

Katalog RDMO

Abschnitt Allgemein

Unterabschnitt Thema

Fragenset

Fragen

- **Text:** Wie lautet die primäre Forschungsfrage des Projektes?
- **Text:** Bitte geben Sie einige Schlagworte zur Forschungsfrage an.

Frage

Text Welcher Disziplin / welchen Disziplinen ist das Projekt zuzuordnen?

Hilfe Die Liste der Disziplinen entspricht der http://www.dfg.de/dfg_profil/gremien/fachkollegien/faecher/ target=_blank">Fachsystematik der Deutschen Forschungsgemeinschaft (DFG).

Optionen

- Geistes- und Sozialwissenschaften / Alte Kulturen
- Geistes- und Sozialwissenschaften / Geschichtswissenschaften
- Geistes- und Sozialwissenschaften / Kunst-, Musik-, Theater- und Medienwissenschaften
- Geistes- und Sozialwissenschaften / Sprachwissenschaften
- Geistes- und Sozialwissenschaften / Literaturwissenschaft
- Geistes- und Sozialwissenschaften / Außereuropäische Sprachen und Kulturen, Sozial- und Kulturanthropologie, Judaistik und Religionswissenschaft
- Geistes- und Sozialwissenschaften / Theologie
- Geistes- und Sozialwissenschaften / Philosophie
- Geistes- und Sozialwissenschaften / Erziehungswissenschaft
- Geistes- und Sozialwissenschaften / Psychologie
- Geistes- und Sozialwissenschaften / Sozialwissenschaften
- Geistes- und Sozialwissenschaften / Wirtschaftswissenschaften
- Geistes- und Sozialwissenschaften / Rechtswissenschaften
- Lebenswissenschaften / Grundlagen der Biologie und Medizin
- Lebenswissenschaften / Pflanzenwissenschaften
- Lebenswissenschaften / Zoologie
- Lebenswissenschaften / Mikrobiologie, Virologie und Immunologie
- Lebenswissenschaften / Medizin
- Lebenswissenschaften / Neurowissenschaft
- Lebenswissenschaften / Agrar-, Forstwissenschaften, Gartenbau und Tiermedizin
- Naturwissenschaften / Molekülchemie
- Naturwissenschaften / Chemische Festkörper- und Oberflächenforschung
- Naturwissenschaften / Physikalische und Theoretische Chemie
- Naturwissenschaften / Analytik, Methodenentwicklung (Chemie)
- Naturwissenschaften / Biologische Chemie und Lebensmittelchemie
- Naturwissenschaften / Polymerforschung
- Naturwissenschaften / Physik der kondensierten Materie

- Naturwissenschaften / Optik, Quantenoptik und Physik der Atome, Moleküle und Plasmen
- Naturwissenschaften / Teilchen, Kerne und Felder
- Naturwissenschaften / Statistische Physik, Weiche Materie, Biologische Physik, Nichtlineare Dynamik
- Naturwissenschaften / Astrophysik und Astronomie
- Naturwissenschaften / Mathematik
- Naturwissenschaften / Atmosphären- und Meeresforschung
- Naturwissenschaften / Geologie und Paläontologie
- Naturwissenschaften / Geophysik und Geodäsie
- Naturwissenschaften / Geochemie, Mineralogie und Kristallographie
- Naturwissenschaften / Geographie
- Naturwissenschaften / Wasserforschung
- Ingenieurwissenschaften / Produktionstechnik
- Ingenieurwissenschaften / Mechanik und Konstruktiver Maschinenbau
- Ingenieurwissenschaften / Verfahrenstechnik, Technische Chemie
- Ingenieurwissenschaften / Wärmeenergie-technik, Thermische Maschinen, Strömungsmechanik
- Ingenieurwissenschaften / Werkstofftechnik
- Ingenieurwissenschaften / Materialwissenschaft
- Ingenieurwissenschaften / Systemtechnik
- Ingenieurwissenschaften / Elektrotechnik
- Ingenieurwissenschaften / Informatik
- Ingenieurwissenschaften / Bauwesen und Architektur

Kommentar for Germany, the classification works, but we might consider also DDC as an option

Unterabschnitt Projektablauf

Fragenset

Fragen

- **Text:** Wann beginnt die Projektlaufzeit?
- **Text:** Wann endet die Projektlaufzeit?

