

Data cartels : the companies that control and monopolize our information / Sarah Lamdan. – Stanford, California : Stanford University Press, 2023. – XVIII, 197 Seiten. – ISBN 978-1-5036-3371-1 : USD 26.00 (auch als E-Book verfügbar)

Motivationen, ein Buch zu schreiben, gibt es viele – hier ist es die Geschichte einer Enttäuschung. Die Juraprofessorin Lamdan, die auch viele Jahre als Rechtsbibliothekarin arbeitete, verbrachte viel Zeit mit der Schulung der Datenbanken Westlaw (Thomson Reuters) und Lexis (RELX), ohne die in der amerikanischen Rechtswissenschaft nichts geht. Dann legte ihr 2017 jemand einen Artikel auf den Tisch, aus dem hervorging, wie Thomson Reuters und RELX darum wetteiferten, für die Einwanderungspolizei U.S. Immigration and Customs Enforcement (ICE) ein datengestütztes Überwachungssystem entwickeln zu können.

In der beginnenden Trump-Ära hatte ICE in Lamdans Darstellung seinen Ruf als brutale Organisation gefestigt, die Familien von Menschen auseinanderriss, die ihr ganzes Leben in den USA gelebt hatten, Razzien für Abschiebungen am Arbeitsplatz durchführte und Kinder in Käfige sperrte. Lamdan begann zu zweifeln. Sowohl für Bibliotheken wie für Rechtsanwält*innen gelten hohe Maßstäbe an Vertrauen und Vertraulichkeit – was geschah da? Schulte sie Einwanderungsanwält*innen an Instrumenten, die ICE halfen, deren Klient*innen zu verhaften? Flossen Daten von Bibliotheksutzer*innen an ICE? Lamdan tat das, was sie immer gemacht hatte, wenn sie Fragen hatte: Sie wandte sich an den Kundendienst. Hatte sich dieser bislang extrem bemüht um die fleißige Multiplikatorin der RELX-Produkte, ließ er nun schlagartig die Jalousie herunter – und mehr als das:

„RELX, especially, exerted its power. A Lexis representative started camping out at my law school, calling my work phone, and my boss, demanding that I speak to her manager. She also started monitoring me through my students, asking them to report back to her if I talked about LexisNexis’s ICE contracts. A blog post that a librarian and I wrote about the issue for the American Association of Law Librarians was erased moments after it was posted, and the organization replaced it with a single sentence: ‚This post has been removed on the advice of General Counsel.‘“ (S. Xf.).

Sarah Lamdan ließ sich nicht einschüchtern und schrieb dieses Buch. Sie nimmt sich darin Thomson Reuters und RELX als Muster eines Firmentyps vor, den sie *Datenkartell* nennt. Diese Kartelle könnten sich über viele Geschäftsfelder erstrecken: RELX und Thomson Reuters sind nicht nur in der Rechtsinformation dominant, sondern relevante Player in der Wissenschaftskommunikation insgesamt, außerdem bei Finanz- und Wirtschaftsinformationen, im Nachrichtengeschäft und – aktuell immer mehr – im Data Brokering. Diese Geschäftsfelder handelt Lamdan nacheinander ab und verdeutlicht dabei, dass die grundlegende Struktur dieser Kartelle immer gleich sei: Sie verwandelten wichtige öffentliche Güter in Klubgüter, um sie privilegiert beherrschen und vermarkten zu können.

