

Richtigstellung durch OCLC zum Aufsatz „Das Lesen der Anderen: Die Auswirkungen von User Tracking auf Bibliotheken“ von Dr. Renke Siems

Bénédict Ripperger, Head of Strategy DACH, OCLC GmbH

20.06.2022

Herr Dr. Siems hat in seinem Aufsatz ‚Das Lesen der Anderen: Die Auswirkungen von User Tracking auf Bibliotheken‘ (veröffentlicht in: O-Bib. Das Offene Bibliotheksjournal / Herausgeber VDB, 9(1), 1–25.¹) mehrmals Bezug auf OCLC genommen und Zusammenhänge hergestellt, Zitate zur Herstellung eines eigenen Narrativ genutzt, die einer Korrektur bedürfen. Sachverhalte sollten nicht einer Meinung untergeordnet werden, sondern den Tatsachen entsprechend präsentiert werden.

Grundsätzlich ist die Gefahr der Nutzerdatenaggregation und -auswertung, auch im Bibliotheksbereich gegeben und wir als OCLC haben uns klar dazu positioniert. Seit der Gründung 1967 ist der Sinn und Zweck von OCLC Bibliotheken (zunächst nur in Ohio, USA, jetzt weltweit) durch Services und Forschung zu unterstützen. Anders als kommerzielle Anbieter hat OCLC, als non-for-profit-Organisation eine bibliotheksgeführte Leitung (das sogenannte Board of Trustees), definiert sich nicht durch Gewinnmaximierung und hat eine langfristige, am Wohl der Bibliotheken, ausgerichtete Planung.

Schon Ende 2021 hat die DFG mit dem Informationspapier ‚Datentracking in der Wissenschaft – Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der DFG vom 28. Oktober 2021‘² auf die Thematik der (Nutzer)-Datenaggregation hingewiesen.

OCLC teilt die im Informationspapier und auch im Artikel von Herrn Dr. Siems dargelegten Bedenken.

Anders als andere Anbieter hat OCLC bereits im Jahre 2010 mit dem Verkauf von NetLibrary (als eBook-Inhalte-Plattform) an EBSCO eine strikte Trennung von eigenen Services und Inhalten vollzogen, um jeden Verdacht von Abhängigkeiten mit Verlagen und Content-Anbietern und unseren eigenen Zielen auszuräumen. Auch können wir einen Fokus auf die Nutzer- sowie Nutzungsdatenaggregation kategorisch ausschließen. Wie jedes Bibliotheksmanagementsystem und/oder Discovery-Service verarbeiten wir personenbezogene Daten (und dies natürlich DSGVO-konform), entgegen den im Informationspapier geschilderten Szenarien, setzen wir allerdings nur technisch notwendige Verfahren der Nutzerdatenverarbeitung ein, auch aggregieren wir weder Nutzungsdaten, noch kaufen oder verkaufen wir Nutzerdaten von/an Dritte. Um dies als ein Geschäftsfeld für OCLC zu definieren, müsste dies im Kreis des Boards of Trustees abgesegnet werden. OCLC veröffentlicht jährlich einen Jahresbericht, in dem wir über alle

¹ (Siems, 2022)

² (Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft, 2021)

strategischen und finanziellen Transaktionen Rechenschaft ablegen³. Der Handel mit Nutzungs- und/oder Nutzerdaten ist nicht mit dem Grundverständnis und den Zielen von OCLC in Übereinkunft zu bringen. Dies ist auch technisch in unseren Anwendungen und Services überprüfbar, sowie unseren Vertragsdokumenten sichtbar (eben da seitens von Bibliotheken OCLC keinerlei Rechte in dieser Hinsicht eingeräumt werden müssen).

Als Partner von Bibliotheken stellen wir mit EZproxy ein Tool zur Verfügung, das neben der Ermöglichung des Fernzugriffs auf eRessourcen auch die Identität von NutzerInnen schützt und den Umfang von Nutzerdaten, die an die Verlage und Inhalteanbieter weitergegeben werden, minimiert (in dem nur eine IP-Adresse und pseudonymisierte Daten weitergegeben werden). Dies ist ein Baustein, um dem Datenhunger von großen Inhalteanbietern auszubremsen.

Das Thema Datenaggregation an sich, gerade im Hinblick auf Forschungsdaten birgt natürlich auch Chancen, wichtig ist die Transparenz und die Freiwilligkeit hierbei. Mit dem Einsatz von OCLC-Software und der Nutzung unserer Services behalten Bibliotheken die Hoheit über Ihre Daten.

