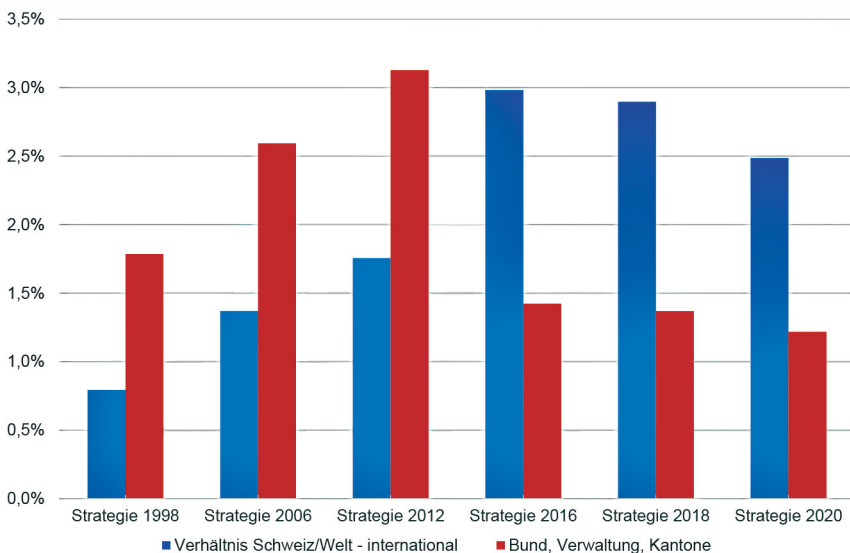
politischer Vorgaben und organisationaler Abläufe nur begrenzt zu beschleunigen sind. Zudem war die Kategorie »Data Science« zum Zeitpunkt (2013) der ersten Überlegungen und Planungen für neue Studiengänge mehrheitlich unbekannt. Epistemologische Transformationen in der Wissenschaft haben Opportunitäten geschaffen, die es zunächst einmal wahrzunehmen und zu stabilisieren galt. Die bereits 2014 angekündigten Investitionen des ETH-Rats für Studiengänge in Datenwissenschaften haben insofern dazu beigetragen, die Erwartungen anderer Akteur*innen zu strukturieren. Das Hochschulfeld hat sich in diesem Sinne als flexibel und anpassungsfähig erwiesen.

6.3.4 Die Internationalisierung des bildungspolitischen Diskurses zur Digitalisierung

»Die Schweiz ist auf Platz 8 weltweit in der Digitalisierung.« (Bundesrätin Doris Leuthard, Konferenz Digitale Schweiz, 20.11.2017)

Die Strategien zur Digitalisierung charakterisiert schliesslich ein verändertes Verhältnis der Selbstreferenzierung des politischen Systems der Schweiz gegenüber politischen Beziehungen zur »Aussenwelt« wie anderen Staaten oder internationalen Organisationen. Adressierten die Strategien zur Informationsgesellschaft 1998 und 2006 noch in erster Linie die politischen Akteur*innen der Schweiz, steigt die Anzahl internationaler Referenzen nach 2012 markant an. Die Analyse der Häufigkeiten der Themenbereiche »Verhältnis Schweiz-Welt« und »Bund, Kantone, Verwaltung« macht dies deutlich: Während zunächst die Prävalenz beider Themenbereiche zunimmt, sinkt die Häufigkeit der innerhelvetischen Bezüge (»Bund, Kantone, Verwaltung«) stark ab, während diejenigen zu externen Akteur*innen wie anderen Staaten, der Europäischen Union sowie internationalen Organisationen signifikant ansteigt (vgl. Abbildung 9).

Abbildung 9: Relative Häufigkeiten der Themenbereiche »Verhältnis Schweiz-Welt« sowie »Bund, Kantone, Verwaltung« in den Strategiedokumenten 1998–2020



Dies gilt ausgeprägt für Bildung und Forschung: Diagnosen über den gegenwärtigen Zustand des Schweizer Bildungs- und Forschungssystems im internationalen Vergleich werden mit einer Analyse der Chancen und Herausforderungen der Digitalisierung kombiniert. Die Berichte untersuchen den Ist-Zustand und stellen ihn mit anderen Nationalstaaten oder der Europäischen Union in Beziehung. Die Vergleiche erfolgen anhand bestimmter Metriken (etwa schulischer Leistungstests wie PISA), Rankings und bibliometrischer Verfahren wie der Häufigkeiten und des Impacts von Publikationen, Zitaten oder Patenten (EconSight 2019; SBFI 2016, 2017; IDAG KI 2019: 42ff.).

Die daraus gezogenen Schlussfolgerungen lauten in der Regel wie folgt: Die Schweiz bzw. ihr Bildungssystem ist sehr gut bis hervorragend aufgestellt, findet internationale Anerkennung und leistet in gewissen Bereichen weltweite Spitzenforschung (Bundesrat 2016; SBFI 2017; Economiesuisse & W.I.R.E. 2017; SECO 2017a). Obwohl auch Schwächen angesprochen werden (wie die mangelnde Chancengleichheit, die tiefe MINT-Quote allgemein und der niedrige Frauenanteil in technischen Studiengängen im Speziellen), findet eine kontinuierliche Selbstversicherung über die eigene »Spitzenposition« statt. Das Ziel, »[den] Forschungs- und Innovationsstandort Schweiz auf einem internationalen Spitzenplatz [zu] halten«, wird ab 2012 zum zentralen Kernpostulat der bildungs- und forschungspolitischen Bemühungen.

Demgegenüber werden die Herausforderungen der »digitalen Transformation« offen und allgemein gehalten. Die bildungspolitischen Akteur*innen imaginieren Innovation sowie Wissens- und Technologietransfer als zentrale Instrumente, um dem vermeintlichen »Rückfall« der Forschungs- und Entwicklungsleistung von Schweizer Universitäten, Hochschulen und Unternehmen im internationalen Vergleich zu begegnen. Das Szenario eines drohenden Verlusts des weltweiten »Spitzenplatzes« bildet den Negativhorizont, vor dessen Hintergrund es »rasch« und »koordiniert« zu handeln gilt (SBFI 2017: 41ff.). Obwohl beispielsweise anerkannt wird, dass die beiden ETH gemessen an ihrer Grösse besonders viele und einflussreiche Publikationen in den für die Digitalisierung zentralen Forschungsbereichen aufweisen, begründet allein die absolute Anzahl Professuren im internationalen Vergleich (»mangelnde Kapazitäten«) zusätzliche finanzielle Mittel in Millionenhöhe für die beiden technischen Hochschulen.

Die zunehmende Orientierung an internationalen Referenzen ist nicht alleine durch den politischen Diskurs zur Digitalisierung induziert: Im Bereich der obligatorischen Schulbildung haben beispielsweise die PISA-Studien und andere Referenzrahmen internationale Vergleichbarkeit und normativen Handlungsdruck geschaffen (Münch 2009). Im universitären Feld werden Hochschulen durch Rankings in ein globales Feld integriert, das spezifische Hierarchien und Visibilitäten erzeugt (Heintz 2008, 2010), die wiederum zu Rückkopplungen auf bildungs- und forschungspolitische Aktivitäten führen. So legitimieren die Akteur*innen ihre Aktivitäten und Förderinitiativen stets durch Verweise auf ähnliche Programme in anderen Bildungs- und Forschungssystemen (vornehmlich Staaten der EU, Nordamerikas sowie Ostasiens) (SNF 2015, 2017; SBFI 2017; Expertengruppe 2018; IDAG KI 2019).

Der politische Diskurs und die beschlossenen Massnahmen tragen ferner dazu bei, das statistische Dispositiv in Bezug auf die Digitalisierung neu auszurichten (GDS 2016: 4): Die öffentliche Statistik war lange nicht in der Lage, die notwendigen Daten für internationale Vergleiche zur Digitalisierung von Bildung und Forschung (wie bei-

spielsweise der Anteil Personen mit einem Tertiärabschluss im IKT-Bereich oder die »digitalen Kompetenzen« der Bevölkerung) zur Verfügung zu stellen. Erst der Aufbau eines Indikatorensystems²¹ zur Erfassung der Digitalisierung sowie die kontinuierliche Aktualisierung der Indikatoren ermöglichen es der Schweiz, sich in Bezug auf bestimmte Metriken mit anderen Staaten in Beziehung zu setzen, diese in Ranglisten zu überführen und zu bewerten. Die wechselseitige Beobachtung befördert somit die Koordination von verteilten Aktivitäten über verschiedene Bildungs- und Forschungssysteme hinweg. Im Ergebnis trägt die Anwendung ähnlicher Strategien und Massnahmen dazu bei, die weitere Entwicklung entstehender Wissensgebiete – wie der Datenwissenschaften – über nationalstaatliche und Feldgrenzen hinaus zu strukturieren und zu stabilisieren.

Die Analyse verdeutlicht, inwiefern die untersuchten hochschul- und forschungspolitischen Massnahmen sowie die diskursiven Mittel Ziele und Inhalte der Strategien in bildungs- und forschungspolitische Konzepte transformieren. Kompetenzen, Innovation, Adaption und Internationalität übersetzen die Kontingenzen der Digitalisierung in prozessierbare Grössen, die für die Akteur*innen der Bildungs- und Forschungspolitik anschlussfähig sind. In den Zukunftsentwürfen werden die Datenwissenschaften als ein neues Wissensgebiet gerahmt, um einerseits gesellschaftliche Problemstellungen mithilfe »digitaler Technologien« zu bearbeiten, andererseits um im internationalen Standortwettbewerb mithalten zu können. Die damit einhergehende instrumentalistische Inanspruchnahme und ein einseitiges, technikdeterministisches Fortschrittsdenken blenden indes vielseitige soziale, politische oder organisationale Aspekte tendenziell aus. Die Orientierung an technischen und ökonomistischen Rationalitäten wiederum ist nicht neu, sondern stellt vielmehr eine zentrale, historische Leitlinie helvetischer Bildungs- und Forschungspolitik seit der zweiten Hälfte des 20. Jahrhunderts dar. Insofern bleibt der bildungs- und forschungspolitische Diskurs trotz veränderter Begriffe und Transformationsrhetorik der Digitalisierung an Stabilität und Kontinuität orientiert.

6.4 »Data turn«: Auf dem Weg zur Datengesellschaft?

»Mit anderen Worten vollzieht sich das Soziale in steigendem Maße datenvermittelt und datenbasiert, wenn nicht gar datengetrieben, und zentrale gesellschaftliche Bereiche reproduzieren sich mit Rückgriff auf Daten. Eine Gesellschaft, die fortwährend reflektiert und reflexiv Daten produziert, sich mittels dieser Daten in ihren zentralen Bereichen reproduziert, Lebenschancen datenbasiert verteilt, sich zunehmend anhand von Daten selbst beschreibt, das Thema Daten intensiv diskutiert und sich in vielerlei Hinsicht in eine (un-)bewusste Abhängigkeit gegenüber ihren eigenen Datenbeständen begibt, wollen wir hier als Datengesellschaft bezeichnen.« (Prietl & Houben 2018: 7)

21 Das aktuelle Messsystem entspricht den Handlungsfeldern der SDS mit den Indikatoren Bildung, Forschung und Innovation; Infrastruktur; Sicherheit; Natürliche Ressourcen und Energie; Politische Partizipation und E-Government; Wirtschaft; Daten, digitale Inhalte und künstliche Intelligenz; Soziales, Gesundheit und Kultur. Vgl. online: <https://www.bfs.admin.ch/bfs/de/home/statistiken/kultur-medien-informationsgesellschaft-sport/informationsgesellschaft/strategieindikatoren.html> (Zugriff: 03.02.2022).

Die Zukunftsvisionen, die sich mit der Digitalisierung von Bildung und Forschung auseinandersetzen, schreiben Daten eine fundamentale Rolle zu: Sie werden als »Rohstoff« oder als »Treiber« für wissenschaftlichen Fortschritt, Innovation oder Wirtschaftswachstum betrachtet, sind also fundamental in die Funktionsweise zeitgenössischer Gesellschaften eingeschrieben. Parallel dazu werden die ökonomischen und technischen Potenziale und Herausforderungen der »exponentiell wachsenden [...] Datenbestände« vermehrt thematisiert. Wie erwähnt, gehören Daten und künstliche Intelligenz zu jenen Themenfeldern, die im untersuchten Zeitraum kontinuierlich an Bedeutung gewinnen (vgl. Abbildung 6). Anfangs kaum relevant, tauchen »Daten« seit der SIG 2006 zunächst vor allem in Bezug auf ethische und rechtliche Aspekte wie Datenschutz und die Privatsphäre von Individuen auf (Abun-Nasr 2009). Mit der SDS 2016 wird der Datenbegriff ubiquitär, während gleichzeitig die subjektorientierte Perspektive auf Daten in den Hintergrund rückt. Politische Akteur*innen beginnen, soziale Felder anhand von Daten zu beschreiben und zu rahmen. Zudem werden Konturen einer eigenen »Datenpolitik« formuliert und in den Kontext bestehender politischer Ziele gestellt: Es sind dies – wie ich in der Analyse zeigen werde – wiederum die internationale »Wettbewerbsfähigkeit« des schweizerischen Bildungs- und Forschungssektors einerseits sowie die Nichtverfügbarkeit eines Arbeitskräftenreservoirs mit den entsprechenden Ausbildungen und Fähigkeiten zur Bearbeitung grosser Datenmengen (»Fachkräftemangel«) andererseits. Ich analysiere im Folgenden die Rolle von Daten in kollektiven Zukunftsentwürfen (Guenduez et al. 2020) der Schweizer Bildungs- und Forschungspolitik und wie Vorstellungen sozialer Ordnung mit wissenschaftlich-technologischem Fortschritt einerseits und ökonomischen Opportunitäten andererseits verknüpft werden.

6.4.1 Metaphern von Daten und Datenwissenschaften in soziotechnischen Zukunftsvisionen

Soziotechnische Imaginationen von Wissenschaft und Technologie stützen sich auf kognitive Formate wie Metaphern, Narrative oder Bilder, die ihren Sinn durch Einbettung in spezifische räumlich-zeitliche und kulturelle Kontexte erhalten (Jasanoff 2015; Mische 2009). Insbesondere die bundesrätlichen Strategien, aber auch andere Dokumente im Korpus formulieren unterschiedliche Visionen, wie Daten bzw. Datenwissenschaften zukünftig die gesellschaftlichen Verhältnisse strukturieren und prägen werden. Im Folgenden skizziere ich vier wichtige Zukunftsvisionen in Bezug auf die gesellschaftliche Rolle von Daten bzw. Datenwissenschaften, die sich im Material empirisch rekonstruieren lassen, und diskutiere die darin implizierten Datenverständnisse.

Eine erste zentrale Rahmung von Daten ist die Extraktionsmetapher, der zufolge Daten die »Rohstoffe« der Informations- bzw. Wissensökonomie oder das »neue Öl des 21. Jahrhunderts« darstellen (ETH-Rat 2016a: 2; Schwab et al. 2011). Die Metapher erweist sich als umso bedeutender für ein Land wie die Schweiz, das sich selbst wiederkehrend als »rohstoffarm« beschreibt (SDS 2018: 3; Bundesrat 2016: 3138; SBFI 2017: 2; Economiesuisse & W.I.R.E. 2017) und infolgedessen besonders auf Bildung, Forschung und Innovation angewiesen sei. Der resultierende Wettbewerbsnachteil verwandelt sich infolge der intensiven Durchdringung der »Informations- und Wissensgesellschaft Schweiz« mit »Digitalisierungstechnologien« sowie den bei deren

Nutzung anfallenden Daten gewissermassen in einen Wettbewerbsvorteil: Es bietet sich nun die Möglichkeit, quasi durch Selbstschöpfung der digitalisierten Praktiken und Prozesse an Daten als »Rohstoffe« zu gelangen. Die Datenflüsse sind da, man muss sie nur »anzapfen«. Ausserdem böten Daten gegenüber natürlichen Rohstoffen den Vorteil, dass sie sich »heute in einem bis vor kurzem undenkbar Ausmass und mit Grenzkosten nahe dem Nullwert kopieren, versenden, weiterverarbeiten, speichern und allen zur Verfügung stellen« lassen (Expertengruppe 2018: 35). Unabhängig von ihren topographisch-geologischen Gegebenheiten ist die Schweiz demnach »herausragend aufgestellt, um die Geschäftsmöglichkeiten zu nutzen«, die sich durch die Verfügbarkeit von Daten »als einem zentralen Wirtschaftsgut« der Zukunft ergeben (ETH-Rat 2014: 51).

Die Metapher beschreibt hier einen Prozess der »digitalen Durchdringung« sozialer Lebensbereiche, in der digitale Aktivitäten sowie computerisierte Prozesse und Systeme zur Grundlage von »Rohstoffen« werden. Vor diesem Hintergrund ist die politische Lösung für eine intensivierete Digitalisierung der Gesellschaft als ein zirkulärer, sich selbst beschleunigender Prozess zu sehen: Durch vermehrte Vernetzung und Nutzung digitaler Technologien fallen zusätzliche Daten an, die wiederum neue »Geschäftsmöglichkeiten« erlauben und so weiter.

Eine zweite in den Dokumenten oft verwendete Rahmung bezieht sich auf Daten zur Prozessoptimierung sowie Steigerung von Effizienz und Transparenz. Trotz der Popularität der »Rohstoff«-Metapher wird betont, dass erst die Weiterverarbeitung, d. h. – wiederum in der Terminologie des Bergbaus – die »Anreicherung« und »Verfeinerung« von Daten zu einer »Wertsteigerung« führt (ETH-Rat 2016a: 2). Den Datenwissenschaften als »Basiswissenschaft« der »Digitalisierung« (SBFI 2017: 70) fällt insofern die Aufgabe zu, aus den Unmengen an potenziell verfügbaren Daten »aussagekräftige Modelle zu erstellen und mit Hilfe der Datenanalytik intelligente Entscheidungen zu ermöglichen« (ETH-Rat 2014: 51). Verallgemeinert sollen die gewonnenen Daten dabei helfen, »Produktions-, Betriebs-, Vertriebs- und Verwaltungsprozesse aller Art zu verbessern, Risiken zu reduzieren, beim Eintretenfall schneller zu reagieren, Verhaltensweisen und Ereignisse besser voraussehen zu können und entsprechend zu handeln« (Expertengruppe 2018: 37). Prozessoptimierung, Risikominimierung und Antizipation sind demnach die zentralen Managementstrategien, um organisationale und gesellschaftliche Frage- und Problemstellungen datenbasiert zu bearbeiten.

Die Überführung des Sozialen in digitale datenverarbeitende Maschinen und Systeme ist nicht wirklich neu, sondern erfolgt kontinuierlich seit Mitte des 20. Jahrhunderts (Gugerli 2018). Dennoch erlebt dieser (industrielle) Rationalisierungsstos in den untersuchten Dokumenten eine erneute Konjunktur. Dabei beinhalten soziotechnische Systeme der Datenverarbeitung nicht mehr nur Versprechungen von Objektivität und Effizienz, sondern auch von Offenheit, Transparenz und Nachvollziehbarkeit als zentralen Funktionsprinzipien (SDS 2018: 14; Expertengruppe 2018: 137). Die Verknüpfung von Rationalisierungs- und Visibilitätsnarrativen durchzieht zudem mehrere an die SDS anknüpfende Strategien.²²

22 Es sind dies die Strategien für offene Verwaltungsdaten 2019–2023 (Bundesrat 2018: 2), Open Access (Swissuniversities 2017: 2), Open Science (Swissuniversities 2019: 9) und Dateninnovation (BFS 2017: 9).

Drittens sind »Daten als Rohstoffe« eingebunden in umfassendere soziotechnische Transformationsprozesse der »4. industriellen Revolution«, die insbesondere die Rhetorik von ökonomischen Akteur*innen in industriellen, technologiegetriebenen Feldern (Dickel & Schrape 2015; Nachtwey & Seidl 2017) kennzeichnen²³: Mit der fortschreitenden Automatisierung industrieller Produktionsabläufe werde die »Fertigung von personalisierten Massenkonsumgütern möglich«, wodurch sich ein »grundlegender Wandel der Produktions- und Wertschöpfungsketten [abzeichnet]« (Bundesrat 2016: 3131). Neben der Sicherheit und Zuverlässigkeit solcher Produktionssysteme be ruht »[das] Konzept von Industrie 4.0 [...] auf der Möglichkeit, leistungsfähige Analysen sehr grosser Datenmengen durchzuführen« (ebd.). Die Datenwissenschaften gelten insofern als »das neue Forschungsparadigma, mit dem diese Vision [die 4. industrielle Revolution, PS] realisiert werden soll« (ETH-Rat 2016a: 2). Darüber hinaus wird eine Parallelität von Zukunftsvisionen der »4. industriellen Revolution« (Hirsch-Kreinsen et al. 2018; Meyer 2020) und datenintensiver Wissenschaft als »4. wissenschaftliches Paradigma« (Hey et al. 2009) (vgl. Kap. 7.2) postuliert. Den Datenwissenschaften kommt die Rolle als gesellschaftliches *boundary object* zu, indem sie die Analyse grosser Datenmengen in beiden Praxisfeldern ermöglichen, was Grundlage für epochalen Wandel (»Revolution«, »Paradigma«) bzw. den »nächsten Entwicklungsschritt der Gesellschaft« (Expertengruppe 2018: 25) ist.

Die auf die industrielle Produktion zugeschnittene Zukunftsvision wird in den Dokumenten von der Feststellung begleitet, wonach auch frühere Revolutionen zu ökonomischen Verlagerungen und Arbeitsplatzverlusten geführt hätten, dass aber langfristig sowohl die Zahl der Arbeitsplätze als auch der gesellschaftliche Wohlstand gestiegen seien (SECO 2017a: 27; SECO 2017b: 35; Economiesuisse & W.I.R.E. 2017: 24). Zu beobachten ist ferner, dass insbesondere die bundesrätlichen Strategiedokumente, die sich an die breite Öffentlichkeit richten, bei der Verwendung dieses Narrativs auf Terme wie »Revolution« oder »Disruption« verzichten. Vielmehr übersetzen sie die Rahmung in bildungs- und forschungspolitische Begrifflichkeiten, indem sie diese mit bekannten Kategorien des politischen Diskurses wie dem Upskilling von Berufen, lebenslangem Lernen oder »Fachkräftemangel« verknüpfen.

Schliesslich ist das vierte oft verwendete Zukunftsnarrativ – wiederum mit den vorangegangenen zusammenhängend – dasjenige der »Vernetzung« und »Intelligenz« von soziotechnischen Systemen: Die »4. Industrielle Revolution wird angetrieben von der Vernetzung intelligenter, miteinander verknüpfter Systeme« (ETH-Rat 2016a: 2). Seit der SIG 2012 stützt sich der politische Diskurs zur Digitalisierung vermehrt auf Begriffe wie Interdisziplinarität, Vernetzung und Intelligenz. Insbesondere der »intelligenten« bzw. »smarten« Steuerung, Nutzung oder Verwendung bestimmter Technologien oder Infrastrukturen (Energie, Mobilität, Ressourcen etc.) kommt eine wichtige Rolle zu. Erst die »künstliche Intelligenz« soziotechnischer Systeme ermögliche es, aus der stetig anwachsenden »Datenflut« sinnvolle Erkenntnisse zu generieren (SDS 2018: 14; Tschan 2019; IDAG KI 2019). Daten als »Strukturelement« von intelligenten Systemen wird dabei eine fundamentale Notwendigkeit attestiert:

23 Auch das Narrativ der »Revolution« stellt Daten in Analogie zur Rolle von Energieträgern in früheren sozioökonomischen Transformationsprozessen (wie der Kohle, der Elektrizität oder des Erdöls) (Bundesrat 2016: 3131; Economiesuisse & W.I.R.E. 2017: 90).

»Es besteht eine zwingende Abhängigkeit zu Daten, welche in mancher Hinsicht die Achillessehne heutiger [Künstlicher-Intelligenz-]Systeme (Aufwand für initiales Training, Sicherstellen der Aktualität der trainierten Systeme, Gefahr von Bias basierend auf den Daten) darstellen. Denn [Künstliche-Intelligenz-]Systeme können nicht nur gut mit grossen Datenmengen umgehen, die meisten Methoden benötigen umgekehrt auch grosse Datenmengen, um erfolgreich trainiert zu werden. Dabei nimmt der Nutzen von immer mehr Daten, die zum Trainieren zur Verfügung stehen, nicht ab« (IDAG KI 2019: 25f.).

Die Metapher der letalen Verletzlichkeit soziotechnischer Systeme künstlicher Intelligenz in ihrer und durch ihre Abhängigkeit von Daten geht folglich über die zuvor geschilderten Extraktions-, Rationalisierungs- und Visibilitätsnarrative hinaus: Daten werden fortan als kritische »Infrastruktur«²⁴ konzipiert und schreiben sich somit in die Funktionsweise der Schweiz als einer Datengesellschaft ein, indem sich diese »in eine (un-)bewusste Abhängigkeit gegenüber ihren eigenen Datenbeständen begibt« (Prietl & Houben 2018: 7).

Als Zukunftsvision wird die Kollaboration von Mensch und Maschine präsentiert, so »dass zunehmend komplexe Tätigkeiten vollständig durch Roboter oder in enger Zusammenarbeit mit dem Menschen erledigt werden können« (SECO 2017a: 34). Daraus resultiert das Bild einer hybriden, »vernetzten« Gesellschaft, in der elementare ökonomische, soziale oder politische Prozesse datenbasiert und weitgehend autonom ablaufen. Damit einher gehen grundlegende Freiheits- und Emanzipationsversprechungen (Economiesuisse & W.I.R.E. 2017: 93), indem autonome intelligente Systeme den Grossteil repetitiver Arbeit übernehmen würden, während man sich als Nutzerin bzw. Arbeitnehmerin fortan auf die kreativen und sozialen Aspekte von Arbeit konzentrieren könne (IDAG KI 2019: 61).

Wie die verwendeten Metaphern und Narrative von Daten zeigen, wird der Datenbegriff in den Dokumenten unterschiedlich spezifiziert, bleibt aber – analog zum Begriff der Digitalisierung – insgesamt eher vage und unterbestimmt. Daten werden in den Zukunftsvisionen primär mit positiven Attributen gerahmt: Nebst ihrer Funktion als »Rohstoffe« für die Datenwirtschaft sollen sie transparent, offen, vertrauenswürdig, intelligent oder »von hoher Qualität« sein, die Datenbewirtschaftung »nachhaltig und ressourcenschonend« (SDS 2018: 14). Dadurch werden ihre Eigenschaften mit übergeordneten politischen Zielen wie Innovation, Transparenz oder Nachhaltigkeit in Verbindung gebracht. Im Sinne des Risikodiskurses werden zwar auch mögliche negative Effekte adressiert; diese werden allerdings oft nicht den Daten selbst, sondern den sie verarbeitenden soziotechnischen Systemen bzw. Algorithmen zugeschrieben, was Aspekte wie Unvollständigkeit, Fehlerhaftigkeit oder Verzerrung, die alle auf die Konstruiertheit von Daten verweisen, weitgehend neutralisiert (mit Ausnahme der instruktiven Ausführungen in IDAG KI 2019: 33f.).²⁵

24 »Daten als Infrastruktur zu bezeichnen, bedeutet, sie als Teil der Basisausstattung für das gute Funktionieren von Gesellschaft, Politik und Wirtschaft zu verstehen« (Bundesrat 2018: 3).

25 Zudem können Daten und Algorithmen konzeptuell kaum getrennt werden (Dalton & Thatcher 2014; Kitchin & Lauriault 2014): So lassen sich zum Beispiel die fehlende Transparenz und Nachvollziehbarkeit von algorithmischen Systemen auch durch Daten »von hoher Qualität« weder adressieren noch beheben.

Durch die Verwendung spezifischer Imaginationen und positiv konnotierter Attribute sprechen die untersuchten Dokumente den Daten einen eigenen Status, eine eigene Wertigkeit zu; sie signalisieren damit, dass Daten – wie andere »Rohstoffe« – *an sich* existieren. Dabei handelt es sich bei Daten immer um eine »methodisch generierte Repräsentation« sozialer Realität (Prietl & Houben 2018: 16): Daten sind nicht an sich transparent, offen oder intelligent (Gitelman & Jackson 2013), sondern werden erst dazu gemacht: einerseits durch Praktiken, Methoden und soziotechnische Systeme, durch die sie generiert, gespeichert und prozessiert werden, andererseits durch die sie begleitenden Imaginationen und Narrative in technischen, medialen, ökonomischen oder politischen Diskursen (Dourish & Gómez Cruz 2018). Datafizierung beinhaltet also neben den »Prozesse[n] der Überführung sozialer Wirklichkeit in (vorgelblich) objektive Datenstrukturen und die Nutzung dieser Daten« (Prietl & Houben 2018: 7) stets auch die narrative Konstruktion und Repräsentation solcher Prozesse.²⁶ Die identifizierten Metaphern zeigen ferner, dass Datenpraktiken in der politischen Diskussion reinterpretiert, d. h. umgedeutet und mit etablierten Narrativen gesellschaftlicher Selbstbeschreibung verknüpft werden. So schliesst etwa die Metapher der »Daten als Rohstoffe« an die wiederholte Erzählung der Schweiz als »rohstoffarmes Land« an und transformiert diesen vermeintlichen Wettbewerbsnachteil in einen Vorteil. Dadurch verbindet sie einen zentralen Topos der Schweizer Bildungs- und Forschungspolitik mit einer der bedeutendsten Datenmetaphern (»Big Data«) des wissenschaftlichen und ökonomischen Diskurses (Mayer-Schoenberger & Cukier 2013). In diesem wird die soziotechnische Herausforderung der anfallenden »Datenflut« im digitalen Zeitalter in eine volkswirtschaftliche Chance für die Schweiz umgedeutet – und durch die Betonung der technologischen und ökonomischen Opportunitäten die sozialen, ethischen und rechtlichen Implikationen von Daten an andere Instanzen verschoben.

6.4.2 Daten als neues Politikfeld

Die verwendeten Datenmetaphern (Dourish & Gómez Cruz 2018) in der politischen Diskussion der letzten Jahre indizieren einen diskursiven Wandel, der die Relevanz von Daten zunehmend jenseits der rechtlich-ethischen Debatten um Datenschutz und die Wahrung der Privatsphäre reflektiert: Neben ökonomischen und wissenschaftlichen Akteur*innen beginnen sich auch der Staat und seine Verwaltungseinheiten für »neue soziale Daten« und ihre Verarbeitung in soziotechnischen Systemen zu interessieren (Jarchow & Estermann 2015; BFS 2017; Tschan 2019; IDAG KI 2019). Seit 2016 enthalten die Strategien einen eigenen Abschnitt »Daten und digitale Inhalte«. Zudem formuliert der Bund Ansätze für eine »kohärente und zukunftsorientierte Datenpolitik« (BAKOM 2018), die massgeblich von einer Logik der Zugänglichkeit und Transparenz von Daten (Open Access und Open Data) geprägt wird.

Trotz dieses »data turn« bleiben die bundesstaatlichen Aktivitäten von einer gewissen Zurückhaltung gekennzeichnet (Guenduez et al. 2020): Es existieren zwar erste

26 Anders als Digitalisierung eignet sich Datafizierung schlecht als politische Lösung zur gesellschaftlichen Selbstbeschreibung: Zu offensichtlich sind die Analogien von »Big Data« zu den zahlreichen Überwachungsskandalen sowie Verstössen gegen Datenschutzregeln der letzten Jahre (Elish & boyd 2018).

Projekte in staatlichen Institutionen (BAKOM 2018; IDAG KI 2019: 88; Tschan 2019), diese werden allerdings überwiegend nicht als datafiziert repräsentiert. Darin offenbart sich eine gewisse Diskrepanz in der gesellschaftlichen Wahrnehmung und Bedeutungszuschreibung von Digitalisierung und Datafizierung: Die Schweiz wird zwar als hochdigitalisiert und mit »Digitalisierungstechnologien« durchdrungen charakterisiert (ISA IG 2016: 9ff.). Zahlreiche soziale Prozesse, von kommunikativen Praktiken und Medienkonsum über ökonomische Transaktionen bis hin zu staatlichen Dienstleistungen, operieren teilweise seit Jahrzehnten digital.²⁷ Sie werden allerdings nicht als datafizierte Prozesse beschrieben. Die Prävalenz von Datenpraktiken in zentralen gesellschaftlichen Bereichen ist inhärent mit ihrer gesellschaftlichen Wahrnehmung verknüpft. Dies manifestiert sich auch darin, dass die Bundespolitik die technologischen, wissenschaftlichen und ökonomischen Diskurse zu Daten und künstlicher Intelligenz mit einiger Verzögerung aufnimmt: Ein erstes Beispiel sind die legislativen Arbeiten zur Revision der Datenschutzgesetzgebung (aus dem Jahre 1993), die über ein ganzes Jahrzehnt andauern und Ende 2020 abgeschlossen wurden.²⁸ Die Diskussionen drehen sich im Wesentlichen um die Frage, ob neben Aktualisierungen des existierenden Gesetzes in der Schweiz die Vorgaben der europäischen Datenschutz-Grundverordnung (DSGVO) nachvollzogen werden sollen oder ob darüber hinausgegangen werden soll.

Ein zweites Beispiel bildet das »Zukunftsfeld« (Krause 2019: 8) künstliche Intelligenz: Obwohl künstlicher Intelligenz als wissenschaftlichem Grundlagen- und Robotik als Anwendungsfeld bereits in den 1990er-Jahren signifikante Förderung durch den SNF im Rahmen eines NFP zukam,²⁹ findet das Thema bis 2018 kaum Berücksichtigung im untersuchten politischen Diskurs. 2019 bildet insofern ein Wendejahr, da – auch in Anlehnung an zahlreiche internationale Initiativen (Christen et al. 2020: 93ff.) – verschiedene Berichte erscheinen, die die Rolle von Daten und Systemen künstlicher Intelligenz in unterschiedlichen Politikfeldern thematisieren (EconSight 2019; IDAG KI 2019; SATW 2019; SEFRI 2019; Christen et al. 2020). Jenseits einer »bundesinterne[n] Arbeitsgruppe«, die »Empfehlungen für einen transparenten und verantwortungsvollen Einsatz von künstlicher Intelligenz erarbeiten [soll]« (SDS 2018: 14), die in einem umfassenden Bericht zum Einsatz von Technologien künstlicher Intelligenz in verschiedenen Politikfeldern resultieren (IDAG KI 2019), werden keine weitergehenden Massnahmen getroffen, etwa eine eigene KI-Strategie, wie sie viele andere Länder erarbeitet haben. Vielmehr soll »die KI-Politik als wesentliche[r] Bestandteil der Strategie »Digitale Schweiz« berücksichtig[t]« werden (IDAG KI 2019: 15).

Trotz der identifizierten Diskrepanz in der Bedeutungszuschreibung von Digitalisierung und Datafizierung und der verzögerten Reaktion auf aktuelle Debatten im »Zukunftsfeld« Daten und künstliche Intelligenz indiziert die Emergenz von Daten als zentrales Politikfeld (Bieber 2015), dass die Politik die gesellschaftliche Rolle von

27 Neben der Wissenschaft sind auch die für die gesellschaftliche Selbstwahrnehmung zentralen ökonomischen Felder wie Pharmaunternehmen, Banken und Versicherungen hochgradig datafiziert.

28 Bundesamt für Justiz (2019): Stärkung des Datenschutzes. Online: <https://www.bj.admin.ch/bj/de/home/staat/gesetzgebung/datenschutzstaerkung.html> (Zugriff: 03.02.2022).

29 Schweizerischer Nationalfonds (2020): NFP 23 »Künstliche Intelligenz und Robotik«. Online: <https://www.snf.ch/de/JXxClR51SN5RS0Y3/page/fokusForschung/nationale-forschungsprogramme/nfp23-kuenstliche-intelligenz-robotik> (Zugriff: 03.02.2022).

Daten in unterschiedlichen Bereichen zunehmend als konstitutives Funktionsprinzip für deren Operieren anerkennt.

6.4.3 Bildungs- und forschungspolitische Massnahmen als Teil der »Datenpolitik«

Die neue »Datenpolitik« ist insofern weniger technologieorientiert, als die verwendeten Datenmetaphern vermuten lassen: Sie umfasst neben der Revision der Datenschutzgesetzgebung primär Aus- und Weiterbildungsangebote sowie Massnahmen zur »Stärkung der Forschung und Innovation«. Einen Schwerpunkt bildet die »Initiative for Data Science in Switzerland«, lanciert durch den ETH-Bereich im Rahmen der »Strategischen Planung 2017–2020« (ETH-Rat 2014, 2016a). Die vom Bund mit 50 Mio. CHF alimentierte Initiative beinhaltet das SDSC, zusätzliche Professuren sowie zwei neue Masterstudiengänge in Datenwissenschaften an der ETHZ und der EPFL (ETH-Rat 2014: 63).³⁰ Die Investitionen werden vor dem Hintergrund eines drohenden »Rückfall[s]« in den Forschungsleistungen getätigt, was als »zentrales Risiko« für den Forschungs- und Innovationsstandort Schweiz gerahmt wird (SBFI 2017: 53). Die Initiative orientiert sich explizit an homologen Aktivitäten anderer Nationalstaaten bzw. Akteur*innen in Forschung und Wissenschaft (SBFI 2017; ETH-Rat 2016a). Ferner wird sie als wichtiger Beitrag zur Lösung des »Fachkräftemangels« in der Datenökonomie beschrieben, »anhand derer die von der Schweizer Wirtschaft dringend benötigten Spezialisten in diesem Bereich ausgebildet werden sollen« (ebd.: 97). Ebenfalls Teil der neuen »Datenpolitik« sind die durch den Bundesrat bewilligten und durch den SNF lancierten NFP 75 »Big Data« (SNF 2015) sowie NFP 77 »Digital Transformation« (SNF 2018). Ersteres erforscht die Möglichkeiten von Big Data in der Informatik und Informationstechnologie einerseits sowie in Anwendungsfeldern andererseits und adressiert gleichzeitig deren gesellschaftliche, ökonomische, regulatorische sowie bildungsbezogene Herausforderungen. Bei Letzterem handelt es sich um die in der SDS angekündigte Untersuchung der gesellschaftlichen Auswirkungen der Digitalisierung auf Wirtschaft und Gesellschaft (GDS 2018: 3).

Die Förderung der Datenwissenschaften geht mit einer organisationalen Innovation, einem hochschul- und disziplinenübergreifenden Forschungszentrum mit enger Anlehnung an die Industrie, einher: Das SDSC nimmt eine »Schnittstellenfunktion« zwischen wissenschaftlichen Disziplinen, Anwendungsfeldern und der Industrie ein. Es stellt eine klassische »boundary organization« (Åm 2013; Guston 1999) dar, die feldspezifische Praktiken und Logiken wissenschaftlicher Forschung mit Innovations- und Wettbewerbspolitik integriert (Blümel 2018: 25). Daneben hat der Bund mit dem nationalen »Kompetenzzentrum für Datenwissenschaften (Data Science Competence Center)« eine eigene Vermittlerinstitution zwischen seinen verschiedenen administrativen Einheiten, der Wissenschaft und der Industrie geschaffen (Bundesrat 2020). Das Kompetenzzentrum soll einerseits technisch-methodische Expertise aufbauen, andererseits bundeseigene Datenbestände verwaltungsübergreifend erfassen, verknüpfen und zugänglich machen, »um das Potenzial der KI in der Bundesverwaltung

30 Weiter unterstützt der Bund neue Studiengänge in Cyber Security (ETHZ/EPFL), einen berufsbegleitenden Master in Artificial Intelligence (FernUni Schweiz) und das Massnahmenprogramm »Digital Skills«.

auszuschöpfen« (IDAG KI 2019: 89). Beide Zentren verfolgen eine primär technologiezentrierte Perspektive, indem sie rechtliche, ethische, ökonomische oder soziale Aspekte der Datenwissenschaften und künstlicher Intelligenz explizit nicht zum Gegenstand machen.

Auch andere bildungs- und forschungspolitische Massnahmen werden in den Kontext der neuen »Datenpolitik« gestellt: Beispielsweise koordinieren das SBFI, der SNF und Swissuniversities ihre Bemühungen und integrieren den offenen Datenzugang in der Wissenschaft in den Strategien zu Open Science und Open Access. Zudem haben Bundesrat und Parlament eine sogenannte »Wissenschaftsschranke« im revidierten Urheberrechtsgesetz eingeführt, »die sicherstellen [soll], dass insbesondere einer automatisierten Auswertung grosser Datenmengen (mittels sog. Text- und Data-Mining) keine urheberrechtlichen Klauseln entgegenstehen« (BAKOM 2018: 2) – wobei dies wiederum explizit als Mittel zur »Stärkung des Forschungs- und Wirtschaftsstandortes Schweiz« (ebd.) begründet wird.

Was die identifizierten Aktivitäten der entstehenden »Datenpolitik« kennzeichnet, ist die Vermengung und Einbindung von Standort- und Wettbewerbsrhetorik in bildungs- und forschungspolitische Massnahmen: Insbesondere die Datenwissenschaften gelten in den relevanten politischen, ökonomischen und wissenschaftlichen Bezugskontexten als bedeutender »Wettbewerbsfaktor«. Die Reduktion der epistemischen Transformationen, ökonomischen Opportunitäten und politischen Konflikte, die Daten als Politikfeld markieren, auf dieses zentrale Argument des untersuchten Diskurses macht eine Abgrenzung zwischen den relevanten Bezugsfeldern schwierig und porös (Ribes et al. 2019: 297). Damit ordnet sich die Diskussion in eine historische Kontinuität von bildungs-, forschungs- und wirtschaftspolitischen Verflechtungen und Interessen im schweizerischen Hochschulfeld ein. Sie reproduziert dabei aber eine ökonomistische Deutung, die weder der gesellschaftlichen Bedeutung noch der Komplexität und Multidimensionalität des Phänomens Datenwissenschaften gerecht wird.

Die Analyse zeigt, dass die hochschul- und forschungspolitischen Akteur*innen Daten und Datenwissenschaften durch spezifische Metaphern rahmen, die sich eng an existierende Rationalisierungs-, Visibilitäts- und Fortschrittsnarrative anlehnen. Durch die Verknüpfung mit etablierten Narrativen gesellschaftlicher Selbstbeobachtung machen sie Zukunftsvisionen anschlussfähig für die Gegenwart und binden sie in aktuelle politische Diskurse ein. Über den untersuchten Zeitraum hinweg manifestiert sich eine zurückhaltende und verzögerte Berücksichtigung von Daten als zentrales Politikfeld. Erst in den letzten Jahren anerkennen Politik und Verwaltung zunehmend die konstitutive Bedeutung von Daten als zentrale Funktions- und Operationsprinzipien in unterschiedlichen Feldern. Damit einher geht die Repräsentation der Schweiz als einer Datengesellschaft. Die bildungs- und forschungspolitischen Massnahmen der entstehenden »Datenpolitik« sind in die existierende Standort- und Wettbewerbsrhetorik eingebunden. Diese sind primär durch technologiezentrierte Perspektiven auf Daten bzw. Datenwissenschaften markiert, wodurch deren Multidimensionalität erheblich reduziert wird. Darin offenbart sich eine Diskrepanz zwischen der diskursiven Bedeutungskonstruktion und organisationalen Implementierungen, insbesondere in den neuen Kompetenzzentren für Datenwissenschaften in den ETH sowie der Bundesverwaltung.

6.5 Diskussion

Das Kapitel hat in drei Schritten untersucht, wie diskursive Strategien und Massnahmen der Digitalisierung operieren, wie Bildung und Forschung in diesem Diskurs gerahmt werden und welche Zukunftsvorstellungen über Daten und Datenwissenschaften in diesen Dokumenten kursieren. Eine erstaunliche Kontinuität und Stabilität markiert den politischen Diskurs zur Digitalisierung: Trotz einer ausgeprägten Prozess- und Transformationssemantik schliesst die Strategie »Digitale Schweiz« an etablierte politische Ziele wie Wohlstandsvermehrung und Wettbewerbsfähigkeit von Wirtschaft und Wissenschaft an. Die »digitale Transformation« wird als eine kontinuierliche, lineare Entwicklung imaginiert, die dadurch gewissermassen planbar wird – und somit zur Reduktion der Unsicherheiten in Bezug auf die künftige Entwicklung der Digitalisierung sozialer Felder wie Bildung, Forschung oder Wirtschaft beiträgt. Die Prozessesemantik signalisiert insofern Stabilität und Kontinuität gerade durch den eingeschlagenen Transformationsprozess, in dem alles gleich bleibt, weil es sich verändert (Brunsson 2017; Esposito 2016).

Bildung und Forschung werden im politischen Diskurs sowohl als Triebkräfte als auch als »Getriebene« der Digitalisierung imaginiert, indem sie permanent Innovationen produzieren, synchron dazu aber immer auch die Neuheiten anderer Felder verarbeiten und adaptieren müssen. Diverse bildungs- und forschungspolitische Aktivitäten werden als Massnahmen zur Innovationsförderung getroffen und sind entsprechend darauf ausgerichtet, Feldgrenzen zu überschreiten und insbesondere das wissenschaftliche stärker mit dem ökonomischen Feld zu verlinken. Dabei sind die Offenheit und Mehrdeutigkeit zentraler Begriffe wie »Digitalisierung«, »Kompetenzen« oder »Innovation« weniger eine Schwäche des untersuchten Diskurses als vielmehr ein strategisches, weil strukturelles und verbindendes Element: Sie erlauben Akteur*innen unterschiedlicher Felder, sich strategisch darauf zu beziehen, um feldübergreifende Kollaborationen zu etablieren. Insofern tragen die Multidimensionalität und Ambiguität solcher Praktiken von Begriffsarbeit zur Koordination der Akteur*innen über das hochschul- und forschungspolitische Feld hinaus bei.

Die Analyse der Zukunftsvisionen und Narrative verdeutlicht, dass den Datenwissenschaften in der politischen Diskussion mehr zukommt als der Status einer neuen interdisziplinären Wissenschaft: Sie werden als grundlegende »Basiswissenschaft« gerahmt, um die Herausforderungen und Probleme der Zukunft datengestützt zu bearbeiten. Zusammen mit anderen hybriden wissenschaftlichen Grundlagenfeldern wie künstliche Intelligenz oder Anwendungsbereichen wie Robotik gelten sie als zentrale Faktoren für die »Wettbewerbsfähigkeit«, nicht nur von Bildung und Forschung, sondern auch der Wirtschaft. Die Investitionen der Bundespolitik tragen insofern zur Stabilisierung solcher Wissensgebiete bei, was wiederum Anreize für andere, vor allem ökonomische und akademische Akteur*innen setzt, sich ebenfalls in diesen feldübergreifenden Bereichen zu betätigen. Die Stellungnahmen, Massnahmen und Investitionen von Akteur*innen im Feld der Bildungs- und Forschungspolitik haben insofern fundierende Effekte in den Räumen zwischen den Feldern von Politik, Wissenschaft und Ökonomie, in denen solche Arrangements entstehen und sich entwickeln.

Der Rückgriff auf Datenmetaphern stellt ein zentrales Element in der Konstruktion und Bedeutungszuschreibung soziotechnischer Systeme dar. Datenmetaphern

erweisen sich als besonders hilfreich, weil sie es erlauben, die Verbindung herzustellen zwischen Systemen und Prozessen der Datenverarbeitung einerseits und den bekannten kulturellen Formaten des politischen Diskurses andererseits. Trotz der unterschiedlichen Metaphern und Narrative, die zur Beschreibung von Daten verwendet werden, dominieren Perspektiven, die Daten bzw. deren Prozessierung in erster Linie als technische Probleme rahmen, die es mittels Engineering-Lösungen zu bearbeiten und ökonomisch zu valuieren gilt. Die Multidimensionalität und Ambiguität des Gegenstandes, die sich in der Analyse der Stellenanzeigen präsentiert hat, wird dadurch erheblich reduziert, wenn rechtliche, ethische oder soziale Aspekte von Datenwissenschaften und künstlicher Intelligenz in den organisationalen Implementierungen ausgeblendet bzw. an andere Institutionen delegiert werden.

Die Förderung bestimmter, als »Zukunftstechnologien« gerahmter Bereiche wie Datenwissenschaften, künstliche Intelligenz und Robotik ist ein zentrales Element der neuen »Datenpolitik«. Darin manifestiert sich auch die enge Anlehnung an das Narrativ des Standortwettbewerbs, das Bildungs- und Forschungspolitik in wirtschaftspolitische Interessen einbindet, wenn nicht gar gleichsetzt. Dies verweist auf historische Kontinuitäten der Verknüpfung von wissenschaftlichen, arbeitsmarkt- und bildungspolitischen Argumentationen im Wissenschafts- und Hochschulfeld der Schweiz. Gerade durch die Verwendung offener, ambiguer Begrifflichkeiten und der ubiquitären Transformationsrhetorik markiert der Diskurs somit eine nur vermeintlich widersprüchliche Orientierung an Stabilität und Kontinuität. Indem die Digitalisierung im politischen Diskurs durch die aufgezeigten grundlegenden Orientierungen und Ziele als kontinuierlich und linear gedacht und somit plan- und berechenbar gemacht wird, verbleibt sie als eine »gezähmte« Transformation zurück.

Teil III – Konstruktionen der Datenwissenschaften im akademischen Feld

Im dritten Teil der Arbeit untersuche ich die Rolle von Akteur*innen des akademischen Feldes sowie ihrer Praktiken bei der Genese der Datenwissenschaften. Zu den zentralen Akteur*innen im akademischen Feld zähle ich sowohl individuelle (wie Lehrende, Studiengangleitende und Forschende) als auch kollektive Akteur*innen (wie Studiengänge, Institute oder Forschungszentren). Beide tragen durch ihre jeweiligen Praktiken, Positionierungen und Investitionen zur Herausbildung organisations- und feldspezifischer Ausprägungen der Datenwissenschaften bei. So können beispielsweise interdisziplinär zusammengesetzte Curricula als Kompromissprodukte betrachtet werden, die das Ergebnis individueller wie organisationaler Kollaborationen und Konflikte sind. Darüber hinaus sind sie nicht unabhängig von Interessen ausserhalb des akademischen Feldes zu denken: Wie ich zeigen werde, prägen insbesondere Akteur*innen in ökonomischen Feldern (Unternehmen) sowie im Feld der Hochschulpolitik durch ihre Bezugnahmen und Investitionen die Emergenz und Implementierung des Wissensfeldes. Kollaborationen mit Unternehmen und Organisationen anderer Felder sind integrierter Bestandteil in verschiedenen Curricula und die darin skizzierten Kompetenzprofile lehnen sich eng an ausserwissenschaftliche Referenzfelder an. Datenwissenschaftliche Curricula als Netzwerke verschiedener Expertisen verbinden somit gewissermassen Akteur*innen aus multiplen Feldern. Die Effekte dieser vielfältigen Kollaborationen und Konflikte sind somit nicht nur im akademischen Feld zu lokalisieren, sondern auch in einem Raum zwischen den verschiedenen involvierten Feldern.

Die Ergebnisse in den folgenden Kapiteln markieren eigentümliche Verschränkungen und Synchronizitäten vermeintlich widersprüchlicher Prozesse der Suche nach Neuheit und Transformation auf der einen Seite gegenüber Prozessen von Institutionalisierung bzw. der Reproduktion von bestehenden Strukturen auf der anderen Seite: So verorten die untersuchten Akteur*innen die Genese der Datenwissenschaften im Kontext eines fundamentalen Wandels wissenschaftlicher Erkenntnisproduktion, der durch Begriffe wie Paradigmenwechsel oder Interdisziplinarität umschrieben wird. In der Suche nach adäquaten Begrifflichkeiten und Kategorien für diese Transformationsprozesse offenbaren sich Kollaborationen und disziplinäre Grenzüberschreitungen, die interdisziplinäre Inhalte in institutionellen Innovationen wie Forschungszen-

tren oder Curricula implementieren und organisieren. In Letzteren aktualisieren sich bestehende, insbesondere ingenieurwissenschaftliche Ausbildungstraditionen, während gleichzeitig neue Kompetenzprofile und Berufsfelder geschaffen werden.

Den Innovationen und Veränderungsprozessen stehen allerdings synchrone Prozesse von Institutionalisierung und Reproduktion im akademischen Feld entgegen: Trotz der neuen, feldübergreifenden Kategorien und veränderten epistemischen Praktiken deuten die Ergebnisse auf die Persistenz disziplinärer Grenzziehungen und Konflikte sowie die Kontinuität feldspezifischer Differenzierungen hin. Die strukturell ungleichen Kapitalverhältnisse sowie die Beobachtungs- und Wahrnehmungsweisen bleiben im untersuchten Feld weitgehend erhalten. Der Dynamik der Transformation sowie der beschleunigten Implementierung neuer Curricula steht die Trägheit politischer Vorgaben und organisationaler Aushandlungsprozesse entgegen. Inhaltlich ist das Feld bereits in einem frühen Stadium seiner Entwicklung durch Kanonisierungsprozesse im interdisziplinären Kernbereich gekennzeichnet und es bilden sich Strukturlogiken heraus, die die untersuchten Studiengänge durchziehen. Die resultierenden Curricula unterscheiden sich – je nach Positionierung der verantwortenden Hochschulen und Universitäten im Feld – insofern nur geringfügig voneinander. Schliesslich reproduzieren die neuen Skillsets und kulturellen Narrative im untersuchten Berufsfeld trotz Versprechungen von Offenheit und Demokratisierung existierende soziale Ungleichheiten.

Die qualitative Inhaltsanalyse von Interviews mit Studiengangleitenden und Dozierenden sowie von Curricula an Schweizer Universitäten und Hochschulen fördert vier fundierende Praktiken zutage, durch welche Akteur*innen im akademischen Feld die Datenwissenschaften als neues Wissensgebiet konstruieren und in denen sich die beschriebenen vermeintlich widersprüchlichen Entwicklungen manifestieren. Die identifizierten Praktiken bilden die Struktur der Kapitel 7 bis 10 dieses dritten Teils der Arbeit: Stellungnahmen und Reflexionen zur Transformation des akademischen Feldes sind ebenso wie politische und ökonomische Rahmenbedingungen auf einer gesellschaftlichen Makroebene verortet. Dies zeigt sich zunächst in der Suche nach neuen Begriffen sowie in disziplinären Grenzziehungen und -überschreitungen, mit denen individuelle wie kollektive Akteur*innen die Neuheit des Phänomens Datenwissenschaften im akademischen Feld bearbeiten (Kapitel 7). Die daran anknüpfenden hochschulstrategischen Überlegungen adressieren in erster Linie die Meso-Ebene von Universitäten bzw. Hochschulen. Sie machen organisationale Verhandlungen darüber sichtbar, wie die Inhalte eines interdisziplinären Wissensfeldes in curriculare Inhalte überführt werden sollen (Kapitel 8). Ebenfalls auf Ebene der Organisation stellt sich die Frage, wie die Curricula durch spezifische Strukturlogiken die Wissensinhalte und deren Vermittlung organisieren (Kapitel 9). Schliesslich übersetzen auf der Mikroebene Kompetenzprofile und Zuschreibungen individueller Eigenschaften die curricularen Wissensinhalte in operationalisierbare Anforderungen an gegenwärtige und zukünftige Praktiker*innen der Datenwissenschaften (Kapitel 10).

Kapitel 7 – Die Konstruktion der Datenwissenschaften im akademischen Feld durch Begriffsarbeit und *boundary work*

»We are witnessing an Abbott (1988) style disciplinary turf war, with computer science, statistics, mathematics, information science, and various collaborations with scientific, state, and industrial sectors each making claims to the expertise, funding, and regulatory sphere of data science. No single configuration seems to hold, and at each academic institution I have visited or inspected, I have found a different arrangement of interlinked disciplines and sectors (and those left out or cut out) of the game« (Ribes 2019: 515f.).

7.1 Einleitung

Das akademische Feld ist eine zentrale gesellschaftliche Instanz für die Genese und Konstituierung neuer Wissensfelder. Sowohl individuelle als auch kollektive Akteur*innen entwerfen unterschiedliche disziplinäre, epistemische, organisationale und weitere Perspektiven auf entstehende Wissensgebiete. Die dabei entstehenden vielfältigen Bedeutungen rahmen und strukturieren einerseits Forschungsschwerpunkte, Professuren oder Studiengänge und positionieren diese andererseits im akademischen Feld. So materialisieren sich in der Struktur und im Aufbau neuer Curricula, aber auch in organisationalen Implementationen wie Forschungsinitiativen, -zentren oder -netzwerken insofern je nach epistemologischer, disziplinärer und institutioneller Verortung multiple Verständnisse der jeweiligen Wissensformation.

Daran anknüpfend untersuche ich, wie die Datenwissenschaften im akademischen Feld eingeführt werden und in welchem Verhältnis sie zu verwandten Wissensfeldern und Disziplinen stehen. Der Fokus der Analyse liegt dabei neben den Akteur*innen auf den Diskursen und Praktiken, die die Kategorie »Data Science« im akademischen Feld konstruieren. Zur Beantwortung der Fragestellungen stütze ich mich auf die inhaltsanalytische Auswertung qualitativer Interviews mit Lehrenden an Schweizer Hochschulen und Universitäten und stelle sie in Relation zu akademischen Diskursen über die Emergenz der Datenwissenschaften. Die Analyse identifiziert zwei zentrale Praxismodi in der Konstruktion der Datenwissenschaften im akademischen Feld. Erstens stellt Begriffsarbeit eine zentrale diskursive Praxis dar: Datenwissenschaften werden im akademischen Feld durch Definitionen, Kategorisierungen und Narrative

eingeführt und hergestellt. In den angeführten Begriffen artikulieren sich einerseits Unklarheiten bzw. divergierende Ansichten darüber, ob es sich um einen fundamental neuen Modus wissenschaftlicher Erkenntnisproduktion handelt, der alle Disziplinen transformieren wird, oder aber um eine neue Disziplin, die sich zunächst gegenüber anderen Disziplinen behaupten muss und sich schliesslich ins disziplinäre Gefüge einpassen wird. Andererseits manifestieren sich darin Suchprozesse nach den »richtigen« Begriffen zur Beschreibung der beobachtbaren Veränderungen im akademischen Feld.

Zweitens sind Praktiken von *boundary work*, d. h. Grenzziehungen und -überschreitungen, fundierend in der Genese der Datenwissenschaften im akademischen Feld. Sie finden sich im empirischen Material in zwei unterschiedlichen Dimensionen, die für die involvierten Disziplinen und Akteur*innen eine wichtige Orientierungsfunktion haben. Zentral ist in einer ersten Dimension das Verständnis als Wissensfeld, das sich aus unterschiedlichen wissenschaftlichen Disziplinen zusammensetzt. Dabei wird ein interdisziplinärer »Kern« in den Computerwissenschaften, Statistik, Mathematik und Engineering verortet, der sich auf Methoden, Technologien und andere Formalisierungen abstützt. Wie ich bereits im Forschungsstand gezeigt habe, erheben diese – und einige weitere – Disziplinen Anspruch auf die epistemische Autorität für das Feld der Datenwissenschaften. Im empirischen Material artikuliert sich dies in intensiven Konflikten und Grenzziehungen zwischen den involvierten Disziplinen, insbesondere zwischen der Informatik und der Statistik. Gleichzeitig äussern sich darin allerdings auch Praktiken der Grenzüberschreitung, indem ein interdisziplinärer Kern, der auf jedes Forschungsfeld oder -problem anwendbar und in diesem Sinne universal ist, durch verschiedene Anwendungsbereiche, sogenannte *domains*, in der Wissenschaft, der Industrie, im Non-Profit-Sektor, in der staatlichen Verwaltung und in weiteren Feldern komplementiert wird. Die Differenzierung in eine interdisziplinär zusammengesetzte methodisch-technische Expertise und eine domänenspezifische Fachexpertise stellt das zentrale Organisationsprinzip vieler Institutionen der Datenwissenschaften dar und transzendiert somit aktiv etablierte Grenzen zwischen verschiedenen Disziplinen.

Eine zweite Dimension betrifft die grundlegende Unterscheidung zwischen »Wissenschaft« und »Industrie«,¹ die auf unterschiedliche Handlungslogiken in sozialen Feldern im Sinne gesellschaftlicher Differenzierung abzielt. Im empirischen Material findet die Dichotomie eine Übersetzung in akademische Grundlagenforschung einerseits und angewandte Forschung in der Praxis andererseits. Obwohl die Grenzziehung historisch stets porös war (Ribes et al. 2019: 297), bleibt sie für viele interviewte Datenwissenschaftler*innen eine wichtige Referenz. Gleichzeitig wird durchaus anerkannt, dass die Differenzierung aufgrund transversaler Forschungskollaborationen und Karrierewege sowie eines beschleunigten Transfers von Wissen, das in Methoden, Instrumenten und Technologien kodifiziert ist, über die unterschiedlichen Felder hinweg noch instabiler wird. Zu beobachten sind insofern unterschiedliche Spielarten von Anwendungsorientierung, die sich in der Unterscheidung von echten und konstruierten Daten und Problemen einerseits und einer an ausserwissenschaftlichen Referenzen orientierten »operativen Exzellenz« andererseits manifestieren.

1 »Industrie« wird dabei über die wirtschaftsstatistische Einteilung hinaus breit verstanden und umfasst sowohl traditionelle Technologiekonzerne (wie ABB) als auch Unternehmen der IT- und Softwarebranche.

Das Kapitel ist entlang der beiden zentralen Praktiken strukturiert, die Datenwissenschaften im akademischen Feld konstruieren: Zunächst untersuche ich, inwiefern sich am Begriff des Paradigmenwechsels multiple Verständnisse von Datenwissenschaften zwischen Universalisierung und Disziplinierung dokumentieren (Kap. 7.2). Anschliessend zeigt die Suche nach den richtigen Kategorien, wie unterschiedliche Begriffe angeführt werden, um die veränderten Modi der Wissensproduktion adäquat zu beschreiben (Kap. 7.3). »Data Science« als inklusiver Kategorie kommt dabei die Funktion der Verknüpfung unterschiedlicher Disziplinen und Bezugsfelder zu. Die nachfolgenden Teilkapitel behandeln den zweiten identifizierten Praxismodus: Praktiken von *boundary work* äussern sich zum einen in disziplinären Konfliktlinien, die Datenwissenschaften als Wissensfeld durchziehen und strukturieren (Kap. 7.4). Zum anderen werden durch die kontinuierliche Ausrufung von Interdisziplinarität und die Logik der Domänen aber auch etablierte Grenzen überschritten (Kap. 7.5). Schliesslich tragen die verschiedenen Spielarten von Anwendungsorientierung zur Differenzierung zwischen unterschiedlichen Positionen im akademischen Feld bei (Kap. 7.6). In einer abschliessenden Diskussion werden die zentralen Erkenntnisse der Analyse zusammengefasst und in Bezug zu den übergeordneten Fragestellungen der Arbeit gesetzt (Kap. 7.7).

7.2 Paradigmenwechsel? »Data Science« zwischen Universalisierung und Disziplinierung

Ein erstes zentrales Moment in der Konstruktion der Datenwissenschaften ist die Frage nach dem grundlegenden Verständnis der Wissensformation. Eine Vielzahl an Definitionen, Positionsnahmen und Begriffsverständnissen kursiert in wissenschaftlichen, ökonomischen, medialen und weiteren Diskursen, die alle spezifische Perspektiven auf die epistemische Formation werfen (vgl. für einen Überblick Kap. 3.1). Ein wichtiger Teil der Diskussion bezieht sich auf die veränderten Bedingungen der Wissensproduktion, die mit der Verfügbarkeit grosser Datenmengen und neuen Analysemöglichkeiten einhergehen. Diese kristallisieren sich anhand des Begriffs Paradigmenwechsel heraus, der multiple Verständnisse von Datenwissenschaften zwischen Universalisierung und Disziplinierung prägt.

