Das Lesen der Anderen

Die Auswirkungen von User Tracking auf Bibliotheken

Renke Siems, Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg, Stuttgart

Zusammenfassung
Die großen Wissenschaftsverlage entwickelten sich in den vergangenen Jahren weg von einem verlegerischen Inhaltsanbieter hin zu einem Data Analytics Business. Als Plattformunternehmen erzielen sie hohe Margen und nutzen dieses Kapital, um aus der Wissenschaftscommunity entstehende Alternativangebote aufzukaufen und sich in weitere Geschäftsfelder auszudehnen. Ziel ist es, sich in allen zentralen Prozessen der Wissenschaftssteuerung unverzichtbar zu machen, sodass dann wie im Informationsbereich von einem *vendor lock-in* gesprochen werden muss. Zu diesem Zweck haben die Verlage ihre Plattformen mit Instrumenten für ein umfassendes *User Tracking* ausgestattet. Zugleich versuchen sie, die Zugangsauthentifizierung unter ihre Kontrolle zu bringen, um den personalisierten Zugriff auf alle Nutzenden sicherzustellen. Einige Verlage oder deren Mutterkonzerne verflechten sich auch mit der Sicherheitsindustrie und (halb-)staatlichen Akteuren zu undurchsichtigen Datengeschäften, bei denen auch die Hochschulnetze in den Blick geraten. Der Aufsatz versucht, diese Entwicklung zu analysieren und Konsequenzen zu formulieren.

Summary
In recent years, the major science publishers have evolved away from publishing content providers to data analytics businesses. As platform companies, they generate high margins and use this capital to buy up alternative offers emerging from the science community and to expand into other business areas. The goal is to make themselves indispensable in all central processes of science control, so that we should see this as a vendor lock-in, just as it is known from the information sector. To this end, publishers have equipped their platforms with tools for comprehensive user tracking. At the same time, they are trying to bring access authentication under their control in order to ensure personalized access to each user. Some publishers or their parent corporations also intertwine with the security industry and (semi-)government actors in opaque data deals that also bring university networks into view. This essay attempts to analyze this development and outline the consequences.

Zitierfähiger Link (DOI): https://doi.org/10.5282/o-bib/5797

Autorenidentifikation:
Siems, Renke: GND: 1241986576; ORCID: https://orcid.org/0000-0002-9824-5449

Schlagwörter: Wissenschaftlicher Verlag; Elektronische Medien; Kapitalismus; Informationsgesellschaft; Technologieunternehmen; Monopol; Big Data; Personenbezogene Daten; Datenanalyse; Elektronische Überwachung; Manipulation

1. Tracking

Digitale Medien dominieren in wissenschaftlichen Bibliotheken und ein Ende des Wachstums ist nicht zuletzt angesichts des durch die Pandemie ausgelösten Schubs kaum absehbar. Die weiträumige Schließung von Bibliotheken und die Umstellung auf digitale Lehre haben Informationsverhalten und Medienpraktiken auch in den Disziplinen verändert, die in dem Bereich nicht als *early adopters* aufgefallen waren. Da eine völlige Rückabwicklung etwa im Bereich der Lehre nicht zu erwarten sein wird und auch Erwerbungsentscheidungen der Bibliotheken angepasst wurden, wird diese Veränderung nachhaltig sein und sich fortschreiben – umso mehr, als digitale Medien in der Wissenschaft sich in den breiten Trend eines technisch assistierten Lesens und Schreibens integrieren – vom Gebrauch kollaborativer Schreibwerkzeuge bis hin zu Textmining. Die Digitalisierung der Wissensarbeit bedarf dabei als Grundlage ein neues Verständnis von guter wissenschaftlicher Praxis, weil diese veränderte Medienpraxis anderen Abhängigkeiten ausgesetzt ist als ehedem und daher anders gelagert auf Redlichkeit vertrauen können muss – leider ist allzu oft das Gegenteil der Fall.

Dass aus einem assistierten Lesen schnell ein ausgeforschtes Lesen wird, ist im Fall von E-Book-Readern schon sehr lange in der Diskussion. Auswahl der Lektüre, Lesedauer, Lesegeschwindigkeit, Markierung von Stellen – die Intransparenz darüber, was mit diesen gesammelten Daten geschieht, wird nur vom Verschwinden einiger Bücher von den Geräten ihrer Leserinnen überboten.[1] Clifford Lynch hat diese Entwicklung vor einigen Jahren analysiert[2] und bei seinen Ausführungen zur wissenschaftlichen Literatur dabei auf einen Blogpost von Eric Hellman von 2015 verwiesen, in dem dies bereits thematisiert wird.[3] Man muss deshalb davon ausgehen, dass es eine Überwachung des Informationsverhaltens von Wissenschaftler*innen und Studierenden wahrscheinlich kaum weniger lang gibt als eine Auswertung des privaten Lesens in elektronischer Form. Über etliche Jahre werden also wohl schon individualisierte Nutzungsprofile von wissenschaftlicher Fachinformation erstellt, die Daten ausgewertet und auf bislang meist noch unbekannte Weise gehandelt.

Im Gegensatz zum Geschehen bei den digitalen Medien für private Endkunden blieb die Überwachung des Informationsverhaltens im Wissenschaftsbereich jedoch längere Zeit eher unterm Radar. Erst der Streit um veränderte Authentifizierungsmethoden für die Verlagsplattformen brachte die für das Thema grundlegenden einschlägigen Beiträge von Cody Hanson[4] hervor, weitere Initiativen aus dem Wissenschaftsbereich[5] sowie eine offizielle Stellungnahme seitens des Ausschusses für

1 Vgl. Stone, Brad: Amazon erases Orwells Books from Kindle, in: New York Times 17.07.2009. Online: <https://www.nytimes.com/2009/07/18/technology/companies/18amazon.html?_r=2>, Stand: 20.02.2022.

2 Vgl. Lynch, Clifford: The Rise of Reading Analytics and the Emerging Calculus of Reader Privacy in the Digital World, First Monday 4 (22), 2017. Online: <https://doi.org/10.5210/fm.v22i4.7414>.

3 Vgl. Hellman, Eric: 16 of the top 20 research journals let ad networks spy on their readers, Go to Hellman 12.03.2015. <https://go-to-hellman.blogspot.com/2015/03/16-of-top-20-research-journals-let-ad.html>, Stand: 20.02.2022.

4 Vgl. Hanson, Cody: User Tracking on Academic Publisher Platforms, <https://www.codyh.com/writing/tracking.html> sowie die dort verlinkten aufgezeichneten Vorträge. Stand: 20.02.2022.

5 Vgl. SPARC: Adressing the Alarming Systems of Surveillance Systems built by Library Vendors, <https://sparcopen.org/news/2021/addressing-the-alarming-systems-of-surveillance-built-by-library-vendors/> und den von der ZBMed getragenen Aufruf: „Stop Tracking Science!"; <https://stoptrackingscience.eu/> Dort findet sich auch reichhaltig weitere Literatur zum Thema. Beide Stand: 20.02.2022.

Wissenschaftliche Bibliotheken und Informationssysteme (AWBI) der Deutschen Forschungsgemeinschaft (DFG).[6]

Das Informationspapier des AWBI ist grundlegend für die Thematik: es erläutert knapp den aktuellen Wandel bei den großen Wissenschaftsverlagen[7] und ihre Hinwendung zu einem Data Analytics Business, einige Methoden der Datengewinnung und zieht ein erstes Fazit, welches zu einem breiten Diskurs über die möglichen Rechtsverletzungen durch das Tracking und die Gefährdungen für Wissenschaftler*innen aufruft. Die hier gegebene Darstellung der Trackingmethoden basiert auf Vorträgen auf dem Bremer Bibliothekartag von 2021. Sie schließt an das Informationspapier an und versucht, weitere Punkte zu ergänzen.[8]

1.1. First Party Data

Sogenannte *First Party Data* sind die Daten aus erster Hand: Login-Daten, E-Mail-Verteiler für Newsletter und Autorenverzeichnisse wären Beispiele aus dem Bereich der Informationsversorgung. Solche Daten sind aus Sicht der Data Analytics besonders wertvoll, weil es echte Daten echter Menschen sind und nicht nur – wie bei manch anderen Trackingverfahren – bloß errechnete Zuordnungen. Entsprechend intensiv ist im kommerziellen Internet das Bemühen um Zugriff auf diese Art Daten. Das können z.B. Identifikationsnummern von Geräten sein oder auf Webseiten Features wie „Remember me on this device“ oder „Login with…“. Ziel ist immer eine klare Identifizierung zu Beginn, um von dort aus die Onlinebiographie der betreffenden Person bruchlos verfolgen und mit weiteren Informationen synchronisieren zu können.

Diese Intention muss man auch hinter den Bemühungen der Verlage vermuten, durch technische Veränderungen die Nutzerauthentifizierung in ihren Zugriff zu bekommen. Den von den Verlagen vorgestellten Initiativen Resource Access (RA21), Seamless Access und Get Full Text Research (GetFTR) ist gemeinsam, dass sie wie die gebräuchliche Shibboleth-Anwendung auf der SAML-Technologie[9] aufsetzen. Jedoch integrieren sie diese viel stärker als bislang in ihre Plattformen und nehmen damit zunehmend Einfluss auf die Frage, welche Daten zur Authentifizierung übertragen werden oder eben nicht. Eine gemeinsame Erklärung von DBV und der Allianz der deutschen Wissenschaftsorganisationen kam daher hinsichtlich RA21 zu einem klaren Ergebnis:

„Diesen Vorteilen steht potentiell die Gefahr gegenüber, dass Anbietende im Rahmen der Weiterentwicklung von SSO [Single Sign On; d. Verf.] die Weitergabe von personenbezogenen Informationen durch den institutionellen Identity Provider über das erforderliche Maß hinaus als Vorbedingung für die Nutzung festlegen. Auch um Einrichtungen nicht zu benachteiligen, die keinen eigenen Identity

6 DFG: Datentracking in der Wissenschaft. Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Bonn 2021. Online: <https://www.dfg.de/download/pdf/foerderung/programme/lis/datentracking_papier_de.pdf>, Stand: 20.02.2022.

7 Ausführlicher hierzu ist SPARC: Landscape Analysis, 29.03,2019, <https://infrastructure.sparcopen.org/landscape-analysis> und die Updates hierzu: <https://sparcopen.org/news/2020/sparc-releases-update-to-landscape-analysis-and-accompanying-interactive-website/>, Stand: 20.02.2022.

8 Für eine konzentrierte Überblicksdarstellung von Web Tracking insgesamt vgl. Reuben Binns: Tracking on the Web, Mobiles and Internet-of-Things, arXiv 26.01.2022, <https://arxiv.org/abs/2201.10831>, Stand: 20.02.2022

9 Security Assertion Markup Language, ermöglicht Single-Sign-On-Verfahren zur Authentifizierung.

Provider betreiben, ist eine ausschließliche Festlegung auf SSO-basierte Verfahren aus Sicht der Informationsinfrastruktureinrichtungen nicht wünschenswert. Grundsätzlich sollte immer eine IP-basierte Zugangskontrolle als Alternative angeboten werden. Bei der Umsetzung webbasierter SSO-Verfahren muss sichergestellt werden, dass der Datenschutz auch im Sinne des Prinzips der Datensparsamkeit in vollem Maße umgesetzt wird (privacy by design)."[10]

Die Verlage verfolgen ihre Strategie davon unbeeindruckt und mit fragwürdigen Methoden weiter: So setzte die American Chemical Society US-amerikanischen Bibliotheken Opt-Out-Fristen zur Umstellung auf Seamless Access mitten in den ersten Corona-Lockdown.[11] In Deutschland spielt Seamless Access bislang keine Rolle, allerdings wird von den Verlagen immer häufiger Google CASA (Campus Activated Subscriber Access) angeboten. Google CASA ermöglicht als Teil des „Google Scholar Subscriber Links"-Programms, dass bei Recherchen auf dem Campus über Google Scholar und bestehender Anmeldung im Google-Account Google sich die Zugriffsrechte merken kann und den Zugriff dann auch off-campus ermöglicht unter Umgehung der sonst nötigen Remote Access-Systeme.[12] Was dabei wie lange wo gespeichert wird, ist genauso unklar wie wer die Daten alles sieht. Da eine Reihe von Verlagen den Service benutzen, wäre denkbar, dass sie auf diese Weise Informationen zu verlagsübergreifenden Nutzungsmustern einzelner Personen erhalten.

