

# **Über den Tellerrand geschaut: „Data Communities“ jenseits von generischen / fachspezifischen Ansätzen**

## **Vorstellung und Diskussion eines Konzeptes**

### **Asef, Esther Marie**

esther.asef[at]fu-berlin.de  
Freie Universität Berlin, Deutschland  
ORCID-iD: 0000-0003-2411-4953

### **Huber, Elisabeth**

ehuber[at]uni-bremen.de  
Universität Bremen, Deutschland  
ORCID-iD: 0000-0003-4084-4315

### **Imeri, Sabine**

sabine.imeri.1[at]ub.hu-berlin.de  
Humboldt-Universität zu Berlin, Deutschland  
ORCID-iD: 0000-0002-8844-4014

### **Ommert, Eva**

eva.ommert[at]uni-leipzig.de  
Universität Leipzig, Deutschland  
ORCID-iD: 0000-0001-9093-3477

### **Rizzolli, Michaela**

michaela.rizzolli[at]fu-berlin.de  
Freie Universität Berlin, Deutschland  
ORCID-iD: 0000-0001-8154-6563

### **Wagner, Cosima**

cosima.wagner[at]fu-berlin.de  
Freie Universität Berlin, Deutschland  
ORCID-iD: 0000-0003-4957-3478

**Zusammenfassung.** Alternativ zu einer vereinfachenden Unterscheidung von generischem und fachspezifischem Forschungsdatenmanagement (FDM) diskutiert der Workshop das Potenzial des Konzepts der „Data Communities“<sup>1</sup> für die geistes- und sozialwissenschaftlichen Disziplinen. Mit Blick auf die STEM-Disziplinen formuliert, werden „Data Communities“ als fluide und mehr informelle Gruppen von Forschenden verstanden, die über Praktiken der Arbeit mit und des Teilens von Daten eines bestimmten Typs miteinander verbunden sind – und nicht unbedingt durch disziplinäre Zuordnung. Entsprechend sollten die Praktiken und Interessen solcher „Data Communities“ der Ausgangspunkt sein, um das Teilen von Daten zu unterstützen und anzuregen. Der Workshop thematisiert angesichts heterogener Methoden und Forschungspraktiken in den Geistes- und Sozialwissenschaften Grenzen, Schnittstellen und übergreifende FDM-Praktiken mit Beiträgen aus unterschiedlichen Praxisfeldern. Dabei interessieren insbesondere drei Ebenen: die potenzielle Reichweite des Konzepts „Data Communities“, etwa hinsichtlich der Definition von Datentypen, die strukturellen und politischen Rahmenbedingungen sowie die Ausdifferenzierung der Archivierungslandschaft. Ziel ist es, Raum für die konzeptionelle Diskussion zu schaffen und zu eruieren, welche – strategischen wie operativen – Kriterien sich daraus für ein forschungsadäquat unterstützendes Datenmanagement ableiten lassen.

## 1 Hintergrund

Forschungsdatenmanagement (FDM) muss neben generischen Lösungen, die nur für einige Aspekte sinnvoll erscheinen, vor allem solche Praktiken und Infrastrukturen anbieten und etablieren, die den spezifischen Besonderheiten einzelner Forschungsbereiche möglichst umfassend gerecht werden – zu diesem Punkt besteht weitgehend Einigkeit. Auch die (bereits bewilligten oder geplanten) NFDI-Konsortien gehen in diese Richtung: Forschungsdatenmanagement ist vor allem spezifisch und richtet sich entlang von Disziplinen, Teildisziplinen oder Gruppen von Disziplinen aus.

Bei näherer Betrachtung ist diese Zuordnung aber keineswegs gleichmäßig oder eindeutig. Vielmehr ist in den Geistes- und Sozialwissenschaften die Ausdifferenzierung innerhalb einzelner Forschungsdisziplinen mitunter größer als es die disziplinäre Ordnung nahelegt – gerade in Wissenschaftsbereichen mit vielen „kleinen Fächern“, breit gefächertem Methodenspektrum und heterogener Forschungspraxis. Das gilt auch für interdisziplinäre Forschungszusammenhänge sowie interdisziplinär angelegte Fächer. Wie spezifisch fachspezifisches Datenmanagement hier sein kann bzw.

