

Forschungsdatenmanagement in den Wirtschaftswissenschaften

VDB-Fortbildungsveranstaltung in der Universitätsbibliothek Frankfurt am Main, 19./20. September 2016

Das Thema Forschungsdatenmanagement stand am 19./20. September im Mittelpunkt der Fortbildungsveranstaltung für Fachreferentinnen und Fachreferenten der Wirtschaftswissenschaften.¹ Die bereits seit 1997 von der Kommission für Fachreferatsarbeit des Vereins Deutscher Bibliothekarinnen und Bibliothekare (VDB) ausgerichtete Tagung fand in diesem Jahr in der Universitätsbibliothek Johann Christian Senckenberg in Frankfurt am Main statt. Etwa 50 Teilnehmende aus Deutschland, Österreich und der Schweiz zeigten reges Interesse an dem sehr aktuellen Thema. Viele Kolleginnen und Kollegen nutzten dabei bereits das traditionelle Get-together am Vorabend der Veranstaltung, um sich in einem informellen Rahmen kennenzulernen und auszutauschen.



Park des IG-Farben-Hauses (Foto: Goethe-Universität Frankfurt)

Peter Bachofner (VDB) eröffnete die Tagung mit einem Dank an die Frankfurter Kolleginnen Ellen Haß und Dagmar Gärtner für die Vorbereitung der Veranstaltung. Der Direktor der Universitätsbibliothek Frankfurt, Heiner Schnelling, gab in seiner Begrüßungsansprache zunächst einen Überblick über die

1 Das Tagungsprogramm ist auch auf der Website der Universitätsbibliothek Frankfurt am Main einsehbar: „Fachreferententagung Wirtschaftswissenschaft“, Universitätsbibliothek Johann Christian Senckenberg, zuletzt geprüft am 01.11.2016, <http://www.ub.uni-frankfurt.de/tagung/wiwi2016/programm.html>.

wechselvolle Geschichte des Veranstaltungsgebäudes.² Enrico Schleiff, Vizepräsident der Universität, hob in seiner Begrüßung sodann vor allem die Bedeutung von Frankfurt als Ort des Austausches von Wissenschaft und Wirtschaft hervor. In diesem Zusammenhang stellte er das SAFE Data Center vor.³ Dessen Anliegen ist es, Forschenden – neben den bisher dominierenden amerikanischen Datenquellen – verstärkt auch europäische Datensätze zur Verfügung zu stellen.

Stephan Büttner (FH Potsdam) betrachtete im ersten Vortrag des Tages Bibliotheken und Bibliothekar/inn/e/n im Forschungsdatenmanagement aus informationswissenschaftlicher Perspektive. Er legte dar, dass sich Forschungsdaten allgemein als digitale und elektronisch speicherbare Daten definieren lassen, die im Zuge eines wissenschaftlichen Vorhabens zum Beispiel durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen. Der Begriff Forschungsdatenmanagement fasse alle Aktivitäten zusammen, „die mit der Aufbereitung, Speicherung, Archivierung und Veröffentlichung von Forschungsdaten verbunden sind“.⁴ Über die Grundlagen des Themas informiert auch das kollaborativ verfasste *Lehrbuch Forschungsdatenmanagement*.⁵ Der Referent näherte sich mittels eines historischen Exkurses der Rolle von Bibliotheken und Informationseinrichtungen im Forschungsdatenmanagement an. Insbesondere in den USA wurde bereits seit den 1970er Jahren immer wieder über das Berufsbild eines Data Librarian diskutiert. 2008 erstellte das Digital Curation Centre schließlich eine Übersicht der *Core Skills for Data Management*.⁶ In der Praxis würden innovative Konzepte, wie das des Embedded Librarian, in Forschungsgruppen gegenwärtig jedoch noch relativ selten umgesetzt. Eine zentrale Frage sei zudem, inwieweit Wissenschaftlerinnen und Wissenschaftler bereit seien, ihre Forschungsdaten zu teilen. Es sei zu beobachten, dass diese Bereitschaft in der Regel von der Karrierestufe des Forschenden abhängt. Anreize zur Publikation der Forschungsdaten gingen dabei oftmals von den Geldgebern aus. Hier wären etwa die Vorgaben des Horizon-2020-Programms der Europäischen Union zu nennen.⁷ Die Anerkennung der Publikation von referenzierbaren Forschungsdaten (mit guten Metadaten) als eine relevante wissenschaftliche Leistung, könne eine weitere Motivation sein.

