



HAL
open science

Professionalisierung mittels Ambiguität. Die diskursive Konstruktion von Data Scientists in Wirtschaft und Wissenschaft

Robert Dorschel, Philipp Brandt

► To cite this version:

Robert Dorschel, Philipp Brandt. Professionalisierung mittels Ambiguität. Die diskursive Konstruktion von Data Scientists in Wirtschaft und Wissenschaft. Zeitschrift für Soziologie, 2021, 50 (3-4), pp.193 - 210. 10.1515/zfsoz-2021-0014 . hal-03405863

HAL Id: hal-03405863

<https://hal.science/hal-03405863>

Submitted on 3 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Robert Dorschel*, Philipp Brandt

Professionalisierung mittels Ambiguität. Die diskursive Konstruktion von Data Scientists in Wirtschaft und Wissenschaft

Professionalization via Ambiguity. The Discursive Construction of Data Scientists in Higher Education and the Labor Market

<https://doi.org/10.1515/zfsoz-2021-0014>

Zusammenfassung: Der Beitrag fragt, welche Deutungsmuster mit der Konstruktion und Subjektivierung der aufsteigenden Berufsgruppe der ‚Data Scientists‘ einhergehen. Aus soziologischer Perspektive beruht ihre fortschreitende Institutionalisierung maßgeblich auf Objektivierungsprozessen sozialer Wissensbestände. Anhand einer diskursanalytischen Auswertung von Stellenanzeigen und Studiengangsbeschreibungen zeigen wir eine ambige Logik in der sozialen Konstruktion von Data Scientists auf. In Wirtschaft und Wissenschaft besteht einerseits Dissens über die Klassifikation und Einordnung von Data Scientists, während andererseits basale Deutungsmuster Data Scientists als GrenzgängerInnen, Anti-Nerds und WeltverbesserInnen diskursivieren. Die ambige Logik aus Dissens und Konsens interpretieren wir als einen strategischen Vorteil für die Professionalisierung: Ambiguität schafft Integrationsfähigkeit für unterschiedliche Erwartungshaltungen und Imaginationen. Der Beitrag legt die Subjektivierung ebenjener Professionals frei, die eine Schlüsselposition in der objektivierenden Digitalwirtschaft einnehmen.

Schlüsselwörter: Data Science; Digitalisierung; Diskursanalyse; Professionalisierung Professionen; Wissensarbeit; Tech Workers; Ambiguität; Bourdieu; Foucault.

Danksagung: Für Anregungen und konstruktive Kritik danken wir den HerausgeberInnen, den drei anonymen GutachterInnen, sowie Jonas Ferdinand, Steffen Mau, Philippe Saner, Philipp Staab, Anna Thieser und Teresa Völker.

***Korrespondenzautor:** Robert Dorschel, University of Cambridge, Department of Sociology, 16 Mill Lane, Cambridge, CB2 1SB, E-Mail: rcd49@cam.ac.uk

Philipp Brandt, Sciences Po, Department of Sociology/Centre for the Sociology of Organizations, 19 rue Amélie, 75007 Paris, E-Mail: philipp.brandt@sciencespo.fr

Abstract: This article studies the social construction of the emerging „data scientist” profession. Data science’s ongoing institutionalization rests on the objectification of social stocks of knowledge. A discourse analysis of job ads and university programs reveals an ambiguous logic in the construction of data scientists. Initially, we show divergent classifications of data science work in the academic and economic fields. We then reveal deeper interpretation patterns across both fields that define data scientists as frontierspersons, anti-nerds, and do-gooders. We present this simultaneous divergence and consensus as a strategic advantage for professionalization: the underlying meaning integrates varying expectations and ideas around data scientists, who become broadly salient. The article thus presents the subjectivation of professionals in an objectifying digital economy.

Keywords: Data Science; Digitalization; Discourse Analysis; Professionalization; Professions; Knowledge Work; Tech Workers; Ambiguity; Bourdieu; Foucault.

1 Einleitung

Die Soziologie widmet sich zunehmend der Erforschung des digitalen Wandels. Insbesondere gesellschaftstheoretische Diagnosen, wie die eines sich ausbreitenden metrischen Wir (Mau 2017), eines digitalen Kapitalismus (Schiller 2000; Staab 2019), eines Überwachungskapitalismus (Zuboff 2019) oder eines Datenkolonialismus (Coudry & Mejias 2019), haben sich kritisch mit den Auswirkungen der Digitalisierung auf das Soziale auseinandergesetzt und damit ein lautes Echo innerhalb und außerhalb der Soziologie hervorgerufen. In den deutschen wie auch den angloamerikanischen Debatten existieren jedoch bis dato kaum Studien über die berufliche Hinterbühne dieser makrosoziologischen Beobachtungen. Digitale Arbeit bzw.

Arbeit im digitalen Kapitalismus wurde bisher hauptsächlich im Hinblick auf die prekären Beschäftigungsgruppen untersucht (Graham & Anwar 2018). Nur vereinzelt nehmen sich Studien der Berufsgruppen an, die den ‚digitalen Informationsraum‘ (Boes et al. 2015) und seine Technologien konfigurieren, verwalten und ausweiten (Marwick 2013; Mützel et al. 2018; Wajcman 2018).

Wir wenden uns in diesem Artikel den ‚Data Scientists‘ zu: einer im Kontext der Digitalisierung entstehenden Berufsgruppe, die sozialstrukturell der neuen Mittelklasse (Reckwitz 2019: 86) zuzurechnen ist und mittlerweile zentrale Funktionen innerhalb datenbasierter Organisationen erfüllt. Data Scientists nehmen die objektivierende Hinterbühne einer Digitalwirtschaft ein, deren Algorithmen und KI-Systeme das private und soziale Leben zunehmend beeinflussen. Der Beitrag fragt, welche Deutungsmuster mit der diskursiven Konstruktion und Subjektivierung von Data Scientists einhergehen. In der Regel werden Data Scientists schlicht rudimentär als Experten definiert, die gute Statistikkennntnisse und grundlegende Programmierfähigkeiten vorweisen können; und somit geeignet dafür sind, die in digitalisierten Systemen fortlaufend anfallenden Datenmengen nach ‚Mustern‘ (Nassehi 2019) zu analysieren. Intuitiv einleuchtend, legt diese Erklärung doch implizit einen professionssoziologisch unbefriedigenden, funktionalistischen Erklärungsansatz nahe (Abbott 1988). Die rudimentären Definitionen, die von praktizierenden Data Scientists über Data Science formuliert werden, gehen mit Zelebrierungen und Warnungen in der medialen Öffentlichkeit einher. So wird Data Science im Harvard Business Magazine als ‚Sexiest Job of the 21st century‘ (Davenport & Patil 2012) gepriesen, während die ehemalige Mathematik-Professorin und Aktivistin Cathy O’Neil vor den ‚Weapons of Math Destruction‘ (2016) von Data Scientists warnt. Wenn wir davon ausgehen, dass Berufsfelder sich im Zuge ihrer Institutionalisierung auch mit externen Akteuren in ‚verbundenen Ökologien‘ (Abbott 2005) bewegen, kann von einem solchen Blickwinkel eine Reihe weiterer externer Akteure identifiziert werden, die an der sozialen Konstruktion von Data Science beteiligt sind. Der Prozess, ein neues Arbeitsfeld zu definieren und zu okkupieren, wird dementsprechend nicht nur von, sondern auch *für* Data Scientists vorangetrieben. Dieser Aufsatz untersucht die Phänomenklassifikationen und Deutungsmuster, die mit ebendieser externen Konstruktion und Subjektivierung von Data Scientists einhergehen.

Wir analysieren die sozialfigurative Bestimmung und Ausbreitung von Data Scientists in zwei Arenen, die zentrale Konsekrationsinstanzen für aufstrebende Berufsfelder darstellen: Wissenschaft und Wirtschaft (Abbott 1988: 55, 64). Wir konzeptualisieren diese Arenen im Licht

der Differenzierungs- und Herrschaftstheorie Bourdieus weitergehend als soziale Felder mit Benennungsmacht (Bourdieu 2004). Mittels einer wissenssoziologischen Diskursanalyse von US-amerikanischen Data Science-Stellenanzeigen und Studiengangbeschreibungen, dem Entstehungsraum von Data Science, untersuchen wir die Diskursivierungen und Subjektivierungen von Data Scientists innerhalb dieser Felder. Im Anschluss an Foucault begreifen wir diskursive Konstruktionen nicht primär als Repräsentationen, sondern als Akte des performativen Mithervorbringens (Foucault 1981: 74). Aus dieser Perspektive sind Diskurse nicht nur Ressourcen von Professionen, sondern sie gestalten die Subjektformen und Kapitalwechselkurse sowie damit verbundene Möglichkeiten der ‚Kompetenzdarstellungskompetenz‘ (Pfadenhauer 2003) aktiv mit.

Im Zuge unserer Analyse von Data Science-Studiengangbeschreibungen und Stellenanzeigen wird ersichtlich, dass die Konstruktion von Data Science einer ambigen sozialen Logik folgt. Es besteht Dissens, aber auch Konsens bei der Konstruktion von Data Scientists in Wissenschaft und Wirtschaft. So wird in den Studiengangbeschreibungen ein Bild von Data Scientists gezeichnet, welches deren multipolare akademische Expertise betont, sie als ‚leader‘ tituliert und ihnen eine kritische Datenethik attestiert. In den Stellenausschreibungen hingegen werden Data Scientists als Schnittstellenprofession beschrieben, die als ‚team player‘ agieren sollen und von denen in erster Linie eine Datenverwertungsethik verlangt wird. Parallel zu diesen kontrasthaften sozialfigurativen Feldkonstruktionen lassen sich jedoch gemeinsame basale Deutungsmuster identifizieren. So werden Data Scientists von Wissenschaft und Wirtschaft als *GrenzgängerInnen* gedeutet: als entdifferenzierende Arbeitssubjekte, die heterogene Wissensgebiete oder Projektglieder miteinander verkoppeln können, um dadurch zu neuen und innovativen Lösungen für Probleme zu gelangen. Und obwohl der Grenzgänger ein technisches Arbeitssubjekt impliziert, wird dieses Muster von beiden Feldern mit der Deutung von Data Scientists als *Anti-Nerds* flankiert. Ob sie nun als ‚leader‘ oder ‚team player‘ klassifiziert werden: beides fordert soziale Kompetenzen. Dieses Bild zeigt einen scharfen Kontrast zu bestehenden Code-versierten ExpertInnen, die mit Data Scientists um das Tätigkeitsfeld der digitalen Datenarbeit buhlen, auf, insofern Computerarbeit als charakteristische Tätigkeit für ‚sozial unbeholfene‘ Nerds angesehen wird (siehe Kendall 1999: 262). Das dritte gemeinsame Deutungsmuster ist die basale Klassifikation von Data Scientists als *WeltverbesserInnen*. Es wird von beiden Feldern eine Welt konstatiert, die sich von zunehmenden Datenmengen überflutet sieht und in der Data

Scientists in solutionistischer Manier die Welt verbessern wollen, können und sollen.

Die hier beobachtete diskursive Ambiguität – eine Subjektformation, die bisher nur vereinzelt für Professionen in Anschlag gebracht wurde (z. B. Stichweh 1996 für LehrerInnen) – verstehen wir im Anschluss an Foucault als Professionalisierungsstrategie ohne Strategen (Foucault 1978: 132). Ambiguität fungiert als eine spezifische Form kulturellen Kapitals, womit die soziale Position verbessert werden kann. Die ambige diskursive Konstruktionsweise von Data Scientists ähnelt dabei der bekannten Logik eines Kippbilds (Gregory 2000): sie vermag vom Standpunkt der Betrachterin variierende ‚Data Science-Imaginationen‘ (Saner 2019) in einem Bild einzufangen. Wir theoretisieren, dass so eine breite gesellschaftliche Anschlussfähigkeit ermöglicht und damit ein zentraler Erklärungsfaktor für die fortschreitende Institutionalisierung und Professionalisierung von Data Science erkennbar wird. Dieses Argument schließt an Befunde zur strategischen Effektivität von Ambiguität in anderen Feldern an (Davenport & Leitch 2005) und unterstreicht im Fall der Data Scientists, dass Ambiguität auch in Arbeitsfeldern, die eine vermeintliche datenbasierte oder anderweitig formale Objektivität versprechen, Wirkung entfaltet. Diese Strategien, so argumentieren wir, finden insbesondere im Kontext technologischer Transformationen reichhaltige Nährböden.

