

# Barcamp Open Data – Forschen mit verfügbaren Daten

Alice Wagner und Tamara Heck

Mal angenommen, alle Forschende teilen ihre Forschungsdaten und können auf eine gut vernetzte Infrastruktur zugreifen, um Daten zu finden. Ist das der Schlüssel, um Open Data in der Forschung zu realisieren? Oder fehlt ein wichtiger Aspekt, der bei Open Data mitunter vernachlässigt wird? Denn neben dem Willen, Forschungsdaten zu teilen und einer guten Infrastruktur, die dies erst ermöglicht, stehen offene Forschungsdaten auch für die Nachnutzung bereit – und diese wird nicht automatisch durch das Teilen in Infrastrukturen realisiert. Die Nachnutzung, das heißt „Forschung mit verfügbaren Daten“, war daher Fokus des Barcamp Open Data.

Zentrale Diskussionspunkte bei der Nachnutzung von Forschungsdaten sind unter anderem das Potential der Nachnutzbarkeit. Hierbei spielt unter anderem die Dokumentation und technische Bereitstellung der Daten eine Rolle, deren Leitlinien unter dem Stichwort „FAIR Data“<sup>1</sup> diskutiert werden. Daneben sind fachspezifische Anforderungen an Forschungsdaten relevant, wie Zielgruppen, an denen Daten erhoben wurden. Kontroverse Ansichten gibt es weiterhin bezüglich Transparenz und Flexibilität bei induktivem und deduktivem Erkenntnisgewinn bei der Nachnutzung. Folglich entscheiden sich Forschende in der Tendenz

doch eher dazu, eigene Daten neu zu erheben. Auch das Teilen der eigenen Forschungsdaten und dessen Nutzen sehen viele Forschende noch skeptisch. Demgegenüber verspricht das Teilen und Nachnutzen von Forschungsdaten eine verbesserte Qualität und Transparenz von Forschungsdaten sowie Zeit und Ressourcenersparnis, die bei der Erhebung von neuen Daten anfallen würden. Das Barcamp Open Data bot eine disziplinübergreifende Plattform zur Reflektion über Chancen, Herausforderungen und Hand-

lungsoptionen beim Teilen und vor allem Nachnutzen von Forschungsdaten.

Der Impulsvortrag<sup>2</sup> zum Auftakt des Barcamps kam von Thomas Lösch, Mitarbeiter am Forschungsdatenzentrum Bildung am DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation. Lösch begann mit einer Begriffsklärung zur Bedeutung von „Nachnutzung von Daten“. Denn das Nachnutzen kann vielseitig verstanden werden: Forschende können Daten anderer nachnutzen, sie können eigene ältere Daten nachnutzen, sie können mit verfügbaren Daten wissenschaftliche Erkenntnisse überprü-

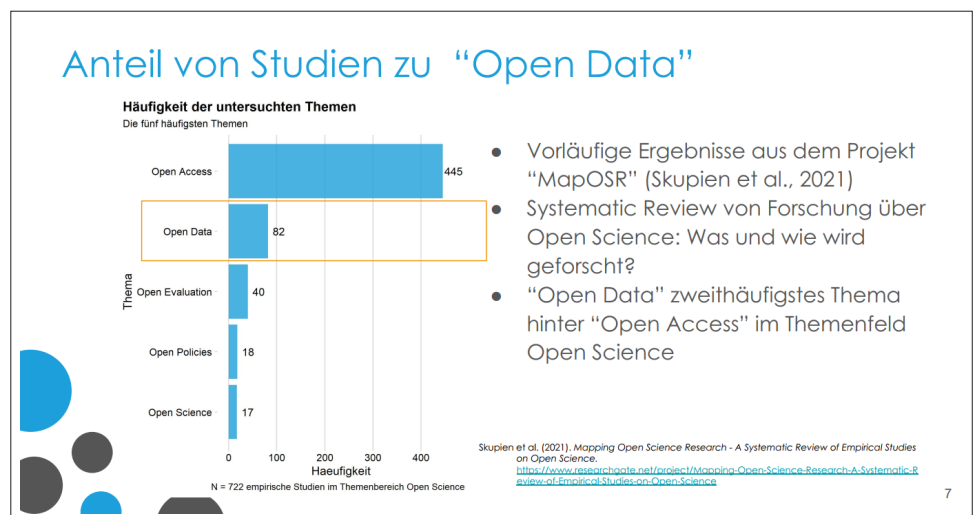


Abbildung 1: Folie aus dem Impulsvortrag von Thomas Lösch zu einer Reviewstudie über empirische Arbeiten zu Open Science. Link zur Studie von Skupien et al. (2021):