Alle Interviewpartner*innen teilen die sehr allgemein gehaltene Definition, wonach sich »Data Science« vorrangig mit »Methoden und Techniken zur Analyse von Daten in verschiedenen Anwendungsgebieten« beschäftigt, wie es ein Professor formuliert (vgl. unten). Die Befragten identifizieren die dahinterstehenden soziotechnischen Veränderungen wie die permanente Verfügbarkeit von grossen Datenmengen aus unterschiedlichen Quellen sowie neue Systeme zu deren Speicherung, Prozessierung, Analyse oder Visualisierung als die Treiber einer Entwicklung, die gesellschaftsdiagnostisch als Datafizierung (Prietl & Houben 2018; Schäfer & van Es 2017; van Dijk 2014) beschrieben wird. Manche der befragten Personen waren bereits vor Jahrzehnten mit grossen Datenmengen konfrontiert, lange bevor Begriffe wie »Big Data« oder »Data Science« überhaupt bekannt waren:

»Als ich an der [Grossforschungsinstitution] war, haben wir brutal viele Daten produziert [...]. Heute produziert nicht mehr nur die Forschung, sondern überall werden sehr

viele Daten produziert. Unternehmen möchten diese nutzen können« (Prof_FH_A: 8:26).

Die Verfügbarkeit von Daten, neuen Technologien zu deren Speicherung und Prozessierung und methodische Innovationen kennzeichnen jene veränderte Epistemologie, d. h. Art und Weise, wie Wissen und Erkenntnisse produziert werden. Die Diagnose eines »computational turn« (Berry 2011) beschreibt nicht in erster Linie die Genese einer neuen Disziplin oder eines neuen Feldes, sondern eine umfassende Transformation wissenschaftlicher Erkenntnisproduktion an sich.² Im empirischen Material wird diese Entwicklung vor allem durch den Begriff des Paradigmenwechsels auf den Punkt gebracht: Er bezeichnet den Übergang zu einem neuen Modus der Wissensproduktion, der auf der Verfügbarkeit von »Big Data« und neuen Möglichkeiten zu deren Analyse basiert (Gray 2009: xviii; Kitchin 2014).

Ein Professor für Mathematik umschreibt den Paradigmenwechsel anhand seines eigenen Werdegangs wie folgt:

»[Ich] könnte sagen, es geht um Methoden und Techniken zur Analyse von Daten in verschiedenen Anwendungsgebieten. Das gab es früher nicht, es ist eine extreme Beschleunigung von mathematischen und statistischen Innovationen, erstmals seit Jahrzehnten. Plötzlich [gibt es] sehr grosses Interesse einer riesigen Zahl von Leuten, [das] habe ich in meiner 25-jährigen Tätigkeit in diesem Feld noch nie erlebt, vor allem in der Mathematik und deren Verknüpfung mit technologischen Aspekten und Fragen. Heute beschäftigen sich Armeen von Leuten damit und schaffen völlig neue Angebote, nicht nur in der Forschung, sondern vor allem in der Lehre, und das grösstenteils gratis! Ich denke wirklich, dass wir in aussergewöhnlichen Zeiten leben, es ist nicht nur ein weiterer Hype! [lacht] [...]. Es ermöglicht ganz viele neue Fragestellungen in verschiedenen Gebieten und bietet Möglichkeiten für Anwender in der Industrie, aber auch Akademiker« (Prof_UH_C: 3:44).

Kennzeichnend für den Paradigmenwechsel ist nicht nur die Art und Weise, wie wissenschaftliches Wissen heutzutage produziert wird, die der Mathematikprofessor im Zitat euphorisch beschreibt, sondern die artikulierte Verknüpfung von wissenschaftlicher Forschung, Methoden der Datenanalyse und Anwendungen »in der Industrie« bzw. in »Unternehmen«, die in beiden Zitaten deutlich wird. Die Immanenz feldüberschreitender Bezüge ist ein wiederholtes Charakteristikum der Konstruktion der Datenwissenschaften, sowohl in den Interviews als auch im feldspezifischen Diskurs: Methoden und Techniken zur Analyse von Daten sind zwar im wissenschaftlichen Feld entwickelt worden, ihre eigentliche Bedeutung erhalten sie allerdings erst in Anwendungsmöglichkeiten der *domains* – wissenschaftliche und vor allem industrielle Anwendungsgebiete,³ in denen es »echte Probleme« zu lösen gilt. Indem der datenintensive Modus der Erkenntnisproduktion zur Grundlage des Paradigmenwechsels gemacht wird, wird diesem gleichzeitig eine universelle Relevanz für alle Disziplinen

2 Der Wandel äussert sich unterschiedlich in wissenschaftlichen Feldern wie den Sozialwissenschaften (Lazer et al. 2009; Chang et al. 2014), der Biologie (Leonelli 2016) oder der Medizin (Trajanoski 2012).

3 So beantworten diverse Studienprogramme die Frage nach der Verortung des Paradigmenwechsels dahingehend, dass »Data Science« in erster Linie »aus den Unternehmen« komme (EPFL 2017: 3).

attribuiert. Damit positionieren sich die Datenwissenschaften in diesem Sinne als »universelle Wissenschaft« (Ribes 2019; Ribes et al. 2019), die nicht nur für die Wissenschaften, sondern für *alle* gesellschaftlichen Bereiche fundamental ist.

Die Repräsentation von »Data Science« als einem Paradigmenwechsel und – daran anknüpfend – einer »universellen Wissenschaft« ist eine wichtige Lesart des Feldes. Parallel dazu begleitet die Datenwissenschaften von Beginn weg die Frage, wie sich ein transversales Wissensgebiet in ein disziplinär strukturiertes Feld einpassen lässt. Akademische Diskussionen über die Inklusion und Exklusion bestimmter Disziplinen in bzw. aus dem Kernbereich der Datenwissenschaften führen zu intensiven epistemischen und organisationalen Konfliktlinien. Neben dem akademischen Feld beabsichtigen in erster Linie politische und industrielle Akteur*innen, durch eine Differenzierung in verschiedene Subgruppen und Kompetenzprofile die Konturen des Feldes zu schärfen und die Bedürfnisse des Arbeitsmarktes mit den verfügbaren Kompetenzen in Übereinstimmung zu bringen (BHEF 2016; Demchenko et al. 2016; PwC & BHEF 2017). In den Interviews äussert sich dies in Form wiederholter Desiderata nach »Landkarten« oder »Mappings« der verschiedenen Kompetenzprofile – nicht zuletzt, um die Ausrichtung der eigenen Studiengänge zu justieren und Kopplungen zwischen Universitäten und der Nachfrage in der Industrie herzustellen.

Solche Bestrebungen nach einer Differenzierung in spezifische Subgruppen und klaren Grenzziehungen stehen einer diskursiv als universal konstruierten Wissensformation entgegen. Sie führen letztlich zu einer Disziplinierung »im Sinne einer Unterwerfung unter Regeln« und »mit einer deutlichen Einschränkung der individuellen Gestaltungsfreiheit zu Gunsten kollektiver Autonomie« (Schultheis 2005: 67f.). Als disziplinierte Wissensformation stehen die Datenwissenschaften in Widerspruch zur permanenten Akklamation von Interdisziplinarität, die diesen Raum diskursiv als einen mit offenen Grenzen herstellt.

7.3 Die Suche nach den richtigen Kategorien

Mit den Debatten über Universalisierung und Disziplinierung der Datenwissenschaften wird die Frage nach den passenden Kategorisierungen und Begrifflichkeiten virulent, um die fundamentalen epistemologischen Veränderungen zu beschreiben, die sich in Diagnosen eines Paradigmenwechsels artikulieren. Es stellt sich deshalb die Frage, welche Begriffe zur Charakterisierung des neuen Wissensfeldes verwendet werden, inwiefern sie sich unterscheiden und weshalb sich »Data Science« schliesslich als Kategorie durchsetzen kann.

Verschiedene Begriffe werden wiederholt angeführt, um die veränderten Bedingungen der Wissensproduktion zu fassen. Dazu zählen neben »Data Science« insbesondere »Big Data«, »Machine Learning« sowie »künstliche Intelligenz«. Die Kategorien repräsentieren nicht nur divergente Wissensformationen, sondern weisen je eigene Historizitäten und Konjunkturen auf (Aronova et al. 2017; Nilsson 2009). Für die interviewten Lehrenden ist »Data Science« als Kategorie in dieser Hinsicht von zweitrangiger Bedeutung: Einerseits wird der Begriff als »Google-fähig« betrachtet, was heutzutage wichtig sei, und »[von] dem her ist das Label eigentlich nicht schlecht« (Prof_FH_D: 12:66). Zudem sei »Data Science« hilfreich für die »berufliche Positionierung« der Studierenden in einem stark expandierenden Sektor des Arbeitsmark-

tes. Andererseits wird der Begriff aber von manchen eher kontrovers betrachtet im Sinne eines diffusen »Container for everything« (Prof_UH_B: 3:45) oder gar als »ein Marketinginstrument, das nichts über die Inhalte aussagt« (Prof_UH_A: 6:20), abgestempelt.⁴ Wieder andere schliesslich sehen die Begriffswahl als pragmatisches Mittel zum Zweck, worauf ein Interviewter hinweist: »Vielleicht ist Data Science [einfach] ein besserer Begriff dafür, dass man irgendwie an der Schnittstelle von Informatik, Statistik, Mathematik und Business arbeitet« (Prof_FH_A: 8:29).

Diverse Befragte weisen darauf hin, dass »Data Science« bis 2014 kaum verbreitet war. Ein Professor an einer Fachhochschule beschreibt die Situation in jener Zeit so:

»Für die [erste Swiss Data Science] Konferenz 2014 fand ich auf LinkedIn keine Schweizer Data Scientists, die ich hätte anfragen können. [Wir] haben mit 40 Leuten gerechnet, es kamen 120 Leute« (Prof_FH).

Das neue Label stiess auf deutlich grösseres Interesse, als die Organisator*innen der Konferenz erwartet hatten. Dies wirkte sich in den folgenden Jahren auch auf die Verfügbarkeit von Bildungsangeboten aus: Gab es vor 2014 noch kein Programm unter diesem Titel, können heute an fast allen Schweizer Universitäten und Fachhochschulen Studiengänge, Vertiefungen oder Weiterbildungen in Datenwissenschaften belegt werden (vgl. Tabelle 13 im Anhang). Angesichts des zeitlichen Planungsvorlaufs von zwei bis drei Jahren indiziert das Angebot eine rapide Diffusion und Implementation der Kategorie »Data Science« innerhalb des schweizerischen Hochschulfeldes.

Die Vieldeutigkeit der zirkulierenden Begriffe zur Beschreibung des Feldes verdeutlicht sich anhand des feldübergreifenden Korpus der Stellenanzeigen, auf das ich im Detail im Kapitel 5 eingehe. Hier greife ich darauf zurück, um zu zeigen, inwiefern das untersuchte Wissensgebiet durch unterschiedliche Kategorien repräsentiert wird. Als Kategorien weisen »Data Science« und »Machine Learning« relativ ähnliche Inhalte und Verteilungen auf: »Machine Learning« wird enger gefasst und bezieht sich in erster Linie auf ein breites Methodenspektrum, das von linearen Regressionsanalysen bis zu neuronalen Netzwerken reicht. Gleichzeitig wird »Machine Learning« in nicht-technischen Diskussionen oft als Synonym für »künstliche Intelligenz« verwendet (Engemann & Sudmann 2017; Krause 2019). »Data Science« selbst enthält sowohl methodische als auch disziplinspezifische Begriffe und findet Anwendung auf eine Vielfalt unterschiedlicher wissenschaftlicher und ausserwissenschaftlicher Felder.

Demgegenüber unterscheiden sich die Anzeigen für den Suchbegriff »Big Data« erheblich von jenen für »Data Science« und »Machine Learning«. Sie beziehen sich explizit auf ein technisch-ingenieursorientiertes Vokabular und stellen neben Softwarelösungen Design und Architektur von Environments zur Prozessierung von Daten in den Vordergrund. Disziplinäres Vokabular ist mit Ausnahme computerwissenschaftlicher Begriffe selten. Die Inserate für »Big Data« weisen bis Anfang 2018 ein starkes Wachstum auf und sind seither rückläufig (vgl. Abbildung 4). Auf der diskursiven Ebene deutet sich ein Wandel weg von »Big Data« hin zu »Data Science« und »Machine Learning« an: Infolge der zunehmend negativ konnotierten öffentlichen Wahrnehmung

4 Die Debatte über die Nützlichkeit des Begriffs ist so alt wie die jüngere Verwendungsweise von »Data Science« selbst, vgl. den frühen Text von Pete Warden (2011): Why the term »data science« is flawed but useful. Online: <http://radar.oreilly.com/2011/05/data-science-terminology.html> (Zugriff: 03.02.2022).

von »Big Data« im Nachgang zu Überwachungs- und Spionageaffären (»Big Brother«) verzichtenden Unternehmen und Organisationen zunehmend auf den Begriff und verwenden andere Kategorien (Elish & boyd 2018).⁵

Die Verschiebungen auf der Makroebene finden ihr Äquivalent in Studiengängen, die im Entwicklungsverlauf ihre Bezeichnung geändert bzw. »präzisiert« haben, wie es ein Befragter ausdrückt. Diverse Befragte attribuieren dem Begriff »Big Data« nur geringe Wissenschaftlichkeit und taxieren diesen als nicht praktikabel. So grenzt ein Fachhochschulprofessor die Wahl für »Data Science« als neue Studiengangbezeichnung explizit von »Big Data« ab:

»Data Science scheint nicht schlecht als Begriff, besser als Big Data, weil damit noch mehr Themen assoziiert werden. Data Science impliziert Systematik und Umgang mit Technologie. Mit Big Data assoziiert man gesellschaftliche Daten, gesellschaftliche Veränderung insgesamt, also ein sehr weites Feld« (Prof_FH_C: 9:35).

Die Gegenüberstellung von »Systematik und Umgang mit Technologie«, d. h. einer wissenschaftlich-technischen Dimension, und der »gesellschaftliche[n]« Dimension, d. h. der politischen, ökonomischen, rechtlichen oder sozialen Implikationen der Datennutzung, erfährt im empirischen Material nur wenig Resonanz. Die zwei Dimensionen koexistieren mehr oder weniger unabhängig voneinander und werden in den Curricula selten in Beziehung gebracht. Indem der Bedeutungsgehalt von »Big Data« nur schwierig zu fixieren ist (Kitchin & McArdle 2016), manifestiert sich in der Verwendung der Kategorie eine Ambiguität bzw. »konzeptuelle Vagheit« (Favaretto et al. 2020).

»Machine Learning« wiederum wird zwar feldübergreifend verwendet, dient aber insbesondere zur Charakterisierung eines breiten Spektrums methodischer Praktiken. Die Kategorie findet in einer Vielzahl von Beschreibungen methodenbezogener Kurse Verwendung, aber auch in Stelleninseraten, die disziplinenübergreifend angeboten bzw. gesucht werden. Viele Disziplinen bedienen sich eines Konvoluts an statistischen Methoden, ohne allerdings damit ein neues Feld zu begründen; vielmehr erfolgt die Transformation der methodischen Praxis – etwa die »mathematischen und statistischen Innovationen« bzw. »ganz viele[n] neue[n] Fragestellungen in verschiedenen Gebieten« (vgl. Zitat oben) – innerhalb der existierenden wissenschaftlichen Disziplinen bzw. institutionellen Gefäße.

»Data Science« kommt insofern die Funktion der Verknüpfung unterschiedlicher Bezugsfelder und Disziplinen zu: Neue Forschungszentren oder Studiengänge werden mit »Data Science« gelabelt, weil andere Kategorien als zu eng (»Machine Learning«) bzw. als zu weit (»Big Data«, »Künstliche Intelligenz«) taxiert werden. Gleichzeitig wurde »Data Science« als Wissenskategorie in den letzten zwei Jahrzehnten massgeblich durch Protagonisten der US-amerikanischen Tech- und Software-Industrie (Hammerbacher 2009; Varian 2009) gerahmt und symbolisch aufgeladen (vgl. Kap. 10.7). Die Attraktivität des Labels trifft darüber hinaus auf die Bereitschaft einer Vielzahl von Akteur*innen (vgl. das Zitat des Mathematikprofessors), nicht nur wissenschaftliche Forschungsergebnisse, sondern auch Daten, methodische Innovationen oder Software zur öffentlichen Verfügung zu stellen, was wiederum im Einklang steht

5 Um die These zu überprüfen, ist allerdings eine Betrachtung über einen längeren Zeitraum notwendig.

mit wissenschaftsinternen Entwicklungen wie der Reproduzierbarkeitskrise und der *open-science*-Bewegung. Vor diesem Hintergrund kann »Data Science« in der Suche nach den richtigen Begriffen folglich als offene, transversale Kategorie interpretiert werden: Sie erlaubt es Akteur*innen aus verschiedenen Feldern und Disziplinen, sich unter diesem Label zu positionieren, aber dennoch im Herkunftsbereich tätig zu bleiben – was oft als die Interdisziplinarität der Datenwissenschaften umschrieben wird. Die Inklusivität der Kategorie ermöglicht es legitimerweise, multiple Rollen (als Forschende, Unternehmerinnen, Hacker etc.) in verschiedenen Feldern wahrzunehmen und darüber hinaus an jenen materiellen und symbolischen Ressourcen teilzuhaben, die im Rahmen der gegenwärtigen epistemologischen Transformation verteilt werden.

7.4 Innerwissenschaftliche Konfliktlinien

Neben der Begriffsarbeit und der Suche nach den passenden Kategorien bilden Praktiken von *boundary work* einen zweiten zentralen Praxismodus in der Konstruktion der Datenwissenschaften im akademischen Feld. Grenzziehungs- und Grenzüberschreitungsprozesse stellen ein konstituierendes Moment in der Herausbildung der Datenwissenschaften und insbesondere ihrer feldspezifischen Konstruktion in der Wissenschaft dar. Eine wesentliche Herausforderung der Datenwissenschaften seit ihrer Formierung besteht darin, wie ein Wissensgebiet, das auf unterschiedlichen technischen, methodischen und inhaltlichen Expertisen basiert, in ein disziplinär strukturiertes Feld überführt werden kann. In den Interviews wird oft problematisiert, dass solche Disziplinierungsbemühungen bestehende disziplinäre Konfliktlinien im akademischen Feld aktualisieren würden. Ein Professor an der ZHAW schildert den Konflikt der zwei »Kerndisziplinen« Statistik und Informatik anlässlich der Gründung eines interdisziplinären Forschungsnetzwerks in »Data Science«, des Datalab⁶, wie folgt:

»[Wir] haben an der ZHAW schnell gemerkt, dass es Zeit ist, etwas Neues zu gründen, interdisziplinär unter dem Stichwort Data Science: Wie können wir aus Daten Mehrwert schaffen, indem wir mit unterschiedlichen Brillen darauf schauen? [...] [Die] Kerngruppe waren Statistiker und Informatiker, die ähnliche Methoden verwenden, aber sich sonst spinnefeind sind. Später kamen Life Sciences, Juristen, Business-Leute oder auch quantitativ arbeitende Soziologen, Linguisten und Psychologen hinzu. [Es ist] ein interdisziplinär aufgestelltes Netzwerk, aber kollaborieren tun echte Menschen. Ziel ist es, dass sich echte Menschen zusammensetzen und gemeinsam Projekte entwickeln, Anträge schreiben und so weiter.« (Prof_FH)

Die zitierte Passage verweist auf die bisweilen widersprüchlichen Verschränkungen der Konfliktlinien zwischen Universalisierung und Disziplinierung: Darin äussert sich zum einen die Erkenntnis, dass Forschung und Wissensproduktion in Datenwissenschaften stets in Auseinandersetzung mit anderen Disziplinen durchgeführt werden. Die Offenheit und Multiperspektivität auf denselben Gegenstand, der hier hinter

6 Online: <https://www.zhaw.ch/de/forschung/departementsuebergreifende-kooperationen/datalab/> (Zugriff: 03.02.2022).

den »Daten« verschwindet, eröffnet somit einen Raum, damit sich das Forschungsnetzwerk entfalten kann.⁷ Die Inklusion weiterer Disziplinen über die »Kerngruppe« hinaus ermöglicht es, die im Netzwerk zirkulierende Expertise – im Zitat verkürzt als ›quantitatives Arbeiten mit Daten‹ gefasst – zu stabilisieren, ohne dass dadurch eine neue Entität entsteht. Im Sinne einer Universalwissenschaft können so Forscher*innen der beteiligten Disziplinen teilhaben, ohne ihre jeweilige Fachidentität aufgeben zu müssen.

Zum anderen wird darin gleichzeitig die Problematik manifest, dass zwar verschiedene Disziplinen mit »ähnlichen« Methoden Daten bearbeiten, aber den jeweiligen Bestandteilen des Forschungsprozesses wie Daten, Modellen oder Theorien sehr unterschiedliche Bedeutungen zuschreiben. Das Netzwerk unterschiedlicher Expertisen droht dabei durch disziplinäre Konflikte und »Feindschaften« jederzeit instabil zu werden. Über die epistemologischen Unterschiede, »die Konkurrenz der beiden Herangehensweisen« hinaus geht es in der Auseinandersetzung letztlich auch um Drittmittel, Forschungsprojekte sowie das »grössere Stück vom Big Data Kuchen« (Kauer mann & Küchenhoff 2016: 143). Obwohl mit dem Datalab eine neue organisationale Kollaborationsform zur Akkumulation von wissenschaftlichem und ökonomischem Kapital geschaffen wurde, bleiben die unterschiedlichen zur Verfügung stehenden Ressourcen also Gegenstand für soziale Konfliktlinien und Distinktionsprozesse.

Allen Betonungen von Kollaboration und Interdisziplinarität zum Trotz werden also wiederkehrend epistemologische Grenzen zwischen den Disziplinen artikuliert (Gieryn 1983, 1999), die auf den zentralen Elementen wissenschaftlicher Forschungsprozesse wie Methoden und Daten beruhen. Explizit wird dies in den Interviews – auch aufgrund der einseitigen Verteilung der Interviewpartner*innen – vor allem anhand der Disziplinen der Informatik und Statistik, die sich »spinnefeind« (vgl. oben) seien. In einem weiteren sehr prägnanten Beispiel schreibt ein Informatikprofessor der Statistik »eine sehr enge Sicht auf Daten« zu:

»Ob es dann Data Science heisst, ist mir egal. Aber es ist eben auch nicht Statistik! Weil die Statistiker haben eine sehr enge Sicht auf Daten und was man mit Daten tun darf und was nicht. Aber wir tun es trotzdem, weil es funktioniert. Weil wir so viele Daten haben. Viele der statistischen Verfahren, die Sie in Ihrem Studium gelernt haben, basieren auf Datensätzen mit 50 Datenpunkten. Aber bei 50'000 Datenpunkten sind andere Verfahren besser geeignet« (Prof_UH_A: 6:69).

Der springende Punkt ist hier weniger die effektive Grösse des Datensatzes, sondern »was man mit Daten tun darf und was nicht«. Angesprochen werden hier grundsätzliche epistemologische Differenzen zwischen den involvierten Disziplinen, die nun unter der Prämisse der interdisziplinären Zusammenarbeit reartikuliert werden: Das Narrativ einer ›normativen‹, modellgebundenen Statistik wird dabei einer ›explorativen‹, datengetriebenen Computerwissenschaft gegenübergestellt. Indem andere Disziplinen Anspruch auf einen traditionellen Hoheitsbereich des Feldes wie die (Weiter-)Entwicklung von statistischen Methoden zur Datenanalyse erheben, stellen die

7 »Datalab connects several departments and institutes at the ZHAW, creating an internal collaboration network for researchers and also an interface for external partners and organisations« (ZHAW 2016: 7).

Datenwissenschaften insofern für die Statistik eine spezifische Herausforderung dar (Grommé et al. 2018).

Umgekehrt sind epistemische Praktiken, die die Computer- und Informationswissenschaften für sich in Anspruch nehmen oder die diesen zugeschrieben werden, unter Statistiker*innen verpönt. So distanzieren sich der Statistik zugehörige Befragte beispielsweise explizit von »Data Mining« aufgrund »mangelnder Systematik« (Prof_UH_E: 13:2): Eine kontrollierte wissenschaftliche Arbeitsweise durchforstet ihrer Ansicht nach nicht nur Daten und sucht nach Korrelationen, wie dies prominent Chris Anderson (2008) postuliert hat, sondern formuliert gezielt Fragestellungen an die Daten, überprüft Hypothesen und kommuniziert Ergebnisse (Schutt & O’Neil 2013: 44) – eine Selbstbeschreibung, der sich auch die »Data Science«-Community regelmäßig bedient (Bowne-Anderson 2018; Strachnyi 2017).

An diese Grenzziehung – nun von der entgegengesetzten Seite her – anknüpfend, verwehrt sich ein Professor einer Fachhochschule, der sich selbst nicht als Statistiker beschreibt, sondern im »Analytics-Bereich« positioniert, einer Vereinnahmung der Datenwissenschaften durch die Informatik ebenfalls sehr pointiert:

»Noch heute glauben Informatiker, Data Science habe etwas mit Informatik zu tun. [Das] Verhältnis zwischen Informatik und Data Science ist wie das zwischen Limonade und einer Flasche: [...] Der Flasche kann man Verschiedenes abfüllen, der ist das egal. Der Limonade ist es auch egal, wo sie aufbewahrt wird, oder? Und [...] es gibt natürlich eine Schnittstelle: Ein Data Scientist muss eigentlich sagen können, was er will: Welche Daten, in welcher Form und so weiter. All diese Parameter, die wichtig sind. Er muss es auch gegenüber dem Informatiker kommunizieren können, aber er ist nicht derjenige, der das System kennt. Er muss auch nur begrenzt Prozesse kennen. Er muss einfach diese Schnittstelle beherrschen, oder?« (Prof_FH_D: 12:63)

Trotz der anfangs scharfen Grenzziehung gegenüber der Informatik verortet der Interviewte »Data Science« im Zitat weniger (inter-)disziplinär, sondern als ein Wissensfeld, das quer zu bestehenden wissenschaftlichen Einteilungen liegt und sich in verschiedenen »Flaschen«, d. h. Disziplinen, wiederfinden kann. Der Begriff der »Schnittstelle«, der sowohl in den Interviews als auch datenwissenschaftlichen Diskursen allgemein angeführt wird, verdeutlicht die räumliche Positionierung vieler Befragter, die ihre wissenschaftliche Praxis nicht in abgeschlossenen Bereichen, sondern an Übergangszonen situieren, wo sich verschiedene Felder überlappen und miteinander in Kommunikation stehen. Solche diskursiven Praktiken sind insofern Ausdruck jener Bemühungen, die Grenzen einer stark disziplinär strukturierten Wissenslandschaft zu transzendieren und gleichzeitig die nach wie vor existierenden internen Konfliktlinien latent zu halten.

Trotz der vielfältigen interdisziplinären Kollaborationen und Arrangements in Forschung und Lehre erweisen sich feldspezifische Grenzziehungen und Markierungen in solchen Settings als äusserst persistent. Die theoretisch-konzeptuelle Annahme von Interdisziplinarität, Zusammenarbeit und Austausch existiert parallel zu konflikthafter Beziehungen, die durch die Existenz gemeinsamer Kategorien und institutioneller Gefässe (wie interdisziplinärer Institute, Netzwerke oder Studiengänge) nicht einfach verschwinden. Wie insbesondere die Analyse zum Datalab zeigt, halten die involvierten Akteur*innen die Grenzziehungen latent: Sie beteiligen sich am

Aufbau eines interdisziplinären Forschungsnetzwerks, ohne zwingend die Positionen und Perspektiven der anderen Beteiligten zu teilen. Das resultierende Netzwerk von Expertisen ist sodann durch die Kopräsenz unterschiedlicher disziplinärer Wissensbestände, aber auch Interessen, Strategien und Machtverhältnisse charakterisiert. Somit münden die identifizierten gemeinsamen Kategorien und institutionellen Gefässe in synchrone Relationen von Konkurrenz und Kollaboration.

7.5 Die Logik der Domänen: Interdisziplinarität im Kernbereich und *domain knowledge*

Trotz der wechselseitigen Grenzziehungen teilen beide Positionen – sowohl Informatik als auch Statistik – die Feststellung, dass die nach wie vor »stark disziplinär getriebene« Verortung der Datenwissenschaften ein grosses Problem darstellt. Darin würden sich Kompetenzbereiche und Methoden (wie beispielsweise Datenmanagement, Datenanalyse, Scripting, Programmierung etc.) ergänzen, die sowohl aus der Statistik als auch aus der Informatik stammten. Obwohl an der Schnittstelle der beiden wichtigsten Bezugsfelder seit längerem praktiziert, existierten bis vor wenigen Jahren an den meisten Universitäten kaum explizite curriculare Gefässe und Strukturen dafür. Die Implementation der Datenwissenschaften in Forschungszentren, Netzwerke oder Studienprogramme kann in diesem Sinne als *boundary crossing* betrachtet werden, als Übertreten der etablierten Grenzziehung existierender Disziplinen.⁸ Im Sample manifestiert sich dies ferner auch auf personaler Ebene: Bei mehreren Befragten changieren die beruflichen Karrieren zwischen dem akademischen und industriellen Feld. Datenwissenschaftler*innen können zwischen den involvierten Feldern hin und her wechseln und multiple Rollen einnehmen (Safavi et al. 2018). Grenzüberschreitende Rollen haben eine lange Kontinuität in den besagten Feldern (Ribes 2019) und erfahren gegenwärtig in der Schnittstellenfunktion von Data Scientists eine erneuerte Konjunktur.

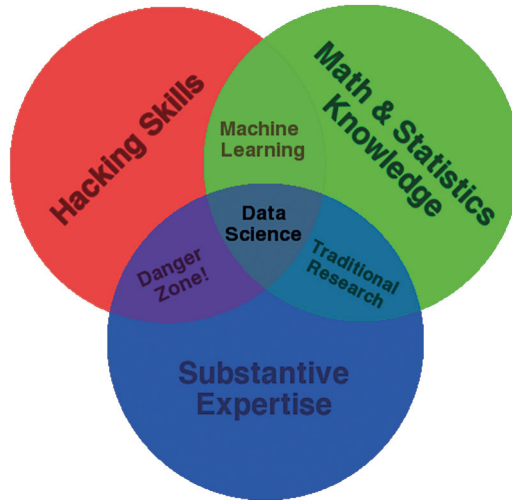
Interdisziplinarität zwischen den genannten informatischen und statistisch-mathematischen Kompetenzbereichen und Methoden alleine reicht für die Praxis als Data Scientist nicht aus: Es gilt sich in bestimmten Anwendungsbereichen, sogenannten *domains*, zu vertiefen und zu fokussieren. Domänen werden sehr breit definiert und können prinzipiell jede wissenschaftliche Disziplin oder gesellschaftliches Feld umfassen. Das »Data-Science-Venn-Diagramm« (vgl. Abbildung 10) repräsentiert die Datenwissenschaften als Schnittfläche einer interdisziplinär zusammengesetzten mathematisch-statistischen (»Math & Statistics Knowledge«), methodisch-technischen (»Hacking Skills«) sowie einer inhaltlichen Fachexpertise (»Substantive Expertise«).

Ribes et al. (2019) bezeichnen die Differenzierung in einen universellen, interdisziplinär zusammengesetzten Kern und spezifische Anwendungsdomänen als »the logic of domains«. Sie hat eine lange Geschichte in den Computerwissenschaften, im Engineering sowie der Forschung zu künstlicher Intelligenz (Dreyfus & Dreyfus 2005;

8 Andere Beispiele aus jüngerer Zeit sind Neuro- und Klimawissenschaften oder Cultural Studies (Brint et al. 2009; Holley 2009; Knight et al. 2013), wobei keines dieser Felder eine ähnlich rasche Institutionalisierung erfahren hat wie die Datenwissenschaften.

Malazita 2019; Ribes 2019; Ribes et al. 2019). Weil den Datenwissenschaften ein eigenes Anwendungsgebiet grösstenteils fehlt, positioniert sich das Feld durch die Verwendung der Logik der Domänen – analog den Computerwissenschaften im letzten Viertel des 20. Jahrhunderts – als allgemeine, universelle Wissenschaft, die für jeden gesellschaftlichen Bereich relevant ist.

Abbildung 10: Das Data-Science-Venn-Diagramm



Quelle: Drew Conway (2010) – eigenes Werk, CC BY-NC 3.0,
<http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>

In den Interviews präsentiert sich das Denken in der Logik der Domänen als zentrales Ordnungs- und Klassifizierungsprinzip von Lehrenden und Forschenden der Datenwissenschaften, das sich in unterschiedlichen Ausprägungen artikuliert. Ein Professor beschreibt dies wie folgt:

»Ich bin überzeugt, dass Interdisziplinarität ein Kern ist, wenn wir jetzt auf Data Science zurückkommen. Deshalb haben wir das Ganze so strukturiert, dass man Data Science entweder als Haupt- oder als Nebenfach wählen kann, das war eigentlich unser Ziel. Es geht um domain knowledge, die man entweder schon mitbringt, oder daneben als weitere Kompetenz sich aneignen kann« (Prof_UH_A: 6:70)

Interdisziplinarität wird nicht nur als Kombination der genannten Kompetenzbereiche, sondern in Verknüpfung mit einem Anwendungsfeld konzipiert. Indem »domain knowledge« in den Vordergrund rückt, wird die methodisch-technische Expertise, die gemeinhin den »Kern« von Definitionen und Curricula der Datenwissenschaften ausmacht, unsichtbar und universal zugleich. Zudem sind – wie ich noch zeigen werde – die untersuchten Studienprogramme in Konzeption und Aufbau (vgl. Kap. 9.5) als auch die Forschungsförderung in der Logik der Domänen strukturiert.

Die Konstruktion der Datenwissenschaften als universal und domänenunabhängig erlaubt es ansatzweise, die vorgängig diskutierten Limitationen von Disziplinie-

rung, adäquaten Begrifflichkeiten und innerwissenschaftlichen Konfliktlinien zu überwinden und zu neutralisieren. Indem der Schwerpunkt auf die Verknüpfung der unterschiedlichen Bezugfelder gelegt wird, überschreitet das Denken in der Logik der Domänen existierende disziplinäre und soziale Grenzziehungen und präsentiert sich somit als pragmatisch und anwendungsorientiert.

7.6 Spielarten von Anwendungsorientierung: »Echte Probleme« und »operative Exzellenz«

Die in der Logik der Domänen implizierte Anwendungsorientierung wird als ein Ansatzpunkt betrachtet, um die beschriebenen Spannungen zwischen quantitativ-methodischer Expertise (Universalisierung) einerseits und spezifischen Themenfeldern (*domains*) andererseits zu überbrücken. In den konkreten Anwendungsmöglichkeiten der Datenbanktechnologien, Algorithmen und Methoden in den Unternehmen wird – wie bereits erwähnt – »das eigentlich Neue« des Feldes verortet. Grosse Datenmengen würden keine Fragestellungen an sich beinhalten; solche könnten erst durch »Anwendungsorientierung und konkrete Projekte« gefunden werden. Je nach Verortung und Profil der untersuchten akademischen Institutionen äussern sich unterschiedliche Spielarten von Anwendungsorientierung: Im Falle der technischen Hochschulen sind Kollaborationen der Daten- und Computerwissenschaften mit spezifischen Domänen keineswegs neu. So bestehen etwa an der ETH datenbezogene Kooperationen mit Unternehmen, aber auch zwischen den involvierten Departementen über den neu geschaffenen Masterstudiengang hinaus schon sehr viel länger:

»Das passiert laufend, dass andere Projekte und Leute auf uns zukommen mit spezifischen Problemen bezüglich der Verarbeitung grosser Daten. [...] An der ETH wird täglich mit Firmen zusammengearbeitet. Ich habe das in meiner Promotion selbst gemacht, das wird nicht spezifisch sein. Denn die Unternehmen bringen die echten Probleme. Aber wir arbeiten auch mit anderen Departementen zusammen, die ebenfalls mit Datenproblemen zu uns kommen« (Prof_ETH_B: 2:9).

In dieser ersten Ausprägung von Anwendungsorientierung, die nicht nur bei Professor*innen an den ETH zu beobachten ist, sind es Forschungspartner*innen aus anderen wissenschaftlichen Disziplinen, vor allem aber aus Unternehmen, die die »echten Probleme« mitbringen. Das Narrativ der echten oder realweltlichen Probleme (*real-world problems*), die es im Rahmen der praxisorientierten Lehre zu bearbeiten gilt, findet eine Entsprechung in echten Daten (*real-world data*), die in unzähligen Kursauschreibungen, Webseiten und Werbematerialien von Studiengängen angeführt werden. Das evozierte Bild referiert einerseits auf einen Typus von Daten, die unabhängig von ihrer Konstruktion durch soziotechnische Systeme existieren und die es entsprechend bloss zu erheben und zu analysieren gilt, um unmittelbare Erkenntnisse für die eigene Forschung, das wissenschaftliche Feld oder »die Gesellschaft« zu ziehen. Die Metapher der »Roh-Daten« liegt hier nahe (Gitelman 2013) und markiert ein essentialistisches Datenverständnis, wie es insbesondere den politischen Diskurs

kennzeichnet.⁹ Zum anderen impliziert die Rahmung, die sonstige wissenschaftliche Arbeit mit Daten sei nicht echt, d. h. konstruiert, praxisfern oder sonst wie weniger bedeutend (Dreyfus & Dreyfus 2005). Die Betonung des angewandten, realweltlichen Bezugs von akademischer Forschung und Lehre ist eng gekoppelt an historisch gewachsene Vorstellungen und Ideale der Engineering-Ausbildung (Malazita 2019), innerhalb derer sich die meisten Studiengänge in Datenwissenschaften positionieren.

Die Grenzziehung zwischen konstruierten wissenschaftlichen und echten realweltlichen Daten und Problemen wurde in den letzten drei Jahrzehnten zu einem zentralen Organisationsprinzip der Forschungsförderung (Ribes et al. 2019). In der Schweiz folgen die Förderprogramme zu »Big Data« (SNF 2015) und »Data Science« (SDSC 2017) dieser Logik. Die Projektausschreibung des SDSC (2017) macht dies explizit, indem der Ausschreibungstext die zwei Module »Data science meets domain science« und »Data science methods for the real-world [sic!]« (ebd.: 5f.) einander gegenüberstellt: Bei datenwissenschaftlichen Anwendungen und Methoden in *domain science* handelt es sich um spezifische Probleme, die lediglich innerhalb des wissenschaftlichen Feldes relevant sind. Demgegenüber weisen Technologien und Methoden der Datenwissenschaften einen Bezug zur »real world« auf und ermöglichen Anwendungen über die Wissenschaft hinaus, sind also von universeller Relevanz. Die darin artikulierten Betonung des gesellschaftlichen »Impacts« von »Data Science« evoziert wiederum eine Trennung in wissenschaftliche und nichtwissenschaftliche Anwendungsgebiete.

Eine zweite Ausprägung von Anwendungsorientierung, die der ersten gegenübergestellt wird, zeigt sich in einer divergierenden Verortung der relevanten Probleme und deren Verknüpfung mit bestimmten methodischen Vorgehensweisen innerhalb der Datenwissenschaften: Während an den ETH und Universitäten die »echten Probleme« durch Kooperationen mit anderen Disziplinen oder externen Partner*innen adressiert werden, sind es an den Fachhochschulen die Studierenden, die »echte Probleme« aus ihrer beruflichen Praxis mitbringen. Ein Professor einer Fachhochschule verknüpft die unmittelbare Anwendungsorientierung im beruflichen Umfeld mit den notwendigen Kompetenzen und Vorgehensweisen von Data Scientists wie folgt:

»Ein Grundstock an klassischer sowie deskriptiver und schliessender Statistik ist notwendig, damit sie fundiert arbeiten können. Algorithmen- oder Methodenentwicklung als solche sind aber nicht Ziel des Studiengangs. Es geht vor allem um die Optimierung von Arbeits- und Betriebsabläufen in ihren Unternehmen, beispielsweise anhand von Logfiles die Performance von Systemen zu verbessern. Es geht nicht darum, Hypothesentests durchzuführen oder Forschungsfragen zu beantworten, sondern darum, operative Exzellenz zu schaffen« (Prof_FH_C: 9:29).

Die befragte Person markiert hier Differenzen zwischen verschiedenen epistemischen Praktiken innerhalb der Datenwissenschaften. Dabei assoziiert sie bestimmte

9 Die Konzeption als »real-world data« unterschlägt zudem, dass Forschung – sowohl in wissenschaftlichen als auch anderen Kontexten – Probleme, Modelle oder Daten immer zuerst bestimmen, auswählen und bearbeiten, d. h. als solche *konstruieren* muss: Nicht bloss die Daten sind nicht einfach gegeben, sondern auch Fragestellungen und zu lösende Probleme müssen als solche anerkannt und bestimmt werden (Mützel et al. 2018).

Vorgehensweisen (wie »Algorithmen- oder Methodenentwicklung«, »Hypothesentests durchführen oder Forschungsfragen zu beantworten«) mit den beiden ETH, während sie für die Fachhochschulen eine angewandte Form der Datenwissenschaften (wie beispielsweise die »Optimierung von Arbeits- und Betriebsabläufen«) reklamiert, bei der »operative Exzellenz« geschaffen werden soll. In den Zuschreibungen unterschiedlicher Problembearbeitung im Rahmen der Anwendungsorientierung manifestieren sich divergierende Konzeptionen der Datenwissenschaften, die wiederum unterschiedliche Anknüpfungspunkte für feldexterne Akteur*innen ermöglichen.

Die Anwendungsorientierung infolge der postulierten Universalität datenwissenschaftlicher Expertise stellt einen zentralen *Modus Operandi* dar. Daraus resultiert eine gewisse Notwendigkeit für alle Akteur*innen im Feld, die eigenen Lehr- und vor allem Forschungsaktivitäten als anwendungsorientiert zu markieren: Alle Akteur*innen kollaborieren mit internen und externen Partner*innen und verstehen sich als angewandt und praxisorientiert. Trotzdem fördert die Analyse des empirischen Materials unterschiedliche Spielarten von Anwendungsorientierung zutage: Zum einen zeigt sich eine Varietät von Anwendungsorientierung, die sich am realweltlichen Bezug von Forschungsproblemen und Daten – gegenüber »rein wissenschaftlich« induzierten – unabhängig von der feldspezifischen Verortung (Wissenschaft, Industrie, Verwaltung etc.) orientiert. Zum anderen manifestiert sich eine Spielart von Anwendungsorientierung als »operative Exzellenz«, die Praxisbezüge ausschliesslich in ausserwissenschaftlichen Referenzen verortet. Auch wenn beide Ausprägungen unter dem gemeinsamen Label »Data Science« koexistieren können, zeigen sich unterschiedliche institutionelle Verortungen im Feld (Erstere vor allem an den ETH, Letztere an den Fachhochschulen). Sie äussern sich in Zuschreibungen divergierender methodischer Vorgehensweisen und epistemischer Praktiken, die somit die jeweiligen Positionen im akademischen Feld charakterisieren.

7.7 Diskussion

Die Analyse hat die Bedeutung von zwei zentralen Praktiken, diskursiver Begriffsarbeit einerseits und *boundary work* andererseits, in der Konstruktion der Datenwissenschaften sichtbar gemacht. Neben disziplinären Wissensbeständen, Daten, Methoden und Instrumenten sind auch Definitionen, Kategorien und Narrative fundierend für die Herausbildung eines neuen Wissensfeldes. So repräsentiert das akademische Feld die Datenwissenschaften als ein Wissensfeld, das durch technologische und epistemologische Verschiebungen an der Schnittstelle unterschiedlicher Disziplinen entstanden ist und gleichzeitig verschiedene Anwendungsbereiche transformiert. Dem Begriff des Paradigmenwechsels und der Universalität der Datenwissenschaften stehen Differenzierungen in spezifische Subgruppen und Desiderata nach klaren Grenzziehungen gegenüber anderen Disziplinen und Feldern entgegen. Betonen Erstere die Offenheit und Transversalität des Wissensgebiets, so zielen Letztere auf die strukturelle Einpassung und inhaltliche Engführung in ein disziplinäres Gefüge.

In der Suche nach den adäquaten Kategorien zur Charakterisierung der identifizierten epistemologischen Verschiebungen hat sich im Untersuchungsfeld die Bezeichnung »Data Science« gegenüber anderen Begrifflichkeiten (wie »Big Data« oder »Machine Learning«) etabliert und durchgesetzt. Werden Letztere als zu weit bzw. zu

eng taxiert, stellt »Data Science« eine inklusive Kategorie dar, die es Akteur*innen erlaubt, multiple Rollen in unterschiedlichen Feldern und Organisationen wahrzunehmen und somit existierende Grenzen zu überschreiten. Solche Suchprozesse haben indes auch sinnstiftenden Charakter, indem sie zur Kohäsion und zum Zusammengehörigkeitsgefühl der unterschiedlichen Akteur*innen beitragen und womöglich gar die Herausbildung einer distinkten feldübergreifenden Identität befördern.

Neben der konstitutiven Rolle von Begriffen und Kategorien bilden synchrone Praktiken von *boundary work*, d. h. Grenzziehungen und -überschreitungen, den zweiten Praxismodus, der in die Konstruktionsweise der Datenwissenschaften im akademischen Feld eingeschrieben ist: Durch die Konzeption der Datenwissenschaften als ein interdisziplinär zusammengesetztes Wissensgebiet erheben diverse Disziplinen Ansprüche auf die epistemische Autorität für das Feld der wissenschaftlichen Datenbearbeitung, was sich im empirischen Material in intensiven Konfliktlinien und Grenzziehungen zwischen einzelnen Disziplinen artikuliert.

Umgekehrt werden durch Interdisziplinarität und die Logik der Domänen aber auch etablierte Grenzen überschritten: So produziert die permanente Formulierung von Interdisziplinarität, Anwendungsorientierung oder »echten Problemen« einen diskursiven Raum zwischen disziplinären und sozialen Feldern, der vielfältige Schnittstellen und Anknüpfungspunkte schafft. Indem die Datenwissenschaften als universal und domänenunabhängig konstruiert werden, ermöglicht dies ansatzweise die Limitationen, die sich durch Disziplinierung sowie innerwissenschaftliche Konfliktlinien ergeben, zu überwinden und zu neutralisieren. Dies gelingt den Akteur*innen vor allem durch diskursive Strategien der Grenzüberschreitung, indem sie die gemeinsamen Interessen am Aufbau des neuen Wissensfeldes betonen und gleichzeitig interne Konflikte latent halten.

Entgegen den Bestrebungen des akademischen Feldes um Autonomisierung und Immunisierung, die sich auf Prinzipien wie Domänenunabhängigkeit, Reproduzierbarkeit oder der universellen Anwendbarkeit von Methoden und Technologien abstützen, ist das Soziale stets in Datenwissenschaften eingeschrieben: Es äussert sich in sozialen Kategorien (wie Paradigmenwechsel, Domänen, *real-worldness*, Anwendungsorientierung etc.), disziplinären Konfliktlinien sowie Grenzüberschreitungen, die für die Sinnhaftigkeit des Phänomens Datenwissenschaften ebenso fundierend sind wie die Praxis mit Daten, Methoden und Algorithmen selbst. Die Synchronizität und die Kontingenzen von Suchprozessen einerseits, von Abgrenzung und Zusammenarbeit andererseits generieren in diesem Sinne äusserst widersprüchliche Dynamiken, die neben wissenschaftlich-technischen Artefakten (Projekte, Publikationen, Patente etc.) auch auf der sozialen und symbolischen Ebene produktive Potenziale freisetzen können. Die Datenwissenschaften stellen insofern einen Fall dar, bei dem es den beteiligten Disziplinen und Akteur*innen gelang, einen solchen Raum der Möglichkeiten zwischen etablierten disziplinären Feldern zu schaffen sowie – zumindest bis zum jetzigen Zeitpunkt – offen und durchlässig zu halten.

Kapitel 8 - Die Verhandlung der Datenwissenschaften in Universitäten und Hochschulen

8.1 Einleitung

Die Datenwissenschaften konfrontieren Universitäten und Hochschulen mit der Frage, wie sie ein Wissensfeld, das als interdisziplinär, transversal oder universal beschrieben wird, in ihre bestehenden organisationalen Strukturen, die stark disziplinär verfasst sind, einpassen und implementieren können (Wing et al. 2018: 5f.). Somit bilden die Datenwissenschaften parallel zur Frage ihrer epistemischen und disziplinären Verfasstheit ein Verhandlungsobjekt innerhalb von Universitäten und Hochschulen. Als Organisationen stellen diese soziale Felder dar, die sich aus multiplen organisationalen Einheiten (Verwaltung, Fakultäten, Fachbereiche etc.) sowie disziplinär strukturierten Wissensfeldern (meist in Instituten organisiert) zusammensetzen. Als kollektive Akteur*innen sind ihre Positionierungen und Investitionen, wie strategische Forschungsschwerpunkte oder der Aufbau neuer organisationaler Einheiten, insofern als Kompromisse der Interessen und Strategien der verschiedenen Teilfelder zu verstehen. Neue, interdisziplinäre Curricula wie in den Datenwissenschaften stellen somit »Kompromissprodukte« verschiedener Disziplinen dar und bilden Verhandlungen und Positionierungen innerhalb von Universitäten oder Hochschulen als organisationalen Feldern ab. Die resultierenden Curricula repräsentieren insofern auch die Vielstimmigkeit der involvierten disziplinären Wissensfelder. Ich untersuche deshalb in diesem Kapitel, inwiefern wissenschaftliche und ausserwissenschaftliche Faktoren die Planung, Ausgestaltung und Implementierung der Curricula in Universitäten und Hochschulen beeinflussen. Und welche Akteur*innen und Disziplinen sind an der Ausarbeitung beteiligt?

Die Analyse von Curricula und Interview verdeutlicht, dass es nicht ausschliesslich wissenschaftliche, d. h. epistemologische und disziplinär-fachliche Kriterien sind, die den Aufbau der Curricula charakterisieren. Ebenso zentral sind hochschul- und wissenschaftspolitische, ökonomische sowie organisationale Einflussfaktoren, die die Einführung der neuen Studienprogramme rahmen: Erstens ist, wie bereits gezeigt, das entstehende Feld der Datenwissenschaften massgeblich durch die hohe Nachfrage nach Data Scientists vonseiten der »Industrie«, d. h. verschiedenen ökonomischen Feldern geprägt. So legitimieren und forcieren insbesondere »Nützlichkeitsabwägungen« die Schaffung neuer Studienprogramme in den Datenwissenschaften. Zweitens sind

es gesetzliche Regelungen und politische Vorgaben bezüglich neuer Studienprogramme, die deren Aufbau bzw. Implementierung vor allem in der temporalen Dimension beeinflussen, d. h. verlangsamen oder beschleunigen können. Drittens sind es die zur Verfügung stehenden Ressourcen, d. h. ökonomisches Kapital, die die Möglichkeitsräume von Hochschulen und Universitäten als Akteur*innen im akademischen Feld erweitern oder einschränken können. Schliesslich strukturieren auch die Relationen zu anderen Akteur*innen die Strategien und Investitionen von Universitäten und Hochschulen: Ausgehend von der Positionierung aufgrund der Struktur und des Volumens der verfügbaren sozialen und symbolischen Kapitalia (Prestige) dienen Bezugnahmen zu anderen Akteur*innen als Orientierungsgrössen bzw. Abgrenzungspunkte und somit zur Profilbildung des eigenen Tätigkeitsbereichs.

Die Wechselwirkungen dieser organisationalen, feldinternen und -externen Einflussfaktoren strukturieren und prägen demnach die Möglichkeiten und Handlungsspielräume der Universitäten und Hochschulen beim Aufbau neuer Studiengänge. Die Analyse der empirisch beobachtbaren Implementationsprozesse macht deutlich, dass diese in sehr unterschiedlichen Verständnissen der Datenwissenschaften und damit zusammenhängender Expertisen resultieren, die in divergierenden Lehrinhalten einerseits und symbolischen Grenzziehungen gegenüber anderen Akteur*innen im Feld andererseits zutage treten. Dabei werden die epistemischen und disziplinären Distinktionen durch solche im organisationalen Feld der Hochschulen gekreuzt: Die sozialen Konstruktionen der Datenwissenschaften verschränken sich mit jenen wissenschaftlichen Definitionen, Kategorisierungen sowie disziplinären Praktiken von Grenzarbeit und prägen dadurch ein akademisches Feld, das durch die Gleichzeitigkeit von Kooperations- und Konkurrenzverhältnissen gekennzeichnet ist.

Das Kapitel ist wie folgt strukturiert: Zunächst untersuche ich, wie Anforderungen und Ansprüche aus dem Feld der Ökonomie die Schaffung neuer Studienprogramme in den Datenwissenschaften begründen und forcieren (Kap. 8.2). Weiter widme ich mich dem Zusammenspiel von politischen Bewilligungsverfahren und organisationalen Abläufen, die den Aufbau neuer Studienprogramme insbesondere in zeitlicher Hinsicht beeinflussen (Kap. 8.3). Das Wechselspiel von Hochschulpolitik und organisationalen Absichten und Bemühungen äussert sich weiter in spezifischen Finanzierungsmodi der Studiengänge, die die Möglichkeitsräume der organisationalen Akteur*innen im akademischen Feld in finanzieller Hinsicht rahmen (Kap. 8.4). Anschliessend untersuche ich, in welchen organisationalen Einheiten die untersuchten Studiengänge in den Hochschulen und Universitäten konkret verortet sind und welche Koordinationsprobleme damit einhergehen (Kap. 8.5). Nach der Analyse der ökonomischen, hochschulpolitischen und organisationalen Faktoren lege ich den Fokus auf Strategien, Positionierungen und Investitionen, die sich am Beispiel der Datenwissenschaften in synchronen Konkurrenz- und Kooperationsverhältnissen im akademischen Feld manifestieren (Kap. 8.6). In einer Diskussion fasse ich die zentralen Erkenntnisse des Kapitels zusammen und setze sie in Bezug zu den übergeordneten Fragestellungen der Arbeit (Kap. 8.7).

8.2 »Nützlichkeitsabwägungen«: Die Nachfrage auf dem Arbeitsmarkt

»Die Nachfrage ist riesig! Sie wissen nicht, wie sie unter den Angeboten aussuchen sollen. Im Moment wird nur nach ihnen gesucht. [In] der New York Times stand, dass momentan wegen mangelnden Angebots circa zwei Drittel der Stellen mit Informatikern gefüllt werden, die Datenanalyse-Flair haben. Sie arbeiten bei Kreditkartenfirmen, Versicherungen, überall. Einzelne gründen auch Firmen oder Startups und machen Data Science. Auch quer durch alle wirtschaftlichen Branchen« (Prof_UH_A: 6:27).

Die Feststellung einer stetig steigenden Nachfrage nach Data Scientists in verschiedenen ökonomischen Feldern bildet einen zentralen Topos der Diskussion über die Datenwissenschaften im akademischen Feld. Die Bedürfnisse von Unternehmen nach kompetenten Absolvent*innen können als wesentliche Treiber neuer Studienprogramme in den Datenwissenschaften und verwandten Studienbereichen gelten. So bestätigen denn auch alle Interviewten übereinstimmend, dass die Nachfrage vonseiten der »Industrie« die Errichtung von Studiengängen massgeblich beeinflusst hat:

»Allgemein entstand das Programm, weil Big Data und Data Science immer wichtiger wurden, gerade auch in den Unternehmen. Die Industrie braucht Data Scientists, die stark sind in Informatik, Mathematik und Statistik sowie Elektrotechnik und Information Retrieval. [...] [E]s steht in den Statuten der ETH, dass sie einen Austausch mit der Industrie und der Gesellschaft pflegt, dass sie Studierende ausbildet, die später nützlich für die Industrie und die Gesellschaft sind« (Prof_ETH_B: 2:16).

Solche »Nützlichkeitsabwägungen«, die sich in der zitierten Antwort eines ETH-Professors auf die Frage äussern, was die Motivation für ein eigenständiges Studienprogramm in Datenwissenschaften war, stehen exemplarisch für eidgenössische Bildungspolitiken, insbesondere im technischen und naturwissenschaftlichen Bereich (Honegger et al. 2007: 11). Hochschultypen- und disziplinenübergreifend wird wiederholt die Annahme geäussert, dass in den Kompetenzen, die durch Studiengänge in »Data Science« vermittelt würden, »die Zukunft liege«. Solche Diagnosen sind komplementär zu politischen Begründungen für die Bedeutung datenwissenschaftlicher Studiengänge als auch zu empirischen Beobachtungen im Arbeitsmarkt selbst. Sie verweisen auf umfassendere Transformationsprozesse, in denen alle gesellschaftlichen Bereiche datenwissenschaftliche Expertise in ihre eigenen Praktiken, Organisationen und Systeme integrieren.

Im Falle der Universitäten zeigt sich, dass Diagnosen einer hohen Nachfrage an Forderungen nach stärkerer Anwendungsorientierung oder Industriepartnerschaften gekoppelt werden. An der Universität Genf hat die wirtschaftswissenschaftliche Fakultät eine Initiative lanciert, um die diagnostizierte »Lücke« zwischen dem wachsenden Bedarf und fehlendem Studienangebot mittels neuer Studienprogramme zu »überbrücken«:

»The Geneva School of Economics and Management (GSEM) at the University of Geneva has launched a strategic initiative to strengthen its ties to the city, and align its curriculum with the longterm [sic] demands of corporations and institutions across industries. The *Bridge the Gap* programme was born from this initiative and will offer two profes-

sional masters, including a new specialisation in Business Analytics starting in fall 2017« (GSEM 2016: 2; Hervorhebung im Original).

Obwohl die Orientierung am Arbeitsmarkt historisch betrachtet eine zentrale gesellschaftliche Funktion der Hochschulbildung darstellt (Clark 1983: 136ff.), ist die enge Kopplung von universitären Studiengängen und Berufslaufbahnen wie in diesem Fall beachtenswert: Es geht demnach weniger um die Bildung professioneller Praktiker*innen mit universell anwendbarer Expertise, sondern um die unmittelbare Ausrichtung von Curricula an »den langfristigen Bedürfnissen von Unternehmen und Institutionen« (GSEM 2016: 2; meine Übersetzung).

Inhaltlich strebt die GSEM durch ihren neuen Master in Business Analytics eine Kombination von Statistik, Management sowie »Data Science« (bzw. Analytics in ökonomischen Teilgebieten) an, um damit die »Lücke« zwischen »akademischer Bildung« und den Bedürfnissen von Unternehmen zu schliessen. Die drohende »Lücke« wird durch den Bezug auf eine Studie des McKinsey Global Institute (Manyika et al. 2011) wie folgt quantifiziert:

»Since 2011, the McKinsey Global Institute predicts a lack of specialists: it estimates that by 2018, the United States may lack between 140,000 and 190,000 employees with deep analytical skills, as well as 1.5 million business managers and analysts who know how to use analytics to make more effective decisions« (GSEM 2016: 3).

Die Studiengangbeschreibung signalisiert potenziellen Studierenden in dem Szenario eine verheissungsvolle Zukunft, weil in ihrem Studienbereich viele Stellen nicht besetzt werden könnten. Die McKinsey-Studie wird von verschiedenen Hochschulen wiederkehrend als Beleg für die hohe Nachfrage zitiert. Obwohl mittlerweile aktualisierte Prognosen vorliegen, die wiederum neue Zukunftsversprechungen für Data Scientists entwerfen, entfaltete die Studie innerhalb des akademischen Feldes eine richtungsweisende Bedeutung und strukturiert die Einführung der neuen Studienprogramme.¹

Neben der hohen Nachfrage wird die Einführung der neuen Studienangebote zudem als Anpassung an veränderte Bedürfnisse des Arbeitsmarktes gerechtfertigt. Die diskutierte Distinktionslinie zwischen Grundlagen- und Anwendungsorientierung verschiebt sich durch die epistemologische Transformation, die mit den Datenwissenschaften einhergeht, hin zu einer grundlegenden Neuordnung des Verhältnisses von Bildung, Wissenschaft und Ökonomie. Dieses artikuliert sich in der Figur der »scientist-entrepreneurs« (Powell & Sandholtz 2012; Berman 2012), die zwischen den involvierten Feldern changieren können. Die diagnostizierte Nachfrage in ökonomischen Feldern und die symbolische Konstruktion der Profession der Data Scientists sind wesentliche Treiber für die Konstruktion der Angebotsseite der Datenwissenschaften im akademischen Feld. Sie legitimieren und beschleunigen organisationale Abläufe und Praktiken zur Implementation eines Wissensfeldes, das noch vor wenigen Jahren ausserhalb wissenschaftlicher Spezialdiskurse kaum wahrgenommen wurde.

¹ Des Weiteren spielen Zukunftsentwürfe und Szenarien von Beratungsunternehmen und Think-Tanks eine wichtige Rolle bei der Implementierung von Studiengängen in Datenwissenschaften (Saner 2019).

8.3 Das Zusammenspiel politischer Bewilligungsverfahren und organisationaler Planungsprozesse

Der Aufbau neuer Studienprogramme stellt akademische Institutionen nicht nur vor inhaltliche, sondern auch vor diverse strategische und hochschulpolitische Herausforderungen (Lattuca & Stark 2009). Dabei stellt sich insbesondere die Frage, inwiefern das Wechselspiel politischer und organisationaler Faktoren die Ausgestaltung und die Geschwindigkeit bei der Implementierung neuer Studiengänge in den Datenwissenschaften beeinflusst. Trotz der verfassungsrechtlich verankerten Freiheit der wissenschaftlichen Lehre und Forschung sind öffentlich finanzierte Universitäten und Hochschulen der Schweiz gegenüber ihren politischen Trägern (Bund, Kantone, Hochschulräte etc.) rechenschaftspflichtig. Dies betrifft insbesondere finanzielle und organisationale Aspekte, aber auch den Aufbau von Studienprogrammen oder Professuren in neuen Wissensfeldern, die gegenüber politischen Akteur*innen legitimiert werden müssen. Im Folgenden untersuche ich anhand von drei Fallbeispielen, wie durch unterschiedliche Strategien die Möglichkeitsräume zwischen politischen Vorgaben und organisationaler Autonomie genutzt werden.