Im Folgenden wird zuerst der Forschungsstand über Data Scientists rekapituliert. Daran schließt die Erörterung unseres theoretischen Rahmens an, der sich, neben der Professionssoziologie, aus einer Verknüpfung von Pierre Bourdieus Feldtheorie und der Diskurstheorie Michel Foucaults speist. Im Weiteren erfolgt eine Einführung in die Wissenssoziologische Diskursanalyse als Methodik und es wird die Datengrundlage unserer Untersuchung näher erläutert. Der empirische Analyseteil gliedert sich in drei Abschnitte, in denen die drei zentralen Deutungsmuster rekonstruiert werden. Die Herleitung erfolgt ausgehend von den konfliktiven Klassifikationspraxen. Daran knüpft die Analyse der ambigen Logik, die den diskursiven Konstruktionen zugrunde liegt, an. Abschließend diskutieren wir die Bedeutung unserer Befunde für die Professionssoziologie sowie für die soziologische Erforschung des digitalen Wandels.

2 Forschungsstand

Es überwiegt aktuell eine außerakademische Auseinandersetzung mit Data Scientists als einer neuen Berufsgruppe; medial wirksame Beiträge versuchen, Data

Science etwa durch Entstehungsanekdoten in ihrer Existenz zu erklären (siehe etwa Hammerbacher 2009). An anderen Stellen bemühen sich Data Scientists, ihre neue Rolle zu erklären, wobei sie primär die technischen Grundlagen betonen (siehe etwa O’Neil & Schutt 2013). Außenstehende Autoren, wie González-Bailón (2017), vermitteln eine historische Perspektive, die die Nähe zwischen Data Science und den Sozialwissenschaften betont. Salganik (2017) liefert den technisch und reflexiv umfangreichsten Beitrag in seinem Versuch, Data Scientists und Sozialwissenschaftlern eine gemeinsame Zukunft zu skizzieren. Andere Beiträge finden bezüglich dieser Causa größere Spannungen bzw. schreiben Data Science weniger Legitimation zu (Donoho 2015; Ribes 2019).

Dezidiert akademische Forschungen, die sich Data Science als Untersuchungsgegenstand annehmen, sind rar. Eine Ausnahme stellt Philippe Saners (2019) Untersuchung zur Implementierung von Data Science in die Schweizer Hochschullandschaft dar. Der Kernbefund seiner Analyse lautet, dass insbesondere ökonomische Rechtfertigungen für die Implementierung von Data Science-Studiengängen aufgeführt wurden. In einem anderen Beitrag nimmt sich Saner, zusammen mit Mützel und Unternährer, der praktischen Arbeit von Data Scientists an. In Anlehnung an Science and Technology Studies (Latour & Woolgar 1986) zeigen die AutorInnen auf, dass Daten niemals roh existieren, sondern sozial produziert und sinnbehaftet weiterverarbeitet werden (Mützel et al. 2018: 122). Dabei nehme der Prozess der Datenaufbereitung bzw. -reinigung von ‚messy data‘ hin zu verheißungsvollen ‚tidy data‘ die meiste Zeit der Data Science-Tätigkeit ein – dies spiegele sich in der glanzvollen Außendarstellung des Berufes jedoch nicht wider.

Unklar bleibt in den diversen Abhandlungen, wie es sich mit Data Scientists und der Professionalisierung verhält. Zwar weist Data Science eine noch sehr junge Entwicklungsgeschichte auf, trotzdem sehen wir bereits in den wenigen Data Science-Abhandlungen bekannte Muster aus der Forschung zu Professionen in der Entstehung. Wie etwa bei den von Stichweh (1984: 255) untersuchten englischen Physikern zeigt sich eine Vermischung von Experten und Amateuren. Des Weiteren erkennen wir sowohl die von Abbott vorhergesagten Konflikte zwischen benachbarten Expertengruppen über ‚Jurisdiktionen‘ (Abbott 1988: 40) als auch die tiefer liegenden, konstruktivistischen Prozesse um Experten, Nichtexperten und materielle Objekte, die sich sonst erst in der historischen und kulturellen Vielfalt von ‚Wissensarbeit‘ erkennen lassen (Eyal 2013). Gleichzeitig stellt der Data Science-Fall eine

Herausforderung für die Professionssoziologie und ihre Fokussierung auf historische Prozesse dar. Zwar hat sich, seit Abbotts allgemeiner Theorie zum systemischen Zusammenhang von Professionen, eine wachsende Debatte den Problemen von neueren Berufsgruppen im Organisationskontext gewidmet (Manske & Schnell 2010; Muzio et al. 2011; Noordegraaf 2007; Suddaby et al. 2009), jedoch beziehen sich die von dieser Forschung hervorgebrachten Befunde auf Probleme, denen Data Scientists erst begegnen können, nachdem sie die ersten Schritte im (diskursiven) Professionalisierungsprozess bereits durchlaufen haben. Ein weiteres Defizit der dominanten Professionssoziologie ist eine Konzentration auf Professionals, ohne deren Einbettung in einen weiteren gesellschaftlichen Kontext systematisch in den Blick zu nehmen (Atzeni 2016; Stichweh 2000). Abbotts Konzept der verbundenen Ökologien bietet einen wichtigen Referenzpunkt, ohne jedoch eine ausformulierte Differenzierungstheorie mit entsprechenden Analysekonzepten bereitzustellen, um die Emergenz einer neuen Profession rekonstruieren zu können. Seine knappen Beobachtungen zur Institutionalisierung der benachbarten Computerwissenschaften etwa stützen sich auf eine Kluft zwischen praktischer und wissenschaftlicher Arbeit (siehe Abbott 2005: 266), eine Situation, die bei Data Science zu diesem frühen Zeitpunkt lediglich eines von unterschiedlichen möglichen Szenarien darstellt.

Um dem im Fall der Data Scientists fehlenden Verständnis der Fremdbeschreibung entgegenzuwirken, wählen wir den Ansatz der Diskursanalyse zur Untersuchung der Konstruktion von Data Science in zwei ihrer institutionellen Entstehungskontexte: Wissenschaft und Wirtschaft. Fremdbeschreibungen haben auch für etablierte Berufsgruppen wie die ärztliche Profession eine signifikante Rolle gespielt (Atzeni 2016). Unser diskursanalytischer Ansatz soll im Fall der Data Scientists ermöglichen, die frühen sprachlich-institutionellen Zuschreibungen von Kompetenzen in verschiedenen sozialen Feldern einfangen zu können, die der Generierung einer professionalen Kompetenzdarstellungskompetenz vorausgehen. Vereinzelt Studien haben einen diskursiven Ansatz zur Professionsanalyse bereits verfolgt – sowohl in neueren Erscheinungen, wie Managementberater (Schmidt-Wellenburg 2009) oder Trennungs- und Scheidungsberater (Halatcheva-Trapp 2018), als auch etablierteren Berufen, wie der Medizin (Wetterer 2002; Atzeni 2016). Ziel dieses Beitrags ist es damit, auch weiter Aufschluss über die diskursive Beschaffenheit professionalisierter Felder zu geben.

3 Theoretischer Rahmen

Diese Untersuchung analysiert also die sozialen Logiken, die die diskursiven Konstruktionsweise und Subjektivierung von Data Scientists strukturieren. In der professionssoziologischen Literatur werden vor allem zwei Bereiche wiederholt behandelt, von denen angenommen wird, dass dortigen Klassifikations- und Deutungspraktiken eine besonders hohe Relevanz bei der Institutionalisierung von Berufen zukommt: die Arena der Universitäten sowie die Arena der Unternehmens- bzw. Arbeitswelt (Abbott 1988; Freidson 1988). Wir ziehen weitergehend Pierre Bourdieus Feldtheorie heran, um jene professionellen Arenen differenzierungs- und herrschaftstheoretisch in den Blick zu nehmen. Die Feldtheorie Bourdieus liefert einen Ansatz, der auch an den Beziehungen zwischen den Feldern interessiert ist (Bourdieu 2004) und im Vergleich zu Abbotts Ansatz stärker auf die Analyse von Kapitalformen, Strategien und sozialen Positionen geeicht ist (Liu & Emirbayer 2016). Somit hilft er, die Genese von Data Science als Beruf innerhalb und zwischen dem ökonomischen und akademischen Machtfeld zu verfolgen. Im Sinne der Feldtheorie werden Wissenschaft und Wirtschaft als akademisches und ökonomisches Konfliktfeld verstanden (Bourdieu 2002, 2014), welches wiederum selbst in ein erweitertes, gesamtgesellschaftliches Konkurrenzfeld eingebettet ist, das Feld der Macht (Bourdieu 2004). Soziale Felder sind eigenlogische Mikrokosmen, in denen bestimmte Regeln und Werte gelten und Akteure mittels verschiedener Kapitalformen versuchen, ihre Sichtweisen der Welt durchzusetzen, um so ihre soziale Position (welche auf der anerkannten Wertigkeit ihrer Kapitalformen beruht) zu sichern oder zu verbessern (Bourdieu & Wacquant 2013: 127). Felder sind also einerseits Horte des Konflikts, andererseits aber immer auch des Konsenses. Streit über eine Ressource kann nur entbrennen, wenn die Relevanz der Ressource von beiden Seiten anerkannt ist, wenn also basale Wissensordnungen in Form eines gemeinsamen Glaubens für das Spiel (*illusio*) sowie unhinterfragte Weltannahmen (*doxa*) existieren (Bourdieu 2015: 112).

Wie in der Einleitung skizziert, findet die Entstehung von Data Science maßgeblich im Diskurs über Data Scientists statt. Wir ziehen Michel Foucaults Diskurstheorie heran, um die Emergenz von Data Science hinsichtlich ihrer eigenlogischen, diskursiven Aushandlung im akademischen und ökonomischen Feld zu beleuchten. Foucault (1981) definiert den Diskurs als eine Menge von Aussagen, die in Beziehung zueinander stehen, die ein Netz bilden und deren Verbindungen durch Formationsregeln und eigendynamische Logiken bestimmt werden

(Foucault 1981: 39). Das sozialtheoretisch Relevante am Diskurs ist seine überrepräsentative Bedeutung, denn der Diskurs ist aufzufassen:

„nicht [...] als Gesamtheit von Zeichen (von bedeutungstragenden Elementen, die auf Inhalte oder Repräsentationen verweisen), sondern als Praktiken [...] die systematisch die Gegenstände bilden, von denen sie sprechen. Zwar bestehen diese Diskurse aus Zeichen; aber sie benutzen diese Zeichen für mehr als nur zur Bezeichnung der Sachen. Dieses mehr macht sie irreduzibel auf das Sprechen und die Sprache. Dieses mehr muß man ans Licht bringen und beschreiben.“ (Foucault 1981: 74)

Als Ordnungen des Denk- und Sagbaren kreieren Diskurse also aktiv den symbolischen gesellschaftlichen Bezugsrahmen, durch den soziale Phänomene erst hervorgebracht werden (Diaz-Bone 2006; Keller 2011a: 127). Mit dem Diskurskonzept und der korrespondierenden Diskursanalyse sucht Foucault nach Grundmustern des Wissens, nach den historisch spezifischen Wissensordnungen, die Subjektformen (re-)produzieren, indem sie die Grenzen des Sagbaren ziehen. Foucault interessierte sich insbesondere in seinem Frühwerk dafür, wie historisch spezifische Subjekte, z. B. ‚der Wahnsinnige‘, durch diskursive Praktiken bestimmter gesellschaftlicher Institutionen (etwa der Wissenschaft) hergestellt wurden (Foucault 2003). Daran, und an neuere Entwicklungen in der empirischen Subjektivierungsforschung (siehe Bosančić et al. 2021) angelehnt, interessieren wir uns für die Diskursivierung des Data Scientist-Subjekts.

