

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

## Version 4.2

## Dokumentation

Torsten Rathmann

FoDaKo (Forschungsdatenmanagement im Kooperationsverbund)<sup>1</sup>

Universitätsbibliothek und Zentrum für Informations- und Medienverarbeitung

Bergische Universität Wuppertal

Gaußstr. 20

42119 Wuppertal



Dieses Dokument und die beschriebenen Fragenkataloge dürfen frei verwendet werden<sup>2</sup>.

## 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)<sup>3</sup> entwickelt. Dabei wurde der im RDMO-Projekt<sup>4</sup> 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. Ausnahme ist der Fragenkatalog „DFG Biodiversitätsforschung“, denn die DFG hat für diesen Wissenschaftszweig festgelegt, dass Kosteninformationen auch im Datenmanagementplan abgelegt werden sollen.<sup>5</sup> Das vorliegende Downloadpaket des FoDaKo-Projekts enthält in dieser Weise zusammengestellte Fragenkataloge für alle Fächer, für die es spezifische Anforderungen und Empfehlungen gibt, die über

---

<sup>1</sup> <https://fodako.nrw>

<sup>2</sup> 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-Projekt<sup>3</sup> ebenfalls unter CC0-Lizenz angeboten wird, siehe <https://rdmorganiser.github.io/dokumentation/> [aufgerufen am 23.04.2019]

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

<sup>4</sup> RDMO, <https://rdmorganiser.github.io/> [aufgerufen am 16.04.2019]

<sup>5</sup> *Richtlinien zum Umgang mit Forschungsdaten in der Biodiversitätsforschung*, [https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien\\_forschungsdaten\\_biodiversitaetsforschung.pdf](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten_biodiversitaetsforschung.pdf), englisch: *Guidelines on the Handling of Research Data in Biodiversity Research*, [https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines\\_biodiversity\\_research.pdf](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_biodiversity_research.pdf) [aufgerufen am 21.08.2020]

die die fachübergreifenden *DFG-Leitlinien zum Forschungsdatenmanagement*<sup>6</sup> hinausgehen. Ein fachübergreifender, nach den *Leitlinien zum Forschungsdatenmanagement* entwickelter Katalog ist mit enthalten und der um die neu hinzugekommenen Fragen erweiterte generische Katalog, der hier den Namen „Alle Fragen“ trägt. Alle Fragenkataloge sind an RDMO 1.3 angepasst.

## Entstehung

Im DFG-geförderten Projekt RDMO<sup>4</sup> ist nicht nur die gleichnamige Software<sup>3</sup> entwickelt worden, sondern auch ein Fragenkatalog mit dem Namen „RDMO“, in den die WissGrid-Checkliste<sup>7</sup> und viele Hinweise und Überlegungen eingeflossen sind. Dieser Fragenkatalog enthält zurzeit 123 Fragen. Im Projekt FoDaKo (Forschungsdatenmanagement im Kooperationsverbund, 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 einige Nutzer und Berater verwirrt hat.

## Veränderungen gegenüber Version 4.1

Korrektur eines Fehlers in der Dokumentation. Die Fragenkataloge haben sich gegenüber Version 4.1 nicht verändert.

## Veränderungen gegenüber Version 4.0

Anpassung an RDMO 1.3, Korrektur kleinerer Fehler

## Veränderungen gegenüber Version 3.0

- Die folgenden Unterkataloge sind neu hinzugekommen:
  - DFG 101 Alte Kulturen
  - DFG 104 Mündlicher Korpus

---

<sup>6</sup> DFG, *Leitlinien zum Umgang mit Forschungsdaten*, [http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien\\_forschungsdate\\_n.pdf](http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdate_n.pdf) [aufgerufen am 09.07.2018]

<sup>7</sup> 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](https://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf) [aufgerufen am 16.04.2019]

- DFG 104 Textkorpus
- DFG 105 Edition
- DFG 106 Sozial- u. Kulturanthropologie, Außereurop. Kulturen, Judaistik u. Religionswiss.
- DFG Biodiversitätsforschung
- Unterkataloge jetzt der DFG-Fachsystematik folgend benannt
- Erweiterung des Katalogs „Alle Fragen“ um die Fragen aus den neuen Unterkatalogen, die bisher nicht enthalten waren
- Korrekturen, z.B. von nicht mehr funktionsfähigen Weblinks

### Veränderungen gegenüber Version 2.3

- Unterkatalog für die Bildungswissenschaften hinzugekommen
- Das DFG Fachkollegium 112 „Wirtschaftswissenschaften“ hat inzwischen eigene Empfehlungen<sup>8</sup> 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 abgehängten 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

---

<sup>8</sup> 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\\_de\\_1811.pdf](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/fachkollegium112_forschungsdatenmanagement_de_1811.pdf) [aufgerufen am 27.08.2020]

Unterkatalogen herausgenommen. Die beiden Fragen waren nur in den Unterkatalogen enthalten, um die Frage „Gibt es von 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.

### dfg.xml, 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 stellt<sup>9</sup>. Die DFG wurde ausgewählt, weil sie der größte Forschungsförderer in Deutschland ist und außerdem *Leitlinien zum Forschungsdatenmanagement*<sup>6</sup> 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

---

<sup>9</sup> DFG, *Umgang mit Forschungsdaten*, [http://www.dfg.de/foerderung/antrag\\_gutachter\\_gremien/antragstellende/nachnutzung\\_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.

