

RDMO-Fragenkataloge für Datenmanagementpläne, die als Beigabe zu Anträgen auf Forschungsförderung an die DFG dienen sollen

Version 3.0 vom 20.04.2019

Dokumentation

Torsten Rathmann
Projekt FoDaKo (Forschungsdatenmanagement im Kooperationsverbund)
Universitätsbibliothek und Zentrum für Informations- und Medienverarbeitung
Bergische Universität Wuppertal
Gaußstr. 20
42119 Wuppertal



Dieses Dokument und die beschriebenen Fragenkataloge dürfen frei verwendet werden¹.

Einführung und Kurzfassung

Für Datenmanagementpläne, die Anträgen auf Forschungsförderung an die DFG beigelegt werden sollen, wurden Fragenkataloge zur Benutzung mit der Webapplikation RDMO (Research Data Management Organiser) entwickelt. Dabei wurde der im RDMO-Projekt entwickelte generische Fragenkatalog „RDMO“ durch Weglassen und Hinzufügen von Fragen so angepasst, dass nur die Empfehlungen der DFG für Antragstellende durch Fragen abgedeckt sind. Fragen zu den Kosten des Forschungsdatenmanagements wurden dabei nicht berücksichtigt, weil Kosten im vorgesehenen DFG-Formular eingetragen werden müssen, um erstattet werden zu können. Das Downloadpaket des FoDaKo-Projekts enthält in dieser Weise zusammengestellte Fragenkataloge für die Bildungs-, Sozial- und Wirtschaftswissenschaften. Außerdem sind ein fachübergreifender, nach den DFG „Leitlinien zum Forschungsdatenmanagement“ entwickelter Katalog enthalten und der um die neu hinzugekommenen Fragen erweiterte generische Katalog, der hier den Namen „Alle Fragen“ trägt.

Entstehung

Im DFG-geförderten Projekt RDMO² ist nicht nur die gleichnamige Software³ entwickelt worden, sondern auch ein Fragenkatalog mit dem Namen „RDMO“, in den die WissGrid-Checkliste⁴ und viele Hinweise und Überlegungen eingeflossen sind. Dieser Fragenkatalog enthält zurzeit 123 Fragen. Im Projekt FoDaKo (Forschungsdatenmanagement im Ko-

¹ Lizenz: CC0 1.0 Universal, <https://creativecommons.org/publicdomain/zero/1.0/legalcode>; die in dieser Dokumentation beschriebenen Fragenkataloge sind Weiterentwicklungen des Fragenkatalogs „RDMO“, der vom RDMO-Projekt² ebenfalls unter CC0-Lizenz angeboten wird.

² RDMO, <https://rdmorganiser.github.io/> [aufgerufen am 16.04.2019]

³ RDMO, <https://github.com/rdmorganiser/rdmo> [aufgerufen am 16.04.2019]

⁴ Ludwig, J., & Enke, H. (Editoren) (2013), „Leitfaden zum Forschungsdaten-Management“, Verlag Werner Hülsbusch, ISBN: 978-3-86488-032-2, Seiten 83-100, https://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf [aufgerufen am 16.04.2019]

operationsverbund, ein vom BMBF gefördertes Projekt der Universitäten Düsseldorf, Siegen und Wuppertal) wurde überlegt, wie diese Fülle für Wissenschaftler und Schulungsteilnehmer erschließbar gemacht werden kann. Zunächst wurde eine Ansicht (View, Template) entwickelt, da Fragenkataloge zu der Zeit noch nicht über die Web-Oberfläche in RDMO importiert werden konnten und das Kopieren eines Fragenkatalogs mittels Neueingabe mühsam gewesen wäre. Inzwischen ist der Export und Import von Fragenkatalogen über die Web-Oberfläche einfach möglich und die Ansicht konnte durch passende Unterkataloge des RDMO-Standard-Fragenkatalogs ersetzt werden. Unter einem Unterkatalog soll dabei ein Fragenkatalog verstanden werden, der nur Fragen eines größeren Katalogs enthält.

Ansichten für die hier genannten Kataloge werden im beschriebenen Download-Paket nicht mehr mitgeliefert, weil Ansichten Fragen und Antworten filtern und diese Filterfunktion Nutzer auch schon verwirrt hat.

