

Stefan Klein, Joschka Hüllmann

Datenkapitalismus akademischer Wissenschaftsverlage

Elsevier ist ein global führender akademischer Wissenschaftsverlag, der sich auf die Bereiche Wissenschaft, Bildung und Gesundheit spezialisiert hat. Wie die gesamte Medienbranche sind die Wissenschaftsverlage von der Digitalisierung unmittelbar betroffen. Während die physische Produktion von Büchern und Zeitschriften auf dem Rückzug ist, bauen die Verlage digitale Archive ihrer herausgegebenen Beiträge auf, die zugleich als Vertriebsplattformen dienen. Dabei sehen sich Unternehmen wie Elsevier zunehmend nicht mehr als Verlag, sondern als „global information analytics company“¹. Die Analyse einiger Bausteine des Geschäftsmodells des Marktführers Elsevier veranschaulicht die Mechanismen dieser spezifischen Variante des Datenkapitalismus, der in ähnlicher Form auch von den anderen akademischen Verlagsgruppen praktiziert wird.²

Geschäftsmodell

Aus ökonomischer Sicht betreibt Elsevier eine spezielle mehrseitige Plattform:³

- 1 Elsevier: Editor Welcome Package, September 2017, S. 5, https://www.elsevier.com/_data/assets/pdf_file/0015/90123/EditorWelcomePack.pdf (9.7.2018).
- 2 Z. B. SpringerNature und Wiley.
- 3 Vgl. M. Kenney, J. Zysman: The Rise of the Platform Economy, in: Issues in Science and Technology, 2016, S. 61-69.

1. Auf der einen Marktseite stehen Autoren, Herausgeber und Gutachter, welche die von Elsevier vertriebenen Inhalte produzieren (Produzenten).
2. Auf der anderen Seite stehen die Leser, die von den reichhaltigen, digital verfügbaren Informations- und Serviceangeboten profitieren. Da sie die Kosten der Nutzung in der Regel nicht tragen, führen der einfache Zugang, kombiniert mit personalisierten Empfehlungen zu steigender Nutzungsintensität, d. h. zum Aufruf oder Download von Beiträgen (Konsumenten).
3. Bibliotheken von Forschungsinstitutionen und Universitäten finanzieren die Produzenten (1) und bezahlen für den Zugang der Konsumenten (2).

Gemäß der Logik mehrseitiger Märkte subventioniert (3) die Leser (2) und finanziert die Medienproduktion (1). Seitenübergreifende positive Netzwerkexternalitäten gibt es zwischen Marktseite (1) und (2): (1) schafft Angebot für (2), (2) schafft positive Reputationseffekte für (1). Entgegen der Logik mehrseitiger Märkte finden allerdings keine direkten Transaktionen zwischen den Marktseiten statt, sondern nur zwischen Marktseiten und Plattform. Gruppe (3) gewährt den Zugang für die Leser (2) (soweit es sich um Angehörige der eigenen Institution handelt) als Teil ih-

res Leistungsangebots, profitiert aber nur indirekt im Sinne der Abwendung negativer Reputationseffekte. Gruppe (3) profitiert durch positive Reputationseffekte von Publikationsleistungen von (1) aus der eigenen Institution. Insgesamt gibt es keine oder minimale Erträge für Gruppe (3) und negative seitenübergreifende Externalitäten durch die (auch) nachfrageorientierte Preispolitik der Plattformbetreiber (Verlage).

Digitale Archive wie ScienceDirect weisen starke Skaleneffekte auf: Die Betriebskosten pro Artikel sinken mit der Zahl der Beiträge, und die Attraktivität der Sammlung steigt mit ihrem Umfang (Bündelungseffekte). Nutzerzahl und Nutzungsfrequenz bilden die Basis von Empfehlungen, welche die Nutzungsfrequenz wiederum erhöhen und damit für Autoren, deren Institutionen und Zeitschriften den Impact-Faktor verbessern. Gold Open Access gegen Zahlung von Bearbeitungsgebühren (article processing charges, APC), ein Euphemismus für die Rückübertragung der Verwertungsrechte an publizierten Beiträgen, ist nicht nur äußerst lukrativ für die Verlage, es ist auch kompatibel mit dem Interesse der Autoren an erhöhten Downloadzahlen.