---

<sup>1</sup> Cooper und Springer 2019.

sollte oder mit Blick auf Infrastrukturen und andere Ressourcen so überhaupt realisierbar ist, scheint keineswegs klar. Womöglich sind in der Praxis die Forschungsgegenstände oder die angewandten Methoden entscheidender für Notwendigkeiten und Verfahren im Umgang mit Daten als die Zugehörigkeit zu einer Fachdisziplin?

Ein alternativer Ansatz der Strukturierung von Services und Infrastrukturen wird mit Blick auf die STEM-Fachdisziplinen unter dem Label „Data Communities“ in einem Issue Brief der US-amerikanischen Non-Profit-Organisation Ithaka S+R diskutiert<sup>2</sup>.

Als „emergent data communities“ begreifen die Autorinnen „a loosely connected group of scholars who all work with a particular type of data, often linked by professional relationships“<sup>3</sup>. Data Communities zeichnen sich dadurch aus, dass es bereits gewisse formelle oder informelle Praktiken des Teilens / Nachnutzens von Daten gibt (bspw. auch die Nutzung von öffentlichen Daten). Folglich bestehe unter den Forschenden auch eine höhere Motivation, die Praktiken des Datenteilens voranzutreiben und zu institutionalisieren (bzw. daran mitzuwirken). Als wichtige Eigenschaften von Data Communities sehen die Autorinnen bereits vorhandene Standards, etwa in Bezug auf Dateiformate, Metadaten und ggf. Dokumentation / Kontextualisierung. Als Beispiele werden u.a. eine Datenbank mit spezifischen Informationen zu einer Fliegenart sowie die Forschung zum Thema Luftverschmutzung angeführt. Die Autorinnen argumentieren, dass in einem Bottom-up-Ansatz geeignete Communities identifiziert und gezielt gefördert werden sollten, etwa durch den Aufbau von an spezifische Bedürfnisse angepassten Repositorien.

## **2 Inhalte des Workshops**

Der Workshop möchte das Konzept der Data Communities in die Diskussion im deutschsprachigen Raum einbringen, sein Potenzial für die Geisteswissenschaften eruieren und aus verschiedenen Perspektiven diskutieren. Er richtet sich gleichermaßen an FDM-Praktizierende wie an Forschende und schließt ausdrücklich auch die Sozialwissenschaften mit ein, zumal eine klare Abgrenzung in vielen Fällen ohnehin schwierig ist. Idealerweise sind möglichst verschiedene Forschungsbereiche / Fächer im Workshop vertreten. Beiträge können

---

<sup>2</sup> Cooper und Springer 2019.

<sup>3</sup> Cooper und Springer 2019, 16.

einen oder mehrere der folgenden Aspekte aufgreifen oder eigene Schwerpunkte setzen:

**Data Communities in den Geistes- und Sozialwissenschaften:** Hier ist von Interesse, wie die Teilnehmenden das Potenzial des Konzepts einschätzen und ob sie in ihren Arbeitsfeldern solche Communities identifizieren können. Was könnte z.B. als „a particular type of data“ angesehen werden? Wären dies eher methodisch definierte Kategorien, wie z.B. Sprachkorpora, digitale Quelleneditionen, qualitative Interviewdaten, ethnografische Forschungsmaterialien, oder wäre hier eher an eine thematische Orientierung zu denken, wie z.B. Daten aus dem Bereich der Migrationsforschung? Oder auch beides, etwa Interviewdaten in Bezug auf Migration?

**Strukturelle Bedingungen und politische Rahmungen:** Unseres Erachtens besteht eine Diskrepanz zwischen einerseits vorhandenen Strukturen, die auf disziplinären Grenzen aufbauen (repräsentiert z.B. durch die Fachgesellschaften als „natürliche“ Ansprechpartner) und hier weitgehend unhinterfragt reproduziert werden, und andererseits komplexen Forschungsfeldern, multiperspektivischer sowie inter-/transdisziplinärer Bearbeitung von Forschungsfragen und daraus entstehenden interdisziplinär orientierten Forschungsverbänden oder Fächern (z.B. Game Studies oder Regionalstudienfächer). Hieraus resultiert ein fast schon anachronistischer Zuordnungsdruck, gerade für die „kleinen Fächer“, die durch eine Heterogenität an Methoden und Forschungsgegenständen gekennzeichnet sind. Es stellt sich die Frage, inwieweit das Konzept der Data Communities geeignet ist, um der Vielfalt der Forschungsrealität gerecht zu werden und das Forschungsdatenmanagement auf strategisch-operativer Ebene voranzubringen.

