

Anja Müller, Christian Winterhalter

Tagungsbericht –Open Science Conference, 13.-14.3.2018, Berlin

Vom 13.–14.3.2018 fand die International Open Science Conference, nun bereits zum fünften Mal, im Herzen von Berlin in der Friedrichstraße statt. Die Aktualität des Themas und die internationale Bedeutung drückte sich nicht allein durch die Internationalität der Teilnehmer (aus 35 Ländern), sondern auch dadurch aus, dass die Konferenz innerhalb kürzester Zeit ausgebucht war, noch bevor ein stärkeres Bewerben der Konferenzanmeldung stattgefunden hatte. Organisiert wurde die Konferenz erneut von der ZBW – Leibniz-Informationszentrum Wirtschaft und dem Leibniz-Forschungsverbund Science 2.0 unter der Regie von Klaus Tochtermann. Die Konferenz bot eine ausgewogene Mischung aus überblickgebenden und praxisorientierten Präsentationen, interessanten Podiumsdiskussionen und informativen Poster Talks. Den diesjährigen Schwerpunkt bildeten Präsentationen und Diskussionen zur European Open Science Cloud (EOSC) sowie zur Bedeutung der FAIR-Prinzipien im Umgang mit Forschungsdaten.

Eröffnet wurde die Konferenz durch drei *Opening Talks* von Georg Schütte, Staatssekretär im Bundesministerium für Bildung und Forschung (BMBF), Jean-Claude Burgelman, Leiter der Abteilung Open Data Policy and Science Cloud, DG Research and Innovation, Europäische Kommission sowie Klaus Tochtermann, Direktor der ZBW – Leibniz-Informationszentrum Wirtschaft.

Als Gastgeber begrüßte Tochtermann das Auditorium und erläuterte – ausgehend vom Feedback des letzten Jahres – den in diesem Jahr stärkeren Fokus auf interaktiveren Formaten wie Podiumsdiskussionen und Poster-Talks. Er akzentuierte dabei insbesondere drei inhaltliche Aspekte: die Frage nach der Motivation für Wissenschaftler, im Sinne von Open Science zu agieren; die Reflexion über Open Science; und das Vorhaben der Konferenz, die vielen Akronyme rund um das Forschungsdatenmanagement und Open Science aufzulösen.

Jean-Claude Burgelman ging in seiner Rede auf wichtige Open-Science- Meilensteine auf EU-Ebene im Verlauf des Jahres 2018 ein. Er zeigte kürzlich Erreichtes, wie die überarbeitete *Recommendation on Access to and Preservation of Scientific Information*, auf und ging weiter auf wichtige Termine und Dokumente zur Entwicklung der EOSC ein, so u.a. den voraussichtlichen Launch der EOSC Governance-Strukturen im November und die Veröffentlichung der Empfehlungen der High Level Expert Group on the European Open Science Cloud im Dezember. In Bezug auf FAIR wies er u.a. auf die Veröffentlichung des „Final proposal FAIR Data Action Plan“ durch die FAIR data Expert Group im

September hin. Insgesamt bot der Vortrag einen guten Überblick über die Open-Data-Aktivitäten auf Ebene der EU.

Die Perspektive der nationalen Forschungsförderorganisation zeigte Georg Schütte auf und ging anhand der Leitfragen: *Are we strategic enough? Are we fast enough? Are we relevant enough?* auf den Stand und Perspektiven der Open-Science-Aktivitäten in Deutschland ein. Hinsichtlich der strategischen Fragen zeichnete er ein differenziertes Bild, das die Bemühungen in den Kontext durchaus gegenläufiger Tendenzen der Globalisierung verortete (Öffnung und Teilen auf der einen vs. Abschottung und Monopolisierung auf der anderen Seite). Betont wurde die Wichtigkeit, Anreize in den wissenschaftlichen Communities zur Nutzung entsprechender Infrastrukturen zu schaffen, die als Top-Down-Prozess nicht funktionieren können. Hervorgehoben wurde auch, eine Beschleunigung der Prozesse auf nationaler Ebene erreichen zu müssen, sowie der Aspekt, die Relevanz der Wissenschaft für die Gesellschaft in einer Perspektive zu betrachten, die über „Citizen Science“ hinausgehen sollte. Die EOSC sei, so Schütte, dabei als wichtiger Teil dieser Bemühungen anzusehen.