Primärdaten, die Grundlage einer Publikation sind, langfristig (für mindestens 10 Jahre) in der eigenen Einrichtung oder in einer fachlich einschlägigen, über-regionalen Infrastruktur gespeichert werden sollen (letzte Frage).

**Welche Forschungsdaten entstehen im Verlauf des Vorhabens?**

*Um was für einen Datensatz handelt es sich?*

project/dataset/description

**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.

Für eine Reihe von Fächern hat die DFG die Richtlinien bestimmter Fachorganisationen aufgelistet<sup>9</sup>. Aus diesen wurden weitere fachspezifische Fragenkataloge entwickelt.

## **101\_dfg.xml, der Unterkatalog „DFG 101 Alte Kulturen“**

Das DFG-Fachkollegium 101 „Alte Kulturen“ hat Empfehlungen veröffentlicht<sup>10</sup>, von denen einige über die hinausgehen, die in den fachübergreifenden Leitlinien<sup>6</sup> der DFG formuliert sind. Daraus wurden Stichwörter extrahiert (**fett und kursiv**) und passende Fragen zugeordnet und ergänzt (*kursiv*). In der auf die zugeordnete Frage folgenden Zeile ist das Attribut angegeben, auf das die Frage zeigt (verkleinert). Der Unterkatalog enthält alle Fragen, die auch im Unterkatalog „DFG“ enthalten sind, erweitert um die folgenden Fragen:

### **Verantwortlichkeiten für das Datenmanagement**

Welche Personen oder Institutionen sind verantwortlich für die Projektkoordination?

project/coordination/name

---

<sup>10</sup> Fachkollegium „Alte Kulturen“, Handreichung zum Umgang mit Forschungsdaten, [https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/handreichung\\_fachkollegium\\_101\\_forschungsdaten.pdf](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/handreichung_fachkollegium_101_forschungsdaten.pdf), [aufgerufen am 21.08.2020]

### *Projektpartner*

project/partner/name

*Wer ist bei diesem Partner der/die Ansprechpartner/in für das Datenmanagement?*

project/partner/contact\_person/name

*Wer ist verantwortlich für die Pflege der PIDs und die Objektpflege (d. h. die Langzeitarchivierung des Objekts und somit dafür, dem PID-Service einen Objektumzug und die neue Adresse mitzuteilen)?*

project/dataset/pids/responsible\_person/name

***Es wird empfohlen, frühzeitig Kontakt zu Serviceeinrichtungen (Rechenzentren, Bibliotheken, Datenarchive, eScience-Center) aufzunehmen***

*Wer leistet Unterstützung in Fragen des Forschungsdatenmanagements für das Projekt?*

project/support

### ***Umfang der Daten***

*Was ist die tatsächliche oder erwartete Größe des Datensatzes?*

project/dataset/size/volume

### ***Aufbereitung***

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

project/dataset/creation\_methods

### ***Dokumentation der Daten***

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

project/dataset/documentation (neues Attribut ab 3.0)

### ***eindeutigen Referenzierung der Daten (Zitierbarkeit digitaler Daten)***

*Sollen für diesen Datensatz persistente Identifikatoren (PIDs) genutzt werden?*

project/dataset/pids/yesno

*Welches System von persistenten Identifikatoren soll genutzt werden?*

project/dataset/pids/system

*Welche (Sub-)Entitäten / Untereinheiten sollten sinnvollerweise eigene Identifikatoren erhalten? Welche dieser Identifikatoren sollten dauerhaft und zitierfähig sein?*

project/dataset/pids/subentities

### ***Erlaubnis zur Bearbeitung ... (Welche Genehmigung steht noch aus? Gibt es ggf. Fristen oder weitere Bedingungen für Erlaubniserteilung?)***

*Wird für das Forschungsvorhaben eine offizielle Genehmigung benötigt?*

project/legal\_aspects/official\_approval/statutory\_approval/yesno

*Wenn ja, um welche Genehmigung handelt es sich?*

project/legal\_aspects/official\_approval/statutory\_approval/title

*Wenn ja, welche ist die ausstellende Behörde?*

project/legal\_aspects/official\_approval/statutory\_approval/agency

*Wenn ja, was ist der Stand des Genehmigungsverfahrens?*

project/legal\_aspects/official\_approval/statutory\_approval/status (neues Attribut ab 4.0)

*Wurde der/die Rechteinhaber/in recherchiert?*

project/dataset/ipr/owner/name

*Welche Rechte wurden eingeholt bzw. was ist der Stand der Rechteeinholung?*

project/dataset/ipr/status

***Welcher Teil der Daten als archivwürdig angesehen wird***

*Auf Basis welcher Kriterien / Regeln werden die Daten zur Archivierung (nach Projektende) ausgesucht?*

project/preservation/selection\_criteria

***Verantwortung für die Archivierung und Bereitstellung nach Ende des Vorhabens. Bei der Ablage in zertifizierten Repositorien, in Archiven und Datenzentren in öffentlich-rechtlicher Trägerschaft reicht eine Erklärung der Institution aus.***