Veränderungen gegenüber Version 2.3

- Unterkatalog für die Bildungswissenschaften hinzugekommen
- Das DFG Fachkollegium 112 „Wirtschaftswissenschaften“ hat inzwischen eigene Empfehlungen⁵ herausgegeben. Daraus wurde ein Unterkatalog für die Wirtschaftswissenschaften entwickelt und vom bisherigen Unterkatalog „RatSWD + DFG“ abgetrennt, der jetzt „Soziologie + DFG“ heißt
- Die neuen Unterkataloge enthalten vier Fragen, die im Katalog „RDMO“ keine Entsprechung haben. Die Domäne wurde um entsprechende Attribute erweitert. Auch Katalog „RDMO“ wurde um diese Fragen erweitert. Der erweiterte und nach den folgenden Punkten angepasste Katalog ist unter dem Namen „Alle Fragen“ mit enthalten. „Alle Fragen“ wurde – wie die übrigen Kataloge im Download-Paket – gemäß der vier letzten Punkte angepasst
- Da die DFG den Begriff „Langzeit“ schon für einen Zeitraum von zehn Jahren verwendet, wurden vorsichtshalber drei Fragen zur Langzeitspeicherung in den Katalog „DFG“ aufgenommen
- Die Frage *„Für welche Personen, Gruppen oder Institutionen könnte dieser Datensatz (für die Nachnutzung) von Interesse sein? Für welche Szenarien ist dies denkbar?“* wurde um die Teilfrage *„Welche Konsequenzen hat das Nachnutzungspotential später für die Bereitstellung der Daten?“* erweitert. In der Hilfe befinden sich Hinweise, wie diese Frage erörtert werden kann.
- Anpassung an die Datenschutz-Grundverordnung und die neuen Datenschutzgesetze
- Zahlreiche (meist fachspezifische) Hinweise in den beim Ausfüllen sichtbaren Hilfen
- Korrektur von abgehangenen Links, Schreib- und Zuordnungsfehlern

Veränderungen gegenüber Version 2.2

Die Fragen *„Projektpartner“* und *„Gibt es an Ihrer Einrichtung Regeln oder Richtlinien zum Umgang mit den im Projekt erhobenen Forschungsdaten? Wenn ja, skizzieren Sie diese kurz und verweisen Sie ggf. auf weiterführende Informationen. Geben Sie bitte auch an, welchen Grad an Verbindlichkeit sie haben.“* wurden aus beiden Unterkatalogen herausgenommen. Die beiden Fragen waren nur in den Unterkatalogen enthalten, um die Frage *„Gibt es von*

⁵ DFG, Management von Forschungsdaten: Was erwartet das Fachkollegium 112 „Wirtschaftswissenschaften“ von Antragstellenden?(Stand November 2018), https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/fachkollegium112_forschungsdatenmanagement_1811.pdf [aufgerufen am 16.04.2019]

weiteren Seiten (z. B. von der Fachcommunity) Anforderungen an das Datenmanagement, die beachtet werden müssen?“ besser verstehen zu können. Durch Umformulierung dieser Frage in „Gibt es von Seiten Ihrer Fachcommunity Anforderungen an das Datenmanagement, die beachtet werden müssen?“ konnte erreicht werden, dass die Fragen „Projektpartner“ und „Gibt es an Ihrer Einrichtung...“ ganz entfallen konnten. Außerdem wurden einige Fehler korrigiert.

Veränderung gegenüber Version 2.1

Alle Unterkataloge wurden an RDMO 0.11.0, d.h. an das neue Datenmodell, angepasst.

Der Unterkatalog „DFG“

Die Unterkatalog „DFG“ ist konzipiert für Datenmanagementpläne, die in ihrer ersten Version als Anhang zu Anträgen auf Forschungsförderung bei der Deutschen Forschungsgemeinschaft (DFG) dienen sollen und für die DFG keine zusätzlichen fachspezifischen Anforderungen stellt⁶. Die DFG wurde ausgewählt, weil sie der größte Forschungsförderer in Deutschland ist und außerdem *Leitlinien zum Forschungsdatenmanagement*⁷ verfasst hat. Aus diesen Leitlinien sei hier der Abschnitt „Projektplanung und Antragstellung“ zitiert:

Bereits in die Planung eines Projekts sollten Überlegungen einfließen, ob und welche der aus einem Vorhaben resultierenden Forschungsdaten für andere Forschungskontexte relevant sein können und in welcher Weise diese Forschungsdaten anderen Wissenschaftlerinnen und Wissenschaftlern zur Nachnutzung zur Verfügung gestellt werden können. In einem Antrag sollten die Antragstellenden daher ausführen, welche Forschungsdaten im Verlauf eines wissenschaftlichen Forschungsvorhabens entstehen, erzeugt oder ausgewertet werden. Dabei sollten fachspezifisch angemessene Konzepte und Überlegungen für die Qualitätssicherung, für den Umgang mit und die langfristige Sicherung der Forschungsdaten zugrunde gelegt werden. Die einschlägigen Erläuterungen müssen Informationen zu Datentypen, falls vorhanden zu disziplinspezifischen Standards und zur Wahl geeigneter Repositorien enthalten, sofern diese für ein bestimmtes Fachgebiet oder bestimmte Datentypen vorhanden sind. Zusätzlich werden Angaben zu ggf. betroffenen Rechten Dritter sowie erste Planungen zum zeitlichen Rahmen der Datenveröffentlichung erbeten.