Neben einer ähnlichen relationalen und machtfokussierten Epistemologie kann der gemeinsame Fluchtpunkt der Foucault’schen mit der Bourdieu’schen Sozialtheorie in der Frage lokalisiert werden, welche Subjektivierungs- bzw. Habitualisierungsweisen zu einer spezifischen Zeit und an einem bestimmten Ort vorherrschen (Diaz-Bone 2002; Kajetzke 2008; Maeße & Hamann 2016; Schmidt-Wellenburg 2014). Zwar bestehen wichtige Differenzen in der Theoretisierung von Subjekt bzw. Habitus, grundsätzlich jedoch zielen beide Konzepte auf die Analyse der Macht- und Herrschaftsverhältnisse, die die Herausbildung von spezifischen Denk-, Wahrnehmungs- und Handlungsweisen bedingen (Reckwitz 2015). Durch die Kombination der Theorien entsteht eine Heuristik, die an der feldspezifischen Hervorbringung und Brechung von Diskursen und einhergehenden Subjektivierungs- bzw. Habitualisierungsweisen interessiert ist. Die Kombination aus Diskurs- und Feldtheorie eignet sich somit, um das frühe Entwicklungsstadium von Professionen zu untersuchen, denn sie legt die diskursive Beschaffenheit professionalisierter Felder frei und ergänzt damit das existierende professionssoziologische Instrumentarium.

Aus dieser Perspektive nehmen wir Data Science als potenzielles Professionsfeld in den Blick, dessen Emergenz und Okkupation der Datenarbeit im Hinblick auf Diskursivierungen des akademischen und ökonomischen Feldes untersucht werden müssen.

4 Wissenssoziologische Diskursanalyse

Zur Operationalisierung der Foucault’schen Methodologie ziehen wir das Forschungsprogramm der Wissenssoziologischen Diskursanalyse (WDA) von Reiner Keller (2011a, 2011b) heran.¹ Dieses bietet sich nicht nur durch konkrete methodische Überlegungen zur Umsetzung der Foucault’schen Diskurstheorie an, sondern auch, weil es mit der Bourdieu’schen Sozialtheorie insofern vereinbar ist, als das Programm ebenfalls die Akteure systematisch in den Blick nehmen möchte, die an der eigendynamischen diskursiven Konstruktion der Wirklichkeit beteiligt sind (Keller 2011a: 49). Im Anschluss an Foucault geht Keller davon aus, dass in den endlichen und verstreuten Äußerungsereignissen von sozialen Akteuren Aussagen enthalten sind, die als Manifestation der strukturierten Prozessierung kontingenter gesellschaftlicher Wissensvorräte von Interesse sind (Keller 2011a: 68). Äußerungen bergen einen typisierbaren Inhalt, eine Aussage, die es zu entschlüsseln gilt. Diskurse entstehen demnach aus zusammenhängenden Aussagen in Aussagensystemen. Sie materialisieren sich im Kontext sozialer Felder in sprachlichen und nichtsprachlichen Interaktionen, visuellen Medien und Textdokumenten.

Die zwei Konzepte der Phänomenstruktur und Deutungsmuster stellen die Ankerpunkte in Kellers Forschungsprogramm dar. Das Konzept der Phänomenstruktur verweist auf die unmittelbaren inhaltlichen Themen eines Diskurses sowie deren unterschiedliche Elemente und Dimensionen. Es geht hier also um die Untersuchung eines Korpus hinsichtlich der regelmäßig auftretenden Themen, Problemdimensionen und Wertungen (Keller 2011a: 249). Eine Sondierung der Phänomenstruktur steht insbesondere zu Beginn einer wissenssoziologischen Dis-

¹ Foucault erarbeitete in seiner archäologischen Phase zwar elaboreierte methodologische und sozialtheoretische Überlegungen über Diskurse, jedoch keine konkrete Handlungsanleitung zur praktischen Umsetzung des diskursanalytischen Forschungsprogramms. Reiner Kellers Entwurf einer wissenssoziologischen Diskursanalyse stellt einen Verfahrensvorschlag im Anschluss an Foucault dar, an dem sich diese Untersuchung orientiert.

kursanalyse an. Das Herzstück der WDA stellt das Konzept der Deutungsmuster dar. Der Zusatz ‚Muster‘ verweist bereits darauf, dass es hier um die sprachlich-materiale Gestalt mehrerer Sinngehalte geht. Ein Deutungsmuster verknüpft und verdichtet mehrere Phänomenklassifikationen und Narrative in basalen Interpretationsschemata, die individuelle und kollektive Erfahrungen organisieren, indem sie Sinn stiften (Keller 2011a: 240). Während die Phänomenstruktur also auf die thematischen Dimensionen eines Diskurses verweist und unmittelbar durch eine ‚Oberflächenanalyse‘ erschließbar ist, ist das Konzept der Deutungsmuster auf einer tiefer liegenden Ebene zu verorten und erfordert eine stärker interpretativ operierende ‚Feinanalyse‘ des Materials.

5 Datengrundlage und Auswertung

Als Datenmaterial dieser Untersuchung fungieren Studienangangsbeschreibungen und Stellenanzeigen. Hiermit wollen wir uns der Einbettung von Data Science-Professionals innerhalb eines ausdifferenzierten Feldes der Macht zuwenden und die gesellschaftlichen Erwartungshaltungen einfangen. Aus diskurstheoretischer Perspektive werden die Materialien der akademischen und ökonomischen Felder nicht nur als Repräsentation von sozial Gegebenem verstanden, sondern als Praktiken, die (mit) hervorbringen, wovon sie sprechen. Hier wird maßgeblich definiert, wer dazugehört und wer nicht (z. B. Atzeni 2016; Duffy & Schwartz 2017). Studienangangsbeschreibungen und Stellenanzeigen stellen demnach zentrale Instanzen der Wissensproduktion und Subjektanrufung dar. Data Scientist-AnwärterInnen müssen durch diesen ‚obligatory passing point‘ (Callon 1984) hindurch. Gleichwohl ist die Aussagekraft von den zwei Textmaterialien kritisch zu reflektieren: Die Materialien stellen lediglich einen Ausschnitt der Diskursaktivitäten in den jeweiligen Feldern dar (als weitere Materialsarten könnten etwa Blogs, Podcasts, Mitteilungen von Akkreditierungsinstanzen oder Lehrbücher fungieren). Des Weiteren sind die ausgewählten Materialien speziell in der Hinsicht, dass es sich um kuratierte Außenpräsentationen von institutionalisierten Feldakteuren handelt. Zudem ist zu unterstreichen, dass dieser Beitrag lediglich die Anrufung des Data Science-Subjekts untersucht, nicht aber seine nichtdiskursive Manifestierung. Unter Vorbehalt dieser Einschränkungen ist unser Ansatz relevant und ausreichend robust, um eine Grundlage zur Erforschung des Emergenzprozesses von Data Science aus diskurstheoretischer Perspektive zu schaffen.

Geografisch konzentrieren wir uns auf den US-amerikanischen Kontext. Wie bereits erwähnt, wurde hier erstmals die Data Science-Profession ausgerufen (Daven-

port & Patil 2012; Hammerbacher 2009). Zwar haben sich seither Data Science-Studiengänge und Positionen global verbreitet, es bleibt jedoch unklar, inwieweit die Rolle der Data Scientists im Zuge dieser Verbreitung an lokale kulturelle Kontexte angepasst wurde. Dieser Prozess selbst stellt einen relevanten Forschungsgegenstand dar. Für ein solches Vorhaben bedarf es allerdings in erster Instanz eines fundierten Verständnisses für die Diskursivierung von Data Scientists im Entstehungsraum.

Die feldtheoretische Kontextualisierung verlangt ein Vorgehen in der Datenerhebung, das Rücksicht auf die Feldpositionen innerhalb der Arenen der Wissenschaft und Wirtschaft nimmt. Für das akademische Feld wurden jeweils fünf hoch rangierende, fünf mittelhoch rangierende und fünf niedrig rangierende Universitäten (siehe Tab. 1) verwendet.² So sollte eine partielle relationale Kontextualisierung der diskursiven Konstruktionen mit der Struktur und dem Positionengeflecht des akademischen Feldes ermöglicht werden. Innerhalb dieser drei Statuskategorien wurden solche Universitäten für das Korpus ausgewählt, die das Data Science-Studium mit-diskursivieren anstatt sich lediglich des neuen Begriffs zu bedienen. Dieses Verfahren ergab 15 Universitäten mit 25 Data Science-Studiengangsbeschreibungen als Korpusmaterial (manche Universitäten³ bieten mehr als eine Data Science-Studienoption an). Für das ökonomische Feld fungieren Stellenanzeigen des führenden Internetportals ‚glassdoor‘ als Korpusmaterial. Da Stellenanzeigen um ein Vielfaches zahlreicher auftreten als Studiengänge wurde hier das stratifizierte Erhebungsverfahren mit Zufallsstichproben erweitert. Es wurden zufällig jeweils vier Stellenanzeigen von den auf dem Börsenmarkt höchstdotierten Technologie-Unternehmen ausgewählt. Neben diesen 16 Stellenanzeigen wurden 16 weitere Annoncen von sechs führenden Start-up-Unternehmen zufällig ausgewählt (siehe Tab. 1, zweite Spalte).⁴ Die Erhebung der Stellenanzeigen sowie der Studienangangsbeschreibungen fand zwischen März und Mai 2019 statt.

² Hierfür wurde auf das zwar durchaus umstrittene, aber dennoch weitreichend saliente Hochschulranking von Times Higher Education zurückgegriffen. Als ‚high ranking‘-Universitäten wurden solche eingestuft, die innerhalb der ersten 30 Plätze des US-spezifischen Rankings kategorisiert wurden. Als ‚middle high ranking‘ solche, die zwischen den Plätzen 31–100 rangierten. Als ‚low ranking‘ zählten schließlich alle Universitäten auf den dahinter liegenden Plätzen.

³ Um die Statuskategorien trotz des frühen Zeitpunkts unseres Beobachtungszeitraumes ausgeglichen zu gewichten, haben wir eine Universität aus Kanada in unser Sample mit aufgenommen (Carleton University). Die geographische, kulturelle und institutionelle Nähe im akademischen Bereich führen uns zu der Überzeugung, dass diese Beobachtung zu keiner Verzerrung der Analyse führt.

⁴ Als Start-up wurden solche Unternehmen definiert, die im Erhebungszeitraum jünger als 10 Jahre alt waren und mittels innovativer Technologien, Produkten und/oder Geschäftsmodelle versuchen, Märkte zu erschließen (Startup-Monitor 2018).

Tab. 1: US-Data Science Studiengangsbeschreibungen und Stellenanzeigen

Studiengänge	Stellenausschreibungen
Hochplatziert (N) Brown University (2) – Master Data Science Columbia University (3) – Master Data Science Duke University (2) – Master Interdisciplinary Data Science Harvard University (2) – Master Data Science Stanford University (2) – Master Statistics & Data Science	Etablierte Techfirmen (N) – Amazon (4) – Apple (4) – Google (4) – Microsoft (4)
Mittelpatziert (N) Carleton University (1) – Master Data Science Boston University (2) – Master Computer Science – Data Analytics Georgetown University (2) – Master Data Science for Public Policy Northwestern University (1) – Master Data Science University of Virginia (3) – Master Data Science	Start-up Techfirmen (N) – DoorDash (2) – InstaCart (2) – JUUL Labs (4) – Slack (2) – Stripe (2) – WeWork (4)
Niedrigplatziert (N) DePaul University (1) – Master Data Science Illinois Tech Institute (1) – Master Computer Science in Data Science Oklahoma University (1) – Master Data Science and Analytics San Francisco (1) – Master Data Science Southern Methodist University (1) – Master Data Science	
Gesamt N = 25	Gesamt N = 32

Anmerkung: Die Erhebung der Materialien (n = 57) fand zwischen März und Mai 2019 statt.

Die Datenauswertung orientierte sich an Kellers (2011b) Programm zum diskursanalytischen Forschungsvorgehen. In der Phase der Oberflächenanalyse wurden die Texte allesamt mehrmals gelesen, um insbesondere die Phänomenstruktur herauszuarbeiten. Parallel dazu wurden Codes für Sinnabschnitte vergeben und im Laufe des Forschungsprozesses unter abstrakteren Kategorien subsumiert. Memos dienten zur ersten Theoretisierung der Befunde. Ausgehend von den ersten diagnostizierten Regelmäßigkeiten zwischen den Aussagen wurden Schlüsseldokumente bestimmt, in denen die Wissensmuster als Bündelungen von Klassifikationsschemata besonders konzentriert auftraten. Diese Dokumente wurden dann einer Feinanalyse unterzogen, mit dem Ziel, von den bestehenden Codes auf zugrunde liegende Organisationsmechanismen und Deutungsmuster zu schließen. Während der

abschließenden Ergebnisaufbereitung erfolgte eine Systematisierung und Interpretation der herausgearbeiteten regelhaften Phänomenstrukturen und Deutungsmuster.