Ergänzt werden solche Bestrebungen des Sammelns von First Party Data durch die Einbettung eines Hash in die Metadaten eines Artikel-PDFs, der sich bei jedem Download verändert. In Verbindung mit dem Zeitstempel des Downloads wird dadurch jedes PDF individualisierbar. So werden nicht nur Login-Daten und Informationsverhalten erfasst, sondern potenziell auch die Verbreitung von Dokumenten z.B. durch *author sharing*.[13]

1.2. Third Party Data

Unter *Third Party Data* versteht man, dass nicht nur die eigenen, selbst erhobenen Daten ausgewertet werden, sondern eine externe Drittpartei diese mit Daten aus anderen Quellen gemeinsam analysiert. Google CASA bildet somit eine Gelenkstelle zwischen First und Third Party Data, da es die direkte Identifizierung leistet wie bei einer Authentifizierung, aber eigentlich zu den Services von Drittparteien gehört. Diese Third Parties mit ihren integrierten Datenanalysen sind das gängige Erlösmodell im kommerziellen Internet. Wann immer ein Cookiebanner auf einer Webseite aufploppt (und oft genug auch, wenn nicht), ist Tracking durch das Arsenal von Advertising Technology (AdTech) im Spiel.

10 Empfehlungen zu Methoden zur Kontrolle des Zugriffs auf wissenschaftliche Informationsressourcen. Ein gemeinsames Papier von Deutscher Bibliotheksverband e.V. (dbv) und Schwerpunktinitiative Digitale Information der Allianz der deutschen Wissenschaftsorganisationen, 27.11.2019, Online: <https://dbv-cs.e-fork.net/sites/default/files/2020-11/2019_11_27_dbv_Stellungnahme_Empfehlungen%20zu%20Methoden%20zur%20Kontrolle%20des%20Zugriffs%20auf%20wissenschaftliche%20Informationsressourcen.pdf>, Stand: 20.02.2022, vgl. aktuell auch McLean, Jaclyn; Stregger, Elizabeth: Sounding the Alarm. Scholarly Information and Global Information Companies in 2021, in: Partnership. The Canadian Journal of Library and Information Practice and Research 2 (16), 2021, S. 1–7. Online: <https://doi.org/10.21083/partnership.v16i2.6692>.

11 Vgl. Hanson, Cody: E-resource librarians, Twitter, @codyh, 26.03.2020, <https://twitter.com/codyh/status/1243250490403483648>, Stand: 20.02.2022.

12 Vgl. Vogel, Christian: Kennen Sie Google CASA? medinfo. Informationen aus Medizin, Bibliothek und Fachpresse, 08.07.2020, <https://www.medinfo-agmb.de/archives/2020/07/08/6880>, Stand: 20.02.2022.

13 Vgl. Saunders, Jonny: More fun publisher surveillance, Twitter, @json_dirs, 26.01.2022, <https://twitter.com/json_dirs/status/1486120144141123584>, Stand: 20.02.2022.

Immer geht es darum, die durchgehende Geschichte des Onlineverhaltens individualisierter Personen verfolgen zu können, um entweder aggregierte Informationsverhaltensprofile zu vermarkten oder durch manipulative Nutzerlenkung (nudging) Einfluss auf Verhaltensentscheidungen zu nehmen.[14] Auf den großen Verlagsplattformen konnte Cody Hanson 139 solcher Third Parties nachweisen. Er erläutert deren Problematik:

„The reason I was interested in third-party assets being loaded on these sites is that any JavaScript loaded on these pages has access to the entire DOM, or document object model, meaning it can read the address and contents of the page. It also has access to every user action that happens on that page, and can itself load additional scripts from additional sources. So when, for example, a publisher puts JavaScript from Google on its pages, Google can record any information from the page about the article being sought, or search terms from a library user in the publisher platform. Fourteen of the fifteen publisher platforms included Google code on the article page."[15]

Neben den großen Firmen im Trackinggeschäft wie Google und Facebook sind eine Reihe weiterer Tracker zu finden sowie „Meta-Tracker" in Form von datenaggregierenden Audience-Tools zur Zielgruppenanalyse („targeting") von Adobe, Neustar, Oracle, AddThis usw., die Daten aus vielen Quellen zusammenführen und damit die getrackten Personen in Merkmalsgruppen klassifizieren (Alter, Geschlecht, Interessen, sexuelle Orientierung, Einkommen usw.). Hinzu kommen sogenannte Finger Printer bzw. Canvas Printer wie Double Click, die Personen anhand ihrer Gerätespuren identifizieren, obwohl diese das durch ihre Browsereinstellungen zu vermeiden versuchen. Da z.B. Audience Tools mit einer Vielzahl weiterer Data Broker zusammenarbeiten, fließen die so gesammelten Daten in ganze Ökosysteme der Datenverwertung, wie es Wolfie Christl am Beispiel des Marketing-Dienstleisters Acxiom/liveramp visualisiert:[16]

14 Zur Entwicklung dieser Problematik vgl. Crain, Matthew: Profit over Privacy. How Surveillance Advertising Conquered the Internet. Minneapolis 2021.

15 Hanson 2019; s. Anm. 4.

16 Vgl. Christl, Wolfie: Corporate Surveillance in Everyday Life. How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions. Wien 2017, S. 55. Online: <https://crackedlabs.org/en/corporate-surveillance>, Stand: 20.02.2022.

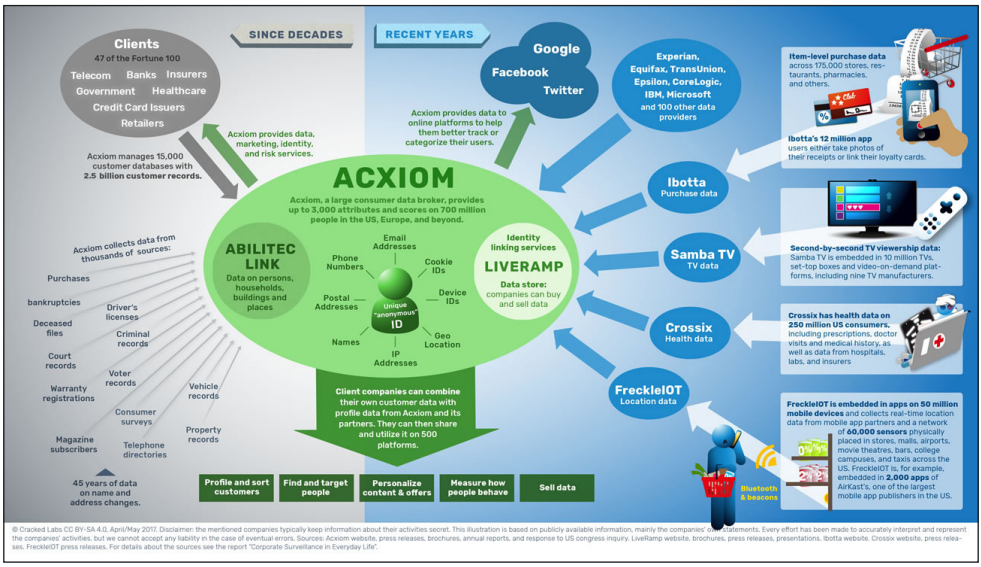

Abb. 1: Datenflüsse Acxiom (CC-BY-SA 4.0 Wolfie Christl)

Acxiom/liveramp ist u.a. auf der *Nature*-Website nachweisbar. Aufgrund dieser Zusammenarbeit der Verlage mit AdTech ist damit das bisher eher geschützte „Sonderbiotop" der wissenschaftlichen Informationswelt aufgelöst. Was eine Wissenschaftlerin oder ein Wissenschaftler recherchiert, verbindet sich in einer ungebrochenen Online- (und teils Offline-)Biographie mit dem Stöbern im Online-Shop, der Twitter-Timeline, den bevorzugten Zeitungen, der Information, ob man Kinder hat, der Joggingstrecke und den Einkäufen im Supermarkt um die Ecke.

1.3. Bid-Streaming und Portscanning

Das Geschäft der Third Parties ist seit langem als *surveillance advertising* (Überwachungswerbung) in der Kritik. Wichtige Internetunternehmen gehen langsam auf Distanz und auch der Digital Services Act der EU unternimmt erste, wenn auch unzureichende Regulierungen.[17] Die Entwicklung ist aber uneinheitlich: Auf wirtschaftlichen Druck hin wurden in Deutschland im § 26 Telekommunikation-Telemedien-Datenschutz-Gesetz (TTDSG) Regelungen zu Personal Information Management Services (PIMS) aufgenommen, die aufgrund der mangelhaften Anlage letztendlich Third Party Data zu First Party Data zu machen drohen, indem nach Anmeldung bei einer zentralen Webseite Einwilligungen „auf Vorrat" gegeben werden können, damit hinterher die Cookie-Banner entfallen. Dadurch wird der AdTech-Industrie der Spielraum wieder zurückgegeben, den Browser-, Geräte- und Diensteanbieter

17 Vgl. Röttger, Tania: Auf dem Weg zum Digital Services Act. Wie die EU Gesetze gegen Desinformation macht, Correctiv 26.03.2021, <https://correctiv.org/faktencheck/hintergrund/2021/03/26/auf-dem-weg-zum-digital-services-act-wie-die-eu-gesetze-gegen-desinformation-macht/>, Stand: 20.02.2022.

ihr gerade nehmen.[18] PIMS könnten bei entsprechender Modellierung helfen, Datenschutzverbesserungen für Nutzende effizient zu gestalten, doch der jetzige Stand wird in der Fachdiskussion als „Regulieren am Problem vorbei" kritisiert.[19] Auch der zunehmende Rückgriff auf First Party Data, was die Aktivitäten zur Authentifizierung und die Etablierung von Google CASA seitens der Verlage motiviert, ist entsprechend ein Ziel der AdTech-Industrie:

„Leaving the cookie behind means that brands and publishers will have more accurate data that they collect from known individuals, but it will be a smaller footprint because it misses the many visitors that never log-in or opt-in. Perhaps 20% of users are logged in on a site, and so audience addressability from the buyer's perspective would be insufficient without data sharing between publishers, or without the help of sophisticated, scaled third parties offering probabilistic modeling. Getting confidence in this modeling, is an exercise in ‚degrees of determinism'. Everyone benefits if buyers can gain confidence that they are trading on real, good data."[20]

