

3 Methode

3.1 Die Untersuchung von Datafizierung und Algorithmen

Wie ich bei Earlybird beobachten konnte, gelten Daten immer schon als inhärent wertvoll. In der Praxis zeigt sich aber, dass diese Annahme sich eher als eine notwendige produktive Fiktion erweist: Daten sind keineswegs einfach so gegeben, sie müssen in Kooperation mit Nutzerinnen erzeugt werden. Zudem braucht es zahlreiche Bearbeitungs- und Verarbeitungsschritte, um den Daten ihre Geheimnisse zu entlocken, aus Daten produktive Ressourcen zu machen und sie in Geschäftsprozesse zu integrieren.

In meiner Arbeit frage ich, wie Unternehmen überhaupt an Daten kommen – d.h. wie sie Nutzerinnen dazu motivieren, sich innerhalb von Datafizierungsinfrastrukturen zu »verhalten«, um daraus Daten zu gewinnen. Zudem ergründe ich, durch welche Prozesse der Quantifizierung, Kategorisierung, Formalisierung und des Vergleichs aus diesen Daten profitgenerierende Ressourcen gemacht werden. Was braucht es, um mit Daten möglich zu machen, was sie versprechen? Welche Probleme sollen Daten und Algorithmen lösen? Wie werden diese Probleme benannt und reformuliert? Wie werden bestehende Algorithmen ausgewählt, angepasst und verworfen? Wie wird ein algorithmisches System zu einem Algorithmus, der als vermeintlich unzweifelhaftes, technisch-rationales Orakel in einer Blackbox verschwindet? Wie gewinnen findige Unternehmen aus den Verhaltensweisen ihrer Nutzerinnen Daten? Und wie machen sie aus diesen Daten wertvolle und (automatisch) umsetzbare Einsichten?

In der empirischen Untersuchung dieser Fragen stütze ich mich auf qualitative, ethnografische Methoden (O'Reilly, 2005; Mason, 2002; Emerson et al., 2011; Hine, 2017; Pink et al., 2016). Um zu verstehen, wie Daten entstehen und wie sie ihren Wert entfalten, kombiniere ich zwei Ansätze der »critical algorithm studies« (siehe Kitchin 2016 für einen Überblick verschiedener Ansätze), die an die Wissenschafts- und Technikforschung anschließen (Seaver, 2019). Zum einen ein »unpacking« soziotechnischer As-

semblagen der Datafizierung/Algorithmen, das durch informelle Interviews und teilnehmende Beobachtung in einem Unternehmen ermöglicht wurde. Und zweitens eine Untersuchung von Bildungsmaterial hinsichtlich der Frage, wie Empfehlungssysteme quantifizieren, kategorisieren, vergleichen und Relationen zwischen Nutzerinnen und Dingen herstellen.

Die Untersuchung von Datafizierungsprozessen und Algorithmen ist mit verschiedenen Zugangsproblemen konfrontiert (Kitchin, 2016). Datenbanken und Algorithmen sind oftmals in den Händen privater Unternehmen, die Algorithmen als proprietäre Güter betrachten, die es vor neugierigen Blicken der Konkurrenz oder der Öffentlichkeit zu schützen gilt. Dateninfrastrukturen sind Blackboxes: Was im Hintergrund passiert, während wir im Web shoppen oder uns von TikTok ablenken lassen, ist unserem Blick unzugänglich. Selbst wenn die Operationen der Datenbanken und Empfehlungsalgorithmen sichtbar wären, bräuchte es eine Menge an technischer Expertise, um zu verstehen, was vor sich geht. Es braucht also spezifische Berechtigungen und Kenntnisse, um Zugang zu erhalten (siehe auch: Burrell 2016). Dieses »blackboxing« fußt auf der problematischen Annahme, dass es sich bei Algorithmen um singuläre, technische Dinge handle, die von Expertinnen vollständig überblickt und beherrscht werden. Das ist gemäß Nick Seaver nur bedingt der Fall. Algorithmen sind in ihrem Kontext zu untersuchen: Er schlägt vor, Algorithmen als soziotechnische Systeme, »composed of collective human practices« zu konzeptualisieren (Seaver 2017, S. 5, auch: Seaver 2019).