6 Analyse und Ergebnisse: Diskursivierung von Data Scientists zwischen Kontrast und Konsens

Die empirische Analyse ist in drei Abschnitte gegliedert. Der erste Abschnitt befasst sich mit der diskursiven Konstruktion der Fachexpertise, der zweite Abschnitt mit den sozialfigurativen Elementen sowie der zugeschriebenen

Arbeitsrolle und der dritte Abschnitt mit der postulierten Ethik der Data Scientists. Die Reihenfolge der Analyseabschnitte korrespondiert mit einer Zunahme an Kontrasten zwischen den Konstruktionsweisen des akademischen und ökonomischen Feldes. Während bei der Benennung der Fachexpertise mit Hinblick auf die Phänomenstruktur lediglich feine Unterschiede zutage treten, bestehen auf der ethischen Ebene grobe Unterschiede. In allen drei Bereichen stellen wir jedoch auf einer epistemischen bzw. doxischen Ebene gemeinsame Deutungsmuster fest. Abschließend erfolgt eine Systematisierung der verschiedenen Befunde, die verdeutlicht, dass externe Diskursivierung von Data Scientists einer ambigen sozialen Logik folgt.

6.1 Data Scientists als GrenzgängerInnen

Die Eingrenzung der Wissenskompetenz von Data Scientists ist einer der zentralen Themenkomplexe, um den Aussagen in den untersuchten Dokumenten kreisen. Im akademischen Feld lässt sich diesbezüglich eine Klassifikation von Data Scientists als wissenschaftlich breit geschulten Multitalenten finden. Der akademischen diskursiven Praxis nach sind Data Scientists in der Lage, die Vielzahl von Aufgaben, die in der professionellen Datenanalyse anfallen, eigenständig und auf innovative Art und Weise zu lösen. In allen drei Status-Schichten des akademischen Feldes wird das Data Science-Studium damit beworben, dass eine Vermittlung von tiefer Wissensexpertise für eine Breite von Anforderungsprofilen erfolgt. Data Science, so lässt es sich aus den beworbenen Studiengangsinhalten und den Lehrplänen rekonstruieren, verbinde die Lehre von technischen, statistischen, kommunikativen und betriebswirtschaftlichen Kenntnissen:

„Students master subjects from computer science, statistics, and management such as regression, web scraping, SQL and NoSQL database management, natural language processing, business communications, machine learning, cluster analysis, application development, and interviewing skills.” (University of San Francisco 2019)

Als technische Fähigkeiten werden immer wieder ‚data cleaning‘, ‚programming‘ und ‚data visualization‘ in den Lehrplänen der Studiengänge erwähnt. Zusätzlich zu diesen praktischen ‚skills‘ umfassen die statistischen Fähigkeiten, die vermittelt werden sollen, die Anwendung und Entwicklung statistischer Modelle. Häufig müssen daher, um in die Studiengänge aufgenommen zu werden, Grundkenntnisse im Umgang mit Programmiersprachen und Statistik vorgewiesen werden. Das Studium selbst

zielt also auf der einen Seite auf die Kombinierung von Programmierung mit statistischen Analysemethoden. Dies soll angehenden Data Scientists ermöglichen, vor allem zweierlei Aufgabenbereiche innovativ zu verknüpfen: (1) große Datenmengen aufzubereiten und systematisch nach Mustern zu durchforsten sowie (2) mittels ‚machine learning‘, ‚deep learning‘ und der Programmierung von Algorithmen Produkte bzw. Programme aus den hergestellten Erkenntnissen zu generieren.

Um diese Kombinierung von Programmier- mit Statistikkenntnissen in der Datenarbeit sinnvoll einsetzen zu können, sollen auf der anderen Seite neben den technischen und statistischen Fähigkeiten auch ökonomisches und kommunikatives Wissen vermittelt werden – etwa in Kursen, die ‚management skills‘ oder das ‚business thinking‘ der Studierenden trainieren. Data Scientists sollen so in die Lage versetzt werden, Probleme nicht nur technisch rigoros anzugehen, sondern in ihrer Arbeit betriebswirtschaftliche Vorgaben oder Ziele nicht aus den Augen zu verlieren. Data Scientists sollen also nach ihrer Ausbildung, so versprechen und verkünden es zumindest die Universitäten, als *multipolare AkademikerInnen* auftreten können.⁵

Die Form der Multipolarität, mit der eine Singularisierung der Fachexpertise angestrebt wird, korrespondiert mit der Institutionalisierung von Data Science innerhalb der Universität als Organisation. Data Science ist in der Regel keine eigenständige Disziplin, sondern existiert eingebettet in bestehende Institute und Fakultäten. Anders als in den klassischen ‚interdisziplinären Projekten‘, in denen sich häufig bestehende Disziplinen als Selbstzweck oder um spezielle inhaltliche Fragen zusammenschließen (Abbott 2001), widmet sich Data Science keinem speziellen Problem, und etablierte Disziplinen versuchen, ihre Eigenständigkeit trotz der Emergenz von Data Science zu bewahren (z. B. Donoho 2015). An den meisten Universitäten ist Data Science innerhalb der Institute der Statistik, der Informatik (Computer Science) oder der Ingenieurwissenschaften angesiedelt. Damit knüpft Data Science an die abwechslungsreiche Geschichte sowohl der US-amerikanischen Universitäten als auch der Statistik selbst an (Ben-David 1971). Es existieren bisher nur an einigen mittel- oder höher rangierenden Universitäten eigene Data Science-Institute (siehe Columbia University 2019a) bzw. Data Science Schools (siehe University of Virginia 2019c).

⁵ Den Terminus ‚Akademiker‘ ziehen wir als deutsches Pendant zum US-amerikanischen Terminus ‚professional‘ heran (Freidson 1986). Als Akademikerin ist in diesem Sinne eine universitär geschulte Wissensarbeiterin gemeint, nicht aber eine Wissenschaftlerin oder ‚academic‘.

Diese mitunter schwache Form der Institutionalisierung wird als Stärke des Faches gerahmt. In den Studiengangsbeschreibungen präsentiert sich ein Narrativ, welches eine produktive Verbindung zwischen Data Science als Querschnittsprogramm und der datenweltlichen Durchdringung sämtlicher Bereiche der sozialen Welt proklamiert:

„Our faculty represents the fundamental multidisciplinary nature of the big data industry.” (University of San Francisco 2019)

„Our core courses draw on expertise and involve faculty from different disciplines across Duke. By doing so, they reflect the multiple quantitative disciplines that contribute skill sets to data science.” (Duke University 2019a)

Data Science durchziehe die Fakultäten in gleichem Maße, wie Daten die Wirtschaft und Gesellschaft durchziehen, lautet die homologisierende Narration – eben, weil die Kurse größtenteils von bestehenden Fakultäten angeboten werden. Der Anspruch an Data Scientists ist somit ein anderer als an die ausbildenden Institutionen und Wissenschaftler selbst. Somit kann auch Wirtschaftsnähe als ein Verkaufsargument eines Studiums dienen, ohne Anlass für Verlustängste wissenschaftlicher Autonomie zu bieten.⁶ Gleichzeitig ist die Bewerbung von Data Science als Querschnittsprogramm wohl auf organisationsökonomische Aspekte zurückzuführen; es dürfte in vielerlei Hinsicht unproblematischer sein, ein Programm in bestehende Disziplinen einzugliedern, statt ihm eigene Organisationseinheit zukommen zu lassen. Diese Positionierung weist für Data Science als Disziplin freilich das Risiko auf, eine fachliche Selbstverwaltung nicht zu erreichen, sondern in organisationaler Abhängigkeit zu verbleiben. Außer Frage steht jedoch, dass die Positionierung von Data Science als Querschnittsdisziplin einen konfliktärmeren Eintritt in das akademische Feld ermöglicht hat.

Durch die oszillierende Diskursaktivität zwischen Disziplinen entsteht also ein Bild von Data Scientists als multipolaren AkademikerInnen. Data Scientists wird die Rolle zugeschrieben, disparate Wissenssysteme ‚mastern‘ und kombinieren zu können. Data Scientists sollen die Arbeitswelt dadurch bereichern, dass sie eine methodisch versierte Verbindung bestehender Fachexpertisen betreiben, so das Mantra. Die Konstruktion von Data Scientists

erfolgt gewissermaßen in Abgrenzung zum ‚Fachidiot‘ oder ‚Nerd‘, obwohl sie die für diese Bezeichnungen charakteristische Kompetenz in der Arbeit mit Computern (siehe Kendall 1999) auch für sich beanspruchen.

In den Stellenanzeigen, bei denen in der Regel kein spezielles Data Science-Studium vorausgesetzt wird, sondern lediglich ein Studium in einem quantitativen Fach, finden wir eine etwas anders gelagerte Diskursivierung von Data Scientists. Hier dominiert anstelle des Bildes von Data Scientists als multipolaren AkademikerInnen jenes einer *Schnittstellenprofession*. Data Scientists werden bei den etablierten Techunternehmen wie auch den Start-ups als eine Art Mediatoren konstruiert, die sich dadurch auszeichnen, praktische Vermittlungs- und Übersetzungsarbeit innerhalb und zwischen Arbeitsgruppen leisten zu können. Den Stellenanzeigen zufolge sollen Data Scientists primär über die berufliche Eignung verfügen, mit verschiedenen Abteilungen und Berufsgruppen – und somit auch zwischen unterschiedlichen Akteuren, Konzepten oder Zielen – zusammenarbeiten und zwischen ihnen vermitteln zu können. Es wird das Bild von Data Scientists als ‚broker‘ bzw. einer Brücke konturiert:

„Our team drives the cross-company strategy for healthcare and life sciences, and applies Microsoft’s deep technical assets on improving healthcare and health equity around the world. We focus on impacting technology and the industry in the 3–10 year horizon, and act as a *bridge* [eigene Hervorhebung] between Microsoft Research and product teams, and have landed numerous technical innovations into shipping products.” (Microsoft 2019a)

„Work with cross-functional partners in product, engineering, design, and operations to achieve business goals.” (InstaCart 2019a)

Data Scientists werden also von den Unternehmen als Schnittstellenprofession zwischen individuellen und organisationalen Akteuren konstruiert. Diese zugeschriebene Rolle geht einerseits mit Anforderungen im Betrieb einher, andererseits ist sie zugleich Anerkennung und Legitimation einer Verantwortung, die Data Scientists zuteilwerden soll.⁷

Diese Figuration der Arbeitsrolle schließt zwar intuitiv an den Befund von Data Scientists als multipolaren AkademikerInnen an, weicht jedoch in ihrer sozialen Form ab.

⁶ Womöglich manifestiert sich mit der Wirtschaftsnähe in dem akademischen Diskurs auch der Umstand, dass das US-amerikanische Hochschulsystem stark ökonomisiert ist. Entsprechende Befunden beziehen sich auf die Hochschulverwaltung (z. B. Espeland & Sauder 2007), wohingegen die wissenschaftlichen Disziplinen eine streckenweise kontraproduktive Opposition zur Ökonomisierung ihrer Arbeit einnehmen (Owen-Smith & Powell 2001).

⁷ Im nächsten Abschnitt wird deutlich, dass sich trotz dieser Schnittstellenpositionierung nicht automatisch personelle Führungsverantwortung für Data Scientists ergibt. Womöglich dürfte eine Schnittstellenprofession ohne Führungsverantwortung für ein gewisses Maß an Anomie in bestehenden Organisationsstrukturen und -logiken sorgen.

Es wird in den Stellenanzeigen weniger auf die Bündelung von Wissen auf einer höheren Ebene referiert, sondern vielmehr auf die Funktion des vermittelnden Mitarbeiters innerhalb eines datenökonomischen Projekts gepocht. Die Stellenanzeigen küren Data Scientists als solche Akteure, die Verbindungen herstellen und zwischen heterogenen TeammitgliederInnen ‚brokern‘. Während das akademische Feld Data Scientists ein multipolares Kompetenzprofil zuschreibt und als Akteure positioniert, die scholastisch in mehreren Arenen des Wissens verankert sind, führt das ökonomische Feld die Rolle von Data Scientists als praktischen VermittlerInnen ins Feld, die Probleme vielleicht nicht eigenständig lösen können oder in ‚sauberer‘ wissenschaftlicher Manier abarbeiten (Evans 2010), es dafür aber schaffen, grenzgängerisch Brücken zwischen spezialisierten Akteuren bzw. Experten zu schlagen.