*Handelt es sich dabei um ein zertifiziertes Repository oder Datenzentrum (z.B. durch das Core Trust Seal, nestor-Siegel oder ISO 16363)?*

project/dataset/preservation/certification

*Wurden mit dem Repository oder Datenzentrum bereits angemessene Archivierungslösungen besprochen?*

project/dataset/preservation/repository\_arrangements

*Wer ist verantwortlich für die Archivierung und Bereitstellung nach Ende des Vorhabens?*

project/preservation/responsible\_person/name

## **spokencorpus\_dfg.xml, der Unterkatalog „DFG 104 Mündlicher Korpus“**

Das DFG-Fachkollegium 104 „Sprachwissenschaften“ hat spezielle Empfehlungen für korpuslinguistisch arbeitende Forschende veröffentlicht<sup>11</sup>. Einige dieser Empfehlungen gehen über die hinaus, die in den fachübergreifenden Leitlinien<sup>6</sup> der DFG formuliert sind, so dass auch dieser Fragenkatalog eine Erweiterung des Kataloges „DFG“ ist. Die Empfehlungen sind sehr detailliert, so dass zwei Fragenkataloge daraus entwickelt wurden, einer für mündliche Korpora, der andere für Textkorpora. Die Empfehlungen wurden für Fragen verwendet und/oder als Ausfüllhilfe zitiert. Außerdem hat das Fachkollegium 104 auch *Informationen zu rechtlichen Aspekten bei der Handhabung von Sprachkorpora*<sup>12</sup> herausgebracht, aus denen auch Texte als Ausfüllhilfe zitiert werden.

***Transkription und weitere Annotation (d.h. große Teile von Abschnitt 1.2<sup>11</sup>)***

*Welche Instrumente, Software, Technologien oder Verfahren werden zur Erzeugung*

---

<sup>11</sup> *Empfehlungen des DFG-Fachkollegiums 104 „Sprachwissenschaften“, Empfehlungen zu datentechnischen Standards und Tools bei der Erhebung von Sprachkorpora*, [https://www.dfg.de/download/pdf/foerderung/grundlagen\\_dfg\\_foerderung/informationen\\_fachwissenschaften/geisteswissenschaften/standards\\_sprachkorpora.pdf](https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/standards_sprachkorpora.pdf), [aufgerufen am 21.08.2020]

<sup>12</sup> *Informationen zu rechtlichen Aspekten bei der Handhabung von Sprachkorpora*, [https://www.dfg.de/download/pdf/foerderung/grundlagen\\_dfg\\_foerderung/informationen\\_fachwissenschaften/geisteswissenschaften/standards\\_recht.pdf](https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/standards_recht.pdf), englische Übersetzung: Guidelines for Building Language Corpora under German Law, [https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines\\_review\\_board\\_linguistics\\_corpora.pdf](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_review_board_linguistics_corpora.pdf) [aufgerufen am 21.08.2020]

*oder Erfassung der Daten genutzt?*

project/dataset/creation\_methods

**Ein grundständiges Backup muss gewährleistet sein**

*Wie und wie oft werden Backups der Daten erstellt?*

project/dataset/data\_security/backups

**Metadaten**

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

project/dataset/metadata/standards

**Erfolgt eine solche Bearbeitung, so ist es für Zwecke der Archivierung vorteilhaft, wenn der Bezug zur Originalaufnahme nachvollziehbar dokumentiert wird**

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

project/dataset/documentation (neues Attribut ab 3.0)

**Für die Universitäten und andere öffentlichen Stellen der Länder gelten die jeweiligen Landesdatenschutzgesetze.**

*Welches Gesetz ist bezüglich der Fragen des Datenschutzes für das Projekt maßgeblich neben den unmittelbar geltenden Teilen der EU Datenschutz-Grundverordnung?*

project/legal\_aspects/sensitive\_data/privacy\_law

**besondere Arten personenbezogener Daten**

*Enthält der Datensatz „Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen“ oder genetische oder biometrische „Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung“ (Art. 9 DSGVO)?*

project/dataset/sensitive\_data/personal\_data/bdsg\_3\_9

**schriftliche Einwilligungserklärung**

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

project/dataset/sensitive\_data/personal\_data/consent/extent

**Einige LDSGs (gültig für Universitäten, s.o.) erlauben auch eine Verarbeitung personenbezogener Daten ohne Einwilligung für ein bestimmtes Forschungsvorhaben, wenn gewisse Voraussetzungen gegeben sind**

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

project/dataset/sensitive\_data/personal\_data/consent/statement

**Rechteinhaber nicht immer vollständig bekannt oder ermittelbar**

*Wurde recherchiert, wer die Rechte innehat?*

project/dataset/ipr/owner/name

**Vervielfältigung, Bearbeitung, besonderes Persönlichkeitsrecht**

*Welche Rechte wurden eingeholt bzw. was ist der Stand der Rechteeinholung?*

project/dataset/ipr/status (neues Attribut ab 4.0)



## **textcorpus\_dfg.xml, der Unterkatalog „DFG 104 Textkorpus“**

Die Empfehlungen für Textkorpora unterscheiden sich von den Empfehlungen für mündliche Korpora so sehr, dass für beide je ein gesonderter Fragenkatalog aufgenommen worden ist. Die Vorgehensweise beim Unterkatalog „DFG 104 Textkorpus“ ist analog der beim Unterkatalog „DFG 104 Mündlicher Korpus“.