Dieser Absatz wurde in die Fragenliste unten umformuliert (**fett und kursiv**) und zu diesen Fragen dann möglichst gut passende aus dem generischen Fragenkatalog „RDMO“ gesucht (*kursiv*). Zusätzlich angegeben ist das Attribut, auf das die Frage zeigt (verkleinert). Inhaltlich kommt noch die Vorgabe der DFG hinzu, dass Primärdaten, die Grundlage einer Publikation sind, langfristig (für mindestens 10 Jahre) in der eigenen Einrichtung oder in einer fachlich einschlägigen, überregionalen Infrastruktur gespeichert werden sollen (letzte Frage).

Welche Forschungsdaten entstehen im Verlauf des Vorhabens?

Um was für einen Datensatz handelt es sich?

project/dataset/description

⁶ DFG, *Umgang mit Forschungsdaten*,

http://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/ [aufgerufen am 20.11.2018]; auf dieser Webseite befindet sich eine Liste von Links zu fachspezifischen Empfehlungen.

⁷ DFG, *Leitlinien zum Umgang mit Forschungsdaten*,

http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf [aufgerufen am 09.07.2018]

Mit welchen Datentypen wird umgegangen werden?

In welchen Formaten liegen die Daten vor?

project/dataset/format

Welche Maßnahmen sind zum Zwecke der Qualitätssicherung der Daten geplant?

Welche Maßnahmen zur Qualitätssicherung werden für diesen Datensatz ergriffen?

project/dataset/quality_assurance

Welche disziplinspezifischen Standards sind zu beachten?

Gibt es von Seiten Ihrer Fachcommunity Anforderungen an das Datenmanagement, die beachtet werden müssen?

project/additional_rdm_policy/yesno

Welche Anforderungen an das Datenmanagement sind dies?

project/additional_rdm_policy/requirements

Welche Repositorien sind für die Speicherung vorgesehen?

Wo wird der Datensatz während des Projektes gespeichert?

project/dataset/storage/type

Wo werden die Daten (einschließlich Metadaten, Dokumentation und ggf. relevantem Code bzw. relevanter Software) nach Projektende gespeichert bzw. archiviert?

project/dataset/preservation/repository

Welche der aus dem Vorhaben resultierenden Forschungsdaten könnten für andere Forschungskontexte relevant sein?

Für welche Personen, Gruppen oder Institutionen könnte dieser Datensatz (für die Nachnutzung) von Interesse sein? Für welche Szenarien ist dies denkbar? Welche Konsequenzen hat das Nachnutzungspotential später für die Bereitstellung der Daten?

project/dataset/reuse_scenario

In welcher Weise werden die Forschungsdaten anderen Wissenschaftlerinnen und Wissenschaftlern zur Verfügung gestellt?

Soll dieser Datensatz veröffentlicht oder geteilt werden?

project/dataset/sharing/yesno

Wenn ja, unter welchen Nutzungsbedingungen oder welcher Lizenz sollen die Daten veröffentlicht bzw. geteilt werden?

project/dataset/sharing/conditions

Welche Hinderungsgründe gibt es, die Daten anderen Wissenschaftlerinnen und Wissenschaftlern zur Verfügung zu stellen?

Wenn nicht, begründen Sie dies bitte und unterscheiden Sie dabei zwischen rechtlichen und/oder vertraglichen Gründen und freiwilligen Einschränkungen.

project/dataset/sharing/explanation

Falls eine Zugänglichmachung der Daten geplant ist, für wann?

Wann werden die Daten veröffentlicht?

project/dataset/data_publication_date

Welche Rechte Dritter werden berührt?

Enthält dieser Datensatz personenbezogene Daten?

project/dataset/sensitive_data/personal_data_yesno/yesno

Werden die Daten anonymisiert oder pseudonymisiert?

project/dataset/sensitive_data/personal_data/anonymization

Werden Daten genutzt und/oder erstellt, die durch Urheber- oder verwandte Schutzrechte geschützt sind?

project/legal_aspects/ipr/yesno

Be- oder entstehen an diesem Datensatz Urheberrechte?

project/dataset/ipr/copyrights

Be- oder entstehen an diesem Datensatz andere Schutzrechte?

project/dataset/ipr/other_rights

Warum, wie lange und wo werden die Daten langfristig aufbewahrt?

Muss dieser Datensatz langfristig aufbewahrt werden?

project/dataset/preservation/yesno

Aus welchen Gründen müssen die Daten langfristig aufbewahrt werden?

project/dataset/preservation/purpose

Wie lange müssen die Daten aufbewahrt werden?

project/dataset/preservation/duration

Die Frage „Welche Repositorien sind für die Speicherung vorgesehen?“ oben deckt das „wo“ ab.