Ob Data Scientists nun in der Praxis das Generalistentum tatsächlich meistern können oder doch eher eine Berufsgruppe darstellen, die ‚von allem ein bisschen kann‘ und wertvolle Mediatisierungsarbeit leistet, fällt aus dem Rahmen dieser diskursanalytischen Untersuchung. Was wir jedoch diskursanalytisch zeigen – indem wir die Befunde der feldspezifischen Phänomenbeschreibung interpretieren und abstrahieren –, ist eine gemeinsame basale Wissensstruktur, die auf der Deutung des *Grenzgängertums* beruht. Data Scientists werden gewissermaßen als ‚marginal man‘ (Park 1928: 881) diskursiviert, als GrenzgängerInnen, die durch die Überschreitung bestehender Disziplinen und Organisationsrollen – mitsamt ihren behüteten Methoden, Expertisen und Verantwortungsbereichen – zur Weiterentwicklung des Wissens bzw. Profits beitragen. Im akademischen Feld bewegt sich die Data Scientist zwischen abgetrennten Fachgebieten hin und her; hier wird die Entdifferenzierung gewissermaßen in das Arbeitssubjekt selbst hineinverlagert: Der Data Scientist ‚meistert‘ und kombiniert eigenständig verschiedene Fachexpertisen und entwickelt ein Querschnittsdenken. In der Konstruktion als Schnittstellenprofession im ökonomischen Feld ist die Data Scientist eher die Brücke; doch auch hier wird sie als Grenzgängerin gedeutet, indem ihr die Expertise zugeschrieben wird, im Kollektiv von Arbeitssubjekten Übersetzungen und Entdifferenzierungen herstellen zu können. So wie die Grenzgängerin letztlich dank ihrer multiplen Erfahrungswelt eine eigene Innovationsader verspricht (Bargatzky 1981: 155), wird nun die Data Scientist von Wissenschaft und Wirtschaft als solches Arbeitssubjekt konstruiert, das Wissensgebiete oder arbeitsrelevante Projektglieder organisch verkoppeln könne, um so zu neuen und kreativen Lösungen für Probleme zu gelangen. Die grenzgängerische Fähigkeit, die Data Scientists von beiden Feldern zugeschrieben wird,

ist also, temporäre Entdifferenzierungen zwischen heterogenen Entitäten innerhalb von Arbeitsprozessen herstellen zu können.

6.2 Data Scientists als Anti-Nerds

Ein weiterer zentraler Aussagenkomplex in der untersuchten diskursiven Praxis betraf die Einordnung von Data Scientists in die Arbeitswelt sowie damit verbundene sozialfigurative Charakteristika. Hier herrschte auf der Phänomenebene größerer Dissens als zuvor bei den Kenntnissen und Aufgaben. So werden Data Scientists in Studiengangsbeschreibungen über die Statusgrenzen des akademischen Feldes hinweg als ‚leader‘ tituiert, wobei die Führungsrolle hier in Bezug auf Expertise zu verstehen ist und nicht in Bezug auf Personen. Die University of Virginia beschreibt ihr Studienangebot etwa damit, dass „[n]ow more than ever, the field of data science demands leaders“ (University of Virginia 2019b). Die Duke University verspricht, mit ihrem Programm gar „a new type of quantitative thought leader“ (Duke University 2019a) zu kultivieren.

Nun ist sicherlich zu erwarten, dass renommierte Universitäten damit werben, Führungskräfte zu produzieren; allerdings verweist der Terminus ‚leader‘ unseres Erachtens auf eine etwas andere Sozialfigur. Der Begriff ist nicht mit dem gängigen Verständnis von Führungskräften gleichzusetzen, sondern zielt vielmehr auf die Bestimmung einer neuen Art von Führung – eine Führung, die weniger auf der Disziplinarmacht beruht (Foucault 2014), sondern kollaborativ agiert und Autorität dementsprechend informeller und subtiler ausübt. So führt die Duke University im Anschluss an die Ausrufung eines ‚thought leaders‘ aus: „Duke’s Master in Interdisciplinary Data Science combines rigorous computational and technical training with field knowledge and repeated practice in critical thinking, teamwork, communication, and collaborative leadership to generate data scientists who can add value to any field“ (Duke University 2019a). Letztlich bleibt in den diskursiven Praktiken der Universitäten nebulös, ob mit der Rolle des ‚leader‘ etwa auch Personalverantwortung einhergehen muss (womöglich wird der ‚leader‘-Begriff als Update des Manager-Begriffes ins Spiel gebracht, der aufgrund einer inflationären Verwendung zunehmend in Verruf gerät). Es ist aber feststellbar, dass eine Verbindung zwischen ‚leadership‘ und sozialen Kompetenzen wie etwa ‚communication‘ oder ‚collaboration‘ im Diskursnetz gezogen wird.

Insbesondere die mittel- und höher prestigehaften Universitäten bewerben ihre Studiengänge mit verpflicht-

tenden Kursen, die soziale Kompetenzen trainieren sollen. Um erfolgreich in die Rolle des Data Scientists schlüpfen zu können, benötigt es ‚hard skills‘ und ‚soft skills‘, so der diskursive Kodex. An den weniger renommierten Universitäten steht die Vermittlung von sozialen Kompetenzen seltener in den Curricula. Das liegt gegebenenfalls daran, dass die Studiengangsprogramme häufig nur online erfolgen und keine Präsenz vor Ort erfordern. Mit der diskursiven Anrufung von Data Scientists als kommunikativen Führungspersönlichkeiten kann also zeitgleich eine Distinktion der renommierten von den weniger angesehenen Instituten im akademischen Feld festgestellt werden – eine Distinktion, die sich wohl bis zu nichtuniversitären Anbietern von Data Science-Lernprogrammen, wie etwa *datacamp*, erstreckt. Die Phänomenbeschreibung von Data Scientists als ‚leader‘ liegt dementsprechend nicht einheitlich über dem Feld – hier wird der Diskurs entlang der Positionen im Statusgefüge gebrochen.

Das zumindest bei mittel- und höher prestigehaften Universitäten verbreitete Bild vom Data Scientist als ‚leader‘ ist in den Stellenausschreibungen der Digitalkonzerne kaum wiederzufinden. Stattdessen lässt sich ein Bild ausmachen, das auf einer anderweitig ‚engineered culture‘ (Kunda 1992) aufbaut, dessen Elemente zwar keinen Antagonismus zu der universitär konstruierten Rolle zeigen, allerdings in Konturen einer differenten Sozialfigur münden. Anstelle vom ‚leader‘ ist in den Stellenausschreibungen viel häufiger vom ‚team player‘ die Rede, dem nicht Führungskompetenzen, sondern ein spielerischer Habitus anhaftet. Data Scientists werden im ökonomischen Feld als passionierte Arbeitssubjekte und Entdecker in der datafizierte Welt konstruiert. Das Ziel des Spiels ist das Finden von Mustern und Regelmäßigkeiten in großen Datenmengen. Dieses anspruchsvolle Ziel soll jedoch nicht innerhalb eines Terrains von Einzelkämpfern verfolgt werden, sondern im Kontext eines ‚teams‘ und harmonischen Habitats. Die Arbeit des Data Scientists mag harte Arbeit beinhalten und technische Fähigkeiten voraussetzen, sie ist dafür kollaborativ und ‚spaßig‘ zugleich, so das Mantra:

„We’re working hard, having fun, and making history. Come join our team! You will have an enormous opportunity to impact the customer experience, design, architecture, and implementation of a cutting edge product used every day by people you know.“ (Amazon 2019c)

„Our mix of thoughtful, inventive and neighborly employees work together to deliver our common goal, to make grocery shopping effortless, and give valuable time back to our customers. We believe that just as meals are best shared together, success is best shared together. If this excites you, then Instacart just might be the place for you. Welcome home.“ (InstaCart 2019a)

Hier wird also nicht nur, wie für immer mehr Berufsfelder üblich, ein Kreativitätsversprechen artikuliert (Reckwitz 2012: 133), sondern auch ein affektueller Schwur an den Teamgeist geleistet. Die Diskursaktivitäten versprechen eine spielerische Arbeitssituation mit der Aussicht auf eine transformative Arbeit innerhalb eines kollaborativen – manchmal gar familiären – Teams. Besonders junge Techunternehmen betonen diesen Gedanken. Data Scientists sind, so die diskursive Anrufung bzw. die diskursive ‚engineered culture‘, keine ‚unternehmerische Selbste‘ (Bröckling 2007), sondern passionierte TeamspielerInnen. Auf der untersuchten Job-Plattform ‚glassdoor‘ etwa werden Führungskompetenzen in der Regel nur für Ausschreibungen sogenannter ‚Senior Data Scientists‘ verlangt. In anderen Worten: Das subjektivierende Moment des ökonomischen Diskursfeldes ist, anstelle der individualisierenden Assoziierung von Data Scientists als ‚thought leader‘, ein kollektivierendes.

Auf einer tieferen Ebene hegen diese zwei Rollenzuschreibungen, die augenscheinlich aneinander vorbeireden, ein gemeinsames epistemisches Profil.⁸ So bauen beide Phänomenstrukturen auf dem Deutungsmuster auf, dass Data Scientists eine ausgeprägte Sozialität besitzen sollen. Ob als ‚leader‘ oder kollegiale TeamspielerIn, beide Rollen setzen kommunikatives Talent voraus. Data Scientists, so wird in beiden Diskurssträngen impliziert, sollten kommunikativ sein, benötigen Empathie und eine Verbindung von intellektuellem und sozialem Gespräch.

Aus diesem Befund leiten wir die Vermutung ab, dass in beiden Feldern eine Distinktion des Data Scientists von benachbarten Sozialfiguren wie den StatistikerInnen oder den InformatikerInnen erfolgt; jenen zwei etablierten Berufen, die als zentrale Konkurrenten um die Jurisdiktion der Datenarbeit gelten dürften, auf die es Data Scientists abgesehen haben. Die konkurrierenden Berufsgruppen, die seit den 1980er-Jahren das Bild des Nerds als solches geprägt haben, werden mit hohem technischen und digitalen Kapital, dafür aber wenigen sozialen Kompetenzen in Verbindung gebracht (Kendall 1999). Data Scientists hingegen werden von beiden Feldern als selbstaktive Anti-Nerds diskursiviert, als Akteure mit einer Libido für Austausch und Interaktion; Data Scientists müssen dem Diskurs nach nicht mehr disziplinarisch geführt werden, sie werden vielmehr als Verantwortungsträger und Initiator ins Spiel gebracht.

Im Anschluss an Foucault kann diesbezüglich von einem Übergang von disziplinarischen zu selbstkontrol-

⁸ Damit will nicht gesagt sein, dass die Ebene der Deutungsmuster die ‚wahre‘ oder ‚objektivere‘ Ebene darstellt.

lierenden Macht- und Subjektformen gesprochen werden (Deleuze 1992). Mit diesem Übergang wird keine Verpuffung von Machtverhältnissen postuliert, wohl aber eine Veränderung ihrer Form, ihrer Wirkungsweisen und Richtungen. Als Anti-Nerds sind Data Scientists zwar durchaus in Kontakt zum Metadiskurs der Quantifizierung (Mau 2017: 29) – jenem Diskurs, dem auch Statistiker- und InformatikerInnen nahestehen –, verkörpern diesen jedoch in neuer Manier. Denn, ob nun als ‚leader‘ oder ‚team player‘, beides klassifiziert ein hochkommunikatives Arbeits-subjekt und lässt die Rolle als ‚number crunchers‘ oder ‚code monkey‘ hinter sich: „you do more than just crunch numbers. You work with Engineers, Sales Associates and Marketing teams“ (Google 2019a). Die zwei unterschiedlichen Subjektbildnisse hält also die eigensinnige Identität (Schnell 2012) von Data Scientists als sozialkompetente und selbstaktive Anti-Nerds zusammen.

