

Der unsichtbare Dritte

Zum Datentracking in den Wissenschaften

| GERHARD LAUER | Persönliche Daten von Forscherinnen und Forschern dienen in immer größerem Ausmaß der Gewinnmaximierung von global agierenden Verlagen. Das Datentracking spielt dabei eine entscheidende Rolle.

Die Betriebsamkeit der akademischen Welt wurde in diesem Sommer 2024 kurz durch die Nachricht unterbrochen, dass der britische Verlagskonzern Taylor & Francis den Zugang zu Millionen von Forschungsaufsätzen an Microsoft verkauft habe, die Microsofts KI-Unternehmen „Open AI“ inzwischen zum Training seiner KI-Modelle nutze. Der Verlag soll dafür im ersten Jahr zehn Millionen Dollar erhalten haben, die Autorinnen und Autoren der Papiere dagegen wurden noch nicht einmal über den Datenhandel informiert. Ein Unternehmen, das neben Elsevier/RELX, Springer Nature und Wiley zu den weltweit größten Fachzeitschriftenverlagen zählt, agiert hier als alleiniger Rechteinhaber über die Beiträge in seinen mehr als 2 700 Fachzeitschriften und jährlich mehr als 2 000 Büchern – das ist jedenfalls das Selbstverständnis von Taylor & Francis. Das Wissen gehört weder den Wissenschaften noch der Öffentlichkeit, sondern allein dem Verlag.

Die Nachricht zeigt einmal mehr, um wie viel es in der Welt der Daten längst auch für die Wissenschaften geht. Wer entscheidet darüber, was mit dem Wis-

sen der Hochschulen und Universitäten geschieht, wer darf wie mit den Daten weiterarbeiten, wer verdient daran und wer wird noch nicht einmal gefragt? Wie Sarah Lamdan in ihrem Buch „Data Cartels“ eindringlich gezeigt hat, muss man nur dem Geld folgen, um die Zusammenhänge zwischen Daten, Macht und Geld zu sehen. Die Wissenschaften stehen bei diesem Spiel bestenfalls am Rand.

Asymmetrie im Verhältnis von Verlagen und Wissenschaft

Die Asymmetrie im Verhältnis von Verlagen und Wissenschaft ist keine Neuigkeit. Eine fairere Ausgestaltung des Verhältnisses war eins der Ziele, mit der die Allianz der deutschen Wissenschaftsorganisationen neue Vertragsmodelle mit den drei großen Wissenschaftsverlagen Elsevier, Springer Nature und Wiley 2014 auszuhandeln begonnen hat. Die Konsortialverträge des Projekts Deal sollten die Rechte und damit auch die Kosten wissenschaftsfreundlicher und einer offenen Gesellschaft angemessener gestalten. Die Bilanz fällt zehn Jahre später nüchtern aus. Bei den zuletzt geschlossenen Transformationsverträgen 2022/23 haben sich die Großverlage die Rechte insbesondere für KI-Anwendungen zu sichern versucht, wenn auch nicht erfolgreich. Auch beim Datentracking konnte keine für alle Seiten befriedigende Einigung gefunden werden.

Datentracking und Datenoligopole hängen eng zusammen, denn die Großunternehmen verstehen sich schon länger nicht mehr als Wissenschaftsverlage, sondern als Data-Analytics-Unternehmen, die datenbasierte Informationsdienste

anbieten. Im jüngsten Halbjahresbericht von RELX vom Juni 2024 wird beispielsweise die Notwendigkeit betont, persönliche Daten der Nutzerinnen und Nutzer auch von öffentlichen Einrichtungen wie Bibliotheken umfangreich zu erheben. Das ist schon länger gängige Praxis, nicht nur bei Elsevier. So werden Echtzeitdaten über Interessen und das Verhalten von Universitäten, Instituten und einzelnen Forscherinnen und Forschern gesammelt und von den Verlagen genutzt. Die durch die Erstanbietertracker der Verlage und durch die Drittanbietertracker erhobenen und gespeicherten Daten, etwa bei Literaturrecherchen in Hochschulbibliotheken, werden in verschiedenen Kontexten weiterverwertet, etwa bei Sicherheitsbetrachtungen, aber auch bei weiteren Analysen zusammengeführt, so dass daraus personalisierte Profile, Zugriffs- und Nutzungsdaten und Verweildauer bei Informationsquellen über einzelne Wissenschaftlerinnen und Wissenschaftler, Gruppen und Institute entstehen. Die Betroffenen können selbst nur sehr eingeschränkt einsehen, ob und wie lange diese Verhaltensdatenprofile gespeichert beziehungsweise an andere Interessenten weitergeleitet werden, denn zum einen sind die meisten Tracker in den Medienbereitstellungsdiensten der Verlage bereits eingebunden, zum anderen werden die Daten automatisiert an Dritte weitergeleitet. Datentracking ist auch in den Wissenschaften innerhalb weniger Jahre zu einer Routine der Verlage geworden und verstärkt die Machtasymmetrie zwischen Verlagen und Wissenschaften zusätzlich, statt sie abzubauen.