Vor diesem Hintergrund haben uns die formulierten Zusammenhänge und die Behandlung der Fakten durch Herrn Dr. Siems in seinem Aufsatz doch sehr verwundert. Hierzu einige notwendige Berichtigungen:

In Kapitel ‚3. Bibliothekssysteme‘ des Aufsatzes wird OCLC genannt und eine unsere ‚User Stories‘ zitiert. Die einleitende Prämisse vor dem Zitat ist die Aussage, dass die Bibliothekssysteme aller großen Anbieter mit dem Modell einer library analytics entwickelt worden seien. Suggestiert wird die grundsätzliche Annahme solche neuen Systeme würden Personendaten mit Mediennutzungsdaten sowie Finanzinformationen gleichermaßen als Analyseobjekte einer übergreifenden Data Analytics implementieren. Der Überbau wird dann mit der Referenz auf die vorangegangenen Kapitel noch weitergetrieben, indem das 3. Kapitel mit dem Satz „[d]ie kommerziellen Produkte im Bereich der Bibliotheksmanagementsoftware werden immer mehr Teil solcher Ökosysteme“⁴ beginnt. Die Gefahr ist nicht von der Hand zu weisen, allerdings ist OCLC als illustrierendes Beispiel hierfür mehr als unpassend. Wir betreiben kein Ökosystem, eben da OCLC nicht noch andere Betätigungsfelder neben dem Fokus auf bibliothekarische Services hat. Wir bieten nicht umfangreiche Inhalte über eigene Plattformen an und integrieren diese in eigene Softwarelösungen, wir sind nicht Verleger, wir sind nicht Teil eines, in der Hand von Private-Equity-Unternehmen klar ausgerichtetes, Unternehmenskonglomerats, das neben den Bibliotheken, in der Branche der analytischen Informationssysteme die Wissenschaftsproduktion als Ganzes im Auge hat und Milliardenumsätze verbucht.

So wird nun die User Story der University of Gloucestershire zitiert, in dem es um die Analyse von Nutzungsdaten von eMedien sowie Ausleihstatistiken im Hinblick auf die Studierenden der Universität geht.

Durch Auslassen der entscheidenden Informationen suggestiert das Zitat, sowie die weiteren Ausführungen von Herrn Dr. Siems, dass wir, OCLC, mit unserem cloudbasierten

³ Aktueller Jahresbericht für das Financial Year 21 hier zu finden: <https://www.oclc.org/en/annual-report/2021/home.html>

⁴ (Siems, 2022)

Bibliothekssystem WMS diese Analyse forcieren, überhaupt ermöglichen, um dann zu behaupten, „dass nicht nur Forschende und Studierende, sondern auch Bibliotheken und ihre Beschäftigten selbst Ziel einer datengestützten Ausbeutung und Enteignung sind“⁵.

Die Fakten sind anders und klar im ersten Absatz der User Story genannt:

„Die University of Gloucestershire ist stolz auf das hohe Niveau an Unterstützung, das sie ihren Studierenden bietet, und setzt sich selbst hohe Maßstäbe in Bezug auf Studierendenengagement und personalisierte Unterstützung. Diese Ziele passen hervorragend zur Vision von Jisc, deshalb haben OCLC und die University of Gloucestershire in Zusammenarbeit mit Jisc Datensätze aus WorldShare® Management Services (WMS) und auf EZproxy® gehosteten Diensten zur Nutzung freigegeben.“⁶

Es handelte sich hier um ein Projekt der University of Gloucestershire zusammen mit Jisc, der gemeinnützigen britischen Organisation zur Förderung digitaler Technologien in Forschung und Lehre. OCLC hat dazu beigetragen Daten aus den bei der University of Gloucestershire genutzten OCLC-Diensten dem Bibliotheksmanagementsystem WMS sowie EZproxy bereitzustellen. Diese Daten wurden mit anderen Daten - durch die Institution selbst - aggregiert und analysiert. OCLC selbst hat und hatte keinen Zugriff auf die Gesamtheit der Daten und auch keine Analysen durchgeführt. Es ging im vorliegenden Fall, um das Gegenteil von Ausbeutung und Enteignung, nicht um predictive analytics, sondern um ein wissenschaftliches Interesse einer Institution einen Sachverhalt zu untersuchen. Wir als OCLC unterstützen natürlich unsere Mitglieder bei Forschungsvorhaben. Der große Unterschied ist, dass nicht wir in unseren Systemen Daten aggregieren, mit oder ohne Zustimmung der betroffenen Personen, seien es Studierende oder BibliotheksmitarbeiterInnen, sondern eine Institution ihre eigenen Daten analysiert und wir hierbei unterstützen, damit die notwendigen Daten – soweit vorhanden – aus den genutzten Systemen verwendet werden können. Über den Ansatz, den im anglo-amerikanischen Raum als ‚student success‘ bezeichneten, Lernerfolg von Studierenden messbar zu machen, kann diskutiert werden, dies hat allerdings im vorliegenden Fall nichts mit dem Bibliothekssystem und OCLC zu tun. Ganz allgemein könnte man hier noch anführen, dass in Ländern mit (höheren) Studiengebühren, das Interesse, den Einfluss von Bibliotheken auf den Lernerfolg von Studierenden quantifizieren zu können, größer ist.