Die Verlage stellen den Bibliotheken nicht nur standardisierte Downloadstatistiken (COUNTER)⁴ und über Nutzerabfragen nicht lizenzierter Inhalte (turn aways) zur Verfügung, sondern bieten darüber hinaus weitreichende, kostenpflichtige Nutzungsstatistiken, z. B. über SciVal (Elsevier), an. Aus Sicht der Bibliotheken entspricht dies der Auslagerung der Datenanalyse an den Leistungsanbieter. Die Nutzungsstatistiken kombinieren anonyme Abfragen aus der von Elsevier registrierten und freigegebenen IP-Adresse, z. B. der WWU Münster, mit personenbezogenen Informationen der Nutzer, sofern diese sich für irgendeinen Dienst von Elsevier registriert haben.

Die Nutzungsstatistiken der Verlage ermöglichen den Bibliotheken nutzungsbezogene interne Abrechnungen, z. B. gegenüber Fachbereichen. Gleichzeitig sehen die Verlage diese Statistiken jedoch als Rechtfertigung steigender Preise.

Aufbau einer integrierten digitalen Plattform

Zwar bilden die digitalen Archive den Kern von Elseviers Plattform, sie werden aber erweitert durch ein Netzwerk digitaler Lösungen,⁵ welches die Wertschöpfungskette akademischer Wissensproduktion unterstützt. Dieses

4 The COUNTER Code of Practice, <https://www.projectcounter.org/code-of-practice-sections/usage-reports/> (4.7.2018).

5 Elsevier: All Elsevier digital solutions, <https://www.elsevier.com/solutions> (4.7.2018).

Netzwerk an Lösungen, darunter SCOPUS (Referenzdatenbank und SCImago Journal Rank), Mendeley (Literaturverwaltung) und SSRN (Archiv für graue Literatur und Pre-Prints), unterstreicht die strategische Reorientierung vom Anbieter von Inhalten zu einem Anbieter von Tools und Informationsdienstleistungen entlang der Produktion, Analyse, Archivierung und Distribution akademischer Publikationen, mithin für Autoren, Leser, Herausgeber und Gutachter sowie Forschungsinstitutionen und Bibliotheken.⁶ Die Dienstleistungen bedienen effizient die Bedürfnisse von Lesern und Autoren und ändern damit zugleich die akademische Wissensproduktion.

Für den Aufbau ihrer digitalen Archive akquirieren die Wissenschaftsverlage Verwertungsrechte der zu publizierenden Beiträge im Gegenzug und als Vorbedingung für die Veröffentlichung der Beiträge, d. h., zumeist ohne Vergütung der Autoren. Sie nutzen dabei die Bedeutung des Publikationserfolgs in namhaften wissenschaftlichen Zeitschriften als Bedingung akademischer Reputation und damit akademischer Karrierechancen aus. Die erzwungene Aufgabe der Rechte am intellektuellen Eigentum als Preis für die Publikation stellt dabei potenziell einen Missbrauch der Marktmacht dar.

Während die Reputation einer Zeitschrift in der akademischen Gemeinschaft und die damit einhergehende Bereitschaft von Herausgebern und Gutachtern erhebliche Arbeitszeit – ohne oder gegen marginale Vergütung – für die Publikation von Beiträgen zu investieren im Vordergrund der akademischen Aufmerksamkeit steht, sind Eigentum und Kontrolle an eben diesen Zeitschriften durch den Verlag im Hintergrund. So gelingt es den Verlagen, ein System freiwilliger und weitgehend unbezahlter Arbeit aufrechtzuerhalten⁷ und Umsatzrenditen von 40 %⁸ zu erzielen.