Unterabschnitt Projektpartner

Frage

Text Welche Personen oder Institutionen sind verantwortlich für die Projektkoordination?

Fragenset

Fragen

- **Text:** Projektpartner
- **Text:** 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.

Hilfe: Immer mehr Hochschulen und wissenschaftliche Einrichtungen verabschieden Leitlinien oder Richtlinien (oft auch als „Policies“ bezeichnet) zum Forschungsdatenmanagement. Diese enthalten unter anderem Empfehlungen und/oder Vorgaben zum Umgang mit Forschungsdaten, die die Wissenschaftlerinnen und Wissenschaftler der Einrichtung beachten sollten oder müssen (je nach Grad der Verbindlichkeit). Beispiele für solche Richtlinien sind die [>Grundsätze zu Forschungsdaten der Universität Bielefeld oder die \[>Leitlinien zum Umgang mit Forschungsdaten der Georg-August-Universität Göttingen.\]\(http://www.uni-goettingen.de/de/01-juli-2014-forschungsdaten-leitlinie-der-universitaet-goettingen-einschl-umg/488918.html\)](https://data.uni-bielefeld.de/de/policy)

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

Hilfe: Bitte geben Sie den Namen und eine Email Adresse an.

Unterabschnitt Förderung

Fragenset

Fragen

- **Text:** Wer fördert das Projekt?
- **Text:** In welcher Förderlinie und/oder welchem Förderprogramm wird das Projekt gefördert?
- **Text:** Gibt es von Seiten des Forschungsförderers Vorgaben oder Richtlinien bezüglich des Umgangs 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.

Hilfe: Auch Forschungsförderer stellen zunehmend Anforderungen an das Datenmanagement in von ihnen geförderten Projekten. Beispiele sind die [>DFG-Leitlinien zum Umgang mit Forschungsdaten oder die \[>Guidelines on Data Management in Horizon 2020 der Europäischen Kommission.\]\(http://ec.europa.eu/research-oa-data-mgt_en.pdf\)](http://www.dfg.de/foerderung/antragstellung_begutachtung_e)

Unterabschnitt Weitere Anforderungen

Frage

Text Gibt es von weiteren Seiten (z. B. von der Fachcommunity) Anforderungen an das Datenmanagement, die beachtet werden müssen?

Hilfe Beispiele für fachspezifische Empfehlungen und Richtlinien sind: - [>Richtlinien zum Umgang mit Forschungsdaten in der Biodiversitätsforschung - \[>Empfehlungen zur Bereitstellung und Nutzung quantitativer Daten in der Bildungsforschung - \\[>Bereitstellung und Nutzung quantitativer Daten in der Bildungsforschung - \\\[>Förderkriterien für wissenschaftliche Editionen in der Literaturwissenschaft - \\\\[>Empfehlungen zu datentechnischen Standards und Tools sowie zu - \\\\\[>Empfehlungen zu datentechnischen Standards und Tools sowie zu - \\\\\\[>rechtlichen Fragen bei der Erhebung von Sprachkorpora\\\\\\]\\\\\\(http://www.dfg.de/download/pdf/foerderung/rechtliche_fragen\\\\\\)\\\\\]\\\\\(http://www.dfg.de/download/pdf/foerderung/rechtliche_fragen\\\\\)\\\\]\\\\(http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachliche_standards_tools\\\\)\\\]\\\(http://www.dfg.de/download/pdf/foerderung/antragstellung_forschungsdaten/foerderkriterien_editionen_literaturwissenschaft\\\)\\]\\(http://www.dfg.de/download/pdf/foerderung/antragstellung_forschungsdaten/foerderkriterien_editionen_literaturwissenschaft\\)\]\(http://www.dfg.de/download/pdf/foerderung/antragstellung_forschungsdaten/richtlinien_forschungsdaten\)](http://www.dfg.de/download/pdf/foerderung/richtlinien_forschungsdaten)

Optionen

- Ja
- Nein
- Noch zu klären

Frage

Text Welche Anforderungen an das Datenmanagement sind dies?

Hilfe Bitte skizzieren Sie sie kurz und verweisen Sie ggf. auf weiterführende Informationen. Geben Sie bitte auch an, welchen Grad an Verbindlichkeit sie haben.