Das erste Beispiel betrifft den bundesstaatlich geförderten ETH-Bereich: Dieser lancierte im Jahr 2016 die bereits erwähnte »Initiative for Data Science in Switzerland«, die die Gründung des SDSC, die Etablierung zusätzlicher Professuren sowie neue Masterstudiengänge an den ETH Lausanne und Zürich umfasst. Die Initiative ist Teil der ETH-Strategie 2017–2020 und geht mit zusätzlichen finanziellen Ressourcen in der Mehrjahresplanung einher (ETH-Rat 2014). Sie ist eingebettet in die Strategie »Digitale Schweiz« des Bundesrates, mit welcher die globale Wettbewerbsfähigkeit von Wirtschaft und Wissenschaft der Schweiz erhalten werden soll. Die Schaffung zusätzlicher Forschungs- und Lehrkapazitäten in den Datenwissenschaften wird folglich explizit mit den wirtschafts- und wissenschaftspolitischen Interessen der Schweiz begründet. Entsprechend sei, wie es ein Professor an der ETH ausdrückt, eine »[strategische] Positionierung in einem so wichtigen Bereich wie den Datenwissenschaften natürlich zentral für die Schweiz« (Prof_ETH_A: 4:7), sowohl was den Forschungs- und Wissenschaftsstandort als auch die Ausbildung von Arbeitskräften betrifft. Der Planungsprozess sei »nicht top-down« zu verstehen, sondern die verantwortlichen Departemente von ETHZ und EPFL brachten im Sinne eines »wechselseitigen Dialogs« mit dem ETH-Rat und dem Bund ihre Schwerpunkte mit ein und erarbeiteten das Konzept für die Initiative gemeinsam (Prof_ETH_A: 4:4).²

Die zwei ETH-Studiengänge werden als Resultate eines mehrjährigen Abstimmungs- und Koordinationsprozesses präsentiert (ETH-Rat 2014, 2016a), der eine Vielzahl von Akteur*innen im akademischen und politischen Feld involvierte. In den Schilderungen der befragten ETH-Professor*innen kommen weder disziplinäre Grenzziehungen zur Sprache noch werden Rivalitäten oder Konflikte zwischen den verschiedenen Akteur*innen genannt; vielmehr werden sie vor dem Hintergrund eines globalen »Talent-Turns« (Aratnam 2016), eines eklatanten »Fachkräftemangels« (Herberg 2018a) sowie intensiver Konkurrenzbeziehungen zurückgestellt. Insofern ordnen sich die neuen ETH-Studiengänge in den Dienst übergeordneter nationalstaatlicher

2 Trotz der Autonomie der beiden Hochschulen liegt eine hohe Überschneidung sowohl auf der strukturellen (Aufbau Curricula) als auch auf der inhaltlichen Ebene (Kursangebote) vor (vgl. Kap. 9.3).

Interessen ein (Jessop 2008). Das erste Fallbeispiel verweist demnach auf eine Situation, in der organisationale und politische Interessen ineinander übergehen und sich gegenseitig befördern.

Ein zweites Fallbeispiel ist eine kantonale Universität, bei der ebenfalls parallel zu einem neuen Studienprogramm zusätzliche Professuren aufgebaut werden. Während die Einrichtung einer neuen Vertiefung in »Data Science« unter dem Radar der politischen Behörden erfolgt,³ wird die Besetzung neuer Professuren nicht nur mit der universitären Strategie, sondern auch mit den politischen Verhältnissen im Trägerkanton in Verbindung gebracht:

»Wir haben einige spezifische Vorhaben in Absicht, und einige Dinge hängen auch davon ab, wie die Stimmbevölkerung des [Kantons] entscheiden wird. Es wird vier neue Professuren in IT geben, wovon mindestens eine in Data Science sein soll. Es gibt einen gewissen Druck der Wirtschaftsorganisationen, die eher in die Vergangenheit als in die Zukunft schauen und sagen: ›Wir brauchen mehr Programmierer und Software-Entwickler und so weiter.‹ Es gibt dort noch kaum Verständnis für Data Science oder Artificial Intelligence. Das heisst bei der Besetzung dieser Stellen gibt es ein gewisses Risiko, dass eher den aktuellen Bedürfnissen der lokalen Wirtschaftsorganisationen gefolgt wird, als dass der Blick in die Zukunft gerichtet wird.« (Prof_UH_B: 3:23).

Der Befragte deutet hier an, dass die universitäre Autonomie in der Besetzung neuer Professuren durch politische Mehrheiten und kurzfristige ökonomische Bedürfnisse limitiert wird. Die disziplinären Konfliktlinien, die bei der Besetzung von Professuren artikuliert werden, werden demnach zusätzlich durch feldübergreifende Grenzziehungen zwischen dem akademischen und dem industriellen Feld überlagert. Ferner schreibt er den »lokalen Wirtschaftsorganisationen« zu, dass sie die gegenwärtig relevanten Kategorien noch nicht erfasst hätten, sondern »eher in die Vergangenheit als in die Zukunft schauen« würden. Der »Blick in die Zukunft«, das Erkennen der zukünftig wichtigen Kategorien, strukturiert die Erwartungen von Akteur*innen in verschiedenen gesellschaftlichen Kontexten hinsichtlich der Implementierung neuer Kategorien sowohl im wissenschaftlichen als auch im ökonomischen Feld (Saner 2019). Das zweite Fallbeispiel beschreibt eine Situation, in der sich die Interessen von Universität auf der einen und politische sowie ökonomische Bedingungen auf der anderen Seite nur bedingt kombinieren lassen. Vielmehr sind die universitären Bedürfnisse mehr oder minder Letzteren unterworfen.

Ein drittes Beispiel betrifft ebenfalls eine kantonale Universität, die sich mit der Herausforderung hochschulpolitischer Bewilligungs- und organisationaler Planungsverfahren konfrontiert sah und sich in der Folge für eine Umgehungsstrategie entschied:

»Wir sind zu 95 % ein Studiengang Master of Science in Data Science, ausser im Namen. Das war eine technische Frage, weil Vertiefungen viel schneller implementiert werden können. Neue Studiengänge dauern 2–3 Jahre und landen in der Gesetzessammlung. Eine Vertiefung war einfach zu realisieren, und wir waren damals sowieso daran, die

3 »Das [Programm] ist viel weniger politisch, das interessiert den Kanton sowie die Industrie- und Handelskammer gar nicht.« (Prof_UH_B: 3:24).

Studienordnung der gesamten Fakultät zu überarbeiten und stark modular aufzubauen. Die Struktur wurde bewusst gewählt, um eben diese Durchlässigkeit zu ermöglichen. Früher oder später wird es eventuell ein Programm mit solchem Namen geben« (Prof_UH_A: 6:19).

Die Visibilität eines Studiengangs stellt vor allem in der Frühphase (2013–2015) der Datenwissenschaften an Schweizer Universitäten und Hochschulen eine bedeutende Grösse dar. In der Konkurrenzsituation zwischen Universitäten und Hochschulen war ein Wettlauf um die Frage entbrannt, wer denn »die Ersten« sind, die einen solchen Studiengang anbieten, was denn auch mehrere Befragte für sich in Anspruch nehmen. Der Aufbau neuer Studiengänge dauert, wie übereinstimmend betont wird, aufgrund hochschulpolitischer Bewilligungsverfahren zwei bis drei Jahre. Da Vertiefungen und Spezialisierungen sehr viel schneller etabliert werden können, wählten mehrere Universitäten und Hochschulen diese Strategie, um rasch ein Programm zu implementieren und mit dem Label »Data Science« zu verknüpfen.

Das dritte Beispiel indiziert, inwiefern politische Vorgaben und Rechtsetzungen einer »schnelle[n]« Implementierung« der Datenwissenschaften als neues Studienfach entgegenstehen können. Anders als in den ersten zwei Fällen werden die Auswirkungen der politischen Bestimmungen als limitierend betrachtet, weshalb direkt eine Umgehungs- bzw. Anpassungsstrategie angewandt wird, um sich die organisationalen Handlungsspielräume bezüglich der politischen Bestimmungen zunutze zu machen.

Die drei Fallbeispiele illustrieren, wie politische Vorgaben die universitären Strategien und Planungsprozesse bei der Ausgestaltung von Studienprogrammen beeinflussen. Obwohl die Befragten die Planungsprozesse mehrheitlich als kollaborativ oder »bottom-up« skizzieren, werden Gestaltungsspielräume primär in inhaltlicher Hinsicht verortet. Demgegenüber bringen vor allem die hochschulpolitischen Rahmenbedingungen formelle und strukturelle Limitierungen für die Hochschulen und Universitäten mit sich. Die Analyse der Aufbau- und Implementierungsprozesse der untersuchten Studiengänge verdeutlicht insofern die Bedeutung des für die schweizerische Bildungs- und Wissenschaftspolitik typischen Föderalismus (Jost 2007), in dem verantwortlichen politischen Entitäten signifikante Relevanz zukommt. Demgegenüber bleibt die Autonomie der Universitäten und Hochschulen stets eine relative: Die Möglichkeitsräume, die sich durch feldübergreifende Kollaborationen in den Datenwissenschaften oder durch die hohe Nachfrage im Feld der Ökonomie ergeben, bleiben beschränkt. Diese eröffnen sich insbesondere dann, wenn organisationale und hochschulpolitische Interessenlagen sich gegenseitig befördern.

8.4 Die Mechanik hinter den Kulissen: Zur Finanzierung der Studiengänge

Neben politischen Bewilligungsverfahren und organisationalen Planungsprozessen stellen die finanziellen Ressourcen, auf die Universitäten und Hochschulen zurückgreifen können, ebenfalls einen strukturellen Einflussfaktor dar. Grössere, finanziell gut ausgestattete Universitäten und Hochschulen können rascher auf Entwicklungen im akademischen Feld reagieren und neue Studienangebote schaffen (Brint et al.

2009). Trotz der konstatierten hohen Nachfrage sowohl vonseiten der Unternehmen als auch der Studierenden selbst stehen allerdings nicht an allen Universitäten und Hochschulen genügend finanzielle Ressourcen zur Verfügung, wie ein Befragter ausführt: »[Es] ist sehr *small scale*, wir hatten 40 Leute. Nächstes Jahr erwarten wir 60 Leute. Für mehr brauchen wir aber zusätzliche Ressourcen« (Prof_UH_B: 3:15). Insbesondere an Institutionen mit limitierten finanziellen Mitteln werden diese nicht unmittelbar zur Verfügung gestellt, sondern unterliegen organisationalen Aushandlungsprozessen. An kleineren Universitäten spielen zudem Überlegungen, wie angesichts limitierter finanzieller Möglichkeiten ein attraktives Studienangebot offeriert werden kann, eine wichtige Rolle:

»[Der gemeinsame Master] wurde gemacht, um Ressourcen zwischen relativ kleinen Universitäten bzw. Informatik-Departementen zusammenzulegen. Dadurch [gibt es] mehr Angebote für die Studierenden, sonst könnten wir nicht alles anbieten« (Prof_UH_D: 1:4).

Die befragte Person schildert im Zitat, dass trotz der rechtlichen und politischen Herausforderungen ein gemeinsames Studienangebot über die Organisationsgrenzen hinaus mit anderen Universitäten aufgebaut wurde. Die daraus resultierenden Koordinationsprobleme werden deshalb zugunsten der Sichtbarkeit und Profilbildung im akademischen Feld in Kauf genommen. Demgegenüber kennen grössere, gut ausgestattete Institutionen dieses Problem weniger, wie ein anderer Interviewter ausführt:

»Wir haben keinen [Numerus Clausus]; wenn es zu viele Leute gibt, dann melden wir uns bei der Uni-Leitung und melden zusätzlichen Personalbedarf an. Aber wir haben jetzt dieses Semester über 100 Leute im Nebenfach Data Science. Ich kann nicht sagen, welchen Anteil dies an der Gesamtzahl aller MA-Studierenden ausmacht. [Aber] das Interesse ist signifikant« (Prof_UH_A: 6:52).

Die Analyse deutet an, dass die Hochschul- und Studiengangfinanzierung einen zentralen Faktor, eine Art »Mechanik hinter den Kulissen« (Seefranz & Saner 2012: 82ff.) für den Aufbau und die Implementierung neuer Programme darstellt. Sie prägt die Möglichkeitsräume, die den Akteur*innen im akademischen Feld zur Verfügung stehen, und kann insofern die Notwendigkeit von interorganisationalen Kooperationen begründen oder solche ausschliessen. Regelstudiengänge auf Bachelor- und Masterstufe werden von den Trägern der jeweiligen Universitäten und Hochschulen sowie durch den Bund finanziert. Die Finanzierung der Lehre bemisst sich durch interkantonale Konkordate sowie das Hochschulförderungs- und Koordinationsgesetz des Bundes primär an der Anzahl Studierender.⁴ Da die Kosten pro studierende Person nicht an allen Universitäten und Hochschulen identisch sind, können Beiträge anderer Kantone höher sein als diejenigen des Trägerkantons. Dies stellt einen finanziellen Anreiz dar, ein neues Studienprogramm in Eigenverantwortung aufzubauen oder nicht (Rostetter 2020).

4 Vgl. SBFI (2020): Finanzierung der kantonalen Hochschulen. Online: <https://www.sbf.admin.ch/sbf/de/home/hs/hochschulen/finanzierung-kantonale-hochschulen.html> (Zugriff: 03.02.2022).

Schliesslich implizieren auch die Modalitäten der Hochschulfinanzierung aufgrund der Interdisziplinarität der Wissensformation eine Herausforderung. Die Frage nach der Verortung des neuen Studienfeldes wird in Bezug auf die zur Verfügung gestellten Ressourcen nach Ansicht eines Befragten zu Problemen mit der gegenwärtigen Hochschulfinanzierung führen:

»Wir würden gerne einen Master-Studiengang in Data Science anbieten, aber ob das ein so interdisziplinärer sein kann, ist letztlich eine politische Frage. [...] Wenn man sich die Hochschulfinanzierung anschaut, dann werden Programme aus Töpfen finanziert, die quasi [...] auf einer Landkarte der Disziplinen beruhen, die 50 Jahre alt ist. Und dann heisst es: »Ah, das ist ein Ingenieur, dann braucht der Student so und so viel Geld, und das ist ein Wirtschaftsingenieur, der braucht nur ein Drittel davon. Und ein Physiker braucht dreimal so viel, weil der braucht noch seine ganzen Labore und so.« Das heisst, es ist 'ne Finanzierungsfrage, wo hängst du einen solchen Studiengang auf. Und wenn du sagst, du machst einen interdisziplinären [Studiengang], dann gucken die mich an und fragen: »Ja, wo jetzt?« [...] Ich bin mal gespannt, ob die Strukturen, die wir typischerweise so an Hochschulen haben, mithalten können. Weil es dreht sich schon schnell, das Rad« (Prof_FH_B: 7:19).

Der Befragte verknüpft im Zitat mit der Interdisziplinarität der Datenwissenschaften und den politischen Bedingungen der Hochschulfinanzierung zwei Aspekte, die nur selten gemeinsam thematisiert werden. Das Beispiel offenbart somit eine potenzielle Inkompatibilität der Datenwissenschaften – und von anderen interdisziplinären Wissensfeldern – mit den gegenwärtigen organisationalen Hochschulstrukturen, vor allem was die Finanzierung von Studiengängen betrifft.⁵

Die finanziellen Ressourcen der Universitäten und Hochschulen sowie die Modalitäten der Hochschulfinanzierung können insofern die Spielräume für curriculare Innovationen erheblich erweitern oder einschränken. Die Verknüpfung von Finanzierungsmechanismen mit der Einteilung neuer Studiengänge in disziplinär strukturierte Wissensfelder stellt gewissermassen eine hochschulpolitische Hinterbühne dar, die die organisationalen Verhandlungen neuer Studiengänge nachhaltig prägt. Zudem wurde deutlich, dass sehr unterschiedliche Kapitalverhältnisse im akademischen Feld existieren: Obwohl ökonomisches Kapital zwar nicht die primäre »Währung« darstellt, beeinflusst es dennoch als sekundäres Merkmal die Handlungsoptionen von Universitäten und Hochschulen. Die Verfügbarkeit ökonomischen Kapitals kann folglich die Notwendigkeit von interorganisationalen Kooperationen begründen oder solche ausschliessen.

5 Die artikulierte Skepsis gegenüber den existierenden Hochschulstrukturen, die mit der dynamischen Entwicklung in neuen Wissensfeldern (»es dreht sich schon schnell, das Rad«) nicht mehr »mithalten können«, findet ihre Entsprechung im hochschulpolitischen Diskurs, indem für das gesamte Bildungssystem eine permanente Adaptivität und Flexibilität von Strukturen und Inhalten eingefordert wird, um der digitalen Transformation gerecht zu werden (Economiesuisse & W.I.R.E. 2017: 79; SBFI 2017: 39).

8.5 Datenwissenschaften als Organisationsproblem – zur Verortung und Implementierung der Studiengänge in den Hochschulen

Neben ökonomischen Motiven, politischen Bewilligungsverfahren und Finanzierungsmodi rahmen auch organisationale, d. h. hochschulinterne Eigenheiten die Einführung neuer Studiengänge. Indem Datenwissenschaften als ein interdisziplinäres Wissensfeld beschrieben werden, stellt sich die grundlegende Frage, welchem bestehenden oder neu zu schaffenden Organisationsbereich (Fakultäten, Institute oder spezialisierte Zentren) die primäre institutionelle Zuständigkeit für die Durchführung und Implementierung eines solchen Studiengangs zukommt (Small 1999). Die Schwierigkeit der organisationalen Verortung interdisziplinärer Curricula markiert zentrale Koordinationsprobleme innerhalb von Universitäten als Organisationen (Meier 2009; Krücken & Meier 2006), deren operative Strukturen auf disziplinären Klassifikationen aufbauen. Ein Universitätsprofessor beschreibt dies so:

»Wir führten Gespräche innerhalb der Fakultät, dass es [Data Science] braucht. Wir führten auch Gespräche ausserhalb, aber es hat sich gezeigt, dass [es] zu komplex [ist]. [...] Wenn Sie an der Universität rasch etwas aufbauen möchten, dann sollten Sie möglichst Fakultätsgrenzen nicht überschreiten. Deshalb haben wir das erst einmal innerfakultär aufgezogen. Jetzt aber beginnen die Gespräche mit anderen Instituten und Fakultäten [...]« (Prof_UH_A: 6:5).

In diesem Beispiel wurde demnach die »rasche« Implementierung des neuen Studienangebots gegenüber einer fakultätsübergreifenden Kooperation priorisiert, obwohl Letztere als adäquates organisationales Format betrachtet wird. Gleichzeitig ist das Begehren nach einer »rasch[en]« Implementation der Datenwissenschaften als universitäres Studienfach vor dem Hintergrund der existierenden Konkurrenzsituation im untersuchten Feld zu verstehen. Daraus ergibt sich die Frage, in welchen organisationalen Leistungsbereichen die untersuchten Studiengänge jeweils verortet sind und ob bzw. welche Koordinationsprobleme aus den jeweiligen Arrangements resultieren.

An der ETH Zürich wird der Masterstudiengang als Kooperation der drei Departemente für Informatik, Mathematik sowie Informationstechnologie und Elektrotechnik angeboten, wobei das Departement für Informatik federführend ist. Der Masterstudiengang der EPFL hingegen wird von der »School for Computer and Communication Science« alleine getragen, die auch mit einer Ausnahme alle Kurse im Kernbereich anbietet. An den Universitäten wiederum präsentiert sich die fakultäre Einbettung der Datenwissenschaften unterschiedlich: In Zürich, Lausanne und Neuchâtel besteht ein gemeinsames Angebot an Fakultäten, die Informatik und Wirtschaftswissenschaften integrieren.⁶ An anderen Universitäten werden die Studiengänge hingegen durch Informatikdepartemente alleine getragen (BENEFRI und USI). An den Fachhochschulen schliesslich wird das Angebot – sowohl in der grundständigen Lehre als auch in der Weiterbildung – durch Departemente für Technik (FHNW, ZHAW, SUPSI), Informatik (HSLU – Bachelor), Wirtschaft (HSLU – Master) oder für Technik und Informatik (BFH) gemeinsam bestritten. Bei den Weiterbildungsstudiengängen bieten die Mo-

⁶ Die HSG als Business School ist die Ausnahme, weil das Angebot dort übergreifend aufgebaut wird.

dul-Lösungen die Möglichkeit, Lehrinhalte aus verschiedenen ökonomischen Feldern in ein MAS-Studium zu integrieren (vgl. Kap. 9.8).

Die Analyse der institutionellen Verortung der neuen Studiengänge an Schweizer Universitäten und Hochschulen zeigt, dass für die Datenwissenschaften kaum neue organisationale Gefässe ausgebildet wurden, anders als etwa an US-amerikanischen Forschungsuniversitäten, die oft ihre organisationalen Strukturen weitreichend neu konzipieren (Moore-Sloan Data Science Environments 2018; Wing et al. 2018).⁷ Die Datenwissenschaften als Wissensfeld werden vielmehr grösstenteils innerhalb der existierenden Gefässe konzeptualisiert, d. h. als neue Kategorie in eine stark disziplinär strukturierte Landschaft integriert. In der Mehrheit der Fälle sind dies Fakultäten bzw. Departemente für Informatik und Technik, womit die Datenwissenschaften primär als technikwissenschaftliches Fachgebiet dargestellt werden. Trotzdem wirken die interdisziplinären Curricula als »Kompromissprodukte« verschiedener Disziplinen und repräsentieren insofern auch die Vielstimmigkeit der involvierten Wissensfelder.

Einige Universitäten und Hochschulen haben strategische Initiativen lanciert, die sich über Datenwissenschaften hinaus Themen wie künstliche Intelligenz, Robotik oder vernetztes Lernen annehmen. Lokale Hochschulinitiativen zur Etablierung von Studiengängen und Einrichtung von Professuren treffen auf die generelle politische Bereitschaft, die Digitalisierung von Bildung und Forschung stärker zu fördern.⁸ In der Selbstbeschreibung reagieren sie damit in erster Linie auf die soziotechnischen Veränderungen, die unter Digitalisierung gefasst und als Triebkräfte gesellschaftlichen und organisationalen Wandels gerahmt werden.⁹ Demgegenüber bleibt die Rolle der Datenwissenschaften in den strategischen Initiativen enger auf wissenschaftlich-technische Diskurse beschränkt.

Die untersuchten Beispiele zeigen, dass der »raschen« Etablierung von Studiengängen zur Herstellung von Sichtbarkeit in einem zunächst noch kaum bekannten Themenfeld gegenüber – mittel- bis langfristigen – intraorganisationalen Kooperationen Priorität eingeräumt wurde. Disziplinäre Konfliktlinien über den Aufbau der Studienangebote innerhalb von Hochschulen und Universitäten als organisationalen Feldern bleiben im empirischen Material weitgehend latent. Intensive Auseinandersetzungen um die disziplinäre Vorherrschaft über die Datenwissenschaften hätten die beschleunigte Expansion und Diffusion, wie sie im Schweizer Hochschulfeld zu beobachten ist, gehemmt und wurden insofern den hochschulstrategischen Bestrebungen, »die Ersten« zu sein, untergeordnet. Somit wurde die Interdisziplinarität der Studienangebote in Datenwissenschaften gewissermassen auf die Ebene der curricularen Inhalte transferiert.

7 Eine Ausnahme stellt die diskutierte »Initiative for Data Science in Switzerland« des ETH-Bereichs dar, die mit dem SDSC ein disziplinen- und hochschulübergreifendes Forschungszentrum umfasst und transversale Kollaborationen zwischen Universitäten, Forschungsinstitutionen und der Industrie unterstützt.

8 Sie wird auf Bundesebene in erster Linie in innovations- und wettbewerbspolitischen Diskursen artikuliert, wobei das Argument, im globalen Wettbewerb mithalten zu können, zentral ist, während die Kantone in erster Linie die »Ausbildung von Fachkräften« ins Zentrum stellen (Krummenacher 2019).

9 Beispiele sind die Strategie ETH+ der ETH Zürich, die Digital Society Initiative der Universität Zürich, die Digitalisierungsinitiative der Zürcher Hochschulen und die IT-Bildungsoffensive des Kantons St. Gallen.

8.6 Die Synchronizität von Kooperations- und Konkurrenzverhältnissen im akademischen Feld

Ökonomische Motive, hochschulpolitische Vorgaben sowie organisationale Mechanismen und Abläufe prägen die Verhandlung der Datenwissenschaften in den untersuchten Hochschulen und Universitäten. Sie schränken die relative Autonomie, die den Organisationen im akademischen Feld zukommt, ein und zeigen exemplarisch, wie Akteur*innen anderer Felder deren Möglichkeitsräume und Strategien zu beeinflussen vermögen. Des Weiteren rahmen die Relationen zwischen den Hochschulen und Universitäten die Positionierung einzelner Studiengänge. Ich zeichne deshalb im Folgenden die Entwicklung der neuen Kategorie »Data Science« nach und setze sie in Bezug zu den Positionen und Hierarchien, die das Feld der Schweizer Hochschulen und Universitäten strukturieren (Kap. 8.6.1). Anschliessend arbeite ich die wechselseitigen Bezugnahmen zwischen den Studiengängen und weiteren Akteur*innen im empirischen Material heraus. Dabei treten zwei entgegengesetzte Entwicklungen zutage: Zum einen wird eine Differenzierung nach innen deutlich, die sich im nationalstaatlichen Kontext des untersuchten Feldes sowohl in affirmierenden als auch abgrenzenden Beziehungen zu den beiden ETH äussert (Kap. 8.6.2). Zum anderen ordnen sich die untersuchten Studiengänge in ein globales akademisches Feld ein, indem sie sich an Empfehlungen von Fachgesellschaften, Vorbildinstitutionen und Rankings orientieren (Kap. 8.6.3).

8.6.1 Entwicklung der Datenwissenschaften im akademischen Feld der Schweiz

Die zurzeit angebotenen Studiengänge in Datenwissenschaften an Schweizer Universitäten und Hochschulen finden sich sowohl im Bereich der akademischen Grundbildung (Bachelor und Master) an den ETH, Universitäten und Fachhochschulen als auch in der akademischen Weiterbildung (CAS, DAS, MAS). Es existieren 42 Studiengänge und Vertiefungen auf allen Stufenniveaus, mit Ausnahme von Bachelor-Studiengängen an Universitäten und den ETH (vgl. Tabelle 13 im Anhang). Obwohl bestimmte Kurse schon lange Bestandteil von Curricula der Mathematik, Statistik oder technischen Wissenschaften waren, werden erste Angebote, die explizit auf »Data Science« referieren, ab 2013 zunächst an den Fachhochschulen ZHAW und BFH auf Stufe Weiterbildung eingeführt.¹⁰ Die Implementierung auf Stufe Weiterbildung markiert die starke Nachfrageorientierung der Studiengänge in der Genese des Feldes: Den berufsorientierten Weiterbildungsstudiengängen wird eine Art »Pilotfunktion« zugeschrieben: Während darin einerseits ein »Experimentiercharakter« gesehen wird, um neue Inhalte und Formate zu testen, wird andererseits die Marktcompatibilität der Studienangebote betont. Den Pionier*innen unter den Fachhochschulen bot sich dadurch die Möglichkeit, den entstehenden Zwischenraum rasch mit ihrer Kernexpertise zu besetzen, indem sie »Data Science« weniger als wissenschaftliches, sondern in erster Linie als berufliches Praxisfeld definierten. Dadurch gelang es ihnen, den ETH und Universitäten ein Stück weit zuvorzukommen, feldübergreifende Netzwer-

¹⁰ In verschiedenen Hochschulsystemen sind synchrone Bemühungen zu beobachten: So existierte in den USA 2014 lediglich ein halbes Dutzend Studienprogramme, die »Data Science« im Titel trugen (Aasheim et al. 2014).

ke und wichtige Konferenzen wie die Swiss Conference on Data Science zu etablieren und massgeblich zu prägen. Sie erarbeiteten sich so ein gewisses Renommee im Feld, das auch von den ETH und Universitäten anerkannt wird.¹¹

An den Universitäten und ETH verlief die Entwicklung umgekehrt: Universitäten und ETH bauten zunächst forschungsbasierte Masterprogramme auf, bevor sie Studienangebote für die berufsorientierte Weiterbildung anboten. Regelstudiengänge existieren auf Masterstufe an den Universitäten seit 2016 (Zürich und BENEFRI) und an den ETH seit 2017. Andere Universitäten wie Genf, Lausanne oder USI implementierten Studiengänge in verwandten Themenfeldern wie »Business Analytics« oder »Artificial Intelligence«. Angesichts des beschriebenen Planungsvorlaufes von zwei bis drei Jahren indiziert dies hochschulstrategische Überlegungen von Positionierung und Differenzierung. Während die Mehrheit der Masterstudiengänge auf 90 ECTS-Punkte ausgelegt ist, handelt es sich bei den zwei ETH-Programmen um spezialisierte Masterstudiengänge à 120 ECTS-Punkte. Den Umfang begründen die ETH damit, »dass sich Studierende Kompetenzen in mehr Bereichen aneignen müssen« (Prof_ETH_B: 2:5). Mit der Wahl bauten die ETH auf die disziplinäre Grundierung im Bachelorstudium und kombinierten wissenschaftliche Interdisziplinarität, Forschungsorientierung sowie eine starke Methodenfokussierung in ihren Masterstudiengängen. Indem sie diese mit langer Vorlaufzeit ankündigten und explizit in den Dienst nationalstaatlicher Interessen stellten, schufen sie gleichzeitig eine Orientierungs- und Ordnungsfunktion, was anderen Akteur*innen im Feld Spielräume zur eigenen Profilbildung und Abgrenzung bot.

Solche Möglichkeitsräume hatten die Fachhochschulen aus finanziellen, organisationalen und strukturellen Gründen nicht: Sie schufen im Bereich der Regelstudiengänge zunächst erste Vertiefungsrichtungen sowie Spezialisierungen innerhalb bestehender Informatikstudiengänge im Bachelor (HSLU, HSR und FHNW) sowie dem hochschulübergreifend angebotenen Master of Engineering und reagierten damit auf die rasch anwachsende Zahl neuer Studienangebote an den Universitäten und ETH. 2018 startete der erste Masterstudiengang »Applied Information and Data Science« an der HSLU. 2019 lancierte die FHNW bereits den ersten eigenständigen Studiengang in Data Science auf Bachelorstufe, gefolgt von drei weiteren in den Folgejahren (HSLU und FHGR). Da die Übertrittsquoten in die technischen Fachbereiche – wo die Datenwissenschaften verortet werden – auf Masterstufe an den Fachhochschulen sehr niedrig sind,¹² war die Implementation auf Bachelorstufe ein strategisches Mittel, um sich in ihrem Hauptsegment – der praxisorientierten Ausbildung von Berufstätigen – zu positionieren, was mit angepassten Anforderungs- und Kompetenzprofilen einhergehen musste. Mit Ausnahme eines Lehrprogramms an der Universität St.

11 Dies manifestiert sich auch in subfeldübergreifenden Kooperationen, wie dem seit 2018 durch swissuniversities geförderten PhD-Netzwerk in Data Science in Kooperation mehrerer Fachhochschulen und Universitäten, das Promotionsmöglichkeiten auch für Masterabsolvent*innen und Mitarbeitende an Fachhochschulen eröffnet, vgl. <https://phd-data-science.ch/> (Zugriff: 03.02.2022).

12 Sie lagen 2018 für den Fachbereich Technik & IT an Fachhochschulen bei lediglich 16% gegenüber 97% im Bereich Technische Wissenschaften an den ETH und Universitäten, vgl. BFS (2021): Übergänge und Verläufe auf der Tertiärstufe: <https://www.bfs.admin.ch/bfs/de/home/statistiken/bildung-wissenschaft/uebertritte-verlaue-bildungsbereich/tertiaerstufe.assetdetail.18964734.html> (Zugriff: 03.02.2022).

Gallen gibt es an den Universitäten und ETH derzeit keine vergleichbaren Studiengänge auf Bachelorstufe. Gegenüber ihrer anfänglichen Rolle als Vorreiterinnen blieb den Fachhochschulen im Bereich der Regelstudiengänge insofern bloss die Rolle als Nachzüglerinnen, was sich neben finanziellen Möglichkeiten durch die Komplexität der Entscheidungsstrukturen und eine – im Vergleich zu den Universitäten – geringere Autonomie gegenüber politischen Akteur*innen und ökonomischen Anspruchsgruppen erklärt.

Schliesslich ist in jüngster Zeit zunehmend die Verknüpfung mit anderen Kategorien zu beobachten: So heissen kürzlich etablierte Studiengänge bzw. Vertiefungen beispielsweise »Financial Entrepreneurship and Data Science« (Universität Lausanne), »Mobility, Data Science and Economics« (HSLU), »Data Science for Industry 4.0« (SUPSI) oder »Data Analytics and Economics« (Universität Fribourg). Sie verbinden datenwissenschaftliche Expertise mit spezifischen Anwendungsfeldern. Darin äussern sich zwei widersprüchliche Entwicklungen: Zum einen die Abkehr von Studiengängen, die durch ihre Denomination einen universalen Anspruch auf die Wissensformation erheben, gleichzeitig jedoch einen methodisch klar umrissenen Kernbereich formulieren und somit zu einer epistemischen Schliessung beitragen. Zum anderen erweist sich die Offenheit des Wissensfeldes als inhaltlich äusserst adaptiv an heterogene Anwendungsfelder. So bieten sich durch die Verknüpfung mit weiteren Themenfeldern Opportunitäten zur Spezialisierung und Nischenbildung, was insbesondere für kleinere Universitäten und Hochschulen (oder solche ohne technikwissenschaftliche Departemente) eine Strategie darstellt, um ihren Platz im Feld zu beanspruchen.

Die Diffusion der Kategorie »Data Science« markiert eine spezifische Dynamik, die durch die institutionelle Verortung der Akteur*innen sowie der Hierarchie im Feld betrachtet werden muss: Während die etablierten, gut kapitalisierten und renommierten Akteur*innen auf eine epistemische und soziale Schliessung des Feldes zuarbeiten, versuchen die Newcomer durch inhaltliche und curriculare Innovationen und Spezialisierung die Grenzen des Feldes offen zu halten. Sie suchen und etablieren neue, feld- und disziplinenübergreifende Kollaborationen, um sich die Unterbestimmtheit der Datenwissenschaften zunutze zu machen.

8.6.2 Die Differenzierung des Feldes nach innen: Orientierung und Abgrenzung gegenüber den ETH

Die Koexistenz und Parallelität verschiedener Initiativen und Bemühungen, um die Multidimensionalität der Datenwissenschaften an Schweizer Hochschulen und Universitäten in konkrete organisationale Strukturen zu übersetzen, ist in den Kontext hochschulpolitischer, ökonomischer und epistemologischer Transformationsprozesse des akademischen Feldes zu stellen: New-Public-Management-Reformen (Felouzis et al. 2013; Grande et al. 2013), der Bologna-Prozess (Jessop 2008; Maesse 2010) und Internationalisierungsprozesse (Meyer & Ramirez 2006) haben in den vergangenen Jahrzehnten zur Integration eines Feldes beigetragen, das sich durch multiple, synchrone Kooperations- und Konkurrenzverhältnisse auszeichnet. So konkurrieren Universitäten und Hochschulen als »Wettbewerbsakteure« (Hasse 2015; Krücken 2017) um Forschungsmittel, Visibilität, Forschende und Studierende (Rostetter 2020). Synchron dazu sind sie in vielfältige Kollaborationsformen und Netzwerke hochschulpolitischer

Gremien und zivilgesellschaftlicher Aktivitäten eingebunden sind (Jongbloed et al. 2008).

Im lokalen Kontext des Untersuchungsfeldes artikuliert sich das institutionelle Arrangement von Zusammenarbeit und Wettbewerb in kontinuierlichen wechselseitigen Beobachtungsverhältnissen der jeweiligen Hochschulstrategien, Planungsaktivitäten und Curricula. Mehrere Befragte gaben an, sich im Rahmen von Planungsprozessen bei anderen Angeboten informiert und »inspirier[t]« zu haben:

»[K]ategorisch gab es das nicht, dass man sich von anderen Hochschulen bewusst abgrenzen wollte; aber natürlich sind [wir] alle vernetzt im In- und Ausland und man kriegt mit, was bei anderen so läuft und lässt sich davon inspirieren« (Prof_ETH_A: 4:12).

Vor diesem Hintergrund ist es interessant, dass niemand von einem »Konkurrenzverhältnis« spricht; vielmehr wird die Notwendigkeit betont, »dass die Schweiz mehr Data Scientists ausbildet« (Prof_UH_D: 1:15). Die unterschiedlich ausgeprägten Profile werden im Sinne einer »Arbeitsteilung« als miteinander komplementär gelesen. Dies erfolgt beispielsweise durch die Zuschreibung bestimmter Schwerpunkte und Expertisen, die durchaus gewürdigt werden. So werden etwa den unterschiedlichen Studiengruppen an ETH, Universitäten und Fachhochschulen spezifische Interessen attribuiert, was sich wiederum in der Besetzung bestimmter Studieninhalte äussert.

Für das Untersuchungsfeld bilden primär die beiden ETH die massgebende Referenz. So werden die Leistungen und Beiträge der ETH in den Datenwissenschaften von Lehrenden an Universitäten und Fachhochschulen durchgehend anerkannt und hochgeschätzt:

»Ein Student, den ich betreute, geht nun nach Zürich für den [Master]. Habe die Kursliste gesehen und das ist super, exzellent. Sie werden sich die Studierenden auswählen können, [...] vom Format her ist es ein Top-Programm. In den USA gibt es die Programme schon länger und dort unterrichten Leute, die diese Methoden erst entwickelt haben. Aber auch an der ETH gibt es Leute, die viel Grundlagenarbeit in dem Bereich geleistet haben. [Das] ETH-Programm ist natürlich nicht mit unserem Programm vergleichbar« (Prof_UH_C: 3:31).

Indem der Befragte dem ETH-Master sowohl inhaltliche als auch personale Exzellenz zuschreibt und zugleich das eigene Studienprogramm als »nicht [...] vergleichbar« taxiert, reproduziert er die etablierte Hierarchie des Feldes. Auch wenn die Mehrheit der Befragten angibt, die neuen ETH-Programme nicht *en détail* zu kennen, sind doch alle davon überzeugt, dass diese sicherlich »auf höchstem Niveau« sein werden. Entsprechend selbstbewusst spricht ein Professor an der ETH auf die Frage nach möglichen Vorbildern denn auch von der eigenen »Vorbildfunktion«, die man einnehmen wolle, um »weltweite Visibilität« zu erreichen (Prof_ETH_B: 2:6). Die ETH-Professor*innen begründen die privilegierte Position der ETH mit den Forschungsleistungen der beteiligten Departemente sowie der Interdisziplinarität des Studiengangs. Parallel dazu ist auch eine Differenzierung vonseiten der beiden ETH gegenüber anderen Programmen erkennbar:

»Data Science ist [...] ein breites Feld von unterschiedlich ausgebildeten Data Scientists. Und ich glaube, da ist es sehr zu begrüßen, dass es wirklich dieses Spektrum an möglichen Programmen gibt, eben von den Fachhochschulen bis zu den Universitäten und ETHs. Wir versuchen natürlich die Studenten da auszubilden, wo die ETH besonders stark ist, natürlich halt in den methodologischen Grundlagen [...] in den genannten Feldern. Das heisst, die Algorithmik selbst zu verstehen und weiter entwickeln zu können. Also die Methoden nicht nur anwenden können, sondern auch wirklich weiter entwickeln können. Das ist ein grosser Schwerpunkt hier und das umfasst natürlich viele Facetten, weil viele Methoden hier im Einsatz sind« (Prof_ETH_A: 4:21).

Das Verständnis von »Data Science« als einem »breiten Feld«, das unterschiedliche hochschulspezifische, disziplinäre und methodische Zugänge umfasst, beschreibt die synchronen Kollaborations- und Konkurrenzverhältnisse: Trotz der Anerkennung von »unterschiedlich ausgebildeten Data Scientists«, was eine Arbeitsteilung zwischen den verschiedenen Institutionen begünstigt, grenzt sich der Befragte durch wiederholten Verweis auf die »methodologischen Grundlagen« sowie »Algorithmik« von einer reinen Methodenanwendung ab – eine Aussage, die umso deutlicher wird, wenn man sie mit der betonten »Anwendungsorientierung« in den Äusserungen der Fachhochschulprofessor*innen kontrastiert. Eine ebenso deutliche Distinktion erfolgt gegenüber jenen Programmen, die sich auf die berufliche Aus- bzw. Weiterbildung von Berufstätigen (»vocational training«) fokussieren (EPFL 2016: 46).

Trotz der allseitigen Anerkennung sind es insbesondere Lehrende an Fachhochschulen, die sich pointiert von den ETH distanzieren: Sie argumentieren, dass die Fachhochschulen einen politischen Auftrag zur Zusammenarbeit mit der Industrie hätten und somit stärker die Anwendungsorientierung der vermittelten Wissensinhalte fokussieren müssten. Die Distinktion vonseiten der Fachhochschulprofessor*innen äussert sich in zwei unterschiedlichen Dimensionen: Eine erste Dimension bildet die diskutierte Trennlinie Grundlagenforschung vs. Anwendungsorientierung, wobei den ETH eine klare »Forschungsausrichtung« zugeschrieben wird, was für die Fachhochschulen »nicht relevant« sei:

»Die neuen Master-Studiengänge an der ETH haben ganz klar eine Forschungsausrichtung. [Die] sind für uns nicht relevant, beispielsweise was die EPFL im Bereich der Hirnforschung macht, und das ist absolut auch Data Science [...]. Aber das sind nicht Bereiche, in denen wir unterwegs sind. Unser Fokus liegt auf operationeller Data Science« (Prof_FH_C: 9:21).

In diesem Zitat reproduziert der Befragte gewissermassen die epistemischen Grenzbeziehungen des Kollegen der ETH aus entgegengerichteter Warte und führt sie zur Verteidigung des eigenen Tätigkeitsbereichs an. Eine zweite Dimension bilden die unterschiedlichen Ressourcen und Kapitalverhältnisse, die die Hierarchien des Feldes abbilden. So haben die Bundesmillionen, die im Rahmen der »Initiative for Data Science in Switzerland« zusätzlich für den ETH-Bereich gesprochen wurden, zu gewissen Unsicherheiten über die eigene Positionierung im Feld geführt, was von einer interviewten Person wie folgt verhandelt wird:

»Es hat mich ein bisschen erstaunt, dass angesichts des ganzen Rummels, der diese Summe ausgelöst hat: ›Oh, der Bund zahlt jetzt 30 Millionen, damit wir es endlich mal richtig machen‹. Und am Ende kommt nur ein Tool dabei raus? Finde ich irgendwie eine saukleine Vision für 30 Millionen! [lacht] [...] Das ist vielleicht eine Frage der Befindlichkeit, aber es gab so: ›Hey, wir müssen uns nicht verstecken!‹ Man denkt manchmal so: Wenn man was *grassroots*-mässig startet und dann etwas [...] richtig gemacht [wird]: ›Ah, jetzt können wir einpacken, das sind wir am gleichen Platz.‹ Nein, ich glaube nicht« (Prof_FH_B: 7:10).

Und weiter:

»Ich glaube, das wird schon gut sein, was da läuft. Da sind super Leute dabei. Die Leute [...] [werden] sicher viel für Interdisziplinarität an der ETH beitragen. Aber die Ausrichtung ist so unterschiedlich: Das [Swiss Data Science Center] hilft seinen Forschern, etwas besser in Daten zu werden. Und an den Fachhochschulen sind wir halt doch mehr darauf ausgerichtet, Cases zu lösen, die einen grösseren Impact haben, also im Sinne von Breite auch in die Wirtschaft. Und ich glaube, das wird sich gut ergänzen« (Prof_FH_B: 7:11).

Obwohl die Person keinen Zweifel an der fachlichen Qualität der ETH-Forschenden hat und die »Frage der Befindlichkeit« letztlich der Annahme einer Komplementarität der beiden Angebote (»das wird sich gut ergänzen«) weicht, verweist der Vorwurf von mangelndem »Impact« in der »Breite« auf die fundamentale Konfliktlinie zwischen akademischer Forschungsorientierung und Anwendungsorientierung in der Praxis. Die ungleiche Kapitaldistribution zwischen den involvierten Hochschulen im Feld wird somit in unterschiedliche Praktiken und Positionierungen im Feld der Datenwissenschaften übersetzt. Indem die Kapitalverteilung den Hochschulen und Universitäten als Akteur*innen Möglichkeitsräume für eigene Investitionen eröffnet bzw. limitiert, trägt sie gleichzeitig zur Reproduktion der etablierten Feldstrukturen bei.

Die Rolle des ETH-Bereichs als selbsternannter »Leuchtturm« (ETH-Rat 2014: 6) des schweizerischen Hochschulfeldes besteht somit im Falle der neuen Studiengänge verallgemeinert in einer Orientierungs- und Ordnungsfunktion für andere Universitäten und Fachhochschulen. Trotz der grundsätzlich positiv-affirmierenden Einstellungen gegenüber den Leistungen der ETH in den Datenwissenschaften ist gerade keine Anpassung organisationaler Strukturen oder inhaltlicher Ausrichtungen zu beobachten, sondern vielmehr divergierende Positionsbezüge und die Konfirmation der eigenen, konkurrierenden Profile. Der Fokus auf Anwendungsorientierung, der für das akademische Feld der Datenwissenschaften insgesamt kennzeichnend ist, bietet demnach in seinen multiplen Spielarten Möglichkeiten zur Diversifizierung und Abgrenzung.

8.6.3 Die Integration in ein globales Feld: Empfehlungen, Vorbilder und Rankings

Über den ›lokalen‹ Kontext des akademischen Feldes hinaus sind die untersuchten Universitäten und Hochschulen in transnationale Wahrnehmungs- und Beobachtungsstrukturen eingebunden: Als neue Kategorie ist »Data Science« ein globales Phä-

nomen, das sich in einer Vielzahl neuer Studiengänge rund um den Globus manifestiert (Demchenko et al. 2016; Song & Zhu 2017; Stadelmann et al. 2013; Tang & Sae-Lim 2016; Zhang et al. 2017). Die Curricula sind in diesem Sinne nicht nur Stellungnahmen in einem nationalstaatlich konfigurierten Feld, sondern orientieren sich zudem an strukturell ähnlichen Positionen eines globalen Feldes der Hochschulbildung (Hedmo et al. 2007; Zapp & Lerch 2020; Zapp & Ramirez 2019). Die Analyse der Curricula und Interviews identifiziert drei Mechanismen, die eine Ausrichtung an Akteur*innen und Institutionen jenseits der Schweizer Hochschullandschaft befördern und somit zur Integration in ein globales akademisches Feld der Datenwissenschaften beitragen.

Ein erster Mechanismus ist die Orientierung an curricularen Empfehlungen von professionellen Fachgesellschaften, Akademien oder Forschungsprojekten, die einen gewissen normativen Zwang auf die Ausgestaltung und die Implementation neuer Studiengänge ausüben. So flossen bei Studiengängen, die von Informatikdepartementen getragen werden, die curricularen Empfehlungen, die gemeinsam von den computer- bzw. ingenieurwissenschaftlichen Fachgesellschaften ACM (Association for Computing Machinery) und der IEEE (Institute of Electrical and Electronics Engineers) herausgegeben werden (ACM & IEEE Computer Society 2013), in die Planung mit ein.¹³

Auch die im Rahmen von EDISON entwickelten Kompetenzprofile und Modell-Curricula (Demchenko et al. 2016) entfalten strukturierende Effekte auf die Planung und Entwicklung der Curricula, wie ein Befragter erläutert:

»Wir wollten eigentlich auch gar nicht EDISON hier realisieren, wir wollten eigentlich zunächst so einen Data-Science-Studiengang machen und dann haben wir eigentlich dieses EDISON Framework entdeckt. Was ein gutes Projekt ist aus meiner Sicht. [...] Das lief bis Ende 2017. Und ähm ja, wir haben versucht, Inputs zu leisten. Wir haben auch unseren Studiengang dort diskutiert. Also wir haben das bereits ein wenig als Sounding Board genutzt und unsere Ideen ein wenig abgestellt. [...] Aber das ist jetzt nicht in dem Sinne EDISON-approved oder so. (lacht) [...] Aber das trägt sehr stark den Geist des Projekts« (Prof_FH).

Das Zitat macht deutlich, dass es weniger um formelle Zertifizierungen oder Akkreditierungen der untersuchten Studiengänge geht, obwohl solche Institutionen einen nicht zu unterschätzenden Einfluss auf die Herausbildung eines globalen Feldes der Hochschulbildung haben (Zapp & Ramirez 2019; Bloch et al. 2018). Vielmehr stellt die Orientierung an EDISON eine eigenständige Übernahme und Implementierung zentraler Leitideen (»Geist des Projektes«) und curriculärer Elemente dar. Ohne unmittelbaren normativen Zwang auf das Studienprogramm auszuüben, prägen solche Empfehlungen sowohl die Strukturen als auch die Inhalte neuer Studiengänge und leisten ihren Beitrag zur Vereinheitlichung des Feldes.

Eine zweite zentrale Orientierungsgrösse stellt für die untersuchten Akteur*innen das Feld der US-amerikanischen Forschungsuniversitäten dar. Diese werden

13 Die Empfehlungen der US-amerikanischen Akademien der Wissenschaften für Bachelor-Studiengänge in Datenwissenschaften (NASEM 2018) waren den Befragten hingegen zum Interviewzeitpunkt noch nicht bekannt. In der Schweiz existieren gemäss derzeitigem Kenntnisstand keine vergleichbaren curricularen Empfehlungen für technikkwissenschaftliche Studiengänge im Hochschulbereich.

zusammen mit Unternehmen im »Silicon Valley« von vielen Befragten entweder als der eigentliche Ausgangspunkt oder als der dynamische Treiber der Etablierung von »Data Science« als eigenes Feld betrachtet (oder beides). So bezogen sich alle befragten Professor*innen bei der Frage, ob es »Vorbilder« bei der Planung der eigenen Curricula gab, direkt oder indirekt auf die Situation an US-amerikanischen Hochschulen:

»Unsere Vorbilder waren eigentlich ausserhalb von Europa, weil es praktisch nichts gab, als wir 2014 mit der Planung begannen. [Wir orientierten uns] an amerikanischen Hochschulen, auch weil wir gute Kontakte haben« (Prof_UH_A: 6:41).

Auf Nachfrage nach konkreten Vorbildinstitutionen nennen die Befragten einige wenige leistungsstarke US-amerikanische Forschungsuniversitäten wie die Columbia University, Carnegie Mellon University, Stanford University, University of California Berkeley oder das Massachusetts Institute of Technology. Auch der ETH-Rat (2016: 1) verweist anlässlich der Lancierung der »Initiative for Data Science in Switzerland« auf die »führenden internationalen Forschungsinstitute und Lehranstalten« als massgebende Referenzgrössen. Der angestrebte »Austausch der Best Practices in der Lehrplanentwicklung sowie gemeinsame Forschungsprojekte zur Weiterentwicklung der Datenwissenschaften« (ebd.) bezeichnet dabei den – für die Positionierung der ETH charakteristischen – Anspruch, nicht nur die Curricula der anderen wahrzunehmen, sondern selbst eine globale Referenzgrösse mit »Vorbildfunktion« zu werden.

Trotz der divergierenden Konzeptualisierung und Implementierung der Datenwissenschaften an US-amerikanischen Forschungsuniversitäten, die sich intensiv mit der Transformation und Innovation organisationaler Strukturen auseinandersetzen, was in der Schweiz noch kaum zu beobachten ist, nehmen die untersuchten Studienprogramme die Curricula an US-amerikanischen Forschungsuniversitäten und technischen Hochschulen als erfolgreiche Modelle und Vorbilder wahr.¹⁴

Einen dritten Mechanismus stellen Listen und Rankings als Dispositive des Vergleichs dar (Hamann 2018; Heintz 2008, 2016). Um sich im Dickicht des rasch expandierenden akademischen Feldes der Datenwissenschaften zurechtzufinden, orientieren sich die Akteur*innen auch an solchen (hierarchisierenden) Aufzählungen von Studienprogrammen und Hochschulen. Eine wichtige Referenz bildet die Internetplattform »www.mastersindatascience.org«, die von mehreren Auskunftspersonen als eine wichtige Informationsquelle über die US-amerikanische Hochschullandschaft angegeben wird:

»[Ich] habe mich damals stark am amerikanischen Modell [...] orientiert. Es gibt so Webseiten, die diese Profile sammeln [...] wie bspw. die Plattform mastersindatascience.org. [...] Dort gibt es Definitionen: Was ist ein Data Scientist? Wo sind die Betätigungsfelder, in welchen Industrien und so weiter. [...] Aber wie gesagt, das war nicht in Reichweite, einen Studiengang mit stark mathematisch-statistischer Ausrichtung zu machen. Aber wir haben uns auch nicht an den wissenschaftlichen Universitäten in den USA orien-

14 Zu einigen der US-amerikanischen Vorbildinstitutionen bestehen überdies Austauschbeziehungen, die von den Befragten und anderen Lehrenden über persönliche Kontakte oder Karrierewege etabliert wurden.

tiert, sondern eher an anwendungsorientierten Domänen, die bei uns die Fachhochschulen ansprechen« (Prof_FH_C: 9:12).

Die Webseite beschreibt Definitionen und »Karriereprofile« für Data Scientists und führt Studiengänge in den Datenwissenschaften in Form unterschiedlicher Listen auf: So führt etwa die Liste »23 Great Schools with Master's Programs in Data Science« Studienprogramme forschungstarker privater und öffentlicher Universitäten auf, aber auch solche, die explizit als Werbung ausgewiesen werden.¹⁵ Trotz der strukturellen und inhaltlichen Diversität der Curricula werden sie durch ihre Qualität in dieselbe Gruppe kategorisiert, obwohl die Auswahlkriterien nicht weiter spezifiziert werden. Es fehlen somit für die Liste der »grossartigen Hochschulen« explizite Vergleichsdimensionen, um die Unterschiede zwischen äquivalenten Studiengängen feststellen zu können (Heintz 2010: 164).¹⁶ Trotz der Verschiedenheit und Arbitrarität der angeführten Listen verweist deren Verbreitung und Bekanntheit im Untersuchungsfeld auf die zentrale Ordnungs- und Orientierungsfunktion solcher Listen in digitalen Welten, indem sie »ganz heterogene und vor allem verstreute und unzusammenhängende Elemente in zumeist zeitlich limitierte Zusammenhänge bringen« (Mämecke et al. 2018: 7f.).

Schliesslich sind neben Listen auch Rankings, d. h. hierarchisierte Ranglisten, Teil des Vergleichsdispositivs: Studiengänge führen diese an, um ihre Ausrichtung und Positionierung zu legitimieren (GSEM 2016: 3). Rankings ebnen die Differenzen zwischen unterschiedlichen Ausprägungen und Positionierungen im Feld ein, indem sie durch Vergleiche und Ordnungskonstruktionen eindimensionale Homogenität herstellen (Sauder & Espeland 2009). Sie reduzieren somit gewissermassen die Multidimensionalität der Datenwissenschaften als Wissensfeld und führen gleichzeitig zur Reproduktion sozialer Hierarchien, die die Positionen von Universitäten und Hochschulen im akademischen Feld insgesamt strukturieren.

Die drei identifizierten Mechanismen rahmen und beeinflussen die Verortung und Implementierung einer neuen Kategorie innerhalb des akademischen Feldes. Die Orientierung an Empfehlungen, Vorbildinstitutionen und Rankings reduziert Unsicherheit über die eigene Verortung in einem rapide expandierenden Teil des Hochschulfeldes und ermöglicht gleichzeitig die Anlehnung an etablierte Studienprogramme, insbesondere an renommierten US-amerikanischen Forschungsuniversitäten. Damit tragen sie zur Integration der untersuchten Studiengänge in ein global strukturiertes akademisches Feld der Datenwissenschaften bei.

15 Vgl. <https://www.mastersindatascience.org/schools/23-great-schools-with-masters-programs-in-data-science/> (Zugriff: 03.02.2022). Des Weiteren gehören Listen nach geographischen (US-amerikanische Bundesstaaten), feldspezifischen (ökonomische Berufsfelder) oder technologischen (»Data Science Technologies« wie R, Python oder Hadoop) Kriterien zum Vergleichsrepertoire der Webseite.

16 Vergleiche mit anderen Listen von Studiengängen in Datenwissenschaften legen nahe, dass die Webseite »www.mastersindatascience.org« nicht vollständig ist oder nicht (mehr) regelmässig aktualisiert wird.

8.7 Diskussion

Die Analyse hat gezeigt, dass nicht nur wissenschaftsinterne, epistemische Transformationen, sondern auch hochschulpolitische und ökonomische Rahmenbedingungen sowie organisationale Abläufe die Datenwissenschaften innerhalb des akademischen Feldes prägen. Dabei offenbart sich ein Gegensatz zwischen den diskutierten Diagnosen eines fundamentalen gesellschaftlichen Wandels (»Digitalisierung«, »Paradigmenwechsel«) und konkreten Praktiken an Universitäten und Hochschulen, die sich mit der Herausforderung konfrontiert sehen, wie ein neues Wissensfeld in das bestehende disziplinäre Gefüge und organisationale Settings eingefügt werden kann.

Die Analyse der Diffusion der neuen Kategorie »Data Science« macht deutlich, dass diese zunächst in unterschiedlichen gesellschaftlichen Kontexten (innerhalb wie ausserhalb der Hochschulbildung) etabliert und verbreitet werden musste, ehe sie in Studienprogrammen institutionalisiert werden konnte. Die Implementation in Studienangebote folgte auf die Konstruktion der Nachfrage in anderen gesellschaftlichen Bereichen, insbesondere im ökonomischen Feld. Politische Vorgaben befördern und beschleunigen den Aufbau und die Ausgestaltung neuer Studienprogramme vor allem dann, wenn sich organisationale und hochschulpolitische Interessenlagen gegenseitig befördern. Auch die zur Verfügung stehenden finanziellen Ressourcen sowie die Modalitäten der Hochschulfinanzierung können die Möglichkeitsräume für curriculare Innovationen erheblich erweitern oder einschränken. Das Wechselspiel von hochschulpolitischen Vorgaben und organisationalen Abläufen stellt gewissermassen die Hinterbühne dar, die die organisationalen Verhandlungen und Klassifikationen neuer Studiengänge in eine disziplinär strukturierte Wissenslandschaft nachhaltig prägt.

Die Einführung der Datenwissenschaften erfolgt primär innerhalb der bestehenden organisationalen Strukturen und Orientierungen. Mit Ausnahme der beiden ETH bilden die untersuchten Universitäten und Hochschulen kaum neue organisationale Gefässe aus. Strukturelle Veränderungen werden – wenn überhaupt – eher in politischen und hochschulstrategischen Überlegungen um Digitalisierung verortet und diskutiert. Die Emergenz neuer Kategorien wie »Data Science« wird demnach grundlegend durch die Orientierung an etablierten Strukturen des akademischen Feldes verarbeitet.

Dabei ist die organisationale Verhandlung der Datenwissenschaften, anders als die intensiven epistemischen Auseinandersetzungen über die disziplinäre Deutungshehoheit vermuten lassen, kaum durch Auseinandersetzungen und Konflikte geprägt, da diese weitgehend latent gehalten werden. Dennoch koexistieren Konkurrenzverhältnisse und Praktiken der Abgrenzung mit multiplen Formen der Kollaboration und Anerkennung. So sind klare Bemühungen um Abgrenzung und Positionierung gegenüber anderen Akteur*innen im Feld zu beobachten, insbesondere der Fachhochschulen gegenüber den ETH. Die Universitäten und Hochschulen stehen in Konflikt um finanzielle Ressourcen, zusätzliche Studierende und letztlich über die verschiedenen Perspektiven der Datenwissenschaften. Letztere manifestieren sich in divergierenden Forschungs- und Lehrpraktiken einerseits und symbolischen Grenzziehungen gegenüber anderen Akteur*innen im Feld andererseits. Gleichzeitig tragen die wechselseitige Anerkennung, normative Empfehlungen und die Orientierung an Vorbildinstitutionen aber auch zur Koordination von Planungs- und Implementierungsaktivitäten der unterschiedlichen Akteur*innen im Feld bei. Dies äussert sich, wie ich noch zei-

gen werde, insbesondere in der Herausbildung strukturell ähnlicher Curricula (vgl. Kap. 9). Darüber hinaus begünstigen die unterschiedlichen Devices die Integration der untersuchten Studiengänge in ein globales akademisches Feld der Datenwissenschaften. Neben wechselseitigen Orientierungen und Differenzierungen des Feldes ist es die Synchronizität von Kooperations- und Konkurrenzverhältnissen im akademischen Feld, die die rasche Expansion und Diffusion an Schweizer Universitäten und Hochschulen erst ermöglicht hat.

Die empirisch beobachtete Struktur des Feldes entspricht im Wesentlichen der Differenzierung in die drei Hochschultypen ETH, Universitäten und Fachhochschulen: Sie basiert sowohl auf divergierenden organisationalen Eigenheiten und fachlich-disziplinären Traditionen als auch auf Volumina und Verteilung der feldspezifisch relevanten Kapitalsorten. Während bereits kapitalstarke, grosse Universitäten und Hochschulen mit zusätzlichen Ressourcen ausgestattet werden, bleibt kleineren bzw. weniger gut situierten Universitäten und Hochschulen lediglich die Möglichkeit, durch Profilbildung inhaltliche Nischen zu besetzen sowie durch organisationale Kooperationen Angebote zu etablieren, die sonst nicht möglich wären. In den Studiengängen und Curricula als organisationalen Stellungnahmen reproduzieren sich somit die sozialen Hierarchien des akademischen Feldes.

Entgegen den vielfach geäusserten Annahmen von Innovation und Transformation der organisationalen und sozialen Strukturen des akademischen Feldes durch die Datenwissenschaften verweist die Analyse vielmehr auf die Erhaltung und Perpetuierung der etablierten Arbeitsteilungen und Hierarchien. Der diskursiv konstruierte Möglichkeitsraum, der sich anhand der Datenwissenschaften zwischen den etablierten Feldern eröffnen sollte, erweist sich für die Akteur*innen an Universitäten und Hochschulen als hochgradig durch hochschulpolitische Vorgaben, organisationale Abläufe sowie durch die soziale Hierarchie des akademischen Feldes bestimmt und limitiert. Somit stehen der diskursiven Vielstimmigkeit auf der gesellschaftlichen Makroebene organisationale Engführungen auf bestimmte Positionierungen im Feld gegenüber.

Kapitel 9 – Die Strukturlogik datenwissenschaftlicher Curricula

9.1 Einleitung

Universitäten und Hochschulen reagieren mit der Einrichtung neuer Studienprogramme einerseits auf epistemologische Veränderungen sowie disziplinäre Konfliktlinien in der Wissenschaft. So wird »Data Science« als ein Konvolut von Wissensbeständen, Methoden und epistemischen Praktiken beschrieben, die unterschiedlichen wissenschaftlichen Disziplinen und anderen sozialen Feldern entstammen. Im empirischen Material manifestiert sich dies in je nach disziplinärer Prägung, universitärer Verortung sowie Hochschultyp unterschiedlichen Visionen und Vorstellungen darüber, was die Datenwissenschaften charakterisiert sowie wo und wie sie im akademischen Feld situiert werden sollen. Andererseits berücksichtigen Universitäten und Hochschulen nicht nur wissenschaftliche bzw. disziplinäre Entwicklungen und Überlegungen, sondern auch die Einbettung und Anbindung an ausserwissenschaftliche Referenzen, wie die Nachfrage des Arbeitsmarktes, politische Strategien oder organisationale Finanzierungsmodelle. Entsprechend prägen und strukturieren nicht nur wissenschaftliche, sondern auch ökonomische und politische Vorstellungen die Implementation neuer Wissensgebiete in Studiengängen und Curricula.