Ausgehend von dieser Liste wurde der Unterkatalog „DFG“ zusammengestellt.

Der Unterkatalog „Soziologie + DFG“

Für eine Reihe von Fächern hat die DFG die Richtlinien bestimmter Fachorganisationen aufgelistet⁶. Ein Beispiel hierfür sind die Sozial- und Wirtschaftswissenschaften, für die der Rat für Sozial- und Wirtschaftsdaten (RatSWD) die *Basisinformationen zum Forschungsdatenmanagement*⁸ herausgegeben hat.

Die sechs Varianten und Fallunterscheidungen in den *Basisinformationen*⁸ lassen sich nicht direkt auf einen Fragenkatalog für RDMO abbilden, aber — ähnlich wie im Falle der DFG-Leitlinien auch — lassen sich aus dem Text Stichwörter extrahieren (fett und kursiv gesetzt) und zu diesen Fragen passende Fragen (nur kursiv gesetzt) im generischen Fragenkatalog „RDMO“ finden, der mit der RDMO-Software mitgeliefert wird. Die Entscheidung, welcher der sechs Varianten im Projekt gefolgt wird, kann in der Antwort auf die Frage

Für welche Personen, Gruppen oder Institutionen könnte dieser Datensatz (für die Nachnutzung) von Interesse sein? Für welche Szenarien ist dies denkbar? Welche Konsequenzen hat das Nachnutzungspotential später für die Bereitstellung der Daten?,

die schon im Fragenkatalog „DFG“ vorhanden ist, mitgeteilt und begründet werden. Die Hilfe zu dieser Frage gibt nähere Hinweise im Sinne einer Ausfüllanleitung.

Im Folgenden sind die Fragen aufgelistet, die noch nicht durch den Fragenkatalog „DFG“ abgedeckt sind.

Sekundärdatenanalyse

Wird der Datensatz selbst erzeugt oder nachgenutzt?

project/dataset/origin

Wenn nachgenutzt, wer hat den Datensatz erzeugt?

project/dataset/creator/name

Wenn nachgenutzt, unter welcher Adresse, PID oder URL ist der Datensatz verfügbar?

project/dataset/uri

⁸ Rat für Sozial- und Wirtschaftsdaten, *Basisinformationen zum Forschungsdatenmanagement, Orientierungshilfen für die Beantragung und Begutachtung datengenerierender und datennutzender Forschungsprojekte*,
http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/basisinformationen_forschungsdatenmanagement.pdf [aufgerufen am 20.11.2018]

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 Projektablauf bereitgestellt werden (z. B. mit Programmfiles zur Datenaufbereitung).

Welche Instrumente, Software, Technologien oder Verfahren werden zur Erzeugung oder Erfassung der Daten genutzt?

project/dataset/creation_methods

Datensicherheit, Angaben zu den technischen und organisatorischen Maßnahmen zur Umsetzung der Datensicherheit (Sicherung vor Verlust, Manipulation/Verfälschung, unberechtigtem Zugang)

Wer darf auf den Datensatz zugreifen?

project/dataset/data_security/access_permissions

Wie und wie oft werden Backups der Daten erstellt?

project/dataset/data_security/backups

Welche Maßnahmen zur Gewährleistung der Datensicherheit werden getroffen (z. B. Schutz vor unbefugtem Zugriff, Datenwiederherstellung, Übertragung sensibler Daten)?

project/dataset/data_security/security_measures

Wie wird die Identität von Personen, die auf die Daten zugreifen, verifiziert?

project/dataset/preservation/access_authentication

Bis wann werden die (unanonymisierten bzw. unpseudonymisierten) Originaldaten spätestens sicher vernichtet?

project/dataset/sensitive_data/personal_data/deletion

Angaben zur Einverständniserklärung

In welchem Umfang wird die "informierte Einwilligung" der Betroffenen eingeholt?

project/dataset/sensitive_data/personal_data/consent/extent

Wenn keine "informierte Einwilligung" eingeholt wird, begründen Sie dies bitte

project/dataset/sensitive_data/personal_data/consent/statement

Wo und wie sind die "informierten Einwilligungen" abgelegt?

project/dataset/sensitive_data/personal_data/consent/record

Ethikkommission

Wurde das Forschungsvorhaben von einer Ethikkommission begutachtet?

project/legal_aspects/official_approval/ethics_committee

Angaben zur methodischen, technischen und bibliographischen Dokumentation und der dabei ggf. verwendeten Standards (z. B. den technischen Formaten, den Metadatenformaten)

Welche Software, Verfahren oder Technologien sind notwendig, um die Daten zu nutzen?

project/dataset/usage_technology

Wird die Dokumentation von ggf. zur Nutzung notwendiger Software benötigt, um die Daten zu nutzen?

project/dataset/software_documentation

Welche Standards, Ontologien, Klassifikationen etc. werden zur Beschreibung der Daten und Kontextinformation genutzt?

project/dataset/metadata/standards

Werden Metadaten und Kontextinformation auf Korrektheit und Vollständigkeit geprüft?

project/dataset/metadata/quality_assurance

Die hier angeführten Fragen sind zusammen mit denen aus dem Fragenkatalog „DFG“ im Fragenkatalog „Soziologie + DFG“ vereint.