If we care about the logic of these systems, we need to pay attention to more than the logic and control associated with singular algorithms. We need to examine the logic that guides the hands, picking certain algorithms rather than others, choosing particular representations of data, and translating ideas into code [...] My point is that when our object of interest is the algorithmic system, »cultural details *are* technical details – the tendencies of an engineering team are as significant as the tendencies of a sorting algorithm. (Seaver, 2019, S. 419, Hervorhebung im Original)

Algorithmen sind Teil von weitaus komplexeren, heterogenen soziotechnischen »Assemblagen« und in soziale wie technische Kontexte eingebettet (Kitchin, 2016). »Algorithmus« bezeichnet in diesem Sinne nicht nur eine technische Prozedur, sondern steht als Stellvertreter für ein ganzes soziotechnisches System (Gillespie, 2016), das in der Praxis viel »messier« (Ziewitz, 2016) ist als die Mythologisierung einer »power of the algorithm« (Beer, 2017) erah-

nen lässt. Algorithmische Systeme als soziotechnische Assemblagen bestehen aus zahlreichen, unterschiedlichen, ineinander gefalteten Verfahren, Datenquellen, Standards, Kategorien, etc. Die Folge ist, dass in vielen Fällen nicht einmal ihre Erzeugerinnen überblicken oder komplett verstehen können, wie bestimmte Resultate zustande kommen (Burrell, 2016).

Algorithms cannot be adequately studied as stand-alone processes if we are to start understanding the roles they now play. But they are also more than technical infrastructures – algorithms also need to be recognised more broadly as both situated artefacts and generative processes that engage in complex ways with their surrounding ecosystems. This is an ecosystem that involves technical – software, code, platforms and infrastructure – and human designs, intents, audiences and uses more broadly. (Willson, 2017, S. 141)

Für das »unpacking« dieser soziotechnischen Arrangements ist das Hauptproblem nicht, dass Algorithmen in Blackboxes verschwinden. Es geht ebenso um das allgemeinere Problem der Unsichtbarkeit und der fehlenden »Transparenz« einer Infrastruktur, die sich als vermeintliche Selbstverständlichkeit der Wahrnehmung weitgehend entzieht (Star, 1999). Krisen und Störungen, aber auch eine »methodische Befremdung« (Amann & Hirschauer, 1997), können zu einer »infrastructural inversion« beitragen, d.h. zu einer Aufmerksamkeit für die Arbeit, welche die Infrastruktur erzeugt und erhält (Bowker & Star 1999, siehe auch: Hine 2017, S. 23f.). Auch die in Infrastrukturen eingelassenen Praktiken der Quantifizierung, Kategorisierung, Formalisierung, Bewertung und Vergleiche zeichnen sich dadurch aus, dass sie bestimmte Dinge unsichtbar und andere dafür sichtbarer machen (Bowker & Star, 1999): Unsicherheiten, »messiness« (siehe Mützel et al. 2018), Selbstverständlichkeiten und Annahmen (Law, 2009) verschwinden hinter »pristine numerical output[s]« wie beispielsweise Credit-Scores (Fourcade & Healy, 2017a, S. 289).

Susan Leigh Star und Martha Lampland (2009) fordern dazu auf, auf die »Infrastruktur zu hören«. Für mich bedeutet das, den Daten während ihrer verschiedenen Transformationsschritte zu folgen und diesen Weg durch die Infrastruktur zu beschreiben. Welchen Akteuren begegnen die Daten, wie verändern sie sich in diesen Begegnungen, wie nehmen sie auf andere Einfluss und vermehren sich?