In dem Corporate-Responsibility-Bericht des Elsevier-Mutterkonzerns erhebt das Unternehmen Anspruch auf die Legitimität der Verwertungsrechte und bestätigt auch die strategische Neuausrichtung hin zu plattformübergreifender Datenanalyse: „... underpins our business strategy to deliver improved outcomes for our customers by combining content and data with analytics and technology across global platforms. It helps us build leading positions

6 R. C. Schonfeld: When is a Publisher not a Publisher? Cobbling Together the Pieces to Build a Workflow Business, in: The Scholarly Kitchen, 9.2.2017, <https://scholarlykitchen.sspnet.org/2017/02/09/cobbling-together-workflow-businesses> (4.7.2018).

7 G. Monbiot: Academic publishers make Murdoch look like a socialist | Comment is free, in: The Guardian, 29.8.2011, <https://www.theguardian.com/commentisfree/2011/aug/29/academic-publishers-murdoch-socialist> (9.7.2018).

8 RELX Group: Corporate Responsibility Report 2017, London 2017, S. 5, <https://www.relx.com/-/media/Files/R/RELX-Group/documents/responsibility/download-center/relx2017-cr-report.pdf> (9.7.2018).

in our markets by leveraging *our* skills and assets.“⁹ Die Analyse von Soros trifft das sorgfältig aufgebaute Selbstbild: „The exceptional profitability of these companies is largely a function of ... avoiding paying for the content on their platforms. They claim they are merely distributing information. But the fact that they are near-monopoly distributors makes them public utilities and should subject them to more stringent regulations, aimed at preserving competition, innovation, and fair and open universal access.“¹⁰

Datensammlung und -aggregation

Elsevier sammelt zunehmend umfassendere und detaillierte Daten über die Nutzer seiner digitalen Infrastruktur: über das Lese- bzw. Downloadverhalten, die Sammlung und Annotation von Quellen in Mendeley, die Nutzung von SSRN (Upload und Download), die Reviewtätigkeit (Inhalt der Reviews, Umfang der Reviewtätigkeit, Termintreue etc.), aber auch institutionelle Forschungsschwerpunkte. Die von Elsevier gesammelten Informationen sind potenziell für Wissenschaftsspionage (Wer forscht an welchen Themen?) wie für Überwachung (Welche Literatur wird gelesen und genutzt?) geeignet.¹¹

Die Datenschutzerklärung von Elsevier dokumentiert Umfang und Zwecke der Datensammlungen über Nutzer, deren Verhalten und deren Weitergabe an Dritte: „We collect information about you in three ways: directly from your input, from third-party sources, and through automated technologies. ... Content that you upload and share or store in your account, such as annotations, comments, contributions and replies ... We are committed to delivering a relevant and useful experience to you. Depending on how you interact with us and the Service, we use your personal information to: ... Deliver targeted advertisements, promotional messages, and other information related to the Service and to your interests; ... Identify usage trends and develop data analysis, including for the purposes of research, audits, reporting and other business operations, such as to pay royalties and license fees to third-party content providers, ... evaluate our business performance.“¹² Dazu kommen „details and other information about you from other entities within the RELX Group and from third parties, including: Social networks ..., Service providers ..., Partners ... and/or publicly

available sources and data suppliers“ sowie technische Informationen wie IP-Adresse, Browser, Betriebssystem, benutztes Gerät, Referrer und mehr.

Die Menge und Reichhaltigkeit der Informationen lässt sich mit denen von Facebook oder Google vergleichen. Die Informationen werden aktuell verwendet, Elseviers eigene Angebote zu verbessern oder zu erweitern, etwa für Forschungsmonitoring und Impact-Analysen (z.B. Plum Analytics). Informationen über weitergehende Verwendung, z.B. Verkauf an Dritte, sind spärlich. Es gibt keine technische Transparenz oder Möglichkeit, interne Algorithmen nachzuvollziehen bzw. die gemachten Angaben zu überprüfen. Die Nutzer müssen den Aussagen Elseviers vertrauen oder auf die Nutzung der Dienste verzichten.