Die in der Liste *Fachspezifische Empfehlungen zum Umgang mit Forschungsdaten*⁶ verlinkte Stellungnahme der Deutschen Gesellschaft für Soziologie (DGS)⁹ ist ein Diskussionsbeitrag und enthält keine zusätzlichen Empfehlungen, mit denen sich Antragsteller in einem Datenmanagementplan auseinandersetzen müssen. Die Stellungnahme kann aber für Datenmanagementpläne als Ideenlieferant für Gründe genutzt werden, wenn die Nachnutzung sozialwissenschaftlicher Primärdaten nicht angezeigt ist.

Der Unterkatalog „Ökonomie + DFG“

Das DFG Fachkollegium 112 „Wirtschaftswissenschaften“ hat eigene Empfehlungen⁵ herausgebracht, die in der Liste „Fachspezifische Empfehlungen zum Umgang mit Forschungsdaten“ verlinkt sind⁶. Deshalb sollten Antragsteller aus den Wirtschaftswissenschaften mindestens die drei Dokumente

- *Was erwartet das Fachkollegium 112 „Wirtschaftswissenschaften“ von Antragstellenden?*⁵
- RatSWD, *Basisinformationen zum Forschungsdatenmanagement*⁸
- *Leitlinien zum Umgang mit Forschungsdaten*⁷

beachten. Da die beiden letzten Quellen schon durch den Fragenkatalog „Soziologie + DFG“ abgedeckt sind, werden an dieser Stelle nur die durch die Empfehlungen des Fachkollegiums „Wirtschaftswissenschaften“ neu hinzugekommenen Stichwörter und Fragen aufgelistet (gleiches Verfahren wie beim Katalog „Soziologie + DFG“).

Eine Mindestanforderung ist hierfür die Bereitstellung der verwendeten Programme und eine aussagekräftige Beschreibung der Datensätze, deren Ablage entweder bei den Zeitschriften selbst oder in Repositorien (an Universitäten, Forschungsinstituten oder an zentralen fachspezifischen Informationszentren) erfolgen sollte.

Welche Informationen sind für Außenstehende notwendig, um die Daten zu verstehen (d. h. ihre Erhebung bzw. Entstehung, Analyse sowie die auf ihrer Basis gewonnenen Forschungsergebnisse nachvollziehen) und nachnutzen zu können?

project/dataset/metadata/scope

Welche Komponenten der Datendokumentation stehen zusammen mit dem Datensatz zur Verfügung?

project/dataset/documentation (neues Attribut)

Wo wird die Dokumentation zur Verfügung gestellt?

project/dataset/documentation/where (neues Attribut)

Wahrung von Geschäftsgeheimnissen

Enthält dieser Datensatz nicht-personenbezogene sensible Daten?

project/dataset/sensitive_data/other/yesno

Wenn ja, um welche nicht personenbezogenen sensiblen Daten handelt es sich?

project/dataset/sensitive_data/other/description

Es könnte eine Ausschlussfrist für die Nachnutzung von Primärdaten durch Dritte von Antragstellenden festgelegt und begründet werden.

Sollen die Daten erst nach Ablauf einer Sperrfrist zugänglich gemacht werden? Falls ja, warum? Wie lang soll die Sperrfrist sein?

project/dataset/preservation/embargo_period

⁹ Deutsche Gesellschaft für Soziologie (DGS), *Bereitstellung und Nachnutzung von Forschungsdaten in der Soziologie*,
https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/dgs_stellungnahme_forschungsdaten.pdf [aufgerufen am 16.04.2019]

Der Unterkatalog „Bildungswissensch. + DFG“

Das Fachkollegium „Erziehungswissenschaft“ hat Empfehlungen veröffentlicht¹⁰, von denen einige über die hinausgehen, die in den fachübergreifenden Leitlinien⁷ der DFG formuliert sind. Daraus wurden ein Zitat und ansonsten wieder Stichwörter extrahiert und passende Fragen zugeordnet.

Bei Projektanträgen, die eine Erhebung neuer Daten beinhalten, sollte grundsätzlich geprüft werden, ob nicht bereits vergleichbare Datensätze vorliegen und verfügbar sind, die sich zur Untersuchung der Forschungsfragen eignen könnten. Sofern dies der Fall ist, wäre ggf. der Nachweis zu führen, warum eine Nachnutzung nicht sinnvoll ist und neue Daten erhoben werden sollen.