International sind in der Branche überdies Ausweichbewegungen zu sehen auf Technologien wie das schon lange bekannte *Bid Streaming*. Mit Real Time Bidding kann man automatisiert und in Echtzeit auf Werbeplätze im Internet bieten. Es ist die technologische Basis hinter dem, was als „Programmatic Advertising" unsere Interneterfahrung prägt: die „zufällig" passenden Werbebanner zu dem, was wir im Netz tun. Unsere Interaktionen wie z.B. Suchanfragen werden in Echtzeit verauktioniert, um in Millisekunden die auf die Person zugeschnittene Werbung auszuspielen. Was ursprünglich im Kontext von Google Adwords entstanden war, kann auch ohne Cookies durch die Verknüpfung mit einem Identifier eine Vielzahl von Daten zu einer Person, zu ihrem Gerät und ihrem Aufenthaltsort übertragen.[21] Auch auf Verlagsplattformen tauchen Hinweise hierfür auf. Real Time Bidding ist gegenwärtig Gegenstand einer Klage vor dem Hamburger Landgericht.[22]

Portscanning, also das Suchen nach offenen Zugängen zu einem Rechner(system), ist nach deutschem Rechtsverständnis eine Technologie am Rande der Legalität, wenn sie auf fremde Rechner und Netzwerke gerichtet ist. Das Suchen nach offenen Ports kann als Vorstufe zu den in den „Hacker-Paragraphen" 202c und 303b StGB sanktionierten Handlungen betrachtet werden. Gleichwohl wird Portscanning international im Bereich der *Risk Solutions* eingesetzt, also der Branche, die sich u.a. der

18 Vgl. Engeler, Malte: Das neue Telekommunikation-Telemedien-Datenschutzgesetz. Was es über das regulatorische Klima der deutschen Datenschutzpolitik verrät, Telemedicus 14.07.2021, <https://www.telemedicus.info/soko21-das-neue-telekommunikation-telemedien-datenschutzgesetz-was-es-ueber-das-regulatorische-klima-der-deutschen-datenschutzpolitik-verraet/>, Stand: 20.02.2022.

19 Vgl. den Vortrag von Louisa Specht-Riemenschneider auf dem Datentag 03.11.2021; <https://stiftungdatenschutz.org/veranstaltungen/unsere-veranstaltungen-detailansicht/ttdsg-einwilligungsverwaltung-234#lg=1&slide=6>, Stand: 20.02.2022.

20 Baron, Andrew: Identity Innovations Will Create Complexity, Shared Standards Can Help, Pubmatic 08.05.2020, <https://pubmatic.com/blog/identity-innovations-will-create-complexity-shared-standards-can-help/>, Stand: 20.02.2022.

21 Vgl. die Darstellung auf Folie 45 bei Ryan, Johnny: Briefing on adtech, RTB, and the GDPR at dmexco Brave Event, 16.09.2019, <https://de.slideshare.net/JohnnyRyan/briefing-on-adtech-rtb-and-the-gdpr-at-dmexco-brave-event>, und für mehr Details <https://brave.com/wp-content/uploads/2019/02/3-bid-request-examples.pdf>, beide Stand: 20.02.2022. (Ich danke Matthias Eberl für den Hinweis.)

22 Vgl. die Klageschrift Dr. Johnny Ryan (Kläger) gegen IAB Technology Laboratory, Inc. (Beklagte), 24.03.2021. Online: <https://www.iccl.ie/wp-content/uploads/2021/06/GERMAN-LANGUAGE-ORIGINAL-Schriftsatz-an-das-Landgericht-Hamburg.pdf>, Stand: 20.02.2022.

Betrugsprävention widmet, aber eben auch mit vertieften Datenanalysen *predictive analytics*, „Vorhersage-Produkte", entwickelt und für Big Data Policing an die Sicherheitsindustrie verkauft (s. dazu Abschnitt 4.3.). Vor einiger Zeit wurde z.B. öffentlich, dass ebay in einem Bündel von Maßnahmen auch Portscanning der Seitenbesucher*innen einsetzt.[23] Eingesetzt wird hier vielfach ThreatMetrix, eine Firma, die nach eigenen Angaben 4,5 Milliarden Geräte identifizieren kann und zum Elsevier-Mutterkonzern RELX gehört. Wolfie Christl wertet ThreatMetrix als „global mass surveillance system", das Daten von 1,4 Milliarden Menschen aus 185 Ländern verarbeitet,[24] und konnte dabei auch den Einsatz auf der Elsevier-Plattform ScienceDirect nachweisen. Eine weitere im Bibliotheksbereich bekannte Firma, die Risk Solutions anbietet, ist Clarivate.

Da der RELX-Konzern auch Daten-Produkte für Behörden und Sicherheitsindustrie anbietet, stellt sich damit die Frage, ob solche Datenerhebungen auch innerhalb des Konzerns für entsprechende Produkte wie LexisNexis Risk Solutions weiterverarbeitet werden. Mangels Transparenz kann dazu noch nicht viel gesagt werden (vgl. Abschnitt 5.3.).

2. Workbenches

„Workflow is the new content"[25] – eine Parole, die die Entwicklung in den Informationsinfrastrukturen tatsächlich gut beschreibt. Die Pfadabhängigkeit,[26] die in der Informationsversorgung zu beobachten ist und Alternativen schwer entwickelbar macht,[27] füllte den Verlagen in den vergangenen Jahren mit Gewinnmargen von teils weit über 30 %[28] die Kriegskasse für eine Vielzahl von Firmenübernahmen – nicht nur, aber auch mit Anbietern im Wissenschaftsbereich. So wie die großen Plattformen mit strategischen Übernahmen sowohl das eigene Geschäftsfeld erweitern wie generell verhindern, dass sich eine Konkurrenz zu ihnen entwickeln kann, so decken die großen Verlage durch die Zukäufe mittlerweile den ganzen Research Life Cycle ab.[29]

Was Roger Schonfeld in Bezug nur auf das Publizieren als entstehenden *Supercontinent* diskutierte,[30] realisiert sich wenige Jahre später als Ökosystem, das in vernetzten Workbenches alle Stufen des

23 Vgl. Nemec, Dan: Ebay is port scanning visitors to their website. And they aren't the only ones, nem.ec 24.05.2020, <https://blog.nem.ec/2020/05/24/ebay-port-scanning/>, Stand: 20.02.2022.

24 Vgl. Darstellung und Screenshots bei Christl, Wolfie: ThreatMetrix, Twitter, @WolfieChristl, 23.07.-18.08.2020, <https://twitter.com/wolfiechristl/status/1286341387718397952>, Stand: 20.02.2022.

25 Dempsey, Lorcan: Workflow is the new content ..., Twitter, @lorcanD, 17.05.2021, <https://twitter.com/lorcand/status/1394276710519087104>, Stand: 20.02.2022.

26 Begriff aus der Organisationspsychologie, der die Entwicklung einer zunehmenden Verengung der anfangs vielfältigen Handlungsmöglichkeiten hin zu einem ganz engen „Pfad" beschreibt, der eine Umkehr oder einen Ausbruch kaum möglich macht. Klassisches Bespiel sind Schreibtastaturen: Es gibt Layouts, mit denen man bis zu 30 % schneller schreiben kann, allein die Menge der bestehenden Geräte und der auf das jetzige Layout trainierten Schreiber*innen macht einen Wechsel jedoch schier unmöglich.

27 Vgl. Brembs, Björn u.a.: Replacing academic journals, Zenodo 24.09.2021, <https://doi.org/10.5281/zenodo.5564003>, Stand: 20.02.2022, und Abschnitt 7.

28 Vgl. SPARC 2019, S. 11.

29 Vgl. die Grafik auf <https://stoptrackingscience.eu/background-information/> und den Artikel von Posada, Alejandro; Chen, George: Inequality in Knowledge Production. The integration of Academic Infrastructure by Big Publishers, ELPUB 2018, <10.4000/proceedings.elpub.2018.30>, Stand: 20.02.2022.

30 Schonfeld, Roger C.: The Supercontinent of Scholarly Publishing?, The Scholarly Kitchen 03.05.2018, <https://scholarlykitchen.sspnet.org/2018/05/03/supercontinent-scholarly-publishing/>, Stand: 20.02.2022.

Forschungsprozesses von Informationsrecherche, Laborarbeit, Schreiben, Publizieren, Wissenschaftskommunikation und Kennzahlenerstellung übergreift und das die einzelnen Forschenden an sich gar nicht mehr verlassen müssen – und sicher auch gar nicht sollen. Wie immer in einer Aufmerksamkeitsökonomie bedeutet ein längerer Aufenthalt, dass mehr Datenspuren entstehen, es also mehr zu beobachten und zu verwerten gibt. Claudio Aspesi bezeichnet dies als "the next battleground" mit dringendem Handlungsbedarf:

"If it doesn't invest in alternative solutions, the academic community may find itself beholden to a small number of vendors for managing communities, data flows, research assessment, and learned society communications, all within digital silos that could hinder the growth of cross-disciplinary collaboration and discovery."[31]

Dies ist das, was die Vorsitzende des Rats für Informationsinfrastrukturen Petra Gehring als „Lebendfalle für Forschende" bezeichnete: „Große Player greifen absichtsvoll die Integrität des wissenschaftlichen Austauschs an. Sie betrachten den gesamten intellektuellen Zyklus staatlich getragener und damit freier Forschung als ihr künftiges Produkt."[32]

3. Bibliothekssysteme

Die kommerziellen Produkte im Bereich der Bibliotheksmanagementsoftware werden immer mehr Teil solcher Ökosysteme. Bibliothekssysteme sind aus der Sicht von Data Analytics hochinteressante Objekte, weil auch in ihnen echte Daten echter Menschen stecken – nämlich individualisierte Personendaten aller Nutzer*innen, verbunden mit ihrem Mediennutzungsverhalten und zwar auch dessen Teil, der sich von außen nicht messen lässt (Ausleihe), sowie den darauf gründenden Finanzierungsströmen der Bibliothek. Entsprechend haben alle großen Anbieter Modelle einer *library analytics* entwickelt und ihre Systeme auf Cloudbetrieb umgestellt, sodass es strukturell für die Anwender keinen abschließenden Überblick über die Datenflüsse mehr geben kann. Neben den Forschenden kommen hier verstärkt die Studierenden als Zielobjekt in den Blick, wenn man sich ansieht, was OCLC als User Story für ihr System WorldShare Management Services (WMS) aus der Universität Gloucestershire zusammenträgt – die Verantwortlichen auf dem Campus sind jedenfalls begeistert:

„Die Einführung von WMS an der University of Gloucestershire hat uns die Möglichkeit geboten, Lernanalytikfunktionen von Grund auf in ein Bibliotheksmanagementsystem zu integrieren. In dem Projekt wurden Benutzertransaktionen sowohl mit gedruckten als auch mit elektronischen Ressourcen-Datensätzen aus der Ausleihedatenbank von WMS und den detaillierten Nutzungsprotokollen von EZproxy analysiert. Die Integration des Datenfeeds in das Studierendenaktensystem der Universität ermöglicht weitere Einblicke. […] Über die Benutzer-ID können wir jetzt auf das Studienprogramm und den Fachbereich des Studierenden zugreifen […]. Auf diese Weise können wir beispielsweise erkennen, wie viele Studierende eines bestimmten Programms auf ScienceDirect zugegriffen haben.

31 Aspesi, Claudio; Brand, Amy: In pursuit of open science, open access is not enough, Science 368, 2020, <https://www.science.org/doi/10.1126/science.aba3763>, Stand: 20.02.2022.

32 Gehring, Petra: Das Schicksal von Open Science steht auf dem Spiel, Forschung & Lehre 02.08.2021, <https://www.forschung-und-lehre.de/politik/das-schicksal-von-open-science-steht-auf-dem-spiel-3902>, Stand: 20.02.2022.

