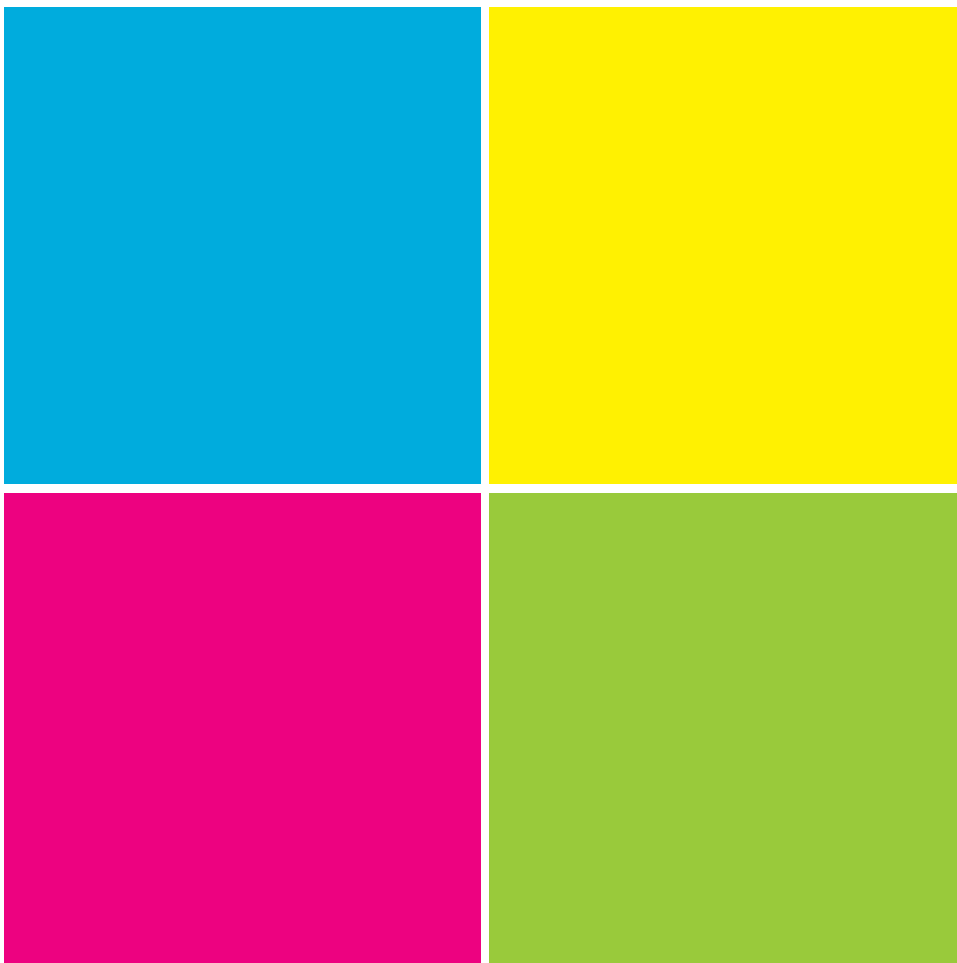


Basisinformationen zum Forschungsdatenmanagement*

Orientierungshilfen für die Beantragung
und Begutachtung datengenerierender
und datennutzender Forschungsprojekte



*Kurzfassung des RatSWD OutPut Papers Nr|3 „Forschungsdatenmanagement in den
Sozial-, Verhaltens- und Wirtschaftswissenschaften“
www.ratswd.de/dl/RatSWD_Output3_Forschungsdatenmanagement.pdf

Basisinformationen zum Forschungsdatenmanagement in den Sozial-, Verhaltens-, und Wirtschaftswissenschaften

Der RatSWD möchte mit dieser Empfehlung Antragstellenden und Gutachtenden eine Orientierungshilfe zu den grundlegenden Fragen eines modernen Forschungsdatenmanagements in den Sozial-, Verhaltens- und Wirtschaftswissenschaften geben.

Im Folgenden werden die Basisinformationen für Projekte vorgestellt und dabei danach unterschieden, ob die Projekte selbst Daten generieren (1.1) oder im Rahmen einer Sekundärdatenanalyse nachnutzen (1.2).

