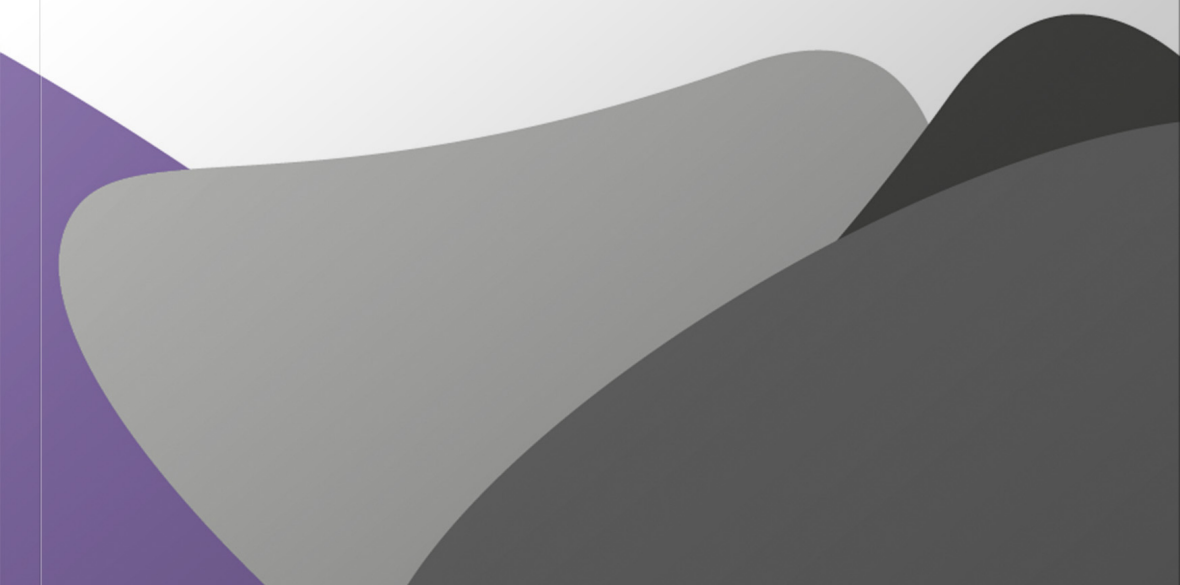




# Wissenschaftstracking und Lock-in-Effekt

Ein Deal gegen digitale Souveränität

*Yuliya Fadeeva*





**Wissenschaftstracking und Lock-in-Effekt  
Ein Deal gegen digitale Souveränität**

Yuliya Fadeeva

Publiziert in Melusina Press (Universität Luxemburg), 2026  
11, Porte des Sciences  
L-4366 Esch-sur-Alzette  
<https://www.melusinapress.lu>

Verlagsleitung: Niels-Oliver Walkowski, Johannes Pause  
Lektorat: Carolyn Knaup, Niels-Oliver Walkowski  
Gestaltung: Valentin Henning, Erik Seitz

Die digitale Version dieser Publikation steht unter <https://www.melusinapress.lu>  
frei zur Verfügung.

Bibliografische Information der Nationalbibliothek Luxemburg: Die  
Nationalbibliothek Luxemburg verzeichnet diese Publikation in der  
Luxemburgischen Nationalbibliografie; detaillierte bibliografische Daten sind im  
Internet über [bnl.public.lu](http://bnl.public.lu) abrufbar.

DOI (Publication): 10.26298/1981-5920  
ISBN (Web): 978-2-919815-92-0  
ISBN (PDF): 978-2-919815-94-4  
ISBN (Epub): 978-2-919815-93-7

Das vorliegende Werk steht unter einer CC BY-SA 4.0 Lizenz. Informationen zu  
dieser Lizenz finden Sie unter [https://creativecommons.org/licenses/by-sa/4.0/  
deed.de](https://creativecommons.org/licenses/by-sa/4.0/deed.de). Die in diesem Werk enthaltenen Bilder und Ressourcen unterliegen der  
selben Lizenz, sofern sie keiner anderen Quelle entnommen wurden oder mit  
einer anderen Lizenz versehen sind.





2026

# Danksagung

Diese Arbeit basiert auf der master thesis im berufsbegleitenden Masterstudiengang Bibliotheks- und Informationswissenschaft (MALIS) an der Technischen Hochschule Köln, die im Juli 2024 eingereicht wurde.

Die Idee dazu entstand während meiner Beschäftigung an der Universitätsbibliothek Duisburg-Essen 2020-2023. Für zahlreiche Förderung, Begleitung, Hilfsbereitschaft und eine enorme Perspektiverweiterung bin ich meiner Universitätsbibliothek sehr verbunden.

Für die ausgezeichnete Betreuung als externer Erstgutachter danke ich herzlich Dr. Renke Siems, ohne dessen Hilfe und Unterstützung es diese Arbeit nicht gegeben hätte. Ich habe sehr von fachlichen Beiträgen von Dr. Robert Altschaffel und Dr. Joschka Selinger profitiert. In Bezug auf die Veröffentlichung danke ich insbesondere Prof. Dr. Gerhard Lauer, Prof. Dr. Björn Brembs und Prof. Dr. Michael Beurskens für wertvollen Austausch. Für die jahrelange Unterstützung, produktive Diskussionen und zahllose Anregungen danke ich Prof. Dr. Jens Loenhoff und dem Institut für Kommunikationswissenschaft der Universität Duisburg-Essen. Last but not least danke ich den vielen hervorragenden Menschen um mich, die monatelang klaglos ausgehalten haben, dass ich fast ausschließlich über das im Anschluss Folgende sprach.

10.12.2025,  
Yuliya Fadeeva

# Inhalt

|                                                                                          |            |
|------------------------------------------------------------------------------------------|------------|
| <b>EINLEITUNG</b>                                                                        | <b>14</b>  |
| <b>1. DIGITALE SOUVERÄNITÄT</b>                                                          | <b>23</b>  |
| 1.1 Abhängigkeiten und Bedrohungen                                                       | 28         |
| 1.2 Datensouveränität und Empfehlungen für eine souveräne Wissenschaft                   | 30         |
| 1.3 Privatheit                                                                           | 35         |
| 1.4 Daten(schutz)recht und die Anwendung auf Wissenschaft                                | 44         |
| <b>2. LOCK-IN I ZEITSCHRIFTENKRISE</b>                                                   | <b>55</b>  |
| 2.1 Entwicklung der Pfadabhängigkeit im wissenschaftlichen Publizieren                   | 60         |
| 2.2 Bibliometrisierung der Wissenschaft                                                  | 66         |
| 2.3 Transformationsprozess der Verlage:<br>Transformationsverträge und Plattformisierung | 76         |
| <b>3. LOCK-IN II INSTITUTION - ELSEVIER UND RELX</b>                                     | <b>89</b>  |
| 3.1 Adaptionsprozess an Open Access: Elseviers<br>Durchdringung der Institutionsebenen   | 96         |
| <b>4. LOCK-IN II INSTITUTION - TRACKING</b>                                              | <b>131</b> |
| 4.1 Definition und Einbettung des Trackings in die Werbe- und Analytikindustrie          | 132        |
| 4.2 Formen des Trackings in der Wissenschaft                                             | 159        |
| 4.3 RELX Lock-In II Institution                                                          | 165        |
| 4.4 (Digitale) Souveränität in der Wissenschaft                                          | 187        |
| 4.5 Privatheit und Wissenschaft                                                          | 189        |
| <b>SCHLUSS</b>                                                                           | <b>200</b> |

**LITERATUR**

**202**

**ANHANG**

**229**



# Abbildungen

|                      |                                                                                                                  |     |
|----------------------|------------------------------------------------------------------------------------------------------------------|-----|
| <b>ABBILDUNG 1:</b>  | Umstrittene Dimensionen von Privatheit bei Mulligan, Koopman und Doty                                            | 42  |
| <b>ABBILDUNG 2:</b>  | Drei-Phasen-Modell der „organisational path dependence“                                                          | 60  |
| <b>ABBILDUNG 3:</b>  | RELX (Elsevier) mergers & acquisitions                                                                           | 98  |
| <b>ABBILDUNG 4:</b>  | Forschungsprozess-Präsenz Elsevier                                                                               | 100 |
| <b>ABBILDUNG 5:</b>  | Elsevier Presence Throughout the Research Lifecycle                                                              | 101 |
| <b>ABBILDUNG 6:</b>  | Elsevier CRIS                                                                                                    | 107 |
| <b>ABBILDUNG 7:</b>  | SPARC map data tools and services: Elsevier, Clarivate, DigitalScience                                           | 124 |
| <b>ABBILDUNG 8:</b>  | Beschleunigtes Drei-Phasen-Modell der „organisational path dependence“                                           | 126 |
| <b>ABBILDUNG 9:</b>  | Acxiom and its data providers, partners, and clients                                                             | 145 |
| <b>ABBILDUNG 10:</b> | Broadcasting personenbezogener Daten                                                                             | 148 |
| <b>ABBILDUNG 11:</b> | Datensatz-Ausschnitt mit Berufsgruppen-Angaben zu Wissenschaftler:innen, geliefert u. a. von LexisNexis an Xandr | 154 |
| <b>ABBILDUNG 12:</b> | Schematisches Web-Tracking und Cross-Domain-Tracking                                                             | 230 |
| <b>ABBILDUNG 13:</b> | Übersicht der Datentypen durch Datenbroker Hum                                                                   | 231 |
| <b>ABBILDUNG 14:</b> | Timeline Elsevier Rankings Involvement                                                                           | 232 |
| <b>ABBILDUNG 15:</b> | Server-Side-Tracking                                                                                             | 233 |
| <b>ABBILDUNG 16:</b> | Schematische Darstellung des Third-Party-Trackings                                                               | 233 |

|                                                                                                        |     |
|--------------------------------------------------------------------------------------------------------|-----|
| <b>ABBILDUNG 17:</b> Tägliche Datenübertragungen pro Person und pro Land im weiteren europäischen Raum | 234 |
| <b>ABBILDUNG 18:</b> LexisNexis Digital ID Netzwerk                                                    | 235 |
| <b>ABBILDUNG 19:</b> RELX Risk-Datenaufbereitung                                                       | 236 |
| <b>ABBILDUNG 20:</b> RELX Business Services                                                            | 236 |
| <b>ABBILDUNG 21:</b> RELX Insurance – holistic assessment of risk                                      | 237 |
| <b>ABBILDUNG 22:</b> LexisNexis BehavioSec behavioral biometrics                                       | 237 |
| <b>ABBILDUNG 23:</b> LexisNexis Precise View of Identity                                               | 238 |
| <b>ABBILDUNG 24:</b> Cookie Notice Elsevier engl.                                                      | 239 |
| <b>ABBILDUNG 25:</b> Cookie Notice Elsevier dt.                                                        | 239 |
| <b>ABBILDUNG 26:</b> Zusammenhang der Elsevier Datenschutz-Dokumente                                   | 240 |

# Tabellen

|                    |                                                                               |     |
|--------------------|-------------------------------------------------------------------------------|-----|
| <b>TABELLE 1:</b>  | Analysetool in Anwendung auf Wissenschaft                                     | 43  |
| <b>TABELLE 2:</b>  | Kriterien digitaler Souveränität in Bezug auf Plattformen                     | 85  |
| <b>TABELLE 3:</b>  | Ausprägung digitaler Souveränität in Bezug auf wissenschaftliche Plattformen  | 87  |
| <b>TABELLE 4:</b>  | Beispiele für Anwendungen der Professional Services                           | 115 |
| <b>TABELLE 5:</b>  | Übersicht Tracking- und Überwachungsmethoden in der Wissenschaft (Teil I)     | 162 |
| <b>TABELLE 6:</b>  | Übersicht Tracking- und Überwachungsmethoden in der Wissenschaft (Teil II)    | 163 |
| <b>TABELLE 7:</b>  | Cookie-Erklärung, deutsch (mit ThreatMetrix) und englisch (ohne ThreatMetrix) | 178 |
| <b>TABELLE 8:</b>  | RELX-Informationsbestand Risk                                                 | 179 |
| <b>TABELLE 9:</b>  | RELX-Informationsbestand Elsevier (STM).                                      | 182 |
| <b>TABELLE 10:</b> | Analyse tool in Anwendung auf die Wissenschaft                                | 191 |
| <b>TABELLE 11:</b> | Privatheitsproblematik für die Wissenschaft (Teil I)                          | 195 |
| <b>TABELLE 12:</b> | Privatheitsproblematik für die Wissenschaft (Teil II)                         | 197 |



# Einleitung

Internationale Großverlage bzw. Medienkonzerne profitieren enorm von der Digitalisierung an der Schnittstelle zwischen Publikationswesen und Wissenschaftsevolution (Holzer, 2022, 164). Sie profitieren, obwohl Wissenschaftler:innen und Institutionen mit ihren Leistungen und Geschäftspraktiken zunehmend unzufrieden sind: “Five for-profit publishing companies—Elsevier, Springer Nature, Wiley Blackwell, Taylor and Francis, and Sage Publications—dominate the market, generating more than 50 percent of revenues (over \$7 billion in 2022) with profit margins up to 38 percent—higher than big tech companies.” (Gulliver & Drake, n.d.) Verschiedene Entwicklungen der letzten Jahrzehnte haben die Situation im wissenschaftlichen Publizieren und zunehmend im gesamten Wissenschaftssektor negativ beeinflusst und führen langsam dazu, das „Sondermilieu Wissenschaft aufzulösen“ und in die Datenökonomie der Web-Konzerne zu integrieren (Siems, 2024c). Dazu zählen neben der genannten Marktkonzentration vor allem etablierte Strukturen im wissenschaftlichen Publizieren, die einseitig die großen Konzerne begünstigen und zu Pfadabhängigkeiten bei Wissenschaftsorganisationen geführt haben. Das asymmetrische und „staggeringly profitable“ Modell wissenschaftlichen Publizierens beschreibt Stephen Buranyi in *The Guardian*:

[S]cientific publishers manage to duck most of the actual costs. Scientists create work under their own direction – funded largely by governments – and give it to publishers for free; the publisher pays scientific editors who judge whether the work is worth publishing and check its grammar, but the bulk of the editorial burden – checking the scientific validity and evaluating the experiments, a process known as peer review – is done by working scientists on a volunteer basis. The publishers then sell the product back to government funded institutional and university libraries, to be read by scientists – who, in a collective sense, created the product in the first place. (Buranyi, 2017)

Zugleich haben sich die Erwartungen an positive Strukturveränderungen durch Open Access nicht bestätigt. Die Konzerne haben mit APCs (Article Processing Charges), die ein Vielfaches der tatsächlichen Kosten betragen und mit Transformationsverträgen Geschäftsmodelle für sich gefunden, die für sie nicht weniger lukrativ sind als die Subskriptionsmodelle. Zugleich werden damit die Abhängigkeiten der Wissenschaftseinrichtungen auf internationaler oder gar globaler Ebene in die Zukunft zementiert:

Global systems for research publishing are in bad shape. In 2022, the five biggest publishers had a combined revenue of US\$7.7 billion, with profit margins of up to 38 percent. They made only 31 percent of their articles Open Access (without paywall) and only 25 percent of their journals were fully Open Access. The median publication charge was \$2,860 but could be as high as \$11,690. (Drake et al., n.d.)

Competition in the scholarly publishing market is inhibited by the non-substitutability of journals, lack of transparency and high levels of market concentration. Meanwhile, the revenue gap between subscription and OA business models limits the rate of transition. The case for market intervention rests on the status of scientific knowledge as a public good, disseminated by private actors. (Johnson et al., 2017, 27)

Hinzu kommen interne Transformationen der einstigen Großverlage zu Datenanalyse- und Brokering-Konzernen, deren Geschäftsmodelle auf der Sammlung und Verwertung von Daten beruhen. Diese Daten sammeln die Konzerne durch Interaktionen mit Wissenschaftler:innen und wissenschaftlichen Einrichtungen, aber auch auf anderen Wegen. Folgendes Zitat Jeff Pooleys umreißt die digital-politische Dimension der Problematik und nennt zwei Schlüsselwerke der letzten Jahre in der Diskussion digitaler Überwachungs- und Brokering-Ökonomie, Sarah Lamdan mit *Data Cartels* (2022) und Shoshanna Zuboffs *The Age of Surveillance Capitalism* (2019):

Elsevier and the other big publishers are, to borrow Sarah Lamdan's phrase, data cartels. I've called this drive to extract profit from researchers' behavior *surveillance publishing*—by analogy to Shoshanna Zuboff's notion of surveillance capitalism, in which firms like

Google and Meta package user data to sell to advertisers. The core business strategy is the same for Silicon Valley and Elsevier: extract data from behavior to feed predictive models that, in turn, get refined and sold to customers. In one case it's Facebook posts and in the other abstracts and citations, but either way the point is to mint money from the by-products of (consumer or scholarly) behavior. One big difference between the big tech firms and the publishers is that Google et al entice users with free services like Gmail: If you're not paying for it, the adage goes, then you're the product. In the Elsevier case we're the product and we're paying (a lot) for it. (Pooley, 2024)

Die Open-Access-Transformation ist ein dominierendes Thema im wissenschaftlichen Publikationsfeld, das hier nur in Auszügen diskutiert wird, die der gesamten Komplexität und Ambivalenz nicht gerecht werden können. Die vorliegende Arbeit betrifft nur einen Aspekt der Transformation und zwar den Abschluss des dritten bundesweiten Transformationsvertrags (DEAL-Vertrag) im September 2023 (bis Ende 2028) mit Elsevier. Der DEAL-Vertrag wurde nach einem jahrelangen Elsevier-Boykott der deutschen Einrichtungen geschlossen, der Abschluss wurde durchaus kritisch gesehen.<sup>1</sup> Es gibt zahlreiche diskussionswürdige Aspekte des Vertrags, die an anderer Stelle zu besprechen wären bzw. bereits besprochen werden. Mir dient der DEAL-Vertrag als ein Anlass dafür, einen übergeordneten Fragenkomplex zu beleuchten: Einerseits, wie das aktuell diskutierte Thema des umfassenden und invasiven Trackings der einstigen Wissenschaftsverlage, am Beispiel Elseviers, in den Datenschutz-Konditionen des DEAL-Vertrags umgesetzt wurde. Hier gibt es Gründe zur Annahme, dass es sich um eine unglückliche Umsetzung handelt. Dies spiegelt die bisher fehlende Adressierung des Datentracking in Lizenzverträgen wider, jedenfalls laut der Einschätzung des AWBI (Altshaffel *et al.*, 2024, 27):

In den meisten der jüngeren Verträge zwischen Verlagen und Bibliotheken behalten sich die Verlage Kontrolle über die Daten vor, die bei der Nutzung ihrer Medien generiert werden. Die Bibliotheken bleiben von den Verlagen und deren Geschäftsmodellen abhän-

---

<sup>1</sup> Brembs *et al.* (2023a); Brembs (2023b); Dirnagl (2023); Herb (2023); Siems (2023a).

gig und scheinen daher nur begrenzt Maßnahmen zum Schutz der Nutzerdaten ergreifen zu können. (Altschaffel *et al.*, 2024, 27)

Andererseits geht es um das Aufkommen von Anwendungen<sup>2</sup> für die Wissenschaftsevaluation und -verwaltung, vor allem bei Elsevier und anderen Konzernen wie Clarivate und Digital Science. In Kombination mit Anwendungen für den Forschungskreislauf derselben Anbieter können diese Entwicklungen herbeiführen, die eine ernsthafte Gefahr für die Wissenschaft, Wissenschaftler:innen und wissenschaftliche Institutionen darstellen. Die Argumentation in den Kapiteln zum Lock-In II hat in Bezug auf die mögliche Nutzung dieser weiteren Anwendungen einen dystopischen bzw. warnenden Charakter; es ist keine Aussage über konkrete Einrichtungen<sup>3</sup>. Im Kapitel Lock-In II – Tracking geht es nicht um Enthüllungen oder (neuartige) Nachweise der Weitergabe von Daten zwischen den RELX-Tochterunternehmen Elsevier und LexisNexis® Risk Solutions<sup>4</sup>. Es geht um die Zusammenführung der bereits bekannten und in Elseviers Privacy Policy festgehaltenen Formen der Sammlung und Verwertung von Daten. Mein Ziel ist, die grundsätzliche Eignung und Legitimität der Zusammenarbeit mit dem Konzern (stellvertretend auch für andere Konzerne dieser Art) zu hinterfragen. Für dieses Ziel liegen in der Literatur bereits hinreichend viele hinreichend deutliche Nachweise über Praktiken und Interessenkonflikte vor. Diese Literatur sowie die öffentlich verfügbaren Informationen von und über RELX und seine Tochterunternehmen sind die wesentlichen Quellen dieser Arbeit. Die hier gezo-

---

<sup>2</sup> Ich folge häufig der englischen Ausdrucksweise oder ihren deutschsprachigen Entsprechungen, um Programme zu benennen: „tool“, „solution“, „Anwendung“, „Produkt“, „Instrument“, „Ranking“.

<sup>3</sup> Generell treffe ich hier keine Aussagen darüber, welche deutsche Einrichtung welche Anwendungen nutzt und welche Auswirkungen dadurch konkret erwartet oder beobachtet werden könnten. Trotzdem sollte angemerkt werden, dass zahlreiche deutsche Universitäten mindestens einige der im weiteren Verlauf genannten Anwendungen wie Pure verwenden. Die Kommission für Forschungsinformationen in Deutschland erfasst aktuell 130 deutsche Einrichtungen, die CRIS nutzen. Führend ist die Anwendung von HISinOne-RES mit 31,1% Anteil und 47 nutzenden Institutionen, gefolgt von Pure mit 12,6% (17 Institutionen). Converis von Clarivate liegt mit 5,9% (8 Institutionen) deutlich dahinter. (Stand Juli 2025) [https://kfid-online.de/fis\\_landkarte\\_visual/fis\\_landkarte\\_statistik.php](https://kfid-online.de/fis_landkarte_visual/fis_landkarte_statistik.php) [zuletzt geprüft am 4.7.2025]

<sup>4</sup> Ab jetzt meist abgekürzt als ‚LexisNexis‘, womit nicht die gleichnamige juristische Datenbank-Anwendung gemeint ist. Letztere wird in dieser Arbeit nicht thematisiert.

genen Schlüsse sind nicht neu, sondern wurden sowohl von Akteur:innen (u.a. Björn Brembs, Gerhard Lauer, Renke Siems) als auch von Institutionen (u.a. Deutsche Forschungsgemeinschaft DFG, Wissenschaftsrat, Scholarly Publishing and Academic Resources Coalition SPARC) formuliert.

Ich verteidige keine digitalisierungsfeindliche Position, sondern eine aktive Gestaltung digitaler Souveränität in der Wissenschaft, ganz im Sinne der Ausschöpfung digitaler Potenziale<sup>5</sup>. Damit stellt sich die grundsätzliche Frage, ob die unabhängige, integre Wissenschaft (Merton, 1973) im Verständnis moderner demokratischer Gesellschaften weiterhin mit den „big five“, speziell Elsevier, betrieben werden kann. Dabei nehme ich an, dass Elsevier zustimmt, die aktuell als wahrscheinlich rechtswidrig eingestuften Tracking-Praktiken zu unterlassen. Ich formuliere diese Frage in Form zweier konkurrierender Thesen und prüfe im Verlauf der Arbeit, welche These sich bestätigen lässt. Beispiel meiner Argumentation ist der Konzern RELX, genauer das Tochterunternehmen Elsevier, weil Elsevier als führender Konzern unter den „big five“ die kritisierten Aspekte und Praktiken in der Verbindung von Datenanalyse und Publikationstätigkeit sowie in der Beziehung zur Sicherheitsindustrie (LexisNexis) am meisten exemplifiziert. Die konkurrierenden Thesen lauten:

1. Unabhängige Wissenschaft ist in Zusammenarbeit mit Elsevier möglich, weil Elsevier seine Angebote ohne inakzeptable Tracking-Praktiken betreiben könnte.
2. Unabhängige Wissenschaft ist in Zusammenarbeit mit Elsevier nicht möglich, weil Elsevier seine Angebote nicht ohne inakzeptable Tracking-Praktiken betreiben könnte.

In Anschluss an Zuboffs Rede von „Überwachung auf der Basis von Asymmetrien von Wissen und Macht“ (Zuboff, 2019, 103 (dt. Ausgabe)) argumentiere ich, dass durch das kontinuierliche und umfassende Sammeln von Daten verschiedene Asymmetrien entstehen. Bekannt ist die *Machta-*

---

<sup>5</sup> DFG Deutsche Forschungsgemeinschaft (2020); Wissenschaftsrat (2022); Schwerpunktinitiative „Digitale Information“ der Allianz der Deutschen Wissenschaftsorganisationen (2019).

*symmetrie* zwischen den Großverlagen einerseits und Wissenschaftseinrichtungen andererseits (hier diskutiert als Lock-In I). Durch die massiven Aufkäufe wissenschaftlicher Anwendungen und fehlende Investitionen in digitale Infrastrukturen kann daher auch eine *funktionale* Asymmetrie verzeichnet werden. Am schwersten wiegt jedoch die Gefahr einer *epistemischen* Asymmetrie, die durch einen systematischen Wissensvorsprung entsteht, den Elsevier und andere durch verschiedene Praktiken und Anwendungen aufbauen. Die Zusammenführung dieser Asymmetrien durch institutionsdurchdringende und -formende Anwendungen, Tracking und Transformationsverträge führt, so argumentiere ich, zu einem nicht nur ökonomischen und systematischen, sondern auch epistemischen institutionsweiten Lock-In. Diese Entwicklung sollte möglichst vermieden werden, um das Ziel einer unabhängigen, handlungsfähigen Wissenschaft und digitaler Souveränität zu erreichen.

Im nächsten Kapitel diskutiere ich verschiedene Konzepte digitaler Souveränität, der Datensouveränität und des Datenschutzes. Mit Hilfe des Begriffs der Privatheit versuche ich, spezifische schädliche Aspekte für die Wissenschaft zu erfassen, die durch das systematische Tracking entstehen und durch die Konzepte der Souveränität und des Datenschutzes nicht ganz erfasst werden. Einen Überblick der multidimensionalen Prozesse, die zur Entwicklung der heutigen Situation in der Publikations- und Wissenschaftsevaluationspraxis geführt haben, bietet das Kapitel Lock-In I. Das Erklärungsmuster einer organisatorischen Pfadabhängigkeit, die zum Phänomen des Lock-In führt, wende ich auf die Situation im wissenschaftlichen Publikationssystem an, die als sogenannte Zeitschriftenkrise mit zur Open-Access-Bewegung führte. Im Kapitel Lock-In II Institution – Elsevier und RELX werden der Konzern RELX, das Tochterunternehmen Elsevier und LexisNexis genauer untersucht. Die Ausrichtung Elseviers als Datenanalyse-Unternehmen zeigt sich an den vielschichtigen Ansätzen, die wissenschaftlichen und institutionellen Strukturen der Wissenschaft zu durchdringen, Einflussmechanismen auszubauen und miteinander systematisch zu verzahnen. Die zahlreichen Trackingformen, die für den Wissenschaftsbereich relevant sind, diskutiere ich im Kapitel Lock-In II Institution – Tracking, wo auch die Diskussion der digitalen Souveränität wieder aufgenommen wird. Die Anwendung der unterschiedlichen Argumentationslinien auf den DEAL-Vertrag mit Elsevier findet durchgehend statt.

Einige wichtige Themen sind aus Platzgründen zu kurz gekommen, u. a. die Diskussion der Folgen des massiven Einsatzes von Artificial Intelligence (AI) in akademischen Prozessen (wie ein weiteres Anschwellen des Publikations-Outputs, steigende Unübersichtlichkeit und der bereits verzeichnete Qualitätsverlust der Forschung), Untersuchungen der weiteren politischen Involvierung von LexisNexis, die gesamte medizinisch-pflegerische Branche Elseviers oder auch ein Vergleich zwischen den Big-five-Konzernen.





1.

# **Digitale Souveränität**

Die Rede von digitaler Souveränität hat sich in den letzten Jahren verbreitet, z. B. in Kontexten der digitalen und infrastrukturellen (nationalen) Selbständigkeit und Sicherheit vor fremder Überwachung bzw. Einflussnahme sowie allgemein als ein Ziel der europäischen Digitalisierung. Beispiel für das letztere Ziel ist der Abschnitt zum Ausbau der digitalen Souveränität in Europa im Programm der deutschen EU-Ratspräsidentschaft 2020<sup>6</sup>. In Deutschland findet sich der Bezug auf digitale Souveränität in Strategie- und Positionspapieren der Regierung, aber auch auf wirtschaftlicher und gesellschaftlicher Ebene wird sie gefordert. Vor allem in Bezug auf die öffentliche Verwaltung scheint das Thema sehr präsent zu sein — u. a. im Bundesministerium des Innern und für Heimat<sup>7</sup> (IT-Planungsrat & IT-Rat, 2020) mit dem Zentrum Digitale Souveränität ZenDiS<sup>8</sup> und im IT-Planungsrat<sup>9</sup>.

Eine allgemein akzeptierte Definition digitaler Souveränität existiert bislang nicht, auch keine terminologische Unterscheidung in ‚digital‘ und ‚technologisch‘<sup>10</sup>. Gehring (2022, Abschnitt 4) moniert ein durchgehendes begriffliches Durcheinander zwischen „digitaler Souveränität“ und „Datensouveränität“, besonders in der Datenstrategie der Bundesregierung (Bundesregierung, 2021).

Kurz zur Begriffsnutzung: Unter ‚Daten‘ werden gemeinhin unverarbeitete, rohe, nicht interpretierte Elemente verstanden, aus denen durch Interpretation, Strukturierung oder weitere Art von Bearbeitung ‚Informationen‘ werden. Ich verwende jedoch in manchen Fällen beide Ausdrücke

---

6 <https://www.auswaertiges-amt.de/blob/2360342/b44ce5e6c00598f286cd43c6a1e8dd93/nat-programm-rp-data.pdf> [zuletzt geprüft am 25.07.2024]

7 <https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/digitale-souveraenitaet/digitale-souveraenitaet-node.html>. [zuletzt geprüft am 25.07.2024]

8 <https://www.zendis.de/>. [zuletzt geprüft am 25.07.2024]

9 IT-Planungsrat und IT-Rat: Stärkung der Digitalen Souveränität der Öffentlichen Verwaltung, Eckpunkte – Ziel und Handlungsfelder (2020), <https://www.cio.bund.de/SharedDocs/downloads/Webs/CIO/DE/digitale-loesungen/eckpunktpapier-digitale-souveraenitaet.pdf?blob=publicationFile&v=2>. [zuletzt geprüft am 25.07.2024]

10 Siehe Pohle (2020) und Gehring (2022) für Diskussionen der Unterscheidung, als politikwissenschaftlich-soziologischer bei Pohle und (rechts)philosophischer Perspektive bei Gehring.

synonym, da diese Unterscheidung in meiner Argumentation keine tragende Rolle spielt.

In der unklaren Begriffslage finden sich entsprechend der heterogenen Perspektiven unterschiedliche Schwerpunkte. Häufig findet ein Lavieren zwischen und Vermischen von nicht explizierten Vorverständnissen (vgl. Hummel *et al.*, 2021) statt. Die Schwerpunktstudie digitale Souveränität des Bundesministeriums für Wirtschaft und Energie aus dem Jahr 2021 (BMW, 2021) sammelt 18 Definitionen digitaler Souveränität im Zeitraum 2015-2021 aus den Bereichen Wirtschaftsverbände<sup>11</sup> und internationale Verbände<sup>12</sup>, aus der IT-Branche<sup>13</sup>, öffentlichen (Forschungs-) Einrichtungen und dem Stiftungswesen<sup>14</sup>. Auch das BMW und das Bundesministerium für Bildung und Forschung BMBF bieten eigene Definitionen. Unterschiede gibt es in den Interessen der vertretenen Akteur:innen, dem Differenzierungsgrad der Beschreibung und Kategorisierung der digitalen Souveränität im Gemenge individueller/staatlicher/nationaler Aufgabe(n), Kompetenzen und Ziele. Die kürzeste Definition gibt das BMW: „Die Fähigkeit zu selbstbestimmten Handeln und Entscheiden im digitalen Raum.“ (BMW, 2017)

Deutlich anders setzen sich das Deutsche Institut für Vertrauen und Sicherheit im Internet mit der Kombination aus vier verschiedenen Definitionen aus den Bereichen Nationalstaaten, Wirtschaft und souveräne Systeme und Individuen (Gräf *et al.*, 2018) oder das BMBF mit dem Thema auseinander:

Unter technologischer Souveränität versteht das BMBF den Anspruch und die Fähigkeit zur kooperativen (Mit-)Gestaltung von Schlüsseltechnologien und technologiebasierten Innovationen. Dies

**11** Verband der Elektrotechnik VDE; Bundesverband der Deutschen Industrie BDI.

**12** CESifo, ein unabhängiges globales Netzwerk, wirtschaftlich-wissenschaftlicher Schwerpunkt.

**13** Digital-Gipfel & Nationaler IT-Gipfel; Bitkom; Acatech.

**14** Deutsches Institut für Vertrauen und Sicherheit im Internet DIVSI; Institut für Innovation und Technik (iit) Berlin; Gesellschaft für Informatik; Fraunhofer-Institut für System- und Innovationsforschung, Sachverständigenrat für Verbraucherfragen SVRV; Kompetenzzentrum öffentliche IT; Bertelsmann Stiftung; Konrad-Adenauer-Stiftung KAS.

umfasst die Fähigkeiten, Anforderungen an Technologien, Produkte und Dienstleistungen entsprechend der eigenen Werte zu formulieren, Schlüsseltechnologien entsprechend dieser Anforderungen (weiter) zu entwickeln und herzustellen sowie Standards auf den globalen Märkten mitzubestimmen. Technologische Souveränität kann dabei auch erfordern, Schlüsseltechnologien und technologiebasierte Innovationen in Europa eigenständig zu entwickeln und hierfür eigene Produktionskapazitäten aufzubauen, wenn es zum Erhalte der staatlichen Handlungsfähigkeit oder zur Vermeidung einseitiger Abhängigkeiten notwendig ist. (BMBF, 2021, 3)

Die Dreiteilung der Handlungsebenen in *Staat* (bzw. Europa, besonders in Bezug auf kritische Infrastruktur), *Privatwirtschaft* und *Individuum* (Kagermann *et al.* (2021); bitkom (2015); Gräf *et al.* (2018); Fokusgruppe (2019); GI (2020); Pohle & Thiel (2019) in der Darstellung des BMWi (2021, 5.1, Übersichtstabelle: Definitionen Digitaler Souveränität)) zeigt, trotz der allgemein wirkenden Formulierung, die standortgebundene Perspektive auf Deutschland und Europa. Diese Standortbindung stellt das Kompetenzzentrum Öffentliche IT am Fraunhofer-Institut für Offene Kommunikationssysteme (Goldacker 2017) als „gesamtgesellschaftliche“ digitale Souveränität dar, die in (geopolitisches) „innen“ und „außen“ zerfällt. Entlang dieser Einteilung unterscheidet das Kompetenzzentrum in Bezug auf Technik bzw. Dienstleistung sowie in organisatorische, wirtschaftliche sowie gesellschaftliche digitale Souveränität:

–die *innere* digitale Souveränität als Gesamtheit der digitalen Souveränität aus individueller, institutioneller und Produzenten-/Anbieterperspektive aller gesellschaftlichen Teile und Gruppen, gepaart mit der Fähigkeit von Staat und Verwaltung, die digitale Souveränität zielführend zu gestalten [...]

–die nach *außen* gerichtete digitale Souveränität, die darauf abzielt, die Gesellschaft gegen Eingriffe in die digitale Souveränität (einschließlich des schleichenden Verlustes derselben) durch wirtschaftliche oder äußere politische Einflüsse zu schützen und möglichst die Abhängigkeit der Gesellschaft von Wirtschaftsunternehmen und anderen Staaten in Bezug auf IT-Komponenten – Hardware und Software – und in Bezug auf IT-Dienstleistungen – insbe-

sondere Plattformen und Clouddienste – zu verringern. (Goldacker, 2017, 9-10)<sup>15</sup>

Es sind vor allem wirtschafts- und industriepolitische Fragestellungen, die im Umgang mit digitaler Souveränität – in welcher Bedeutung sie dann jeweils verstanden wird – eine zentrale Rolle spielen, so Julia Pohle<sup>16</sup> und Thorsten Thiel, die zu diesem Thema forschen (Pohle, 2020; Pohle & Thiel, 2019, 2021; Thiel, 2019). Sie warnen vor dem Hintergrund der aktuellen Popularität des Begriffs davor, Demokratie, europäische Werte und Souveränität zu verschmelzen (Pohle & Thiel, 2021). Thiel (2021) sieht die ideologische Einfärbung des Begriffs und erstarkende Kontrollmechanismen kritisch, die selbst zum Widerspruch mit Freiheitswerten führen könnten.

Pohle (2020) verfolgt die Entwicklung des Souveränitätsbegriffs aus der politischen Philosophie und der Entstehung des Internets als vermeintlich frei-anarchischen Raums<sup>17</sup>. Während ursprünglich der Souverän als absolutistischer Herrscher Macht und Zuständigkeit bündelte – repräsentiert in Thomas Hobbes' Leviathan –, verändert sich der Souveränitätsbegriff später in eine demokratische Richtung. Dabei wird stärker zwischen der Abgrenzung von Nationalstaaten nach außen und der Legitimation

**15** Auch die restlichen Punkte, bis möglicherweise auf den letzten, haben mindestens einen mittelbaren Bezug auf die Funktionsfähigkeit der Wissenschaft:

„–die digitale Souveränität im technischen bzw. Dienstleistungsbereich, die beschreibt, in welchem Umfang die Gesellschaft, als Ganzes betrachtet, in der Lage ist, nachgefragte bzw. Erfolg versprechende IT-Komponenten selbst zu entwickeln und zu produzieren bzw. entsprechende IT-Dienstleistungen anzubieten  
–die organisatorische digitale Souveränität, die bestimmt, in welchem Maß die digitale Souveränität unabhängig von anderen Staaten bzw. marktbestimmenden Wirtschaftsunternehmen durchgesetzt werden kann  
–die wirtschaftlich bestimmte digitale Souveränität als Maß dafür, ob die digitale Souveränität im technischen und im Dienstleistungsbereich unter wirtschaftlichen Gesichtspunkten konkurrenz- und durchsetzungsfähig ist [...]  
–die gesellschaftliche Gestaltung der Digitalisierung, die Chancen und Risiken, Stärken und Herausforderungen für die Gesellschaft durch die Digitalisierung identifiziert und unter Einbeziehung aller gesellschaftliche Gruppen diskutiert [...]“ (vgl. Goldacker, 2017, 9-10)

**16** <https://www.wzb.eu/de/forschung/digitalisierung-und-gesellschaftlicher-wandel/politik-der-digitalisierung/projekte/digitale-souveraenitaet>.  
[zuletzt geprüft am 25.07.2024]

**17** Siehe Pohle & Thiel (2021); Pohle (2020) für den historischen Abriss dieser Entwicklung mit zahlreichen Literaturhinweisen.

des Staates durch die Selbstbestimmung der Bürger:innen nach innen unterschieden. Der Wert der Souveränität und somit die Wichtigkeit, diese zu schützen, ist aus dieser Sicht kaum zu überschätzen, was auch die Zugkraft im öffentlich-politischen Diskurs erklärt: „Die Souveränität eines demokratischen Staates besteht in der Sicherung der Selbstbestimmungsfähigkeit seiner Bürgerinnen und Bürger mit ihren unveräußerlichen Rechten.“ (Pohle, 2020, 6)

Pohles Formulierung zeigt die Zuständigkeit und Verantwortung des Staates in Bezug auf die Selbstbestimmungsfähigkeit und die Rechte seiner Rechtssubjekte am klarsten. Mit diesem klassischen Verständnis von Souveränität ist der Rahmen gezeichnet, in welchem auch der Bereich des Digitalen eingeordnet wird. Es ist dasjenige Gebiet, das die Selbstbestimmung der Bürgerinnen mit ihren unveräußerlichen Rechten betrifft, wie immer dessen Ausprägungen – z. B. sich rasant entwickelnde digitale Technologien – gestaltet sein mögen.

## 1.1 Abhängigkeiten und Bedrohungen

Welche Gefahren bestehen für die digitale Souveränität in ihren verschiedenen Bedeutungen und Handlungsfeldern? Pohle & Thiel (2021, 327) nennen drei Abhängigkeiten im europäischen Diskurs um digitale Souveränität, von denen die ersten zwei in fast gleicher Weise für die Wissenschaft bestehen:

1. Abhängigkeit von digitalen Infrastrukturen, „wonach unsere Gesellschaften auf einen materiellen und immateriellen technologischen Unterbau angewiesen sind. Von Serverfarmen bis Softwarepatenten fußt die digitale Gesellschaft darauf, dass technologische Apparate und Protokolle, die zuverlässig und sicher zu sein haben, ineinandergreifen.“
2. Abhängigkeit von privaten Akteur:innen. Hier treten die großen Plattformunternehmen als „Quasi-Souveräne“ auf, „da sie etwa im Hinblick auf Größe und wirtschaftliche Macht ebenbürtig mit den Staaten

erscheinen und ihre Geschäftsmodelle, zum Beispiel das Kuratieren von Öffentlichkeit oder das Erzeugen proprietärer Märkte, staatsähnlicher sind als frühere Erlösmodelle. [...]“.<sup>18</sup> (Pohle & Thiel, 2021, 327)

Drei Gründe verstärken laut Pohle die Wahrnehmung der digitalen Transformation als einer Bedrohung der (digitalen) Souveränität:

1. Die Vormachtstellung eines Oligopols an Unternehmen (GAFAM)<sup>19</sup> bei zentralen Infrastrukturen des sozialen Lebens, in einem Milieu von aggressiven Werbestrategien, Datenhandel und Netzwerk-Effekt;
2. Die internationalen Reaktionen auf die Enthüllungen von Edward Snowden hätten zwar eine große Beunruhigung angesichts umfassender Überwachung ausgelöst, jedoch keine politisch-rechtlichen Gegenmaßnahmen nach sich gezogen. Stattdessen wurden digitale Schutzräume und Enklaven persönlicher Kontrolle gesucht bzw. geschaffen;
3. Suche nach alternativen normativen Orientierungspunkten für die europäische und deutsche Digitalpolitik, um sich einerseits von der „stark markt-libertären, individualistischen und technopositivistischen“ Perspektive der USA zu lösen und andererseits einen „internationalen Systemkonflikt“ mit China zu vermeiden. (Pohle, 2020, 8)

Diese Punkte machen zwar deutlich, weshalb der Ruf nach digitaler Souveränität bezogen auf den staatlichen Handlungsrahmen so präsent ist, aber Pohle und Thiele sehen das Leitkonzept „digitale Selbstbestimmung“ nicht als Lösung. Sie vergleichen den deutschen Diskurs digitaler Souveränität mit dem deutlich weniger normativ aufgeladenem europäischen, wo eher „strategische Autonomie“ und „technologische Souveränität“ genannt werden (Pohle & Thiel, 2021, 333) und wirtschaftliche Gründe im Vordergrund stehen: „Als primäre Merkmale europäischer Digitalpolitik

---

**18** Die dritte Abhängigkeit betrifft die geopolitische Dimension, insbesondere die Rivalität der USA und Chinas. (Ebd.)

**19** Die Abkürzung steht für die fünf Techgiganten Google, Apple, Facebook, Amazon, Microsoft, wobei es eigentlich korrekter wäre, die Mutterkonzerne zu benennen: Alphabet, Apple, Meta Platforms, Amazon und Microsoft.

gelten [...] wirtschaftliche Stärke und die Bereitstellung einer Infrastruktur, die diese ausbaut und schützt. Diese sollen durch gezielte Wettbewerbsförderung erreicht werden, weniger durch Protektionismus. [...]“ (ebd., 334). Im außereuropäischen Raum gelten nochmal andere politische Richtlinien, weil die Rhetorik (nationaler) Souveränität bei autoritären und hybriden Regimen, die ihre digitalen Kontrollmöglichkeiten der eigenen Bevölkerung legitimieren wollen, schnell ins Negative umschlägt.

Wichtigster Kritikpunkt von Pohle und Thiel ist die Gefahr, die für demokratische Gesellschaften von steigender staatlicher Kontrolle ausgeht. Diese Gefahr sehen sie durch den positiv-präskriptiven Stellenwert des Begriffs digitaler Souveränität, wodurch auch staatliche Kontrollmaßnahmen aufgewertet und legitimiert werden. Im digitalen Kontext bedeuteten „Souveränitätsforderungen fast immer auch [...] das Schaffen und Rechtfertigen neuer Eingriffs- und Beobachtungsbefugnisse [...]“ (ebd., 339). So hätte auch schon der erste „Schub des Souveränitätsdiskurses vor 2010“ eine Katalysatorwirkung auf das Ausmaß und die Formen der „Eingriffspunkte für Staaten und private Dritte in digitale Netzwerke und ihre gesellschaftliche Nutzung“ gehabt (ebd.). Heute sehen Pohle und Thiel den Diskurs digitaler Souveränität „begleitet von immer expliziteren Bemühungen, Datenverkehr und Kommunikationsinfrastrukturen für staatliche Institutionen leichter zugänglich zu machen.“ (Pohle & Thiel, 2021, 339) Diese Kritik steht neben der positiven Wirkung des Begriffs digitaler Souveränität als politische Selbstbestimmung für Demokratie und Rechtsstaat.

## 1.2 Datensouveränität und Empfehlungen für eine souveräne Wissenschaft

Eine andere Kritik verbindet Petra Gehring mit dem Begriff der ‚Datensouveränität‘, sofern es denn gelinge, ihn von digitaler Souveränität zu trennen (vgl. Gehring, 2022). Die Unterschiede sieht sie in der Betonung des ökonomischen Status‘ von Daten, die von Bürger:innen in vermeintlich souveräner Entscheidung als Ressource oder Kapital eingesetzt wer-

den (sollen). Diese Entwicklung gehe mit dem Rückgang politischer Sicherung des Datenschutzes einher. Gehring sieht auch die positive Rolle, die der Begriff als abstraktes Ziel und Bewältigung der Digitalisierung spielt, dass nämlich „die Rede von Datensouveränität [...] wie ein Lauffeuer funktioniert“ und deutsche wie europäische Öffentlichkeit „an demokratische Partizipation und politische Mitgestaltungsansprüche [denken lässt]. Vielleicht sogar an Volkssouveränität.“ Das täusche aber über die Tatsache hinweg, dass es im „Ruf nach Datensouveränität oder digitaler Souveränität“ nicht um nationalstaatliche Interaktion gehe, sondern um das Bedürfnis, „den digitalen Wandel nicht mehr aus einer Position reiner Ohnmacht erleben zu müssen.“ (Gehring, 2022, 42)

Die beschriebene „Abhängigkeiten und Ohnmacht im Zusammenhang mit Digitalisierung“ der Bürger:innen ist durchaus auf die Wissenschaft anwendbar, wenn „Trolle und Netzgiganten“ im folgenden Zitat durch entsprechende Verlagsnamen ersetzt werden:

Viele erleben eben nicht nur Fortschritt, immer schnellere und bessere digitale Vermittlungen, sondern [...] ein Leerlaufen des geltenden Rechts, einen zu schwachen Schutz kritischer Infrastrukturen und eine Ohnmacht von Kartell- und Strafverfolgungsbehörden [...] Wer sich im Netz als Freiwild fühlt, will nicht, wo der Staat eingreift, erneut eher die inländischen Nutzerinnen und Nutzer reglementiert sehen als Trolle und Netzgiganten. Hierauf reagieren auch die Subjekte eines im Kleinen überakribisch, im Großen nachlässig wirkenden Datenschutzes, der sich auf das ‚Klick‘ des immer wieder neuen Zustimmens beschränkt. (Gehring, 2022, 42f.)

An der Digitalen Strategie des Bundeswirtschaftsministeriums von 2016 (BMWi, 2016) kritisiert Gehring, dass die Rede von Datensouveränität genutzt wird, um die staatliche Verpflichtung zur Regelung der Datenerhebung und -verwendung abzulösen und die Verantwortung auf Individuen zu verlagern. Diese sollen ‚souverän‘ die Handhabung und Verantwortung der gesamten Problematik übernehmen:

„Der ‚Datenschutz‘ fehlt, er wird durch das Duo Souveränität und Sicherheit ersetzt. ‚Souverän‘ schütze ich mich allenfalls nun selber. Der datenschutzrechtliche Grundsatz der sogenannten Datensparsamkeit (das Gebot, unnötige Datenerfassung zu unterlassen) wird mit der Forderung,

Datenerhebung nicht länger zu vermeiden, sogar geradewegs für obsolet erklärt.“ (Gehring, 2022, 24)<sup>20</sup> Die scharfe Kritik am Aufweichen des gesetzlichen Datenschutzes zugunsten einer persönlich geregelten ökonomischen Verwertung eigener Daten als Ressource teilt auch die Datenschutzkonferenz: „Datensouveränität‘ verstanden als eigentumsähnliche Verwertungshoheit kann daher nur zusätzlich zum Recht auf informationelle Selbstbestimmung greifen, dieses jedoch keinesfalls ersetzen.“ (Datenschutzkonferenz, 2017)

Es gibt eine dauerhafte Spannung zwischen zwei Polen, die beide nicht beim Individuum und dessen Datenschutz-Verantwortung, sondern innerhalb der staatlichen Verantwortung liegen. Auf der einen Seite steht die geforderte staatliche Sicherstellung der digitalen Souveränität seiner Bürger:innen, die Schutz und Handlungsfreiheit innerhalb des rechtlichen Rahmens gewährleistet. Auf der anderen Seite besteht – auch in demokratischen Gesellschaften – die Gefahr steigender staatlicher Eingriffs- und Kontrollmacht. In beiden Bereichen entsprechen die aktuell praktizierten Formen des Umgangs mit (mindestens) personenbezogenen Daten und die Verfügbarkeit von Infrastrukturen nicht dem über die vergangenen Jahrhunderte erkämpften Verhältnis von Rechten und Pflichten zwischen Individuen und Staaten. In dieser komplexen Problematik ist die Wissenschaft auf mehreren Ebenen betroffen und gefährdet.

Der Wissenschaftsrat hat sich in seinen *Empfehlungen zur Souveränität und Sicherheit der Wissenschaft im digitalen Raum* (2023) ebenfalls zum Thema digitale Souveränität geäußert:

Die meisten öffentlichen Stellungnahmen zur digitalen Souveränität weisen eine wirtschafts- und technologiepolitische, teils auch geo- bzw. sicherheitspolitische Stoßrichtung auf. So werden sowohl Defizite hinsichtlich der selbstbestimmten Steuerung und Ausgestaltung von digitalen Prozessen aller Wirtschafts- und Gesellschaftsbereiche benannt als auch Abhängigkeiten von wenigen privatwirtschaftlichen Unternehmen angeprangert, die noch dazu aus

<sup>20</sup> Vgl. Lisker (2023) für eine detaillierte Darstellung, weshalb die Strategie, die Verantwortung auf ‚digital mündige‘ Bürger:innen zu verschieben schon deshalb nicht gelingen kann, weil bestimmte Formen der Erhebung und Verarbeitung von Daten gar nicht durch Benutzer:innen erreichbar sind.

dem außereuropäischen Ausland stammen. Die vielfach artikulierten Forderungen nach (mehr) digitaler Souveränität sind somit Teil breiterer Bestrebungen, die neben einem (Rück-)Gewinn an individueller, gesellschaftlicher und staatlicher Gestaltungshoheit und -freiheit vor allem darauf zielen, die Handlungsfähigkeit sowie die Innovations- und Wettbewerbsfähigkeit der heimischen Wirtschaft zu stärken. Wesentliche Bezugs- und Rechtfertigungsrahmen bilden geopolitische Handels- und Ressourcenkonflikte sowie der globale Wettbewerb um Technologieführerschaft. Digitale Souveränität überschneidet sich, ist aber nicht identisch mit ‚Technologiesouveränität‘, einem Begriff, der vor allem geopolitische Verhältnisse adressiert und mit dem vielfach eine Schwerpunktsetzung auf so genannte Schlüsseltechnologien, wie Künstliche Intelligenz oder Quantencomputer, einhergeht. (Wissenschaftsrat, 2023, 10)

Auch wenn die Formulierungen recht allgemein ausfallen, ist der Bezug zur Wissenschaft gegeben – Handlungs- und Gestaltungsfreiheit im wissenschaftlichen Raum zurückgewinnen und bestehende Abhängigkeiten (von übermächtigen Unternehmen) verringern, Innovations- und Wettbewerbsfähigkeit der heimischen Wissenschaft stärken. Wissenschaft agiert zwar global und kooperativ, gleichzeitig aber auch kompetitiv. Aus nationalen Interessen geht es vordergründig um kompetitive Vorteile des ‚eigenen‘ Wissenschaftsstandortes.

Seine Aufmerksamkeit richtet der Wissenschaftsrat vor allem auf die Ebene wissenschaftlicher Einrichtungen und Infrastrukturen. Er identifiziert vier Handlungsdimensionen – Kapazitäten und Fähigkeiten, Kooperation auf verschiedenen Ebenen, Auswahl- sowie Gestaltungsmöglichkeiten und schließlich die Sensibilisierung und Reflexionsfähigkeit. Er mahnt, die „digitale Selbstbefähigung“ von Wissenschaftseinrichtungen zu stärken, in deren Cybersicherheit „dringender Handlungsbedarf“ (ebd. 6f.) bestehe. Die Empfehlungen zielen darauf ab, künftiges Handeln so zu gestalten, dass (erneute) Abhängigkeiten von kommerziellen Anbietern vermieden werden. Das soll unter anderem durch Kooperation und Vernetzung im digitalen Raum geschehen, so dass „Synergie- und Skaleneffekte“ genutzt werden und durch den Ausbau öffentlich finanzierter Infrastrukturen und Plattformen (Beispiele sind NFDI auf nationaler und EOSC auf europäischer Ebene). (Ebd. 7) Mehrfach spricht der WR von „Mindeststandards für Pluralität und Offenheit“ und betont die Notwendig-

keit von Open-Source-Software. Gleichzeitig wird herausgestellt, dass es sich um Aufgaben handelt, welche die Wissenschaft nun dauerhaft selbst in die Hand nehmen muss, inklusive Pflege und Entwicklung: „Dies beinhaltet, verlässliche Strukturen und Finanzierungsmöglichkeiten für die nachhaltige Bereitstellung, Pflege und Entwicklung zu etablieren.“ (Ebd. 8) Zur Vermeidung eines erneuten Lock-In sollen künftige Angebote interoperabel sein und „möglichst nahtlos an bestehende Strukturen anknüpfen und nutzerfreundlich ausgestaltet sein. Ihre Bekanntheit gilt es zu erhöhen.“ (Ebd.)<sup>21</sup> Die Aufforderung lautet, „die Gestaltung des digitalen Raumes als Daueraufgabe von Wissenschaftseinrichtungen zu verankern“. (Ebd.) Auf organisatorischer Ebene heißt das auch, „Beschaffungs- und Vergabeprozesse sind so zu gestalten, dass ein Anbieterwechsel möglich bleibt sowie definierte Mindeststandards für Pluralität und Offenheit eingehalten werden.“ (Wissenschaftsrat, 2023, 7)

Die Abkehr von der Nutzung kommerzieller Anbieter und eine klare Empfehlung zur Einrichtung, Nutzung und dauerhafter Bereitstellung nicht aufkaufbarer, transparenter ‘Scholar-led’-Angebote ist ein wichtiger Schritt für die Zukunft der Digitalisierung der Wissenschaft. Die Verringerung bestehender Abhängigkeiten kann in diesem Konzept durch Wechsel zu einer eigenen digitalen Infrastruktur und eigenen, kooperativen Formaten erfolgen. Es ist bezeichnend und ein wichtiges Symbol, dass der WR nicht über die Nutzung kommerzieller Angebote spricht, obwohl Konzerne wie RELX, Clarivate oder Digital Science ihr Portfolio technisch führender Produkte immer weiter ausbauen. Auch der Rat für Informationsinfrastrukturen fordert die Stärkung öffentlich getragener Infrastrukturen: „Bedarfsgerechte Daten- und Informationsinfrastrukturen, die durch öffentliche Mittel (mit)finanziert werden, sollten grundsätzlich fest in den Forschungsalltag von Hochschulen und außeruniversitären Forschungseinrichtungen integriert sein.“ (RfII, 2023, 40)

---

<sup>21</sup> Die Erhöhung der Bekanntheit, genauso wie die Aufforderung zu mehr Kooperation und Vernetzung reflektieren die Tatsache, dass zahlreiche digitale Angebote (z. B. Repositorien) auf der Ebene einzelner Einrichtungen sowie europaweit (Open Research Europe) zwar entwickelt und bereitgestellt, aber kaum in einem nennenswerten Umfang genutzt werden. Das hat verschiedene Ursachen, neben fehlender Vernetzung speziell in Bezug auf digitale Formate auch die Tatsache, dass sie nicht (in einem renommeerelevanten Sinne) etabliert sind.

Die Empfehlungen des WR weisen in Richtung einer souveräne(re)n Wissenschaft hin und lassen sich systematisch in die grundsätzliche Zielsetzung digitaler Souveränität (wie sie oben diskutiert wird) einordnen. Die Abhängigkeitsproblematiken von übermächtigen kommerziellen Playern im aktuellen Wissenschaftssystem sind gravierend. Ob diese ausschließlich durch gesetzliche Regulierungen erfolgen können, kann hier nicht beantwortet werden. In Bezug auf die kritischen Perspektiven auf den Diskurs digitaler Souveränität bzw. Datensouveränität lässt sich jedoch festhalten, dass es für den Bereich der Wissenschaft nicht sinnvoll ist, von einer Bedrohung durch den Ausbau staatlicher Kontrolle oder einer unregulierten Ökonomisierung von privaten Daten sowie Forschungsdaten auszugehen. Gerade Wissenschaft stellt einen Bereich dar, in dem staatliche Unterstützung und Schutz erwartet werden. Daher ist die Rede von digitaler Souveränität im Wissenschaftsbereich sinnvoll und nötig, insofern sie die unterschiedlichen Bereiche und Handlungsebenen einschließt. Für eine genauere Analyse braucht es weitere Differenzierung, zu der ich in den nächsten Abschnitten komme.

## 1.3 Privatheit

Eine andere Perspektive bietet der hauptsächlich philosophische Diskurs von Privatheit (privacy) (Nissenbaum, 2010; Rössler, 2005; Finn et al., 2013; Lane et al., 2014; Post, 2001; Solove, 2009). Dieser, zumindest zeitweise durch die US-amerikanische Gesetzgebung geprägte Diskurs hatte seine Anfänge am Ende des 19. Jahrhunderts mit dem breiteren Einsatz der Fotografie (Warren & Brandeis, 1890). Die Ansätze unterscheiden sich darin, ob sie Privatheit als einen eigenständigen Wert sehen (Rössler, 2005) oder denken, dass diese sich auf andere Werte – z. B. Sicherheit, Autonomie, Demokratie, Würde – reduzieren lässt (Thomson, 1975). Weitere Ansätze, sogenannte cluster accounts (Solove, 2006; Allen, 2011; Nissenbaum, 2010), argumentieren dafür, dass Privatheitsansprüche sich aus verschiedenen Quellen speisen, jedoch keinen einzelnen gemeinsamen Kern haben. Die Debatte eint, dass bereits die Bedeutung des Konzepts ‚Privatheit‘ umstritten und unklar ist und sich mit der Entwicklung der Informationstechnologien wandelt:

The combination of increasing power of new technology and the declining clarity and agreement on privacy give rise to problems concerning law, policy and ethics. [...] Information being a cornerstone of access to individuals, the privacy debate has co-evolved with – and in response to – the development of information technology. It is therefore difficult to conceive of the notions of privacy and discussions about data protection as separate from the way computers, the Internet, mobile computing and the many applications of these basic technologies have evolved. (van den Hoven *et al.*, 2020)

Aktuelle Fragestellungen betreffen auch Gruppen (group privacy) (Floridi, 2017; Mittelstadt, 2017) und die soziale Dimension von Privatheit. Der systematische Fokus liegt aber vorwiegend auf dem Individuum und dessen Beziehung zu Privatheit. Wichtige Themen sind die Beziehung zwischen dem Individuum und dem Staat, die Rolle von Privatheit für soziale Beziehungen, Überwachung bei der Arbeit oder Umgang mit social media. Um den Wissenschaftssektor und seine Beziehung zu Konzernen wie RELX herauszugreifen, scheint eine andere Herangehensweise fruchtbarer, die nicht ursprünglich aus dem Privatheitsdiskurs stammt.

Deirdre Mulligan, Colin Koopman und Nick Doty argumentieren 2016 in “Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy”, dass es sich bei Privatheit um einen “essentially contested concept” handelt, also einen wesentlich umstrittenen Begriff. Das Konzept wesentlich umstrittener Begriffe von Walter Bryce Gallie aus dem Jahr 1956 fand Eingang in verschiedene philosophische Diskurse (Ästhetik, Religions-, politische und Geschichtsphilosophie). Gallie führt den Begriff ein, um eine spezifische Begriffsklasse zu beschreiben, der bisherige Analysen und Klassifizierungen nicht gerecht wurden. Diese Begriffe gehören zu menschlichen Handlungsfeldern, die einen hohen gesellschaftlichen Stellenwert haben, Errungenschaften oder Ziele darstellen. Gallies Beispiele sind ‚Kunst(werk)‘, ‚Demokratie‘ oder ‚ein christliches Leben‘ (später werden auch ‚(soziale) Gerechtigkeit‘ und ‚Macht‘ genannt). Solchen abstrakten Begriffen wird unumstritten eine positive Bewertung zugesprochen, umstritten ist ihre weitere Erfassung. Für Gallie besteht das Spezifische an dieser Begriffsklasse nicht darin, dass sie mehrdeutig oder verworren ist, sondern dass diese wesentlich umstrittene Begriffe *notwendigerweise* dauerhafte und unauflösbare Dispute provozieren.

Ein wesentlich umstrittener Begriff soll laut Gallie folgende notwendige Eigenschaften aufweisen:<sup>22</sup> Erstens, der Begriff muss positiv wertend sein, damit wird eine Errungenschaft ausgedrückt. Die Errungenschaft muss, zweitens, eine komplexe interne Struktur besitzen, der die positive Wertung im Ganzen zu- bzw. ausgesprochen wird, so dass, drittens, Erklärungen der Wertschätzung einen Bezug auf die entsprechenden Teile oder Eigenschaften des Gegenstandes nehmen müssen. Viertens, die zugesprochene Errungenschaft muss auf eine Weise *offen* sein, die ihre ggf. beträchtliche und nicht vorhersehbare oder festlegbare Anpassung an veränderte Umstände erlaubt.