Wird der Datensatz selbst erzeugt oder nachgenutzt?

project/dataset/origin

Wenn nachgenutzt, wer hat den Datensatz erzeugt?

project/dataset/creator/name

Wenn nachgenutzt, unter welcher Adresse, PID oder URL ist der Datensatz verfügbar?

project/dataset/uri

Wenn selbst erzeugt, sind bereits existierende, ähnliche Forschungsdaten verfügbar und warum ist deren Nachnutzung hier nicht möglich bzw. sinnvoll?

project/dataset/reuse_existing (neues Attribut)

Datendokumentation

Welche Informationen sind für Außenstehende notwendig, um die Daten zu verstehen (d. h. ihre Erhebung bzw. Entstehung, Analyse sowie die auf ihrer Basis gewonnenen Forschungsergebnisse nachvollziehen) und nachnutzen zu können?

project/dataset/metadata/scope

Welche Komponenten der Datendokumentation stehen zusammen mit dem Datensatz zur Verfügung?

project/dataset/documentation

Auf Nachfrage benutzerfreundliche Bereitstellung

Welche Komponenten der Datendokumentation werden erst auf Anfrage bereitgestellt?

project/dataset/documentation_on_request (neues Attribut)

Fristen für die Datenübergabe

Sollen die Daten erst nach Ablauf einer Sperrfrist zugänglich gemacht werden? Falls ja, warum? Wie lang soll die Sperrfrist sein?

project/dataset/preservation/embargo_period

Der Spezialfall eines Forschungsprojekts, das mit bereits erhobenen Datensätzen arbeitet, die aber erst aufbereitet werden müssen¹¹, ist in diesem Fragenkatalog nicht berücksichtigt. Eine Berücksichtigung würde eine weitere Fallunterscheidung und zusätzliche Fragen erfordern. Der Fall dürfte nicht so häufig vorkommen und lässt sich mit dem Katalog „Alle Fragen“ behandeln.

Der Katalog „Alle Fragen“

Der generische Fragenkatalog „RDMO“ wurde um vier Fragen erweitert, damit er eine Obermenge der fachspezifischen Kataloge bleibt. Für die vier neuen Fragen wurden

¹⁰ Petra Stanat, *Bereitstellung und Nutzung quantitativer Forschungsdaten in der Bildungsforschung: Memorandum des Fachkollegiums „Erziehungswissenschaft“ der DFG*, https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten_bildungsforschung.pdf [aufgerufen am 16.04.2019]

¹¹ Seite 6 in [10]

entsprechend vier neue Attribute definiert. Es handelt sich dabei um die folgenden Fragen und Attribute:

Wenn selbst erzeugt, sind bereits existierende, ähnliche Forschungsdaten verfügbar und warum ist deren Nachnutzung hier nicht möglich bzw. sinnvoll?

project/dataset/reuse_existing

Diese Frage ist dem Fachkollegium „Erziehungswissenschaft“ der DFG offenbar besonders wichtig¹⁰, denn „Dieser Punkt wäre dann auch als Gegenstand der Begutachtung einzubeziehen.“ Eigentlich soll diese Frage schon im Antrag selbst erörtert werden, aber — falls Fragen des Datenmanagements in den Datenmanagementplan ausgelagert sein sollten — vergessen werden sollte die Beantwortung dieser Frage auf keinen Fall. Deshalb wurde sie hier neu aufgenommen.

Welche Komponenten der Datendokumentation stehen zusammen mit dem Datensatz zur Verfügung?

project/dataset/documentation

Eine Dokumentation wird vom Fachkollegium „Erziehungswissenschaft“ der DFG für solche Forschungsdaten gefordert, die bereitgestellt werden sollen¹⁰. Das Fachkollegium „Wirtschaftswissenschaften“ fordert sogar „die Bereitstellung der verwendeten Programme und eine aussagekräftige Beschreibung“ aller Forschungsdaten, die Grundlage einer Publikation sind⁵. Die Frage wurde deshalb aufgenommen.

Wo wird die Dokumentation zur Verfügung gestellt?

project/dataset/documentation/where

Das Fachkollegium „Wirtschaftswissenschaften“ fordert „deren Ablage entweder bei den Zeitschriften selbst oder in Repositorien (an Universitäten, Forschungsinstituten oder an zentralen fachspezifischen Informationszentren)“⁵. Aus diesem Grund wurde die Frage mit aufgenommen.