Christiane Martin (Universitätsbibliothek Bern) betrachtete ebenfalls wissenschaftliche Bibliotheken als Akteure im Forschungsdatenmanagement mit einem Fokus auf der Situation in der Schweiz.⁸ Wie Büttner ist sie überzeugt, dass diese für die meisten Phasen des Lebenszyklus der Forschungsdaten Dienstleistungen anbieten können: beginnend bei der Planung und der Projektphase über die

- 2 Das 1931 fertiggestellte Gebäude diene zunächst als Zentralverwaltung der I.G. Farben. Nach dem Zweiten Weltkrieg hatte dort die US-Armee ihre Europazentrale. Seit 2001 wird das Gebäude von der Goethe-Universität Frankfurt genutzt.
- 3 SAFE-Data-Center-Website, zuletzt geprüft am 01.11.2016, <http://safe-frankfurt.de/de/forschung/data-center.html>.
- 4 Elena Simukovic, Maxi Kindling und Peter Schirnbacher, *Forschungsdaten an der Humboldt-Universität zu Berlin (Berlin: 2013)*, 6, zuletzt geprüft am 01.11.2016, <http://nbn-resolving.de/urn:nbn:de:kobv:11-100213001>.
- 5 Stephan Büttner et al., *Lehrbuch Forschungsdatenmanagement (2016)*, zuletzt geprüft am 01.11.2016, https://handbuch.tib.eu/w/Lehrbuch_Forschungsdatenmanagement.
- 6 „Core Skills for Data Management“ (2008), Digital Curation Centre, zuletzt geprüft am 01.11.2016, <http://data-forum.blogspot.com/2008/12/rdmf2-core-skills-diagram.html>.
- 7 European Commission. Directorate-General for Research & Innovation, *H2020 Programme. Guidelines on FAIR Data Management in Horizon 2020 (2016)*, zuletzt geprüft am 01.11.2016, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- 8 Siehe dazu auch: Christiane Martin, „Wissenschaftliche Bibliotheken als Akteure im Forschungsdatenmanagement“, *LIBREAS. Library Ideas* 23 (2013), zuletzt geprüft am 01.11.2016, <http://libreas.eu/ausgabe23/03martin>.

Aufbereitung und Speicherung der Daten bis hin zur Veröffentlichung sowie der Sicherstellung von Auffindbarkeit und Nachnutzbarkeit. Martin sprach sich dafür aus, auch die Forschungsdaten von Projekten, die ihre Zielsetzung verfehlt haben, zur Nachnutzung zu veröffentlichen. Wie ihr Vorredner unterstrich sie die Notwendigkeit, verstärkt Anreize für die Publikation von Forschungsdaten zu schaffen. In der Schweiz fehlten dazu noch Richtlinien vom Schweizerischen Nationalfond (SNF). Eine Top-Down-Unterstützung sei jedoch zur Etablierung eines Forschungsdatenmanagements notwendig. Eine Lösung auf nationaler Ebene strebt zurzeit das Projekt Data Life-Cycle Management (DLCM) der Rektorenkonferenz der schweizerischen Hochschulen an.⁹

Thorsten Meyer (ZBW – Leibniz-Informationszentrum Wirtschaft) stellte die Projekte der Deutschen Zentralbibliothek für Wirtschaftswissenschaften (ZBW) zum Forschungsdatenmanagement vor. Seit 2010/11 ist das Thema ein wichtiger Arbeitsbereich der ZBW, die anstrebt, in direkter Zusammenarbeit mit der wirtschaftswissenschaftlichen Forschung zu agieren und sich bei infrastrukturellen Entwicklungen am Bedarf der Community zu orientieren. Im Rahmen ihrer Projekte arbeitet die ZBW mit unterschiedlichen Institutionen und Organisationen zusammen. Das Hauptaugenmerk ist dabei weniger auf das Hosting gerichtet, als auf die Unterstützung beim Auffinden, der Zitierfähigkeit und der Zugänglichkeit von wirtschaftswissenschaftlichen Forschungsdaten. In Zusammenarbeit mit GESIS betreibt die ZBW *dajra*, eine DOI-Registrierungsagentur für Sozial- und Wirtschaftsdaten.¹⁰ Gemeinsam mit dem Rat für Sozial- und Wirtschaftsdaten (RatSWD) publizierten beide den Wegweiser für Forschungsdaten *Auffinden, Zitieren, Dokumentieren*.¹¹ Das 2016 gestartete Projekt GeRDI (Generic Research Data Infrastructure) zielt darauf ab, eine national vernetzte Forschungsdateninfrastruktur aufzubauen.¹² Existierende und zukünftige Forschungsdaten sollen virtuell verknüpft werden, um eine disziplinübergreifende Nachnutzung zu erleichtern. Im Fokus stehen vor allem Daten aus universitären Projekten und von Forschungsinstituten („long tail“), für die zumeist keine Infrastruktur zur Speicherung, Archivierung und Dokumentation vorhanden ist.