Monopolisierung

Nach Einschätzung des Bundeskartellamts¹³ ist Facebook ein marktbeherrschendes Unternehmen, sodass das Ausmaß der Sammlung und Verwendung von Nutzerdaten aus Drittquellen möglicherweise den Tatbestand des Missbrauchs einer marktbeherrschenden Stellung (in der Gestaltung der Nutzungsbedingungen) und Datenschutzverstöße erfüllt. Vergleichbar Elsevier: als Marktführer im Bereich akademischer Publikationen, für die es in den meisten Fällen keinen alternativen Zugang und keine Substitutionsmöglichkeit gibt, werden den Nutzern Bedingungen oktroyiert, denen sich die Nutzer, für deren Arbeit der Zugang zu spezifischen Beiträgen essenziell ist, kaum entziehen können. Sammlung und Nutzung von Fremddaten wie auch die Weitergabe von Nutzerdaten an Dritte sind intransparent und könnten Datenschutzverstöße darstellen. Viele Autoren haben angesichts der Reputation und Spezialisierung der von Elsevier vertriebenen Zeitschriften nur begrenzte Alternativen, z.B. Open-Access-Zeitschriften, ohne dabei Einbußen an ihren Karriereaussichten zu riskieren. Zwar ist diese Situation auch eine Folge von Fehlentwicklungen des akademischen Systems (einseitige Fokussierung auf Publikationserfolg in hoch gerankten Zeitschriften), aber sie wird durch das Geschäftsmodell von Elsevier systematisch gefördert, ausgenutzt und verschärft. Die potenzielle Interessenkollusion der Rollen Verlag und Rankingdienstleister widerspricht den Regeln guter Governance, ist aber Teil von Elseviers Geschäftsmodell.¹⁴

9 Ebenda, Hervorhebung nicht im Original.

10 G. Soros: Remarks delivered at the World Economic Forum, Davos, Schweiz, 25.1.2018, <https://www.georgesoros.com/2018/01/25/remarks-delivered-at-the-world-economic-forum> (9.7.2018).

11 S. Zuboff: Big Other: Surveillance Capitalism and the Prospects of an Information Civilization, in: Journal of Information Technology, 30. Jg. (2015), H. 30, S. 75.

12 Elsevier: Privacy Policy, 22.5.2018, S. 1-2, https://www.elsevier.com/_data/assets/pdf_file/0006/732795/Elsevier_Global_Privacy_Policy-2018-05-22.PDF (9.7.2018).

13 Bundeskartellamt: Vorläufige Einschätzung im Facebook-Verfahren: Das Sammeln und Verwerten von Daten aus Drittquellen außerhalb der Facebook Website ist missbräuchlich, Pressemitteilung vom 19.12.2017, https://www.bundeskartellamt.de/SharedDocs/Meldung/DE/Pressemitteilungen/2017/19_12_2017_Facebook.html (4.7.2018).

14 Ein aktuelles Beispiel einer Interessenkollusion: J. Tenant: Elsevier are corrupting open science in Europe, 29.6.2018, <https://www.theguardian.com/science/political-science/2018/jun/29/elsevier-are-corrupting-open-science-in-europe> (4.7.2018).

Geschäftsgebaren

So wie Facebook die Datenintegration mit WhatsApp vorangetrieben hat, hat Elsevier Mendeley und SSRN, beides vormals eigenständige Lösungen, die sich den Zielen der Forschungscommunity verpflichtet sahen, integriert. Elsevier setzt die eigenen Rechte im Zweifel gegen die eigenen Autoren durch und sendet Abmahnungen, wenn Artikel auf Academia.edu oder ResearchGate verfügbar gemacht wurden.¹⁵

2016 haben Eve, Lawson und Tennant Anklage gegen die RELX-Gruppe bei der britischen Kartellbehörde (UK Competition and Markets Authority) wegen Missbrauchs einer dominanten Marktposition und wettbewerbswidriger Praktiken erhoben.¹⁶ Hintergrund sind unter anderem die Geheimhaltungsklauseln in den Verträgen mit Bibliotheken, die den EU-Transparenzregeln bei öffentlicher Beschaffung widersprechen. Die Verhandlungen über ein wissenschaftsfreundliches Lizenzmodell für Bibliotheken in Deutschland im Rahmen des DEAL-Konsortiums haben auch nach zwei Jahren noch keine sichtbaren Fortschritte gemacht.