Curricula stellen mehr als die Summe der relevanten Wissensbestände eines Feldes dar: Sie artikulieren stets auch Annahmen über die technologischen, ökonomischen, politischen und weiteren Verhältnisse eines Wissensfeldes. Somit repräsentieren Curricula – ähnlich wie Stellenanzeigen – durch die Kombination von Strukturen und Inhalten eine bestimmte Deutung eines Wissensgebiets. Indem verschiedene Curricula miteinander koexistieren und um Studierende konkurrieren, resultieren demnach multiple, vielstimmige Perspektiven auf ein Wissensfeld. Durch die Schaffung neuer Studiengänge versuchen Universitäten und Hochschulen, ihre Absolvent*innen mit dem entsprechenden Rüstzeug auszustatten, um den veränderten technologischen, ökonomischen und sozialen Bedingungen zu begegnen. Gleichzeitig entwerfen, fördern und entwickeln sie damit andererseits aber auch bestimmte Berufsprofile, die bis dato nicht existieren. In diesem Sinne positionieren die Studiengänge ihre Studierenden bzw. ihre Absolvent*innen in neuen Wissensfeldern, was diese interessant und attraktiv für verschiedene umgebende Felder macht. Gleichzeitig prägen sie dadurch wiederum die Entwicklungen von Arbeitsmärkten oder neuen Technologien mit (Jasanoff 2004).

Auf diesen Einsichten aufbauend untersuche ich anhand der Curricula und Interviews mit Lehrenden aktuell verfügbarer Studiengänge an Schweizer Universitäten und Hochschulen, wie die Datenwissenschaften im akademischen Feld organisiert werden, d. h. wie die heterogene Wissensformation in formale und inhaltliche Strukturen übersetzt, in Curricula implementiert und gelehrt wird. Zur Beantwortung der Fragestellungen stütze ich mich auf die Inhaltsanalyse der Curricula in Datenwissenschaften an Schweizer Hochschulen und Universitäten. Trotz der Heterogenität der Curricula identifiziert die Analyse folgende fundierende Strukturlogik: Als zentral für die Studiengänge erweist sich die Differenzierung in einen Kernbereich, der in der Regel die Pflichtveranstaltungen umfasst, einen Wahlbereich, praxisorientierte Kurse (inkl. Praktika) sowie eine Abschlussarbeit. Manche Studiengänge ergänzen diese curricularen Einheiten durch sogenannte komplementäre Inhalte. Von den Regelstudiengängen fällt lediglich das Masterprogramm Künstliche Intelligenz der Fernuniversität Schweiz ab, das weder einen Wahlbereich noch eine Abschlussarbeit umfasst.¹ Alle anderen Studienprogramme und Vertiefungen auf Bachelor- und Masterniveau weisen unabhängig vom Hochschultyp zumindest drei der vier Elemente auf.

Der Aufbau der untersuchten Curricula schliesst insofern an technikwissenschaftliche Wissensfelder wie Engineering an, in denen die Differenzierung in einen Kernbereich und komplementäre Inhalte als »Strukturmuster« (Paulitz & Prietl 2017: 158ff.) beschrieben wird (Grimson & Murphy 2015). Charakteristisch für die Datenwissenschaften ist, dass das Strukturmuster insbesondere durch die Gefässe Wahlbereich sowie die Praxisorientierung ausdifferenziert wird: Die Curricula beschränken sich nicht auf die Ausbildung methodisch und technisch versierter Datenwissenschaftler*innen, die durch »ergänzende Massnahmen« zu entscheidungsfähigen Praktiker*innen vervollständigt werden sollen. Vielmehr verknüpfen sie die Interdisziplinarität eines universellen methodisch-technischen Kerns der Datenwissenschaften mit inhaltlicher Fachexpertise sowie einer ausgeprägten Anwendungsorientierung in einem bestimmten Feld; Letztere erfolgt vor allem durch gemeinsame Projekte mit – akademischen und ausserakademischen – Praxispartner*innen sowie in den Abschlussarbeiten der Studierenden. Die Curricula übersetzen insofern die datenwissenschaftliche Logik der Domänen in die Organisationsstruktur von Universitäten und Hochschulen.

Das Kapitel ist wie folgt aufgebaut: Zunächst untersuche ich die Zulassungsvoraussetzungen, die eine bestimmte Position der Studiengänge im Feld artikulieren (Kap. 8.2). Anschliessend gehe ich auf die einzelnen Bestandteile datenwissenschaftlicher Curricula ein: die Definition eines Kernbereichs (Kap. 8.3) und damit verknüpft die Bedeutung der Statistik (Kap. 8.4), Wahlbereiche und Wahlfächer (Kap. 8.5), praxisorientierte Kurse (Kap. 8.6) und komplementäre Inhalte (Kap. 8.7). Danach analysiere ich das Subfeld der Weiterbildung (Kap. 8.8), das sich von den Regelstudiengängen erheblich unterscheidet. In einer abschliessenden Diskussion (Kap. 8.9) fasse ich die zentralen Erkenntnisse der Analyse zusammen und setze sie in Bezug zu den übergeordneten Fragestellungen der Arbeit.

¹ Der Studiengang bildet gewissermassen einen Hybrid zwischen universitärer Allgemeinbildung und praxisorientierter Weiterbildung. Vgl. <https://fernuni.ch/kuenstliche-intelligenz/master/studieninhalte/> (Zugriff: 03.02.2022).

9.2 Zulassungsvoraussetzungen

Einen ersten zentralen Bestandteil von Curricula bilden Zulassungsvoraussetzungen oder -beschränkungen. Die Studiengänge können grundsätzlich in solche mit starken und solche mit schwachen Zulassungsverfahren unterschieden werden, die mit deren Positionierung im akademischen Feld korrespondieren. So verdeutlicht etwa die ETHZ durch ihre Zulassungsvoraussetzungen die »wissenschaftlich-technische Ausrichtung« des Masters in Datenwissenschaften: Studienanfänger*innen benötigen dafür einerseits einen Bachelorabschluss in den Studienbereichen Informatik, Elektro- oder Maschineningenieurwissenschaften, Mathematik oder Physik. Andererseits müssen zusätzlich die »fachlichen Voraussetzungen« mit spezifischen Kompetenzen in Statistik und Mathematik sowie Informatik (jeweils 30 Kreditpunkte) sowie »fachspezifische Kenntnisse und Fertigkeiten in den Bereichen Datenanalyse und Datenmanagement« (16 Kreditpunkte) nachgewiesen werden, auf denen die Kursinhalte im Master aufbauen.² An der EPFL werden nur Studierende mit einem Bachelorabschluss des federführenden Departements für Computer Science und Communication Systems zugelassen, während bei Interessent*innen mit Abschlüssen in Engineering, Mathematik oder Physik eine Prüfung *sur dossier* erfolgt.³ Aufgrund der hohen Nachfrage von Studieninteressierten werden zudem an beiden ETH die Bewerbungsdossiers individuell geprüft und nach Herkunftsuniversitäten sowie Noten ausgewählt. Die Zulassungsvoraussetzungen führen somit neben einer disziplinären Engführung auf bestimmte technik- und naturwissenschaftliche Studienfelder, die als »Kern« von »Data Science« betrachtet werden, auch zu einer sozialen Schliessung des Feldes, indem sich die – in Rankings – führenden Universitäten in erster Linie aneinander orientieren.

Die meisten Universitäten haben niedrigere Zugangshürden. Ausnahmen sind die USI sowie BENEFRI, die Bachelorabschlüsse in Informatik oder entsprechende Äquivalente verlangen. In der Regel befähigt ein Bachelorabschluss in einem quantitativen Fach der Technik-, Natur- oder Wirtschaftswissenschaften zur Aufnahme des Masterstudiums, wobei bestimmte Studienleistungen wie fortgeschrittene Statistikveranstaltungen sowie ein Einführungskurs in Programmierung (und gegebenenfalls Datenbanken) als Voraussetzung genannt werden. So ist etwa an der Universität Zürich explizit kein spezifischer Bachelorabschluss notwendig, um das Masterstudium beginnen zu können:

-
- 2 Im Fachbereich Mathematik und Statistik müssen Veranstaltungen in Analysis, Lineare Algebra, Numerische Methoden sowie Wahrscheinlichkeit und Statistik erfolgreich absolviert worden sein. Im Fachgebiet Informatik sind dies Algorithmen und Datenstrukturen, Datenbanken und Datenmodellierung, Komplexitätstheorie sowie Programmierung. Vgl. dazu den Anhang zum Studienreglement 2017 für den Masterstudiengang Data Science vom 18. Oktober 2016, S. 2f. Online: <https://rechtssammlung.sp.ethz.ch/Dokumente/324.1.1600.20.pdf> (Zugriff: 03.02.2022).
 - 3 Ausgewählte Informatik- und Mathematikvorlesungen (wie »Functional programming«, »Introduction to Machine Learning« oder »Introduction to database systems«) sind verbindlich für die Zulassung zum Masterstudium vorausgesetzt bzw. müssen während des ersten Mastersemesters nachgeholt werden. Vgl. den Art. 4 – Conditions d'admission im Règlement d'application du contrôle des études de la section de systèmes de communication pour le master en data science pour l'année académique 2019–2020 du 21 mai 2019. Online: https://academique.epfl.ch/plansEtude/2019/SC_DSreg2019-2020.docx (Zugriff: 03.02.2022).

»Das ist vielleicht das Spezielle an diesem Data Science Master: Wir machen keine Einschränkung, welchen Bachelor sie haben. [...] Was wir machen ist: wir geben Auflagen. [Es] gibt die Auflage, dass die Studierenden Programmieren können müssen, weil sie sonst die Inhalte in den Veranstaltungen nicht verstehen würden. Mit einem Kurs in Programmierung und Datenbanken werden sie aufgenommen, ansonsten müssen sie dies im Verlauf des ersten Jahres bestanden haben« (Prof_UH).

Die schwachen Zulassungsbedingungen korrespondieren mit einem höheren Anteil an Wahlfächern in universitären Studienprogrammen, als dies an den ETH oder den Fachhochschulen der Fall ist. Etwas anders gelagert ist das »Data Science Fundamentals« an der Universität St. Gallen, das lediglich für interne Studierende zugänglich ist. Das Lehrprogramm wählt interessierte Studierende der existierenden Fächer auf Bachelorstufe aufgrund von Motivationsschreiben sowie auf Basis der Noten im Einführungsjahr aus.⁴

An den Fachhochschulen erfordern die Masterstudiengänge keine spezifischen Bachelorabschlüsse für eine Studienzulassung. Beispielsweise kennt der Master in Applied Information and Data Science der HSLU ein stark individualisiertes Prüfungsverfahren mittels Self-Assessment-Test, den Nachweis ausreichender Englischkenntnisse sowie gegebenenfalls eines persönlichen Eignungsgesprächs.⁵ Auf BA-Stufe erfordern der Bachelor Data Science der FHNW oder der Bachelor Artificial Intelligence & Machine Learning der HSLU eine Berufsmaturität technischer oder kaufmännischer Richtung, eine gymnasiale Maturität oder Fachmaturität mit jeweils einjähriger Arbeitswelterfahrung.⁶ Inhaltlich-fachspezifische Studienvoraussetzungen existieren bei diesen Studiengängen keine.

In der Einrichtung von starken bzw. schwachen Zulassungsbedingungen manifestieren sich Grenzziehungen, die auf die Hierarchisierung bestimmter Positionen im akademischen Feld zurückzuführen sind. Hochschulen wie die ETH haben aufgrund ihres internationalen Prestiges selbst dann hohe Zahlen von Bewerber*innen, wenn sie die Offenheit und Multidimensionalität der Studiengänge – und damit der Datenwissenschaften insgesamt – sowohl aufgrund der disziplinären als auch der sozialen Voraussetzungen stark einschränken. Demgegenüber sinken die Möglichkeiten einer selektiven Auswahl für jene Institutionen mit niedrigerem Status im akademischen Feld, deren Studienangebote sich primär an lokal orientierte Studieninteressierte sowie die regionalen Arbeitsmärkte richten. In diesem Sinne organisieren Zulassungsvoraussetzungen die Curricula, indem sie bestimmte Studierendensubjekte ein- oder ausschliessen. Sie bilden die Grundlage, damit die nachfolgenden beschriebenen Elemente der curricularen Strukturlogik ihre Effekte auf die Vermittlung der Datenwissenschaften entfalten können.

4 Vgl. <https://www.unisg.ch/en/studium/bachelor/zusatzabschluesse/datasciencefundamentals/application-procedure-regulations> (Zugriff: 03.02.2022).

5 Vgl. <https://www.hslu.ch/de-ch/wirtschaft/studium/master/applied-information-and-data-science/zulassung/> (Zugriff: 03.02.2022).

6 Vgl. <https://www.fhnw.ch/de/studium/technik/data-science> sowie <https://www.hslu.ch/de-ch/informatik/studium/bachelor/artificial-intelligence-and-machine-learning/> (Zugriffe: 25.11.2020).

9.3 Kanonisierungsprozesse im Kernbereich

Die Studieninhalte im Kernbereich repräsentieren zentrale Werte und Normen eines Wissensfeldes und schaffen damit die Bedingungen für Forschungs- und Lehrpraktiken, über die sich (inter-)disziplinäre Gemeinschaften reproduzieren. Entsprechend wichtig ist die Zusammensetzung des Kernbereichs für die Ausrichtung und Positionierung von Studiengängen. Datenwissenschaftliche Curricula explizieren ihre Interdisziplinarität oft in der Verknüpfung von Wissensbeständen und Methoden im Kernbereich mit einer fachlichen Vertiefung in Anwendungsfeldern. Die relativen Anteile der verschiedenen Disziplinen des Kernbereichs bilden folglich eine zentrale Vergleichsdimension, da sie die interdisziplinäre Ausprägung eines Curriculums indizieren. Ich untersuche deshalb im Folgenden die inhaltliche Prägung des Kernbereichs am Beispiel der Masterstudiengänge in Data Science an den beiden ETH und gehe anschliessend auf die Bedeutung der Statistik ein (Kap. 9.4).

An der ETHZ ist der Studiengang entlang der Schwerpunkte der drei involvierten Departemente für Informatik, Mathematik und Elektrotechnik aufgebaut:⁷ Der Kernbereich ist durch die Module *Data Management and Processing* sowie *Data Analysis* (das wiederum die zwei Teile *Information & Learning* sowie *Statistics* umfasst) strukturiert, in denen jeweils mindestens eine Vorlesung belegt werden muss (mind. 32 ECTS). Hinzu kommen Wahlfächer (*core electives*) im Kernbereich (mind. 10 ECTS). Das heisst, über ein Drittel aller zu belegenden ECTS-Punkte müssen in diesem statistisch-technologischen Kernbereich belegt werden. Ein Befragter begründet den Aufbau des Masterprogramms, indem er die drei Schwerpunkte im Kernbereich direkt mit den Kompetenzanforderungen an Data Scientists verknüpft:

»Ein Data Scientist muss verstehen, wie die Daten gespeichert werden, wie die Daten abgefragt werden und darauf aufbauend Algorithmen ausgeführt werden, um daraus Informationen zu generieren. Dazu braucht es sehr viele theoretische Grundlagen in Informatik, Statistik, Mathematik und Wahrscheinlichkeitstheorie« (Prof_ ETH_B: 2:2).

An der EPFL verantwortet die *School of Computer and Communication Sciences* den Studiengang alleine.⁸ Die EPFL bietet den Studierenden mehr Wahlfreiheit als die ETH, indem nur ein Viertel aller Punkte (mind. 30 ECTS) im Kernbereich absolviert werden muss. Darüber hinaus existiert im Kernbereich keine weitergehende modulare Einteilung. Trotz dieser Unterschiede bezüglich Studienstruktur und fakultärer Einbettung ist eine grosse inhaltliche Überschneidung der beiden Studiengänge feststellbar (vgl. Tabelle 7).

7 Vgl. <https://inf.ethz.ch/studies/master/master-ds.html> (Zugriff: 03.02.2022).

8 Vgl. <https://www.epfl.ch/education/master/programs/data-science/> (Zugriff: 03.02.2022).

Tabelle 7: Curricula der Masterstudiengänge in Data Science der ETHZ und EPFL

Master in Data Science ETHZ		Master in Data Science EPFL	
Kernbereich (min. 32 ECTS)	ECTS	Kernbereich (min. 30 ECTS)	ECTS
Data Management and Processing (min. 16 ECTS)			
Advanced Algorithms	8	Advanced Algorithms	7
Big Data	8	Systems for Data Science	6
Optimization for Data Science	8	Optimization for Machine Learning	4
Data Analysis: Information & Learning			
Mathematics of Information	8	Information Theory and Signal Processing	6
Advanced Machine Learning	8	Machine Learning	7
Data Analysis: Statistics (min. 16 ECTS)			
Fundamentals of Mathematical Statistics	10	Statistics for Data Science	6
Computational Statistics	10	Applied Data Analysis	6
		Information security and privacy	6
+ Core Electives + Interdisciplinary Electives	>10 8-12	+ Optional courses (+ Optional minor)	~42 ~30)
+ Data Science Lab + Seminar + Science in Perspective	14 2 2	+ Semester project in data science + Project in social and human sciences	12 6
Master Thesis	30	Master Thesis	30
Total	120	Total	120

Die Kurse tragen zwar teilweise unterschiedliche Labels, können aber aufgrund der Kursbeschreibungen als komplementär bis identisch bezeichnet werden. Einen Unterschied im Kernbereich bildet der Kurs »Information security and privacy« an der EPFL, der gemäss Ausschreibung die informationstechnologischen Aspekte von Sicherheit und Privatsphäre sowie ethische Fragen abdeckt. Ein ähnlicher Kurs findet sich unter dem Titel »Big Data, Law, and Policy« an der ETHZ im wählbaren Kernbereich, wobei die Veranstaltung gemäss Programm mehr die gesellschaftlichen Implikationen von Big Data fokussiert (vgl. Kap. 9.7).

Insgesamt ist trotz der zahlreichen Überschneidungen das Angebot im Kernbereich der EPFL etwas weniger auf mathematisch-statistische, dafür stärker auf computerwissenschaftliche Grundlagen ausgerichtet. Die Analyse der beiden Curricula verweist auf Kanonisierungsprozesse im Kernbereich, d. h. auf die Herausbildung eines methodisch-technischen Kerns, der sich primär auf Wissensbestände der Statistik, Mathematik und Informatik abstützt. Für den spezifischen Hochschultyp ETH können die untersuchten Curricula insofern als Versuche der »Disziplinierung« (Schultheis 2005) der Datenwissenschaften gelesen werden, indem sie diese einerseits von anderen Wissensfeldern wie Informatik oder Statistik differenzierbar machen und somit Autonomie im Sinne einer Disziplin begründen. Andererseits führt der

Prozess der Disziplinierung aber auch zu einer »Unterwerfung unter Regeln« (ebd.), da sie eine spezifische curriculare Struktur und Inhalte verknüpft, die fortan als bestimmend für das Wissensfeld repräsentiert und interpretiert werden.

Bemerkenswert ist die Etablierung des datenwissenschaftlichen Lehrkanons an den ETH auch insofern, als die Studieninhalte in einem mehrjährigen Vorbereitungs- und Planungsprozess erarbeitet und festgelegt wurden – also zu einem Zeitpunkt, bevor die grundlegende Frage beantwortet wurde, ob und inwiefern es sich bei der neuen Kategorie »Data Science« überhaupt um eine eigenständige wissenschaftliche Disziplin handelt (NASEM 2018: 2–1). Während in zahlreichen wissenschaftlichen Diskussionen und Beiträgen immer noch um diese Frage gerungen wird, indizieren die beobachteten Kanonisierungsprozesse im Kernbereich demnach eine beschleunigte Ausarbeitung und Implementierung der Curricula auf organisationaler Ebene der beiden ETH.

9.4 Zur Bedeutung der Statistik im Curriculum

Eine zentrale Dimension für die Analyse datenwissenschaftlicher Curricula sind die relativen Anteile der verschiedenen Disziplinen, also die erforderlichen Studienleistungen in Informatik, Statistik, Mathematik und in weiteren Fächern. In der Ausrichtung und Zusammensetzung interdisziplinärer Curricula äussern sich jene disziplinären Konfliktlinien, die die Datenwissenschaften im akademischen Feld seit Beginn begleiten. Besondere intensive Diskussionen charakterisieren dabei das Verhältnis von Datenwissenschaften und Statistik, was auf eine widersprüchliche Wahlverwandtschaft hinweist: Einerseits beklagen manche Statistiklehrenden den Verlust der eigenen Fachidentität sowie eine Usurpation ihres Anspruchsgebiets durch andere Disziplinen, und hier insbesondere die Informatik. Andererseits manifestiert sich darin zugleich eine Universalisierung statistischer Expertise unter neuen Bezeichnungen, die Anwendung auf zahlreiche Felder findet. Ich untersuche solche Widersprüche im Folgenden anhand des Umfangs und der Bedeutung, die Statistikinhalte in den Curricula einnehmen.

An den beiden ETH mit ihrer starken Fokussierung auf die »methodologischen Grundlagen« zählen je zwei Statistikkurse explizit zum Kernbereich: An der ETH Zürich ist dies das Modul »Data Analysis: Statistics« mit den zwei Vorlesungen »Fundamentals of Mathematical Statistics« sowie »Computational Statistics« (davon muss mindestens ein Kurs erfolgreich abgeschlossen werden); an der EPFL sind die grundlegenden Kurse »Statistics for Data Science« sowie »Applied Data Analysis« Teil des Kernbereichs. Auch andere Studiengänge markieren statistische Methoden und Verfahren als zentral. Der Master in Computational Science an der USI umfasst einen einführenden Kurs »Introduction to Statistics« im Wahlbereich, der die notwendigen statistischen Grundlagen bietet für fortgeschrittene Kurse wie »Data Analytics«, »Stochastic Methods« oder »Simulation & Data Science«.⁹ Die GSEM der Universität Genf fokussiert mit ihrem »Research Center for Statistics« statistische Modellierungen und Analysemethoden, was einen der wissenschaftlichen Schwerpunkte im Master Business Analytics ausmacht. Die wirtschaftswissenschaftliche Fakultät der Universität

9 Vgl. <https://www.usi.ch/en/node/5558> (Zugriff: 03.02.2022).

Zürich, die die Vertiefung Data Science im Master Informatik verantwortet, bietet jährlich verschiedene Statistikkurse an.¹⁰ Der Master »Statistics and Data Science« der Universität Bern ist eine eigentliche Erweiterung des bestehenden Masterprogramms in Statistik. Schliesslich weisen auch die Fachhochschulen in ihren Masterstudiengängen jeweils eine fundierende Veranstaltung in statistischer Datenanalyse aus.

Demgegenüber weisen Studienangebote, in denen »Data Science« eine Vertiefung innerhalb eines Masters (vor allem in Computerwissenschaften, aber auch in Management oder Finance) darstellt, oft keine gesonderten Kurse in Statistik auf: So tragen beispielsweise weder das Kursangebot der Vertiefung Data Science an den BENEFRI-Universitäten noch die einzelnen Kursbeschreibungen explizit Statistikinhalte im Titel – obwohl statistische Methoden wie etwa Clustering, Support Vector Machines, Machine Learning oder neuronale Netzwerke in verschiedenen Veranstaltungen gelehrt werden.¹¹ Dasselbe gilt ferner auch für den Master in Information Systems oder die Vertiefung Financial Entrepreneurship & Data Science im Master of Finance an der Universität Lausanne. Obwohl alle drei Programme mehrere Veranstaltungen mit Bezeichnungen wie Big Data Analytics, Machine Learning oder Business Analytics umfassen, bleibt auch hier Statistik grösstenteils ungenannt.

Die Analyse zeigt, dass Statistik jeweils in jenen Universitäten und Hochschulen Bestandteil des Curriculums ist, in denen Professuren oder Institute für Statistik oder Mathematik existieren und an der Planung und Konzeption der Curricula beteiligt waren. Dies ist an den beiden ETH sowie an technick- oder naturwissenschaftlichen Fakultäten von Universitäten der Fall. An Business Schools werden (mit Ausnahme der GSEM der Universität Genf) Themen- und Methodenfelder wie Data Analytics, Machine Learning oder neuronale Netzwerke nicht explizit mit Statistik in Verbindung gebracht, sondern von Lehrenden anderer Disziplinen, vorwiegend der Informatik sowie der Wirtschaftswissenschaften angeboten – in der Regel von Professuren, deren Denomination Datenanalyse enthält.¹²

Die Stellung der Statistik in den untersuchten Curricula präsentiert sich als ambivalent: Da nur wenige Schweizer Universitäten und Hochschulen eigenständige Statistikdepartemente kennen, wird die Mehrheit der untersuchten Studiengänge durch die Informatik oder die Wirtschaftswissenschaften verantwortet und organisiert. Der Statistik kommt dabei – mit wenigen Ausnahmen – die Rolle als Juniorpartnerin, etwa bei Kursen zu Statistik und Datenanalyse, zu. Trotzdem werden Statistikinhalte in praktisch allen Studiengängen gelehrt, wenn auch teils unter anderen disziplinären Bezeichnungen. Insofern zeichnet sich weniger ein relativer Bedeutungsverlust ab als vielmehr die Persistenz der strukturell schwachen Stellung der Statistik im disziplinären Gefüge. Durch die disziplinäre Engführung der Datenwissenschaften auf einzelne Fachkombinationen (vor allem in der Form von Informatik und Wirtschafts-

10 Die Suche nach Modulen und Veranstaltungen erfolgt über das Vorlesungsverzeichnis der Universität Zürich. Online: <https://studentservices.uzh.ch/uzh/anonym/vvz/index.html#/search> (Zugriff: 03.02.2022).

11 Vgl. die Kursliste und Kurzbeschreibungen der Vertiefung Data Science an den BENEFRI-Universitäten. Online: <https://mcs.unibn.ch/program/courses-timetable/course-list/> (Zugriff: 03.02.2022).

12 Umgekehrt erfordern Vertiefungen an Business Schools, die Management und Data Science kombinieren, in der Regel keine Kurse in Datenbankmodellierung und nur Basiskenntnisse im Programmieren.

wissenschaften) reartikulieren die Curricula die epistemologischen und disziplinären Auseinandersetzungen um die Deutungshoheit über die Expertise in der Datenanalyse. Durch die Implementierung in die jeweiligen Curricula werden die Konfliktlinien nicht nur in organisationale Gefässe überführt, sondern bilden auch die Grundlage für Lehr- und Forschungspraktiken und somit die Reproduktion etablierter Disziplinen unter neuen Bezeichnungen.

9.5 Wahlfächer: Die Übersetzung der Logik der Domänen in die Curricula

Ein weiteres wichtiges Element der Curricula bilden die Wahlbereiche in den Curricula. Die Wahlfächer dienen dazu, die universellen theoretischen und methodischen Wissensbestände, die sich die Studierenden in den Kernbereichen angeeignet haben, mit inhaltlicher Expertise in bestimmten Domänen, der sogenannten *domain knowledge*, zu erweitern (vgl. Kap. 7.5). Die Art und Weise, wie Kern- und Wahlbereich curricular miteinander verknüpft werden, unterscheidet sich dabei je nach Hochschultyp, fakultärer und inhaltlicher Ausrichtung: In den Studienprogrammen der technischen Hochschulen zeigt sich die proklamierte Interdisziplinarität in der Kombination des Kernbereichs von »Data Science« mit spezifischen technik- und naturwissenschaftlichen Anwendungsfeldern im Wahlbereich. An der EPFL können die Studierenden einen Minor (30 ECTS) eines anderen Departements¹³ belegen oder die Punkte ohne Vertiefung aus dem Wahlbereich (bis zu 42 ECTS) des Masterstudiengangs auswählen. An den ETHZ erfolgt die Vertiefung nicht in einem gesonderten Nebenfach, sondern ist curricular in den Masterstudiengang eingebettet: Studierende wählen im Rahmen des Moduls »Interdisciplinary Electives« Kurse im Umfang von 8–12 ECTS in einem wissenschaftlichen Themengebiet (vgl. Tabelle 7).¹⁴ Bei den Wahlangeboten handelt es sich – mit einzelnen Ausnahmen – um themenzentrierte, interdisziplinäre Wissensbereiche, die sich kaum auf einzelne Fächer reduzieren lassen. Vielmehr sind es stark forschungsorientierte Themengebiete, die auf Masterstufe angeboten werden, aber keine eigene departementale Entsprechung haben und sich eng an ausseruniversitären Akteur*innen, primär in der Industrie, orientieren (Abbott 2005: 265). Bei BENEFRRI wird »Data Science« selbst zur Spezialisierungsmöglichkeit innerhalb des Masters Informatik, d. h., hier fehlt eine inhaltliche Domäne. An anderen technikwissenschaftlich ausgerichteten Studienprogrammen wie der USI erfolgt die Wahl freier Leistungen im übrigen Studienangebot der verantwortlichen Institute bzw. Fa-

13 Zur Auswahl stehen Biocomputing; Biomedical Technologies; Computational Science and Engineering; Management, Technology and Entrepreneurship oder Space Technologies (EPFL 2017: 4).

14 Gegenwärtig werden Kurse in Computational Biology; Bioinformatics, and Biomedicine; Computer Networks; Finance and Insurance; Geographic Information Systems; Law, Policy, and Innovation; Neural Information Processing; Social Networks; Transport Planning and Systems sowie Weather and Climate Systems angeboten. Für jeden Schwerpunkt werden Veranstaltungen sowohl auf Basis als auch auf fortgeschrittenem Niveau angeboten, wobei in der Regel mindestens zwei einführende Kurse pro Vertiefung empfohlen werden. Vgl. <https://www.inf.ethz.ch/studies/master/master-ds/interdisciplinary-electives.html> (Zugriff: 03.02.2022).

kultäten, wobei es sich meist um weitere technologie- oder methodenzentrierte Kurse handelt.

Studiengänge und Vertiefungen an Business Schools sind in der Regel so strukturiert, dass sie die technisch-wissenschaftlichen Methodenveranstaltungen im Kernbereich mit domänenspezifischen Inhalten in Economics, Finance, Management, Marketing oder Personalentwicklung kombinieren. An der Universität Zürich erfolgt eine fachliche Vertiefung in einem Nebenfach sowohl innerhalb wie ausserhalb der Fakultät. Umgekehrt kann »Data Science« selbst zu einer methodischen Vertiefung und einem Nebenfach werden. Der Master Business Analytics der Universität Genf bietet im Wahlmodul die Möglichkeit einer praxisorientierten Business- (inkl. Internship) oder einer Forschungsorientierung, in denen Veranstaltungen aus dem wirtschaftswissenschaftlichen Kursangebot der Fakultät belegt werden. In den Masterstudiengängen der Fachhochschulen bestehen Vertiefungs- und Spezialisierungsoptionen in gesonderten Wahlpflichtmodulen: So ermöglichen an der HSLU die Wahlpflichtmodule den Studierenden, »ein eher Business, Technik oder Analytik orientiertes Profil auszubilden«. Unter »Domain Experience« werden Veranstaltungen in Bereichen wie Marketing, Mobilität, Finance, Energiesysteme und IoT, Gesundheit, Tourismus oder Sport offeriert.

Die Breite der angebotenen Wahlbereiche und Wahlfächer verdeutlicht die thematische Offenheit und damit Universalität datenwissenschaftlicher Anwendungen und Methoden. Es existiert kaum noch ein Feld – ob wissenschaftlich oder nicht –, das nicht zu einer Anwendungsdomäne geworden ist. Die Vielfalt der angebotenen Wahl- und Vertiefungsbereiche spiegelt sich in den untersuchten Curricula wider, die somit Anschlussfähigkeit über den eigentlichen Kernbereich hinaus ermöglichen. Die Überführung der Logik der Domänen (Ribes et al. 2019) in die Organisationsstruktur datenwissenschaftlicher Curricula erlaubt es somit, die Limitationen, die sich durch Disziplinierung und innerwissenschaftliche Konfliktlinien vor allem im Kernbereich ergeben (vgl. Kap. 9.3), zu neutralisieren. Da es sich bei den angebotenen Wahlfächern oft um Gegenstände und Themenfelder handelt, die sich nur in der Minderheit als eigenständige Disziplin ausdifferenziert haben, können rasch und ohne Gefahr neuer Konflikte solche Wahlmöglichkeiten aufgebaut, adaptiert oder aufgegeben werden. Dies erlaubt es den Studiengängen, unmittelbar auf veränderte Themenkonjunkturen zu reagieren, ohne den politischen Restriktionen zu unterliegen oder latente – organisationale wie disziplinäre – Konfliktlinien zu aktivieren.

9.6 Praxisorientierte Kurse zur Überschreitung feldspezifischer Differenzierungen

Eine zentrale Komponente der untersuchten Curricula ist die Verknüpfung der wissenschaftlich-technischen Kern- und Wahlfächer mit sogenannten praxisorientierten Formaten. Diesen Kursen wird besonders das Potenzial attribuiert, die entstehende Lücke zwischen der Vermittlung von theoretischem und methodischem Wissen und dessen Anwendung in konkreten Projekten mit »echten Daten« und Fragestellungen

überbrücken zu können (Ridsdale et al. 2015: 20; NASEM 2018).¹⁵ Bezüglich Kursaufbau und -inhalten äussert sich dies je nach Hochschultyp und Studienprogramm auf unterschiedliche Weise: Häufig zu beobachten sind sowohl individuelle als auch Gruppenprojekte, in denen Studierende innerhalb der Universitäten und Hochschulen über längere Zeit (drei bis sechs Monate) zu bestimmten Fragestellungen mit Datensätzen arbeiten, die als unstrukturiert, »messy« oder »real-world« beschrieben werden. Beispiele dafür sind etwa das »Data Science Lab« an der ETH Zürich, das »Semester Project in Data Science« an der EPFL oder das »Master's project« an der Universität Zürich, die Bestandteil von Regelmodulen sind.

Neben Labs, Industrieprojekten und Praktika werden zunehmend auch Bootcamps, Hackathons und *hack weeks* Teil der Hochschulbildung in den Datenwissenschaften (Huppenkothen et al. 2018). An der Universität St. Gallen ist ein zweiwöchiges Bootcamp Bestandteil des aussercurricularen Lehrprogramms »Data Science Fundamentals«, das parallel zum regulären Studium absolviert wird. Bootcamps und Hacking Events erfüllen sich zunehmender Beliebtheit in der Vermittlung von »Data Science« in Unternehmen, Hochschulen und anderen Organisationen, da sie mit einer raschen Immersion in die Wissensbestände, Methodologien und Praktiken von »Data Science« assoziiert werden (NASEM 2018: 3–14ff.). Ein Professor der Universität St. Gallen skizziert den Kurs denn auch als eine Einführung in die spezifischen Arbeits- und Herangehensweisen von Data Scientists:

»Wir machen ein zweiwöchiges Bootcamp mit Data Science und Machine Learning [und] werfen sie von Anfang an ins kalte Wasser: Daten sind messy, Daten können nicht eingelesen werden, Fragestellungen müssen gefunden werden usw. Bereits nach 2–3 Tagen beginnen sie die Arbeit von Data Scientists zu verstehen. Dann beginnen sie mit der Ausarbeitung eines eigenen Projektes mit Fragestellungen, Daten und Methoden« (Prof_UH).

Die Ausgangslage und Arbeitssituation von Data Scientists wird in solchen Formaten wie im Zitat als eine grundlegend chaotische präsentiert: Daten sind zwar (im Überfluss) vorhanden, können aber aufgrund von Problemen bei der Prozessierung nicht verarbeitet werden. Ausserdem ist vielfach gar nicht klar, worum es inhaltlich überhaupt gehen soll, da die Fragestellungen noch gar nicht bekannt sind. Die Beschreibung lehnt sich an jenen Modus der Wissensproduktion an, der nicht Theorien oder Modelle, sondern Daten als Ausgangspunkt nimmt (Anderson 2008): »Big Data erfordert oft erst das Finden von neuen, innovativen Fragestellungen, was nur über Anwendungsorientierung und konkrete Projekte möglich ist« (Prof_ETH_A: 4:25). Der Norm zur Anwendungsorientierung entsprechend wird das Bild der »echten Daten« bzw. »echten Probleme« evoziert (vgl. Kap. 7.6), etwa indem auf die »Messyness« von Daten oder die Inkompatibilität bestimmter Formate hingewiesen wird. Ein spezifisches Praxiswissen ergibt sich hier durch die Anforderungen an Data Scientists, die sich immer wieder von neuem mit einer strukturellen Orientierungslosigkeit (»Wir [...] werfen sie von Anfang an ins kalte Wasser«) konfrontiert sehen.

15 Die Vorbilder für solche Kurse werden in den USA verortet, wo sogenannte Capstone Projects in sehr unterschiedlichen Studiengängen existieren. Auch curriculare Empfehlungen betonen oft die Bedeutung von Capstone Projects (NASEM 2017: 2–5; De Veaux et al. 2017: 2.8).

Die Curricula entwerfen praxisorientierte Kurse als das zentrale »Scharnier« zwischen den wissenschaftlich-theoretischen Kursen im »Kernbereich« und den spezifischen Anwendungsdomänen. Die Verknüpfung von datenwissenschaftlicher Problembearbeitung »von Anfang bis zum Ende«, »Teamarbeit« und Interdisziplinarität (»Leuten mit unterschiedlichen Hintergründen«) bilden die Ingredienzen der Praxis. Diese wird stets als kollaborativ und dynamisch gedacht. Sie findet gewissermassen quer zu den disziplinären Gefässen statt, die Universitäten und Hochschulen kennzeichnen, und transzendiert diese gleichsam.

Daneben pflegen alle untersuchten Institutionen Kollaborationen im Bereich der Lehre mit Organisationen ausserhalb der Hochschulen. Die EPFL schreibt ein obligatorisches »Ingenieurspraktikum« bei einem Industrieunternehmen vor, das mindestens acht Wochen bis zu sechs Monaten dauert. Die Ziele dieses Praktikums lauten wie folgt:

»As a Master student, it is mandatory to undertake an engineering internship in a company during your studies. This internship aims to facilitate your immersion in the professional world, aware you of the teamwork and familiarize yourself with the industry procedures.«¹⁶

Die Ziele des Praktikums sind quasi deckungsgleich mit der Selbstbeschreibung des oben erwähnten Data Science Labs, mit dem Unterschied, dass dieses in einem industriellen Kontext stattfindet. Es verortet insofern die zukünftige professionelle Praxis von Data Scientists explizit ausserhalb der Universität. Auch andere universitäre Masterstudiengänge (wie Information Systems der Universitäten Neuchâtel und Lausanne sowie der Master in Business Analytics an der Universität Genf) rechnen Praktika als Studienleistungen an. Es besteht überdies an allen untersuchten Universitäten und Hochschulen die Option, Masterarbeiten oder Studierendenprojekte in Kollaboration mit Unternehmen zu verfassen, die von Professor*innen der jeweiligen Institution betreut werden.

Solche feldübergreifenden Kollaborationen werden als »Win-win-Situationen« gerahmt: Die Studierenden in regulären Studiengängen erhielten Einblicke in organisationale Praktiken und könnten Kontakte zu potenziellen Arbeitgebern knüpfen. Der Vorteil für die Unternehmen bestünde in einem exklusiven Zugang zu Wissensbeständen und Methoden an Hochschulen und Universitäten, die *State of the art* seien und die sie über »ihre« Studierenden abholen könnten. Umgekehrt liege der Vorteil für Hochschulen darin, über spezifische Kollaborationen in die Praxis von Organisationen eingebunden zu werden und damit Einsicht in unternehmensinterne Problemstellungen, Prozesse und gegebenenfalls auch Daten zu erhalten. Die Studierenden in solchen Arrangements nehmen die Rolle von *boundary crossers* ein und festigen damit die transversalen Relationen zwischen den verschiedenen Feldern.

Die fokussierte Praxisorientierung der untersuchten Studiengänge garantiert die enge Anlehnung an ausseruniversitäre, industrielle Verfahren und Wissensbestände. Sie stützt sich auf instrumentalistische Vorstellungen, wie sie insbesondere die Engineering-Ausbildung und die Technikwissenschaften (Malazita 2019; Cech 2014) insgesamt kennzeichnen: Die Relevanz bestimmter Wissensformationen, die sich

16 Vgl. <https://www.epfl.ch/schools/ic/internships/> (Zugriff: 03.02.2022).

wie im Falle der Datenwissenschaften vor allem auf quantitative, technologische und metrikenbasierte Instrumente, Messverfahren und Methoden abstützen, ergibt sich demnach erst durch Applikationen in der »Praxis«. Die »echten«, »realweltlichen« Probleme, die damit adressiert werden, werden dabei zumeist ausserhalb der akademischen Forschung situiert. Dies ermöglicht es vor allem den technischen und Fachhochschulen, aber auch Universitäten, Forderungen nach Wissenstransfer, Innovation und Wettbewerbsfähigkeit, die im politischen Diskurs artikuliert werden, einzulösen.

9.7 Das Prinzip der Komplementarität

Schliesslich umfasst ein Teil der Studienprogramme neben Kern- und Wahlbereich sowie praxisorientierten Kursen sogenannte »komplementäre Inhalte«. Dazu werden verschiedene Kurse, insbesondere zu rechtlichen, sicherheitsbezogenen, ethischen und ökonomischen Aspekten und Fragestellungen von Systemen und Praktiken der Datenverarbeitung sowie deren Implementation in Produkte und Dienstleistungen gezählt. Die Rahmung als »komplementär« verweist darauf, dass diese Inhalte als Ergänzung zu den methodisch-wissenschaftlichen Praktiken, Werkzeugen und Grundkompetenzen in statistischen Verfahren, Programmierung und Datenbankmodellierung betrachtet werden. Dieses Prinzip referenziert ingenieurwissenschaftliche Ausbildungstraditionen, die das Soziale weitgehend ausserhalb von Technologie und Wissenschaft verorten (Malazita 2019): Geistes- und sozialwissenschaftliche Studieninhalte werden in dieser Traditionslinie als ergänzende (eben »komplementäre«) Massnahmen zur Ausbildung von kommunikativen Fähigkeiten, einer reifen, entscheidungsfähigen Persönlichkeit oder kultureller Vervollständigung betrachtet (Wisnioski 2009; Paulitz & Prietl 2017) – eine Vorstellung, die auch in schweizerischen Ingenieurstudiengängen und insbesondere an der ETHZ historisch verankert ist (Gurgerli et al. 2005; Honegger et al. 2007).

In ihrer zeitgenössischen Ausgestaltung stützt sich die geistes- und sozialwissenschaftliche Lehre an den ETH stärker auf die Befähigung zu Reflexivität, kritischem Denken sowie Kontextwissen, die in den Kompetenzprofilen von Data Scientists als zentral erachtet werden (vgl. Kap. 10).¹⁷ Sowohl das Modul »Wissenschaft im Kontext« an der ETHZ als auch die Veranstaltungen im Modul »Project in social and human sciences« an der EPFL bieten solche Studieninhalte für Masterstudierende in Data Science an den ETH (vgl. Tabelle 7). So findet sich die Vorstellung der Komplementierung von Data Scientists durch eine Art »Verantwortungsethik« (Weber 1993[1919]) auch prägnant in den Interviews wieder:

»Es ist sehr wichtig, einen Dialog führen zu können [mit] den Anwendern und möglichen Einsatzbereichen von Datenwissenschaften. Das versuchen wir zu stärken durch diese Lab-Komponente, andere Aspekte wie Teamarbeit werden natürlich gefördert, ja. Natürlich ein Bewusstsein, ein Verantwortungsbewusstsein zu schärfen für diese Inhalte; in den Kursen ist es natürlich wichtig, dass die Leute eine Sensibilität haben, wenn sie in den Arbeitsmarkt eintreten« (Prof_ETH_A: 4:24).

17 Vgl. <https://gess.ethz.ch/studium/science-in-perspective.html> (Zugriff: 03.02.2022).

Die komplementären Inhalte sollen das »Verantwortungsbewusstsein« schärfen und eine »Sensibilität« unter Studierenden für datenwissenschaftliches Arbeiten durch die Verknüpfung mit Projekt- und Teamarbeit schaffen. Somit werden auch ethische, rechtliche und soziale Aspekte durch die Brille der Anwendungs- und Praxisorientierung gerahmt, ohne die sie ihre Wirkung nicht entfalten könnten. Durch die Fokussierung auf die Praxis, die stets lokal gebunden ist, entfernt sich die Zuschreibung zukünftiger Data Scientists allerdings auch von dem Anspruch nach allgemeiner Reflexionskompetenz (vgl. oben). Die Konzeption vernachlässigt auch das partizipative Moment, das seit Jahrzehnten Bestandteil der Computer- und Technologiegestaltung ist (Kensing & Blomberg 1998; Muller & Kuhn 1993; Schuler & Namioka 1993). Zugleich schreibt der Befragte eine »Sensibilität« für komplementäre Inhalte primär jenen Absolvent*innen zu, die »in den Arbeitsmarkt eintreten«. Die Verantwortung von akademischen Data Scientists, die angesichts zahlreicher umstrittener Forschungskollaborationen nicht minder bedeutsam ist (Bird et al. 2016; Verma 2014; Watson 2014), bleibt in dieser Perspektive ungenannt.

Das Studienangebot im Bereich der »komplementären Inhalte« hängt unmittelbar mit der Grösse, der finanziellen Ausstattung sowie der thematischen Breite und Ausrichtung der verantwortenden Hochschulen und Universitäten zusammen. So können die ETH Zürich, die EPFL oder die Universität Zürich spezifische Vorlesungen anbieten, die in der Regel durch vertiefende Seminare begleitet werden. An kleineren Universitäten werden meist aus »Ressourcengründen« keine besonders ausgewiesenen Veranstaltungen zu rechtlichen oder ethischen Fragen durchgeführt; obwohl auch dort kein Zweifel besteht, dass sich »Data Scientists [...] mit den Themen auskennen [müssen]«. Eine häufige Strategie besteht insofern darin, die Inhalte »in andere Kurse [zu integrieren]« (Prof_UH_D: 1:11). Der Master in »Applied Information and Data Science« an der Hochschule Luzern – Wirtschaft ist insofern eine Ausnahme, da er als einziger Studiengang im Sample explizit ein verpflichtendes Modul in »Ethical and Legal Issues of Big Data« enthält – obwohl auch hier kritisch angemerkt wird, dass »[dies] eine Bedeutung [hat], die man gar nicht genug hoch einschätzen kann. Wir haben das jetzt in einem Modul abgebildet. Ich bin ziemlich sicher, dass dies nicht reichen wird« (Prof_FH).

Die Analyse des Kursangebots verdeutlicht auch, dass der Schwerpunkt der existierenden Veranstaltungen in technisch orientierten Studiengängen primär auf sicherheitsspezifischen und rechtlichen Themen (vor allem Datensicherheit und Datenschutz) und nur marginal auf ethischen Fragestellungen liegt. Die sozialen Implikationen von Datenpraktiken, Algorithmen und soziotechnischen Systemen wie Fragen sozialer Gerechtigkeit (Barocas et al. 2017; Green 2018), von Fairness (Greene et al. 2019; Zook et al. 2017) oder Repräsentation (Berman & Bourne 2015; Drosou et al. 2017) werden zwar in der Forschung an manchen Standorten adressiert.¹⁸ Die untersuchten Curricula reflektieren allerdings die Frage, wie sich solche Diskussionen produktiv in die Hochschullehre integrieren lassen (Bates et al. 2020; Fiesler et al. 2020; Malazita 2019; Mayer & Malik 2019), erst ansatzweise.

18 Die Diskussion findet unter den Stichworten Fairness, Accountability, Transparency, and Ethics (kurz: FATE) von Algorithmen, Data Science oder Machine Learning statt und hat sich mittlerweile zu einem Subfeld innerhalb der Computer- und Datenwissenschaften mit eigenen Konferenzen und Journals entwickelt (Selbst et al. 2019).

Trotz Rekonzeptualisierung des »komplementären« Teils der Curricula an technisch orientierten Hochschulen bleiben die Wahlveranstaltungen auf die Ausbildung »verantwortungsbewusst[er]« Praktiker*innen fokussiert. Während Aspekte von Informationssicherheit und Datenschutz in die untersuchten Curricula integriert und gelehrt werden, gilt dies für ethische und soziale Implikationen nur begrenzt. Durch die Verknüpfung mit sozialen und interaktiven Kompetenzen werden ergänzende Inhalte als anwendungsorientierte Fragestellungen gerahmt und vermittelt, die zukünftige Data Scientists zu verantwortungsbewussten Praktiker*innen komplementieren sollen. Bestehende disziplinäre und epistemische Grenzziehungen gegenüber geistes- und sozialwissenschaftlichen Perspektiven und Inhalten stehen somit einer Erweiterung und Öffnung etablierter Ausbildungstraditionen entgegen.

9.8 Das Subfeld der Weiterbildungsstudiengänge

Schliesslich gehe ich in diesem letzten Teilkapitel auf die Weiterbildungsstudiengänge ein. Diese stellen ein spezifisches Subfeld innerhalb der datenwissenschaftlichen Studiengänge dar, das sich grundlegend von den Regelstudiengängen an Universitäten, ETH und Fachhochschulen unterscheidet. Wie bereits erwähnt, waren berufsorientierte Weiterbildungen ab 2014 die ersten Studienangebote in Datenwissenschaften auf Tertiärniveau in der Schweiz (vgl. Kap. 8.6.1): An der ZHAW und der BFH existierten Certificates of Advanced Studies (CAS) Programme in statistischer Datenanalyse oder Business Intelligence. Diese wurden zunächst durch neue CAS in Big Data, Machine Intelligence oder Datenvisualisierung zu Diploma of Advanced Studies (DAS) ergänzt und schliesslich zu Master of Advanced Studies (MAS) erweitert. In der Zwischenzeit hat sich das Subfeld erheblich ausdifferenziert und es können an allen Fachhochschulen, den beiden ETH sowie mehreren Universitäten berufliche Weiterbildungen im Bereich Datenwissenschaften, Machine Learning oder künstliche Intelligenz absolviert werden (vgl. Tabelle 13 im Anhang).

Ein Spezifikum der Weiterbildungsstudiengänge liegt darin, dass es sich bei der grossen Mehrheit dieser Studierenden um bereits hochqualifizierte Personen handelt:¹⁹

»80 % der Leute haben bereits ein Hochschulstudium absolviert, manche haben einen Master oder sogar [ein] Doktorat in Physik oder Informatik. [Die] Leute sind überdurchschnittlich ausgebildet im Vergleich zu anderen MAS« (Prof_FH_A: 8:23).

Diese Einschätzung trifft sich mit den Ergebnissen anderer Studien und Erhebungen, wonach der Anteil Personen mit Masterabschlüssen und PhD in »Data Science« überdurchschnittlich hoch ist (Burtch 2018; H. D. Harris et al. 2013). Es verweist aber auch auf die sich verändernde Sozialstruktur der Profession der Data Scientists: Waren

19 Es existieren sehr unterschiedliche Zugangsvoraussetzungen für den Besuch der CAS: Neben einem Hochschulstudium bzw. einem mehrjährigen Praxisnachweis sowie einer Einführung in wissenschaftliches Arbeiten als Äquivalent werden beispielsweise in manchen CAS Programmierkenntnisse vorausgesetzt, während andere so konzipiert sind, dass diese im Laufe des Unterrichts erworben werden.

Weiterbildungen für hoch qualifizierte Berufstätige zunächst die einzigen direkten Ausbildungsmöglichkeiten, so haben sich in der Zwischenzeit rund zwei Dutzend Regelstudiengänge in der Schweizer Hochschullandschaft etabliert. Es ist deshalb davon auszugehen, dass sich Alter, Lebens- und Arbeitserfahrungen der zweiten Generation von Data Scientists erheblich von jenen der ersten Generation unterscheiden werden.

Weiterbildungsstudiengänge sind sehr stark nachfrageorientiert, können relativ rasch aufgebaut und implementiert werden und sind entsprechend von hoher Volatilität gekennzeichnet. Da die Studiengänge aufgrund regulatorischer Vorgaben ihre Kosten vollumfänglich selbst decken müssen, verlangen sie hohe Gebühren.²⁰ Stärker als im Bereich der grundständigen Lehre entscheidet die Nachfrage, ob sich ein Studienangebot durchsetzt und Abnehmer findet. Dies wird allerdings in der gegenwärtigen Situation nicht als Problem betrachtet:

»Das anfangs angebotene CAS war so stark nachgefragt, dass wir das CAS in drei separate CAS aufgesplittet haben. Teilweise haben wir jetzt Wartezeiten für einzelne CAS, d. h. wir könnten die Kurse problemlos doppelt anbieten. [...] Normal sind so 20 Leute pro CAS, aufgrund der hohen Anzahl Bewerbungen haben wir die Zahl auf 32 erhöht. [...] [Ich] könnte selbst nur noch unterrichten und hätte ein volles Pensum damit« (Prof_FH_A: 8:6).

Den Befragten gilt die rapid steigende Nachfrage nach Weiterbildungen in Datenwissenschaften als Beleg für die zunehmende Orientierung von Organisationen unterschiedlicher Felder an datengetriebenen Prozessen und Praktiken. Sie attestieren den Weiterbildungsmodulen insofern auch eine »Pilotfunktion« mit experimentellem Charakter: Neue Inhalte könnten rasch in ein Gefäss implementiert werden, das sich auf dem Weiterbildungsmarkt zu bewähren hat und – im Falle der Fachhochschulen – schliesslich in die grundständige Lehre aufgenommen wird.

»[Die] Weiterbildung ist sehr nahe am Puls der Leute im Job, [sie] hat eine Art Pilotfunktion; darin kann experimentiert werden mit verschiedenen Formaten. Je nach Erfolg und Nachfrage wird das Thema dann in die normale Lehre aufgenommen« (Prof_FH_C: 9:14).

Trotz der hohen Nachfrage und spezifischen Grenzziehungen gegenüber anderen Hochschulprofilen existiert in diesem Teilbereich bloss eine schwache Konkurrenz um potenzielle Studieninteressierte (Prof_FH_A: 8:18; Prof_FH_C: 9:20). Grund dafür ist die stark regionale Verhaftung der Angebote, die kaum Studierende aus anderen Landesteilen oder dem Ausland anziehen können, auch weil häufig Abend- oder Wochenendunterricht für die in der Regel voll berufstätigen Teilnehmenden stattfindet.

Studienangebote der Weiterbildung sind hochschulübergreifend modular strukturiert, was eine baukastenförmige Studienwahl ermöglicht: In der Regel können unterschiedliche CAS belegt und zu einem DAS oder MAS zusammengeführt werden

20 Ein einzelner CAS-Baustein kostet je nach Hochschule zwischen 4'500 CHF (Fernfachhochschule Schweiz) und 17'000 CHF (für das CAS »Data Science & Management«, das gemeinsam von der Universität Lausanne und der EPFL angeboten wird), ein gesamter MAS schlägt mit 25'000–30'000 CHF zu Buche.

(BFH, FFHS, HSLU oder ZHAW). Dabei bestehen innerhalb der CAS nur geringe Wahlmöglichkeiten, da es sich um äusserst strukturierte Kurse mit vorgegebenen Themen und Stundenplänen handelt. Bisweilen unterscheiden sich die Angebote in den Pflichtmodulen: An der ZHAW sind die zwei Module »Datenanalyse« und »Information Engineering« Voraussetzung, um die anschliessenden Module »Statistical Modelling« und »Machine Intelligence« belegen zu können. Des Weiteren existiert ein optionales CAS in »Data Product Design«.²¹ An der BFH ist lediglich das Modul »Datenanalyse« Pflicht. Dafür liegt der Fokus der BFH durch das CAS Data Visualization stärker auf Visualisierungsaspekten, während diese an der ZHAW in die CAS Datenanalyse und CAS Statistical Modeling integriert sind. Daneben erhalten auch rechtliche Fragestellungen für Praktiker*innen aus den Unternehmen immer grössere Bedeutung, was sich allerdings noch kaum in eigenen Kursen ausgebildet hat. Ausnahme sind gewisse rechtswissenschaftliche Veranstaltungen wie beispielsweise zum Datenschutzrecht.

In den Weiterbildungsstudiengängen ist die Immersion in die Praxis anders gelagert, als dass die Studierenden als professionelle Praktiker*innen in die Hochschulen eintreten. Sie bringen ihre eigenen, d. h. unternehmens- bzw. organisationsspezifischen Frage- und Problemstellungen selbst mit, die sie im Laufe der Weiterbildung kontinuierlich bearbeiten. Während die MAS-Studiengänge dafür eigene Module für Masterarbeiten etabliert haben, umfasst das CAS Data Science & Management von Universität Lausanne und EPFL ein Individual Data Project.

»The capstone project of the Data Science and Management Executive Certificate is the ›Individual Data Project - IDP‹. Participants are encouraged to start the program with a concrete ›data transformation‹ ambition in mind (usually for their employer). They will develop their project throughout the program with the support of HEC Lausanne and EPFL experts.«²²

Die Idee dieses und vergleichbarer Praxismodule liegt darin, ein »Datenprojekt« auszuarbeiten und über die gesamte Studienzeit (meist 1–2 Jahre) zu verfolgen. Ziel ist es, bestimmte Prozesse und Verfahren in der Herkunftsorganisation der Studierenden zu adaptieren, sprich: zu datafizieren und somit maschinenlesbar zu machen, oder Ideen für neue »Datenprodukte« auszuarbeiten und zu implementieren. Das Netzwerk aus Hochschule, Studierenden und Unternehmen lehnt sich dabei an das Verständnis von Digitalisierung als digitaler Transformation bzw. »Datentransformation« an, schreibt sich somit in organisationale Strategien ein und möchte diese Entwicklung aktiv gestalten.

Trotz der Dynamik und der organisationalen Innovationen, die erste Weiterbildungsangebote an Schweizer Hochschulen und Universitäten im akademischen Feld der Datenwissenschaften darstellten, offenbart die Analyse sowohl strukturelle als auch inhaltliche Konvergenzen: Die Weiterbildungsangebote sind allesamt modu-

21 Das CAS thematisiert die Umsetzung von datenwissenschaftlichen Methoden und Vorgehensweisen in konkrete Anwendungen, Produkte oder Dienstleistungen. Zudem fokussiert es neben unternehmerischen Fragen auch »Datenschutz und Datensicherheit«. Es ist laut einem Befragten entwickelt worden aus der Erkenntnis, dass die Studierenden »zwar technisch top« seien, aber den Nutzen von datenintensiven Lösungen nicht gut erklären, d. h. »sich nicht gut verkaufen können« (Prof_FH).

22 Vgl. <https://execed.unil.ch/en/open-program/certificate-data> (Zugriff: 03.02.2022).

lar aufgebaut und trotz Profildbildungen und Schwerpunktsetzungen inhaltlich sehr ähnlich konzipiert. Analog zu den Regelstudiengängen verweist dies auf Kanonisierungsprozesse sowie eine inhaltliche Engführung auf bestimmte Aspekte der Datenwissenschaften, denen von Praktiker*innen in Unternehmen Relevanz zugesprochen wird. Trotz der geringen Konkurrenz um Studierende in diesem Subfeld zeigt sich ausserdem eine ausgeprägte Orientierung der Anbieter von Weiterbildung untereinander, was letztlich auf eine isomorphe Entwicklung hin zu einem strukturell ähnlich segmentierten Teilfeld innerhalb der Datenwissenschaften hinweist.

9.9 Diskussion

Die Analyse der formalen Strukturen und inhaltlichen Ausprägungen von Curricula der Datenwissenschaften im Schweizer Hochschulfeld offenbart eine Vielfalt unterschiedlicher Studiengänge, Vertiefungen, Programmbezeichnungen und Kursangebote auf allen Stufenniveaus, die vor allem von technikwissenschaftlichen Hochschulen und Fakultäten einerseits und wirtschaftswissenschaftlichen Business Schools andererseits organisiert und durchgeführt werden. Dabei neigen Studiengänge an technikwissenschaftlichen Hochschulen bzw. Fakultäten eher zum starken Pol im Kontinuum interdisziplinärer Curricula (Knight et al. 2013): Sie sind in der Regel stärker strukturiert, weisen einen höheren Anteil verpflichtender Studienleistungen im Kernbereich sowie weniger Wahlmöglichkeiten für die Studierenden auf. Datenwissenschaftliche Programme an Business Schools hingegen sind eher am schwächeren Pol orientiert, da sie weniger computerwissenschaftliche und statistische Methodenveranstaltungen im Kernbereich vorschreiben und gleichzeitig etwas mehr Optionen in der Wahl wirtschaftswissenschaftlicher Anwendungsfelder lassen.

Die Analyse legt eine Strukturlogik offen, die die Mehrheit der untersuchten Curricula durchzieht: Zentral sind die Differenzierung und ein modulartiger Aufbau in einen Kernbereich, der in der Regel die Pflichtveranstaltungen umfasst und der gemeinhin die Ausrichtung des Studiengangs prägt; in einen Wahlbereich, in dem eine inhaltliche Expertise in einer bestimmten Fachdomäne erworben wird; in Kurse und Praktika, welche die Anwendbarkeit der erworbenen theoretischen und methodischen Wissensbestände in der Praxis garantieren; in eine Abschlussarbeit, in der die erworbenen Kenntnisse anhand einer eigenen Forschungsarbeit nachgewiesen werden; sowie schliesslich in sogenannte »komplementäre Inhalte«, in denen durch Vermittlung rechtlicher, ethischer und sozialer Aspekte »verantwortungsbewusste« Praktiker*innen ausgebildet werden sollen. Insofern unterscheiden sich die Curricula nicht nur bezüglich ihrer Struktur, sondern auch ihrer inhaltlichen Ausgestaltung nur geringfügig voneinander. Inhaltliche Differenzen erklären sich primär durch die Positionierung und Ausrichtung der verantwortenden Hochschulen und Universitäten im akademischen Feld.

Die Curricula übersetzen die Universalität und die Logik der Domänen, die die Datenwissenschaften im akademischen Feld insgesamt charakterisiert, in die organisationalen Strukturen von Universitäten und Hochschulen. Im Aufbau der Curricula artikuliert sich eine Aktualisierung und Erweiterung existierender technikwissenschaftlicher Ausbildungstraditionen: Technisch-methodische Veranstaltungen werden als Kernbereich der Datenwissenschaften und somit als universal und domä-

nenunabhängig gerahmt. Durch die ergänzenden Module werden sie durch inhaltliche Fachexpertise, Anwendungsorientierung sowie eine »Sensibilität« vor allem für rechtlich und – in wesentlich geringerem Ausmass – ethisch heikle Aspekte datenwissenschaftlicher Praxis erweitert. Sozialwissenschaftliche Fragestellungen und Kategorien finden bis dato noch kaum Eingang in die Studieninhalte. Relevanz wird den Curricula in erster Linie durch die Bearbeitung »echter«, »realweltlicher« Probleme in der Praxis zugeschrieben, die zumeist ausserhalb der akademischen Forschung situiert werden und sich eng an ausserakademische, industrielle Verfahren und Wissensbestände anlehnen.

In der identifizierten curricularen Strukturlogik und den beobachtbaren Implementierungen der Studiengänge manifestieren sich widersprüchliche Verschränkungen von Innovation sowie Interdisziplinarität einerseits und Kanonisierung und Reproduktion andererseits. So stellen die Curricula zum einen organisationale Innovationen dar: In der Regel kollaborieren zwei oder mehr Fachbereiche in der Planung und Umsetzung interdisziplinärer Studiengänge. Es handelt sich um disziplinäre Grenzüberschreitungen, was angesichts der teilweise intensiven Konflikte über die Disziplinierung der Datenwissenschaften bzw. die Deutungshoheit über die Expertise in Bezug auf Datenbearbeitung im akademischen Feld erstaunen mag. Die Verknüpfung spezifischer disziplinärer Wissensbestände in gemeinsamen curricularen Gefässen formuliert zudem einen neuen Anspruch auf besagte Expertise. Insofern stellen die Curricula auch Rekombinationen bestimmter existierender Wissensfelder dar, gegenüber denen sie sich neu positionieren und autonomisieren können.