Unsere Daten aus Studierendenakten sind ziemlich umfangreich, daher können wir sogar die Nutzung von Bibliotheksressourcen in unterschiedlichen demografischen Gruppen vergleichen."[33]

Solche Analysemöglichkeiten sind geeignet, Material für eine *predicitive analytics* von Studierenden zu liefern, die international von Hochschulen aus finanziellen Gründen in Erwägung gezogen wird und jede Möglichkeit der Diskriminierung eröffnet.[34] Auch außerhalb des Campus dürfte dies Interesse wecken, denn viel „genauer und umfassender als ein Zeugnis oder ECTS-Punkte geben unsere Datenprofile Aufschluss darüber, was wir wissen (Expertise), wie gut und wie schnell wir lernen (Talent), wie wir Situationen lösen (Kompetenz), wie oft wir etwas versuchen (Frustrationstoleranz) oder ob wir auch einmal ungewöhnliche Lösungswege wagen (Kreativität)."[35] Mindestens potentielle Arbeitgeberinnen und Arbeitgeber, Versicherungen sowie Bonitäts- und Ratingagenturen werden deshalb versuchen, Zugriff zu erhalten, sobald solche Daten verfügbar sind.

Beim Blick auf Bibliothekssysteme wurde dabei schon seit Längerem festgehalten, dass nicht nur Forschende und Studierende, sondern auch Bibliotheken und ihre Beschäftigten selbst Ziel einer datengestützten Ausbeutung und Enteignung sind:

„Die Bibliotheken sehen sich in einer zunehmenden Abhängigkeit von den bibliotheksanbietenden Systemen wie Alma, Exlibris oder OCLC mit der Konsequenz, dass sie späterhin ihre eigenen Katalogdaten zurückkaufen müssen. Sie haben selbst die Rohdaten lizenziert, aber die Bedingungen, zu denen lizenziert wird, deuten darauf hin, dass seitens der einschlägigen Verlage die Vorbereitungen laufen, dieses strategische Asset unter Kontrolle zu bekommen. Was das für die zukünftige Wissenschaft und für das Urheberrecht bedeutet, ist ungewiss."[36]

Wer aus der Bibliothekspraxis das zähe Geschäft der Metadatenlieferungen kennt, sieht hier das Muster, wie Bibliotheksbeschäftigte gerade in den notorischen Fällen als Clickworker wie bei Amazons *Mechanical Turk*[37] Daten erstellen und reparieren, sodass sie überhaupt systemtauglich und nutzbar

33 Erfahren Sie, welche Bibliotheksressourcen den größten Einfluss haben, <https://www.oclc.org/de/member-stories/gloucestershire.html>,Zur Situation bei ExLibris Alma vgl. z.B. Sarah Lamdan: Libraries are part of the ed tech surveillance ecosystem, Twitter, @greenarchives1, 02.06.2021, <https://twitter.com/greenarchives1/status/1400067993929912326>, zu EBSCO vgl.: The Evolving Landscape of Library Data Analytics, <https://www.ebsco.com/blogs/ebscopost/evolving-landscape-library-data-analytics>, alle Stand: 20.02.2022.

34 Vgl. Shea Swauger: The next normal: Algorithms will take over college, from admissions to advising. The Washington Post 12.11.2021, Online: <https://www.washingtonpost.com/outlook/next-normal-algorithms-college/2021/11/12/366fe8dc-4264-11ec-a3aa-0255edc02eb7_story.html>, Stand: 20.02.2022.

35 Lentsch, Justus: Unsere Bildungsdaten gehören uns! Wiarda Blog 16.02.2021, <https://www.jmwiarda.de/2021/02/16/unsere-bildungsdaten-geh%C3%B6ren-uns/>, Stand: 20.02.2022.

36 Weingart, Peter; Taubert, Niels (Hrsg.): Wissenschaftliches Publizieren. Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung. Berlin 2016, S. 109. Online: <https://edoc.bbaw.de/files/2662/00_FB38_WissenschaftlichesPublizieren_gesamt_edoc.pdf>, Stand: 20.02.2022.

37 Amazons „Mechanical Turk" (benannt nach dem historischen „Schachtürken") ist ein Dienst, bei dem Auftraggeber digital noch nicht völlig automatisierbare Aufgaben einstellen können, die dann verteilt in kleinen Paketen von freiberuflichen Crowdworkern abgearbeitet werden. Als „letzte Meile der Automatisierung" sind diese Programme gegenwärtig viel gefragt, um z.B. die Objekterkennung von KI oder Audio-Transskription zu verbessern, Textkorrekturen zu erledigen, problematische Inhalte auf Social Media zu kennzeichnen usw. Ähnlich wie bei Uber und anderen Plattformen sind bei dieser Art von „Mikrowork" dann häufig Lohndumping, Behinderung von gewerkschaftlicher Organisation und Dauerüberlastung an der Tagesordnung, vgl. Gray, Mary L., Suri, Siddharth: Ghost Work. How to Stop Silicon Valley from Building a New Global Underclass. Boston 2019.

sind – offenbar nur, damit die Bibliotheken dann in der gleichen Weise ihre Rechte daran verlieren sollen, wie die Autor*innen beim klassischen Verlagsmodell.

Die Verlage ihrerseits haben für ihre Zwecke ebenfalls Interesse an Bibliothekssystemen gefunden. Innerhalb eines Webinars der Scholarly Network Security Initiative (SNSI),[38] welche von einer Reihe von Verlagen getragen wird, wurden in einem Vortrag Überlegungen angestellt, wie sich Bibliothekssoftware zur Unterstützung des Kampfes gegen Schattenbibliotheken einsetzen ließe. Ein Proxy-Plugin für eine „Modern Library Architecture" soll dabei eine ganze Reihe Daten zugänglich machen: Zeitstempel, detaillierte Browser-Informationen, Nutzername, Account-Informationen, IP-Adresse, aufgerufene URLs, Geräteinformationen, Lokalisierungsdaten, Nutzerverhalten (wenn die Person andere Ressourcen aufruft als ihre Fachaffiliation vermuten ließe) und biometrische Daten. Letzteres ist zentral und bezieht sich auf die individuelle Art, wie ein*e Nutzer*in mit ihrem Gerät interagiert: Tippgeschwindigkeit, Art der Mausbenutzung oder der Touchsteuerung – das ergibt in der Summe einen individuellen Fingerabdruck, mit dem ein*e Hochschulangehörige*r auch dann als Nutzer*in von SciHub wiederzuerkennen ist, wenn er oder sie meint, technische Sicherheitsvorkehrungen getroffen zu haben. Um Bibliotheken zur Installation einer solchen „Analysis Engine" zu motivieren, sollten ihnen Rabatte angeboten werden – gleichwohl dürften solche Eingriffe in die hochschulische Netzsicherheit bei den dortigen Rechenzentren auf Bedenken stoßen. Diese Überlegungen erregten einige öffentliche Aufmerksamkeit,[39] denn auch wenn sie als hypothetisch vorgetragen wurden, eröffnete sich dadurch doch ein Einblick in Mentalitäten. Verlage und Systemanbietende scheinen sich hierbei einig zu sein, denn bei einem späteren SNSI-Webinar stellte Don Hamparian von OCLC „Libraries as Security Advocates" vor, die selbstverständlich „Publisher Assets" beschützten.[40]

4. Geschäftsmodelle

4.1. Open Access

Open Access ist kein Tracking-Geschäftsmodell, aber man wird umgekehrt zugeben müssen, dass Open Access gegen Tracking auch nicht hilft – jedenfalls solange die Publikationsinfrastrukturen die bleiben, die sie sind und nur die Laufrichtung des Geldes umgekehrt wird, um an Stelle des Lesens das Publizieren mit Gebühren zu belegen. Ein Login ist für die meisten Trackingtechnologien keineswegs nötig, weshalb Cody Hanson auch ein Beispiel von PLoS One untersuchen konnte. Solange der Datenverkehr auf den Verlagsplattformen stattfindet, solange kann er auch beobachtet, ausgewertet und monetarisiert werden. Es ist bei der Umstellung auf kommerziellen Open Access eher sogar noch von einer höheren Motivation für Tracking auszugehen: Denn durch die Transformation entwickelt

38 Onlinepräsenz unter: <https://www.snsi.info/>, Stand: 20.02.2022.

39 Vgl. Dobusch, Leonhard: Neues vom Großverlag Elsevier. Kein Open-Access-Deal, dafür mit Spyware gegen Schattenbibliotheken?, Netzpolitik, 26.10.2020, <https://netzpolitik.org/2020/neues-vom-grossverlag-elsevier-kein-open-access-deal-dafuer-mit-spyware-gegen-schattenbibliotheken/>; Metha, Gautama: Proposal to install spyware in university libraries to protect copyrights shocks academics, Coda 13.11.2020, <https://www.codastory.com/authoritarian-tech/spyware-in-libraries/>, die Aufzeichnung des Webinars auf Vimeo: <https://vimeo.com/623425480> (ab ca. Min. 27) sowie Saunders, Johnny: like the cartel they are, Twitter, @json_dirs, 14.12.2021, <https://twitter.com/json_dirs/status/1470633210581192705>, alle Stand: 20.02.2022.

40 Vgl. die Aufzeichnung des Webinars auf Youtube: <https://www.youtube.com/watch?v=HEBQyg_ezHI> (ab Min. 24:45). Stand: 20.02.2022.

sich hinsichtlich der Nutzung eine neue Unübersichtlichkeit im Vergleich zu vorher, wo sich klar definierte Zielgruppen hinter einheitlichen Authentifizierungssystemen versammelten. Damit entsteht eine neue Form von Double Dipping:[41] Autor*innen bezahlen mit Article Processing Charges (APC), Lesende mit persönlichen Daten – eine weitere Variante der Volumenoptimierung des Umsatzes, so wie die APC-basierte Transformation insgesamt über einen längeren Zeitraum betrachtet sich schon jetzt als deutlich kostenträchtiger erweist.[42]

Vor allem aber ist zu beachten, dass insbesondere der APC-gesteuerte Open Access den Großverlagen hilft, die oben skizzierten Workbenches über den gesamten Research Life Cycle hinweg zu etablieren. Aspesi weist darauf hin, dass es in der Vergangenheit neben technischen und finanziellen Limitierungen auch das Copyright war, das den Aufbau richtig großer multifunktionaler Portale behinderte, weil es den Content rechtlich parzellierte. Dieses Hindernis falle, wie auch Schonfeld mit seiner Darstellung der Übernahme von ProQuest durch Clarivate erläutert,[43] mit der Transformation in der gleichen Weise zunehmend weg, wie der technische Fortschritt auch die finanziellen Hemmnisse mindert. Der Bericht, den Hubertus Neuhausen von der APE2020 gibt, zeigt den Stand der Diskussion, in der die Verlage ganz offen zeigen, wie sie den Forschungszyklus in der Gänze dominieren wollen und dabei untereinander kooperieren. Die selbstkritischen Fragen, die er dort stellt, sind es sehr wert, noch mal nachgelesen zu werden[44] – umso mehr, als mittlerweile der erste „Multiverlags-Pool" online ist.[45]

4.2. People Analytics in der Wissenschaft

Der von der Pandemie ausgelöste Digitalisierungsschub in der Arbeitswelt förderte sehr rasch auch eine Vielzahl anekdotischer Berichte über Kontrollwünsche von Arbeitgeber*innenseite zu Tage. Aktuell hat Wolfie Christl eine detailreiche Studie hierzu vorgelegt, die die Entwicklung der zurückliegenden Jahre analysiert:[46] Präparierte IT-Arbeitsplätze, Sensoren in der Arbeitskleidung, Auswertung von Kassendaten, Ortungssysteme zur Erfassung von Bewegungsmustern von Mitarbeiter*innen indoor wie outdoor, Software zum *Process Mining* von Arbeitsabläufen sowohl zu deren Optimierung wie zur Identifizierung von „unerwünschten Aktivitäten" – all das prägt den Arbeitsalltag von immer mehr Menschen. Dies betrifft nicht nur die Arbeit in der Produktion oder die mobile Pflegekraft, deren Minutenkontingente sich nicht ändern durch die jeweilige Verkehrssituation und ob die betreute

41 Begriff für die doppelte Abrechnung von Leistungen wie z.B. bei Hybrid Open Access-Zeitschriften, wo einzelne Artikel offen zugänglich sind nach Bezahlungen von Publikationsgebühren, die Zeitschrift insgesamt aber weiterhin im Lizenzmodell angeboten wird.