Das Ausmaß der Geheimhaltung und der Mangel an Transparenz widerspricht den Regeln und dem Geist einer kooperativen Beziehung zwischen Autor und Verlag, wie sie noch von Elseviers Logo symbolisiert und unter Verweis auf die Geschichte reklamiert werden: „Publishers and scholars cannot do it alone. They need each other. This remains as apt a representation of the relation-

15 S. Robinson: Elsevier issues academics thousands of takedown notices, doesn't want any friends anyway, 13.1.2014, <https://www.mh-pbooks.com/elsevier-issues-academics-thousands-of-takedown-notices-doesnt-want-any-friends-anyway> (4.7.2018).

16 M. P. Eve: Anti-competitive practices of RELX Group, 2016, <https://www.martineve.com/images/CMA-RELX.pdf> (4.7.2018).

onship between Elsevier and its authors today – neither dependent, nor independent, but interdependent.“¹⁷

Notwendige Debatte

Die Situation der dominierenden Wissenschaftsverlage ist vergleichbar mit der von Facebook oder Google: Die Erstellung und Nutzung der Kernleistung (wissenschaftliche Veröffentlichungen) wird durch ein wachsendes Repertoire digitaler Tools erweitert; beides bildet die Grundlage zur Sammlung von Nutzerdaten, die aus externen Quellen ergänzt werden – angeblich im Interesse der Nutzer. Soweit erkennbar werden die Daten bisher überwiegend zur Verbesserung des Angebots, zur Intensivierung der Nutzung (Empfehlungen) und zum Aufbau kostenpflichtiger statistischer Informationen für Bibliotheken und akademische Institutionen verwendet. Die digitale Wissensproduktion ermöglicht so die „Datafication“ der Wissenschaft, die als analytisches Wissenschaftsmanagement vermarktet wird und zugleich ungeahnte Potenziale der Überwachung einzelner Wissenschaftler und des Wissenschaftsbetriebs erzeugt.

Systematische Intransparenz und das Geschäftsgebaren von Elsevier begründen die Sorge, dass Elsevier seine Marktposition missbraucht, Datenschutzregeln missachtet und sensible Nutzerdaten ausbeuten könnte. Mehr Transparenz und wirksamere Governancestrukturen der marktführenden Wissenschaftsverlage sind daher erforderlich, aber auch eine breitere Diskussion möglicher Regeln zur Begrenzung der Ausbeutung von Autoren und Gutachtern sowie zur Sicherstellung eines angemessen bezahlten Zugangs zu den Archiven der Wissenschaftsverlage sind im Interesse der Gesellschaft.

17 Elsevier: History, <https://www.elsevier.com/en-gb/about/our-business/history> (4.7.2018).

Title: *Data Capitalism – an Economic Approach*

Abstract: *Companies aggregate, process and monetise data about consumers by leveraging actively shared information as well as the data they glean from tracking offline transactions. Data-driven business models come in two forms: Firms either use the personal data of users to tailor their own offerings to them, or they sell the data to others who use it for their own ends. These models have different implications for the protection of consumers. While they are typically reasonably well-informed about the data they transmit to a particular firm, they have less control about whom their data is sold on to. Consumers are regularly confronted with a lot of privacy-relevant information and choices, but they struggle to make decisions that are in line with their stated privacy preferences. Therefore, regulation should address privacy design concepts that actually foster control. The right to data portability is one of the fundamentally new elements of the General Data Protection Regulation. However, its economic implications are potentially very complex and, under certain circumstances, data portability may not even be in the interest of consumers. Nevertheless, the strengthening of consumer rights may likewise support competition. The monetary value of data is difficult to calculate. But if consumers are interested in a fair profit share of personal data markets, policy makers may disregard their claims. One article highlights the transformation in the academic publishing industry (STM) from a traditional focus on publishing services to a new emphasis on data analytics by reconstructing and illustrating the mechanisms of data capitalism. It offers an explanation for the growing profitability and concentration of the STM industry despite the growing open science movement.*

JEL Classification: D43, K24, L10, L40