**FDM-Praxisschwerpunkt Archivierung von Forschungsdaten:** Zumeist werden – nicht zuletzt vor dem Hintergrund struktureller Bedingungen – drei Ansätze zur Archivierung bzw. zur Schaffung von Repositorien diskutiert und umgesetzt: generisch, institutionell und fachspezifisch. Auch hier stellt sich die Frage, ob nicht weitere Ansätze verfolgt werden sollten, die sich z.B. an Data Communities orientieren könnten. Könnten etwa Formen verteilter Archivierung von Forschungsdaten weiterführen? Für die Kuration und die Qualitätssicherung bei der Datenarchivierung ist ein umfassendes Verständnis für das jeweilige Material und den Entstehungskontext wesentlich. Verschiedene Data Communities formulieren unterschiedliche Bedingungen und

Anforderungen an Archivierung und Nachnutzung. Die ursprünglich datenproduzierende und datenarchivierende Data Community kann sich dabei durchaus auch von nachnutzenden Data Communities unterscheiden<sup>4</sup>.

### **3 Ablauf**

Der Workshop beginnt im ersten Teil mit einem Input-Vortrag von Danielle Cooper und einem Kommentar von Erzsébet Tóth-Czifra (in Englisch) zum Konzept der Data Communities, mit Raum für Fragen und Diskussion im Anschluss. Im zweiten Teil diskutieren verschiedene Expert:innen, wie Forschungsdatenmanagement für eine Data Community in Bezug auf die drei genannten Schwerpunkte praktisch aussehen kann (in Deutsch). In Kleingruppenarbeit sollen im Anschluss die Chancen und Herausforderungen, die sich aus der Praxis ergeben, herausgearbeitet und für einen partizipativen Debattenbeitrag aufbereitet werden.

### **4 Beitragende**

- Danielle Cooper (New York), Managerin für Kollaborationen und Forschung bei Ithaka S+R.
- Erzsébet Tóth-Czifra (Berlin), Open Science Officer bei DARIAH-EU, verantwortlich für die Konzeption und Stärkung der Dissemination von Forschungsergebnissen in den Geisteswissenschaften.
- Maria Effinger (Heidelberg), Leiterin der Abteilung "Publikationsdienste" der Universitätsbibliothek Heidelberg, Co-Sprecherin NFDI4Culture, Teilprojektleiterin eines INF-Projektes am SFB933.
- Dirk Goldhahn (Leipzig), Wissenschaftlicher Mitarbeiter Digital Humanities an der Sächsischen Akademie der Wissenschaften zu Leipzig.

---

<sup>4</sup> siehe z.B. Steiner, Zumbühl und Bauder 2008.

- Christoph Schmitt (Rostock), Leiter der Wossidlo-Forschungsstelle für Europäische Ethnologie/Volkskunde an der Universität Rostock und des digitalen Archivs WossiDiA.
- Samuel Mund (Hildesheim und Hannover), Wissenschaftlicher Mitarbeiter Sammlungsmanagement und Erschließung am Center for World Music und zuständig für Digital Humanities, EDV und Technik am Europäischen Zentrum für Jüdische Musik.

## Bibliografie

- Cooper, Danielle, and Rebecca Springer. 2019. "Data Communities: A New Model for Supporting STEM Data Sharing." *Ithaka S+R* (May 2019). <https://doi.org/10.18665/sr.311396>.
- Steiner, Daniel, Heinz J. Zumbühl, and Andreas Bauder. "Two Alpine Glaciers over the Past Two Centuries: A Scientific View Based on Pictorial Sources." In *Darkening Peaks: Glacier Retreat, Science, and Society*, edited by Ben Orlove, Ellen Wiegandt, and Brian H. Luckman, 83–99. Berkeley: University of California Press, 2008.