

Wissenschaftliche Daten sind kein Privateigentum einzelner Forschender, sondern ein kollektives Gut. Die Bereitstellung von Forschungsdaten zur Nachnutzung und Replikation muss auch in der Soziologie die Norm sein

Positionspapier der Akademie für Soziologie

Januar 2019

Produktive Wissenschaft lebt ganz entscheidend davon, dass sich die Wissenschaftler/inn/en an der Produktion von Kollektivgütern beteiligen. Das entstandene Wissen wird geteilt – genauso wie die Methoden und Daten, mit denen Wissen geschaffen wird. Kollektivgüter werden dann erfolgreich erstellt, wenn kluge institutionelle Regeln ihre Schaffung unterstützen. In diesem Zusammenhang wird derzeit über die Nutzung von Forschungsdaten debattiert und im Wissenschaftssystem entschieden, welche Infrastrukturen und Institutionen man aus- bzw. aufbauen soll. Erst der möglichst offene Zugang zu Daten und allen weiteren Materialien für die Gemeinschaft der Forschenden macht den Forschungsprozess intersubjektiv nachvollziehbar. Durch Transparenz wird wissenschaftliches Weiterdenken stimuliert.

Seit einigen Jahren bemüht sich in Deutschland die „Allianz der Wissenschaftsorganisationen“, darunter die Deutsche Forschungsgemeinschaft (DFG), die Norm der nachhaltigen Sicherung und Bereitstellung von Forschungsdaten zu stärken und geeignete Verfahren zur Durchsetzung dieser Norm zu entwickeln.¹ Um die praktische Umsetzung zu erleichtern, hat die DFG die wissenschaftlichen Fachgesellschaften aufgefordert, sich an der Diskussion zu beteiligen und fachspezifische Standards zu formulieren.

Die Akademie für Soziologie unterstützt diese Initiativen zur Bereitstellung von Forschungsdaten ausdrücklich. Öffentliche Zugänglichkeit zu wissenschaftlichen Produkten ist ein integraler Bestandteil von offenen Gesellschaften und ethisch zu verantwortender Forschung. Offene, nachnutzbare Daten ermöglichen die Korrektur von Fehlern, sie bilden

¹ Grundsätze zum Umgang mit Forschungsdaten, verabschiedet im Jahr 2010:

https://www.wissenschaftsrat.de/download/archiv/Allianz_Grundsätze_Forschungsdaten.pdf.

Dort heißt es etwa: „Die nachhaltige Sicherung und Bereitstellung von Forschungsdaten dient daher nicht nur der Prüfung früherer Ergebnisse, sondern in hohem Maße auch der Erzielung künftiger Ergebnisse. Sie bildet eine strategische Aufgabe, zu der Wissenschaft, Politik und andere Teile der Gesellschaft gemeinsam beitragen müssen.“

ein wesentliches Potenzial für inter- und multidisziplinäre sowie internationale Forschungszusammenarbeiten, und damit besonders gute Chancen für einen kumulativen Erkenntnisgewinn. Vielfach ist das Potenzial reichhaltiger, in der Regel mit öffentlichen Mitteln finanzierten Daten noch längst nicht ausgeschöpft, und dies wird einzelnen Forschenden auch gar nicht möglich sein. Daraus folgt: Die Bereitstellung und Nachnutzung von Forschungsdaten muss auch im Fach Soziologie die Regel sein.

Es ist aus Sicht der Akademie für Soziologie unverständlich, dass diese Grundposition innerhalb des Fachs Soziologie nicht allgemein geteilt wird. So hat die Deutsche Gesellschaft für Soziologie (DGS) jüngst in einer Stellungnahme erklärt, „dass (...) die Sekundärnutzung von Daten nicht als ‚Normalmodell‘ von Forschung gelten (...)“ darf. Vor allem rechtfertigten gravierende Probleme bei nicht-standardisiert erhobenen Daten, Forschungsdaten im Normalfall nicht weiterzugeben.²

