

Leitlinie zum Umgang mit Forschungsdaten

(Mainz, 01.04.2019)

Allgemeines

Mit der Menge an Forschungsdaten in allen Wissenschaftsdisziplinen wächst auch die Herausforderung im Umgang mit diesen Daten, zumal vor dem Hintergrund des allgemeinen und von der EU im Forschungsrahmenprogramm „Horizon 2020“ eingeforderten Anspruchs, Daten öffentlich zugänglich zu machen (Open Data bzw. Open Access).

Daten sollen nachvollziehbar, die mit ihnen erzielten Forschungsergebnisse überprüfbar und schließlich für die wissenschaftliche Community verfügbar sein. Handelt es sich um personenbezogene Daten, wie dies an der Katholischen Hochschule in den weitaus meisten Forschungsvorhaben der Fall ist, müssen diese besonders sensibel gehandhabt werden. Ethische und rechtliche Beschränkungen müssen – neben den spezifischen Anforderungen der Drittmittelgeber – berücksichtigt werden.

Vorliegende Leitlinie dient den Forschenden an der Katholischen Hochschule Mainz als Grundlage für die Definition von Forschungsdaten und den verantwortungsvollen Umgang mit diesen. Sie soll darüber hinaus auch anregen, von Beginn eines Forschungsprozesses an die Möglichkeiten einer (begrenzten) Veröffentlichung der gewonnenen Forschungsdaten zu erwägen.

Das Institut für angewandte Forschung und internationale Beziehungen (ifb) unterstützt und berät zum Thema Forschungsdatenmanagement. Es wird dringend empfohlen, das ifb bereits in der Planungsphase eines Forschungsvorhabens beratend einzubeziehen. Darüber hinaus bietet das ifb Schulungen zum Thema Forschungsdatenmanagement an.

Was sind Forschungsdaten?

Forschungsdaten sind sämtliche, insbesondere digitale Daten, die im Verlauf eines Forschungsvorhabens bei Experimenten, Messungen, Erhebungen, Umfragen, Surveys, mittels Fragebögen oder audiovisuellen Aufzeichnungen etc. generiert werden. Zu den Forschungsdaten gehören auch deren Dokumentation (Metadaten) und die zu ihrer Erhebung oder Verarbeitung erforderliche Software. Forschungsdaten können in unterschiedlichen Aggregationsstufen und digitalen Formaten vorliegen.

Was bedeutet Forschungsdatenmanagement?

Der Begriff des Forschungsdatenmanagements (FDM) umfasst sämtliche Aufgaben, die sich aus dem sog. Lebenszyklus von Daten ergeben: von der Datenerhebung in einem Forschungsprojekt über die Datenbearbeitung bzw. -analyse, die Archivierung der Forschungsdaten in dafür vorgesehenen hochschulischen Dateninfrastrukturen oder (fach-) spezifischen externen Repositorien bis hin zur Löschung der Daten:

- Planung der Datenerhebung,
- Erhebung und Analyse der Daten,

- Dokumentation der Daten (Metadaten),
- rechtssichere und forschungsethisch unbedenkliche Verwendung von Forschungsdaten,
- Archivierung der generierten Daten,
- nachhaltige Nutzbarmachung der Daten für andere Forscherinnen und Forscher.

Wie werden Daten organisiert und verwaltet?

Nationale und internationale Drittmittelgeber erwarten zunehmend die Vorlage von sogenannten Datenmanagementplänen (DMP) schon bei der Beantragung von Forschungsprojekten. Datenmanagementpläne dienen der Berücksichtigung „methodenangemessener Standards zur Sicherstellung einer hohen Datenqualität – auch mit Blick auf die Wahrung geistigen Eigentums, die Innovationsfähigkeit des Forschungssystems, datenschutzrechtliche und forschungsethische Anforderungen“.¹

Datenmanagementpläne sollen auch Regelungen zur Archivierung und ggf. Veröffentlichung von Forschungsdaten treffen. Die langfristige Sicherung von Daten in geeigneten Repositorien ist die Grundlage für die Nachvollziehbarkeit, Replizierbarkeit und letztlich Überprüfbarkeit von Forschungsergebnissen. Sie garantiert also Transparenz und Qualitätssicherung in der Wissenschaft. Darüber hinaus wird eine Nachnutzung von Forschungsdaten, die im Kontext öffentlich geförderter Forschungsvorhaben generiert wurden, zunehmend von Drittmittelgebern und Scientific Community auch für andere Forschungskontexte erwartet.

Datenmanagementpläne dienen der Einhaltung der allgemeinen Grundsätze guter wissenschaftlicher Praxis und sollen sich an den Leitlinien der DFG zum Umgang mit Forschungsdaten orientieren:

http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf

Datenmanagementpläne sind keine statischen, sondern dynamische Dokumente. Sie sind dem jeweils aktuellen Verlauf eines Forschungsvorhabens anzupassen.

Wer ist verantwortlich für die Forschungsdaten?

Während der gesamten Laufzeit eines Forschungsvorhabens sind die Projektleitungen für das Management aller generierten Forschungsdaten, ihre Korrektheit und den Umgang mit diesen verantwortlich. Sie sind verpflichtet, die Einhaltung der fachbezogenen Standards sicherzustellen, die spezifischen Anforderungen der Drittmittelgeber im Hinblick auf die gewonnenen Forschungsdaten zu berücksichtigen und, sofern dies sinnvoll und möglich ist, die langfristige Archivierung der Forschungsdaten zu gewährleisten. Dabei beachten sie insbesondere ethische, datenschutzrechtliche sowie urheberrechtliche Belange. Die Beachtung der allgemeinen Grundsätze zur Sicherung guter wissenschaftlicher Praxis, wie sie von der DFG formuliert wurden,² wird vorausgesetzt.