6.3 Data Scientists als WeltverbesserInnen

Der dritte zentrale Aussagenkomplex in der diskursiven Praxis des akademischen und ökonomischen Feldes betrifft die postulierte gesellschaftliche Bedeutung von Data Science sowie die Ethik der Data Scientists. In den Studiengangsbeschreibungen findet sich das Narrativ, dass die Welt mit gigantischen Mengen an Daten konfrontiert sei, wobei nun das Data Science-Studium die Werkzeuge bereitstelle, um diese Flut an Daten zu bewältigen: „The world is being transformed by data and data-driven analysis is rapidly becoming an integral part of science and society“ (Stanford University 2019b). Die Datafizierung begründet eine Notwendigkeit von Data Science. Dies wird weitergehend aufgegriffen, indem die Studiengangsprogramme Data Scientists als solche AnalystInnen positionieren, die mittels der richtigen Ausbildung jene neue Datenwelt nicht nur meistern können, sondern durch ihre Arbeit sogar ‚Gutes‘ zu leisten vermögen:

„Our rigorous and forward-thinking online MSDS program shows students how to make sense of the vast amounts of data available now, and turn insights into action. We believe in the good that data science can do. Our students go on to make monumental strides in discovery, analysis, and problem-solving in diverse fields.“ (University of Virginia 2019b)

Ethik ist ein hinlänglich bekanntes Element der Entwicklung von Berufsgruppen (Abbott 1983). Eine Schlüsselkompetenz, die die Studiengangsbeschreibungen der mittel- und hoch prestigehaften Universitäten in ihren Ausbildungscurricula immer wieder betonen, ist die einer *Dataethik*. Die Erhebung, Aufbereitung und Auswertung von

Datenmengen müssen von ethischen Reflexionen begleitet werden, so das Mantra. Daher sind spezielle Ethikkurse verpflichtend – so etwa das Seminar ‚Critical Thinking in Data Science‘ an der Harvard University. Der Kurs soll lehren, welchen „wide-ranging impact data science has on the world and how to think critically about issues of fairness, privacy, ethics, and bias while building algorithms and predictive models that get deployed in the form of products, policy and scientific research“ (Harvard University 2019a).

Es ist anzumerken (wie auch bei den ‚communication skills‘), dass Ethik nur an mittel- und hoch prestigehaften Universitäten als Bestandteil in den Lehrprogrammen gesetzt ist. Insbesondere in den nur online erfolgenden Data Science-Studienprogrammen steht Ethik in der Regel nicht in den Curricula.⁹ Steht Datenethik auf dem Stundenplan, wird insbesondere die Möglichkeit von ‚bias production‘ oder Diskriminierung thematisiert, welche von Datenarbeit ausgehen kann: „Some pressing issues include: the ways in which using data can subtly exacerbate existing systemic prejudices, such as through implicit algorithmic bias“ (Georgetown University 2019a). Häufig werden in den codierten Sinnabschnitten Algorithmen als ‚black boxes‘ beschrieben, die (ungewollt) Diskriminierung auf Grundlage quantitativer Indikatoren (re-)produzieren würden.

Mit der Datenethik positionieren die dominanten Universitäten Data Scientists als eine reflexive Berufs- bzw. Professionsgruppe, die Rücksicht auf die Auswirkungen von Datenarbeit auf marginalisierte soziale Gruppen nehmen soll. Mit dieser Debatte knüpft Data Science an Professionen allgemein an (Abbott 1983). Hervorzuheben ist hier, dass die Rolle von ‚bias production‘ hauptsächlich in Bezug auf ethnische Gruppen ohne Bezug auf genuin ökonomische Dimensionen gedacht wird.¹⁰ Womöglich

⁹ Es ist dementsprechend auch hier eine Tendenz zur hierarchischen Differenzierung des akademischen Feldes und des Berufsfeldes Data Science zu vermuten, denn, indem prestigehafte Universitäten Data Scientists mit zusätzlichen Kompetenzen ausbilden, produzieren sie nicht nur eine Differenzierung des Berufsfeldes Data Science in reflexive und weniger reflexive Data Scientists, sondern reproduzieren zugleich ihre eigene dominante Stellung innerhalb des akademischen Feldes. Die genaue Entfaltung dieser Heterogenität müsste jedoch mittels eines eigenen Forschungsdesigns empirisch überprüft werden. Die Beobachtung von Variation in Konformität ist allerdings von dem Feld der Anwaltskanzleien bekannt, gerade im Hinblick auf moralisch sensible Themen (Phillips et al. 2013; Phillips & Zuckerman 2001).

¹⁰ Ohne Zweifel hegen etwa ethnische Diskriminierungen auch eine ökonomische Ungleichheitsdimension. Nichtsdestotrotz ist zu kon-

können hiermit feldspezifische Anzeichen einer sogenannten ‚Identitätspolitik‘ (Fraser 2009: 74) und damit eines neuen Legitimationsmusters des Kapitalismus (Boltanski & Chiapello 2013) konstatiert werden. Anstelle einer postulierten ‚Gemeinwohlorientierung‘ als strategischer Ideologie und als ‚professional project‘ (Larson 1977) erscheint ‚Gruppenwohlorientierung‘ die Wahl der Stunde, um einem Beruf ein Professionsethos zu verleihen.

Wenn wir uns jedoch den Stellenanzeigen zuwenden, so ist weder bei den etablierten Techunternehmen noch bei den Start-ups die Rede von einem ethischen Umgang mit Daten. Die diskursive Praxis innerhalb des ökonomischen Feldes kippt in ein anderes Bild. In den Stellenanzeigen werden Daten nicht als potenziell unheilvolles Medium betrachtet, sondern als uneingeschränkt positive Ressource für Wirtschaft und Gesellschaft gesehen. Die vermeintlich allgegenwärtigen Daten werden als Rohstoff narrativiert, den es zu nutzen gilt: „As a Data Science Program Manager, you will utilize the power of Google data, machine learning, and artificial intelligence, to deliver insights at scale that drive both long term strategy and daily insights“ (Google 2019a). Es offenbart sich in den Stellenanzeigen anstelle einer Datenethik eine *Verwertungsethik*.

Doch die Arbeit mit Daten soll nicht nur ökonomischen Segen für Techkonzerne einbringen, sondern auch hier die Welt ein Stück weit verbessern, so die Bewerbung der Data Science-Stellen. Microsoft wirbt etwa damit, dass Data Scientists, als professionelle Übersetzer in einem datenökonomischen Projekt, kein geringeres Problem als die Krebsforschung voranbringen können:

„In your role as Senior Data Scientist, you will work on real-world problems with customers to create new meaning out of patient data. For example, one of your major projects will be to collaborate with a small cross-functional team of researchers and engineers to use NLP [= Natural Language Processing (Anmerkung der Autoren)] to derive new meaning out of the cancer journey.“ (Microsoft 2019a)

Während sich die Phänomenstrukturen der Datenethik und Verwertungsethik fast schon inkommensurabel gegenüberstehen, ist mit einer Spielart des Solutionismus (Morozov 2013), nämlich eines *Datensolutionismus*, ein gemeinsames Glaubenssystem implizit vorhanden. Es genügt eine Menge an Daten und eine hungrige Data Scientist, und schon kann die Welt zu einer besseren gemacht

statieren, dass die Reflexionsimperative ökonomische Kategorien nicht direkt adressieren. Dass also Algorithmen auch dezidiert ökonomische Deprivation registrieren und verschlimmern können (Eubanks 2018), wird nicht thematisiert.

werden – so die Lösung in beiden Feldern. Die diskursive Kopplung zwischen großen Datenmengen und dem gesellschaftlichen Potenzial der Datenauswertung findet sich im ökonomischen Feld besonders ausgeprägt bei den etablierten Konzernen. Amazon etwa betont: „[o]ur petabytes of data has the ability to improve the shopping experience for hundreds of millions of consumers worldwide [...] Those of us who love to work with data see this as the pinnacle of opportunities that you cannot find anywhere else in the world“ (Amazon 2019b). Es lässt sich daraus schließen, dass der datenkulturelle Solutionismus – der Glaube an die ökonomische und soziale Kraft der Daten – von den großen Technologiekonzernen zu vereinnahmen versucht wird. Diese implizieren, dass etwa nur ‚petabytes of data‘ wirklich wirksame Arbeit eines Data Scientists zuließen. Das ökonomische Feld stellt gewissermaßen eine Korrelation zwischen Weltverbesserungsmöglichkeiten und der Größe von Datensammlungen diskursiv her.¹¹

Beide Felder propagieren also die Möglichkeit der Weltverbesserung mittels Datenarbeit. Dem Kontrast zwischen dem Postulat einer notwendigen kritischen Reflexion der Arbeit mit Daten einerseits und dem unhinterfragten Aufruf zur Auswertung von großen Datenmengen andererseits liegt ein gemeinsamer basaler Glaube zugrunde. Im untersuchten Diskursfeld werden Daten von beiden Seiten als natürliche Trägerinnen von Macht und gesellschaftlicher Bedeutung angesehen. Auch die Studiengangsbeschreibungen, die einen kritischen Umgang mit Daten von Studierenden einfordern, implizieren schließlich eine hohe soziale Relevanz von Daten. Nicht nur kapitalistisches, auch akademisches Handeln bedarf Rechtfertigung (Abbott 2001; Collins 1998; Espeland & Sauder 2007). Daten wird gewissermaßen eine übermenschliche Kraft zugesprochen. Die Rolle, die Data Scientists einhellig zugewiesen wird, ist die einer verwaltenden Profession für die ausgerufene, neuartige Produktivkraft. Es wird eine „illusio“ für BerufseinsteigerInnen kreiert, die in dem Glauben mündet, neben der ökonomischen Wertschöpfung auch eine gesellschaftlich sinnvolle Tätigkeit auszuüben. Von beiden Feldern wird auch hier die Professionalisierung vorangetrieben. Denn

¹¹ Aus feldanalytischer Sicht ist des Weiteren zu vermerken, dass Akteure, die die natürliche Relevanz von Daten hervorheben, die dominanten Akteure im Status quo begünstigen. Denn wer an die Macht der Daten glaubt, der erkennt in der Regel das symbolische Kapital nicht im Datum, sondern in Datenmengen bzw. möglichst großen Datenmengen. Im Diskursfeld Data Science findet somit eine symbolische Konsekration der großen Datenkonzerne statt, besitzen diese doch die größten Schätze an Heiligtümern, an deren magische Fähigkeiten alle Akteure in doxischer Manier glauben.

wenn wir eine Profession als eine solche Berufsgruppe begreifen, die es (mitunter) vermag, sich als gesellschaftlich relevante sowie altruistisch veranlagte Berufsgruppe darzustellen (Abbott 1988; Freidson 1988), dann wird dies von Wissenschaft und Ökonomie auf widerspruchshafte, aber nicht widerspruchsvolle Art und Weise erreicht.

6.4 Die diskursive Logik der Ambiguität als Professionalisierungsstrategie

Was ist denn nun ein Data Scientist? Die Studiengangsbeschreibungen und Stellenanzeigen sagen, jeweils im Widerspruch zueinander, multipolare AkademikerInnen und Schnittstellenprofession, ‚thought leader‘ und ‚team player‘ sowie DatenethikerIn und VerwertungsethikerIn. Unsere Diskursanalyse sagt mit Hinblick auf die Ebene der Deutungsmuster: Grenzgängerin, die sich von Nerds abgrenzt und für Weltverbesserung einsetzt.

Aus der generellen Analyse einer Grammatik aus Differenz und basalem Konsens zwischen den feldspezifischen diskursiven Praxen leiten wir unseren Kernbefund dieser Untersuchung ab, der lautet, dass eine soziale Logik der Ambiguität die Konstruktion von Data Scientists in Wissenschaft und Wirtschaft strukturiert. Während auf der Ebene der Phänomenstruktur Differenzen zwischen den Feldern bei der Klassifikation von Data Scientists zutage treten, finden sich auf einer tieferen Ebene gemeinsame Deutungsmuster. Es ist jedoch nicht davon auszugehen, dass eine detaillierte Koordination zwischen dem akademischen und ökonomischen Feld die ambige diskursive Konstruktion von Data Scientists bestimmt. Im Anschluss an Foucault begreifen wir die Ambiguität im Diskurs vielmehr als ‚Strategie ohne Strategie[n]‘ (Foucault 1978: 132).

Die Metapher des Kippbildes veranschaulicht die ambige diskursive Konstruktion von Data Science. Kippbilder, vor allem der psychoanalytischen Forschung bekannt, bergen zwei unterschiedliche Figuren in ein und demselben Bild. Sie zeigen etwa Abbilder von zwei verschiedenen Tieren, wie das von Wittgenstein vielfach aufgegriffene Bild einer Ente oder eines Hasen (Wittgenstein 2001: 1025), in einem Bild. Kippbilder lassen also unterschiedliche subjektive Wahrnehmungen von Phänomenen im Rahmen eines diffusen Phänomens zu (Gregory 2000); keine der beiden möglichen Figuren ist die eine richtige oder wahre Figur.