***Möglichst früh in der Planungsphase sollte Kontakt zu einem Zentrum aufgenommen werden, das mit der Lizenzierung des jeweiligen Ressourcentyps Erfahrung hat. Dies kann Hilfestellung geben oder u.U. auch die Lizenzierung übernehmen und gleichzeitig sicherstellen, dass die Lizenzbedingungen so gestaltet sind, dass es nach der Projektlaufzeit Datengrundlage und Projektergebnisse in die eigenen Kurationsprozesse aufnehmen und sie langfristig zugänglich machen kann.***

*Wer leistet Unterstützung in Fragen des Forschungsdatenmanagements für das Projekt?*

project/support

### **Standards und Tools (d.h. große Teile von Abschnitt 2.2<sup>11</sup>)**

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

project/dataset/creation\_methods

### **Empfehlungen zur Textauswahl**

*Auf Basis welcher Kriterien / Regeln werden die Daten zur Archivierung (nach Projektende) ausgesucht?*

project/preservation/selection\_criteria

***Aus Gründen der Zitierbarkeit und der Replizierbarkeit von Ergebnissen wird empfohlen, die in einem Projekt erstellten Ressourcen zu versionieren und mit einer entsprechenden Versionsangabe zu veröffentlichen.***

*Werden verschiedene Versionen des Datensatzes erzeugt?*

project/dataset/versioning\_yesno

### **Separate, dynamische Komponenten**

*Welche Versionierungsstrategie wird für diesen Datensatz angewandt?*

project/dataset/versioning\_strategy

### **Metadaten**

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

project/dataset/metadata/standards

### **Richtlinien zur Transskription, Tagset zur Annotation**

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

project/dataset/documentation (neues Attribut ab 3.0)

### **PID**

*Sollen für diesen Datensatz persistente Identifikatoren (PIDs) genutzt werden?*

project/dataset/pids/yesno

### **Annotationsschicht durch einen PID verweisbar**

Welche (Sub-)Entitäten / Untereinheiten sollten sinnvollerweise eigene Identifikatoren erhalten? Welche dieser Identifikatoren sollten dauerhaft und zitierfähig sein?

project/dataset/pids/subentities

### **Für die Universitäten und andere öffentlichen Stellen der Länder gelten die jeweiligen Landesdatenschutzgesetze.**

Welches Gesetz ist bezüglich der Fragen des Datenschutzes für das Projekt maßgeblich neben den unmittelbar geltenden Teilen der EU Datenschutz-Grundverordnung?

project/legal\_aspects/sensitive\_data/privacy\_law

### **besondere Arten personenbezogener Daten**

Enthält der Datensatz „Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen“ oder genetische oder biometrische „Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung“ (Art. 9 DSGVO)?

project/dataset/sensitive\_data/personal\_data/bdsg\_3\_9

### **schriftliche Einwilligungserklärung**

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

project/dataset/sensitive\_data/personal\_data/consent/extent

### **Einige LDSGs (gültig für Universitäten, s.o.) erlauben auch eine Verarbeitung personenbezogener Daten ohne Einwilligung für ein bestimmtes Forschungsvorhaben, wenn gewisse Voraussetzungen gegeben sind**

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

project/dataset/sensitive\_data/personal\_data/consent/statement

### **Die für Textkorpora relevanten Regelungen des Urheberrechtsgesetzes**

Wurde recherchiert, wer die Rechte innehat?

project/dataset/ipr/owner/name

Welche Rechte wurden eingeholt bzw. was ist der Stand der Rechteeinholung?

project/dataset/ipr/status (neues Attribut ab 4.0)

### **Regelung des Zugangs zu Daten durch ein Passwort ... oder über eine DFN-AAI-Authentifizierung ...**

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

project/dataset/preservation/access\_authentication

Kostenfragen wurden fortgelassen, da die Projektkostenkalkulation gewöhnlich in separaten Dokumenten aufgestellt wird, Software-Management-Fragen ebenfalls.

## **edition\_dfg.xml, der Unterkatalog „DFG 105 Edition“**

Das Fachkollegium Literaturwissenschaft der Deutschen Forschungsgemeinschaft hat Kriterien für die Förderung wissenschaftlicher Editionen als Grundlage für die Antragstellung und Begutachtung im Bereich der Literaturwissenschaft herausgebracht<sup>13</sup>. Die sollten auch im Datenmanagementplan Berücksichtigung finden. Entsprechend wurden die Fragen des Katalogs „DFG“ für literaturwissenschaftliche Editionsprojekte um die folgenden Fragen erweitert.

### ***projektspezifisch verwendeten XML-Elemente und Attribut-Wert-Paare im TEI-Header dokumentiert***

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

project/dataset/documentation (neues Attribut ab 3.0)

### ***Das Repositorium sollte durch ein DINI-Zertifikat oder ein vergleichbares Qualitätssiegel ausgezeichnet sein.***

*Handelt es sich dabei um ein zertifiziertes Repositorium oder Datenzentrum (z.B. durch das Core Trust Seal, nestor-Siegel oder ISO 16363)?*

project/dataset/preservation/certification

### ***im Projektteam keine ausreichende Expertise***

*Wer leistet Unterstützung in Fragen des Forschungsdatenmanagements für das Projekt?*

project/support

## **106\_dfg.xml, der Unterkatalog „DFG 106 Sozial- und Kulturanthropologie, Außereurop. Kulturen, Judaistik und Religionswiss.“**

Auch das Fachkollegium Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft hat unter Berücksichtigung fachspezifischer Herausforderungen Empfehlungen und Erwartungen zum Umgang mit Forschungsdaten sowohl in Anträgen als auch in geförderten Vorhaben formuliert<sup>14</sup>. Ausgehend vom Katalog „DFG“ wurde deshalb ein passender Katalog durch Hinzunahme der folgenden Fragen zusammengestellt.