Die Projektleitungen entscheiden im Benehmen mit den Projektmitarbeiterinnen und -mitarbeitern über die Auswahl der zu archivierenden Daten, den Speicherort und die Speicherkonditionen (Fachrepositorium, Nachnutzbarkeit etc.) sowie ggf. über deren Veröffentlichung bzw. den Zeitpunkt der Veröffentlichung sowie geeignete Lizenzbedingungen (etwa sog. Creative Commons-

¹ Rat für Sozial- und Wirtschaftsdaten (RatSWD 2018): Forschungsdatenmanagement in den Sozial-, Verhaltens- und Wirtschaftswissenschaften - Orientierungshilfen für die Beantragung und Begutachtung datengenerierender und datennutzender Forschungsprojekte. RatSWD Output 3 (5). Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD). https://www.ratswd.de/dl/RatSWD_Output3_Forschungsdatenmanagement.pdf (Zugriff 01.04.2019), S. 5.

² http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf (Zugriff 01.04.2019).

Lizenzverträge; mehr dazu unter <https://de.creativecommons.org/index.php/was-ist-cc/>). Eine Veröffentlichung soll unter dem Namen der Projektverantwortlichen erfolgen.

Die Projektleitungen treffen überdies frühzeitig Regelungen für das Ausscheiden aus dem Projekt durch Arbeitgeber- oder Ortswechsel etc. von Projektbeteiligten.

Wie werden Daten dokumentiert?

Eine nachhaltige Sicherung von Forschungsdaten kann in unterschiedlichen Stufen und abhängig von der eventuellen Nachnutzung und/oder Veröffentlichung der Daten erfolgen:

- ohne nutzungsfreundliche Dokumentation und Nachnutzungsmöglichkeiten,
- mit nutzungsfreundlicher Dokumentation und restriktiven Nutzungsmöglichkeiten und schließlich
- mit einer Bereitstellung von Forschungsdaten einschließlich nutzungsfreundlicher Dokumentation für eine Nutzung auch außerhalb des Aufbewahrungsortes.

Über den Dokumentationsgrad entscheiden die Projektverantwortlichen nach Maßgabe des Drittmittelgebers und unter Berücksichtigung insbesondere datenschutzrechtlicher und urheberrechtlicher Aspekte.

Zu dokumentieren sind der Entstehungskontext der Forschungsdaten, die Kontextinformationen zu Werkzeugen sowie zur verwendeten Software, der Methodenbericht, die Analyseprotokolle sowie der Forschungsprozess an sich.

Dabei können bestehende disziplinäre Dokumentations- und Metadatenstandards (z.B. DDI, Dublin Core etc.) hilfreich sein. Technische Informationen zu den verwendeten Programmcodes, die für die Datenaufbereitung und Datenanalyse in den jeweils genutzten Statistikprogrammen (z.B. SAS, SPSS, STAT, R) erzeugt werden, müssen ebenfalls in die Dokumentation einfließen.

Wie werden Daten gespeichert?

Die kurzfristige Speicherung digitaler bzw. digitalisierter Forschungsdaten erfolgt entweder auf Servern und IT-Strukturen der Katholischen Hochschule Mainz oder alternativ in von der IT der KH Mainz ausgewählten Cloud-Diensten, um den dezentralen Zugriff aller Projektbeteiligten auf synchronisierte Daten jeweils aktuell zu gewährleisten (bspw. Seafile, B2DROP).

Die langfristige Speicherung und Archivierung von Forschungsdaten erfolgt in der IT-/Informationsinfrastruktur der Katholischen Hochschule Mainz oder in anerkannten externen Fachrepositorien (bspw. beim Leibniz-Institut für Sozialwissenschaften GESIS). Die Konditionen der Archivierung und ggf. der späteren Veröffentlichung der Daten werden vertraglich mit dem gewählten Fachrepositorium vereinbart. Eine frühzeitige Abstimmung wird hier dringend empfohlen. Besteht die Möglichkeit, Kosten des Datenmanagements gegenüber dem Drittmittelgeber geltend zu machen, sollte die Abstimmung mit dem Fachrepositorium und eine Angebotseinholung schon während der Antragstellung erfolgen, um die Kosten gegenüber dem Drittmittelgeber geltend zu machen.

Für das Format der Daten werden vorzugsweise freie Standardformate empfohlen, um Interoperabilität und langfristige Lesbarkeit sicherzustellen.

Die Aufbewahrungsfristen richten sich nach allgemeinen gesetzlichen oder vertraglichen Vorschriften, insbesondere den Maßgaben der Drittmittelgeber. In der Regel und gemäß den Leitlinien der DFG zum Umgang mit Forschungsdaten (s.o.) müssen Forschungsdaten für mindestens zehn Jahre aufbewahrt werden.

Werden Nachnutzungsrechte oder Veröffentlichungsrechte an Dritte übertragen, so müssen die Projektverantwortlichen sicherstellen, dass die Daten für wissenschaftliche Zwecke frei verfügbar bleiben.

Bei Fragen zum Forschungsdatenmanagement wenden Sie sich bitte an:

Dr. phil. Elke Bruck
Institut für angewandte Forschung und internationale Beziehungen (ifb)
Katholische Hochschule Mainz
Saarstr. 3
D-55122 Mainz
Telefon: +49 6131 289 44-730
E-Mail: elke.bruck@kh-mz.de