Eine solche Figuration tritt unseres Erachtens auch auf der gesellschaftlichen Ebene auf, in der diskursiven Formation von Data Scientists in Wissenschaft und Wirtschaft. In ihr wirken multiple und widersprüchliche Konstruktionen und Imaginationen, ohne dass dabei eine tiefere Grundlage

bezweifelt werden muss. Das kippbildhafte Subjektbildnis von Data Scientists kann nun dahingehend als Machtstrategie gewertet werden, dass insbesondere in Zeiten tiefgreifenden Wandels, etwa durch Urbanisierung (Abbott 1988) oder eben Digitalisierungsprozesse, eine ambige Konstruktionsweise von sozialen Phänomenen mit Etablierungs- und Diffusionsvorteilen einhergehen dürfte. Data Scientists erscheinen als geeignete Profession für die ‚Überleitung‘ hin zu einer digitalen Gesellschaft, in der Mustererkennung innerhalb von ‚big data‘ einen immer wichtiger werdenden Funktionsbereich darstellt (Nassehi 2019). Das Diskursfeld Data Science erfährt Vorteile gegenüber konkurrierenden Berufs- und Professionsgruppen, indem es differente ‚Data Science-Imaginationen‘ (Saner 2019) über sich selbst aufnimmt und so eine breite gesellschaftliche Anschlussfähigkeit herstellt. Auch Diskurse beeinflussen also Digitalisierungsprozesse (siehe Büchner 2018). Der Diskurs um Data Science findet die Balance: Mittels der ambigen Konstruktion können verschiedene Erwartungshaltungen angesprochen werden, ohne dabei eine ambivalent-zerrissene Sozialfigur abzugeben. Die diagnostizierten Deutungsweisen von Data Scientists integrieren konträre Phänomenklassifikationen und entwerfen ein ambiges Subjekt, einen produktiven ‚habitus clivé‘ (Bourdieu betrachtete die Möglichkeit eines gespaltenen Habitus als Ausnahme und tendenziell als Belastung für die jeweiligen TrägerInnen). Jene ambige Subjektivierung verspricht, mehrere und widersprüchliche ‚cultural holes‘ sowie ‚structural holes‘ (Pachucki & Breiger 2010) in Wissenschaft, Wirtschaft und Gesellschaft überbrücken zu können.

Ambiguität stellt also eine spezifische Form kulturellen Kapitals dar; sie wird durch ein Bespielen mehrerer Feldlogiken möglich und scheint eine Distinktion der Data Scientists von anderen Berufs- und Professionsgruppen zu erlauben. Ambiguität begreifen wir als (mit-)verantwortlich für den bisherigen Professionalisierungsprozess von Data Scientists sowie fördernd für weitere Professionalisierungsanstrengungen. Bemerkenswert ist hierbei die bisherige Zurückhaltung von data scientists im Institutionalisierungsprozess, welche Fragen für die Professionssoziologie aufwirft, beispielsweise zu dem von Atzeni (2016) für Ärzte herausgearbeitetem professionellen Erwartungsmanagements in der Wechselwirkung von Fremd- und Selbstbeschreibung. Dass Ambiguität generell als Strategie fungieren kann, schließt an bisherige Befunde an (siehe etwa Bauer 2011; Leifer 1988; Suckert 2018). Die Politikwissenschaften etwa konstatieren eine ‚strategic ambiguity‘ in diskursiven politischen Aushandlungsprozessen (S. Davenport & Leitch 2005). In der Kulturosoziologie wird im Anschluss an die ‚feinen Unterschiede‘ heute von ‚kulturellen Allesfressern‘ als Leitfi-

guren des Konsums gesprochen (Peterson & Kern 1996). Andreas Reckwitz scheint die gesellschaftliche Bedeutung der Ambiguität zu betonen, wenn er vom ‚hybriden Subjekt‘ als dominanter Subjektform der Moderne spricht und diese in der Spätmoderne auf eine paradoxe Heirat von Romantik und Bürgerlichkeit zurückführt (Reckwitz 2006). Wir haben aufgezeigt, wie Ambiguität im Kontext der Digitalisierung eine Strategie zur Professionalisierung darstellt. Ambiguität strukturiert die Subjektform der objektivierenden digitalen DatenarbeiterInnen.

7 Schluss

In den Studiengangbeschreibungen und Stellenausschreibungen aus dem US-amerikanischen Raum besteht nicht nur Dissens, sondern auch basaler Konsens um die Klassifikation und Deutung von Data Scientists. Während auf der Ebene der Phänomenstruktur unterschiedliche Beschreibungen bestehen – die wir als eigenständige soziale Tatsachen herausgearbeitet haben –, lassen sich auf der Ebene der Deutungsmuster feldübergreifende Imaginationen über Data Scientists herauschälen. Data Scientists werden, zusammengefasst, als GrenzgängerInnen, Anti-Nerds und WeltverbesserInnen gedeutet. Im Anschluss an den soziologischen Forschungsstand kann diese ambige Konstruktion von Data Scientists als professionsartige Diskursivierung gewertet werden. So impliziert die Deutung von Data Scientists als GrenzgängerInnen eine wissensbasierte Autonomie – eine zentrale Voraussetzung für den Weg zur Profession (Abbott 1988; Freidson 1988). Die Deutung von Data Scientists als Anti-Nerds kann als Monopolisierungsstrategie (Larson 1977) eines aufstrebenden Berufsfeldes verstanden werden, das sich von anderen Berufsgruppen distinguert, die ebenfalls um die Jurisdiktion der professionellen Datenarbeit konkurrieren. Das Deutungsmuster der WeltverbesserInnen lässt sich schließlich als Form eines gemeinwohlorientierten Ethos verstehen (Abbott 1983; Merton 1982). Alle drei basalen Sinnmuster können somit als Grundlagen für Kompetenzdarstellungskompetenzen (Pfadenhauer 2003) betrachtet werden. Diskurstheoretisch ist anzunehmen, dass sich die ambige, diskursiv zugeschriebene Kompetenzdarstellungskompetenz auf gewisse Weise performativ in die Praxis der Data Scientists einschreiben wird. Inwieweit und auf welche Art und Weise dies geschieht, ist eine empirische Frage. Des Weiteren ist anzumerken, dass Professionalisierung auch nach Anfangserfolgen kein Selbstläufer ist (Mieg 2018). Die Entwicklung von Data Science hin zu einer Profession ist zwar in vollem Gange,

eine staatliche Konsekration, in Form von Autonomiezusicherungen etwa, steht aber noch aus. Unser Befund der Ambiguitätsstrategie weist, unter Berücksichtigung ihrer erfolgreichen Rolle bei der Entwicklung von Ärzten (Atzeni 2016), jedoch darauf hin, dass eine diskursive Grundlage gegeben ist, welche die Data Science-Professionals bei ihren Institutionalisierungsbemühungen befördert.

Aktuell bietet sich also die seltene Chance, das Entstehen einer neuen Profession im Prozess mitzuverfolgen. Anschließende Forschungen sind notwendig, um die Institutionalisierung von Data Science weiter zu ergründen. Insbesondere ethnografische Forschungen scheinen angebracht, um zu untersuchen, inwiefern und auf welche Weise die Diskursivierungen durch Wissenschaft und Wirtschaft auf die ‚data science communities‘ einwirken. Hier wäre auch zu ergründen, inwiefern die diskursive Subjektivierung bzw. Habitualisierung sich in körperlichen Dispositionen niederschlägt. Weitergehend scheint uns relevant, unsere Befunde aus dem US-amerikanischen Raum mit den Konstruktionsweisen von Data Scientists in anderen nationalen Kontexten zu vergleichen. Quantitative Erhebungen empfehlen sich, um zu erfahren, welche sozialen Hintergründe Data Scientists in der Regel aufweisen. Aus Bourdieu'scher Perspektive wäre insbesondere von Interesse, welche Kapitalformen Data Scientists mit sich bringen. Erst kürzlich konnten Friedman und Laurison aufzeigen, dass ‚vererbte‘ Formen kulturellen Kapitals höhere Renditen in solchen Berufsfeldern abwerfen, in denen die Regeln des Spiels eher diffus sind (Friedman & Laurison 2019: 133). Im Anschluss an diese Untersuchung wirft dies die Frage auf, inwiefern die Beherrschung von kulturellen Codes der höheren Schichten auch ein Vorteil im diskursiv-ambigen Professionsfeld Data Science ist. Nicht zuletzt könnte die strategische Rolle von Ambiguität in Bezug auf weitere digitale und analoge Phänomene inspiziert werden.

Literatur

- Abbott, A., 1983: Professional Ethics. *American Journal of Sociology* 88: 855–885.
- Abbott, A., 1988: *The System of Professions*. Chicago: University of Chicago Press.
- Abbott, A., 2001: *Chaos of Disciplines*. Chicago: University of Chicago Press.
- Abbott, A., 2005: Linked Ecologies: States and Universities as Environments for Professions. *Sociological Theory* 23: 245–274.
- Atzeni, G., 2016: *Professionelles Erwartungsmanagement: Zur soziologischen Bedeutung der Sozialfigur Arzt*. Baden-Baden: Nomos.