Welche Komponenten der Datendokumentation werden erst auf Anfrage bereitgestellt?

project/dataset/documentation_on_request

Das Fachkollegium „Erziehungswissenschaft“ der DFG fordert nicht, dass die Dokumentation schon vollständig vorliegt, wenn die Daten archiviert werden, sondern hält eine „auf Nachfrage nutzerfreundliche Bereitstellung direkt durch die Datenproduzenten“ für ausreichend, wenn die Forschungsdaten bei der Institution selbst archiviert werden¹⁰. Es soll hier nicht diskutiert werden, wie brauchbar eine nachträgliche Dokumentation von Forschungsdaten sein wird, aber um diese Variante im Datenmanagementplan beschreiben zu können, wurde die obige Frage mit aufgenommen.

Der erweiterte Fragenkatalog wurde „Alle Fragen“ genannt. Das soll nicht heißen, dass der Katalog nun alle Fragen enthält, die in Datenmanagementplänen irgendwann einmal auftreten und behandelt werden könnten, sondern soll Nutzern die Suche nach einem Katalog erleichtern im Sinne einer Obermengen-Untermengen-Relation. Beispielsweise ist „DFG“ ein Unterkatalog von „Bildungswissensch. + DFG“, und der wiederum ein Unterkatalog von „Alle Fragen“. Das Diagramm in der Abbildung unten soll den Zusammenhang anschaulich verdeutlichen.

Dass Katalog „Soziologie + DFG“ ein Unterkatalog von „Ökonomie + DFG“ ist, ist jedoch kein Konstruktionsprinzip, sondern hängt damit zusammen, dass aus dem Positionspapier der Wirtschaftswissenschaften⁵ zusätzliche Empfehlungen an die Antragsteller extrahiert werden konnten, aus dem Positionspapier der Soziologie⁹ jedoch nicht.

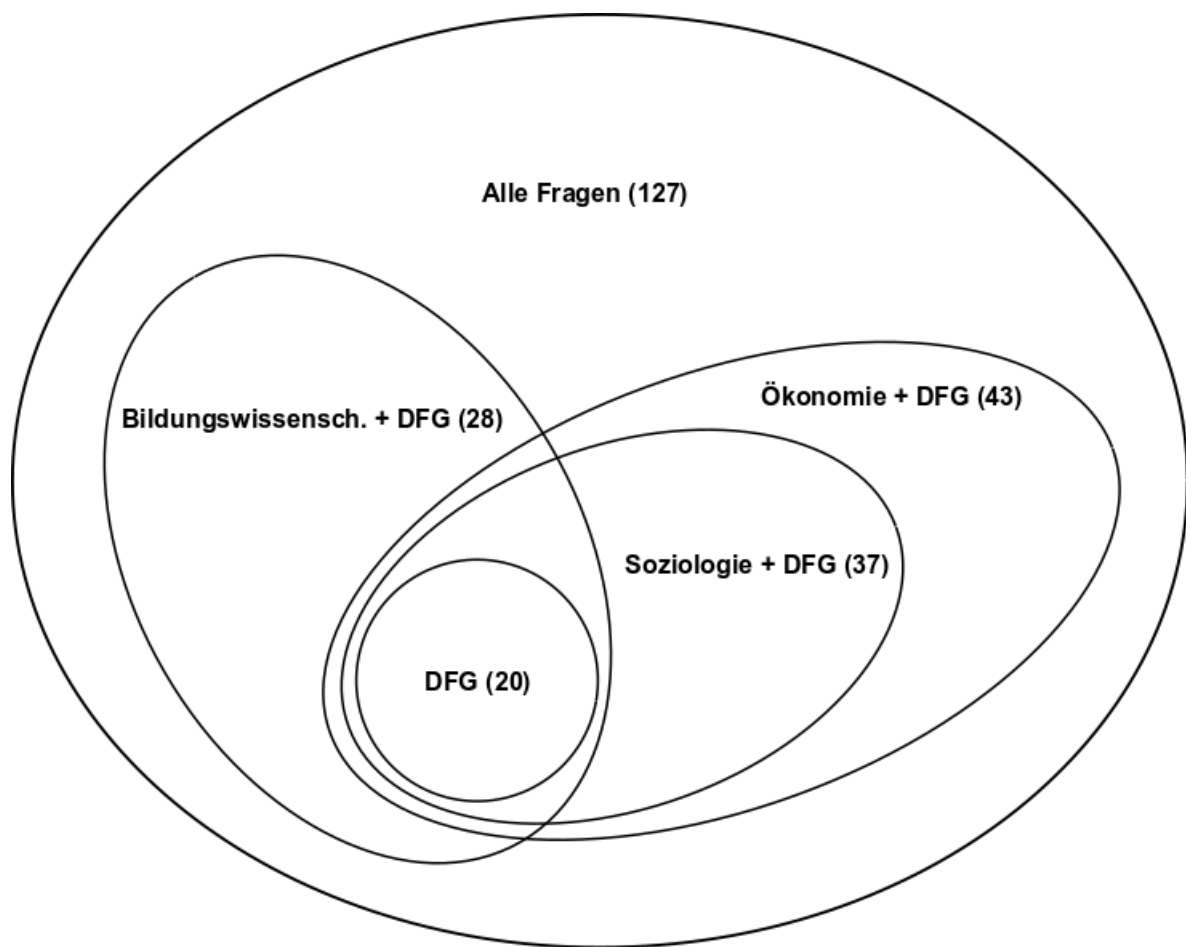


Abbildung 1: Mengendiagramm der Fragenkataloge dieses Download-Pakets. In Klammern steht die Zahl der Fragen, die der jeweilige Katalog enthält.