Klubgüter privatisieren öffentliche Güter, indem sie den Zugang zu ihnen begrenzen und dadurch Exklusivität schaffen. Diejenigen, denen der Zugang möglich ist, haben deshalb nicht unbedingt ein

Interesse daran, die Zugangsschranke zu entfernen, denn die Exklusivität eines Beach Resorts hängt ja gerade daran, dass Hinz und Kunz dort eben nicht sein können – obwohl vielleicht Platz genug wäre. Klubgüter gibt es auch im digitalen Bereich und sie sind dort besonders umkämpft, weil es im Digitalen in der Regel keinen konkurrierenden Konsum gibt. Steht ansonsten das Schnitzel, das Hinz isst, für den Teller von Kunz nicht mehr zur Verfügung, können digitale Informationen verlustfrei kopiert und parallel konsumiert werden. Dennoch sorgen auch im Digitalen Klubgüter für Exklusivität durch Zugangsbeschränkung. Eine Methode hierfür ist uns in den Bibliotheken wohlvertraut: Es ist die Paywall. Das E-Book könnte viel mehr parallele Leser*innen haben, aber die Verlagslizenz begrenzt den Zugriff und stimuliert dadurch den Preis, was gerade bei Lehrbüchern momentan ein Thema ist,¹ und die Zustände bei den Journals wurden selbst vom Investmentbanking der Deutschen Bank schon als bizarr eingestuft.²

Wenn wir den Begriff Datenkartelle hören, denken wir an die großen Internetkonzerne mit ihren zweifelhaften Praktiken, die mit jahrzehntelanger Verspätung nun langsam Gegenstand der Regulierung werden wie in der aktuellen europäischen Datengesetzgebung, die explizit GAFAM, das Konglomerat von Google, Apple, Facebook, Amazon und Microsoft, in den Blick nimmt. Die Kartelle, die Lamdan behandelt, sind kleiner und segeln damit unter den Schwellenwerten, wie sie z.B. der *Digital Markets Act* der EU festlegt, hindurch – und oft genug auch unterhalb der Aufmerksamkeitsschwelle. Dagegen schreibt Sarah Lamdan ihr Buch: Ihre Kernaussage ist, dass die Datenkartelle, die sie anhand von RELX und Thomson Reuters behandelt, das ‚GAFAM der kuratierten Information‘ sind. Ihr Ziel ist ein ‚Google ohne Schrott‘ (vgl. S. 12) mit kondensierten Informationen für die Entscheidungsebene. Deshalb der grundsätzliche Hang zum Klubgut: Der Entscheider baut auf Informationen, die eben nicht allen zur Verfügung stehen. Das GAFAM der kuratierten Information ist folglich nicht weniger kritisch zu betrachten als etwa Meta Platforms Inc., bleibt aber weiterhin im Windschatten regulatorischer Aufmerksamkeit.

„The truth is paywalled, but the lies are free“³ – Datenkartelle schaffen damit für Sarah Lamdan einen mächtigen gesellschaftlichen *Bias*, eine Verzerrung, wie die Sozialwissenschaften sagen, da sie für die breite Masse der Bürger*innen die Möglichkeiten einschränken, informierte Entscheidungen über ihre Finanzen, ihre Gesundheit oder ihre Rechte zu treffen. Stattdessen werden über sie durch das Tracking Informationen gesammelt, damit sie dann Objekte der Entscheidungen anderer werden, statt selbst zu entscheiden. „Wer entscheidet, wer entscheidet?“ war schon für Shoshana Zuboff die prominente Frage im Überwachungskapitalismus.⁴ Lamdan folgt ihr in der Analyse, dass es sich hier um eine besondere Form der Ausbeutung handelt, weil es nicht irgendwelche Produkte sind, sondern geradezu intime Ausflüsse unserer Persönlichkeit, die monetarisiert werden:

1 Campaign to investigate the ebook market for libraries. Online: <<https://academicebookinvestigation.org/>>, Stand: 31.10.2022.

2 Vgl. Klein, Samuel: Turning the Supertanker. Deutsche Bank on Elsevier's excess, Notes from the Knowledge Futures Group, 05.08.2019, <<https://notes.knowledgefutures.org/pub/supertanker/release/3>>, Stand: 31.10.2022. Vgl. die Einordnung bei Gerhard Lauer: Datentracking in den Wissenschaften. Wissenschaftsorganisationen und die bizarre Asymmetrie im wissenschaftlichen Publikationssystem, in: o-bib 9 (1), 2022, S. 1–13. Online: <<https://doi.org/10.5282/o-bib/5796>>.