### ***Erlaubnis der Rechte-Inhaber (Archive, Bibliotheken, Museen, Privatpersonen, Sammlungen) geklärt und dokumentiert***

*Welche Komponenten der Datendokumentation stehen zusammen mit dem*

---

<sup>13</sup> Informationen für Geistes- und Sozialwissenschaftler/innen, Förderkriterien für wissenschaftliche Editionen in der Literaturwissenschaft,

[https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/foerderkriterien\\_editione\\_n\\_literaturwissenschaft.pdf](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/foerderkriterien_editione_n_literaturwissenschaft.pdf) [aufgerufen am 21.08.2020]

<sup>14</sup> Handreichung des Fachkollegiums 106 Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft zum Umgang mit Forschungsdaten,

[https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/handreichung\\_fachkollegium\\_106\\_forschungsdaten.pdf](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/handreichung_fachkollegium_106_forschungsdaten.pdf) [aufgerufen am 21.08.2020]

*Datensatz zur Verfügung?*

project/dataset/documentation (neues Attribut ab 3.0)

***informierte Einwilligungserklärungen***

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

project/dataset/sensitive\_data/personal\_data/consent/extent

***Falls auf eine informierte Einwilligung verzichtet werden soll (z.B. wegen Fristenregelungen, Begrenzung der Zugänglichkeit, Anonymisierung im Erhebungsprozess oder im Anschluss), ist dies zu begründen.***

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

project/dataset/sensitive\_data/personal\_data/consent/statement

***Wann brauche ich ein Ethik-Votum?***

*Wurde das Forschungsvorhaben von einer Ethikkommission begutachtet?*

project/legal\_aspects/official\_approval/ethics\_committee

***Verfahren für die Gewährung von Zugangsmöglichkeiten für Dritte***

*Sollte die Nachnutzung dieses Datensatzes Einschränkungen unterliegen, erläutern Sie bitte wie der eingeschränkte Zugriff auf die Daten ermöglicht wird und die Gründe für die Einschränkungen.*

project/dataset/sharing/restrictions\_explanation

**edu\_dfg.xml, der Unterkatalog „DFG 109 Bildungswiss.“**

Das Fachkollegium „Erziehungswissenschaft“ hat Empfehlungen veröffentlicht<sup>15</sup>, von denen einige über die hinausgehen, die in den fachübergreifenden Leitlinien<sup>6</sup> 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 ab 3.0)

---

<sup>15</sup> 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](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten_bildungsforschung.pdf) [aufgerufen am 16.04.2019]

### **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 (neues Attribut ab 3.0)

### **Auf Nachfrage benutzerfreundliche Bereitstellung**

*Welche Komponenten der Datendokumentation werden erst auf Anfrage bereitgestellt?*

project/dataset/documentation\_on\_request (neues Attribut ab 3.0)

### **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 müssen<sup>16</sup>, 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.

## **ratswd\_dfg.xml, der Unterkatalog „DFG 111 Sozialwiss.“**

Für die Sozial- und Wirtschaftswissenschaften hat der Rat für Sozial- und Wirtschaftsdaten (RatSWD) die *Basisinformationen zum Forschungsdatenmanagement*<sup>17</sup> herausgegeben.

Die sechs Varianten und Fallunterscheidungen in den *Basisinformationen* 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?,*

---

<sup>16</sup> Seite 6 in [15]

<sup>17</sup> 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](http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/basisinformationen_forschungsdatenmanagement.pdf) [aufgerufen am 20.11.2018]

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

***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 „DFG 111 Sozialwiss.“ vereint.

Die in der Liste *Fachspezifische Empfehlungen zum Umgang mit Forschungsdaten*<sup>9</sup> verlinkte Stellungnahme der Deutschen Gesellschaft für Soziologie (DGS)<sup>18</sup> 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.

### **economics\_dfg.xml, der Unterkatalog „DFG 112 Wirtschaftswiss.“**

Das DFG Fachkollegium 112 „Wirtschaftswissenschaften“ hat eigene Empfehlungen<sup>8</sup> herausgebracht. Deshalb sollten Antragsteller aus den Wirtschaftswissenschaften mindestens die drei Dokumente

- *Was erwartet das Fachkollegium 112 „Wirtschaftswissenschaften“ von Antragstellenden?*<sup>8</sup>
- *RatSWD, Basisinformationen zum Forschungsdatenmanagement*<sup>17</sup>
- *Leitlinien zum Umgang mit Forschungsdaten*<sup>6</sup>

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 „DFG 111 Sozialwiss.“).

---

<sup>18</sup> 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](https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/dgs_stellungnahme_forschungsdaten.pdf) [aufgerufen am 16.04.2019]

***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 ab 3.0)

*Wo wird die Dokumentation zur Verfügung gestellt?*

project/dataset/documentation/where (neues Attribut ab 3.0)

### ***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

## **biodiversity\_dfg.xml, der Unterkatalog „DFG Biodiversitätsforschung“**

Die Mitglieder der AG Daten der DFG Senatskommission für Biodiversitätsforschung haben Richtlinien<sup>5</sup> herausgebracht. Daraus wurde der Fragenkatalog „DFG Biodiversitätsforschung“ entwickelt, wieder durch Erweiterung des Katalogs „DFG“. Kostenfragen sind hier mit enthalten, da Kosteninformationen in den Richtlinien ausdrücklich für den Datenmanagementplan vorgesehen sind. Der Unterkatalog hat keine DFG-Fachnummer im Namen, weil die Biodiversitätsforschung nicht als Fach in der DFG-Fachsystematik aufgeführt ist.