Vorbereitung des Katalogimports

Ergänzung der Domäne

Die Kataloge „Ökonomie + DFG“, „Bildungswissensch. + DFG“ und „Alle Fragen“ können nur genutzt werden, wenn die Domäne im RDMO erweitert wird. Die folgenden vier Attribute werden zusätzlich benötigt:

```
project/dataset/reuse_existing
project/dataset/documentation
project/dataset/documentation_on_demand
project/dataset/documentation/where
```

Gehen Sie dafür mit Admin-Rechten auf Management → Domäne. Suchen Sie Attribut `project/dataset`, klicken Sie auf das zugehörige „+“ rechts, tragen Sie als Schlüssel *reuse_existing* ein und klicken auf „Sichern“. Entsprechend ist das Vorgehen für die nächsten beiden Attribute.

Um das letzte Attribut zu erzeugen, suchen Sie das neu erzeugte Attribut `project/dataset/documentation`, klicken wieder auf das „+“ rechts und tragen als Schlüssel *where* ein.

Empfehlung: Optionenset data_protection_laws aktualisieren

Im *Bundesstatistikgesetz* sind einige Paragraphen gestrichen worden. Bestimmungen über die Nachnutzung von Zensusdaten gibt es dort nicht mehr. Die Option kann gelöscht werden.

Sozialgesetzbuch X: Das Beispiel sollte aus § 75 „Übermittlung von Sozialdaten für die Forschung und Planung“ entnommen werden. Vorschlag: „Sozialgesetzbuch X (z.B. für Daten aus dem Sozialleistungsbereich und für die wissenschaftliche Arbeitsmarkt- und Berufsforschung)“

Alte Fragenkataloge umbenennen oder löschen

Falls Sie bereits Fragenkataloge mit den Schlüsseln dfg, ratswd_dfg, economics_dfg, edu_dfg oder all haben, sollten sie die Schlüssel der schon vorhandenen Kataloge umbenennen. Durch die Umbenennung werden Ihre Kataloge weder abgehängt noch inhaltlich verändert, sondern alle Projekte, die diesen Katalog ausgewählt haben, behalten ihn unverändert.

Alternativ können Sie Kataloge aus früheren Versionen dieses Download-Pakets auch löschen. Projekte, die einen Katalog in ihren Projektinformationen ausgewählt hatten, der inzwischen gelöscht wurde, verlieren nur die Zuordnung zu einem Katalog und müssen einen neuen auswählen. Der Datenmanagementplan, d.h. alle gespeicherten Antworten, bleiben aber erhalten.

Import der Unterkataloge

Gehen Sie mit Admin-Rechten auf Management → Fragen. Dort klicken Sie in der rechten Spalte unter Punkt „Import“ auf „XML-Datei auswählen“, wählen einen Fragenkatalog als XML-Datei aus und laden diesen hoch. Falls es vor dem Hochladen schon einen Fragenkatalog mit dem gleichen Schlüssel gegeben hat, sollten Sie den Fragenkatalog nach dem Hochladen auf doppelt vorhandene Fragen prüfen.

Komponenten des Downloadfiles

dfg.xml	Fachübergreifender Katalog von Fragen, der die <i>Leitlinien zum Umgang mit Forschungsdaten</i> ⁷ abdeckt
ratswd_dfg.xml	Fragenkatalog für die Sozialwissenschaften, enthält alle Fragen aus dfg.xml ergänzt um Fragen, die zu den Empfehlungen in den <i>Basisinformationen zum Forschungsdatenmanagement</i> ⁸ des RatSWD gut passen
economics_dfg.xml	Fragenkatalog für die Wirtschaftswissenschaften, enthält alle Fragen aus ratswd_dfg.xml ergänzt um Fragen, die zu den Empfehlungen des DFG Fachkollegiums 112 „Wirtschaftswissenschaften“ ⁹ gut passen
edu_dfg.xml	Fragenkatalog für die Bildungswissenschaften, enthält alle Fragen aus dfg.xml ergänzt um Fragen, die zu den Empfehlungen des Fachkollegiums „Erziehungswissenschaft“ der DFG ¹⁰ gut passen
all.xml	Vereinigung aller 4 Fragenkataloge mit dem generischen Katalog rdmo.xml
Dokumentation.pdf	Diese Dokumentation