<https://www.researchgate.net/project/Mapping-Open-Science-Research-A-Systematic-Review-of-Empirical-Studies-on-Open-Science>

fen, aber auch neue Fragestellungen mit diesen Daten beantworten. Hinsichtlich dieser Vielfältigkeit und in Bezug auf wissenschaftliche Fragestellungen und Methoden, sind die Potentiale des Nachnutzens von Daten nicht in allen Forschungskontexten gleich bzw. werden von Forschenden anders wahrgenommen. Dass Open Data und dessen Facetten in vielen Studien erforscht wird, zeigte Lösch am Beispiel eines aktuellen Reviews (Abb 1). Nach empirischen Studien zu Open Access ist Open Data das zweithäufigste

<sup>1</sup> <https://www.go-fair.org/fair-principles>

<sup>2</sup> Verfügbar unter <https://osf.io/xh2sa>

12:00	SESSIONS I	FORSCHEN MIT "LINKED OPEN DATA" Daria Tisch	MAL ANGENOMMEN... OPEN DATA IST SCHON REALISIERT Alexia Meyermann	FORSCHEN MIT VERFÜGBAREN QUALITATIVEN DATEN Christoph Leser	WORAUF ACHTEN BEIM TEILEN VON DATEN Sami Nenno, Christian Blötner
13:00	LUNCH BREAK				
13:45	UPDATE SESSIONPLANUNG				
14:00	SESSIONS II	KONKRETE VORTEILE VON DATA SHARING Robert Lipp, Katja Sternitzke	WORAUF ACHTEN: DATEN (INTERDISZIPLINÄR) NACHNUTZBAR MACHEN? Matthias Fuchs	OPEN GOVERNMENT DATA Jesper Zedlitz	DATENDOKUMENTATION Erich Weichselgartner
15:00	SESSIONS III	WERKSTATTBERICHT MOBILITY KOMPASS Matthias Fuchs	WIE KANN ICH DATEN ZUR SEKUNDÄRNUTZUNG ZU BESTIMMTEN VARIABLEN/ NACH DATENSTURKTUR FINDEN? Elisabeth Fischer	REPLICATION WIKI SOZIALWISSENSCHAFTEN Jan H. Höfler	

Thema im Themenfeld Open Science. Lösch stellte in diesem Zusammenhang aktuelle Studien vor, die den Einfluss von Faktoren auf das Teilen und Nachnutzen von Daten untersuchen. Neben persönlichen Faktoren wie Motivation und Einstellung sind dies u.a. das Verhalten innerhalb einer Disziplin sowie das Angebot an Daten, bspw. über gut strukturierte Infrastrukturen, die das Auffinden von Daten erleichtern.

Gemeinsam mit Kolleginnen und Kollegen aus dem DIPF und den weiteren Leibniz-Instituten GESIS, ZPID und ZBW geht Lösch im aktuellen Projekt Fonik<sup>3</sup> der zentralen Frage nach, ob die aktuelle Pandemie die Forschungspraktiken verändert hat und Forschende mehr verfügbare Daten nutzen. Hintergrund ist zum einen die zunehmende Verfügbarkeit von Forschungsdaten, einerseits in Online-Repositoryn, andererseits in professionellen Forschungsdatenzentren, die Daten und Dokumentation langfristig verfügbar machen. Zum anderen waren Datenerhebungen während COVID-19 für viele Forschende mitunter nur eingeschränkt oder gar nicht möglich, sodass diese Forschungsprozesse und Methoden anpassen oder ändern mussten. Das Projekt Fonik untersucht mithilfe einer Umfrage die Auswirkung der Pandemie auf die Nachnutzung von Daten sowie für Forschende ausschlaggebende Aspekte, Daten nachzunutzen.