Nach den Eröffnungsvorträgen war eine Klärung des Zusammenhangs zwischen FAIR-Prinzipien und Open Science sowie die Vorstellung der GO FAIR-Initiative das Anliegen des Vortrags *GO FAIR helps Open Science go far* von Karl Luybens (National Coordinator for Open Science, Niederlande). So skizzierte er, wie sich die GO FAIR-Initiative zum einen an der Vision der EOSC orientiert, zum anderen aber durch ihren globalen und spartenübergreifenden Ansatz zugleich darüber hinausgeht. Mit der Schaffung eines gemeinsamen Rahmens für den Zusammenschluss von Implementierungsnetzwerken mit Fokussierung auf die Bereiche Policy, Training oder Technik zielt die Initiative darauf ab, ein Internet of Fair Data and Services zu entwickeln. Mit der gemeinsamen Nutzung von Daten und Services innerhalb des Referenzrahmens der FAIR-Prinzipien soll zugleich ein wesentlicher Grundstein für die Realisierung der EOSC gelegt werden.

Sarah Jones (Associate Director Digital Curation Centre, University of Glasgow, Großbritannien) führte, durch umfangreiches Datenmaterial belegt, unter dem Titel *Open data, FAIR data & RDM: the ugly duckling* aus, wie sich der Umgang der Wissenschaftler mit Daten gestaltet (Anteil derjenigen, die Daten publizieren), wie Daten geteilt werden (zumeist auf Nachfrage als E-Mail-Anhang) und welches die primäre Motivation für das Teilen von Daten ist (Sichtbarkeit). Gut strukturiert legte sie zudem dar, wie sich der Diskurs im Umgang mit Daten über die Jahre in Policies und Positionspapieren verändert hat (hin zu „open“ und FAIR) und welche Etappen dabei als Vorläufer der FAIR-Prinzipien gelten können. Terminologiekritisch ging sie jedoch auch auf die Probleme ein, die sich durch die Vielzahl von Begriffen, Akronymen und technischem Vokabular ergeben haben und die zu Verwirrung bei der Zielgruppe der Wissenschaftler führen. Folgerichtig mahnte sie ein größeres Bewusstsein bei der Formulierung von Begrifflichkeiten im Bereich des

Forschungsdatenmanagements an, grenzte sehr klar die Ebenen voneinander ab (Forschungsdaten, Forschungsdaten nach FAIR-Prinzipien, offene Daten) und legte dar, wie auf der lokalen Ebene die Wissenschaftler in den Dialog zu Forschungsdaten eingebunden werden können.

Um eine Klärung der Ebenen von Open-Science-Strategien ging es auch im Vortrag von Wolfram Horstmann (Direktor der Staats- und Universitätsbibliothek Göttingen): *Reflections on Open Science strategies at local, national and European level*. So legte er systematisch die europäische, nationale und lokale Ebene mit ihrer jeweiligen Zielstellung dar (was auch für ein Publikum mit Kenntnis der Wissenschaftsorganisationen in Deutschland gleichwohl an vielen Stellen erhellend war) und skizzierte die Verbundenheit aber auch Unterschiede der vielen Teilbereiche von Open Science. Er plädierte dafür, Open Science nicht als Selbstzweck zu begreifen oder als reine Infrastrukturmaßnahme, sondern als Kulturwandel, der letztlich das aktive Handeln des einzelnen Wissenschaftlers erfordere. Gerade um die „letzte Meile“ hin zur konkreten Praxis der Wissenschaftler zu überbrücken, sei die Bedeutung der lokalen Ebene nicht zu unterschätzen und stärker in den Blick zu nehmen, d.h. die Unterstützung der Wissenschaftler auf lokaler Ebene durch ihre Universitäten, Bibliotheken und Rechenzentren.

„Open Science as a Service“ – dieses Fazit der Schlussfolie umfasst treffend den Ansatz der Aktivitäten des OpenAIRE-Netzwerks, den Natalia Manola (Department of Informatics and Telecommunications, University of Athens & Athena Technology and Innovation Research Center, Griechenland) in ihrer Präsentation *Mind the gap: Open Science in transit* referierte. Wie ihre Vorredner legte sie den Fokus dabei insbesondere auf die Verschränkung eines pragmatischen und partizipativen Ansatzes und einfach zu nutzender Tools. Bezogen auf Forschungsdatenmanagement lag der Fokus ihrer Präsentation insbesondere auf den Aktivitäten im Bereich des Trainings, der Service-Tools für das Forschungsdatenmanagement wie Amnesia zur Anonymisierung personenbezogener Daten oder OpenDMP sowie der noch im Aufbau befindlichen Unterstützung im Bereich Lizenzen und Rechtsfragen.