3 Robinson, Nathan J.: The truth is paywalled but the lies are free. The political economy of bullshit, Current Affairs, 02.08.2020, <<https://www.currentaffairs.org/2020/08/the-truth-is-paywalled-but-the-lies-are-free/>>, Stand: 31.10.2022.

4 Zuboff, Shoshana: Das Zeitalter des Überwachungskapitalismus, Frankfurt am Main; New York 2018, S. 382.

„The goods and services data analytics companies provide aren't merely products we buy and sell, they are products derived from us. Data analytics companies capitalize on the invisible, ethereal products of our minds. Industrials profited from mining, logging, and other trades that strip value from our land, but unlike historical extractive ventures, data isn't cut from trees or extracted from earth, it's harvested from humans. Companies like RELX or Thomson Reuters appropriate our informational labour. They siphon intellectual and social output, co-opting our artistic and scholarly work and vacuuming up the digital exhaust we release through our wearable devices and online activities. They take things we make, turning our output into something they can own and store like wheat in silos.“ (S. 6).

Diese Art von Transformation verdeutlicht zweierlei: Was ins Silo des Produkteigentums eingelagert werden kann, muss zum einen nicht die Form haben, die den tatsächlichen Bedürfnissen entspricht. Es gibt viele Formen der digitalen Wissenschaftskommunikation, die den Bedürfnissen der Forschung viel näherliegen als das PDF eines Papiermediums – aber damit übersteht man keine Evaluation, solange die beginnende Reform der Forschungsbewertung⁵ noch nicht weiter vorangeschritten ist. Zum anderen steckt hinter der technologischen Illusion immer eine soziale Beziehung, wie Meredith Whittaker es aktuell für die Künstliche Intelligenz (KI) verdeutlicht:⁶ Irgendjemand wählt die Trainingsdaten aus, über irgendwen sollen sie etwas aussagen, irgendwer zieht seine Schlüsse daraus – es ist eine Frage der Hierarchie und der Macht, die fortlaufend wächst und gestärkt wird auch durch die Entwicklungen früherer Jahre, wo aus Lamdans Sicht das Wettbewerbsrecht immer weiter eingeschränkt wurde, selbst wenn das Wettbewerber und die Demokratie schädigte.

Es ist auch diese Entwicklung, die die gegenwärtige Transformation hin zu einem APC-gesteuerten Open Access so fragwürdig macht: Diese lässt den Traffic auf den entsprechenden privaten Firmenplattformen, und damit kann das Informationsverhalten weiter beobachtet werden. Zusammen mit dem Content ist es für die Kartelle das, was für Google der stetig wachsende Berg vergangener Suchanfragen ist – Ausgangsmaterial für Vorhersageprodukte, wie Zuboff es nennt, die künftiges Handeln und damit Trends verkäuflich machen.⁷ Größe ist dabei für Lamdan der entscheidende Faktor (und der Treiber für die zahllosen Aufkäufe im Informationssektor). RELX und Thomson Reuters hätten dabei mittlerweile ein hohes Level erreicht, auf dem sie aus der Masse der ihnen zur Verfügung stehenden Daten wie Matjroschka-Puppen immer neue Datenprodukte hervorziehen könnten, ein *double dipping* mit Data Assets.

Die Datenprodukte erwachsen dabei nicht nur aus einem Größenvorteil, sie nützten auch den Großen am meisten: Konzernen und Regierungsinstitutionen, die entscheiden, wer versichert wird, wer überwacht, wer Sozialleistungen bekommt, wer das Sorgerecht für die Kinder. Datenkartelle reichten weit hinein in die staatliche Entscheidungsfindung: Thomson Reuters und RELX hätten beide hunderte an Verträgen mit öffentlichen Stellen allein in den USA und verfügten über Drehtüren beim Personal. Damit seien sie die „informational landlords that decide *who can swim in the ocean of knowledge*.“ (S. 16).

