

# 1 Einleitung

---

Vor nicht allzu langer Zeit war Web-Surfing die zentrale Metapher, die die Anwendung und Erfahrung des frühen Internets bezeichnete. Noch ohne Lotsen, aber mithilfe der Web-Browser Netscape Navigator und Microsoft Explorer, erschlossen sich abenteuerlustige Web-Surferinnen dieses neue Reich. Die Erfahrung des frühen Webs war geprägt von Vielheit, Überfluss, Irrwegen, zufälligen Entdeckungen und zielloser, aber doch manchmal zur Ankunft führenden, Navigation. Evgeny Morozov beschreibt diesen Moment der Netzgeschichte – als man Internet im Englischen noch groß schrieb – als (kurze) Renaissance des Flanierens: Neugier, Exploration und Offenheit für glückliche Zufallsentdeckungen zeichneten den »Cyberflâneur« aus, der durch die Straßen des Netzes schlenderte:

The idea of exploring cyberspace as virgin territory, not yet colonized by governments and corporations, was romantic [...] Online communities like Geo-Cities and Tripod were the true digital arcades of that period, trading in the most obscure and the most peculiar, without any sort of hierarchy ranking them by popularity or commercial value [...] For a brief moment in the mid-1990s, it did seem that the Internet might trigger an unexpected renaissance of flânerie. (Morozov, 2012)

Einen anschaulichen Bericht über das Flanieren im Web liefert der Anthropologe Daniel Miller in einer seiner Arbeiten zum Webgebrauch der Einwohnerinnen von Trinidad. Er beschreibt eingehend, wie er sich diesen digitalen Raum erschließt – als Rhythmik des Suchens und Findens, von Ablenkung, Aufmerksamkeit und Verführung (siehe auch Paasonen 2016). Miller wird nicht absichtlich zum Flâneur. Die Linkstruktur und die ästhetische Gestaltung der Webseiten machen ihn erst dazu.

Even as a researcher I found that almost on a daily basis I would start with the intention of following one particular route of investigation and then find myself seduced by the aesthetics of one of the websites visited and moved

by the simplicity of clicking to follow a link proffered by that site. A few more clicks would send me hurtling down some channels carved out of cyberspace by the sculptured links of these website creators, often to such a degree that it was hard to retrieve the original place from which this diversion had began, but often grateful that my lack of determination had in fact led me to view some unexpected vistas and delight in some other creations than those I would otherwise have encountered. (Miller, 2000, S. 18)

Obwohl der Browser auch heute noch eine zentrale Rolle spielt (zumindest auf Desktop-Computern), hat Web-Surfing als zentraler Begriff zur Beschreibung der Nutzungs- und Erfahrungsmodi des Netzes an Stellenwert eingebüßt: Metaphern werden alt und »uncool«, wie Dewey (2015) nahelegt. »Hardly anyone ›surfs‹ the Web anymore« (Morozov, 2012). Doch sind nicht nur andere Metaphern an die Stelle des Web-Surfs getreten, sondern auch neue Ordnungsformen, die mit neuen Erfahrungsmodi verbunden sind. Morozov führt den Niedergang des Pariser Flaneurs auf architektonische und technologische Veränderungen im Paris der zweiten Hälfte des 19. Jahrhunderts zurück: Das kleinräumige mittelalterliche Paris musste offenen Boulevards, künstlicher Beleuchtung, Verkehr und großen Warenhäusern weichen. Gleichmaßen gilt für das Web, dass technologische und soziale Veränderungen die Web-Surferin – beziehungsweise den »Cyberflâneur« – zu einer immer unwahrscheinlicheren Figur werden lassen. Das zeigt sich an der zunehmenden Irrelevanz des Browsers zugunsten von Apps (siehe zum Beispiel Morris & Murray 2018) und Plattformen (zum Beispiel Van Dijck et al. 2018; Bucher & Helmond 2017), aber auch wegen der Ökonomisierung und Monopolisierung (Srnicek, 2017; Doctorow, 2020) des Netzes. Die architektonischen beziehungsweise technologischen Veränderungen des Webs bestehen vor allem in einer weitgehenden Vorstrukturierung möglicher Verhaltensweisen (click, like, share) und der algorithmischen Vorsortierung von »Content« für spezifische Nutzerinnen oder Kategorien von Nutzerinnen.

Knapp 20 Jahre nach Miller schreibt Jia Tolentino (2019) im *New Yorker* über die App TikTok, die aufgrund ihrer unzimmerlichen Fokussierung auf algorithmische Empfehlungen als wegweisend gilt (siehe zum Beispiel Herrman 2019) – wenn auch nicht unbedingt in einem positiven Sinn. Ich zitiere hier als Gegenüberstellung zu Daniel Millers autoethnografischer Beschreibung relativ ausführlich drei Abschnitte, in denen Tolentino ihre Nutzung der App beschreibt:

I opened the app, and saw a three-foot-tall woman making her microwave door squeak to the melody of ›Yeah‹, by Usher, and then a dental hygienist and her patient dancing to ›Baby Shark‹. A teen-age girl blew up a bunch of balloons that spelled ›pussy‹ to the tune of a jazz song from the beloved soundtrack of the anime series ›Cowboy Bebop‹. Young white people lip-synched to audio of nonwhite people in ways that ranged from innocently racist to overtly racist. A kid sprayed shaving cream into a Croc and stepped into it so that shaving cream squirted out of the holes in the Croc. In five minutes, the app had sandblasted my cognitive matter with twenty TikToks that had the legibility and logic of a narcoleptic dream. [...] After I had watched TikTok on and off for a couple of days, the racist lip-synchs disappeared from my feed. I started to see a lot of videos of fat dogs, teen-agers playing pranks on their teachers, retail workers making lemonade from the lemons of being bored and underpaid. I still sometimes saw things I didn't like: people in horror masks popping into the frame, or fourteen-year-old girls trying to be sexy, or rich kids showing off the McMansions where they lived. But I often found myself barking with laughter, in thrall to the unhinged cadences of the app. [...] I found it both freeing and disturbing to spend time on a platform that didn't ask me to pretend that I was on the Internet for a good reason. I was not giving TikTok my attention because I wanted to keep up with the news, or because I was trying to soothe and irritate myself by looking at photos of my friends on vacation. I was giving TikTok my attention because it was serving me what would retain my attention, and it could do that because it had been designed to perform algorithmic pyrotechnics that were capable of making a half hour pass before I remembered to look away. (Tolentino, 2019, k.S.)

TikTok erfordert kein Login. Die App generiert einen zufälligen Nutzernamen. Man muss sich mit niemandem befreunden und man braucht keine Angaben über Interessen und Vorlieben zu machen. Die App startet schon beim ersten Mal automatisch im »For You«-Modus und die Nutzerin wird direkt mitten in einen unendlichen Strom von Videos versetzt.

Millers Beispiel des Web-Surfinds und Tolentinos Bericht über ihre Erfahrung mit TikTok scheinen vordergründig ähnlich zu sein. Beide beschreiben eine digitale Welt, welche die Nutzerin immer weiter in den Kaninchenbau hineinzieht. Doch unterscheiden sie sich auf grundlegende Weise.

Daniel Miller traversierte eine emergente und zufällige Ordnung anhand von Hyperlinks auf Webseiten: Die unüberschaubare Menge an Inhalten des

frühen Netzes mit all seinen persönlichen Webseiten und obskuren Onlinegemeinschaften bedingt gerade das Versprechen erlebnisreicher Exploration.<sup>1</sup> Überfluss und Chaos wird zunehmend zum Problem erklärt, vor dem die Nutzerinnen bewahrt werden sollen. Miller war konfrontiert mit der relativ zufälligen, ungeordneten Welt der privaten Homepages der 90er-Jahre, die für die zeitgenössischen Surferinnen eine aus heutiger Sicht möglicherweise erstaunliche Faszination ausübten. Auch Tolentino ist einer kuriosen Welt ausgesetzt: Sie befindet sich aber in der (relativ) geschlossenen Welt einer App, welche die Nutzerinnen möglichst lange unterhalten beziehungsweise ihr »Engagement« erhalten will und dabei nichts dem Zufall überlässt. Obwohl die TikTok-Videos nur so auf die Nutzerin einzuprasseln scheinen, ist doch Millers Surfen im Netz »direkter«. Er springt von Seite zu Seite und nutzt dabei die von den Seitenbetreibern angebotenen Links. Sein Ausflug ins Netz ist von den vielen unterschiedlichen persönlichen Webseiten kuratiert, welche die Links zu anderen Seiten nach eigenen Kriterien der Relevanz auswählen. Tolentinos TikTok-»Binge« ist hingegen kuratiert von automatisierten Empfehlungssystemen. Der zentrale Unterschied liegt nicht so sehr darin, was der Web-Surfer Miller oder die App-Nutzerin Tolentino »navigieren«, sondern in der Ordnungsweise dieser Dinge. Eine (relativ) zufällige Ordnung der Hyperlinks bei Miller. Oder eine nach spezifischen, wenn auch nicht weniger opaken, Kriterien geordnete Liste von Inhalten bei Tolentino.

Daniel Miller surfte in einem Web, das noch weitgehend auf die Beobachtung von Nutzerverhalten verzichtete. Jia Tolentinos Erfahrung ist dagegen in eine datafizierte und datafizierende Infrastruktur eingelassen, die zahlreiche ihrer Verhaltensweisen registriert und in Beziehung setzt zu vergangenem Verhalten und zu den Verhaltensweisen anderer Nutzerinnen. Tolentino hat nicht nur explizite »Interaktionsmöglichkeiten« wie »like«, »follow«, »share« oder »comment«. Auch die Art und Weise, wie sie den nie endenden Strom von Videos navigiert, erzeugt Nutzungsdaten. Wie oft sie sich ein Video anschaut und ob sie ein Video zu Ende schaut oder schon vorher weiterscrollt,

---

1 Obwohl dieses Versprechen womöglich von bestimmten stillschweigenden Annahmen abhängig ist, deren Bewährung je länger je mehr in Zweifel gezogen werden kann, falls sie überhaupt einmal gegolten haben sollten: Dass die stetige Expansion des Webs nicht als Problem erscheint, ist auch abgänglich von einem Vertrauen in eine Wissenskulturl, in der man sich »implizit darauf verlassen [kann], dass nur sinnvolle und richtige Informationen eingegeben werden« (Schröter, 2004, S. 127).