Abschnitt Inhaltliche Einordnung

Unterabschnitt Daten

Fragenset

Hilfe Die nächsten Fragen dienen zur Beschreibung der Datensätze, die im Projekt erzeugt und/oder verwendet werden. Sie helfen zudem, den Wert der Daten hinsichtlich der potentiellen Nachnutzung und einer späteren Archivierung einschätzen zu können. Vor der Erzeugung von Daten empfiehlt es sich zu prüfen, ob bereits vorhandene Daten nachgenutzt werden können. Die Vermeidung doppelter Erhebungen spart Aufwand und Kosten. Handelt es sich um personenbezogene Daten, ist dies zudem durch den Grundsatz der Erforderlichkeit ([>BDSG\) sowie das Prinzip der Datenvermeidung und Datensparsamkeit \(\[>§3a BDSG\\) geboten. Die Angaben zu den im Projekt erzeugten oder verwendeten Daten sind nach „Datensätzen“ strukturiert. Die Definition dessen, was jeweils ein Datensatz ist, ist eine wichtige konzeptionelle Entscheidung, die für jedes Vorhaben bzw. Projekt individuell getroffen und sorgfältig abgewogen werden muss.\]\(https://www.gesetze-im-internet.de/bdsg_1990/__3a.html\)](https://www.gesetze-im-internet.de/bdsg_1990/)

Fragen

- **Text:** Um was für einen Datensatz handelt es sich?

Hilfe: Bitte beschreiben Sie hier kurz, um welchen Datentyp es sich handelt und mit welcher Methode die Daten erhoben oder erstellt wurden, z.B.: * quantitative Online-Befragung * 3D-Modellierung / digitale Rekonstruktion einer steinzeitlichen Siedlung * Software, die im Projekt entwickelt wird

Fragenset

Fragen

- **Text:** Wird der Datensatz selbst erzeugt oder nachgenutzt?

Optionen:

- Erzeugt
- Nachgenutzt

- **Text:** Wenn nachgenutzt, wer hat den Datensatz erzeugt?
- **Text:** Wenn nachgenutzt, unter welcher Adresse, PID oder URL ist der Datensatz verfügbar?

Fragenset

Fragen

- **Text:** Für welche Personen, Gruppen oder Institutionen könnte dieser Datensatz (für die Nachnutzung) von Interesse sein? Für welche Szenarien ist dies denkbar?

Fragenset

Fragen

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

Hilfe: Manche Daten können im Prinzip jederzeit neu erstellt werden. Beispiele hierfür sind etwa naturwissenschaftliche Experimentdaten oder auch Digitalisate analoger Objekte (solange die Originale nicht verlorengehen). Der Aufwand und die Kosten hierfür können natürlich durchaus beträchtlich sein. Im Hinblick auf die Frage der Notwendigkeit einer späteren Langzeitarchivierung sollte in diesen Fällen der Aufwand einer erneuten Erstellung gegen den Aufwand der Langzeitarchivierung abgewogen werden. Andere Daten wiederum lassen sich per se nicht erneut erheben. Dies ist etwa bei jeglicher Art von episodischer Beobachtungen, sei es sozialwissenschaftlicher oder naturwissenschaftlicher Art, der Fall, da diese ein bestimmtes Phänomen an einem bestimmten Zeitpunkt und/oder Ort abbilden und somit i.d.R. nicht wiederholbar sind. Ihr Wert für die Nachnutzung durch andere wie auch der Verlust bei einer nicht erfolgten oder misslungenen Langzeitarchivierung ist ungleich höher als bei reproduzierbaren Daten.

Optionen:

- ja, mit geringem Aufwand
- ja, mit mäßigem, aber vertretbarem Aufwand
- nein bzw. nur mit unverhältnismäßig großem Aufwand
- nein, die Daten sind per se nicht reproduzierbar

Abschnitt Technische Einordnung

Unterabschnitt Daten

Fragenset

Fragen

- **Text:** Wann beginnt die Erhebung bzw. Erstellung der Daten?
- **Text:** Wann endet die Erhebung bzw. Erstellung der Daten?
- **Text:** Wann beginnt die Datenbereinigung / -aufbereitung?
- **Text:** Wann endet die Datenbereinigung / -aufbereitung?
- **Text:** Wann beginnt die Datenanalyse?
- **Text:** Wann endet die Datenanalyse?

Fragenset

Fragen

- **Text:** Was ist die tatsächliche oder erwartete Größe des Datensatzes?