Zu den vier notwendigen Eigenschaften nennt Gallie drei weitere, mit deren Hilfe eine Unterscheidung von “radically confused” (Gallie, 1956, 180) Begriffen möglich wird: Die fünfte Bedingung verlangt, dass die streitenden Parteien zwar jeweils für sich die Deutungshoheit und die Funktionen beanspruchen, die der Begriff für theoretische Schulen, Bewegungen oder Traditionen erfüllt, einander aber im gewissen Sinne als peers anerkennen. Sie sehen einander als peers im Streit um den fraglichen Begriff, dessen Anwendungskriterien oder Exemplifizierungen. Ferner, sechstens, muss der Begriff aus einem autoritativen Beispiel („original exemplar“) abgeleitet werden, das von allen Parteien als solches anerkannt wird. Die letzte Bedingung drückt aus, dass der fortdauernde Disput über den Begriff, seine Kriterien und Exemplifizierungen eine fortschreitende Weiterentwicklung der durch den Begriff ausgedrückten Errungenschaft bewirkt.<sup>23</sup>

---

22 “(I) it must be appraisive in the sense that it signifies or accredits some kind of valued achievement. (II) This achievement must be of an internally complex character, for all that its worth is attributed to it as a whole.” (Gallie, 1956, 171f.) (III) „Any explanation of its worth must therefore include reference to the respective contributions of its various parts or features [...]. (Gallie, 1956, 172) „(IV) The accredited achievement must be of a kind that admits of considerable modification in the light of changing circumstances; and such modification cannot be prescribed or predicted in advance. For convenience I shall call the concept of any such achievement ‘open’ in character“ (ebd.).

23 “(V) that each party recognizes the fact that its own use of it is contested by those of other parties, and that each party must have at least some appreciation of the different criteria in the light of which the other parties claim to be applying the concept in question.” (Gallie, 1956, 172) “These are (VI) the derivation of any such concept from an original exemplar whose authority is acknowledged by all the contestant users of the concept, and (VII) the probability or plausibility, in appropriate senses of these terms, of the claim that the continuous competition for

Was ist gewonnen mit dieser recht künstlich wirkenden Klassifizierung? Der negative Mehrwert besteht in der Abgrenzung normativer abstrakter Begriffe mit hoher gesellschaftlicher Relevanz von Geschmacksfragen und Vulgärrelativismus<sup>24</sup>; beide Kategorien verfehlen die Rolle und Funktion solcher Begriffe. Gallie sieht auch einen entscheidenden Unterschied zu konkurrierenden wissenschaftlichen Theorien, weil es bei wesentlich umstrittenen Begriffen kein “definite or judicial knock-out” durch Formalisierung oder Rückgriff auf anerkannte Methoden gibt. Das liegt daran, dass die Wissenschaft durch Methoden und Terminologien viel stärker formalisiert ist als die Alltagssprache, in der die wesentlich umstrittenen Begriffe operieren.

Der positive Mehrwert des Konzepts zeigt sich einerseits in der Möglichkeit, den Begriffsgehalt nach verschiedenen Kriterien zu kontextualisieren und andererseits darin, die zentrale Rolle des kontinuierlichen Diskurses für die Bedeutungskonstitution herauszustellen. Der argumentative Bezug zu konkurrierenden Ansätzen bietet die Möglichkeit, künftige Veränderungen in den Begriff zu integrieren.

### 1.3.1 Wesentlich umstrittene Privatheit

Mulligan, Koopman und Doty betrachten unterschiedliche Ansätze innerhalb des Privatheitsdiskurses, mit der elusiven Natur des Privatheitsbegriffs umzugehen, z. B. bei Nissenbaum (2010) durch formale (statt inhaltliche) Kriterien von Privatheit, die Information, Kontext und Sozialgruppe hinsichtlich jeweiliger Adäquatheitsstandards relativieren. Auch der Ansatz von Solove (2009), Privatheit mit Wittgensteins Familienähnlichkeiten-Begriff zu erfassen, biete die Möglichkeit, flexibel und offen für verschiedene inhaltliche Bestimmungen zu sein, ohne sich auf notwendige Kriterien festlegen zu müssen (Mulligan *et al.*, 2016, 4).

---

acknowledgement as between the contestant users of the concept, enables the original exemplar's achievement to be sustained and/or developed in optimum fashion.” (Gallie, 1956, 180)

<sup>24</sup> Darunter verstehe ich undifferenzierte Vorwürfe von Beliebigkeit, kognitiver Anarchie oder Aufgabe von Geltungskriterien, der häufig mit (ebd., 179) ‚dem‘ Relativismus assoziiert werden. Vgl. z.B. Hollis & Lukes (1982); Stich (1998).

Aber diese Ansätze schaffen es laut Mulligan *et al.* nicht, einer Eigenschaft gerecht zu werden, die sie für zentral am Privatheitsbegriff und die praktische Arbeit damit halten: “[N]either engages head-on with privacy’s essential contestability. Thus, their efforts advance privacy work, but in ways that risk diminishing the generative power of contestability.” (Ebd.)

Die Autor:innen argumentieren, dass Schutzstrategien des Wertes “informational self-determination” durch gemeinsam entwickelte bzw. geteilte Praktiken der letzten Dekaden des 20. Jahrhunderts im Lichte der heutigen Umstände nicht mehr greifen (ebd., 2). Mehr noch, “informational self-determination” sei nicht mehr geeignet, um den Wert selbst zu erfassen, weil die Beschreibung angesichts früherer Umstände zwar passend, aber nicht erschöpfend war. Aus heutiger Perspektive, mit technischen Möglichkeiten von data analytics und machine learning, sei der Fokus auf individuelle Kontrolle nicht mehr hinreichend, um neu entstandene privatheitsbezogene Gefahren zu erfassen:

Informational self-determination can hardly be considered a sufficient objective, nor individual control a sufficient mechanism, for protecting privacy in the face of this new class of technologies and attendant threats. Privacy harms may arise for which individual control offers no protection or remedy, for example, when actions are taken based on group classifications [...], or new and unexpected insights are inferred from data that individuals have intentionally disclosed [...], or an individual’s sensitive personal information is derived through analysing data revealed by others in their social network [...], their behaviour on social media [...] or by cross-referencing sets of ‘de-identified’ data in which they are included [...]. Such privacy harms stem from information, yet they are not addressed by strategies centred on individual control over information. (Mulligan *et al.*, 2016, 2)

Das Problem besteht für die Autor:innen eben nicht nur in der Entwicklung einer adäquaten Gesetzgebung oder Steigerung individueller Kompetenz (data literacy), sondern auch in der sich ändernden Konzeption dessen, worin der zu schützende Wert besteht. Hier muss sozusagen an beiden Enden gearbeitet werden – die Gefahr durch neue Technologien und Handlungsmöglichkeiten zeigt uns neue Facetten bzw. Eigenschaften des Wertes, die erst durch diese Gefahr denk- und greifbar und zum Ge-

genstand unserer Überlegungen werden. In anderen Worten: Privatheit entwickelt sich relativ zu Veränderungen der gesellschaftlichen, sozialen und technischen Praxis. Dabei, so Mulligan et al., müssen neue Konzepte erstritten werden, sie entstehen diskursiv:

As we bump up against the limits of informational self-determination, we must reflect on what gets lost when we reify privacy as just one thing—one principle, one formalization, one method of protection. We must engage with the whole tangled, ambiguous and essentially contested terrain of privacy. (Mulligan *et al.*, 2016, 2)

Ein neues, subtileres Verständnis von Privatheit entsteht in der Konfrontation mit potenziellen Bedrohungen und ist dadurch adäquater in Bezug auf die Dimensionen, die Mulligan et al. identifiziert haben, nämlich ‚Prinzip‘, ‚Formalisierung‘ und ‚Schutzmethode‘:

[S]cholars and practitioners have advanced alternative conceptions of privacy that address the new risks of data science, machine learning and other technological innovations. Scholars have argued for privacy concepts and approaches that unburden the individual by attending to the contextual norms of spheres of social life [...], address risks posed by ancillary data [...] and attend to the emerging semantics of machine learning [...]. Professionals have developed new concepts, such as ‘trust’ and ‘meeting expectations’, and contrast concepts such as ‘creepiness’, in an effort to address the shifting privacy concerns of customers and citizens [...]. Privacy regulators have expanded their work as well, branching out into information ethics in response to the challenges posed by big data [...]. (Mulligan *et al.*, 2016, 2)

Hier liegt der Vorteil, Privatheit im Sinne eines wesentlich umstrittenen Begriffs aufzufassen, denn durch den offenen und diskursiven Vollzugscharakter dieser Norm werden wechselnde Umstände eingefangen. Durch die Bindung des Begriffs an einen Auslegungszwang entsteht ein Mechanismus, Privatheit nicht nur praktisch anwendbar zu machen, sondern auch an künftige technologische Entwicklungen anzupassen: “Recognizing privacy’s essential contestedness is key to securing its generativity for generations to come. Privacy’s ability to respond to problems crea-

ted by technology and societal change is essential to its ongoing relevance.” (Mulligan *et al.*, 2016, 4)

Als wesentlich umstrittener Begriff ist Privatheit kein vages Konzept, sondern ein multidimensionales. Innerhalb von “privacy’s multi-faceted and open-textured meaning” unterscheiden die Autor:innen so heterogene Dimensionen der Privatheit wie ‚Prinzip‘, ‚Rechtfertigung‘, ‚schädigender Handlung‘ und ‚Schutzmethode‘ (ebd., 3) und machen gleichzeitig plausibel, warum sich Privatheit nicht in Rückgriff auf Datenschutz-Regelungen, Privatsphäre oder Daten erschöpft. In ihrer Argumentation weisen Mulligan *et al.* (2016, 5-9) zumindest Gallies notwendige Bedingungen überzeugend für den Privatheitsbegriff nach und ziehen praktische Schlüsse für die angewandte Forschung im Bereich data science und Mensch-Maschine-Interaktion. Der Anspruch der Autor:innen, eine Übersetzung zwischen Privatheitsdiskurs und technischen Vorgaben zu befördern, wird zudem durch eine kleinteilige Analyse der Dimensionen von Privatheit umgesetzt (ebd. 10-15).

Mulligan *et al.* identifizieren folgende inhärent kontroverse Metadimensionen von Privatheit, die sie jeweils unterteilen und mit Beschreibungen, Fragen und Beispielen exemplifizieren (Abbildung 1): Theorie (Objekt, Rechtfertigung, Kontrast-/Gegenbegriff, Exemplar), Schutz (Ziel, Subjekt), Schaden (schädigende Handlung<sup>25</sup>, Täter, Gefährdungsquelle), Bereitstellung/provision (Mechanismus, Bereitsteller), Umfang/scope (soziale Eingrenzung, zeitlicher und quantitativer Umfang).

In der Anwendung auf Wissenschaft kann nun mit weiteren Kategorien (Schaden, Bereitstellung, Umfang) differenzierter gefragt werden, in Bezug auf was wer in welchem Umfang geschützt werden soll (s. Tabelle 1).

---

<sup>25</sup> Hier verweisen die Autor:innen auf Soloves sechszehnteilige Taxonomie privatheitsschädigender Handlungen als Beispielliste.

Table 1. Dimensions of contests over privacy.

| privacy dimension               | description                                                                                                                                                    | interrogation                            | examples                                                                      |
|---------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|-------------------------------------------------------------------------------|
| <b>dimensions of theory</b>     |                                                                                                                                                                |                                          |                                                                               |
| object                          | that which privacy provides to those protected, i.e. <i>privacy provides protected agents with X</i>                                                           | 'What's privacy for?'                    | dignity; control over personal information                                    |
| justification                   | the motivation and basis for providing privacy, i.e. <i>privacy is justified because of X</i>                                                                  | 'Why should this be private?'            | individual liberty; social welfare                                            |
| contrast concept                | that which contrasts to privacy, i.e. <i>that which is private is mutually exclusive with that which is X</i>                                                  | 'What's not private?'                    | public; open; transparent                                                     |
| exemplar                        | the archetypal threat to this concept of privacy, i.e. <i>privacy is violated by X</i>                                                                         | 'What's an example?'                     | identity theft; intrusive surveillance; gossiping neighbours                  |
| <b>dimensions of protection</b> |                                                                                                                                                                |                                          |                                                                               |
| target                          | that which privacy protects, i.e. <i>privacy protects things of type X</i>                                                                                     | 'What's privacy about? Privacy of what?' | personal information; body or likeness; private space                         |
| subject                         | actor(s) or entity(ies) protected by privacy, i.e. <i>privacy protects agent X</i>                                                                             | 'Whose privacy is at stake?'             | myself, my child; social groups (e.g. teens); roles (e.g. students)           |
| <b>dimensions of harm</b>       |                                                                                                                                                                |                                          |                                                                               |
| action                          | the act or behaviour that initiates or constitutes a privacy harm, i.e. <i>staring at him while he was dressing in the locker room violated his privacy</i>    | 'What act violated privacy?'             | Solove's four meta-harms (collection, processing, dissemination and invasion) |
| offender                        | actor(s) violating privacy, i.e. <i>privacy violated by agent X</i>                                                                                            | 'Who violated privacy?'                  | government; business entity; peeping tom                                      |
| from-whom                       | actor(s) against-whom privacy is a protection, i.e. <i>privacy provides protection against agent X</i>                                                         | 'Who is privacy protecting against?'     | everyone; Government; 'friends of friends'                                    |
| <b>dimensions of provision</b>  |                                                                                                                                                                |                                          |                                                                               |
| mechanism                       | that which instrumentally secures privacy, i.e. <i>the lock on her door protected her privacy</i>                                                              | 'How is privacy provided?'               | legal regulations; technical design; social norms                             |
| provider                        | actor(s) charged with securing privacy, i.e. <i>the telecommunications provider was responsible for technically securing the privacy of her communications</i> | 'Who is supposed to provide privacy?'    | Government; business entity; technology                                       |
| <b>dimensions of scope</b>      |                                                                                                                                                                |                                          |                                                                               |
| social boundaries               | that wherein privacy applies, i.e. <i>privacy applies in domain, situation, field, or site X</i>                                                               | 'Where is privacy found?'                | hospital or university; nation-state or globally                              |
| temporal scale                  | the time span at which privacy applies, i.e. <i>privacy applies for a span of X time</i>                                                                       | 'How long is privacy required?'          | permanent; fixed expiration; variable expiration                              |
| quantitative scope              | extent of application of privacy, i.e. <i>privacy should be applied with a scope of X</i>                                                                      | 'How widely does privacy apply?'         | universally as strict rule; casuistically as per-case                         |

ABBILDUNG 1: Umstrittene Dimensionen von Privatheit bei Mulligan, Koopman und Doty.

TABELLE 1: Analysetool in Anwendung auf Wissenschaft.

| Privacy Dimension        | Interrogation                      | Wissenschaftsaspekt                                                                                                 |
|--------------------------|------------------------------------|---------------------------------------------------------------------------------------------------------------------|
| theory: object           | What's privacy for?                | Selbstbestimmung, Arbeit als Wissenschaftler:in, persönliche Sicherheit                                             |
| theory: justification    | Why should this be private?        | Persönlichkeitsrecht, Wissenschaftsfreiheit                                                                         |
| theory: contrast concept | What's not private?                | kompromittiert/manipuliert, gefährdet/vulnerabel, abhängig                                                          |
| theory: exemplar         | What's an example?                 | Beeinflussung (Verhalten, Karriere), Gefährdung der Forschungsfreiheit                                              |
| protection: target       | Privacy of what?                   | Persönliche Daten, Forschungsergebnisse, wissenschaftliche Profile, Kontakte                                        |
| protection: subject      | Whose privacy is at stake?         | Wissenschaftler:innen, Studierende, Angehörige von Wissenschaftseinrichtungen                                       |
| harm: action             | What act violated privacy?         | Tracking, Aggregation von Informationen, Profilerstellung, Intransparenz: Vorenthalten/Weitergabe von Informationen |
| harm: offender           | Who violated privacy?              | Verlage als Interaktionspartner, kommerzielle Drittparteien, data broker                                            |
| harm: from-whom          | Who is privacy protecting against? | Fremde Interessen, externer Einfluss, Verlust von Unabhängigkeit                                                    |

|                              |                                     |                                                                                                                         |
|------------------------------|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| provision:<br>mechanism      | How is privacy provided?            | DSGVO (Einhaltung nicht nachprüfbar), Bibliotheken, Schutz- & Fürsorgepflicht (Beamte), technische Kontrolle            |
| provision:<br>provider       | Who is supposed to provide privacy? | Gesetzgebung (national, EU), Wissenschaftseinrichtungen, vertrauenswürdige Technologien                                 |
| scope:<br>social boundaries  | Where is privacy found?             | Bibliothek, Universität, Forschungseinrichtung, Förderer, Software für wissenschaftliche Arbeit/Wissenschaftsmanagement |
| scope:<br>temporal scale     | How long is privacy required?       | permanent, mindestens während der Zugehörigkeit zur Institution                                                         |
| scope:<br>quantitative scope | How widely does privacy apply?      | (universell/einzelfallabhängig)                                                                                         |

# 1.4 Daten(schutz)recht und die Anwendung auf Wissenschaft

Für eine Darstellung der Rechtsgrundlage für den Bereich Datenschutz und Datensouveränität in Anwendung auf die Wissenschaft beziehe ich mich insbesondere auf Diskussionen auf der 15. Göttinger Urheber- und Datenschutzrechtstagung 2023 sowie Selinger (2024). Selinger nennt folgende Bereiche als einschlägig: Verfassungsrecht, Datenschutzrecht, Vertragsrecht, Urheberrecht, Kartellrecht und Prozessrecht sowie Rechtsbehelfe. Nachfolgend gehe ich auf die ersten drei Bereiche in ihrer Relevanz für die generelle Problematik der Datenerhebung und -verarbeitung durch Großverlage ein.

## 1.4.1 Verfassungsrecht

Direkt betroffen sind auf der obersten Ebene (Grundgesetz und EU-Recht) die Wissenschaftsfreiheit sowie das Persönlichkeitsrecht auf informationelle Selbstbestimmung:

- Art. 5 Abs. 3 GG Kunst und Wissenschaft, Forschung und Lehre sind frei. [...]
- Art. 13 EU-GrCharta – Freiheit der Kunst und der Wissenschaft
- Das Recht auf informationelle Selbstbestimmung: Art. 2 Abs. 1 GG i.V.m. Art. 1 Abs. 1 GG

Im Rahmen des Volkszählungsurteils BVerfGE 65, 1<sup>26</sup> 1983, deutlich vor heutigen technischen Möglichkeiten von big data, wurde durch höchst-richterliche Grundsatzentscheidung geurteilt, dass es keine belanglosen Daten über Personen gibt. Ungehindertes Datensammeln und technische Möglichkeiten der Verknüpfung ermöglichen eine besondere Bedrohung für das Persönlichkeitsrecht des Einzelnen (individuelle Handlungsfreiheit) und zugleich das gesamte demokratische Gemeinwesen. Diese Rechtsprechung betrifft direkt die gängigen Praktiken (speichern, erheben, sammeln, verknüpfen personenbezogener Daten) der Wissenschaftsverlage.

---

**26** „1. Unter den Bedingungen der modernen Datenverarbeitung wird der Schutz des Einzelnen gegen unbegrenzte Erhebung, Speicherung, Verwendung und Weitergabe seiner persönlichen Daten von dem allgemeinen Persönlichkeitsrecht des Art. 2 Abs. 1 GG in Verbindung mit Art. 1 Abs. 1 GG umfaßt. Das Grundrecht gewährleistet insoweit die Befugnis des Einzelnen, grundsätzlich selbst über die Preisgabe und Verwendung seiner persönlichen Daten zu bestimmen.

2. Einschränkungen dieses Rechts auf ‚informationelle Selbstbestimmung‘ sind nur im überwiegenden Allgemeininteresse zulässig. Sie bedürfen einer verfassungsgemäßen gesetzlichen Grundlage, die dem rechtsstaatlichen Gebot der Normenklarheit entsprechen muß. Bei seinen Regelungen hat der Gesetzgeber ferner den Grundsatz der Verhältnismäßigkeit zu beachten. Auch hat er organisatorische und verfahrensrechtliche Vorkehrungen zu treffen, welche der Gefahr einer Verletzung des Persönlichkeitsrechts entgegenwirken.“

## 1.4.2 Datenschutzrecht

Hier sind vor allem drei Gesetze einschlägig, auf nationaler wie EU-Ebene:

- Art. 5 ePrivacy-RL (EU-Richtlinie, die sogenannte Cookie-Richtlinie)<sup>27</sup> – Vertraulichkeit der Kommunikation

(3) Die Mitgliedstaaten stellen sicher, dass die Speicherung von Informationen oder der Zugriff auf Informationen, die bereits im Endgerät eines Teilnehmers oder Nutzers gespeichert sind, nur gestattet ist, wenn der betreffende Teilnehmer oder Nutzer [...] seine Einwilligung gegeben hat. [...]

- § 25 TTDSG Telekommunikation-Telemedien-Datenschutz-Gesetz – Schutz der Privatsphäre bei Endeinrichtungen

Die Speicherung von Informationen in der Endeinrichtung des Endnutzers oder der Zugriff auf Informationen, die bereits in der Endeinrichtung gespeichert sind, sind nur zulässig, wenn der Endnutzer auf der Grundlage von klaren und umfassenden Informationen eingewilligt hat.

Obige Gesetze sind spezifischer und enger in der Anwendung als die DSGVO und beziehen sich auf Fälle, wenn Informationen in den Endgeräten der Verbraucher:innen gespeichert werden: EU-Richtlinie e-Privacy, umgesetzt in Deutschland als sogenannte Cookie-Richtlinie.

---

<sup>27</sup> RICHTLINIE 2002/58/EG DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 12. Juli 2002 über die Verarbeitung personenbezogener Daten und den Schutz der Privatsphäre in der elektronischen Kommunikation (Datenschutzrichtlinie für elektronische Kommunikation).

- Art. 6 DSGVO – Rechtmäßigkeit der Verarbeitung

(1) Die Verarbeitung ist nur rechtmäßig, wenn mindestens eine der nachstehenden Bedingungen erfüllt ist:

a) Die betroffene Person hat ihre Einwilligung zu der Verarbeitung der sie betreffenden personenbezogenen Daten für einen oder mehrere bestimmte Zwecke gegeben;

b) die Verarbeitung ist für die Erfüllung eines Vertrags, dessen Vertragspartei die betroffene Person ist, oder zur Durchführung vorvertraglicher Maßnahmen erforderlich, die auf Anfrage der betroffenen Person erfolgen;

c) die Verarbeitung ist zur Erfüllung einer rechtlichen Verpflichtung erforderlich, der der Verantwortliche unterliegt; [...]

f) die Verarbeitung ist zur Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten erforderlich, sofern nicht die Interessen oder Grundrechte und Grundfreiheiten der betroffenen Person, die den Schutz personenbezogener Daten erfordern, überwiegen, [...]

Da die Datenerhebung durch Verlage einen erheblichen Grundrechtseingriff darstellt, wird durch den Art. 6 eine anspruchsvolle Einwilligung verlangt. Strittig ist in Anwendung auf den Spezialbereich Wissenschaft und Wissenschaftseinrichtungen, unter welchen Voraussetzungen betroffene Personen, also Wissenschaftler:innen – aber auch Studierende und andere Bibliotheksnutzende – überhaupt in einer Situation sind, wirksam einzuwilligen. Selinger diskutiert folgende Überlegungen, warum es sich nicht um die verlangte Art von Einwilligung handeln kann.

- Art. 7 DSGVO: Bedingungen für die Einwilligung (4)

ad a) Bei der Beurteilung, ob die Einwilligung freiwillig erteilt wurde, [...]. Wirksamkeit der Einwilligung setzt Freiwilligkeit und Informiertheit des Datensubjekts sowie Transparenz der Verarbeitung voraus

Die hohen Anforderungen an die Freiwilligkeit seien nicht erfüllt, wenn der Datenverarbeitende aus der Position der wirtschaftlichen Stärke agiert oder rechtliche und praktische Abhängigkeitsverhältnisse bestehen. Das sei gleich auf mehreren Ebenen der Fall: Im Verhältnis zwischen Verlagen und Wissenschaftseinrichtungen sowie zwischen Verlagen und Einzelpersonen. Zudem ist die Einwilligung nur wirksam, wenn das Datensubjekt hinreichend informiert ist über die Datenverarbeitung. Dies setze umfangreiche Informationspflichten seitens der Verlage voraus und werde in der Regel nicht eingehalten. Ähnlich sehe es mit der Forderung von Transparenz der Verarbeitung aus, diese werde in der Praxis ebenfalls nicht erfüllt.

ad b) Verarbeitung der Daten sei notwendig zur Erfüllung des Vertrags (Verlagsseitig)

Selinger bezweifelt, dass es sich bei der umfangreichen Datenerhebung um Informationen handelt, die für Verlage notwendig für die Vertragserfüllung sein sollen. Jedenfalls dann, wenn es sich um Aufgaben eines Wissenschaftsverlags handelt, gehen die allermeisten erhobenen Daten über das Benötigte hinaus und können nicht gerechtfertigt werden.

ad c) rechtliche Verpflichtung der Vertragspartei (Verlag) zur Erhebung der Daten

Selinger argumentiert, dass es unklar sei, aus welchem Grund Verlage dazu verpflichtet sein müssten. (Eine gesetzliche Regelung, dass Verlage dazu verpflichtet sind, liegt nicht vor.)

ad f) Wahrung des berechtigten Interesses des Verlags zur Verwertung von Daten

Hier stellt sich die Frage, weshalb es sich beim Interesse des Verlags an Daten um ein berechtigtes Interesse handelt. Selinger fragt zurecht, warum ein ökonomisches Interesse den Schutz personenbezogener Daten überwiegen soll und verweist auf umfangreiche Rechtssprechung in ähnlichen Fragen, u. a. des EU-Gerichtshof im Fall von social media.

Selinger schlussfolgert, dass es sich bei der aktuellen Praxis der Datenerhebung und -verarbeitung durch große Verlage häufig um Datenschutzverstöße handelt und diese Praxis damit entsprechend rechtswidrig ist. Sie stelle meistens auch eine Verletzung des Rechts auf informationelle Selbstbestimmung und damit eine Grundgesetzverletzung sowie eine Verletzung der Wissenschaftsfreiheit dar. In Bezug auf die Wissenschaftsfreiheit zählt Selinger folgende Eingriffe:

- Individuelles Freiheitsrecht (als Wissenschaftler:in ungehindert zu arbeiten)
- Wirkung zwischen Privaten (auch in Privatrechtverhältnissen müssen sich Private, also sowohl natürliche Personen wie auch Unternehmen, an grundrechtliche Werte halten; anderenfalls begehen sie Rechtsverstöße)
- Objektiver Schutzauftrag\* (die Wissenschaftsfreiheit verpflichtet den Staat, die Wissenschaft von Behinderungen durch Dritte zu schützen, worunter auch privatrechtliche Behinderungen der Wissenschaftsfreiheit durch kommerzielle Unternehmen fallen; Universitäten seien als staatliche Einrichtungen verpflichtet, vorbeugende Vorkehrungen gegen die Behinderung der Wissenschaftsfreiheit zu treffen)
- chilling effects (darunter versteht Selinger eine negative verhaltenslenkende Wirkung auf das Forschungsverhalten durch Wissen darum, dass personenbezogene Daten gesammelt werden, z. B. Abstand nehmen von der Forschung an politisch heiklen Themen)

## Erläuterungen zum objektiven Schutzauftrag\*

### BEAMTENRECHTLICHE FÜRSORGEPFLICHTEN

- (Landesbeamte) § 45 BeamtStG

Der Dienstherr hat im Rahmen des Dienst- und Treueverhältnisses für das Wohl der Beamtinnen und Beamten und ihrer Familien, auch für die Zeit nach Beendigung des Beamtenverhältnisses, zu sorgen. Er schützt die Beamtinnen und Beamten bei ihrer amtlichen Tätigkeit und in ihrer Stellung.

Selinger macht darauf aufmerksam, dass verbeamtete Angehörige von Wissenschaftseinrichtungen Anspruch auf Schutz durch ihre Einrichtungen haben. Darunter fallen auch Persönlichkeitsrechtsverletzungen. Mögliche Konsequenzen für Universitäten könnten Klagen auf Schadensersatz und Unterlassungsansprüche sein.

### GEMEINSAME VERANTWORTLICHKEIT FÜR DIE DATENVERARBEITUNG

- Art. 25 DSGVO

Legen zwei oder mehr Verantwortliche gemeinsam die Zwecke der und die Mittel zur Verarbeitung fest, so sind sie gemeinsam Verantwortliche. Sie legen in einer Vereinbarung in transparenter Form fest, wer von ihnen welche Verpflichtung gemäß dieser Verordnung erfüllt, [...]

Dieser Bereich ist in der praktischen Diskussion besonders aktuell, weil mindestens einer der drei deutschlandweit laufenden DEAL-Verträge der aktuellen juristischen Einschätzung nach die bisher genannten rechtswidrigen Praktiken aufweist. In der Frage, ob gemeinsame Verantwortung zwischen den Wissenschaftseinrichtungen und dem Konzern besteht und welche Konsequenzen das für die möglichen juristischen Folgen hat, ist der im Juli 2024 veröffentlichte Abschlussbericht des AWBI ausschlaggebend, der sich auch auf juristische Gutachten bezieht, welche die DFG in Auftrag gegeben hatte. Dem zufolge besteht eine gemeinsame Verantwortung<sup>28</sup>

i.S.v. Art. 26 DS-GVO der Wissenschaftseinrichtung (welche die Verträge abgeschlossen hat und insoweit durch ihre Bibliotheken handelt) bereits dann, wenn diese (1) die Datenverarbeitung durch die Verlage ermöglicht und (2) selbst von der Datenverarbeitung (auch nur mittelbar) profitiert. Die Bibliotheken vermitteln für ihre Einrichtungen den Zugang der Forschenden zu Plattformen und Datenbanken, wodurch erst deren personenbezogene Daten durch die Verlage genutzt werden können. Die Wissenschaftseinrichtungen erhalten von den Verlagen Statistiken, die teilweise nur auf Grundlage einer Identifikation der Nutzenden erhoben werden können, also das Datentracking voraussetzen. Zwar werden nach allen Landesdatenschutzgesetzen gegen Behörden (und damit Einrichtungen in rein staatlicher Hand) keine Bußgelder verhängt, jedoch müssen privatrechtlich organisierte Einrichtungen (etwa in Stiftungsform) Bußgelder nach Art. 83 DS-GVO fürchten. Gegen alle Einrichtungen können Betroffene Schadensersatzansprüche (Art. 82 DS-GVO) und Unterlassungsansprüche (Art. 79 DS-GVO) geltend machen. Darüber hinaus kommen auch Schutzpflichten der Einrichtungen zugunsten ihrer Beschäftigten in Bezug auf deren personenbezogene Daten in Betracht (§ 241 Abs. 2 BGB bzw. § 45 S. 2 BeamStG), also der Einrichtungen gegenüber den Forscherinnen und Forschern. Und dies

<sup>28</sup> „[N]ach einer Entscheidung des EuGH aus dem Jahr 2019 (EuGH Urt. v. 05.06.2018, Rs. C-210/16 – Facebook-Fan-Pages Rn. 36 f.; EuGH Urt. v. 10.07.2018, Rs. C-25/17 Zeugen Jehovas Rn. 73; EuGH, Urt. v. 29.07.2019, Az. C-40/17 – Fashion ID Rn. 80; EuGH Urt. v. 05.12.2023, Rs. C-683/21 – Nacionalinis visuomenes sveikatos centras Rn. 43 f.; EuGH Urt. v. 07.03.2024, Rs. C-604/22 – IAB Europe Rn. 77) [...]“ (Altschaffel et al., 2024, 28)

gilt dennotwendig vor allem, wenn der Zugriff auf für die Forschung unverzichtbare Quellen ausschließlich über elektronische Datenbanken ermöglicht wird, die über das unbedingt Erforderliche hinausgehend personenbezogene Daten verarbeiten. (Altschaffel *et al.*, 2024, 28)

Auch wenn also die öffentlichen Einrichtungen keine Bußgelder befürchten müssen und die Folgen von möglichen Schadensersatzansprüchen speziell im Elsevier-DEAL-Vertrag nicht tragen müssten<sup>29</sup>, besteht die Möglichkeit von Unterlassungsansprüchen. Vor allem letztere könnten dazu führen, den Zugang zu den Plattformen zu unterbinden, denn die Einrichtungen stehen

gegenüber Behörden und Betroffenen natürlich weiterhin in der Verantwortung, gegebenenfalls auch in der Öffentlichkeit und auch vor Gericht. Eine vorübergehende oder gar endgültige Untersagung der betreffenden Verarbeitungsvorgänge durch die Aufsichtsbehörde (Art. 58 Abs. 2 lit. f DS-GVO), die faktisch eine Sperrung des Zugangs zur Datenbank bedeutet, schwebt freilich weiter als nicht ausgleichsfähiges Damoklesschwert über den Verträgen. (Altschaffel *et al.*, 2024, 32f.)

---

<sup>29</sup> Die Formulierung in 7.6. Data Protection lautet: “Elsevier shall defend, indemnify and hold harmless DEAL Operating Entity and each Participating Institution for any violation or alleged violation of Art. 26 GDPR.” (Elsevier & MPDL Services GGmbH, 2023) Das bestätigt auch der Bericht des AWBI: „Die Einrichtungen müssen im Ergebnis finanziell die Folgen etwaiger Datenschutzverstöße nicht tragen. [...] mögliche finanzielle Folgen [der Abschaltung des Zugangs] würden im Nachhinein durch die Verlage ausgeglichen werden.“ (Altschaffel *et al.*, 2024, 32)





2.

## **Lock-In I Zeitschriftenkrise**

Um die Entwicklung innerhalb des wissenschaftlichen Publikationssektors zu analysieren, möchte ich die Begriffe ‚Pfadabhängigkeit‘, ‚Lock-In‘ sowie ‚Netzwerkeffekt‘ („network effect“) einführen.

Als “network externality” bezeichnen Katz & Shapiro (1985) einen spezifischen Mehrwert eines Produktes für einzelne Nutzer:innen, der sich aus der Anzahl der Nutzer:innen ergibt: “There are many products for which the utility that a user derives from consumption of the good increases with the number of other agents consuming the good.” (Katz & Shapiro, 1985, 424) Der Mehrwert steigt mit steigender Nutzendenzahl und sinkt, wenn nur wenige Nutzer:innen dieses Produkt verwenden. Die gängige Verwendung folgen Shapiro & Varian (2010) darin, “network effects” als *positive network externality* zu bezeichnen, also als einen positiven Effekt des Wertzuwachses eines Produktes durch den Zuwachs an Nutzenden.<sup>30</sup> Das heißt, dass es *unabhängig* von der Qualität des Produktes einen Vorteil für Nutzer:innen hat, wenn möglichst viele andere das Produkt ebenfalls verwenden.

Als indirekte Auswirkung des network effects nennen Katz und Shapiro einen möglichen Einfluss auf individuelle Kaufentscheidungen zugunsten eines verbreiteteren Produkts. Die Annahme von Nutzer:innen wäre dabei, dass die Menge möglicher Anwendungen des Produkts mit der Anzahl verkaufter Exemplare zusammenhängt. Es wären laut Katz und Shapiro (1985, 424) demnach mehr Anwendungen des verbreiteteren Produkts verfügbar, als für vergleichbare, aber weniger verbreitete Produkte.

In den späten 1980ern war die Nutzung von Faxgeräten und PCs ein häufiges Beispiel. Aus heutiger Sicht sind social media (Meta) oder die Nutzung nicht interoperabler Software (Microsoft und Apple) naheliegende Belege sowie die Produkte von Elsevier oder Clarivate. Allein die Größe des Nutzendenkreises von Facebook oder Twitter/X und die Vorstellung der Nachteile aus dem Verzicht auf diese Medien reichen aus, damit auch unzufriedene Nutzer:innen das Produkt weiterhin verwenden.<sup>31</sup>

---

<sup>30</sup> Vgl. Xiong et al. (2018) für eine kurze Darstellung der anderen Verwendungen.

<sup>31</sup> Siehe hier auch Zuboff (2019).

Pfadabhängigkeit ist ein häufig genutzter Erklärungsansatz für organisationale Trägheit. Sie beschreibt eine „spezifische Dynamik organisatorischer Prozesse“ (Schreyögg, 2014) und hebt die Rolle kontingenter Ereignisse und Prozesse gegenüber einem ökonomischen Rationalitätsideal hervor. Kontingente Entscheidungen führen kumuliert zu Festlegungen für zukünftige Entscheidungen und begünstigen die Fortführung einmal gewählter Strukturen, Abläufe, Produkte etc. selbst dann, wenn diese den verfügbaren Alternativen unterlegen sind: “A path-dependent sequence of economic changes is one of which important influences upon the eventual outcome can be exerted by temporally remote events, including happenings dominated by chance elements rather than systematic forces.” (David, 1985, 332)

Der Begriff wird vor allem mit zwei Arbeiten verbunden, “Clio and the Economics of QWERTY” (David, 1985) und “Competing Technologies, Increasing Returns, and Lock-In by Historical Events” (Arthur, 1989). Paul A. David argumentiert anhand der Einführungsgeschichte der Tastaturanordnung („QWERTY“, weil dies die ersten fünf Buchstaben oben links auf der englischsprachigen Tastatur sind) bei Schreibmaschinen im 19. Jahrhundert. Die Anordnung der Buchstaben diene mechanischen Zwecken. Um die interne Verkeilung des Mechanismus zu vermeiden, wurden solche Buchstaben nebeneinander platziert, die lexikalisch selten nebeneinander vorkommen. Diese rein funktionale Entscheidung war der Schreibmaschinenkonstruktion geschuldet, hat sich jedoch in den digitalen Bereich durchgesetzt. David argumentiert, dass sich die QWERTY-Tastatur sogar gegen die überlegenere Tastatur von August Dvorak durchsetzte, die 1932 patentiert wurde und Vorteile in der Erlernbarkeit und Tippgeschwindigkeit versprach.

Viele kleine, für sich unbedeutende Schritte in einem “contingent, non-reversible dynamical [process], including a wide array of biological and social processes that can properly be described as ‘evolutionary’” (David, 2000, 2) führten dazu, immer mehr QWERTY-Tastaturen zu verkaufen sowie immer mehr Schreibkräfte an QWERTY-Tastaturen und nicht an Dvorak-Tastaturen auszubilden. Diese Dynamiken verstärkten sich gegenseitig. David nennt drei Eigenschaften, die QWERTY zum dominierenden Design machten und die Grundbestandteile einer „QWERTY-nomics“ bilden: “technical interrelatedness, economies of scale, and quasi-irreversibility of investment”. (David, 1985, 334)

Der Erklärungsansatz ist nicht unumstritten. Liebowitz & Margolis (1995) kritisieren sowohl Aspekte der These von Lock-In und Pfadabhängigkeit als auch die Standardbeispiele und argumentieren für eine stärkere Regulierungsfähigkeit des Marktes. Ich werde hier nicht weiter auf diese Kritiken eingehen. Zum einen ist die Überzeugungskraft der markt-libertären Position gering. Zum anderen handelt es sich bei der Pfadabhängigkeit im Wissenschaftssektor (in seiner Zusammenarbeit mit hier besprochenen kommerziellen Anbietern) aus ökonomischer Sicht gerade nicht um einen echten Markt.<sup>32</sup>

Erreicht die Pfadabhängigkeit ein bestimmtes Maß, spricht man von einem Lock-In, einer „verriegelte[n] Lösung“ (Schreyögg, 2014, 4), einem negativen Zustand, in dem ein Pfadwechsel erschwert bzw. verunmöglicht wird/erscheint. Die Verfestigungstendenz bestehender Strukturen kommt in unterschiedlichster Art vor, sowohl in Form „einer bestimmten strategischen Ausrichtung, eines etablierten Geschäftsmodells oder einer spezifischen Kooperationsform“ (Schreyögg, 2014, 2). Diese Tendenz kann bei technologischen wie rechtlichen Lösungen exemplifiziert werden. Dadurch erhalten einmal getroffene Entscheidungen eine überproportionale Rolle für die künftige Entwicklung:

[Es] bildet sich in einem Unternehmen meist unbemerkt und ungewollt ein Pfad heraus, das heißt, die Organisationsmitglieder replizieren unter bestimmten Umständen in ihrem Handeln zunehmend ein und dasselbe Muster, bis schließlich dieser Pfad trotz veränderter Bedingungen nicht mehr oder nur noch mit sehr großen Anstrengungen verlassen werden kann. (Schreyögg, 2014, 2)

Wichtig ist, dass der Zustand der Verriegelung trotz möglicher Nachteile wie Ineffizienz aufrechterhalten wird, weil die einmal gewählte Option nicht mehr so einfach geändert werden kann: „Im Extremfall wird angenommen, dass sich ein zunächst noch offener Prozess schließlich zu einem deterministischen Verlauf verfestigt, an den die beteiligten Ent-

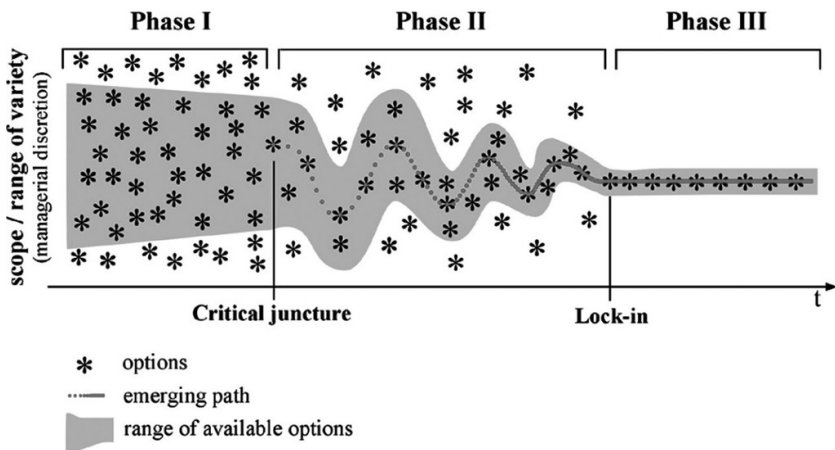
<sup>32</sup> Vgl. dazu Aspesi (2022); SPARC (2019); “Market forces alone are not sufficient to deliver widespread access to scientific information. There are clear indications that the subscription market is not functioning effectively, due to nonsubstitutability, excessive concentration, lack of transparency and perverse incentives.” Johnson et al. (2017, 5)

scheidungsträger gewissermaßen gefesselt sind.“ (Schreyögg, 2014, 4) Die „hartnäckigen Beharrungstendenzen“ gegenüber Änderungsbestrebungen werden durch „*selbstverstärkende Effekte*“ erklärt, „die eine Eigendynamik entfalten, die ab einem bestimmten Punkt kaum mehr zu stoppen ist.“ (Schreyögg, 2014, 2-3, Herv. im Original)

Ab einem bestimmten Zeitpunkt schließt sich der Pfadbildungsprozess, es entsteht die Situation des Lock-in. Von da an gibt es entweder keine Alternativen mehr (wie im Falle QWERTY) oder die Handlungsmuster sind so verriegelt, dass effektivere Alternativen zwar verfügbar sind, aber nicht mehr ergriffen werden können. Es liegt dann vollkommene Pfadabhängigkeit vor, sie wirkt in dem Sinne deterministisch, als sie andere Handlungsmuster ausschließt. (Schreyögg, 2014, 9)

Diese Entwicklung stellt Schreyögg im Drei-Phasen-Modell (Schreyögg & Sydow, 2011) dar (Abbildung 2). Die erste Phase ist durch eine Fülle alternativer Optionen (Sterne) gekennzeichnet, von denen eine verfügbar erscheinende (grauer Bereich) ausgewählt wird. Bereits in Phase I gibt es einen kritischen Moment („critical junction“), ab dem die weitere/wiederholte Nutzung der Option selbstverstärkende Effekte zeigt und sich ein Pfad ausbildet (Phase II). Dieser Pfad wird zunehmend fester, es zeigen sich „positive Rückkopplungseffekte“ (Schreyögg, 2014, 6), während in der zweiten Phase die Menge der Optionen zu einem bestimmten Zeitpunkt auf eine schrumpft. Das ist der Lock-In-Moment. Ab da befindet sich die Organisation in Phase III, in dem die vormalige *Option* nicht mehr optional, sondern zwingend scheint.

Bemerkenswert an diesem sozialen Phänomen ist, dass Entscheidungsprozesse und einmal etablierte Praktiken “tend to continue to reproduce this and only this particular outcome”, auch wenn sich äußere Bedingungen ändern oder effizientere Lösungen verfügbar werden (Schreyögg & Sydow, 2011, 325). Wie sich dies im Wissenschaftskontext zeigt, bespreche ich im Folgenden.



**ABBILDUNG 2:** Drei-Phasen-Modell der "organisational path dependence" (Schreyögg & Sydow, 2011, 323).

## 2.1 Entwicklung der Pfadabhängigkeit im wissenschaftlichen Publizieren

Die zentralen Entwicklungen im wissenschaftlichen Publizieren der letzten Jahrzehnte, die zum Lock-In im Publikationsbereich führten, werden unter dem Begriff ‚Zeitschriftenkrise‘ in ihrer Entstehung in die 1990er Jahre verortet. Die zugrundeliegenden Probleme sind eine Verflechtung aus mindestens vier Elementen: erstens, der Kommerzialisierung der Wissenschaft; zweitens, der Konzentration auf dem Verlagsfeld, die es wenigen Verlagen ermöglichte, viele zentrale Zeitschriften zu bündeln und unverhältnismäßige Preissteigerungen zu verlangen; drittens, dem Status wissenschaftlicher Literatur, der sie – in ökonomischer Terminologie – zu einem nicht-ersetzbaren Gut macht und viertens, der Digitalisierung und Plattformisierung der kommerziellen Angebote. Die heutige Situation beschreibt Jeff Pooley (2024) so: *“Commercial scholarly publishing*

is a mechanism to transfer value from scholars, taxpayers, universities to the world's most profitable companies.”

Wissenschaftliche Kommunikation durch Zeitschriften geht knapp 360 Jahre zurück, auf die Entstehung der ersten Zeitschriften aus wissenschaftlichen Fachgesellschaften in England und Frankreich 1665.<sup>33</sup> Individuelle Korrespondenz wurde durch das Zeitschriftenformat im Laufe der Zeit institutionalisiert und für Fachkreise verbreitet, die so am wissenschaftlichen Diskurs teilnehmen und Forschungsergebnisse schnell mitteilen konnten. Mit wachsender Zahl der Zeitschriften wuchs auch ihre Bedeutung, sie trugen zur Professionalisierung und Spezialisierung der Forschung bei, vor allem in naturwissenschaftlichen und medizinischen Disziplinen (Larivière *et al.*, 2015). Auch wenn schon in der Viktorianischen Epoche wirtschaftliche Akteur:innen Zeitschriften publizierten, wurden die meisten Zeitschriften von Fachgesellschaften herausgegeben (Larivière *et al.*, 2015). Zu einem einträglichen Geschäft wurden wissenschaftliche Zeitschriften erst nach 1945, in beträchtlichem Maße dank Robert Maxwell. Dieser kommerzialisierte die politische Förderung der wissenschaftlichen Kommunikation nach dem zweiten Weltkrieg, stieß die Gründung zahlreicher Zeitschriften an und machte ein Modell attraktiv für Wissenschaftler:innen, bei dem sie ihre Forschungsergebnisse und Arbeitsleistung umsonst an die Verlage gaben und alle Rechte daran abtraten. In Folge dessen hatte sich ein System etabliert, das

die Kosten allein auf die Seite der Öffentlichkeit und die Gewinne allein auf die Seite der Verlage [verteilt]. Mit dem rasanten Wachstum der Universitäten entstand aus diesem – nach Einschätzung der Deutschen Bank ‚bizarren‘ (Deutsche Bank 2005; Klein 2019) – Geschäftsmodell ein Oligopol einiger Verlage vor allem im Sektor der wissenschaftlichen Zeitschriften mit drastischen Preissteigerungen für die Bibliotheken und geringen Kosten für die Verlage. (Altshaffel *et al.*, 2024, 23)

Da wissenschaftliche Literatur nicht-substituierbar ist und Wissenschaftler:innen sowohl auf bestimmte Literatur angewiesen sind als auch selbst

<sup>33</sup> Für eine ausführlichere Auseinandersetzung mit der Entstehungsgeschichte der modernen Wissenschaftskommunikation s. Price (1963); Biagioli (2002).

bestimmen, wo sie ihre eigenen Arbeiten publizieren, müssen wissenschaftliche Einrichtungen genau diese Literatur, also bestimmte Zeitschriften, bereitstellen. Deshalb ist es also nicht sinnvoll, von einem Markt in wirtschaftlichem Sinne zu sprechen, denn Bibliotheken (Nachfrage) können nicht unter verschiedenen Zeitschriften (Angebot) wählen und dadurch marktregulierende Mechanismen wie Konkurrenz nutzen.<sup>34</sup>

Bibliotheken müssen bestimmte Zeitschriften anbieten, auch wenn deren Preise steigen. Die üblichen Vergabekriterien des öffentlichen Dienstes, die einen Preisvergleich verlangen und unverhältnismäßige Verdienste aus öffentlichen Geldern ausschließen sollen, werden dadurch umgangen. Wissenschaftliche Literatur hat ein Alleinstellungsmerkmal, so dass für sie die sogenannte Single-Source-Regelung gilt, die es erlaubt, das benötigte Produkt auch zu sehr schlechten Bedingungen zu beziehen.<sup>35</sup> Im Gegensatz zu anderen Produkten, für die der Bedarf mit dem Erwerb gedeckt wäre, gilt für wissenschaftliche Zeitschriften, dass dieses Alleinstellungsmerkmal für *jede* benötigte Zeitschrift zutrifft. Damit waren und sind Bibliotheken in einer äußerst schlechten Verhandlungsposition gegenüber den Zeitschriftbesitzer:innen, den Großverlagen. Bibliotheken benötig(t)en das, was die Verlage zu immer schlechteren Subskriptions-

---

34 "Competition in the scholarly publishing market depends on two primary variables: barriers to entry and market concentration. Barriers to market entry do not arise from financial or legislative constraints, but from cultural inertia. Top-tier academic journals are non-substitutable goods for both authors and readers, and operate as mini-monopolies within a discipline or field. Career incentive structures that reward publishing in established journals with a high 'impact factor' reinforce the dominant position of large publishers. A cultural bias against open access publications in certain disciplinary and national contexts stifles growth among smaller OA publishers. Competition is further hindered by excessive market concentration, and a lack of transparency due to widespread use of non-disclosure clauses. While scholarly publishing is a global market with over 5,000 journal publishers, five commercial publishers accounted for more than 50% of all articles published in 2013." (Johnson et al., 2017, 4)

35 "The possibility to increase profits in such an extreme fashion lies in the peculiarity of the economics of scholarly publishing. Unlike usual suppliers, authors provide their goods without financial compensation and consumers (i.e. readers) are isolated from the purchase. Because purchase and use are not directly linked, price fluctuations do not influence demand. Academic libraries, contributing 68% to 75% of journal publishing revenues [31], are atypical buyers because their purchases are mainly controlled by budgets. Regardless of their information needs, they have to manage with less as prices increase. Due to the publisher's oligopoly, libraries are more or less helpless, for in scholarly publishing each product represents a unique value and cannot be replaced [...]." (Larivière et al., 2015, 10-12)

Bedingungen anboten, unbedingt und stießen mit ihren Budgets an Grenzen.<sup>36</sup>

Das obige Modell der Pfadabhängigkeit anwendend, kann man sagen, dass durch den Status wissenschaftlicher Literatur bereits die „critical junction“ erreicht wurde, so dass für jede benötigte Zeitschrift eine Pfadabhängigkeit besteht. Das Problem wurde präsenter, als es für Bibliotheken nicht mehr möglich wurde, die benötigten, immer teurer werdenden Zeitschriften einzeln zu bestellen. Das geschah durch die Konzentration der Verlagslandschaft.<sup>37</sup> Verlage brachten durch den Aufkauf konkurrierender Häuser die wichtigsten Zeitschriften in ihren Besitz. Aber so erwarben sie auch sehr viele weniger wichtige Titel, die sie auch vertreiben wollten. Am Beispiel der Profitentwicklung von Elsevier bzw. damals Reed-Elsevier und heute RELX, zwischen 1991 und 2013 zeigen Larivière *et al.* (2015, 10), wie die Konsolidierung der Publikationsindustrie die rasante Entwicklung der Gewinnmargen der großen Verlage bewirkte. Diese Entwicklung wurde nicht einmal durch die Wirtschaftskrise 2008 aufgehalten:

[B]etween 1991 and 1997, both the profits and the profit margin increased steadily for the company as a whole. While profits more

<sup>36</sup> “Individual research institutions are often non-specialised, meaning that they have to cater for the need of researchers operating in various disciplines – each having their own must-have journals. RPOs also have very different priorities and financial availability, meaning that it is often difficult to negotiate subscription deals collectively at the national level. Similarly, countries rarely coordinate negotiations with publishers at international level. Large publishers are global companies, and for them even a research-intensive country the size of the UK or Germany is only part of a much larger commercial market. Protecting the integrity and profitability of that market is the publishers’ legitimate aim, and driving a hard bargain with them is a tall order for dispersed buyers.” (Johnson *et al.*, 2017, 31)

<sup>37</sup> “The majority of academic journals are published by commercial publishers. Over the decades, some (not all) publishers have increased subscription fees and/or APCs at rates much higher than inflation, and some track and spy on research activities [...]. The Big Deals publishers, Elsevier, Springer Nature, Wiley, Taylor & Francis and American Chemistry Society (ACS), each publishes over 2000 journals [...] and together they occupy over 50% of the market share (Morais *et al.*, 2019; Stoy *et al.*, 2019). Together, their subscription costs exceed 75% of total expenditures on journal publications in Europe, with the median price per article range from €1,344 to €2,658 [...]. In 2021, the Association of Scientific, Technical and Medical Publishers reported that the estimated growth of new scholarly journals is 2–3% annually and the global market is expected to reach the value of \$28 billion by 2023 [...]” (Ma, 2023b, 4)

than doubled over that period—from 665M USD to 1,451M USD—profit margin also rose from 17% to 26%. [After a stagnation in 2008-2009] profits reach[ed] an all-time high of more than 2 billion USD in 2012 and 2013. The profit margin of the company's Scientific, Technical & Medical division is even higher [...]. Moreover, its profits increased by a factor of almost 6 throughout the period, and never dropped below 30% from one year to another. The profit margin of this division never decreased below 30% during the period observed, and steadily increased from 30.6% to 38.9% between 2006 and 2013. (Ebd.)<sup>38</sup>

Diese Profitsteigerung lag auch daran, dass Verlage den Verkauf von einzelnen Zeitschriften durch den Verkauf von Paketen ersetzten. Es konnten nur noch ganze Bündel von Zeitschriften („Big Deals“) bezogen werden – eine Entwicklung, die mit der Digitalisierung des Zeitschriftenwesens einherging und den Verlagen zwei sehr positive Umstellungen ermöglichte. Die erste bestand aus der neuen Lösung, um weniger attraktive Titel dauerhaft an Bibliotheken zu verkaufen, nämlich durch teilweise nichtoptionale und intransparente Paketlösungen, die den Bezug einzelner Zeitschriften

teurer als der Bezug des Gesamtpakets [machten], wobei die Art und Weise, wie die Angebote dafür errechnet wurden, intransparent blieb, und die Verlage auch keine Vergleiche der Preise erlaubten, da Non Disclosure Agreements in den Verträgen dies verhin-  
derten. Das gleiche Zeitschriftenpaket konnte also unterschiedlich

---

**38** Siehe auch “Overall, the market has significantly consolidated since 2000 — when the top 5 publishers held 39% of the market of articles to 2022 where they control 61% of it. Looking at larger sets of publishers makes the consolidation even more extreme, as the top 10 largest publishers went from 47% of the market in 2000 to 75% in 2023, and the top 20 largest publishers from 54% to controlling 83% of the corpus. [...] These data show two main waves of market consolidation. The first wave aligns with the rise of The Big Deal journal subscription package model, roughly 2000 to 2006. During this period, it became increasingly difficult to remain independent, as large journal packages began to swallow up more and more of each library's budget. Getting into a Big Deal package provided financial stability as independent journals became much easier for libraries to cancel than large packages of hundreds or thousands of journals. Over this period, the top 5 largest publisher increased their share of the market from 39% to 49%, the top 10 largest increased from 47% to 58%, and the top 20 from 54% to 65%.” (Crotty, 2023)

viel kosten, je nachdem, wie viel Geld eine Einrichtung zur Verfügung hatte. (Holzer, 2022, 167)

Diese Änderungen, die wissenschaftliche Einrichtungen an ihre finanziellen Grenzen, aber Verlagen signifikante Mehrgewinne und Netzwerkeffekte brachten und keine Alternativen zu wissenschaftlichen Zeitschriften als Format zeigten, können als signifikantes Kennzeichen des *Lock-In im Publizieren* klassifiziert werden. Bezeichnend ist, dass diese Entwicklung so unausgewogen nachteilig für die Seite der Wissenschaft ausfiel wie sie vorteilhaft für die großen Verlage war und immer noch ist. Dazu Holzer (2022, 165): „Das Geschäftsmodell der Big Deals resultiert bis heute darin, dass die wissenschaftlichen Bibliotheken anschaffen müssen, was sie nicht benötigen, und aufgrund von Budgetbindungen dann nicht anschaffen können, was sie brauchen.“

Manfred Boni fasst die Situation zusammen:

Der Publikationsmarkt ist durch verfestigte, asymmetrische Strukturen gekennzeichnet, die aus ökonomischen und sozialen Gründen den großen Wissenschaftsverlagen eine nahezu monopolartige Stellung verschaffen. Ihrer Preis- und Technologieführerschaft folgen mehr oder weniger ungebrochen auch Universitätsverlage und Fachgesellschaften. Der Marktasymmetrie haben die Bibliotheken trotz Konsortialbildung keine wirksame Nachfragemacht entgegen stellen können. Klein- und mittelständische Verlage sind mit ihrem Programmangebot existentiell bedroht. (Boni, 2010, 38f.)

Die zweite Änderung bzw. Einsparung brachte die Ablösung vom Printgeschäft. Viele Schritte und Ausgaben auf verlegerischer Seite entfielen oder wurden ausgelagert, wie die Drucksetzung von Manuskripten oder bedarfsdeckendes Drucken und Versenden von Zeitschriften. Während Druckexemplare nur an ein Individuum verkauft werden können, kann eine Online-Zeitschrift auf der Website der Zeitschrift ohne Mehrkosten für den Verlag gleichzeitig von einer unbegrenzten Menge an Individuen und Institutionen aufgerufen werden. Damit sinken die Kosten für zusätzliche Abonnements für die Verlage gegen Null und ihre Gewinnmöglichkeiten steigen in dramatischer Weise:

In a system where the marginal cost of goods reaches 0, their cost becomes arbitrary and depends merely on how badly they are needed, as well as by the purchasing power of those who need them. In addition, costs are strongly influenced by the power relations between the buyer and seller, i.e. publishers and academic libraries. In such a system, any price is good for the seller, as the additional unit sold is pure profit. (Larivière *et al.*, 2015, 12)

Einen Ausgang aus dieser Problematik sollte die Open-Access-Bewegung bieten. Mit den drei Erklärungen – Budapester 2002 (BOAI, 2000), Bethesda und Berliner Erklärung (2003) – wurde der Entschluss der Wissenschaftsgemeinschaft festgehalten, den Zugang zu wissenschaftlicher Literatur rechtlich und finanziell zu reformieren. Gleichzeitig sollte wissenschaftliches Publizieren an die neuen digitalen Möglichkeiten der schnellen und günstigen Zugänglichkeit angepasst werden. Vor allem sollte dadurch die standardmäßige Abtretung von Rechten an die Verlage abgeschafft werden, die sich seit Robert Maxwell etabliert hatte. Wissenschaftler:innen sollten durch Open Access die Hoheit über Verwertungen ihrer eigenen Arbeit erlangen. Bibliotheken sollten unabhängig von den Gebühren der Verlage werden für die Bereitstellung der bereits öffentlich bezahlten Arbeit. Ich werde zunächst auf weitere problematische Entwicklungen des wissenschaftlichen Publizierens eingehen und dann auf einen Aspekt der Open-Access-Transformation zu sprechen kommen, den DEAL-Vertrag mit Elsevier.

## 2.2 Bibliometrisierung der Wissenschaft

Die Bibliometrie ist eine vergleichsweise junge Disziplin, für die neben Arbeiten aus dem Beginn des 20. Jahrhunderts z. B. von Alfred Lotka (1880-1948) und Samuel C. Bradford (1878-1948) vor allem Eugene Garfield (1925-2017) als Begründer steht.<sup>39</sup> Garfields Arbeit an der Datenbank Science Citation Index (SCI), am von ihm gegründeten Institute for Scien-

---

<sup>39</sup> Siehe auch Godin (2006).

tific Information (ISI) in den 1950er Jahren in Philadelphia sowie die Bücher *Little Science, Big Science* von Derek de Solla Price (1963) und Genady Dobrov's *Nauka o Nauke* (1966) zählen zu den bekanntesten Begründungswerken der Disziplin. Alan Pritchard schlägt 1969 in einem zweiseitigen Paper den Begriff „Bibliometrics“ (Pritchard, 1969) für die Anwendung mathematischer und statistischer Methoden auf Kommunikationsmedien vor. Zu den wissenschaftlichen Kommunikationsmedien gehören heute weitere Publikationsformen wie Forschungsdaten, Software, Code, Patente und multimediale Beiträge.<sup>40</sup>

Der von Garfield entwickelte SCI sollte die Bezüge zwischen einzelnen Forschungselementen (vor allem Zeitschriftenartikel) analysieren. Dazu wurden im SCI Artikel und die in ihnen referierten Arbeiten bibliographisch erfasst, um dann Verbindungen und Gewichtungen zwischen Autor:innen, Zeitschriften und später auch vielen anderen Faktoren herzustellen. Für die Wirkungsmessung wurde neben dem SCI der Journal Citation Reports (JCR) eingesetzt. Der SCI wurde, zusammen mit der später entstandenen Datenbank Social Sciences Citation Index (SSCI) und JCR zu Web of Science (WoS). Hierzu noch einige Entwicklungen: 1992 kaufte Thomson Reuters WoS, 2016 ging es für den Preis von 3,55 Mrd. US-Dollar zu den Privat-Equity-Eigentümern Onex und Baring Asia (SPARC, 2019, 26-29) und heißt seitdem Clarivate Analytics.<sup>41</sup> Neben WoS ist Scopus<sup>42</sup> ähnlich einflussreich, ein Product von Elsevier, seit 2024 KI-gestützt<sup>43</sup> und mit eigenem Index CiteScore und Analysefunktionen ausgestattet. Gefolgt werden diese beiden einflussreichsten Anwendungen von Dimensions<sup>44</sup>

---

**40** Folgende Bereiche werden heute mengentheoretisch zueinander in Beziehung gesetzt: Informetrie als die Lehre der Quantifizierung bzw. Messung von Informationen enthält als Obermenge unter anderem die Szientometrie (Lehre vom Messen wissenschaftlicher Aktivitäten), in Bezug auf welche die Bibliometrie (Lehre vom Messen von Publikationen) wiederum als echte Teilmenge gesehen werden kann, da Publikationen nur einen Teil der wissenschaftlichen Aktivitäten darstellen.