Ralf Toepfer (ZBW – Leibniz-Informationszentrum Wirtschaft) zeigte, wie das Projekt SoWiDataNet (2014–2017) im Austausch mit Wissenschaftlerinnen und Wissenschaftlern aus den Sozial- und Wirtschaftswissenschaften entwickelt wird.¹³ In Zusammenarbeit der Projektpartner ZBW, GESIS, Wissenschaftszentrum Berlin für Sozialforschung (WZB) und Deutsches Institut für Wirtschaftsforschung (DIW Berlin) entsteht eine eng an den Bedürfnissen der Scientific Community orientierte Infrastruktur für die eigenständige Archivierung und Distribution von quantitativen Forschungsdaten. Zielgruppe sind Forschungseinrichtungen ohne eigene Forschungsdateninfrastruktur. Toepfer ging auf drei Spannungsfelder im Forschungsdatenmanagement ein und erläuterte die im Projekt vorgesehenen Lösungsansätze. So sollen die Daten grundsätzlich frei zugänglich sein. Es besteht jedoch die Möglichkeit, in SoWiDataNet für bestimmte Zeiträume Zugriffsbeschränkungen einzustellen.

9 DLCM-Website, zuletzt geprüft am 01.11.2016, <http://www.dlcm.ch>.

10 *dajra*-Website, zuletzt geprüft am 01.11.2016, <http://www.da-ra.de>.

11 ZBW, GESIS und RatSWD (Hrsg.), *Auffinden, Zitieren, Dokumentieren: Forschungsdaten in den Sozial- und Wirtschaftswissenschaften* (2014), zuletzt geprüft am 01.11.2016, doi: 10.4232/10.fsuzida2015.2.

12 Projektpartner sind das Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften, die Christian-Albrechts-Universität zu Kiel, das Deutsche Forschungsnetz sowie die Technische Universität Dresden.

13 SoWiDataNet-Website, zuletzt geprüft am 01.11.2016, <http://sowidatanet.de>.

Um eine qualitativ hochwertige Dokumentation der Daten bei geringem Arbeitsaufwand für die Forschenden zu erreichen, wird es Kuratorinnen und Kuratoren geben, die auch bei Fragen der Anonymisierung und des Lizenzrechts beraten. Die Forschenden werden zudem durch die Einbindung von Normdaten und kontrollierten Vokabularen (STW, TheSoz und GND) unterstützt. Die Beta-Version von SoWiDataNet wird bis März 2017 fertiggestellt. Im Laufe des Jahres soll es dann mit datorium¹⁴ zu einem Repositorium für Forschungsdaten fusionieren.

Über den Aufbau eines Forschungsdatenarchivs für Fachzeitschriften referierte Sven Vlaeminck (ZBW – Leibniz-Informationszentrum Wirtschaft). Ebenso wie SoWiDataNet entstand das Projekt European Data Watch Extended (EDaWaX) (2011–2016) in Zusammenarbeit mit der wirtschaftswissenschaftlichen Community.¹⁵ Auch Vlaeminck berichtete von den fehlenden Anreizen zum Data Sharing (Zeitaufwand vs. fehlende Anerkennung). Im Rahmen des Projekts wurden unter anderem die Data Policies von wirtschaftswissenschaftlichen Fachzeitschriften untersucht. Die Studie zeigte, dass von 346 Publikationen im Sample 79,5 Prozent über keine entsprechende Policy verfügten (VWL: 68,2 Prozent, BWL: 87,5 Prozent). Die Anforderungen der 71 Zeitschriften mit Data Policies waren zudem sehr unterschiedlich.¹⁶ Darüber hinaus setzte nur ein Teil der Fachzeitschriften seine Data Policies auch wirklich konsequent um.¹⁷ Ein weiteres Problem ist die Zugänglichkeit der Forschungsdaten. Vlaeminck demonstrierte am Beispiel der *American Economic Review*, dass diese oftmals nur in Form von Zip-Dateien verfügbar sind. Solche Datensätze sind jedoch einerseits schwer zu zitieren, andererseits bleiben Lizenzfragen offen. Um die Bereitstellung der Daten zu verbessern, wurden im EDaWaX-Projekt ein geeignetes Metadatenschema und eine Softwarelösung konzipiert. In die Pilotapplikation wurden sukzessive weitere Funktionen integriert (Normdaten und semi-automatisierte DOI-Vergabe). Zurzeit laufen die Vorbereitungen für die Umwandlung des Projekts in einen Service der ZBW.