In den Kapiteln 5 sowie 6 bis 8 versuche ich Prozesse der Datafizierung zu rekonstruieren, indem ich Operationen der Quantifizierung, Kategorisierung, Formalisierung, Bewertung und des (algorithmischen) Vergleichs in ih-

ren technischen, sozialen, ökonomischen und unternehmerisch-praktischen Kontexten berücksichtige. Daraufhin untersuche ich, wie dies zur In-Wert-Setzung von Personendaten beiträgt. Es geht dabei um die alltäglichen und banalen Tätigkeiten, welche das vermeintlich magische Funktionieren algorithmischer Systeme ermöglichen. Darüber hinaus geht es um die Selbstverständlichkeiten, die (unhinterfragten) Annahmen und die manchmal dubiosen¹ Praktiken, welche in Datafizierungsinfrastrukturen eingefaltet werden. Wie bereits in Kapitel 2 angedeutet, sind Daten und Algorithmen nicht nur in verschiedene Kontexte eingebettet. Diese Einbettung ist auch dafür wichtig, wie Daten generiert werden und ihren Wert in der weiteren Anwendung entfalten können.

Ein solcher Fokus geht davon aus, dass Algorithmen nicht einfach auf »die Realität« angesetzt werden können, sondern dass Realität und algorithmische Funktionserfordernisse aufeinander abgestimmt werden müssen (Ziewitz, 2017). Daten und Algorithmen repräsentieren die Welt nicht einfach. Sie öffnen kein Fenster auf die soziale Welt. Daten und Algorithmen sind Teil soziotechnischer Arrangements, die soziales Handeln strukturieren (Marres 2017, S. 22). Algorithmische Technologien sind »interaktiv« (vgl. Hacking 2001 zum Verhältnis von Sozialforschung und sozialen Kategorien der Beobachtung; MacKenzie 2006 zum Verhältnis von Ökonomie und ökonomischer Theorie bzw. Modellierung): »When the monitoring and analysis of everyday activities is used as a basis for intervention into these activities, a complex set of exchanges between knowledge and behaviour is set in motion« (Marres, 2017, S. 9).

Datafizierung und Algorithmen sind produktiv: Sie produzieren Einheiten wie zum Beispiel »Konsumentinnen« (siehe Cluley & Brown 2015; Zwick & Denegri Knott 2009), mögliche Verhaltensweisen (Alaimo & Kallinikos, 2017) oder Dinge wie »Interessen« (Degeling, 2017). Zudem machen sie »latente« Relationen sichtbar und produzieren neue prospektive Relationen (siehe Kapitel 8). In Bezug auf diese Produktivität von Datafizierungsprozessen und Algorithmen verfolge ich eine zweite empirische Strategie. Darin geht es darum, die technisch-abstrakte Logik von Empfehlungssystemen zu erfassen (Kitchin, 2016; Burrell, 2016). Die Frage, wie Standard-Algorithmen wie Collaborative Filtering idealtypisch funktionieren, fußt auf der Untersuchung von Ausbildungsmaterial (ein Massive Open Online Course, Handbücher sowie

¹ Siehe zum Beispiel Crawford & Paglen (2019) zu Trainingsdatensätzen, die für Bilderkennungssoftware verwendet werden.

Übersichtsartikel zu Empfehlungssystemen) und einem theoretisch geleiteten Nachvollzug von »pseudo-code«. Das Ziel besteht darin, die technische Funktionsweise anhand des analytischen Vokabulars der Quantifizierungs-, Kategorisierungs- und Vergleichssoziologie aufzuschlüsseln, um zu zeigen, wie verschiedene Empfehlungssysteme (formale) Relationen zwischen Nutzerinnen und Dingen (Filme, Artikel, Produkte, etc.) herstellen.