**41** Clarivate bildet unter den führenden Konzernen eine Ausnahme, weil er selbst keine wissenschaftlichen Zeitschriften besitzt.

**42** <https://www.elsevier.com/products/scopus>. [zuletzt geprüft am 25.07.2024]

**43** <https://blog.scopus.com/posts/scopus-ai-release-may-2024>. [zuletzt geprüft am 25.07.2024]

**44** <https://www.dimensions.ai/>. [zuletzt geprüft am 25.07.2024]

der Holtzbrinck Publishing Group<sup>45</sup> (zu der auch Altmetric<sup>46</sup> und ein Mehrheitsanteil an Springer Nature gehören) und Google Scholar sowie von der nicht-kommerziellen Organisation CrossRef<sup>47</sup>.

Vor allem WoS, Scopus und GoogleScholar dominieren seit Langem den Markt für Literaturrecherche und sind damit zentral für den Forschungszyklus. Das ist jedoch nicht mehr ihre Hauptfunktion. Was die quantitative Beschäftigung mit Zitationen und Publikationen so relevant für diverse Akteur:innen im Wissenschaftssektor macht und bibliometrische Anwendungen zunehmend in den Vordergrund rückt, sind die unterschiedlichen Schlussfolgerungen, die aus der Analyse, berechtigt oder unberechtigt, gezogen werden. Solche Schlussfolgerungen betreffen unter anderem die Verbreitung und Relevanz von Publikationsorganen und die Rezeption durch und von Wissenschaftler:innen, die als Maßstab ihrer Qualität als Wissenschaftler:innen bzw. ihrer Leistungen dienen. Wie kommt das? Die Zitation der Publikationen einer Wissenschaftlerin steht für den Einfluss und die Wirkung ihrer Forschung und ist ein wichtiges Zeichen der Anerkennung in der Fachgemeinschaft. Zitation durch peers begründet Renommee und Aufmerksamkeit, die nicht-monetäre Währung der meritokratisch organisierten Wissenschaft. Wer viel zitiert wird, gilt als relevant und hat bessere Bedingungen für weitere wissenschaftliche Tätigkeit, z. B. bei der Stellen- und Fördermittelvergabe. Hier liegt der Grund für das Interesse an der Zitationsanalyse. Auch wenn es unterschiedliche Ansätze gibt, erfolgt die Evaluation meist nicht direkt, sondern vermittelt über Publikationsorgane. Deren Renommee wird stellvertretend auf die in diesen Publikationsorganen, vorwiegend Zeitschriften und Buchverlagen, publizierenden Wissenschaftler:innen übertragen: “In which journal a scientist publishes is considered one of the most crucial factors determining their career. The underlying common assumption is that only the best scientists manage to publish in a highly selective tier of the most prestigious journals.” (Brembs, 2018, 37)

---

45 <https://holtzbrinck.com/>. [zuletzt geprüft am 25.07.2024]

46 <https://www.altmetric.com/>. [zuletzt geprüft am 25.07.2024]

47 <https://www.crossref.org/>. [zuletzt geprüft am 25.07.2024]

Für die bibliometrische Analyse werden heute verschiedene Indikatoren verwendet, allen voran der ebenfalls am ISI ins Leben gerufene (Journal) Impact Factor (JIF oder IF). Ursprünglich diente der IF der bibliothekarischen Relevanzanalyse von Zeitschriften, um Erwerbs- und Bestandsüberlegungen zu stützen. Der IF wird aus der durchschnittlichen jährlichen Zitationshäufigkeit einer Zeitschrift in anderen Zeitschriften auf Grundlage der Daten der vergangenen zwei Jahre berechnet. Jedoch erfolgt die autoritative Berechnung des IF durch Clarivate, berücksichtigt nur die im WoS indexierten Zeitschriften und ist nicht transparent. Alle drei Faktoren sind jeweils hinreichend, um den IF als objektives Instrument der Qualitätsbeurteilung zu disqualifizieren. Trotz zahlreicher Kritiken<sup>48</sup> ist der IF in vielen Disziplinen, insbesondere der Medizin und den Naturwissenschaften, häufig der zentrale Indikator für die Evaluation von Wissenschaftler:innen. Daneben gibt es noch weitere, mehr oder weniger verbreitete Indikatoren für Personen und Organisationseinheiten – *einfache*, wie den h-Index (Hirsch 2005)<sup>49</sup>, absolute Publikations- und Zita-

---

**48** Hirschi (2018); Larivière & Sugimoto (2019); Wouters & Costas (2012); Lewandowski (2006). Fühles- Ubach & Tunger (2021, 6) explizieren das mit Verweis auf eine frühere Arbeit (Ball et al. 2009): „Der Impact Factor ist kein Maß für wissenschaftliche Wahrnehmung von Personen oder Institutionen, denn er lässt keine Aussage zu Einzelveröffentlichungen zu. Der Impact Factor ist die durchschnittliche Zitationsrate für Journals. Eine Einzelveröffentlichung, selbst in absoluten TOP-Journals wie Science oder Nature, kann am Ende unzitieren bleiben oder die Publikation ist das absolute TOP-Paper und ist um ein Vielfaches häufiger zitiert. Zitationsverteilungen sind extrem schiefe Verteilungen. Es braucht den Einsatz durchaus existierender, aber noch wenig bekannter Indikatoren, wie zum Beispiel Feld- oder Journalnormalisierungen, die weitaus belastbarere Aussagen zu Output und Impact ermöglichen als beispielsweise einfache Indikatoren wie der h-Index.“ Die DFG schreibt 2022: „Generell weisen Metriken hinsichtlich der Produktivitätsmessung von Wissenschaft viele Schwächen auf [...]. Auf die Schwächen hinsichtlich der Messung der Qualität von Artikeln hat bereits C. Hoeffel (Hoeffel, 1998) hingewiesen; Eugene Garfield, einer der Begründer des Journal Impact Factor, verweist bereits 1955 in seinen grundlegenden Gedanken darauf, dass Wissenschaftlerinnen und Wissenschaftler den Anspruch haben sollten, Literatur nicht nur auf der Basis bibliografischer Indizes wahrzunehmen, sondern durch organisiertes und umfassendes Lesen (Garfield, 1955). Vor diesem Hintergrund erscheint es als nicht zielführend, wenn Metriken über ihre Funktion als Orientierungshilfe hinaus bei der Einschätzung einer wissenschaftlichen Leistung dominieren oder sogar die alleinige Basis für eine solche Einschätzung darstellen.“ (DFG Deutsche Forschungsgemeinschaft | AG Publikationswesen, 2022, 35, Herv. YF)

**49** Der h-Index, Hirsch-Index, auch Hirsch-Faktor oder -Koeffizient, benannt nach Jorge Hirsch, misst die Zitierhäufigkeit einer Person (oder auch größere Einheiten mit gemeinsamen Merkmalen, Arbeitsgruppen, Institutionen etc.). Der h-Index gibt die höchste Anzahl derjenigen Publikationen  $x$  an, die mindestens  $x$  Mal zitiert wurden. Er ist damit Aussagekräftiger als allein die Publikationszahl und will sta-

tionszahlen und *relationale* wie den CPP (Zitationsraten im Vergleich zum internationalen Journal-Durchschnitt JCSm oder im Feldvergleich FCSm), den J-Factor (Zusammenhang zwischen Publikationstätigkeit von Institutionen, Zeitschriften und Einfluss), welche ungleich größere Datenmengen und komplexere Interpretationen und Gewichtungen erfordern, aber dafür auch aussagekräftigere Analysen ermöglichen.

Einige methodische, technische<sup>50</sup> und systematische Kritikpunkte an der Eignung von Metriken als Evaluationsgrundlage wissenschaftlicher Qualität fasst Manfred Boni zusammen:

Technisch wie methodisch lassen sich Repräsentativität und Validität von Zitatanalysen anzweifeln. Denn trotz aller verfeinerten Methoden bleiben die Ergebnisse statistische Artefakte, die als Indikatoren von Exzellenz gewertet werden. Zeitschriftenpublikationen haben in den Ingenieur-, Sozial- und Geisteswissenschaften nicht dieselbe hohe Bedeutung wie in den Naturwissenschaften und der Medizin. Zeitschriften werden im Citation Index unterschiedlich breit und tief erfasst, was nicht nur für die verschiedenen Fachgebiete gilt, sondern als US-bias vor allem auch europäische Zeitschriften in Deutsch, Französisch oder Spanisch benachteiligt. [...] Die im Web of Science grundsätzlich auf die Initialen verkürzten Namensansetzungen der Autoren und Ko-Autoren stellen Identifizierungsprobleme dar, ebensowenig sind die Institutionen eindeutig bestimmt, an denen die Autoren arbeiten oder zu früheren Zeitpunkten gearbeitet haben [...]. Schließlich ist die Zitierpraxis selbst, wer wen und was aus welchen Gründen zitiert, ein sozialer Prozess, der einfache Interpretationen verbietet [...]. (Boni, 2010, 305)

---

tistische Ausreißer einfangen. Umstritten ist der h-Index unter anderem durch seine leichte Manipulierbarkeit und die Eigenschaft, im Laufe der Zeit, also der wissenschaftlichen Tätigkeit der Person, anzusteigen. (Sidiropoulos et al. (2007); Batista et al. (2006); Gadd (2018)) Auch der h-Index sollte nicht als einziger Indikator verwendet werden.

**50** Siehe zum Beispiel Mongeon & Paul-Hus (2016) für die Abdeckungsschwierigkeiten von Disziplinen in WoS und Scopus.

## 2.2.1 Bibliometrie und Ranking

Die Disziplin Bibliometrie ist multidimensional und arbeitet mit verschiedenen Methoden und Indikatoren, die verglichen und interpretiert werden. Diese Komplexität geht verloren, wenn eindimensionale und verkürzte Anwendungen bibliometrischer Methoden zur Grundlage für Entscheidungen im Wissenschaftsmanagement werden. Trotz Kritik – ob allgemein (bei Holzer, 2022; Münch, 2014b, 2015; Retzlaff, 2022) oder spezifisch (bei Mongeon & Paul-Hus, 2016; Herb, 2018; Straumsheim, 2016) – nimmt der Trend zur Bibliometrisierung der Wissenschaft zu. Das liegt unter anderem an der Suggestion, durch Metriken ein Mittel der Komplexitätsreduktion für zwei zentrale Probleme zu geben: Erstens ein einfaches und objektives Verfahren im Umgang mit exponentiell steigenden Publikationsmengen<sup>51</sup> und zweitens eine Möglichkeit der Einordnung in ein globales Wettbewerbsgeschehen in der Wissenschaft. Nicht nur im kommerziellen Duktus der Datenanalyse, auch im bibliothekarischen Diskurs ist teilweise explizit von Bibliometrie als „neuem Geschäftsfeld“ (Ball & Tunger, 2006) für Bibliotheken<sup>52</sup> die Rede, mindestens aber als von einer festen neuen Dienstleistung<sup>53</sup>. Damit ist die Bibliometrie für hochschulisches Government ein – nach innen gerichtetes – Instrument für Leistungsmessung, Qualitätsentwicklung, für das eigene Controlling und den Strategieprozess. Zugleich ist damit eine – nach außen gerichtete – Orientierung im universitären Wettbewerb geboten (Jansen *et al.*, 2007; Janßen & Sondermann, 2016; May, 2014): „Eine strategische Orientierung dieser Art kann im Benchmark<sup>54</sup> mit anderen Einrichtungen dabei unterstützen, das Engagement dieser anderen Einrichtungen zu analysieren

---

**51** In den 1960er Jahren bezeichnete De Solla Price exponentielles Wachstum als ein “fundamental law of any analysis of science” Price (1963, 5). Er prognostizierte eine Verdoppelung der Publikationszahlen bzw. die Zahl der in Wissenschaft Tätigen alle zehn bis fünfzehn Jahre (ebd., 6, Differenzierung der Verdoppelungszahlen für verschiedene Wissenschaftsformate auf den Seiten 6-7).

**52** Z. B. das ZBIW-Seminar „Bibliometrie als Geschäftsfeld in Bibliotheken: Grundlagen von quantitativen Publikationsauswertungen“ im März 2023 an der TH Köln, zu dem auch ein Vertreter von Elsevier eingeladen wurde, um SciVal zu präsentieren. Alle Teilnehmenden der Veranstaltung erhielten danach Emails von Elsevier und zwar ohne Kenntnis über die Weitergabe ihrer Kontaktdaten.

**53** S. May (2014) für eine Übersicht der Bibliotheken, die bibliometrische Analysen durchführen bzw. in ihren Servicekatalog aufgenommen haben. 10 Jahre später haben deutlich mehr Bibliotheken bibliometrische Angebote für ihre Angehörigen.

und so die eigenen Aktivitäten im Wettbewerb entsprechend auszurichten und diesen gegebenenfalls Ressourcen bereitzustellen.“ (Fühles-Ubach & Tunger, 2021, 3)

Die Kehrseite dieses Instruments ist die Tendenz zu einer metrisierten Wissenschaft, an der es bereits einige Kritik gibt (Brembs et al., 2013; Chan, 2019; Janßen & Sondermann, 2016; Chen & Chan, 2021; Holzer, 2022; Torny, 2020). Obwohl auch zentrale Förderer zunehmend zu einem ‚verantwortungsvollen‘ Einsatz von Metriken anhalten<sup>55</sup>, hat diese Tendenz eine selbst-reproduzierende Wirkung, die zum *mehr oder weniger systematischen Nachteil der Wissenschaft* wird. Holzer bezeichnet es als einen „wesentliche[n] Fehlanreiz“, wenn wissenschaftsinterne

Gratifikationssysteme (Mittelverteilung, Karrierefortschritt, Förderung) nicht mehr in erster Linie auf den Inhalten, Ergebnissen und Aktivitäten von einzelnen Wissenschaftlern aufbauen, sondern vermeintlich objektivere, externalisierte Metriken und/oder das symbolische Kapital von renommierten Publikationsorten als ausschließliches Mittel nutzen, um – verkürzend – Entscheidungen über die Qualität und den Wert der publizierten Ergebnisse zu treffen. (Holzer, 2022, 178)

Boni beschreibt, welche Attraktivität von der Idee metrisierbarer Wissenschaft ausgeht, die in einer unübersichtlichen und extrem komplexen Wissenschaftswelt für das Wissenschaftsmanagement Klarheit, Struktur und Messbarkeit suggeriert. Gleichzeitig entwickeln sich daraus selbstverstärkende Mechanismen, die wissenschaftsinterne Prozesse dominieren:

<sup>54</sup> „Benchmarking“ steht für den Vergleich mit anderen in Bezug auf bestimmte Leistungen oder Aspekte.

<sup>55</sup> „Grundsätzlich gilt für Metriken genau wie für die Reputation des Publikationsortes: [...] Die Aussagekraft solcher sekundären Qualitätszuschreibungen für Bewertung und Vergleich wissenschaftlicher Leistung ist begrenzt und bei deren Anwendung sind bestimmte Rahmenbedingungen zu beachten. So sind sie nicht in allen und oft nur in eng begrenzten Fachgebieten akzeptiert und überhaupt vergleichbar. Bibliometrische Indikatoren wiederum sind leicht zu manipulieren, was insbesondere dann problematisch wird, wenn sie als Kennzahlen zur Grundlage von Zuweisungssystemen gemacht werden.“ (DFG Deutsche Forschungsgemeinschaft | AG Publikationswesen, 2022, 27) S. auch European Commission. Directorate General for Research and Innovation. et al. (2017).

[M]it standardisierten bibliometrischen Indikatoren lassen sich Personen, Fächer und Institute leicht vergleichen. [...] Mit ihnen wird die Forschungsleistung einzelner Lehrstühle, Institute, Fachbereiche und ganzer Universitäten kommensurabel. Sie stellen für die Hochschulverwaltungen und die Wissenschaftsadministration einen einfachen Indikator dar, an dem sich Ressourcenentscheidungen technokratisch orientieren können. Ein Kreislauf beginnt, in dem aus der Akkumulation symbolischen Kapitals durch Zitate die sozialen und ökonomischen Kapitals folgt. Eine selbst-verstärkende Wirkung führt nach der Veröffentlichung in A-Journals zu verbesserter Finanzausstattung dank zusätzlicher staatlicher Mittelzuweisungen, erhöhter Drittmittelförderung [...] und zusätzlicher Finanzmittel nicht-staatlicher Stellen. Der Zugang zu den A-Journals stellt also einen entscheidenden Wettbewerbsvorteil dar. Die Chancen weiterer Veröffentlichungen in A-Journals steigen [...] Renommee reproduziert sich wechselseitig. (Boni, 2010, 305f.)

Welcher Zusammenhang besteht zwischen Metrisierung und Ranking(s), bzw. was ist letzteres? Unter ‚Ranking‘ (vom englischen ‘to rank’ für ‚ein-stufen‘ oder ‚einordnen‘) wird der kompetitive Vergleich mit dem Zweck einer Rangordnung zwischen Einrichtungen, typischerweise Universitäten, verstanden. Besonders einflussreich für die Attraktivität von Universitäten und damit für die Zahl der Studierenden sind bestimmte etablierte Anbieter von Rankings wie Times Higher Education University Rankings (THE)<sup>56</sup> oder QS World University Rankings (QS)<sup>57</sup> oder Academic Ranking of World Universities, bekannt als Shanghai-Ranking<sup>58</sup>. Systematisch betrachtet, ist der Fokus auf ein gutes Abschneiden der eigenen Einrichtung in einem der relevanten Rankings die Fortführung der bereits durch Metrisierung skizzierten Problematiken auf höherer Ebene. Diese Probleme betreffen zum Teil die Inhalte (Zeitschriften in WoS oder Scopus) und die Messinstrumente (InCites, JCR; Plummetrics, SciVal, PURE) der involvierten Konzerne wie Clarivate und Elsevier/RELX und das konkrete Ranking als Ergebnis. Trotzdem sehen sich Universitäten gezwun-

56 [www.timeshighereducation.com](http://www.timeshighereducation.com). [zuletzt geprüft am 25.07.2024]

57 [www.topuniversities.com](http://www.topuniversities.com). [zuletzt geprüft am 25.07.2024]

58 <https://www.shanghairanking.com/>. [zuletzt geprüft am 25.07.2024]

gen, sich dem Ranking-System anzupassen, so Chen und Chan (2021).<sup>59</sup> Sie machen darauf aufmerksam, dass das Wettrennen um eine gute Platzierung in den Rankings die wissenschaftlichen Einrichtungen nicht nur vergessen lässt, dass sie in eine neue Pfadabhängigkeit von denjenigen geraten, die die Algorithmen produzieren. Zudem rücken wissenschaftliche Standards weg von inhaltlichen Kriterien und hin zu einem metrisierten System, in welchem wissenschaftliche Einrichtungen weder Kontrolle haben noch eine gute Wissenschaft produzieren:

Rankings are thus a driver for the adoption of particular metrics-emphasizing modalities of research as well as the particular products which facilitate such modalities. Yet the adoption of such modalities represents a ceding of control by universities not only to rankings – which are often the drivers of their decision-making – but also to the algorithms which power the analytics products. As a result, it is argued that the *university shifts away from an intrinsic valuation of research and teaching and is increasingly compelled to pay into the rankings-data analytics system*. Ultimately, the relationship between data analytics firms and rankings is mutually reinforcing, as well as arguably increasingly harmful to the public missions of the university. (Chen & Chan, 2021, 18f. Herv. YF)

University rankings, journal rankings and lists of highly cited researchers are created using metrics. [...] The trust in citations and citation-based metrics entails that the legitimacy of knowledge is held in the hands of commercial indexes largely consisting of English language journals in Western countries, with a strong focus on STEM. At the same time, metrics are also embedded in search algorithms of Google Scholar and other indexing services, which perpe-

---

<sup>59</sup> “Clarivate has a dominant position in international rankings in terms of the volume of data supplied even though it has no direct involvement in rankings. In contrast, Elsevier has had a sustained involvement in rankings since 2007 targeting partnerships with dominant rankings such as QS and THE as a way to present its products to universities as enabling their rankings performance. Given the increasing pressure faced by *universities to perform bibliometrically, universities face pressure to adopt analytics tools promoted by the same data analytics companies which supply rankings*.” (Chen & Chan, 2021, 18 Herv. YF) Siehe die grafische Darstellung der Beteiligung Elseviers an Ranking-Organen im Anhang, Abbildung 15.

tuates the importance of citations and citation-based metrics. (Ma, 2023b, 7)

Wie beim IF ist auch hier offensichtlich, dass Interessenkonflikte gegen die Objektivität der Ergebnisse solcher Rankings sprechen. Trotzdem haben Clarivate und Elsevier aufgrund ihrer dominanten Position in Bezug auf die Mittel der Produktion datenanalytischer Evaluationen einen Vorteil. Man kommt nicht an ihnen vorbei. Universitäten nehmen die Nachteile für die Wissenschaft in Kauf, weil Rankings zu einem Wettbewerb geworden sind, aus dem sie es sich nicht erlauben können auszusteigen:

University rankings, journal impact factors, performance-based funding for universities – these have all affected the culture of academic institutions for years and have progressively led to an erosion of control by academic institutions over their own destiny. Data analytics services have a market and are in demand – in spite of all the issues they pose – because academic life has become a race to secure funding. As a result, research assessment is becoming a business opportunity, and one that commercial vendors appear eager to control, regardless of their possible conflicts of interest. (SPARC, 2020, 29)

The Shanghai ranking has played a crucial role in the construction of universities as entrepreneurial actors [...] and in the de-powering of the scientific and academic communities, the disciplinary and professional associations and, ultimately, the researchers themselves. (Münch, 2014a, 35)

Die Bibliometrisierung der Wissenschaftsevaluation bedient das Bedürfnis nach Komplexitätsreduktion und Übersichtlichkeit von rasant steigenden Publikationsmengen (Hanson et al., 2023) sowie nach Vergleichbarkeit und klaren Kriterien für die Behauptung der Institutionen gegenüber globaler Konkurrenz. Lai Ma (2023b) nennt Wissenschaftsevaluation “the choke point” in der Diskrepanz zwischen wissenschaftlichem Qualitätsprinzipien und der aktuellen Abhängigkeit der Belohnungsstruktur von offenkundig ungeeigneten Indikatoren wie “Journal impact factor (JIF), CiteScore, h-index, field-weighted citation impact (FWCI), and source-normalised impact per paper (SNIP)”, auf deren Grundlage einerseits 40 Rankings und nationale Forschungsevaluationen erhoben und andererseits

Schließungen vorn Fachbereichen und Instituten entschieden werden (Ma, 2023b, 6). Diese Verflechtungen verstärkten die „Krisen der Digitalisierung“ (Holzer, 2022, 176), die dazu führen, dass Wissenschaftler:innen gedrängt sind, sich an Indikatoren (wie den Impact Factor oder Zeitschriftenrankings) zu orientieren, „welche nicht zu diesem Zweck geschaffen worden, aber zu einer Währung im Karrieresystem der Wissenschaft geworden sind.“ (Holzer, 2022, 176)

Diesen mentalen Lock-In, unter dem vor allem junge Wissenschaftler:innen leiden, die sich für eine Zukunft in der Wissenschaft diesem System fügen müssen, und das sich erfolgreich selbst reproduziert, bringen Lari-vière, Haustein und Mongeon auf den Punkt:

Unfortunately, researchers are still dependent on one essentially symbolic function of publishers, which is to allocate academic capital, thereby explaining why the scientific community is so dependent on ‘The Most Profitable Obsolete Technology in History’ [...]. Young researchers need to publish in prestigious journals to gain tenure, while older researchers need to do the same in order to keep their grants, and, in this environment, publishing in a high impact Elsevier or Springer journal is what ‘counts’. In this general context, *the negative effect of various bibliometric indicators in the evaluation of individual researchers cannot be understated.* (Larivière et al., 2015, 13 Herv. YF)

## 2.3 Transformationsprozess der Verlage: Transformationsverträge und Plattformisierung

Die Open-Access-Transformation ist ein dominierendes Thema im wissenschaftlichen Publikationsfeld, das hier in seiner Komplexität und Ambivalenz nicht beurteilt werden kann. Ich werde lediglich einige Aspekte der Transformation thematisieren, und zwar den Abschluss des dritten

bundesweiten Transformationsvertrags (DEAL-Vertrag) im September 2023 (bis Ende 2028) mit Elsevier. Wie auch die ersten beiden Verträge mit Wiley (erster Vertrag 2019-2023, Verlängerung 2024-2028) und Springer Nature (erster Vertrag 2020-2023, Verlängerung 2024-2028) wurde auch der Elsevier-Vertrag durch das DEAL-Konsortium<sup>60</sup> geschlossen, einer Gruppe aus Vorständen der Allianz der deutschen Wissenschaftsorganisationen unter Geschäftsführung der MPDL Services GmbH<sup>61</sup>.

Die DEAL-Verträge regeln die Publikationsmöglichkeiten von Wissenschaftler:innen an deutschen Forschungsinstitutionen, so dass sie in den meisten Zeitschriften der drei Konzerne unter einer offenen Lizenz (CC BY als präferiertes Modell) publizieren können. Die Bibliotheken übernehmen die Gebühren und die Abwicklung. Für alle drei Verträge wurden zudem Archivrechte und der lesende Zugang zu hybriden Zeitschriften vereinbart. Da Deutschland nicht Teil der cOAlition S<sup>62</sup> ist, einer Initiative unterschiedlicher nationaler Fördererorganisationen, der Europäischen Kommission und des European Research Council (ERC) zur vollständigen Umstellung auf Open Access ab Beginn des Jahres 2021, dem Plans S, sind die DEAL-Verträge der größte nationale Schritt zur Umstellung auf Open Access.

Um den Faden der obigen Diskussion der Veränderung im Publikationswesen aufzugreifen bzw. die Frage, warum sich die problematische Konstellation durch wissenschaftsorganisierte Initiativen und die digitalen Distribuierungsmöglichkeiten nicht verbesserte: Laut Hanekop und Wittke (2014) gibt es mehrere Gründe dafür, dass die etablierten Journals ihre dominante Position behalten konnten, obwohl wissenschaftsgeleitete Strukturen Zeitschriften in anderen Funktionen hätten ersetzen können. Das ist zum einen die Qualitätssicherung, welche zwar Wissenschaftler:innen unentgeltlich leiste(te)n, aber die Zeitschriften organisier(t)en. Zum anderen ist es die Verteilung der Reputation durch eine Hierarchisierung von Zeitschriften. Als dann große Forschungsförderer mit Grün-

---

60 <https://deal-konsortium.de/ueber-deal/organisation>. [zuletzt geprüft am 25.07.2024]

61 Die Organisation des DEAL-Konsortiums hat sich seit dem ersten Vertrag 2019 verändert.

62 <https://www.coalition-s.org/>. [zuletzt geprüft am 25.07.2024]

und Gold-Modellen<sup>63</sup> und institutionellen Datenbanken in die Transformation zu Open Access einstiegen, war dies durchaus durch den Wunsch geprägt, die Lock-in-Situation im Publikationsbereich aufzubrechen. Doch auch die Großverlage haben, nach anfänglicher Verweigerungshaltung einen Anpassungswandel an die Open-Assess-Transformation gemacht, wodurch sie sogar Vorteile gewannen. Noch vor der Einführung der Transformationsverträge (TV bzw. TA für “transformative agreement”) beobachteten Hanekop und Wittke, dass statt der erhofften Strukturreform durch Digitalisierungsprozesse eher eine Fortsetzung derselben Machtgefüge durch potentere Mittel erfolgt. Es hätte zwar eine „sektorale“ Transformation im zeitschriftenbasierten wissenschaftlichen Kommunikationssystem gegeben, aber „aufgrund der besonderen Adaptionsfähigkeit der Akteure werden dabei jedoch grundlegende Strukturmerkmale reproduziert“ (Hanekop & Wittke, 2014, 170). Durch die „graduelle“ Transformation erfolgte nur eine „schrittweise Veränderung der institutionellen Ausgangskonfiguration“ (ebd.), bei welcher die jeweils für Verlage vorteilhafteste Konstellation zum Ausgangspunkt für die nächsten Schritte gewählt wird.

Zwar existieren unterschiedliche Publikationsformate nebeneinander, durchaus auch eine OA-community, die scholar-led publishing zum Beispiel in den Digital Humanities vorantreibt und sich von Abhängigkeiten von rein quantitativen Indikatoren abgewendet hat. Der seit 2022 von der DFG geförderte *Action Plan for Diamond Open Access*<sup>64</sup>, der Beschluss des EU-Rats<sup>65</sup> sowie Initiativen wie die Nationale Forschungsdaten-Infrastruktur NFDI<sup>66</sup>, die dauerhaft von der DFG gefördert werden soll, bewegen sich in diese Richtung. Ihre Reichweite kann aber erst in einigen Jahren beurteilt werden. Im Vergleich zu den durch Forschungseinrichtungen finanzierten Publikationsformen (Transformationsverträge und APCs) waren dies zumindest in den letzten Jahren eher geringfügige Grö-

---

<sup>63</sup> ‚Grünes‘ OA steht für die Archivierung bzw. Zweitveröffentlichung von an anderer Stelle nicht OA-publizierte Arbeit; ‚Gold‘ für bereits als OA publizierte Forschung.

<sup>64</sup> <https://www.dfg.de/de/aktuelles/neuigkeiten-themen/info-wissenschaft/2022/info-wissenschaft-22-26>. [zuletzt geprüft am 25.07.2024]

<sup>65</sup> Council of the European Union (2023).

<sup>66</sup> <https://www.nfdi.de>. [zuletzt geprüft am 25.07.2024]

ßen. Zumindest in Bezug auf große Verlage kann keineswegs von einer freiwilligen oder marktgetriebenen Transformation<sup>67</sup> gesprochen werden, was auch durch die Kritik der cOAlition S an der Umsetzung der Transformation bis einschließlich 2023 deutlich wurde<sup>68</sup>.

Seit einigen Jahren sind Transformationsverträge mit größeren Verlagen und den großen Konzernen (Übersicht in der ESAC-Initiative<sup>69</sup> und speziell die Registry aktueller und vergangener Verträge<sup>70</sup>) mit Verlagsfunktion ein Standardmodell der Open-Access-Transformation an Universitäten und Forschungseinrichtungen. Unter TV werden meist Konstellationen verstanden, bei denen das Subskriptionsmodell (als Read-Komponente, R) mit einem Publikationsmodell (als Publish-Komponente, P) verbunden werden, wobei eine der Komponenten als dominierend betrachtet wird: R&P mit einem finanziellen Hauptanteil im Lesen oder, umgekehrt, P&R mit einem Hauptanteil im Publizieren. Praktisch ist der Unterschied nicht ausschlaggebend. Die Verträge unterscheiden sich stark in ihren konkreten Ausprägungen und müssen nach einer ganzen Reihe von Kriterien verglichen und beurteilt werden. Nominelles Ziel der Transformationsverträge ist eine *zeitlich begrenzte* besondere Unterstützung der Verlage dabei, sich vom Subskriptionsmodell auf ein komplettes Publikationsmodell umzustellen. Bei diesem Zielzustand wird nur noch die Publikation als Leistung einmalig bezahlt (Allianz der deutschen Wissenschaftsorganisationen *et al.*, 2022; Schimmer *et al.*, 2015).

---

67 „Open access has made progress where the academic community is receptive (e.g. physics) or where research funders have issued firm mandates in the public interest (e.g. life sciences and medicine). To date, these examples remain the exception rather than the rule, and have led to the emergence of two parallel markets, rather than transformation of the subscription market. Gold-APC journals operate in a small, competitive and buyer-driven market, while the subscription market remains characterised by inelastic demand, and dominated by large commercial publishers. Cases of journals successfully transitioning, or ‘flipping’, from subscription to Gold-APC or Gold no-APC models remain few and far between.“ (Johnson *et al.*, 2017, 4)

68 <https://www.coalition-s.org/blog/transformation-journals-analysis-from-the-2023-reports/>. [zuletzt geprüft am 27.07.2024]

69 <https://esac-initiative.org/>. [zuletzt geprüft am 25.07.2024]

70 <https://esac-initiative.org/about/transformation-agreements/agreement-registry/>. [zuletzt geprüft am 25.07.2024]

TV sind ein Folgeschritt nach einem Hybrid-Modell, das als „double-dipping“ kritisiert wurde. Dabei haben Bibliotheken weiterhin Zeitschriften über Subskriptionen finanziert, aber für die Publikationen von Artikeln in Open Access haben dieselben Zeitschriften zusätzliche Gebühren erhoben. Das sind die sogenannten APCs (Article Processing Charges, manchmal auch Article Publication Charges), die zwischen mehreren Hundert und mehreren Tausend Euro betragen. TV werden für eine bestimmte Laufzeit geschlossen und halten fest, für welchen Preis (Grundpreis und Zusatz-APCs) die corresponding authors der teilnehmenden Einrichtung in welchen Hybrid- oder Open-Access-Zeitschriften (hier gibt es meistens einen geringfügigen Rabatt auf die APCs) des Anbieters im Gold-Open-Access publizieren können. Der lesende Zugang ist – meist als Zugang zur Plattform des Verlags – enthalten. Besonders prestigereiche Titel sind meist von dem Vertrag ausgenommen, für sie müssen volle und häufig besonders hohe APCs bezahlt werden. Eine Konsequenz dieser Konstellation ist eine signifikante finanzielle Mehrbelastung publikationsstarker Einrichtungen.

Eine Analyse aus dem Research Consulting hält 2017 auf Basis von 20 Studien fest, dass europäische „policy interventions“ bei Weitem nicht ausreichen, um Open Access für 2020 zu ermöglichen oder die Wettbewerbsbedingungen im Publizieren zu verbessern (Johnson *et al.*, 2017, 7). Verlage seien nicht motiviert, ihre Zeitschriften auf Open Access zu ‚flippen‘ und obwohl „policy makers“ (ebd.) die aktuellen Probleme durchaus adressieren wollten, gebe es keine Einigung auf konkrete Maßnahmen. Unterschiedliche Disziplinen und Länder ließen keine einheitlichen Lösungen zu und bräuchten angepasste Programme. Die Autor:innen konstatieren eine Spannung zwischen kurzfristigen Erfolgen im Zugang und langfristigen Änderungen in der Marktstruktur (Johnson *et al.*, 2017, 8). Europäische Ansätze würden ohne eine koordinierte globale Initiative keine signifikante Reform der Subskriptionsmechanismen bewirken, dazu sei das aktuelle Konzept zu lukrativ und etabliert (ebd., 6). Die Abhängigkeiten für Bibliotheken seien durch die OA-Transformation zu den Bedingungen der Großverlage nicht verringert, sondern vielleicht sogar gewachsen: „Rather than transforming the market, TAs shift some portion of subscription investment to funding OA without altering the overall business structure.“ (Farley *et al.*, 2021, 298) Zu ernüchternden Urteilen über die reformierende Kraft der Transformationsverträge kommen

auch andere<sup>71</sup>, insbesondere mit Blick auf die Fortsetzung des Bündelungscharakters der Verträge:

The dominant business models for OA are volume based and reward scale, driving the biggest companies to get bigger, and smaller organizations to seek the shelter of a larger partner. Publication volume is the essential measurement of success in an author pays OA market. Transformative agreements (aka, the ‘Bigger Big Deal’) have become the preferred purchasing model for journals, again favoring scale, because the resource intensiveness required to negotiate and administer such deals leads to the benefits accruing to large publishers with large numbers of journals for researchers and scholars to publish in. [...] As the biggest publishers continue their rapid pace of merging, acquiring new partners, launching new journals, and growing their existing journals, it’s unlikely this trend will slow any time soon. (Crotty, 2023)

Weitere Kritikpunkte kommen von SPARC (2021a, 30f.), u. a.: Da alle Transformationsverträge auf APCs basieren, befördern sie weiter die finanzielle Benachteiligung anderer Formate, die Geistes- und Sozialwissenschaften sowie weniger finanzstarker Einrichtungen. Weiter heißt es, die Wahrscheinlichkeit steige, dass Bibliotheken ihre Ausgaben weiter vergrößern müssen, weil die meisten Verlage versuchen werden, ihre aktuellen Profite auch über die Etablierung von Open Access hinaus auf dem gleichen Niveau zu halten. Die Ausgaben für TV würden die Ressourcen binden, die für alternative Anbieter, wissenschaftsgeleitete Infrastrukturen oder Innovationen genutzt werden könnten. Zudem begünstige die TV-Struktur eine weitere Konzentration der Verlage bzw. Konzerne, da sie mehr Ressourcen für Verhandlungen hätten und mehr zusätzliche Services anbieten könnten als kleinere Häuser, die sich tatsächlich auf die verlegerischen Aufgaben konzentrierten. Die bisherigen Ungleichheiten im Zugang zu wissenschaftlicher Literatur würden durch Ungleichheiten in der Publikationsmöglichkeit für Personen und Institutionen mit weniger Finanzen ersetzt sowie für den globalen Süden. Die dazu angedachten „waivers“, die Wissenschaftler:innen die Möglichkeit

<sup>71</sup> Morais et al. (2019); Szprot et al. (2021); Farley et al. (2021); Brayman et al. (2024); Butler et al. (2023); Larsson (2023).

zu publizieren durch spezielle Rabatte oder Erlasse der Gebühren geben sollen, seien diskriminierend “both because they formalize different categories of authors and also because their actual operations are inconsistent, something which is recognized even within STM” (SPARC, 2021a, 31). Neben der begründeten Befürchtung, dass TV nicht den erwarteten Effekt haben werden, das Publizieren zu transformieren, sondern weiterhin das hybride Modell fortsetzen, ist ein weiterer Punkt besonders spannend. Die Gefahr, dass Unternehmen wie Elsevier, deren Geschäftsmodell sich zur Datenanalyse gewandelt hat, “may be handed vast amounts of additional data on grants (and on grants spending patterns), further contributing to consolidation of the data analytics sector into a small number of competitors with excessive market power.” (SPARC, 2021a, 31) An dieser Stelle wird die Diskussion im nächsten Kapitel aufgenommen.

### 2.3.1 Geschäftsfeld wissenschaftliches Publizieren und Bibliotheken

Ein wichtiger Aspekt der Problematik ist die Plattformisierung der Verlagsangebote. Neben und zusammen mit dem Umstieg auf Paketlösungen war die Plattformisierung ein zentraler Schritt in der Kommodifizierung der Wissenschaft und der Involvierung der Bibliotheken in das Geschäftsfeld der Verlage. Hanekop und Wittke skizzieren, wie sich die Anpassung der Verlage an die Open-Access-Bewegung ab der Jahrtausendwende im Zusammenspiel mit der finanziellen und strukturellen Anpassung der Bibliotheken an die neuen Lizenz-Modelle gestaltete. Verlage setzten dabei auf die Einrichtung von Datenbanken (Scopus), und Web-Portalen (ScienceDirect, SpringerLnk) mit Recherchefunktion und Volltextzugriff, die aber im Gegensatz zu wissenschaftlichen OA-Initiativen (wie ArXiv) kostenpflichtig an Bibliotheken lizenziert wurden (Hanekop & Wittke, 2014, 12). Dabei konnten die gleichen Teuerungsmechanismen für neue Online-Paketmodelle eingesetzt und um neue „avanciertere Methoden der Bibliometrie“ (ebd., 15) erweitert werden, die das System mit stabilisierenden Nachweisen eigener Relevanz versorgten. Die Preisentwicklung Elseviers und anderer Großverlage war seit den 1990er Jahren rasant, was bereits die universitären Budgets auszureizen begann:

In 1998, Elsevier rolled out its plan for the internet age, which would come to be called ‘The Big Deal’. It offered electronic access to bundles of hundreds of journals at a time: a university would pay a set fee each year– according to a report based on freedom of information requests, Cornell University’s 2009 tab was just short of \$2m– and any student or professor could download any journal they wanted through Elsevier’s website. Universities signed up en masse. (Buranyi, 2017)

Hier habe den Verlagen eine institutionelle Eigenschaft des wissenschaftlichen Publikationssystems in die Hände gespielt, so Hanekop und Wittke (2014, 15): Zum einen konnte diese Umstellung des Erlösmodells, das für Verlage deutlich vorteilhafter war, eingeführt werden, weil der Wert der Publikationen relativ mediumunabhängig ist und an den erwarteten Leistungen der Zeitschriften hängt, nicht an der Printform. Zum anderen sehen die Autor:innen die Rollenverteilung des Subskriptionsmodells als relevant, denn Verlage hätten für die Wissenschaft unvorteilhafte Lizenzen nicht an die Wissenschaftler:innen verkaufen müssen, die so unattraktive Optionen womöglich abgelehnt hätten und ihre Literaturbeschaffung anders zu organisieren suchen würden:

Bibliotheken subskribieren die Journale auch dann, wenn die dort erschienen Aufsätze auch über Self-Archiving Archive zugänglich sind. Die Erlösströme der Verlage sind in dieser besonderen Konstellation auch dann institutionell abgesichert, wenn der Zugang zu den Aufsätzen über Open Access Archive möglich ist. (Hanekop & Wittke, 2014, 15)

Die institutionelle Unflexibilität hat demnach ihren Anteil daran, dass ein immer teurer werdendes Subskriptionsmodell auch dann fortbestehen konnte, als die Digitalisierung der Wissenschaftspraxis durch wissenschaftlich verwaltete digitale Plattformen und die Open-Access-Bewegung sie technisch überflüssig machte. Wie oben diskutiert, ist das Argument der Verlage für ihre Notwendigkeit die Qualitätssicherung, die sie durch verschiedene Indikatoren herausstellen. Diese Funktion ist – im Gegensatz zu Journal-Paketlösungen – sehr wichtig sowohl für Wissenschaftler:innen als auch für Institutionen, was wieder zum Problem der Wissenschaftsevaluation und der „Vermessbarkeit der Wissenschaft“ (Holzer, 2022) führt.

Was eindeutig gemessen werden kann, ist die extreme Profitabilität des wissenschaftlichen Publikationssektors. Buranyi (2017) führt vor, welche erstaunlichen Gewinne die Konzerne erzielen:

With total global revenues of more than £19bn, it weighs in somewhere between the recording and the film industries in size, but it is far more profitable. In 2010, Elsevier's scientific publishing arm reported profits of £724m on just over £2bn in revenue. It was a 36% margin – higher than Apple, Google, or Amazon posted that year. (Buranyi, 2017)

Holzer hält 2022 fest, dass der Umsatz der globalen Wissenschaftsverlage inzwischen über 26 Mrd. US-Dollar beträgt (Holzer, 2022, 167).

Durch die Plattformisierung gelangen die Konzerne neben der Publikationstätigkeit an einen weiteren Weg der Profitgenerierung, nämlich den direkten Kontakt zu den Endkunden ihrer Angebote, den Wissenschaftler:innen, Studierenden und anderen Nutzer:innen der Plattform. Während im Printmodell die Bibliotheken als Kontakt zwischen dem Verlag und Leser:innen dienten und für die Anonymität und Sicherheit ihrer Nutzer:innen sorgten, verändert das Plattformmodell diese Konstellation. Die Bibliothek ist die technische Verbindungsstelle, die Nutzer:innen authentifiziert und ihre Verantwortung endet dabei, die Nutzer:innen auf die Plattform zu leiten. Die Problematik dieser Konstellation für den Datenschutz und das Verhältnis zwischen Bibliotheken und ihren Nutzer:innen wird erst seit einiger Zeit thematisiert (Ma, 2023b). In diesem Zusammenhang ist die Verbindung der Kommodifizierung der Wissenschaft und den kommerziellen Überwachungstechniken im „Überwachungskapitalismus“ (Zuboff, 2019) und „surveillance publishing“ (Pooley, 2022) relevant (die in den folgenden Kapiteln diskutiert werden):

The platformisation of scholarly information should be of utmost concerns for research libraries for two main reasons: firstly, the ethical principles concerning information access, as well as privacy and confidentiality of librarians and information professionals [...] are breached; secondly, the open access movement can be sabotaged when commercial platforms take control of what and how scholarly information is organised, disseminated and accessed. (Ma, 2023b, 3)

An dieser Stelle folgt in Anschluss an die Diskussion der digitalen Souveränität oben eine kurze Einschätzung der Grade digitaler Souveränität in Bezug auf die Kriterien: *Daten, Verteilung und Föderierung, Schnittstellen, Soft- und Hardware für den Betrieb der technischen Plattform, technische Kontrolle, Kompetenzen* (i.S.v. Verständnis der Technik) und *Jurisdiktion* (Fokusgruppe, 2019,,16).

**TABELLE 2:** Kriterien digitaler Souveränität in Bezug auf Plattformen.

| <b>Grad digitaler Souveränität/Kategorien digitaler Souveränität</b> | <b>niedrige Ausprägung<br/>(= hohe Abhängigkeit)</b>                                                        | <b>mittlere Ausprägung</b>                                                                                          | <b>hohe Ausprägung<br/>(= keine Abhängigkeit)</b>                                                                                                    |
|----------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Daten</b>                                                         | Der Plattformbetreiber entscheidet, welche Daten er wem zur Verfügung stellt und wie er diese selbst nutzt. | Nutzer/Anwender haben vollständige Kontrolle darüber, wer Zugriff auf Daten hat und können diese jederzeit löschen. | Daten können unabhängig von der technischen Plattform vom Anwender gelesen, geändert und gelöscht werden. Anwender kann den Speicherort frei wählen. |
| <b>Verteilung und Föderierung</b>                                    | Plattform kann nur von einem wirtschaftlichen Betreiber betrieben werden.                                   | Unter bestimmten Bedingungen können Instanzen der Plattform unter eigener Kontrolle betrieben werden.               | Instanzen der Plattform können unabhängig von der Einwilligung einer zentralen Organisation betrieben werden.                                        |
| <b>Schnittstellen</b>                                                | Keine oder nur proprietäre Schnittstellen verfügbar.                                                        | Unterstützung einer hohen Anzahl offener Standards und Schnittstellen.                                              | Zugriff auf alle Daten und Funktionen über offene, frei nutzbare Schnittstellen mit quelloffener Funktionsimplementierung.                           |
| <b>Software für den Betrieb der technischen</b>                      | Software ist weder im Quellcode noch binär verfügbar, sie steht aus-                                        | Software kann unabhängig vom Plattformbetreiber unter eigener                                                       | Quellcode der Software ist unabhängig vom Plattform-                                                                                                 |

|                                      |                                                                                                                     |                                                                                                                                          |                                                                                                                                                       |
|--------------------------------------|---------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>schen Plattform</b>               | schließlich dem Plattformbetreiber zur Verfügung.                                                                   | Kontrolle oder bei einem anderen Betreiber ausgeführt werden.                                                                            | betreiber veränderbar/verändert nutzbar.                                                                                                              |
| <b>Hardware/Technische Plattform</b> | Wirtschaftlicher Betreiber muss technische Lösung außerhalb der EU zukaufen.                                        | Bestehende Non-EU-Lösungen können durch EU-Entwicklungen ergänzt werden.                                                                 | Alle technischen Komponenten können in der EU entwickelt und vorgehalten werden.                                                                      |
| <b>Technische Kontrolle</b>          | Die Plattform ist nur bei einem einzigen Anbieter verfügbar, es gibt keine Kontroll- oder Migrationsmöglichkeiten.  | Nutzer/Anwender kontrollieren wichtige Teile und können diese migrieren; der Aufbau einer selbst betriebenen Lösung ist möglich.         | Nutzer/Anwender haben Kontrolle über alle Komponenten (Quellcode, Hardware, ...).                                                                     |
| <b>Kompetenzen</b>                   | Kaum Verständnis für Prinzip, Voraussetzungen und zugrunde liegende Technik vorhanden; Fokus liegt auf Anwendungen. | Verständnis über Prinzip, Voraussetzungen und Technik ist prinzipiell vorhanden; insb. Kompetenz zur Bewertung existierender Player.     | Kompetenzen für Verständnis, Prinzipien, Voraussetzungen, technologische Zusammenhänge; Beurteilung der existierenden Player vorhanden und gesichert. |
| <b>Jurisdiktion</b>                  | Plattform untersteht Nicht-EU-Recht.                                                                                | Plattform untersteht Nicht-EU-Recht, aber es bestehen verlässliche Verträge, welche die Einhaltung europäischer Standards sicherstellen. | Plattform befindet sich in Deutschland bzw. in der EU und untersteht ausschließlich dieser Jurisdiktion.                                              |

Tabelle 3 hält fest, wie die digitale Souveränität deutscher Einrichtungen im DEAL-Elsevier einzuschätzen wäre. Bei allen zentralen Kategorien wissenschaftlicher Plattformen liegt die Kontrolle der Plattform bei Elsevier.

**TABELLE 3:** Ausprägung digitaler Souveränität in Bezug auf wissenschaftliche Plattformen.

| Grad/ Kategorien digitaler Souveränität            | Niedrige Ausprägung | Mittlere Ausprägung | Hohe Ausprägung |
|----------------------------------------------------|---------------------|---------------------|-----------------|
| Daten                                              | ■                   |                     |                 |
| Verteilung und Förderierung                        | ■                   |                     |                 |
| Schnittstellen                                     | (unklar)            | (unklar)            | (unklar)        |
| Software für den Betrieb der technischen Plattform | ■                   |                     |                 |
| Hardware/ Technische Plattform                     | ■                   |                     |                 |
| Technische Kontrolle                               | ■                   |                     |                 |
| Kompetenzen                                        | ■                   |                     |                 |
| Jurisdiktion                                       |                     | EU-Recht            | EU-Recht        |



3.

## **Lock-In II Institution – Elsevier und RELX**

Elsevier, meist als Verlag bezeichnet, ist seit einiger Zeit kein Verlag mehr, sondern der Sammelbegriff für das Marktsegment Scientific, Technical und Medical business (STM) der RELX Group (vormals Reed Elsevier), einem in London basierten börsennotierten Medienkonzern. RELX beschreibt sich selbst als “global provider of information-based analytics and decision tools”<sup>72</sup>, der “data, and analytical services for corporate, professional, and academic customers” (SPARC, 2019, 10) anbietet. RELX hat vier Marktsegmente, a) risk, b) scientific, technical & medical, c) legal und d) exhibitions. Das Risk Segment hat im Jahr 2023 mit 3,133 Mrd. Pfund zum ersten Mal das STM-Segment mit 3,062 Mrd. Pfund in Bezug auf die Einnahmen überholt. Risk hat doppelt so hohe Wachstumsraten (4% STM und 8% Risk), im bereinigten operativen Gewinn (EBIT) liegen beide Segmente aber gleichauf mit 1,165 Mrd. Pfund (RELX, 2024, 14). Gesamt EBIT für das Jahr 2023 liegt bei 3 Mrd. Pfund (RELX, 2024, 6).

Elseviers führende Position geht einher mit einem außerordentlich schlechten Ruf: “Elsevier has perhaps the single worst reputation” (Tenant, 2018) aufgrund “tone-deaf business practices, which have been viewed as a land grab by the academic community” (ebd.). Larivière, Haustein und Mongeon (2015) zeigen RELX’ (damals Reed Elseviers) beeindruckende finanzielle Entwicklung von Anfang der 1990er Jahre bis 2013 und zwar sowohl für den Gesamtkonzern als auch für den STM-Bereich Elsevier. Sie stellen fest, dass die Profite sich zwischen 1991 und 1997 mehr als verdoppelt haben, von 665 Millionen auf 1,45 Milliarden US-Dollar. Die Gewinnmarge stieg von 17% auf 26%. Nach einem kleinen Rücklauf 1998 bis 2003 stiegen Profit und Profitmargen bis zur Finanzkrise 2008, erreichten dann Profithöhen von über 2 Milliarden US-Dollar in 2012 und 2013. Die STM-Sparte zeigt eine besonders hohe Gewinnmargen-Entwicklung von 2006 mit 30,6% bis 38,9% im Jahr 2013 (Larivière *et al.*, 2015, 10). Während Elsevier 2019 noch die größte “area of activity” für RELX war und mit 33,7% der Einnahmen ganze 40,5% des Profits (SPARC, 2019, 10) für RELX brachte, wird Elsevier mittlerweile vom Risk- und Analytics-Geschäft überholt. Für den STM-Bereich verzeichnet der RELX Jahresbericht 2023 Einnahmen von 3,062 Mrd. Pfund (verglichen mit 2,909 Mrd. Pfund 2022 und 2,649 Mrd. Pfund 2021), mit 47% der Einnahmen aus

---

72 <https://www.relx.com/>. [zuletzt geprüft am 25.07.2024]

den USA, 22% aus Europa und 31% “from the rest of the world” (RELX, 2023, 20).

RELX investiert massiv in technologische Entwicklung: 1,7 Mrd. US-Dollar jährliche Ausgaben für Technologie, über 3000 Techniker:innen in der Sparte Risk und 11.000 bei RELX insgesamt (der Konzern wirbt mit mehr als 15 Jahren Erfahrung mit big data, AI und machine learning) (RELX, 2023)<sup>73</sup>.

Die STM-Sparte beinhaltet Angebote in den Bereichen “Academic & Government, Corporate and Health organisations”. Die im Bericht genannten Anwendungen (“solutions“) “Scopus, SciVal, Pure, ClinicalKey, ClinicalPath, Embase, Engineering Village, Interfolio, Reaxys, SciBite, HESI, Sherpath, Shadow Health, Complete Anatomy, Osmosis and Gravitas” (RELX, 2023) lassen sich folgendermaßen gruppieren:

- *Medizin, Lebenswissenschaften, Pflege*: „Reaxys“, „Sherpath“<sup>74</sup>, „Shadow Health“<sup>75</sup>, „Complete Anatomy“<sup>76</sup>, „Osmosis“<sup>77</sup>, „ClinicalKey“<sup>78</sup>, „ClinicalPath“<sup>79</sup> und „Gravitas“<sup>80</sup>
- *Studiums- und Prüfungssoftware*: „HESI“<sup>81</sup>
- *Anwendungen rund um den Forschungsprozess*: „Scopus AI“<sup>82</sup>, „SciVal“<sup>83</sup>, „Digital commons“ Repositorium und Verlag (beepress)

---

73 <https://www.relx.com/media/press-releases/year-2024/annual-report-2023>. [zuletzt geprüft am 25.07.2024]

74 <https://evolve.elsevier.com/education/sherpath/>. [zuletzt geprüft am 25.07.2024]

75 <https://evolve.elsevier.com/education/simulations/shadow-health/>. [zuletzt geprüft am 25.07.2024]

76 <https://3d4medical.com/>. [zuletzt geprüft am 25.07.2024]

77 <https://www.osmosis.org/>. [zuletzt geprüft am 25.07.2024]

78 <https://www.clinicalkey.com>. [zuletzt geprüft am 25.07.2024]

79 <https://www.clinicpass.net/>. [zuletzt geprüft am 25.07.2024]

80 <https://risk.lexisnexis.com/healthcare>. [zuletzt geprüft am 25.07.2024]

81 <https://www.elsevier.com/products/hesi>. [zuletzt geprüft am 25.07.2024]

- Nicht im Bericht aufgeführte Angebote im Bereich Forschung: „SSRN“, „CellPress“, „Mendeley“, „Plum X Metrics“, „Science-Metrix“<sup>84</sup>
- *Anwendungen im Management*: „Pure“<sup>85</sup>, „Interfolio“<sup>86</sup>, „Funding Institutional“<sup>87</sup>
- *Bereichsübergreifende Anwendungen*: „Analytical Services“<sup>88</sup> für “academic institutions”, “government bodies“ und “funding organisations“; „Data monitor“<sup>89</sup> für die Sammlung und Aufbereitung von Metadaten zu Forschungsdaten aus öffentlichen Repositorien; „Datasets“<sup>90</sup> for AI and digital transformation<sup>91</sup>: Auswertungskapazitäten, predictive modeling und generative AI-Anwendungen für die Lebenswissenschaften und Industrie, “enriched and authoritative scientific Datasets to power data applications that solve R&D challenges” (Elsevier, 2023).

Die Selbstpräsentation der einzelnen Anwendungen ist inhaltlich wenig aussagekräftig und es ist nicht ganz klar, wie sich die Datenanalyse-Angebote unterscheiden. Zu Datasets gibt es eine große Menge (mindestens mehrstellige Millionenangaben) an Quellen aus Volltexten aus peer-reviewed journals, Autor:innen-Profilen, zitierten Referenzen und fachspezifischen Inhalten wie Patenten, chemischen Substanzen u.a. Hier scheint Elsevier intensiv in AI-gestütztes predictive modeling zu investieren:

---

82 <https://www.scopus.com/search/form.uri?display=basic#basic>. [zuletzt geprüft am 25.07.2024]

83 <https://www.scival.com/landing>. [zuletzt geprüft am 25.07.2024]

84 <https://www.science-metrix.com/>. [zuletzt geprüft am 25.07.2024]

85 <https://www.elsevier.com/products/pure>. [zuletzt geprüft am 25.07.2024]

86 <https://www.interfolio.com/news-and-events/press/elsevier-interfolio-acquisition/>. [zuletzt geprüft am 25.07.2024]

87 <https://www.fundinginstitutional.com>. [zuletzt geprüft am 25.07.2024]

88 <https://www.elsevier.com/de-de/products/analytical-services>. [zuletzt geprüft am 25.07.2024]

89 <https://web.archive.org/web/2024011121817/https://www.elsevier.com/products/data-monitor>. [zuletzt geprüft am 25.07.2024]

90 <https://www.elsevier.com/solutions/datasets>. [zuletzt geprüft am 25.07.2024]

91 <https://www.elsevier.com/industry/ai-in-rd>. [zuletzt geprüft am 25.07.2024]

“Spanning all disciplines, Datasets enable KOL [Key Opinion Leader] identification and rising star selection; predictive modeling (e.g., material property predictions or drug-drug interactions); training sets; knowledge graph creation; enterprise, federated and/or semantic search; business intelligence dashboards; and algorithm and neural network training.” (Elsevier, 2023)

Elsevier „Fingerprint Engine“ dient offenbar als AI-Kernfunktion für andere Anwendungen, nämlich Expert Lookup, SciVal, Scopus (Author Profile). Der Unterschied zwischen Fingerprint Engine und AI in Scopus AI ist unklar. „SciBite“<sup>92</sup> scheint als “Semantic analytics software” die Grundlage für Elseviers AI zu sein.

Mein Interesse gilt den Anwendungen im Forschungsprozess, im Wissensschaftsmanagement und den bereichsübergreifenden Anwendungen. Die Bereiche Medizin, Lebenswissenschaften und Pflege sowie HESI werden im Folgenden nicht behandelt.

Eine detaillierte Analyse der finanziellen Entwicklung der marktbeherrschenden Konzerne Elsevier, Wiley, Clarivate und Springer Nature bietet die SPARC Landscape Analysis 2019, welche die Veränderungen im akademischen Publizieren und deren Implikationen für akademische Institutionen untersucht. Es gibt zudem kürzere Updates für die Jahre 2020 und 2021. Die Landscape Analysis sieht Elsevier in einer führenden Position. Zwar stehe Elsevier in der Anzahl der Zeitschriften hinter der Springer Nature Group, mache dafür aber größeren Umsatz mit seinen Zeitschriften und liege mit einer Gewinnmarge von 37% deutlich vor Springer Natures 23% (SPARC, 2019, 11). Die Landscape Analysis 2019-21 sind verfasst vom Marktanalytiker Claudio Aspesi in Kooperation mit dem SPARC-Team und basieren u. a. auf zahlreichen Interviews mit Expert:innen aus den Schlüsselbereichen des Publikations und Wissenschaftsbereichs: “dozens of key stakeholders including provosts, CIO’s, library leaders, students, and higher education administrators in a wide variety of North American institutions, as well as publishers, and other market experts” (SPARC, 2019, 2).

---

92 <https://scibite.com/>. [zuletzt geprüft am 25.07.2024]

Elseviers Neuorientierung in Richtung der Wissenschaftsverwaltung war bereits 2015 in einer Präsentation für Investor:innen angekündigt, die neuen Ziele seien Services für “university administrations, funding bodies, and governments with tools aimed at estimating and improving the productivity of research and optimizing funding decisions” (SPARC, 2019, 14). Aspesi rekonstruiert aus Interviews und eigener Analyse vier Kernpunkte als Elseviers Zukunftsstrategie: *Erstens*, die Verteidigung des Zeitschriften-Kerngeschäfts, *zweitens* die Verbesserung der Zeitschriften-Führung durch mehr Einblick in Wissenschaftsprozesse. Durch seine unterschiedlichen Analysefunktionen könnte Elsevier darauf hinarbeiten, Trends frühzeitig zu erkennen und selbst zu setzen (durch Entwicklung passender Publikationsformate für aufkommende Forschungsthemen und Disziplinen, oder durch Anwerbung von vielversprechenden Forscher:innen für eigene Zeitschriften):

Elsevier could also identify which segments of various disciplines are likely to evolve into the next growth area for research by looking (for example) at project participation patterns, size and quality of teams, and funding bodies’ decisions, targeting these segments with new, dedicated journals ahead of other publishers. Similarly, Elsevier could isolate in advance new trends in interdisciplinary studies, allowing it to establish publication forums where none exist today and even driving funding decisions which lead to accelerated growth for these types of research. (SPARC, 2019, 15)

Für solche Analysen sei Elsevier besonders gut aufgestellt, da der Konzern Zugang zu sehr vielen Daten über Forschungsmuster habe und zudem einen taktischen Vorteil durch seine Größe und Marktposition (ebd.). Der *dritte* Punkt ist, diese wertvollen Einsichten an drei Hauptgruppen zu verkaufen, nämlich an Universitäten, Forschungsförderer und Regierungen. Es gibt natürlich schon einen Markt für Produktivitätsmessung von Forschung in bestimmten Bereichen, auf dem Clarivate, Elsevier und kleinere Unternehmen wie Academic Analytics vertreten sind.

Analysen sind bei Universitätsleitungen durch ihre Funktion als Grundlage für wissenschaftliche Personal-Entscheidungen und Ressourcenverteilung begehrt, da sie z. B. besser einschätzen wollten, wie sich Promovierende und Wissenschaftler:innen früh in ihrer Karriere entwickeln (SPARC, 2019, 16). Außerm lassen sie sich sehr gut in die weitere Metri-

sierung und Produktivitätssteigerung für Rankings einpflegen bzw. werden selbst zum konstitutiven Teil davon. So ergeben sich für Elsevier weitere Möglichkeiten der Einflussnahme auf zentrale Prozesse in Wissenschaftsorganisationen. Von hier aus könnte für Universitäten ein Lock-In bereits aus der intransparenten Funktionsweise der Anwendungen und der entstehenden Abhängigkeit folgen, die Elseviers Position aufgrund spezifischer Netzwerkeffekte und Pfadabhängigkeit verstärken. Sind die Anwendungen bereits in universitäre Strukturen integriert, sei es schwierig, sie zu ersetzen:

In general, the experience of the corporate sector is that users tend to rely on specific third-party data and find it difficult to abandon once it is embedded in their core processes. This means that these services can generate recurring revenues with strong pricing power, and Elsevier is uniquely positioned to offer these services given its access to content and underlying data. (SPARC, 2019, 16f.)