dürfte ebenso im Hintergrund von TikTok registriert und ausgewertet werden. Zu Beginn wirkt es so, als würden die verschiedenen Videos willkürlich in Tolentinos Feed erscheinen. Unpassende, irrelevante oder uninteressante Videos werden ihr jedoch immer seltener angezeigt. Dank »[A]lgorithmic pyrotechnics« kann TikTok seinen Nutzerinnen genau das vorsetzen, was sie mögen: »All you have to do is watch, engage with what you like, skip what you don't, and you'll find an endless stream of short videos that feel personalized just for you«, wie TikTok in seiner Beschreibung im App Store verlautet. Und bevor man es merkt, ist eine halbe Stunde vergangen. Tolentino navigiert durch eine gefilterte Ordnung von Videos, die auf Vorhersagen beruht, wie Tolentino – qua Verhaltensweisen – mit bestimmten Videos »zusammenpasst«. Wenn Web-Surfen die Metapher ist, welche die dominante Erfahrungsform des frühen Netzes beschreibt, stehen »binging« oder »gorking out« (McKelvey & Hunt, 2019, S. 5) für aktuelle Erfahrungsmodi digitaler Inhalte. Der Begriff »binge« steht zwar für den ununterbrochenen Konsum von TV-Serien. Er kann aber leicht adaptiert werden, um einen zeitgenössischen Konsummodus zu kennzeichnen, in dem ohne großes Zutun der Konsumentin nach dem Konsum einer Einheit sogleich die nächste, von Empfehlungssystemen vorselektierte, Einheit erscheint und sich zum Konsum oder zum Kauf anbietet. Die Nutzerin soll dabei weniger Zeit mit Suchen verbringen müssen: Verschiedene Formen von Technologieunternehmen und ihre Algorithmen explorieren das Angebot für uns, so dass das Navigieren durch unbekannte Gewässer eine Erfahrung der Vergangenheit wird. Wir müssen nicht mehr selbst nach interessanten, relevanten oder uns unterhaltenden Dingen suchen. Sie sollen wie von selbst zu uns kommen. Wie Bucher nahelegt, drehen sich die Verhältnisse von Suchen und Finden um: »[U]sers do not merely browse the content that they find interesting; the ›interesting‹ content increasingly finds them« (Bucher, 2012, S. 12).

In den zwanzig Jahren zwischen Miller und Tolentino hat sich das Internet grundlegend gewandelt. Seit der ersten Bannerwerbung 1994 hat sich eine digitale Ökonomie entwickelt, die maßgeblich auf Personalisierung und der Kommodifizierung von Nutzerdaten basiert (Crain, 2018; West, 2017). Tech-Unternehmen wählen für uns das nächste Video, »Produkte, die dich auch interessieren könnten« oder »Songs extra für dich« aus. Sie sprechen uns oftmals als einzigartige Individuen an und offerieren uns einen Spiegel auf die eigene Identität: Weshalb denkt TikTok, dass ich Videos mag, in denen sich Menschen gegenseitig Streiche spielen? Was habe ich in der Vergangenheit angeklickt, dass mir im Webshop eine aufblasbare Badewanne empfoh-

len wird? Und: Würde ich mich eigentlich auch für Anleitungen von Fitnessübungen interessieren, wenn mir Instagram nie solche Posts angezeigt hätte? Empfehlungssysteme ziehen uns als Nutzerinnen in ihre Rückkopplungsschlaufen und versuchen, unsere Aufmerksamkeit – unser »engagement« in der Sprache des Marketings aber auch der Computerwissenschaften (Seaver, 2018) – möglichst lange aufrechtzuerhalten.

Miller und Tolentino beschreiben, wie ihnen digitale »Inhalte« auf der Vorderbühne präsentiert werden. In diesem Buch werfe ich einen Blick auf die Hinterbühne der digitalen Ökonomie. Ich gehe dabei insbesondere zwei Fragen nach. Erstens die *Frage der Datafizierung*: Wie entstehen Personendaten beziehungsweise wie produzieren und verarbeiten Unternehmen Personendaten? Wie kommt es, dass Personendaten in der digitalen Ökonomie zu einer wertvollen »commodity« werden? Aus meiner Feldforschung drängte sich daran anschließend die zweite *Frage nach Empfehlungssystemen* auf, die in der digitalen Ökonomie eine zentrale Funktion einnehmen: Was ist die Funktion, die Empfehlungssysteme in der digitalen Ökonomie erfüllen und wie operieren sie?

## Datafizierung

Was wir im Internet tun, ist vermeintlich immer schon datenförmig: »Everything is made of data these days« (zitiert in: Sadowski 2019). Alles, was in der Welt des Internets passiert, findet im Medium des Digitalen statt: Bits und Bytes, die zwischen Computern hin- und hergesendet, gespeichert und wieder abgerufen werden können. Aber nur weil etwas digital ist, heißt das nicht, dass wir es mit verrechenbaren Daten zu tun haben. Daniel Miller verweist auf ein beinahe schon historisches Beispiel eines Internetnutzers, dessen Verhaltensweisen (noch) nicht Gegenstand einer weitgehenden Beobachtung und Datafizierung waren. Es musste erst jemand auf die Idee kommen, dass die auf den ersten Blick trivialen und unbedeutenden Klicks zu Daten gemacht, ausgewertet und möglicherweise sogar »motiviert« werden können. »[C]oming to think of something as »data« in the first place« (Dourish & Gómez Cruz, 2018, S. 2), ist nicht so selbstverständlich, wie es uns im Zeitalter von Big Data und Algorithmen erscheinen mag.