5 Vgl. z.B. die Coalition for Advancing Research Assessment (CoARA), <<https://coara.eu/>>, Stand: 31.10.2022, sowie den Paris Call on Research Assessment, <<https://osec2022.eu/paris-call/>>, Stand: 31.10.2022.

6 Wolfangel, Eva: „KI ist eine Technik der Mächtigen“ (Interview mit Meredith Whittaker), in: MIT Technological Review 2022 (6), S. 22–25.

7 Zuboff: Zeitalter des Überwachungskapitalismus, 2018, S. 22.

Lamdan sieht dabei durch den Übergang vom früheren Publisher zum jetzigen Data-Analytics-Unternehmen durchaus einen Qualitätsabbau für Nutzende bei den Datenkartellen, da die Technik bei allen Investments mit der Datenmasse überfordert sein könne. Zudem herrsche Intransparenz: Man sehe nur noch unendlich viele Daten hineinlaufen und Datenprodukte wieder herauskommen (vgl. S. 17). *Overpolicing*, die übermäßige polizeiliche Verfolgung bestimmter Personen oder Milieus, komme auch aufgrund von fehlerhaften Dossiers vor und dies nicht nur bei den bekanntermaßen benachteiligten Gruppen. Mit dem dynamischen Wachstum der Risk-Assessment-Geschäftssparten der Datenkartelle würden datenbasierte Entscheidungssysteme aber trotzdem in immer weitere Bereiche eingeführt wie Jugendämter, Vermietungen, Banken und Versicherungen. Das unscharfe Gerede von ‚Sicherheit‘, wenn eigentlich Überwachung gemeint sei, bewirke keinesfalls ein Sinken der Kriminalität, nur die Diskriminierung und die Zahl Inhaftierter stiegen dadurch an. Die öffentliche Hand, die den Aufstieg der Datenkartelle durch das schiere Verschenken ihrer Datenbanken in den 1990er-Jahren ermöglichte (vgl. S. 28–30), sei nun der größte Kunde – RELX allein verkaufe in den USA Datenprodukte an mehr als 7.500 Einrichtungen etwa der Polizei – und nutze die Möglichkeit, durch das ‚Data Shopping‘ rechtliche Prinzipien wie die Unverletzlichkeit der Wohnung zu umgehen. Für eine Datenbank braucht man keinen richterlichen Durchsuchungsbefehl, man fragt sie eben ab:

„Today, RELX runs its own private, third-party data center where thousands of law enforcement agencies consolidate and share their data. The company’s Public Safety Data Exchange compiles federal, state, and local law enforcement data, links it to our LexIDs, and is available to RELX customers in products with names like ‚Accurint Virtual Crime Center.‘“ (S. 37).

Aufgrund des sozialen Bias reagieren die Datenkartelle somit meist recht gleichgültig auf Fehler in ihren Systemen: Betroffen und benachteiligt sind im Zweifel die anderen, die 99 % – insofern sind Fehler nur technisch, nie eine Frage der Empathie oder der Angemessenheit. Schrieb Anatole France noch vom „Gesetz, das den Reichen wie den Armen verbietet, unter Brücken zu schlafen, in den Straßen zu betteln und Brot zu stehlen“,⁸ so hat jetzt nicht einmal mehr dieser Spott noch eine Grundlage.

Im Kapitel ‚Academic Research‘, das uns mit dem Bereich des Informationswesens am nächsten liegt, resümiert Lamdan viele allzu vertraute Dinge: die Machtposition der Verlage, die gerade durch die Metrifizierung der Wissenschaft nochmal deutlich gesteigert wurde, die Taktgeberfunktion, die Elsevier für die Branche in der Vergangenheit vielfach hatte, und die rechtlichen Zwänge aus der Digitalisierung, wenn Bibliotheken nicht mehr besitzen, sondern nur noch lizenzieren. Sie erinnert daran, dass der Zugang zu wissenschaftlichen Informationen laut UN-Charta ein Menschenrecht ist und Open Science mit Open Society zusammengedacht werden müsse – was vom praktizierten Klubgut-Modell der Verlage und deren Hinwendung zum Data Analytics Business komplett negiert werde: „Instead of focusing on selling critical scientific information at an affordable price, the company is concentrating on developing software that sifts through the company’s ‚vast corpus‘ of academic data to draw ‚insights‘ from – and monetize – the entire research process.“ (S. 51).