1 Datengenerierende Projekte (Primärdatenanalyse):

A. Angaben zu Datensicherheit, Archivierung und Weitergabe der Forschungsdaten

Angaben zu den technischen und organisatorischen Maßnahmen zur Umsetzung der Datensicherheit (Sicherung vor Verlust, Manipulation/Verfälschung, unberechtigtem Zugang) bei Primärerhebungen und bei Sekundäranalysen sollten dokumentiert werden. Dies beinhaltet u.a. Angaben zur Einverständniserklärung, zur Privilegierung von Forschung wie auch zu Dokumentationspflichten.

Für alle Daten gilt eine in den Regeln guter wissenschaftlicher Praxis festgelegte 10-jährige Aufbewahrungsfrist (vgl. DFG 2015). Diese gilt sowohl für im Projekt generierte Primärdaten wie auch für – im Fall von sekundäranalytischen Projekten – an den Originaldaten vorgenommene Modifikationen (Syntaxfiles). Eine Sicherung kann in einer der drei folgenden Varianten erfolgen:

Variante 1:

Sicherung der Daten in einem Repository der Forschungseinrichtung des Datenproduzenten (Universität, Forschungsinstitut) oder Übergabe an eine Einrichtung der Forschungsdateninfrastruktur ohne Zugriffs- oder Nutzungsrechte für Dritte (<http://www.ratswd.de/forschungsdaten/fdz>).

Ziel ist es, die erhobenen Daten im Sinne der datenschutzrechtlichen Anforderungen und vor dem Hintergrund des informationstechnischen Wandels zu sichern. Ein Zugriff für andere Wissenschaftlerinnen und Wissenschaftler im Rahmen sekundäranalytischer Nachnutzungen ist aus datenschutzrechtlichen oder auch anderen, im Antrag darzustellenden, Gründen nicht vorgesehen. Die Dokumentation muss nicht nutzerfreundlich sein.

Variante 2:

Sicherung der Daten und des Datenzuganges auch für andere Wissenschaftlerinnen und Wissenschaftler zwecks sekundäranalytischer Nachnutzung unmittelbar beim Datenproduzenten oder einer Einrichtung der Forschungsdateninfrastruktur.

Die Forschungsdaten (Primärdaten) werden wie in Variante 1 archiviert, können jedoch vor Ort (an der Forschungseinrichtung selbst oder einer Infrastruktureinrichtung) auch von anderen Wissenschaftlerinnen und Wissenschaftlern im Rahmen von Gastaufenthalten genutzt werden. Aufgrund dieser Nutzungsmöglichkeiten entsteht ein höherer Anonymisierungs-, Aufbereitungs- und Dokumentationsaufwand, der in der Ressourcenausstattung des Projekts berücksichtigt werden sollte.

Variante 3:

Sicherung und Bereitstellung der Daten zwecks sekundäranalytischer Nachnutzung durch Übergabe an eine Einrichtung der Forschungsdateninfrastruktur

Die Daten werden an eine Infrastruktureinrichtung übergeben und dort zur Nutzung in Form eines Scientific Use Files (SUF), Public Use Files (PUF), einer Variante der Datenfernverarbeitung oder im Rahmen von Gastaufenthalten bereitgestellt. Hierzu sind besondere Maßnahmen der Anonymisierung, der Datensatzaufbereitung und der Dokumentation vorzusehen.

Falls eine Übergabe an eine bestehende Einrichtung der Forschungsdateninfrastruktur geplant ist, sollte diese im Förderantrag benannt und bereits vorab kontaktiert werden, um die Übergabemodalitäten und insbesondere den Ressourcenbedarf zu klären.

B. Angaben zu Datenschutz und Forschungsethik

Bei Primärerhebungen sollten zum Zeitpunkt der Antragstellung Angaben gemacht werden, auf welche der drei Sicherungsvarianten sich die verwendeten Einwilligungserklärungen für Studienteilnehmende beziehen. Angaben zu datenschutzrechtlichen Vorkehrungen während der Projektlaufzeit und nach Ablauf des Projektes sollten kurz benannt werden, ebenso wie die Information, ob das Votum einer Ethikkommission eingeholt wird.