42 Vgl. Morrison, Heather u.a.: Open access article processing charges 2011–2021, Sustaining the Knowledge Commons 24.06.2021, <https://sustainingknowledgecommons.org/2021/06/24/open-access-article-processing-charges-2011-2021/>, Stand: 20.02.2022.

43 Vgl. Schonfeld, Roger C.: The New Clarivate Science. A Second-Order Consequence of Open Access, The Scholarly Kitchen 9.12.2021, <https://scholarlykitchen.sspnet.org/2021/12/09/new-clarivate-science/>, Stand: 20.02.2022.

44 Vgl. Neuhausen, Hubertus: Open Access? – Ist durch! Aber was ist mit den Daten? ABI Technik 40 (3), 2020, S. 277–291. Online: <https://www.degruyter.com/document/doi/10.1515/abitech-2020-2022/html>, Stand: 20.02.2022.

45 Vgl. Lisa Janicke Hinchliffe: Elsevier's ScienceDirect as Content Supercontinent?, The Scholarly Kitchen 10.01.2022, Online: <https://scholarlykitchen.sspnet.org/2022/01/18/sciencedirect-as-content-supercontinent/>, Stand: 20.02.2022.

46 Vgl. Christl, Wolfie: Digitale Überwachung und Kontrolle am Arbeitsplatz. Von der Ausweitung betrieblicher Datenerfassung zum algorithmischen Management? Wien 2021. Online: <https://crackedlabs.org/daten-arbeitsplatz/info>, Stand: 20.02.2022.

Person gerade einen guten oder eher schlechten Tag hat – in der Entwicklung von Wissensarbeit sorgte dies in der Vergangenheit schon für Veränderungen und Konflikte auf dem Campus.[47]

Die in Abschnitt 3. beschriebenen Workbenches haben nun das Potential, eine People Analytics der Wissenschaft zu ermöglichen. Gerade die übergreifenden Teile wie Forschungsinformationssysteme werden explizit damit beworben, „vertiefte Einsichten" zu ermöglichen, und Bibliothekssysteme werden, wie Schonfeld in seinem Kommentar zum Aufkauf von ProQuest durch Clarivate formulierte, „a component in a larger suite."[48] Vorangetrieben werden die Entwicklungen auch durch Kopplungsgeschäfte wie in den Niederlanden, wo Elsevier einen Vertrag abschließen konnte, der ohne Zusatzkosten zu den bisherigen Subskriptionsaufwendungen Open-Access-Publizieren für die Wissenschaftler*innen ermöglichte – wenn die Hochschulen im Gegenzug hierzu bei „Metadaten-Projekten" kooperierten und Elseviers Forschungsinformationssysteme lizenzierten.[49] Dies ist ein Wiedergänger der *Browser Wars* der 1990er Jahre, als Microsoft Netscape aus dem Markt drängte, indem es den Internet Explorer fest mit Windows koppelte – denn schließlich wird allen Wissenschaftler*innen sofort klar sein, dass ihr Forschungsoutput nur dann vollständig in den Evalutionsintrumenten verzeichnet sein wird und „zählt", wenn er in der Publikationsplattform desselben Anbieters erschienen ist. Der Markt für Forschungsinformation wird sich somit rasch aufräumen und ein weiterer Schritt in den Vendor Lock-In ist getan.

Wer die Debatten um *#IchbinHanna* verfolgt hat,[50] kann daher durchaus zum Schluss kommen, dass hier noch Luft nach oben ist: haben sich diese Workbenches erst einmal durchgängig gekoppelt und vernetzt, bilden diese Ökosysteme zum einen für die Hochschulen einen Vendor Lock-In weit über den Bereich der Literaturversorgung hinaus, wo dieser Zwang bereits gegeben ist. Zum anderen wird man dann mindestens in technischer Hinsicht Wissenschaftler*innen ähnlich behandeln können wie den „Picker" im Amazon-Lager mit seiner engmaschigen Überwachung. Es muss dann nicht mehr so grobschlächtig zugehen wie kürzlich an der Universität Liverpool, wo man zur Deckung eines Finanzlochs einfach den weniger zitierten Wissenschaftler*innen gekündigt hat.[51]

All diese Entwicklungen echoen einen kleinen Halbsatz aus Christl Studie: „Beschäftigtendaten werden zum Produkt."[52] Davon haben die Beschäftigten in der Regel nichts, aber auch ansonsten muss sich der Nutzen des Produkts nicht unbedingt dort materialisieren, wo manche sich das vorstellen.

47 Vgl. Siems, Renke: Unser industrielles Erbe. Bibliotheken und die digitale Transformation, in: o-bib 4 (3), 2017. <https://doi.org/10.5282/o-bib/2017H3S1-15>, Stand: 20.02.2022.

48 Schonfeld, Roger C.: Clarivate to acquire ProQuest, The Scholarly Kitchen 18.05.2021, <https://scholarlykitchen.sspnet.org/2021/05/18/clarivate-to-acquire-proquest/>, Stand: 20.02.2022.

49 Vgl. Knecht, Sicco de: Leaked document on Elsevier negotiations sparks controversy, ScienceGuide 06.11.2019, <https://www.scienceguide.nl/2019/11/leaked-document-on-elsevier-negotiations-sparks-controversy/>, Stand: 20.02.2022.

50 Unter diesem Hashtag organisierte sich die breite Protestbewegung gegen die erdrückende Abhängigkeit vieler Wissenschaftler*innen gegen das System sehr kurz befristeter Arbeitsverträge. Vgl. die Dokumentation unter <https://ichbinhanna.wordpress.com/>, Stand: 20.02.2022.

51 Vgl. Bishop, Dorothy: University staff cuts under the cover of a pandemic. The cases of Liverpool and Leicester, BishopBlog 03.03.2021, <http://deevybee.blogspot.com/2021/03/university-staff-cuts-under-cover-of.html> – nach Protesten wurden die Kürzungen weitgehend zurückgezogen, vgl.: University of Liverpool staff call off strike action, BBC News 01.10.2021, <https://www.bbc.com/news/uk-england-merseyside-58766202>, beide Stand: 20.02.2022.

52 Christl 2021, S. 11.

Irlands Science Foundation hatte sich etwa kürzlich entschlossen, die strategische Neuausrichtung datenbasiert voranzutreiben, und wählte Elsevier dazu als Partner aus:

„They have access to a vast array of data, and this will help us to establish where Ireland is good and nearly good in emergent and convergent technologies. It will also help us decide on the actions we need to take to make us really good. For example, we might see a certain field where we are nearly good at present and find that we need to recruit top talent or run new funding calls to support it."[53]

Wo dieser wundersame Datenschatz denn so herkommt, wurde offenbar nicht hinterfragt. Aber man muss wohl konstatieren: User Tracking hat begonnen, in den wissenschaftlichen Wettbewerb einzugreifen.

4.3. „Risk Solutions"

Der Wissenschaftssektor ist nicht der einzige Bereich, dem sich die großen Verlage bzw. deren Mutterkonzerne sowie vertraute Serviceanbieter mit einem Data Analytics-Geschäftsmodell zuwenden. Immer mehr Player werden wie gesagt im „Risk Solutions"-Bereich tätig wie dem Big Data Policing. Bei letzterem sind seit einigen Jahren der Elsevier-Mutterkonzern RELX und Thomson Reuters aktiv. Mit den Produkten LexisNexis Risk Solutions und Clear bieten sie aggregierte Datenprodukte für die Sicherheitsindustrie an. Diese Aktivitäten gehen weit zurück: RELX war 2006 einer der frühen Investoren in die Data Analytics-Firma Palantir,[54] die von Peter Thiel u.a. mit Mitteln von In-Q-Tel, dem Risikokapitalzweig der CIA, gegründet worden war und seitdem eine Kundschaft von Militär und Nachrichtendiensten, aber auch Finanzdienstleistungen versorgt.[55] Seit 2010 kooperiert auch Thomson Reuters mit Palantir.[56]

Risk Solutions sind offenbar ein boomendes Geschäft: Thomson Reuters hatte 2020 allein mit der militarisierten Grenzpolizei ICE (Immigrants and Customs Enforcement) Verträge im Volumen von über $ 60 Millionen,[57] LexisNexis schloss im Frühjahr 2021 einen Vertrag mit ICE über $ 16,8 Millionen ab.[58] Die Datenprodukte aggregieren Daten aus vielerlei öffentlichen wie privaten Quellen mit Informationen, wie sie hierzulande die Schufa sammelt, aber auch mit Datenbanken von Autokennzeichen und Kundinnen und Kunden von Mobilfunknetzen. Die in Abschnitt 2.3. erwähnte Firma

53 McCall, Barry: Propelling Ireland to first-mover status in research and innovation, in: The Irish Times 26.08.2021. Online: <https://www.irishtimes.com/sponsored/innovation-partner-profiles/propelling-ireland-to-first-mover-status-in-research-and-innovation-1.4655112>, Stand: 20.02.2022.

54 Vgl. die Eigendarstellung: Investing in disruptive data & analytics technologies, <https://www.relx.com/our-business/our-stories/rev-venture-partners>, Stand: 20.02.2022.

55 Vgl.: A (Pretty) Complete History of Palantir, Social Calculations 11.08.2015, <https://web.archive.org/web/20190724214959/http://www.socialcalculations.com/2015/08/a-pretty-complete-history-of-palantir.html>, Stand: 20.02.2022.

56 Vgl.: Thomson Reuters and Palantir Technologies Enter Exclusive Agreement to Create Next-Generation Analytics Platform for Financial Clients, MarketWired 12.04.2010, <https://web.archive.org/web/20180708032604/http://m.marketwired.com/press-release/thomson-reuters-palantir-technologies-enter-exclusive-agreement-create-next-generation-nyse-tri-1167653.htm>, Stand: 20.02.2022.

57 Vgl. Lyons, Kim: Thomson Reuters faces pressure over ICE contracts, The Verge 21.05.2020, <https://www.theverge.com/2020/5/21/21266431/thomson-reuters-ice-clear-software>, Stand: 20.02.2022.

58 Vgl. Biddle, Sam: LexisNexis to Provide Giant Database of Personal Information to ICE, The Intercept 02.04.2021, <https://theintercept.com/2021/04/02/ice-database-surveillance-lexisnexis/>, Stand: 20.02.2022.