8 France, Anatole: Die rote Lilie. Roman. Aus dem Französischen übersetzt von Caroline Vollmann, Zürich 2003, S. 147.

Ein solches Vorgehen erfordert natürlich Investitionen, und Lamdan rechnet vor, wo die nötigen Gelder dafür herkommen: „If it only costs Elsevier \$ 600 to publish an article and Elsevier makes \$ 4.000 from selling it, what happens with the other \$ 3.400?“ (S. 62). Selbst abzüglich einer Renditerate von 30 % verbleibt eine Menge Geld, die von Elsevier und genauso von den anderen großen Playern genutzt wird, um den Markt an Innovationen im Feld der *scholarly communication* leerzukaufen⁹ und ganze Daten-Ökosysteme zu bauen. Bibliotheken bezahlen damit nicht nur zu hohe Preise, sondern subventionieren den ‚Extraktionsimperativ‘, wie Zuboff es nennt: Aus der Menge an Daten werden neue Produkte geschaffen, die wiederum noch mehr Daten generieren, die wiederum extrahiert und monetarisiert werden – während das Gesamtsystem fortlaufend immersiver wird im Bemühen, die Anreizsysteme und damit den Verlauf der Forschung gemäß der eigenen Prämissen, nicht denen der Wissenschaft, zu modellieren. Diese Verkettung von einzelnen Anwendungen zu ganzen Werkbänken ist nach Lamdan an den Hochschulen kaum jemandem bewusst, da niemand den Überblick über alle Workflows habe (vgl. S. 66).

Mit der Hinwendung zu KI-Systemen steigerten die Verlage ohnehin bestehende Ungleichheiten in der Wissenschaft nochmals erheblich, denn „machine-generated journal impact metrics and researcher ranking systems incorporate the same biases, white supremacy, and systemic racism that permeate academia, favoring the work of white men at elite institutions. Academic data analytics don’t repair the racist and misogynistic decisions made in university tenure committees and college boards, they datafy them.“ (S. 63f.).

Diese durch die Einführung von KI-Systemen produzierte Verzerrung macht sich im folgenden Kapitel für Lamdan auch in der Rechtsinformation bemerkbar. Die Plattformen seien durch Assistenzsysteme und Algorithmen darauf gepolt, den Nutzenden, die sie ja bestens kennen, Lesarten des Rechts vorzuschlagen, erfolgversprechende Fälle und Gegenargumente – eine eigenständige Suche und Urteilsbildung durch die Nutzenden findet damit immer seltener statt. Durch dieses Anleiten gewinnen die von RELX und Thomson Reuters vertriebenen Datenbanken eine rechtsformende Kraft – und das vor dem Hintergrund, dass die ordnende Grundsystematik der Systeme aus der Zeit des Amerikanischen Bürgerkriegs stamme. „It’s not surprising that products based on John B. West’s 1800s-era legal classification system ‚reflect a nineteenth century worldview‘ that institutionalizes the white, male, heteronormative, upper-class, able-bodied and political conservative perspective.“ (S. 76).