Wer dies als Beispiel für ‚datengestützte Ausbeutung und Enteignung mit/in Bibliothekssystemen‘ anführt, hat den Boden des Faktischen verlassen. Auch das in diesem Zusammenhang angeführte Zitat aus (Weingart, 2016) entbehrt jeder Grundlage: „Die Bibliotheken sehen sich in einer zunehmenden Abhängigkeit von den bibliotheks anbietenden Systemen wie Alma, Exlibris oder OCLC mit der Konsequenz, dass sie späterhin ihre eigenen Katalogdaten zurückkaufen müssen.“⁷

Ohne jetzt weiter auf die ursprüngliche Publikation, aus der dieses Zitat stammt und die Fehler dort einzugehen, ist die Behauptung – in Bezug auf OCLC - schlicht falsch. Aus

⁵ (Siems, 2022)

⁶ (OCLC, 2022)

⁷ (Weingart, 2016)

zwei Gründen kann es diese Gefahr, eigene Katalogdaten wieder zurückkaufen zu müssen, nicht geben, das sollte Herr Dr. Siems auch bekannt sein.

Ganz unabhängig von ‚bibliotheksanbietenden Systemen‘ gibt es in Deutschland sechs Bibliotheksverbände, die aktuell vier Verbundkataloge betreiben. Die Katalogdaten werden von dort bsp. auch in den WorldCat importiert (der durch OCLC seit mehr als 45 Jahren betriebene weltgrößte bibliographische Katalog) bleiben jedoch immer in der Hand der Verbände und damit der angeschlossenen Institutionen.

Der zweite Grund ist schlicht die vertragliche Zusicherung in unseren Verträgen mit unseren Mitgliedern, die jederzeit den Export der bibliographischen Daten festlegt, sowie die einfache Möglichkeit dies bsp. in WMS zu konfigurieren.

Dieses Zitat und der selbst geführte Vergleich mit Clickworkern bei Amazon ist für Herrn Dr. Siems ausreichend, um zu behaupten, dass es offenbar sei, dass „die Bibliotheken dann in der gleichen Weise ihre Rechte daran verlieren sollen, wie die Autor*innen beim klassischen Verlagsmodell“⁸.

Dies entbehrt jeder Grundlage, passt allerdings in den Grundtenor des Artikels, der eine Meinung transportiert, diese jedoch nicht mit Fakten stützt.

Das Vermengen von Zitaten, das mutmaßlich gezielte Zusammenbringen von nicht zusammengehörenden Fakten und Aussagen ist auch am Ende des 3. Kapitels festzustellen. Auch hier werden Aussagen seitens OCLC in ein Narrativ eingebaut, das schlicht falsch ist.

Zusammengefasst wird im letzten Abschnitt des Kapitels ein Zusammenspiel zwischen Verlagen und Anbietern von Bibliothekssystemen attestiert. Dies wird dann mit einer Aussage von Don Hamparian von OCLC untermauert, wobei auch hier in Bezug auf OCLC genau das Gegenteil zum zuvor skizzierten Interesse, der Fall ist.

Zunächst ist die Einschätzung nicht falsch, dass die Analyse von Nutzerdaten auf Verlagsplattformen kritisch zu sehen ist und auch die Lösung zur Ermöglichung des Fernzugriffs durch Bibliotheken auf dem Prüfstand stehen sollte und natürlich Verlage/Inhalteanbieter hier Ihre ganz eigenen Interessen vertreten.