Den zweiten Tag eröffnete Uwe Risch (HeBIS-Verbundzentrale) mit seinem Vortrag über das Mitte 2015 gestartete Projekt GOETHE (Germany's Old Equity Timeseries for Historical Evaluation).¹⁸ Forschenden soll damit eine Datenbank für die Analyse von historischen Aktienkursen an deutschen Regionalbörsen aus der Zeit von 1914 bis 1939 zur Verfügung gestellt werden. Aus zeitgenössischen Zeitungen, Kursblättern und weiteren Veröffentlichungen wurden dazu die Daten für die Börsen in Berlin, Breslau und Frankfurt am Main manuell extrahiert. In einem zweiten Schritt kamen Verknüpfungen von Körperschaften und Personen mit der GND hinzu. Zurzeit sind etwa 1,5 Millionen

14 datorium-Website, zuletzt geprüft am 01.11.2016, <http://datorium.gesis.org>.

15 EDaWaX-Website, zuletzt geprüft am 01.11.2016, <http://www.edawax.de>. Projektpartner sind der Rat für Sozial- und Wirtschaftsdaten, das Forschungsdatenzentrum des Sozio-ökonomischen Panels am DIW Berlin und das Max-Planck-Institut für Innovation und Wettbewerb.

16 49 Zeitschriften hatten eine Data Availability Policy, 30 eine verpflichtende Policy und 25 eine Policy, die auch den Programmcode einfordert. 20 Zeitschriften verfügten über eine Policy mit Regeln für die Verwendung proprietärer Daten.

17 Sven Vlaeminck und Lisa-Kirstin Herrmann, "Data policies and data archives: A new paradigm for academic publishing in economic sciences?" in *New Avenues for Electronic Publishing in the Age of Infinite Collections and Citizen Science: Scale, Openness and Trust (Proceedings of the 19th International Conference on Electronic Publishing)* (Amsterdam: IOS Press, 2015): 145–155, hrsg. von Birgit Schmidt und Milena Dobрева, zuletzt geprüft am 01.11.2016, doi:10.3233/978-1-61499-562-3-145.

18 Projektpartner sind SAFE Data Center, Universitätsbibliothek Johann Christian Senckenberg und HeBIS.

tagesgenaue Aktienkurse aus den Jahren 1925–1930 verfügbar sowie 700.000 monatliche Kurse für den Zeitraum 1871–1914. Die Unternehmensdatenbank umfasst mehr als 1.200 Einträge. Risch stellte zudem das Projekt Forschungsdateninfrastruktur für Finanzdaten (FiF) vor, dessen Ziel eine verbesserte Aufbereitung und Zugänglichmachung von finanzwirtschaftlichen Forschungsdaten ist. In dem FiF-Repositorium sind auch die GOETHE-Daten erfasst. Organisatorisch ist eine Arbeitsteilung vorgesehen zwischen Forschungsdatenmanager/inne/n, Forschungsdatenbibliothekar/inn/en und Repositoriumsbetreuer/inne/n. Aufgaben für die Bibliothekarinnen und Bibliothekare wären etwa die Festlegung von Erschließungsstandards, die Langzeitarchivierung und Fragen des Datenschutzes.

Wie Forschungsdatenmanagement in der Praxis funktioniert, zeigte Jessika Rücknagel (SUB Göttingen) an den Beispielen des interdisziplinären Forschungsprojekts EFForTS sowie der Göttingen eResearch Alliance. Im Projekt EFForTS erforschen Wissenschaftlerinnen und Wissenschaftler der Universität Göttingen gemeinsam mit Partnern aus Indonesien ökologische und sozioökonomische Aspekte des Tieflandregenwalds auf Sumatra.¹⁹ Bei ihren Feldstudien begleiten sie Embedded Data Manager.²⁰ Rücknagel berichtete, dass die Datenspezialisten und die Forschenden dabei zunächst eine gemeinsame Sprache finden mussten. Die Data Manager vermitteln den Wissenschaftlerinnen und Wissenschaftlern Kenntnisse im Forschungsdatenmanagement und unterstützen sie in der Nutzung des Informationssystems EFForTS-IS (BExIS) sowie bei der Datenaufbereitung.²¹ Ebenso haben sie bei der Erstellung einer Data Policy mitgearbeitet. Im Anschluss an EFForTS stellte Rücknagel mit der eResearch Alliance (eRA) ein Projekt der Universität Göttingen vor, das sich seit 2014 zum Ziel setzt, Forschungsinfrastrukturangebote für die 13 Einrichtungen des Göttingen Campus auszubauen.²² Aktuell wird zum Beispiel ein Template zur Erstellung von Data-Management-Plänen entwickelt. Weitere Services, wie ein Großgeräteportal, ein DOI-Service und ein Forschungsdaten-Repositorium, sind geplant.