Rechtsinformation sei dabei genauso als kritische Infrastruktur zu werten wie die Wissenschaft, und die urheberrechtsbewehrte Monopolisierung zum Klubgut für alle Abhängigen verheerend: etwa 70 % des Budgets einer *law firm* gingen an RELX und/oder Thomson Reuters. Für Menschen, die sich keine*n Anwält*in leisten können, werde sogar die einfachste Ausübung ihrer Rechte schier unmöglich, denn der Staat ließ die Epoche der Papierdokumente auslaufen und migrierte auf digitale Formate – schuf aber keine angemessene Infrastruktur für deren Benutzung. In der Folge sind diejenigen, die in Abschiebezentren einsitzen, weil ICE sie auf Basis der Daten, die RELX und Thomson Reuters liefern,

9 Vgl. die Visualisierungen in Universität Utrecht: Innovations in scholarly communication. Changing research workflows, <<https://101innovations.wordpress.com/>>, Stand: 31.10.2022.

aufgespürt hat,¹⁰ völlig abhängig von den digitalen *Prison Products* der gleichen Firmen, mit stark eingeschränkten Funktionen und in einer Sprache, die viele der Betroffenen nicht verstehen. Dass diese Kiosk-Rechner ihrerseits vollgestopft sind mit Tools zur Überwachung aller Online-Aktivitäten derer, die sie benutzen, versteht sich dabei von selbst (vgl. S. 85f).

Finanzdaten sind für Lamdan wiederum anders gelagert als Fach- oder Rechtsinformation: Es handle sich dabei nicht um *publisher-grade information*, aber auch nicht um unstrukturierte Informationen. Dieser Zwischenbereich ist teils schwierig zu bearbeiten; im Rahmen der Nationalen Forschungsdateninfrastruktur unternimmt es in Deutschland das BERD-Konsortium¹¹. Die Produkte der Datenkartelle würden von ihnen dabei sehr verwaschen beschrieben, dafür aber premiumbepreist. Fehlerhafte Daten ‚korrigieren‘ sich laut Lamdans Darstellung, da alle sie benutzten, als seien sie richtig. Es entsteht gerade in der Digitalisierung eine immer extremere Asymmetrie zwischen Datennutzer*innen und der allgemeinen Öffentlichkeit in Bezug auf die Möglichkeit, ökonomische Vorteile zu erzielen.

Dabei hätten Finanzdaten in der Vergangenheit schon eine Ähnlichkeit mit der Rechtsinformation besessen, denn der Zugang zu ihnen sei über lange Zeit gesetzlich geschützt gewesen, um z.B. die Gefahr einer Bankenpanik aufgrund von gestreuten Gerüchten einzudämmen – jede*r sollte sich selbst informieren können. Die Aufweichung dieses Informationsschutzes geschah dann nach Lamdan zum einen durch den stillen Boykott der Firmen, die die Informationen, die sie zu berichten verpflichtet waren, in PR-Schutt vergruben, und zum anderen durch die Datafizierung, die diejenigen belohnt, die den Geschwindigkeitsvorteil ausspielen können. Der geschützte Informationszugang wird unter Digitalbedingungen nahezu unbrauchbar, denn der hochpreisige Service liefert die beste Zusammenschau schon per Knopfdruck und begründet damit Abhängigkeiten und halbseidene Praktiken z.B. zur Umgehung der Regulierungen gegen Insiderhandel (vgl. S. 109).

Lamdan thematisiert nicht im Detail, dass die Finanzdatenprodukte auch eine enge Verwandtschaft mit dem Risk Assessment haben, also dem Data Brokering für die Zielgruppe der Entscheider. Deutlich ist das bei Datenbanken wie Thomson Reuters *World-Check* (mittlerweile *Refinitiv*), die sogenannte *PEPs*, ‚Politically Exposed Persons‘, verzeichnen. Offizieller Sinn dieser Produkte ist u.a. die Minimierung von Geldwäsche: Erscheint in einer Filiale der Credit Suisse eine Person, die ein diskretes Konto wünscht und viel Geld in Aussicht stellt, so sollen die Bankangestellten anhand solcher Instrumente prüfen, ob sie den Schwiegersohn eines Kleptokraten oder eine Oligarchentochter vor sich haben. Nun haben u.a. die Recherchen zu den *Panama Papers*¹² und den *Swiss Secrets*¹³ gezeigt, dass das Konto mit hoher Wahrscheinlichkeit eröffnet wird, egal was in diesen Datenbanken steht. Den investigativen Journalisten offenbarten sich beim Blick in von solchen Data Brokern bereitgestellten Produkten aber auch jede Menge Trackingdaten: Welche Mobilgeräte stehen z.B. mit welcher Person in Verbindung? Mit

10 Vgl. dazu auch: Georgetown Law Center on Privacy & Technology: American Dragnet. Data-driven deportation in the 21st century, 10.05.2022, <<https://americandragnet.org/>>, Stand: 31.10.2022.