Datenunternehmer (Beauvisage & Mellet, 2020) wie Nik von Earlybird Digital verdienen ihr Geld damit, anderen Unternehmen das Potenzial von Personendaten zu offenbaren. Obwohl viele Unternehmen Daten über ihre

Verkäufe und Kunden registrieren und aufbewahren, sei nicht allen bewusst, auf was für einem »Datenschatz« sie sitzen würden. Nik sieht es als seine Aufgabe an, solchen Unternehmen dabei zu helfen, ein »data mindset« zu entwickeln, in ihren Datenbanken die Datenschätze zu heben und dadurch möglich zu machen, »was mit Daten möglich ist« (siehe Kapitel 2). Er knüpft dabei an ein verbreitetes Deutungsmuster an, demgemäß Big Data (zum Beispiel Mayer-Schönberger & Cukier 2013) und Personendaten (wichtig: Schwab 2011) zu einer wertvollen neuen Ressource werden, um Profite zu generieren. Doch obwohl mit Daten im Sinne eines absoluten Mittels vieles möglich zu sein scheint, ist in der Praxis klar, dass es ein schwieriges Unterfangen ist, Personendaten für Firmen produktiv zu machen.

Um zu untersuchen, wie Verhaltensweisen zu wertvollen und produktiven Personendaten werden, verfolge ich ein zweigleisig komplementäres Forschungsdesign (siehe Kapitel 3): Einerseits hatte ich die Gelegenheit, im schweizweit bekannten Kundentreue-Unternehmen Earlybird und in dessen Tochterfirma Earlybird Digital ethnografische Feldforschung zu betreiben. Dabei konnte ich untersuchen, wie sie im Rahmen ihrer Digitalisierungsstrategie die Verhaltensweisen, Interessen und Vorlieben ihrer Nutzerinnen als eine wertvolle und profitgenerierende Ressource mobilisieren.<sup>2</sup> Earlybird hat sich im Rahmen eines strategischen Digitalisierungsprojekts zum Ziel gesetzt, möglichst viele Daten über ihre »Member« zu sammeln und ihr Businessmodell an diesen Daten auszurichten. Die Tochterfirma Earlybird Digital unterstützt Earlybird (und andere Unternehmen) in diesem Prozess.

Andererseits ergab sich aus dieser Feldforschung eine weitere empirische »fieldsite«: Der Onlinekurs »Introduction to Recommender Systems«, der von der University of Minnesota auf der e-Learning-Webseite Coursera (2017) angeboten wurde. Dieser Kurs vermittelt die technische Funktionsweise und die grundsätzliche Logik verschiedener Empfehlungssysteme. Diese beiden empirischen Fieldsites ermöglichen einen komplementären Zugang zur unternehmerischen Praxis und zur computerwissenschaftlichen Theorie von Empfehlungssystemen.<sup>3</sup>

---

2 Die Namen der Unternehmen sowie alle Namen der Mitarbeitenden wurden anonymisiert.

3 Der komplementäre Einbezug von Bildungsmaterialien zu Empfehlungssystemen drängte sich auch deshalb auf, weil Earlybird zum Zeitpunkt der Forschung noch nicht über ein produktives Empfehlungssystem verfügte.

Auf dieser empirischen Basis beantworte ich zwei Fragen: Wie entstehen Personendaten? Und: Wozu dienen sie? Etwas abstrakter formuliert: Über welche sozialen und technischen Prozesse werden individuelle Verhaltensweisen in ökonomisch verwertbare Daten transformiert?

Die kurze Antwort: Personendaten werden aus relationalen Verwicklungen von Unternehmen und Nutzerinnen erzeugt und genutzt, um Unternehmen, Nutzerinnen und eventuell Dritte weiter miteinander zu verwickeln. Dazwischen wird formalisiert, quantifiziert, kategorisiert, geputzt, gerechnet, verglichen und sortiert.

In Kapitel 4 gebe ich einen Überblick über drei Momente der Datafizierung, indem ich eine Parallele zu Anna Tsings (2013; 2018) Arbeit über die Warenkette von Kiefern-pilzen herstelle. Wie Tsing darlegt, transformieren verschiedene Akteure »Matsutake«-Pilze in ihrem Lebenszyklus von Geschenken zu Waren und wieder zu Geschenken. Die Pilze generieren gute Beziehungen zwischen untereinander bekannten Sammlerinnen und Aufkäufern. Diese Beziehungen werden in den Warenlagern von Großhändlerinnen entfernt, um die Pilze zu standardisierten und profitablen Waren zu machen. Sobald sie in Japan ankommen, beginnen die »relationalen Tentakel« der Pilze wieder zu wuchern: Traditionellerweise schenken sich Japanerinnen und Japaner die Pilze, um ihre guten Beziehungen zu erhalten und zu vertiefen.

Der Lebens- und Warenzyklus des Pilzes erweist sich als gute Metapher, um über Personendaten nachzudenken: Welche Rolle spielen *Relationen* zwischen Nutzerinnen, Unternehmen und Dingen bei der Erzeugung und In-Wert-Setzung von Personendaten? Der Wert von Personendaten geht aus der Spannung zwischen heterogenen Wertregimes von Gabe und Ware, zwischen echten, spurenhafte Verhaltensweisen der Nutzerinnen, der Entfernung dieser Spuren und Sinninvestitionen und der Erzeugung neuer, vielversprechender Relationen hervor.