ThreatMetrix ist Teil von LexisNexis Risk Solutions. Der Einsatz dieser Datenbanken bei ICE dient u.a. der Vorbereitung von Razzien, um Einwandernde aufzufinden und zu deportieren – nach Meinung von Bürgerrechtsorganisationen unter laufenden Menschenrechtsverletzungen.[59] Auch in der US-amerikanischen Politik wird der Einsatz dieser Produkte teils als rechtsmissbräuchlich eingestuft.[60]

Thomson Reuters und LexisNexis sind nun in den USA gleichzeitig die maßgeblichen Anbieter von juristischer Fachinformation, ohne Westlaw und LexisNexis ist ein Arbeiten nicht möglich. Der Albtraum, dem Aktivist*innen wie die Juristin Sarah Lamdan nachgehen,[61] wäre entsprechend, dass Anwält*innen durch Recherchen in Fachdatenbanken dazu beitragen, dass ihre Klientinnen und Klienten gefasst werden – was die betreffenden Firmen natürlich bestreiten.[62] Nichtsdestoweniger tragen aber die Erlöse aus dem Wissenschaftssektor dazu bei, solche Aktivitäten weiter ausbauen zu können. Und da die Daten wie von ThreatMetrix weltweit gesammelt und die Datenprodukte ebenso weltweit verkauft werden, müssen Einwandernde ohne gültige Papiere auch nicht die einzigen und letzten Personen sein, gegen die sie verwendet werden.

Der amerikanische Bibliothekar Shea Swauger hat sich im Selbstversuch mit Thomson Reuters Clear auseinandergesetzt, bis er seine Akte (so muss man das wohl nennen) in Händen hielt. Auf 41 Seiten waren seine persönlichen und familiären Verhältnisse festgehalten, wann er wo wohnte, Immobilienbesitz, wann er von wem ein gebrauchtes Auto kaufte, Wahlbeteiligung und politische Präferenz, Finanzsituation – und jede Menge fehlerhafte Angaben dabei. Diese zu korrigieren oder die Daten zu löschen, war jedoch keineswegs möglich:

„To wrap it up, @Westlaw, through CLEAR, collects a shit ton of data about you. They share it with law enforcement, including @ICEgov, and anyone who has enough money to buy CLEAR. And for most people, there's nothing you can do about it[.]"[63]

Mittlerweile gibt es in diesem Bereich auch eine gegenläufige Entwicklung: nicht nur Informationsdienstleister drängen in die Sicherheitsindustrie, sondern das Ganze geschieht auch umgekehrt: Palantir nutzt die Pandemie, um mit teils rüden Methoden in das Forschungsdatengeschäft einzudringen und sich dem lukrativen Markt der Gesundheitsdaten zu widmen.[64]

59 Vgl. die Aktivistenseite »No Tech for ICE«, <https://notechforice.com/>, Stand: 20.02.2022.

60 Vgl. House Committee of Oversight and Reform: Oversight Subcommittee Launches Investigation into Sale of Utility Customer Info to ICE for Deporting Immigrants, Press Release 26.02.2021, <https://oversight.house.gov/news/press-releases/oversight-subcommittee-launches-investigation-into-sale-of-utility-customer-info>, Stand: 20.02.2022.

61 Vgl. Lamdan, Sarah: When Westlaw Fuels ICE Surveillance. Legal Ethics in the Era of Big Data Policing, N.Y.U. Review of Law & Social Change 43 (2), 2019, S. 255–293. Online: <https://socialchangenyu.com/review/when-westlaw-fuels-ice-surveillance-legal-ethics-in-the-era-of-big-data-policing/>. Dies.: Librarianship at the Crossroads of ICE Surveillance, In the library with the lead pipe 13.11.2019, <https://www.inthelibrarywiththeleadpipe.org/2019/ice-surveillance/>, Stand: 20.02.2022.

62 Vgl. Jung, Justin: UCLA School of Law holds contracts with companies selling personal data to ICE, Daily Bruin 17.07.2020, <https://dailybruin.com/2020/07/17/ucla-school-of-law-holds-contracts-with-companies-selling-personal-data-to-ice>, Stand: 20.02.2022.

63 Swauger, Shea: Some of my co-workers, Twitter, @SheaSwauger, 13.12.2019, <https://twitter.com/SheaSwauger/status/1205587676172144641>, Stand: 20.02.2022.

64 Vgl. Max Chafkin: Peter Thiel. München 2021, S. 278f.

5. Stellungnahmen

Gegenwärtig ist die Stellungnahme des AWBI noch die maßgebliche Äußerung zum Thema. Das Ziel des Informationspapiers ist es, „einen breiten Diskurs in der Wissenschaft sowohl auf Ebene der wissenschaftlichen Entscheidungsträgerinnen und -träger als auch unter Wissenschaftlerinnen und Wissenschaftlern sowie in Einrichtungen der Informationsinfrastruktur zu führen, um die Praxis des Trackings, dessen Rechtmäßigkeit, Maßnahmen zur Einhaltung des Datenschutzes und Konsequenzen der Aggregation von Nutzungsdaten zu reflektieren und geeignete Maßnahmen zu ergreifen."[65]

Diese Motivation wurde vom Rat für Informationsinfrastrukturen in seinem aktuellen Empfehlungspapier zu Datendiensten aufgenommen mit der Ankündigung, gegebenenfalls „zu Einsatz und Folgen entsprechender Technologien im Forschungsprozess zu einem späteren Zeitpunkt Stellung zu beziehen."[66] Fachgesellschaften[67] und Universitätsvereinigungen[68] haben dies bereits getan und auch Wissenschaftler*innen äußern sich vielfach, sei es einzeln, sei es in den Initiativen „Stop Tracking Science!"[69] und „Take action to stop the lock up of research and learning".[70] Wer sich in Deutschland bislang allerdings nicht in einer Stellungnahme äußerte, sind die bibliothekarischen Gremien und Verbände – im Gegensatz etwa zur American Library Association, die in der „Resolution on the Misuse of Behavioral Data Surveillance in Libraries" deutliche Worte findet.[71]

6. Schlussfolgerungen

6.1. Überwachungskapitalismus

Das User Tracking auf den Verlagsplattformen, die Entwicklung der Workbenches und die Ausweitung der Geschäftsfelder sind Musteranwendungen[72] von Shoshana Zuboffs Analysen des Überwachungskapitalismus, wie sie ihn zu Beginn ihrer großen Studie definiert:

„Überwachungskapitalismus beansprucht einseitig menschliche Erfahrung als Rohstoff zur Umwandlung in Verhaltensdaten. Ein Teil der Daten dient der Verbesserung von Produkten und Diensten,

65 DFG 2021, S. 13.

66 RfII: Nutzung und Verwertung von Daten im wissenschaftlichen Raum. Empfehlungen zur Ausgestaltung von Datendiensten an der Schnittstelle zwischen Wissenschaft und Wirtschaft. Göttingen 2021, S. 15. Online: <https://rfii.de/download/nutzung-und-verwertung-von-daten-im-wissenschaftlichen-raum-september-2021/>, Stand: 20.02.2022.

67 Vgl.: DGPs-Vorstand und Kommission Open Science unterstützen „Stop Tracking Science"-Initiative, 7.12.2021, <https://www.dgps.de/aktuelles/details/dgps-vorstand-und-kommission-open-science-unterstuetzen-stop-tracking-science-initiative/> und: DPG-Positionspapier zur Zukunft des wissenschaftlichen Publikationswesens, 13.11.2021, <https://www.dpg-physik.de/veroeffentlichungen/publikationen/stellungnahmen-der-dpg/wissenschaftssystem/dpg-positionspapier-zur-zukunft-des-wissenschaftlichen-publikationswesens>, beide Stand: 20.02.2022.

68 Vgl.: LERU Data Statement, <https://www.leru.org/publications/is-university-autonomy-threatened-by-eu-data-policy-and-law>, Stand: 20.02.2022.

69 Vgl. <https://stoptrackingscience.eu/>, Stand: 20.02.2022.

70 Vgl. <https://investinopen.org/blog/take-action-to-stop-the-lock-up-of-research-and-learning/>, Stand: 20.02.2022.

71 Vgl. Resolution on the Misuse of Behavioral Data Surveillance in Libraries, 26.1.2021, <https://www.ala.org/advocacy/intfreedom/datasurveillanceresolution>. Andreas Degkwitz als Vorsitzender des DBV äußerte sich in vergleichbarer Weise auf einer Podiumsdiskussion von Wikimedia Deutschland, vgl. den Mitschnitt auf <https://av.tib.eu/media/55690>, Stand: 20.02.2022.

72 Vgl. die Darstellung von Pooley, Jefferson D.: Surveillance Publishing. Working Paper 16.11.2021, <https://osf.io/preprints/socarxiv/j6ung/download>, Stand: 20.02.2022.

den Rest erklärt man zu proprietärem *Verhaltensüberschuss*, aus dem man mithilfe fortgeschrittener Fabrikationsprozesse, die wir unter der Bezeichnung „Maschinen- oder künstliche Intelligenz" zusammenfassen, *Vorhersageprodukte* fertigt, die erahnen, was Sie jetzt, in Kürze oder irgendwann tun. Und schließlich werden diese Vorhersageprodukte auf einer neuen Art von Marktplatz für Verhaltensvorhersagen gehandelt, den ich als *Verhaltensterminkontraktmarkt* bezeichne. So erpicht wie zahllose Unternehmen darauf sind, auf unser künftiges Verhalten zu wetten, haben Überwachungskapitalisten es mittels dieser Operationen zu immensen Wohlstand gebracht."[73]

Der Ausgangspunkt dieser Entwicklung war, als Googles Ingenieurinnen und Ingenieure früh begriffen, dass man mit den bei Suchanfragen in Unmengen anfallenden Kollateraldaten „die Suchmaschine in ein rekursives Lernsystem verwandeln könnte".[74] Dies wurde zum Ausgangspunkt für *Adwords*, womit Google die Dotcom-Krise überstand und begann, die Rentabilität in schwindelerregende Höhen zu treiben. Die Datenextraktion wurde dabei zum Handlungsimperativ, denn fehlende Daten bedeuten nicht vorhersagbares Verhalten und dies entgangene Einnahmen. Die Diversifizierung von Google von der Suchmaschine zu vielerlei Services und bis hin zu Betriebssystemen und Hardware hat hier ihren Ausgangspunkt und war Vorbild sowohl für die übrigen Plattformen des GAFAM-Komplexes der großen Internetkonzerne Google, Apple, Facebook, Amazon und Microsoft wie auch vielerlei kleinere Firmen z.B. im AdTech-Bereich. Sie alle setzen als letzten Schritt Vorhersageprodukte in Verhaltensmodifikation um. Das *Nudging*, das kaum merkliche „Anstupsen" des Nutzerverhaltens in die gewünschte Richtung, mit dem dann die Kartenapp die Fahrroute den Konsumgewohnheiten anpasst, die ganz zufällig in dem Moment passenden Empfehlungen – all diese Vorhersageprodukte aus immensen proprietären Datenbanken, die die einzelne Person bald besser kennen als sie sich selbst, werden von einem der vielen Interviewpartner Zuboffs knapp zusammengefasst: „Wir lernen, die Musik zu schreiben, und sorgen dann dafür, dass sie die Leute tanzen läßt."[75]

Dies heißt in der Konsequenz dann eben auch, dass die Hinwendung zu einem Data-Analytics-Geschäftsmodell, wie es bei den großen Verlagen zu sehen ist und was zu einer intensiven Zusammenarbeit mit den großen Akteuren des Überwachungskapitalismus führte, eine Grundsatzentscheidung für ein spezifisches Paradigma ist, welches seine Spielregeln in den Vordergrund stellt und nur danach funktioniert. Dies ist nichts, was man im Detail verhandelt und in Richtlinien, Verträgen oder AGBs nach Kund*innenwünschen festlegt, denn dieses Paradigma bedeutet eine grundsätzliche Asymmetrie an Wissen, Handlungsmöglichkeit und Macht. Gegenüber diesen Playern etwa auf Schutz der Privatsphäre zu pochen, ist für Zuboff daher vergleichbar damit, „als würde man von Henry Ford verlangen, jedes Modell T von Hand zu fertigen – oder von einer Giraffe einen kürzeren Hals. Derlei Forderungen bedrohen die Adressaten in ihrer Existenz. Sie verletzen die Grundmechanismen und Bewegungsgesetze, die hinter der Konzentration von Wissen, Markt und Wohlstand dieses Marktungetüms stehen."[76]

73 Zuboff, Shoshana: Das Zeitalter des Überwachungskapitalismus. Frankfurt/M. 2018, S. 22.
74 Zuboff 2018, S. 90.
75 Zuboff 2018, 337f.
76 Zuboff 2018, S. 224.