11 BERD@NFDI (NFDI Consortium for Business, Economic and Related Data), <<https://www.berd-nfdi.de/>>, Stand: 31.10.2022.

12 Vgl. Obermayer, Bastian; Obermaier, Frederik: Panama Papers. Die Geschichte einer weltweiten Enthüllung, Köln 2021.

13 Vgl. Munzinger, Hannes; Obermaier, Frederik; Obermayer, Bastian: Schweizer Geheimnisse. Wie Banker das Geld von Steuerhinterziehern, Foltergenerälen, Diktatoren und der katholischen Kirche versteckt haben – mit Hilfe der Politik, Köln 2022.

welchem Gerät ist der WhatsApp-Account verknüpft? In welche Funkzelle bucht sich das Gerät ein?¹⁴ So wenig sich Datenkartelle in ihrem sozialen Bias für die 99 % interessieren, so sehr entwickeln sie sich offenbar zu einer Service-Infrastruktur für das 1 %.

Sowohl Thomson Reuters wie RELX (vormals Reed Elsevier) haben auch starke Wurzeln im Nachrichtengeschäft. In ihrem abschließenden Kapitel widmet sich Lamdan der Situation in diesem Feld, das aus ihrer Sicht von einem Dualismus gekennzeichnet ist: Sind Nachrichten einerseits eine essenzielle öffentliche Ressource für eine aufgeklärte demokratische Gesellschaft, so dienen sie andererseits zugleich als Futter für Data Analytics Tools. Waren Nachrichten wie Rechtsinformationen lange geschützt, hier durch einen öffentlich-rechtlichen Bereich, so können News heute ebenso wie Rechts- und Forschungsinformationen mit anderen Daten verknüpft werden für Vorhersageprodukte (vgl. S. 112).

In ihrer historischen Ableitung sieht Lamdan die Publics News als wichtigen Teil der Informationsinfrastrukturen von vitalen Demokratien mit entsprechendem Vertrauenskapital; die Digitalisierung im News-Bereich habe dagegen katastrophale Folgen gehabt: In weniger als einem Jahrzehnt habe die US-Zeitungsindustrie mehr als die Hälfte der Angestellten verloren, wobei hauptsächlich die ‚Provinz‘ betroffen gewesen sei, die jeden Zugang zu den Lokalnachrichten verlor. Zuvor hätten seit den 1970ern unter dem Einfluss der Chicago School die Public News bereits erheblich an Boden verloren (vgl. S. 114). Mit der Verbreitung des Internets sei dann der Schutz öffentlich-rechtlicher Nachrichtensender nochmals bewusst geschwächt worden, bis selbst die *Sesamstraße* hinter der Paywall verschwand.

Der News-Bereich verhielte sich unter der Herrschaft der Datenkartelle damit nicht anders als andere Informationsbereiche: „private companies swoop in and take over information infrastructure that the public fails to adequately support.“ (S. 117). Das Tracking im Newsbereich ist noch ausgebauter als im Wissenschaftsbereich;¹⁵ es zählten Nutzerdaten und Klicks, nicht die Produktqualität, und der demokratische Diskurs verblasst, wie die letzten Jahre gezeigt haben.