In Kapitel 5 beschreibe ich diese Spannung anhand eines »Datenexperiments«, das ich bei Earlybird mitverfolgen konnte: Ein Versuch, Daten zu generieren und diese Daten für die Kategorisierung der Nutzerinnen zu verwenden. Dabei zeigten sich drei Probleme. Erstens erwies es sich keineswegs als einfach, überhaupt an Personendaten zu kommen. Es wurde für die Beteiligten bei Earlybird klar, dass sie den Nutzerinnen als Gegenleistung etwas bieten müssen, um sie in datengenerierende Beziehungen zu verwickeln. Zweitens stieß ich auf einen seltsamen Widerspruch: Earlybird war es doch gelungen, Personendaten zu generieren, die sie als »schöne Daten« bezeichneten. Die Daten galten vor allem deshalb als »schön«, weil die Nutzerin-



nen viel von sich preisgegeben hatten: ihre Wünsche, Vorlieben und Interessen. Diese Bedeutungen, welche die Nutzerinnen mit ihren Äußerungen und Handlungen verbinden, schienen zwar zentral für den Wert und die Schönheit der Daten, wurden in deren weiteren Lebensverlauf aber unter großem Aufwand »ignoriert«, um Nutzerinnen und ihre Verhaltensweisen vergleichbar und verrechenbar zu machen. Das *dritte* Problem bestand darin, neue Marketing-Kampagnen oder »Empfehlungen« für die Nutzerinnen des Systems aus den Daten abzuleiten, um die weitere Nutzung des Angebots zu motivieren. So weit kam es aber nicht ganz. Stattdessen wurde der entwickelte »Classifier« zu einem Geschenk, um neue, an Data-Science-Dienstleistungen interessierte, Businesskunden für Earlybird Digital zu gewinnen.

An meiner theoretischen Figur, die ich »Momente der Datafizierung« nenne, zeige ich anschliessend, dass die Datafizierung von Verhaltensweisen aus drei, in ihrer Logik sequentiellen, aber empirisch überlappenden, Momenten besteht, in denen Personendaten zwischen den verschiedenen Wertregimes von Gabe und Ware wabern (Kapitel 6 bis 8).

Personendaten sind nicht einfach gegeben – aber sie können auch nicht ohne Weiteres genommen werden. Während die Techgiganten vermeintlich immer schon über Nutzerdaten verfügen, müssen kleine oder neue digitale Unternehmen das Problem des Kaltstarts lösen: Ohne Nutzerdaten können sie den (neuen) Nutzerinnen keine passenden Empfehlungen aussprechen, um sie weiter an das Unternehmen zu binden. Im ersten Moment (siehe Kapitel 6) sollen »inducement gifts« (Elder-Vass, 2016) die Nutzerinnen dazu verlocken, digitale Dienstleistungen oder Plattformen zu nutzen und im Gegenzug den Unternehmen ihre Daten zu überlassen. Dieser Austausch ist gerahmt als »give-to-get« (Fourcade & Klutetz, 2020) und technisch so organisiert, dass die Annahme des Geschenks bereits das Gegengeschenk konstituiert. In Momenten des »clicks« oder des »taps« werden Geschenk und Gegengeschenk gleichzeitig ausgetauscht – und so die Produktion von Personendaten gestartet.

Das zweite Moment der Datafizierung (Kapitel 7) dreht sich um einen produktiven Widerspruch von Personendaten. Auf der einen Seite gelten sie als Spuren von echten, und für die Nutzerinnen bedeutungsvollen, Handlungen. Auf der anderen Seite müssen diese Sinnüberschüsse aber technisch unsichtbar gemacht oder weggearbeitet werden, damit Nutzerinnen und ihre Verhaltensweisen überhaupt vergleichbar und verrechenbar werden. Ich beschreibe zwei Varianten, wie Sinnüberschüsse operativ ignoriert werden: Durch *Encoding* stellen Unternehmen ihren Nutzerinnen bereits genuin digi-

tale Verhaltensweisen wie beispielsweise »liking«, »friending« oder »retweet« zur Verfügung, die sich in der tatsächlichen Gebrauchsweise unterschiedlicher Nutzerinnen und Nutzergemeinschaften in ihrer Bedeutung maßgeblich unterscheiden können. In der Datenbank ist ein Like aber immer ein Like, unabhängig von den Sinninvestitionen der Nutzerinnen.

Die zweite Variante besteht darin, Äußerungen von Nutzerinnen *ex post* zu vergleichbaren und verrechenbaren Daten zu machen. Das Verfahren der *Vektorisierung* ermöglicht es beispielsweise, Texte oder Wörter in einen gemeinsamen geometrischen Raum zu projizieren, um sie in Bezug auf Ähnlichkeiten und Differenzen vermessen zu können. Einerseits sollen sich Nutzerinnen in digitalen Infrastrukturen (soziale Netzwerkplattformen, Streaming, Onlineshops, Suchmaschinen, etc.) auf authentische Weise ausdrücken und verhalten können; andererseits abstrahieren Prozesse der Datafizierung von den subjektiven Bedeutungen, welche die Nutzerinnen mit ihren Handlungen und Ausdrucksweisen verbinden, und von den sozialen Kontexten, in denen diese Handlungen stattfinden. Wenn ich schreibe, dass die Deutungsweisen der Nutzerinnen »ignoriert« werden, beabsichtige ich keine Methodenkritik. Damit Daten als Daten funktionieren können, *müssen* sie von den lokalen, situativen Bedeutungen, die sie mittragen, abstrahieren – erst so öffnet sich das Potenzial ihrer weiteren Verwertung.