Den Abschluss der Fachreferententagung bildeten von Natascha Schumann (GESIS) moderierte Lightning Talks zum Thema Forschungsdatenmanagement. Stefan Bender (Forschungsdaten- und Servicezentrum, Deutsche Bundesbank) griff das wiederholt angesprochene Problem einer mangelnden Anreizstruktur für die Dokumentation und Bereitstellung von Forschungsdaten auf. In diesem Zusammenhang stellte er die Frage, ob Metadaten-Standards, wie der der Data Documentation Initiative (DDI), zu kompliziert seien. Für Bender stellt das Forschungsdatenmanagement eine schwer realisier- und finanzierbare Querschnittsaufgabe dar. Ohne diese werde sich die Forschung jedoch – zugespitzt formuliert – abschaffen, während private Unternehmen das Datenmanagement übernehmen. „Data is the new oil“ – mit diesem Bild einer wertvollen schwarzen Masse, aus der die nützlichen Daten aber erst extrahiert werden müssen, beschrieb Bender sowohl die Bedeutung

19 EFForTS-Website, zuletzt geprüft am 01.11.2016, <https://www.uni-goettingen.de/de/310995.html>. EFForTS: Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems.

20 Fabian Cramer, Claudia Engelhardt und Heike Neuroth, „Embedded Data Manager – Integriertes Forschungsdatenmanagement: Praxis, Perspektiven und Potentiale“, *Bibliothek Forschung und Praxis* 39, Nr. 1 (2015): 13–31, zuletzt geprüft am 01.11.2016, doi:10.1515/bfp-2015-0006.

21 BExIS-Website, zuletzt geprüft am 01.11.2016, <http://www.biodiversity-exploratories.de>. BExIS: Biodiversity Exploratories.

22 Website der eResearch Alliance, zuletzt geprüft am 01.11.2016, <http://www.eresearch.uni-goettingen.de>. Website des Göttingen Campus, zuletzt geprüft am 01.11.2016, <http://grc.uni-goettingen.de>.

des Forschungsdatenmanagements als auch die damit verbundenen Herausforderungen. Zu deren Bewältigung seien ein verbesserter Informationsaustausch und die Gründung von Kompetenzzentren ein guter Weg. Stephan Büttner rief die Wissenschaft ebenfalls dazu auf, die Entwicklungen im Bereich Forschungsdatenmanagement aktiv mitzugestalten. Wolfgang König (Goethe-Universität, Frankfurt) ging auf den Wandel des Wirtschaftssystems hin zu einer Dienstleistungsgesellschaft ein, in der die Datenproduktion und -verwertung zentrale Faktoren seien. Das Datenmanagement müsse daher in der Europäischen Union und in Deutschland stärker gefördert werden. Uwe Risch betonte die Bedeutung von Einrichtungen wie dem Rat für Informationsinfrastrukturen (RfII) und dem Rat für Sozial- und Wirtschaftsdaten für den Austausch und die Koordinierung lokaler und regionaler Initiativen. Thorsten Meyer hob die Bedeutung der Zusammenarbeit und der Kommunikation von Bibliotheken mit den Datenproduzenten hervor. Die Stärken der Bibliotheken sieht er in der Normierung und im Metadatenmanagement.

In der Abschlussdiskussion stand die Frage im Mittelpunkt, was die Entwicklungen im Forschungsdatenmanagement für die Arbeit von Fachreferentinnen und Fachreferenten bedeuten. Viele Projekte im Bereich Wirtschaftswissenschaften befinden sich noch im Aufbau. Vor diesem Hintergrund ermutigten die Forschungsdatenmanagement-Expertinnen und -Experten die Fachreferentinnen und Fachreferenten, in ihren jeweiligen Institutionen an der Entwicklung von Forschungsdaten-Policies mitzuarbeiten oder solche zu initiieren. Die Tagung in Frankfurt hat gezeigt, dass es sich beim Forschungsdatenmanagement um ein zukunftsweisendes Feld für Bibliotheken handelt, für das weitere Kompetenzen aufgebaut werden müssen.

Thomas Gerdes, ZBW – Leibniz-Informationszentrum Wirtschaft, Kiel/Hamburg

Zitierfähiger Link (DOI): <http://dx.doi.org/10.5282/o-bib/2016H4S316-321>