Dem Impulsvortrag folgte eine gemeinsame Planung der Barcamp-Sessions unter aktiver Beteiligung aller Teilnehmenden. Auch als Online-Veranstaltung aufgrund der COVID-Situation behielt das Barcamp seinen typischen Charakter. Das heißt, die Beiträge und Themen der einzelnen Sessions des Barcamps wurden nicht vorab bestimmt, sondern von den Teilnehmenden zu Beginn der Veranstaltung vorgeschlagen und in einer gemeinsamen Abstimmung festgelegt. Am Ende konnten die Teilnehmenden zwischen 11 Sessions wählen (Abb 2).

**Abbildung 2: Screenshot des Barcamp-Sessionplans mit den hauptverantwortlichen Teilnehmenden, die die Themen einbrachten.**

Link: <https://barcamps.eu/barcamp-open-data/events/6353521d-f2cf-430e-9758-80fcab7a0362#sessions>

Eine zentrale Frage auf dem Barcamp war, wie Open Data, sprich das Teilen und Nachnutzen von Forschungsdaten, in Disziplinen bzw. bei Forschenden etabliert werden kann, wenn dies noch nicht etablierte Praxis und noch kein Teil der Forschungskultur ist. Um Open Data zu realisieren, müssten Anreize für die datengebenden Forschenden geschaffen und Vorteile sichtbar werden (siehe z.B. Webseite der ZBW<sup>4</sup>). Diskutiert wurden Vorteile wie ein Zitationsgewinn durch Datenpublikationen, ein Reputationsgewinn für die Forschungskarriere sowie neue Kooperationsmöglichkeiten. Die Vorteile für die Wissensgemeinschaft sind die verbesserte Zugänglichkeit zu Daten und daraus folgend eine effizientere Forschung und Lehre. Auch die Öffentlichkeit profitiert von Open Data, indem Möglichkeiten für Citizen Science eröffnet werden und wirtschaftliche Vorteile für die Privatwirtschaft entstehen. Das Vertrauen der Öffentlichkeit in die Wissenschaft wird gestärkt und Forschungsergebnisse werden für diese transparent und nachprüfbar gemacht.

Um diese Vorteile zu nutzen, ist bei der Veröffentlichung von Forschungsdaten eine gute Dokumentation essentiell. Für eine qualitativ hochwertige Dokumentation, zu der auch ein forschungsprozessbegleitendes Forschungsdatenmanagement sowie ein Datenmanagementplan gehören, bedarf es laut Teilnehmenden Ressourcen für das Forschungsdatenmanagement von Drittmittelgebern. Diese Ressourcen sind in erster Linie Personalmittel für die Datenaufbereitung, außerdem Sachkosten für die Archivierung der Daten bei (fach-)spezifischen Forschungsdatendienstleistern und Repositoryn, wie bspw. Forschungsdatenzentren. Des Weiteren werden Dokumentations- und Metadatenstandards gebraucht, um verfügbare Daten transparent und nachnutzbar zu machen. In diesem Zusammenhang wurde auf das vom Bundesministerium für Bildung und Forschung geförderte Forschungsprojekt PsyCuraDat<sup>5</sup> aufmerksam gemacht, in dessen Rahmen nutzerorientierte Kurationskriterien entwickelt werden, welche auf die Bedürfnisse der Forschenden in ihrer Rolle als Datengebende und Datennehmende eingehen. So soll ein Standard für die Dokumentation von psychologischen

3 <https://www.leibniz-openscience.de/forschung/projekte/laufende-projekte/reusing-research-data-in-a-time-of-crisis-a-change-in-research-practices-in-the-covid-19-pandemic-fonik>

4 <https://openeconomics.zbw.eu/open-data>

5 <https://leibniz-psychology.org/institut/drittmittelprojekte/psycuradat>

Forschungsdaten entstehen und die Dokumentation sowie die Nachnutzung dieser Daten effizienter und effektiver gemacht werden. Im Gegensatz dazu wurde von Teilnehmenden angemerkt, dass in anderen Disziplinen eine ausreichende Infrastruktur zur Realisierung von Open Data fehle.