Die *vierte* Richtung hält Aspesi für die spekulativste, sie ist jedoch interessant für die Diskussion von Tracking und Data Brokering im nächsten Kapitel. Hier geht es um die Möglichkeit, Erkenntnisse aus der Forschung, aus Trends und Einflussmöglichkeiten als ein neues Produkt an die Industrie und den Investmentmarkt zu verkaufen. Dies könnte zum Beispiel durch die Vermarktung des umfassenden Zugangs zu Wissenschaftstrends oder durch eigene Investitionen in Venture-Capital-Firmen geschehen, die aufkommende Themen der Forschung frühzeitig aufgreifen. Dass Elsevier text und data mining durch Einrichtungen umfassend unterbinden will, sei ein Zeichen dafür, dass der Konzern sich dieser Möglichkeiten bewusst sei (SPARC, 2019, 17). Hier zum Vergleich die obige Textstelle zur Anwendung der Datasets, die Aspesis Vermutungen bestätigt:

Spanning all disciplines, Datasets enable KOL identification and rising star selection; predictive modeling (e.g., material property predictions or drug-drug interactions); training sets; knowledge graph creation; enterprise, federated and/or semantic search; business intelligence dashboards; and algorithm and neural network training. (Elsevier, 2023)

Die wichtigsten systematischen Ereignisse nach den SPARC-Analysen sind der Fokus auf AI, mit der Einführung von “AI building blocks”<sup>93</sup> und den AI-Komponenten in Anwendungen (Scopus AI, SciVal, Expert Lookup) sowie die beiden Transformationsverträge in den Niederlanden (2020-2024) und in Deutschland (2023-2028). Im Folgenden diskutiere ich, wie Elsevier seine Präsenz im Wissenschaftssektor ausweitet und welche Konsequenzen dies mit sich bringt bzw. bringen könnte.

## 3.1 Adaptionprozess an Open Access: Elseviers Durchdringung der Institutionsebenen

Die enormen Preiserhöhungen der Zeitschriftenkrise, für die Open Access noch keine finanzielle Erleichterung bei den Wissenschaftseinrichtungen bewirken konnte, haben nicht nur zum weiteren finanziellen Aufstieg der Oligopol-Unternehmen beigetragen. Brembs *et al.* (2023b), Grossmann & Brembs (2021) und andere haben vorgeführt, dass die Mehreinnahmen zur Finanzierung einer Neuausrichtung der ehemaligen Wissenschaftsverlage zu Data-Analytics-Konzernen beigetragen haben:

We have seen that Elsevier is effectively reinvesting some of these gains in new products and services which offer no benefit to libraries, as they are targeted to other users within universities (and even to completely different customers). In other words, librarians are asked to bear the financial burden of investments which will benefit some other category of customers, as well as the shareholders and management of Elsevier. (SPARC, 2019, 11f.)

Diese Strategie ist teilweise eine Reaktion auf die Veränderungen des Publikationsparadigmas hin zu Open Access und Open Science. Hauptsäch-

---

93 <https://www.elsevier.com/industry/ai-in-rd>. [zuletzt geprüft am 25.07.2024]

lich ist sie von der Motivation geleitet, eigene Investitionen – das eigene Zeitschriften-Portfolio – zu schützen und weiterhin hohe Gewinne zu erzielen. Elseviers Handeln beschreiben Posada und Chen (2018, 38) als “a form of rent-seeking simultaneously complimenting its adoption of Open Access”.

Über die letzten Jahre hat Elsevier – und andere zentrale Player im Oligopol wie Springer Nature, Digital Science, Wiley oder Clarivate – für eine beträchtliche Marktkonzentration gesorgt, indem Elsevier nicht nur Zeitschriften, sondern auch Anwendungen innerhalb des Forschungszyklus aufgekauft hat. Darunter fallen Start-ups, kleine Unternehmen oder aus der Wissenschaft entwickelte Tools. Die Abbildung 3 von Chen et al. (2019) zeigt, dass Elsevier in den letzten 15-20 Jahren die Käufe von Zeitschriftentiteln immer weiter zugunsten von Zukäufen an „academic services“ zurückbaut. Von 2009 bis 2017 gibt es nur ein Jahr (2014), in dem nicht mindestens so viele Produkte im Bereich des Forschungszyklus erworben wurden wie im Bereich academic content. Die Zielsetzung hinter dieser schleichenden wie massiven Veränderung, bei der Elsevier nicht der einzige, aber der führende Player ist, und die systematischen Konsequenzen nicht nur für den Publikationsmarkt, sondern für den gesamten Wissenschaftssektor, beschreibt Jeff Pooley:

Led by Elsevier, commercial publishers have, for about a decade, layered another business on top of their legacy publishing operations. That business is to mine and process scholars' works and behavior into prediction products, sold back to universities and research agencies. Elsevier, for example, peddles a dashboard software, Pure, to university assessment offices—one that assigns each of the school's researchers a Fingerprint® of weighted keywords. The underlying data comes from Elsevier's Scopus, the firm's propriety database of abstracts and citations. Thus the scholar is the product: Her articles and references feed Scopus and Pure, which are then sold back to her university employer. That same university, of course, already shells out usurious subscription and APC dollars to Elsevier—which, in a painful irony, have financed the very acquisition binge that transformed the firm into a full-stack publisher. (Pooley, 2024)

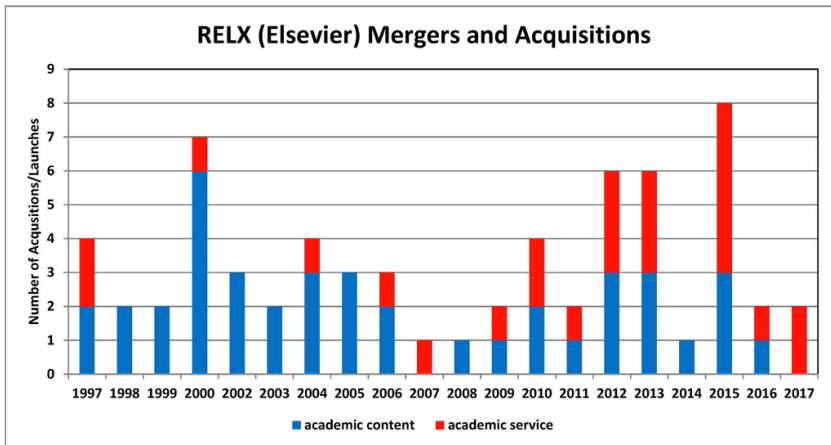


ABBILDUNG 3: RELX (Elsevier) mergers & acquisitions (Chen *et al.*, 2019).

Der RELX-Jahresbericht zeigt, dass diese Anwendungen für Elsevier eine steigende Rolle spielen und mittlerweile ein Gegengewicht zum Publikationsgeschäft darstellen (das nur noch 10 % Umsatz aus Printverkäufen erhält): “Databases, Tools and Electronic Reference, together with Corporate Primary Research, accounts for around 45% of STM revenues, with Academic & Government Primary Research accounting for a similar amount, all in electronic format.” (RELX, 2023, 20)

### 3.1.1 Elseviers aktuelle und potenzielle Beteiligung im Forschungszyklus, in der Wissenschaftsevaluation und im Wissenschaftsmanagement

Im Folgenden teile ich in analytische Einzelschritte, was praktisch zunehmend ineinanderfließt. Die Ebenen der Integration von „academic services“, Workflows, Anwendungen oder Produkten in den Forschungszyklus, die Wissenschaftsbewertung sowie das Wissenschaftsmanagement (als Sammelbegriff für unterschiedliche Vorgänge und Organisationseinheiten der Institution Wissenschaft) haben gerade die Charakteristik, überlappende Effekte zu schaffen. Anwendungen verbinden bisher dis-

junkte Funktionen zwischen Forschung und Management und decken so Bereiche ab, die vorher zu getrennten Organisationseinheiten wissenschaftlicher Institutionen gehörten. Vor allem handelt es sich um Bereiche und Funktionen, die vormals ohne die Involvierung kommerzieller Parteien (die Produkte und Interessen in mehreren dieser Ebenen haben) stattfanden. Dazu gehören Forschung und Evaluation, Förderung, Dissemination sowie diverse lokale Prozesse in Fakultäten und Bibliotheken: Stellenvergabe und Förderentscheidungen, Erfolgsbeurteilung größerer Einheiten, Entscheidungen über den Bezug von Zeitschriften und Publikationsprozesse.

## Schritt I Forschungszyklus

Elseviers Anwendungen erstrecken sich über alle Forschungsphasen und enthalten Tools, die Wissenschaftler:innen in vielen Schritten ihrer Arbeit nutzen. Legt man ein einfaches Modell des Forschungsablaufs an (Discovery, Analysis, Writing, Publication, Outreach, Assessment) sind je nach Disziplin mehrere Produkte von Elsevier in den meisten oder allen Phasen beteiligt. Scopus, ScienceDirect, SSRN, DataSearch und Mendeley unterstützten die Literatursuche und die Erarbeitung der Forschungsfrage, disziplinspezifisch kann mit hivebench ein elektronisches Laborbuch verwendet werden; beim Schreiben und Publizieren begegnen Wissenschaftler:innen den hilfreichen Funktionen von Mendeley, Scopus und Expert Lookup (für die Unterstützung des Peer Review), der Preprint-Plattform SSRN und dem Repositorium Digital Commons, dem Publikations-Workflow-System Aries Systems und natürlich den über 2900 Zeitschriften. Newsflo analysiert die Medienwirksamkeit der Wissenschaft, die Evaluation erfolgt durch Journal Metrics und mit Plum X werden Altmetrics erfasst. Auf der Fakultäts- bzw. Institutionsebene können Leistungserhebungen mit SciVal und dem Informationsmanagement von Pure erfolgen. Vgl. Abbildung 4 zur Elseviers Präsenz im Forschungsprozess (Kramer & Bosman, n.d.).

Einen etwas anderen Fokus legt die Grafik von Chen, Posada und Chan (2019) bzw. Posada & Chen (2018), die zum einen vergleichen, inwiefern sich die Rolle Elseviers gegenüber der eines traditionellen Verlags gewandelt hat. Zum anderen untersuchen sie, wie im Oligopolssystem die Konzerne Elsevier und Wiley Bereiche untereinander aufgeteilt haben<sup>94</sup>, die



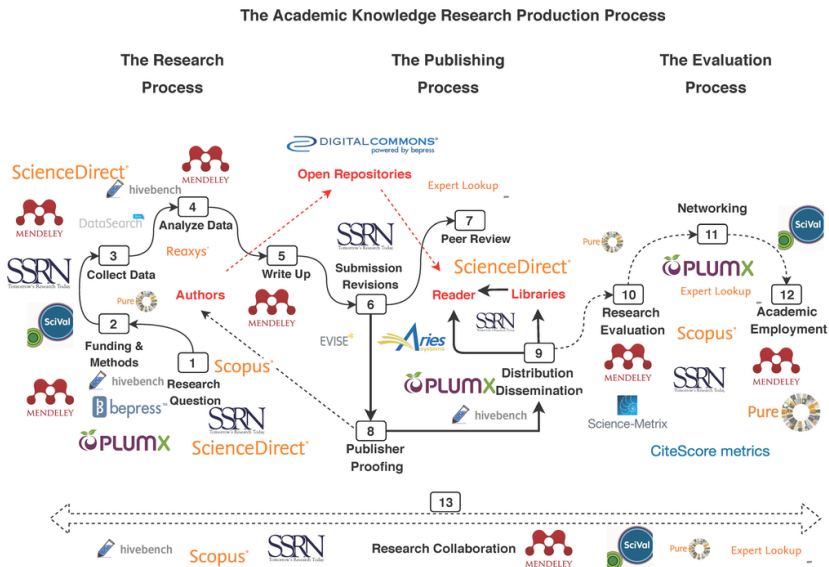
ABBILDUNG 4: Forschungsprozess-Präsenz Elsevier.

sie im Rahmen ihrer transformierten Strategien verfolgen. Für Elsevier sind es die Einflussnahme im Wissenschaftsmanagement und die Integration in Workflows im Forschungszyklus sowie in der Wissenschaftsbewertung. Für Wiley ist es die Dominanz in der Infrastruktur im “higher-level education cycle”, also in Anwendungen und Inhalten rund um Organisation und Durchführung des Studiums. Die Autor:innen zeigen, wie beide Konzerne eine möglichst lückenlose (und alternativlose) Abdeckung der relevanten Prozesse durch ihre Anwendungen anstreben, um so eine Integration in Hochschulstrukturen zu erreichen. So würden die Konzerne, das ist die Konsequenz aus den genannten Studien, durch ihre Anwendungen Einflussmöglichkeiten erlangen und zwar auf individueller wie auf institutioneller Ebene.

## Schritt II Wissenschaftsevaluation

Die potenziellen Einflussmöglichkeiten können sich mit bereits bestehenden Abhängigkeiten des Publikationsbereichs verbinden. Auf der *individuellen* Ebene generieren die Produkte durch ihre Integration mit anderen ‘in-house products’ eine Abhängigkeit, die die Nutzung des Gesamtpakets nicht nur bequem und naheliegend macht. Alternativen zu einem der Produkte werden gleichzeitig zu einer Entscheidung gegen das Gesamtpaket und damit weniger attraktiv. Eine zusätzliche Anreizstufe entsteht durch die Kollaboration mit anderen Wissenschaftler:innen. So wird die Nutzung des Produkts zusätzlich als Kommunikationsmedium

<sup>94</sup> “We believe in particular that the differing expansion paths for the two companies may be a form of oligopolistic behavior to not only leverage competencies but also enhance monopoly power within each lifecycle, whether being education or academic production.” (Posada & Chen, 2018, Absatz 44)



**ABBILDUNG 5:** Elsevier Presence Throughout the Research Lifecycle (Chen *et al.*, 2019).

nahegelegt und die mögliche Ablehnung mit sozialen Hemmungen verbunden: “Beyond merely encouraging the use of the suite, products such as Mendeley [...] and SSRN [...] further promote [...] collaboration with other researchers on the platform, entrenching dependency by serving as the infrastructural basis for collaborative communication.” (Posada & Chen, 2018, Absatz 26) Das ergibt einen Netzwerkeffekt in direkter Form (Kollaboration mit anderen Nutzenden) sowie auch indirekt auf der Ebene der Integrationsmöglichkeit zwischen den Einzelanwendungen: “Mendeley’s Reference Manager [...] and Hivebench’s lab notes [...] features encourage the use of other Elsevier products for which there is integration.” (Posada & Chen, 2018, Absatz 26)

Ein weiterer Punkt ist die mögliche Einflussnahme auf Forschungsvorhaben, Kooperationen und Publikationsorte, die bestimmte Produkte in Kombination ausüben. Damit entstehen suggestive Angebote, die einen

Wettbewerbsvorteil um Mittel und Stellen durch ihre jeweiligen Indikatoren und Sichtbarkeitsvorteile nahelegen:

With Elsevier, SSRN's rankings for journals [...] and curated ejournal [...] may further influence the direction of research for individuals merely through visibility. Yet it is SciVal Funding which has the greatest influence on researchers through its ability to recommend targeted funding opportunities for researchers as well as potential collaborators to enhance the possibility for maximum finding [...]. Similarly, Elsevier Journal Finder [...], using Elsevier Fingerprint Engine [...], recommends the most suitable journals for work publication, journals with the greatest opportunities for success. [...] the crossservice integration of Elsevier services encourages the collective use of services due to additional convenience, thereby maximizing Elsevier's influence at the individual level. (Posada & Chen, 2018, Absatz 29)

Im kompetitiven Rennen, das durch bibliometrische Instrumente der Konzerne selbst weiter zugespitzt wird, sind die Angebotspakete für Wissenschaftler:innen eine naheliegende Möglichkeit, einen Vorteil zu erhalten. Gleichzeitig bietet sich für Wissenschaftler:innen so die Möglichkeit, einen Nachteil zu vermeiden, den sie durch Verzicht auf die Angebote erleiden könnten: "funding competition and the need for publication success drives this adoption." (Posada & Chen, 2018, Absatz 29)

Ein entscheidender Schritt dieses Abhängigkeitskomplexes ist die Abgabe der Kontrolle über einen Teil wissenschaftlicher und organisationsinterner Kernprozesse an einen Player, der prinzipiell unter einem Befangenheitsverdacht steht, was jedoch nicht problematisiert wird:

[C]ontrol is transferred to the publisher as their recommendations/ consulting becomes increasingly 'crucial' to achieving the goals of institutions/ individuals, whether funding/publication or enhanced enrollment in education; with the motivations of the institutions in particular tied into the global university rankings. (Ebd., Absatz 32)

Dabei verbinden sich einerseits die Einflüsse lokaler, institutioneller und globaler Konkurrenz, die durch bibliometrische Indikatoren befeuert wird, und andererseits die Integration verschiedener Funktionen in ein-

zeln Elsevier-Produkten. So werden alternative Produkte verdrängt und der Anreiz für ein Elsevier-Gesamtpaket wächst zu einer Pfadabhängigkeit bereits auf individueller Ebene: “The vertical integration is affecting individual decision making not only as a result of convenience but also as a result of the increasing progression towards a lack of choice especially as the convenience of the services and their increasing user base makes the use of such services increasingly mandatory.” (Posada & Chen, 2018, Absatz 32).

Die Option der individuellen Verweigerung – wie das Projekt *The Cost of Knowledge*<sup>95</sup> –, die die individuelle Publikations-, Begutachtungs- oder Herausgeber:intätigkeit als einzelne Handlungen anspricht (und bereits karriererelevante Einschnitte bedeuten kann), wird ungleich schwieriger, weil es vermehrt um gesammelte Kerntätigkeiten geht. Die Entscheidung gegen einen Konzern wird damit zu einer Entscheidung gegen die gesamte vom Konzern aufgekaufte “integrated academic knowledge infrastructure” (ebd.). Wissenschaftler:innen müssten damit auch die negativen Konsequenzen für ihre Jobsuche, Bewerbung um Förderung und Kollaboration in Kauf nehmen. Die Einflussnahme der Konzerne auf Abhängigkeiten im wissenschaftlichen Bereich steigt damit nochmals dramatisch: “Through their rebranding of information analytics, Elsevier and Wiley have used these technologies to produce and reproduce systems of dependence.” (Posada & Chen, 2018, Absatz 32)

Die Einbindung in die Entscheidungsprozesse und Workflows führt auch auf *institutioneller* Ebene in einen Lock-In, der nahe an die Beschreibung der Pfadabhängigkeit von David (1985, 334) kommt, nämlich “technical interrelatedness, economies of scale, and quasi-irreversibility of investment”:

[T]he products themselves build a clear dependency through its integration into the institutional workflow. Elsevier’s further move to offering metrics-based decision making is simultaneously a move to gain further influence in the entirety of the university process, as well as to monetize its disproportionate ownership of content, in

---

95 <http://thecostofknowledge.com/>. [zuletzt geprüft am 27.07.2024]

addition to cementing an increasing supply of content through such dependence. (Posada & Chen, 2018, 39)

Dass es bei einer Kombination von (Elsevier-)Produkten zur Evaluation der Publikationstätigkeit von Elsevier um einen Interessenkonflikt geht, ist offensichtlich (siehe auch das folgende Zitat). Das ist bereits ein hinreichender Grund gegen ihre Verwendung. Posada und Chen zeigen zudem an der Interaktion zwischen “showcasing” Data Analytics mit Plum X und Bepress (mit “Impact Analytics”) die Gefahren der Einflussnahme, die mit dem Verlust von Kontrolle einhergehen:

Elsevier further serves to influence university decision making by *acting as the creator of source data through its algorithms*. Here, in adopting Elsevier products, universities simultaneously *cede a portion of their control to Elsevier*. For the source of their decisions are no longer the publications and actions themselves, but statistics processed and provided by Elsevier, a clear conflict of interest due to else of the proprietary algorithms used in the generation of not only statistical data but also in recommendations pushed by its software. (Posada & Chen, 2018, 36, Herv. YF)

So kann Elsevier sich selbst in mehrfacher Hinsicht als Ersteller von Zitationsmetriken und Anbieter von Datenanalyse vermarkten<sup>96</sup> und zugleich mit SciVal benchmarking oder Empfehlungen für Kooperationen anbieten. Damit unterstütze SciVal eine Form von Entscheidungsfindung, die im Interesse der Verlagskonzerne liegt und biete die Möglichkeit, Metriken für den Leistungsvergleich zwischen Universitäten zu nutzen (ebd.). So entstünden Empfehlungen für institutionelle und individuelle Kooperationen (ebd.). Was mit dem Versprechen einer state-of-the-art (intransparenter und nicht-vergleichbarer) Analyse-Technologie anfängt, führt über die Überlappung von Funktionen zu Ergebnissen, die mindestens in einem klaren Interessenkonflikt stehen und gleichzeitig in unterschiedlichen Weisen in die Entscheidungsworkflows von Universitäten eingespeist werden:

<sup>96</sup> “Promoting the use of metrics in the decision making process under the promises of efficiency [...], Elsevier, as the generator of data analytics, entrenches itself in the process of university decision making.” (Posada & Chen, 2018, Absatz 36)

*Decisions made as such are fundamentally influenced by Elsevier through the use of metrics created by Elsevier.* The epitome of Elsevier's attempts to influence university decision making comes through Elsevier Analytical Services, whereby Elsevier's in-house team provides a report of the university's performance amongst its peers [...]. Here, Elsevier describes to the university the university's performance and makes the corresponding recommendations, cementing their influence in university decision making. (Posada & Chen, 2018, 36 Herv. YF)

Wenn diese Entscheidungen High-impact-Produkte von Elsevier favorisieren, erscheint dies als eine natürliche und objektive Folge ihrer Qualität. Die Objektivität des Ergebnisses ist jedoch nicht überprüfbar: “[T]here is the potential for internal direction to favour Elsevier: a casualty of the proprietary algorithmic nature of search engines functions as reflected in the antitrust google search result manipulation case [...]” (Posada & Chen, 2018, 41) Allein dieser Aspekt sollte hinreichend sein, damit wissenschaftliche Institutionen solche Produkte ausschließen.

Ein Beispiel, bei dem die individuelle und institutionelle Ebene in innerwissenschaftlicher Transformation zusammengewirkt haben, beschreibt Stephen Buranyi, als “publishers began to meddle with the practice of science itself, starting down a path that would lock scientists’ careers into the publishing system, and impose the business’s own standards on the direction of research.” (Buranyi, 2017) Buranyi bezieht sich auf die Zeitschrift *Cell*. Sie hatte mit Ben Lewin begonnen, ein Vorgehen zu zeigen, das mit einem hohen Renommee und besonderem Profil *assoziiert* wurde. Die Zeitschrift wurde zu einer Marke, an die sich Wissenschaftler:innen anpassten, weil sie dort publizieren wollten. Auch die Herausgeber:innen anderer Zeitschriften übernahmen charakteristische inhaltliche Vorgaben – hohe Ablehnungsquote, große Fragen, lange, ausgereifte Forschungsvorhaben:

Suddenly, *where* you published became immensely important. Other editors took a similarly activist approach in the hopes of replicating *Cell*’s success. Publishers also adopted a metric called ‘impact factor’, [...] it became a way to rank and advertise the scientific reach of their products. The new-look journals, with their emphasis on big results, shot to the top of these new rankings, and scientists

who published in ‘high-impact’ journals were rewarded with jobs and funding. Almost overnight, a new currency of prestige had been created in the scientific world. [...] *It is difficult to overstate how much power a journal editor now had to shape a scientist’s career and the direction of science itself.* (Buranyi, 2017, Herv. YF)

### Schritt III Wissenschaftsmanagement und -politik

Mit der Digitalisierung der Wissenschaft wächst auch die Rolle von FIS (Faculty Information System) und RIMS (Research Information Management System) bzw. CRIS (Current Research Information System) für wissenschaftliche Einrichtungen. Die Abbildung 6 *Why you need CRIS*<sup>97</sup> zeigt, wie Elsevier sich im Wissenschaftsmanagement positioniert. Elsevier bildet darin ein vermittelndes Glied zwischen “institutional customers”, zu denen alle Leitungspositionen von Universitäten gehören – Rektorate, senior research officers, Dekan:innen und Institutsleitungen, Fakultäten mit ihren Wissenschaftler:innen, Bibliothekar:innen und Wissenschaftsservice-Einrichtungen. Elseviers CRIS-Lösung (enabling systems & services als verbindendes analytisches Element) vermittelt laut eigener Darstellung eine erschöpfende Bandbreite an relevanten Erkenntnissen: eine Forschungsstrategie für Universitätsleitungen; Expertise und Kollaborationshilfen für Forschungsreferent:innen; Unterstützung in der Fördermittelbeschaffung und Forschung auf der Ebene der Dekanate und Institute; Forschungsmanagement für Bibliotheken und Impact-Messung für Forschungsservices. Im Folgenden beschreibe ich diese Funktionen genauer.

Zu den Aufgaben von FIS gehören die Organisation wissenschaftlicher Leistungsmessungen für die fakultätsinterne Berichterstattung, Akkreditierung, Stellenvergabe etc.:

FIS allow institutions to electronically collect, store, showcase, analyze, and report on faculty activities and accomplishments using a single system of record. Once data is input into the FIS, faculty can use tools to create biosketches, curriculum vitae (CV), and dossiers. Institutional leaders are also using FIS data to facilitate decision-ma-

<sup>97</sup> <https://www.elsevier.com/academic-and-government/why-you-need-cris>. [zuletzt geprüft am 25.07.2024]

king by gaining deeper insights on what activities faculty are engaged in, the impact of those activities, areas of institutional strength, and institutional gaps that need to be addressed. (Cahill, 2020)

FIS verbinden zwei Ebenen, eine Forschungsebene mit Informationen über Forschungsaktivitäten und eine lokale Ebene. Auf letzterer werden Informationen zum aktiven Personal der Institute bzw. Fakultäten gesammelt und ausgewertet. Für eine Institution haben CRIS strategischen Wert, weil sie klar erscheinende Kriterien der Leistungsmessung (ob es sich um gute Kriterien handelt, ist damit nicht gesagt) und des Vergleichs mit konkurrierenden Einheiten bieten und so Entscheidungen erleichtern sollen. Darüber hinaus bieten sie weitere bequeme Funktionen an: “tools that facilitate the collaboration and publication of research, automate the population of faculty profiles and management of curriculum vitae (CV), track compliance across policies (such as open access policies), and facilitate research analytics, benchmarking, and metrics tools.” (Cahill, 2020) Eine solche Erweiterung, die parallel zur Überlappung von



ABBILDUNG 6: Elsevier CRIS.

Funktionen in Forschungsanwendungen läuft, sieht Cahill künftig auch in der Verbindung von FIS- und CRIS-Funktionen. Das würde folgen, wenn Anbieter von CRIS auch FIS-Funktionen integrieren, z. B. angepasste Workflow-Möglichkeiten und Berichte über die Aktivität innerhalb der Fakultät, Akkreditierungen, Jahresreviews, Beförderungen und tenure (ebd.). In beiden Anwendungen gibt es einen Trend zur Erweiterung des erfassten Bereichs auf die gesamte Institution und nicht nur auf Institute oder Fakultäten, um so einen kompetitiven Vergleich auch innerhalb von Einrichtungen zu ermöglichen. Ein wachsender Bedarf nach “research and faculty intelligence and established performance management frameworks” spiegelt den Wettbewerb im universitären Feld wider. Auf diese Suche nach einfachen Kriterien und Kennzahlen für Bewertungen und Entscheidungsgrundlagen geben vermeintlich objektive “analytics and business intelligence and performance metrics” (Cahill, 2020) eine Antwort. Passenderweise kaufte Elsevier im Juni 2022 das Faculty Information System Interfolio<sup>98</sup>, was SPARC scharf kritisierte (Coalition, 2022b,a; SPARC, 2022a, c, b). Besonders problematisiert wurden die damit verbundene Marktkonzentration mit Folgen wie Preisanstiege, Bündelungen<sup>99</sup> und Gefahren für die digitale und wissenschaftliche Souveränität und den Datenschutz.

Informationen über faculty activity und damit frühzeitige Einblicke in Forschungsthemen, neue Projekte und damit neue Publikationen oder über vielversprechende junge Wissenschaftler:innen seien ein strategischer Vorteil für Elseviers verlegerisches Geschäft: “These advantages could significantly increase Elsevier’s power in the academic publishing market in which it is already the dominant player and could reduce the possibility of future deconcentration in this market.” (SPARC, 2022a) Auch

---

<sup>98</sup> <https://www.interfolio.com/news-and-events/press/elsevier-interfolio-acquisition/>. [zuletzt geprüft am 25.07.2024]

<sup>99</sup> “Given Elsevier’s history of bundling other products, the acquisition raises the possibility of predatory pricing by increasing the incentives to and improving the company’s position for doing so. Unlike its academic journals, Elsevier’s research analytics and Faculty Information Systems (FIS) products have direct substitutes, making predatory pricing a potentially tempting strategy for cementing market share in these areas. For example, Elsevier may choose to bundle its analytics tools with Interfolio’s Faculty Information System at a below-market price. This discounting would disincentivize Interfolio customers from using a competitor like Clarivate. Once its competitors are weakened, Elsevier would face significantly reduced constraints to increase prices to above-market rates.” (SPARC, 2022b)

für die Forschungsanalyse könnten fakultätsinterne Informationen, die Interfolio sammelt, für die Entwicklung neuer Metriken der Produktivitätsmessung genutzt werden, die eng mit FIS verknüpft sind. Der Vorteil interner Informationszugänge sei enorm, so könnte Elsevier interne Produkte an der Konkurrenz vorbei platzieren oder die Informationen selbst im Datenhandel verwerten (ebd.).

SPARC sieht durch dieses Produkt im Elsevier-Portfolio ein Datenleck mit dem Potenzial “[to] compromise faculty privacy” (2022b). Elsevier komme in der Lage, mehr und genauere, nicht-öffentliche und sensible Informationen über Wissenschaftler:innen abzugreifen, weil FIS wertvolle geschützte Informationen enthalten. Dazu gehören Informationen über lokale Forschung und Arbeitspläne, Förderungen, Gutachten für Stellenbesetzungen, Beförderungen und tenure sowie über Fakultätsmitglieder, die sich gerade wegbewerben (SPARC, 2022a). Zudem fördere dieser Kauf die weitere Festlegung auf metrikenbasierte Entscheidungskriterien zum “solely data-driven commodified business” (Ma, 2023b, 5) und „erode academic freedom“ (SPARC, 2022b). Auch hier geht bei der Implementierung von Produkten die Abgabe von Kontrolle an Elsevier einher, wovon auch de Rijke (2020) eindrücklich warnt. Ein Lock-In auf Institutionsebene wird durch die Verbindung von Produkten wahrscheinlicher, weil konkurrierende Produkte für *einzelne* Funktionen nicht mit dem *Paket* an Elsevier-Produkten mithalten können “locking [institutions] into Elsevier’s products” (SPARC, 2022b).

Elseviers CRIS-Produkt ist Pure, das den Workflow im universitären Forschungsmanagement unterstützen und mit einem Dashboard die Entscheidungsfindung erleichtern soll. Pure enthält Informationen über die Forschungstätigkeit des wissenschaftlichen Personals der Einrichtung und gibt gezielte Förderempfehlungen auf Basis der Fingerprint Engine. So greift Elsevier über Pure in den Entscheidungsprozess ein: “Pure enables a top-down flow of recommendations, and thus control, with Elsevier through Pure recommending to institutions.” (Posada & Chen, 2018, 33) Problematisch sind nicht nur die Intransparenz der Anwendung, sondern auch die fehlende Kontrolle oder Einsicht darüber, welche Daten über die einzelnen Personen gesammelt und wie diese verwendet werden. Damit geht die Abgabe der Kontrolle über einen wichtigen Aspekt akademischer Praxis einher, nämlich die selbstverantwortete Präsentation in der wissenschaftlichen Gemeinschaft, durch CV, eigene Websites

oder Publikationslisten, die nach eigenem Ermessen gestaltet werden. Diese Eigenständigkeit wird Wissenschaftler:innen aus der Hand genommen. Gleichzeitig gibt es für Einrichtungen keine Überprüfungsmechanismen, da die Algorithmen nicht zugänglich sind: “citation trends cannot be directly controlled by the researcher and the ‘Fingerprint Engine’ is purely driven by text mining and pattern matching by the algorithms programmed by Elsevier’s engineers.” (Chen & Chan, 2021, 18). Chen und Chan (ebd.) warnen davor, dass die Ergebnisse solcher Empfehlungen, mit ihrer “proclaimed accuracy and insightfulness of the platforms and the profiles they develop” zu selbsterfüllenden Prophezeiungen werden, deren Grundlage nicht wissenschaftliche Leistungen sind, sondern die rekursive Verstärkung innerhalb eines intransparenten Systems.

Bepress und Digital Commons stellen eine institutionelle Repositoriums-Struktur, über die Einrichtungen auch Abschlussarbeiten und Dissertationen publizieren können. Damit bietet Elsevier, flankierend für die CRIS-Anwendungen und die Elsevier-Zeitschriften, eine Publikationsmöglichkeit für Einrichtungen ohne eigene Repositorien. Elsevier wirbt mit Sichtbarkeitsvorteilen durch die Manipulation der Suchfunktion: “Bepress further attempts to optimize institutional output by moving academic output to the top/near-top of search results [...]” (Posada & Chen, 2018, 33). Da die meisten Einrichtungen mittlerweile entweder eigene Repositorien haben oder Kapazitäten größerer Einrichtungen nutzen können, dürfte die Attraktivität des Angebots von Beepress gerade in der Sichtbarkeit liegen. In der Kombination der versprochenen Sichtbarkeit mit Publikationsmöglichkeiten sehen Posada und Chen (ebd.) eine weitere Schneise für Abhängigkeiten.

Mit Funding Institutional<sup>100</sup> bietet Elsevier die nächste Anwendung für eine lückenlose Abdeckung des Wissenschaftsökosystems, das möglichst zu einem “Elsevier research ecosystem” werden soll. Die Fördermittelsuche soll – wie schon die Forschung und das Institutionsmanagement – *vermittelt* von Elseviers Plattform ablaufen, also beobachtet und ausgewertet von Elsevier. „A single solution“ für “discovery, funding insight and decision-making” (ebd.), die auch die anderen “unique Elsevier assets, such

---

<sup>100</sup> Zitiert von <https://www.fundinginstitutional.com/>. [zuletzt geprüft am 25.07.2024]

as Pure, Scopus, Scival and the Elsevier Fingerprint Engine” (ebd.) integriert, verleitet dazu, Wissenschaft komplett im Elsevier-System zu betreiben, ohne andere Quellen, Anwendungen oder Perspektiven. Da Elsevier durch Funding Institutional auch frühzeitig Informationen darüber erlangen könnte, welche Förderlinien gesucht werden, welche Personen oder Institutionen welche Projekte planen und wann einreichen, hätte das Unternehmen einen wertvollen Vorteil gegenüber möglicher Konkurrenz.

Ein Fall von Interessenkonflikt, der zu politischen Wirklichkeit wurde, ist Elseviers Vertragabschluss als Subunternehmen der EU für den Open Science Monitor im Jahr 2018<sup>101</sup>. Der Vertrag wurde auch außerhalb des akademischen Kreises scharf kritisiert, sowohl wegen Elseviers einstiger Opposition gegen Open Access als auch durch die Involvierung einer offensichtlich profitierenden Partei. Kritisiert wurde zudem die mit der Position einhergehende Macht, künftige politische Entwicklungen im EU-Bereich zu beeinflussen:

The fact that Elsevier will be paid to help monitor the dysfunctional publishing world it has helped to create and strives to sustain seems an insensitive decision. Moreover, the contract specifically calls for the ‘socio-economic impacts’ to be evaluated in order to ‘facilitate policy making’. This means that Elsevier will be providing data to guide EU policy decisions that it stands to gain from materially in significant ways. *The obvious conflict of interest here should have disqualified the company immediately.* But the main contractors seem to have no issues with ignoring this glaring problem, or with the fact that many EU researchers will regard Elsevier as the last organization on the planet that should be involved in any way. (Moody, 2018)

Tennant bemängelt nicht nur die generelle Unzulässigkeit einer solchen Kooperation als auch Elseviers konkretes Favorisieren eigener Produkte:

<sup>101</sup> [https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/about-open-science-monitor\\_en#documents](https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science/open-science-monitor/about-open-science-monitor_en#documents). [zuletzt geprüft am 25.07.2024]

[T]he conflict of interest of having an organisation that stands to benefit from the monitor by using its own services is so blindingly apparent that you have to wonder why Elsevier were subcontracted in the first place. How is it reasonable for a multi-billion dollar publishing corporation to not only produce metrics that evaluate publishing impact, but also to use them to monitor Open Science and help to define its [future direction]<sup>102</sup>? Elsevier will be providing data through the monitor that will be used to help facilitate future policy making in the EU that it inevitably will benefit from. That's like having McDonald's monitor the eating habits of a nation and then using that to guide policy decisions. (Tennant, 2018)

### Schritt IV Interessenkonflikt – „Dutch Deal“

Bekannt gewordene Details aus Verhandlungen zwischen der Association of Universities in the Netherlands (VSNU), der Netherlands Federation of University Medical Centres (NFU) und dem Dutch Research Council (NWO) und Elsevier weckten im Dezember 2019 Sorgen (Aspesi, 2019; de Rijcke, 2020) über die Folgen für den künftigen Umgang mit wissenschaftlicher Infrastruktur. Dieser Vertrag könnte, so Claudio Aspesi, „a very insidious precedent for the academic community“ (Aspesi, 2019) darstellen. Im Vertrag (SURF & Elsevier, n.d.) wurden für 16,4 Mio. Euro jährlich für 5 Jahre (Laufzeit 2020-2024) zwei unterschiedliche Leistungen aneinander gebunden. Zum einen war da ein Read&Publish-Transformationsvertrag mit lesendem Zugang zu Elsevier-Zeitschriften (Read-Teil) so wie Publizieren im Gold Open Access *ohne* APCs in den meisten Elsevier-Zeitschriften (Publish-Teil); zum anderen eine unklare Vereinbarung über die Implementierung von Elseviers „Professional Services“ an niederländischen Universitäten bzw. über die gemeinsame Entwicklung neuer „research intelligence services“ (de Rijcke, 2020). Welche Services genau gemeint waren, wie sie zum Einsatz kommen sollten und welche Art von Verbindlichkeiten damit verbunden sein würde, lässt der Vertrag offen.

---

<sup>102</sup> Verlinkung im Original auf: <https://www.techdirt.com/articles/20180402/09131239544/hated-science-publisher-elsevier-to-help-eu-monitor-open-science-including-open-access.shtml>. [zuletzt geprüft am 25.07.2024]

Die unklaren Formulierungen sowie mangelnde Transparenz darüber, ob und in welcher Form die darin angesprochenen Pilotprojekte an welchen Institutionen umgesetzt wurden, stießen auf Kritik (Knecht, 2020, 2019; de Rijcke, 2020; Aspesi, 2019). Die wichtigsten betreffen die *Verbindung* zweier unterschiedlicher Services in einem Vertrag, wodurch bestehende Abhängigkeiten (Publizieren und Zugang zur Literatur) aufrechterhalten und neue geschaffen werden (“Professional Services”) sowie den *Zugang* eines Datenkonzerns zu Daten aus der Forschung und über Forschende:

Elsevier managed to gain access to researchers’ data through the offering of new analytic products and services, essentially bundling together two previously separately procured university contracts into one. [...] Elsevier’s new contract accepts a form of Open Access while enabling the extraction of data from Dutch universities to enrich their analytics products aimed at Open Science. Through such a lock-in bundled contract, Elsevier has the potential to generate infrastructural dependencies as the contract describes the creation of a Pure Community Module to aggregate and duplicate data from existing Current Research Information Systems in the Netherlands (de Rijcke, 2020). (Chen & Chan, 2021)

Dass die Umorientierung vom Subskriptionsmodell auf ein datenbasiertes Portfolio für Elsevier sehr vorteilhaft sei, steht außer Frage, so Aspesi (SPARC, 2020). Dadurch erhalte Elsevier den Zugang zu wissenschaftlichen Informationen und die konkurrenzfreie Sicherung ihrer Produkte auf weiteren Ebenen des Wissenschaftssektors. Es sei mehr als fraglich, ob die Institutionen diesen Schritt mitgehen müssen, der ein deutliches Risiko künftigen Lock-Ins in verschiedenen Organisationsebenen der Einrichtungen berge, um an den anderen Bestandteil des Vertrags (Zeitschriften wie *Cell* oder *The Lancet*) zu gelangen (ebd.). Die Parallele zur Bündelung von Zeitschriften zu Paketen liegt nahe, dieses Mal mit einem noch größeren Paket. Es zwingt Institutionen nicht nur unnötige Zeitschriften auf, sondern eigene Infrastruktur-Elemente, die tief in ihre Funktionen, Wissensbestände und Entscheidungsautonomie eingreifen (SPARC, 2020).

Im Unterschied zum Publizieren neige das Geschäft der Datenanalyse bereits intrinsisch zur Konzentration auf wenige Anbieter, “because users want access to the best data and analytics or to the broadest reach, re-

ardless of cost” (SPARC, 2020, 19). Plattform-Effekte würden dadurch gesteigert, dass es sich um bezahlte, zudem sehr kostspielige Services handelt. Der Vertrag hat zwar eine Nicht-Exklusivitäts-Klausel (s. SURF & Elsevier, Schedule 5, Section 1, 1.2 Collaboration Principles), Universitäten hätten aber wenig administrative Möglichkeiten oder finanzielle Anreize, mehrere Produkte gleichzeitig zu lizenzieren. Elsevier sei sowohl der größte Vertragspartner für Zeitschriften für die meisten Bibliotheken als auch potenziell der größte Anbieter für Analyse-Services, der auch potentere Konkurrenzprodukte wie Digital Science oder Clarivate durch einen „Bigger Deal“ (also Open Access und Data Analytics) ausstechen würde. Die Nachteile eine Monopolstellung eines Anbieters seien schon aus den Publikationserfahrungen bekannt und könnten durch daten-spezifische Probleme ergänzt werden: “the transparency of algorithms, the retention of ownership of the data, the retention of the right to use outputs when contracts are discontinued, non-disclosure agreements and other clauses that have historically penalized academic institutions in their dealings with scholarly publishers.” (SPARC, 2020, 20)

Neben den bereits genannten Lock-in-Effekten entsteht eine spezifische Gefahr für die digitale Souveränität wissenschaftlicher Einrichtungen und eine unabhängige, selbstbestimmte Wissenschaft. “[T]his deal may effectively transfer crucial means to influence Dutch science policy to a monopolistic private enterprise.” warnt Sarah de Rijcke (2020), Professorin für Science, Technology, and Innovation Studies an der Leiden University. Auch Aspesi sieht die größte Gefahr im wachsenden Einfluss eines Konzerns, ohne ausgleichende oder korrigierende Alternativen: “[O]ne company (and one algorithm) may heavily influence decisions on which departments should grow in size and budget, which research projects should be funded, who should be promoted, etc. As we know, algorithms contain errors and biases, and those errors and biases could affect a vast amount of academic research.” (SPARC, 2020, 20) Die Entwicklung zu einem ranking-orientierten Handeln unter verstärktem Einsatz von intransparenter Datenanalyse könnte zu einer Machtasymmetrie über zentrale wissenschaftliche Fragen zugunsten Elseviers führen:

[L]ack of diversity may influence ‘what’ academic institutions measure, not just ‘how.’ If data analytics becomes a quasi-monopoly, what is for sale may well become the metrics that academic institutions use to evaluate their research, as happened with the Impact

Factor. [...] Elsevier is already attempting to play a leading role in defining standards for research assessment. If unchecked, Elsevier may well find itself in the position of defining, *on behalf of the academic community*, what constitutes good research and then selling the tools to perform the actual assessment. (SPARC, 2020, 20f., Herv. YF)

Auch bei Chen und Chan (2021) gibt es ähnliche Befürchtungen. Sie warnen in Anschluss an de Rijcke und Aspesi vor unerwünschten Plattform-Effekten (“undesired platform effects”) und weiteren Kollateralschäden der Abhängigkeit. Dazu gehören auch weitere Entscheidungen gegen Investitionen in lokal entwickelte Scholar-led-Infrastrukturen. Der ‘Dutch-Deal’ stelle einen “complete buy-in into Elsevier’s designed and controlled platform” dar, samt Publizieren, das zwar in Open Access, aber dafür in Elsevier-Zeitschriften stattfindet und von Elsevier-Anwendungen bibliographisch analysiert und kuratiert werde (Chen & Chan, 2021, 15). Die Verschiebung von Kontrolle geschieht durch das Zusammenwirken der einzelnen Ebenen und Services: “[W]e are seeing how enclosure and control over researchers’ data could allow Elsevier to influence funders and university decision making through a new form of algorithmic governance [...]” (Chen & Chan, 2021, 15)

In Weiterführung dieser Argumente möchte ich mögliche Konsequenzen der Beispiel-Anwendungen (Tabelle 4, Unter Schedule 5, Professional service terms, section 1, professional services, 1.2) vertiefen, unabhängig davon, ob sie in dieser Weise umgesetzt wurden oder werden. Diese Beispiele dienen der Verdeutlichung der Problematik. Sie entsteht, indem ein Anbieter auf mehreren institutionellen Ebenen des Wissenschaftsbetriebs und der Wissenschaftsinfrastruktur kritische Positionen besetzt und einen epistemischen Vorteil gegenüber der Institution erhält.

**TABELLE 4:** Beispiele für Anwendungen der Professional Services.

| Use case                                                            | Description                                                                                                                            |
|---------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| 1. Aggregation and deduplication service to based upon CRIS systems | Improve findability and visibility of NL research outputs by aggregating and deduplicating separate CRIS systems into a Pure Community |

|                                |                                                                                                                                                                                      |
|--------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                | module available to all institutions which can serve as a building block to a NL open knowledge base.                                                                                |
| 2. NL Research data            | Link research data from member institutes affiliated researchers in subject or domain specific repositories into Dutch knowledge base.                                               |
| 3. Funding information         | Link NL research outputs to grants and funders (EC, ERC, NWO, RVO, ZonMw), to allow for improved tracking / assessment of impact of funded research.                                 |
| 4. Health Data Management      | Link NL health ‘data silos’ in a secure HDM platform.                                                                                                                                |
| 5. OA compliance as a service  | A proposed service to better use knowledgebase OA publication reminders, meet funder requirements, collect assets + reporting.                                                       |
| 6. Fair recognition and reward | A proposed service to integrate a wider array of metrics and success stories for a better, wider recognition of academics. Inclusion of teaching, society outreach, management, etc. |

Für “Funding information” “Link NL research outputs to grants and funders (EC, ERC, [...]) to allow for improved tracking / assessment of impact of funded research [...]” entsteht für Elsevier die Möglichkeit, eine übergreifende Darstellung des Impacts von Forschung und internationalen Förderinstitutionen auf Landesebene zu geben.

Mit beiden Informationsquellen (CRIS-Systeme und Förderungen) erhielt Elsevier einen epistemischen Vorteil gegenüber sowohl Konkurrenz-Produkten als auch jeder der beteiligten Einrichtungen. Weil der Informationsfluss hin zu den Elsevier-Services erfolgt und dort gebündelt wird, würde ein maßgebendes Instrument mit einem Alleinstellungsmerkmal entstehen. Elseviers Produkt bzw. dessen Ergebnisse wären gegenüber den vergleichbaren Ergebnissen anderer, weniger umfassender

Instrumente oder Institutionen per se besser gestellt. Wissenschaftliche Einrichtungen, Förderorganisationen sowie alle, die an den analysierten Prozessen beteiligt sind – Elseviers Zielgruppe “customers: universities, government, funding organisations” (RELX, 2024, Folie 38) – wissen in einem solchen Szenario weniger als die Stelle, bei der diese Informationen gesammelt werden und die sie mit proprietären Mitteln analysiert. Einzelne Universitäten wären nicht in der Position, die Ergebnisse der Produkte dauerhaft eigenständig zu kontrollieren und befänden sich in Kernbereichen ihrer Management-Entscheidungen in epistemischer Abhängigkeit von diesen Produkten. Wie verlässlich und objektiv so ein Instrument wäre, wie manipulierbar oder beeinflusst von kommerziellen Interessen, wäre in Ermangelung von Instrumenten mit vergleichbar großer Datenmenge nicht überprüfbar. Je mehr Informationen und Funktionen des Wissenschaftsbetriebs bei RELX oder einem anderen Datenanalyse-Konzern zusammenfließen, desto gravierender werden die Abhängigkeit in den Kernentscheidungen und die epistemische Asymmetrie. Damit würden wissenschaftliche Einrichtungen oder ganze Länder als Wissenschaftsstandorte in eine Situation geraten, in der sie wertvolle Aspekte ihres Steuerungswissen über ihre Forschungstätigkeit und Forschungsbewertung, Fördertätigkeit, Management ihrer Einrichtungen und weitere Bereiche ihrer infrastrukturellen Souveränität an kommerzielle Player abgeben. Neben der wirtschaftlichen Abhängigkeit würde damit auch ein epistemischer Lock-In folgen, also das Gegenteil digitaler Souveränität und Wissenschaftsautonomie.

Eine andere Problematik ist mit der Frage verbunden, welche Konsequenzen die Gewährung des Zugangs zu sensiblen Daten hätte. Das Beispiel Gesundheitsdaten-Management (Punkt 4 in Tabelle 4) würde “Link NL health ‘data silos’ in a secure HDM platform” als Service und damit den Zugang zu Gesundheitsdaten auf Landesebene beinhalten. Diese an sich schon bedenkliche Machtposition für einen kommerziellen Player würde potenziert durch Elseviers eigene Aktivitäten im Gesundheitsbereich, die in dieser Arbeit ausgeklammert werden. In welcher Weise dürfte Elsevier/RELX die Daten nutzen, zu denen der Konzern Zugang bekommen würde? Der Vertrag weist darauf hin, dass für die gemeinsamen Produkte zwischen den Services und den Institutionen gesonderte Vereinbarungen zu treffen wären.<sup>103</sup> Mindestens die folgende Passage zu “1.4. Rights of Use and Reuse” legt jedoch nahe, dass Elsevier/RELX auch eine eigene Nutzung der Daten vorsieht, nicht nur ihre Aufbewahrung in

„data silos“: “Elsevier hereby grants to the Institutions a royalty-free and non-exclusive, non-transferable right to use the Deliverables resulting from the Professional Services”.

Was passiert nach Ablauf des Vertrags? Digitale Informationen können weder in der gleichen Weise ‚zurückgegeben‘ werden wie analoge noch kann das Äquivalent der Rückgabe, nämlich die Löschung der Daten, in zuverlässiger Weise kontrolliert werden. Wie sollte gewährleistet werden, dass Elsevier Gesundheitsdaten, auf deren Grundlage gemachte weitere Analyseprodukte etc. nicht mehr als eigenen Datenbestand nutzt (und ist das überhaupt eine plausible Erwartung)? Der Vertrag beinhaltet Passagen, die Elsevier Eigentumsrechte an den in gemeinsamer Erarbeitung entstandenen Produkten gewähren, während die “Customer”-Partei diese nicht-exklusiv nutzen darf:

In the event Customer and Elsevier would jointly decide to enter into a pilot where intellectual property would be jointly created by the parties [...] *Elsevier would become the owner of the intellectual property* and the Customer would be granted a non-exclusive, perpetual, irrevocable, royalty free license to use the jointly created intellectual property for its own internal, non-commercial, purposes. (SURF & Elsevier, n.d., Schedule 5, Section 1.3, Herv. YF)

Im Fall von Gesundheitsdaten ist das eine nicht-triviale Anforderung, deren Vereinbarkeit mit Datenschutzerfordernissen mindestens fraglich sein dürfte. Da nicht klar ist, in welcher Form solche Projekte künftig in Frage kommen werden, ist die generelle Bedingung umso wichtiger: Um einem Data Broker Zugang zu Gesundheitsdaten zu geben, müsste eine zufriedenstellende Lösung dieser, hier nur angeschnittenen praktischen Fragen sowie auch des generellen Interessenkonflikts vorausgesetzt werden.

---

103 Vgl. Schedule 5, Section 1, 1.2 und 1.3.

### 3.1.2 Institutionelle und epistemische Abhängigkeit

Die Problempunkte der Entstehung von ineinandergreifenden Abhängigkeiten auf wissenschaftlicher, institutioneller und finanzieller Ebene und der inhärente Interessenkonflikt in der Nutzung von Elsevier-Angeboten in den Bereichen research assessment, research dissemination (wo bereits Dominanz besteht) und research management hängen zusammen. Dieser grundsätzliche Konflikt bzw. die Nicht-Thematisierung der Zusammenarbeit mit Elsevier wurde auch von SPARC formuliert:

Public interactions with Elsevier's management during the first part of 2021 suggest that Elsevier itself continues to publicly downplay the conflicts of interest among its portfolio of activities. More broadly, little attention seems to be given to the conflicts that arise when Elsevier collects data from researchers and then sells research assessments to academic institutions, funding bodies, and governments. (SPARC, 2021a, 7)

Im Update zur *Landscape Analysis* (SPARC, 2020) nennt Aspesi mehrere Folgen des Interessenkonflikts zwischen der Bewertung und Verbreitung von Forschung durch Elsevier (exemplarisch für das Problem), die ich hier etwas anders als Aspesi aufteile und auf die aktuelle Situation anpasse:

1. absichtliche oder unabsichtliche Beeinflussung zugunsten der eigenen Produkte; (SPARC, 2020, 15)

Zur Erläuterung: Die Argumentation der vorliegenden Arbeit ist zum großen Teil an den Gefahren ausgerichtet, die durch die Zusammenarbeit mit einem kommerziellen Konzern entstehen können und vielleicht unter bestimmten Umständen entstehen müssen. Hier geht es nicht um den Nachweis konkreter Beeinflussung oder die vollständige Sammlung aller Hinweise auf Beeinflussung. Punktuelle Hinweise sollten aber als hinreichende Gründe betrachtet dafür werden, dass es sich um reale Gefahren handelt, z. B. bei Elseviers Rolle beim Open Science Monitor: "Throughout the methods, you can see that there is

an overwhelming bias towards Elsevier products and services, such as Scopus, Mendeley, and Plum Analytics.” (Tennant, 2018) Tennant verweist auf eine Untersuchung von Bergstrom & West von Eigenfactor.org im Dezember 2016, die den Impact Factor und Scopus Site Score vergleicht. Das Ergebnis lautet, dass SiteScore-Analysen deutlich im Interesse Elseviers ausfallen, denn Springer-Nature-Zeitschriften “scored 40% lower and Elsevier titles 25% higher when using CiteScore rather than previous journal impact factors.” (Tennant, 2018)

2. zu enge Bindung an den Hauptförderer für STM-Forschung in den USA erhöht noch die dominierende Position Elseviers und minimiert die Chancen anderer Anbieter und damit Pluralität. (SPARC, 2020, 16)

Dieser Punkt ist genauso gültig für Europa, auch hier ist Elseviers Status als Subunternehmen der EU im Open Science Monitor und als DEAL-Partner zweifellos dominant. Bereits der Abschluss des DEAL-Vertrags wurde stark kritisiert (z. B. von Torny, 2023; Siems, 2023a; Reitz, 2023; Knoche, 2021a; Kaden, 2023; Herb, 2023; Haucap *et al.*, 2021; Brembs, 2023b; Brembs *et al.*, 2023a), weil dadurch unter anderem eine weitere Stärkung Elseviers und Schwächung der Bibliodiversität als Folgen zu erwarten sind. Mit dieser Pfadabhängigkeit – die anderen Argumente sind hier ausgeklammert – werden positive Effekte von Open-Access-Publikationsmöglichkeiten mindestens getrübt.

3. exzessive Kontrolle, die Elsevier durch das 2019 etablierte International Center for the Study of Research (ICSR) erhält. Neben der gebotenen Skepsis gegenüber dem Anspruch einer Neubestimmung von Methodologien der Wissenschaftsbewertung und der Durchführung von Bewertungen “when they have clear conflicts of interest” (SPARC, 2020, 16) warnt Aspesi insbesondere vor entstehenden Asymmetrien. Die Ankündigung einer “computational platform, ICSR Lab, which will host researchers wanting to analyze large data sets (including SCOPUS and PlumX, but also their own)” (SPARC, 2020, 15f.) lässt – da ICSR den gesamten Datenumfang von Elsevier nutzen kann – die Attraktivität dieses Analyse-Labs als hoch einschätzen. Hier bestehe besondere Gefahr für wissenschaftliche Objektivität. Zwar dürften diejenigen, die die Dienste des cloud-basierten ICSR Lab in Anspruch nehmen, die Rechte an ihren Datensätzen behalten,

| but they must grant ICSR and Elsevier the right to use the data fed by researchers ‘for the purpose of developing and maintaining the

ICSR Lab and related services,’ a formulation that is vague enough to allow ICSR (and possibly Elsevier) a wide and potentially unlimited range of uses. At a minimum, Elsevier and ICSR will have access to a significant amount of analytical work on science assessment and be in a position to affect the outcome of research projects on science assessment. (SPARC, 2020, 16)

Es ist schwierig, den Status des ICSR Lab in Bezug auf seine Nutzung der Daten aktuell zu beurteilen. Außerhalb der Berichte von PR Newswire über den Start des ICSR (2019)<sup>104</sup> und des ICSR Labs (2020) (PR Newswire, 2020)<sup>105</sup> stellte es sich als problematisch heraus, Aussagen über die Verwendung der involvierten Daten durch das ICSR Lab selbst zu finden. An einer Stelle ist die Rede von einer *“free-to-use sandbox”*<sup>106</sup>, die Webpräsenz erwies sich jedoch als unergiebig: Der Link zur “Elsevier’s Research Collaborations data protection policy” führt auf die Elsevier Privacy Policy Seite<sup>107</sup>, ohne weitere Informationen zum ICSR. Die Links <https://www.elsevier.com/icsr>, und <https://www.elsevier.com/icsrlab> leiten auf <https://www.elsevier.com/> um, der Link zu Richtlinien “ICSR and responsible research evaluation”<sup>108</sup> führt zum HTTP-Fehler 404, genau wie <https://elsevier.international/icsr/icsrlab.html>. Aus der Perspektive von 2024 scheint eine exorbitante Nutzung des ICSR ausgeblieben zu sein – die Unterseite auf der Elsevier “Research Collaborations”<sup>109</sup> verzeichnet 5 beendete und 10 aktive Projekte.

**104** “Elsevier is the owner of several data sources that are widely used in research assessment (for example, SciVal, Scopus and PlumX Metrics),’ explained Andrew Plume, Chair of the ICSR Advisory Board, and Senior Director Research Evaluation at Elsevier.” (PRNewswire, 2020)

**105** “By allowing the *sharing* of summary results alongside publications *without a license* the ICSR demonstrates its support of open research and open science. ICSR Lab also explicitly supports replication studies. ICSR Lab is powered by the platform Databricks in which analyses are coded in Python or SQL, allowing researchers complete control over their calculations.” (PRNewswire, 2020)

**106** <https://www.elsevier.com/connect/advancing-responsible-research-assessment-elseviers-story>. [zuletzt geprüft am 27.07.2024]

**107** <https://www.elsevier.com/legal/privacy-policy>. [zuletzt geprüft am 25.07.2024]

**108** <https://elsevier.international/icsr/responsible-research-evaluation.html>. [zuletzt geprüft am 25.07.2024]

**109** <https://researchcollaborations.elsevier.com/en/organisations/international-center-for-the-study-of-research>. [zuletzt geprüft am 25.07.2024]

Der grundsätzliche Punkt der Nutzung einer intransparenten proprietären Software, durch die kommerzielle Akteur:innen Zugang zu und die Möglichkeit der eigenen Nutzung von Forschungsdaten erhalten, bleibt von der konkreten aktuellen Umsetzung unberührt. Neben der Machtasymmetrie gibt es eine epistemische Asymmetrie durch einseitige Intransparenz und einen grundsätzlichen Vorteil des Anbieters gegenüber den konkreten ‚Kund:innen‘, den Wissenschaftler:innen. Diese Asymmetrien versetzen ICSR und damit Elsevier prinzipiell in eine Position “to affect the outcome of research projects on science assessment”. (SPARC, 2020, 16)

In welcher Weise fördern die einzelnen Anwendungen Abhängigkeitsverhältnisse und einen Einfluss Elseviers/RELX auf Wissenschaftsinstitutionen? Die Problematik wird durch ein Zusammenwirken mehrerer Faktoren befördert. Der Druck auf Universitäten, sich in einem durch bibliometrische Indikatoren und Rankings geprägten Konkurrenzkampf zu behaupten und auch organisationelle Prozesse nach privatwirtschaftlichen Kriterien auszurichten, macht Elseviers Angebote im Wissenschaftsmanagement attraktiv. Durch die Option, einen der Konzerne zu engagieren, die die relevanten Kennzahlen erheben und Empfehlungen aufgrund dieser Kennzahlen aussprechen, begründet Elseviers Angebot von “research intelligence” eine selbstverstärkende Pfadabhängigkeit. Elsevier profitiert gleich in mehrfacher Sicht davon. Gleichzeitig verlieren die Universitäten weiter Kontrolle und epistemische Hoheit:

As universities increasingly feel the need to out-perform/keep up and thus choose to adopt Elsevier’s services, control is ceded to Elsevier either indirectly through decision making utilizing Elsevier generated data, or directly with Elsevier’s recommendations on funding/collaboration options. Here, Elsevier not only extracts rent through the ownership of data but also through the promotion of high impact journals with the increasing adoption of citation metrics, metrics being simultaneously promoted by publishers such as Wiley and Taylor Francis. (Posada & Chen, 2018, 46)

Sind Anwendungen einmal in den institutionellen Workflow integriert, sinkt die Wahrscheinlichkeit, dass andere Produkte lizenziert werden oder vorhandene ersetzen. Diese ‚natürliche‘ Trägheit steigt weiter durch die Integration der verschiedenen Anwendungen, die es naheliegender

machen, Elsevier Produkte weiter zu nutzen. Diese Produkte verbinden sich mit neuen integrierten Bestandteilen zu einem Ökosystem. So entstehen Netzwerkeffekte, die ihre Nutzung weiter befördern, für Wissenschaftler:innen wie für die Wissenschaftsverwaltung. Das Elsevier-Ökosystem durchdringt und durchleuchtet durch integrierte Anwendungen die Einrichtung. Die verschiedenen Weisen der Einflussnahme werden durch das Sammeln von Daten auf allen Ebenen ergänzt. Da Elseviers Zielgruppe nicht auf Universitäten beschränkt ist, sondern auch “government, funding organisations” (RELX, 2024, Folie 38) enthält, scheint die Ausweitung einer Elsevier-In-House-Solution auf die *Institution* Wissenschaft und ihre Kernfunktionen nicht übertrieben. Diese Entwicklung nennen Chen und Chan eine ‚vertikale Integration‘:

[T]he vertical integration and bundling of the various Elsevier products into an end to-end platform [...] signals a clear strategy [...] to focus on data analytics as a key revenue stream and to monetize the ‘new oil’ of researchers’ personal information and digital ‘fingerprint’, based on publication records and citation counts, research focuses, institutional affiliation, funding sources, and collaborative network. Combining data from these sources and adding machine learning to the mix allows Elsevier developers to create new forms of predictive ‘research intelligence’ at scale and with great efficiency. The sales pitch around these tools is that they allow institutions to create in-depth research profiles of their faculty, make visible research collaborations both statistically and graphically, promote more efficient searchers for knowledge and expertise, and above all, allow administrators to make decisions regarding resources allocation and even hiring decisions. (Chen & Chan, 2021, 16)

Die SPARC *Map of Data Tools and Services* (Abbildung 7) zeigt in einer Gegenüberstellung von Elsevier, Clarivate und Digital Science, wo sich die drei konkurrierenden Konzerne im Gefüge der institutionellen Schlüsselstellen sehen. Mit nur vergleichsweise wenigen Anwendungen sollen bei allen Anbietern die Produktivität erhoben und Wissenschaftsevaluation und Beratung angeboten werden. Diese Funktionen bieten Elsevier, Clarivate und Digital Science gleichermaßen für miteinander konkurrierender Parteien (z. B. Universitäten untereinander und in Bezug auf Förderung) an.