Synchron zu den grenzüberschreitenden Kooperationen und Innovationen manifestiert sich in den Curricula allerdings auch eine Orientierung an bzw. die Aktualisierung von etablierten Strukturen: So zeigen sich inhaltlich bereits in einer frühen Phase der Institutionalisierung der Datenwissenschaften im akademischen Feld Kanonisierungsprozesse im interdisziplinären Kernbereich, noch bevor überhaupt abschliessend klar ist, was der Gegenstandsbereich des neuen Feldes sein soll und welche Disziplinen daran beteiligt sind. Bereits die Studienzulassung schränkt die Offenheit der Studiengänge – und damit der Datenwissenschaften als Wissensfeld insgesamt – sowohl disziplinär als auch sozial erheblich ein. Durch die disziplinäre Engführung auf ein bestimmtes Konzept der Datenwissenschaften reartikulieren die Curricula zudem die epistemologischen und disziplinären Auseinandersetzungen um Deutungshoheit über die Expertise in der Datenanalyse. Dadurch werden existierende Konfliktlinien in organisationale Gefässe überführt und die Grundlagen für Lehr- und Forschungspraktiken gelegt, die die Reproduktion etablierter Disziplinen unter neuen Bezeichnungen erlauben.

Kapitel 10 – Die Suche nach den richtigen Kompetenzen

»Was eben fehlt [...], sind so ein bisschen diese Auslegeordnungen: Uns fehlen Kompetenzmodelle, uns fehlen sozusagen [...] Landkarten, wie diese Data Economy, was für unterschiedliche Arten von Playern es gibt, welche Berufsprofile damit einhergehen und so weiter« (Prof_UH_E: 13:90).

10.1 Einleitung

Die organisationalen Verhandlungen über die Etablierung von Studiengängen sind in den Curricula mit der Frage nach den jeweiligen Anforderungen, Fähigkeiten und Kompetenzen verknüpft, die für ein bestimmtes Studium und eine daran anschließende berufliche Tätigkeit als notwendig erachtet werden. Dabei artikulieren die Curricula sowohl Gegenwartsdiagnosen als auch Zukunftsvorstellungen darüber, wie eine bestimmte Praxis entworfen und imaginiert wird. Curricula können damit als Produkte einer feldübergreifenden Konstruktion sowohl von Gegenwart als auch Zukunft (Young 1998) konzipiert werden: Die Bildungsangebote schaffen durch ihre Auswahl, Kombination und Fokussierung ganz bestimmter Inhalte neue Möglichkeitsräume und Zukünfte für Studierende und Absolvent*innen.

Ich widme mich deshalb in diesem Kapitel den zugeschriebenen »Kompetenzen« und individuellen Eigenschaften von (zukünftigen) Praktiker*innen der Datenwissenschaften. Der Kompetenzbegriff hat sich zur Bearbeitung bildungsbezogener Problem- und Fragestellungen im Zusammenhang mit gegenwärtigen und künftigen Inhalten, Praktiken und Qualifikationen im Bildungsfeld sowie im Arbeitsmarkt etabliert. Der Begriff orientiert sich an psychologischen und pädagogischen Arbeiten der Bildungsforschung, die Kompetenzen als individuelle Fähigkeiten bzw. Eigenschaften konzeptualisieren, die über bestimmte Indikatoren und Metriken operationalisiert und gemessen werden können. Die konstruktivistische Wissenssoziologie fasst – wie bereits beschrieben – Kompetenzen als »sozial zugeschriebene Qualitäten« (Kurtz 2010: 8). An diese Perspektive anknüpfend untersuche ich, wie die Lehrinhalte in individuelle Zuschreibungen an Datenwissenschaftler*innen übersetzt werden. Wie äussern sie sich in Interviews und Curricula? An welche bestehenden Kategorien und Wissensbestände lehnen sie sich?

Die inhaltsanalytische Auswertung von Curricula und Interviews zeigt, dass sich die untersuchten Studiengänge und Curricula elementar auf den Kompetenzbegriff stützen. Die Suche nach den »richtigen« Kompetenzen ist inhärent mit der Genese

und Entwicklung des untersuchten Feldes verknüpft: Akteur*innen in der Ökonomie, Wissenschaft sowie Bildungs- und Hochschulpolitik versuchen in unzähligen Kompetenzdefinitionen und -profilen, die richtige Rezeptur für die Komposition eines »Data Science Skillsets« zu finden und zu begründen. Damit einher geht allerdings gleichzeitig auch eine uneinheitliche Verwendung des Kompetenzbegriffs: Je nach zugrundeliegendem Modell, gewählter Forschungsstrategie oder Ausrichtung eines Studiengangs unterscheidet sich, was überhaupt als Kompetenzen erfasst und verstanden wird.

Ich argumentiere, dass sich in den Diskussionen über die jeweiligen Anforderungen, Qualifikationen und Fähigkeiten, die für ein Studium oder eine berufliche Tätigkeit als notwendig erachtet werden, organisationale Suchprozesse über die Bedeutung und Sinnhaftigkeit der Datenwissenschaften manifestieren. Trotz unterschiedlicher Herangehensweisen verdeutlichen die intensive Suche und das Ringen um das »richtige« Kompetenzprofil bzw. die adäquaten Curricula den hohen – bildungspolitischen genauso wie ökonomischen – Stellenwert, den solche Konstrukte im Untersuchungsfeld einnehmen. Suchprozesse nach Kompetenzprofilen und Zuschreibungen individueller Eigenschaften bilden somit einen weiteren Praxismodus in der Konstruktion der Datenwissenschaften zwischen etablierten Feldern: Praktiken von *boundary work* schärfen das Profil nach innen und binden es synchron in das bestehende disziplinäre Gefüge ein. Die (zukünftigen) Data Scientists werden als Intermediäre (Bessy & Chauvin 2013) mit einer zentralen Schnittstellenfunktion zwischen verschiedenen disziplinären Wissensbeständen und organisationalen Einheiten konzipiert, woraus sich auf der curricularen Ebene eine Komplementierung ihres Kompetenzprofils jenseits ihrer technisch-methodischen Expertise durch soziale, kommunikative und mentale Eigenschaften ableiten lässt. Die Aushandlung und Festlegung mess- und überprüfbarer Kompetenzbereiche befördert und festigt insofern die Selbst- und Fremdwahrnehmung eines interdisziplinären Wissensfeldes *in the making*.

Das Kapitel ist wie folgt strukturiert: Zuerst beschreibe ich die allgemeinen Kompetenzprofile in den Datenwissenschaften anhand der untersuchten Curricula (Kap. 10.2). Anschließend diskutiere ich die den Subjekten attestierten Kompetenzen und persönlichen Eigenschaften jenseits der statistisch-mathematischen und technischen Fähigkeiten, die als bedeutend für eine datenwissenschaftliche Praxis eingestuft werden. Der Begriff der Datenkompetenz (Kap. 10.3) markiert ein inklusives Verständnis der Wissensbestände und Praktiken des entstehenden Feldes. Die geforderte Problemlösungskompetenz (Kap. 10.4) schliesst in vielfältiger Hinsicht an einen ingenieurialen Wissensmodus und das Primat der Praxis in technologieorientierten Wissensgebieten an. Kommunikation und Teamarbeit (Kap. 10.5) hingegen verweisen auf die Bedeutung von Kollaborationen multipler Expertisen, insbesondere an Schnittstellen unterschiedlicher organisationaler Einheiten. Denkweisen und persönliche Einstellungen (Kap. 10.6) bilden in unterschiedlichen Ausprägungen die kognitiven Grundlagen dafür, das Soziale mit und durch Daten wahrzunehmen und zu repräsentieren. Schliesslich markieren veränderte kulturelle Rahmungen den zukünftigen Praktiker*innen vielversprechende Zukünfte (Kap. 10.7), die sich in der Verknüpfung der technischen und analytischen Praktiken, Kompetenzen und Denkweisen mit veränderten Wertvorstellungen manifestieren und insofern zur Demokratisierung der Wissensformation »Data Science« beitragen sollen. Abschliessend werden die wichtigsten

Ergebnisse der Analyse vor dem Hintergrund der übergeordneten Fragestellungen der Arbeit diskutiert (Kap. 10.8).

10.2 »Jack of all trades and a master of some«: Kompetenzprofile in den Curricula

Die datenwissenschaftlichen Kompetenzprofile lassen sich zunächst entlang der Unterscheidung zwischen berufsbezogenen und allgemeinen Kompetenzen bei Green (2013) betrachten:¹ So umfassen die Curricula zum einen methodisches Wissen und Fertigkeiten wie statistische Modellierung, maschinelle Lernverfahren oder neuronale Netzwerke und technische Fertigkeiten wie objektorientierte Programmierung oder Datenbank-Engineering. In ihrer Kombination können diese als berufsbezogene Kompetenzen betrachtet werden, die sie beispielsweise von »reinen« Statistiker*innen und Informatiker*innen differenzieren. Solche anwendungsorientierten Fähigkeiten gelten diversen Interviewten als der eigentliche Kern von Data Science, der von umliegenden Feldern abgrenzt. Zum anderen setzen die Curricula kognitiv-abstraktes Wissen der mathematischen, statistischen und computerwissenschaftlichen Grundlagen als theoretische Basis. Dabei handelt es sich um generische Fähigkeiten, die auch für quantitative Berufsgruppen der Mathematik, Informatik oder des Engineerings charakteristisch sind. Sie taugen entsprechend kaum zur Ausbildung eines distinkten Berufsprofils.

Der Masterstudiengang in Data Science an der EPFL beschreibt die Kombination, wie sie in vielen Studiengang- und Kursbeschreibungen zu finden ist, beispielhaft:

»This new, multidisciplinary style of analysis, utilizes techniques from computer science, mathematics and statistics, while requiring astute problem understanding and good communication skills. It aims to answer questions such as, »how can the data generated be useful?« These anticipation and prediction abilities have made Data Science one of the most active fields in industry with a continuing shortage of talent« (EPFL 2017: 3).

Zusätzlich zu den Kernbausteinen (multidisziplinäre Analysemethoden sowie Techniken der Informatik, Mathematik und Statistik) werden demnach Problemverständnis (»astute problem understanding«) und kommunikative Fähigkeiten als grundlegend für »Data Science« gesetzt. Die Curricula betrachten also den »Kernbereich« datenwissenschaftlicher Expertise als nicht ausreichend. Vielmehr erfolgt eine Komplementierung durch eine ganze Reihe weiterer interaktiver, kommunikativer und mentaler Eigenschaften (wie Datenkompetenz, Kommunikation oder eine unternehmerische Haltung). Neben dem allgemeinen Bedeutungszuwachs dieser »social skills« in zeitgenössischen Arbeitsmärkten (Salvisberg 2010) kann dies gerade im akademischen Feld

1 Green (2013) unterscheidet zwischen berufsbezogenen und allgemeinen Kompetenzen. Er fasst Fähigkeiten, die für eine bestimmte Berufsgruppe relevant sind, unter berufsbezogenen Kompetenzen zusammen (ebd.: 22). Demgegenüber umfassen allgemeine Kompetenzen jene Aufgabenbereiche, die berufsgruppenübergreifend bedeutsam sind. Als Subkategorien davon unterscheidet Green zwischen kognitiven, interaktiven sowie physisch-körperlichen Fähigkeiten (ebd.).

auch als Reaktion auf zunehmende wissenschaftsexterne Bezüge und Anforderungen verstanden werden (Maasen 2010; Maasen & Dickel 2016).

Parallel zu dieser vergeschlechtlichten Einteilung in ›hard‹ und ›soft skills‹ (Guerrier et al. 2009) erarbeiten sich die Datenwissenschaftler*innen, so die Konzeption der Curricula, praxisorientiertes Wissen in bestimmten Anwendungsfeldern, ohne allerdings zu Expert*innen in diesen Bereichen zu avancieren. Vielmehr geht es darum, mit den Forschungs- bzw. Praxispartner*innen in einen Dialog treten zu können, ohne dass disziplinäre Grenzen aufgehoben würden. In diesem Zusammenhang bezeichnet ein ETH-Professor einen Data Scientist als »Jack of all trades and a master of some«.² Eine breite Ausbildung in den verschiedenen methodischen und technischen Bereichen der Datenwissenschaften wird durch Spezialisierung und Verknüpfung mit einem bestimmten Fachgebiet komplementiert. Für Data Scientists im wissenschaftlichen Feld sei entsprechend vor allem das Verständnis für die Arbeit von Forschenden aus den verschiedenen Domänen wichtig (Prof_ETH_D: 5:1).³ Die Kombination unterschiedlicher Wissensbestände, Qualifikationen und persönlicher Einstellungen situiert datenwissenschaftliche Expertise an der Schnittstelle disziplinärer und organisationaler Einheiten, wodurch ihr eine universelle Relevanz und Anwendbarkeit zugesprochen wird. Beides zugleich macht sie so interessant für unterschiedliche soziale Felder und trägt gleichzeitig zu ihrer gesellschaftlichen Wahrnehmung und Konstruktion als einer emergierenden technischen Eliteprofession (Avnoon 2021) bei. Ich untersuche im Folgenden die weiteren Bausteine, die neben der technisch-methodischen Expertise als grundlegend für die Zusammenstellung der ›richtigen‹ Kompetenzprofile identifiziert werden.

10.3 Datenkompetenzen: Die Suche nach neuen Begrifflichkeiten

Ein erster zentraler Baustein bilden Datenkompetenzen bzw. Data Literacy.⁴ Sie gelten den interviewten Professor*innen übereinstimmend als eine »Grundkompetenz, über die jeder verfügen sollte. Deshalb ist Data Science so wichtig, auch als Nebenfach oder als allgemeine Grundvorlesung« (Prof_UH_A: 6:68). Über das Beherrschen verschiedener Techniken und Methoden zur Bearbeitung von Daten hinaus involviert ein umfassendes Verständnis von Datenkompetenzen diverse weitere Praktiken:

»Data Literacy ist die Fähigkeit, Daten auf kritische Art und Weise zu sammeln, zu managen, zu bewerten und anzuwenden (Ridsdale et al. 2015). Dazu gehören eine ganze

2 Dabei handelt es sich um eine gängige Abwandlung des englischen Sprichworts »Jack of all trades, and master of none«, das auf Deutsch mit »Hansdampf in allen Gassen« übersetzt werden kann.

3 Die Studienprogramme der ETH bieten den Studierenden die Möglichkeit von thematischen Nebenfächern bzw. Vertiefungen, während an den Universitäten eher fachspezifische Master angeboten werden.

4 Die Abgrenzung von Datenkompetenz und Data Literacy (wörtlich ›Datenlesefähigkeit‹) ist unscharf. Bisweilen wird Data Literacy mit Datenkompetenz gleichgesetzt (Heidrich et al. 2018; Lübcke & Wanemacher 2018), während umgekehrt Literacy als die *kritische* Anwendung von Kompetenzen betrachtet wird (vgl. Livingstone 2014).

Reihe von Einzelkompetenzen – vom Schaffen des grundlegenden Bewusstseins bis hin zu ethischen und rechtlichen Fragestellungen« (Heidrich et al. 2018: 106).

Auffallend an dieser Definition, die sich mit den Konzeptualisierungen im empirischen Material deckt, ist vor allem die Breite der »Einzelkompetenzen«, die neben methodisch-analytischen Fähigkeiten von der Schaffung eines »grundlegenden Bewusstseins« bis zu den »ethischen und rechtlichen Fragestellungen« reicht. Datenkompetenzen werden insofern sowohl umfassend für den gesamten Prozess der Datenbearbeitung verwendet als auch für spezifische Teilschritte davon (Dichev & Dicheva 2017; Gould 2017; Heidrich et al. 2018; Schuff 2018). Dies wird in einem Interview besonders anschaulich: Nachdem ein Befragter zunächst die Praxis »eines [sic!] Data Scientists« als die Summe unterschiedlicher, »interdisziplinärer« Kompetenzen beschreibt,⁵ expliziert er in der Frage nach den Herausforderungen datenwissenschaftlicher Praxis in Organisationen »Datenkompetenz« wie folgt:

»Sprich: sie stellen die falschen Fragen oder können gar nicht die richtigen Fragen an die Daten stellen. Und das ist eigentlich [...] eine Grundkompetenz eines Data Scientists. Weil er heisst ja auch Data Scientist, d. h. er muss auch gute Fragen stellen können. Und [...] das ist wirklich die Quintessenz: Dass Data Scientist wirklich ein Job sein muss, [der] Datenkompetenz mit Anwendungskompetenz verbindet« (Prof_FH_D: 12:4).

Die »richtigen« bzw. »gute[n] Fragen« an die Daten werden hier nicht nur zum Kern einer auszubildenden Datenkompetenz erhoben, sondern zur Praxis von Data Scientists an sich. Die Analyse der Definitionen und Konzeptionen im empirischen Material und der Literatur verdeutlicht, dass es sich um einen kontingenten Begriff handelt, der universell ist, gleichzeitig aber stets auch nur einen spezifischen Teilaspekt bezeichnen kann. Die breite Verwendung in unterschiedlich ausgerichteten Studiengängen und institutionellen Settings markiert die Suche nach »neuen« Begrifflichkeiten und Konzepten in einer Phase unsicherer Entwicklung und des Zusammenfließens und Überlagerns unterschiedlicher Expertisen. Dies stellt keine Inkompetenz des Untersuchungsfeldes dar, sondern ist elementarer Bestandteil jener synchronen Prozesse von Grenzziehung und -überschreitung, von Kooperation und Konkurrenz, die die Wissensformation in ihrer Entwicklung charakterisieren.

Die Notwendigkeit breit vermittelter Datenkompetenzen im Sinne eines kollektiven »Upgrading« (Dröge 2019: 23f.) oder Upskilling digitaler Wissensarbeit muss vor dem Hintergrund der zunehmenden Verbreitung automatisierter Systeme der Datenverarbeitung, wie sie eine expandierende Data-Analytics-Industrie für alle möglichen Betätigungsfelder anbietet (Beer 2019; Gehl 2015), betrachtet werden. Dadurch werden Datenanalysen zwar einerseits für ein breiteres Publikum zugänglicher, andererseits bleiben fundamentale Prozesse wie Datenerhebung und -konstruktion, Methodenauswahl, Qualitätsmessgrößen oder Ergebniskontrollen für die Nutzer*innen solcher Systeme tendenziell undurchsichtig und opak.

5 »Dort hat es [...] Informatik-Kompetenzen, dort hat es analytische Kompetenzen, sprich auch statistische Kompetenzen, aber es gibt auch Visualisierungskompetenzen, Kommunikationskompetenzen. Und die Verbindung mit einer Anwendungsdomäne« (Prof_FH_D: 12:7).

Eine interviewte Person greift dieses Spannungsverhältnis zwischen allgemeinen Anforderungen an alle Studierenden und der Spezialisierung ganz bestimmter Professionen wie der Datenwissenschaften wie folgt auf:

»Also... insofern also die Spezialisierung auf ein Profil [...] ist schon noch interessant, und da kann ich nur sagen: [...] ein Stück weit werden wir da hochspezialisierte Personen [ausbilden], die da die Algorithmik machen, und ganz viele, die eigentlich drumherum letztendlich Probleme eingeben, oder Daten interpretieren müssen. Und das ist dann so diese Literacy-Thematik.« (Prof_UH_E: 13:40)

Die Differenzierung in »hochspezialisierte Personen« und jene, die »letztendlich Probleme eingeben, oder Daten interpretieren müssen«, wird im Interview mit einem vorsichtig warnenden Unterton artikuliert. Der Befürchtung zufolge würde auch die breite Vermittlung von Datenkompetenzen nicht zur Einebnung bestehender Hierarchien führen, sondern vielmehr neue Konfliktlinien zwischen den involvierten Berufsgruppen produzieren. Diese Position schliesst an die Diagnose beruflicher Polarisierung an, wobei die Data Scientists zu den »Gewinnerinnen« und die breite Masse datenkompetenter Studierender relativ betrachtet zu den »Verlierer[n]« zählen dürften (Dröge 2019: 24f.) – wobei die geschlechterspezifischen Folgen solcher Polarisierungs- und Hierarchisierungsprozesse noch zu überprüfen sind (vgl. Kap. 10.7). Die Hierarchisierung unterschiedlicher Spielarten von Datenkompetenzen kontrastiert zudem jene Schilderungen, die in den kommunikativen und interaktiven Kompetenzen der Datenwissenschaften den Schlüssel für fächer- und disziplinenübergreifende Kollaborationen verorten.

Unabhängig von zukünftigen beruflichen Differenzierungs- und Hierarchisierungsprozessen wird Datenkompetenz als elementarer Baustein anerkannt und bildet gemeinsam mit den methodisch-analytischen und technischen Fähigkeiten das Fundament datenwissenschaftlicher Kompetenzprofile. Verallgemeinert formuliert bezeichnen Datenkompetenzen also weniger die statistische Analyse von (grossen) Datensätzen, sondern vielmehr eine Vertrautheit und positive Grundeinstellung gegenüber dem Modus datenintensiver Wissensproduktion. Ihre breite Vermittlung zielt vor allem auf die Fähigkeit, bestimmte Aspekte der sozialen Realität überhaupt als Daten erfassen zu können und zu *wollen*. In der datengetriebenen digitalen Transformation müssen nicht alle Daten-Cracks werden, sondern Datenkompetenzen als kognitive Einstellung werden gewissermassen zur Grundlage dafür, die Welt durch die »Brille« von Daten wahrzunehmen, zu erfassen und zu beurteilen (Beer 2019).

10.4 Problemlösungskompetenzen: Die Übersetzung der Anwendungsorientierung in die Praxis

Indem eine zentrale Erwartung an datenwissenschaftliche Praxis in der Adressierung und Bearbeitung gesellschaftlicher Herausforderungen besteht, bilden Kompetenzen zur Problemlösung ein weiteres Puzzlestück in der Zusammenstellung der »richtigen« Kompetenzprofile (Lowrie 2018). Die Vermittlung bzw. Aneignung von Problemlösungskompetenzen ist in allen untersuchten Studiengängen vorgesehen, wenn auch je nach Ausrichtung des Programms in unterschiedlicher Umsetzung: In den tech-

nikwissenschaftlich orientierten Studiengängen äussert sich die Erwartungshaltung im bereits diskutierten Narrativ der »echten Probleme« (*real-world problems*): Dazu gilt es »echte Daten« (*real-world data*) zu erheben und analysieren, um grundlegende Problemstellungen und »globale Herausforderungen« wie Klimawandel, Hungerkatastrophen oder die Effekte transkontinentaler Migrationsströme methodisch kontrolliert zu bearbeiten (ETH-Rat 2014, 2016a). Es ist in erster Linie das politische Feld, das den Datenwissenschaften solche weitreichenden Kapazitäten zur Problemlösung zuschreibt, die eher solutionistische Technikutopien (Morozov 2013) referieren als die kontingenten Realitäten wissenschaftlicher Erkenntnisproduktion.

Die befragten Leitungspersonen technikwissenschaftlicher Studiengänge äussern sich diesbezüglich zurückhaltender: Sie beziehen das Potenzial der Datenwissenschaften auf konkrete Fragestellungen innerhalb ihrer Forschungsbereiche. Dabei könne zwar durchaus auch gesellschaftlicher »Impact« anfallen, allerdings eher als Nebenprodukt wissenschaftlicher Tätigkeit. Dennoch schätzen auch sie Problemlösungskompetenzen als bedeutend für die professionelle Praxis der Studierenden ein.

In Business Schools sowie Weiterbildungsstudiengängen, in denen eher ein unternehmensorientiertes Verständnis von Datenwissenschaften praktiziert wird, werden die zu lösenden Probleme in erster Linie in Unternehmen und andere Organisationen transponiert. So schildert ein Universitätsprofessor die Herausforderung, Studierende in datenwissenschaftlichen »Methoden und Denkweisen« zu schulen, um dadurch organisationale Problemstellungen zu bearbeiten:

»Die Idee ist es, die Produktivität und die Problemlösungskompetenz der Leute, die in Unternehmen oder [die] Verwaltung gehen, zu erhöhen. Um Probleme in Organisationen zu lösen, muss man heutzutage auch gut mit Daten umgehen können. Die Herausforderung ist, dass diese beiden Welten nicht verknüpft sind: Leute im Management und Leute in Data Science verstehen sich nicht und fühlen sich jeweils durch die andere Seite missverstanden. Wir möchten, dass die Studierenden beide Seiten verstehen: Unsere Hochschule hat einen Schwerpunkt in Management, Business und Recht, aber kaum Verständnis für Datenprozesse, Analytics, Algorithmen und so weiter. [...] [Das] Programm dient dazu, die Leute in diesen Methoden und Denkweisen, die sie hier bis jetzt nicht kriegen würden, zu komplementieren« (Prof_UH_B: 3:8).

Der Interviewte imaginiert in diesem Zitat organisationale Probleme primär als datengetrieben, wobei den traditionellen Entscheidungsträger*innen (»Leute im Management«) bezüglich Daten die methodisch-technische Expertise fehle. Im Gegensatz zu den technikwissenschaftlichen Studiengängen (vgl. oben) erfolgt die »Komplementierung« demnach umgekehrt: Die Studierenden verfügen zwar über ökonomisches bzw. betriebswirtschaftliches Domänenwissen, haben »aber kaum Verständnis für Datenprozesse, Analytics, Algorithmen und so weiter«. Darin impliziert ist ein instrumentalistisches Verständnis, das die Komplexität sozialer Organisation auf operationalisierbare, mittels Datenanalysen lösbare Prozesse reduziert. Betriebswirtschaftliche und andere Expertisen werden durch datenwissenschaftliche »Methoden und Denkweisen« transformiert: Entscheidungen des Managements werden nicht mehr als kontingent gerahmt, sondern als Ergebnis methodisch kontrollierter Datenanalysen unternehmensinterner und -externer Datenbestände, wodurch sie objektiviert und neutralisiert erscheinen. Eine solche Rationalisierung unternehmerischen Entschei-

Handlungshandelns ist nicht neu, sondern hat eine lange Kontinuität in der Wissenschaftsgeschichte der Betriebswirtschaftslehre (Burren 2010). Dennoch aktualisiert sich durch die Betonung von Problemlösungskompetenz das »Spannungsfeld von Wissenschaftlichkeit und Praxisorientierung« (ebd.: 12) in Business Schools als zeitgenössische Ausprägung einer datenintensiven Anwendungsorientierung.

10.5 Kommunikation und Teamarbeit: Zur »Moderationsrolle« von Data Scientists an organisationalen Schnittstellen

Die Fähigkeit, organisationale Problemstellungen datenbasiert zu bearbeiten, wird ferner mit kommunikativen und sozialen Kompetenzen von Datenwissenschaftler*innen verknüpft, die in empirischen Studien wiederkehrend als bedeutend identifiziert werden. Eine weitere interviewte Person verknüpft diese Bausteine wie folgt:

»[Die] Probleme von Organisationen zu verstehen und die entsprechenden Daten zu finden, d. h. in die Organisationen zu gehen und diese gemeinsam zu identifizieren, aufzubereiten und [zu] analysieren, ist zentral. Noch lange bevor es zu den ganzen induktiven Verfahren kommt. [...] Verstehen, was das Problem genau ist, also was Herbert Simon als Intelligenz bezeichnet, und Lösungsvorschläge für diese Probleme formulieren zu können, [ist] absolut zentral. Deshalb ist uns eine interdisziplinäre Zusammensetzung wichtig, weil dies unterschiedliche Leute mit unterschiedlichen Fähigkeiten anzieht« (Prof_UH_A: 6:35).

Problemlösungsstrategien basieren diesem Verständnis zufolge auf kollektiver »Intelligenz«, die sich auf Kollaboration und Interdisziplinarität verschiedener Expertisen abstützt. Den unterschiedlichen Ausprägungen von »Data Science« gemein ist die Betonung grundlegender kommunikativer Kompetenzen, wodurch Interdisziplinarität erst möglich wird: Wenn die Phase der »disziplinäre[n] Eigenbrötler« (Prof_UH_A: 6:71) bzw. der »Informatiker-Autismus«⁶ (Expertengruppe 2018: 51) tatsächlich überwunden werden soll, dann müsse mit den Forschungs- bzw. Praxispartner*innen aus anderen Domänen zunächst eine gemeinsame sprachliche Basis gefunden werden. Während die Verknüpfung kommunikativer und sozialer Kompetenzen inhaltlich naheliegend sein mag, stimmt dies für das Untersuchungsfeld nur bedingt: Bis heute werden Data Scientists noch häufig als »Einzelkämpfer« (sic!) bzw. »unicorns« (Davenport 2020; Hermida & Young 2017) imaginiert – was nicht zuletzt ihre symbolische Aufladung mitbefördert. Da solche qualifikatorischen Ausnahmerecheinungen aber auf dem Arbeitsmarkt nicht verfügbar waren, hat sich mittlerweile die Wahrnehmung als Teamplayer durchgesetzt und die Zusammensetzung von »Data Science Teams« in den Vordergrund gerückt.

6 »Die heutige Informatik-Ausbildung geht hingegen immer noch davon aus, dass Informatik-Spezialisten im geschlossenen Silo für digitale Probleme digitale Lösungen suchen und den Austausch nach aussen nicht brauchen. In einer Welt des [Internet of Everything] stellt ein solcher »Informatiker-Autismus« eine Gefahr dar. Die Lehrgänge müssen deshalb entsprechend ergänzt werden mit Schulung im Bereich der mündlichen und schriftlichen Kommunikation, damit die Informatiker ihre Gestaltungskompetenz in Zusammenarbeit mit allen anderen wahrnehmen können« (Expertengruppe 2018: 51).

Kommunikative und soziale Kompetenzen sind auch notwendig, um zunächst die Struktur und die Entstehung von Datensätzen verstehen zu können, bevor organisationale Problemstellungen adressiert werden, wie ein anderer Professor betont:

»Social Skills sind für Data Scientists sehr wichtig: Zum einen muss ein Data Scientist alles über die Daten wissen: wie sind diese strukturiert, wie sind sie entstanden und so weiter. [Weil] sonst ist kein gutes Modell möglich. Zum andern muss ein Data Scientist, um dies in Erfahrung zu bringen, mit den Leuten kommunizieren, die diese erhoben [und] verarbeitet haben. Auf PhD-Level haben wir Kurse, auf Master-Niveau machen wir Projekte innerhalb von Kursen oder gemeinsam mit Unternehmen. [Dies] ist eher eine Lehre als eine Theorie, wir erlernen das praktisch« (Prof_UH_D: 1:14).

Das Zitat verdeutlicht zwei wichtige Dimensionen der sozialen und kommunikativen Kompetenzen: Zum einen konzipiert der Befragte im ersten Teil Datenkompetenzen als Teil sozialer Kompetenzen. Das Wissen über Daten wird nur durch Kommunikation und Kollaboration mit anderen Akteur*innen möglich, was die Relationalität von Daten andeutet: Daten sprechen nicht für sich oder sind irgendwie selbst-evident, sondern nur im Verhältnis zu den Kontexten ihrer Produktion und Konstruktion zu verstehen (Dourish & Gómez Cruz 2018; Gitelman 2013). Gegenüber Konzeptionen, die die methodisch-technischen Dimensionen von Datenkompetenzen als Kern identifizieren (vgl. oben), bilden in diesem Fall umgekehrt kommunikative und soziale Kompetenzen die Grundlage für darauf aufbauende Daten- und Analysekompetenzen.

Andererseits lässt sich der Praxisbezug von Teamarbeit und Kommunikation in den Curricula beobachten: Teamarbeit kommt im Studium insbesondere in den praxisorientierten Kursen sowie Studierendenprojekten eminente Bedeutung zu. Alle untersuchten Universitäten und Hochschulen kennen Unterrichtsformate (vgl. Kap. 9.6), in denen die Studierenden längere Zeit kollaborativ an Projekten und Fragestellungen arbeiten, die sie mittels der verschiedenen Schritte eines Analyseprozesses zu bearbeiten und zu lösen versuchen. Die Curricula werden insofern auf die Realitäten der künftigen Arbeitsorganisation (Boes et al. 2018) synchronisiert.

Schliesslich äussert sich in den diskutierten »Social Skills« die Erwartung einer engeren Zusammenarbeit von Analytics und anderen organisationalen Einheiten innerhalb von Unternehmen. So betont ein Befragter einer Fachhochschule die Notwendigkeit der Einnahme einer »schwierige[n] Moderationsrolle«, da sich Data Scientists an der Schnittstelle verschiedener Unternehmenseinheiten bewegen:

»[Ein] Data Scientist spielt eigentlich eine ganz wichtige Moderationsrolle in Unternehmen, weil er auch ein wenig zwischen IT und Business, zwischen Strategie und Technologie steht. Er spielt eine schwierige Moderationsrolle. [...] Wenn jemand gegenüber Laien [...], gegenüber [einem] Fachpublikum, das sind so unterschiedliche Kommunikationsformate, die man beherrschen muss und das wollen wir auffangen. [...] Das sind nicht die Leute, die irgendwo im dritten Untergeschoss von morgens bis abends vor dem Bildschirm sitzen, sondern das sind Leute, die etwas mobilisieren, bewegen können in Unternehmen. Und das braucht einfach gewisse Meta-Skills.« (Prof_FH_D: 12:65)

Entsprechend vielfältig sind die Bezeichnungen, die sich sowohl im Fachdiskurs als auch im empirischen Material für die vermittelnde Rolle (*interstitial role*) bzw. die ge-

nannten »Meta-Skills« herausgebildet haben: »Data translators« (Marr 2018), »business translators« (Manyika et al. 2017), »Data Manager« (Prof_UH_E: 13:4), »Data Allrounder« (Prof_FH_D: 12:53) oder »Insight Providers« (Kim et al. 2016) beziehen sich auf die Rollen von Data Scientists als Intermediäre (Bessy & Chauvin 2013) zwischen verschiedenen Funktionen (»zwischen IT und Business, zwischen Strategie und Technologie«), aber auch zwischen der Repräsentation von Daten in soziotechnischen Systemen und deren organisationaler Wirklichkeit. Die Positionierung von Data Scientists als Moderator*innen an zentralen organisationalen Schnittstellen verdeutlicht, dass jenseits der technologisch-methodischen Expertise die Ausbildung unternehmerischer Subjekte und Identitäten im Zentrum steht.

10.6 Denkweisen und persönliche Einstellungen: Das »Data Mindset« als Grundlage für Datafizierungsprozesse

Die Curricula umfassen als letzten Baustein Denkweisen und persönliche Einstellungen, die dem Berufsprofil von Data Scientists attribuiert werden. Die Analyse zeigt, dass es bei diesen Attributen wenig Divergenz gibt, sowohl was die institutionelle Verortung der Studiengänge als auch die wissenschaftlichen Fachbereiche der Befragten betrifft. Sie stimmen darin überein, dass die Ausbildung und Verinnerlichung einer datenorientierten bzw. datengetriebenen Denkweise zentral sei.

In ihrer Diskussion der gesellschaftlichen Implikationen von Big Data betonen Mayer-Schönberger und Cukier, dass Datafizierung nicht nur einen soziotechnischen Prozess – die »Überführung sozialer Wirklichkeit in (vorgeblich) objektive Datenstrukturen« (Prietl & Houben 2018: 7) – darstellt, sondern auch einer mentalen Einstellung, eines Begehens oder Willens dazu bedarf:

»In order to capture quantifiable information, to datafy, we need to know how to measure and how to record what we measure. This requires the right set of tools. It also necessitates a *desire* to quantify and to record. Both are prerequisites of datafication [...]« (Mayer-Schönberger & Cukier 2013: 78; eigene Hervorhebung)

Ein »Data Mindset«, d. h. ein Bewusstsein dafür, ein soziales Phänomen mit und durch Daten wahrzunehmen und zu repräsentieren, wird demnach zur Grundlage von Datafizierung und der Arbeit mit Daten insgesamt. Viele Beiträge, die sich an Mayer-Schönberger und Cukier anknüpfend mit der Datafizierung des Sozialen beschäftigen, interessieren sich primär für die inkludierenden wie exkludierenden Effekte auf das Soziale (boyd & Crawford 2012; Dalton & Thatcher 2014; Kitchin & Lauriault 2018). Die Frage jedoch, wie ein »Data Mindset«, wie eine Denkweise für Daten entsteht, was sie auszeichnet, wie und wo sie vermittelt wird und in andere gesellschaftliche Bereiche diffundiert, wird weit weniger thematisiert. Es wird gemeinhin immer angenommen und als Bedingung für Datafizierungsprozesse vorausgesetzt, ohne es selbst zum Gegenstand zu machen (Carrillo 2016).

Das »Data Mindset« ist ein schillerndes Konzept und zirkuliert in diversen Spielarten sowohl in wissenschaftlichen, ökonomischen und weiteren, oft hybriden Anwendungskontexten. Trotz seiner ambigen und flexiblen Verwendungsweise ist es ein zentraler Feldbegriff, der zur Grundlage für die Wirkmächtigkeit datenwissenschaft-

licher Praxis wird: Das Denken und Wahrnehmen mit und durch Daten ist stets mit spezifischen Praktiken – wie dem Beurteilen, Sortieren oder Klassifizieren mit und durch Daten – verknüpft, die für sich eine spezifische Objektivität und Neutralität (Porter 1995) suggerieren und in Anspruch nehmen.

In den folgenden zwei Abschnitten werde ich zeigen, inwiefern das »Data Mindset« in den Studiengängen – je nach Ausrichtung – sowohl als »forschende Grundhaltung« (Kap. 10.6.1) als auch als »unternehmerische Haltung« (Kap. 10.6.2) operationalisiert wird. Neben den Praktiken, Tools und Kompetenzbegriffen fluktuiert es als *boundary object* zwischen den gesellschaftlichen Feldern der Wissenschaft und Ökonomie und trägt insofern trotz seiner diffusen Bedeutung zu deren Verknüpfung im Bereich der Datenwissenschaften bei.

10.6.1 »Eine forschende Grundhaltung zu entwickeln ist absolut zentral!«

»Kreativität, Neugier und wissenschaftliche Denkweise fördern neuartige Erkenntnisse zu Tage« (Stockinger et al. 2016: 62).

Praxisorientierte Handbücher, die sich explizit an Data Scientists in datenverarbeitenden Organisationen richten, beschreiben den »data science process« oft als Ensemble verschiedener Teilschritte der Datenerhebung, Datenaufbereitung, Datenbereinigung, Modellierung, Analyse und Evaluation, die in einem zirkulären Verhältnis zueinander stehen (Provost & Fawcett 2013: 27; Schutt & O’Neil 2013: 41). Dies setzt allerdings voraus, dass bereits klar ist, was überhaupt die Fragestellungen sind, die sich an die »Daten« richten. Die Arbeit an und mit Daten bedarf grundlegender kontingenter Entscheidungen: So stellt sich etwa die Frage, was überhaupt »die Daten« sind. Woher stammen sie, wie setzen sie sich zusammen? Wie ist ihre Qualität? Wie kann diese bewertet werden? Je nach Beantwortung dieser Fragen und entsprechenden Praktiken resultieren andere Datenkonstruktionen, die wiederum Grundlage werden für daran anschließende Kategorisierungen oder Klassifikationen (Mützel et al. 2018). Datenwissenschaftliches Arbeiten als epistemische Praxis besteht demnach zuallererst einmal im grundlegenden Suchen und Finden von Fragen an den Gegenstand.

Als Voraussetzung, um überhaupt auf Antworten zu diesen Fragen zu kommen, erachten die Studiengänge persönliche Eigenschaften wie eine forschende Grundhaltung, Kreativität oder Neugier als weiteren elementaren Baustein eines datenwissenschaftlichen Kompetenzprofils. Wie gezeigt, wird die zentrale »Datenkompetenz« beispielsweise mit der Fähigkeit, »gute Fragen stellen« zu können, operationalisiert. Was aber sind »gute Fragen« und (wie) können diese erlernt werden? Im Gegensatz zu methodischen und analytischen Fähigkeiten, die sich als ein Bündel unterschiedlicher Praktiken im Sinne eines Handwerks erlernen lassen, sind Wissenschaftlichkeit, Kreativität oder Neugier diffuse, auf die Persönlichkeit von Studierenden abzielende Konzepte. Im empirischen Material lassen sich kaum Spezifizierungen dazu finden. Obwohl alle Befragten übereinstimmen, dass diese »absolut zentral« seien, bleibt in der Artikulation dieser Konzepte offen und unterbestimmt, wie dies gelingen kann.

Auf die Frage, was denn »eine forschende Grundhaltung« ausmacht, skizziert eine befragte Person zumindest eine Stossrichtung, wie eine solche befördert werden kann:

»Aber allgemein sind es einfach so Grundlagen, die sehr wichtig sind: Kreativität [...] wenn ich grosse, komplexe Daten habe, ist oft gar nicht so klar, welche Fragen ich an die Daten stelle, [...] also Neugier, die man einfach wecken muss. Und das kann man einfach nur im Kontext von praktischen Projekten [...] machen« (Prof_ETH_A: 4:25).

Die Zentralität des Begriffs manifestiert sich einerseits in der Bedeutungszuschreibung (»einfach so Grundlagen, die sehr wichtig sind«), andererseits aber auch im Ungesagten, d. h. im Ausweichen auf andere Begriffe (»Kreativität«, »Neugier«) sowie im Ringen nach einer Antwort, die durch die Sprechpausen indiziert werden. Der Befragte verknüpft eine »forschende Grundhaltung« mit Kreativität und Neugier – zwei ebenfalls schwierig fassbare Begriffe, die auf inkorporierte Talente und Begabungen verweisen, die die jeweilige Forscherin inkorporiert hat und gewissermassen aus sich selbst schöpfen kann (»also Neugier, die man einfach wecken muss«). In beiden Begriffen artikuliert sich eine Verwandtschaft zum künstlerischen Genie-Diskurs des 19. Jahrhunderts, der sich in seiner aktualisierten Fassung eines Kreativitätsimperativs präsentiert (Raunig et al. 2011; Reckwitz 2012; Wuggenig 2016). Die Suche nach Neuheit als datenwissenschaftliche Kompetenz erfordert somit das Überschreiten des Kalkulier- und Erwartbaren und beinhaltet insofern immer auch etwas Unbeschreibbares.

Interessant ist schliesslich, dass die interviewte Person eine »forschende Grundhaltung« insbesondere in Projekten mit Praxispartner*innen situiert: Erst durch konkrete Projekte mit wissenschaftlichen oder industriellen Partner*innen setzt ein Verständnis für die Potenzialität von Daten ein, »was mit Daten möglich ist« (Unternährer 2021). Es ist das Tun, das der Denkweise in diesem Fall vorausgeht und die Voraussetzung bildet, datenorientierte Denkweisen einzuüben und zu internalisieren.

10.6.2 Entrepreneurship als Kommodifizierung der datenorientierten Denkweise

Die Anrufung verinnerlichter Dispositionen wie Kreativität und Neugier wirkt widersprüchlich bei der gleichzeitigen Feststellung, dass das Erlernen und Einüben von Denkweisen in der Praxis erfolgt. Technikwissenschaftlich orientierte Studiengänge geben insofern keine Antworten auf das Problem, sondern verschieben die Lösung, indem sie datenwissenschaftliches Arbeiten stets als interdisziplinär und kollaborativ rahmen. Das »schöpferische Element« wird in die Aushandlung verschiedener disziplinärer Logiken transformiert.

Studienprogramme an Wirtschaftsfakultäten bzw. Business Schools wählen einen anderen Ansatz: Sie kombinieren – wie erwähnt – datenwissenschaftliche Techniken und Methoden mit verschiedenen betriebswirtschaftlichen Praxisfeldern (wie Management, Marketing, HR etc.) (Carrillo 2016). Die Studiengänge schreiben insbesondere einer datenorientierten unternehmerischen Haltung bzw. Business-Orientierung ihrer Studierenden eine grosse Bedeutung zu. Ein Interviewpartner einer Fachhochschule verknüpft auf die Frage nach der disziplinären Verortung der Datenwissenschaften methodisch-analytische Fähigkeiten mit einer unternehmerischen Grundhaltung:

»Ganz am Anfang haben wir uns hingesetzt und zusammen überlegt: was ist für uns eigentlich Data Science? Das war ein mühsamer Prozess von ein paar Wochen und wir haben gemerkt: Analytics ist wichtig. Wir können nicht von einem Data Scientist sprechen, wenn der nicht weiss, wie das geht [...] Man muss verstehen, was machen diese Verfahren, was ist das Ergebnis. [...] Es braucht unternehmerisches Denken, als Data Scientist baue ich Data Products, ich baue nicht einfach einen Algorithmus. Ich setze den ein und muss im Kopf haben, was da rauskommen könnte, d. h. da muss ich eine Idee haben. Es kommt kein Business-Mensch und sagt: ›Oh, lass uns doch mal den schicken Algorithmus mit den Daten kombinieren, um das zu bauen«. Die Idee kommt typischerweise von mir, d. h. ich brauche das Entrepreneurship [...]« (Prof_FH_B: 5:26).

Neben der bereits diskutierten interdisziplinären Konstitution der Wissensformation werden im Zitat die Kombination von mathematischen, statistischen und informatischen Grundlagen mit unternehmerischem Denken und Kreativität angesprochen. »Data Science« wird hier nicht in erster Linie als »numbers crunching«, sondern als eine kreative Tätigkeit charakterisiert, als das Generieren von »Ideen« und das Schaffen von neuen »Data Products«, etwa in Form von Visualisierungen, Applikationen oder Dashboards, die wiederum in ordentliche Geschäftsprozesse einfließen oder diese repräsentieren.

Datenorientiertes Entrepreneurship wird in diesen Studiengängen explizit nicht als Prädisposition verstanden, sondern als lernbar in verschiedenen Kursen implementiert. Anschauliche Beispiele sind Module wie »Data-Driven Business-Models«, »Managing Data Science Projects« oder »Data Ideation« an der Hochschule Luzern – Wirtschaft.⁷ Die Studierenden lernen in solchen Lehrveranstaltungen verschiedene »Innovationstechniken kennen und erkunden Ökosysteme, welche innovative Formen der Datennutzung hervorbringen«, um organisationale Problemstellungen datenbasiert zu bearbeiten. Es handelt sich um die Einführung und Vermittlung in eine datenorientierte Denkweise, wobei das Ziel – neben Analyse, Interpretation, Visualisierung und Kommunikation – darin besteht, »an die Daten die richtigen Fragen zu stellen«.⁸ Die Vorgehensweise orientiert sich explizit an einem »Design-Zyklus«: Die Studiengänge übersetzen somit exemplarisch die in den Textbüchern beschriebene Vorgehensweise (vgl. oben) datengetriebener Forschungsprozesse in organisationale Kontexte.

Die Ausbildung einer datenorientierten Denkweise zielt insofern nicht nur auf den betrieblichen Mehrwert, den sich Organisationen von »Data Science« erhoffen, sondern festigen auch die Orientierung an Transparenz und Objektivität (Porter 1995) datenbasierter gegenüber traditionellen organisationalen Entscheidungsprozessen. Genauso wie die datenwissenschaftliche Praxis selbst befördert eine datengetriebene Denkweise also, wie Mayer-Schönberger und Cukier postulieren, die digitale Trans-

7 Vgl. die Modulbeschreibungen online: <https://www.hslu.ch/de-ch/wirtschaft/studium/master/applied-information-and-data-science/module/> (Zugriff: 03.02.2022).

8 »Ziel ist es, dass die Studierenden lernen, an die Daten die richtigen Fragen zu stellen, Daten zielgerichtet zu analysieren, zu interpretieren und für die Optimierung von Design-Ergebnissen zu nutzen. Schliesslich müssen die Ergebnisse wirkungsvoll aufbereitet und kommuniziert werden. Das Modul orientiert sich eng an einem Design-Zyklus« (HSLU Master Applied Information and Data Science, Modulbeschreibung »Data Ideation«).

formation und Datafizierung von zeitgenössischen Organisationen (Beer 2019; Boes et al. 2018).⁹

10.7 Vielversprechende Zukünfte: Zur »Sexyness« von Data Scientists

»Data Science is an attractive name which makes Data Science sound young, exciting, innovative, and partially mysterious. This may endow those entering this field with a particularly creative and less conservative mindset than in other, more established disciplines« (Ley & Bordas 2018: 172).

Technologische Innovationen artikulieren stets kollektive Imaginationen und sind somit Teil von kulturellen Narrativen (Jasanoff 2015). Die zu beobachtende Konjunktur von »Data Science« erklärt sich insofern nicht alleine durch epistemologische, politische oder organisationale Transformationen, sondern operiert ebenso auf der symbolischen Ebene: Subjekte mit spezifischen Skillsets werden als zukünftige Studierende angerufen und adressiert, indem ihnen vielversprechende – berufliche wie soziale – Opportunitäten in Aussicht gestellt werden. Die kulturelle Rahmung von »Data Science« wurde fast ausschliesslich durch männliche US-amerikanische Informatiker, Statistiker und Ökonomen geprägt – mit einigen wenigen Ausnahmen wie Hilary Mason (Mason & Wiggins 2010; Patil & Mason 2015) –, die an der Schnittstelle von Technologieunternehmen und Wissenschaft ihre eigenen Visionen des Feldes formulierten und gleichzeitig in Anspruch nehmen, den Begriff »Data Scientist« überhaupt erst erschaffen zu haben (Hammerbacher 2009; Varian 2009; Davenport & Patil 2012). Sie entwarfen ein spezifisches Verständnis einer neuen Profession, das sich gegenüber existierenden Rahmungen, wie etwa dem weit verbreiteten negativen Zerrbild von »Leute[n], die irgendwo im dritten Untergeschoss von morgens bis abends vor dem Bildschirm sitzen« (Prof_FH_D: 15:65), etablieren musste.

Der Artikel »Data Scientist: The Sexiest Job of the 21st Century« des Harvard-Professors Thomas H. Davenport und des Informatikers Dhanurjay »DJ« Patil, der als erster Chief Data Scientist im Weissen Haus in der zweiten Amtszeit von Barack Obama fungierte, steht paradigmatisch für diese Konzeption von »Data Science«. Der Artikel kombiniert ökonomische Deutungen (hohe Nachfrage bei gleichzeitigem »Fachkräftemangel«) mit dem Narrativ der »Sexyness« von hochqualifizierten Professionellen in quantitativen Disziplinen. Zudem wird die Attraktivität der neuen Profession mit Werten wie Offenheit, Demokratisierung und Dezentralität in Verbindung gebracht (Cornelissen 2018; Kross et al. 2020), die einst als antikapitalistische Kritik von sozialen Bewegungen formuliert, mittlerweile aber in solutionistischen Technikutopien von Unternehmen, Nationalstaaten oder Stiftungen usurpiert wurde (Dickel & Schrape 2015; Haven & boyd 2020; Nachtwey & Seidl 2017).

9 Das CAS »Data Science and Management«, das gemeinsam von der EPFL und der Universität Lausanne angeboten wird, fasst dies unter der Losung »Transformation towards the data-driven enterprise« wie folgt: »Data is a strategic asset nowadays. Everyone agrees with the principle, but we are far away from an effective use of data in business. In order to create value from Big Data and analytics, companies need to transform into Data-Driven Enterprises« (EPFL & HEC Lausanne 2019: 2; eigene Hervorhebung).

Das durch den Artikel von Davenport und Patil popularisierte Narrativ wird nicht nur in medialen Repräsentationen von »Data Science« breit rezipiert, sondern findet auch in Studienprogrammen oder Mission Statements von Forschungsinstituten Anwendung. Die Referenz auf die hohe Nachfrage nach datenwissenschaftlichen Kompetenzen und die daraus resultierende soziale wie symbolische Attraktivität wird in den Kontext des ökonomischen Potenzials von Daten (»Daten als Rohstoffe«) gestellt, das es zunächst durch geeignete Methoden freizulegen und auszuschöpfen gilt. Es sind in erster Linie Studienprogramme in Business Schools, die auf diese Rahmung zurückgreifen. Sie kombinieren – wie bereits erwähnt – ein statistisch-technikwissenschaftliches Methodenrepertoire mit einem ausgeprägten Fokus auf Management, Kommunikation und Visualisierung:

»What kind of person does all this? What abilities make a data scientist successful? Think of him or her as a hybrid of data hacker, analyst, communicator, and trusted adviser. The combination is extremely powerful – and rare« (Davenport & Patil 2012: 73).

Die Konstruktion einer »Sexyness« kann insofern als Versuch eines revidierten gesellschaftlichen Rollenmodells für technische Berufe gelesen werden, das die traditionell männlich kodierte Ingenieurslogik durch nicht-technische, manageriale Skills zu erweitern und zu demokratisieren versucht. Damit wird eine Abkehr von den angeführten »disziplinäre[n] Eigenbrötler[n]« bzw. dem »Informatiker-Autismus« angestrebt, die stellvertretend für »Silodenken« und kommunikative Inkompetenzen erhalten müssen.

Ein geschlechtersensibles *close reading* des Narrativs impliziert allerdings, dass die »Sexyness« über die Attraktivität von hoher Nachfrage und Entlohnung hinausgeht und auf die inhärent vergeschlechtlichte Dimension des Berufsbildes verweist: Wie andere technologiegetriebene Felder bestehen die Datenwissenschaften in der Schweiz gegenwärtig fast ausschliesslich aus männlichen Professoren, die eine mehrheitlich männliche Studierendenschaft unterrichtet. Trotz wissenschaftspolitischer Massnahmen, die Partizipation und Inklusion von Frauen in den Technikwissenschaften und technischen Berufsfeldern insgesamt zu erhöhen (Akademien der Wissenschaften Schweiz 2017; Hasler Stiftung 2018), zeigt sich eine signifikante, sozialstrukturell persistente Unterrepräsentation von Frauen in Engineering- und Informatik-Berufen (Andreoli et al. 2017; Chow & Charles 2019) und neulich auch in den Datenwissenschaften (Duranton et al. 2020; Norén et al. 2019).¹⁰

Sie manifestiert sich im empirischen Material insbesondere darin, dass Data Scientists in den Interviews unisono als »Jack« und nicht als »Jill of all trades and a master of some« adressiert und imaginiert werden. Die selbstverständliche Anrufung als männliche Subjekte (»Weil er heisst ja auch Data Scientist, d. h. er muss auch gute Fragen stellen können«, vgl. Kap. 10.3) widerspricht dabei der visuellen Repräsentation von zukünftigen Data Scientists in vielen untersuchten Materialien und Webseiten, die eine diversere Bildsprache einzunehmen versuchen und über die Abbildung

¹⁰ Zu den Geschlechterverhältnissen wissenschaftlicher Fachbereiche in der Schweiz vgl. Bundesamt für Statistik (2021): Frauen und Wissenschaft. <https://www.bfs.admin.ch/bfs/de/home/statistiken/bildung-wissenschaft/technologie/indikatorsystem/zugang-indikatoren/w-t-input/frauen-und-wissenschaft.html> (Zugriff: 03.02.2022).

mehrheitsschweizerischer, männlicher und weisser Studierendensubjekte hinausgehen. Die Analyse deutet insofern darauf hin, dass sich die Selbstdarstellung der Studiengänge unter dem Einfluss von wissenschaftspolitisch induzierter Inklusion und Internationalisierung nach aussen diversifiziert, während die grundlegenden Orientierungen der Befragten nach wie vor Ausdruck eines trägen, geschlechtsspezifisch strukturierten Habitus (Bourdieu 1997a) sind. Während die Berufsfelder als stark von einem »Fachkräftemangel« betroffen gerahmt werden, der durch die Inklusion weiterer, bisher unterrepräsentierter Gruppen überwunden werden soll, erschweren bzw. verhindern normative Strukturen und Orientierungen weiterhin den Eintritt und die Karrieremöglichkeiten von Frauen (Cech 2014; Charles & Thébaud 2018; Chow & Charles 2019).

Trotz Versuchen der Rekonzeptualisierung und Resituierung schliessen die dominierenden Repräsentationen von Datenwissenschaften an jene »technikwissenschaftliche Business Masculinity« an, die den »Kontext einer entgrenzten Wissenschaft und eines als Wissensökonomie zu bezeichnenden Feldes im Schnittfeld von Wissenschaft und Wirtschaft« (Paulitz & Prietl 2017: 164) charakterisiert.¹¹ Die organisationale Rolle von Data Scientists als Kommunikator*innen, Übersetzer*innen und Manager*innen stellt spezifische Anforderungen nicht nur an vermitteltes und distribuiertes wissenschaftlich-technisches Kapital, sondern auch an kulturelles und soziales Kapital. Die identifizierte vergeschlechtlichte Konstruktion datenwissenschaftlicher Kompetenzprofile reproduziert insofern bestehende Disparitäten und soziale Ungleichheiten.

10.8 Diskussion

Die Analyse macht deutlich, dass sich sowohl die Lehrenden als auch die Curricula elementar auf den Kompetenzbegriff stützen. Mittels Kompetenzdefinitionen und Berufsprofilen wird versucht, eine Rezeptur für die Komposition eines »Data Science Skillsets« zu finden und begründen: Im Baukastenprinzip werden unterschiedliche Anforderungen, Qualifikationen und Fähigkeiten für ein Studium oder eine berufliche Tätigkeit in Datenwissenschaften arrangiert und kombiniert. Während die methodisch-analytischen Fähigkeiten am ehesten zur Ausbildung eines distinkten Profils taugen, sind andere Elemente stärker additiv konzipiert und generisch für viele Berufsgruppen. Da sich zudem im empirischen Material divergierende Vorstellungen darüber finden, welche Bausteine des datenwissenschaftlichen Kompetenzprofils die zentralen sind und wie diese in Curricula organisiert werden sollen, ist kaum eine Hierarchisierung möglich.

Such- und Aushandlungsprozesse über die »richtigen« Kompetenzrahmen und Berufsprofile von Data Scientists auf der Ebene von Studiengängen, Curricula und Berufsklassifikationen bilden in diesem Sinne die Differenzierungsprozesse und die epistemologischen, politischen und ökonomischen Diskussionen auf der Makroebene ab. Es handelt sich um organisationale Antworten auf jene kollektiven Unsicherheiten,

11 Tatsächlich birgt die Mehrdeutigkeit von »Data Science« auch Konturen von Gegenbewegungen: Insbesondere im US-amerikanischen Kontext wurden Gender, Race, Diversity und Social Justice schon früh als Prioritäten definiert und teilweise in Fragen der Kanonbildung miteingezogen (NASEM 2018: 4-1f.).

die sich in gesellschaftlichen Zukunftsdebatten über Digitalisierung und Datafizierung wissenschaftlicher Erkenntnisproduktion, ökonomischer Praktiken oder politischer Prozesse artikulieren.

Analog zur Suche nach adäquaten Begriffen in der Beschreibung der epistemologischen Verschiebungen, die durch die Datenwissenschaften manifest werden, steht die Suche nach den ›richtigen‹ Kompetenzprofilen für Praktiken von Begriffs- und Grenzarbeit, die zur Konstruktion eines Raumes zwischen etablierten Feldern beitragen: Trotz ihrer Vagheit und Unterbestimmtheit formulieren die verwendeten Begrifflichkeiten wie Datenkompetenzen oder »Data-Mindset« Diagnosen und Vorstellungen über den Zustand und die weitere Entwicklung des Feldes. Die damit einhergehenden Bedeutungszuschreibungen verdeutlichen den hohen ökonomischen und politischen Stellenwert, den solche Konstruktionen und Zuschreibungen im Untersuchungsfeld einnehmen. Sie mäandrieren zwischen den verschiedenen involvierten Disziplinen und Feldern und garantieren gleichzeitig, dass die Grenzen dazwischen bestehen bleiben. Ihre Überführung in mess- und überprüfbare Kompetenzen befördert und festigt die Selbst- und Fremdwahrnehmung der Datenwissenschaften als einer Wissensformation in Entstehung. Parallel dazu wird durch die Rekonzeptualisierung der Datenwissenschaften versucht, die Wissensformation zu öffnen und insofern anschlussfähig an veränderte Wertvorstellungen zu machen. Die verwendeten Narrative verweisen allerdings auf tradierte normative Ausbildungs- und Berufsstrukturen, die sich in ausgeprägt maskulinen Imaginationen des Berufsbildes Data Scientist wiederfinden. Diese erschweren Partizipationsmöglichkeiten insbesondere für Frauen, was zur Reproduktion sozialer Ungleichheiten im entstehenden Wissensfeld beiträgt.

Hinsichtlich ihrer Leistung als allgemeine Konstruktionsmechanismen nehmen die identifizierten Kompetenzprofile und individuellen Zuschreibungen eine doppelte Funktion wahr: Einerseits übersetzen sie methodisch-technische Innovationen in operationalisierbare Anforderungen an gegenwärtige und zukünftige Praktiker*innen auf der Mikroebene. Damit schaffen sie ein Dispositiv, durch welches Neuheiten vermittelt werden und in weitere gesellschaftliche Bereiche jenseits des zwischenräumlichen Entstehungskontextes diffundieren können. Gleichzeitig eröffnen sie dadurch vielfältige Anschlussmöglichkeiten für multiple Akteur*innen, die Kompetenzprofile aufnehmen, weiterentwickeln und insofern zu deren Festigung beitragen.

Andererseits rahmen sie Neuheiten durch spezifische Narrative, indem sie Subjekte mit bestimmten Skillsets als zukünftige Praktiker*innen adressieren und ihnen vielversprechende – berufliche wie soziale – Zukünfte in Aussicht stellen. Allerdings sind solche kulturellen Rahmungen nie neutral, sondern stehen in Beziehung zu existierenden Deutungen verwandter Gegenstände, zu denen sie sich – positiv-affirmierend, negativ-abgrenzend oder dazwischen vermittelnd – verhalten müssen. Während dies produktive Dynamiken auslösen und die Nachfrage nach bestimmten Kompetenzprofilen beschleunigen und ankurbeln kann, sind umgekehrt stets auch Prozesse von Selektion und Fokussierung auf bestimmte Aspekte und damit Exklusion anderer verbunden. Als Konstruktionspraktiken zur Übersetzung von Innovationen in individuelle Zuschreibungen machen Kompetenzprofile widersprüchliche Effekte von Inklusions- und Exklusionsprozessen sichtbar.

Teil IV - Schlussbetrachtungen

Kapitel 11 – Synthese

Der letzte Teil bildet eine Synthese der Resultate der empirischen Untersuchung mit den verschiedenen theoretischen Bausteinen des analytischen Modells sowie der methodischen Vorgehensweise. Im Zentrum stehen die Rekapitulation und Diskussion der zentralen Ergebnisse sowie deren Einbettung in das analytische Modell von Räumen zwischen Feldern vor dem Hintergrund des Erkenntnisinteresses und der Fragestellungen der Arbeit. Als Erstes beschreibe ich die in den empirischen Kapiteln identifizierten Praktiken von Begriffs- und Grenzarbeit und untersuche deren Beitrag zur transversalen Konstruktion der Datenwissenschaften über verschiedene Felder hinweg (Kap. 11.1). Daran anknüpfend werden zweitens Widersprüche in der Herausbildung der Datenwissenschaften manifest, die in synchronen Such- und Bestimmungsprozessen einerseits und Institutionalisierungs- und Kanonisierungsprozessen andererseits resultieren und für die soziale Öffnung bzw. Schliessung des sich etablierenden Wissensgebiets stehen (Kap. 11.2). Anschliessend diskutiere ich drittens die synchronen Relationen von Konkurrenz und Kollaboration, die die Datenwissenschaften als Raum zwischen etablierten sozialen Feldern strukturieren (Kap. 11.3). Ein vierter Punkt betrifft die zeitliche Dimension, da die Datenwissenschaften in allen untersuchten Feldern in einer beschleunigten Art und Weise als neue Kategorie bearbeitet wurden (Kap. 11.4). In einem fünften Schritt gehe ich sodann auf die Frage ein, inwiefern sich die Datenwissenschaften als hybride Wissensformation dank datengestützter Denk- und Wahrnehmungsweisen derart rasch und breit etablieren konnten (Kap. 11.5). Abschliessend diskutiere ich die Generalisierbarkeit des analytischen Modells, skizziere mögliche Anwendungsbeispiele über den konkreten Untersuchungsfall der Datenwissenschaften hinaus sowie inhaltliche und methodische Erweiterungen (Kap. 11.6).