Optionen:

- weniger als ein GB

- 1 GB bis 1 TB
 - 1 TB bis 100 TB
 - mehr als 100 TB
 - genau: _____
 - noch nicht bestimmt
- **Text:** Wie hoch ist die erwartete Erzeugungsrate der Daten pro Jahr?
- Hilfe:** Optional. Dies ist nur relevant, wenn das Wachstum die TB-Größenordnung erreicht.

Fragenset

Fragen

- **Text:** In welchen Formaten liegen die Daten vor?
- Hilfe:** Bei der Wahl des Dateiformates sollten auch die Konsequenzen für die kollaborative Nutzung, die Langzeitarchivierung sowie die Nachnutzung beachtet werden. Es empfiehlt sich, möglichst standardisierte, nicht-proprietäre und allgemein bzw. in der spezifischen Community verbreitete Formate zu nutzen. Weitere Kriterien sowie detaillierte Erläuterungen sind z.B. im http://www.wissgrid.de/publikationen/Leitfaden_Data-Management-WissGrid.pdf target=_blank">WissGrid-Leitfaden, S. 22 f.) zu finden.

Fragenset

Fragen

- **Text:** Welche Instrumente, Software, Technologien oder Verfahren werden zur Erzeugung oder Erfassung der Daten genutzt?
- Hilfe:** Diese Informationen sind für alle Arten von Daten relevant, um ihre Genese nachvollziehen zu können. Bei reproduzierbaren Daten kommt ein weiterer Aspekt hinzu. Diese müssen nicht notwendigerweise aufbewahrt werden - allerdings müssen alle Geräte, Software und auch Informationen über die Vorgehensweise erhalten bleiben, die notwendig sind, um die Daten erneut erstellen zu können.
- **Text:** Welche Software, Verfahren oder Technologien sind notwendig, um die Daten zu nutzen?
- Hilfe:** Um Daten nachnutzen zu können, bspw. für die Replikation von Studien, Metaanalysen oder die Beantwortung neuer Forschungsfragen, werden neben den Daten selbst auch die Software, Geräte etc. und das Wissen über spezielle Verfahren zur Nutzung benötigt. Ebenso wie bei den Formaten gilt hier: je standardisierter, offener und etablierter diese sind, desto einfacher ist i.d.R. eine Nachnutzung möglich.
- **Text:** Wird die Dokumentation von ggf. zur Nutzung notwendiger Software benötigt, um die Daten zu nutzen?

Fragenset

Fragen

- **Text:** Werden verschiedene Versionen des Datensatzes erzeugt?
 - **Text:** Welche Versionierungsstrategie wird für diesen Datensatz angewandt?
- Hilfe:** Bitte beschreiben Sie hier kurz projektinterne Regelungen zur Versionierung von Datensätzen (z.B.: Welche Änderungen machen eine neue Version erforderlich? Wie werden die Änderungen dokumentiert? Wie werden die verschiedenen Versionen benannt?)

- **Text:** Welche Technologie bzw. welches Tool wird zur Versionierung verwendet?

Optionen:

- Einfaches Kopieren
- Versionskontrollsystem: _____
- Sonstiges: _____
- Noch nicht entschieden

Abschnitt Datennutzung

Unterabschnitt Nutzungsszenarien

Fragenset

Hilfe Die folgenden Fragen dienen dazu einzuschätzen, welche Ressourcen für die geplante Nutzung der Daten während der Projektlaufzeit nötig sind. Dabei kann es sich um technische bzw. IT-Ressourcen handeln, aber auch um Expertise, die z.B. durch Datenmanagement- oder IT-ExpertInnen eingebracht wird.

Fragen

- **Text:** Wozu / wie wird dieser Datensatz während des Projektes genutzt?
- **Text:** Wie häufig wird dieser Datensatz genutzt?
- **Text:** In welchem Umfang werden Infrastrukturressourcen benötigt (CPU-Stunden, Bandbreite, Speicherplatz etc.)?

Optionen:

- Die üblichen Infrastrukturressourcen des Arbeitsplatzes reichen aus.
 - Es werden folgende Infrastrukturressourcen benötigt: _____
 - **Text:** Gibt es beabsichtigte (ggf. auch potentielle) Nutzungsszenarien, für die die Unterstützung durch Datenmanagement- oder IT-ExpertInnen sinnvoll oder notwendig ist?
- Optionen:**
- Ja: _____
 - Nein

Unterabschnitt Datenspeicherung und -sicherheit

Fragenset

Fragen

- **Text:** Wo wird der Datensatz während des Projektes gespeichert?
- **Text:** Unter welcher URL kann der Datensatz während des Projektes abgerufen werden?
- **Text:** Gibt es projektinterne Richtlinien zur einheitlichen Organisation der Daten? Wenn ja, wo sind diese festgehalten?