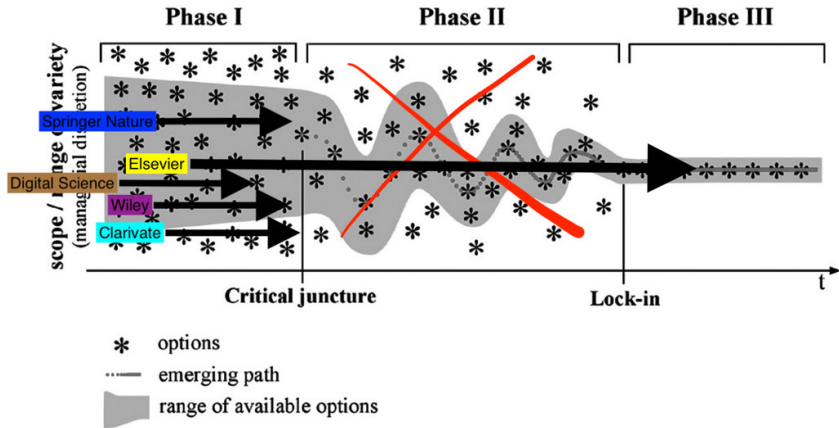
| Exhibit 7: Map of Data Tools and Services Offered by Clarivate, Digital Science, and Elsevier |                                                             |                                        |                             |
|-----------------------------------------------------------------------------------------------|-------------------------------------------------------------|----------------------------------------|-----------------------------|
|                                                                                               | ↓ CLARIVATE                                                 | ↓ DIGITAL SCIENCE                      | ↓ ELSEVIER                  |
| Research Intelligence Solutions for Universities                                              |                                                             |                                        |                             |
| Performance Assessment                                                                        | InCites                                                     | Dimensions                             | Pure, SciVal                |
| Researchers Profiling                                                                         | InCites                                                     | Symplectic                             | Pure                        |
| Researchers Networking                                                                        |                                                             |                                        | Pure                        |
| Expertise Discovery                                                                           | InCites                                                     | Dimensions                             | Pure                        |
| Productivity Analysis                                                                         | InCites                                                     |                                        | SciVal                      |
| Identification of New Trends                                                                  |                                                             |                                        | SciVal                      |
| Identification of partners                                                                    | InCites                                                     | Figshare                               | SciVal                      |
| Benchmarking                                                                                  | InCites                                                     | Dimensions                             | SciVal                      |
| Funding                                                                                       |                                                             |                                        | Funding Institutional       |
| Ad Hoc Consultation                                                                           | Professional Services                                       | Digital Science Consultancy            | Analytical Services         |
| Results Promotion                                                                             | WoS Author Connect                                          |                                        |                             |
| Research Intelligence for Funders                                                             |                                                             |                                        |                             |
| Ad Hoc Consultation                                                                           | InCites, Essential Science Indicators                       |                                        | Analytical Services         |
| Applications Peer Review                                                                      | Professional Services                                       | Digital Science Consultancy            | Expert Lookup/Review Finder |
| Research Intelligence for Governments                                                         |                                                             |                                        |                             |
| Ad Hoc Consultation                                                                           | InCites, Essential Science Indicator, Professional Services | Digital Science Consultancy            | Analytical Services         |
| Research Productivity Tools                                                                   |                                                             |                                        |                             |
|                                                                                               | EndNote, EndNote Site License, Kopernio, Converis           | Readcube, Altmetric, Overleaf, BioRAFT | Mendeley, SSRN, PlumX       |
| Discipline-Specific Solutions                                                                 |                                                             |                                        |                             |
| Chemistry                                                                                     |                                                             |                                        | Reaxys                      |
| Life Sciences                                                                                 |                                                             | Labguru, BioRAFT                       |                             |
| Source: Lit search, interviews                                                                |                                                             |                                        |                             |

**ABBILDUNG 7:** SPARC map data tools and services: Elsevier, Clarivate, DigitalScience.

Elseviers dominante Stellung auf dem Publikationssektor zeigt sich nicht nur im Druck auf Bibliotheken, den DEAL-Vertrag abzuschließen oder Scopus zu lizenzieren und im Druck auf Wissenschaftler:innen, in High-Impact-Journals zu publizieren. Auch für Universitäten würde ein Netz-

werkeffekt entstehen, weitere Produkte der research intelligence zu nutzen und den Empfehlungen der jeweiligen Anwendungen im CRIS- und FIS-Bereich zu folgen. Wissenschaftler:innen haben durch den DEAL-Vertrag eine weitere Motivation, in Elsevier-Zeitschriften zu publizieren. Alternative Anwendungen und kleinere Verlage können mit dem integrierten Angebot eines Elsevier-Ökosystems nicht mithalten. Größere Konkurrenten mit Verlagsfunktion wie Springer Nature und Wiley und ohne Verlagsfunktion wie Clarivate und Digital Science können zwar dazu beitragen, dass das Oligopol nicht zu einem Monopol wird. Beides ist aber nicht im Interesse der Bibliodiversität und einer souveränen Wissenschaft. Zu befürchten wäre zudem, dass Elseviers integriertes Angebot, verstärkt durch die Konzentrationsaffinität datenintensiver Anwendungen (s. Abschnitt zu „Dutch Deal“) sehr schnell zu einem *Lock-In auf mehreren Ebenen* führen würde, sobald die Anwendungen auf Institutionsebene einmal integriert wären. Diese Gefahr lässt sich im Phasen-Modell (s. Teil III) durch Überspringen der zweiten Phase darstellen: Die Abbildung 8 zeigt in der Phase I Elsevier mit den konkurrierenden Konzernen (mit und ohne Verlagsfunktion). Da in der Kombination von Analysefunktionen, Management und Verlagsfunktion Elsevier als Ökosystem aus organisationspezifischer Sicht das stärkste Angebot machen würde, würde ein Institutions-Lock-In deutlich schneller drohen als in Situationen mit stärkerer Konkurrenz.

Neben der Größe beruht Elseviers Dominanz auf einem Wissensvorsprung: „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.“ (Siems, 2022, 1) Diese Asymmetrie besteht zum einen gegenüber der Konkurrenz, also anderen Playern im Publikationsbereich und in der Datenanalyse (manche Konkurrenten sind in beiden Bereichen vertreten, manche nur in einem). Sie besteht andererseits auch gegenüber der heterogenen Gruppe, die der Konzern als seine Zielgruppe bestimmt – s. Abbildung 6 in Schritt III – Rektorate, Senior research officers, Dekan:innen und Leitungen von Instituten, Fakultäten mit ihren Wissenschaftler:innen, Bibliothekar:innen, Wissenschaftsservice-Einrichtungen. Diesen Wissensvorsprung erzielt Elsevier durch unterschiedliche Handlungen:



**ABBILDUNG 8:** Beschleunigtes Drei-Phasen-Modell der “organisational path dependence” (Elsevier).

- Die Plattformisierung des wissenschaftlichen Angebots (ScienceDirect) und Zusammenführung und Integration der Anwendungen aus dem *Forschungszyklus*, Anwendungen im Bereich CRIS und FIS, Anwendungen für die *Forschungsförderung*. Resultat wäre eine überlappende Durchdringung der Institutionsebenen, durch die Elsevier sowohl einen *Zugang zu Informationen* als auch *Einflussmöglichkeit auf wissenschaftsinterne und institutionsinterne Prozesse* erhält.
- Integration der vorhandenen offenen und öffentlichen Strukturen, Forschungsdaten, Metadaten, Open-Science-Ergebnisse in Elseviers eigene Plattformen und Anwendungen, die diverse Tracking-Methoden<sup>110</sup> nutzen.
- Bereichsübergreifende Anwendungen wie Analytical Services, Data monitor oder Datasets, für deren Ausführung Institutionen weitere eigene Daten in Elsevier-Anwendungen einpflegen.
- Pilotprojekte zur Entwicklung gemeinsamer Produkte, über die Elsevier weitreichende Befugnisse bekommt, wie im Dutch Deal oder durch Zugang zu Daten im ICSR.

<sup>110</sup> Mehr dazu im nächsten Kapitel.

Die Wirkung ist dann am größten, wenn möglichst viele Anwendungen genutzt werden, um einen lückenlosen Informations- und Interaktionskreislauf zu erreichen. Durch den überlappenden Charakter vieler Anwendungen ist es aber möglich, diesen Kreislauf auch mit wenigen zentralen Anwendungen zu erreichen. Die Abbildungen 4 und 7 zeigen, dass Mendeley, SciVal und Pure<sup>111</sup> solche disziplinunabhängigen Anwendungen sind. Sie können im Forschungsprozess (Mendeley, Pure), der Wissenschaftskommunikation (SciVal, Mendeley) und der Evaluation auf institutioneller (Pure, SciVal) und individueller (Mendeley) Ebene sowie in der akademischen Stellensuche und -besetzung (Pure, Mendeley) genutzt werden. Die Nutzung allein dieser Anwendungen gibt Elsevier umfassende Informationen über mehrere Aktivitäten wissenschaftsinterner wie wissenschaftsorganisatorischer und institutioneller Art und gewährt zugleich regulierenden Einfluss, z. B. durch Gewichtung, Empfehlungen und eigene Indikatoren.

Der Lock-In II wäre – im Unterschied zum Lock-In I auf Publikationsebene – tiefer in den Workflows der Institution verankert, da er in miteinander verknüpften Anwendungen manifestiert wäre. Letztere verweisen durch ihre inhaltlichen Überlappungen zwischen Forschung und Wissenschaftsverwaltung, Förderung und Personalentwicklung aufeinander. So lassen sich die Bereiche nicht mehr klar voneinander trennen. Diese Überlappung schafft eine gewisse Redundanz innerhalb der Produktpalette (z. B. zwischen Pure und Interfolio), so dass mehrere Konstellationen von Elsevier-Anwendungen hinreichend wären, die „critical junction“ des Drei-Phasen-Modells (Schreyögg & Sydow, 2011) zu beschleunigen.

Zur organisatorischen Abhängigkeit käme erschwerend hinzu, dass Elsevier selbst diejenigen Metriken mitgeneriert, aufgrund derer Elseviers Anwendungen Empfehlungen für Entscheidungen erstellen. Das darin implizite Potenzial einer Institution, ihre eigene Stellung in Rankings zu verbessern, bringt Erpressbarkeit mit sich, eben diese Anwendungen (weiterhin) zu nutzen. Elsevier gibt Beispiele für Ranking-„Aufstiege“ derjenigen Institutionen, die ihre Dienste verwenden. Rankings üben Druck

---

**111** “The vertical integration of both firms have the effect of influencing university decision making, with the adoption of Elsevier products, especially Bepress, Pure, Plum, SciVal influencing university decision making through their promotion of data analytics [...]” (Posada & Chen, 2018, 35)

auf den gesamten Wissenschaftsbetrieb aus, aber die Aussicht auf schnelle Erfolge durch Datenanalyse ist für Institutionen im globalen Süden bzw. ressourcenarme Universitäten besonders verlockend:

Elsevier showcases how Arizona State University (ASU) identifying opioids research as a priority collaboration topic internally through the use of the Prominence indicator [...]. The tools promise to bring algorithmic-based analytics to universities in an integrated package. Pure is promoted as being able to suggest collaboration partners [...] while the algorithmically-generated grant recommendations of funding bodies can be directly imported into the Pure workflow [...]. According to the Pure website, the tool ‘facilitates an evidence-based execution of strategy to unlock your full research potential, leading to increased funding, improved international collaboration and greater visibility.’ [...] Elsevier sees universities in other parts of the world as ripe for their marketing. They have promoted its Fingerprint Engine, a text-mining algorithm that generates smart keywords, in their products Pure and Expert Lookup [...], to National Cheng Kung University (NCKU) in Taiwan. Their promotional material shows how NCKU uses the algorithmically-generated keyword feature in Pure to search for international collaborators and to increase international visibility [...]. In another example, Elsevier showcases how Covenant University (Nigeria) uses Scopus and Scival to guide their research decisions [...]. In Egypt, Elsevier took part in the creation of a national rankings committee to assist universities in the country with visibility in world rankings. (Chen & Chan, 2021, 14-16)

Der Wissensvorsprung wird auch auf eine andere Weise erzielt und aufrechterhalten, die eine technische Fortsetzung der bereits beschriebenen Zugangsformen zu Informationen darstellt. Damit ist umfassendes Tracking gemeint, das auch diejenigen Personen und Institutionen erfasst, die keine der genannten Anwendungen nutzen. Die bisher beschriebenen Abhängigkeitseffekte und Asymmetrien werden so auf einer tieferen Stufe fundiert. Diese im wissenschaftlichen Alltag ubiquitäre Problematik bringt Brembs scharf auf den Punkt: „Nutzen Sie beispielsweise Scopus oder Web of Science für die Literaturrecherche, Mendeley zum Zitieren, Authorea zum Schreiben, FigShare, um Ihre Daten zugänglich zu machen, oder LabGuru für Ihr Labormanagement? Vielleicht nutzt ja Ih-

re Uni-Leitung Plum Analytics, Pure oder Elements, um die Forschungsleistung Ihrer Fakultät mit der von anderen Fakultäten zu vergleichen? Alle diese Werkzeuge sind mittlerweile Überwachungstools für die Data-Analytics-Konzerne, die sich früher ‚Verlage‘ nannten.“ (Brembs, 2023a) Dieser Themenkomplex wird im nächsten Kapitel besprochen. In diesem Zusammenhang treten andere Geschäftsfelder und -praktiken des RELX-Konzerns in den Vordergrund.



4.

## **Lock-In II Institution – Tracking**

## 4.1 Definition und Einbettung des Trackings in die Werbe- und Analytikindustrie

Dieser Teil beinhaltet technische Details, die in den Bereich der Informatik hineinragen. Meine Argumentation braucht jedoch nur ein sehr generelles Verständnis der Vorgänge. Bei der Darstellung beziehe ich mich für die Anschaulichkeit – soweit möglich – auf eine Beispielnutzerin *N*. Zunächst beschreibe ich einige wenige Trackingformen, den Zusammenhang zwischen Tracking und der Werbe-Analytikindustrie und diskutiere die Rolle und Wirkung des Tracking in der Wissenschaft. Vorab sollte geklärt werden, was Tracking im Gegensatz zu anderen Formen der Handhabung von Daten von und über *N* charakterisiert, wofür ich die Definition von Binns (2022) verwende: “‘Tracking’ is the collection of data about an individual’s activity across multiple distinct contexts and the retention, use, or sharing of data derived from that activity outside the context in which it occurred.” Tracking hat in diesem Verständnis folgende Merkmale:

1. Datensammlung über die Aktivitäten einer Person (unabhängig davon, ob es sich um ausschließlich digitale Aktivitäten handelt),
2. über mehrere, voneinander getrennte Kontexte hinweg,
3. Handlungen mit diesen Daten in Form von Zurückbehalten, Gebrauch oder Teilen außerhalb des Kontextes der Datenentstehung.

Diese Definition möchte ich um das Merkmal der Asymmetrie ergänzen. Dadurch soll eingefangen werden, dass es sich zum einen um eine einseitige Ausübung von Macht und Kontrolle *über* eine Person handelt, auch wenn es automatisierte Vorgänge sind, die große Personengruppen betreffen. Zum anderen ist Tracking asymmetrisch, weil eine Person im Normalfall nicht im vollen Umfang wissen oder in Erfahrung bringen kann, welche Informationen über sie von wem zusammengetragen, wie

verarbeitet, wem gegeben werden und zu welchen weiteren Zwecken künftig verwendet werden könnten.<sup>112</sup> In den meisten Fällen haben Nutzer:innen eher eine vage Ahnung darüber, dass ihre Aktivitäten im Netz dem Interesse Dritter ausgesetzt sind, z. B. durch das sogenannte Cookie-Banner im Browser. Durchschnittliche Nutzer:innen können von *einzelnen* Formen der Datensammlung oder -verarbeitung wissen. Aber welche Formen des Trackings ihnen ubiquitär begegnen, wie invasiv und weitreichend die Eingriffe sind und welche Konsequenzen dies für ihre Privatsphäre und selbstbestimmtes Handeln im digitalen (und zunehmend auch analogen) Raum hat, sind Elemente von Spezialwissen. Selbst dieses Spezialwissen bleibt in seiner Verallgemeinerung der technischen Komplexität notwendigerweise unvollständig. Dieses weitere Merkmal schließe ich an die obige Definition an als:

4. diese Handlungen sind asymmetrisch, weil sie für Personen im digitalen Raum einseitig *unausweichlich*; in ihrem Ausmaß, Zusammenhang und in Bezug auf aktuelle wie künftige Konsequenzen für Personen *nicht überblickbar* sind, insbesondere in Bezug auf die Machtposition der trackenden Seite gegenüber den getrackten Personen.

So verstanden ist Tracking nicht gleichzusetzen mit einer einzigen Methode oder Technologie und prinzipiell offen für verändernde Formen der Umsetzung: „Potenzial für Tracking bieten alle technischen Ansätze, die dritten Parteien über längere Zeitintervalle Informationen liefern, anhand derer man Verbraucher bzw. die von ihnen verwendeten IT-Geräte wiedererkennen und voneinander unterscheiden kann“ (Schneider *et al.*, 2014, 36).

Die hier besprochenen Beispiele sind also weder in ihrer Anzahl noch in der technischen Form erschöpfend. Die technischen Voraussetzungen wandeln sich und neue Formen sind zu erwarten. Wichtig ist festzuhalten, welche Art von Nachweis über Tracking von Verbraucherseite bzw. der IT-Forensik erbracht werden kann, nämlich eine *indirekte*. Fest-

---

<sup>112</sup> Insbesondere bei RTB, auf das ich weiter unten eingehe, ist es nahezu ausgeschlossen, die Verteilung nachzuvollziehen.

gestellt wird, dass es Bemühungen gibt, bestimmte Informationen über Nutzer:innen zu erlangen.<sup>113</sup> Was mit diesen Informationen weiter geschieht, kann nur aus Indizien geschlussfolgert werden oder wird durch Enthüllungen (z.B. Microtargeting durch Cambridge Analytica<sup>114</sup>, die Berichte von Cody Hanson<sup>115</sup> oder Wolfie Christl<sup>116</sup>) publik. Diese Enthüllungen stehen jedoch als Einzelfälle einer allgegenwärtigen, extrem lukrativen<sup>117</sup> und größtenteils unsichtbaren Praxis entgegen. Erkenntnisse über Tracking stehen also immer im Kontext weiterer Handlungen, i) deren Datenschutz- und juristischer Status zumindest fragwürdig ist, ii) die vielfältige kommerzielle und politische Interessen Dritter verfolgen und iii) intransparent und asymmetrisch sind.

Vorab die wichtigsten Begriffsklärungen in einer vereinfachten Darstellung: Wenn von 'First Party' die Rede ist, kann das in etwa analog zur realen Interaktion zwischen der Person N (mit ihrem Endgerät) und der besuchten Website auf der zugehörigen Domain verstanden werden. Indirekt und meist ohne Ns Wissen an dieser N-Website-Interaktion beteiligte Parteien sind ‚Drittparteien‘, ‚Third Party‘ bzw. ‚Parties‘. Wenn N die Website <https://stoptrackingscience.eu/> aufruft, könnte N ihre besonders wertvollen ‚First-Party-Daten‘ auf der Seite hinterlassen, wenn sie zum Beispiel ihre Email-Adresse angibt oder die Petition signiert. Die Seite <https://stoptrackingscience.eu/> ist in diesem Beispiel die ‚First-Party-Website‘ (und ‚Host‘ bzw. ‚Domain‘). Die von ihr gesetzten Cookies (die zum Beispiel die Interaktion von N mit der Website analysieren) bzw. genutzte Ressourcen wären ‚First-Party-Cookies‘ bzw. ‚First-Party-Ressourcen‘, wenn sie von derselben Domain (oder Subdomain) gehostet werden und ‚Third-Party-Ressourcen‘, wenn sie von einer anderen Domain stammen.

---

**113** „Von außen betrachtet lässt sich nur beobachten, ob dritte Parteien durch die Anwendung von Trackingmethoden versuchen, Verbraucher wiederzuerkennen. Für welche Zwecke diese Wiedererkennungsmöglichkeit ausgenutzt wird, erkennt man aus der bloßen Erkenntnis, dass Trackingmethoden angewendet werden, nicht.“ (Schneider et al., 2014, 39)

**114** Bildung (2019).

**115** Hanson (2019).

**116** Christl (2021); Christl & Toner (2024a); Ryan & Christl (n.d.).

**117** Statista (n.d.); <https://de.statista.com/themen/4554/programmatic-advertising/>. [zuletzt geprüft am 25.07.2024]

Würde N danach zur Website <https://www.springernature.com/de> wechseln, würde zum Beispiel ein Tracker von [www.google.com](https://www.google.com), in dem Fall als Third Party ein Third-Party-Cookie auf Ns Gerät speichern wollen.<sup>118</sup>

Vereinfachend gesprochen betreibt ein First-Party-Cookie ‚Same-Site-Tracking‘<sup>119</sup>, weil er von derselben Domain oder Subdomain gesetzt und ausgelesen wird. Ein Third-Party-Cookie betreibt entsprechend ‚Cross-Site-Tracking‘ (vgl. Abbildungen 12 und 13 im Anhang), weil er Ns Verhalten über mehrere Websites hinweg verfolgt (nämlich überall, wo die Domain hinter dem fraglichen Cookie ihre Cookies platziert hat, die N dann jeweils wiedererkennen).<sup>120</sup>

Würde im späteren Verlauf die Third Party (Google oder zahlreiche andere) Informationen über N bekommen, die dann erschlossen, kombiniert, aggregiert und weiterverwendet werden, wären es ‚Third-Party-Daten‘ über N. Die meisten im Internet kursierenden Daten sind Third-Party-Daten. S. Abbildung 14 im Anhang für eine Übersicht der Daten durch den Datenbroker Hum, der sich auf den Wissenschaftssektor konzentriert (wobei Hum zusätzlich zwischen Zero und First Party unterscheidet).

---

**118** Diese Angabe bezieht sich auf den 20. Juni 2024.

**119** “If a cookie is stored with a first-party resource as host, it enables Same-Site tracking as the domain of the cookie will refer to the website the user is visiting. First-party cookies are practical for tracking repeat visits to the same site. Such cookies can also help collect analytics and understand user engagement. They can be set directly by the visited website, or by a third-party service running in the context of the visited website. First-party cookies are commonly known to be deployed for analytics purposes. However, in the last few years, trackers are increasingly relying on first-party cookies for tracking purposes [...]” (Fouad et al., 2024, 2)

**120** “If a cookie is stored with a third-party resource as host, it enables Cross-Site tracking as the domain and path of the cookie will refer to an external domain that the user may not be aware of. Third-party cookies are used to track the browsing activity of a user on all the sites where this resource is present.” (Fouad et al., 2024, 2)

## 4.1.1 Cookie

Das heutige Tracking lässt sich in seinen Anfängen technisch auf das Cookie zurückverfolgen, welches zunächst einen funktionalen und nicht primär monetären Hintergrund hatte:

As the web developed, technology was invented to support tracking. Web developers wanted a way to be able to remember users. In 1994 John Montully, a developer at Netscape, an early web browser, created the ‘cookie’; a way to store a small amount of information associated with a user. (Binns, 2022, 5)

Ein Cookie ist also ein kleiner N-spezifischer Code<sup>121</sup>, den die Host-Domain der von N aufgerufenen Seite beim HTTP(S)-Aufruf des Browsers oder systematisch durch Javascript an Ns Endgerät schickt. Wenn Ns Gerät das Cookie nicht blockiert, wird es dort gespeichert. Die Cookie-Datei enthält verschiedene Felder, u. a. ein Ablaufdatum, nach welchem das Cookie nicht mehr gesendet werden soll und einen URL-Pfad innerhalb der spezifizierten Domain zur Einschränkung des Sendebereichs. Bei Ns erneutem Aufruf dieser Domain bzw. Subdomain wird das Cookie im Browser automatisch mitgesendet und vom Host ausgelesen. Dieser Vorgang unterscheidet sich nicht bei First-Party-Cookies und Third-Party-Cookies, denn Cookies werden in beiden Fällen auf Ns Endgerät gespeichert und anschließend ausgelesen. Der relevante Unterschied ist, wer die verschiedenen Cookies auf Ns Gerät auslesen wird – der Host der aufgerufenen Seite, also die First Party oder verschiedene Third Parties, also Drittparteien (Binns, 2022, 7).<sup>122</sup>

Eine Bemerkung zum Auslesen von Cookies und Cross-Site-Tracking: Dass ein Cookie nur von der Domain ausgelesen werden kann, von der es gesetzt wurde, wird durch die Same-Origin-Policy verlangt. “An origin is defined by the scheme, host, and port of a URL. Generally speaking, do-

<sup>121</sup> Genauer ist es ein Triplet aus „host, key, value“, bei dem der Host sich auf die Domain bezieht, welche Zugang zum Cookie hat (Fouad et al., 2024, 4) und die das Cookie dann auslesen kann.

<sup>122</sup> Für eine grafische Darstellung der Interaktion vgl. die Abbildung 17 im Anhang.

cuments retrieved from distinct origins are isolated from each other.” (w3.org, n.d.) Um trotzdem mehreren Interessierten Zugang zu einem Cookie zu gewähren, gibt es das sogenannte Cookie Syncing (Piwik, n.d.). Dabei handelt es sich um einen Prozess zwischen den trackenden Parteien, die den N-spezifischen Code, den „unique user identifier“, untereinander abgleichen und so ihre jeweiligen Informationen über N austauschen bzw. synchronisieren. Das heißt also, dass unüberblickbar viele Informationen über N auch völlig ohne Ns Web-Aktivität trackenden Drittparteien und deren Kundschaft vorliegen, eben durch die Synchronisation ihrer gesammelten Trackingdaten. Zu dieser Kundschaft zählen AdTech-Firmen und Data Broker wie LiveRamp, ehemals Acxiom – nach eigener Bezeichnung eine „Datenkollaborations Plattform“<sup>123</sup> mit dem „Cookie Sync Tag“<sup>124</sup> (Christl & Toner, 2024a, 31).

In den 2000er Jahren verlagerte sich einerseits ein großer Teil der Dienste und Ressourcen der First-Party-Websites auf die Drittanbieter, die auf der einen Seite technisch-funktionale Kernaufgaben für sehr viele Websites übernahmen:

This included everything from widgets for embedding and serving video content, user analytics, the loading of special fonts, comment functions below news articles, buttons which allowed content to be shared via the user’s account on a social network, anti-spam measures which separate bots from genuine human users, and much else besides (Binns, 2022, 9).

Auf der anderen Seite verknüpften Drittanbieter ihre Inhalte und Dienste mit Analyse-Technologien und domainübergreifender Datensammlung: “From the distribution of third-party advertiser code on first-party websites, it was a small step to the use of ‘third-party’ cookies.” (Binns, 2022, 7) Durch die Third-Party-Cookies konnten Adtech-Unternehmen über das

<sup>123</sup> <https://liveramp.de> und <https://liveramp.com/>. [beide zuletzt geprüft am 25.07.2024]

<sup>124</sup> Siehe die Liste von LiveRamps “Sync Partners”, Figure 14 (Christl & Toner, 2024a, 32) mit u. a.: Google Cloud, Google Analytics, Amazon, Adobe, Facebook, BlueKai, Xandr, Microsoft Bing.

ursprüngliche Ziel, die Messung der Sichtung ihrer Anzeigen hinausgehen, indem sie “*also track those users from one site to the next*” (ebd.).<sup>125</sup>

Eine weitere Form des Trackings, die hier nicht eigens behandelt wird, entstand aus dieser Veränderung der Ressourcen-Verteilung, nämlich das Tracken durch Webfonts. Fonts von Google, die auf sehr vielen Websites eingesetzt werden, werden nicht auf der Domain dieser Websites gehostet, sondern bei Google. Jedes Aufrufen der Website durch N führt zum Abrufen der Fonts bei Google. Dabei werden Informationen über die aufrufende Seite, unsere Nutzerin N, an Google gesendet.

## 4.1.2 Post-Cookie-Tracking

Mit Googles Abkehr von Third-Party-Cookies, die nach mehreren Verzögerungen nun für 2025 geplant ist, sind neue Formen des Tracking zu erwarten und werden auch schon teilweise implementiert. Eine davon ist Googles Folgekonzept zu Cookies, die “Privacy Sandbox” bzw. Sandbox API. Dieses Verfahren, bei dem der Browser (Chrome) und nicht mehr die verschiedenen Third-Party-Cookies Tracking und Analysen durchführt, präsentiert Google als datenschutzbewusste Alternative. Scharfe Kritik lässt vermuten, dass es sich dabei um ‘privacy-washing’ handelt, bei dem Nutzer:innen aufgrund irreführender Darstellung des Features als einer Datenschutz-Funktion “ad privacy feature” in die Aktivierung einwilligen (noyb, 2024). Das Europäische Zentrum für digitale Rechte (noyb) hat deshalb am 13. Juni 2024 eine Beschwerde gemäß Artikel 80(1) DSGVO bei der österreichischen Datenschutzbehörde eingereicht (ebd.).

Welche Konsequenzen für Verbraucher:innen aus dieser und anderen Nachfolgetechnologien zum Third-Party-Cookie zu erwarten sind, ist noch nicht absehbar. Zu befürchten sind zumindest anfängliche Nachteile für den Datenschutz, da die Anpassungen der Werbeindustrie entsprechenden Anpassungen der Gesetzgebung meist vorausgehen. Eine solche Neuentwicklung, die sehr negative Prognosen für die Transparenz

---

<sup>125</sup> “In the early 2000s, no single tracker was present on more than 10% of top sites; but by 2016, google-analytics.com was present on nearly a third of top sites. As mentioned above, it is now present on a majority of websites.” Binns (2022, 11)

und Regulierbarkeit hervorruft (Fouad *et al.*, 2024), ist das seit kurzer Zeit bekannte Server-Side-Tracking (SST). Es verändert das Verhältnis zwischen First-Party- und Third-Party-Tracking dahingehend, dass die aktuelle Gesetzeslage – sowohl national mit § 25 TTDSG Telekommunikation-Telemedien-Datenschutz-Gesetz als auch auf EU-Ebene mit dem Art. 5 ePrivacy-Richtlinie<sup>126</sup> –, die auf dieser Unterscheidung beruht, das Problem nicht erfassen kann. SST kann sowohl als Third-Party-Auftrag durchgeführt werden als auch durch die Einbettung als First-Party-Subdomain (Fouad *et al.*, 2024, 2) bzw. indem die Domain selbst zum Tracker wird, was die bis dahin hilfreiche Unterscheidung zusätzlich unterläuft.

Zudem entsteht bei SST ein Transparenzproblem, da es Tracker ‚hinter‘ dem jeweiligen Zielserv­er versteckt: “[K]nown methods to identify more traditional tracking are not well suited to detect SST. The reception of a user ID does not necessarily imply that the server will share it directly with different partners.” (Fouad *et al.*, 2024, 1) Statt der z. B. durch Ad-blocker und tools wie uBlock Origin eruierbaren Präsenzen von Trackern auf einer Website übernimmt nun der Server die Trackingaktivität und verschleiert diese für Nutzer:innen wie auch für die professionelle bzw. IT-forensische Überprüfung:

Instead of the client directly reporting its tracking data to different companies around the world, it is the role of the SST server to do it. This approach is currently hampering online privacy as it transforms what used to be very clear tracking requests performed by the browser on the client side into masked and hidden ones on the server side. For users, it becomes complicated to protect themselves as their favorite blocking tools are not adapted to this hidden form of tracking. For regulators, it is simply a nightmare as they cannot see clearly which companies are performing tracking in the background. (Fouad *et al.*, 2024, 1) (Vgl. Abbildung 16 im Anhang)

Diese Tracking-Form ist noch wenig untersucht und ihre Implikationen für das Wissenschaftstracking sind noch nicht absehbar.

---

<sup>126</sup> Vgl. dazu das Kapitel Digitale Souveränität.

Dem nächsten Abschnitt der Anwendung der Trackingformen auf die Wissenschaft vorweggreifend sei hier der Abschlussbericht des Ausschusses für wissenschaftliche Bibliotheken und Informationssysteme der DFG (Altschaffel *et al.*, 2024) erwähnt. Der Bericht deckt weitere technische Aspekte auf, die Tracking besonders tiefgreifend und invasiv machen. Das ist zum einen die versteckte Weiterleitung beim Aufrufen einer Adresse, die automatisch aktiviert wird und wodurch weitere Datenauswertungen stattfinden. Zum anderen geht es um das sogenannte CNAME-Hiding, das die Aufdeckung und Nachverfolgung des Trackings erschwert.<sup>127</sup>

Fouad *et al.* (2024) nennen neben SST noch weitere Post-Cookie-Ansätze der Identifizierung und Profilerstellung wie „contextual advertising“, bei dem Kontextsignale mit First-Party-Daten des jeweiligen Advertisers und Informationen aus eigenen Datenbanken Third-Party-Cookies ersetzen, „identity graphs“<sup>128</sup>, oder „interest-based advertising“<sup>129</sup>. Daneben beschreibt Freiberg (2022) noch den Kohorten-Ansatz von Permutive, bei dem die „Profilerstellung anhand von First-Party-Daten mit anschließender lokaler Speicherung im Browser“ erfolgt. Weil das Profil durch Interaktionen mit dem Browser entsteht und lokal (verschlüsselt) gespeichert wird – im Gegensatz zu einer Cloud – „bezeichnet Permutive diesen Ansatz als dezentrales ‘Edge Computing’ [...]“ (ebd.).<sup>130</sup> Dies sind jedoch nur skizzenhaft genannte weitere Trackingformen.

---

**127** „Ein CNAME ist ein alternativer Name für einen Adresseintrag im Internet – ein Dienst, der die eigentliche Adresse xyz.de hätte, könnte auch unter abc.de dargestellt werden, welches das Tracking aus Sicht des Benutzers verbergen würde (dies ist ein fiktives Beispiel, es bezieht sich jedoch auf ein konkretes Beispiel, das am 07.10.2023 dokumentiert wurde).“ (Altschaffel *et al.*, 2024, 26)

**128** [Beispiele von Amazon <https://aws.amazon.com/neptune/identity-graphs-on-aws/>, und LiveRamp <https://liveramp.com/identity-resolution/>. [beide zuletzt geprüft am 25.07.2024]

**129** [<https://webkit.org/blog/11529/introducing-private-click-measurement-pcm/>, <https://blog.mozilla.org/en/mozilla/privacy-preserving-attribution-foradvertising/>, <https://blog.google/products/chrome/get-know-new-topics-api-privacy-sandbox/>] [zuletzt geprüft am 25.07.2024]

**130** In Freibergs Test mit der Website von SpringerNature zeigte sich, „dass in der IndexedDB, einer seit HTML5 gegebenen Option zur lokalen Speicherung von strukturierten Daten, eine mit einem Hashwert benannte Datenstruktur mit dem Ursprung ‚nature.com‘ abgelegt wurde, die nach Interaktionen wie Scrollen oder

### 4.1.3 Fingerprinting

Ein Ansatz zur Identifizierung, der nicht über Cookies funktioniert und so nutzer:innenseitige Versuche der Anonymitätswahrung durch z. B. Ad-blocker umgeht, ist Fingerprinting. Es existiert als Device (Browser- oder auch Canvas-) und als Personen-Variante. Bei der ersten Variante werden über verschiedene Schnittstellen „(JavaScript, HTML-CanvasElemente, WebGL, Web Audio API) browserspezifische Merkmale abgefragt, deren individuelle Kombination so eindeutig wie ein Fingerabdruck sein kann [...]“ (Freiberg, 2022, 98). Die spezifischen Browser-Einstellungen und Interaktionen zwischen Soft- und Hardware können als individuierbare Grafiken dargestellt werden, die einem Fingerabdruck gleichen. Solche Fingerabdrücke werden in speziellen Datenbanken, den Fingerprinting-Bibliotheken, von Data Brokern angeboten. Freiberg beschreibt den Vorgang einer Grafikzeichnung bzw. Audio-Variante des Fingerprinting im Detail:

JavaScript liefert unter anderem Informationen über installierte Schriftarten, die Bildschirmauflösung und zahlreiche Browser-Einstellungen. Eine weitere Methode des Fingerprinting ist die Erstellung von verborgenen Grafiken, die je nach Software- und Hardware-Ausstattung des Geräts kleine Unterschiede aufweisen. Das in HTML 5 eingeführte Canvas-Element zum Zeichnen von 2D-Grafiken auf Webseiten wird zum Beispiel für das Fingerprinting missbraucht, indem unbemerkt auf einer nicht-sichtbaren Zeichenoberfläche (Canvas) Text- und Bildelementen erstellt werden [...]. Je nach Hardware, Betriebssystem und Browser fällt diese Grafik leicht anders aus. In der Folge wird die Grafik ebenfalls unbemerkt heruntergeladen, aus ihr ein individueller Hashwert berechnet, der schließlich zur Speicherung und zum späteren Abgleich in der Fin-

---

dem Anklicken von Links zunehmend anwächst. Dieses Datenbankobjekt enthält technisch gesehen eine 'events' benannte Reihe von Ganzzahlen, die wahrscheinlich die Interaktionen in verschlüsselter und daher anonymisierter Form enthält. Diese Bewegungs- und Interessenprofile werden sogenannten Kohorten zugewiesen, also Gruppen mit übereinstimmenden Profilmerkmalen, für die zielgruppenspezifische Inhalte bereitgestellt werden können. Welche Mitglieder der Kohorte aktuell auf der analysierten Seite unterwegs sind, kann laut Permutive übrigens in Echtzeit abgefragt werden.“ (Freiberg, 2022, 102)

gerabdruck-Datenbank des Anbieters gesichert wird. Mit der WebGL-Programmierschnittstelle zur Darstellung von 3D-Grafiken können noch präzisere Fingerabdrücke generiert werden, da die Berechnung von 3D-Grafiken stärker von der jeweiligen Grafik-Hardware geprägt wird. Auch Hersteller und Modell der Hardware können über die WebGL-Schnittstelle abgefragt werden. Relativ neu ist das Fingerprinting über die WebAudio-API. Über diese Audio-Schnittstelle können nicht nur Informationen über die Audio-Hardware ausgelesen, sondern auch ein Oszillator generiert werden, dessen je nach Hardware anders ausfallende, stumme Ton-Wiedergabe im Hintergrund anschließend abgetastet und zur Generierung eines Hashwerts herangezogen wird. (Freiberg, 2022, 98)

Wie Schneider *et al.* (2014) bereits vor 10 Jahren festhalten, ist es neben dem Geräte-Fingerprinting auch möglich, Personen-Fingerprinting zu erstellen, um Nutzer:innen auch an verschiedenen Geräten mit unterschiedlichen Konfigurationen geräteübergreifend zu identifizieren. So können mit „Fingerprinting Bibliotheken wie z.B. von BlueCava[...], Iovation[...] oder ThreatMetrix [...] nochmal deutlich bessere Trennschärfen erreichen [werden]“ (Schneider *et al.*, 2014, 50). Laut Schneider *et al.* würden „Fingerprints sehr nahe an eindeutige Identifikationen herankomm[en]“ (ebd.). ThreatMetrix und aktuelle Formen der Identifikation werden weiter unten wieder aufgegriffen.

#### 4.1.4 Marketing, Data Brokering, Verknüpfungsdienstleister

Die Entwicklung der Technologien und Interaktionsmöglichkeiten im Web 2.0 ist stark von der Entwicklung der webbasierten Werbeindustrie geprägt. Die Ausdifferenzierung und Ausweitung von Werbestrategien begann analog zu Printmedien. Eine Werbeanzeige wurde in einem vermieteten Slot auf der Website für einen festgelegten Zeitraum geschaltet, sichtbar für alle Besucher:innen dieser Website. Dieses Vorgehen wurde mit dem Aufkommen von Möglichkeiten weiterentwickelt, immer mehr domainübergreifende Informationen über das Verhalten von Nutzer:innen zu sammeln, zu Profilen zu clustern (Keegan & Eastwood, 2023) und an diese Nutzer:innen zu binden. So wurden letztere wiedererkennbar

und an ihren Geräten nachverfolgbar gemacht. Damit wurde es auch möglich, Werbeanzeigen gezielter auf angestrebte Kund:innen auszurichten und die Anzeigen auch direkt auf der aufgerufenen Seite anzuzeigen. Dort befindet sich nunmehr nur ein Platzhalter, während die Werbeanzeige selbst vom Server des jeweiligen Werbetreibenden geladen wird. Diese Strategie, das Microtargeting, und ihre politischen Dimensionen wurden vor allem im Zusammenhang der U.S.-Wahlen 2016 und Cambridge Analytica öffentlich bekannt.

Elaboriertere technische Möglichkeiten des Cross-Site-Tracking, eine immer engmaschigere Profilbildung und Überwachung der Aktivitäten von Nutzer:innen liefen parallel zur zunehmend komplexeren Beziehung zwischen First und Third Party:

Among the third-party resources commonly referenced by first-parties are those belonging to the largest web giants [...]. This means that users of their services can be tracked as they move around the web, often even when not logged in to the respective service. This has further enriched the already rich profiles these platforms had built on their users activities beyond their walled gardens. As a result, a social network might be able to infer sensitive private information about a user not because of what they do on that social network, but as a result of the websites they visit outside of it. These trackers are sometimes called ‘personal trackers’ because while they are ‘third-parties’ in a technical sense, the user may also have a personal account with them because they also operate as user-facing first-party services like Google or Facebook [...] (Gatti, 2020, 10)

Unabhängig davon, wie adäquat die Unterscheidung zwischen First und Third Party künftig sein wird – bereits an diesen wenigen Beispielen zeigt sich, dass für durchschnittliche Nutzer:innen eine nahezu *unausweichliche* Abdeckung ihrer Internet-Aktivitäten durch eine Kombination verschiedener Tracking-Formen besteht. Ob direkt durch die besuchte Domain oder durch eingebundene Dritte, die Datensammlung und die zahlreichen Arten der Datenverwertung gehen in sehr lukrative Geschäftsformen ein, die über reine Werbetätigkeit hinausgehen. Aber bereits der Markt für digitale Werbung ist beeindruckend: Investitionen in Internetwerbung beliefen sich für 2021 weltweit auf über 417 Mrd. US-Dollar (Statista, 2024), Tendenz steigend.

Ein Beispiel für sehr umfassendes, kombiniertes und invasives Tracking sind Verknüpfungsdienstleister, deren Produkte für die AdTech-Industrie, aber auch für weitere Zwecke genutzt werden können. Ihr primäres Geschäft (dessen Konformität mit der Gesetzgebung der Länder, in denen sie operieren, unklar ist) besteht in der Erstellung von möglichst vollständigen Profilen und eine nahtlose Beobachtung von ganzen Bevölkerungen:

Companies have started to combine and link data from the web and smartphones with the customer data and offline information that they have been amassing for decades. Many businesses try to record and measure every interaction with a consumer, including on websites, platforms, and devices they do not control themselves. They can seamlessly collect rich data about their customers and others in real-time, enhance it with information from third parties, and utilize the enriched profiles for digital advertising, which is often based on real-time auctions of consumer profiles. The lawfulness of many of these data practices under the GDPR is disputed and subject to ongoing debate and litigation. (Christl & Toner, 2024b, Summary)

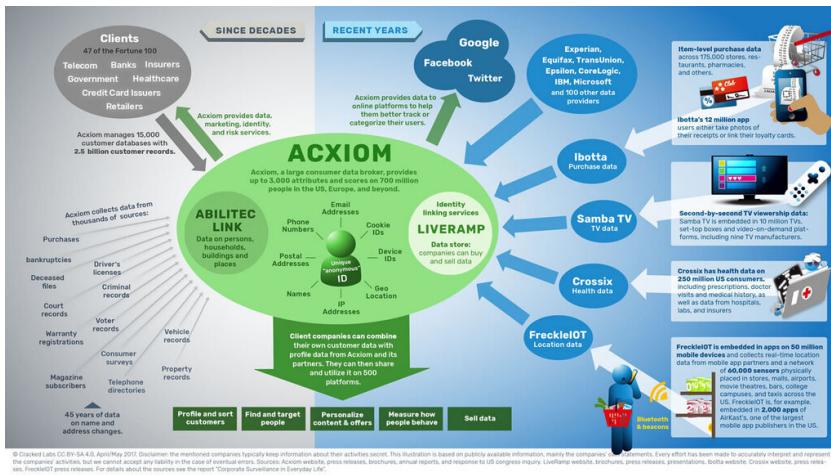
LiveRamp, ein „consumer data broker“ verkauft auf seinem „data marketplace“ weltweit Daten über 700 Millionen Verbraucher:innen von 150 Datenlieferanten:

LiveRamp operates a massive identity surveillance system that assigns every person a proprietary identifier, which is tied to identifying attributes such as names, postal addresses, email addresses, phone numbers and digital IDs referring to browsers, smartphones and other devices. Its AbiliTec and RampID systems maintain and constantly update comprehensive identity records about whole populations: the address where they live, the devices they use, and people they share a household with. (Christl & Toner, 2024b, Summary)

Die Gefahr, die von einem Privatkonzern mit eigenen bevölkerungsweiten Identitätsdatenregistern mit Personen- und Haushalts-IDs ausgeht, ist enorm, wie auch Christl & Toner (2024b) und Reuter (2024) kritisieren:

LiveRamp's identity graph systems can be considered private population registers, and their identity databases and proprietary identifiers facilitate the exchange of personal data across databases and companies. Many businesses in the digital marketing industry utilize LiveRamp's identity surveillance technology to recognize, track, follow, profile and target people across the digital world and trade profile information about them. As of 2023, LiveRamp operates in many countries across the planet including the UK, France, Germany, Belgium, Spain, Italy, Poland and Romania. (Christl & Toner, 2024b, Summary)

Es ist wichtig zu verstehen, dass es sich nicht nur um Informationen über das Online-Verhalten und digital gesammelte Aktivitäten oder Eigenschaften von Personen handelt. Es geht auch um die Verknüpfung von bereits über lange Zeiträume gesammelten weiteren analogen Informationen mit den digitalen Daten, wie die Abbildung 9 zeigt.



**ABBILDUNG 9:** Acxiom and its data providers, partners, and clients.

Diese Bestände werden zusammengebracht und durch die RampID zu einem Identitätsgraphen verknüpft, der laufend aktualisiert wird, auch über Wechsel von Geräten, Konten oder Wohnorten hinweg:

The AbiliTec identity graph links different ‘offline’ identifiers to each other, including names, postal addresses, email addresses and phone numbers [...] The RampID system links digital identifiers to a person’s AbiliTec record, including mobile device IDs, cookie IDs, connected TV IDs and other proprietary IDs [...] LiveRamp claims to have data about 14 billion devices globally. (Christl & Toner, 2024b, Summary)

LiveRamp beschreibe seine RampID und das Verfahren der Anonymisierung als “secure, one-way deidentification” zwischen der „Offline“-Identität (AbiliTec ID) und der RampID mit „digital identifiers“, doch eine hinreichende Anonymisierung sei damit keineswegs gegeben<sup>131</sup>, so Christls und Toners Studie:

[T]his still potentially allows thousands of companies to send names, postal addresses or email addresses to LiveRamp, receive the corresponding RampIDs, and then match them with records that are also linked to RampIDs from many other companies. (Ebd.)

Google beschreibt die RampID als „Verbindungsschlüssel“ zwischen den google-eigenen Daten mit den Werbedaten (Reuter, 2024).

Dass solche Informationsbestände über EU-Bürger:innen jedem oben diskutierten Verständnis von digitaler Souveränität widersprechen und ein beträchtliches Missbrauchspotenzial für kommerzielle, politische und kriminelle Ziele bergen, bedarf an dieser Stelle keiner zusätzlichen Argumentation, sondern kann konstatiert werden.

## 4.1.5 RTB Broadcasting

Der heute dominierende Ansatz, Werbung und Microtargeting zusammenzubringen, erschien 2007 mit dem Programmatic Advertising, das

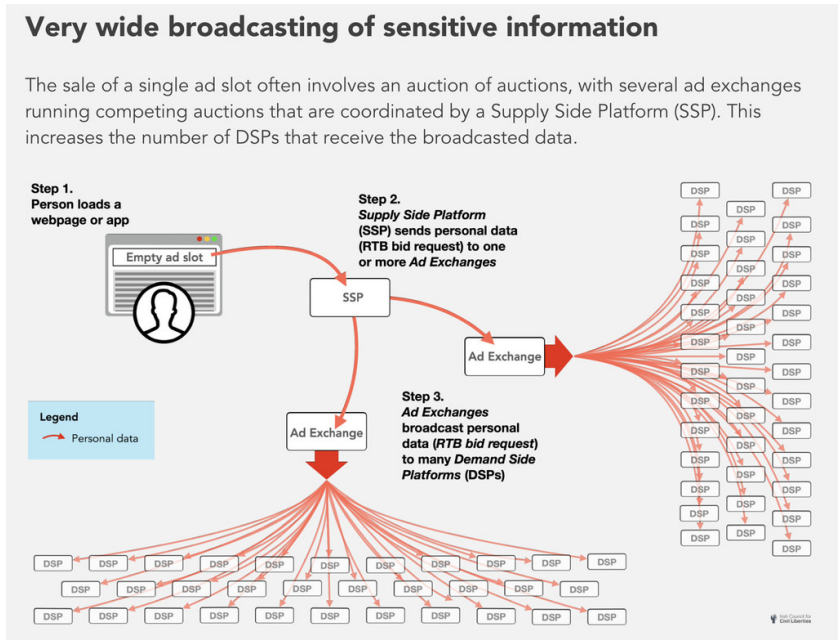
<sup>131</sup> “[T]he RampID system stores and processes personal data about the links and associations between ‘offline’ information (such as names, postal addresses, email addresses and phone numbers) and ‘online’ information (such as cookie IDs, mobile device IDs, connected TV IDs and custom IDs).” (Christl & Toner, 2024b, 27)

auf der Technologie des ‘Bidstreaming’, bzw. ‘Real-Time Bidding’ basiert. Hier gibt es weitere Zwischenschritte bzw. Intermediäre gegenüber dem früheren Modell. Ruft N eine Website auf (vgl. die Abbildung 10), verschickt eine ‘Supply Side Platform’ (SSP) Ns Profildaten an mindestens eine ‘ad exchange company’ (größter Anbieter dieses Dienstes ist Google). Letztere gibt diese Daten vollautomatisiert an eine unbekannt große Menge an ‘Demand Side Platforms’ (DSP) weiter – bei diesem Vorgang spricht man von ‘broadcasting’ bzw. ‘bidstreaming’ – die gegeneinander auf den freien Werbeplatz bieten. Die DSPs sind Werbetreibende, es können sich aber auch beliebige Akteur:innen mit weiteren Interessen als DSP registrieren und Zugang zu Daten erhalten. All das passiert vollautomatisiert, programmiert auf eine optimale Passung zwischen dem Werbeziel und dem Nutzerprofil, in der kurzen Zeit, die das Laden der aufgerufenen Seite benötigt. Diese Auktionen passieren in Echtzeit bzw. dauern unter einer Sekunde. So sieht N wahrscheinlich eine andere Anzeige, als ihr Büronachbar M, obwohl beide dieselbe Seite aufrufen – eine personalisierte Werbung bei jedem Seitenaufruf.

Dieser Ansatz ist äußerst einträglich: “RTB ‘Programmatic’ including display and video ads accounted for an estimated \$99bn in 2021, more even than search advertising (\$78.3bn)”, zitieren Ryan & Christl (n.d., FN1) den PwC IAB Internet Advertising Revenue Bericht (PwC, 2022, 21). Für die digitale Werbeindustrie ist der Ansatz durch eine neue Form der Kundenzentrierung attraktiv, die aber weitreichende und unübersichtliche Konsequenzen für die Sicherheit der Nutzer:innen hat:

This promised advertisers something quite unique; the ability to bid against each other for the opportunity to show a particular advert to a particular user based on a profile of that user. Within this tiny window, a surprisingly large number of actors are involved in a complex, automatically choreographed bidding war, involving the processing of lots of data from multiple sources. (Binns, 2022, 8)<sup>132</sup>

**132** Diese Dynamik führte dazu, dass alle Parteien zu dateninvasiveren Geschäftsmodellen übergingen: “As a result, users were not only being tracked directly via the third-party scripts belonging to advertising technology providers, but also by other third-parties who were providing other, non-advertising services to first-party websites. But before long, those other third-parties also took up the opportunity to become enmeshed in the digital advertising industry too, for instance by operating as data management platforms, supplementing demand-side platforms with



**ABBILDUNG 10:** Broadcasting personenbezogener Daten Ryan & Christl (n.d., 5).

Klar ist aber, dass RTB zur Verbreitung von sehr umfangreichen Profilen an eine nahezu nicht abschätzbare Menge an nahezu nicht nachverfolgbarer DSPs führt. Ryan und Christl weisen darauf hin, dass dies mindestens eine enorme Sicherheitslücke darstellt: “RTB data about the average person are broadcast to many entities, hundreds of times a day. There is no way to limit or know what happens to RTB data after they are broadcast.” (Ryan & Christl, n.d., 7) Die im letzten Abschnitt erwähnte RampID von LiveRamp dient als ein „universal identifier“ (Christl & Toner, 2024b,

additional personal data about individuals to better inform ad targeting and bidding strategies. Third-parties shared some of the profits of this data monetisation with the first-parties on whom they depended, and soon the dependency was reversed; many websites came to rely on these third-parties not only for functionality (analytics, content distribution, interactive elements), but also as their main revenue source.” (Gatti, 2020, 10)

Summary) für die AdTech-Industrie und speziell auch das RTB. Die Ramp ID wird über RTB übertragen und dient im digitalen Werbemarkt der Verknüpfung personenbezogener Daten “across a large number of publishers, advertisers and adtech intermediaries. This occurs 60 billion times a day, according to LiveRamp [...]” (Christl & Toner, 2024b, Summary).

Die Gefahren für die EU lassen sich analog auf die Wissenschaft anwenden und beinhalten bereits die Wissenschaft als eine der betroffenen Kategorien (vgl. die Abbildung 17 im Anhang).

## 4.1.6 Prädiktive Analytik und Risk Solutions

Ein wachsendes Einsatzfeld für Tracking und Datenhandel ist prädiktive Analytik. Dieses Verfahren ist nicht neu und wird in zahlreichen Disziplinen verwendet. Dabei handelt es sich um ein statistisches Verfahren, bei dem aus vorhandenen (personenbezogenen oder nicht-personenbezogenen) Daten umfangreiche Schlüsse auf unbekannte Eigenschaften, künftiges Verhalten von Personen oder generell künftige Ereignisse und Trends gezogen werden. Die genutzte Datenbank enthält historische Daten und wird durch neu hinzukommende und generierte Daten erweitert und modifiziert. Durch Aggregation von Einzeldaten und zunehmend auch maschinelles Lernen werden Modelle berechnet, die in Anwendung auf Einzelfälle Vorhersagen über künftige Handlungen machen oder Aussagen über *noch unbekannte* (Gruppen-)Eigenschaften treffen. Dies können auch sensible Informationen wie Gesundheitszustand, Ethnie oder Religionszugehörigkeit sein.<sup>133</sup> Kritische Stimmen warnen, dass mit prädi-

---

<sup>133</sup> “[E]asily accessible digital records of behavior, Facebook Likes, can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender” (Kosinski et al., 2013, 5802)

„Zwischen den zunächst wenig Aufschluss gebenden Hilfsdaten von [N] und sensiblen Informationen über [N] können mittels prädiktiver Analytik Korrelationen berechnet werden, weil die historischen Daten sämtlicher Personen [über die Daten vorliegen] so mannigfaltig sind, und neben vielen Hilfsdaten von vielen Personen auch einige sensible Informationen über einige Personen beinhalten.“ (Lisker, 2023, 52)

kitver Analytik „erhebliche Missbrauchspotenziale verbunden sind, welche sich als soziale Ungleichheit, Diskriminierung und Ausgrenzung manifestieren. Diese Missbrauchspotenziale werden vom geltenden Datenschutzrecht (EU DSGVO) nicht reguliert [...]“ (Mühlhoff, 2022, 31). Mühlhoff warnt davor, dass riesige Datenmen, wenn auch anonymisiert, „in einem weitestgehend rechtsfreien Raum“ (ebd.) genutzt werden.

Wichtig sind die Veränderungen durch neue Technologien der Verarbeitung großer Datenmengen und Verfahren des maschinellen Lernens. Auch wenn ein statistisches Verfahren mit anonymisierten Daten arbeiten kann und es unterschiedliche Verfahren der Anonymisierung gibt, können sowohl die zugrundegelegten Daten als auch die Einzelfälle auf verschiedenen Ebenen personenbezogene und sensible Daten involvieren. Die Problematik besteht bereits in der Tatsache, dass kommerzielle Player im Besitz dieser Daten sind, nicht erst in den einzelnen Aktivitäten mit den Daten. Eine Geschäftsform prädiktiver Analytik, die ein besonderes Interesse an personenbezogenen und sensiblen Daten hat und eigene Datenbanken auf unterschiedlichen Wegen erweitert, ist – neben AdTech – die Sicherheitsindustrie.

### LexisNexis® Risk Solutions

LexisNexis® Risk Solutions, ein Tochterunternehmen von RELX, das als Sicherheitsbereich von Elsevier agiert, wirbt mit einem besonders umfangreichen Datenvorrat, u. a. „[b]ig data platforms“, “[p]redictive modeling” und “[m]achine learning and artificial intelligence” (RELX, 2024, Folie 23, s. Abbildung 19 im Anhang). Die Risk-Sparte von RELX bietet auf Grundlage ihrer Datenanalyse eine Reihe unterschiedlicher Leistungen an, darunter Risikomanagement, Betrugsentdeckung und -prävention und Beratung von Versicherungen (RELX, 2024, 14):

LexisNexis Risk Solutions comprises the following market-facing industry/sector verticals: Business Services<sup>134</sup>, Insurance<sup>135</sup>, Specialised Industry Data Services<sup>136</sup> and Government Solutions<sup>137</sup>. (Ebd.)

LexisNexis' Kundschaft ist umfangreich: "We do business with 92% of the Fortune 100; 84% of the Fortune 500; nine of the world's top ten banks and 21 of the world's top 25 insurers". (Ebd.)

Die Datenbasis von LexisNexis erlaubt ein so breites Leistungsportfolio, um neben predictive modelling die Kreditwürdigkeit einzuschätzen oder für diverse Zwecke Identitäten zu prüfen und erschöpfende Profile zu erstellen. Laut eigener Darstellung verfügt LexisNexis über eine umfangreiche "identity database" mit u. a. „9 billion active devices“, „2.5 billion digital identities“ und "[c]overage of nearly 100% of the U.S." (LexisNexis, 2023, 6), die ihnen ein "[c]lear understanding of identities across systems, agencies, and states [...]" (LexisNexis, 2023, 2) ermögliche. In welcher Form die Identität von Personen staatenübergreifend ein "clear understanding" erlauben kann, wird durch weitere Eckdaten der Selbstpräsentation angedeutet: "Tens of billions of public records" (RELX, 2024, Folie 23, s. Abbildung 19 im Anhang), zu denen „Bankruptcies, loans, judgments, court records, DMV records, census data, etc.“ (RELX, 2024, Folie 27, s. Abbildung 21 im Anhang) gehören. Für 2021 wirbt LexisNexis in den Bereichen "Physical Identity Intelligence" mit u. a. „285m US unique identities“, „12bn unique name / addresses“ und „Digital Identity Intelligence“ u. a. „5bn devices“, „2bn digital identities“, „2bn unique email addresses“, „1.4bn unique IP addresses“, „1.3bn unique cell phones“ (RELX, 2024, Folie 26, s. Abbildung 20 im Anhang). Im Bereich „Insurance solutions“ könne ein "Holistic Assessment of Risk" in den Bereichen „Auto“, „Property“, „Life“ und „Commercial“ erstellt werden, mit "policy and carrier history data", der Kreditwürdigkeit und Schadenhisto-

**134** Darunter spezifiziert RELX "Risk assessment of individuals, digital devices and transactions to help prevent fraud and financial crime" (RELX, 2024, Folie 18).

**135** "Risk assessment tools for insurance underwriting, pricing, and claims processing". (RELX, 2024, Folie 18)

**136** "[I]ntelligence, aviation, HR, etc.". (RELX, 2024, Folie 18)

**137** "Tax and identity fraud prevention, enabling access to social benefits, disaster relief, crime prevention." (RELX, 2024, Folie 18)

rie sowie Fahrzeuginformationen: “telematics”, also Informationen zum Fahrverhalten; der Fahrzeughistorie, Fahrzeugregistrierungen, dem Fehlverhalten im Straßenverkehr und mit Geodaten (RELX, 2024, Folie 27, s. Abbildung 21 im Anhang).