3.2 Feldzugänge

Wie oben beschrieben ist die soziologische Untersuchung von Algorithmen durch mehrere »opacities« (Burrell, 2016) erschwert. Einerseits sind Algorithmen technische Konstrukte privater Unternehmen, welche die Funktionsweisen als Geschäftsgeheimnisse betrachten, die es vor den Blicken der Konkurrenz und möglicherweise auch vor Kritik zu schützen gilt (Kitchin, 2014). Andererseits braucht es zum Verständnis von Algorithmen eine »technical literacy« (Burrell, 2016), über die normalerweise nur Fachpersonen verfügen. Diese Art der Unzugänglichkeit wird durch Heterogenität und Verteiltheit verschärft: Algorithmische Systeme bestehen aus Konglomeraten verschiedener Algorithmen und anderer technischer Komponenten, welche von verschiedenen Akteuren erzeugt, betreut und betrieben werden (Seaver, 2019; Kitchin, 2016).

Mein Forschungsdesign ist dementsprechend als »multisited« (Burrell, 2009), digitale (Pink et al., 2016; Hine, 2017) Ethnografie konzipiert. Wie Marcus (2010) hervorhebt, kann sich die zeitgenössische Ethnografie kaum mehr um »das Ganze« kümmern, sondern muss strategisch aus den verschiedenen Teilen des Feldes selektieren. Angesichts der »Privatheit« und der technischen Unzugänglichkeit von Algorithmen, entschied ich mich für den Fokus auf zwei »fieldsites«. Der erste Zugang besteht darin, Daten und Algorithmen in der unternehmerischen Praxis zu untersuchen. Der zweite befasst sich mit einem dazu komplementären Zugang, welcher die gängigsten algorithmischen Empfehlungssysteme durch die pädagogische Vermittlung computerwissenschaftlicher Theorie und praktischer Anwendungsfälle erschließt. Angesichts der im Feld angetroffenen Praxis, Computerwissen-

schaftler in Ausbildung für algorithmische Projekte zu rekrutieren, ist dieser zweite Zugang angebracht.²

Earlybird und Earlybird Digital

Earlybird wurde 2005 gegründet. Das Geschäftsmodell basierte schon damals darauf, die Zielgruppe der Jugendlichen mit verschiedenen Partnerunternehmen zusammenzubringen, d.h. Jugendlichen gegen eine kleine Mitgliedschaftsgebühr Rabatte bei diesen Unternehmen zu verschaffen. Einige Jahre später ging Earlybird Partnerschaften mit mehreren Banken ein und wurde damit zum Beziehungsbroker zwischen Jugendlichen, Banken und Partnerunternehmen. Die Banken erhalten ein attraktives Programm, um Neukundinnen zu akquirieren. Die Partnerunternehmen erhalten Zugang zur Zielgruppe der Jugendlichen. Die Jugendlichen erhalten Vergünstigungen. Seither konnte Earlybird noch mehr Banken als Kundinnen gewinnen und seinen »Membern« attraktive »Deals« bei über 600 Partnerunternehmen sichern. Zum Zeitpunkt meiner Feldforschung waren mehr als 200 000 aktive Member bei Earlybird registriert und 15 Angestellte arbeiteten im Unternehmen.

Earlybird ist mit mehreren anderen Firmen verbunden. Am wichtigsten ist die 2015 gegründete Tochterfirma Earlybird Digital, die mit der Programmierung des administrativen Systems, der Internetseite und der Mobile App beauftragt ist. Zum Zeitpunkt meiner Forschung waren dort 14 Personen beschäftigt (8 davon als Programmierer und Programmiererinnen in Kosovo). Der wichtigste Kunde von Earlybird Digital ist Earlybird. Die Tochterfirma offeriert ihre Dienstleistungen (Frontend- und Backendprogrammierung, Data-Science-Dienstleistungen) aber auch anderen Kundinnen.

Zwischen Oktober 2016 und April 2017 war ich als bezahlter »Forschungspraktikant« bei Earlybird und Earlybird Digital angestellt.³ Bei der Aushandlung des Zugangs war es für den operativen Geschäftsführer von Earlybird wichtig, dass ich meine soziologische Expertise im Unternehmen einbringe,

² Eine dritte empirische »fieldsite« war der Besuch von vier Konferenzen der Organisation »Swiss Cognitive«. Diese versucht, Technologien der künstlichen Intelligenz in Unternehmen zu etablieren. Dieses Material habe ich nicht ausgewertet.