Dieser Aspekt knüpft an die Herausforderung an, dass Daten für alle gut such- und auffindbar sein müssen, um überhaupt eine Nachnutzung zu ermöglichen. Neben der sich im Aufbau befindenden nationalen Forschungsdateninfrastruktur<sup>6</sup> können Metasuchmaschinen Abhilfe schaffen, wie die Angebote B2FIND<sup>7</sup> und das GESIS-Portal<sup>8</sup>. Portale für Open Government Data sind bspw. das nationale<sup>9</sup> und das europäische Datenportal<sup>10</sup> sowie Bundesländerportale, wie das von einem Teilnehmer vorgestellte Open Data-Portal Schleswig Holstein<sup>11</sup>.

Auch Lehrende sollten Studierende bei der Suche nach passenden Datensätzen für Abschluss- oder Seminararbeiten unterstützen. Voraussetzung hierfür ist natürlich, dass Forschungsdaten für Studierende oder andere nicht-beruflich Forschende zugänglich sind. Im weiteren Verlauf der Diskussion ging es dabei auch um rechtliche Aspekte beim Zweck der Nachnutzung von Daten. Hier muss zwischen Forschung und Lehre unterschieden werden. Sollen Forschungsdaten für die Lehre genutzt werden dürfen, muss explizit das Einverständnis von Studienteilnehmenden vorliegen. Der einfachste Weg ist, dieses Einverständnis schon vor der Datenerhebung einzuholen, was jedoch aktuell noch oft vergessen wird.

Neben zahlreichen weiteren Diskussionspunkten, sammelten die Teilnehmenden dank gemeinschaftlicher Dokumentation via Online-Pads Links und Tipps zu relevanten Datenportalen, Best Practices und Vorlagen zum Datenmanagement sowie aktuellen Projekten und Datenrichtlinien. Die Dokumentationen der einzelnen Sessions sind mit Klick auf den Sessio- nentitel im Sessionplan (Abb. 2) einsehbar.

Der rege Austausch unter den etwa 100 Teilnehmenden des Barcamps kam auch zustande, weil sowohl Forschende, Studierende als auch Mitarbeitende aus Forschungsinfrastruktur-Instituten aus unterschiedlichen Disziplinen teilnahmen. So wurde das Thema Open Data aus verschiedenen Perspektiven betrachtet und diskutiert. Der Netzwerkcharakter eines Bar-

## Infokasten:

Das Barcamp Open Data am 21.09.21 wurde gefördert durch den Leibniz Forschungsverbund Open Science und veranstaltet in Kooperation mit dem Barcamp Open Science, das jährlich im Frühjahr stattfindet (wieder am 7.03.2022). Organisiert wurde das Barcamp vom DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation und dem Leibniz-Institut für Psychologie (ZPID), in Kooperation mit der Medienpädagogin Kristin Narr, die die Moderation übernahm. Eventseite und Dokumentation der Session-Diskussionen:

<https://barcamps.eu/barcamp-open-data>

camp ist online sicherlich schwerer umzusetzen als bei physischen Treffen, dennoch gab es digitale Meet- and Greet-Optionen zum Austauschen und Plaudern – und um in dieser digitalen Welt zumindest den Hauch eines Social Events nachzuahmen, wurden die Teilnehmenden am Vorabend zu einem Online-Spiel eingeladen und lösten gemeinsam einen Sherlock Holmes-Fall. Die Idee war sehr gelungen und eine gute Alternative zum physischen Netzwerken. Barcamp-Treffen vor Ort ersetzt die Onlinevariante jedoch nicht vollends.

Aufgrund der digital stattfindenden Open Science Conference wird aber auch das Barcamp Open Science<sup>12</sup> am 7. März 2022 wieder digital stattfinden. Für das Community Building soll es dann nochmal verstärkt Optionen zum Netzwerken und Austausch geben. |

### Dr. Tamara Heck

Informationszentrum Bildung  
DIPF | Leibniz-Institut für Bildungsforschung  
und Bildungsinformation  
heck@dipf.de

### Alice Wagner

unterstützte das Barcamp als studentische  
Hilfskraft am DIPF.

<sup>6</sup> <https://www.nfdi.de>

<sup>7</sup> <http://b2find.eudat.eu>

<sup>8</sup> <https://datasearch.gesis.org>

<sup>9</sup> <https://govdata.de>

<sup>10</sup> <https://data.europa.eu>

<sup>11</sup> <https://opendata.schleswig-holstein.de>

<sup>12</sup> <https://www.open-science-conference.eu/barcamp>