Wie kommt LexisNexis an so umfangreiche Informationsbestände, die selbst im Vergleich zur RampID von LiveRamp beeindruckend sind? Das Unternehmen machte einen wichtigen Entwicklungsschritt, als Reed Elsevier 2004 für 775 Mio. US-Dollar Hank Ashers<sup>138</sup> Firma Seisint Inc. kaufte. Seisint hat(te) mit Accurint einen direkten Zugang zu öffentlichen Verzeichnissen und US-Behördendaten: “Seisint’s Accurint product provides online access and analysis of public records to help governments and law firms with tasks including *due diligence*, *debt recovery*, *pre-employment screening and identity checks*”<sup>139</sup> heißt es in einer Pressedarstellung der Elsevier-Repräsentantin Susanna Smart bei Spectrum Equity (Herv. YF).

Diese personenbezogenen Informationen standen dann Reed Elsevier zur Verfügung und „aus dem früheren Accurint Data Link wurde die LexID“ (Siems, 2024a, 5). Das Konzept der LexID, einem umfassenden Identifier privatwirtschaftlicher Provenienz und “a kind of shadow Social Security number” (Funk, 2023), ist also maßgeblich Hank Asher zu verdanken. Die LexID ist aber, so McKenzie Funk (ebd.) nicht der einzige “[persistent identifier]”, as they are called, that have been issued not by the government but by data companies like Acxiom, Oracle, Thomson Reuters, TransUnion — or, in this case, LexisNexis.”

Auf welchen Ebenen die LexID in das Leben der Menschen in den USA eingreift, wie verbreitet und sie ist und welche Macht mit dieser Klassifizierung ausgeübt werden kann, fasst Siems Rezension von *The Hank Show* zusammen:

<sup>138</sup> McKenzie Funks Buch über Hank Asher zeigt eine schillernde Figur, die privatwirtschaftliches, massenhaftes Sammeln und Auswerten von Daten, zunächst öffentlicher, durch aufkommende Überwachungstechnologien zunehmend erhobener und gehandelter, sozusagen erfand.

<sup>139</sup> <https://www.spectrumequity.com/news/reed-elseviers-lexisnexis-acquires-seisint-for-775-million>. [zuletzt geprüft am 25.07.2024]

Die LexID macht die Lebensumstände [eines] Menschen sichtbar, seinen Biorhythmus, die Vorlieben, die Hobbies und das Netzwerk an Menschen, die teils ephemere, teils intensiv das soziale Umfeld bilden. Sie beeinflusst damit die ärztliche Versorgung, aber auch, wie teuer die Autoversicherung ist, welche Kreditkarte zur Verfügung gestellt wird und wie lange McKenzie Funk in der Hotline warten muss. Die LexID und ihre konkurrierenden Identifier ermöglichen somit, was Shoshana Zuboff *prediction products* nennt, nämlich die maschinelle Auswertung und Analyse von menschlichem Verhalten durch Polizei, Banken, Nachrichtendiensten, Versicherungen, politischen Parteien und die Marketingindustrie, um das Verhalten vorhersagen und ausbeuten zu können. Identifier wie die LexID oder eben die Voter\_id von Cambridge Analytica stützten in den USA die Last Minute-Kampagnen in den Swing States, die ihrerseits möglicherweise Trump die Präsidentschaft sicherten. In der Pandemie ermöglichten sie das contact tracing, bei den Protesten gegen die Ermordung von George Floyd sollen sie die Gesichtserkennung der Menschen in der Menge unterstützt haben, während die über ihnen kreisenden Überwachungsdrohnen die IMSI-Nummern, also die Mobilnetzkennungen, aus den Handys der Protestierenden saugten. (Siems, 2024a, 1)

Eine weitere Datenquelle und Geschäftssparte ("Government Solutions" oben) für LexisNexis sind Behörden-Kooperationen im Sicherheitsbereich. Das 2003 gegründete LNSSI (LexisNexis Special Services Inc.)<sup>140</sup> arbeitete an teils als geheim eingestufteten Projekten mit den US-Behörden zusammen. Reed Elsevier und die CIA waren die wichtigsten Investoren in Palantir<sup>141</sup> (Siems, 2024a, 5), die umstrittene Überwachungsfirma, die von Peter Thiel<sup>142</sup> mitbegründet wurde. Aufsehen erregte auch der Vertragsabschluss 2021 mit der ICE (Biddle, 2021; Lamdan, 2019), der amerikanischen Einwanderungsbehörde, der LexisNexis die Datengrundlage für Abschiebungen lieferte.

<sup>140</sup> <https://www.lexisnexisspecialservices.com/>. [zuletzt geprüft am 25.07.2024]

<sup>141</sup> [www.palantir.com](http://www.palantir.com). [zuletzt geprüft am 25.07.2024]

<sup>142</sup> <https://www.nytimes.com/2021/09/21/books/review/the-contrarian-peter-thiel-max-chafkin.html>. [zuletzt geprüft am 25.07.2024]

Ein weiterer bekannt gewordener Fall, bei dem LexisNexis mit Daten handelt, ist Xandr. Der Data Broker veröffentlichte (vermutlich versehentlich) kurzzeitig eine Datei<sup>143</sup> mit Targeting Groups, welche über LexisNexis und LiveRamp (vgl. Abbildung 11) an Xandr gelangten und deren Entdeckung zu einer weiteren Enthüllung führte (Keegan & Eastwood, 2023)<sup>144</sup>. Die Profile in den Targeting Groups enthielten u. a. geschützte Berufsgruppen, in Kombination mit

sensibleren Themen wie Religionszugehörigkeit, Gesundheitszustand, LGBTQ, ‘climate change believer’/‘climate change non-believer’, ‘vaccine propensity’, ‘don’t trust vaccine’ und ‘pro life high-income Republican voter’, die es jedem mit genug Geld ermöglichen, eben auch Datensegmente mit entsprechenden Wissenschaftler\*innen zu kaufen, oft inklusive Geofencing- bzw. Geolocation-Daten. (Brembs, 2023a)

```
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597312,Branded Data - Stixista - Fusion - LexisNexis - Home Value 1 Year Change - Home Depreciated 30-35% Last Year (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597313,Branded Data - Stixista - Fusion - LexisNexis - Home Value 1 Year Change - Home Appreciated 18-15% in the Last 5 Years (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597313,Branded Data - Stixista - Fusion - LexisNexis - Home Value 1 Year Change - Home Depreciated 15-20% Last Year (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597314,Branded Data - Stixista - Fusion - LexisNexis - Home Value 1 Year Change - Home Depreciated 1-5% Last Year (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597315,Branded Data - Stixista - Fusion - LexisNexis - Time Since Last Business Association - 15-20 Years (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597316,Branded Data - Stixista - Fusion - LexisNexis - Time Since Last Business Association - 1-2 Years (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597317,Branded Data - Stixista - Fusion - LexisNexis - Time Since Last Business Association - 3-4 Years (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597318,Branded Data - Stixista - Fusion - LexisNexis - Time Since Last Business Association - 20+ Years (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597319,Branded Data - Stixista - Fusion - LexisNexis - Time Since Last Business Association - 18-15 Years (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597320,Branded Data - Stixista - Fusion - LexisNexis - Professional License Type - Physicians Surgeons Dentists Psychologists and Optometrists (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597321,Branded Data - Stixista - Fusion - LexisNexis - Professional License Type - Licensed Practical Nurses and Pharmacists (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597322,Branded Data - Stixista - Fusion - LexisNexis - Professional License Type - Accountants Chiropractors and Registered Nurses (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597323,Branded Data - Stixista - Fusion - LexisNexis - Professional License Type - Veterinarians Real Estate Agents and Hair Stylists (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597324,Branded Data - Stixista - Fusion - LexisNexis - Own Current Residence - First and Last Name Property Record (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597325,Branded Data - Stixista - Fusion - LexisNexis - Tenager in Household Age 13-18 (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597326,Branded Data - Stixista - Fusion - LexisNexis - Time Since Last Business Association (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597327,Branded Data - Stixista - Fusion - LexisNexis - Own Current Residence (BlueMail)
"Audiences by Oracle (BlueMail, DataLogic, AdFull11)",1262,2597328,Branded Data - Stixista - Fusion - LexisNexis - Financial - Original Mortgage Amount - $50k-$75k (BlueMail)
```

**ABBILDUNG 11:** Datensatz-Ausschnitt mit Berufsgruppen-Angaben zu Wissenschaftler:innen, geliefert u. a. von LexisNexis an Xandr.

Data Fusion – die Zusammenführung verschiedener Informationsteile aus verschiedenen Quellen, der Austausch zwischen Datenhändlern, die Vergabe eindeutiger Identifikatoren und die Anwendung von Analyseverfahren – greift nicht nur tief in die Privatsphäre und den Rechtsraum von Millionen von Menschen ein. Unternehmen wie LexisNexis oder das Kon-

<sup>143</sup> [https://raw.githubusercontent.com/the-markup/xandr-audience-segments/main/data\\_marketplace\\_public\\_segments\\_pricing\\_05212021.csv](https://raw.githubusercontent.com/the-markup/xandr-audience-segments/main/data_marketplace_public_segments_pricing_05212021.csv). [zuletzt geprüft am 25.07.2024]

<sup>144</sup> “A spreadsheet on ad platform Xandr’s website reveals a massive collection of ‘audience segments’ used to target consumers based on highly specific, sometimes intimate information and inferences.” <https://github.com/the-markup/xandr-audience-segments>. [zuletzt geprüft am 25.07.2024], Datensatz zum Artikel von Keegan and Eastwood (2023).

kurrenzprodukt CLEAR von Thompson Reuters erhalten durch Data Fusion und Identifikatoren wie LexID auch einen sehr großen Wissensvorsprung: „Jetzt kann man so viel über jemanden wissen, dass man sein typisches Verhalten vorhersagen kann, ohne ihn in irgendeiner Weise zu kennen oder je getroffen zu haben.“ (Siems, 2024a, 2)

Last but not least handelt es sich bei Data Fusion um einen „Multimilliardenmarkt, der nach Marktprognosen mit 15-20 % jährlich wächst [...]“ (Siems, 2024a, 7). Gerade hier überschneiden sich die entscheidenden Player aus dem Wissenschaftsbereich mit dem Wirtschaftsbereich, und das zum Nachteil der Wissenschaft, wie Siems warnt:

Marktbestimmend ist eine äußerst überschaubare Anzahl Firmen, und neben Palantir kennt man drei davon im Wissenschaftsbereich schon sehr lange: RELX, Thomson Reuters und Clarivate. Paranoid computing hat sich demnach teils auch aus dem Wissenschaftsbereich heraus entwickelt, teils durch die Vielzahl von Aufkäufen, durch Fusionen und die Erschließung neuer Geschäftsfelder sich dort festgesetzt – und mit Scoring kennt man sich im Wissenschaftsbereich ebenso schon lange aus. (Siems, 2024a, 7)

## ThreatMetrix® und BehavioSec®

Weitere Details der Identifikationstechnologien aus der Selbstpräsentation von ThreatMetrix®<sup>145</sup> zeigen die Reichweite und den quasi militärisch-nachrichtendienstlichen Charakter des Unternehmens. Personen könnten samt Standort und IP-Adresse auch dann identifiziert werden, wenn sie versteckte Proxy-Server oder VPNs (durch „TrueIP“) nutzten. „ThreatMetrix SmartID®“ könne weitere Wege der Anonymitätswahrung wie privates Browsing, Cookie-Entfernung und „traditionelle“ Formen der Umgehung von Device-Fingerprinting-Tools durchdringen, heißt es in der Selbstdarstellung (LexisNexis, 2019, 4). „Trust Tags“, ein weiteres Produkt, „ermöglicht Unternehmen, vertrauenswürdige und dubiose Nutzer, Geräte, Standorte und Personas zu definieren, zu kategorisieren, zu taggen und zu differenzieren“ (ebd.). Wie präzise und unausweichlich die

<sup>145</sup> LexisNexis (2019) und <https://risk.lexisnexis.com/global/de/products/threatmetrix>. [zuletzt geprüft am 25.07.2024]

Identifikationsmöglichkeiten arbeiten, zeigt der „Unique, Tokenized Customer Identifier“ LexID® Digital, “[which] [c]onsistently identifies a person irrespective of changes in devices, locations or behavior. Intelligence from the LexisNexis® Digital Identity Network® helps reliably recognize the same returning user behind multiple devices, email addresses, physical addresses and account names.” (LexisNexis, n.d.a, 1)

Eine relativ neue, AI-gestützte Identifikationstechnologie ist BehavioSec. Damit werden verhaltensindividuelle Analysen, wie sie schon im Fingerprinting existieren, perfektioniert, “transforming Digital Behavior into Actionable Intelligence”<sup>146</sup>. BehavioSec wirbt mit biometrischer Identifizierung durch kombinierte Interaktionsformen der Nutzer:innen mit den unterschiedlichen Geräten. Diese Identifizierung wurde unter anderem durch Smartphones und Tablets erst möglich, z. B. Tastendruck („keystroke“), Beschleunigungsmesser („accelerometer“), Gyroskop, also die Bewegung des Geräts im Raum („gyroscope“), Mausbewegung („mouse movement“), Display-Berührung („surface“), Bewegungsmuster (“motion“), Trefferbereich („hit zone“) und Druckmuster („pressure“) (ebd.). Laut Werbe-Broschüre erfolgt die Datensammlung „passively“, also unbemerkt von Nutzer:innen und “invisibly authenticates users by analyzing user device interactions when they transact online” (LexisNexis, n.d.b, 2). Zusätzlich zu anderen Formen des Datensammelns und der Analyse ist mit BehavioSec nun auch eine Technologie gegeben, die verhaltensbiometrische Analysen und die verschiedenen neuen Formen der Interaktionsformen mit sogenannten smart devices kombiniert, um digitale Anonymität zu verhindern: “The combination of typing rhythm, touch pressure, swiping patterns, device orientation and other metrics help confirm identity with pinpoint precision [...]” (Ebd.)

Eine Abbildung (23, im Anhang) der Broschüre von LexisNexis (2023) zeigt unter der Überschrift “Discover a Precise View of Identity” das Foto einer Person im Mittelpunkt verschiedener Identifikationsebenen: *digitaler* (Email), *physischer* (ein EKG-Kurven-Symbol für Gesundheitsdaten, Standort, Wohnadresse), *behavioraler* Erkennung (Symbol, das im Bereich BehavioSec als „Pressure“ beschrieben wird; ein Maussymbol, bei-

<sup>146</sup> <https://risk.lexisnexis.com/products/behaviosec>. [zuletzt geprüft am 25.07.2024]

des Zeichen für biometrische Identifikation) und einer Geräteidentifikation („device“) (Symbole für lokale und mobile Geräte – Bildschirmcomputer, Mobiltelefon, Tablet). LexisNexis wirbt mit der Kompetenz, die Person aus diesen vier Kategorien verlässlich zu identifizieren und Betrugsfälle ausschließen zu können. Die Darstellung der „Comprehensive Identity Database“ mit „2.5 billion digital identities“ und „Coverage for nearly 100% of the U.S [...]“ macht außerdem klar, dass all diese Analysen auf eine unbestimmt große Menge der Nicht-Kunden von LexisNexis angewendet werden. Diese Normalverbraucher:innen wissen nichts darüber, dass sie durch kombinierte Verhaltensmuster, eine Geräteidentifikation, ihren Wohnort, ihre Gesundheitsdaten oder ihren Standort identifiziert oder in einer „Comprehensive Identity Database“ geführt werden.

Eine andere Grafik (Abbildung 18 im Anhang) zeigt zwei Funktionen von LexID Digital, den „Digital ID View“ mit den unterschiedlichen personenbezogenen Daten einer ID (14 Aspekte, darunter sind z. B. Telefonnummer, Email-Adresse und Adresse, SmartID der Endgeräte, Credit Card Hash, zu denen es jeweils mehrere Werte geben kann; lokale und globale Entitäten, Angaben zu „first seen“ und „last seen“) und „Relationships“. Letztere Analyse betritt laut Beschreibung<sup>147</sup> „a level deeper“ und erfasst die Interaktionen zwischen den einzelnen Entitäten und früherem Verhalten (das entsprechend aufgezeichnet werden muss, um diese retrospektive Analyse zu ermöglichen). Das gesamte Ausmaß der unterschiedlichen Informationen über riesige Menschengruppen, ob als private Bevölkerungsregister bei LiveRamp oder LexisNexis‘ Datenbank und damit einhergehende Gefahren ist kaum zu überschätzen:

The company's databases offer an oceanic computerized view of a person's existence; by consolidating records of where you've lived, where you've worked, what you've purchased, your debts, run-ins with the law, family members, driving history, and thousands of other types of breadcrumbs, even people particularly diligent about their privacy can be identified and tracked through this sort of digital mosaic. LexisNexis has gone even further than merely aggregating all this data: The company claims it holds 283 million dis-

---

147 LexisNexis (n.d.a, 2).

tinct individual dossiers of 99.99% accuracy tied to ‘LexIDs.’ (Biddle, 2021)

Diese Überwachungstechnologie wird durch die Weiterentwicklung von AI und Verarbeitungskapazitäten kontinuierlich verbessert: “The behavioral data is compared to the unique profile of the user in near real time” (LexisNexis, n.d.b, 2) sowie “Deep intelligence [AI] The scores are delivered with detailed intelligence on human device interaction. This is combined with digital identity intelligence to quickly achieve situational awareness.” (LexisNexis, n.d.b, 2).

Die Frage, welche explizite Datenschutz-Policy LexisNexis hat, wurde bisher nicht gestellt, da der Fokus darauf liegt, welche Datenbestände sich bereits in Besitz der RELX-Tochterfirma befinden und laufend dazukommen. Allein die hier zugrunde gelegten öffentlichen Informationen bzw. eigenes Werbematerial machen in ihrem nachrichtendienstartigen Format der Datenbeschaffung eine eigene Untersuchung der Datenschutz-Policy sinnfrei. LexisNexis wirbt mit “Privacy by Design” durch “providing dynamic risk assessment of identities while maintaining data privacy through the use of tokenization” (LexisNexis, n.d.a, 3). Die Abbildung 18 im Anhang enthält den Hinweis auf „privacy“ durch „tokenization“: “Entities local to an organization can be decrypted for review while the global entities remain tokenized” (ebd.). Jedoch werden diese Maßnahmen auf *bereits vorhandenes* Datenmaterial angewendet, so dass es sich hier um Sicherheitsmaßnahmen gegen den Verlust von Daten bei LexisNexis handeln dürfte. Die Rechtmäßigkeit des Erwerbs des Datenmaterials ist damit nicht adressiert. Nicht adressiert werden auch das Machtpotenzial dieser epistemischen Vogelperspektive von Konzernen sowie die Tatsache, dass diese Technologie und Geschäftsform das Gegenteil digitaler Souveränität oder der Selbstbestimmungsfähigkeit von Personen exemplifiziert.

## 4.2 Formen des Trackings in der Wissenschaft

Es ist keine neue Erkenntnis, dass ein Großteil des Webgeschehens durch Werbung geprägt ist und hierfür Überwachungs- und Datensammelpraktiken in massiver Weise eingesetzt werden. Spätestens seit der Diskussion um das Cookie-Banner, das uns bei Aktivitäten im Netz begleitet, sind diese Themen weitestgehend bekannt, wenn auch nicht im Detail. Der relevante neue Teil daran ist die Anwendung der Überwachungsökonomie auf den Wissenschaftsbereich, der als ein grundgesetzlich geschützter Bereich vom Geschäftsmodell der Finanzierung durch trackingbasierte Verwertungsketten ausgeschlossen sein sollte.

Welche dieser allgemeinen Trackingmechanismen treffen auf den Wissenschaftssektor zu, welche Akteur:innen sind dabei relevant und welche Folgen sind zu erwarten? Eine Reihe von Autor:innen, Untersuchungen und Institutionen haben sich kritisch mit dieser Thematik beschäftigt, angefangen mit den Arbeiten von Clifford Lynch, dem Director of the Coalition for Networked Information (CNI) (Lynch, 2017, 2019), dem Vortrag des IT-Leiters der University of Minnesota Libraries Cody Hanson (2019) bei der CNI-Konferenz 2019 und seiner folgenden Verbreitung des Themas<sup>148</sup>. Hanson erzählte zum einen, wie im Rahmen der Diskussion um die Einführung der RA21-Authentifizierung zufällig ans Licht kam, dass die digitalen Plattformen von Wissenschaftsverlagen längst invasive Datensammlung und Identifikation von Nutzer:innen betreiben. Diese Praktiken wurden im Wissenschaftsbereich aus guten Gründen nicht erwartet und entsprechen dem auf Überwachung und Datenhandel basierenden AdTech-Geschäftsmodell.

Zum anderen führte Hanson eigene, technisch niederschwellige Untersuchungen der größten Wissenschafts-Plattformen durch und zeichnete einen ersten inhaltlichen Rahmen der Trackingaktivitäten nach. Die Palette der Trackingoptionen wurde seitdem ergänzt und vertieft<sup>149</sup> und die

---

<sup>148</sup> <https://www.codyh.com/writing/tracking.html>. [zuletzt geprüft am 25.07.2024]

<sup>149</sup> Siems (2021, 2022); Freiberg (2022); Dittmann et al. (2023).

kritische Diskussion der Legitimität, Konsequenzen und des Ausmaßes der Aktivitäten und damit verbundenen Handlungsanschlüsse initiiert. Gleichzeitig wurde eine neue Brisanz in die Reflexion der seit einigen Jahren beobachteten Transformation der Wissenschaftsverlage (sowie Clarivate ohne Verlagsfunktion) in Datenanalyse-Konzerne deutlich<sup>150</sup>. In dieser Debatte wurde der Zusammenhang zwischen den Überwachungsaktivitäten der Konzerne und der Fortführung der DEAL-Verträge mit Springer-Nature, Wiley und dem Vertragsabschluss mit Elsevier kritisch reflektiert<sup>151</sup>. Die finanzielle Förderung des Datenanalytik- und Brokering-Business durch Universitäten und die Involvierung von Bibliotheken in ihrer Schnittstellen-Position, die Nutzer:innen autorisiert und den Plattformzugang herstellt, haben eine ganze Reihe negativer Konsequenzen.

Der im Juli 2024 publizierte zweite Teil des AWBI-Berichts hat weitere Erkenntnisse zum Tracking mit Schwerpunkt auf die DEAL-Verträge gebracht und bestätigt, auf den Plattformen bzw. in den Produkten der Wissenschaftsverlage „sowohl First-Party-Tracking als auch Third-Party-Tracking [zu] finden, letzteres vor allem in Form eingebundener Medienbereitstellungsdienste (Content Delivery Networks, CDN), eingebundener sozialer Medien oder externer Analytics-Plattformen.“ (Altschaffel *et al.*, 2024, 26)

Die Trackingtechniken lassen sich nur teilweise durch IT-Forensik und technischen Datenschutz erfassen. Beispiele sind ein spezifischer digitaler Fingerabdruck “Science-Tracking Fingerprint (STF)” (Altschaffel *et al.*, 2023) und weitere Maßnahmen, wie sie die Arbeitsgruppe *Multimedia and Security* (AMSL) an der Fakultät für Informatik der Otto-von-Guericke Universität Magdeburg durchführt. Wie bereits oben ausgeführt, werden dabei nur bestimmte Aspekte des Trackings aufgespürt, die ein notwendigerweise unvollständiges Bild ergeben.

---

<sup>150</sup> Siehe unter anderen Siems (2022, 2023b, 2021); SPARC (2023, 2021b); Scholl (2021); Roßnagel (2022); Ronzheimer (2021); Kunz (2022); Knoche (2021b); Hanson (2019); Freiberg (2022); Eric (2015); Förstner (n.d.); Brembs (2023a); Brembs & Siems (2020); DFG AWBI (2021).

<sup>151</sup> Brembs *et al.* (2021a); Brembs & Siems (2020); Brembs *et al.* (2023a); Brembs (2023a); Förstner (n.d.); Brembs *et al.* (2021b); Förstner & Brembs (2022); Siems (2022, 2024b, 2023c, 2021).

Eine nicht-erschöpfende Übersicht der in der Debatte genannten Formen des Wissenschaftstracking bietet die *Übersicht Tracking- und Überwachungsmethoden in der Wissenschaft* in den Tabellen 5 und 6. Zu beachten ist, dass es sich bei den jeweiligen Nachweisen der Präsenz von Trackingmethoden<sup>152</sup> um Momentaufnahmen handelt, die zu spezifischen Zeitpunkten festgehalten wurden, nicht um eine Dauerbeobachtung. Die trackende Seite kann die Konfiguration der eigenen Website *jederzeit* verändern, zum Beispiel durch die zeitnahe Entfernung aktuell kritisierter Anwendungen. In welcher Weise die über Audience Tools oder durch LexisNexis gesammelten Daten genutzt werden, ist nicht unabhängig nachweisbar, da diese Information im System der trackenden Seite bzw. der vielen unterschiedlichen Third Parties liegen. Über die Formen und Konsequenzen solcher schwierig nachverfolgbaren Trackingmethoden kann nur spekuliert werden, ebenso über das oben beschriebene, schwierig festzustellende Server-Site-Tracking. Hier werden außerdem nur Formen des Webtrackings aufgelistet; für App- und Email-Tracking s. Altschaffel *et al.*(2023).

---

**152** Hanson (2019); Siems (2021, 2024b); Freiberg (2022); Dittmann et al. (2023); SPARC (2023).

**TABELLE 5:** Übersicht Tracking- und Überwachungsmethoden in der Wissenschaft (Teil I).

| Tracking Methode             | Beschreibung                                                                                                                                                                                                        | Formen, Beispiele                                                              | Anwendung                                                                | Konsequenzen                                                                                                                                                                                                                                                        |
|------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Authetifizierungsform</b> | Anbieter stipulieren die Nutzung bestimmter Authentifizierungsmethoden, die eine anonymisierte Nutzung ihrer Angebote, z.B. nur durch die Authentifizierung als Institutionsmitglied, verunmöglichen <sup>153</sup> | SAML basierte Verfahren: Single-Sign On, RA21 zu Seamless Access.org zu GetFTR | Kontrolle über Nutzer:innen von Bibliotheksdiensten bzw. Hochschulnetzen | Online-Verhalten von eingeloggten Nutzenden wird beobachtet, sie werden auf Verlagsplattformen umgeleitet „The architecture of GetFTR is designed to take place behind the scenes, with no ability for libraries or other users to opt out [...]“ (SPARC, 2021a, 8) |
| <b>Cookies</b>               | Code auf Geräten der Endnutzer:innen, der von first und third party gesetzt wird. Reicht von einfacher Analyse bis zu sitzungsübergreifender Verfolgung und praktisch nicht entfernbaren Versionen                  | first/third party cookie, evercookie (Quant Cast, KISS-metrics)                | Erstellung und Verwertung von Profilen                                   | Profilerstellung, Überwachung digitaler und analoger Aktivitäten                                                                                                                                                                                                    |
| <b>Audience Tools</b>        | anwendungsübergreifende Aggregation von Daten zur Erstellung umfassender Profile „Im Fall von Anbietern                                                                                                             | Adobe (Audience Manager), Neustar, Oracle                                      | Erstellung und Verwertung von Profilen Überwachung                       | „Da viele Anbieter teils die gleichen Third Parties einbinden, teils diese ihre Daten aus-                                                                                                                                                                          |

<sup>153</sup> „[T]hey want to replace the authentication options previously supported by libraries and academic institutions, such as IP range activation, VPN, proxy servers, or anonymous authentication to neutral third parties, as with the Shibboleth service, in favor of their own initiatives [‘Research Access 21 (RA21)’, ‘Seamless Access’, or ‘Get Full Text Research (GetFTR)’] (Siems, 2021)

|                                                                                                                                                                                                                                         |                  |  |                                                                                                                                                                                                                                                                |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| wie Acxiom/Liveramp ist auch eine Synchronisation von Online- mit Offlineleben möglich, da dort Daten auch über Einkäufe, Führerscheine, TV-Konsum, Wählerverzeichnisse, Straffälligkeiten und anderes vorliegen.“ (DFG AWBI, 2021, 11) | BlueKai, AddThis |  | tauschen kann das Informationsverhalten von Hochschulangehörigen plattformübergreifend erhoben werden und im Fall von Google, Facebook oder Twitter mit dem bereits vorliegenden Wissen über ihr sonstiges Onlineleben verknüpft werden.“ (DFG AWBI, 2021, 10) |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

TABELLE 6: Übersicht Tracking- und Überwachungsmethoden in der Wissenschaft (Teil II).

| Tracking Methode          | Beschreibung                                                                                                                                                                                                                            | Formen, Beispiele | Anwendung                                  | Konsequenzen                                                                                                                                                                                               |
|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|--------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Real-Time Bidding         | „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 [...]“ (Siems, 2022, 7) | BidSwitch         | Programmatic Advertising                   | “RTB data about the average person are broadcast to many entities, hundreds of times a day. There is no way to limit or know what happens to RTB data after they are broadcast.” (Ryan & Christl, n.d., 7) |
| Geofencing + IP targeting | Ineinandergreifende Verfahren: „Im [...] Hochschulnetz ist die Lokalisierung genau-                                                                                                                                                     |                   | Lokalisierung von Institutionen und Perso- | „Die dadurch mögliche gezielte, standortbezogene Übermittlung von Inhalten über die IP-                                                                                                                    |

|                             |                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                |                                                |                                                                                                                                                                                                                                                                                                                    |
|-----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                             | <p>er, da hier die Adresse des Einwahlknotens öffentlich bekannt ist und mit dem Gerätestandort übereinstimmt. Der geographische Zaun (engl. fence) um die Nutzenden ist also deutlich enger gezogen. Die Zuordnung lässt sich über mobile Geräte noch verfeinern, da sie von Mobilfunk und WLAN-Netzknoten, RFID-Antennen und je nach Geräteeinstellungen über GPS-Daten registriert werden können.“ (Freiberg, 2022, 98)</p> |                                                                | <p>nen<br/>Übelbachung</p>                     | <p>Adresse wird als IP-Targeting bezeichnet und kann nicht nur bestimmte Geräte, sondern auch hinter den IP-Adressen stehende Haushalte, Betriebe, Institutionen und deren Subnetze wie auch Fachbereiche und Bibliotheken adressieren.“ (Freiberg, 2022, 98)</p>                                                  |
| <p><b>Port Scanning</b></p> | <p>Identifikation durch Ausnutzung von Sicherheitslücken in Browsern/Servern</p>                                                                                                                                                                                                                                                                                                                                               | <p>Suche nach Schwachstellen in Endgeräten oder Netzwerken</p> | <p>Form der Individualisierung<br/>Hacking</p> | <p>„Schon die Suche nach offenen Ports auf fremden Rechnern und/oder Netzwerken, um dort dann etwa Schad- oder Überwachungssoftware einzuschleusen, ist nach deutschem Verständnis am Rand der Legalität, da sie als Vorstufe von entsprechend sanktionierten Tatbeständen (§§202c, 303b StGB) gewertet werden</p> |

|                                                                        |                                                                                                                                                               |                            |                                                                          |                                                                                                                                                                                                                                   |
|------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|--------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                        |                                                                                                                                                               |                            |                                                                          | kann.“ (DFG AWBI, 2021, 11)                                                                                                                                                                                                       |
| <b>Device/<br/>Browser-&amp;<br/>Personen-<br/>Finger<br/>printing</b> | Individuierung von Geräten bzw. geräteübergreifend von Personen über die Erstellung von Profilen ihrer Geräte bzw. biometrischer Merkmale der Gerätebedienung | Double Click<br>BehavioSec | Identifizierung von Personen durch behavioral biometrische Profilbildung | Erfassung auch derjenigen, die Maßnahmen gegen Tracking ergreifen bzw. durch institutionelle Sicherheitsmaßnahmen geschützt sind; Geräteübergreifende Profilbildung                                                               |
| <b>History<br/>Sniffing</b>                                            | Zugriff auf die Verbindungshistorie im Browser, um das Surfverhalten zu verfolgen. S. SPARC (2023, Appendix C-3 und C-4)                                      | über Java-Script           | Spionage und Überwachung                                                 | „Der Javascript-Code der Third Parties [...] kann [...] auslesen, mit welchem Text der Nutzer sich beschäftigt, zu welchem er als nächstes weiterbrowsst und welche Suchworte er auf der Plattform eingibt.“ (DFG AWBI, 2021, 10) |

## 4.3 RELX Lock-In II Institution

In Bezug auf Elsevier gibt es zwei tracking-relevante Bereiche: zum einen die Frage, welche Produkte und Services von Elsevier in welcher Weise problematische Datenaktivitäten (z. B. von LexisNexis und ThreatMetrix) implementieren könnten und inwiefern der DEAL-Vertrag diese Aktivitäten umsetzt. Zum anderen stellt sich die Frage, inwiefern RELX durch seine verschiedenen Sparten jetzt oder künftig einen Einfluss auf den Wissenschaftsbetrieb ausüben könnte und welche Konsequenzen daraus zu erwarten wären.

Ich beziehe mich hauptsächlich auf privacy-bezogene Darstellungen der Elsevier-Website sowie SPARCs Analyse (2023) der problematischen Aspekte von ScienceDirect. Die Ergebnisse dieser Analyse sind alarmierend:

Based on our findings, many of ScienceDirect's data privacy practices directly conflict with library privacy standards and guidelines [...] The widespread data collection, user tracking and surveillance, and disclosure of user data inherent to these business models run counter to the library's commitment to user privacy as specified in the ALA Code of Ethics, Library Bill of Rights, and the IFLA Statement on Privacy in the Library Environment. (SPARC, 2023, 4f.)

Die Untersuchung wurde 2022 und 2023 auf ScienceDirect-Seiten (ScienceDirect home page, Browse Journals and Books sowie Suchergebnisse für keyword search; eine Seite mit einem Open-Access-Zeitschriftenartikel) mit folgenden Werkzeugen durchgeführt: Website Evidence Collector, Blacklight, NoScript, Ghostery, uBlock Origin, Web Developer Tools in Firefox und Chrome (SPARC, 2023, 13). Es geht hier nicht darum, die SPARC-Analyse erschöpfend zusammenzufassen, sondern gezielt Punkte der Analyse aufzugreifen. Diese Analyse war im Jahr 2023, kurz vor dem Abschluss des Elsevier-DEAL-Vertrags besonders brisant, weil sich die meisten kritisierten Punkte auch im DEAL-Vertrag finden.

### 4.3.1 Elseviers Privacy Policy und der DEAL-Vertrag

Ein wichtiger Punkt für die Evaluierung des Datenschutzaspekts des DEAL-Vertrags wird auch von SPARC moniert: Relevante Datenschutzpassagen sind nicht eigens Teil des Vertrags, sondern als „Privacy Policy (PP)“ und „Data Processing Addendum (DPA)“ lediglich zu externen Dokumenten verlinkt und nicht datiert.<sup>154</sup> So sind sie jederzeit einseitig veränderbar. Diese Dokumente enthalten unterschiedliche Richtlinien und nehmen nicht aufeinander Bezug. Unter diesen weiteren Dokumenten<sup>155</sup> sind

<sup>154</sup> Es gibt nur einen allgemeinen Hinweis auf das letzte Änderungsdatum und keinen Zugriff auf vorherige Versionen. Hier können nur Web-Archive oder eigene Screenshots verwendet werden.

viel spezifischere Angaben zu datenrelevantem Verhalten, die in der PP selbst nicht bzw. unter “other technologies” genannt werden. Dazu gehören “web beacons” und „unique identifier“, die in der ScienceDirect Cookie Notice nicht erwähnt werden, “even though the tools used to investigate data collection on ScienceDirect indicate the use of web beacons and local storage” (SPARC, 2023, 15).

Es folgen sind einige Auszüge aus der Elsevier PP, da es keine eigene ScienceDirect-Policy gibt. An dieser Policy und damit dem DEAL-Vertrag sind viele Punkte problematisch, ich werde nur einige von ihnen näher betrachten, nämlich *Datensammlung*, *(Un-)Freiwilligkeit*, *Interessenkonflikt*, *Datentransfer* und *Datenweitergabe*. Besonders relevante Abschnitte sind kursiviert.

- **Device and usage data**

The Service may automatically collect information about how you and your device interact with the Service, such as:

- Computer, device and connection information, such as IP address, browser type and version, operating system and other software installed on your device, unique device identifier and other technical identifiers, error reports and performance data;
- Usage data, such as the features you used, the settings you selected, URL click stream data, including date and time stamp and referring and exit pages, and pages you visited on the Service; [...]
- Usage data, such as the features you used, the settings you selected, URL click stream data, including date and time stamp and referring and exit pages, and pages you visited on the Service; [...]
- For location-aware Services, *the physical location of your device.*
- *We collect this data through our servers and the use of cookies and other technologies.* You can control cookies through your browser’s settings and other tools. However, *if you choose to block certain cookies, you may not be able to register, login, or access certain parts or*

make full use of the Service. For more details, visit the cookie notice of the Service. [...]

- **Grounds for processing**

When we process any personal information within the scope of certain privacy and data protection laws, we do so: [...]

- *Where applicable, with your consent; and/or*
- *As necessary to operate our business, protect the security of our systems, customers and users, detect or prevent fraud, or fulfill our other legitimate interests as described in the sections above, except where our interests are overridden by your privacy rights.*
- *Where we rely on your consent to process personal information, you have the right to withdraw your consent at any time, and where we rely on legitimate interests, you may have the right to object to our processing.*
- *[...] We may also share your personal information with our group companies and with sponsors, joint venture partners and other third parties, including entities for which we are acting as an agent, licensee, application host or publisher, such as societies, that wish to send you invitations, surveys, programs or information about their products and services that may be of interest to you as determined by your choices in managing your communications preferences and other settings.*
- *[...] We also will disclose your personal information if we have a good faith belief that such disclosure is necessary to: [...]*
- *Detect, investigate and help prevent security, fraud or technical issues; and/or*
- *Protect the rights, property or safety of Elsevier, our users, employees or others; and as part of a corporate transaction, such as a transfer of a journal or other assets to or an acquisition by or merger with another company. [...]*

- **Information we collect**

- *We also may obtain contact details and other information about you from other entities within RELX and from third parties, including:*
- *Social networks that you grant permission to the Service to access your data on one or more networks;*

- Service providers that help us determine a location in order to customize certain products to your location;
- Partners with which we offer co-branded services *or* engage in joint marketing activities; *and/or*
- Publicly available sources *and data suppliers to validate or supplement the information we hold* or to support the Service.“<sup>156</sup>

## Datensammlung

Der letzte Abschnitt der Datensammlung “Information we collect” beinhaltet eine vage und sehr umfangreiche Sammeltaktik. Sowohl andere Bereiche innerhalb von RELX als auch “third parties” dienen als Quellen von Informationen über Nutzer:innen der Elsevier-Services. Spezifiziert, aber nicht ausschließlich, wird von folgenden vier Quellen gesprochen:

- a. “social networks”, zu denen “the Service” nutzer:inseitig der Zugang gewährt wird;
- b. „service providers“, die die Lokalisierung von Nutzer:innen festlegen;
- c. Geschäftspartner in “co-branded services or [...] joint marketing activities” von Elsevier/RELX. Bedenkt man die Tiefe und den Umfang der personenbezogenen und wissenschaftlichen Daten, die bei RELX zusammenfließen, ist dies zusätzlich problematisch.

---

<sup>156</sup> <https://www.elsevier.com/legal/privacy-policy>, Herv. YF.

Zum Vergleich, hier die entsprechende Passage der RELX-Website „Privacy“:

„3. Information We Share

We may share your personal information with certain of our corporate affiliates and with service providers as necessary to provide the services and fulfill the transactions you have requested or otherwise on our behalf based on our instructions. These third parties are required to use appropriate measures to protect such information in accordance with this privacy policy.

We will also share your personal information with third parties if we have your consent or if we have a good faith belief that such disclosure is necessary to (a) meet any applicable law, regulation, legal process or other legal obligation; (b) detect, investigate and help prevent fraud, security or technical issues, and/or (c) protect the rights, property or safety of RELX, our users, employees or others, and as part of a corporate transaction, such as a sale or transfer of assets or an acquisition by or merger with another company.“ <https://www.relx.com/site-services/privacy>.  
[beide zuletzt geprüft am 25.07.2024]

- d. Der letzte Punkt nennt öffentlich zugängliche Quellen und „data suppliers“ (worunter Data Broker fallen), um die bereits bei Elsevier/RELX vorhandene Information zu erweitern, zu prüfen oder “the Service” weiter zu unterstützen.

Punkte b)-d) sind nicht optional und bedürfen keiner Zustimmung seitens des Vertragspartners, es wird nicht auf ein Widerspruchsrecht gegen diese Datensammlung hingewiesen. Es gibt weder eine Legitimierung einer so breit angelegten Datensammlung als solcher noch der einzelnen Punkte. Weshalb ist es notwendig, “contact details and other information” einzuholen, die für beide Vertragspartner vorliegen? Weshalb ist eine Ortung von Nutzenden notwendig, wenn die Rechtmäßigkeit der Nutzung des “Service” bereits durch die Ressourcen des universitären Netzwerks bzw. Authentifizierungsverfahrens erfolgt? Weshalb sollten unbenannte Geschäftspartner von Elsevier/RELX, deren genaue inhaltliche und geschäftliche Ausrichtung sowie der Bezug zu wissenschaftlichen Einrichtungen unklar ist, in einer Informationsbeziehung zu wissenschaftlichen Einrichtungen und deren Angehörigen stehen? Welche Informationen sind mit “other information” gemeint und weshalb sollte es für das Erbringen der Vertragsleistung durch Elsevier/RELX notwendig sein, diese unbestimmt gehaltenen Informationen einzuholen? Welcher Art ist diese Notwendigkeit?

SPARCs Untersuchung hält weiter fest, dass Elsevier

collected and disclosed specific types of personal information, such as:

*Internet or other electronic network activity information*, such as browsing history, search history, online behavior, interest data, and interactions with our and other websites, applications, systems, and advertisements.

*Geolocation data*, such as approximate device location.

*Inferences* drawn from any of the personal information listed above to create a profile or summary about, for example, a consumer’s preferences and characteristics

*Sensitive personal information*, such as personal information that reveals a consumer's passport number, racial or ethnic origin, and account log-in and password. (SPARC, 2023, 15)

Spätestens das Sammeln derart sensibler Informationen wie Ausweisnummer oder "racial or ethnic origin" ist nicht nachvollziehbar und für wissenschaftliche Zwecke nicht gerechtfertigt.

## (Un)Freiwilligkeit

An der Formulierung "[...] if you choose to block certain cookies, you may not be able to register, login, or access certain parts or make full use of the Service" im Abschnitt "Device and usage data" wird beispielhaft deutlich, dass die DSGVO-Forderung der Freiwilligkeit der Zustimmung unterlaufen wird. Der DEAL-Vertrag wird durch die Institutionen geschlossen, aber Nutzer:innen müssen persönlich den Datenschutz-Vereinbarungen zustimmen, für welche gesetzlich hohe Anforderungen gelten: „Freiwilligkeit“ fehlt bereits, wenn die Erbringung einer Dienstleistung zwingend an eine Einwilligung gekoppelt wird (Art. 7 Abs. 4 DS-GVO).“ (Altschaffel *et al.*, 2024, 28) Von Freiwilligkeit kann keine Rede sein, wenn es nicht möglich ist, ein *bereits bezahltes* Produkt zu nutzen, das Wissenschaftler:innen und Studierende für ihre Arbeit brauchen (und darauf beruht die Single-Source-Exemption der Verlage), weil die Registrierung oder der Login ohne Zustimmung verwehrt werden. Wie Siems bemerkt, können Angehörige von Universitäten „sich hier nicht schützen, es sei denn, sie rufen die Seite gar nicht auf. Das können die Betroffenen sich aber nicht aussuchen, denn sie brauchen die Artikel, die sie zitieren wollen, die auf der Literaturliste stehen usw. [...]“. (Siems, 2023a)

## Interessenkonflikt

Es besteht ein grundsätzliches Problem mit der aktuellen Situation bzw. Praxis: Wissenschaftler:innen, Bibliothekar:innen oder Organisationen wie SPARC müssen bei so relevanten Verdachtsmomenten gegen Elsevier bzw. RELX nachweisen, dass Daten der Wissenschaftler:innen und wissenschaftsrelevante Daten *weiterhin* an verschiedene weitere Sparten, wie LexisNexis oder ThreatMetrix oder weitere Third Parties fließen. Da es sich bei der Institution Wissenschaft um ein hochsensibles, durch das

Grundgesetz und die EU-GrCharta geschütztes Gut handelt, ist die Beweislast aber umgekehrt. Spätestens nach dem ersten Nachweis von Praktiken, welche die Wissenschaftsfreiheit und die Rechte von Wissenschaftler:innen, Studierenden und Angehörigen von Wissenschaftseinrichtungen verletzt, müssten die “publishers [who] reinvent themselves as platform businesses” (SPARC, 2023, 9) restlos alle Zweifel ausräumen, dass ihre Aktivitäten im Advertising, Data Brokering und in weiteren Geschäftsbereichen, die direkte Interessenkonflikte implizieren, den Kriterien wissenschaftlicher Integrität und des europäischen Datenschutzes genügen. Das ist bisher nicht passiert. Die Tätigkeit als Datenanalyse-Anbieter stellt einen gravierenden *prima facie* Interessenkonflikt dar, der bisher nicht adressiert wird. Die Untersuchungen von AWBI, SPARC, Hanson, Siems, Freiberg und die nachgewiesenen Tracking-Aktivitäten (s. Tabellen 5 und 6) sind *hinreichend*, um Konzerne wie RELX, “a global provider of information-based analytics”<sup>157</sup> oder Clarivate, “a leading global provider of transformative intelligence”<sup>158</sup> für die Wissenschaft untragbar zu machen. Diesen Punkt macht auch SPARC deutlich:

The collection and disclosure of data about who someone is, where they are, and what they search for and read by the same overarching company that provides sophisticated surveillance and data brokering products to corporations, governments, and law enforcement should be alarming. These practices raise the question of whether simultaneous ownership of key academic infrastructure alongside sophisticated surveillance and data brokering businesses should be permitted at all—by users, by institutions, or by policy makers and regulatory authorities. (SPARC, 2023, 5)

Auch ohne einen technischen Beleg für den Gebrauch personenbezogener Daten in den Bereichen Brokering oder Risk sollte die Forderung nach einer wesentlich gründlicheren und zuverlässigeren Datenschutz-Regelung als der aktuell im DEAL-Vertrag geltenden erfüllt werden. Eine zufriedenstellende Regelung wäre mindestens unabhängig von einseitigen Änderungen der Privacy Policy oder mündlichen Aussagen ohne juristisch bindende Kraft: “*Privacy policies can be changed unilaterally, and de-*

157 <https://www.relx.com/>. [zuletzt geprüft am 25.07.2024]

158 <https://clarivate.com/>. [zuletzt geprüft am 25.07.2024]

*nials are not legally binding.* To be meaningful, any privacy guarantee a vendor makes must be durable, verifiable, and not limited to a particular jurisdiction.” (SPARC, 2023, 5, Herv. YF)

Der aktuelle Stand aller drei DEAL-Verträge mit Springer Nature, Wiley und Elsevier zeigt leichte Verbesserungen, dass nämlich „[e]inige der problematischen Datentrackingtechniken [...] inzwischen von den Verlagen zurückgefahren, gesetzlich *konformere* Policy-Regelungen etabliert, eine Data Policy Governance etabliert und eine Cookie-Management-Plattform *zugesagt bzw. teils schon umgesetzt* [wurden].“ (Altschaffel et al., 2024, 32) Ob das in Anbetracht der inzwischen seit Jahren bestehenden DEAL-Verträge mit Wiley und Springer Nature sowie der jahrelangen Verhandlungen um einen Elsevier-DEAL überzeugt, weiterhin am DEAL-Format festzuhalten, ist eine andere Frage.

## Datentransfer

Der Bericht des AWBI enthält Informationen aus den Verhandlungen aller DEAL-Verträge als zusätzliche Quelle für die Diskussion. Eine bisher hier nicht eigens behandelte Frage ist der Datentransfer in Nicht-EU-Länder, insbesondere in die USA. Dieser Transfer ist insofern problematisch, als er

nicht nur gegen geltendes Daten- und Persönlichkeitsschutzrecht verstößt, sondern auch das Datentracking ausweitet. Zu Beginn der Gespräche war als Folge des ‚Schrems II‘-Urteils (EuGH, Urt. v. 16.07.2020 – Az. C-311/18) die Weitergabe personenbezogener Daten in die USA nur sehr eingeschränkt erlaubt. Doch die Verlage Elsevier und Wiley bestanden mit Verweis auf die (IT-)Strukturen in ihren globalen Konzernen auf der Möglichkeit eines solchen Datentransfers außerhalb Europas. (Altschaffel et al., 2024, 32)

Hier ist nicht nur die konkrete Verhandlungsführung interessant, sondern auch die rechtlichen Änderungen, die nun auch den Datentransfer in Länder zulassen, die deutschen bzw. europäischen Datenschutzregelungen *nicht entsprechen*. Diese konkreten Abläufe möchte ich in größerem Umfang wiedergeben:

Im Lauf der Verhandlungen ergab sich aufgrund der Veröffentlichung eines Angemessenheitsbeschlusses der Europäischen Kommission zum Datenschutzrahmen EUUSA (EU-US Data Privacy Framework vom 10.07.2023, Abl. L 231 vom 20.9.2023, S. 118 ff., [https://eur-lex.europa.eu/eli/dec\\_impl/2023/1795/oj?locale=de&date=2023.12.01](https://eur-lex.europa.eu/eli/dec_impl/2023/1795/oj?locale=de&date=2023.12.01)) allerdings eine neue Situation, denn der Datentransfer in die USA ist seit dem nicht mehr grundsätzlich verboten (Europäische Kommission 2024). Die Verträge mit Elsevier und Wiley sehen vor, dass der Datentransfer außerhalb der EU nur bei Vorliegen der Voraussetzungen der Art. 44 ff. DS-GVO erlaubt ist. Dies umfasst neben den Angemessenheitsbeschlüssen (Art. 45 DS-GVO) auch bloß vertragliche Garantien der Verantwortlichen (Art. 46 DS-GVO) – der Vertrag von Wiley hebt insoweit explizit die EU-Standardvertragsbedingungen hervor (Art. 46 Abs. 3 DS-GVO, EU SSCs), die aber auch für Elsevier aufgrund des weiten Verweises ein zulässiger Weg sind. *Damit kann kaum noch von einem dem europäischen Rechtsrahmen angemessenen Datenschutzniveau gesprochen werden, denn die Daten der deutschen Wissenschaftseinrichtungen werden nun auch in Ländern wie Indien, der VR China, den USA oder auch den Philippinen verarbeitet.* Die Verlage regeln das im Detail jeweils unterschiedlich. So lässt Elsevier gemäß seiner Privacy Policy Informationen in ‘Australia, China, France, Germany, India, Ireland, the Netherlands, the Philippines, Singapore, the United Kingdom, and the United States’ (<https://www.elsevier.com/de-de/legal/privacy-policy>, Stand 06.12.2023) Daten verarbeiten. [...] Allerdings wird bei beiden Verlagen [Elsevier und Wiley] nicht ohne weiteres klar, *ob personenbezogene Daten aus Deutschland tatsächlich in diese Länder transferiert werden, und wenn doch, unter welchen Rahmenbedingungen*, da etwa das Open Access Agreement Team von Elsevier in Chennai, Indien, sitzt. Da der Vertrag auf der Grundlage einer fehlenden gemeinsamen Verantwortlichkeit i.S.v. Art. 26 Abs. 1 DS-GVO geschlossen wurde, sieht er weder Kontroll- noch Auskunftsrechte der Einrichtungen bzw. der MPDLS GmbH vor. Allerdings haben Wissenschaftlerinnen und Wissenschaftler die Möglichkeit, über Datenweiterverarbeitung und das Datenschutzniveau detaillierte Auskünfte einzufordern (Art. 15 DS-GVO). Daher sind Auskunftsanfragen bzw. Klagen durch Wissenschaftlerinnen und Wissenschaftler ein möglicher Weg, die rechtliche Grauzone aufzuheben. (Altschaffel et al., 2024, 32f. Herv. YF)

Festhalten lässt sich, dass die DEAL-Vertragsbedingungen *keine* Möglichkeit der „Kontroll- noch Auskunftsrechte der Einrichtungen bzw. der MPDLS GmbH“ beinhalten. Die Frage, ob Daten europäischer Wissenschaftler:innen und Studierenden außerhalb Europas verarbeitet werden und unter welchen Bedingungen, bleibt der Selbstauskunft überlassen. Eine solche Anfrage nach Artikel 15 DS-GVO bei Elsevier habe ich am 28.03.2024 gestellt, die Auskunft steht auch zum Abgabezeitpunkt dieser Arbeit (Juli 2024) noch aus, obwohl sie spätestens innerhalb von vier Wochen erfolgen muss. Die Dokumentation der Anfrage und des Austauschs mit Elsevier findet sich im Anhang.

## 4.3.2 Involvierung von Risk-Services

### Datenweitergabe

Die kursivierten Passagen der Privacy Policy oben zeigen, an welchen Stellen Formulierungen mit einem geschickten Gebrauch von „and“ und „and/or“ (bei dem unklar bleibt, ob es sich um ein einschließendes oder ausschließendes „und“ bzw. „oder“ handelt) die Beteiligung von Drittanbietern („obtain [...] other information about you from other entities within RELX oder from third parties“) und die Weitergabe von Daten („share your personal information with our group companies“) implizieren oder nur nahelegen. Wenn man bedenkt, dass die Risk-Sparte von RELX auch die Sicherheitsfunktion erfüllt (s. Ausführungen über die Präsenz auf der ScienceDirect-Website unten), ist es naheliegend, aber nicht explizit ausgedrückt, dass es sich bei den Sicherheitsmaßnahmen um LexisNexis-Services handelt. LexisNexis' Selbstdarstellung, wie sie oben ausgeführt wurde, erlaubt eine weite Auffassung davon, welche Instrumente und Technologien für die Aufgaben „prevent security, fraud or technical issues“ oder „protect the security of our systems“, eingesetzt werden, z. B. durch Anlegen sehr umfangreicher Profile für Hunderte Millionen Menschen in LexIDs, in „distinct individual dossiers of 99,99% accuracy“.

Handelt es sich um Beweise, dass RELX seiner Risk-Sparte vollen Zugang zu den Daten gibt, die durch Elseviers Produkte und ScienceDirect gesammelt werden? Nein. Wie bereits festgehalten wurde, ist ein solcher Beweis ohne den Zugang zum System der Gegenseite nicht möglich. Die von verschiedenen Stellen, u. a. im SPARC-Bericht gesammelten Indizien

sprechen jedoch dafür, dass personenbezogene Daten nicht bei Elsevier bleiben. Sie werden für “behavioral advertising and targeted marketing purposes” (SPARC, 2023, 24) an Dritte weitergegeben sowie auch an andere Elsevier-Services. Daten, die zur Identifizierung von Personen führen können, werden durch Cookies und Zählpixel an Drittparteien (Google, Adobe, Cloudflare, New Relic und andere) weitergegeben (ebd.). Spannend ist hier Wolfie Christls Twitter-Beitrag am 18.8.2020 (ebd., FN31) „Does RELX, the Scientific Publisher, Use Personal/Behavioral Data on Academic Scholars Who Access Publications via Elsevier for Its Data Brokerage and Risk Management Services? I Don’t Know. In Any Case, ScienceDirect (RELX/Elsevier) Embeds ThreatMetrix (RELX/LexisNexis Risk).<sup>159</sup>.”

Mittlerweile hat Elsevier ThreatMetrix aus der Liste der notwendigen Cookies auf ScienceDirect entfernt. Jedoch schließt zum einen die Begründung Elseviers<sup>160</sup>, ThreatMetrix habe *primär* eine Schutzfunktion der Website gegen Betrug, nicht aus, dass ThreatMetrix oder andere LexisNexis-Unternehmen gleichzeitig *auch andere* Funktionen erfüllen. Zum anderen weist der SPARC-Bericht zurecht darauf hin, dass die Abwesenheit bestimmter Domains, die mit ThreatMetrix – oder eben anderen LexisNexis Unternehmen – assoziiert werden, noch lange nicht beweist, dass es *nicht* zur Weitergabe von Daten über ScienceDirect kommt:

The 2022 analysis did not include domain names associated with ThreatMetrix—nav.sciencedirect.com, content.id.elsevier.com, or the h.online-metrix.net—in their first and third-party host lists. The 2023 analysis did find one domain name that is associated with ThreatMetrix (h-elsevier-fame.online-metrix.net; see Figure C-9 in Appendix C) on the login page. Regardless, the absence of these domains from the home, browsing, search results, and article pages does not necessarily mean that Elsevier does not feed ThreatMetrix data collected from ScienceDirect users by other means. (SPARC, 2023, 22)

<sup>159</sup> <https://t.co/BqoO38tojQ>; <https://twitter.com/WolfieChristl/status/129565504074144563>. [beide zuletzt geprüft am 25.07.2024]

<sup>160</sup> [https://web.archive.org/web/20221017174021/https://service.elsevier.com/app/answers/detail/a\\_id/28037](https://web.archive.org/web/20221017174021/https://service.elsevier.com/app/answers/detail/a_id/28037). [zuletzt geprüft am 25.07.2024]

Dieses Argument ist mit Bezug auf die Server-Side-Tracking-Möglichkeiten umso überzeugender, wenn man beachtet, wie umfangreich der Bestand der First-Party-Cookies mit Elsevier-Domain<sup>161</sup> auf der Cookie-Management-Plattform von Elsevier ist.

Auch nach der Entfernung des expliziten Bezugs auf ThreatMetrix aus der Liste unbedingt notwendiger Cookies im Juni 2022 taucht in späteren Nutzungsvorgängen der weiteren Wissenschafts-Services ThreatMetrix wieder auf: „users who wish to log into ScienceDirect encounter ThreatMetrix on the login page at id.elsevier.com. [FN33] The id.elsevier.com domain is used when authenticating user accounts for other Elsevier products, such as Mendeley and Scopus. [FN34]<sup>162</sup>“ (SPARC, 2023, 22) Die Fußnote 33 dokumentiert zwei Versionen der ScienceDirect „Use of Cookies“-Seite<sup>163</sup>, eine mit der Erwähnung von ThreatMetrix<sup>164</sup> und eine ohne<sup>165 166</sup>. Aktuell gibt es keine eigene „Use of Cookies“-Seite mehr auf ScienceDirect, hinter dem Link folgt eine Weiterleitung auf die Elsevier Cookie Notice.<sup>167</sup> Hier lohnt sich ein Vergleich der deutschen und englischen Versionen in Tabelle 7. Auch hier liegt die Vermutung nahe, dass ThreatMetrix aktiv an den Elsevier-Produkten beteiligt ist.

<sup>161</sup> <https://www.elsevier.com/>. [zuletzt geprüft am 25.07.2024]

<sup>162</sup> Die FN34 dokumentiert die Weiterleitung von der Mendeley-Seite [https://web.archive.org/web/20221102212747/https://service.elsevier.com/app/answers/detail/a\\_id/19574/supporthub/mendeley/kw/cookies/p/16075/https://www.scopus.com/signin.uri?origin=&zzone=TopNavBar](https://web.archive.org/web/20221102212747/https://service.elsevier.com/app/answers/detail/a_id/19574/supporthub/mendeley/kw/cookies/p/16075/https://www.scopus.com/signin.uri?origin=&zzone=TopNavBar) auf eine <https://id.elsevier.com> login page.

<sup>163</sup> <https://www.sciencedirect.com/legal/use-of-cookies>. [zuletzt geprüft am 25.07.2024]

<sup>164</sup> <https://web.archive.org/web/20220617003826/https://www.sciencedirect.com/legal/use-of-cookies>. [zuletzt geprüft am 25.07.2024]

<sup>165</sup> <https://web.archive.org/web/20220617234129/https://www.sciencedirect.com/legal/use-of-cookies>. [zuletzt geprüft am 25.07.2024]

<sup>166</sup> Dort blieb zunächst noch Cloudflare, das aber in der aktuellen Version auch nicht mehr vorhanden ist, dafür aber im Cookie-Management zu finden ist.

<sup>167</sup> Ähnlich sieht es auch bei einem anderen Verweis auf ThreatMetrix aus, der in SPARCs Landscape-Analyse (2021a, 15) mit Referenz auf ScienceDirect (jedoch ohne Einzelnachweis) beschrieben wird: „ScienceDirect links directly to ThreatMetrix’s (a RELX company) processing notice to describe how collected data are used: ‘As explained in our privacy policy, sciencedirect.com and Linkinghub.elsevier.com collect information through the use of cookies or similar technologies...Security cookies and related technologies, such as those provided by ThreatMetrix, are used to maintain online security and protect our website against fraud and abuse.’”

TABELLE 7: Cookie-Erklärung, deutsch (mit ThreatMetrix) und englisch (ohne ThreatMetrix).

| Elsevier-Website | Cookies und ähnliche Technologien                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Cookie Notice                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Wortlaut</b>  | <p>„Verwendung von Zählpixeln und anderen Technologien</p> <p>Einige unserer Websites, Anwendungen und elektronischen Mitteilungen enthalten elektronische Kennzeichnungen, die als Zählpixel, Gifs oder Pixel-Tag bekannt sind, eindeutige Identifikatoren und ähnliche Technologien, die dabei helfen, Cookies zu liefern, die Online-Aktivität zu messen, zielgerichtete Werbeanzeigen bereitzustellen oder die Effektivität unserer Werbekampagnen oder anderer Vorgänge zu analysieren. Wir verwenden außerdem Technologien von Drittanbietern von Datensicherheit wie zum Beispiel Threatmetrix, um Online-Sicherheit zu gewährleisten und unsere Websites und andere Dienste vor Betrug und Missbrauch zu schützen.“</p> | <p>„Use of web beacons and other technologies</p> <p>Some of our websites, applications and electronic communications contain electronic tags known as web beacons, gifs or pixel tags, unique identifiers and similar technologies that help deliver cookies, measure online activity, provide more relevant advertising, or analyze the effectiveness of our promotional campaigns or other operations. We may also use technologies from third-party data security providers to maintain online security and protect our websites and other services against fraud and abuse.“</p> |
| <b>Zeitpunkt</b> | <p>Zuletzt aktualisiert: 7. Dezember 2021 <a href="https://www.elsevier.com/legal/cookie-notice-de-de">https://www.elsevier.com/legal/cookie-notice-de-de</a> am 01.07.2024 [Herv. YF] Vgl. Screenshot in Abbildung 24 im Anhang</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | <p>Last updated: 15 June 2020 <a href="https://www.elsevier.com/de-de/legal/cookie-notice">https://www.elsevier.com/de-de/legal/cookie-notice</a> am 6.7.2024 Vgl. Screenshot in Abbildung 25 im Anhang</p>                                                                                                                                                                                                                                                                                                                                                                           |

Das Data Privacy Assessment von SPARC hält auch Punkte fest, bei denen ScienceDirect einen “Minimum Viable Privacy MVP” nicht einhalten kann. Darunter sind große Mengen gesammelter Daten und diverse Sammlungsorte samt der damit verbundenen Gefahren sowie die Notwendigkeit der Aktivierung von JavaScript, um die Website nutzen zu können und Open-Access-Artikel herunterzuladen<sup>168</sup>. Ein weiterer Punkt ist die mögliche Sammlung der verhaltensbiometrischen und weiteren

personenbezogenen Daten durch ThreatMetrix beim Account-Login-Prozess sowie die Default-Aktivierung für Recommendations Service, Reading History und Search History (SPARC, 2023, 23).