Im dritten Moment der Datafizierung geht es darum, das latente Potenzial von Personendaten auszuschöpfen (Kapitel 8). Was im Normalfall »Auswertung« oder in der automatisierten Version »Algorithmus« heißt, besteht in einer Verdichtung von Relationen: Einerseits werden über Operationen des Faltens unterschiedliche Annahmen, Normalitäts- und Ordnungsvorstellungen eingefaltet und neue Sichtbarkeiten und Unsichtbarkeiten produziert. Andererseits werden über algorithmische Operationen der Relationierung die nun vergleichbaren Nutzerinnen zu anderen Nutzerinnen (bzw. zu deren Daten) in Beziehung gesetzt. Es kommt zu einer Verdichtung der Relationen von Nutzerinnen und Dingen, so dass daraus die »passendsten«, vielversprechendsten neuen Relationen abgeleitet werden können. Ziel dieser Art von Auswertung ist nicht die Produktion von Erkenntnis *per se*, sondern die Produktion von neuen, prospektiven Relationen und mehr Daten: »In the digital context, social relations are not treated as ›given‹, but as something that can be ›enhanced‹, or ›generated‹ and ›deployed‹ towards practical purposes« (Marres, 2017, S. 47). Der algorithmische Vergleich ist produktiv: Er stellt aus bestehenden (Daten-)Relationen neue Relationen zwischen Nutzerinnen und Dingen her. Welches TikTok-Video, welcher Earlybird-Deal oder welcher Song

auf Spotify soll der Nutzerin als nächstes angezeigt werden? Diese neuen Relationen zwischen Nutzerinnen und Dingen sollen als »good matches« die Beziehung von Nutzerin und Unternehmen aufrecht erhalten.

Momente der Datafizierung funktionieren tendenziell zirkulär: Sobald die Datenproduktion in die Gänge kommt, treten prospektive Relationen in der Form von (personalisierten) Empfehlungen an die Stelle von digitalen Geschenken.

## Empfehlungssysteme

Empfehlungssysteme verkörpern die drei Momente der Datafizierung und geben die Antwort auf die Frage, zu was Personendaten dienen bzw. worin ihr Wert besteht. Die automatisierte Empfehlung basiert auf (Personen-)Daten und generiert neue Verhaltensweisen und Personendaten (wenn sie funktioniert).

Empfehlungssysteme sind so gesehen eine zentrale Form der Digitalität: Eine Datafizierungs- und eine Relationierungsmaschine, die eine Brückenfunktion zwischen Unternehmen, Nutzerinnen, Dingen (z.B. »Deals« bei Earlybird, Filme, Songs, Produkte, etc. in anderen Fällen) und unter Umständen auch Drittparteien wie Werbekundinnen, Entwicklerinnen oder »content providers« (wie beispielsweise Musikerinnen im Fall von Spotify) einnimmt und dabei Verhaltensweisen und Daten produziert, die sie gleich weiter nutzt.

In der computerwissenschaftlichen Literatur gelten Empfehlungssysteme als Lösung für das Problem des »information overload«: Sie reduzieren im Namen der Nutzerin ein überbordendes Angebot an digitalen Inhalten auf eine »mundgerechte«, auf den Geschmack der Nutzerin abgestimmte Menge, die eine Auswahl erleichtert oder sogar unnötig macht. Empfehlungen verbessern das Nutzererlebnis – so das Versprechen. Im Gegenzug brauchen die Nutzerinnen bloß ihre Verhaltensdaten zur Verfügung zu stellen: Personalisierte Empfehlungen und Personendaten bilden Geschenk und Gegengeschenk, das die Beziehung erhält und vertieft.

Personalisierende Empfehlungssysteme verdichten die Relationen zwischen Nutzerinnen, Dingen und Unternehmen. Aufgrund bereits registrierter Verhaltensdaten prüfen sie alle möglichen, paarweisen Kombinationen von Nutzerinnen und Dingen, um diejenige auszuwählen, welche die Beziehung von Nutzerin und Unternehmen mit größter Wahrscheinlichkeit in die

Zukunft verlängert. Eine erfolgreiche Empfehlung generiert nicht nur »conversions« (zum Beispiel in Form von Verkäufen), sondern auch neue Verhaltensweisen – i.e. »engagement« –, die wiederum in die Berechnung neuer Empfehlungen eingehen. Empfehlungssysteme erweisen sich als Datafizierungsmaschinen und als Generatoren von »good matches« (Kapitel 9).

Anhand von verschiedenen theoretischen und praktischen Beispielen aus dem Onlinekurs »Introduction to Recommender Systems« lege ich vier Typen von Empfehlungssystemen dar und beschreibe, auf welche Weise sie die Brücke zwischen Nutzerinnen und Dingen schlagen (siehe Kapitel 10). Ich identifiziere drei unterschiedliche Weisen der Relationierung, d.h. wie aus bestehenden Relationen auf neue, prospektive Relationen geschlossen wird. *Popularitätsmetriken* relationieren pauschal: Sie unterstellen, dass dem Populären eine allgemeine, für alle gleichermaßen geltende, Relevanz zukommt. *Stereotypisierende Recommender* verwenden soziale Kategorien: Zeigen Nutzerinnen einer bestimmten Kategorie eine Vorliebe für bestimmte Dinge, wird die Kategorie zur Stellvertreterin für diese Vorliebe. *Inhaltsbasierte Empfehlungen* und *Collaborative Filtering* beruht auf paarweisen »matchings«: Jede Einheit wird zu jeder anderen Einheit in Beziehung gesetzt, um sie in Bezug auf ihre Ähnlichkeit mit der interessierenden Einheit vergleichen zu können. Während die inhaltsbasierte Empfehlung ein einfaches paarweises Matching von Dingen durchführt, besteht Collaborative Filtering aus einem doppelten Matching-Verfahren: Erst werden Nutzerinnen untereinander, dann Nutzerinnen und Dinge paarweise relationiert.