Regulierung des Datentrackings in den Wissenschaften

Das Datentracking widerspricht dem Grundauftrag der Bibliotheken und Hochschulen, den Schutz der Privat-

AUTOR



Gerhard Lauer ist Professor für Buchwissenschaft an der Johannes Gutenberg-Universität Mainz.

Foto: Dietrich Kühne



Foto: mauritius images / imageBROKER

sphäre der Nutzerinnen und Nutzer und ihrer personenbezogenen Daten zu gewährleisten. Da die Weitergabe von Daten das Vertrauensverhältnis zwischen den Einrichtungen und ihren Nutzerinnen und Nutzern verletzt, haben Bibliothekare und Wissenschaftler und die DFG deutliche Kritik an der etablierten Datentracking-Praxis formuliert. Das DFG-Informationspapier „Datentracking in der Wissenschaft“ von 2021 hat die Problematik ins allgemeine Bewusstsein der Wissenschaftsorganisationen gehoben, und das hat unter anderem dazu geführt, dass bei den letzten Deal-Verhandlungen vorgeschaltete Beratungen zur Frage der Regulierung des Datentrackings geführt worden sind. In der Europäischen Union ist das Datenschutzrecht zwar dank der DSGVO weitestgehend harmonisiert und verlangt zwingend eine umfassende Auskunftspflicht der Datentracker gegenüber den Betroffenen sowie eine ausdrückliche und freiwillige Zustimmung der Betroffenen zur Verarbeitung personenbezogener Daten, sobald ein Schritt über die urheberrechtlichen Verwertungsrechte und die Systemsicher-

»Bibliotheken, die Datenverarbeitung durch Verlage ermöglichen und selbst davon profitieren, können mit Bußgeldern belegt werden.«

heit hinausgegangen wird. Aber die Einhaltung der DSGVO obliegt sowohl den Verlagen als auch den Wissenschaftseinrichtungen. Bibliotheken, die Datenverarbeitung durch Verlage ermöglichen und selbst – sei es auch nur mittelbar – davon profitieren, können daher mit Bußgeldern belegt werden. Betroffene können sogar Schadensersatzansprüche und Unterlassungsansprüche geltend machen. Schließlich sind Wissenschaftseinrichtungen auch angreifbar, wenn

Deal-Verhandlungen 2022/23

Vonseiten der Wissenschaften und Bibliotheken wurden für die vorgeschalteten Verhandlungen zum Datentracking zwei

Dinge vorgeschlagen. Erstens sollte eine Musterklausel einer geteilten Verantwortlichkeit für das Datentracking zwischen Verlagen und Bibliotheken vereinbart und zweitens das Datentracking auf ein technisches und rechtliches Minimum reduziert werden. Eine geplante Muster-Joint-Controller-Vereinbarung sollte Fälle der gemeinsamen Verantwortung so weit wie möglich begrenzen, indem sie genau festlegt, in welchen Fällen die Verlage Daten an die Einrichtungen übermitteln dürfen.

Zudem wurden Anforderungen an die von den Verlagen beauftragten auftragsverarbeitenden Drittunternehmen, etwa für das Hosting oder beispielsweise eingebettete Programmbibliotheken, formuliert,

welche unter anderem Informationspflichten, Sicherheit und Vertraulichkeit nach der DSGVO spezifizieren. Im gleichen Zug sollten die Verlage verpflichtet werden, unabhängig von den gesetzlichen Pflichten jährlich eine Datenschutzfolgeabschätzung durchzuführen und das Ergebnis den Einrichtungen zur Verfügung zu stellen. Das hätte dann die Vorlage für weitere Verhandlungen auch mit an-

deren als nur den Großverlagen werden können, konnte aber nicht erfolgreich verhandelt werden. Weder eine Musterklausel noch eine Reduktion des Datentrackings noch eine Berichtspflicht konnten in den vorgelagerten Verhandlungen zum Datentracking erreicht werden.