11.1 Transversale Konstruktionen der Datenwissenschaften durch Praktiken von Grenz- und Begriffsarbeit

Die verteilte empirische Analyse der drei Untersuchungsfelder Arbeitsmarkt, Hochschul- und Forschungspolitik sowie Universitäten und Hochschulen fördert als erstes zentrales Ergebnis eine Vielfalt von Akteur*innen zutage, die in der Konstruktion, Verbreitung und Etablierung der Datenwissenschaften als Wissensfeld involviert sind. Dies zeigt sich zunächst in der elementaren Frage nach Herkunft bzw. Bedeu-

tung des Begriffs: Dem medial vermittelten Narrativ, das den Entstehungskontext in Unternehmen der amerikanischen Technologieindustrie im Silicon Valley lokalisiert, steht eine jahrzehntelange Vermischung und Vermengung der Praktiken, Interessen und Strategien von Akteur*innen in Wissenschaft, Industrie und Forschungspolitik gegenüber. Zugleich indiziert die Begriffskonstruktion »Data Science« eine primäre Verankerung in der Wissenschaft, die in dieser Reinform nie existierte und die längst transzendiert wurde, indem Methoden, Wissensbestände und Einstellungen in die unterschiedlichsten sozialen Praxisfelder diffundiert sind. Insofern macht bereits die gewählte Begrifflichkeit auf die synchrone Herausbildung der Datenwissenschaften aufmerksam.

Entsprechend diesen vielfältigen Entstehungskontexten zeigt die vorliegende Analyse anhand der untersuchten Stelleninserate, Strategiedokumente der Hochschul- und Forschungspolitik sowie der Curricula und Interviews mit Professor*innen, dass es nicht die eine Auslegung gibt, sondern vielmehr verschiedene Verständnisse, Interpretationen sowie organisationale und curriculare Umsetzungen miteinander koexistieren. Die so identifizierten Bedeutungskonstruktionen verweisen auf die elementare Rolle von zwei zentralen Praxismodi in der transversalen Herausbildung der Datenwissenschaften: Erstens generieren und befördern die Akteur*innen in den untersuchten Feldern durch Praktiken von Begriffsarbeit multiple Verständnisse der Inhalte und der Zusammensetzung des Wissensgebiets. Charakteristisch für den Begriff »Data Science« sind breit formulierte, universelle Definitionen, denen unterschiedliche Felder, Disziplinen oder Professionen zustimmen können. Durch sprachliche Festlegungen versuchen die beteiligten Akteur*innen eine bestimmte Deutung des Phänomens zu fixieren und privilegieren diese gegenüber anderen möglichen Bedeutungen. Als Produkte kollektiver Aushandlungsprozesse in Unternehmen, Universitäten oder Institutionen der Hochschul- und Forschungspolitik eröffnen Stelleninserate, Strategiedokumente oder Curricula somit ein breites »Möglichkeitsfeld« (Bourdieu 1987: 188), was alles unter »Data Science« subsumiert werden kann bzw. was nicht, und erweitern so die Handlungsoptionen für andere – individuelle wie kollektive – Akteur*innen.

Zweitens beruhen solche diskursiven Praktiken auf symbolischen und sozialen Grenzziehungen: So ziehen Akteur*innen in den untersuchten Feldern diskursive Grenzen zwischen den Datenwissenschaften und anderen Wissensbereichen wie Informatik und Statistik, was beispielsweise in eingeschränkten Handlungsoptionen oder mangelnden Kooperationsmöglichkeiten münden kann. So existieren an einigen Universitäten die neu geschaffenen Einheiten parallel zu bestehenden Gefäßen weiter, was mitunter in einer eingeschränkten Bedeutung für beide resultieren kann. In der Mehrheit der untersuchten Fälle transzendieren die Studiengänge und Forschungsschwerpunkte jedoch solche etablierten Grenzen und halten die neu geschaffenen Gefäße diskursiv offen und durchlässig, um damit disziplinenübergreifende Kooperationen zu ermöglichen.

Die Analyse weist im Falle der Datenwissenschaften auf diverse Interdependenzen von Begriffs- und Grenzarbeit hin: Im Falle des Arbeitsmarkts hat das Topic Modeling der Stelleninserate unterschiedliche Konnotationen der Datenwissenschaften identifiziert, die die inserierenden Organisationen je nach Tätigkeitsfeld durch distinkte Kombinationen von Begriffen, Methoden, Analyseinstrumenten oder Qualifikationsanforderungen repräsentieren. Die Listen und Aufzählungen äquivalenter

Alternativen signalisieren dabei primär Inklusivität und eröffnen Möglichkeitsräume für potenzielle Bewerber*innen auch aus anderen Tätigkeits- und Herkunftsfeldern. Insofern dominieren in Stelleninseraten Praktiken von Grenzüberschreitung, die das soziale Phänomen Datenwissenschaften als offen, flexibel und adaptiv konstruieren.

Ähnlich artikulieren Akteur*innen im Feld der Hochschul- und Forschungspolitik kollektive Visionen und Szenarien, wie durch die engere Verknüpfung von wissenschaftlich-technologischem Fortschritt einerseits und Hochschulbildung, Forschung und Ökonomie andererseits nationalstaatliche Ziele wie die Erhaltung des materiellen Wohlstands oder der Wettbewerbsfähigkeit von Bildung und Forschung erreicht werden sollen. Praktiken von Begriffs- und Grenzarbeit präsentieren sich als inhärent miteinander verknüpft: Die Offenheit und Ambiguität zentraler Konzepte wie »Digitalisierung«, »Kompetenzen« oder »Innovation« macht die untersuchten Strategien zugänglich für multiple Interpretationen und Bedeutungskonstruktionen und somit anschlussfähig für unterschiedliche Akteur*innen. Sie zielen insbesondere darauf ab, eine Vielzahl von Akteur*innen verschiedener Felder zu involvieren und symbolische wie soziale Grenzen zwischen den Feldern zu überschreiten. Durch die Massnahmen und Investitionen im Rahmen der Strategie »Digitale Schweiz« sowie der neuen »Datenpolitik« werden die Datenwissenschaften als Wissensgebiet stabilisiert, finanziell alimentiert und in den Dienst nationalstaatlicher Interessen gestellt.

Im akademischen Feld wiederum koexistieren verschiedene Begriffe (wie »Big Data«, »Data Science« oder »Data Analytics«) miteinander, die nur bedingt voneinander differenziert werden, was zunächst ebenfalls eine inklusive Konzeption des Wissensgebiets ermöglicht. Bisweilen werden solche symbolisch-begrifflichen Grenzbeziehungen in soziale Grenzen übersetzt, etwa wenn neue Studiengänge oder interdisziplinäre Forschungszentren lanciert werden, die neben der Inklusion bzw. Exklusion bestimmter Disziplinen und Personen mit der Zuteilung von finanziellen Mitteln einhergehen. Diese stehen dann wiederum anderen organisationalen Einheiten nicht zur Verfügung. Die Analyse hat auch gezeigt, dass die neuen Gefässe oft von existierenden disziplinären Konkurrenzverhältnissen (insbesondere zwischen Statistik und Informatik) geprägt sind, die durch die Emergenz und Implementierung der Datenwissenschaften im akademischen Feld eine erneute Virulenz erhalten. Allerdings werden solche Konflikte in vielen Fällen strategisch latent gehalten und es wird Einheit nach aussen signalisiert, um die sich eröffnenden Opportunitäten wie zusätzliche Finanzmittel, organisationales Wachstum oder feldübergreifende Kollaborationen nicht aufs Spiel zu setzen.

Die identifizierten Konstruktionsleistungen in den drei Untersuchungsfeldern verbindet insofern der Umstand, dass die Praktiken von Begriffs- und Grenzarbeit transversal angelegt und ineinander verschränkt sind: Durch Bezüge auf Konzepte, Praktiken und Wissensinhalte multipler Felder übernehmen die untersuchten Stelleninserate, Strategiedokumente und Curricula nicht nur bestimmte Begriffe und Kategorien, sondern etablieren auch Beziehungen zwischen den einzelnen Bereichen, überschreiten etablierte Grenzen und schaffen so Möglichkeitsräume für vielfältige Kooperationen zwischen den involvierten Akteur*innen. Somit resultieren die Datenwissenschaften durch die jeweiligen Konfigurationen in Arbeitsmarkt, Hochschul- und Forschungspolitik, Wissenschaft und Hochschulbildung in der Summe als transversaler Raum zwischen den verschiedenen gesellschaftlichen Praxisfeldern.

11.2 Widersprüche in der Herausbildung der Datenwissenschaften zwischen Innovation und Reproduktion

Ein weiteres zentrales Ergebnis der Arbeit sind widersprüchliche Prozesse, die die feldübergreifende Emergenz der Datenwissenschaften charakterisieren. In den empirischen Analysen manifestiert sich dies in eigentümlichen Gleichzeitigkeiten von Such- und Bestimmungsprozessen einerseits und Institutionalisierungs- und Kanonisierungsprozessen andererseits: So wird auf der einen Seite die beschriebene Herausbildung der Datenwissenschaften und damit einhergehend die Transformation existierender disziplinärer und organisationaler Gefässe und Strukturen erkennbar. Andererseits deutet die Analyse der Implementierung von Studiengängen sowie der curricularen Inhalte auf die Institutionalisierung, Perpetuierung und Reproduktion disziplinärer Wissenskulturen in organisationalen Einheiten und Settings hin. Die Synchronizität von Neuheit und Veränderung einerseits, Verfestigung und Reproduktion andererseits durchzieht einem roten Faden gleich die einzelnen Untersuchungsfelder.

Zunächst wird diese Widersprüchlichkeit im schweizerischen Arbeitsmarkt anhand von divergierenden Repräsentationen deutlich, durch die bestimmte Sichtweisen auf das Phänomen Datenwissenschaften konstruiert werden: So steht Annahmen einer universalen Diffusion der Datenwissenschaften oder der Durchdringung aller Lebensbereiche eine ausgeprägte Konzentration der Stelleninsetrate auf einige wenige ökonomische Felder und geographische Standorte in der Schweiz gegenüber. Während dies einerseits als Indikator für eine ungleich distribuierte und fortschreitende Datafizierung ökonomischer und anderer Praxisphären verstanden werden kann, relativiert sich dadurch andererseits die beschriebene diskursive Konstruktion als strategischer Wissensbereich mit hoher Nachfrage und andauerndem »Fachkräftemangel«.

Des Weiteren manifestieren sich in den untersuchten Stelleninseraten anhand der Aufzählungen von Qualifikationen, Wissensbeständen, Methoden und Tools Suchprozesse über die richtigen und adäquaten Begrifflichkeiten. Auch hier erweist sich die Offenheit und Ambiguität der verwendeten Begriffe als konstitutiv für das Arbeitsmarktsegment Datenwissenschaften, indem ein breites Segment von Praktiker*innen angesprochen werden soll. Synchron dazu deutet das Topic Modeling auch an, dass sich in Stelleninseraten gewisse Begriffe etabliert haben, die feld- und sprachübergreifend genutzt werden, obwohl in den Listen unterschiedliche Praktiken, Instrumente oder Konzepte zusammengefasst sind, die ein teilweise äusserst breites Spektrum abdecken und arbiträr erscheinen. Daraus resultiert und kennzeichnet eine zunehmende Spannung zwischen Prozessen sozialer Öffnung bzw. Schliessung die Repräsentation der Datenwissenschaften im Arbeitsmarkt.

Im Feld der Hochschul- und Forschungspolitik führen Strategiedokumente Begriffe und Kategorien wie »Digitalisierung«, »Innovation« oder »Kompetenzen« ein, die an existierende Diskurse der Wissens- und Informationsgesellschaft anschliessen, diese aber auch neu kontextualisieren. Akteur*innen der Hochschul- und Forschungspolitik skizzieren darin Szenarien einer vielversprechenden Zukunft, wie sich durch die beschleunigte Adaption und Diffusion datenwissenschaftlicher Analysemethoden und Instrumente, aber auch durch ein spezifisches datengestütztes Mindset (vgl. unten) aktuelle gesellschaftliche Herausforderungen bearbeiten und zudem mone-

tarisieren lassen. Trotz der neuen Begrifflichkeiten sowie einer ubiquitären Transformationsrhetorik verweist die Inhaltsanalyse der Strategiedokumente und Berichte gleichzeitig auf eine bemerkenswerte Permanenz etablierter politischer Konzepte wie »Wohlstand«, »Standortwettbewerb« oder »Fachkräftemangel«. Das Sprechen über kontinuierliche Transformation und Innovation wird dabei selbst zu einem elementaren Bestandteil des politischen Diskurses, der sich in seinen Zielsetzungen und gewählten Mitteln jedoch zentral an Kontinuität und Stabilität orientiert.

Im akademischen Feld wiederum verweist die Inhaltsanalyse der Curricula und Interviews auf widersprüchliche Entwicklungen bei der Implementierung von Studiengängen der Datenwissenschaften an Schweizer Universitäten und Hochschulen: So wird in wissenschaftlichen wie nichtwissenschaftlichen Publikationen, Konferenzen und anderen Diskussionsformaten über die Frage gestritten, welche disziplinären Inhalte, Methoden oder persönlichen Eigenschaften Teil der Datenwissenschaften sein sollen. Daneben artikulieren Unternehmen, aber auch Organisationen der Hochschul- und Forschungspolitik nicht nur eine hohe Nachfrage nach Data Scientists, sondern auch das Bedürfnis nach klar abgrenzbaren und definierten Fähigkeits- und Qualifikationsprofilen. Ferner sind auf der Ebene der Curricula anhaltende Suchprozesse und ein »Ringeln« nach den adäquaten Begrifflichkeiten und der Zusammensetzung von Kompetenz- bzw. Qualifikationsprofilen zu beobachten.

Als Lösung für die identifizierten Suchprozesse präsentieren die Universitäten und Hochschulen die Modularität der Studiengänge, die auf den Baukasten als Organisationsprinzip zeitgenössischer Curricula verweist. Die neu geschaffenen Curricula und Studiengänge sind durch spezifische Kombinationen epistemischer Wissensinhalte mit einem sogenannten Kernbereich, Anwendungsfeldern (*domains*) sowie »komplementären Inhalten« strukturiert. Damit bieten die Studiengänge eigene Antworten auf die kollektiven Suchprozesse nach der Bedeutung, Abgrenzung und Verortung der Datenwissenschaften. Diese curriculare Strukturlogik, die eng an technikkwissenschaftliche Ausbildungstraditionen angelehnt ist, macht somit bereits in einer frühen Phase der Institutionalisierung der Datenwissenschaften im akademischen Feld Kanonisierungsprozesse manifest.

Die weitere Entwicklung wird zeigen, ob und inwiefern sich die Datenwissenschaften im Spannungsverhältnis von Innovation und Reproduktion bewegen, ob sie ihre Offenheit bewahren oder sich stärker »disziplinieren« und zu sozialer Schliessung tendieren werden. Die Implementation der Datenwissenschaften in den drei Untersuchungsfeldern hat jedoch trotz der empirisch zu beobachtenden Widersprüche produktive Dynamiken in Gang gesetzt und Feldgrenzen überschreitende Kollaborationen begünstigt.

11.3 Konkurrenz- und Kooperationsverhältnisse strukturieren die Datenwissenschaften als Raum zwischen Feldern

Die Transversalität der Datenwissenschaften wird nicht nur bei deren Herausbildung, sondern auch bei der Situierung zwischen verschiedenen Feldern sichtbar: So umfassen die Datenwissenschaften Akteur*innen, Wissensbestände, Methoden, Tools, zugeschriebene Fähigkeiten (»Kompetenzen«), Qualifikationen und Persönlichkeitsmerkmale diverser disziplinärer und sozialer Felder. Das Konvolut unterschiedlicher

Expertisen und Mindsets eröffnet jenen Zwischenraum (Eyal 2013b; Furnari 2014), der durch poröse Grenzen Kollaborationen unterschiedlicher Akteur*innen und multiple Rollen (als Forscher*innen, Unternehmer*innen, Aktivist*innen, Hacker*innen etc.) ermöglicht.

Die Analyse hat offengelegt, inwiefern der Zwischenraum der Datenwissenschaften durch synchrone Konkurrenz- und Kooperationsverhältnisse strukturiert wird. Deutlich wird dies insbesondere im akademischen Feld, in dem eine Vielzahl von Akteur*innen mit unterschiedlichen Interessen und Strategien differenziert werden kann. Die beiden ETH, die Universitäten und Fachhochschulen sowie weitere Akteur*innen positionieren sich durch Profilbildung und neue Forschungsschwerpunkte, Studiengänge oder Weiterbildungsangebote im entstehenden Zwischenraum. Sie stützen sich dazu einerseits auf ihre institutionellen Traditionen, inhaltlichen Schwerpunkte und disziplinären Aufstellungen. Andererseits operieren sie durch Praktiken von Begriffs- und Grenzarbeit, indem sie bestimmte begriffliche Festlegungen (wie etwa die Benennung von Studienprogrammen) treffen, die sie durch spezifische Kombinationen interdisziplinärer Expertisen, d. h. von Wissensinhalten, Methoden, Tools und Lehrveranstaltungen, in Form von Curricula implementieren. Damit entwickeln sie distinkte Profile, die sie aktiv von denjenigen anderer Akteur*innen abgrenzen, und tragen so zur Differenzierung der Datenwissenschaften im akademischen Feld bei.

Parallel zu solchen Bemühungen um Abgrenzung und Differenzierung sind im akademischen Feld zahlreiche Formen disziplinärer, intra- und interorganisationaler Kooperation zu beobachten. Zusammenarbeit und Grenzüberschreitung sind oft von Beginn zentral in die neuen Gefässe wie Studiengänge oder Forschungszentren eingeschrieben. Dabei entwerfen die involvierten Akteur*innen ihre eigenen Visionen, also darüber, was im Zwischenraum verhandelt wird, wer dazu gehören soll und wer nicht oder welche Formen der Kooperation möglich sind. Das Primat der Offenheit und Inklusion, das diesen Raum charakterisiert, eröffnet dabei Anschlussfähigkeit für verschiedene Akteur*innen, Disziplinen und Expertisen der umgebenden etablierten Felder.

Ferner hat die Analyse auch eine wechselseitige Orientierung an anderen Akteur*innen aufgezeigt: Die Orientierung an Vorbildinstitutionen, insbesondere an renommierten US-amerikanischen Forschungsuniversitäten, an Empfehlungen von Fachgesellschaften sowie an Rankings trägt zur Verortung in einem rapide expandierenden Teil des Hochschulfeldes bei und reduziert damit assoziierte Unsicherheiten. Solche wechselseitigen Beobachtungsverhältnisse begünstigen somit die Angleichung von Planungs- und Implementierungsaktivitäten der unterschiedlichen Akteur*innen im Feld, was insbesondere in der Herausbildung strukturell ähnlicher Curricula resultiert. Sie entfalten eine koordinierende und strukturierende Wirkung und unterstützen darüber hinaus die Integration der Studiengänge an schweizerischen Hochschulen und Universitäten in das entstehende globale akademische Feld der Datenwissenschaften.

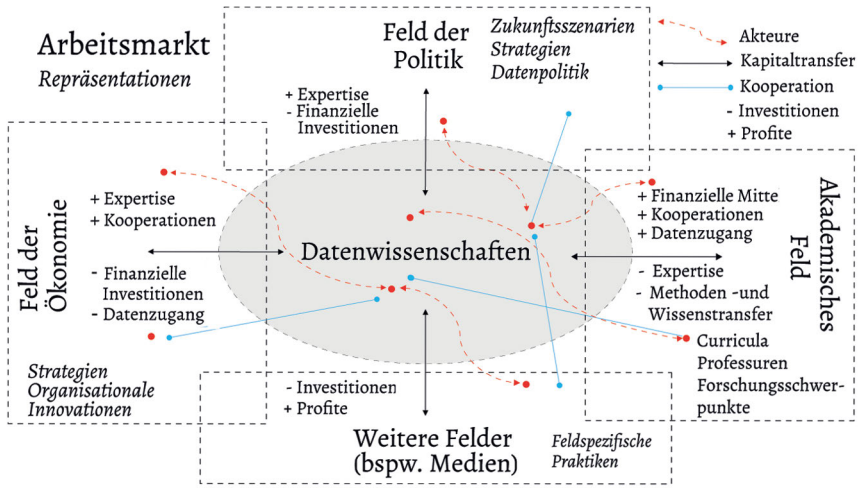
Die Datenwissenschaften resultieren somit nicht als ein durch unterschiedliche Kapitalausstattung und -verhältnisse strukturiertes hierarchisches Feld, in dem sich die dominanten Akteur*innen gegenüber den Herausforderern durchsetzen können (beispielsweise die Informatiker*innen gegen die Statistiker*innen, etablierte Technologieunternehmen gegenüber Start-ups oder die ETH gegen die Fachhochschu-

len), sondern vielmehr als ein multiple Akteur*innen und Disziplinen umfassender Zwischenraum, in dem Konkurrenz-, Kooperations- und Beobachtungsverhältnisse parallel zueinander existieren (vgl. Abbildung 11). Entsprechend profitieren viele Akteur*innen von der Offenheit und Durchlässigkeit des Zwischenraumes und haben ein Interesse daran, diese spezifische Konfiguration möglichst durabel zu halten. Sie können sich für temporär begrenzte Kooperationen oder dauerhaft in den Raum begeben und so ihre Positionen in den etablierten Feldern erhalten oder verbessern.

Da im Zwischenraum nur geringe Eintrittshürden existieren, eröffnet dies neue Opportunitäten und Möglichkeitsräume: So bietet sich etwa Universitäten und Hochschulen die Gelegenheit, durch neue Professuren und Forschungsschwerpunkte, die zunächst eigene finanzielle Investitionen bedingen, zusätzliche Forschungsgelder und Drittmittel zu akquirieren; auch etablieren sie neue Studiengänge und ziehen damit zusätzliche Studierende an. Beides ermöglicht organisationales Wachstum. Parallel erhalten sie durch Kooperationen mit Unternehmen und anderen Organisationen Zugang zu »echten« Datenquellen (*real-world data*), was die Erarbeitung neuer Modelle und Verfahren begünstigt und somit wissenschaftlich-technologisches Kapital ermöglicht, das sich wiederum in zusätzlichen finanziellen Mitteln auszahlen kann. Dazu müssen Universitäten und Hochschulen umgekehrt Expertisen mit den Kooperationspartner*innen teilen und so Methoden und Wissen in andere Organisationen bzw. Felder transferieren.

Die Analyse der Konfiguration im akademischen Feld hat auch gezeigt, dass Hochschulen und Universitäten in ihren jeweiligen Tätigkeitsfeldern eine Aussenwirkung erzielen und so anderen Akteur*innen signalisieren, dass sich Kollaborationen und Investitionen in die Datenwissenschaften lohnen. Diese werden dadurch motiviert, sich ebenfalls in den Zwischenraum zu begeben, aktiv zu werden, mögliche Kooperationen aufzubauen und zu investieren. Davon können auch kollektive Akteur*innen profitieren, die über weniger Möglichkeiten für finanzielle Investitionen verfügen und sich entsprechend positionieren und distinkte Strategien ausgebildet haben. So kann beispielsweise das diskutierte Datalab trotz seiner schwachen Institutionalisierung innerhalb einer grossen Fachhochschule, beschränkten Ressourcen und disziplinären Konfliktlinien dennoch intra- und interorganisationale Kooperationen über disziplinäre Felder hinweg stimulieren, die in erweiterten persönlichen Netzwerken, Datenzugang in externen Organisationen und schliesslich in materiellen Effekten (beispielsweise der Drittmittelinwerbung) münden. Indem die Verortung in den Datenwissenschaften zum gemeinsamen inklusiven Merkmal wird, eröffnet der Zwischenraum somit trotz persistenter Konkurrenzverhältnisse jenen Möglichkeitsraum, in welchem durch relativ geringe personale und wissenschaftlich-methodische Investitionen hohe Profite in Form von ökonomischem, sozialem und symbolischem Kapital möglich werden.

Abbildung 11: Die Datenwissenschaften als transversaler Raum zwischen Feldern



11.4 Die beschleunigte Verarbeitung neuer sozialer Kategorien

Ein weiterer Punkt betrifft die Temporalität des Phänomens (Adams et al. 2009; Wajcman 2019): Die Herausbildung der Datenwissenschaften ist in allen untersuchten Feldern durch Geschwindigkeit und Beschleunigung gekennzeichnet (Beer 2017, 2019). Während sich vergleichbare hybride Wissensgebiete wie Material- (Bensaude-Vincent 2001; Bensaude-Vincent & Hessenbruch 2004), Nano- (Biniok 2013) oder Klimawissenschaften (Allan 2017; Edwards 2001) über mehrere Jahrzehnte ausdifferenziert haben, vollzieht sich die Genese der Datenwissenschaften und vor allem deren Adaption in den untersuchten Feldern in einer beschleunigten Art und Weise: Organisationen im ökonomischen Feld datafizieren ihre Abläufe und Entscheidungsprozesse, schaffen neue »Data Science«-Einheiten und rekrutieren Data Scientists, um ihre Frage- und Problemstellungen datengestützt zu bearbeiten.

Die Hochschul- und Forschungspolitik hat in beschleunigter Weise auf die wissenschaftlichen und ökonomischen Diskurse über die Bedeutung der Datenwissenschaften reagiert und vergleichsweise schnell umfangreiche Investitionen in Förderprogramme und neue organisationale Einheiten gesprochen. Durch Strategien wie die »Digitale Schweiz« oder Prognosen über die zukünftige Entwicklung bestimmter Technologiefelder haben hochschul- und forschungspolitische sowie ökonomische Akteur*innen normativen Handlungsdruck erzeugt, um Antworten auf die Herausforderungen komplexer soziotechnischer Phänomene wie Digitalisierung, Datafizierung oder künstliche Intelligenz und daraus resultierende gesellschaftliche Unsicherheiten zu geben und die Ausbildung der damit zusammenhängenden Expertisen sicherzustellen.

Die in Strategien und Zukunftsprognosen diskursiv eingesetzten Mittel warnen entsprechend vor Wohlstandsverlusten, mangelnder Wettbewerbsfähigkeit oder einem »Fachkräftemangel«. Als soziotechnische Imaginationen (Jasanoff 2015) schaffen solche Strategiepapiere allerdings nicht nur kontingente und fiktive Zukunftsent-

würfe und -szenarien, sondern zeitigen ebenso strukturierende Effekte auf die Strategien und Investitionen anderer Akteur*innen in der Gegenwart: Das Festlegen von strategischen Zielsetzungen und Leitprinzipien schafft Erwartungssicherheit und ermöglicht es Akteur*innen in unterschiedlichen Feldern, sich affirmativ auf diese zu beziehen. Strategien und Visionen sind zudem mehr als nur diskursive Massnahmen: Sie werden von materiellen und weiteren Investitionen der politischen Institutionen, Unternehmen oder Hochschulen begleitet, die somit nicht-diskursive Effekte auf die Emergenz von Wissensfeldern in Zwischenräumen entfalten.

Schliesslich haben Universitäten und Hochschulen rasch auf solche Ansprüche und Forderungen umliegender Felder, vor allem aus der Ökonomie sowie der Hochschul- und Forschungspolitik, reagiert: So sind im akademischen Feld innerhalb weniger Jahre Dutzende neue Studiengänge, interdisziplinäre Zentren und andere organisationale Einheiten etabliert worden. Darüber hinaus beschäftigen sich eine kaum mehr überschaubare Zahl an Forschungsgruppen und Wissenschaftler*innen mit den entsprechenden Fragen nach der »digitalen Transformation« ihrer jeweiligen Arbeitsbereiche. Die empirisch zu beobachtende beschleunigte Bearbeitung innerhalb von Forschung und Hochschulbildung verdeutlicht somit am Beispiel der Datenwissenschaften auch die enge Verzahnung von wissenschaftlicher Grundlagenforschung, angewandter Technologieentwicklung, ökonomischer Verwertung und politischer Steuerung.

11.5 Die Durchsetzung neuer Denk- und Wahrnehmungsweisen fundiert die Mächtigkeit datenwissenschaftlicher Expertise

Wie aber war es möglich, dass sich die Datenwissenschaften als Wissensgebiet in derart beschleunigter Art und Weise von einer hybriden wissenschaftlich-industriellen Schnittstelle aus etablieren konnten? Die Analyse hat hier neben den Investitionen, Interessen und Strategien der Akteur*innen insbesondere auf spezifische Denk- und Wahrnehmungsweisen aufmerksam gemacht: Die Mächtigkeit von »Data Science« als Expertise liegt weniger in der Ubiquität von datenwissenschaftlichen Anwendungen und Methoden, sondern vielmehr in der Etablierung und Durchsetzung eines »Data Mindsets« (Mayer-Schoenberger & Cukier 2013), einer »Datalogie« (Thornham & Gómez Cruz 2016) oder eines »Daten-Blicks« (Beer 2019). Der gemeinsame springende Punkt dieser unterschiedlichen Begriffskonzeptionen liegt darin, dass sie davon ausgehen, dass die soziale Welt und deren Komplexität in erster Linie durch und vermittelt über Daten wahrgenommen wird. So verwenden Akteur*innen aus multiplen Feldern, Disziplinen oder Professionen ähnliche oder gar dieselben Daten, Werkzeuge und Methoden und beginnen, das Soziale durch die Linse von Daten zu betrachten.

Die empirische Analyse macht die Durchsetzung datenwissenschaftlicher Expertise insbesondere im akademischen Feld deutlich: So zeigen sich auf der organisationalen Ebene, anders als bei den intensiven Auseinandersetzungen um die epistemische Deutungshoheit über die Datenwissenschaften, kaum Konflikte über deren disziplinäre Verortung. Die Ergebnisse deuten eher darauf hin, dass diese strategisch latent gehalten werden, um die Kohäsion des Netzwerks unterschiedlicher Expertisen nicht zu destabilisieren. Trotz bestehender Konkurrenzverhältnisse und beobachtbarer Praktiken der Abgrenzung werden auch vertiefte Formen der Zusammenarbeit

mit anderen Disziplinen, organisationalen Einheiten oder anderen Hochschulen in diesem Rahmen möglich. In dieser Offenheit manifestiert sich die Mächtigkeit (Rose 1992) der Datenwissenschaften als Expertise: Zu starke Konflikte um die disziplinäre Vorherrschaft über das neue Wissensgebiet hätten die beschleunigte Expansion und Diffusion im Schweizer Hochschulfeld verzögert bzw. erschwert. Somit blieb eine rasche Implementation innerhalb nur einer Disziplin (z. B. in den Computerwissenschaften) nur in kleineren Universitäten bzw. als Vertiefungen eine Option.

Auf der Ebene von Studiengängen und Curricula hat die Analyse ferner gezeigt, dass die Datenwissenschaften als Konvolut unterschiedlicher Expertisen in spezifische Ausbildungs- und Qualifikationsprofile übersetzt werden müssen, die wiederum in individuellen Zuschreibungen als »Datenkompetenzen« münden. Trotz der intensiven Bemühungen und Suchprozesse hat sich bis dato keine allgemein akzeptierte Version solcher »Datenkompetenzen« etabliert. Vielmehr funktionieren die oft arbiträr verwendeten Konzepte als doppelte Zuschreibungen: Einerseits sollen sie Wissen darüber vermitteln, wie datengestützte Prozesse strukturiert sind und funktionieren, wie Daten erhoben, analysiert oder visualisiert werden können. Andererseits enthalten sie stets auch normative Zuschreibungen, dass solche datenfokussierten Wahrnehmungsweisen des Sozialen etabliert werden *müssen*. Die Bezeichnung »Data Mindset« verdeutlicht somit den normativen Anspruch, ökonomische, politische und wissenschaftliche Fragestellungen nicht nur datengestützt bearbeiten zu *können*, sondern auch zu *wollen*.

Der feldübergreifende Konsens über die Notwendigkeit einer datengestützten Sichtweise auf die soziale Welt, der die breite Diffusion von Datenkompetenzen in unterschiedliche Disziplinen und Felder fundiert, trägt schliesslich auch zur Normalisierung von Datafizierung, d. h. zur Problemlösung mit und durch Daten, als gesellschaftlichem Phänomen bei. Darin manifestiert sich letztlich eine Ko-Produktion (Jasanoff 2004) von Datenwissenschaften und Gesellschaft: Während sich die Gesellschaft als »Datengesellschaft« (Houben & Prietl 2018) zunehmend über und durch Daten – und damit auch Dispositive und Expertisen zu deren Analyse – selbst beschreibt und reproduziert, sind umgekehrt epistemologische, disziplinäre, ökonomische und politische Kategorisierungen und Differenzlinien sowohl in die datenverarbeitenden Systeme als auch in die Konstruktions- und Organisationsprinzipien der Datenwissenschaften eingeschrieben. Die Datenwissenschaften sind demnach – trotz ihrer einschlägigen Bezeichnung – gerade kein ausschliesslich wissenschaftliches Phänomen, sondern umfassen ebenso ökonomische, politische und andere Deutungsmuster und stützen sich auf ein breites Spektrum unterschiedlicher Wissensinhalte und Expertisen, die diversen Feldern entstammen. Es gehört allerdings auch zu ihren zeitgenössischen Ausprägungen, die Implikationen einer an Transparenz und Objektivität (Porter 1995) orientierten datenbasierten Vorgehensweise – mit wenigen Ausnahmen – kaum zu thematisieren, geschweige denn in die Konstruktionsprinzipien und Steuerungsmechanismen der entstehenden organisationalen Gefässe und Strukturen einzubinden.

11.6 Zur Generalisierbarkeit des analytischen Modells

Zunächst lässt sich festhalten, dass sich die Kombination des analytischen Modells mit dem empirischen Vorgehen einer verteilten Analyse von feldspezifischen Konfigurationen als tauglich erwiesen hat, die Emergenz der Datenwissenschaften als transversalen Raum zwischen Feldern zu untersuchen. Die analytische Perspektive legte den Blick weniger auf die Analyse einzelner Bezugsfelder, sondern fokussierte die feldspezifischen Konfigurationen der Datenwissenschaften. Das empirische Vorgehen ermöglichte es, die unterschiedlichen Bedeutungs- und Sinnkonstruktionen der Datenwissenschaften im Arbeitsmarkt, in der Hochschul- und Forschungspolitik sowie im akademischen Feld zu identifizieren und einer inhaltlichen Analyse zugänglich zu machen. Dadurch war das Modell in der Lage, ein entstehendes, unterbestimmtes Wissensgebiet theoretisch-analytisch zu fassen, das (noch) keine etablierten Grenzen, keine klaren Spielregeln für die Teilnahme oder hierarchische Strukturierungsprinzipien aufweist, wie dies die soziologische Feldtheorie als zentrale Charakteristika annimmt. Insofern eignet sich das analytische Modell spezifisch für die Analyse interdisziplinärer Wissensfelder oder soziotechnischer Systeme, die multiple Bedeutungen aufweisen und Akteur*innen verschiedener sozialer Felder involvieren.

Indem die Arbeit ein empirisches Vorgehen für die verteilte Analyse von Zwischenräumen konzipiert und am Beispiel der Datenwissenschaften methodisch umgesetzt und geprüft hat, eröffnet dies die Möglichkeit zur Übertragbarkeit und Plausibilisierung an anderen empirischen Gegenständen. Als naheliegend und besonders vielversprechend können soziotechnische Phänomene wie künstliche Intelligenz oder Digitalisierung gelten: Wie die Analyse des politischen Diskurses gezeigt hat, wird Digitalisierung mit verschiedenen Bedeutungen versehen, die von einem Technologiefeld, dem das Potenzial zugeschrieben wird, andere etablierte Wissenspraktiken zu transformieren, bis hin zu einem Synonym für gesellschaftliche Transformation an sich reichen. Ähnlich gelagert ist dies im Falle von künstlicher Intelligenz, die im engeren Sinne ein breites Kompendium unterschiedlicher Methoden und Technologien umfasst, die parallel dazu aber auch als Chiffre für einen umfassenden Wandel von Gesellschaft verwendet wird.

Solche soziotechnischen Phänomene sind nicht ausschliesslich als technologisch oder wissenschaftlich zu konzipieren. Sie sind nie bloss Resultat ›fortscheidender technologischer Entwicklung‹, wie eine rein technikdeterministische Sichtweise glauben macht. Vielmehr unterliegen sie auch ökonomischen und politischen Konjunkturen und sind durch vielfältige Kollaborations- und Kooperationsverhältnisse sowie durch die daraus resultierenden Spannungen und Widersprüche gekennzeichnet. Die empirisch beobachtbaren und soziologisch rekonstruierbaren Sinnkonstruktionen solcher komplexer Phänomene ergeben sich somit erst aus der Vielfalt divergierender Perspektiven, was eindeutig zuordenbare Handlungs- und Betrachtungslogiken unzureichend macht. Sie verweisen zudem darauf, dass die gesellschaftliche Bedeutung solcher Wissensfelder nie selbsterklärend ist, sondern nur durch ein komplexes Zusammenspiel von Interessen, Strategien und Investitionen ökonomischer, politischer, wissenschaftlicher und weiterer Akteur*innen zu fassen ist.

Die zugrunde gelegte theoretisch-analytische Perspektive bietet sich darüber hinaus allerdings auch für andere soziale Phänomene an, da sie eine empirische Herangehensweise ermöglicht, einen Gegenstand in seinen unterschiedlichen feldspezifischen

Konfigurationen zu untersuchen. Dies zielt insbesondere darauf ab, das interessierende soziale Phänomen nicht eindimensional, d. h. aus der handlungspraktischen Perspektive eines bestimmten Feldes, sondern als multiperspektivisch und durch teilweise konkurrierende Bedeutungen konstruiert zu verstehen. Beispielsweise können Krankheitsbilder auf ihre medizinischen, psychologischen, sozialarbeiterischen, ökonomischen oder rechtlichen Dimensionen (Eyal et al. 2010; Martin 1993) hin untersucht werden, die eine eindimensionale Perspektive nicht adäquat erschliessen kann. Dadurch wird es auch möglich, die mannigfaltigen, mitunter widersprüchlichen konflikthafteren und kooperativen Relationen zwischen den involvierten Akteur*innen bzw. sozialen Feldern zu fassen und der Analyse zugänglich zu machen.

Sozialtheoretisch leistet die Arbeit dank dieser perspektivischen Verschiebung einen Beitrag zum Verständnis entstehender transversaler Wissensgebiete als Räumen zwischen Feldern. Dadurch schafft das analytische Modell einen Platz für (emergierende) soziale Phänomene innerhalb der soziologischen Feldtheorie jenseits der dichotomen Konzeption von Autonomisierung und Etablierung einerseits sowie Heteronomisierung und Verfall andererseits. Die Arbeit bettet sich wiederum ein in vielfältige Bemühungen in verschiedenen sozialtheoretischen Ansätzen, solche hybriden räumlichen Formationen sozialer Differenzierung adäquater zu berücksichtigen und in existierende Theorien des Sozialen zu integrieren. Indem die Arbeit eine stärker akteur- und praxisorientierte Perspektive einnimmt, als dies existierende Konzeptionen zwischenräumlicher Formationen (Eyal 2013b; Furnari 2014) bis dato getan haben, richtet sich der Fokus auf die kollektiven Praktiken der Bedeutungskonstruktion sowie von Grenz- und Begriffsarbeit in den verschiedenen Feldern.

Trotz der geographischen Beschränkung der empirischen Analyse auf die spezifischen Konfigurationen der Herausbildung der Datenwissenschaften im schweizerischen Arbeitsmarkt, in der schweizerischen Hochschul- und Forschungspolitik sowie an Schweizer Universitäten und Hochschulen hat die Arbeit transnationale Relationen, Beobachtungs- und Abhängigkeitsverhältnisse identifiziert, deren Effekte sich auf einen und in einem Raum zwischen Feldern zeigen, der nationalstaatliche Grenzen transzendiert. Der hohe Anteil englischsprachiger Stelleninserate und die Dominanz international tätiger Unternehmen im Sample des Arbeitsmarktes verweisen darauf, dass nicht nur Organisationen und Felder, sondern auch Staaten und Sprachregionen sehr unterschiedlich datafiziert, d. h. durch Datenpraktiken strukturiert und geprägt sind. Analog weist die Inhaltsanalyse des hochschul- und forschungspolitischen Diskurses auf eine zunehmende, enge Orientierung an Aktivitäten und Förderinitiativen »vergleichbarer« Bildungs- und Forschungssysteme hin, die zur Koordination und Stabilisierung entstehender Wissensgebiete über nationalstaatliche Grenzen hinweg beigetragen hat. Schliesslich sind technische Hochschulen und Universitäten durch verschiedene Mechanismen in ein globales akademisches Feld eingebunden, das sich nicht zuletzt auf die rasche Diffusion der Datenwissenschaften abstützt. Die beschleunigte Herausbildung der Datenwissenschaften im nationalstaatlichen Kontext der Schweiz ist deshalb vor dem Hintergrund solcher transnationaler Austauschbeziehungen und Dynamiken zu verstehen. Daran anknüpfend könnte eine weitergehende Forschung die ökonomischen und politischen Bedingungen, die die Diffusion datenwissenschaftlicher Expertise über nationalstaatliche Grenzen hinweg ermöglicht hat, vertieft untersuchen. Dies würde auch dazu beitragen, das analytische Modell in das zunehmend transnational verfasste Forschungsprogramm der aktuellen sozio-

logischen Feldtheorie (Buchholz 2016; Krause 2018; Schmidt-Wellenburg & Bernhard 2020a) zu integrieren.

Eine weitere Stärke des analytischen Modells liegt darin, dass es methodisch flexibel und adaptiv an den empirischen Gegenstand erweiterbar ist. Die Analyse operierte zunächst auf einer makroskopischen Beobachtungsebene und untersuchte das Teilsegment der Datenwissenschaften des Arbeitsmarktes auf strukturelle Verteilungen und diskursive Bedeutungskonstruktionen. Anschliessend wechselte ich auf die Ebene sozialer Felder, indem ich kollektive Zukunftsvisionen in Diskursen der Hochschul- und Forschungspolitik einerseits und akteuriale Strategien und Handlungen im akademischen Feld andererseits analysierte. Dieses Vorgehen ermöglichte es, die feldübergreifenden Praktiken und Konstruktionsleistungen zu identifizieren und deren Beiträge zur Herausbildung der Datenwissenschaften als transversalen Raum zwischen Feldern sichtbar zu machen. Zudem ist das Modell offen für den Einbezug weiterer sozialer Felder, wie beispielsweise des Feldes der Medien (vgl. Abbildung 11), dessen Bedeutung in der vorliegenden Arbeit nicht adressiert werden konnte.

Daran anknüpfend stellt die Integration einer mikrosoziologisch fundierten Analyse der Praktiken individueller oder organisationaler Akteur*innen (Brandt 2016; Nax & Draheim 2019; Unternährer 2021), medialer Repräsentationen (Dahl 2020) oder professioneller Identitäten (Avnoon 2021; Dorschel & Brandt 2021) eine aussichtsreiche Erweiterung des analytischen Modells dar. Damit können zusätzlich Bedeutungs- und Sinnkonstruktionen auf der Praxisebene einzelner Organisationen, insbesondere im ökonomischen Feld (Mützel et al. 2018), oder feldspezifischer Settings in die Konzeption sozialer Phänomene als Zwischenräume eingebunden werden. Komplementär dazu lassen sich methodische Erhebungs- und Analyseverfahren der qualitativen Sozialforschung wie ethnographische Verfahren oder biographische Interviews in den empirisch-analytischen Rahmen integrieren. Zudem stellen auch weitere computergestützte text- und netzwerkanalytische Methoden ein vielversprechendes Vorhaben dar, um Bedeutungs- und Sinnkonstruktionen der Datenwissenschaften in unterschiedlichen Publikationsformen (wie Handbücher, Online-Foren, Beiträge in sozialen Medien etc.) zu erschliessen.

Die Arbeit hat gezeigt, dass die Datenwissenschaften nicht nur die methodisch-technische Expertise, sondern auch das kognitiv-kulturelle Dispositiv einer sich etablierenden Datengesellschaft bilden. Sie stellen den zentralen Baustein dar, damit sich sowohl Individuen als auch Kollektivformationen durch die ›Brille‹ von Daten wahrnehmen, beobachten und kommunizieren können. Dabei verdeckt der Umstand, dass diese Form von Expertise in Methoden und Technologien programmiert und somit zwar dauerhaft und übertragbar wird, aber gleichzeitig für die meisten Menschen abstrakt und unnahbar bleibt, die weitgehende Durchdringung der sozialen Welt. Um die Datenwissenschaften für eine gesellschaftliche (Selbst-)Reflexion zu öffnen, fehlt es nicht nur an notwendigen und breit vermittelten Kenntnissen und Fähigkeiten (die vielbeschworenen digitalen Skills), sondern auch an der Thematisierung der sozialen, politischen oder ökonomischen Implikationen datengestützter Entscheidungsprozesse und Verfahren innerhalb der Datenwissenschaften (im Sinne der sehr viel seltener thematisierten *critical data studies*). Um allzu einseitigen Orientierungen an technologischer Machbarkeit, politischen Strategiezielen und vor allem ökonomischer Valorisation vorzubeugen, müssen die Datenwissenschaften als ein Netzwerk unterschiedlicher Expertisen auf jene Perspektiven erweitert werden, die sich den histori-

schen Entstehungsbedingungen, gesellschaftlichen Spannungsverhältnissen und Implikationen solcher Orientierungen widmen. »Data Science« ist mit anderen Worten nicht nur als Expertise, als ein Tool für Gesellschaft zu betrachten und zu erforschen, sondern als Gesellschaft selbst.

Die vorliegende Arbeit hat die grundlegenden Kontexte der Genese und Diffusion, die Ausprägungen und Funktionsweisen der Datenwissenschaften untersucht und somit einer gesellschaftlichen Auseinandersetzung zugänglich gemacht. Es bleibt nun den Leser*innen überlassen, wie sie davon für eine Debatte über das Verhältnis von Datenwissenschaften und Gesellschaft Gebrauch machen.

Bibliografie

Alle Internetverweise wurden vor der Drucklegung geprüft.

- Aasheim, Cheryl, Susan Williams, Paige Rutner und Adrian Gardiner. 2014. »Big Data Analytics and Data Science Undergraduate Degree Programs.« S. 1–22 in *Annual Meeting of the Decision Sciences Institute Proceedings (DSI 2014)*. Tampa.
- Aasheim, Cheryl, Susan Williams, Paige Rutner und Adrian Gardiner. 2015. »Data Analytics vs. Data Science: A Study of Similarities and Differences in Undergraduate Programs Based on Course Descriptions.« *Journal of Information Systems Education* 26(2): 103–15.
- Abbott, Andrew. 1988. *The System of Professions. An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.
- Abbott, Andrew. 1995. »Things of Boundaries.« *Social Research* 62(4): 857–82.
- Abbott, Andrew. 2001. *Chaos of Disciplines*. Chicago: University of Chicago Press.
- Abbott, Andrew. 2005. »Linked Ecologies: States and Universities as Environments for Professions.« *Sociological Theory* 23(3): 245–74.
- Abraham, Martin und Thomas Hinz (Hg.). 2018. *Arbeitsmarktsoziologie* (3. Auflage). Wiesbaden: Springer.
- Abun-Nasr, Sonia. 2009. *Strategie für eine Informationsgesellschaft in der Schweiz. Eine zehnjährige Politik im Lichte empirischer Analyse*. Lausanne: IDHEAP.
- Acker, Amelia und Joan Donovan. 2019. »Data Craft: A Theory/Methods Package for Critical Internet Studies.« *Information, Communication & Society* 22(11): 1590–1609.
- ACM und IEEE Computer Society. 2013. *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*, hg. von ACM Computing Curricula Task Force. New York: ACM.
- Adams, Vincanne, Michelle Murphy und Adele E. Clarke. 2009. »Anticipation: Technology, Life, Affect, Temporality.« *Subjectivity* 28(1): 246–65.
- Adhikari, Ani und John DeNero. 2016. *Computational and Inferential Thinking. The Foundations of Data Science*. Online: <https://www.inferentialthinking.com/>
- Akademien der Wissenschaften Schweiz. 2017. »Ausserschulische MINT-Angebote in der Schweiz, Übersicht und Analyse 2016.« *Swiss Academies Communications* 12(6).
- Akkerman, Sanne F. und Arthur Bakker. 2011. »Boundary Crossing and Boundary Objects.« *Review of Educational Research* 81(2): 132–69.
- Allan, Bentley B. 2017. »Producing the Climate: States, Scientists, and the Constitution of Global Governance Objects.« *International Organization* 71(1): 131–62.

- Äm, Heidrun. 2013. »Don't Make Nanotechnology Sexy, Ensure Its Benefits, and Be Neutral: Studying the Logics of New Intermediary Institutions in Ambiguous Governance Contexts.« *Science and Public Policy* 40(4): 466–78.
- Anderson, Chris. 2008. »The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.« In *Wired*, 23.06.2008. Online: <https://www.wired.com/2008/06/pb-theory/>
- Anderson, Paul, James Bowring, Renée McCauley, George Pothering und Christopher Starr. 2014. »An Undergraduate Degree in Data Science: Curriculum and a Decade of Implementation Experience.« S. 145–50 in *Proceedings of the 45th ACM technical symposium on Computer science education – SIGCSE '14*. New York: ACM Press.
- Andreoli, Laura, Bernadette Baumberger und Anja Umbach-Daniel. 2017. *Ingenieur-Nachwuchs Schweiz 2017. Entwicklung des Ingenieur-Nachwuchses an universitären Hochschulen und Fachhochschulen*. Rüşchlikon: Rütter Soceco AG.
- Aratnam, Ganga Jey. 2012. *Hochqualifizierte mit Migrationshintergrund. Studie zu möglichen Diskriminierungen auf dem Schweizer Arbeitsmarkt*. Basel: edition gesowip.
- Aratnam, Ganga Jey. 2016. »Der Talent-Turn und seine staatliche Begleitmusik. Migration und die Politik der Hochqualifizierung.« S. 341–366 in *Staatlichkeit in der Schweiz. Regieren und Verwalten vor der neoliberalen Wende*, hg. von L. Criblez, C. Rothen und T. Ruoss. Zürich: Chronos.
- Aronova, Elena, Christine von Oertzen und David Sepkoski. 2017. »Introduction: Historicizing Big Data.« *Osiris* 32(1): 1–17.
- Asamoah, Daniel, Derek Doran und Shu Schiller. 2015. »Teaching the Foundations of Data Science: An Interdisciplinary Approach.« *Pre-ICIS SIGDSA Workshop*, Fort Worth, Texas. arXiv: 1512.04456.
- Augsburg, Tanya und Stuart Henry. 2009. *The Politics of Interdisciplinary Studies : Essays on Transformations in American Undergraduate Programs*. Jefferson, NC: McFarland & Co.
- Avnoon, Netta. 2021. »Data Scientists' Identity Work: Omnivorous Symbolic Boundaries in Skills Acquisition.« *Work, Employment and Society* 35(2): 332–49.
- Baier, Christian und Andreas Schmitz. 2012. »Organisationen als Akteure in sozialen Feldern – Eine Modellierungsstrategie am Beispiel deutscher Hochschulen.« S. 191–220 in *Feldanalyse als Forschungsprogramm 1*, hg. von S. Bernhard und C. Schmidt-Wellenburg. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bail, Christopher A. 2014. »The Cultural Environment: Measuring Culture with Big Data.« *Theory and Society* 43(3): 465–82.
- Barlow, Mike. 2013. *The Culture of Big Data*. Sebastopol, CA: O'Reilly Media.
- Barocas, Solon und danah boyd. 2017. »Engaging the Ethics of Data Science in Practice.« *Communications of the ACM* 60(11): 23–25.
- Barocas, Solon, Elizabeth Bradley, Vasant Honavar und Foster Provost. 2017. »Big Data, Data Science, and Civil Rights.« *arXiv*: 1706.03102.
- Bates, Jo, David Cameron, Alessandro Checco, Paul Clough, Frank Hopfgartner, Suvodeep Mazumdar, Laura Sbaffi, Peter Stordy und Antonio de la Vega de León. 2020. »Integrating FATE/Critical Data Studies into Data Science Curricula: Where Are We Going and How Do We Get There?« S. 425–435 in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT* '20*. New York: ACM Press.
- Baumer, Eric P. S., David Mimno, Shion Guha, Emily Quan und Geri K. Gay. 2017. »Comparing Grounded Theory and Topic Modeling: Extreme Divergence or Un-

- likely Convergence?» *Journal of the Association for Information Science and Technology* 68(6): 1397–1410.
- Baur, Nina, Cristina Besio und Maria Norkus. 2016. »Organisationale Innovation am Beispiel der Projektifizierung der Wissenschaft. Eine figurationssoziologische Perspektive auf Entstehung, Verbreitung und Wirkungen.« S. 373–402 in *Innovationsgesellschaft heute*, hg. von W. Rammert, A. Windeler, H. Knoblauch und M. Hutter. Wiesbaden: Springer VS.
- Beauchamp, Catherine und Lynn Thomas. 2011. »New Teachers' Identity Shifts at the Boundary of Teacher Education and Initial Practice.« *International Journal of Educational Research* 50(1): 6–13.
- Beckert, Bernd, Susanne Bührer und Ralf Lindner. 2008. »Verläufe und Motive von ›Seitenwechseln‹: Intersektorale Mobilität als Form des Wissenstransfers zwischen Forschung und Anwendung.« S. 313–340 in *Wissensproduktion und Wissenstransfer*, hg. von R. Mayntz, F. Neidhardt, P. Weingart und U. Wengenroth. Bielefeld: transcript.
- Beckert, Jens. 2013. »Imagined Futures: Fictional Expectations in the Economy.« *Theory and Society* 42(3): 219–40.
- Beckert, Jens. 2016. *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Cambridge, MA: Harvard University Press.
- Beckert, Jens und Richard Bronk. 2018. »An Introduction to Uncertain Futures.« S. 1–36 in *Uncertain Futures: Imaginaries, Narratives, and Calculation in the Economy*, hg. von J. Beckert und R. Bronk. Oxford: Oxford University Press.
- Bedenlier, Svenja und Markus Deimann. 2020. »›Bildung‹ und ›Digitalisierung‹ im Spiegel von Digitalisierungsstrategien.« *Zeitschrift für Hochschulentwicklung* 15(1): 41–59.
- Beer, David. 2017. »The Data Analytics Industry and the Promises of Real-Time Knowing: Perpetuating and Deploying a Rationality of Speed.« *Journal of Cultural Economy* 10(1): 21–33.
- Beer, David. 2019. *The Data Gaze. Capitalism, Power and Perception*. London/Thousand Oaks/New Delhi: Sage.
- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller und Akitaka Matsuo. 2018. »Quanteda: An R Package for the Quantitative Analysis of Textual Data.« *Journal of Open Source Software* 3(30): 774.
- Bensaude-Vincent, Bernadette. 2001. »The Construction of a Discipline: Materials Science in the United States.« *Historical Studies in the Physical and Biological Sciences* 31(2): 223–48.
- Bensaude-Vincent, Bernadette. 2014. »The Politics of Buzzwords at the Interface of Technoscience, Market and Society: The Case of ›Public Engagement in Science.« *Public Understanding of Science* 23(3): 238–53.
- Bensaude-Vincent, Bernadette und Arne Hessenbruch. 2004. »Materials Science: A Field about to Explode?« *Nature Materials* 3(6): 345–47.
- Berman, Elizabeth Popp. 2012. *Creating the Market University. How Academic Science Became an Economic Engine*. Princeton: Princeton University Press.
- Berman, Francine D. und Philip E. Bourne. 2015. »Let's Make Gender Diversity in Data Science a Priority Right from the Start.« *PLoS Biology* 13(7): 1–5.
- Berman, Francine, Victoria Stodden, Alexander S. Szalay, Rob Rutenbar, Brent Hailpern, Henrik Christensen, Susan Davidson, Deborah Estrin, Michael Franklin,

- Margaret Martonosi und Padma Raghavan. 2016. *Realizing the Potential of Data Science. Final Report from the National Science Foundation Computer and Information Science and Engineering Advisory Committee Data Science Working Group*. Washington, D.C.: National Science Foundation.
- Berman, Francine, Victoria Stodden, Alexander S. Szalay, Rob Rutenbar, Brent Hailpern, Henrik Christensen, Susan Davidson, Deborah Estrin, Michael Franklin, Margaret Martonosi und Padma Raghavan. 2018. »Realizing the Potential of Data Science.« *Communications of the ACM* 61(4): 67–72.
- Bernhard, Stefan und Christian Schmidt-Wellenburg. 2012. »Feldanalyse als Forschungsprogramm.« S. 27–56 in *Feldanalyse als Forschungsprogramm 1*, hg. von S. Bernhard und C. Schmidt-Wellenburg. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bernstein, Basil. 1975. *Class, Codes and Control. Volume III: Towards a Theory of Educational Transmissions*. London: Routledge.
- Berry, David M. 2011. »The Computational Turn: Thinking about the Digital Humanities.« *Culture Machine* 12: 1–22.
- Bessy, Christian und Pierre-Marie Chauvin. 2013. »The Power of Market Intermediaries: From Information to Valuation Processes.« *Valuation Studies* 1(1): 83–117.
- Betschon, Stefan. 2019. »Der Aufstand der Elektrosensiblen.« *Neue Zürcher Zeitung*, 15.07.2019: 10.
- Beunza, Daniel und David Stark. 2004. »Tools of the Trade: The Socio-Technology of Arbitrage in a Wall Street Trading Room.« *Industrial and Corporate Change* 13(2): 369–400.
- BHEF. 2016. *Competency Map for the Data Science and Analytics-Enabled Graduate*. Washington, D.C.: Business and Higher Education Foundation.
- Bieber, Christoph. 2015. »Daten vs. Netz. Bausteine und Perspektiven eines unfertigen Politikfeldes.« S. 173–98 in *Die Gesellschaft der Daten. Über die digitale Transformation der sozialen Ordnung*, hg. von F. Süssenguth. Bielefeld: transcript.
- Binder, Amy J. 2002. *Contentious Curricula: Afrocentrism and Creationism in American Public Schools*. Princeton: Princeton University Press.
- Biniok, Peter. 2013. *Wissenschaft als Bricolage. Die soziale Konstruktion der Schweizer Nowissenschaften*. Bielefeld: transcript.
- Bird, Sarah, Solon Barocas, Fernando Diaz, Kate Crawford und Hanna Wallach. 2016. »Exploring or Exploiting? Social and Ethical Implications of Autonomous Experimentation in AI.« S. 1–4 in *Workshop on Fairness, Accountability, and Transparency in Machine Learning*, 2016. New York.
- Blasius, Jörg, Frédéric Lebaron, Brigitte Le Roux und Andreas Schmitz (Hg.). 2019. *Empirical Investigations of Social Space*. Cham: Springer International Publishing.
- Blei, David M. 2012. »Probabilistic Topic Models.« *Communications of the ACM* 55(4): 77–84.
- Blei, David M., Andrew Y. Ng und Michael I. Jordan. 2003. »Latent Dirichlet Allocation.« *Journal of Machine Learning Research* (3): 993–1022.
- Bloch, Roland, Alexander Mitterle, Catherine Paradeise und Tobias Peter. 2018. *Universities and the Production of Elites. Discourses, Policies, and Strategies of Excellence and Stratification in Higher Education*. Basingstoke: Palgrave Macmillan.

- Blümel, Clemens. 2016. »Der Beitrag der Innovationsforschung zur Wissenschaftspolitik.« S. 1–16 in *Handbuch Wissenschaftspolitik*, hg. von D. Simon, A. Knie, S. Hornbostel und K. Zimmermann. Wiesbaden: Springer.
- Blümel, Clemens. 2018. »Legitimes Sprechen über Innovation? Die Nutzung von Innovationsverständnissen im Wissenschaftspolitischen Feld.« *Zeitschrift für Diskursforschung* (1): 71–102.
- Boes, Andreas, Tobias Kämpf, Barbara Langes und Thomas Lühr. 2018. »Lean« und »agil im Büro. Neue Organisationskonzepte in der digitalen Transformation und ihre Folgen für die Angestellten. Bielefeld: transcript.
- Bongaerts, Gregor. 2011. »Grenzsicherung in sozialen Feldern – Ein Beitrag zu Bourdieus Theorie gesellschaftlicher Differenzierung.« S. 113–33 in *Soziale Differenzierung*, hg. von T. Schwinn. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Borne, Kirk, John Wallin und Robert Weigel. 2009. »The New Computational and Data Sciences Undergraduate Program at George Mason University.« S. 74–83 in *Computational Science – ICCS 2009, Lecture Notes in Computer Science*, hg. von G. Allen, J. Nabrzyski, E. Seidel, G. D. van Albada, J. Dongarra und P. M. A. Sloot. Berlin/Heidelberg: Springer.
- Böschen, Stefan. 2016. *Hybride Wissensregime. Skizze einer soziologischen Feldtheorie*. Baden-Baden: Nomos.
- Boselli, Roberto, Mirko Cesarini, Fabio Mercorio und Mario Mezzanzanica. 2018. »Classifying Online Job Advertisements through Machine Learning.« *Future Generation Computer Systems* 86: 319–28.
- Bott, Peter, Anja Hall und Hans-Joachim Schade. 2000. *Qualifikationsanforderungen im IT-Bereich: Wunsch und Wirklichkeit*. Ergebnisse einer Inserentennachbefragung im Rahmen des Früherkennungssystems Qualifikationsentwicklung. Bonn.
- Bourdieu, Pierre. 1976. *Entwurf einer Theorie der Praxis*. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre. 1983. »Ökonomisches Kapital, kulturelles Kapital, soziales Kapital.« S. 183–98 in *Soziale Ungleichheiten*, hg. von R. Kreckel. Göttingen: Verlag Otto Schwarz & Co.
- Bourdieu, Pierre. 1987. *Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft*. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre. 1988. *Homo Academicus*. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre. 1993. »Über einige Eigenschaften von Feldern.« S. 107–14 in *Soziologische Fragen*. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre. 1997a. »Die männliche Herrschaft.« S. 153–217 in *Ein alltägliches Spiel. Geschlechterkonstruktion in der sozialen Praxis*, hg. von I. Dölling und B. Kraus. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre. 1997b. »Zur Genese der Begriffe Habitus und Feld.« S. 59–78 in *Der Tote packt den Lebenden. Schriften zu Kultur und Politik 2*, hg. von M. Steinrück. Hamburg: VSA-Verlag.
- Bourdieu, Pierre. 1998. »Das ökonomische Feld.« S. 162–204 in *Der Einzige und sein Eigenheim. Schriften zu Politik und Kultur 3*, hg. von M. Steinrück. Hamburg: VSA-Verlag.
- Bourdieu, Pierre. 2000. *Das religiöse Feld. Texte zur Ökonomie des Heilsgeschehens*. Konstanz: UVK.
- Bourdieu, Pierre. 2001a. *Das politische Feld: Kritik der politischen Vernunft*. Edition discours. Konstanz: UVK.