## Anschlüsse

Dieses Buch ist ein Versuch, verschiedene Soziologien (und Ansätze aus der Anthropologie) zu verbinden, um etwas Licht in das Dickicht von Datafizierungsprozessen zu bringen. Von seinem Gegenstand her ist es ein Beitrag zu den »critical algorithm studies« (Gillespie & Seaver, 2015). Ich bediene mich aber einer Reihe von Konzepten aus verschiedenen Forschungs- und Theoriekontexten, zu denen ich einen Beitrag leiste.

Zu Beginn war meine Fragestellung vor allem aus einer Soziologie der Kategorisierung (Bowker & Star, 1999) und des Vergleichs (Heintz, 2010, 2016) motiviert: Wie kategorisieren und vergleichen Unternehmen ihre Nutzerinnen oder Kundinnen? Dem gehe ich vor allem in Kapitel 7 nach, wo ich beschreibe, wie Vergleichbarkeit von Verhaltensweisen ex-ante durch »en-

coding« oder ex-post durch »Vektorisierung« hergestellt wird. In den Kapiteln 8 und 10 erweitere ich den Vergleichsbegriff von Bettina Heintz, indem ich »matching« oder »Relationierung« als spezifische Operation eines algorithmischen Vergleichs definiere, welche die zu vergleichenden Nutzerinnen jeweils paarweise zueinander in Beziehung setzt und diese Relationen vergleicht. Wie ich an verschiedenen Stellen zeige, ist diese Art des algorithmischen Vergleichs in viel geringerem Maß auf klassische Kategorien der Demografie und der Identität angewiesen als traditionelle, statistische Formen der Beobachtung (Heintz, 2021). Wie ich insbesondere in Kapitel 10 darlege, geht es nicht darum, wer die Nutzerinnen *sind*. Viel wichtiger werden (temporäre) Ähnlichkeitsrelationen und prospektive Relationen zwischen Nutzerinnen und Dingen.

In der Feldarbeit stellte sich heraus, dass die Kategorisierungs- und Vergleichsweisen kaum in Isolation betrachtet werden können, sondern sich nur im Kontext der digitalen *Datenökonomie* angemessen verstehen lassen. Zu Beginn ging ich davon aus, dass sich das primäre Interesse von Unternehmen darauf richtet, Wissen über ihre Nutzerinnen zu generieren. Natürlich stimmt es, wenn Earlybird sagt, es sei ihr Ziel, möglichst viel über die Nutzerinnen zu wissen. Das ist aber nur ein Teil des Bildes. Dieses Ziel ließe sich auch mit den Mitteln der traditionellen Marktforschung erreichen. Daten sollen nicht nur Erkenntnisse über Nutzerinnen, sondern immer auch weitere Daten generieren (Sadowski, 2019). Wie auch andere Unternehmen der digitalen Ökonomie versucht Earlybird, eine ganze Datafizierungsinfrastruktur in Stellung zu bringen, in der aus den Verhaltensweisen der Nutzerinnen Daten und aus Daten neue Verhaltensweisen generiert werden sollen (vgl. Zuboff 2018 zu »behavioral surplus«). Momente der Datafizierung sind in diesem Sinne nicht abgeschlossen. Sie zielen darauf, einen ewigen Kreislauf von Verhaltensweisen, Daten und neuen Verhaltensweisen aufrecht zu erhalten. Wenn der Kreislauf von datengenerierenden Verhaltensweisen und verhaltensgenerierenden Daten stoppt (oder gar nicht erst in Gang kommt), zerfällt auch der Wert der Daten (vgl. Hwang 2020 zur Überbewertung der Effektivität personalisierter Werbung).

Der Wert von Daten besteht also darin, dass sie sowohl »engagement« erhalten und steigern (Seaver, 2018) und gleichzeitig neue Daten erzeugen können. Wie gut sie das tatsächlich können, ist eine andere Frage. Personendaten und ihre Verarbeitung sind zentrale Bestandteile, um die ökonomische Zirkulation von Nutzerinnen, Daten und Content am Laufen zu halten. Wie dieser Wert der Daten in ihrer Zirkulation zustande kommt, dürfte für den

jüngeren Forschungszweig der *Valuation Studies* (z.B. Vatin 2013) beziehungsweise der Soziologie der Bewertung (Meier et al., 2016) von Interesse sein. Wie ich zeige, verkörpern die drei Momente der Datafizierung eine mehrstufige, zirkuläre In-Wert-Setzung von Personendaten. Daten gehen aus einer reziproken Beziehung respektive aus einer Art Gabentausch zwischen Unternehmen und Nutzerinnen hervor und sie dienen dazu, diese Beziehungen zu gestalten (siehe Kapitel 6).