### ***Reproduzierbar***

*Ist der Datensatz reproduzierbar, d. h. ließe sich er sich, wenn er verloren ginge, erneut erstellen oder erheben?*

project/dataset/reproducibility

### ***Werkzeuge***

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

project/dataset/creation\_methods

***sind die Daten nur mit spezieller Software lesbar, so muss eine Dokumentation der Software erfolgen***

*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

***Dokumentation und Beschreibung der Daten; diese sollen sich an Standards orientieren***

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

project/dataset/metadata/standards

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

project/dataset/documentation (neues Attribut ab 3.0)

***wie die Daten zur Laufzeit des Projekts verwaltet, aufbewahrt und gesichert werden***

*Gibt es projektinterne Richtlinien zur einheitlichen Organisation der Daten? Wenn ja, wo sind diese festgehalten?*

project/dataset/storage/organisation\_policy

*Gibt es eine projektinterne Richtlinie zur Benennung der Daten? Wenn ja, bitte skizzieren Sie sie kurz und verlinken Sie ggf. zu einer ausführlicheren Dokumentation.*

project/dataset/storage/naming\_policy

*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

***Bezug auf Forschungsobjekte***

*Wie sind die Daten strukturiert? In welchem Verhältnis stehen die einzelnen Komponenten zueinander? In welchem Verhältnis steht der Datensatz zu anderen im Projekt erhobenen oder genutzten Datensätzen?*

project/dataset/structure

***wer außer den Antragstellern mit dem Forschungsdatenmanagement befasst ist***

*Wer leistet Unterstützung in Fragen des Forschungsdatenmanagements für das Projekt?*

project/support

**auffindbar**

*Sollen für diesen Datensatz persistente Identifikatoren (PIDs) genutzt werden?*

project/dataset/pids/yesno

**Nachnutzung ..., bzw. warum diese nur eingeschränkt möglich ist**

*Sollte die Nachnutzung dieses Datensatzes Einschränkungen unterliegen, erläutern Sie bitte die Gründe.*

project/dataset/sharing/restrictions\_explanation

**Kosten**

*Welcher Personalaufwand für das Datenmanagement entsteht im Rahmen der Erhebung, Erstellung oder Akquise der Daten im Projekt?*

project/costs/creation/personnel

*Welche Sachkosten für das Datenmanagement entstehen im Rahmen der Erhebung, Erstellung oder Akquise der Daten im Projekt?*

project/costs/creation/non\_personnel

*Welcher Personalaufwand für das Datenmanagement entsteht im Zusammenhang mit der Nutzung der Daten im Projekt?*

project/costs/usage/personnel

*Welche Sachkosten für das Datenmanagement entstehen im Zusammenhang mit der Nutzung der Daten im Projekt?*

project/costs/usage/non\_personnel

*Welcher Personalaufwand entsteht im Zusammenhang mit der Speicherung der Daten und Maßnahmen zur Datensicherheit während des Projektes?*

project/costs/storage/personnel

*Welche Sachkosten entstehen im Zusammenhang mit der Speicherung der Datensätze während des Projektes?*

project/costs/storage/non\_personnel

*Welcher Personalaufwand entsteht im Zusammenhang mit der Datendokumentation und der Erstellung der Metadaten im Projekt?*

project/costs/metadata/personnel

*Welche Sachkosten entstehen im Zusammenhang mit der Dokumentation und der Erstellung von Metadaten und Kontextinformation im Projekt?*

project/costs/metadata/non\_personnel

*Welcher Personalaufwand entsteht im Zusammenhang mit der Vergabe von persistenten Identifikatoren im Projekt?*

project/costs/pid/personnel

*Welche Sachkosten entstehen im Zusammenhang mit persistenten Identifikatoren im Projekt?*

project/costs/pid/non\_personnel

*Welcher Personalaufwand entsteht im Zusammenhang mit Urheber- oder verwandten Schutzrechten im Projekt?*

project/costs/ipr/personnel

*Welche Sachkosten entstehen im Zusammenhang mit Urheber- und verwandten Schutzrechten im Projekt?*

project/costs/ipr/non\_personnel

*Welcher Personalaufwand entsteht im Zusammenhang mit Langzeitarchivierung für dieses Projekt?*

project/costs/preservation/personnel

*Welche Sachkosten entstehen im Zusammenhang mit Langzeitarchivierung für dieses Projekt?*

project/costs/preservation/non\_personnel

*Wie werden die Kosten für das Datenmanagement im Projekt aufgebracht?*

project/costs/preservation/cover\_how

## **all.xml, der Katalog „Alle Fragen“**

Der generische Fragenkatalog „RDMO“ wurde um sieben Fragen erweitert, damit er eine Obermenge der fachspezifischen Kataloge bleibt. Für die sieben neuen Fragen wurden sechs neue Attribute definiert. Es handelt sich dabei um die folgenden Fragen und Attribute:

*Wer leistet Unterstützung in Fragen des Forschungsdatenmanagements für das Projekt?*

project/support (Attribut in der generischen Domäne von RDMO vorhanden)

In mehreren der fachspezifischen DFG-Papiere wird vorgeschlagen, Kontakt zu Serviceeinrichtungen aufzubauen oder sogar eine weitergehende Zusammenarbeit zu suchen<sup>5 10 11 13</sup>. Die Frage wurde in verallgemeinerter Form aufgenommen, um auch Unterstützung z.B. durch einen Rechtsanwalt oder ähnliches mit zu erfassen.