- Bourdieu, Pierre. 2001b. *Die Regeln der Kunst. Genese und Struktur des literarischen Feldes*. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre. 2002. *Der Einzige und sein Eigenheim. Erweiterte Neuauflage der Schriften zu Politik & Kultur 3*. Hamburg: VSA.
- Bourdieu, Pierre. 2004. *Der Staatsadel*. Konstanz: UVK Verlag.
- Bourdieu, Pierre. 2014. *Über den Staat. Vorlesungen am Collège de France 1989–1992*. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre, Luc Boltanski, Monique de Saint Martin und Pascale Maldidier-Paragamin. 1981. *Titel und Stelle. Über die Reproduktion sozialer Macht*. Frankfurt a.M.: Europäische Verlagsanstalt.
- Bourdieu, Pierre und Jean-Claude Passeron. 1971. *Die Illusion der Chancengleichheit. Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Stuttgart: Ernst Klett Verlag.
- Bourdieu, Pierre und Jean-Claude Passeron. 1973. *Grundlagen einer Theorie der Symbolischen Gewalt*. Frankfurt a.M.: Suhrkamp.
- Bourdieu, Pierre und Jean-Claude Passeron. 2007. *Die Erben. Studenten, Bildung und Kultur*. Konstanz: UVK Verlag.
- Bourdieu, Pierre und Loïc Wacquant. 1996. *Reflexive Anthropologie*. Frankfurt a.M.: Suhrkamp.
- Bowker, Geoffrey C. und Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Bowne-Anderson, Hugo. 2018. »What Data Scientists Really Do, According to 35 Data Scientists.« *Harvard Business Review*, 15.08.2018. Online: <https://hbr.org/2018/08/what-data-scientists-really-do-according-to-35-data-scientists>
- boyd, danah und Kate Crawford. 2012. »Critical Questions for Big Data.« *Information, Communication & Society* 15(5): 662–79.
- Brandt, Philipp S. 2016. *The Emergence of the Data Science Profession*. Doctoral Dissertation. New York: Columbia University.
- Breiman, Leo. 2001. »Statistical Modeling: The Two Cultures.« *Statistical Science* 16(3): 199–215.
- Brint, Steven G., Lori Turk-Bicakci, Kristopher Proctor und Scott Patrick Murphy. 2009. »Expanding the Social Frame of Knowledge: Interdisciplinary, Degree-Granting Fields in American Colleges and Universities, 1975–2000.« *The Review of Higher Education* 32(2): 155–83.
- Brown, Gavin. 2009. »Review of Education in Mathematics, Data Science and Quantitative Disciplines. Report to the Group of Eight Universities.« Canberra: The Group of Eight.
- Brown, Laurie und Ross J. Barnett. 2004. »Is the Corporate Transformation of Hospitals Creating a New Hybrid Health Care Space? A Case Study of the Impact of Co-Location of Public and Private Hospitals in Australia.« *Social Science & Medicine* 58(2): 427–44.
- Brown, Nik, Brian Rappert und Andrew Webster (Hg.). 2000. *Contested Futures. A Sociology of Prospective Techno-Science*. Aldershot: Ashgate.
- Brunsson, Nils. 2017. *Reform as Routine: Organizational Change and Stability in the Modern World*. Oxford: Oxford University Press.
- Brüsemeister, Thomas. 2008. *Bildungssoziologie. Einführung in Perspektiven und Probleme*, hg. von T. Kron. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Buchholz, Larissa. 2016. »What Is a Global Field? Theorizing Fields beyond the Nation-State.« *The Sociological Review* 64(2_suppl): 31–60.
- Buckingham Shum, Simon, Martin Hawksey, Ryan S.J.d. Baker, Ryan, Naomi Jeffery, John T. Behrens und Roy Pea. 2013. »Educational Data Scientists: A Scarce Breed.« S. 278–81 in *3rd International Conference on Learning Analytics and Knowledge, LAK 2013*. Leuven.
- Bukhari, Duaa. 2020. »Data Science Curriculum: Current Scenario.« *International Journal of Data Mining & Knowledge Management Process* 10(3): 1–13.
- Burren, Susanne. 2010. *Die Wissenskultur der Betriebswirtschaftslehre: Aufstieg und Dilemma einer hybriden Disziplin*. Bielefeld: transcript.
- Burtch, Linda. 2016. *The Burtch Works Study. Salaries of Data Scientists*. Evanston, IL: Burtch Works.
- Burtch, Linda. 2018. *The Burtch Works Study Salaries of Data Scientists*. Evanston, IL: Burtch Works.
- Callon, Michel und Vololona Rabeharisoa. 2003. »Research ›in the Wild‹ and the Shaping of New Social Identities.« *Technology in Society* 25(2): 193–204.
- Cao, Longbing. 2017. »Data Science : Challenges and Directions.« *Communications of the ACM* 60(8): 59–68.
- Carillo, Kevin Daniel André. 2016. »How to Engrain a Big Data Mindset into Our Managers' DNA. Insights from a Big Data Initiative in a French Business School.« S. 71–91 in *Managing Big Data Integration in the Public Sector*, hg. von A. Aggarwal. Hershey: IGI Global.
- Carter, Daniel und Dan Sholler. 2016. »Data Science on the Ground: Hype, Criticism, and Everyday Work.« *Journal of the Association for Information Science and Technology* 67(10): 2309–19.
- Cech, Erin A. 2014. »Culture of Disengagement in Engineering Education?« *Science, Technology & Human Values* 39(1): 42–72.
- Chang, Ray M., Robert J. Kauffman und Youngok Kwon. 2014. »Understanding the Paradigm Shift to Computational Social Science in the Presence of Big Data.« *Decision Support Systems* 63: 67–80.
- Charles, Maria und Sarah Thébaud (Hg.). 2018. *Gender and STEM: Understanding Segregation in Science, Technology, Engineering and Mathematics*. Basel: MDPI.
- Chatfield, Akemi Takeoka, Vivian Najem Shlemoon, Wilbur Redublado und Fairuz Rahman. 2014. »Data Scientists as Game Changers in Big Data Environments.« In *25th Australasian Conference on Information Systems*. Auckland: ACIS.
- Chow, Tiffany und Maria Charles. 2019. »An Inegalitarian Paradox: On the Uneven Gendering of Computing Work around the World.« S. 25–45 in *Cracking the Digital Ceiling: Women in Computing Around the World*, hg. von C. Frieze und J. L. Quesenberry. Cambridge: Cambridge University Press.
- Chrisman, Nicholas. 1999. »Trading Zones or Boundary Objects: Understanding Incomplete Translations of Technical Expertise.« S. 28–31 in *Society for Social Studies of Science Annual Meeting*.
- Ciarla, Giulia. 2018. *Von einer Informationsgesellschaft zu einer digitalen Schweiz: Eine Retrospektive der Zukunft*. Unveröffentlichte Masterarbeit. Luzern: Universität Luzern.
- Clark, Burton R. 1983. *The Higher Education System: Academic Organization in Cross-National Perspective*. Berkeley: University of California Press.

- Clarke, Adele. 1991. »Social Worlds/Arenas Theory as Organizational Theory.« S. 119–58 in *Social Organization and Social Process. Essays in Honor of Anselm Strauss*, hg. von D. R. Maines. New York: Aldine De Gruyter.
- Cleveland, William S. 2001. »Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics.« *International Statistical Review* 69(1): 21–26.
- Collins, Randall. 2019. *The Credential Society*. New York/Chichester: Columbia University Press.
- Conway, Drew. 2010. »The Data Science Venn Diagram.« Blog, 30.09.2010. Online: <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>
- Cornelissen, Jonathan. 2018. »The Democratization of Data Science.« *Harvard Business Review*, 27.07.2018. Online: <https://hbr.org/2018/07/the-democratization-of-data-science>
- Coté, Mark. 2014. »Data Motility: The Materiality of Big Social Data.« *Cultural Studies Review* 20(1): 121–49.
- CrowdFlower. 2015. *2015 Data Scientist Report*. Online: <https://visit.figure-eight.com/2015-data-scientist-report>
- Dadzie, Abu-Sah, Elisa M. Sibarani, Inna Novalija und Simon Scerri. 2018. »Structuring Visual Exploratory Analysis of Skill Demand.« *Journal of Web Semantics* 49: 51–70.
- Dagiral, Éric und Sylvain Parasie. 2017. »La « science des données « à la conquête des mondes sociaux : Ce que le « Big Data « doit aux épistémologies locales.« S. 85–104 in *Big data et traçabilité numérique*, hg. von P.-M. Menger und S. Paye. Paris: Collège de France.
- Dahl, Patrik H. 2020. *The Data Scientist Self. Occupational Identities in Data Science Podcasts*. Unpublished Master Thesis. Cardiff: Cardiff University.
- Dalton, Craig und Jim Thatcher. 2014. »What Does a Critical Data Studies Look like, and Why Do We Care?« *Society and Space*, 1–26.
- Dasu, Tamraparni und Theodore Johnson. 2003. *Exploratory Data Mining and Data Cleaning*. Hoboken: John Wiley & Sons.
- Davenport, Thomas H. 2020. »Beyond Unicorns: Educating, Classifying, and Certifying Business Data Scientists.« *Harvard Data Science Review* 2(2). <https://doi.org/10.1162/99608f92.55546b4a>.
- Davenport, Thomas H. und DJ Patil. 2012. »Data Scientist: The Sexiest Job of the 21st Century.« *Harvard Business Review* 90(10): 70–76.
- Debortoli, Stefan, Oliver Müller und Jan vom Brocke. 2014. »Comparing Business Intelligence and Big Data Skills.« *Business & Information Systems Engineering* 6(5): 289–300.
- Debortoli, Stefan, Oliver Müller und Jan vom Brocke. 2014. »Vergleich von Kompetenzanforderungen an Business-Intelligence- und Big-Data-Spezialisten.« *Wirtschaftsinformatik* 56(5): 315–28.
- Dederichs, Andrea Maria und Michael Florian. 2004. »Felder, Organisationen und Akteure – eine organisationssoziologische Skizze.« S. 69–96 in *Bourdieu's Theorie der Praxis*, hg. von J. Ebrecht und F. Hillebrandt. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Demchenko, Yuri, Adam Belloum, Wouter Los, Spiros Koulouzis und Cees De Laat. 2016. »EDISON Project: Building Data Science Profession for European Research

- and Industry.« S. 1–3 in *The Conference for ICT-Research in the Netherlands*. Amsterdam.
- Denny, Matthew J. und Arthur Spirling. 2018. »Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It.« *Political Analysis* 26(2): 168–89.
- Diaz-Bone, Rainer. 2010. *Kulturwelt, Diskurs und Lebensstil. Eine diskurstheoretische Erweiterung der Bourdieuschen Distinktionstheorie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Diaz-Bone, Rainer und Christoph Weischer (Hg.). 2015. *Methoden-Lexikon für die Sozialwissenschaften*. Wiesbaden: Springer VS.
- Dichev, Christo und Darina Dicheva. 2017. »Towards Data Science Literacy.« *Procedia Computer Science* 108(June): 2151–60.
- Dickel, Sascha und Jan-Felix Schrape. 2015. »Dezentralisierung, Demokratisierung, Emanzipation. Zur Architektur des digitalen Technikutopismus.« *Leviathan* 43(3): 442–63.
- van Dijck, José. 2014. »Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology.« *Surveillance and Society* 12(2): 197–208.
- DiMaggio, Paul. 2015. »Adapting Computational Text Analysis to Social Science (and Vice Versa).« *Big Data & Society* 2(2): 1–5.
- DiMaggio, Paul J. und Walter W. Powell. 1983. »The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields.« *American Sociological Review* 48(2): 147–60.
- DiMaggio, Paul, Manish Nag und David Blei. 2013. »Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding.« *Poetics* 41(6): 570–606.
- Djumaliev, Jyldyz, Antonio Lima und Cath Sleeman. 2018. *Classifying Occupations According to Their Skill Requirements in Job Advertisements*. ESCoE Discussion Paper 2018-04. London: National Institute of Economic and Social Research.
- Dobbin, Frank, Daniel Schrage und Alexandra Kalev. 2015. »Rage against the Iron Cage.« *American Sociological Review* 80(5): 1014–44.
- Donoho, David. 2017. »50 Years of Data Science.« *Journal of Computational and Graphical Statistics* 26(4): 745–66.
- Dorschel, Robert und Philipp Brandt. 2021. Professionalisierung mittels Ambiguität. Die diskursive Konstruktion von Data Scientists in Wirtschaft und Wissenschaft». *Zeitschrift für Soziologie* 50(3–4): 193–210.
- Dourish, Paul und Edgar Gómez Cruz. 2018. »Datafication and Data Fiction: Narrating Data and Narrating with Data.« *Big Data & Society* 5(2): 1–10.
- Dreyfus, Hubert L. und Stuart E. Dreyfus. 2005. »Peripheral Vision. Expertise in Real World Contexts.« *Organization Studies* 26(5): 779–92.
- Dröge, Kai. 2019. »Arbeit, Wissen, Digitalisierung – eine soziologische Annäherung.« S. 19–33 in *Digitalisierung der Wissensarbeit. Interdisziplinäre Analysen und Fallstudien*, hg. von K. Dröge und A. Glauser. Frankfurt a.M.: Campus.
- Dröge, Kai und Andrea Glauser (Hg.). 2019. *Digitalisierung der Wissensarbeit. Interdisziplinäre Analysen und Fallstudien*. Frankfurt a.M./New York: Campus.
- Drosou, Marina, Hosagrahar V. Jagadish, Evaggelia Pitoura und Julia Stoyanovich. 2017. »Diversity in Big Data: A Review.« *Big Data* 5(2): 73–84.

- Duranton, Sylvain, Jörg Erlebach, Camille Brégé, Jane Danziger, Andrea Gallego und Marc Pauly. 2020. *What's Keeping Women Out of Data Science?* Paris/Toronto/Sydney/Boston/Frankfurt a.M.: Boston Consulting Group.
- EconSight. 2019. *Künstliche Intelligenz. Globale Entwicklungen, Anwendungsgebiete, Innovationstreiber und Weltklasseforschung.* Basel: EconSight.
- Edwards, Paul N. 2001. »Representing the Global Atmosphere: Computer Models, Data, and Knowledge about Climate Change.« S. 31–65 in *Changing the Atmosphere. Expert Knowledge and Environmental Governance.* Cambridge, MA: MIT Press.
- Eisenberg, Eric M. 1984. »Ambiguity as Strategy in Organizational Communication.« *Communication Monographs* 51(3): 227–42.
- Elish, Madeleine C. und danah boyd. 2018. »Situating Methods in the Magic of Big Data and AI.« *Communication Monographs* 85(1): 57–80.
- EMC. 2011. *Data Science Revealed: A Data-Driven Glimpse into the Burgeoning New Field.* Hopkinton, MA: EMC.
- Emirbayer, Mustafa und Victoria Johnson. 2008. »Bourdieu and Organizational Analysis.« *Theory and Society* 37(1): 1–44.
- Engemann, Christoph und Andreas Sudmann (Hg.). 2017. *Machine Learning – Medien, Infrastrukturen und Technologien der künstlichen Intelligenz.* Bielefeld: transcript.
- EPFL. 2016. *MS in Data Science - starting Fall 2017. Séance information aux étudiants bachelor en IC Dec 13, 2016.* Lausanne: École polytechnique fédérale de Lausanne.
- EPFL. 2017. *Master Data Science.* Lausanne: School of Computer and Communication Sciences, Ecole Polytechnique de Lausanne.
- EPFL & HEC Lausanne. 2019. *Data Science & Management. Certificate of Advanced Studies (CAS).* Lausanne: Swiss Data Science Center/Université de Lausanne.
- Esposito, Elena. 2016. »Pläne und die Zukunft. Das Unvorhersagbare gestalten.« S. 427–36 in *Innovationsgesellschaft heute*, hg. von W. Rammert, A. Windeler, H. Knoblauch und M. Hutter. Wiesbaden: Springer.
- Evans, James A. und Pedro Aceves. 2016. »Machine Translation: Mining Text for Social Theory.« *Annual Review of Sociology* 42: 21–50.
- Eyal, Gil. 2002. »Dangerous Liaison between Military Intelligence and Middle Eastern Studies in Israel.« *Theory and Society* 31(5): 653–93.
- Eyal, Gil. 2013a. »For a Sociology of Expertise: The Social Origins of the Autism Epidemic.« *American Journal of Sociology* 118(4): 863–907.
- Eyal, Gil. 2013b. »Spaces between Fields.« S. 158–82 in *Bourdieu and Historical Analysis*, hg. von P. S. Gorski. Durham/London: Duke University Press.
- Eyal, Gil, Brendan Hart, Emine Onculer, Neta Oren und Natasha Rossi. 2010. *The Autism Matrix: The Social Origins of the Autism Epidemic.* Cambridge: Polity.
- Eyal, Gil und Grace Pok. 2015. »What Is Security Expertise? From the Sociology of Professions to the Analysis of Networks of Expertise.« S. 53–75 in *Security Expertise. Practice, Power, Responsibility*, hg. von C. Villumsen Berling und T. Bueger. London/New York: Routledge.
- Favaretto, Maddalena, Eva De Clercq, Christophe Olivier Schneble und Bernice Simone Elger. 2020. »What Is Your Definition of Big Data? Researchers' Understanding of the Phenomenon of the Decade.« *PLOS ONE* 15(2): e0228987.
- Feldman, Michael, Frida Juldaschewa und Abraham Bernstein. 2017. »Data Analytics on Online Labor Markets: Opportunities and Challenges.« *arXiv: 1707.01790.*

- Felouzis, Georges, Christian Maroy und Agnès van Zanten. 2013. *Les Marchés Scolaires : Sociologie d'une Politique Publique d'éducation*. Paris: Presses universitaires de France.
- Fiesler, Casey, Natalie Garrett und Nathan Beard. 2020. »What Do We Teach When We Teach Tech Ethics?« S. 289–95 in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. New York: ACM.
- Fisher, Nicholas I. 2001. »Crucial Issues for Statistics in the Next Two Decades.« *International Statistical Review* 69(1): 3–4.
- Flick, Uwe. 2016. *Qualitative Sozialforschung. Eine Einführung*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.
- Fligstein, Neil und Doug McAdam. 2012. *A Theory of Fields*. Oxford: Oxford University Press.
- Florian, Michael. 2008. »Felder und Institutionen. Der soziologische Neo-Institutionalismus und die Perspektiven einer praxistheoretischen Institutionenanalyse.« *Berliner Journal für Soziologie* 18(1): 129–55.
- Florian, Michael und Frank Hillebrandt (Hg.). 2006. *Pierre Bourdieu: Neue Perspektiven für die Soziologie der Wirtschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Friedman, Jerome H. 2001. »The Role of Statistics in the Data Revolution?« *International Statistical Review* 69(1): 5.
- Friese, Susanne. 2012. *Qualitative Data Analysis with ATLAS.ti*. London: Sage.
- Fröhlich, Gerhard. 2009. »Feldanalysen: Wissenschaft.« S. 327–37 in *Bourdieu Handbuch. Leben – Werk – Wirkung*, hg. von G. Fröhlich und B. Rehbein. Stuttgart: J.B. Metzler.
- Fröhlich, Gerhard und Boike Rehbein (Hg.). 2009. *Bourdieu Handbuch. Leben – Werk – Wirkung*. Stuttgart: J.B. Metzler.
- Furnari, Santi. 2014. »Interstitial Spaces: Microinteraction Settings and the Genesis of New Practices Between Institutional Fields.« *Academy of Management Review* 39(4): 439–62.
- Furnari, Santi. 2016. »Institutional Fields as Linked Arenas: Inter-Field Resource Dependence, Institutional Work and Institutional Change.« *Human Relations* 69(3): 551–80.
- Galison, Peter. 1997. *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.
- Gardiner, Adrian, Cheryl Aasheim, Paige Rutner und Susan Williams. 2018. »Skill Requirements in Big Data: A Content Analysis of Job Advertisements.« *Journal of Computer Information Systems* 58(4): 374–84.
- Gehl, Robert W. 2015. »Sharing, Knowledge Management and Big Data: A Partial Genealogy of the Data Scientist.« *European Journal of Cultural Studies* 18(4–5): 413–28.
- Geiger, R. Stuart, Orianna DeMasi, Aaron Culich, Andreas Zoglauer, Diya Das, Fernando Hoces de la Guardia, Kellie Ottoboni, Marsha Fenner, Nelle Varoquaux, Rebecca Barter, Richard Barnes, Sara Stoudt, Stacey Dorton und Stefan van der Walt. 2019. *Best Practices for Fostering Diversity and Inclusion in Data Science: Report from the BIDS Best Practices in Data Science Series*. Berkeley: University of California, Berkeley Institute for Data Science.
- Geiger, R. Stuart, Charlotte Mazel-Cabasse, Chihoko Y. Cullens, Laura Norén, Brittaney Fiore-Gartland, Diya Das und Henry Brady. 2018. *Career Paths and Prospects in Academic Data Science: Report of the Moore-Sloan Data Science Environments Survey*. Berkeley: University of California, Berkeley Institute for Data Science.

- Gerlach, Philipp. 2014. *Der Wert der Arbeitskraft. Bewertungsinstrumente und Auswahlpraktiken im Arbeitsmarkt für Ingenieure*. Wiesbaden: Springer VS.
- Geser, Hans. 1983. »Qualifikatorische Anforderungsprofile bei wirtschaftlichen Führungspositionen.« *Soziale Welt* 34(4): 471–99.
- Getoor, Lise, David Culler, Eric de Sturler, David Ebert, Mike Franklin und Hosagrahar V. Jagadish. 2016. *Computing Research and the Emerging Field of Data Science*. Washington, D.C.: Computing Research Association.
- Getto, Barbara und Michael Kerres. 2018. »Wer macht was? Akteurskonstellationen in der digitalen Hochschulbildung.« S. 60–76 in *Digitalisierung und Hochschulentwicklung. Proceedings zur 26. Tagung der Gesellschaft für Medien in der Wissenschaft*. Vol. 74, hg. von B. Getto, P. Hintze und M. Kerres. Münster/New York: Waxmann.
- Giabbanelli, Philippe und Vijay Mago. 2016. »Teaching Computational Modeling in the Data Science Era.« *Procedia Computer Science* 80: 1968–77.
- Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott und Martin Trow. 1994. *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London/Thousand Oaks/New Delhi: Sage.
- Gieryn, Thomas F. 1983. »Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists.« *American Sociological Review* 48(6): 781–95.
- Gieryn, Thomas F. 1999. *Cultural Boundaries of Science. Credibility on the Line*. Chicago: University of Chicago Press.
- Gilmour, Jean A. 2006. »Hybrid Space: Constituting the Hospital as a Home Space for Patients.« *Nursing Inquiry* 13(1): 16–22.
- Gitelman, Lisa (Hg.). 2013. *»Raw Data« is an Oxymoron*. Cambridge, MA/London: MIT Press.
- Gitelman, Lisa und Virginia Jackson. 2013. »Introduction.« S. 1–14 in *»Raw Data« is an Oxymoron*, hg. von L. Gitelman. Cambridge, MA/London: MIT Press.
- Godin, Benoît. 2006. »The Linear Model of Innovation.« *Science, Technology, & Human Values* 31(6): 639–67.
- Godin, Benoît. 2009. »National Innovation System: The System Approach in Historical Perspective.« *Science, Technology, & Human Values* 34(4): 476–501.
- Godin, Benoît. 2016. »Technological Innovation: On the Origins and Development of an Inclusive Concept.« *Technology and Culture* 57(3): 527–56.
- Godin, Benoît. 2018. *Innovation Contested: The Idea of Innovation over the Centuries*. Abingdon/New York: Routledge.
- Gould, Robert. 2017. »Data Literacy Is Statistical Literacy.« *Statistics Education Research Journal* 16(1): 22–25.
- Grande, Edgar, Dorothea Jansen, Otfried Jarren, Arie Rip, Uwe Schimank und Peter Weingart (Hg.). 2013. *Neue Governance der Wissenschaft – Reorganisation – Externe Anforderungen – Medialisierung*. Bielefeld: transcript.
- Gray, Jim. 2009. »Jim Gray on eScience: A Transformed Scientific Method.« S. xvii–xxxi in *The Fourth Paradigm: Data-Intensive Scientific Discovery*, hg. von T. Hey, S. Tansley und K. Tolle. Redmond, WA: Microsoft Research.
- Green, Ben. 2018. *Data Science as Political Action: Grounding Data Science in a Politics of Justice*. Working Paper. *arXiv*: 1811.03435.
- Green, Francis. 2013. *Skills and Skilled Work*. Oxford: Oxford University Press.

- Greene, Daniel, Anna L. Hoffman und Luke Stark. 2019. »Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning.« S. 2122–2131 in *Proceedings of the 52nd Hawaii International Conference on System Sciences* (HICSS 2019). Honolulu: University of Hawaii.
- Griffiths, Thomas L. und Mark Steyvers. 2004. »Finding Scientific Topics.« *Proceedings of the National Academy of Sciences of the United States of America* 101(SUPPL. 1): 5228–35.
- Grimmer, Justin und Brandon M. Stewart. 2013. »Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts.« *Political Analysis* 21(3): 267–97.
- Grimson, William und Mike Murphy. 2015. »The Epistemological Basis of Engineering, and Its Reflection in the Modern Engineering Curriculum.« S. 161–78 in *Engineering Identities, Epistemologies and Values Engineering Education and Practice in Context*, Volume 2, hg. von S. C. Hyldgaard, C. Didier, A. Jamison, M. Meganck, C. Mitcham und B. Newburry. Cham: Springer International Publishing.
- Grommé, Francisca, Evelyn Ruppert und Baki Cakici. 2018. »Data Scientists: A New Faction of the Transnational Field of Statistics.« S. 33–61 in *Ethnography for a data-saturated world*, hg. von H. Knox und D. Nafus. Manchester: Manchester University Press.
- Grün, Bettina und Kurt Hornik. 2011. »Topicmodels: An R Package for Fitting Topic Models.« *Journal of Statistical Software* 40(1): 1–30.
- Grunwald, Armin. 2019. »Digitalisierung als Prozess. Ethische Herausforderungen inmitten allmählicher Verschiebungen zwischen Mensch, Technik und Gesellschaft.« *Zeitschrift für Wirtschafts- und Unternehmensethik* 20(2): 121–45.
- GSEM (2016): *Master in Business Analytics. A new programme for a new approach to management*. Genève: Université de Genève, Geneva School of Economics and Management.
- Guenduez, Ali A., Tobias Mettler und Kuno Schedler. 2020. »Technological Frames in Public Administration: What Do Public Managers Think of Big Data?« *Government Information Quarterly* 37(1): 101406.
- Guerrier, Yvonne, Christina Evans, Judith Glover und Cornelia Wilson. 2009. »Technical, but Not Very...«: Constructing Gendered Identities in IT-Related Employment.« *Work, Employment and Society* 23(3): 494–511.
- Gugerli, David. 2018. *Wie die Welt in den Computer kam. Zur Entstehung digitaler Wirklichkeit*. Frankfurt a.M.: S. Fischer.
- Gugerli, David, Patrick Kupper und Daniel Speich (Hg.). 2005. *Die Zukunftsmaschine: Konjunkturen der ETH Zürich 1855–2005*. Zürich: Chronos.
- Gupta, Babita, Michael Goul und Barbara Dinter. 2015. »Business Intelligence and Big Data in Higher Education: Status of a Multi-Year Model Curriculum Development Effort for Business School Undergraduates, MS Graduates, and MBAs.« *Communications of the Association for Information Systems* 36: 449–76.
- Guston, David H. 1999. »Stabilizing the Boundary between US Politics and Science. The Role of the Office of Technology Transfer as a Boundary Organization.« *Social Studies of Science* 29(1): 87–111.
- Guston, David H. 2000. *Between Politics and Science. Assuring the Integrity and Productivity of Research*. Cambridge: Cambridge University Press.

- Hall, Stuart (Hg.). 1997. *Representations: Cultural Representations and Signifying Practices*. London/Thousand Oaks/New Delhi: Sage.
- Hamann, Julian. 2018. »The Production of Research Elites: Research Performance Assessment in the United Kingdom.« S. 175–99 in *Universities and the Production of Elites*, hg. von R. Bloch, A. Mitterle, C. Paradeise und T. Peter. Cham: Palgrave Macmillan.
- Hammerbacher, Jeff. 2009. »Information Platforms and the Rise of the Data Scientist.« S. 73–84 in *Beautiful Data: The Stories Behind Elegant Data Solutions*, hg. von T. Segaran und J. Hammerbacher. Sebastopol, CA: O'Reilly Media.
- Han, Jiawei, Micheline Kamber und Jian Pei. 2012. *Data Mining. Concepts and Techniques*. Waltham, MA: Morgan Kaufmann.
- Hänzi, Denis. 2015. »Verheissungsvolle Potenziale. Ein neues Ideal der Zukunftssicherheit (Selbst-)Investition?« *Berliner Journal für Soziologie* 25(1): 215–36.
- Harris, Harlan D., Sean Patrick Murphy und Marck Vaisman. 2013. *Analyzing the Analyzers. An Introspective Survey of Data Scientists and Their Work*. Sebastopol, CA: O'Reilly Media.
- Harris, Jeanne G., Nathan Shetterley, Allan E. Alter und Krista Schnell. 2013. *The Team Solution to the Data Scientist Shortage*. Accenture Institute for High Performance.
- Hasler Stiftung. 2018. *Berufe in der Informatik*. Bern: Hasler Stiftung.
- Hasse, Raimund. 2015. »Organisationsforschung und Wettbewerb« S. 109–122 in *Zur Zukunft der Organisationssoziologie*, hg. von M. Apelt und U. Wilkesmann. Wiesbaden: Springer.
- Haven, Janet und danah boyd. 2020. *Philanthropy's Techno-Solutionism Problem*. Miami: Knight Foundation.
- Hayashi, Chikio. 1998. »What Is Data Science ? Fundamental Concepts and a Heuristic Example.« S. 40–51 in *Data Science, Classification, and Related Methods*, hg. von C. Hayashi, K. Yajima, H.-H. Bock, N. Ohsumi, Y. Tanaka und Y. Baba. Tokyo: Springer Japan.
- Hayashi, Chikio, Keiji Yajima, Hans-Hermann Bock, Noboru Ohsumi, Yutaka Tanaka und Yasumasa Baba (Hg.). 1998. *Data Science, Classification, and Related Methods*. Tokyo: Springer Japan.
- Hedmo, Tina, Kerstin Sahlin-Andersson und Linda Wedlin. 2007. »Is a Global Organizational Field of Higher Education Emerging? Management Education as an Early Example.« S. 154–75 in *Towards a Multiversity? Universities between Global Trends and National Traditions*, hg. von G. Krücken, A. Kosmützky und M. Torka. Bielefeld: transcript.
- Heidrich, Jens, Pascal Bauer und Daniel Krupka. 2018. *Future Skills: Ansätze zur Vermittlung von Data Literacy in der Hochschulbildung*. Arbeitspapier Nr. 37 Berlin: Hochschulforum Digitalisierung.
- Heimstädt, Maximilian und Georg Reischauer. 2019. »Framing Innovation Practices in Interstitial Issue Fields: Open Innovation in the NYC Administration.« *Innovation* 21(1): 128–50.
- Heintz, Bettina. 2008. »Governance by Numbers. Zum Zusammenhang von Quantifizierung und Globalisierung am Beispiel der Hochschulpolitik.« S. 110–128 in *Governance von und durch Wissen*, hg. von G. F. Schuppert und A. Vosskuhle. Baden-Baden: Nomos.

- Heintz, Bettina. 2010. »Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs.« *Zeitschrift für Soziologie* 39(3): 162–81.
- Heintz, Bettina. 2016. »Wir leben im Zeitalter der Vergleichung.« Perspektiven einer Soziologie des Vergleichs.« *Zeitschrift für Soziologie* 45(5): 305–23.
- Herberg, Jeremias. 2018a. *Illusio Fachkräftemangel. Der Zwischenraum von Bildung und Wirtschaft in Deutschland und Nordkalifornien*. Wiesbaden: Springer.
- Herberg, Jeremias. 2018b. *Skills Brokers in the San Francisco East Bay: Challenges and Opportunities for Creating Equitable Cross-Sector Collaboration Through Workforce Intermediaries*. Working Paper, Center for Cities + Schools, April 2018, Berkeley: University of California, Berkeley.
- Hermida, Alfred und Mary Lynn Young. 2017. »Finding the Data Unicorn. A Hierarchy of Hybridity in Data and Computational Journalism.« *Digital Journalism* 5(2): 159–76.
- Hey, Tony, Steward Tansley und Kristin Tolle (Hg.). 2009. *The Fourth Paradigm. Data-Intensive Scientific Discovery*. Redmond, WA: Microsoft Research.
- Hirsch-Kreinsen, Hartmut, Peter Ittermann und Jonathan Falkenberg (Hg.). 2018. *Digitalisierung industrieller Arbeit. Die Vision Industrie 4.0 und ihre sozialen Herausforderungen* (2., aktualisierte Auflage). Baden-Baden: Nomos.
- Hofmann, Heike und Susan VanderPlas. 2017. »All of This Has Happened Before. All of This Will Happen Again: Data Science.« *Journal of Computational and Graphical Statistics* 26(4): 775–78.
- Holley, Karri. 2009. »The Challenge of an Interdisciplinary Curriculum: A Cultural Analysis of a Doctoral-Degree Program in Neuroscience.« *Higher Education* 58(2): 241–55.
- Holmes, Douglas R. 2018. »A Tractable Future: Central Banks in Conversation with Their Publics.« S. 173–93 in *Uncertain Futures: Imaginaries, Narratives, and Calculation in the Economy*, hg. von J. Beckert und R. Bronk. Oxford: Oxford University Press.
- Honegger, Claudia, Hans-Ulrich Jost, Susanne Burren und Pascal Jurt. 2007. *Konkurrierende Deutungen des Sozialen. Geschichts-, Sozial- und Wirtschaftswissenschaften im Spannungsfeld von Politik und Wissenschaft*. Zürich: Chronos Verlag.
- Hookway, Branden. 2014. *Interface*. Cambridge, MA/London: The MIT Press.
- Hornik, Kurt, Patrick Mair, Johannes Rauch, Wilhelm Geiger, Christian Buchta und Ingo Feinerer. 2013. »The Textcat Package for N-Gram Based Text Categorization in R.« *Journal of Statistical Software* 52(6).
- Horvitz, Eric und Tom Mitchell. 2010. *From Data to Knowledge to Action: A Global Enabler for the 21st Century*. Washington, D.C.: Computing Community Consortium.
- Houben, Daniel und Bianca Prietl (Hg.). 2018. *Datengesellschaft. Einsichten in die Digitalisierung des Sozialen*. Bielefeld: transcript.
- Huppenkothen, Daniela, Anthony Arendt, David W. Hogg, Karthik Ram, Jacob T. VanderPlas und Ariel Rokem. 2018. »Hack Weeks as a Model for Data Science Education and Collaboration.« *Proceedings of the National Academy of Sciences* 115(36): 8872–77.
- Ignatow, Gabe und Dragomir Radev. 2016. *Text Mining: A Guidebook for the Social Sciences*. London: SAGE.
- Irizarry, Rafael A. 2020. »The Role of Academia in Data Science Education.« *Harvard Data Science Review* (2.1): 1–8.

- Ismail, Nur Amie und Wardah Zainal Abidin. 2016. »Data Scientist Skills.« *IOSR Journal of Mobile Computing & Application* 3(4): 52–61.
- IWGDD. 2009. *Harnessing the Power of Digital Data for Science and Society*. Report of the Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council. Washington, D.C.
- Janning, Frank. 1998. *Das politische Organisationsfeld. Politische Macht und soziale Homologie in komplexen Demokratien*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Jasanoff, Sheila. 2004. *States of Knowledge: The Co-Production of Science and the Social Order*. New York/London: Routledge.
- Jasanoff, Sheila. 2015. »Future Imperfect: Science, Technology, and the Imaginations of Modernity.« S. 1–33 in *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*, hg. von S. Jasanoff und S.-H. Kim. Chicago: Chicago University Press.
- Jasanoff, Sheila und Sang-Hyun Kim (Hg.). 2015. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. Chicago: Chicago University Press.
- Jessop, Bob. 2008. »A Cultural Political Economy of Competitiveness and Its Implications for Higher Education.« S. 13–39 in *Education and the Knowledge-Based Economy in Europe*, hg. von B. Jessop, N. Fairclough und R. Wodak. Rotterdam: Sense Publishers.
- Jessop, Bob, Norman Fairclough und Ruth Wodak (Hg.). 2008. *Education and the Knowledge-Based Economy in Europe*. Rotterdam: Sense Publishers.
- Jongbloed, Ben, Jürgen Enders und Carlo Salerno. 2008. »Higher Education and Its Communities: Interconnections, Interdependencies and a Research Agenda.« *Higher Education* 56(3): 303–24.
- Jost, Hans-Ulrich. 2007. »Sozialwissenschaften als Staatswissenschaften?« S. 81–182 in *Konkurrierende Deutungen des Sozialen. Geschichts-, Sozial- und Wirtschaftswissenschaften im Spannungsfeld von Politik und Wissenschaft*, hg. von C. Honegger, H.-U. Jost, S. Burren und P. Jurt. Zürich: Chronos Verlag.
- Jurafsky, Daniel und James H. Martin. 2018. »N-Gram Language Models.« in *Speech and Language Processing*. London: Prentice Hall.
- Kandel, Sean, Andreas Paepcke, Joseph M. Hellerstein und Jeffrey Heer. 2012. »Enterprise Data Analysis and Visualization: An Interview Study.« *IEEE Transactions on Visualization and Computer Graphics* 18(12): 2917–26.
- Kane, Michael J. 2014. »Cleveland's Action Plan and the Development of Data Science over the Last 12 Years.« *Statistical Analysis and Data Mining: The ASA Data Science Journal* 7(6): 423–24.
- Kastner, Jens. 2002. »Fleischgewordene Höllenmaschine«. Staatlicher Rassismus als Neoliberale Politik.« S. 319–41 in *Theorie als Kampf? Zur politischen Soziologie Pierre Bourdieus*, hg. von U. Bittlingmeyer, R. Eickelpasch, J. Kastner und C. Rademacher. Opladen: Leske + Budrich.
- Kastner, Jens. 2011. »Koloniale Klassifikationen: Zur Genese postkolonialer Sozialtheorie im Kolonialen Algerien bei Frantz Fanon und Pierre Bourdieu.« S. 277–302 in *Pierre Bourdieu und die Kulturwissenschaften. Zur Aktualität eines undisziplinierten Denkens*, hg. von D. Suber, H. Schäfer und S. Prinz. Konstanz: UVK.
- Katz, Luba. 2019. *Evaluation of the Moore-Sloan Data Science Environments*. Final Report. Cambridge, MA: Abt Associates.

- Kauermann, Göran. 2018. »Data Science als Studiengang.« S. 87–95 in *Big Data. Chancen, Risiken, Entwicklungstendenzen*, hg. von C. König, J. Schröder und E. Wiegand. Wiesbaden: Springer.
- Kauermann, Göran und Helmut Küchenhoff. 2016. »Statistik, Data Science und Big Data.« *AStA Wirtschafts- und Sozialstatistisches Archiv* 10(2–3): 141–50.
- Kearnes, Matthew. 2013. »Performing Synthetic Worlds: Situating the Bioeconomy.« *Science and Public Policy* 40(4): 453–65.
- Kellogg, Katherine C. 2009. »Operating Room: Relational Spaces and Microinstitutional Change in Surgery.« *American Journal of Sociology* 115(3): 657–711.
- Kellogg, Katherine C., Wanda J. Orlikowski und JoAnne Yates. 2006. »Life in the Trading Zone: Structuring Coordination Across Boundaries in Postbureaucratic Organizations.« *Organization Science* 17(1): 22–44.
- Kensing, Finn und Jeanette Blomberg. 1998. »Participatory Design: Issues and Concerns.« *Computer Supported Cooperative Work (CSCW)* 7(3–4): 167–85.
- Kim, Miryung, Thomas Zimmermann, Robert DeLine und Andrew Begel. 2016. »The Emerging Role of Data Scientists on Software Development Teams.« S. 96–107 in *Proceedings of the 38th International Conference on Software Engineering – ICSE* 16. Austin: ACM Press.
- Kim, Miryung, Thomas Zimmermann, Robert DeLine und Andrew Begel. 2018. »Data Scientists in Software Teams: State of the Art and Challenges.« *IEEE Transactions on Software Engineering* 44(11): 1024–38.
- King, Gary und Nathaniel Persily. 2020. »A New Model for Industry–Academic Partnerships.« *PS: Political Science & Politics* 53(4): 703–9.
- Kirchner, Corinne und John W. Mohr. 2010. »Meanings and Relations: An Introduction to the Study of Language, Discourse and Networks.« *Poetics* 38(6): 555–66.
- Kitchin, Rob. 2014. »Big Data, New Epistemologies and Paradigm Shifts.« *Big Data & Society* 1(1): 1–12.
- Kitchin, Rob und Tracey Lauriault. 2018. »Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work.« S. 3–20 in *Thinking Big Data in Geography. New Regimes, New Research*, hg. von J. Thatcher, J. Eckert und A. Shears. Lincoln: University of Nebraska Press.
- Kitchin, Rob und Gavin McArdle. 2016. »What Makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets.« *Big Data & Society* 3(1): 1–10.
- Klein, Julie Thompson. 1996. *Crossing Boundaries: Knowledge, Disciplinarity, and Interdisciplinarity*. Charlottesville: University Press of Virginia.
- Kloefkorn, Tyler, Michael Boardman, Nicholas J. Horton und Brandeis Marshall. 2020. »National Academies’ Roundtable on Data Science Postsecondary Education.« S. 956–57 in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. New York: ACM.
- Kluttz, Daniel N. und Neil Fligstein. 2016. »Varieties of Sociological Field Theory.« S. 185–204 in *Handbook of Contemporary Sociological Theory*, hg. von S. Abrutyn. Cham: Springer International Publishing.
- Knight, David B., Lisa R. Lattuca, Ezekiel W. Kimball und Robert D. Reason. 2013. »Understanding Interdisciplinarity: Curricular and Organizational Features of Undergraduate Interdisciplinary Programs.« *Innovative Higher Education* 38(2): 143–58.
- Knorr-Cetina, Karin. 2002. *Wissenskulturen. Ein Vergleich naturwissenschaftlicher Wissensformen*. Frankfurt a.M.: Suhrkamp.

- Köhler, Christoph und Alexandra Krause. 2010. »Kapitel VII Gestaltung von Beschäftigungsverhältnissen: Betriebliche Beschäftigungspolitik.« S. 387–412 in *Handbuch Arbeitssoziologie*, hg. von F. Böhle, G. G. Voss und G. Wachtler. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kolany-Raiser, Barbara, Reinhard Heil, Carsten Orwat und Thomas Hoeren (Hg.). 2018. *Big Data und Gesellschaft. Eine multidisziplinäre Annäherung*. Wiesbaden: Springer.
- Korff, Valeska P., Achim Oberg und Walter W. Powell. 2017. »Governing the Crossroads: Interstitial Communities and the Fate of Nonprofit Evaluation.« S. 85–105 in *Networked Governance*, hg. von B. Hollstein, W. Matiaske und K.-U. Schnapp. Cham: Springer.
- Koselleck, Reinhart. 2006. *Vergangene Zukunft. Zur Semantik geschichtlicher Zeiten*. Frankfurt a.M.: Suhrkamp.
- Krause, Andreas. 2019. »Künstliche Intelligenz: Schweiz forscht an der Spitze.« *Die Volkswirtschaft* (12): 4–8.
- Krause, Monika. 2018. »How Fields Vary.« *The British Journal of Sociology* 69(1): 3–22.
- Kreuter, Frauke, Florian Keusch und Evgenia Samoilova. 2018. »International Program in Survey and Data Science.« S. 1–6 in *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics*. Vol. 10, hg. von M. A. Sorto, A. White und L. Guyot. Kyoto: International Statistical Institute.
- Kriesi, Irene, Marlis Buchmann und Stefan Sacchi. 2010. »Variation in Job Opportunities for Men and Women in the Swiss Labor Market 1962–1989.« *Research in Social Stratification and Mobility* 28(3): 309–23.
- Kross, Sean, Roger D. Peng, Brian S. Caffo, Ira Gooding und Jeffrey T. Leek. 2020. »The Democratization of Data Science Education.« *The American Statistician* 74(1): 1–7.
- Krücken, Georg. 2017. »Die Transformation von Universitäten in Wettbewerbsakteure.« *Beiträge zur Hochschulforschung* 3–4(April 2016): 10–29.
- Krücken, Georg und Frank Meier. 2006. »Turning the University into an Organizational Actor.« S. 241–57 in *Globalization and Organization: World Society and Organizational Change*, hg. von G. S. Drori, J. W. Meyer und H. Hwang. Oxford: Oxford University Press.
- Krummenacher, Jörg. 2019. »Digital lehrt jeder Kanton, wie er will.« *Neue Zürcher Zeitung*, 28.01.2019: 9.
- Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions* (3. Auflage). Chicago: University of Chicago Press.
- Kühner, Angela, Phil C. Langer und Panja Schweder. 2013. »Reflexive Wissensproduktion: Eine Einführung.« S. 7–18 in *Reflexive Wissensproduktion*, hg. von P. C. Langer, A. Kühner und P. Schweder. Wiesbaden: Springer.
- Kurtz, Thomas. 2010. »Der Kompetenzbegriff in der Soziologie.« S. 7–25 in *Soziologie der Kompetenz*, hg. von T. Kurtz und M. Pfadenhauer. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Labrinidis, Alexandros und Hosagrahar V. Jagadish. 2012. »Challenges and Opportunities with Big Data.« *Proceedings of the VLDB Endowment* 5(12): 2032–33.
- Lahire, Bernard. 2014. »The Limits of the Field: Elements for a Theory of the Social Differentiation of Activities.« S. 62–101 in *Bourdieu's Theory of Social Fields. Concepts and Applications*, hg. von M. Hilgers und E. Mangez. London: Routledge.

- Lakatos, Imre. 1978. *The Methodology of Scientific Research Programmes*, hg. von J. Worral und G. Currie. Cambridge: Cambridge University Press.
- Lamont, Michèle und Virág Molnár. 2002. »The Study of Boundaries in the Social Sciences.« *Annual Review of Sociology* 28(1): 167–95.
- Langer, Antje. 2010. »Transkribieren – Grundlagen und Regeln.« S. 515–26 in *Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft*, hg. von B. Friebertshäuser, A. Langer und A. Prenzel. Weinheim/München: Juventa.
- Lattuca, Lisa R. und Joan S. Stark. 2009. *Shaping the College Curriculum: An Academic Plan*. San Francisco: John Wiley & Sons.
- Laux, Henning. 2016. »Hybridorganisationen.« *Soziale Welt* 67(3): 313–32.
- Law, John. 1991. »Introduction: Monsters, Machines and Sociotechnical Relations.« S. 1–25 in *A Sociology of Monsters: Essays on Power, Technology and Domination*, hg. von J. Law. London/New York: Routledge.
- Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy und Marshall Van Alstyne. 2009. »Computational Social Science.« *Science* 323(5915): 721–23.
- Lee, Justin. 2005. *Investigating the Hybridity of »Wellness« Practices*. Working Paper. Los Angeles: Department of Sociology, University of California, Los Angeles. Online: <https://escholarship.org/uc/item/88c4t567>
- Leitch, Shirley und Sally Davenport. 2007. »Strategic Ambiguity as a Discourse Practice: The Role of Keywords in the Discourse on »sustainable« Biotechnology.« *Discourse Studies* 9(1): 43–61.
- Lemaine, Gerald, Roy Macleod, Michael Mulkay und Peter Weingart (Hg.). 1976. *Perspectives on the Emergence of Scientific Disciplines*. Mouton/Den Haag/Paris: Maison des Sciences de l'Homme.
- Leonelli, Sabina. 2014. »What Difference Does Quantity Make? On the Epistemology of Big Data in Biology.« *Big Data & Society* 1(1): 1–14.
- Leonelli, Sabina. 2016. *Data-Centric Biology: A Philosophical Study*. Chicago: University of Chicago Press.
- Lester, Richard K. und Michael J. Piore. 2004. *Innovation. The Missing Dimension*. Cambridge, MA: Harvard University Press.
- Lettkemann, Eric. 2016. *Stabile Interdisziplinarität. Eine Biografie der Elektronenmikroskopie aus historisch-soziologischer Perspektive*. Baden-Baden: Nomos.
- Ley, Christophe und Stéphane P. A. Bordas. 2018. »What Makes Data Science Different? A Discussion Involving Statistics2.0 and Computational Sciences.« *International Journal of Data Science and Analytics* 6(3): 167–75.
- Lindvig, Katrine, Catherine Lyall und Laura R. Meagher. 2019. »Creating Interdisciplinary Education within Monodisciplinary Structures: The Art of Managing Interstitiality.« *Studies in Higher Education* 44(2): 347–60.
- Lipp, Benjamin. 2017. »Analytik des Interfacing. Zur Materialität technologischer Verschaltung in prototypischen Milieus robotisierter Pflege.« *Behemoth. A Journal on Civilisation* 10(1): 107–29.
- Liu, Sida. 2018. »Boundaries and Professions: Toward a Processual Theory of Action.« *Journal of Professions and Organization* 5(1): 45–57.
- Liu, Sida. 2021. »Between social spaces«. *European Journal of Social Theory* 24(1): 123–39.
- Livingstone, Sonia. 2014. »What Is Media Literacy?« *Intermedia* 32(3): 18–20.

- Lowrie, Ian. 2017. »Algorithmic Rationality: Epistemology and Efficiency in the Data Sciences.« *Big Data & Society* 4(1): 1–13.
- Lowrie, Ian. 2018. »Becoming a Real Data Scientist. Expertise, Flexibility and Lifelong Learning.« S. 62–81 in *Ethnography for a data-saturated world*, hg. von H. Knox und D. Nafus. Manchester: Manchester University Press.
- Lübcke, Maren und Klaus Wannemacher. 2018. *Vermittlung von Datenkompetenzen an den Hochschulen: Studienangebote im Bereich Data Science*. Hannover: HIS-Institut für Hochschulentwicklung e.V.
- Maasen, Sabine. 2010. »Transdisziplinarität Revisited – Dekonstruktion eines Programms zur Demokratisierung der Wissenschaft.« S. 247–68 in *Inter- und Transdisziplinarität im Wandel?*, hg. von A. Bogner, K. Kastenhofer und H. Torgersen. Baden-Baden: Nomos.
- Maasen, Sabine. 2019. »Die Transzendenz der Technik – die Immanenz der Religion: Das Beispiel Digitalisierung.« S. 237–54 in *Religion und Gesellschaft. Sinnstiftungssysteme im Konflikt*, hg. von F. W. Graf und J.-U. Hartmann. Berlin/Boston: De Gruyter.
- Maasen, Sabine und Sascha Dickel. 2016. »Partizipation, Responsivität, Nachhaltigkeit.« S. 1–18 in *Handbuch Wissenschaftspolitik*, hg. von D. Simon, A. Knie, S. Hornbostel und K. Zimmermann. Wiesbaden: Springer.
- Mackenzie, Adrian. 2013. »Programming Subjects in the Regime of Anticipation: Software Studies and Subjectivity.« *Subjectivity* 6(4): 391–405.
- Mackenzie, Donald und Yuval Millo. 2003. »Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange.« *American Journal of Sociology* 109(1): 107–45.
- MacKenzie, Donald, Fabian Muniesa und Lucia Siu (Hg.). 2007. *Do Economists Make Markets? On the Performativity of Economics*. Princeton: Princeton University Press.
- Maesse, Jens. 2010. *Die vielen Stimmen des Bologna-Prozesses. Zur diskursiven Logik eines bildungspolitischen Programms*. Bielefeld: transcript.
- Maier, Daniel, A. Waldherr, P. Miltner, G. Wiedemann, A. Niekler, A. Keinert, B. Pfetsch, G. Heyer, U. Reber, T. Häußler, H. Schmid-Petri und S. Adam. 2018. »Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology.« *Communication Methods and Measures* 12(2–3): 93–118.
- Mäkinen, Elina I. 2018. »Action in the Space Between.« *Science, Technology, & Human Values* 43(2): 349–74.
- Mäkinen, Elina I., Eliza D. Evans und Daniel A. McFarland. 2020. »The Patterning of Collaborative Behavior and Knowledge Culminations in Interdisciplinary Research Centers.« *Minerva* 58(1): 71–95.
- Malazita, James W. 2019. »Epistemic Infrastructure, the Instrumental Turn, and the Digital Humanities.« in *Debates in the Digital Humanities: Institutions, Infrastructures at the Interstices*, hg. von A. McGrail, A. D. Nieves und S. Senier. Minneapolis: University of Minnesota Press.
- Mämecke, Thorben, Jan-Hendrik Passoth und Josef Wehner (Hg.). 2018. *Bedeutende Daten: Modelle, Verfahren und Praxis der Vermessung und Verdattung im Netz*. Wiesbaden: Springer.
- Manovich, Lev. 2015. »Data Science and Digital Art History.« *International Journal of Digital Art History* 1(1): 13–35.

- Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh und Angela Hung Byers. 2011. *Big Data: The next Frontier for Innovation, Competition, and Productivity*. Seoul/San Francisco/London/Washington: McKinsey Global Institute.
- Manyika, James, Michael Chui, Susan Lund und Sree Ramaswamy. 2017. *What's Now and next in Analytics, AI and Automation*. San Francisco: McKinsey Global Institute.
- Marcovich, Anne und Terry Shinn. 2011. »Where Is Disciplinarity Going? Meeting on the Borderland.« *Social Science Information* 50(3-4): 582-606.
- Marcus, George E. 1995. »Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography.« *Annual Review of Anthropology* 24(1995): 95-117.
- Markow, Will, Soumya Braganza, Debbie Hughes und Steven Miller. 2017. *The Quant Crunch. How Demand for Data Science Skills Is Disrupting the Job Market*. Boston: Burning Glass Technologies.
- Marquardt, Henning. 2016. »Familial Spaces: ›Yard‹ and ›Matrifocal Family‹ in Pre-Independence Jamaican Literature.« S. 157-75 in *Symbolism 16*, hg. von R. Ahrens, F. Kläger und K. Stierstorfer. Berlin/Boston: De Gruyter.
- Marr, Bernhard. 2018. »Forget Data Scientists and Hire a Data Translator Instead.« In *Forbes Online*, 12.03.2018. Online: <https://www.forbes.com/sites/bernardmarr/2018/03/12/forget-data-scientists-and-hire-a-data-translator-instead/>
- Martin, Emily. 1993. *Flexible Bodies: Tracking Immunity in American Culture from the Days of Polio to the Age of AIDS*. Boston: Beacon Press.
- Mason, Hilary und Chris Wiggins. 2010. »A Taxonomy of Data Science.« In *Dataists*, 25.09.2010. Online: www.dataists.com/2010/09/a-taxonomy-of-data-science/ (Original nicht mehr verfügbar; letzter Zugriff: 25.08.2021)
- Maurer, Andrea. 2006. »Wirtschaftssoziologie als soziologische Analyse ökonomischer Felder? Bourdieus Beitrag zur Wirtschaftssoziologie.« S. 127-46 in *Pierre Bourdieu: Neue Perspektiven für die Soziologie der Wirtschaft*, hg. von F. Hillebrandt und M. Florian. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mayer, Katja und Momin M. Malik. 2019. »Critical Data Scientists at Work: Summary Report of the ICWSM-2019 Workshop on Critical Data Science.« *Thirteenth International AAAI Conference on Web and Social Media (ICWSM-19)*. München.
- Mayer-Schoenberger, Viktor und Kenneth N. Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray.
- Mayring, Philipp. 2010. *Qualitative Inhaltsanalyse* (11. überarbeitete und erweiterte Auflage). Weinheim: Beltz Verlag.
- McAfee, Andrew und Erik Brynjolfsson. 2012. »Big Data: The Management Revolution.« *Harvard Business Review* 90(10): 61-67.
- McFarland, Daniel A., Kevin Lewis und Amir Goldberg. 2016. »Sociology in the Era of Big Data: The Ascent of Forensic Social Science.« *The American Sociologist* 47(1): 12-35.
- McLevey, John. 2015. »Understanding Policy Research in Liminal Spaces: Think Tank Responses to Diverging Principles of Legitimacy.« *Social Studies of Science* 45(2): 270-93.
- McNamara, Amelia, Nicholas J. Horton und Benjamin S. Baumer. 2017. »Greater Data Science at Baccalaureate Institutions.« *Journal of Computational and Graphical Statistics* 26(4): 781-83.
- Medvetz, Thomas. 2012. *Think Tanks in America*. Chicago: University of Chicago Press.

- Meier, Frank. 2009. *Die Universität als Akteur. Zum institutionellen Wandel der Hochschulorganisation*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Meier, Frank und Uli Meyer. 2020. »Organisationen und heterogene Umwelten.« S. 75–100 in *Neo-Institutionalismus*. Bielefeld: transcript.
- Méndez Fernández, Daniel und Jan-Hendrik Passoth. 2019. »Empirical Software Engineering: From Discipline to Interdiscipline.« *Journal of Systems and Software* 148(2): 170–79.
- Merz, Martina und Philippe Sormani. 2016. »Configuring New Research Fields: How Policy, Place, and Organization Are Made to Matter.« S. 1–22 in *The Local Configuration of New Research Fields: On Regional and National Diversity*, hg. von M. Merz und P. Sormani. Cham: Springer.
- Meyer, John W. und Ronald L. Jepperson. 2000. »The ›Actors‹ of Modern Society: The Cultural Construction of Social Agency.« *Sociological Theory* 18(1): 100–120.
- Meyer, John W. und Francisco O. Ramirez. 2006. »Die globale Institutionalisierung der Bildung.« S. 212–35 in *Weltkultur*, hg. von G. Krücken. Frankfurt a.M.: Suhrkamp.
- Meyer, Morgan und Matthew Kearnes. 2013. »Introduction to Special Section: Intermediaries between Science, Policy and the Market.« *Science and Public Policy* 40(4): 423–29.
- Meyer, Uli. 2020. »Industrie 4.0 als sozio-technische Zukunftsvorstellung. Zur Bedeutung von organisationaler Sinnerzeugung und -Stiftung.« S. 349–78 in *Soziologie des Digitalen – Digitale Soziologie?* Baden-Baden: Nomos.
- Mikroyannidis, Alexander, John Domingue, Christopher Phethean, Gareth Beeston und Elena Simperl. 2018. »Designing and Delivering a Curriculum for Data Science Education Across Europe.« S. 540–50 in *Teaching and Learning in a Digital World. Proceedings of the 20th International Conference on Interactive Collaborative Learning – Volume 2*, hg. von M. E. Auer, D. Guralnick und I. Simonics. Cham: Springer International Publishing.
- Mische, Ann. 2009. »Projects and Possibilities: Researching Futures in Action.« *Sociological Forum* 24(3): 694–704.
- Mische, Ann. 2014. »Measuring Futures in Action: Projective Grammars in the Rio + 20 Debates.« *Theory and Society* 43(3–4): 437–64.
- Mohr, John W. 2000. »Introduction: Structures, Institutions, and Cultural Analysis.« *Poetics* 27(2–3): 57–68.
- Mohr, John W. 2013. »Bourdieu’s Relational Method in Theory and in Practice: From Fields and Capitals to Networks and Institutions (and Back Again).« S. 101–35 in *Applying Relational Sociology*, hg. von F. Dépelteau und C. Powell. New York: Palgrave Macmillan.
- Mohr, John W. und Petko Bogdanov. 2013. »Introduction—Topic Models: What They Are and Why They Matter.« *Poetics* 41(6): 545–69.
- Moore-Sloan Data Science Environments. 2018. *Creating Institutional Change in Data Science*. Online: http://msdse.org/files/Creating_Institutional_Change.pdf
- Morozov, Evgeny. 2013. *To Save Everything, Click Here. The Folly of Technological Solutionism*. New York: PublicAffairs.
- Morril, Calvin. 2017. »Institutional Change Through Interstitial Emergence: The Growth of Alternative Dispute Resolution in U.S. Law, 1970–2000.« *Brazilian Journal of Empirical Legal Studies* 4(1): 10–36.

- Muller, Michael J. und Sarah Kuhn. 1993. »Participatory Design.« *Communications of the ACM* 36(6): 24–28.
- Münch, Richard. 2009. *Globale Eliten, lokale Autoritäten. Bildung und Wissenschaft unter dem Regime von PISA, McKinsey & Co.* Frankfurt a.M.: Suhrkamp.
- Münch, Richard. 2014. *Academic Capitalism: Universities in the Global Struggle for Excellence.* New York/Abingdon: Routledge.
- Munzert, Simon, Christian Rubba, Peter Meissner und Dominic Nyhuis. 2015. *Automated Data Collection with R. A Practical Guide to Web Scraping and Text Mining.* Chichester: Wiley.
- Muster, Judith und Stefanie Büchner. 2018. »Datafizierung und Organisation.« S. 253–78 in *Datengesellschaft. Einsichten in die Datafizierung des Sozialen*, hg. von D. Houben und B. Priel. Bielefeld: transcript.
- Mützel, Sophie. 2010. »Koordinierung von Märkten durch narrativen Wettbewerb.« S. 87–106 in *Wirtschaftssoziologie. Sonderheft 49 der Kölner Zeitschrift für Soziologie und Sozialpsychologie*, hg. von J. Beckert und C. Deutschmann. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mützel, Sophie. 2015a. »Facing Big Data: Making Sociology Relevant.« *Big Data & Society* 2(2): 1–4.
- Mützel, Sophie. 2015b. »Structures of the Tasted: Restaurant Reviews in Berlin Between 1995 and 2012.« S. 147–67 in *Moments of Valuation. Exploring Sites of Dissonance*, hg. von A. B. Antal, M. Hutter und D. Stark. Oxford: Oxford University Press.
- Mützel, Sophie. 2015c. »Topic Modeling.« S. 411 in *Methoden-Lexikon für die Sozialwissenschaften*, hg. von R. Diaz-Bone und C. Weischer. Wiesbaden: Springer VS.
- Mützel, Sophie, Philippe Saner und Markus Unternährer. 2018. »Schöne Daten! Konstruktion und Verarbeitung von digitalen Daten.« S. 111–32 in *Datengesellschaft. Einsichten in die Datafizierung des Sozialen*, hg. von D. Houben und B. Priel. Bielefeld: transcript.
- Muzio, Daniel, Damian Hodgson, James Faulconbridge, Jonathan Beaverstock und Sarah Hall. 2011. »Towards Corporate Professionalization: The Case of Project Management, Management Consultancy and Executive Search.« *Current Sociology* 59(4): 443–64.
- Nachtwey, Oliver und Timo Seidl. 2017. *Die Ethik der Solution und der Geist des digitalen Kapitalismus.* IfS Working Paper 11. Frankfurt a.M.: Institut für Sozialforschung.
- Nadai, Eva und Christoph Maeder. 2005. »Fuzzy Fields. Multi-Sited Ethnography in Sociological Research.« *Forum Qualitative Sozialforschung* 6(3): 1–11.
- NASEM. 2017. *Envisioning the Data Science Discipline.* Washington, D.C.: National Academies Press.
- NASEM. 2018. *Data Science for Undergraduates. Opportunities and Options.* Washington, D.C.: National Academies Press.
- NASEM. 2020. *Roundtable on Data Science Postsecondary Education. A Compilation of Meeting Highlights.* Washington, D.C.: National Academies Press.
- NASEM und The Royal Society. 2018. *The Frontiers of Machine Learning.* Washington, D.C.: National Academies Press.
- National Research Council. 2014. *Training Students to Extract Value from Big Data.* Washington, D.C.: National Academies Press.