Die Deutsche Bundesbank gehört zu den größten Datenproduzenten in Deutschland und generiert hochwertige Daten, deren Nutzung über die Jahre stark angewachsen ist. Stefan Bender (Deutsche Bundesbank und RatSWD) führte unter dem Titel *The Role of the Bundesbank Microdata Production in Times of Big Data: The need for Data access and Data Sharing* in die Servicekonzeption des Research Data and Service Centre der Bundesbank ein, das auch externen Nutzern für Forschungszwecke Zugang zu Daten ermöglicht und zugleich Qualitätskontrolle und Sicherheitsprozeduren umfasst. Weiterhin skizzierte er die Aktivitäten der Bundesbank in Hinblick auf den Datenaustausch zwischen den Nationalbanken und die Initiativen auf europäischer und internationaler Ebene, welche in der Nachfolge der Bankenkrise darauf abzielten, Datenlücken zu

schließen sowie Datenqualität und Interoperabilität zu verbessern. Die Aktivitäten des RatSWD bewertete Bender abschließend als ein Best-Practice-Modell, um die Aktivitäten im Bereich Zertifizierung und Vernetzung voranzubringen.

Die Diskrepanz zwischen öffentlicher Verlautbarung und gelebter Praxis der Offenheit war das Thema von Felix Schönbrodts Ausführungen (*Changing incentive structures to foster the actual sharing rate of open data*). Wenngleich die Rhetorik des Teilens und der Offenheit inzwischen fester Bestandteil von (Selbst-)Verpflichtungen wissenschaftlicher Einrichtungen und Förderer sind, so konstatierte Schönbrodt (Ludwig-Maximilians-Universität München) vielfach Ablehnung und Zögern bei Wissenschaftlern in Hinblick auf mehr Offenheit. So werde praktizierte Offenheit nach wie vor als zeitaufwändig wahrgenommen (nicht geeignete Tools, Datenkuratierung, mangelnde Unterstützung seitens der Einrichtung) und derzeit noch nicht als wesentlicher Faktor für die individuelle akademische Laufbahn gesehen. Sehr klar führte Schönbrodt die verschiedenen Ebenen aus, auf denen die Forderung nach einem Kulturwandel praktische Implikationen haben muss, um Open Science effektiv zu befördern. Dies reicht von der Anpassung von Stellenprofilen in der Forschung, Änderung der Kriterien für Berufungsverfahren, besserer Berücksichtigung von (offenen) Datenpublikationen bei der Beantragung von Fördermitteln bis zu ausgeprägten Serviceangeboten auf der lokalen Ebene (Data Stewards, nutzerfreundliche Tools etc.). Er endete damit, dass bei den Aktivitäten letztlich auch die Frage und der Widerspruch beachtet werden muss, ob der Fokus eher auf einer schnelleren offenen Verfügbarmachung ohne größere Hürden liegt oder dem stärkeren Fokus auf hoher Qualität nach den FAIR-Prinzipien.

Der zweite Konferenztag wurde von Matthias Kleiner, Präsident der Leibniz Gemeinschaft, mit einem Opening Talk zum Thema *Science in Openness* eingeleitet. Die Leibniz Gemeinschaft gehört zu den Erstunterzeichnern der Berliner Erklärung und kann damit auf eine lange Tradition im Bereich Openness zurückblicken. Wie alle Teile der Gesellschaft befindet sich auch die Leibniz Gemeinschaft in einem Transformationsprozess, der durch Digitalisierung und Ablösung traditioneller Methoden geprägt ist. Darauf reagiert sie mit neuen Methoden und Serviceentwicklung wie dem Projekt der Plattform „Bürger schaffen Wissen“, an dessen Beispiel die führende Rolle der Leibniz Gemeinschaft im Bereich Citizen Science illustriert wurde. Unter Verweis auf diese Aktivitäten und die insgesamt 50 zur Leibniz Gemeinschaft gehörenden Dateninfrastrukturen (Sozioökonomisches Panel, Leibniz-Zentrum für Agrarlandschaftsforschung etc.) kontextualisierte Kleiner die Aktivitäten und Strukturen der Leibniz Gemeinschaft im Rahmen der nationalen Bemühungen zum Ausbau von Forschungsinfrastrukturen. Dabei ließ er nicht unerwähnt, dass seiner Ansicht nach Open Science als Mittel zur Förderung guter wissenschaftlicher Praxis und auch der Reformierung der Bewertung wissenschaftlicher Leistungen dient und dienen sollte (mehr Qualität statt Quantität).