- Bargatzky, T., 1981: Das „Marginal Man“-Konzept: Ein Überblick. *Sociologus* 141–166.
- Bauer, T., 2011: *Die Kultur der Ambiguität*. Berlin: Insel.
- Ben-David, J., 1971: *The Scientist's Role in Society: A Comparative Study*. Englewood Cliffs NJ: Prentice-Hall.
- Boes, A., T. Kämpf, B. Langes & T. Lühr, 2015: Landnahme im Informationsraum. Neukonstituierung gesellschaftlicher Arbeit in der „digitalen Gesellschaft“. *WSI-Mitteilungen* 68: 77–85.
- Boltanski, L. & É. Chiapello, 2013: *Der Neue Geist des Kapitalismus*. Konstanz: UVK.
- Bosančić, S., F. Brodersen, L. Pfahl, L. Schürmann, T. Spieß & B. Traue, 2021: *Following the Subject. Studies in Subjectivation, Band 1: Grundlagen und Forschungspraxis*. Wiesbaden: Springer.
- Bourdieu, P., 2002: *Der Einzige und sein Eigenheim*. Hamburg: VSA.
- Bourdieu, P., 2004: *Der Staatsadel*. Köln: von Halem.
- Bourdieu, P., 2014: *Homo Academicus*. Frankfurt am Main: Suhrkamp.
- Bourdieu, P., 2015: *Sozialer Sinn. Kritik der theoretischen Vernunft*. Frankfurt am Main: Suhrkamp.
- Bourdieu, P. & L. Wacquant, 2013: *Reflexive Anthropologie*. Frankfurt am Main: Suhrkamp.
- Bröckling, U., 2007: *Das unternehmerische Selbst. Soziologie einer Subjektivierungsform*. Frankfurt am Main: Suhrkamp.
- Büchner, S., 2018: Zum Verhältnis von Digitalisierung und Organisation. *Zeitschrift für Soziologie* 47: 332–348.
- Callon, M., 1984: Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay. *The Sociological Review* 32: 196–233.
- Collins, H.M., 1998: The Meaning of Data: Open and Closed Evidential Cultures in the Search for Gravitational Waves. *American Journal of Sociology* 104: 293–338.
- Couldry, N. & U.A. Mejias, 2019: *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism*. Stanford: Stanford University Press.
- Davenport, S. & S. Leitch, 2005: Circuits of Power in Practice: Strategic Ambiguity as Delegation of Authority. *Organization Studies* 26: 1603–1623.
- Davenport, T.H. & D. Patil, 2012: *Data Scientist*. Harvard Business Review 90: 70–76.
- Deleuze, G., 1992: Postscript on the Societies of Control. *October* 59: 3–7.
- Diaz-Bone, R., 2002: *Kulturwelt, Diskurs und Lebensstil. Eine diskurstheoretische Erweiterung der bourdieuschen Distinktionstheorie*. Opladen: Leske + Budrich.
- Diaz-Bone, R., 2006: Zur Methodologisierung der Foucaultschen Diskursanalyse. *Historical Social Research/Historische Sozialforschung* 31(2): 243–274.
- Donoho, D., 2015 (18. September): 50 years of Data Science. Conference Paper, Princeton NJ.
- Duffy, B.E. & B. Schwartz, 2017: Digital „women's work“: Job Recruitment Ads and the Feminization of Social Media Employment. *New Media & Society* 20: 2972–2989.
- Espeland, W.N. & M. Sauder, 2007: Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology* 113: 1–40.
- Eubanks, V., 2018: *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.
- Evans, J.A., 2010: Industry Induces Academic Science to Know Less about More. *American Journal of Sociology* 116: 389–452.
- Eyal, G., 2013: For a Sociology of Expertise: The Social Origins of the Autism Epidemic. *American Journal of Sociology* 118: 863–907.
- Foucault, M., 1978: *Dispositive der Macht: Über Sexualität, Wissen und Wahrheit*. Leipzig: Merve.
- Foucault, M., 1981: *Archäologie des Wissens*. Frankfurt am Main: Suhrkamp.
- Foucault, M., 2003: *Wahnsinn und Gesellschaft: Eine Geschichte des Wahns im Zeitalter der Vernunft*. Frankfurt am Main: Suhrkamp.
- Foucault, M., 2014: *Überwachen und Strafen. Die Geburt des Gefängnisses*. Frankfurt am Main: Suhrkamp.
- Fraser, N., 2009: Social Justice in the Age of Identity Politics. S. 72–91 in: G.L. Henderson & M. Waterstone (Hrsg.), *Geographic Thought: A Praxis Perspective*. New York: Routledge.
- Freidson, E., 1986: *Professional Powers: A Study of the Institutionalization of Formal Knowledge*. Chicago: University of Chicago Press.
- Freidson, E., 1988: *Profession of Medicine*. Chicago: University of Chicago Press.
- Friedman, S. & D. Laurison, 2019: *The Class Ceiling: Why it Pays to be Privileged*. Bristol: Policy Press.
- González-Bailón, S., 2017: *Decoding the Social World: Data Science and the Unintended Consequences of Communication*. Cambridge: MIT Press.
- Graham, M. & M.A. Anwar, 2018: Digital Labour. S. 177–188 in: J. Ash, R. Kitchin & A. Leszczynski (Hrsg.), *Digital Geographies*. London: Sage.
- Gregory, R.L., 2000: Ambiguity of 'Ambiguity'. *Perception* 29: 1139–1142.
- Halatcheva-Trapp, M., 2018: *Elternschaft im Wechselspiel von Deutungsmustern und Diskurs*. Wiesbaden: Springer.
- Hammerbacher, J., 2009: Information Platforms and the Rise of the Data Scientist. S. 82–94 in: T. Segaran & J. Hammerbacher (Hrsg.), *Beautiful Data: The Stories Behind Elegant Data Solutions*. Sebastopol: O'Reilly Media.
- Kajetzke, L., 2008: *Wissen im Diskurs*. Wiesbaden: Springer.
- Keller, R., 2011a: *Wissenssoziologische Diskursanalyse: Grundlegung eines Forschungsprogramms*. Wiesbaden: VS.
- Keller, R., 2011b: *Diskursforschung. Eine Einführung für SozialwissenschaftlerInnen*. Wiesbaden: Springer.
- Kendall, L., 1999: Nerd Nation: Images of Nerds in US Popular Culture. *International Journal of Cultural Studies* 2: 260–283.
- Kunda, G., 1992: *Engineering Culture: Control and Commitment in a High-tech Corporation*. Philadelphia: Temple University Press.
- Larson, M., 1977: *The Rise of Professionalism: A Sociological Analysis*. Berkeley: University of California Press.
- Latour, B. & S. Woolgar, 1986: *Laboratory Life*. Princeton: Princeton University Press.
- Leifer, E.M., 1988: Interaction Preludes to Role Setting: Exploratory Local Action. *American Sociological Review* 53: 865–878.
- Liu, S. & M. Emirbayer, 2016: Field and Ecology. *Sociological Theory* 34: 62–79.
- Maeße, J. & J. Hamann, 2016: Die Universität als Dispositiv. *Zeitschrift für Diskursforschung* 1: 29–50.
- Manske, A. & C. Schnell, 2010: *Arbeitsmarkt und Beschäftigung: Arbeit und Beschäftigung in der Kultur- und Kreativwirtschaft*.

- S. 699–727 in: F. Böhle, G.G. Voß & G. Wachtler (Hrsg.), *Handbuch Arbeitssoziologie*. Wiesbaden: Springer.
- Marwick, A.E., 2013: *Status Update: Celebrity, Publicity and Branding in the Social Media Age*. New Haven: Yale University Press.
- Mau, S., 2017: *Das metrische Wir: Über die Quantifizierung des Sozialen*. Berlin: Suhrkamp.
- Merton, R.K., 1982: *Social Research and the Practicing Professions*. Cambridge: Abt Books.
- Mieg, H.A., 2018: *Professionalisierung*. Potsdam: Fachhochschule Potsdam.
- Mützel, S., P. Saner & M. Unternährer, 2018: *Schöne Daten! Konstruktion und Verarbeitung von digitalen Daten*. S. 111–132 in: D. Houben & B. Prietl (Hrsg.), *Datengesellschaft. Über die Datafizierung des Sozialen*. Bielefeld: Transcript.
- Muzio, D., D. Hodgson, J. Faulconbridge, J. Beaverstock & S. Hall, 2011: *Towards Corporate Professionalization: The case of Project Management, Management Consultancy and Executive Search*. *Current Sociology* 59: 443–464.
- Nassehi, A., 2019: *Muster: Theorie der digitalen Gesellschaft*. München: Beck.
- Noordegraaf, M., 2007: *From „Pure“ to „Hybrid“ Professionalism: Present-day Professionalism in Ambiguous Public Domains*. *Administration & Society* 39: 761–785.
- O’Neil, C., 2016: *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Broadway Books.
- O’Neil, C. & R. Schutt, 2013: *Doing Data Science: Straight Talk from the Frontline*. Sebastopol: O’Reilly.
- Owen-Smith, J. & W.W. Powell, 2001: *Careers and Contradictions: Faculty Responses to the Transformation of Knowledge and its Uses in the Life Sciences*, S. 109–140 in: S. Vallas (Hrsg.), *The Transformation of Work*. Bingley: Emerald.
- Pachucki, M.A. & R.L. Breiger, 2010: *Cultural Holes: Beyond Relationality in Social Networks and Culture*. *Annual Review of Sociology* 36: 205–224.
- Park, R.E., 1928: *Human Migration and the Marginal Man*. *American Journal of Sociology* 33: 881–893.
- Peterson, R.A. & R.M. Kern, 1996: *Changing Highbrow Taste: From Snob to Omnivore*. *American Sociological Review* 61: 900–907.
- Pfadenhauer, M., 2003: *Professionalität. Eine wissenssoziologische Rekonstruktion institutionalisierter Kompetenzdarstellungskompetenz*. Opladen: Leske + Budrich.
- Phillips, D.J., C.J. Turco & E.W. Zuckerman, 2013: *Betrayal as Market Barrier: Identity-Based Limits to Diversification among High-Status Corporate Law Firms*. *American Journal of Sociology* 118: 1023–1054.
- Phillips, D.J. & E.W. Zuckerman, 2001: *Middle-status Conformity: Theoretical Restatement and Empirical Demonstration in Two Markets*. *American Journal of Sociology* 107: 379–429.
- Reckwitz, A., 2006: *Das hybride Subjekt. Eine Theorie der Subjektkulturen von der bürgerlichen Moderne zur Postmoderne*. Weilerswist: Velbrück.
- Reckwitz, A., 2012: *Die Erfindung der Kreativität: Zum Prozess gesellschaftlicher Ästhetisierung*. Frankfurt am Main: Suhrkamp.
- Reckwitz, A., 2015: *Subjekt*. Bielefeld: Transcript.
- Reckwitz, A., 2019: *Das Ende der Illusionen. Politik, Ökonomie und Kultur in der Spätmoderne*. Berlin: Suhrkamp.
- Ribes, D., 2019: *STS Meet DataScience, Once Again*. *Science, Technology & Human Values* 44: 514–539.
- Salganik, M., 2017: *Bit by Bit: Social Research in the Digital Age*. Princeton NJ: Princeton University Press.
- Saner, P., 2019: *Envisioning Higher Education: How Imagining the Future Shapes the Implementation of a New Field in Higher Education*. *Swiss Journal of Sociology* 45: 359–381.
- Schiller, D., 2000: *Digital Capitalism: Networking the Global Market System*. Cambridge: MIT Press.
- Schmidt-Wellenburg, C., 2009: *Die neoliberale Gouvernementalität des Unternehmens-Management und Managementberatung zu Beginn des 21. Jahrhunderts*. *Zeitschrift für Soziologie* 38: 320–341.
- Schmidt-Wellenburg, C., 2014: *Der Aufstieg der Beratung zur transnationalen Regierungsform im Feld des Managements*. *Berliner Journal für Soziologie* 24: 227–255.
- Schnell, C., 2012: *Eigensinnige Professionalität? Zur Bedeutung „beruflicher Identität“ im Kontext von Subjektivierung*. *ALS-Studien* 5: 21–34.
- Saab, P., 2019: *Digitaler Kapitalismus*. Berlin: Suhrkamp.
- Startup-Monitor, 2018: <https://deutchestartups.org/wp-content/uploads/2019/05/Deutscher-Startup-Monitor-2018.pdf>.
- Stichweh, R., 1984: *Zur Entstehung des modernen Systems wissenschaftlicher Disziplinen: Physik in Deutschland 1740–1890*. Frankfurt am Main: Suhrkamp.
- Stichweh, R. (1996). *Professionen in einer funktional differenzierten Gesellschaft*. Revidierte Fassung, ursprünglich erschienen S. 49–59 in: A. Combe & W. Helsper (Hrsg.), *Pädagogische Professionalität*. Frankfurt am Main: Suhrkamp.
- Stichweh, R., 2000: *Professionen im System der modernen Gesellschaft*. S. 29–38 in: R. Merten (Hrsg.), *Systemtheorie Sozialer Arbeit*. Wiesbaden: Springer.
- Suckert, L., 2018: *Unravelling Ambivalence: A Field-theoretical Approach to Moralised Markets*. *Current Sociology* 66: 682–703.
- Suddaby, R., Y. Gendron & H. Lam, 2009: *The Organizational Context of Professionalism in Accounting*. *Accounting, Organizations and Society* 34: 409–427.
- Wajcman, J., 2018: *How Silicon Valley sets time*. *New Media & Society* 21: 1272–1289.
- Weber, M., 2010: *Die protestantische Ethik und der Geist des Kapitalismus*. München: Beck.
- Wetterer, A., 2002: *Arbeitsteilung und Geschlechterkonstruktion. „Gender at Work“ in theoretischer und historischer Perspektive*. Köln: von Halem.
- Wittgenstein, L., 2001: *Philosophische Untersuchungen*. Kritisch-genetische Edition. Frankfurt am Main: Suhrkamp.
- Zuboff, S., 2019: *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.

Autoren

Robert Dorschel

PhD Candidate, Department of Sociology
University of Cambridge
16 Mill Lane, Cambridge, CB2 1SB, UK
E-Mail: rcd49@cam.ac.uk
Telefon: 01577-9034716

Robert Dorschel ist PhD Candidate an der University of Cambridge. Er hat Soziologie und Politikwissenschaften in Kiel, Berlin und Durham (USA) studiert. Forschungsschwerpunkte: Arbeits-, Professions- und Wirtschaftssoziologie sowie soziologische Theorie. In der Dissertationsschrift wird die Subjektivierung von hochqualifizierten Beschäftigten in der Digitalwirtschaft (sogenannten ‚tech worker‘) untersucht. Die wissenssoziologische Masterarbeit über Data Scientists wurde von der Deutschen Gesellschaft für Soziologie als herausragende Abschlussarbeit ausgezeichnet.

Philipp Brandt

Assistant professor of sociology – SciencesPo
Chercheur – CSO
19 rue Amélie – 75007 Paris
E-Mail: philipp.brandt@sciencespo.fr
Telefon: (33)1 40 62 65 82

Philipp Brandt ist assistant professor of sociology an der Sciences Po Universität und Chercheur am Centre de Sociologie des Organisations. Forschungsschwerpunkte: Professions-, Wirtschafts-, und Organisationssoziologie sowie Methodologie und Wissenstheorie. Analysen zu sozialen Beziehungen in der Fertigungsindustrie sind im Socio-Economic Review und Economic Development Quarterly erschienen (mit J. Whitford & A. Schrank), zu Predigern im frühmodernen England in Poetics (mit M. Hoffman et al.), und zu Abduktion in Computational Social Science in Sociological Science (mit S. Timmermans). Er arbeitet an einer Monographie zur Entstehung von Data Science.