In ihren abschließenden Bemerkungen bündelt Sarah Lamdan einige Punkte, die sie für zentral erachtet, um einen Ausweg aus der jetzigen Situation zu finden:

- Nicht in der eigenen jeweiligen Sparte denken, sondern Datenkartelle als Ganzes in den Blick nehmen. Ihre grundlegende Struktur, die Wucht, die sie aus ihrem Größenvorteil ziehen, würden nicht deutlich, wenn man sie z.B. als Verlage ansehe, nur weil sie es früher u.a. einmal waren.
- Informationen müssten als öffentliches Gut betrachtet werden und Informationsinfrastrukturen damit genauso gefördert und organisiert werden wie öffentliche Straßen und Wasserleitungen.

¹⁴ Vgl. ebd., S. 162ff.

¹⁵ So recherchierte Richard Gutjahr z.B. 470 Tracker auf der Webseite sueddeutsche.de, vgl.: Richard Gutjahr: Der Cookie Wahnsinn. Cookie-Fallen – wir pfeifen auf Ihre Privatsphäre, 11.12.2020, <<https://www.gutjahr.biz/2020/12/wir-pfeifen-auf-ihre-privatsphaere/>>, Stand: 31.10.2022.

- Firmen müssten besser kontrolliert und Oligopole vermieden werden. Die jetzige Situation erinnere an das ‚Gilded Age‘, den Raubtierkapitalismus im ausgehenden 19. Jahrhundert. Der bisherige Exzeptionalismus im Digitalbereich, dass etwas nicht reguliert werden könne, nur weil es Technologie sei, müsse beendet werden.
- Konkurrenz müsse gefördert und geschützt werden, um die Monopolisierung ganzer Märkte zu verhindern. Die Monopolisierung fördere auch Nachteile durch Fehler: Die Masse der einströmenden Informationen überfordere die Techfirmen und bewirke Qualitätsverluste bei deren Produkten.
- Informationszugang müsse möglich sein ohne Datenausforschung. Sonst bliebe auch Open Access ein zwiespältiger Schlüssel: Es müsse eine Trennung zwischen den Informationsinfrastrukturen und den Data-Analytics-Strukturen geben. Dann könne OA auch mit privaten Eignern funktionieren.

Eines der Probleme, das hier aufscheint, ist natürlich dieses: Den Bau einer Straße kann man als öffentlichen Auftrag ausschreiben, genauso wie eine Fährlinie. Was ist aber mit einer so komplexen Infrastruktur wie GitHub? Eine Ersetzbarkeit der Anbieter wird dort an ihre Grenzen kommen. Hinzu kommt, dass die Daten selbst nach Lamdans Darstellung keine Infrastruktur seien, sondern das Internet. Datenunternehmen seien daher keine Versorgungsunternehmen (könnten aber trotzdem entsprechend reguliert werden wie z.B. Radiosender), Daten zudem stets einzigartig im Gegensatz zu Strom und Leitungswasser. Für das Informationswesen ist ihr Vorschlag daher eine digitale Bibliothek, getragen von zentralen öffentlichen Mitteln, die aus einer Besteuerung von Big Tech stammen: Steuerschlupflöcher sollten geschlossen werden zugunsten einer öffentlichen digitalen Infrastruktur – das Internet Archive oder der Hathi Trust seien hierfür Beispiele, denen der öffentliche Bereich nachzueifern sollte.

Data Cartels bündelt eine Menge an Informationen, die mühsam durch Informationsfreiheitsanfragen gesammelt wurden, die in Tweets, Blogs, Firmenschriften und einer Reihe anderer Reports, die man früher unter der Rubrik graue Literatur gefasst hätte, verstreut sind. Es ist Sarah Lamdan zu danken, dass sie dieses weit verstreute Puzzle zusammengefügt und uns damit eine zusammenhängende Darstellung gegeben hat, so unangenehm sie für uns ist. Wir schauen darauf ähnlich fröstelnd wie auf eine Folge von *Black Mirror*. Dabei ist es schon längst unser aller *Truman Show*.

Renke Siems, Universitätsbibliothek Tübingen

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/5902>

Dieses Werk steht unter der Lizenz [Creative Commons Namensnennung 4.0 International](https://creativecommons.org/licenses/by/4.0/).