Optionen:

- Ja: _____

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

Optionen:

- Ja: _____
- Nein
- Noch nicht

Fragenset

Fragen

- **Text:** Wer darf auf den Datensatz zugreifen?
- **Text:** Wie und wie oft werden Backups der Daten erstellt?
Hilfe: Die Frage bezieht sich auf Backups während der Zeit, in denen mit den Daten gearbeitet wird. Fragen der Langzeitarchivierung werden gesondert im entsprechenden Abschnitt behandelt.
- **Text:** Wer ist verantwortlich für die Erstellung der Backups?
Hilfe: Die Frage bezieht sich auf Backups während der Zeit, in denen mit den Daten gearbeitet wird. Fragen der Langzeitarchivierung werden gesondert im entsprechenden Abschnitt behandelt.
- **Text:** Welche Maßnahmen zur Gewährleistung der Datensicherheit werden getroffen (z. B. Schutz vor unbefugtem Zugriff, Datenwiederherstellung, Übertragung sensibler Daten)?

Unterabschnitt Weitergabe und Veröffentlichung

Fragenset

Fragen

- **Text:** Ist der Datensatz interoperabel, d.h. geeignet für den Datenaustausch und die Nachnutzung zwischen bzw. von unterschiedlichen Forschenden, Institutionen, Organisationen und Ländern?

Optionen:

- Für den Datensatz werden standardisierte Formate genutzt: _____
- Der Datensatz ist mit verfügbarer (offener) Software bzw. mit in der jeweiligen Community etablierter und vielgenutzter Software nutzbar: _____
- Der Datensatz kann ohne großen Aufwand mit verschiedenen anderen Datensätzen aus unterschiedlichen Quellen rekombiniert werden: _____
- Weitere für die Interoperabilität wichtige Aspekte: _____

Kommentar: wortlaut noch relativ nah an Original-H2020-Frage

Fragenset

Fragen

- **Text:** Soll dieser Datensatz veröffentlicht oder geteilt werden?

Optionen:

- Ja, intern mit allen, solange sie die Daten nicht veröffentlichen oder nach außen weitergeben
- Ja, extern in begrenztem Umfang mit individueller Freigabe
- Ja, extern für alle
- Nein

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

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

Hilfe: Die Auswahlmöglichkeiten orientieren sich an Lizenzen der <http://de.creativecommons.org/was-ist-cc/> target=_blank">Creative-Commons-Familie.

Optionen:

- Namensnennung (BY)
- keine kommerzielle Nutzung (NC)
- keine Bearbeitung (ND)
- Weitergabe unter gleichen Bedingungen (SA)
- Gemeinfrei (CC0)
- Andere: _____

- **Text:** Sollte die Nachnutzung dieses Datensatzes Einschränkungen unterliegen, erläutern Sie bitte die Gründe.
- **Text:** Wann werden die Daten veröffentlicht?

Unterabschnitt Kollaboratives Arbeiten

Fragenset

Fragen

- **Text:** Werden die Daten kollaborativ genutzt?

Optionen:

- Ja, von mehreren Personen an verschiedenen Institutionen
- Ja, von mehreren Personen derselben Arbeitsgruppe an derselben Institution
- Nein

- **Text:** Welche Plattform, welche Werkzeuge werden zum kollaborativen Arbeiten an Daten und Publikationen genutzt?

- **Text:** Wie ist das kollaborative Arbeiten an denselben Dateien geregelt?

Unterabschnitt Qualitätssicherung

Fragenset

Fragen

- **Text:** Welche Maßnahmen zur Qualitätssicherung werden für diesen Datensatz ergriffen?

Frage

Text Wird die Integration zwischen nachgenutzten und erzeugten Daten gewährleistet? Wenn ja, wie?

Kommentar the stipulation that re-used and created data are of the same type (= and thus capable to be integrated) is problematic, since this implies properties of the data that are not a given. May be this should be reformulated to : Is the connection between re-used and created data ensures (which goes into the provenance realm)

Unterabschnitt Kosten

Fragenset

Fragen

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

Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.