Simon Hodson (Executive Director CODATA, Frankreich, und Chair der EC Expert Group on FAIR Data) berichtete aus Sicht der ältesten internationalen Datenorganisation. Zu Beginn seines Vortrages *Making FAIR data a reality ... and the challenges of interoperability and reusability* stellte er noch einmal die Frage, warum wir eine offene Wissenschaft und FAIR Data benötigen. Er erläuterte weiter die verschiedenen Arbeitsfelder von CODATA wie Policy-Entwicklung, Standardisierung, Curriculumentwicklung und Training und berichtete über die Arbeit und Ziele der Expert Group on FAIR Data der europäischen Kommission, die derzeit u.a. Vorschläge für den FAIR Data Action Plan erarbeitet, der im Herbst 2018 veröffentlicht werden soll. Interessant war weiter die vorgestellte Initiative für Dateninteroperabilität und -integration, die in global wichtigen Feldern wie Infektionskrankheiten und Naturkatastrophenforschung Pilotprojekte entwickeln will.

Auch Juan Bicarregui (Scientific Computing Department, Science and Technology Facilities Council (STFC), Großbritannien) stellte an den Anfang seines Vortrages *The European Open Science Cloud: moving from policy to practice* internationale bedeutende Policies, die die Basis und die Legitimation der Entwicklung der EOSC bilden und die Zielsetzung bestimmten. Im Verlauf seines Vortrags gab er einen Überblick über die Ziele von EOSC, die Projektstruktur des EOSC-Piloten sowie die geplanten Arbeitspakete und den Zeitplan. Weiter stellte er die Herausforderungen des EOSC-Projektes heraus und machte deutlich, dass es zunächst um die Pilotentwicklung geht – „a pilot is a pilot“ – und dies den Beginn der Entwicklung der EOSC darstellt.

Robert Jones (Information Technology Department, CERN, Schweiz) beleuchtete das Thema Open Science aus Sicht einer der datenreichsten europäischen Forschungsorganisation, der European Organisation for Nuclear Research (CERN). Er zeigte in seinem Vortrag *Preparing for High-Luminosity LHC* die Herausforderungen auf, vor denen Organisationen stehen, wenn sie Infrastrukturen für Big Data Sciences aufbauen, und stellte Lösungen des CERN, wie das Worldwide LHC Grid, das LHC Data Portal und das Workflowportal REANA (Reusable analyses) vor. Insbesondere die Vorstellung des europäischen Projektes Helix Nebula Science Cloud machten die gigantischen Dimensionen der zu bewältigenden Technologieentwicklung deutlich. Der Abschluss des Vortrags gestaltete sich interaktiv und bestand in einer kleinen Umfrage unter den Konferenzteilnehmer zum zukünftigen Business-Modell von Zenodo, ein Dienst der bisher durch das CERN als "Nebenaktivität" getragen und entwickelt wird. Die überwiegende Mehrzahl der Teilnehmer sprach sich für ein Modell der kostenlosen Nutzung mit Beschränkung der Datenmenge aus.

Eine ganz andere und anregende Perspektive bot der Vortrag *Translational failure in preclinical research; how open science can help* von Emily Sena (Centre for Clinical Brain Sciences, University of Edinburgh, Großbritannien) zum Thema Translationale Medizin. Sie beleuchtete in ihren Ausführungen die negativen Auswirkungen einer Reihe gängiger Publikationsweisen in der

präklinischen Forschung wie z.B. die Nichtveröffentlichung negativer oder neutraler Ergebnisse, die zu Fehleinschätzungen und Verzerrungen in Metaanalysen führen können. Gleichzeitig zeigte sie in ihrem Vortrag jedoch auch auf, wie die Nutzung der von ihr vorgestellten Open Science Tools, Registries und Guidelines Wissenschaftler bei der Entwicklung offener Workflows unterstützen können, um die präklinische Forschung zu verbessern.