- Navon, Daniel und Gil Eyal. 2014. »The Trading Zone of Autism Genetics: Examining the Intersection of Genomic and Psychiatric Classification.« *BioSocieties* 9(3): 329–52.
- Nax, Christin und Susanne Draheim. 2019. *Datenerzählungen & neue Expertisen – Ein Feldbericht aus Produktionsprozessen maschinenlernender Algorithmen*. Berlin.
- Nelson, Richard R. (Hg.). 1993. *National Innovation Systems. A Comparative Analysis*. Oxford: Oxford University Press.
- Nilsson, Nils J. 2009. *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press.
- Van Noorden, Richard. 2015. »Interdisciplinary Research by the Numbers.« *Nature* 525(7569): 306–7.
- Norén, Laura, Gina Helfrich und Steph Yeo. 2019. *Who's Building Your AI? Research Brief*. Newport Beach, CA: Obsidian Security. Online: <https://www.obsidiansecurity.com/blog/whos-building-your-ai-research-brief/>
- Nowotny, Helga, Peter Scott und Michael Gibbons. 2003. »Introduction: »Mode 2« Revisited: The New Production of Knowledge.« *Minerva* 41(3): 179–94.
- NSB. 2005. *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century*. Washington, D.C.: National Science Board.
- NSF. 2015. *Big Data Regional Innovation Hubs (BD Hubs). Accelerating the Big Data Innovation Ecosystem*. Alexandria, VA: National Science Foundation.
- NSF. 2019. *Harnessing the Data Revolution (HDR): Transdisciplinary Research in Principles of Data Science Phase I (NSF 19-550)*. Alexandria, VA: National Science Foundation.
- Ortiz-Repiso, Virginia, Jane Greenberg und Javier Calzada-Prado. 2018. »A Cross-Institutional Analysis of Data-Related Curricula in Information Science Programmes: A Focused Look at the ISchools.« *Journal of Information Science* 44(6): 768–84.
- Padgett, John F. und Walter W. Powell. 2012. *The Emergence of Organizations and Markets*. Princeton: Princeton University Press.
- Panofsky, Aaron L. 2011. »Field Analysis and Interdisciplinary Science: Scientific Capital Exchange in Behavior Genetics.« *Minerva* 49(3): 295–316.
- Papilloud, Christian und Alexander Hinneburg. 2018. *Qualitative Textanalyse mit Topic-Modellen*. Wiesbaden: Springer.
- Passoth, Jan-Hendrik und Werner Rammert. 2016. »Fragmentale Differenzierung und die Praxis der Innovation.« S. 39–67 in *Innovationsgesellschaft heute*, hg. von W. Rammert, A. Windeler, H. Knoblauch und M. Hutter. Wiesbaden: Springer.
- Patil, DJ. 2011. *Building Data Science Teams. The Skills, Tools, and Perspectives Behind Great Data Science Groups*. Sebastopol, CA: O'Reilly Media.
- Patil, DJ und Hilary Mason. 2015. *Data Driven. Creating a Data Culture*. Sebastopol, CA: O'Reilly Media.
- Paulitz, Tanja und Bianca Prietl. 2017. »Technikwissenschaftliche Business Masculinity als aufstrebender Ingenieurdiskurs.« S. 141–69 in *Macht in Wissenschaft und Gesellschaft*, hg. von J. Hamann, J. Maesse, V. Gengnagel und A. Hirschfeld. Wiesbaden: Springer.
- Pinar, William F., William M. Reynold, Patrick Slattery und Peter M. Taubman. 1995. *Understanding Curriculum. Introduction to the Study of Historical and Contemporary Curriculum Discourses*. New York: Peter Lang.
- Poor, Nathaniel. 2012. »When Firms Encourage Copying: Cultural Borrowing as Standard Practice in Game Spaces.« *International Journal of Communication* 6(1): 689–709.

- Porter, Theodore M. 1995. *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press.
- Powell, Walter W., Achim Oberg, Valeska P. Korff, Carrie Oelberger und Karina Kloos. 2017. »Institutional Analysis in a Digital Era: Mechanisms and Methods to Understand Emerging Fields.« S. 305–44 in *New Themes in Institutional Analysis: Topics and Issues from European Research*, hg. von G. Krücken, C. Mazza, R. E. Meyer und P. Walgenbach. Cheltenham: Edward Elgar.
- Powell, Walter W. und Kurt W. Sandholtz. 2012. »Amphibious Entrepreneurs and the Emergence of Organizational Forms.« *Strategic Entrepreneurship Journal* 6(2): 94–115.
- Prietl, Bianca und Daniel Houben. 2018. »Einführung. Soziologische Perspektiven auf die Datafizierung der Gesellschaft.« S. 7–32 in *Datengesellschaft. Einsichten in die Datafizierung des Sozialen*, hg. von D. Houben und B. Prietl. Bielefeld: transcript.
- Provost, Foster und Tom Fawcett. 2013. *Data Science for Business. What You Need to Know About Data Mining and Data-Analytic Thinking*. Sebastopol, CA: O'Reilly Media.
- PwC und BHEF. 2017. *Investing in America's Data Science and Analytics Talent*. New York/Washington D.C. PricewaterhouseCooper und Business and Higher Education Foundation.
- Rammert, Werner, Arnold Windeler, Hubert Knoblauch und Michael Hutter (Hg.). 2016. *Innovationsgesellschaft Heute: Perspektiven, Felder und Fälle*. Wiesbaden: Springer.
- Raunig, Gerald, Gene Ray und Ulf Wuggenig (Hg.). 2011. *Critique of Creativity: Precarity, Subjectivity and Resistance in the ›Creative Industries‹*. London: MayFlyBooks.
- Rawlings-Goss, Renata. 2018. *Keeping Data Science Broad: Negotiating the Digital and Data Divide Among Higher Education Institutions*. Atlanta: South Big Data Innovation Hub.
- Reckwitz, Andreas. 2012. *Die Erfindung der Kreativität. Zum Prozess gesellschaftlicher Ästhetisierung*. Berlin: Suhrkamp.
- Reed-Danahay, Deborah. 2019. *Bourdieu and Social Space: Mobilities, Trajectories, Emplacements*. New York/Oxford: Berghahn Books.
- Rehbein, Boike. 2003. »»Sozialer Raum« und Felder – Mit Bourdieu in Laos.« S. 77–95 in *Pierre Bourdieus Theorie des Sozialen: Probleme und Perspektiven*. Konstanz: UVK.
- Reichert, Ramón. 2018. »Big Data als Boundary Objects. Zur medialen Epistemologie von Daten.« S. 17–33 in *Bedeutende Daten*, hg. von T. Mämecke, J.-H. Passoth und J. Wehner. Wiesbaden: Springer.
- Reichmann, Werner. 2018. »The Interactional Foundations of Economic Forecasting.« S. 105–23 in *Uncertain Futures: Imaginaries, Narratives, and Calculation in the Economy*, hg. von J. Beckert und R. Bronk. Oxford: Oxford University Press.
- Ribes, David. 2019. »STS, Meet Data Science, Once Again.« *Science, Technology, & Human Values* 44(3): 514–39.
- Ribes, David, Andrew S. Hoffman, Steven C. Slota und Geoffrey C. Bowker. 2019. »The Logic of Domains.« *Social Studies of Science* 49(3): 281–309.
- Ridsdale, Chantel, James Rothwell, Michael Smit, Hossam Ali-Hassan, Michael Bliemel, Dean Irvine, Daniel Kelley, Stan Matwin und Brad Wuetherick. 2015. *Strategies and Best Practices for Data Literacy Education. Knowledge Synthesis Report*. Halifax: Dalhousie University.

- Rip, Arie und Jan-Peter Voss. 2013. »Umbrella Terms as Mediators in the Governance of Emerging Science and Technology.« *Science, Technology & Innovation Studies* 9(2): 39–59.
- Roberge, Jonathan und Robert Seyfert. 2016. »What Are Algorithmic Cultures?« S. 1–25 in *Algorithmic Cultures. Essays on meaning, performance and new technologies*, hg. von J. Roberge und R. Seyfert. Abingdon/New York: Routledge.
- Roberts, Margaret E., Brandon M. Stewart und Dustin Tingley. 2019. »stm: An R Package for Structural Topic Models.« *Journal of Statistical Software* 91(2): 1–40.
- Roberts, Pasha und Greta Roberts. 2013. »Research Brief : Four Functional Clusters of Analytics Professionals.« Cambridge, MA: Talent Analytics.
- Rose, Nikolas. 1992. »Engineering the Human Soul: Analyzing Psychological Expertise.« *Science in Context* 5(2): 351–69.
- Rostetter, Andri. 2020. »Der Wettbewerb unter den Fachhochschulen verschärft sich.« *Neue Zürcher Zeitung*, 23.09.2020: 13.
- Ruppert, Evelyn, John Law und Mike Savage. 2013. »Reassembling Social Science Methods: The Challenge of Digital Devices.« *Theory, Culture & Society* 30(4): 22–46.
- Sacchi, Stefan, Alexander Salvisberg und Marlis Buchmann. 2005. »Long-Term Dynamics of Skill Demand in Switzerland, 1950–2000.« S. 105–34 in *Contemporary Switzerland: Revisiting the Special Case*, hg. von H. Kriesi, P. Farago, M. Kohli und M. Zarin-Nejadan. London: Palgrave Macmillan.
- Safavi, Tara, Maryam Davoodi und Danai Koutra. 2018. »Career Transitions and Trajectories.« S. 675–84 in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining – KDD '18*. New York: ACM Press.
- Salganik, Matthew J. 2017. *Bit by Bit: Social Research in the Digital Age*. Princeton: Princeton University Press.
- Salvisberg, Alexander. 2008. »Sozialkompetenzen und arbeitsmarktliche Ungleichheitsstrukturen.« S. 2563–74 in *Die Natur der Gesellschaft: Verhandlungen des 33. Kongresses der Deutschen Gesellschaft für Soziologie in Kassel 2006*, hg. von K.-S. Rehberg. Frankfurt a.M.: Campus.
- Salvisberg, Alexander. 2010. *Soft Skills auf dem Arbeitsmarkt: Bedeutung und Wandel*. Zürich: Seismo.
- Saner, Philippe. 2019. »Envisioning Higher Education: How Imagining the Future Shapes the Implementation of a New Field in Higher Education.« *Swiss Journal of Sociology* 45(3): 359–81.
- Saner, Philippe und Sophie Mützel (2021a): *Collection of Data Science Job Advertisements 2017–2019 [Dataset]*. Department of Sociology, University of Lucerne. Distributed by FORS, Lausanne. Online: <https://doi.org/10.23662/FORS-DS-1254-1>.
- Saner, Philippe und Sophie Mützel (2021b): *Erhebung von Dokumenten zur Digitalisierung von Hochschulbildung und Forschung in der Schweiz (1998–2020) [Dataset]*. Soziologisches Seminar, Universität Luzern. Distributed by FORS, Lausanne. Online: <https://doi.org/10.23662/FORS-DS-1300-1>.
- Sassen, Saskia. 2013. »When the Center No Longer Holds: Cities as Frontier Zones.« *Cities* 34: 67–70.
- Sassen, Saskia. 2016. *Global Networks, Linked Cities*. London/New York: Routledge.
- Sauder, Michael und Wendy Nelson Espeland. 2009. »The Discipline of Rankings: Tight Coupling and Organizational Change.« *American Sociological Review* 74(1): 63–82.

- Savage, Mike. 2013. »The ›Social Life of Methods‹: A Critical Introduction.« *Theory, Culture & Society* 30(4): 3–21.
- Saxenian, AnnaLee. 1996. *Regional Advantage. Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA/London: Harvard University Press.
- SBFI. 2016. *Bibliometrische Untersuchung zur Forschung in der Schweiz 1981–2013. Bericht des Staatssekretariats für Bildung, Forschung und Innovation SBFI*. Bern: Staatssekretariat für Bildung, Forschung und Innovation.
- Schäfer, Mirko Tobias und Karin van Es (Hg.). 2017. *The Datafied Society. Studying Culture through Data*. Amsterdam: Amsterdam University Press.
- Scheel, Stephan, Francisca Grommé, Evelyn Ruppert, Funda Ustek-Spilda, Baki Cakici und Ville Takala. 2020. »Doing a Transversal Method: Developing an Ethics of Care in a Collaborative Research Project.« *Global Networks* 20(3): 522–43.
- Schmidt-Wellenburg, Christian. 2013. *Die Regierung des Unternehmens: Managementberatung im neoliberalen Kapitalismus*. Konstanz: UVK.
- Schmidt-Wellenburg, Christian und Stefan Bernhard (Hg.). 2020a. *Charting Transnational Fields*. London/New York: Routledge.
- Schmidt-Wellenburg, Christian und Stefan Bernhard. 2020b. »How to Chart Transnational Fields.« S. 1–33 in *Charting Transnational Fields*, hg. von C. Schmidt-Wellenburg und S. Bernhard. London/New York: Routledge.
- Schmitz, Andreas, Daniel Witte und Vincent Gengnagel. 2017. »Pluralizing Field Analysis: Toward a Relational Understanding of the Field of Power.« *Social Science Information* 56(1): 49–73.
- Schreiber, Norbert. 1995. »Wer braucht Sozialwissenschaftler und Soziologen? Eine Inhaltsanalyse von Stellenanzeigen.« *Sozialwissenschaften und Berufspraxis* 18(4): 313–27.
- Schuff, David. 2018. »Data Science for All: A University-Wide Course in Data Literacy.« S. 229 in *Analytics and Data Science, Annals of Information Systems*, hg. von A. V. Deokar, A. Gupta, L. S. Iyer und M. C. Jones. Cham: Springer.
- Schuler, Douglas und Aki Namioka. 1993. *Participatory Design: Principles and Practices*. CRC Press.
- Schultheis, Franz. 2005. »Disziplinierung des Designs.« S. 65–84 in *Forschungslandschaften im Umfeld des Design*, hg. vom Swiss Design Network. Zürich.
- Schulz-Schaeffer, Ingo. 2019. »Innovation als soziale Konstruktion von Technik und Techniknutzung.« S. 1–18 in *Handbuch Innovationsforschung*, hg. von B. Blättel-Mink, I. Schulz-Schaeffer und A. Windeler. Wiesbaden: Springer.
- Schumann, Conny, Patrick Zschech und Andreas Hilbert. 2016. »Das aufstrebende Berufsbild des Data Scientist.« *HMD Praxis der Wirtschaftsinformatik* 53(4): 453–66.
- Schutt, Rachel und Cathy O’Neil. 2013. *Doing Data Science: Straight Talk from the Frontline*. Sebastopol, CA: O’Reilly Media.
- Schwab, Klaus, Alan Marcus, Justin R. Oyola, William Hoffman und Michele Luzi. 2011. *Personal Data: The Emergence of a New Asset Class*. Cologne: World Economic Forum.
- Scott, Richard W. 1994. »Conceptualizing Organizational Fields: Linking Organizations and Societal Systems.« S. 203–221 in *Systems Rationality and Partial Interests*, hg. von H.-U. Derlien, U. Gerhardt und F. W. Scharpf. Baden-Baden: Nomos.
- SDSC. 2017. *Data Science Projects. 1st Call for Proposals*. Fall 2017. Lausanne/Zürich: Swiss Data Science Center.

- Seefranz, Catrin und Philippe Saner. 2012. *Making Differences: Schweizer Kunsthochschulen. Explorative Vorstudie*. Zürich: Institute for Art Education.
- Selbst, Andrew D., danah boyd, Sorelle A. Friedler, Suresh Venkatasubramanian und Janet Vertesi. 2019. »Fairness and Abstraction in Sociotechnical Systems.« S. 59–68 in *Proceedings of the Conference on Fairness, Accountability, and Transparency – FAT** ›19. Vol. 1. Atlanta: ACM Press.
- Shinn, Terry und Bernward Joerges. 2002. »The Transverse Science and Technology Culture: Dynamics and Roles of Research-Technology.« *Social Science Information* 41(2): 207–51.
- Shinn, Terry und Bernward Joerges. 2004. »Paradox oder Potenzial. Zur Dynamik heterogener Kooperation.« S. 77–101 in *Kooperation im Niemandsland. Neue Perspektiven auf Zusammenarbeit in Wissenschaft und Technik*, hg. von J. Strübing, I. Schulz-Schaeffer, M. Meister und J. Gläser. Opladen: Leske + Budrich.
- Simmel, Georg. 1995[1902]. »Der Bildrahmen. Ein ästhetischer Versuch.« in *Aufsätze und Abhandlungen 1901–1908*, hg. von R. Kramme, A. Rammstedt und O. Rammstedt. Frankfurt a.M.: Suhrkamp.
- Small, Mario L. 1999. »Departmental Conditions and the Emergence of New Disciplines: Two Cases in the Legitimation of African-American Studies.« *Theory and Society* 28(5): 659–707.
- Song, Il-Yeol und Yongjun Zhu. 2017. »Big Data and Data Science: Opportunities and Challenges of ISchools.« *Journal of Data and Information Science* 2(3): 1–18.
- Stadelmann, Thilo, Stockinger, Kurt, Braschler, Martin, Cieliebak, Mark, Baudinot, Gerold, Dürr, Oliver und Andreas Ruckstuhl. 2013. Applied data science in Europe: challenges for academia in keeping up with a highly demanded topic. In *9th European Computer Science Summit, Amsterdam, Niederlande, 8.-9. Oktober 2013*.
- Stadelmann, Urs. 2017. »Google – Ein Motor und Konkurrent.« *Neue Zürcher Zeitung*, 31.10.2017: 29.
- Staley, Richard. 2002. »Interstitial Instruments.« *Social Studies of Science* 32(3): 469–76.
- Stampnitzky, Lisa. 2013. *Disciplining Terror*. Cambridge: Cambridge University Press.
- Star, Susan Leigh. 2017. »Was ist kein Grenzobjekt? Reflexionen über den Ursprung eines Konzepts (2010).« S. 213–28 in *Grenzobjekte und Medienforschung*, hg. von S. Giessmann und N. Taha. Bielefeld: transcript.
- Star, Susan Leigh und James R. Griesemer. 1989. »Institutional Ecology, ›Translations‹ and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39.« *Social Studies of Science* 19(3): 387–420.
- Stark, David. 2009. *The Sense of Dissonance: Accounts of Worth in Economic Life*. Princeton: Princeton University Press.
- Steyvers, Mark und Tom Griffiths. 2007. »Probabilistic Topic Models.« S. 427–448 in *Handbook of Latent Semantic Analysis*, hg. von T. K. Landauer, D. S. McNamara, S. Dennis und W. Kintsch. New York: Psychology Press.
- Stichweh, Rudolf. 1993. »Wissenschaftliche Disziplinen: Bedingungen ihrer Stabilität im 19. und 20. Jahrhundert.« S. 235–50 in *Sozialer Raum und akademische Kulturen: Studien zur europäischen Hochschul- und Wissenschaftsgeschichte im 19. und 20. Jahrhundert*, hg. von J. Schriewer, E. Keiner und C. Charle. Frankfurt a.M.: Peter Lang.
- Stichweh, Rudolf. 2005. »Neue Steuerungsformen der Universität und die akademische Selbstverwaltung. Die Universität als Organisation.« S. 123–34 in *Die Idee der Universität heute*, hg. von U. Sieg und D. Korsch. München: K. G. Saur.

- Stichweh, Rudolf. 2013. *Wissenschaft, Universität, Professionen. Soziologische Analysen*. Frankfurt a.M.: Suhrkamp.
- Stockinger, Kurt, Thilo Stadelmann und Andreas Ruckstuhl. 2016. »Data Scientist als Beruf.« S. 59–81 in *Big Data. Grundlagen, Systeme und Nutzungspotenziale*, hg. von D. Fasel und A. Meier. Wiesbaden: Springer.
- Strachnyi, Kate. 2017. *Journey to Data Scientist: Interviews with More than Twenty Amazing Data Scientists*. CreateSpace Independent Publishing Platform.
- Struck, Olaf. 2018. »Betrieb und Arbeitsmarkt.« S. 193–223 in *Arbeitsmarktsoziologie. Probleme, Theorien, empirische Befunde*, hg. von M. Abraham und T. Hinz. Wiesbaden: Springer.
- Süssenguth, Florian (Hg.). 2015. *Die Gesellschaft der Daten. Über die digitale Transformation der sozialen Ordnung*. Bielefeld: transcript.
- Swan, Alma und Sheridan Brown. 2008. *The Skills, Role and Career Structure of Data Scientists and Curators: An Assessment of Current Practice and Future Needs*. A report to the Joint Information Systems Committee (JISC). Truro: Key Perspectives Ltd.
- Swartz, David. 1997. *Culture & Power: The Sociology of Pierre Bourdieu*. Chicago: University of Chicago Press.
- Tang, Rong und Watinee Sae-Lim. 2016. »Data Science Programs in U.S. Higher Education: An Exploratory Content Analysis of Program Description, Curriculum Structure, and Course Focus.« *Education for Information* 32(3): 269–90.
- Thornham, Helen und Edgar Gómez Cruz. 2016. »Hackathons, Data and Discourse: Convolutions of the Data (Logical).« *Big Data & Society* 3(2): 205395171667967.
- Trajanoski, Zlatko (Hg.). 2012. *Computational Medicine. Tools and Challenges*. Wien: Springer.
- Traue, Boris. 2010. »Kompetente Subjekte: Kompetenz als Bildungs- und Regierungsdispositiv im Postfordismus.« S. 49–67 in *Soziologie der Kompetenz*, hg. von T. Kurtz und M. Pfadenhauer. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Tschan, Dieter (2019): KI in der Verwaltung: Chancen und Herausforderungen. In: *Netzwoche*, 16.10.2019. Online: <https://www.netzwoche.ch/news/2019-10-16/ki-inder-verwaltung-chancen-und-herausforderungen>
- Tukey, John W. 1962. »The Future of Data Analysis.« *The Annals of Mathematical Statistics* 33(1): 1–67.
- Ueno, Maomi. 2017. »As the Oldest Journal of Data Science.« *Behaviormetrika* 44(1): 1–2.
- Unternährer, Markus. 2021. *Momente der Datafizierung – Zur Produktionsweise von Personendaten in der Datenökonomie*. Unveröffentlichte Dissertationsschrift. Luzern: Universität Luzern.
- Varian, Hal. 2009. »Hal Varian on How the Web Challenges Managers.« *McKinsey Quarterly* 1(2.2). Online: <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/hal-varian-on-how-the-web-challenges-managers>
- De Veaux, Richard D., Mahesh Agarwal, Maia Averett, Benjamin S. Baumer, Andrew Bray, Thomas C. Bressoud, Lance Bryant, Lei Z. Cheng, Amanda Francis, Robert Gould, Albert Y. Kim, Matt Kretchmar, Qin Lu, Ann Moskol, Deborah Nolan, Roberto Pelayo, Sean Raleigh, Ricky J. Sethi, Mutiara Sondjaja, Neelesh Tiruvilumala, Paul X. Uhlig, Talitha M. Washington, Curtis L. Wesley, David White und Ping Ye. 2017. »Curriculum Guidelines for Undergraduate Programs in Data Science.« *Annual Review of Statistics and Its Application* 4(1): 15–30.

- Verma, Inder M. 2014. »Editorial Expression of Concern and Correction.« *Proceedings of the National Academy of Sciences* 111(29): 10779.
- Vilsmaier, Ulli. 2018. »Grenzarbeit in integrativer und grenzüberschreitender Forschung.« S. 113–34 in *Grenzen. Theoretische, konzeptionelle und praxisbezogene Fragestellungen zu Grenzen und deren Überschreitungen*, hg. von M. Heintel, R. Musil und N. Weixlbaumer. Wiesbaden: Springer.
- Voulgaris, Zacharias. 2014. *Data Scientist: The Definitive Guide to Becoming a Data Scientist*. Basking Ridge, NJ: Technics Publications.
- Wacquant, Loïc. 2004. »Lire ›Le Capital‹ de Pierre Bourdieu.« S. 211–30 in *Pierre Bourdieu, sociologue*, hg. von L. Pinto, G. Sapiro und P. Champagne. Paris: Fayard.
- Wajcman, Judy. 2019. »How Silicon Valley Sets Time.« *New Media & Society* 21(6): 1272–89.
- Walker, Michael A. 2015. »The Professionalisation of Data Science.« *International Journal of Data Science* 1(1): 7–16.
- Wallach, Hanna. 2018. »Computational Social Science ≠ Computer Science + Social Data.« *Communications of the ACM* 61(3): 42–44.
- Watson, Sara M. 2014. »Data Science: What the Facebook Controversy Is Really About.« *The Atlantic*, 01.07.2014. Online: <https://www.theatlantic.com/technology/archive/2014/07/data-science-what-the-facebook-controversy-is-really-about/373770/>
- Weber, Max. 1993. *Politik als Beruf*. Berlin: Duncker & Humblot.
- Weick, Karl E. 1976. »Educational Organizations as Loosely Coupled Systems.« *Administrative Science Quarterly* 21(1): 1–19.
- Weingart, Peter. 2001. *Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft*. Weilerswist: Velbrück Wissenschaftsverlag.
- Weischer, Christoph. 2015. »Daten, prozessproduzierte.« S. 73–74 in *Methoden-Lexikon für die Sozialwissenschaften*, hg. von R. Diaz-Bone und C. Weischer. Wiesbaden: Springer VS.
- Wickham, Hadley. 2014. »Tidy Data.« *Journal of Statistical Software* 59(10): 1–23.
- Williams, Judy. 2013. »Boundary Crossing and Working in the Third Space: Implications for a Teacher Educator's Identity and Practice.« *Studying Teacher Education* 9(2): 118–29.
- Wilson, Courtney. 2017. »The Ugly Truth about Data Science.« In *CloudFactory*, 16.01.2017. Online: <https://blog.cloudfactory.com/ugly-truth-about-data-science>
- Wimmer, Andreas. 2013. *Ethnic Boundary Making: Institutions, Power, Networks*. Oxford: Oxford University Press.
- Wing, Jeannette M. 2006. »Computational Thinking.« *Communications of the ACM* 49(3): 33–35.
- Wing, Jeannette, Vandana P. Janeja, Tyler Kloefkorn und Lucy C. Erickson. 2018. *Data Science Leadership Summit Summary Report*. New York: Columbia University.
- Wisnioski, Matthew H. 2009. »›Liberal Education Has Failed‹: Reading Like an Engineer in 1960s America.« *Technology and Culture* 50(4): 753–82.
- Witte, Daniel und Andreas Schmitz. 2017. »Der Nationalstaat und das globale Feld der Macht, oder: Wie sich die Feldtheorie von ihrem methodologischen Nationalismus befreien lässt.« *Zeitschrift für Theoretische Soziologie* (2): 156–89.

- Witte, Daniel und Andreas Schmitz. 2019. »Netzwerke als transversale Felder.« S. 25–61 in *Netzwerke in gesellschaftlichen Feldern*, hg. von J. Fuhse und K. Krenn. Wiesbaden: Springer.
- Wooten, Melissa und Andrew J. Hoffman. 2017. »Organizational Fields: Past, Present and Future.« S. 55–74 in *The SAGE Handbook of Organizational Institutionalism*, hg. von R. Greenwood, C. Oliver, T. B. Lawrence und R. E. Meyer. London/New York: Sage.
- Wowczko, Izabela A. 2015. »Skills and Vacancy Analysis with Data Mining Techniques.« *Informatics* 2(4): 31–49.
- Wu, C. F. Jeff. 1997. »Statistics = Data Science?« *Identity of Statistics in Science Examined*. Inaugural Lecture for the appointment to the H. C. Carver Professorship at the University of Michigan. Ann Arbor: University of Michigan.
- Wuggenig, Ulf. 2016. »Kreativitätsbegriffe. Vorwort zur Neuauflage.« S. 11–69 in *Kritik der Kreativität*, hg. von G. Raunig und U. Wuggenig. Wien: transversal books.
- Young, Michael F. D. 1998. *The Curriculum of the Future. From the »New Sociology of Education« to a Critical Theory of Learning*. London: Falmer Press.
- Yu, Bin. 2014. »Let Us Own Data Science.« *ASC-IMS Meeting* 48.
- Zapp, Mike und Julia C. Lerch. 2020. »Imagining the World: Conceptions and Determinants of Internationalization in Higher Education Curricula Worldwide.« *Sociology of Education* 93(4): 372–92.
- Zapp, Mike und Francisco O. Ramirez. 2019. »Beyond Internationalisation and Isomorphism – the Construction of a Global Higher Education Regime.« *Comparative Education* 55(4): 473–93.
- Zhang, Jilong, Anna Fu, Hao Wang und Shenqing Yin. 2017. »The Development of Data Science Education in China from the LIS Perspective.« *International Journal of Librarianship* 2(2): 3.
- ZHAW. 2016. *Datalab – the ZHAW Data Science Laboratory*. Winterthur: Zürcher Fachhochschule, School of Engineering.
- Zomer, Arend und Paul Benneworth. 2011. »The Rise of the University's Third Mission.« S. 81–101 in *Reform of Higher Education in Europe*, hg. von J. Enders, H. F. de Boer und D. F. Westerheijden. Rotterdam: Sense Publishers.
- Zook, Matthew, Solon Barocas, danah boyd, Kate Crawford, Emily Keller, Seeta Peña Gangadharan, Alyssa Goodman, Rachelle Hollander, Barbara A. Koenig, Jacob Metcalf, Arvind Narayanan, Alondra Nelson und Frank Pasquale. 2017. »Ten Simple Rules for Responsible Big Data Research«. *PLOS Computational Biology* 13(3): e1005399.

Anhang

Interview Leitfaden

Zu Ihrer Biografie

Können Sie kurz Ihren akademischen Werdegang beschreiben?
Welches sind Ihre Forschungsschwerpunkte? Inwiefern hängen diese mit Data Science zusammen?

Der [Studiengang in Data Science] an der [Universität/Fachhochschule]

Zum Aufbau und Curriculum

Was sind die Schwerpunkte des [Studiengangs in Data Science]? Was ist spezifisch im Curriculum des [Studiengangs in Data Science]?
Wie ist der Studiengang in das [Departement/Fakultät/Universität] »eingebettet«? Ist er »interdisziplinärer« als andere?
Wie wird das Curriculum geplant/zusammengestellt? Welche Personen, welche Departemente sind in die Ausarbeitung involviert?
Gab es Studiengänge/Programme an Universitäten, die als »Vorbilder« genutzt wurden oder von denen man sich bewusst »abgrenzen« wollte?

Zu den Kursinhalten, Kursbeschreibungen etc.

Gibt es Kurse über Ethik, Datenschutz, Privatsphäre o.ä.? Inwiefern sind Inhalte über »verantwortungsvollen Umgang« mit Daten Teil des Curriculums?
In Anlehnung an das Venn Diagramm von Conway: Wie wichtig ist *domain knowledge* im Studium? Worin besteht sie im [Studiengang in Data Science]?
Wie wichtig sind Business Skills, Kommunikation oder Visualisierung?

Zu den Studierenden

Wie viele Studienplätze gibt es? Wie viele Studierende gibt es?
Was ist der Anteil an neuen Studierenden in-house und extern?

Wer sind die Studierenden, woher kommen sie? (Disziplinen, Backgrounds, ggf. Herkunftsländer etc.)

Verhältnis zu anderen Hochschulen

Was unterscheidet Ihrer Meinung nach den [Studiengang in Data Science] von Studiengängen an anderen Universitäten bzw. Hochschulen?

Wie nehmen Sie die kürzlich etablierten Studiengänge in Data Science an der EPFL und der ETHZ wahr?

Wie nehmen Sie das kürzlich etablierte Swiss Data Science Center wahr?

Gibt es Kooperationen mit Universitäten oder Fachhochschulen im Bereich der Lehre oder der Forschung?

Zum Arbeitsmarkt

Wie wird der [Studiengang], wie werden die Absolvent*innen vom Arbeitsmarkt bewertet? Wo landen die Leute nach dem Studium?

Ist der [Studiengang] in erster Linie ein Ausbildungsstudiengang, um Bedürfnisse der Unternehmen bzw. des Arbeitsmarktes abzudecken? Wie wichtig ist die wissenschaftliche Nachwuchsförderung?

Was sind Ihrer Meinung nach die notwendigen Kompetenzen von Data Scientists, heute und in der Zukunft?

Wie wichtig sind Kurse in Zusammenarbeit mit externen Partnern (Unternehmen, Forschungszentren etc.) für die Studierenden? Finden diese regelmässig statt?

Kommen Unternehmen auf Sie zu mit Projektvorschlägen?

Gibt es personelle Verflechtungen mit Unternehmen (z. B. Lehraufträge)?

Definitionen von Data Science

Jetzt haben wir viel über Data Science als Studiengang gesprochen.

Gemäss Ihrem Verständnis, was bedeutet Data Science für Sie persönlich? Wie würden Sie Data Science charakterisieren?

Ist es ein hilfreicher Begriff, um dieses interdisziplinäre Feld zu fassen? Gibt es Alternativen, wenn ja welche?

Sehen Sie Schwierigkeiten mit dem Begriff?

Für die weitere Forschung

Wen würden Sie empfehlen für ein weiteres Gespräch zu Data Science (in der Schweiz/überhaupt)?

Gibt es wichtige Events, Mailinglisten etc. zu verfolgen?

Tabelle 8: Übersicht über Dimensionen, Kodes und Zitate

Dimension	Anzahl Kodes	Anzahl Zitate	Relativer Anteil
Bildungsdimension	14	261	10.4 %
Daten/Datenwissenschaften	18	306	12.2 %
Internationale Dimension	6	171	6.8 %
Kompetenzen	13	156	6.2 %
Ökonomische Dimension	14	291	11.6 %
Organisationale Dimension	12	303	12.1 %
Politische Dimension	15	400	15.9 %
Technische Dimension	15	277	11 %
Wissenschaftliche Dimension	19	344	13.7 %
Total	75*	1144*	100 %

* Hinweis: Aufgrund der Mehrdimensionalität der Kodes übersteigen die Anzahl Kodes sowie die Anzahl Zitate das Gesamttotal

Tabelle 9: Die zehn häufigsten Berufsgruppen im Sample¹

Berufsgruppe	Häufigkeit	Anteil im Sample	Anteil englisch	Anteil deutsch	Anteil französisch	Beispiele aus dem Sample
Scientist	1010	23.3 %	77.9 %	20.8 %	1.3 %	Data Scientist, Research Scientist in Machine Learning
Analyst	763	17.6 %	60.5 %	33.8 %	5.5 %	Data Analyst, Business Analyst
Engineer	595	13.7 %	69.8 %	29.8 %	0.4 %	Data Engineer, Big Data Engineer
Manager	340	7.8 %	69.4 %	28.5 %	2.1 %	Data Manager, IT Project Manager
Specialist	277	6.4 %	54.9 %	44 %	1.1 %	Data Science Specialist, Quantitative Risk Modelling Specialist
Intern/ Praktikant	263	6.1 %	76 %	21.3 %	2.7 %	Data Science Internship, Praktikum im Fachbereich Datenanalyse
Consultant	223	5.1 %	39 %	58.3 %	2.7 %	Data Science Consultant, Consultant Data & Analytics
Developer	210	4.8 %	62.8 %	32.4 %	3.3 %	(Senior) Developer Artificial Intelligence – Data Science, Python Developer
Architect	142	3.2 %	68.3 %	28.2 %	3.5 %	Big Data & Analytics Architect, Data Science Software Architect
Expert	100	2.3 %	86 %	12 %	2 %	Artificial Intelligence Expert, Data Modelling and Visualization Expert
Diverse	436	10 %				Head of Data Science, PhD Student in Computing & Data science, Data Science & AI Solution Director
Total	4341	100 %	65.9 %	31.2 %	2.9 %	

¹ Die Tabelle enthält Mehrfachnennungen.

Abbildung 12: Häufigste Begriffe der Stellentitel

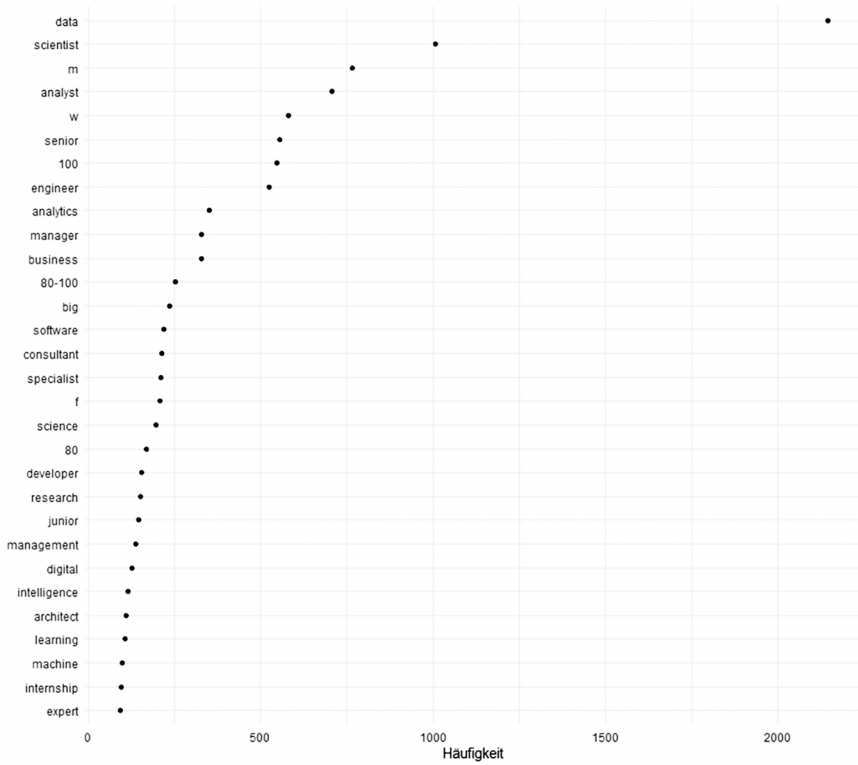
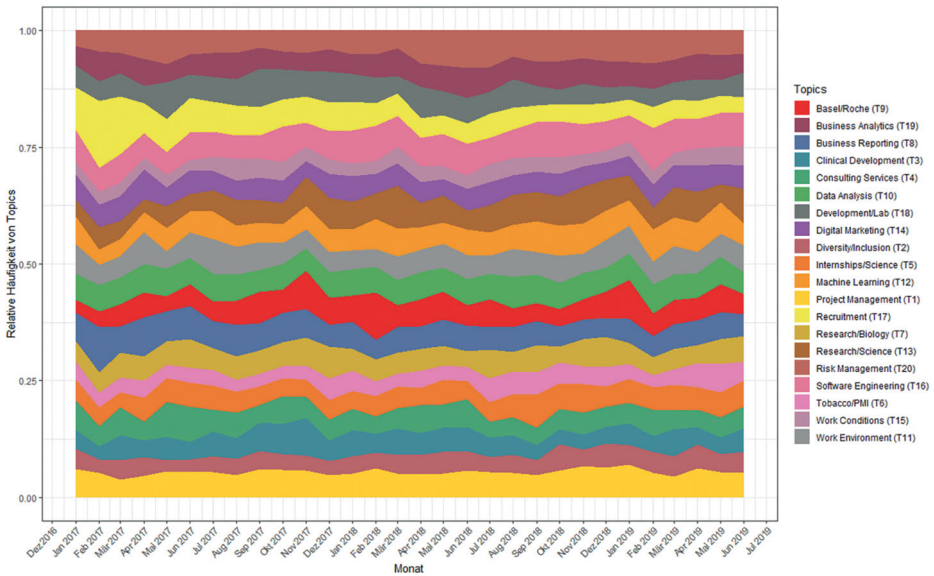


Tabelle 10: Sprachen der drei häufigsten Berufsbezeichnungen

Berufsbezeichnung	Häufigkeiten	Anteil im Sample	Anteil englisch	Anteil deutsch	Anteil französisch
Data Scientist	587	13.5 %	63.2 %	34.6 %	2.2 %
Data Analyst	497	11.4 %	59 %	34.4 %	6.4 %
Data Engineer	128	2.9 %	65.6 %	33.6 %	0.8 %
Total	4341	100 %	67.7 %	29.5 %	2.7 %

Abbildung 13: Geschätzte relative Häufigkeit von Stellenanzeigen nach Topics



Beispielinserat 1: BMW Financial Services

Data Scientist/Analyst (w/m) BMW Financial Services

BMW Group Switzerland — Dielsdorf
Jetzt bewerben//

- *//Dielsdorf
- *//Klein- und Mittelunternehmen
- *//Festanstellung, 100 %
- *//Fachverantwortung

Data Scientist/Analyst (w/m) BMW Financial Services

[1] Mit Leidenschaft fängt bei BMW alles an. Sie macht aus einem Beruf eine Berufung. Sie treibt uns an, Mobilität immer wieder neu zu erfinden und innovative Ideen auf die Straße zu bringen. Nur wenn Expertenwissen, hochprofessionelle Abläufe und Freude bei der Arbeit zusammen kommen, lässt sich gemeinsam Zukunft gestalten.

[2] Die BMW Group Financial Services stellt den BMW, MINI und BMW Motorrad Händlern der Schweiz Finanzdienstleistungen zur Verfügung und ist damit ein integrierter Bestandteil der Vertriebsaktivitäten der BMW Group in der Schweiz. Neben den Finanzprodukten für Endkunden gehören auch Kreditprodukte für die Handels-

betriebe (Lagerwagenfinanzierung, Betriebskredite, Hypothekarprodukte) zum Produktportfolio.

[3] Zur Verstärkung des Vertriebs- und Marketingteams suchen wir per sofort oder nach Vereinbarung eine engagierte Persönlichkeit als

Data Scientist/Analyst (w/m) BMW Financial Services

[4] In dieser abwechslungsreichen und bedeutungsvollen Position verantworten Sie eine KPI basierte Vertriebssteuerung, indem Sie permanent die gesamte Produkt- und Dienstleistungspalette sowie die Vertriebs- und Marketingperformance überwachen, analysieren und bewerten. Sie erstellen komplexe Vertriebs-, Markt- und Wettbewerbsanalysen und erarbeiten anhand dieser Grundlagen konkrete Handlungsanweisungen für die Business Lines und unsere Händlerorganisation. Zudem verantworten Sie die Pricingstrategie der BMW Financial Services. Ebenfalls gehört die Organisation und Durchführung des monatlichen Produkt- und Zinskreises zu Ihrem breiten Aufgabengebiet.

[5] Sie bereiten die notwendigen Datengrundlagen auf und entdecken systematische Muster, Kundensegmente/-clusters, Trends, Kundenwerte und Handlungsbedarf (next best action). Sie identifizieren, welche Personen zu welchem Zeitpunkt für welche Themen/Angebote affin sind und über welche Kanäle kontaktiert werden können. Dazu berücksichtigen Sie auch neuere Ansätze des Information Retrievals und Machine Learnings.

[6] Ihr Profil:

- * Einige Jahre Erfahrung in einer vergleichbaren Position
- * Ausgeprägte analytische und konzeptionelle Fertigkeiten sowie Freude beim Erstellen von Analysen und Strategien
- * Naturwissenschaftliche oder betriebswirtschaftliche Ausbildung (HF, FH, Uni) mit vertieften Kenntnissen in Statistik, Analysen und Controlling
- * Resultatorientierte, engagierte, selbständige aber auch vermittelnde Persönlichkeit mit einer gesunden Balance zwischen Hartnäckigkeit und Pragmatismus
- * Mehrjährige Erfahrung mit statistischer Software, R, Python, SAS oder SPSS
- * Fundierte Kenntnisse der MS-Office Produktpalette, insbesondere Excel
- * Stilsicheres Deutsch, gute Englischkenntnisse, Französisch von Vorteil

[7] Wenn Sie mit uns Ihre Leidenschaft für Innovation und Mobilität teilen möchten, freuen wir uns auf Ihre Online-Bewerbung.

Von Kandidatenvorschlägen durch Personalberater bitten wir abzusehen.

Originalinserat anzeigen

<https://bmwch.rexx-recruitment.com/xml/public/fetchxml.php?agid=19&id=185>²

² Die Internetquelle ist nicht mehr verfügbar.

Tabelle 11: Untersuchte Dokumente

Abkürzung	Autor*innen	Titel	Ort
BAKOM 2018	Bundesamt für Kommunikation	Eckwerte für eine Datenpolitik der Schweiz	Biel
BFS 2017	Bundesamt für Statistik	Dateninnovationsstrategie	Neuchâtel
Bundesrat 2016	Bundesrat	Botschaft zur Förderung von Bildung, Forschung und Innovation in den Jahren 2017–2020 vom 24. Februar 2016	Bern
Bundesrat 2018	Bundesrat	Strategie für offene Verwaltungsdaten in der Schweiz 2019–2023. (Open-Government-Data-Strategie, OGD-Strategie.)	Bern
Bundesrat 2020	Bundesrat	Der Bundesrat schafft ein Kompetenzzentrum für Datenwissenschaft. Medienmitteilung vom 13.05.2020.	Bern
Christen et al. 2020	Markus Christen, Clemens Mader, Johann Čas, Tarik Abou-Chadi, Abraham Bernstein, Nadja Braun Binder, Daniele Dell'Aglio, Luca Fábíán, Damian George, Anita Gohdes, Lorenz Hilty, Markus Kneer, Jaro Krieger-Lamina, Hauke Licht, Anne Scherer, Claudia Som, Pascal Sutter, Florent Thouvenin	Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz. TA-SWISS Publikationsreihe 72/2020	Zürich
Economiesuisse & W.I.R.E. 2017	Economiesuisse & W.I.R.E.	Zukunft digitale Schweiz. Wirtschaft und Gesellschaft weiterdenken.	Zürich
EDK 2018	Konferenz der Kantonalen Erziehungsdirektoren	Digitalisierungsstrategie. Strategie der EDK vom 21. Juni 2018 für den Umgang mit Wandel durch Digitalisierung im Bildungswesen.	Bern
ETH-Rat 2014	Rat der Eidgenössisch-Technischen Hochschulen	Strategische Planung 2017–2020 des ETH-Rats für den ETH-Bereich.	Zürich
ETH-Rat 2016a	Rat der Eidgenössisch-Technischen Hochschulen	Der ETH-Bereich lanciert die »Initiative for Data Science in Switzerland«	Zürich
ETH-Rat 2016b	Rat der Eidgenössisch-Technischen Hochschulen	Internationale Wettbewerbsfähigkeit stärken und in zukunftsweisende Forschungsbereiche investieren	Zürich
Expertengruppe 2018	Expertengruppe zur Zukunft der Datenbearbeitung und Datensicherheit	Bericht der Expertengruppe zur Zukunft der Datenbearbeitung und Datensicherheit	Bern

GDS 2016	Geschäftsstelle Digitale Schweiz, Bundesamt für Kommunikation	Aktionsplan Digitale Schweiz. Massnahmen der Bundesverwaltung. April 2016. Ziel: Bundesamt für Kommunikation.	Biel
GDS 2017	Geschäftsstelle Digitale Schweiz, Bundesamt für Kommunikation	Aktionsplan Digitale Schweiz. Massnahmen der Bundesverwaltung. Stand: November 2017	Biel
GDS 2018	Geschäftsstelle Digitale Schweiz, Bundesamt für Kommunikation	Aktionsplan Digitale Schweiz. Stand: 5. September 2018.	Biel
ICTswitzerland & Economiesuisse 2011	ICTswitzerland & Economiesuisse	ICTswitzerland und Economiesuisse: Digitale Agenda 2020. Auf dem Weg an die Weltspitze.	Zürich
IDAG KI 2019	Interdepartementale Arbeitsgruppe »Künstliche Intelligenz« (SBFI)	Herausforderungen der künstlichen Intelligenz. Bericht der interdepartementalen Arbeitsgruppe »Künstliche Intelligenz« an den Bundesrat	Bern
ISA IG 2016	Interdepartementaler Steuerungsausschuss Informationsgesellschaft, Bundesamt für Kommunikation	Bericht 2012–2015 zur Umsetzung der Strategie des Bundesrates für eine Informationsgesellschaft in der Schweiz vom März 2012. April 2016.	Biel
Jarchow & Estermann 2015	Jarchow, Thomas & Beat Estermann	Big Data: Chancen, Risiken und Handlungsbedarf des Bundes. Ergebnisse einer Studie im Auftrag des Bundesamts für Kommunikation	Bern
SATW 2019	Schweizerische Akademie der Technischen Wissenschaften	Künstliche Intelligenz in Wissenschaft und Forschung	Zürich
SBFI 2017	Staatssekretariat für Bildung, Forschung und Innovation	Herausforderungen der Digitalisierung für Bildung und Forschung in der Schweiz	Bern
SCNAT 2018	Schweizerische Akademie der Naturwissenschaften	Synthese des Workshops vom 8. Februar 2018. Bedeutung der Informatik heute – Visionen für morgen.	Bern
SDS 2016	Bundesamt für Kommunikation	Strategie Digitale Schweiz. April 2016	Biel
SDS 2018	Bundesamt für Kommunikation	Strategie »Digitale Schweiz«. Vom Bundesrat verabschiedet am 5. September 2018	Biel
SECO 2017a	Staatssekretariat für Wirtschaft	Auswirkungen der Digitalisierung auf Beschäftigung und Arbeitsbedingungen – Chancen und Risiken.	Bern
SECO 2017b	Staatssekretariat für Wirtschaft	Bericht über die zentralen Rahmenbedingungen für die digitale Wirtschaft. Bericht des Bundesrats vom 11. Januar 2017.	Bern

SEFRI 2019	Secrétariat d'Etat à la formation, à la recherche et à l'innovation	L'intelligence artificielle dans la formation	Bern
SIG 1998	Bundesrat	Strategie des Bundesrates für eine Informationsgesellschaft in der Schweiz vom 18. Februar 1998	Bern
SIG 2006	Bundesrat	Strategie des Bundesrats für eine Informationsgesellschaft in der Schweiz, Januar 2006	Bern
SIG 2012	Bundesrat	Strategie des Bundesrates für eine Informationsgesellschaft in der Schweiz März 2012	Bern
SNF 2015	Swiss National Science Foundation	NRP 75 Big Data. National Research Programme. Call for Proposals	Bern
SNF 2018	Swiss National Science Foundation	NRP 77 Digital Transformation. National Research Programme. Call Document	Bern
Staatslabor 2018	Staatslabor	Kurzbericht zum Workshop Aktualisierung der Strategie Digitale Schweiz vom 10. April 2018 in Bern	Bern
UVEK 2019	Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK	Bericht zu den Empfehlungen der Expertengruppe zur Zukunft der Datenbearbeitung und Datensicherheit: Kenntnisnahme und weiteres Vorgehen. Stand: 15. Oktober 2019	Bern

Tabelle 12: Die Strategien »Informationsgesellschaft Schweiz« (1998, 2006, 2012) und »Digitale Schweiz« (2016, 2018) im Vergleich

Strategie	Abkürzung	Reichweite bzw. Anspruch	Umfang
Strategie des Bundesrates für eine Informationsgesellschaft in der Schweiz (18. Februar 1998)	SIG 1998	»In einem ersten Schritt werden mit der nachfolgenden Strategie die für die Förderung einer Informationsgesellschaft in der Schweiz einzuhaltenden Grundsätze definiert und die Gebiete mit dem dringendsten Handlungsbedarf bezeichnet. Die zweite Phase besteht aus der Umsetzung der für die bezeichneten Gebiete statuierten Grundsätze durch die zuständigen Departemente.« (S. 1)	1512 Wörter (5 Seiten)
Strategie des Bundesrats für eine Informationsgesellschaft in der Schweiz (Januar 2006)	SIG 2006	»Diese neue Strategie Informationsgesellschaft ist primär für den Bund handlungsrelevant. In Anbetracht des Themas, das herkömmliche Grenzen sprengt, kann sie insbesondere auch für Kantone und Gemeinden ein Orientierungsrahmen sein.« Kein »umfassender gesamtgesellschaftlicher Regelungsanspruch« »Strategie beschreibt primär Ziele und Massnahmen, welche in Bundeskompetenz zu realisieren sind.« (S. 1)	2700 Wörter (11 Seiten)
Strategie des Bundesrates für eine Informationsgesellschaft in der Schweiz (März 2012)	SIG 2012	»Die Strategie des Bundesrates zeigt die Handlungsfelder auf und definiert den Handlungsbedarf der öffentlichen Hand.« (S. 3)	4157 Wörter (20 Seiten)
Strategie Digitale Schweiz (April 2016)	SDS 2016	»Vor diesem Hintergrund gibt die Strategie »Digitale Schweiz« die Leitlinien für das staatliche Handeln vor und zeigt auf, wo und wie Behörden, Wirtschaft, Wissenschaft, Zivilgesellschaft und Politik zusammenarbeiten müssen, damit wir diesen Transformationsprozess gemeinsam zum Nutzen unseres Gemeinwesens gestalten können.« (S. 3)	5200 Wörter (27 Seiten)
Strategie Digitale Schweiz (5. September 2018)	SDS 2018	»Damit alle von den Vorteilen der aktuellen Entwicklungen profitieren können, müssen die Behörden aller föderalen Ebenen, Zivilgesellschaft, Wirtschaft, Wissenschaft und Politik den Wandel gemeinsam vorantreiben (Multistakeholder-Ansatz).« »Vor diesem Hintergrund definiert der Bundesrat mit seiner Strategie die Leitlinien für eine »digitale Schweiz« und fordert alle Anspruchsgruppen der digitalen Schweiz auf, relevante Umsetzungsprojekte und Querschnittsthemen gemeinsam anzugehen.« (S. 1)	7595 Wörter (19 Seiten)

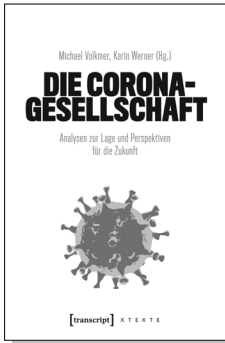
Tabelle 13: Studiengänge und Vertiefungen in Datenwissenschaften an Schweizer Universitäten und Hochschulen (Stand: 03. Februar, 2022)

Universität/Hochschule	Fakultät/Department	Studienstufe	Programmbezeichnung	Jahr
ETH Zürich	Departemente für Informatik, Mathematik sowie Informationstechnologie und Elektrotechnik	Master of Science (MSc)	Data Science	2017
EPF Lausanne	School of Computer and Communication Sciences	Master of Science (MSc)	Data Science	2017
Universität Zürich	Departement Informatik, Fakultät für Wirtschaftswissenschaften	Master of Science (MSc)	Computer Science, Major/Minor in Data Science	2016
Universität Bern, Universität Freiburg, Université de Neuchâtel (BENEFRI)	Departemente für Informatik	Master of Science (MSc)	Computer Science, with specialisation in Data Science	2016
Università della Svizzera Italiana	Fakultät für Informatik	Master of Science (MSc)	Computational Science	2013
Università della Svizzera italiana	Fakultät für Informatik	Master of Science (MSc)	Artificial Intelligence	2017
Université de Lausanne/ Université de Neuchâtel (bis 2019)	Fakultäten für Wirtschaftswissenschaften	Master of Science (MSc)	Information Systems	2007
Université de Lausanne (ab 2019)	Fakultät für Wirtschaftswissenschaften (HEC)	Master of Science (MSc)	Information Systems	2019
Université de Lausanne	Fakultät für Wirtschaftswissenschaften (HEC)	Master of Science (MSc)	Finance, orientation in Financial entrepreneurship and Data Science	2017
Université de Geneva	Fakultät für Wirtschaftswissenschaften	Master of Science (MSc)	Business Analytics	2017
Fernuniversität Schweiz		Master of Science (MSc)	Artificial Intelligence	2019
Universität Bern	Philosophisch-naturwissenschaftliche Fakultät	Master of Science (MSc)	Statistics and Data Science	2019
Universität Freiburg	Wirtschafts- und sozialwissenschaftliche Fakultät	Master of Science (MSc)	Data Analytics & Economics	2020
Universität St. Gallen	School of Management	Certificate (BA)	Data Science Fundamentals	2017
Université de Lausanne & EPFL	Fakultät für Wirtschaftswissenschaften (HEC), Swiss Data Science Center (EPFL)	Certificate of Advanced Studies (CAS)	Data Science & Management	2018

ETH Zürich	Departemente für Informatik, Mathematik sowie Informationstechnologie und Elektrotechnik	Diploma of Advanced Studies (DAS)	Data Science	2018
Universität Zürich	Departement Informatik, Fakultät für Wirtschaftswissenschaften	Certificate of Advanced Studies (CAS)	Big Data and Machine Learning	2018
Universität Bern	Philosophisch-naturwissenschaftliche Fakultät	Certificate of Advanced Studies (CAS)	Applied Data Science	2018
Universität Bern	Philosophisch-naturwissenschaftliche Fakultät	Certificate of Advanced Studies (CAS)	Advanced Machine Learning	2020
Universität Bern	Philosophisch-naturwissenschaftliche Fakultät	Certificate/ Diploma/Master of Advanced Studies (MAS)	Statistical Data Science	2020
Universität Bern	Philosophisch-naturwissenschaftliche Fakultät	Certificate of Advanced Studies (CAS)	Advanced Statistical Data Science	2021
Hochschule Luzern	Wirtschaft	Master of Science (MSc)	Applied Information and Data Science	2018
Fachhochschule Graubünden (bis 2020)	Departement Angewandte Zukunftstechnologien	Master of Science (MSc)	Business Administration, Major in Information and Data Management	2011
Fachhochschule Graubünden (ab 2021)	Departement Angewandte Zukunftstechnologien	Master of Science (MSc)	User Experience Design & Data Visualization, Studienrichtung Data Visualization	2021
Diverse Fachhochschulen (ZHAW, BFH, FHO, SUPSI, HES-SO, HSLU & FHNW)	Schools of Engineering	Master of Engineering (MSE)/ Master of Science	Vertiefungen und Profile in Data Science	2018
Fachhochschule Nordwestschweiz	Hochschule für Technik	Bachelor of Science (BSc)	Data Science	2019
Hochschule für Technik Rapperswil	Hochschule für Technik	Bachelor of Science (BSc)	Informatik mit Studienschwerpunkt Data Engineering & Machine Intelligence	2017
Hochschule Luzern	Informatik	Bachelor of Science (BSc)	(Wirtschafts-)Informatik mit Major in Data Engineering & Data Science	2018
Fachhochschule Nordwestschweiz	Hochschule für Technik	Bachelor of Science (BSc)	Informatik, Vertiefungsrichtung Data Science	2018
Hochschule Luzern	Departemente Informatik, Technik & Architektur und Wirtschaft	Bachelor of Science (BSc)	Mobility, Data Science and Economics	2020

Hochschule Luzern	Departement Informatik	Bachelor of Science (BSc)	Artificial Intelligence & Machine Learning	2020
Fachhochschule Graubünden	Departement Angewandte Zukunftstechnologien	Bachelor of Science (BSc)	Computational and Data Science	2021
Zürcher Hochschule für Angewandte Wissenschaften	School of Engineering	Master of Advanced Studies (MAS)	Data Science	2014
Berner Fachhochschule	Departement Technik und Informatik	Master of Advanced Studies (MAS)	Data Science	2015
Hochschule Luzern	Informatik	Certificate of Advanced Studies (CAS)	Big Data Analytics	2018
Hochschule Luzern	Informatik	Certificate of Advanced Studies (CAS)	Artificial Intelligence/ Künstliche Intelligenz	2019
Fachhochschule Nordwestschweiz	Hochschule für Technik	Certificate of Advanced Studies (CAS)	Data Science	2018
Scuola Universitaria Professionale della Svizzera Italiana (SUPSI)	Department of Innovative Technologies	Certificate of Advanced Studies (CAS)	Big Data Analytics and Machine Learning	2018
Scuola Universitaria Professionale della Svizzera Italiana (SUPSI)	Department of Innovative Technologies	Certificate of Advanced Studies (CAS)	Data Science & Quality by Design	2019
Scuola Universitaria Professionale della Svizzera Italiana (SUPSI)	Department of Innovative Technologies	Certificate of Advanced Studies (CAS)	Data Science for Industry 4.0	2019
Fernfachhochschule Schweiz (FFHS)	Department Informatik	Diploma of Advanced Studies (DAS)	Data Science	2016
Hochschule für Wirtschaft Zürich (HWZ)	Center for Data Science & Technology	Certificates of Advanced Studies (CAS)	Applied Data Analytics (früher: CAS Big Data Analysis)	2017
Hochschule für Wirtschaft Zürich (HWZ)	Center for Data Science & Technology	Certificates of Advanced Studies (CAS)	Customer Intelligence	2016
Hochschule für Wirtschaft Zürich (HWZ)	Center for Data Science & Technology	Certificates of Advanced Studies (CAS)	Machine Learning	2019

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



Kerstin Jürgens

Mit Soziologie in den Beruf Eine Handreichung

September 2021, 160 S., kart.

18,00 € (DE), 978-3-8376-5934-4

E-Book:

PDF: 15,99 € (DE), ISBN 978-3-8394-5934-8



Gabriele Winker

Solidarische Care-Ökonomie

Revolutionäre Realpolitik für Care und Klima

März 2021, 216 S., kart.

15,00 € (DE), 978-3-8376-5463-9

E-Book:

PDF: 12,99 € (DE), ISBN 978-3-8394-5463-3

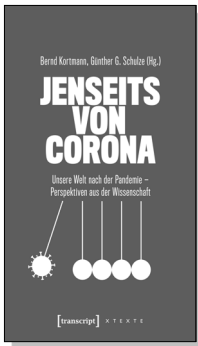
**Leseproben, weitere Informationen und Bestellmöglichkeiten
finden Sie unter www.transcript-verlag.de**

Soziologie



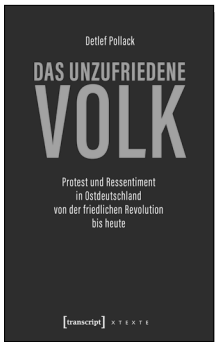
Wolfgang Bonß, Oliver Dimbath,
Andrea Maurer, Helga Pelizäus, Michael Schmid
Gesellschaftstheorie
Eine Einführung

Januar 2021, 344 S., kart.
25,00 € (DE), 978-3-8376-4028-1
E-Book:
PDF: 21,99 € (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



Detlef Pollack
Das unzufriedene Volk
Protest und Ressentiment in Ostdeutschland
von der friedlichen Revolution bis heute

2020, 232 S., Klappbroschur, 6 SW-Abbildungen
20,00 € (DE), 978-3-8376-5238-3
E-Book:
PDF: 17,99 € (DE), ISBN 978-3-8394-5238-7
EPUB: 17,99 € (DE), ISBN 978-3-7328-5238-3

**Leseproben, weitere Informationen und Bestellmöglichkeiten
finden Sie unter www.transcript-verlag.de**