Die Tabelle 8 zeigt, welchen Informationsbestand RELX mit dem Risk-Marktsegment und seinen Produkten hat und welche problematischen Konsequenzen daraus entstehen könnten.

TABELLE 8: RELX-Informationsbestand Risk.

| RELX-Marktsegment                                                     | Personenbezogene Daten                                                                                                   |                                                                                                                                                                           |                                                                                                                                                                                                           | Anwendungsmöglichkeiten, Konsequenzen                                                                                                                      |
|-----------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                       | <i>LexisNexis Identity Verification Solution: LexID® Digital</i> <sup>169</sup> <i>View</i>                              | <i>ThreatMetrix</i> <sup>170</sup> ; <i>ThreatMetrix SmartID®TrueIP</i>                                                                                                   | <i>BehavioSec:</i>                                                                                                                                                                                        |                                                                                                                                                            |
| Risk<br>LexisNexis®<br>Risk Solutions<br>ThreatMetrix®<br>BehavioSec® | LexisNexis® Digital Identity Network® mit einem „Unique, Tokenized Customer Identifier“ der “[c]onsistently identifies a | „erkennt akkurat die Nutzung von Standorten und verschleierten Identitäten, wie versteckte Proxy-Server und VPNs, und bietet [...] die Möglichkeit, die echte IP-Adresse, | eindeutige Identifizierung über Geräte hinweg durch Verhaltensbiometrik z. B. „keystroke“, „accelerometer“, „gyroscope“, „mouse „movement“, „surface“, „motion, „hitzone“, „pressure“ <sup>171</sup> “The | Unerlaubte Verwendungs sensibler Daten;<br><br>verhaltens-biometrische Identifizierung;<br><br>Überwachung mit quasi-nachrichtendienstlichen Technologien; |

<sup>168</sup> Inwiefern es sich bei solchen Hürden um Open Access in einem akzeptablen Sinne handelt, ist eine andere Frage.

<sup>169</sup> <https://risk.lexisnexis.com/global/en/insights-resources/article/lexid-digital>. [zuletzt geprüft am 25.07.2024]

<sup>170</sup> <https://risk.lexisnexis.com/global/de/products/threatmetrix>. [zuletzt geprüft am 25.07.2024]

|                                                                                                                                                                                                                            |                                                                                                                                                                                                                                          |                                                                                                                                                                          |                                                                                                                    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| person irrespective of changes in devices, locations or behavior. [...] reliably recognize the same returning user behind multiple devices, email addresses, physical addresses and account names.” LexisNexis (n.d.a, 1). | den Standort [...] zu sehen.“ LexisNexis (2019, 5) Trust Tags ermöglicht „vertrauenswürdige und dubiose Nutzer, Geräte, Standorte und Personas zu definieren, zu kategorisieren, zu taggen und zu differenzieren.“ LexisNexis (2019, 4). | combination of typing rhythm, touch pressure, swiping patterns, device orientation and other metrics help confirm identity with pinpoint precision [...]” <sup>172</sup> | Datenhandel<br><br>“Clear understanding of identities across systems, agencies, and states.” (LexisNexis, 2023, 2) |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|

DEAL-Datenschutz-Passage

Der DEAL-Vertrag (Elsevier & MPDL Services GgmbH, 2023) beinhaltet mehrere Punkte, die sich von den Elsevier Privacy Policy unterscheiden und als Erfolg der DEAL-Verhandlungsgruppe verbucht werden (s. für eine detaillierte Darstellung Altschaffel *et al.* (2024)). Diese Passagen betreffen zum einen den Punkt 7.6.3 des DEAL-Vertrags und wurden oben von Joschka Selinger diskutiert. Außerdem enthält der DEAL-Vertrag noch folgende Einschränkung, die leider einigermaßen vage bleibt:

This restriction on personal data processing by Elsevier includes but is not limited to the collection, analysis, profiling and aggregation of personal data, such as through utilization of cookies, device finger-

171 <https://risk.lexisnexis.com/products/behaviosec>. [zuletzt geprüft am 25.07.2024]

172 <https://risk.lexisnexis.com/products/behaviosec>. [zuletzt geprüft am 25.07.2024]

printing technology, or similar technologies that track user behavior that are outside the services under the Agreement and the rights and obligations of the Parties under the Agreement.

Zum anderen ist der Punkt 7.6.5. relevant:

a. Elsevier shall refrain from storing any full IP addresses for a time period longer than necessary for the purpose of the processing. Elsevier will utilize encryption technology as appropriate. b. Elsevier shall implement a cookie management platform on the website of the Products to enable Authorized Users to view the nature and duration of cookies and manage their choices with respect to cookies that are not strictly necessary to provide, manage, operate and secure the website of the Products or otherwise for the performance of this Agreement. c. Elsevier shall refrain from using technologies allowing the recovery of cookies or similar technologies deleted by the users in their browsers. (Elsevier & MPDL Services GGmbH, 2023)

Der Punkt 7.6.6 beinhaltet nur die Ankündigung eines gemeinsamen-Workshops “discuss further steps to strengthen data protection and data privacy”. Wie von verschiedenen Stellen kritisiert wurde, sind diese Tracking-Einschränkungen nicht überzeugend, weil es für ihre Implementierung oder Überprüfung keine Mechanismen oder Druckmittel gibt.

Aktuell kann folgendes festgehalten werden: 7.6.5:a. ist nicht überprüfbar; b. es gibt eine Cookie Management-Plattform, die tatsächlich kurze Erklärungstexte zu den Cookie-Arten enthält; c. Ob und inwiefern sich Elsevier an diese Zusage hält, müsste durch IT-forensische Untersuchungen nachgewiesen werden. Dafür müsste aber zunächst hinreichend geklärt werden, was sich hinter „similar technologies“ verbirgt bzw. um welche konkreten Technologien es sich handelt, die unterlassen werden sollen im Gegensatz dazu, welche konkreten Technologien weiterhin angewendet werden sollen. Beides ist intransparent und hinter vagen Ausdrücken versteckt.

### 4.3.3 Einfluss auf den Wissenschaftsbetrieb

Die Tabelle 9 zeigt den RELX-Informationsbestand für einen einzigen Bereich des STM-Marktsegments. Die drei Spalten *personenbezogene Daten*, *wissenschaftliche Daten* und *Wissenschaftsmanagement* halten fest, auf welche Bereiche bzw. institutionelle Ebenen des Wissenschaftssegments sich Produkte von Elsevier und damit Quellen für die Datensammlung erstrecken. Ganz unten werden mögliche negative Konsequenzen sowohl des Trackings als auch Elseviers Präsenz in den unterschiedlichen institutionellen Ebenen skizziert.

TABELLE 9: RELX-Informationsbestand Elsevier (STM).

|                                                       |                                                                                                                                                                                                                      |
|-------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>RELX Marktsegment</b>                              | Science, Technical & Medical, Elsevier (Ohne die Sparten Government, Corporate und Health Organisations)                                                                                                             |
| <b>Personenbezogene Daten</b>                         | First-Party-Daten: Wissenschaftler:innen, Studierende, Angehörige von Wissenschaftseinrichtungen, Bibliotheken<br>ScienceDirect, Scopus, Mendeley, Expert Lookup, Pure                                               |
| <b>Wissenschaftliche Daten, Forschungsdaten</b>       | Scopus, Pure, Mendeley, DataSearch, Journal Metrics, Plum X, SSRN, Cell Press, Expert Lookup, Data Monitor, Digital Commons                                                                                          |
| <b>Wissenschaftsmanagement, Institutionen</b>         | SciVal, Pure, Interfolio, Funding Institutional, Analytical Services                                                                                                                                                 |
| <b>Anwendungsmöglichkeiten für RELX, Konsequenzen</b> | Beeinflussung von Forschung und Wissenschaftskommunikation; Epistemische Überlegenheit durch Steuerungswissen und Trackingwissen; Vernetzung der verschiedenen Sparten („risk“ Produkte); Asymmetrischer Vorteil ge- |

genüber Institutionen; Etablierung von Brokering als Teil des Wissenschaftszyklus

Der Ausschuss der wissenschaftlichen Bibliotheken und Informationsdienste des DFG fasst die Schieflage zusammen:

Die skizzierte Entwicklung hin zu einer privatwirtschaftlichen Wissensindustrie [...] steht im Gegensatz zur Wissenschaftsfreiheit, zum gesetzlich vorgegebenen Umgang mit personenbezogenen Daten und zum Wettbewerbsrecht. Im Einzelnen kann unreguliertes bzw. unerkanntes Datentracking ° eine Verletzung der Wissenschaftsfreiheit und der Freiheit von Forschung und Lehre bedeuten; ° eine Verletzung des Rechts auf den Schutz der eigenen Daten darstellen; ° eine potenzielle Gefährdung von Wissenschaftlerinnen und Wissenschaftlern darstellen, da die Daten auch ausländischen Regierungen und autoritären Regimes zugänglich werden können; ° einen Eingriff ins Wettbewerbsrecht darstellen, da neue Teilnehmer kaum eine Chance auf einen Markteintritt haben; ° eine Wertminderung öffentlicher Forschungsinvestitionen begünstigen, da im Rahmen von Wirtschaftsspionage wissenschaftliche Aktivitätsdaten von kommerziellen Forschungskonkurrenten erhoben oder ihnen gegen Bezahlung zugänglich gemacht werden können. (DFG AWBI, 2021)

Lisker (2023) beschreibt das Datensammeln u. a. als *unausweichlich, kontinuierlich, unsichtbar* und *prädiktiv*. In Anschluss an Lauer (2022a; 2022b), Brems, Siems und SPARC füge ich der bisherigen Argumentation hinzu, dass Tracking in der Wissenschaft *lukrativ, gefährlich, asymmetrisch* und *zersetzend* für die Institution Wissenschaft sowie für die digitale Souveränität (Deutschlands und der EU) ist.

Trackingpraktiken sind *lukrativ*, weil „die Beobachtung des Nutzungsverhaltens [...] der ökonomisch noch viel lohnendere Teil [ist], der sich in immer weitere Bereiche ausdehnt – Rechtsinformationen, Finanzdaten und Nachrichtengeschäft.“ (Lauer, 2022b) Mittlerweile gibt es einen eigenen Markt für die Verwertung der Wissenschaftsdaten. Die Integrationsmöglichkeiten der Anwendungen und Plattformen erlauben wachsende Möglichkeiten der Monetarisierung<sup>173</sup>: “Publishers sit on a goldmine of data. Most of it goes to waste – *but it doesn’t have to!*” (hum.works, n.d.)

Trackingpraktiken sind *gefährlich*, weil die gesammelten Informationen und Profile an eine unbekannte und unkontrollierbare Menge an Drittparteien mit finanziellen, politischen und kriminellen Interessen weitergegeben werden. Zudem involvieren sie speziell bei RELX

die immer weitergehende Verflechtung mit der Sicherheitsindustrie. Mögen Firmen wie Palantir allgemein nach Schwefel riechen, es sind die LexisNexis Risk Solutions von RELX und die Thomson Reuters Special Services, die in Tausenden von Einrichtungen laufen, Millionen Menschen durchleuchten und so das ‘American Dragnet’ schaffen [...]. (Lauer, 2022b)

**173** „Die Plattformen, die diese Metriken zur Verfügung stellen – namentlich das heute von Clarivate betriebene Web of Science – ziehen daraus beträchtliche finanzielle Gewinne [...]. Für die Forschenden kann das zukünftig einerseits heißen, dass sie sich in ihrer gesamten Tätigkeit von der Mittelakquise und der ersten Abspeicherung von Labordaten bis zur Veröffentlichung ihrer Resultate und der Präsentation ihrer metrischen Erfolgsbilanz – vom Datenverwaltungssystem eines einzigen Verlagshauses abhängig machen. Andererseits müssen sie sich darauf einstellen, dass virtuell ihr gesamtes wissenschaftliches Such- und Nutzungsverhalten verfolgt, gespeichert, ausgewertet und gegebenenfalls an interessierte Dritte weitergegeben wird, solange niemand aufwändig nachweist, dass dies ihre Persönlichkeitsrechte verletzt.“ (Battenberg et al., 2015, 13)

„Es geht [beim Tracking] um ein einträgliches Geschäft. Die Pressesprecher großer wissenschaftlicher Verlage räumen das ja selbst ein, wenn sie vom Trend der Branche weg vom Handel mit Inhalten und hin zum Handel mit Daten reden. Und wenn man sich anschaut, welche Unternehmen die großen Verlage in den vergangenen Jahren dazugekauft haben, ob Elsevier, Springer Nature oder Wiley, dann passt das zu der Strategie, künftig Dienstleistungen anzubieten, die den gesamten Workflow von Wissenschaftlern abdecken. Von der Artikelrecherche über das Sammeln und Aufarbeiten von Daten bis hin zum teamübergreifenden Schreiben an einem gemeinsamen Paper: Das Ziel ist, ein auf den einzelnen Wissenschaftler oder seine Institution genau zugeschnittenes Paket an Inhalten und Anwendungen zu verkaufen. Elsevier ist in der Hinsicht am weitesten. Wenn die das durchsetzen, dann hätten wir in der Wissenschaft ein Monopol, das dem von Facebook, Microsoft und SAP in einer Firma entsprechen würde.“ (Brembs & Siems, 2020)

Neben der Gefahr von Wissenschaftsspionage und geopolitischen Nachteilen sind die Verwicklungen von RELX mit ICE problematisch (Siems, 2024c) und nutzen aus, dass die Gesetzgebung staatliche Eingriffe besser reguliert als Konzerne: „Diverse Behörden und Strafverfolgungsorgane zum Beispiel dürfen derartige Daten nur mit einem Durchsuchungsbeschluss selbst erheben; beim Erwerb von Daten sind solche Beschränkungen gewöhnlich nicht vorgesehen.“ (Brembs, 2023a) Siems warnt mit Tim Hwang vor der anstehenden „Subprime Attention Crisis“, der nächsten großen Krise, nur dieses Mal im Bereich der (digitalen) Aufmerksamkeitsökonomie und nicht der Immobilienkredite. Diese Krise sei „ein Vehikel für vielerlei: politische und geopolitische Interessen, Terrorfinanzierung, militärische Aktivitäten, die blühende Sicherheitsindustrie, Radikalisierung, Ausbeutung und Völkermord – ein Instrument für all das schwillt in dunkler Opazität [...] zu immer bedrohlicherer Größe heran.“ (Siems, 2024b, 2)

Die *Asymmetrien* sind *epistemisch*, weil der Wissensvorsprung RELX' umfassend, kontinuierlich und intransparent ist. Der Vorsprung würde dem Konzern im Falle der Implementierung auf Institutionsebene ermöglichen, auf wissenschaftlicher, administrativer und wirtschaftlicher Ebene Fakten zu schaffen, zu deren Überprüfung es nur Konzern-Instrumente geben würde. Bereits länger in der Diskussion sind Asymmetrien auf der *Machtebene*, die Elsevier und den anderen Konzernen durch Marktdominanz und die Kontrolle des Publikationssektors einen ungleichen Vorteil gegenüber Einzelpersonen, Einzelinstitutionen oder sogar einzelnen Ländern geben: „If you control access to the scientific literature, it is, to all intents and purposes, like controlling science.“ (Buranyi, 2017) Es gibt zudem eine *funktionale* Asymmetrie, weil – wie oben argumentiert – Elsevier, Clarivate, Wiley und andere Konzerne mit ihren von den Wissenschaftseinrichtungen finanzierten Gewinnen jahrelang massiv in Anwendungen für den Forschungsbereich investieren konnten, die heute für die Forschungspraxis unentbehrlich sind.

Tracking und die hier beschriebenen institutionsebenenübergreifenden Anwendungen sind *zersetzend* für die Wissenschaft und digitale Souveränität, weil sie in die Struktur und Inhalte<sup>174</sup> der Wissenschaftskommunikation eingreifen<sup>175</sup> und durch die Plattformisierung und Kommodifizierung zum Verlust von Qualität<sup>176</sup> führen. Die Ziele<sup>177</sup> der Konzerne und der Wissenschaft sich grundsätzlich verschieden und geraten notwendig

in Konflikt, wenn Macht- und Wissensasymmetrien vorliegen. Sie führen durch ihre Geschäftsmodelle dazu, dass Brokering zum Teil des Wissenschaftsprozesses wird (Lamdan, 2022) und Wissenschaftseinrichtungen in einem Gefälle aus Abhängigkeiten gegenüber dem Konzern bleiben: Rankings, Research Assessment, Förderung, Forschung, Profile (Erpressbarkeit von Personen und Institutionen).

(Weitere) „Transparenzkrisen“ (Holzer, 2022) drohen in einer ganzen Reihe von Bereichen, in denen die Leitlinien wissenschaftlicher Praxis (DFG Deutsche Forschungsgemeinschaft, 2022) Transparenz verlangen. Darunter fallen die Nachvollziehbarkeit von Methoden (Holzer, 2022, 171) über die Gewinnung und Nutzung von Daten über den Forschungsprozess und Forscherinnen (Holzer, 2022, 171) und Punkte, „wo die digitale Forschungspraxis und die auf digitale Informationen gestützte Bewertungspraxis zusammentreffen.“ (Holzer, 2022, 170) Im weiteren Sinne betrifft das auch die Intransparenz der gesamten trackingbasierten Internetökonomie (Siems, 2024b). Ferner ist ein Vertrauensverlust in Bibliotheken zu erwarten, wenn diese trotz breit angelegter Trackingmaßnahmen

---

**174** “If unchecked, Elsevier may well find itself in the position of defining, on behalf of the academic community, what constitutes good research and then selling the tools to perform the actual assessment.” (SPARC, 2020, 22)

**175** “[I]t is apparent that the platformisation of scholarly information is affecting research culture, research integrity and, most importantly, the authority as to what is knowledge or information.” Ma (2023b, 10)

**176** “As some publishers become platform owners, they boost and boast the quantity of scholarly information with minimal concerns about quality. It is because more publication- and citation-based data can be generated if there are more publications and interactions (Ma, 2023a; Pooley, 2022).” (Ma, 2023b, 3)

**177** Ziel der Plattformen ist “that it is not science, universities or research funders who determine which research is the ‘right’ one and how much it ‘counts’[...] but rather the platform who makes the determination.” (Siems, 2021)  
 Siehe auch SPARC: “As a research analytics provider, Elsevier’s incentive is to extract as much data as possible from users and repackage it into metrics that can be sold back to institutions [...] This dynamic has the potential to undermine academic freedom in a number of critical ways [...] For example, detailed tracking of faculty activity could (understandably) give researchers the sense that someone is always looking over their shoulder, analyzing their activity, and making assessments that could impact their career prospects. Faculty may also reasonably believe that Elsevier analytics products may assess the company’s own publications more favorably than those of competitors. These concerns could harm academic freedom by influencing researchers’ choice of publication venue, away from what is most appropriate and toward what they believe will be assessed most highly by Elsevier.” (SPARC, 2022b)

men mit den spezifischen Verlagen und Konzernen zusammenarbeiten. Siems fragt zurecht, weshalb diese zentralen Punkte unter dem Radar bibliothekarischer Aufmerksamkeit bleiben: „[I]st es dann das in den Grundsätzen niedergelegte Verständnis von Integrität und Fachkompetenz, dass die Bibliotheksverbände sich hierzu nicht äußern und Verhandlungsführer sich mit teils eher halbgaaren Datenschutzklauseln zufrieden geben?“ (Siems, 2024b, 8)

Es droht der Verlust von Steuerungswissen (Holzer, 2022, 174), der Wissenschaftsfreiheit<sup>178</sup> und der Hoheit über eigene Daten (Brembs & Siems, 2020) sowie ein Zustand dauerhafter epistemischer Asymmetrie<sup>179</sup>. Siems hält fest, dass Informationsverhalten von Hochschulangehörigen personalisiert beobachtet wird und diese Daten nicht bei den Verlagen bleiben, sondern von Data Brokern aggregiert und verwertet werden. Das Verhalten von Wissenschaftler:innen werde Teil einer durchgehenden Onlinebeobachtung und Auswertung, in deren Zuge ganzheitliche Onlinebiografien erstellt werden, die demographische Angaben, Social-Media-Verhalten, Einkaufsverhalten etc. beinhalten (Siems, 2024c, 2022).

## 4.4 (Digitale) Souveränität in der Wissenschaft

Die unterschiedlichen Weisen der Einflussnahme auf den Wissenschaftsbetrieb durch aktuelle und mögliche weitere Produkte von Elsevier wurden an mehreren Stellen aufgezeigt. In Anschluss an diese Argumente

---

**178** “Besides these individual consequences, there is a strong systemic risk to free and autonomous research as an ‘institution’. Academic freedom not only secures individual liberties, but, importantly, also requires states to create the conditions and circumstances under which these liberties can be effectively exercised.” Kunz (2022)

**179** Am Beispiel der Plattform GetFTR zeigt sich das Risiko der epistemischen Asymmetrie – es gibt auf der wissenschaftlichen Seite nicht genug Wissen darüber, um dieses kommerzielle Angebot einschätzen zu können: “Ultimately, the academic community does not have the information required to assess whether the unique benefits to researchers outweigh these potential significant risks that we’ve outlined.” SPARC (2020, 18)

möchte ich die Fragestellung der Arbeit aufgreifen, ob die Zusammenarbeit mit Konzernen, für die exemplarisch Elsevier/RELX steht, mit einer unabhängigen Wissenschaft vereinbar ist. Dafür habe ich eingangs zwischen zwei Verständnissen unterschieden, einer schwachen These I und einer starken These II. Welche Ergebnisse lassen sich aus der bisherigen Diskussion in Anwendung auf die Konzepte der Souveränität und Privatheit ziehen?

Die beschriebenen Praktiken der Datensammlung und Verwertung von RELX verletzen die informationelle Selbstbestimmungsfähigkeit (mindestens) deutscher Bürger:innen. Die mögliche Ausweitung des Einflusses des Konzerns auf deutsche Wissenschaftsinstitutionen hätte gravierende negative Folgen für die (digitale) Souveränität in den verschiedenen Verständnissen, die im Teil II diskutiert wurden.

Für das traditionelle Verständnis von Souveränität folgt die Nichtvereinbarkeit der beschriebenen Praktiken mit Wissenschaft bereits aus der Formulierung der möglichen Folgen: Wer keine Kontrolle über und keinen Einblick in die zentralen Infrastrukturen, Prozesse und Bewertungsmechanismen besitzt, ist nicht souverän. Die Gefahren für digitale Souveränität, die Pohle und Thiel skizzieren, betreffen die Abhängigkeiten von digitalen Infrastrukturen und privaten Akteur:innen, die eine Vormachtstellung in einem Oligopol haben. Hier sind negative Auswirkungen der markt-libertären Position der USA und fehlende (staatliche) Regulierung von Überwachungspraktiken zu nennen. Die Gefahr einer Überhand nehmenden Kontrolle und Überwachung durch den Staat ist in diesem Zusammenhang weniger zu befürchten als die aktuell fehlende Regulierung der Anwendung von Überwachungstechnologien durch internationale Konzerne. Letztere verletzen die Rechte der Bürger:innen und können durch ihre Datenbestände einen Wissens- und Machtvorteil über staatliche Institutionen erlangen bzw. haben diesen eventuell bereits.

Eher als Pohle wäre hier Petra Gehrings Kritik zuzustimmen, dass die Rede von Datensouveränität genutzt wird, um die staatliche Verantwortung zur Regelung der Datenerhebung und -verwendung abzulösen und sie auf Individuen zu verlagern. Genau das passiert im DEAL-Vertrag, wo die Verantwortung auf Wissenschaftler:innen, Bibliotheken und Studierende verlagert wird, den jeweiligen Policies ‚freiwillig‘ zuzustimmen.

Wie Freiberg und Siems kritisieren, ist es gar nicht möglich, die Angebote unter Inanspruchnahme der eigenen gesetzlich sicherten Ablehnungsoption für (auch nur manche) Tracking-Formen in hinreichendem Maße zu nutzen. Der grundsätzlichere Punkt ist aber, dass es sich um Nutzungen handelt, die für die Arbeit bzw. Studium notwendig sind. Die Verpflichtung also, für diese notwendigen Handlungen ‚freiwillige‘ Zustimmung zum Datentracking zu geben und sich Trackingformen auszusetzen (durch die bloße Nutzung der Plattform), für die es gar keine Ablehnungsoption gibt, ist zutiefst problematisch. Dieser Punkt wird auch durch den Bericht des AWBI (Altschaffel *et al.*, 2024) gestützt.

## 4.5 Privatheit und Wissenschaft

In dieser Arbeit habe ich nicht eigens für den Wert von Wissenschaftsfreiheit und persönlicher wie wissenschaftlicher Autonomie argumentiert. Diese Werte habe ich als eine gemeinsame normative Grundlage behandelt. Unter dieser Voraussetzung ist es hinreichend nachzuweisen, dass Produkte und Praktiken diese Normen verletzen oder gefährden, um die Forderung zu begründen, entsprechende Produkte oder Praktiken zu meiden oder zu unterlassen.

Zunächst möchte ich etwas genauer auf den Zusammenhang zwischen Privatheit und wissenschaftlicher Autonomie eingehen. Privatheit bezieht sich als Kategorie auf Personen, nicht auf Institutionen. Inwiefern kann sie außerhalb der Datenschutzaspekte zusätzliche Anwendung finden? Über die besprochenen negativen Auswirkungen auf die Institution Wissenschaft hinaus möchte ich einen Aspekt ansprechen, der durch die Privatheitsverletzung gegenüber Wissenschaftler:innen *qua* Wissenschaftler:innen (und nicht z. B. als Angestellte in Unternehmen) eine spezifisch schädigende Wirkung hat. Diese Form der Schädigung kann man etwa wie folgt beschreiben. Als Wissenschaftler:in zu arbeiten bedeutet häufig, in einem hohen Maß frei und kreativ zu arbeiten. Dazu gehört, Themen und Methoden eigenverantwortlich nach eigenem Interesse und eigener fachlicher Einschätzung auszusuchen, Projekte zu konzipieren und zu gestalten, inhaltlich nur dem Wissenschaftsethos und der Ziel-

setzung der Forschungsaufgabe bzw. -förderer verantwortlich zu sein. Weil wissenschaftliche Arbeit häufig in hohem Maße selbstbestimmt verläuft, ist die Trennung zwischen Privatperson und Wissenschaftler:in in einem bestimmten Sinne fließend. Mögliche Einschränkungen und Manipulationen von Wissenschaftler:innen als *Personen* (z. B. durch invasives Tracking) bedeuten in diesem Sinne nicht nur die *direkte* Einschränkung ihrer Arbeit. Sie kompromittieren zudem *indirekt* ihre wissenschaftliche Integrität, die normalerweise die Anforderung an eine Arbeit ist, die – unter Befolgung strenger interner Kriterien – auch große Freiheitsgrade und Autonomie erlaubt. Wer sich als Person exponiert, vulnerabel oder überwacht wähnt, wird nicht in der Weise frei forschen können, wie es dem Wissenschaftsethos und dem Verständnis der Wissenschaft als gesellschaftlichem Wert entspricht. Eine Form von Autonomie, die wir eher mit starken Bürgerrechten im Privatbereich verbinden, kann für zumindest auf manche Bereiche wissenschaftlicher Arbeit angewendet werden:

This notion of autonomy and privacy, however, seems to involve a comparatively strong concept of the unity of the private and the public subject. This subject is autonomous if – in private and public alike – he lives as he has resolved to autonomously and under conditions of authenticity, if he pursues the plans he has developed on his own and with others [...]. (Rössler, 2005, 185)

In diesem Sinne sind Autonomie und Privatheit – zumindest für manche Bereiche wissenschaftlicher Arbeit – eine Voraussetzung für unabhängige Wissenschaft. Rössler identifiziert drei Dimensionen von Privatheit: i) „decisional privacy“, in der es um den Bereich von Handlungen und Entscheidungen eines Subjektes geht, ii) „informational privacy“, bei der es um die Grenzen des Wissens anderer über das Subjekt geht, schließlich die iii) „lokale Privatheit“, die das private Zuhause, den eigenen Rückzugsraum betrifft (Rössler, 2005, 9).

Betrachtet man die oben diskutierten Praktiken der kontinuierlichen Datensammlung und Profilierung, Metrisierung und Bewertung durch intransparente Dritte mit zum Teil nicht-wissenschaftlichen Interessen, durch legales oder illegales Tracking, zeigt sich eine systematische und umfassende Erosion von Privatheit und Autonomie. Es ist gerade Sinn dieser Praktiken, die lokale Privatheit des Rückzugsraums (iii) und die informational privacy (ii) aufzuheben. Wissenschaftler:innen werden als

private Personen erpressbar, so dass ihre decisional privacy im Arbeitsbereich (i) kompromittiert wird.

Das möchte ich an einem Teil der Tabelle 1 vertiefen, die bereits in der Diskussion der Privatheitsdimensionen von Mulligan, Koopman und Doty (2016) und der Anwendung auf die Wissenschaft weiter oben vorgestellt wurde. Hier ist der entsprechende Ausschnitt, die Sortierung erfolgt nach den Dimensionen theory und protection.

**TABELLE 10:** Analyse tool in Anwendung auf die Wissenschaft (basiert auf Tabelle 1).

| <b>Privacy dimension</b>        | <b>Interrogation</b>        | <b>Wissenschaftsaspekt</b>                                                    |
|---------------------------------|-----------------------------|-------------------------------------------------------------------------------|
| <b>theory: object</b>           | what's privacy for?         | Selbstbestimmung, Arbeit als Wissenschaftler:in, persönliche Sicherheit       |
| <b>theory: justification</b>    | why should this be private? | Persönlichkeitsrecht, Wissenschaftsfreiheit                                   |
| <b>theory: contrast concept</b> | what's not private?         | kompromittiert/manipuliert, gefährdet/vulnerabel, abhängig                    |
| <b>theory: exemplar</b>         | what's an example?          | Beeinflussung (Verhalten, Karriere), Gefährdung der Forschungsfreiheit        |
| <b>protection: target</b>       | privacy of what?            | persönliche Daten, Forschungsergebnisse, wissenschaftliche Profile, Kontakte  |
| <b>protection: subject</b>      | whose privacy is at stake?  | Wissenschaftler:innen, Studierende, Angehörige von Wissenschaftseinrichtungen |

In der Theorie-Dimension werden grundsätzliche Fragen (interrogation) geklärt, die den Privatheitsbegriff insgesamt präzisieren, von anderen Bereichen abgrenzen und legitimieren sowie Beispiele geben: what is priva-

cy for? why should sth. be private? what is not private? what is an example?

In Anwendung auf die Wissenschaft ergibt sich folgende Charakterisierung: Privatheit ist relevant für die Selbstbestimmung und Arbeit als Wissenschaftler:in sowie für die persönliche Sicherheit. Begründet wird Privatheit im wissenschaftlichen Bereich zum einen mit der Wissenschaftsfreiheit als einem Abwehrrecht gegen die Eingriffe des Staates. Zum anderen ist das Recht auf informationelle Selbstbestimmung ein Persönlichkeitsrecht. Das Gegenteil von Privatheit im Kontext wissenschaftlicher Arbeit wären Arbeitsbedingungen, in denen sich Wissenschaftler:innen als gefährdet und vulnerabel wahrnehmen. Sie als Wissenschaftler:innen können unter den Verdacht der Manipulierbarkeit geraten und ihre Entscheidungen und Urteile können als nicht mehr *prima facie* den wissenschaftlichen Normen und Zielen verpflichtet angezweifelt werden. Insofern wird ihre wissenschaftliche Integrität durch Handlungen *anderer* kompromittiert, nicht ihre eigenen.

Die Beeinflussungsmöglichkeiten können offensichtlich ausfallen (z. B. bei der Einflussnahme repressiver Regimes) oder wesentlich subtilere Formen annehmen, die schwierig nachzuweisen oder zu ahnden wären. Beachtet man die Mannigfaltigkeit der Quellen, die RELX allein<sup>180</sup> im STM-Bereich Elsevier aus seinen Produkten zusammenfließen lassen könnte, entsteht ein extremes epistemisches Ungleichgewicht. Tabelle 10 oben hält die verschiedenen Quellen fest: Konkrete Wissenschaftler:innen können erpressbar werden über a) personenbezogene Daten über sich selbst oder für sie relevante Kontakte aus dem beruflichen Bereich, b) über wissenschaftliche Daten und Forschungsdaten, die für sie relevant sind oder an denen sie beteiligt sind und vor allem über c) Daten und Einflussmöglichkeiten des Konzerns im Wissenschaftsmanagement, wo ihre Leistung bewertet und ihre Karrierefortschritte verhandelt werden. Zu a) kommt noch ein unabschätzbar großer Fundus an sensiblen Informationen über die Wissenschaftler:innen, der aus dem Brokering- und Risk-Bereich stammt.

---

<sup>180</sup> Wie anfangs erwähnt wird der Gesundheitsbereich in dieser Arbeit ausgespart, aber die Diskussion des „Dutch Deal“ (Schritt IV oben) macht Elseviers Wunschrichtung deutlich.

RELX hat also nicht nur Informationen über konkrete Themen, an denen die Wissenschaftlerin arbeitet und wo sie publiziert, als Herausgeberin oder Gutachterin arbeitet. Die laufend gesammelten Informationen geben RELX Einblick zum Beispiel darin,

- wo und mit wem sie an welchen Projekten welche Arbeitsschritte vollzieht,
- wie sie an ihrer Einrichtung eingeschätzt wird und eventuell welche finanziellen Mittel ihr zur Verfügung stehen,
- ob und wann – sogar von wem – für sie relevante Anträge begutachtet werden,
- wo sie sich um Förderungen oder neue Anstellung bewirbt.

Gleichzeitig besitzt RELX die Kapazitäten, mit LexisNexis ein extrem genaues Profil über das Privatleben dieser Wissenschaftlerin zu erstellen, welches “[c]onsistently identifies a person irrespective of changes in devices, locations or behavior. [...] reliably recognize the same returning user behind multiple devices, email addresses, physical addresses and account names” LexisNexis (n.d.a, 1). Diese Informationen können genutzt werden, um die Wissenschaftlerin zu beeinflussen, u. a.

- in bestimmten Zeitschriften zu publizieren,
- ihre Forschung in eine für den Konzern wünschenswerte Richtung zu lenken,
- bestimmte Bereiche weniger genau zu betrachten,
- bestimmte Arbeiten (Gutachten oder Review-Arbeit) zu übernehmen oder zu vermeiden,
- in Gremien und bei anderen beruflichen Entscheidungen für oder gegen bestimmte für den Konzern oder seine Partner relevanten Projekte, Produkte oder Prozesse abzustimmen.

Folgender Aspekt ist besonders wichtig: Der Schaden für die Wissenschaftlerin entsteht bereits mit dem Bestehen dieser informationellen Auslieferung. Es braucht nicht nachgewiesen zu werden, dass die Wissenschaftlerin eine ihrer professionellen Handlungen durch die Einflussnahme des Konzerns motivieren ließ, zum Beispiel durch die begründete Aussicht auf eine *dadurch* verbesserte personenbezogene Metrik oder Kooperations-Empfehlung der jeweiligen Elsevier-Anwendung (SciVal oder Scopus Expert Lookup). Entscheidend ist, dass der Verdacht der Befangenheit und Erpressbarkeit allein durch die epistemische und Machtasymmetrie begründet werden kann. Durch den hohen Grad an Freiheit in der inhaltlichen wissenschaftlichen Arbeit, der Mitarbeit in Gremien und Beteiligung an folgenreichen Entscheidungen wäre es sehr schwierig, Befangenheit festzustellen. Die Einflussnahme findet nämlich nicht als Bestechung im näheren Sinne des Wortes statt, sondern über Mechanismen, die bereits Teil des Wissenschaftsbetrieb sind und deren Objektivität und Integrität vorausgesetzt werden *müssen*, um diesen Betrieb aufrechtzuerhalten.

Gleichzeitig ist offensichtlich, dass Privatheit im Sinne von Rösslers Dimension informational privacy (ii) und der Privatheit des Rückzugsraums (iii) nicht aufrechterhalten werden kann, wenn die fragliche Wissenschaftlerin davon ausgeht, in jedem privaten Bereich ihres Lebens (und eventuell des Lebens ihrer Familie) für einen in ihrem Beruf wichtigen Player transparent zu sein. Es geht dabei nicht um das viel bemühte Argument, ob sie als Privatperson *oder* Wissenschaftlerin „etwas zu verbergen“ hat. Das ist in vielfacher Weise ein unbefriedigendes und irreführendes Scheinargument. Die Frage ist stattdessen, welche Bedingungen für eine Arbeit mit diesem Maß an Integrität und Autonomie vorliegen müssen. Das sind mindestens die Privatheit der Person in den Dimensionen ii) und iii).

Hier fehlt der Schutz gegen Eingriffe von Konzernen, der analog zum Abwehrrecht gegen staatliche Eingriffe in die Wissenschaft dienen würde. Wenn Privatheit und Autonomie im erörterten Sinne eine Bedingung für wissenschaftliche Arbeit – verstanden als unabhängige, den Zielen und intrinsischen Normen der Wissenschaft verpflichtete Arbeit – darstellen, dann verstoßen die Praktiken von Elsevier und RELX gegen diese Bedingung. Hält man weiterhin an diesem Verständnis von Wissenschaft fest,

ist eine Zusammenarbeit mit RELX und vergleichbaren Konzernen nicht möglich.

Tabellen 11 und 12 zeigen die spezifischen wissenschaftlichen Gefährdungen der Souveränitäts- und Privatheitsdimension.

**TABELLE 11:** Privatheitsproblematik für die Wissenschaft (Teil I).

| Gegenstand            | Beschreibung                                                                                                                                                                             | Privatheitsdimension(en)                                                                                                         | spezifisch wissenschaftliche Gefahren                                                                                                                                                                            |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Souveränität          | „Staatliche Sicherung der Selbstbestimmungsfähigkeit seiner Bürgerinnen und Bürger mit ihren unveräußerlichen Rechten“ (Pohle, 2020, 6)                                                  | theory: object                                                                                                                   | Verlust der Wissenschaftsfreiheit als Teil einer demokratischen Gesellschaft                                                                                                                                     |
| digitale Souveränität | Staatliche Sicherung <i>im digitalen Raum</i> der Selbstbestimmungsfähigkeit seiner Bürger:innen mit ihren unveräußerlichen Rechten, nationaler Wirtschaft, Verwaltung und Institutionen | protection: subject (Personen, juristische Personen)                                                                             | Verlust digitaler Souveränität als Forschungsstandort<br><br>Wissenschaftsspionage<br>Änderung des Forschungsverhaltens auf personale und institutioneller Ebene (Arbeitsgruppen, Institute, Universitäten etc.) |
| Wissenschaftsfreiheit | „Die Freiheit der Wissenschaft, Lehre und Forschung wird in Artikel 5 als ein defensives und konstitutives Individualrecht ohne Gesetzesvorbehalt garantiert. [...] Dieses robuste Ab-   | protection: target (Freiheit von Wissenschaft, Forschung und Lehre)<br><br>protection: subject (Personen, Praktiken und Institu- | Verlust wissenschaftlicher Souveränität<br><br>Verlust von Steuerungswissen Anfälligkeit für kommerzielle                                                                                                        |

|            |                                                                                                                                                                                                                                                                                                                 |                                                                                                                                                                                                                                                                                                                                                               |                                                                                                                                  |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
|            | <p>wehrrecht ist kein Jedermann-Recht. Es schützt die spezifischen Personen, Praktiken und Institutionen wissenschaftlicher Rede, Forschung und Publikation</p> <p>zuvorderst gegen staatliche Einflussnahmen, die auf eine Steuerung, Kontrolle und Sanktionierung der Wissenschaft zielen.“ (Özmen, 2021)</p> | <p>tionen wissenschaftlicher Rede, Forschung und Publikation)</p> <p>harm: offender (Staat(en), Unternehmen, Individuen)</p> <p>harm: action (Einflussnahmen, die auf eine Steuerung, Kontrolle und Sanktionierung der Wissenschaft zielen)</p> <p>provision: provider (Staat)</p> <p>provision: mechanism (Gesetz: Art. 5 GG(3) und Art. 13 EU GrCharta)</p> | <p>und politische Einflussnahme</p> <p>Verlust von Datensouveränität</p> <p>Verlust von Sicherheit für Wissenschaftler:innen</p> |
| Privatheit | <p>multidimensionales, normatives, (möglicherweise wesentlich umstrittenes) Konzept des Umgangs mit Informationen über Personen; nicht unabhängig von Technologien, Wertvorstellungen und gesellschaftlichen Praktiken bestimmbar</p>                                                                           | <p>protection: subject (Personen)</p> <p>protection: target</p>                                                                                                                                                                                                                                                                                               |                                                                                                                                  |

TABELLE 12: Privatheitsproblematik für die Wissenschaft (Teil II).

| Gegenstand                                                  | Beschreibung                                                                                                                                                                                              | Privatheitsdimensi-<br>on(en)                                                                       | spezifisch wissen-<br>schaftliche Gefahren                                                                                                                                                                                                                                                                                                          |
|-------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| informationelle Selbstbestimmung/ Daten-<br>schutzrecht     | Grundrecht/Persön-<br>lichkeitsrecht, perso-<br>nenbezogene Daten                                                                                                                                         | protection: target<br><br>protection: subject<br>(Personen)<br><br>provision: mechanism<br>(Gesetz) | Wissenschaftsunab-<br>hängigkeit leidet                                                                                                                                                                                                                                                                                                             |
| Personen und Institu-<br>tionen im Wissen-<br>schaftssektor | Wissenschaftler:innen<br>(Forschende & Lehren-<br>de), Angehörige von<br>Wissenschaftseinrich-<br>tungen und wissen-<br>schaftlichen Bibliothe-<br>ken, Studierende, Wis-<br>senschaftseinrichtun-<br>gen | protection: subject                                                                                 | Einschränkungen in<br>Arbeit, Forschung, Stu-<br>dium                                                                                                                                                                                                                                                                                               |
| (potenzielle) Bedro-<br>hung durch Akteur:in-<br>nen        | Internationale Konzer-<br>ne, Staat(en), anonyme<br>Bedrohungen,<br>hacking, Wissen-<br>schaftsspionage                                                                                                   | harm: offender                                                                                      | „Wer die Freiheit der<br>Wissenschaft ein-<br>schränkt, missachtet<br>oder verletzt, der ge-<br>fährdet die Wissen-<br>schaft als erkenntnis-<br>bezogene Praxis, aber<br>auch als Institution<br>von Forschung, Lehre<br>und Bildung, und nicht<br>zuletzt als gesellschaft-<br>lich anerkanntes und<br>gefördertes Subsys-<br>tem.“ (Özmen, 2021) |

|                                          |                                                                                                                                                   |                                |                                                                                                                      |
|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|----------------------------------------------------------------------------------------------------------------------|
| (potenzielle) Bedrohung durch Handlungen | Privatwirtschaftliche und „staatliche Einflussnahmen, die auf eine Steuerung, Kontrolle und Sanktionierung der Wissenschaft zielen“ (Özmen, 2021) | harm: action                   | Tracking, Aggregation von Informationen, Profilerstellung, Intransparenz: Vorenthalten/ Weitergabe von Informationen |
| Schutzberechtigt ggü Akteur:in           | Deutschland, EU-Recht                                                                                                                             | provision: provider            | öffentliche Einrichtungen kommen ihrer Schutz- & Fürsorgepflicht nicht nach; Wissenschaftsunabhängigkeit leidet      |
| Schutzmaßnahme                           | Art. 5 GG (3), Art.13 EU-GrCharta, informationelle Selbstbestimmung, DSGVO, (Data Service Act/Data Market Act)                                    | provision: mechanism<br>Gesetz | Stärkere gesetzliche Regulierung der Wissenschaft; bürokratische Hürden                                              |

Die These I steht für die Position, die die Zusammenarbeit mit den legacy publishers fortsetzen würde, wenn Verbesserungen z. B. der Datenschutz-Regelungen der DEAL- oder anderer Transformationsverträge stattfinden, bzw. sogar dann, wenn solche Verbesserungen in Aussicht gestellt werden, aber offenkundig rechtswidrige Tracking-Praktiken unterbunden werden. Tatsächlich entspricht die These I der aktuellen Haltung der Wissenschaftseinrichtungen in der Open-Access-Transformation. Wie hier argumentiert wurde, ist es jedoch nicht möglich, die Publikationspraxis von der internen Transformation Elseviers/RELX' zum Datenanalyse- und Brokering-Konzern zu isolieren. Als solcher verfolgt der Konzern Ziele, die auf einer systematischen Wissensasymmetrie beruhen und setzt sie durch Praktiken um, die mit Wissenschaftsnormen und Grundrechten nicht vereinbar sind. Die strategische Entwicklung von RELX geht zu einer zunehmenden Durchdringung und Beeinflussung der Wissenschaftsinstitutionen durch komplementäre Mittel. Eine *systematische* Verbesserung des Datenschutzes in weiteren Verträgen oder eine im wissenschaft-

lichen Sinne neutrale Leistung von den hier besprochenen Anwendungen zu erwarten hieße, die wesentlichen kommerziellen Zielsetzungen des Konzerns zu verkennen. Das spricht dafür, dass die These I nicht bestätigt werden kann und die These II bestätigt ist. Souveräne Wissenschaft ist mit Elsevier/RELX nicht möglich, weil seine Funktion als Datenanalyse-Konzern auf einer systematischen, u. a. epistemischen Asymmetrie beruht und seine Praktiken nicht mit einer unabhängigen Wissenschaft vereinbar sind.

Die These II entspricht einer radikaleren Position, die nicht an einer weiteren Zusammenarbeit mit den großen Playern des Oligopols festhalten würde. Aus dieser Position heraus ist der Konflikt zwischen den unterschiedlichen Werten und Zielen der beiden Parteien nicht auflösbar; jedenfalls nicht durch naives Vertrauen in die Regulierungskräfte von Märkten oder durch Transformationsverträge, die exorbitante Gewinnmargen und etablierte Machtpositionen der Verlage schützen. In anderen Worten sagt es SPARC (2021a, 7): “The business of publishers is to make money; the ‘business’ of academic institutions is to advance knowledge, not to enable publishers to achieve their commercial goals.” In diesem Sinne wären deutlichere Regulierungen und Vorgaben nötig, um die Machtasymmetrie und Abhängigkeiten im Publikationssektor aufzubrechen und die Empfehlungen des Wissenschaftsrats umzusetzen, Handlungs- und Gestaltungsfreiheit zurückzugewinnen. Dazu gehören dauerhafte Finanzierungen und Entwicklungsmöglichkeiten für wissenschaftliche Infrastrukturen in wissenschaftlicher Verwaltung und die Verpflichtung auf Pluralität und Offenheit (Open-Source-Lösungen). Transparenz, Nachprüfbarkeit und Interoperierbarkeit sollen gewährleisten, dass wissenschaftliche Infrastrukturen und Publikationsformen sich dem Ziel des EU-Rats annähern (Council of the & European Union, 2023). Der Wissenschaftsrat gibt einen Kurs vor, der nicht weiter auf Transformationsverträge setzt. Die Gestaltung des digitalen Raumes und maschinellen Lernens sollten nicht allein wirtschaftlichen Interessen überlassen werden und als Aufgabe der Wissenschaft aufgefasst und entsprechend getragen werden. Dieses Argument lässt sich nahtlos auf die Nutzung der Anwendungen auf allen Ebenen ausweiten, die hier besprochen wurden: Wissenschaftszyklus, Wissenschaftsbewertung und Wissenschaftsmanagement.

## Schluss

Die Integration von Elsevier-Anwendungen in Workflows an Schlüsselstellen wissenschaftlicher Institutionen birgt die konkrete Gefahr, neue Abhängigkeiten auf der gesamten Institutionsebene zu schaffen: im Forschungszyklus (individuell und institutionell), in der Wissenschaftsevaluation (institutionell) und im Management (institutionell). Integrierte Anwendungen könnten diese Bereiche überlappen, kontinuierlich und intransparent für die Nutzenden Daten sammeln und unerwünschte Plattformeffekte auf mehreren Ebenen der Wissenschaft als Institution bewirken. Diese Anwendungen werden als Gesamtpaket angeboten und verdrängen Alternativen, sie schaffen einen geschlossenen Kreis aus ihren eigenen Empfehlungen, Evaluationen und Produkten, sie haben bereits negative inhaltliche Effekte auf die Wissenschaft. Das Geschäftsmodell Elseviers und RELX' benötigt eine asymmetrische Wissensverteilung, u. a. durch Tracking, und ist mit wissenschaftlicher Autonomie unvereinbar, wie anhand der Unterscheidung zwischen einer schwachen These I und einer starken These II argumentiert wurde. Die starke These II wurde bestätigt: Dass eine unabhängige Wissenschaft in Zusammenarbeit mit Elsevier nicht möglich ist, weil Elsevier seine Angebote nicht ohne inakzeptable Tracking-Praktiken betreiben könnte. Diese Angebote sind, wie in dieser Arbeit argumentiert wurde, zu tiefgreifend in das Gesamtktionsfeld des Konzern integriert, um daraus herausgelöst werden zu können. Dasselbe gilt für Elseviers Konkurrenz im Oligopol wissenschaftlicher Verlags- und Datenanalyse-Konzerne. Elsevier zeigt beispielhaft, dass kommerzielle Akteur:innen höchstens in einem extern regulierten Rahmen mit der Wissenschaft zusammenarbeiten sollten. Es handelt sich bei den Großverlagen und Datenanalyse-Konzernen längst nicht mehr um Verlage im traditionellen Sinne. Sie haben ihre Geschäftspraktiken und Strategien auf eine grundsätzlich epistemische Asymmetrie, auf einen konstanten und umfassenden Wissensvorsprung vor der Konkurrenz und den unterschiedlichen Akteur:innen der Wissenschaft umgestellt. Es ist für sie deshalb ein notwendiger Bestandteil ihrer Geschäftspraxis, weiterhin invasiv Daten abzugreifen und immer weiter in die Strukturen der Wissenschaftsinstitutionen vorzudringen. Deshalb kann

die schwache These I, die eine Zusammenarbeit mit Elsevier unter Elseviers Verzicht auf inakzeptable Tracking-Praktiken als möglich behauptete, nicht bestätigt werden. Dies ist jedoch, wie hier argumentiert wurde, nicht im Sinne der Wissenschaft, wo diese Asymmetrien zum Verlust wesentlicher Eigenschaften freier Wissenschaft und zur Deformation (Zuboff 2019) der Wissenschaft führen. Es ist nicht im Sinne der Wissenschaftler:innen und Studierenden, deren Grundrechte verletzt werden. Es ist nicht im Sinne der Universitäten und Bibliotheken, deren Autonomie zugunsten wirtschaftlicher Ziele eingeschränkt wird und nicht im Sinne der Steuerzahler:innen, die bis zum Zehnfachen des errechneten Preises für Konzernprodukte zahlen. Es widerspricht dem politischen Ziel der digitalen Souveränität.

Insbesondere unter den Bedingungen der Digitalisierung, mit nicht absehbaren Folgen der exzessiven Tracking- und Datensammeltätigkeit großer Konzerne und Data Broker sowie des massiven Einsatzes von Überwachungstechnologien durch internationale Unternehmen (RELX und LexisNexis) sollte die Frage ernst genommen werden, „ob die Wissenschaft in der digitalen Welt immer stärker nach den Regeln der Wirtschaft organisiert und betrieben wird und damit ihre eigenen Prinzipien der Freiheit, der Sorgfalt und der Wahrheitsfindung kompromittiert werden.“ (Holzer, 2022, 175) In diesem Sinne ist es geraten, die weitere Zusammenarbeit mit den sogenannten legacy publishers als nicht vereinbar mit zentralen wissenschaftlichen Prinzipien zu betrachten und nicht fortzusetzen. Entscheidenden Einfluss haben hier weniger einzelne Wissenschaftler:innen, die in komplexen Abhängigkeiten von Förderern und Stellenvergabekriterien stehen, sondern vor allem die Akteur:innen, die über die Vergabe von Fördermitteln und Stellen, Beförderungen oder tenure entscheiden. Hier sind die Empfehlungen des Wissenschaftsrats zur Errichtung und Sicherung wissenschaftlich verwalteter (‘scholar-led’) wissenschaftlicher Infrastrukturen, die Empfehlungen des EU-Rats zum wissenschaftlichen Publizieren sowie die Initiativen zur Reform der Wissenschaftsbewertungspraxis wie DORA<sup>181</sup> und CoARA<sup>182</sup> wegweisend.

---

**181** <https://sfdora.org/>. [zuletzt geprüft am 25.07.2024]

**182** <https://coara.eu/>. [zuletzt geprüft am 25.07.2024]

# Literatur

(alle Internetquellen zuletzt geprüft am 25.07.2024)

n.a. *Die Privacy Sandbox: Technologie für mehr Datenschutz im Internet*. [https://privacysandbox.com/intl/de\\_de/](https://privacysandbox.com/intl/de_de/)

Allen, Anita L. 2011. *Unpopular privacy: what must we hide?* Studies in feminist philosophy. New York, N.Y: Oxford University Press.

Altschaffel, Robert, Dittmann, Jana, & Kiltz, Stefan. 2023. Science-Tracker Fingerprinting with Uncertainty: Selected Common Characteristics of Publishers from Network to Application Trackers on the Example of Web, App and Email. *Securware 2023: The seventeenth international conference on emerging security information, systems and technologies*, 88–97.

Altschaffel, Robert, Beurskens, Michael, Dittmann, Jana, Horstmann, Wolfram, Kiltz, Stefan, Lauer, Gerhard, Ludwig, Judith, Mittermaier, Bernhard, & Stump, Katrin. 2024. Datentracking und DEAL – Zu den Verhandlungen 2022/2023 und den Folgen für die wissenschaftlichen Bibliotheken –. *RuZ - Recht und Zugang*, 5 (1), 23–40.

Arthur, W. Brian. 1989. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *The economic journal*, 99 (394), 116.

Aspesi, Claudio. 2019. *Leaked Dutch Contract with Elsevier Raises Significant Alarm Bells*. <https://sparcopen.org/news/2019/leaked-dutch-contract-with-elsevier-raises-significant-alarm-bells/>

Aspesi, Claudio. 2022. *Scholarly publishing is broken. How do we fix it?* <https://av.tib.eu/media/59513>

- Ball, Rafael, & Tunger, Dirk. 2006. *Bibliometric analysis - A new business area for information professionals in libraries? Support for scientific research by perception and trend analysis*. Universität Regensburg. 10.5283/EPUB.5245
- Ball, Rafael, Mittermaier, Bernhard, & Tunger, Dirk. 2009. Creation of journal-based publication profiles of scientific institutions: A methodology for the interdisciplinary comparison of scientific research based on the J-factor. *Scientometrics*, 81 (2), 381–392.
- Batista, Pablo D., Campiteli, Mônica G., & Kinouchi, Osame. 2006. Is it possible to compare researchers with different scientific interests? *Scientometrics*, 68 (1), 179–189.
- Battenberg, Ortrud, Stock, Günter, Weingart, Peter, Frevert, Ute, Rheinberger, Hans-Jörg, Ash, Mitchell G., & der Wissenschaften, Berlin-Brandenburgische Akademie (Hg.). 2015. *Empfehlungen zur Zukunft des wissenschaftlichen Publikationssystems*. Berlin: BBAW.
- Bergstrom, Carl T., & West, Jevin. 2016. *Eigenfactor: Analysis of the CiteScore metric*. <http://eigenfactor.org/projects/posts/citescore.php>
- Biagioli, Mario. 2002. From Book Censorship to Academic Peer Review. *Emergences: Journal for the study of media & composite cultures*, 12 (1), 11–45.
- Biddle, Sam. 2021. *LexisNexis to Provide Giant Database of Personal Information to ICE*. <https://theintercept.com/2021/04/02/ice-database-surveillance-lexisnexis/>
- Bildung, Bundeszentrale für politische. 2019. *Microtargeting und Manipulation: Von Cambridge Analytica zur Europawahl*. <https://www.bpb.de/themen/medien-journalismus/digitale-desinformation/290522/microtargeting-und-manipulation-von-cambridge-analytica-zur-europawahl/>
- Binns, Reuben. 2022. *Tracking on the Web, Mobile and the Internet-of-Things*. <https://arxiv.org/abs/2201.10831>

- bitkom, Telekommunikation und neue Medien, Bundesverband Informationswirtschaft. 2015. *Digitale Souveränität. Positionsbestimmung und erste Handlungsempfehlungen für Deutschland und Europa*. <https://www.bitkom.org/Bitkom/Publikationen/Digitale-Souveraenitaet-Positionsbestimmung-und-erste-Handlungsempfehlungen-fuer-Deutschland-und-Europa.html>
- BMBF. 2021. *Technologisch souverän die Zukunft gestalten: Technologisch souverän die Zukunft gestalten. BMBF Impulspapier zur technologischen Souveränität*. [https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/5/24032\\_Impulspapier\\_zur\\_technologischen\\_Souveraenitaet.html](https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/5/24032_Impulspapier_zur_technologischen_Souveraenitaet.html)
- BMWi, Bundesministerium für Wirtschaft und Energie. 2016. *Digitale Strategie 2025*. [https://www.bmwi.de/Redaktion/DE/Publikationen/Digitale-Welt/digitale-strategie-2025.pdf?\\_\\_blob=publicationFile&v=18](https://www.bmwi.de/Redaktion/DE/Publikationen/Digitale-Welt/digitale-strategie-2025.pdf?__blob=publicationFile&v=18)
- BMWi, Bundesministerium für Wirtschaft und Energie. 2017. *Kompetenzen für eine Digitale Souveränität*. [https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/kompetenzen-fuer-eine-digitale-souveraenitaet.pdf?\\_\\_blob=publicationFile&v=14](https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/kompetenzen-fuer-eine-digitale-souveraenitaet.pdf?__blob=publicationFile&v=14)
- BMWi, Bundesministerium für Wirtschaft und Energie. 2021. *Schwerpunktstudie digitale Souveränität: Bestandsaufnahme und Handlungsfelder*. <https://www.bundesregierung.de/breg-de/service/publikationen/schwerpunktstudie-digitale-souveraenitaet-1981176>
- BOAI. 2000. *Budapest Open Access Initiative – Make research publicly available*. <https://www.budapestopenaccessinitiative.org/>
- Boni, Manfred. 2010. Analoges Geld für digitale Zeilen: der Publikationsmarkt der Wissenschaft. *Leviathan*, 38 (3), 293–312.
- Brayman, Kira, Devenney, Amy, Dobson, Helen, Marques, Mafalda, & Vernon, Anna. 2024. *A review of transitional agreements in the UK*. <https://zenodo.org/doi/10.5281/zenodo.10787392>

- Brembs, Björn. 2018. Prestigious Science Journals Struggle to Reach Even Average Reliability. *Frontiers in human neuroscience*, 12, 37.
- Brembs, Björn. 2023a. *Linné 4.0 – Taxonomie und Beobachtung von Forschenden*. *Laborjournal online* [https://www.laborjournal.de/rubric/essays/essays2023/e23\\_04.php](https://www.laborjournal.de/rubric/essays/essays2023/e23_04.php)
- Brembs, Björn. 2023b. *No Evilsevier DEAL!* <https://bjoern.brembs.net/2023/09/no-evilsevier-deal/>
- Brembs, Björn, & Siems, Renke. 2020. *Eine Gefährdung der Freiheit von Forschung und Lehre*. <https://www.jmwiarda.de/2020/10/26/eine-gefaehrung-der-freiheit-von-forschung-und-lehre/>
- Brembs, Björn, Button, Katherine, & Munafò, Marcus. 2013. Deep impact: unintended consequences of journal rank. *Frontiers in human neuroscience*, 7.
- Brembs, Björn, Förstner, Konrad, Kraker, Peter, Lauer, Gerhard, Müller-Birn, Claudia, Schönbrodt, Felix, & Siems, Renke. 2021a. Auf einmal Laborratte - Suddenly a laboratory Rat. *Journal of science, humanities and arts - josh*, 7 (6).
- Brembs, Björn, Degkwitz, Andreas, Holzer, Angela, Reda, Julia, & Wiarda, Jan-Martin. 2021b. *Wenn du nicht für das Produkt bezahlst, bist du selbst das Produkt?: Eine Podiumsdiskussion zur Kommerzialisierung von offener Wissenschaft*. <https://av.tib.eu/media/55690>
- Brembs, Björn, Förstner, Konrad, Müller-Birn, Claudia, & Dirnagl, Ulrich. 2023a. Kein guter Deal. *FAZ* vom 28. November 2023.
- Brembs, Björn, Huneman, Philippe, Schönbrodt, Felix, Nilsson, Gustav, Susi, Toma, Siems, Renke, Perakakis, Pandelis, Trachana, Varvara, Ma, Lai, & Rodriguez-Cuadrado, Sara. 2023b. Replacing academic journals. *Royal society open science*, 10 (7), 230206.
- Bundesregierung. 2021. *Datenstrategie der Bundesregierung*. <https://www.publikationen-bundesregierung.de/pp-de/publikationssuche/datenstrategie-der-bundesregierung-1845632>

- Buranyi, Stephen. 2017. Is the staggeringly profitable business of scientific publishing bad for science? *The Guardian*. <https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science>
- Butler, Leigh-Ann, Matthias, Lisa, Simard, Marc-André, Mongeon, Philippe, & Haustein, Stefanie. 2023. The Oligopoly's Shift to Open Access. How the Big Five Academic Publishers Profit from Article Processing Charges. *Quantitative science studies*, 1–33.
- Cahill, Mary Beth. 2020. *The Blending of Research Information Management and Faculty Information Management*. <https://www.thetambellinigroup.com/the-blending-of-research-information-management-and-faculty-information-management/>
- Chan, Leslie. 2019. *Platform Capitalism and the Governance of Knowledge Infrastructure*. <https://zenodo.org/record/2656601>
- Chen, George, & Chan, Leslie. 2021. University Rankings and Governance by Metrics and Algorithms. In: Hazelkorn, Ellen, & Mihut, Georgiana (Hg.), *Research Handbook on University Rankings: Theory, Methodology, Influence and Impact*. Edward Elgar Publishing, 424-441.
- Chen, George, Posada, Alejandro, & Chan, Leslie. 2019. Vertical Integration in Academic Publishing. In: Chan, Leslie, & Mounier, Pierre (Hg.), *Connecting the Knowledge Commons — From Projects to Sustainable Infrastructure*. OpenEdition Press. <https://books.openedition.org/oep/9068>
- Christl, Wolfie. 2021. Digitale Überwachung und Kontrolle am Arbeitsplatz. Von der Ausweitung betrieblicher Datenerfassung zum algorithmischen Management? Eine Studie von Cracked Labs. *Cracked labs*. <https://crackedlabs.org/daten-arbeitsplatz/info>

- Christl, Wolfie, & Toner, Alan. 2024a. Pervasive identity surveillance for marketing purposes. A technical report on personal data processing for LiveRamp's RampID identity graph system based on an analysis of software documentation. *Cracked labs*. <https://crackedlabs.org/en/identity-surveillance>
- Christl, Wolfie, & Toner, Alan. 2024b. Pervasive identity surveillance for marketing purposes. Summary. *Cracked labs*. <https://crackedlabs.org/en/identity-surveillance>
- Coalition, Scholarly Publishing and Academic Resources. 2022a. *Elsevier's Acquisition of Interfolio: Risks and Responses*. <https://infrastructure.sparcopen.org/interfolio-acquisition>
- Coalition, Scholarly Publishing and Academic Resources. 2022b. *Executive Summary*. <https://infrastructure.sparcopen.org/interfolio-acquisition/executive-summary>
- Council of the, & European Union. 2023. *Council conclusions on high-quality, transparent, open, trustworthy and equitable scholarly publishing*. <https://data.consilium.europa.eu/doc/document/ST-9616-2023-INIT/en/pdf>
- Crotty, David. 2023. *Quantifying Consolidation in the Scholarly Journals Market*. <https://scholarlykitchen.sspnet.org/2023/10/30/quantifying-consolidation-in-the-scholarly-journals-market/>
- Datenschutzkonferenz. 2017. *Göttinger Erklärung*. [https://www.datenschutzkonferenz-online.de/media/en/20170330\\_en\\_goettinger\\_erklaerung.pdf](https://www.datenschutzkonferenz-online.de/media/en/20170330_en_goettinger_erklaerung.pdf)
- David, Paul. 2000. Path Dependence, its critics, and the quest for 'historical economics'. [https://www.researchgate.net/profile/Paul-David/publication/23740891\\_Path\\_Dependence\\_its\\_critics\\_and\\_the\\_quest\\_for\\_'historical\\_economics'/links/00b7d52678dd6a8c48000000/Path-Dependence-its-critics-and-the-quest-for-historical-economics.pdf](https://www.researchgate.net/profile/Paul-David/publication/23740891_Path_Dependence_its_critics_and_the_quest_for_'historical_economics'/links/00b7d52678dd6a8c48000000/Path-Dependence-its-critics-and-the-quest-for-historical-economics.pdf)

David, Paul A. 1985. Clio and the Economics of QWERTY. *The American economic review*, 75 (2), 332–337.

de Rijcke, Sarah. 2020. *Elsevier and the Dutch Open Science goals*. <https://www.leidenmadtrics.nl/articles/s-de-rijcke-cwts-leidenuniv-nl>

Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen, Pampel, Heinz, Bertelmann, Roland, Hillenkötter, Kristine, Mittermaier, Bernhard, Pieper, Dirk, Schäffler, Hildegard, Seeh, Stefanie, & Tullney, Marco. 2022. *Empfehlungen für transformative*

*Zeitschriftenverträge mit Publikationsdienstleistern: Handreichung der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen vor dem Hintergrund der Umsetzung der Open-Access-Strategie 2021–2025 der Allianz der deutschen Wissenschaftsorganisationen.*

DFG AWBI, Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme. 2021. *Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Ein Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft*. <https://zenodo.org/record/5900759>

DFG Deutsche Forschungsgemeinschaft. 2020. *Digitaler Wandel in den Wissenschaften*. <https://zenodo.org/record/4191344>

DFG Deutsche Forschungsgemeinschaft. 2022. Guidelines for Safeguarding Good Research Practice. Code of Conduct. <https://zenodo.org/record/6472827>

DFG Deutsche Forschungsgemeinschaft | AG Publikationswesen. 2022. Wissenschaftliches Publizieren als Grundlage und Gestaltungsfeld der Wissenschaftsbewertung. <https://zenodo.org/record/6538163>

Dirnagl, Ulrich. 2023. *Wissenschaftler und Bibliothekare, hört die Signale: Keine DEALs mit unseren Papern!* [https://www.laborjournal.de/rubric/narr/narr/n\\_23\\_11.php](https://www.laborjournal.de/rubric/narr/narr/n_23_11.php)

Dittmann, Jana, Altschaffel, Robert, & Kiltz, Stefan. 2023. *Webtracking durch Wissenschaftsverlage - eine Spurensuche: Was wird genutzt, was darf man und was kann/muss/sollte man dagegen tun?*, Vortrag auf der 15. Göttinger Urheber- und Datenschutzrechtstagung am 07.11.2023.