C. Angaben zur methodischen, technischen und bibliographischen Dokumentation der Daten

Entsprechend der geplanten Sicherungsvariante sollten Angaben zur methodischen, technischen und bibliographischen Dokumentation und den dabei ggf. verwendeten Standards (z.B. den technischen Formaten, den Metadatenformaten) gemacht werden. Bei der Projektplanung sollten die dafür notwendigen Arbeitspakete spezifiziert und der Ressourcenbedarf berücksichtigt werden. Ebenfalls sollte kurz erläutert werden, ob im Rahmen des Projekts ein Datenmanagementplan erstellt wird und auf welche Vorlagen hierfür zurückgegriffen wird.

2 Datennutzende Projekte (Sekundärdatenanalyse):

In Forschungsprojekten, in denen selbst keine Daten generiert werden, sondern auf bereits existierende Forschungsdaten zurückgegriffen wird, sollten die im Projekt vorgenommenen Arbeiten zur Datenaufbereitung und Datenmodifikation dokumentiert und nach Projektabschluss bereitgestellt werden (z.B. mit Programmfiles zur Datenaufbereitung).

Die Bereitstellung der für die Projektergebnisse relevanten datenverändernden Maßnahmen ist für eine spätere Rekonstruktion der Datengrundlage und der erlangten Forschungsergebnisse und deren Replikation unabdingbar. Bereits jetzt erwarten einzelne Fachzeitschriften die Bereitstellung der kompletten Programmcodes, um die einzelnen Arbeitsschritte von der Aufbereitung der Grunddaten bis hin zu den in den Publikationen präsentierten Tabellen und Grafiken nachvollziehbar dokumentieren zu können.

In den Forschungsdatenzentren werden gesonderte Repositorien aufgebaut, um den Datennutzenden die Archivierung und Bereitstellung ihrer Programmcodes zu ermöglichen (z.B. datorium bei der GESIS). Projektanträge sollten deshalb Aussagen darüber enthalten, wie die Aufbereitung der Sekundärdaten dokumentiert wird und in welcher Form und in welchem Repository diese Arbeiten nach Projektabschluss bereitgestellt werden. Analog zur Sicherung bzw. Nachnutzung von Primärdaten sind hier drei Varianten möglich:

Variante 1:

Mindestens 10-jährige Aufbewahrung der Programmcodes, der datenmodifizierenden Arbeiten und abgeleiteten Daten in einem Repository der Forschungseinrichtung des Sekundärdatennutzers (Universität, Forschungsinstitut) oder Übergabe an eine Einrichtung der Forschungsdateninfrastruktur ohne Zugriffs- oder Nutzungsrechte Dritter. Die Dokumentation muss dabei nicht zwingend nutzerfreundlich sein.

Variante 2:

Mindestens 10-jährige Aufbewahrung der Programmcodes, der datenmodifizierenden Arbeiten und abgeleiteten Daten in einem Repository der Forschungseinrichtung des Sekundärdatennutzers (Universität, Forschungsinstitut) oder Übergabe an eine Infrastruktureinrichtung mit restriktiven Nutzungsrechten für Dritte nach expliziter Genehmigung durch den Sekundärdatennutzer. Die Dokumentation sollte nutzerfreundlich sein.

Variante 3:

Sicherung und mindestens 10-jährige Aufbewahrung in einem Repository der Forschungseinrichtung des Sekundärdatennutzers (Universität, Forschungsinstitut) bzw. Übergabe an eine Einrichtung der Forschungsdateninfrastruktur mit erweiterten Nutzungsrechten für Dritte. Die Programmcodes, datenmodifizierenden Arbeiten und abgeleiteten Daten, die nutzerfreundlich dokumentiert sein sollten, sind von den Sekundärdatennutzern für die Nutzung auch durch andere Wissenschaftlerinnen und Wissenschaftler unter Einhaltung der üblichen Zitationsregeln freigegeben.

Auch hier sind unterschiedliche Dokumentationsumfänge in den einzelnen Varianten bei der Ressourcenplanung zu berücksichtigen. Bei einer geplanten Übergabe an eine Infrastruktureinrichtung sollten im Zuge der Antragstellung vorab die Möglichkeiten und Bedingungen der Sicherung bzw. Nachnutzung geklärt werden.