³ Nach Abschluss der Feldforschung im engen Sinn folgten weitere Treffen, informelle Gespräche und der gemeinsame Besuch einer Data-Science-Konferenz.

so wie auch andere Praktikanten und Mitarbeiter ihre unterschiedlichen Fähigkeiten einbringen sollen (fast zeitgleich mit mir begann eine Psychologiestudentin mit Interesse für Neuropsychologie ein Praktikum). Im digitalen Tochterunternehmen waren zwei Absolventen der Fachhochschule Informatik angestellt, die sich mit der Datenanalyse und der Programmierung der App beschäftigen sollten. Beide wurden von ihrem Professor für Informatik mit Schwerpunkt künstliche Intelligenz und Informationssicherheit vermittelt. Dieser sitzt als Miteigentümer im Verwaltungsrat des Tochterunternehmens und steuert strategische Beratung bei.

In meiner Rolle als Forschungspraktikant arbeitete ich an verschiedenen Projekten mit. Insbesondere an der Planung, Durchführung und Auswertung eines »Marktforschungstags«, an welchem 25 Jugendliche in Fokusgruppen zu verschiedenen Themen befragt wurden. Als Forschungspraktikant genoss ich aber auch die Möglichkeit, an zahlreichen Sitzungen und Treffen teilzunehmen, welche nicht unmittelbar mit meinen eigenen Aufgaben zu tun hatten. Das Hauptmaterial meiner Forschung besteht aus Sitzungs- und Gesprächsprotokollen. Die Beobachtung der individuellen Arbeit steht im Hintergrund. Mein wichtigster Informant war Nik, der Geschäftsführer von Earlybird Digital, den ich an verschiedene Sitzungen begleiten konnte und der mir in unzähligen informellen Gesprächen geduldig Auskunft gab.

Die Auswahl von Earlybird und Earlybird Digital war zum einen eine Frage des Zugangs: Nur wenige Firmen sind gewillt, ihre Tätigkeiten von externen Personen beobachten zu lassen. Earlybird eignete sich zum anderen aber auch deshalb, weil sich das Unternehmen zum Zeitpunkt meiner Forschung in einem Entwicklungsprozess befand, in dem es eine weitreichende »Digitalisierungsstrategie« verfolgte, um das Geschäftsmodell für die Zukunft abzusichern. Andere Firmen, in denen ich explorative Interviews durchgeführt hatte (eine Versicherung und ein grosser Retailer), waren diesbezüglich bereits weiter fortgeschritten, in ihren Datafizierungspraktiken zudem unzugänglicher und hielten sich bedeckter.⁴ Earlybird befand sich in einem Prozess der Aushandlung, Etablierung und Institutionalisierung von Datafizierungspraktiken, der in den anderen Firmen bereits weiter

4 Siehe auch die Studie von Sami Coll (2015) zu Kundentreueprogrammen der großen Schweizer Retailer. Darin beobachtet er, dass die Möglichkeiten der Datenanalyse größer wären als die tatsächlichen, und für die Kundinnen sichtbaren, Anwendungsfälle. Gemäß Coll wolle keines dieser Unternehmen vorpreschen und negative Öffentlichkeit auf ihre Beobachtungs- und Datafizierungspraktiken lenken.

fortgeschritten war. Das eröffnet für den ethnografischen Zugang die Chancen, Aushandlungs- und Institutionalisierungsprozesse sowie das Ausmerzen von Kontingenz *in situ* mitzuverfolgen.