Dass im Vorfeld der eigentlichen Deal-Verhandlungen zwischen Verlagen und Bibliotheken nach Lösungen gesucht wurde, mag als ein kleiner, aber eigentlich selbstverständlicher Schritt erscheinen. Wie wenig selbstverständlich er ist, macht die einhellige Ablehnung einer gemeinsamen Verantwortlichkeit durch alle drei Großverlage drastisch deutlich. Nach Meinung der Verlage lassen sich die Verarbeitungsschritte der Daten jeweils so klar voneinander abgrenzen und jeweils der Verantwortung der Verlage oder der Bibliotheken eindeutig zuordnen, so dass keine gemeinsame, sondern nur eine getrennte Verantwortung von den Verlagen akzeptiert wurde. Im Ergebnis gehen also die jetzt geschlossenen Verträge davon aus, dass keine gemeinsame Verantwortlichkeit von Wissenschaftsverlagen und Wissenschaftseinrichtungen im Sinne von Art. 26 Abs. 1 S. 1 DSGVO vorliegt. Nur im Innenverhältnis stellen die Verlage die Betreibergesellschaft MPDL Services und die durch den Vertrag gebundenen Einrichtungen von Schadensersatzansprüchen und anderen Folgen frei, falls im Nachhinein doch eine andere Rechtsauffassung von Datenschutzbehörden und/oder Gerichten vertreten werden sollte. Die Einrichtungen müssen also im Ergebnis die Folgen etwaiger Datenschutzverstöße finanziell

nicht tragen, stehen aber gegenüber Behörden und Betroffenen unverändert in der Verantwortung, gegebenenfalls auch in der Öffentlichkeit und auch vor Gericht, so konstatiert der Abschlussbericht der Ad-hoc-Arbeitsgruppe der DFG zum Datentracking in den Wissenschaften. Eine gemeinsame Verantwortung sieht meiner Ansicht nach anders aus.

Ebenso gescheitert sind die Versuche, das Datentracking einzuschränken und den Weiterverkauf von Daten über die Grenzen Europas hinaus zu unterbinden. Die Verlage verweisen zur Rechtfertigung eines Datentransfers etwa in Länder wie die Volksrepublik China auf einen 2023 von der Europäischen Kommission erlassenen Angemessenheitsbeschluss, der die bisherigen strengen Vorschriften im Kontext des Schrems II-Urteils (Übertragen personenbezogener Daten in Länder außerhalb Europas gilt im Zweifelsfall als rechtswidrig) geöffnet hat. Auch konnte keine Berichtspflicht der Verlage ausgehandelt werden.

Am Ende „Deal“

Inzwischen haben dennoch fast alle Wissenschaftseinrichtungen den neuen Deal-Verträgen zugestimmt, obgleich es teils heftige interne Auseinandersetzungen darüber an den Hochschulen gab und noch immer gibt. Die Argumente gegen den Abschluss der Verträge sind durch die jüngsten Verhandlungen nicht weniger geworden. An dem bizarren Geschäftsmodell des Maxwell-Garfield-Systems (Gewinne liegen ausschließlich aufseiten der Verlage) wissenschaftlicher Publikationen hat sich auch weiterhin

grundlegend nichts geändert.

Wie geht es weiter? Sollte es überhaupt zu weiteren Deal-Verhandlungen kommen, sollte Datentracking Teil der Hauptverhandlungen sein und nicht in gesonderte Vorverhandlungen ausgelagert werden, denn das Thema Datentracking wirkt sich auf fast alle Bereiche der Lizenzverhandlungen aus. Aber auch andere Szenarien sind denkbar. Jeder hat nach Art. 15 DSGVO die Möglichkeit, über Datenweiterverarbeitung und das Datenschutzniveau detaillierte Auskünfte einzufordern. Auskunftsanfragen, Beschwerden bei den jeweiligen Datenschutzaufsichtsbehörden oder Klagen von Wissenschaftlerinnen und Wissenschaftlern sind ein möglicher Weg, die rechtliche Grauzone aufzuhellen, in der die Deal-Verträge mit Blick auf das Datentracking geschlossen wurden. So hat jüngst die Gesellschaft für Freiheitsrechte gegen den unzulässigen Einsatz von Cookies durch die Wissenschaftsverlage Springer, Wiley und Nomos Datenschutzbeschwerde beim Landesbeauftragten für den Datenschutz und die Informationsfreiheit Baden-Württemberg erhoben. Ob am Ende die Wissenschaft noch gefragt wird, wenn ihr Wissen und ihre Daten in KI-Zeiten gehandelt werden, hängt auch von den beteiligten Wissenschaftlerinnen und Wissenschaftlern ab.

Eine Textfassung mit Literaturangaben kann bei der Redaktion von Forschung & Lehre angefordert werden.

ONLINE Forschung & Lehre

Publikationen sind eine wesentliche Währung in der Wissenschaft. Was macht eine gute Veröffentlichung aus? Wie verändern sich Publikationswege? Welchen Herausforderungen stehen Forschende gegenüber und wie können sie damit umgehen? Ausgewählte Artikel zum Thema finden Sie auf forschung-und-lehre.de.



Einfach den QR-Code scannen
oder über www.t1p.de/oxitw
aufrufen.