*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 ab 3.0)

Diese Frage ist dem Fachkollegium „Erziehungswissenschaft“ der DFG offenbar besonders wichtig<sup>15</sup>, 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 (neues Attribut ab 3.0)

Alle hier verarbeiteten fachspezifischen Handreichungen enthalten Empfehlungen zur Datendokumentation oder erwähnen sie zumindest. Eine Dokumentation wird beispielsweise vom Fachkollegium „Erziehungswissenschaft“ der DFG für solche Forschungsdaten gefordert, die bereitgestellt werden sollen<sup>15</sup>. Das Fachkollegium „Wirtschaftswissenschaften“ fordert sogar „die Bereitstellung der verwendeten Programme und eine aussagekräftige Beschreibung“ aller Forschungsdaten, die Grundlage einer Publikation sind<sup>8</sup>. Die Frage wurde deshalb aufgenommen.

*Wo wird die Dokumentation zur Verfügung gestellt?*

project/dataset/documentation/where (neues Attribut ab 3.0)

Das Fachkollegium „Wirtschaftswissenschaften“ fordert „deren Ablage entweder bei den Zeitschriften selbst oder in Repositorien (an Universitäten, Forschungsinstituten

oder an zentralen fachspezifischen Informationszentren)<sup>8</sup>. Aus diesem Grund wurde die Frage mit aufgenommen.

### *Welche Komponenten der Datendokumentation werden erst auf Anfrage bereitgestellt?*

project/dataset/documentation\_on\_request (neues Attribut ab 3.0)

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 werden<sup>15</sup>. Es soll hier nicht diskutiert werden, wie brauchbar eine nachträgliche Dokumentation von Forschungsdaten wäre, aber um diese Variante im Datenmanagementplan beschreiben zu können, wurde die obige Frage mit aufgenommen.

### *Wenn ja, was ist der Stand des Genehmigungsverfahrens?*

project/legal\_aspects/official\_approval/statutory\_approval/status (neues Attribut ab 4.0)

Laut *Handreichung zum Umgang mit Forschungsdaten* des Fachkollegiums „Alte Kulturen“ soll „die Zugänglichkeit und die Erlaubnis zur Bearbeitung im Sinn des Forschungsvorhabens vor der Antragstellung verbindlich geklärt sein“. „Offene Fragen oder besondere Herausforderungen müssen ausdrücklich erläutert werden (z.B. welche Genehmigung steht noch aus? Gibt es ggf. Fristen oder weitere Bedingungen zur Erlaubniserteilung?)“<sup>10</sup>. Die Frage scheint also wichtig zu sein und wurde daher aufgenommen.

### *Welche Rechte wurden eingeholt bzw. was ist der Stand der Rechteeinholung?*

project/dataset/ipr/status (neues Attribut ab 4.0)

Diese Frage ist bei sprachwissenschaftlichen Korpora von Bedeutung, da das Urheberrecht mehrere Einzelrechte umfasst und zudem alle Urheber ihre Zustimmung zur Nachnutzung geben müssen<sup>12</sup>. Durch die Aufnahme der Frage soll die Übersicht erleichtert werden.

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 „DFG 109 Bildungswiss.“, und der wiederum ein Unterkatalog von „Alle Fragen“. Das Diagramm in Abbildung 1 soll den Zusammenhang anschaulich verdeutlichen.

Dass der Katalog „DFG 111 Sozialwiss.“ ein Unterkatalog von „DFG 112 Wirtschaftswiss.“ ist, ist jedoch kein Konstruktionsprinzip, sondern hängt damit zusammen, dass aus dem Positionspapier der Wirtschaftswissenschaften<sup>8</sup> zusätzliche Empfehlungen an die Antragsteller extrahiert werden konnten, aus dem Positionspapier der Soziologie<sup>18</sup> jedoch nicht.