6.2. There are no free lunches (and no quick fixes)

Wir sehen in den gegenwärtigen Informationsinfrastrukturen einen erheblichen Teil der Prinzipien des Überwachungskapitalismus am Werk, reflektieren dies aber bislang wenig. Welches *Nudging* schon alles am Werk ist, sei es durch unverstandene Recommendersysteme, sei es durch andere Maßnahmen – wir wissen es nicht und kommen daher wie Zuboff „immer wieder auf die wesentlichen Fragen zurück, die die Wissensteilung definieren: *Wer weiß? Wer entscheidet? Wer entscheidet, wer entscheidet?*“[77] Daher sind auch die Stimmen skeptisch zu beurteilen, die einer *science analytics* grundsätzlich etwas abgewinnen könnten, wenn sie datenschutzgerecht ins Werk gesetzt würde.[78] So lang kann der Löffel überhaupt nicht sein, den man im Sprichwort braucht, um mit dem Teufel aus einem Topf zu essen.

Solche Überlegungen sind auch deshalb fraglich, weil sie implizit voraussetzen, dass sich am Formenrepertoire der Wissenschaftskommunikation im Kern nichts ändert und die Wissenschaft des 21. Jahrhunderts auch weiterhin in Medienformen des 17. Jahrhunderts eingesperrt bleiben soll. Dass dies mindestens für die wachsende Zahl der datenintensiven Disziplinen kaum produktiv sein kann, liegt auf der Hand. Für diese gestaltet sich Forschung unter solchen Umständen immer mehr, wie wenn ein Softwareentwickler seinen Code schreibt, dann einen Aufsatz darüber publiziert, der nächste Entwickler daraus den Quelltext zu erraten versucht, um dann diesen zu ergänzen. Was ist zielführender: solch ein Vorgehen oder Git?[79]

Wir sehen damit in den gegenwärtigen Publikationsinfrastrukturen eine Korrumpierung und Perversion der immer mehr vergangenen Schriftkultur zu Ausbeutungszwecken. Wie eingangs erläutert, entstehen durch die digitale Transformation neue Formen des Lesens und Schreibens, die ihr eigenes Verständnis guter Praxis entwickeln können müssen – aber im Bereich der Wissenschaftskommunikation findet das zugunsten der etablierten Machtverhältnisse bislang nur in homöopathischen Dosen statt. Dabei stehen die Alternativen am Start: Es gibt Open Research Europe, es gibt detaillierte Vorschläge, die traditionellen Zeitschriften durch dezentrale Strukturen, basierend auf offenen Schnittstellen, zu ersetzen.[80] Dadurch könnten auch die digitalen Mehrwerte durch die gegenwärtig entstehenden Datenmärkte wie die Nationale Forschungsdaten-Infrastruktur (NFDI) viel besser gehoben werden, ohne sie in der gleichen Weise externen Interessen auszuliefern wie es jetzt schon die Publikationen sind.

Ein Ausgang aus dem User Tracking ist damit nicht ein Drehen an Stellschrauben, sondern ein Paradigmenwechsel. Vonnöten ist für alle Teile des Research Life Cycle eine vergleichbare Anstrengung, wie wir sie gegenwärtig bei den Forschungsdaten unternehmen mit der NFDI und der European Open Science Cloud (EOSC). Das ist aufwendig und nicht durch ein paar Verfahrensänderungen zu

77 Zuboff 2018, S. 382.

78 Vgl. Wissenschaftsrat: Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access. Bonn 2022, S. 48. Online: <https://www.wissenschaftsrat.de/download/2022/9477-22.html>, Stand: 20.02.2022.

79 Vgl. dazu auch Siems, Renke: When your journal reads you. User tracking on science publisher platforms. Elephant in the Lab, 14.04.2021, <https://doi.org/10.5281/zenodo.4683778>.

80 Vgl. Brembs, Björn u.a.: Replacing academic journals, Zenodo 24.09.2021. <https://doi.org/10.5281/zenodo.5526635>.

erledigen. Es ist sicher auch finanziell zunächst eine Anstrengung - so wie alle Anstrengungen im Kontext digitaler Souveränität. Angesichts der hohen Aufwendungen, die mit dem jetzigen System verbunden sind, sollte dies aber kein Hemmschuh sein. Es liegt wie immer nur am ersten Schritt.

Um diesen zu motivieren, hier noch ein kurzer Vergleich: Als Edward Snowden 2013 in die Öffentlichkeit trat, tat sich für die überraschte Öffentlichkeit ein Blick darauf auf, wie sich die Welt seit dem 11. September 2001 gewandelt hatte. Nachrichtendienste, die das gesamte Internet scannten, die Verschlüsselung von Blackberries knackten und Lieferungen des Netzwerkausrüsters Cisco abfingen, um Backdoors in die Hardware einzubauen – das hatte man zuvor nicht gesehen. Seit Snowdens Whistleblowing hat sich einiges getan, aber eher zum Komfort der *Five Eyes*, der Geheimdienstallianz von Australien, Neuseeland, USA, Kanada und Großbritannien. Denn surveillance advertising ist so ubiquitär geworden, dass die Dienste vieles nicht mehr selbst machen müssen: nicht einbrechen, nicht von Geheimgerichten sich fragwürdige Rechtstitel ausstellen lassen – sie gehen datenshoppen und kaufen quasi auf dem kurzen Dienstweg die Dinge ein, die sie wissen wollen. Allein Lokalisierungsdaten sind ein Milliardenmarkt, und zwar nicht nur bezogen auf Mobilgeräte, sondern z.B. auch bezogen auf Automobile. Die Firma Ulysses wirbt damit, außer in Nordkorea und auf Kuba jedes Fahrzeug auf der Erde in Echtzeit orten zu können. Da die Firma auch ans Militär verkauft, ist also mittlerweile jede*r, der in sein oder ihr Auto steigt, in einem potentiellen Zielobjekt unterwegs.[81]

Die Wissenschaft hatte schon einige Jahre vor Snowden die Erfahrung gemacht, welch destruktive Kraft Datenabgriffe haben können: Im November 2009 wurde das Klimaforschungszentrum der University of East Anglia Opfer eines sorgfältig vorbereiteten Hackerangriffs, bei dem Mails und Dokumente gestohlen wurden. Das Material wurde dann arrangiert, zurechtgeschnitten und selektiv zitiert ins Netz gestellt, um die Klimaforscher als Teil einer Verschwörung und als Betrüger darstellen zu können. Was die Klimawandelleugner „Climategate" nannten, nahm seinen Lauf mit Jahren an heftigsten politischen wie juristischen Auseinandersetzungen und nicht zuletzt persönlichen Bedrohungen. Das Handeln gegen den Klimawandel wurde verschleppt und Vorwürfe geisterten ungeachtet aller Widerlegungen immer weiter durchs Netz und die Medien.[82]

Wissenschaftler*innen sind seitdem immer wieder und verstärkt Angriffen ausgesetzt, gegenwärtig im Kontext der Pandemie.[83] Gleichzeitig liegen auch durch das User Tracking immer mehr Daten über jede und jeden vor, die sich instrumentalisieren lassen: Aus der Kombination von Informationsverhalten, Datenspuren in den kommerziellen Forschungswerkzeugen und der Online-Biographie insgesamt lässt sich jede Bloßstellung ableiten und jeder gezielte Angriff munitionieren. Wer das Kürzel SaaS wie so manche übersetzt mit „Surveillance as a Service" (eigentlich: Software as a Service), kann

81 Vgl. Cox, Joseph: Cars have your Location. This Spy Firm wants to sell It to the U.S. Military, Vice 17.03.2021, <https://www.vice.com/en/article/k7adn9/car-location-data-telematics-us-military-ulysses-group>, Stand: 20.02.2022.

82 Vgl. McKie, Robin: Climategate 10 years on. What lessons have we learned? The Guardian, 09.11.2019, <https://www.theguardian.com/theobserver/2019/nov/09/climategate-10-years-on-what-lessons-have-we-learned>, Stand: 20.02.2022.

83 Vgl. Nogrady, Bianca: 'I hope you die'. How the COVID pandemic unleashed attacks on scientists, Nature 13.10.2021, <https://www.nature.com/articles/d41586-021-02741-x>, Stand: 20.02.2022.

sich den Einbruch ins Rechenzentrum damit künftig vielleicht sparen. Damit gilt in Konsequenz, was Snowden aktuell mit Blick auf die großen Internetkonzerne schrieb:

„It is crucial to bear in mind that the industry's problem is not data "protection," it is data COLLECTION. Mass surveillance must be recognized as a crime, not a business model."[84]

Literaturverzeichnis

- Aspesi, Claudio; Brand, Amy: In pursuit of open science, open access is not enough, Science 368, 2020, <https://www.science.org/doi/10.1126/science.aba3763>, Stand: 20.02.2022.

- Baron, Andrew: Identity Innovations Will Create Complexity, Shared Standards Can Help, Pubmatic 08.05.2020, <https://pubmatic.com/blog/identity-innovations-will-create-complexity-shared-standards-can-help/>, Stand: 20.02.2022.

- Biddle, Sam: LexisNexis to Provide Giant Database of Personal Information to ICE, The Intercept 02.04.2021, <https://theintercept.com/2021/04/02/ice-database-surveillance-lexisnexis/>, Stand: 20.02.2022.

- Bishop, Dorothy: University staff cuts under the cover of a pandemic. The cases of Liverpool and Leicester, BishopBlog 03.03.2021, <http://deevybee.blogspot.com/2021/03/university-staff-cuts-under-cover-of.html>, Stand: 20.02.2022.

- Brembs, Björn u.a.: Replacing academic journals, Zenodo 24.09.2021. <https://doi.org/10.5281/zenodo.5526635>.

- Chafkin, Max: Peter Thiel. München 2021.

- Christl, Wolfie: Corporate Surveillance in Everyday Life. How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions. Wien 2017. Online: <https://crackedlabs.org/en/corporate-surveillance>, Stand: 20.02.2022.

- Ders.: Digitale Überwachung und Kontrolle am Arbeitsplatz. Von der Ausweitung betrieblicher Datenerfassung zum algorithmischen Management? Wien 2021. Online: <https://crackedlabs.org/daten-arbeitsplatz/info>, Stand: 20.02.2022.

- Cox, Joseph: Cars have your Location. This Spy Firm wants to sell It to the U.S. Military, Vice 17.03.2021, <https://www.vice.com/en/article/k7adn9/car-location-data-telematics-us-military-ulysses-group>, Stand: 20.02.2022.