Vorgehen und Analyse

Beobachtungs- und Gesprächsprotokolle, handschriftliche Notizen, analytische Memos, Dokumente und Fotos bilden mein Datenkorpus (Emerson et al., 2011). Eine Schwierigkeit der ethnografischen Vorgehensweise besteht darin, dass im Feld immer mehr vor sich geht als sich beobachten und beschreiben lässt – so auch in meiner Feldforschung bei Earlybird. Die vielen verschiedenen Projekte und Teilprojekte, die gleichzeitig von verschiedenen Personen verfolgt wurden, erforderten es, schon während der Feldforschung analytische, aus meinem Forschungsinteresse begründete, Fokuspunkte zu setzen: Praktiken der Datafizierung, der Kategorisierung von Personen und der Auswertung von Personendaten. Im Sinne eines »iterativ-induktiven« Vorgehens – d.h. einer wechselseitigen Beeinflussung und Abwechslung der Forschungsphasen von Datensammlung, Analyse und ethnografischem »writing up« (O'Reilly, 2005) – legte ich schon während der Feldforschung die beobachterische und analytische Aufmerksamkeit auf einen Komplex von Ereignissen rund um das sogenannte »Datenexperiment« (siehe Kapitel 5). Dieser Fokus rechtfertigte sich aus dem Forschungsinteresse für Datafizierungspraktiken und aus den Relevanzen des Feldes selbst (O'Reilly, 2005).⁵

Eine entscheidende Einsicht, die sich bereits während der Feldforschung einstellte, war, dass Earlybird gemäß eigener Einschätzung nicht über ausreichend Personendaten verfügte und damit kämpfte, solche Daten zu erzeugen. Das »Datenexperiment« erwies sich als thematischer Kristallisierungspunkt, der verschiedene Bemühungen, Daten zu erzeugen, Member zu datengenerierenden Verhaltensweisen zu motivieren und diese Daten irgend-

⁵ Ein solches Vorgehen kann sich als übereilt herausstellen, wenn ein solcher Fokus nicht den Relevanzen der Teilnehmerinnen im Feld entspricht. Ein für mich ausschlaggebener Punkt, der diese Entscheidung rechtfertigte, war eine Entscheidung der Geschäftsleitung, die nach einer Sitzung zum Thema Datenanalyse gefällt wurde. Darin wurde die strategische und praktische Relevanz von Daten diskutiert. Im Anschluss beschloss die Geschäftsleitung, die Sitzung für abwesende Mitarbeiterinnen zu wiederholen, um bei allen ein strategisches Bewusstsein für die Relevanz von Personendaten zu verankern.

wie nutzbar zu machen, für mich, aber vor allem auch für Earlybird, sichtbar machte.

Nach Abschluss der Erhebungsphase erfolgte eine thematisch orientierte Selektion des Materials (Emerson et al., 2011) rund um das Datenexperiment. Dies umfasste Materialien, welche das Datenexperiment beschreiben, aber auch solche, die das Experiment im Kontext von Earlybirds Businessmodell und Datafizierungspraktiken verständlich machen. Daraus resultierte ein stark reduziertes und übersichtlicheres Korpus, das mithilfe theoretischer Begrifflichkeiten und emischer Konzepte und Metaphern aufgeschlüsselt wurde. Im Zentrum stand dabei die beschreibende Rekonstruktion des Datenexperiments in seinem unternehmerischen Kontext über die Zeit. Das resultierende Narrativ findet sich kondensiert in Kapitel 5 und bildet den Ausgangspunkt für eine analytisch-theoretische Aufarbeitung in den Kapiteln 2 sowie 6 bis 8.

In der Darstellung der Ergebnisse aus der Feldforschung versuchte ich, Beschreibung und Analyse so gut wie möglich auseinanderzuhalten. Da Beschreibung und Analyse aber keine trennscharfen Konzepte sind (O'Reilly, 2005) gelingt das nicht immer gleich gut. Die eher beschreibenden, für die Leserin aufbereiteten Passagen sind jeweils umrahmt und die Fussnoten dazu befinden sich am Ende der Box.