An dieser Stelle ist nochmals hervorzuheben, dass OCLC aus Überzeugung eben nicht Anbieter von Bibliothekssystemen und Inhalteanbieter zugleich ist. Don Hamparian sprach in seiner Funktion als Senior Product Manager für EZproxy auf der ‚Scholarly Networks Security Initiative (SNSI)‘ Online Conference (STM - International Association of STM Publishers, 2021). Herr Dr. Siems zitiert nun sehr knapp genau zwei Aussagen von Don Hamparian, die er in einen Zusammenhang stellt, und zwar folgendermaßen: „Verlage und Systemanbieter scheinen sich hierbei einig zu sein, denn bei einem späteren SNSI-Webinar stellte Don Hamparian von OCLC „Libraries as Security Advocates“ vor, die selbstverständlich „Publisher Assets“ beschützten.“⁹ Im Satz zuvor wird von einem ‚Einblick in Mentalitäten‘ gesprochen und im ganzen Abschnitt geht es um die technischen Überlegungen, die Nutzung von sogenannten Schattenbibliotheken bei

⁸ (Siems, 2022)

⁹ (Siems, 2022)

Nutzern nachzuweisen. In diesem Zusammenhang wird nun also der oben wiedergegebene Satz als Ende des Abschnitts gewählt.

Auch hier wird durch Auslassung seitens Herrn Dr. Siems versucht eine Meinung durch Zitate zu unterfüttern. Es hilft schon ein Blick auf den ganzen Titel des Vortrags von Don Hamparian zu werfen, der wie folgt lautet: „Libraries As Security Advocates – Protecting Patron, Institution, Publisher Assets“.

Es geht keineswegs einzig um die Interessen der Verlage, sondern gerade um den Interessenausgleich zwischen NutzerInnen, den Bibliotheken und Verlagen. Im weiteren Verlauf des Vortrags von Don Hamparian wird auch klar verdeutlicht wie EZproxy funktioniert und wie einerseits gerade mit EZproxy eine möglichst ‚anonyme‘ Nutzung von eMedieninhalten der Verlage möglich ist, gleichzeitig Maßnahmen vorgesehen sind, die den Sicherheitsbedenken von Verlagen Rechnung tragen und grundsätzlich Bibliotheken ermöglichen einen datenschutzkonformen, funktional umfassenden und robusten Fernzugriff zu ermöglichen. Dabei ist die Institution mit EZproxy die einzige Stelle, die die Nutzung von eMedien einer spezifischen Person zuordnen könnte. Weder OCLC noch die Inhalteanbieter kommen in den Besitz der dafür notwendigen Daten – per Design, dies ist die Intention unseres Angebots.

Es muss zusammenfassend attestiert werden, dass Herr Dr. Siems OCLC-Inhalte in den Zitaten so ‚kuratiert‘, dass diese in (s)ein Narrativ passen. Dabei werden Aussagen klar entstellend aus dem Kontext gerissen. Dies entspricht nicht den Gepflogenheiten des wissenschaftlichen Arbeitens. Dies ist umso bedauerlicher, als das Thema selbst wichtig ist und von einem solchen Umgang mit Fakten nicht profitiert. Auch sollte man es sich nicht einfach machen und Schwarz-Weiß-Malerei betreiben. In Bezug auf Anbieter von Bibliotheksmanagementsystemen gibt es eben nicht nur die einen und die anderen.

Literaturverzeichnis

- Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft. (2021). *Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage*. Bonn: Deutsche Forschungsgemeinschaft e. V. doi:10.5281/zenodo.5900759
- OCLC. (17. 06. 2022). *Erfahren Sie, welche Bibliotheksressourcen den größten Einfluss haben*. Von <https://www.oclc.org/de/member-stories/gloucestershire.html> abgerufen
- Siems, R. (9(1) 2022). Das Lesen der Anderen - Die Auswirkungen von User Tracking auf Bibliotheken. *O-Bib. Das Offene Bibliotheksjournal / Herausgeber VDB*, S. 1-25. <https://doi.org/10.5282/o-bib/5797>.
- STM - International Association of STM Publishers. (12. 05 2021). *Scholarly Networks Security Initiative SNSI working together to combat the threat of cybercrime*. Von YouTube-Video: https://www.youtube.com/watch?v=HEBQyg_ezHI abgerufen
- Weingart, P. (2016). Zur Situation und Entwicklung wissenschaftlicher Bibliotheken. In N. T. Peter Weingart, *Wissenschaftliches Publizieren : zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung* (S. 103-121). Berlin / Boston: De Gruyter Akademie Forschung.