Der Normalfall sollte aber genau anders herum definiert sein, denn die Bereitstellung von nicht-standardisierten Daten beinhaltet keine unüberwindbaren Hürden. Lösungen sind möglich und sind häufig schon in anderen Fächern implementiert³:

- 1) Datenschutz: Wo eine Anonymisierung von Forschungsdaten zu aufwändig wäre oder gar unmöglich ist, können die Daten einem „Secure Center“ für die Nachnutzung zur Verfügung gestellt werden. Dort können Sekundärnutzer die Daten auswerten, wobei sie nur die Auswertungsergebnisse, aber keine sensiblen Informationen aus dem Secure Center mitnehmen dürfen.
- 2) Rechte der Primärforschung: Es ist übliche Praxis, dass Drittmittelgeber den Primärforscher/inn/en eine „Embargofrist“ (bis zu vier Jahren) einräumen. Ebenfalls existieren bereits Regeln für die korrekte Zitation von Datensätzen.
- 3) Fehlende Ressourcen: Drittmittelgeber wie die DFG ermöglichen es bereits jetzt, Ressourcen für die Datenarchivierung zu beantragen. Beantragt werden können Mittel im Rahmen einzelner Projekte (etwa Personalmittel), aber auch im Rahmen eigener Infrastrukturplattformen. Der dafür notwendige zeitliche Aufwand schafft dann keinen Karrierenachteil für Primärforschende, wenn die Datenbereitstellung allgemeine Regel und nicht nur freiwillige Leistung ist.
- 4) Ungeklärte technische und organisatorische Herausforderungen und Unterstützungsbedarfe: An der Universität Bremen existiert beispielsweise seit 2018 das Datenser-

² „Bereitstellung und Nachnutzung von Forschungsdaten in der Soziologie“ vom 8. Januar 2019; <https://www.soziologie.de/nc/aktuell/stellungnahmen/single-view/archive/2019/01/09/article/bereitstellung-und-nachnutzung-von-forschungsdaten-in-der-soziologie/> . Etwa: „Die Primärforschenden haben jedenfalls das Recht, die Bedingungen für die Nachnutzung einzuschränken [...]. Forderungen z.B. von forschungsfinanzierenden Organisationen oder Publikationsorganen dürfen dieses Recht nicht relativieren.“

³ Vorbild können Disziplinen mit heterogenen Forschungsansätzen sein, siehe etwa die differenzierte Stellungnahme des Fachkollegiums Linguistik: http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_review_board_linguistics_corpora.pdf .

vicezentrum „Qualiservice“ für nicht-standardisierte Daten (www.qualiservice.org). Qualiservice hält diverse Services für die Bereitstellung und Nachnutzung von nicht-standardisierten Daten bereit, unter anderem deren Archivierung, Anonymisierung und sichere Nutzung im „Secure Center“. Damit existiert in Deutschland bereits eine umfassende Infrastruktur für die Bereitstellung und Nachnutzung von nicht-standardisierten Daten.

- 5) Kontextinformationen sind nur den Primärforschenden zugänglich und durch sie interpretierbar: Gerade zur Erhöhung der Transparenz sind relevante Kontextinformationen integraler Bestandteil von Forschungsdaten und wissenschaftlichen Veröffentlichungen. Sollten Primärforschende nicht alle relevanten Kontextinformationen bereitstellen, würden die Möglichkeit intersubjektiver Nachvollziehbarkeit unterlaufen und die Forschungsergebnisse somit entwertet.

Die Akademie für Soziologie erkennt die Prinzipien von FAIR Data – (*Findable, Accessible, Interoperable, and Reusable Data*), wie sie von der Europäischen Kommission etabliert werden,⁴ für die Soziologie als sinnvolle Orientierung an. Abweichungen davon wären in allen Entscheidungsprozessen des Wissenschaftssystems, von der Forschungsförderung bis zur Veröffentlichung begründungspflichtig.

⁴ Siehe für diese Deklarationen z.B. <https://ec.europa.eu/research/openscience/index.cfm?pg=home#>