Generalisierung und Limitationen

Es ist klar, dass die Datenpraktiken eines verhältnismäßig kleinen und lokal gebundenen Unternehmens wie Earlybird nur begrenzt generalisierungsfähig sind. Der stetige Vergleich mit größeren und bedeutenderen Unternehmen der globalen digitalen Ökonomie mag deshalb stellenweise unangebracht wirken. Die von mir entwickelte theoretische Figur der »Momente der Datafizierung« beansprucht nicht, ein allgemeingültiges Modell der Datafizierung zu sein. Es ist ein analytisches Werkzeug, mit dem sich Fragen an die Phänomene der digitalen Ökonomie sowie deren Generierung und In-Wert-Setzung von Daten stellen lassen. Zum Beispiel: Wie motivieren digitale Unternehmen Nutzerinnen dazu, ihre Dienstleistungen zu verwenden (Kapitel 6)?⁶ Wie machen Unternehmen Nutzerinnen und ihre ei-

⁶ Auch wenn »fehlende Daten« für Unternehmen wie Google vermeintlich kein Problem darstellen, lässt sich trotzdem die Frage stellen (und auch tatsächlich empirisch beobachten), wie Google Nutzerinnen zu motivieren versucht, mehr Daten zu hinterlassen: Zum Beispiel durch Aufforderungen, Apps zu bewerten, auf Google Maps Fragen zu

genwilligen Verhaltensweisen kommensurabel (siehe zum Beispiel Paßmann & Gerlitz 2014 zu Twitter) beziehungsweise zu verrechenbaren und vergleichbaren Einheiten (Kapitel 7)? Oder: Welche Prämissen und Weltvorstellungen werden bei der Verwertung von Personendaten algorithmisch eingefaltet? Wie werden Nutzerinnen und Dinge zueinander in Beziehung gesetzt, um neue, profitable Relationen zwischen diesen Einheiten abzuleiten (Kapitel 8)?

»Introduction to Recommender Systems«

An Earlybirds Datafizierungspraktiken wird deutlich, dass das Unternehmen sich immer auch auf autoritatives Wissen und Praktiken der computerwissenschaftlichen Gemeinschaft bezieht. Der Senior Data Scientist sowie die beiden Junior Data Scientists waren während meiner Feldforschungsphase selbst in wissenschaftlicher und Lehrfunktion an der lokalen Fachhochschule tätig. Ein komplementärer empirischer Zugang entstand aus der Feldforschung bei Earlybird Digital (und auch aus meiner »technical illiteracy«). Dani, Junior Data Scientist bei Earlybird Digital, befasste sich in seiner Abschlussarbeit mit der Entwicklung eines Empfehlungssystems für Earlybird. Da es bei Earlybird noch kein produktives Empfehlungssystem gab – aber die Logik ihrer Arbeit um die Entwicklung und Implementierung solcher Systeme kreiste –, verwies mich Dani auf computerwissenschaftliche Literatur und den Kurs »Introduction to Recommender Systems« der University of Minnesota, den er selbst auch besuchte. Dieser wird auf der Lernplattform Coursera angeboten und von Michael Ekstrand und Joseph Konstan durchgeführt. Er vermittelt ein Grundverständnis der technischen Funktionsweise von Empfehlungssystemen und wurde von den Computerwissenschaftlern John Riedl und Joseph Konstan entwickelt. Beide sind für ihre wissenschaftliche und unternehmerische Pionierarbeit zu Empfehlungssystemen bekannt – insbesondere im Bereich des Collaborative Filtering (Cohn,

»locations« zu beantworten oder dem Google Newsfeed schlechte Empfehlungen zurückzumelden. Oder: Weshalb schickt uns Facebook Emails mit dem Hinweis, dass ich eine Nachricht erhalten habe, ohne mir die Nachricht direkt anzuzeigen? Oder stärker auf Geschenke bezogen: Weshalb versuchte Facebook, Partnerschaften mit indischen Mobilfunkbetreibern einzugehen und Kundinnen ohne »mobile data« Zugang zu Facebook zu verschaffen (Bhatia, 2016)?

2019). In verschiedenen Kursmodulen demonstrieren sie die Funktionsweisen und Hintergrundannahmen unterschiedlicher Empfehlungssysteme. In dazugehörigen Aufgaben werden die Teilnehmerinnen dazu aufgefordert, die jeweiligen Berechnungen und algorithmischen Operationen »von Hand«, d.h. mithilfe von Tabellenkalkulationsprogrammen selbst durchzuführen.