Zum Thema *FAIR Data – just an nice idea or a new paradigm shift?* diskutierten auf dem Podium Mercé Crosas (Chief Data Science and Technology Officer, Harvard University, USA), Natalia Manola (Department of Informatics and Telecommunications, University of Athens & Athena Technology and Innovation Research Center, Griechenland), Hans Pfeiffenberger (Alfred-Wegener-Institut. Helmholtz-Zentrum für Polar- und Meeresforschung) und Marc Rittberger (Deutsches Institut für Pädagogische Forschung). Ganz im Sinne der Diskussionen während der Konferenz stand auch in der Podiumsdiskussion erneut das Verhältnis von „Open Data“ und „FAIR Data“ im Mittelpunkt, d.h. ob der Fokus in der Kommunikation mit Wissenschaftlern eher auf dem Open- oder dem FAIR-Aspekt der Daten liegen sollte, ob die Aufgabe der Datenpublikation und der damit einhergehenden Arbeit bei den Wissenschaftlern oder den Infrastruktureinrichtungen liegen sollte und welche Anreizmechanismen für Datenpublikationen sinnvoll erscheinen (Badges etc.). Wenig überraschend war, dass – selbst wenn ein Ideal von Daten, die offen und FAIR sind, unstrittig erscheint – die Bedarfe und Kontexte der jeweiligen Wissenschafts-Communities sehr unterschiedlich ausfallen. Während im Kontext von Pangea bereits eine Umsetzung der FAIR-Prinzipien als erreicht gelten kann und für die Wissenschaftler die Unterscheidung zwischen FAIR und open weniger relevant ist, ging es in anderen in der Diskussion skizzierten Kontexten um noch wesentlich grundlegendere Erfordernisse. Zum Beispiel darum, einzelne Datenproduzenten wie Kultusministerien überhaupt davon zu überzeugen, dass es – auch unter Einhaltung von Datenschutzaspekten – andere Optionen gibt als Daten nach deren Verwendung schlicht zu löschen (z.B. Daten aus bildungswissenschaftlichen Studien in Schulen), auch wenn diese – um Missbrauch vorzubeugen – nicht offen sein können. Konsens der Diskutanten war, dass die FAIR-Prinzipien als Grundlage und Klammer sowie als gut vermittelbare Anforderung gelten können, um in den Wissenschafts-Communities geeignete Prozeduren zur Datenspeicherung und -publikation anzuregen. Begleiten sollten die Infrastruktureinrichtungen dies durch passgenaue, einfach zu bedienende Tools und Beratung zu den Standards. Was Offenheit heißt und heißen kann, ist jedoch – wie deutlich wurde – wiederum sehr stark vom Kontext der Datenentstehung und Möglichkeiten der Nachnutzung abhängig.

Neben Vorträgen und Paneldiskussionen stellten sich 10 Projekte auch im Format der Lightning-Talks vor. Die Bandbreite reichte von konkreten Projekten zur Implementierung der FAIR-Prinzipien, Citizen Science-Projekten bis zu Tool- und Supportentwicklung. Präsent waren diese und weitere

Projekte auch in Form von insgesamt knapp 20 Postern, die in den Pausen angeregte Diskussionen sowohl über strategische Fragen rund um Open Science als auch zu konkreten Einzelprojekten beförderten.

Nicht unerwähnt bleiben darf auch das *Open Science Barcamp*, das – inzwischen schon traditionell – am Tag vor der Open Science Conference stattfand und in 20 Sessions Raum zur intensiven Diskussion und dem Austausch praktischer Erfahrungen im Bereich Open Science diente bzw. als Ausgangspunkt für weitere Reflexionen zu den diskutierten Themen.¹ Ganz im Sinne von Open Science sind die einzelnen Sessions von den Teilnehmern in Etherpads dokumentiert, die über das Metapad des Barcamps erreichbar sind.²

Die Konferenz ist durch die Konferenzwebseite³ sehr gut dokumentiert. Alle Präsentationen wurden kurz nach der Konferenz öffentlich zugänglich gemacht, und interessante Interviews, die während der Konferenz entstanden sind, sind über das Open Science Radio verfügbar. Insgesamt gibt die Fotogalerie und die Begleitung über Twitter die inspirierende Atmosphäre der Konferenz wieder, die Klaus Tochtermann nach der Podiumsdiskussion mit einem positiven Fazit schloss, nicht ohne auf die Fortführung der Diskussionen unter dem Motto „Open Science, Open to the World, Open Innovation“ im nächsten Jahr zu verweisen. Die nächste International Open Science Conference wird vom 19. – 20.3.2019 erneut in Berlin stattfinden.



Anja Müller

Humboldt-Universität zu Berlin

Universitätsbibliothek

Unter den Linden 6

D-10099 Berlin

anja.doreen.mueller@ub.hu-berlin.de

¹ Kaden, Ben (2018): Warum Forschungsdaten nicht publiziert werden. In: *LIBREAS. Library Ideas* 33. Verfügbar unter <https://libreas.eu/ausgabe33/kaden-daten>.

² Open Science Barcamp (2018): <http://www.open-science-conference.eu/barcamp/>. Metapad zu den Session-Dokumentationen verfügbar unter: <https://etherpad.wikimedia.org/p/oscibar2018>.

³ Open Science Conference (2018): <http://www.open-science-conference.eu>.



Christian Winterhalter
Humboldt-Universität zu Berlin
Universitätsbibliothek
Unter den Linden 6
D-10099 Berlin
christian.winterhalter@ub.hu-berlin.de