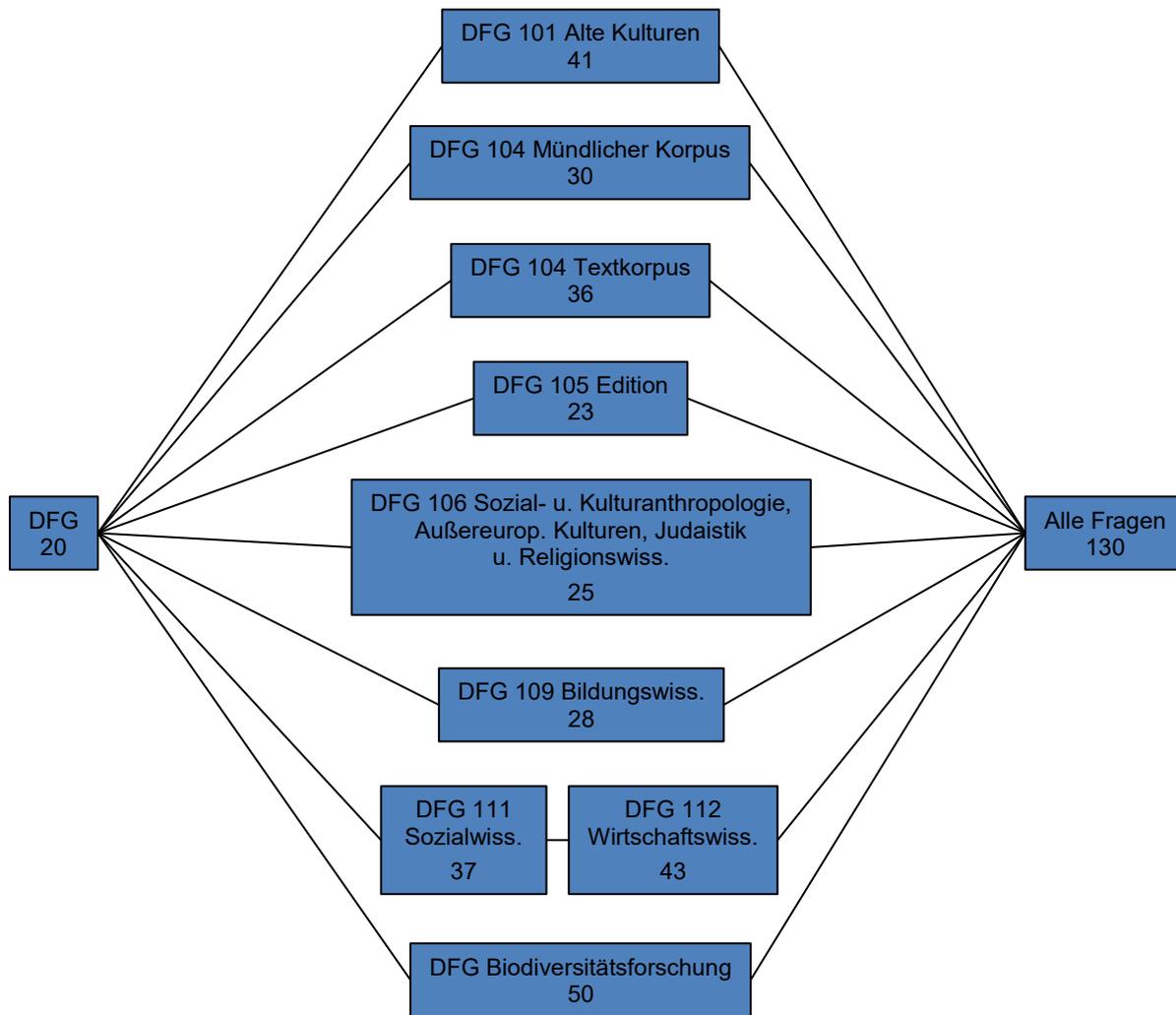


Abbildung 1: Übersicht zu den Fragenkatalogen dieses Download-Pakets. Unter dem Namen steht die Zahl der Fragen, die der jeweilige Katalog enthält. Der schwarze Keil soll die von links nach rechts zunehmende Zahl an Fragen symbolisieren. Eine Verbindungslinie steht für eine Obermengen-Untermengen-Relation, d.h. der Katalog am rechten Ende der Linie enthält immer auch alle Fragen des Katalogs am linken Ende.

## Vorbereitung des Katalogimports

### Ergänzung der Domäne

Nur die Kataloge „DFG“ und „DFG 111 Sozialwiss.“ können ohne Erweiterung der Domäne genutzt werden. Für alle anderen werden zusätzliche Attribute benötigt:

```

project/dataset/reuse_existing
project/dataset/documentation
project/dataset/documentation_on_request
project/dataset/documentation/where
project/legal_aspects/official_approval/statutory_approval/status
project/dataset/ipr/status

```

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 weiteren Attribute.

### **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, 101, spokencorpus\_dfg, textcorpus\_dfg, edition\_dfg, 106\_dfg, ratswd\_dfg, economics\_dfg, edu\_dfg, biodiversity\_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, bleibt aber erhalten.

### **Import der Unterkataloge**

Gehen Sie mit Admin-Rechten in 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.

### **Fehlerbehebung**

#### **„Server Error (500)“ nach Import des Fragenkatalogs**

Beim Versuch, sich in einem Projekt die Antworten anzuschauen, wird nur die Fehlermeldung „Server Error (500)“ ausgegeben. Möglicherweise ist die Domäne unvollständig. Sie können das leicht prüfen, indem Sie mit Administratorrechten in Management → Fragen gehen und sich den Fragenkatalog ansehen, der im betroffenen Projekt gewählt ist. Fragen mit fehlendem Attribut sind mit einem Ausrufezeichen in einem roten Dreieck markiert, siehe Abbildung 2.

**Frage** Wenn ja, was ist der Stand des Genehmigungsverfahrens?



[101\\_dfg/legal-and-ethics/sensitive-data-official\\_approval/status](#) →

**Abbildung 2: Frage mit fehlendem Attribut in der Management-Anzeige**

Ergänzen Sie in diesem Fall das fehlende Attribut erst in der Domäne und tragen es anschließend in das Attribut-Feld der Frage ein. Sie können den Fragenkatalog auch erneut installieren, wichtig ist aber, dass die Domäne vorher ergänzt worden ist.