Dobrov, Gennady. 1966. *Nauka o nauke*. Naukova dumka.

Drake, Tom, Gulliver, Sophie, & Harle, Jon. Research Publishing Is an Under-Recognised Global Challenge. <https://www.cgdev.org/sites/default/files/research-publishing-under-recognised-global-challenge-opportunities-g20-act.pdf>

Elsevier. *Advancing responsible research assessment: Elsevier's story*. <https://www.elsevier.com/connect/advancing-responsible-research-assessment-elseviers-story>

Elsevier. *Cookie Notice | Elsevier Legal*. <https://www.elsevier.com/de-de/legal/cookie notice>

Elsevier. *Cookies und ähnliche Technologien | Elsevier Legal*. <https://www.elsevier.com/legal/cookie notice-de-de>

Elsevier. *Privacy Policy | Elsevier Legal*. <https://www.elsevier.com/de-de/legal/cookie notice>

Elsevier. 2023. *Elsevier introduces authoritative scientific datasets for innovation and business- critical decisions*. <https://www.elsevier.com/about/press-releases/elsevier-introduces-authoritative-scientific-datasets-to-fuel-innovation-and>

Elsevier, B.V., & MPDL Services GGmbH, Max Planck Society. 2023. Projekt DEAL – Elsevier Publish and Read Agreement. [https://pure.mpg.de/pubman/item/item\\_3523659](https://pure.mpg.de/pubman/item/item_3523659)

- European Commission. Directorate General for Research and Innovation, University of Sheffield (United Kingdom), Bar Ilan University (IL), University of North Texas, Graz University of Technology, Leibniz Information Centre for Economics, Kiel University, & Leiden University. 2017. *Next-generation metrics: responsible metrics and evaluation for open science*. LU: Publications Office. <https://data.europa.eu/doi/10.2777/337729>
- Farley, Ashley, Langham-Putrow, Allison, Shook, Elisabeth, Sterman, Leila, & Wach, Megan. 2021. Transformative agreements: Six myths, busted: Lessons learned. *College & research libraries news*, 82 (7), 298.
- Finn, Rachel L., Wright, David, & Friedewald, Michael. 2013. Seven Types of Privacy, in Gutwirth, Serge, Leenes, Ronald, De Hert, Paul, & Poullet, Yves (Hg.), *European Data Protection: Coming of Age*. Dordrecht: Springer Netherlands, 3–32.
- Floridi, Luciano. 2017. Group Privacy: A Defence and an Interpretation, in Taylor, Linnet, Floridi, Luciano, & Van Der Sloot, Bart (Hg.), *Group Privacy*. Cham: Springer International Publishing, 83–100.
- Fokusgruppe, Digitale Souveränität. 2019. *Digitale Souveränität im Kontext plattformbasierter Ökosysteme: Plattform ,Innovative Digitalisierung der Wirtschaft‘*. [https://www.eco.de/wp-content/uploads/2019/11/whitepaper\\_digitale-souveraenitaet-im-kontext-plattformbasierter-oekosysteme\\_digital-gipfel.pdf](https://www.eco.de/wp-content/uploads/2019/11/whitepaper_digitale-souveraenitaet-im-kontext-plattformbasierter-oekosysteme_digital-gipfel.pdf)
- Förstner, Konrad. *Personalisiertes Tracking von WissenschaftlerInnen durch Verlage*. Open Science Radio. <https://www.openscienceradio.org/2020/12/29/osr197-personalisiertes-tracking-von-wissenschaftlerinnen-durch-verlage/>
- Förstner, Konrad, & Brembs, Björn. 2022. *Von Zugang zu Überwachung: neue Geschäftsmodelle der alten Verlage*. [www.youtube.com/watch?v=Y07Fx4D2wQ](https://www.youtube.com/watch?v=Y07Fx4D2wQ)

- Fouad, Imane, Santos, Cristiana, & Laperdrix, Pierre. 2024. The Devil is in the Details: Detection, Measurement and Lawfulness of Server-Side Tracking on the Web. *24th privacy enhancing technologies symposium (pets 2024)*. <https://hal.science/hal-04617727>
- Freiberg, Michael. 2022. Third-Party-Tracking bei Wiley und Springer: Analyse und Ausblick. *Abi technik*, 42 (2), 96–104.
- Fühles-Ubach, Simone, & Tunger, Dirk. 2021. Daten und Metriken als Baustein von Benchmarking und Strategiebildung in der Wissenschaft. *Wissenschaftsmanagement*, 1–8.
- Funk, McKenzie. 2023. The Man Who Trapped Us in Databases. *The New York times*. <https://www.nytimes.com/2023/09/22/magazine/hank-asher-data.html>
- Gadd, Lizzie. 2018 (Mar.). *Is it time to bury the h-index?* <https://thebibliomagician.wordpress.com/2018/03/23/is-it-time-to-bury-the-h-index/>
- Gallie, W. B. 1956. Essentially Contested Concepts. *Proceedings of the Aristotelian society*, 56 (1), 167–198.
- Gatti, Rupert. 2020. *Business Models and Market Structure within the Scholarly Communications Sector*. International Science Council. [https://council.science/wp-content/uploads/2020/06/ISC-Occasional-Paper-Business-Models-and-Market-Structure-within-the-Scholarly-Communications-Sector\\_Rupert-Gatti-3.pdf](https://council.science/wp-content/uploads/2020/06/ISC-Occasional-Paper-Business-Models-and-Market-Structure-within-the-Scholarly-Communications-Sector_Rupert-Gatti-3.pdf)
- Gehring, Petra. 2022. Datensouveränität versus Digitale Souveränität: Wegweiser aus dem konzeptionellen Durcheinander, in Augsberg, Steffen, & Gehring, Petra (Hg.), *Datensouveränität*. Campus Verlag, 19–44. <https://doi.org/10.12907/978-3-593-45194-7>.
- Gesellschaft, Max Planck. 2003. *Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen*. [https://openaccess.mpg.de/68053/Berliner\\_Erklaerung\\_dt\\_Version\\_07-2006.pdf](https://openaccess.mpg.de/68053/Berliner_Erklaerung_dt_Version_07-2006.pdf)

- GI, Gesellschaft für Informatik. 2020. *Schlüsselaspekte digitaler Souveränität, Arbeitspapier*. <https://gi.de/meldung/gi-veroeffentlicht-arbeitspapier-zu-schluesselaspekten-digitaler-souveraenitaet>
- Godin, Benoît. 2006. On the origins of bibliometrics. *Scientometrics*, 68 (1), 109–133.
- Goldacker, Gabriele. 2017. *Digitale Souveränität*. Berlin: Kompetenzzentrum Öffentliche IT, Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS.
- Gräf, Eike, Lahmann, Henning, & Otto, Philipp. 2018. *Die Stärkung der digitalen Souveränität: Wege der Annäherung an ein Ideal im Wandel*. <https://www.divsi.de/wp-content/uploads/2018/05/DIVSI-Themenpapier-Digitale-Souveraenitaet.pdf>
- Grossmann, Alexander, & Brembs, Björn. 2021. Current market rates for scholarly publishing services. *F1000research*, 10.
- Gulliver, Sophie, & Drake, Tom. The \$1 Trillion Paradox: Why Reforming Research Publishing Should be a Global Priority. <https://www.cgdev.org/sites/default/files/why-reforming-research-publishing-should-be-global-priority.pdf>
- Hanekop, Heidemarie, & Wittke, Volker. 2014. Wandel des wissenschaftlichen Publikationssystems durch das Internet: sektorale Transformation im Kontext institutioneller Rekonfiguration, in Dolata, Ulrich, & Schrape, Felix (Hg.), *Internet, Mobile Devices und die Transformation der Medien. Radikaler Wandel als schrittweise Rekonfiguration*. Berlin: edition sigma, 147–176.
- Hanson, Cody. 2019. *User Tracking on Academic Publisher Platforms*. <https://www.codyh.com/writing/tracking.html>
- Hanson, Mark A., Barreiro, Pablo Gómez, Crosetto, Paolo, & Brockington, Dan. 2023. The strain on scientific publishing. <https://arxiv.org/abs/2309.15884>

- Haucap, Justus, Moshgbar, Nima, & Schmal, W. Benedikt. 2021. The impact of the German ‚DEAL‘ on competition in the academic publishing market. *Managerial and decision economics*, 42 (8), 2027–2049.
- Hellman, Eric. 2015. *16 of the top 20 Research Journals Let Ad Networks Spy on Their Readers*. <https://go-to-hellman.blogspot.com/2015/03/16-of-top-20-research-journals-let-ad.html>
- Herb, Ulrich. 2018. Zwangsehen und Bastarde. *Universität des Saarlandes*. <https://publikationen.sulb.uni-saarland.de/handle/20.500.11880/27437>
- Herb, Ulrich. 2023. *The German DEAL contracts*. *Sci:Debug*. <https://doi.org/10.59350/vfqbq-13430>
- Hirsch, J. E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the national academy of sciences*, 102 (46), 16569–16572.
- Hirschi, Caspar. 2018. *Wie die Peer Review die Wissenschaft diszipliniert*. <https://www.merkur-zeitschrift.de/2018/08/27/wie-die-peer-review-die-wissenschaft-diszipliniert/>
- Hollis, Martin, & Lukes, Steven. 1982. Introduction, in Hollis, Martin, & Lukes, Steven (Hg.), *Rationality and Relativism*. Oxford: M.I.T. Press, 1–20.
- Holzer, Angela. 2022. Die Vermessbarkeit der Wissenschaft Digitalisierung, wissenschaftliches Publizieren, Verhaltenstracking und Wissenschaftsbewertung, in Mößner, Nicola, & Erlach, Klaus (Hg.), *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*. Science Studies. Bielefeld, Germany: transcript Verlag, 163–182.
- Hummel, Patrik, Braun, Matthias, Tretter, Max, & Dabrock, Peter. 2021. Data sovereignty: A review. *Big data & society*, 8 (1), 205395172098201.

hum.works. *Turn your disparater data into solid gold - A (nearly) complete guide to data & customer data platforms for scholarly publishers.* <https://www.hum.works/whitepaper-disparate-data>

IT-Planungsrat, & IT-Rat. 2020. *Stärkung der Digitalen Souveränität der öffentlichen Verwaltung, Eckpunkte – Ziel und Handlungsfelder.* [https://www.cio.bund.de/SharedDocs/downloads/Webs/CIO/DE/digitale-loesungen/eckpunktpapier-digitale-souveraenitaet.pdf?\\_\\_blob=publicationFile&v=2](https://www.cio.bund.de/SharedDocs/downloads/Webs/CIO/DE/digitale-loesungen/eckpunktpapier-digitale-souveraenitaet.pdf?__blob=publicationFile&v=2)

Jansen, Dorothea, Wald, Andreas, Franke, Karola, Schmoch, Ulrich, & Schubert, Torben. 2007. Drittmittel als Performanzindikator der Wissenschaftlichen Forschung: Zum Einfluss von Rahmenbedingungen auf Forschungsleistung. *Kzfss Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 59 (1), 125–149.

Janßen, Melike, & Sondermann, Ariadne. 2016. Universitäre Leistungsbewertungen als Bedrohung der akademischen Identität? Ein subjektorientierter Blick auf die Beurteilung und Kontrolle wissenschaftlicher Leistungen im Zeichen von New Public Management. *Berliner Journal für Soziologie*, 26 (3-4), 377–402.

Johnson, Rob, Fosci, Mattia, Chiarelli, Andrea, Pinfield, Stephen, & Jubb, Michael. 2017. *Towards a competitive and sustainable OA market in Europe - a study of the open access market and policy environment.* <https://zenodo.org/record/401029>

Kaden, Ben. 2023. *OA Takeaways: Die Kritik am DEAL-Vertrag mit Elsevier in der FAZ vom 29.11.2023.* <https://open-access-brandenburg.de/oa-takeaways-die-kritik-am-deal-vertrag-mit-elsevier-in-der-faz-vom-29-11-2023/>

Kagermann, Henning, Streibich, Karl-Heinz, & Suder, Katrin. 2021. *Digitale Souveränität Status quo und Handlungsfelder.* acatech Impuls. <https://www.acatech.de/publikation/digitale-souveraenitaet-status-quo-und-handlungsfelder/>

- Katz, Michael L., & Shapiro, Carl. 1985. Network Externalities, Competition, and Compatibility. *The American economic review*, 75 (3), 424–440.
- Keegan, Jon, & Eastwood, Joel. 2023. *From “Heavy Purchasers” of Pregnancy Tests to the Depression-Prone: We Found 650,000 Ways Advertisers Label You – The Markup*. Section: Privacy. <https://themarkup.org/privacy/2023/06/08/from-heavy-purchasers-of-pregnancy-tests-to-the-depression-prone-we-found-650000-ways-advertisers-label-you>
- Knecht, Sicco de. 2019. *Leaked document on Elsevier negotiations sparks controversy - Science- Guide*. <https://www.scienceguide.nl/2019/11/leaked-document-on-elsevier-negotiations-sparks-controversy/>
- Knecht, Sicco de. 2020. *Dutch open science deal primarily benefits Elsevier - ScienceGuide*. <https://www.scienceguide.nl/2020/06/open-science-deal-benefits-elsevier/>
- Knoche, Michael. 2021a. Raus aus den DEAL-Verträgen! Sieben Gründe für den Ausstieg. *Aus der Forschungsbibliothek Krekelborn*, 2. <https://biblio.hypotheses.org/2598>
- Knoche, Michael. 2021b. *Verlage werden Datenkraken – Datentracking in der Wissenschaft*. <https://biblio.hypotheses.org/2493>
- Kosinski, Michal, Stillwell, David, & Graepel, Thore. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences*, 110 (15), 5802–5805.
- Kramer, Bianca, & Bosman, Jeroen. *101 Workflow innovations*. <https://101innovations.wordpress.com/workflows/>
- Kunz, Raffaella. 2022. *The Digital Threat to Science and Academic Freedom*. <https://trafo.hypotheses.org/36984>

- Lamdan, Sarah. 2019. *Librarianship at the Crossroads of ICE Surveillance – In the Library with the Lead Pipe*. <https://www.inthelibrarywiththeleadpipe.org/2019/ice-surveillance/>
- Lamdan, Sarah. 2022. *Data Cartels The Companies That Control and Monopolize Our Information*. Stanford University Press.
- Lane, Julia, Stodden, Victoria, Bender, Stefan, & Nissenbaum, Helen (Hg.). 2014. *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. Cambridge University Press.
- Larivière, Vincent, & Sugimoto, Cassidy R. 2019. The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects, in Glänzel, Wolfgang, Moed, Henk F., Schmoch, Ulrich, & Thelwall, Mike (Hg.), *Springer Handbook of Science and Technology Indicators*. Cham: Springer International Publishing. Series Title: Springer Handbooks, 3–24.
- Larivière, Vincent, Haustein, Stefanie, & Mongeon, Philippe. 2015. The Oligopoly of Academic Publishers in the Digital Era. *Plos one*, 10 (6), e0127502.
- Larsson, Per. 2023. *Open access: Need to move away from transformative agreements - Stockholm University*. <https://www.su.se/english/news/open-access-need-to-move-away-from-transformative-agreements-1.683787>
- Lauer, Gerhard. 2022a. Datentracking in den Wissenschaften. *o-bib. das offene Bibliotheksjournal / Herausgeber vdb*, 1–13.
- Lauer, Gerhard. 2022b. Die Macht der Datenkartelle. *Faz.net*, <https://zeitung.faz.net/faz/geisteswissenschaften/2022-12-07/datenkartelle-im-goldrausch/835425.html>
- Lewandowski, Dirk. 2006. *iQ - Journal Impact Factor*. [https://web.archive.org/web/20070501230234/http://www.forschungsinfo.de/iq/agma/Journal\\_Impact\\_Factor/journal\\_impact\\_factor.asp](https://web.archive.org/web/20070501230234/http://www.forschungsinfo.de/iq/agma/Journal_Impact_Factor/journal_impact_factor.asp)

- LexisNexis, Risk Solutions. (a) *Empowering Smarter Authentication and Fraud Decisioning in an Evolving Digital Landscape*. <https://risk.lexisnexis.com/-/media/files/financial%20services/articles/lexisnexis-lexid-digital-brochure-nxr14092-00-0919-en-us%20pdf.pdf>
- LexisNexis, Risk Solutions. (b) *Transform Human Interactions into Actionable Intelligence*. [https://risk.lexisnexis.com/global/-/media/files/product%20pages/brochure/lnrs\\_behaviosec\\_prodbrochure3\\_en-nxr15807-00-0123-en-us.pdf](https://risk.lexisnexis.com/global/-/media/files/product%20pages/brochure/lnrs_behaviosec_prodbrochure3_en-nxr15807-00-0123-en-us.pdf)
- LexisNexis, Risk Solutions. 2019. *Fallstudie: Großes Finanzinstitut berichtet eine 50%ige Verbesserung bei der Erkennung von Money-Mule-Konten innerhalb des ersten Anwendungsjahres von LexisNexis® Risk Solutions*.
- LexisNexis, Risk Solutions. 2023. *Experience the Power of the One - Secure and Seamless Identity Solution*.
- Liebowitz, Stan J., & Margolis, Stephen E. 1995. Path Dependence, Lock-In, and History. *SSRN electronic journal*. 10.2139/ssrn.1706450
- Lisker, Mareike. 2023. *Von der (Un-)Möglichkeit, digital mündig zu sein Tracking-Infrastrukturen und die Responsibilisierung des Individuums im Internet*. Masterarbeit.
- Lynch, Clifford. 2017. The rise of reading analytics and the emerging calculus of reader privacy in the digital world. *First Monday*. <https://firstmonday.org/ojs/index.php/fm/article/view/7414>
- Lynch, Clifford A. 2019. Reader Privacy: The New Shape of the Threat. *Research library issues*, May, 7–14.
- Ma, Lai. 2023a. Information, platformized. *Journal of the association for information science and technology*, 74 (2), 273–282.
- Ma, Lai. 2023b. The Platformisation of Scholarly Information and How to Fight It. *Liber quarterly: The journal of the association of european research libraries*, 33 (1), 1–20.

- May, Monika. 2014. Bibliometrie – ein Aufgabengebiet von Bibliotheken? *Bibliotheksdienst*, 48 (2), 132–147.
- Merton, Robert K. 1973. The Normative Structure of Science. *Pages 267–278 of: The Sociology of Science. Theoretical and Empirical Investigations*. University of Chicago Press.
- Mittelstadt, Brent. 2017. From Individual to Group Privacy in Big Data Analytics. *Philosophy & technology*, 30 (4), 475–494.
- Mongeon, Philippe, & Paul-Hus, Adèle. 2016. The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106 (1), 213–228.
- Moody, Glyn. 2018. *Hated Science Publisher Elsevier To Help EU Monitor Open Science – Including Open Access*. <https://www.techdirt.com/2018/04/03/hated-science-publisher-elsevier-to-help-eu-monitor-open-science-including-open-access/>
- Morais, Rita, Stoy, Lennart, & Borrell-Damián, Lidia. 2019. *2019 Big Deals Survey Report. An Updated Mapping of Major Scholarly Publishing Contracts in Europe*.
- Mühlhoff, Rainer. 2022. Prädiktive Privatheit: Kollektiver Datenschutz im Kontext von Big Data und KI, in Friedewald, Michael, Roßnagel, Alexander, Heesen, Jessica, Krämer, Nicole, & Lamla, Jörn (Hg.), *Künstliche Intelligenz, Demokratie und Privatheit*. Nomos Verlagsgesellschaft mbH & Co. KG, 31–58.
- Mulligan, Deirdre K., Koopman, Colin, & Doty, Nick. 2016. Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy. *Philosophical transactions of the royal society a: Mathematical, physical and engineering sciences*, 374 (2083), 20160118.
- Münch, Richard. 2014a. *Academic Capitalism*. Routledge.
- Münch, Richard. 2014b. Introduction - Struggling for Excellence in the Academic Field, in *Academic Capitalism*. Routledge, 1–12.

- Münch, Richard. 2015. Alle Macht den Zahlen! Zur Soziologie des Zitationsindexes. *Soziale welt*, 66 (2), 149–160.
- Nissenbaum, Helen Fay. 2010. *Privacy in context: technology, policy, and the integrity of social life*. Stanford, CA: Stanford Law Books.
- noyb, Europäisches Zentrum für digitale Rechte. 2024. *Google Sandbox: Online tracking instead of privacy*. <https://noyb.eu/en/google-sandbox-online-tracking-instead-privacy>
- Özmen, Elif. 2021. *Wissenschaftsfreiheit: Normative Grundlagen und aktuelle Herausforderungen*. <https://www.bpb.de/shop/zeitschriften/apuz/wissenschaftsfreiheit-2021/343220/wissenschaftsfreiheit-normative-grundlagen-und-aktuelle-herausforderungen/>
- Piwik. *Cookie syncing*. <https://piwik.pro/glossary/cookie-syncing/>
- Pohle, Julia. 2020. *Digitale Souveränität: Ein neues digitalpolitisches Schlüsselkonzept in Deutschland und Europa*. Konrad Adenauer Stiftung.
- Pohle, Julia, & Thiel, Thorsten. 2019. Digitale Vernetzung und Souveränität: Genealogie eines Spannungsverhältnisses, in Borucki, Isabelle, & Schünemann, Wolf Jürgen (Hg.), *Internet und Staat: Perspektiven auf eine komplizierte Beziehung*. Nomos Verlagsgesellschaft mbH & Co, 57–80.
- Pohle, Julia, & Thiel, Thorsten. 2021. 2.3.2 Digitale Souveränität: Von der Karriere eines einenden und doch problematischen Konzepts, in Piallat, Chris (Hg.), *Digitale Gesellschaft*. Bielefeld, Germany: transcript Verlag, 319–340.
- Pooley, Jeff. 2022. Surveillance Publishing. *The journal of electronic publishing*, 25 (1).
- Pooley, Jeff. 2024. Large Language Publishing. *Upstream*. 10.54900/zg929-e9595.

- Posada, Alejandro, & Chen, George. 2018. Inequality in Knowledge Production: The Integration of Academic Infrastructure by Big Publishers. In: *22nd International Conference on Electronic Publishing*. OpenEdition Press. 10.4000/proceedings.elpub.2018.30
- Post, Robert C. 2001. Three Concepts of Privacy. *Georgetown law journal*, 89, 2087–2098.
- Price, Derek J. de Solla. 1963. *Little science, big science*. New York: Columbia University Press.
- Pritchard, Alan. 1969. Statistical Bibliography or Bibliometrics? *Journal of documentation*, 25, 348–349.
- PRNewswire. 2019. *Elsevier Announces the International Center for the Study of Research*. <https://www.prnewswire.com/news-releases/elsevier-announces-the-international-center-for-the-study-of-research-300871135.html>
- PRNewswire. 2020. *Elsevier's International Center for the Study of Research Launches ICSR Lab*. <https://www.prnewswire.com/news-releases/elseviers-international-center-for-the-study-of-research-launches-icsr-lab-301024713.html>
- PwC. 2022. *Internet Advertising Revenue Report: Full-year 2021 results*.
- Reitz, Tilman. 2023. Open-Access, DEAL-Verhandlungen und Wissenschaftsfreiheit, in Schimank, Uwe (Hg.), *Open Access - Deal - Wissenschaftlertracking: Das wissenschaftliche Publikationssystem im Wandel*. Eine Schriftenreihe der Berlin-Brandenburgischen Akademie der Wissenschaften. Berlin-Brandenburgische Akademie der Wissenschaften, 9–42.
- RELX. *Acquisition of Seisint*. <https://www.relx.com/media/press-releases/archive/14-07-2004>
- RELX. 2023. *2023 Annual Report including Financial Statements and Corporate Responsibility Report*. <https://www.relx.com/investors/annual-reports/2023>

- RELX. 2024. *RELX The global provider of information-based analytics and decision tools*. <https://www.relx.com/~media/Files/R/RELX-Group/documents/investors/relx-overview-june-2024.pdf>
- Retzlaff, Eric. 2022. Wer bewertet mit welchen Interessen wissenschaftliche Publikationen? Eine Skizzierung des Einflusses kommerzieller Interessen auf die Forschungsbewertung, in Mößner, Nicola, & Erlach, Klaus (Hg.), *Kalibrierung der Wissenschaft: Auswirkungen der Digitalisierung auf die wissenschaftliche Erkenntnis*. Science Studies. Bielefeld, Germany: transcript Verlag, 139–162.
- Reuter, Markus. 2024 (Mar.). *LiveRamp: Datenfirma unterhält „privates Bevölkerungsregister“*. <https://netzpolitik.org/2024/liveramp-datenfirma-unterhaelt-privates-bevoelkerungsregister/>
- RfII, Rat für Informationsinfrastrukturen. 2023. *Föderierte Dateninfrastrukturen für die wissenschaftliche Nutzung*.
- Ronzheimer, Manfred. 2021. Warnung vor den Wissensspionen Forscher wehren sich gegen Datentracking von Verlagen. Sie sehen die Wissenschaftsfreiheit in Gefahr und fordern klare Regeln. *taz* <https://taz.de/Datentracking-bei-Wissenschaftsverlagen/!5798570/>
- Rössler, Beate. 2005. *The value of privacy*. Cambridge, UK ; Malden, MA: Polity.
- Roßnagel, Alexander. 2022. Third-Party-Tracking – ein Problem aus Sicht des Datenschutzes?: Interview mit Prof. Dr. Alexander Roßnagel, Hessischer Beauftragter für Datenschutz und Informationsfreiheit. *Abi technik*, 42 (2), 105–107.
- Ryan, Johnny, & Christl, Wolfie. *Europe's hidden security crisis: How data about European defence personnel and political leaders flows to foreign states and non-state actors*. <https://www.iccl.ie/digital-data/europes-hidden-security-crisis/>

- Schimmer, Ralf, Geschuhn, Kai Karin, & Vogler, Andreas. 2015. Disrupting the subscription journals' business model for the necessary large-scale transformation to open access. Publisher: Max Planck Digital Library. <http://pure.mpg.de/pubman/item/escidoc:2148961>
- Schneider, Markus, Enzmann, Matthias, & Stopczynski, Martin. 2014. *SIT Technical reports. 2014,01: Web-Tracking-Report*. Stuttgart: Fraunhofer-Verlag.
- Scholl, Dominik. 2021. *Datentracking: Die schöne neue Welt der Wissenschaftsverlage*.
- Schreyögg, Georg. 2014. Pfadabhängigkeit und Pfadbruch in Unternehmen. *Schmalenbachs Zeitschrift für betriebswirtschaftliche forschung*, 66 (S68), 1–17.
- Schreyögg, Georg, & Sydow, Jörg. 2011. Organizational Path Dependence: A Process View. *Organization studies*, 32 (3), 321–335.
- Schwerpunktinitiative „Digitale Information“ der Allianz der Deutschen Wissenschaftsorganisationen. 2019. Empfehlungen zu Methoden zur Kontrolle des Zugriffs auf wissenschaftliche Informationsressourcen. <https://doi.org/10.18452/36189>
- Selinger, Joschka. 2024. *Wer zählt das Cookiemonster. Die Folgen von Webtracking für Datenschutz und Wissenschaftsfreiheit*. [https://duepublico2.uni-due.de/receive/duepublico\\_mods\\_00081665](https://duepublico2.uni-due.de/receive/duepublico_mods_00081665)
- Shapiro, Carl, & Varian, Hal R. 2010. *Information rules: a strategic guide to the network economy*. Boston, Mass: Harvard Business School Press.
- Sidiropoulos, Antonis, Katsaros, Dimitrios, & Manolopoulos, Yannis. 2007. Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics*, 72 (2), 253–280.
- Siems, Renke. 2021. When your journal reads you – user tracking on science publisher platforms. Elephant in the Lab. <https://doi.org/10.5281/zenodo.4683778>

10.5281/ZENODO.4683778

Siems, Renke. 2022. Lesen der Anderen. *o-bib. das offene Bibliotheksjournal/Herausgeber vdb*, 1–25.

Siems, Renke. 2023a. *Comments on DEAL-Elsevier*. <https://mastoreader.io/?url=https%3A%2F%2Fscicomm.xyz%2F%4Onicebread%2F111084556968691388>

Siems, Renke. 2023b. »Surveil and Punish« – Tracking and Control of the Research Life Cycle. Publisher: Zenodo. <https://zenodo.org/record/7935017>

Siems, Renke. 2023c. „Überwachen und Strafen“ – Tracking und Kontrolle des Forschungszyklus. *Abi technik*, 43 (2), 86–95.

Siems, Renke. 2024a. Rezension zu: The Hank Show: how a house-painting, drug running DEA informant built the machine that rules our lives / McKenzie Funk. *o-bib. das offene Bibliotheksjournal/Herausgeber vdb*, 1–8.

Siems, Renke. 2024b. Subprime Impact Crisis. Bibliotheken, Politik und digitale Souveränität. *Bibliothek Forschung und Praxis*. <https://doi.org/10.1515/bfp-2024-0008>

Siems, Renke. 2024c. *Subprime Impact Crisis: Bibliotheken, Tracking und digitale Souveränität der Wissenschaften*. [https://duepublico2.uni-due.de/receive/duepublico\\_mods\\_00081664](https://duepublico2.uni-due.de/receive/duepublico_mods_00081664)

Solove, Daniel J. 2006. A Taxonomy of Privacy. *University of Pennsylvania law review*, 154 (3), 477.

Solove, Daniel J. 2009. *Understanding privacy*. Cambridge, Massachusetts London, England: Harvard University Press.

- SPARC, Scholarly Publishing and Academic Resources. 2019. *Landscape Analysis: The Changing Academic Publishing Industry – Implications for Academic Institutions*. [https://www.academia.edu/59578908/SPARC\\_Landscape\\_Analysis\\_The\\_Changing\\_Academic\\_Publishing\\_Industry\\_Implications\\_for\\_Academic\\_Institutions](https://www.academia.edu/59578908/SPARC_Landscape_Analysis_The_Changing_Academic_Publishing_Industry_Implications_for_Academic_Institutions)
- SPARC, Scholarly Publishing and Academic Resources. 2020. *2020 Update - SPARC Landscape Analysis and Roadmap for Action*. <https://infrastructure.sparcopen.org/2020-update>
- SPARC, Scholarly Publishing and Academic Resources. 2021a. *2021 Update: SPARC Landscape Analysis and Roadmap for Action*. <https://sparcopen.org/wp-content/uploads/2021/10/2021-Landscape-Analysis-101421.pdf>
- SPARC, Scholarly Publishing and Academic Resources. 2021b. *Addressing the Alarming Systems of Surveillance Built By Library Vendors*. <https://sparcopen.org/news/2021/addressing-the-alarming-systems-of-surveillance-built-by-library-vendors/>
- SPARC, Scholarly Publishing and Academic Resources. 2022a. *Elsevier's Dominance in the FIS Market Could Threaten Competition in Related Markets*. <https://infrastructure.sparcopen.org/interfolio-acquisition/related-markets>
- SPARC, Scholarly Publishing and Academic Resources. 2022b. *Possible Negative Impacts From Elsevier's Increased Market Power*. <https://infrastructure.sparcopen.org/interfolio-acquisition/possible-negative-impacts>
- SPARC, Scholarly Publishing and Academic Resources Coalition. 2022c. *The Interfolio Acquisition Highly Concentrates the Nascent Market for Faculty Information Systems*. <https://infrastructure.sparcopen.org/interfolio-acquisition/fis-market-concentration>

- SPARC, Scholarly Publishing and Academic Resources. 2023. *Navigating Risk in Vendor Data Privacy Practices - An Analysis of Elsevier ScienceDirect*. <https://sparcopen.org/our-work/privacy-and-surveillance-community-of-practice/navigating-risk-in-vendor-data-privacy-practices/>
- Statista. *Statistiken zum Thema Programmatic Advertising*. <https://de.statista.com/themen/4554/programmatic-advertising/>
- Statista. 2024. *Themenseite: Online-Werbung*. <https://de.statista.com/themen/113/online-werbung/>
- Stich, Stephen. 1998. Epistemic Relativism, in *Routledge encyclopedia of philosophy*. Routledge. Straumsheim, Carl. 2016. *How to Measure Impact*. <https://www.insidehighered.com/news/2016/12/14/exploring-citescore-elseviers-new-journal-impact-metrics>
- SURF, & Elsevier. *UKB-Elsevier\_2020-2024*. <https://esac-initiative.org/about/transformative-agreements/agreement-registry/els2020vsnu>
- Szprot, Jakub, Gruenpeter, Natalia, Rycko, Nikodem, Siewicz, Krzysztof, Krzysztof, Schimmer, Ralf, Campbell, Colleen, Karlstrøm, Nina, Andenæs, Nils, & Lundén, Anna. 2021. *Transformative Agreements: Overview, Case Studies, and Legal Analysis*. <https://zenodo.org/record/4543422>
- Tennant, Jon. 2018. Elsevier are corrupting open science in Europe. *The Guardian*. <https://www.theguardian.com/science/political-science/2018/jun/29/elsevier-are-corrupting-open-science-in-europe>
- Thiel, Thorsten. 2019. Souveränität: Dynamisierung und Kontestation in der digitalen Konstellation, in Hofmann, Jaenette, Kersting, Norbert, Ritzi, Claudia, & Schünemann, Wolf J. (Hg.), *Politik in der digitalen Gesellschaft: Zentrale Problemfelder und Forschungsperspektiven*. transcript, 47–60.

- Thiel, Thorsten. 2021. Das Problem mit der digitalen Souveränität. *Faz*. 26. Januar 2021. <https://www.faz.net/aktuell/wirtschaft/digitec/europa-will-in-derinformationstechnologie-unabhaengiger-werden-17162968.html>
- Thomson, Judith Jarvis. 1975. The Right to Privacy. *Philosophy & public affairs*, 4 (4), 295–314.
- Torny, Didier. 2020. *The absurd race for university rankings...or how publications are transformed into bad data*. <https://polecopub.hypotheses.org/2292>
- Torny, Didier. 2023. *Who wins after a divorce?...or how to interpret the DEAL-Elsevier new agreement*. <https://polecopub.hypotheses.org/3261>
- van den Hoven, Jeroen, Blaauw, Martijn, Pieters, Wolter, & Warnier, Martijn. 2020. Privacy and Information Technology, in Zalta, Edward N. (Hg.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/entries/privacy/>
- w3.org. *Same Origin Policy - Web Security*. [https://www.w3.org/Security/wiki/Same-Origin\\_Policy](https://www.w3.org/Security/wiki/Same-Origin_Policy)
- Warren, Samuel D., & Brandeis, Louis D. 1890. The Right to Privacy. *Harvard law review*, 4 (5), 193–220.
- Wissenschaftsrat. 2022. *Empfehlungen zur Digitalisierung in Lehre und Studium*. <https://www.wissenschaftsrat.de/download/2022/9848-22.html>
- Wissenschaftsrat. 2023. *Empfehlungen zur Souveränität und Sicherheit der Wissenschaft im digitalen Raum*. <https://www.wissenschaftsrat.de/download/2023/1580-23.html>

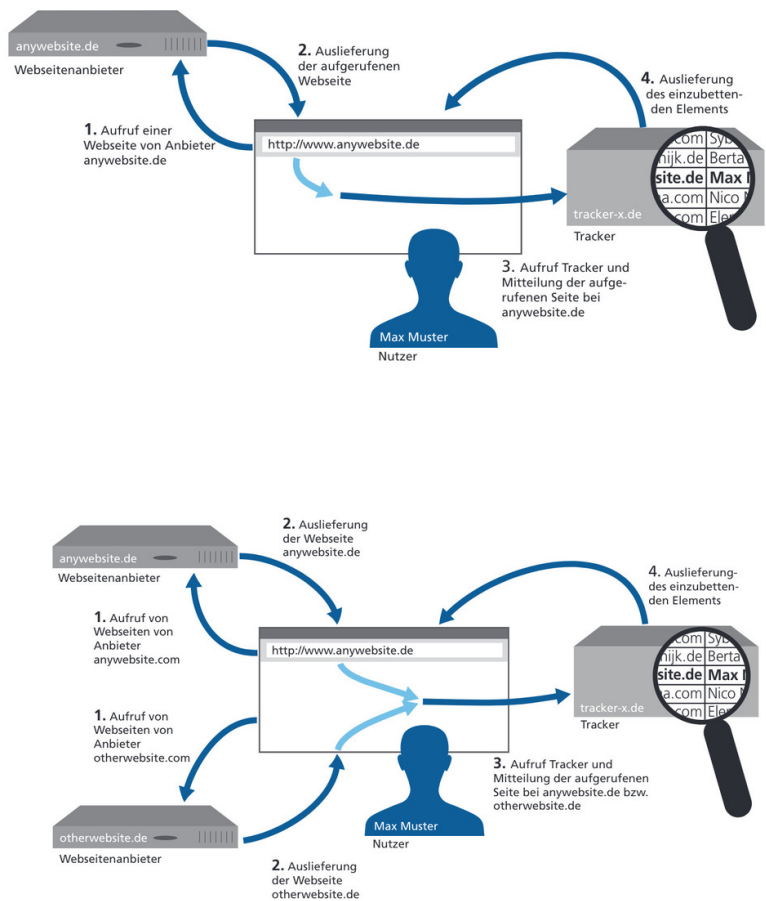
Wouters, Paul, & Costas, Rodrigo. 2012. *Users, Narcissism and Control—Tracking the Impact of Scholarly Publications in the 21st Century*. [https://www.researchgate.net/publication/268271559\\_Users\\_Narcissism\\_and\\_Control-Tracking\\_the\\_Impact\\_of\\_Scholarly\\_Publications\\_in\\_the\\_21\\_st\\_Century](https://www.researchgate.net/publication/268271559_Users_Narcissism_and_Control-Tracking_the_Impact_of_Scholarly_Publications_in_the_21_st_Century)

Xiong, Zehui, Niyato, Dusit, & Wang, Ping. 2018. Network Effects, in Shen, Xuemin, Lin, Xiaodong, & Zhang, Kuan (Hg.), *Encyclopedia of Wireless Networks*. Cham: Springer International Publishing, 1–7.

Zuboff, Shoshana. 2019. *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. London: Profile books.



# Anhänge



**ABBILDUNG 12:** Schematisches Web-Tracking und Cross-Domain-Tracking (Schneider et al. 2014, 22f.)

"\_\_\_\_-Party Data" indicates the source of your data.  
Third party data is on the decline. Zero and first party data is on the rise.

| Type of Data         | How do you get it?                                                            | Example method                                         | Sorts of data you collect                                  |
|----------------------|-------------------------------------------------------------------------------|--------------------------------------------------------|------------------------------------------------------------|
| Zero-Party           | Given to you by the person it relates to                                      | A user signs up for an account                         | Name, email, institutional affiliation                     |
| First-Party          | You observe it on your own properties                                         | A user visits a journal article                        | Interests in content types, formats, and topics            |
| Second-Party         | Provided by trusted third party, such as another publisher in a similar field | A partner shares a spreadsheet of event attendees      | Names and behaviors (who attended x event)                 |
| Third-Party (bought) | You buy it                                                                    | Buy an email list of researchers in a particular field | Name, email, institutional affiliation, job title, address |
| Third-Party (rented) | You rent it (the provider owns the data, you just use it)                     | Buy an ad on Facebook targeting academic researchers   | Minimal: summary statistics (views/clicks)                 |

ABBILDUNG 13: Übersicht der Datentypen durch Datenbroker Hum (hum.works)

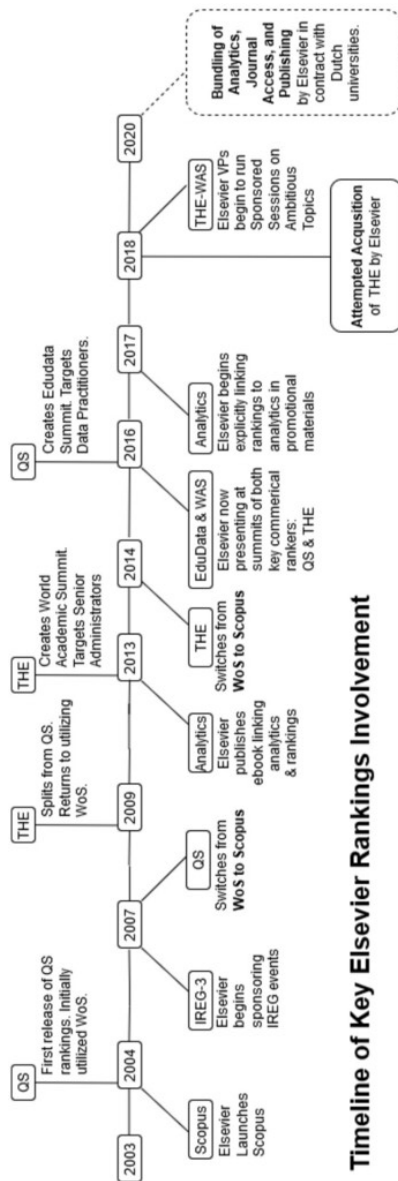
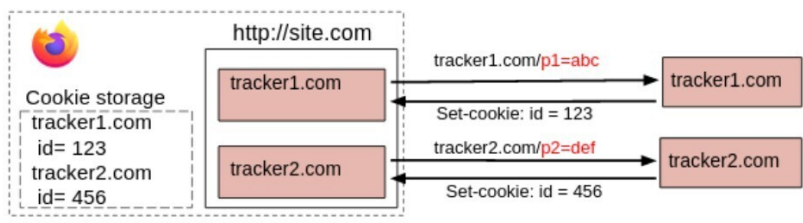


ABBILDUNG 14: Timeline Elsevier Rankings Involvement (Chen & Chan 2021, 12)



(a) Client-side tracking: Direct inclusion of third party services

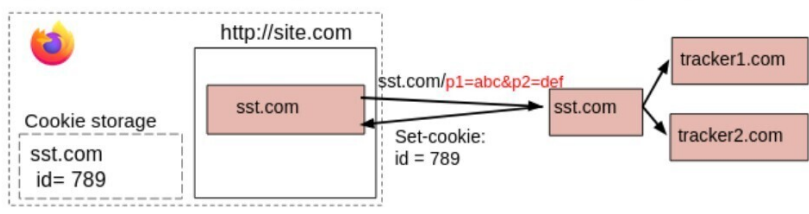


ABBILDUNG 15: Server-Side-Tracking (Fouad et al. 2024, 4)

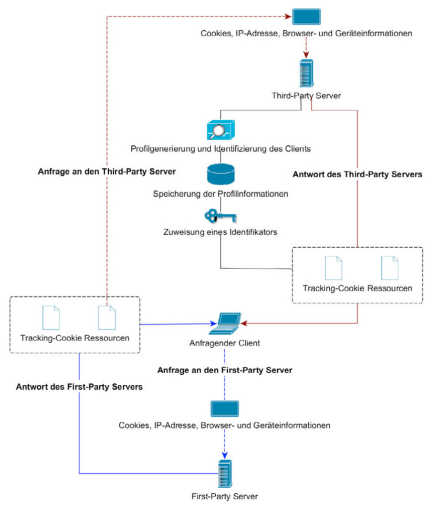
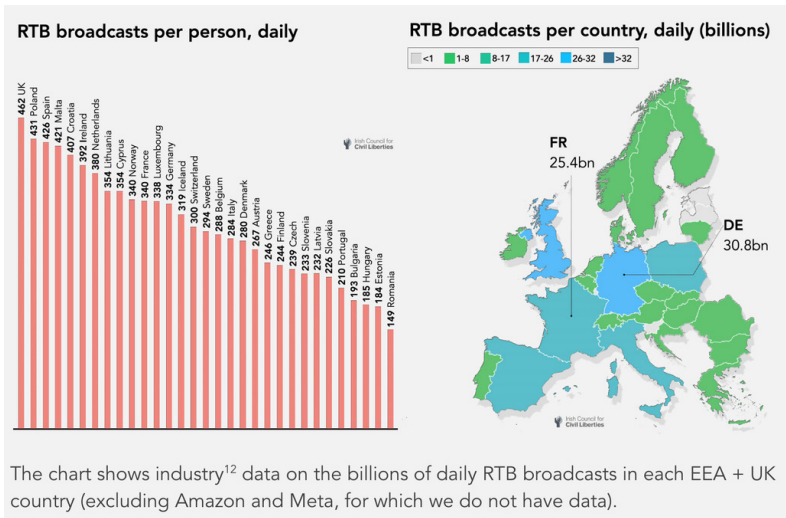


ABBILDUNG 16: Schematische Darstellung des Third-Party-Trackings (Freiberg 2022, 97)



**ABBILDUNG 17:** Tägliche Datenübertragungen pro Person und pro Land im weiteren europäischen Raum (Ryan & Christl n.d.)

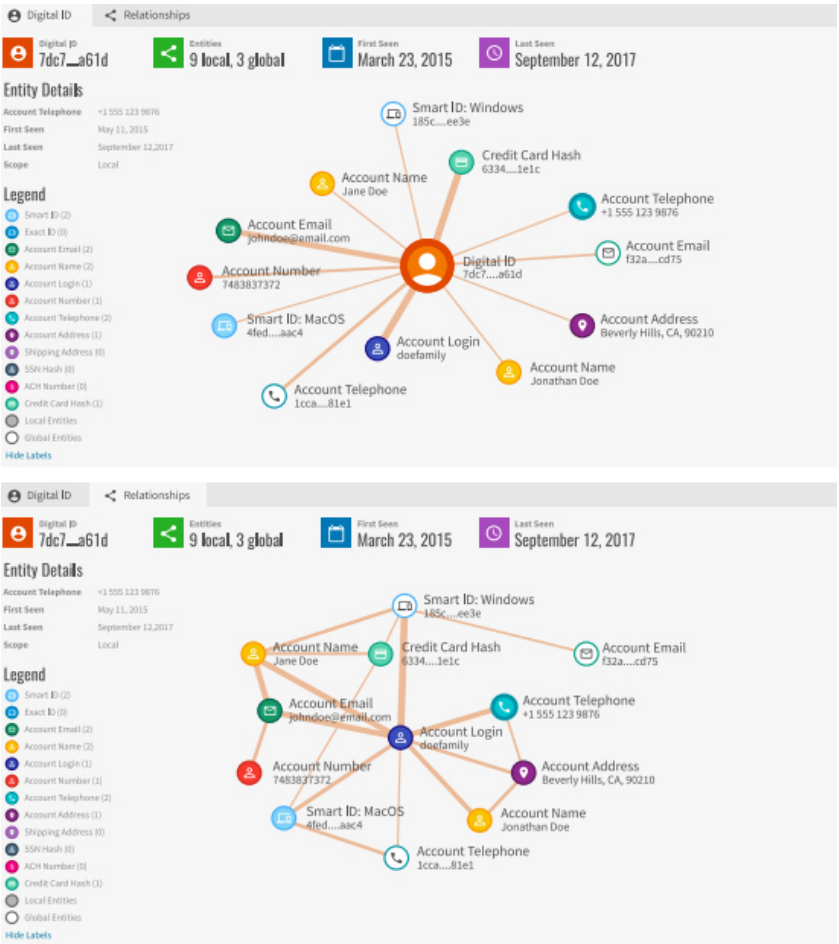


ABBILDUNG 18: LexisNexis Digital ID Netzwerk (LexisNexis n.d.a, 2)

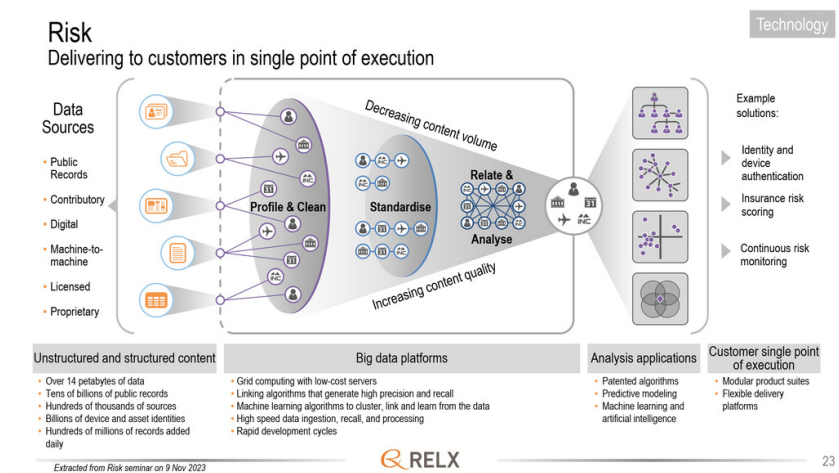


ABBILDUNG 19: RELX Risk-Datenaufbereitung (RELX 2024, Folie 23)

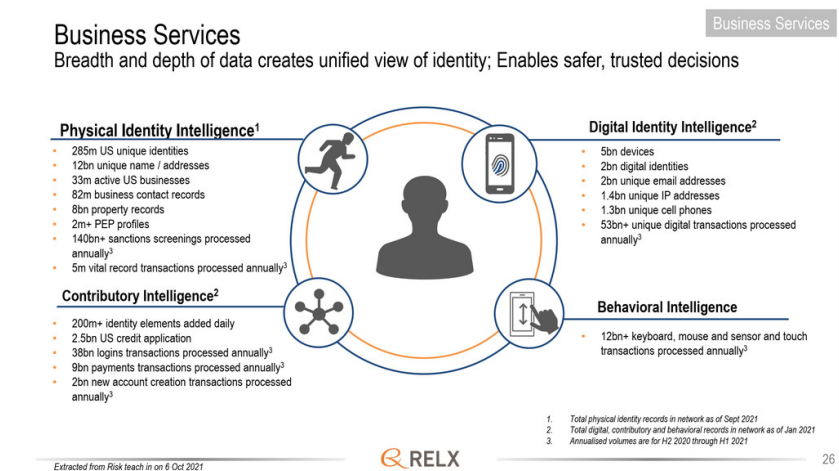
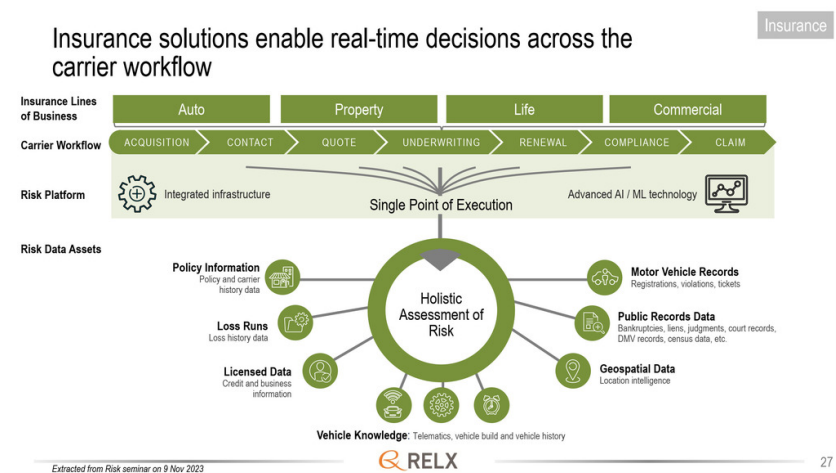
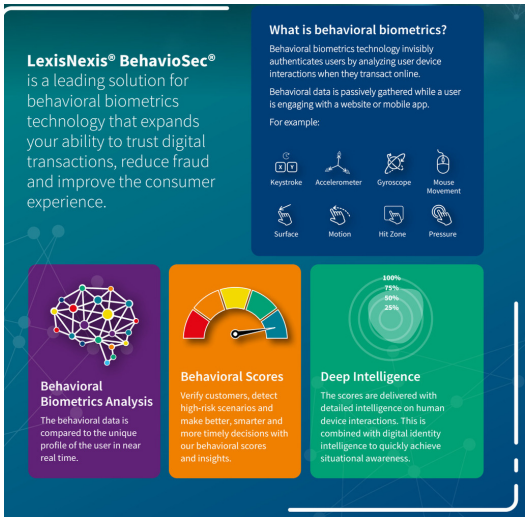


ABBILDUNG 20: RELX Business Services (RELX 2024, Folie 26)



**ABBILDUNG 21:** RELX Insurance – holistic assessment of risk (RELX 2024, Folie 27)



**ABBILDUNG 22:** LexisNexis BehavioSec behavioral biometrics (LexisNexis, 2)

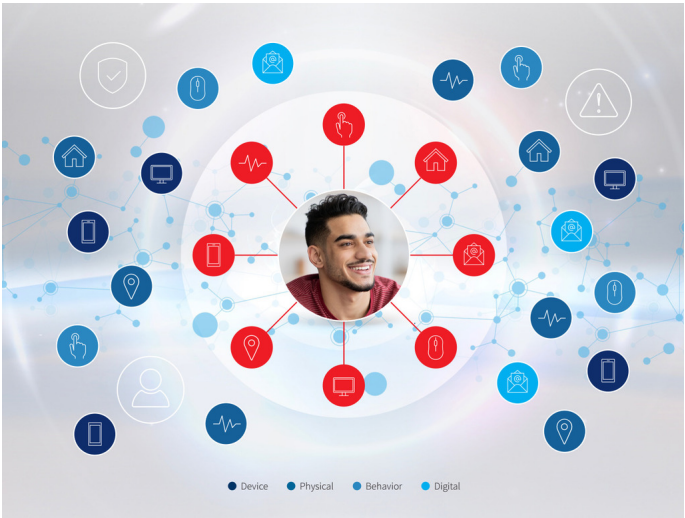
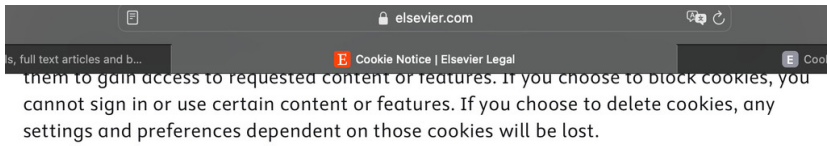


ABBILDUNG 23: LexisNexis Precise View of Identity (LexisNexis 2023, 3)



## Use of web beacons and other technologies

Some of our websites, applications and electronic communications contain electronic tags known as web beacons, gifs or pixel tags, unique identifiers and similar technologies that help deliver cookies, measure online activity, provide more relevant advertising, or analyze the effectiveness of our promotional campaigns or other operations.

We may also use technologies from third-party data security providers to maintain online security and protect our websites and other services against fraud and abuse.

**ABBILDUNG 24:** Cookie Notice Elsevier engl. (<https://www.elsevier.com/de-de/legal/cookie notice>, am 6.7.2024 Elsevier, n.d.a)



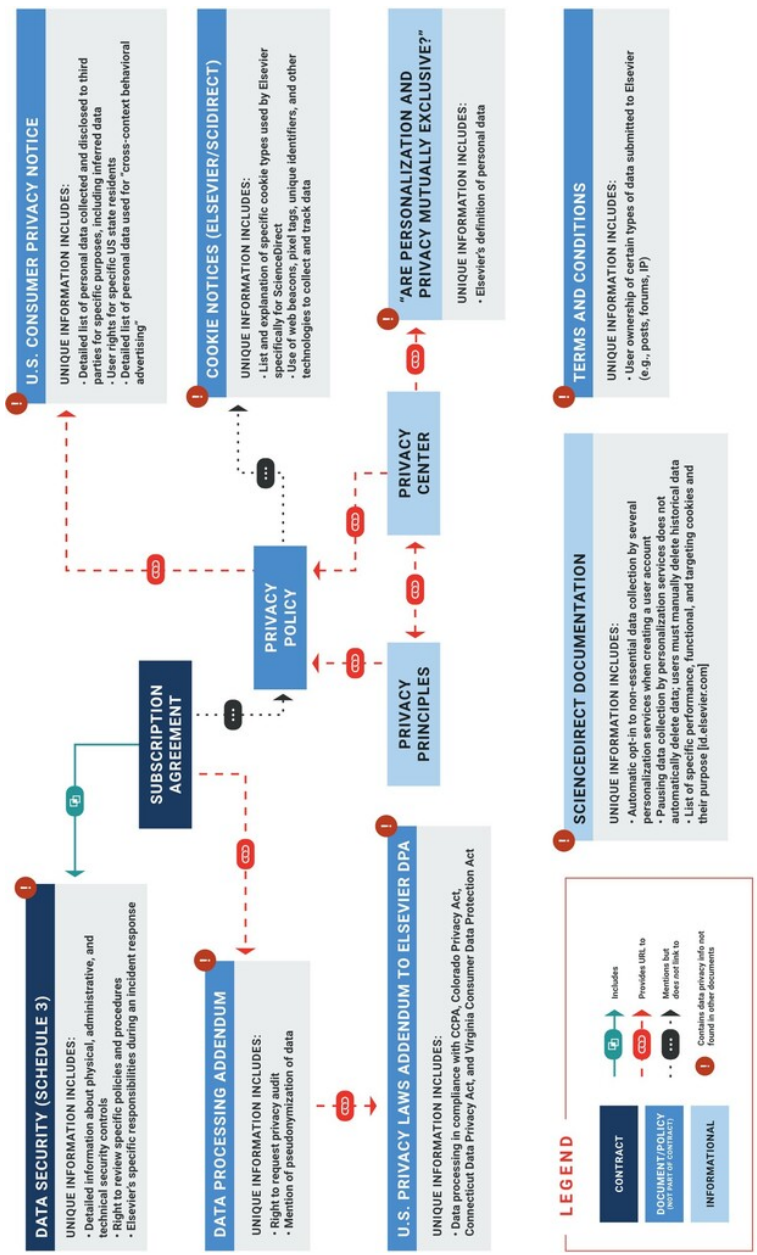
## Verwendung von Zählpixeln und anderen Technologien

Einige unserer Websites, Anwendungen und elektronischen Mitteilungen enthalten elektronische Kennzeichnungen, die als Zählpixel, Gifs oder Pixel-Tag bekannt sind, eindeutige Identifikatoren und ähnliche Technologien, die dabei helfen, Cookies zu liefern, die Online-Aktivität zu messen, zielgerichtete Werbeanzeigen bereitzustellen oder die Effektivität unserer Werbekampagnen oder anderer Vorgänge zu analysieren.

Wir verwenden außerdem Technologien von Drittanbietern von Datensicherheit wie zum Beispiel Threatmetrix, um Online-Sicherheit zu gewährleisten und unsere Websites und andere Dienste vor Betrug und Missbrauch zu schützen.

Zuletzt aktualisiert: 7. Dezember 2021

**ABBILDUNG 25:** Cookie Notice Elsevier dt. (<https://www.elsevier.com/legal/cookie notice-de-de>, am 01.07.2024 Elsevier, n.d.b)



**ABBILDUNG 26:** Zusammenhang der Elsevier Datenschutz-Dokumente (SPARC 2023, 16)