84 Snowden, Edward: It is crucial to bear in mind, Twitter, @Snowden, 23.10.2021, <https://twitter.com/Snowden/status/1451935797301727232>, Stand: 20.02.2022.

- Crain, Matthew: Profit over Privacy. How Surveillance Advertising Conquered the Internet. Minneapolis 2021.

- DBV: Empfehlungen zu Methoden zur Kontrolle des Zugriffs auf wissenschaftliche Informationsressourcen. Ein gemeinsames Papier von Deutscher Bibliotheksverband e.V. (dbv) und Schwerpunktinitiative Digitale Information der Allianz der deutschen Wissenschaftsorganisationen, 27.11.2019, <https://dbv-cs.e-fork.net/sites/default/files/2020-11/2019_11_27_dbv_Stellungnahme_Empfehlungen%20zu%20Methoden%20zur%20Kontrolle%20des%20Zugriffs%20auf%20wissenschaftliche%20Informationsressourcen.pdf>, Stand: 20.02.2022.

- DFG: Datentracking in der Wissenschaft. Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Bonn 2021. Online: <https://www.dfg.de/download/pdf/foerderung/programme/lis/datentracking_papier_de.pdf>, Stand: 20.02.2022.

- Dobusch, Leonhard: Neues vom Großverlag Elsevier. Kein Open-Access-Deal, dafür mit Spyware gegen Schattenbibliotheken?, Netzpolitik, 26.10.2020, <https://netzpolitik.org/2020/neues-vom-grossverlag-elsevier-kein-open-access-deal-dafuer-mit-spyware-gegen-schattenbibliotheken/>, Stand: 20.02.2022.

- Engeler, Malte: Das neue Telekommunikation-Telemedien-Datenschutzgesetz. Was es über das regulatorische Klima der deutschen Datenschutzpolitik verrät, Telemedicus 14.07.2021, <https://www.telemedicus.info/soko21-das-neue-telekommunikation-telemedien-datenschutzgesetz-was-es-ueber-das-regulatorische-klima-der-deutschen-datenschutzpolitik-verraet/>, Stand: 20.02.2022.

- Gehring, Petra: Das Schicksal von Open Science steht auf dem Spiel, Forschung & Lehre 02.08.2021, <https://www.forschung-und-lehre.de/politik/das-schicksal-von-open-science-steht-auf-dem-spiel-3902>, Stand: 20.02.2022.

- Gray, Mary L., Suri, Siddharth: Ghost Work. How to Stop Silicon Valley from Building a New Global Underclass. Boston 2019.

- Hanson, Cody: User Tracking on Academic Publisher Platforms, <https://www.codyh.com/writing/tracking.html>, Stand: 20.02.2022.

- Hellman, Eric: 16 of the top 20 research journals let ad networks spy on their readers, Go to Hellman 12.03.2015, <https://go-to-hellman.blogspot.com/2015/03/16-of-top-20-research-journals-let-ad.html>, Stand: 20.02.2022.

- Hinchliffe, Lisa Janicke: Elsevier's ScienceDirect as Content Supercontinent?, The Scholarly Kitchen 18.01.2022, <https://scholarlykitchen.sspnet.org/2022/01/18/sciencedirect-as-content-supercontinent/>, Stand: 20.02.2022.

- Jung, Justin: UCLA School of Law holds contracts with companies selling personal data to ICE, Daily Bruin 17.07.2020, <https://dailybruin.com/2020/07/17/ucla-school-of-law-holds-contracts-with-companies-selling-personal-data-to-ice>, Stand: 20.02.2022.

- Knecht, Sicco de: Leaked document on Elsevier negotiations sparks controversy, Science-Guide 06.11.2019, <https://www.scienceguide.nl/2019/11/leaked-document-on-elsevier-negotiations-sparks-controversy/>, Stand: 20.02.2022.

- Lamdan, Sarah: When Westlaw Fuels ICE Surveillance. Legal Ethics in the Era of Big Data Policing, N.Y.U. Review of Law & Social Change 43 (2), 2019, S. 255–293. Online: <https://socialchangenyu.com/review/when-westlaw-fuels-ice-surveillance-legal-ethics-in-the-era-of-big-data-policing/>, Stand: 20.02.2022.

- Dies.: Librarianship at the Crossroads of ICE Surveillance, In the library with the lead pipe 13.11.2019, <https://www.inthelibrarywiththeleadpipe.org/2019/ice-surveillance/>, Stand: 20.02.2022.

- Lauer, Gerhard: Lesen im digitalen Zeitalter. Darmstadt 2020.

- Lentsch, Justus: Unsere Bildungsdaten gehören uns! Wiarda Blog 16.02.2021, <https://www.jmwiarda.de/2021/02/16/unsere-bildungsdaten-geh%C3%B6ren-uns/>, Stand: 20.02.2022.

- Lynch, Clifford: The Rise of Reading Analytics and the Emerging Calculus of Reader Privacy in the Digital World, in: First Monday 4 (22), 2017. Online: <https://doi.org/10.5210/fm.v22i4.7414>.

- Lyons, Kim: Thomson Reuters faces pressure over ICE contracts, The Verge 21.05.2020, <https://www.theverge.com/2020/5/21/21266431/thomson-reuters-ice-clear-software>, Stand: 20.02.2022

- McCall, Barry: Propelling Ireland to first-mover status in research and innovation, in: The Irish Times 26.08.2021. Online: <https://www.irishtimes.com/sponsored/innovation-partner-profiles/propelling-ireland-to-first-mover-status-in-research-and-innovation-1.4655112>, Stand: 20.02.2022.

- McKie, Robin: Climategate 10 years on. What lessons have we learned? The Guardian, 09.11.2019, <https://www.theguardian.com/theobserver/2019/nov/09/climategate-10-years-on-what-lessons-have-we-learned>, Stand: 20.02.2022.

- McLean, Jaclyn; Stregger, Elizabeth: Sounding the Alarm. Scholarly Information and Global Information Companies in 2021. Partnership, in: The Canadian Journal of Library and Information Practice and Research 2 (16), 2021, S. 1–7. Online: <https://doi.org/10.21083/partnership.v16i2.6692>.

- Metha, Gautama: Proposal to install spyware in university libraries to protect copyrights shocks academics, Coda 13.11.2020, <https://www.codastory.com/authoritarian-tech/spyware-in-libraries/>, Stand: 20.02.2022.

- Morrison, Heather u.a.: Open access article processing charges 2011–2021, Sustaining the Knowledge Commons 24.06.2021, <https://sustainingknowledgecommons.org/2021/06/24/open-access-article-processing-charges-2011-2021/>, Stand: 20.02.2022.

- Nemec, Dan: Ebay is port scanning visitors to their website. And they aren't the only ones, nem.ec 24.05.2020, <https://blog.nem.ec/2020/05/24/ebay-port-scanning/>, Stand: 20.02.2022.

- Neuhausen, Hubertus: Open Access? – Ist durch! Aber was ist mit den Daten? ABI Technik 40 (3), 2020, S. 277–291. Online: <https://www.degruyter.com/document/doi/10.1515/abitech-2020-2022/html>, Stand: 20.02.2022.

- Nogrady, Bianca: 'I hope you die'. How the COVID pandemic unleashed attacks on scientists, Nature 13.10.2021, <https://www.nature.com/articles/d41586-021-02741-x>, Stand: 20.02.2022.

- Pooley, Jefferson D.: Surveillance Publishing. Working Paper 16.11.2021, <https://osf.io/preprints/socarxiv/j6ung/download>, Stand: 20.02.2022.

- Posada, Alejandro; Chen, George: Inequality in Knowledge Production. The integration of Academic Infrastructure by Big Publishers, ELPUB 2018, <https://dx.doi.org/10.4000/proceedings.elpub.2018.30>.

- RfII: Nutzung und Verwertung von Daten im wissenschaftlichen Raum. Empfehlungen zur Ausgestaltung von Datendiensten an der Schnittstelle zwischen Wissenschaft und Wirtschaft. Göttingen 2021, S. 15. Online: <https://rfii.de/download/nutzung-und-verwertung-von-daten-im-wissenschaftlichen-raum-september-2021/>, Stand: 20.02.2022.

- Röttger, Tania: Auf dem Weg zum Digital Services Act. Wie die EU Gesetze gegen Desinformation macht, Correctiv 26.03.2021, <https://correctiv.org/faktencheck/hintergrund/2021/03/26/auf-dem-weg-zum-digital-services-act-wie-die-eu-gesetze-gegen-desinformation-macht/>, Stand: 20.02.2022.

- Schonfeld, Roger C.: The Supercontinent of Scholarly Publishing?, The Scholarly Kitchen 03.05.2018, <https://scholarlykitchen.sspnet.org/2018/05/03/supercontinent-scholarly-publishing/>, Stand: 20.02.2022.

- Ders.: Clarivate to acquire ProQuest, The Scholarly Kitchen 18.05.2021, <https://scholarlykitchen.sspnet.org/2021/05/18/clarivate-to-acquire-proquest/>, Stand: 20.02.2022.

- Ders.: The New Clarivate Science. A Second-Order Consequence of Open Access, The Scholarly Kitchen 9.12.2021, <https://scholarlykitchen.sspnet.org/2021/12/09/new-clarivate-science/>, Stand: 20.02.2022.

- Siems, Renke: Unser industrielles Erbe. Bibliotheken und die digitale Transformation, in: o-bib 4 (3), 2017. <https://doi.org/10.5282/o-bib/2017H3S1-15>.

- Ders.: When your journal reads you. User tracking on science publisher platforms. Elephant in the Lab, 14.04.2021. <https://doi.org/10.5281/zenodo.4683778>.

- SPARC: Landscape Analysis, 29.03,2019, <https://infrastructure.sparcopen.org/landscape-analysis>, Stand: 20.02.2022.

- Dies.: Update Landscape Analysis, <https://sparcopen.org/news/2020/sparc-releases-update-to-landscape-analysis-and-accompanying-interactive-website/>, Stand: 20.02.2022.

- Dies.: Adressing the Alarming Systems of Surveillance Systems built by Library Vendors, <https://sparcopen.org/news/2021/addressing-the-alarming-systems-of-surveillance-built-by-library-vendors/>, Stand: 20.02.2022.

- Stone, Brad: Amazon erases Orwells Books from Kindle, in: New York Times 17.07.2009. Online: <https://www.nytimes.com/2009/07/18/technology/companies/18amazon.html?_r=2>, Stand: 20.02.2022.

- Vogel, Christian: Kennen Sie Google CASA? medinfo. Informationen aus Medizin, Bibliothek und Fachpresse, 08.07.2020, <https://www.medinfo-agmb.de/archives/2020/07/08/6880>, Stand: 20.02.2022.

- Weinberg, Ulrich: Network Thinking. Was kommt nach dem Brockhaus-Denken? Hamburg 2015.

- Weingart, Peter; Taubert, Niels (Hrsg.): Wissenschaftliches Publizieren. Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung. Berlin 2016, S. 109. Online: <https://edoc.bbaw.de/files/2662/00_FB38_WissenschaftlichesPublizieren_gesamt_edoc.pdf>, Stand: 20.02.2022.

- Wissenschaftsrat: Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access. Bonn 2022. Online: <https://www.wissenschaftsrat.de/download/2022/9477-22.pdf>, Stand: 20.02.2022.

- Zuboff, Shoshana: Das Zeitalter des Überwachungskapitalismus. Frankfurt/M. 2018.