Damit die registrierten Verhaltensweisen ihren Dienst als Daten tun können, müssen sie aber von ihren Kontextbezügen und Sinninvestitionen befreit und in neuen, formalen Kontexten integriert und zu den datafizierten Verhaltensweisen anderer Nutzerinnen in Beziehung gesetzt werden. Um diese Vorgänge auszuformulieren greife ich auf theoretische Konzepte aus der Anthropologie (Mauss, 1923) und der Marktsoziologie (Fourcade & Healy, 2016, 2017b) zurück. Insbesondere die ethnografische Arbeit von Anna Tsing (2013; 2018) zur Warenkette der Matsutake-Pilze und der Aufsatz *A Maussian bargain: Accumulation by gift in the digital economy* von Marion Fourcade und Daniel Kluttz (2020) erweisen sich als produktiv: Personendaten werden wertvoll durch die Transformationen und Sortiervorgänge, die sie zwischen den verschiedenen Wertregimes von Gaben und Waren hin- und herschieben. Sie gelten als Spuren authentischer Verhaltensweisen. Es lässt sich aber auch mit ihnen rechnen: d.h. sie machen die Verhaltensweisen individueller und idiosynkratischer Nutzerinnen »vergleichbar« (Heintz, 2010) beziehungsweise »kommensurabel« (Espeland & Stevens, 1998).

Soziale Relationen zwischen Unternehmen und Nutzerin (oder auch zwischen Gemeinschaften von Nutzerinnen) sind der fruchtbare Boden, auf dem Personendaten als Datenrelationen gedeihen können (Kapitel 6). Diese Relationen – ihre Bedeutungen und die damit verbundenen Verpflichtungen – müssen aber ignoriert werden (Kapitel 7), damit neue, prospektive Relationen errechnet werden können (Kapitel 8). Hier ist eine spezifische Form der Bewertung (siehe zur Unterscheidung von Bewertung und In-Wert-Setzung: Vatin 2013) in das dritte Moment der Datafizierung eingefaltet. Im algorithmischen Vergleich werden Relationen von Nutzerinnen untereinander oder Nutzerinnen und Dingen (siehe Kapitel 8, 10.3 und 10.4) verglichen und in Bezug auf die Frage *bewertet*, welche *Relationen* von Nutzerinnen und Dingen am ehesten das »engagement« der Nutzerinnen erhält und steigert. Das heißt, es werden nicht Personen per se bewertet (wie beispielsweise in Marketing oder Credit-Scoring üblich), sondern Relationen von Nutzerinnen und Dingen.

Im Zentrum steht also nicht das Individuum, sondern Relationen zwischen Nutzerinnen untereinander, Relationen zwischen Nutzerinnen und Dingen und Relationen zwischen Nutzerinnen und Unternehmen. Momente der Datafizierung dezentrieren das Individuum und relokalisieren es in seinen Relationen zu anderen Nutzerinnen, Dingen, interessierten Drittparteien wie Werbekundinnen oder API-Entwicklerinnen. Damit biete ich eine marktsoziologische Perspektive an, die Datafizierung (und Empfehlungssysteme) als unternehmerische Versuche versteht, Beziehungen zu Nutzerinnen/Kundinnen aktiv zu gestalten (Cochoy et al., 2017). »Inducement gifts« (Elder-Vass, 2016) spielen eine zentrale Rolle, um soziale Beziehungen einzugehen, die in ökonomisch profitable Beziehungen transformiert werden können. In der digitalen Ökonomie handelt es sich dabei typischerweise um umsonst nutzbare Dienstleistungen, deren Nutzung Personendaten generiert (z.B. ein Facebook-Konto). Mithilfe dieser Personendaten sollen soziale Beziehungen in ökonomische transformiert werden (wenn z.B. eine kostenfreie Nutzungsperiode zu Ende geht) oder ökonomische Beziehungen durch die richtigen Angebote zur richtigen Zeit für die richtige Nutzerin erschaffen werden. Datenbasierte Empfehlungssysteme sind Technologien, die »good matches« erzeugen – d.h. Relationen zwischen Nutzerinnen und Dingen, welche die Beziehung zwischen Unternehmen und Nutzerin in die Zukunft verlängern.

Wie ich anhand der Untersuchung verschiedener Typen von Empfehlungssystemen darlege, gibt es verschiedene Arten, diese Verknüpfungen herzustellen. Historisch gesehen lässt sich die Tendenz beobachten, Kundinnen immer genauer zu beobachten und sie nicht nur als Teil eines abstrahierten, durchschnittlichen Publikums zu adressieren, sondern als individuell rückverfolgbare Nutzerin. »Personalisierung« ist aber nicht mit einem Interesse für das Individuum per se zu verwechseln. Es geht in Momenten der Datafizierung immer um Relationen. Deshalb ist es treffender, unter Personalisierung formale Verfahren zu verstehen, welche der »entpersonalisierten« Wirtschaft des 20. Jahrhunderts eine »Repersonalisierung« (Hart, 2001, 2005; Moor & Lury, 2018) ökonomischer Relationen entgegenhalten und so die (vermeintliche) Dichotomie von persönlichen und ökonomischen Beziehungen unterlaufen. Unpersönliche Verfahren sollen quasi-persönliche Beziehungen zwischen Nutzerinnen und Unternehmen erzeugen, welche die ökonomischen Beziehungen in Gang halten. In Momenten der Datafizierung werden Personendaten erzeugt und in Wert gesetzt: Aus Daten abgeleitete »good matches« bilden den Kitt, der Unternehmen

und Nutzerinnen (und Dritte) zusammenhalten soll. Aus dieser Perspektive lässt sich Datafizierung und die damit verbundene Personalisierung als eine Verdichtung der Relationen zwischen Unternehmen, Nutzerinnen und Dingen (und möglicherweise Dritten) verstehen.