Jenna Burrell (2016) zeigt in ihrer Arbeit zu Machine-Learning-Algorithmen, dass Bildungsangebote wie Massive Open Online Courses, Handbücher und computerwissenschaftliche Artikel zwar keinen Zugang zu den komplexen Algorithmen von Google und anderen erlauben, aber einen Einblick in die grundsätzliche Logik der technischen Funktionsweise von Algorithmen geben (siehe auch: Kitchin 2016). Wie Mirko Schäfer und Karin Van Es (2017) hervorheben, geht es nicht darum, Algorithmen und deren mathematische Grundlagen bis ins letzte Detail zu verstehen, sondern darum, deren grundlegende Logik zu verstehen, um neue analytische (und kritische) Perspektiven auf algorithmische Phänomene zu ermöglichen.

Hier geht es mir darum, die algorithmischen Funktionsweisen zu verstehen, um die Logiken der Relationierung der unterschiedlichen Empfehlungssysteme sichtbar zu machen: Wie setzen verschiedene Typen von Empfehlungssystemen Nutzerinnen und Dinge zueinander in Beziehung, um in Form von Empfehlungen neue Relationen zwischen Nutzerinnen und Dingen herzustellen?

In einem ersten Schritt arbeitete ich die verschiedenen Kursmodule aus einer immanenten Teilnehmerperspektive durch. Danach sammelte und ordnete ich die Materialien des Kurses und seines nahen Verweisungshorizonts (i.e. computerwissenschaftliche Literatur). In einem dritten Schritt ging ich mit Hilfe eines analytischen Rasters des Vergleichs (Heintz, 2010, 2016) durch die Materialien: Wie wird die Vergleichbarkeit von Nutzerinnen oder von Dingen hergestellt? Wie und wo wird kategorisiert und bewertet? Was sind die (rechnerischen) Operationen des Vergleichs? Wie werden die Resultate dargestellt?

Daraus erstellte ich eine Typologie von Empfehlungssystemen, welche die beschriebenen Systeme hinsichtlich ihrer Kategorisierungs-, Bewertungs- und Vergleichsweisen ordnet. So lässt sich idealtypisch aufzeigen, wie die Systeme Relationen zwischen Dingen, Nutzerinnen und Unternehmen generieren.

Ethnografie ist eine explorative und Theorien generierende Herangehensweise. Auch wenn die genaue Beschreibung, was im Feld vor sich geht, an und für sich schon einen Wert hat, stellt sich immer auch die Frage, wie sich

anhand der ethnografischen Empirie etwas zeigen lässt, das über das spezifische Feld hinaus Relevanz hat. Dies versuche ich in den folgenden Kapiteln zu tun, in denen ich mithilfe verschiedener Ansätze die theoretische Figur der »Momente der Datafizierung« entwickle und an meinem empirischen Material illustriere.

Die ethnografische Analyse von Daten und Algorithmen (und des Digitalen) betont die Notwendigkeit, Technologie zu dezentrieren und ihre Einbettung und Kontextualisierung ernst zu nehmen (Pink et al., 2016, S. 9). Wie ich bereits in Kapitel 2 angedeutet habe, stellen Ethnografinnen immer wieder fest, dass Daten trotz gegenteiliger Behauptung immer Reste solcher Kontexte anhaften, die sich für die Datenproduktion oder die Datenverwertung als problematisch erweisen. Der Kontext der Datenerzeugung und -anwendung ist auch im Feld ein Thema (siehe: Seaver 2015) und als Verweisungshorizont für die In-Wert-Setzung von Daten von zentraler Bedeutung: Dieser Kontext besteht insbesondere in den *datengenerierenden Beziehungen* zwischen Nutzerinnen und Unternehmen, welche für die Verarbeitung der Daten entfernt und unsichtbar gemacht werden müssen – damit sie dann in der Anwendung wieder wuchern können. Wie das geschieht, zeige ich in den folgenden Kapiteln.