Bereich:

- Minimum: 0,0
- Maximum: 12,0
- Schrittgröße: 0,1

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

Hilfe: Bitte schätzen sie die Kosten in ****Euro****.

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

Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.

Bereich:

- Minimum: 0,0
- Maximum: 12,0
- Schrittgröße: 0,1

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

Hilfe: Bitte schätzen sie die Kosten in ****Euro****.

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

Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.

Bereich:

- Minimum: 0,0
- Maximum: 12,0
- Schrittgröße: 0,1

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

Hilfe: Bitte schätzen sie die Kosten in ****Euro****.

Abschnitt Metadaten und Referenzierung

Unterabschnitt Metadaten

Fragenset

Fragen

- **Text:** 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?

Optionen:

- Ort
- Inhalt
- Methodik
- Erzeugungsprozess
- Technologie
- Dokumentation der zur Nutzung notwendigen Software
- Zeit
- Quellen
- Akteure
- Identifikatoren
- Andere: _____

Kommentar: probably we should introduce a separate question for being findable and re-used

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

Optionen:

- Es werden disziplinspezifische Standards, Klassifikationen etc. genutzt: _____
- Es wird ein eigenes Beschreibungssystem genutzt (bitte beschreiben Sie dieses kurz und geben, wenn nötig, an, wo es ausführlicher dokumentiert ist): _____
- Es wird kein festgelegtes System zur Beschreibung genutzt
- Sonstiges: _____
- Es wurde noch nicht entschieden, mit welchem System die Metadaten und Kontextinformationen beschrieben werden

- **Text:** Welche Metadaten werden automatisch erhoben?
- **Text:** Sollte es unvermeidbar sein, projektspezifische oder seltene Ontologien, Metadatenschemata oder Vokabulare zu nutzen, werden Mappings zu gängigen Ontologien etc. erstellt?
Hilfe: Diese Information wird für einen Horizon 2020 Datenmanagementplan benötigt.
Kommentar: original H2020 question
- **Text:** Welche Metadaten werden semi-automatisch erhoben?
- **Text:** Welche Metadaten werden manuell erhoben?
- **Text:** Werden Metadaten und Kontextinformation auf Korrektheit und Vollständigkeit geprüft?
Optionen:
 - Automatische Prüfung auf Vollständigkeit
 - Manuelle Prüfung auf Korrektheit
 - Manuelle Prüfung auf Vollständigkeit
 - Sonstiges: _____
- **Text:** Wer ist verantwortlich für die Dokumentation und Prüfung der Metadaten und Kontextinformationen auf Richtigkeit und Vollständigkeit?

Fragenset

Fragen

- **Text:** Welcher Personalaufwand entsteht im Zusammenhang mit der Erstellung von Metadaten und Kontextinformation im Projekt?
Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.
Bereich:
 - Minimum: 0,0
 - Maximum: 12,0
 - Schrittgröße: 0,1
- **Text:** Welche Sachkosten entstehen im Zusammenhang mit der Erstellung von Metadaten und Kontextinformation im Projekt?
Hilfe: Bitte schätzen sie die Kosten in ****Euro****.

Unterabschnitt Objektstruktur, Granularität und Referenzierung

Fragenset

Fragen

- **Text:** 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?

Fragenset

Hilfe Persistente Identifikatoren (PIDs) sollen die dauerhafte Referenzierung von (insbesondere) digitalen Objekten wie Publikationen oder Forschungsdaten ermöglichen. Statt, wie i.d.R. bei der Angabe eines Hyperlinks als Referenz der Fall, direkt auf den Speicherort des Objektes zu verweisen, fungiert die PID als eine Zwischeninstanz, von der aus zum Objekt weitergeleitet wird (dies nennt man „Auflösen“ der PID). Die PID bleibt gleich, auch wenn der Speicherort des Objektes sich ändert. Während ein Hyperlink in diesem Fall ins Nichts führen würde, ist das Objekt über die PID weiterhin erreichbar. Mehr Informationen zur Funktionsweise, Verwendung und verschiedenen Arten von PIDs gibt es z.B. in einem vom Projekt DASISH erstellten [Online-Tutorial](http://training.dasish.eu/training/3/) oder in den [Informationsmaterialien des Australian National Data Service \(ANDS\)](http://www.andis.org.au/guides#identify) (nach unten scrollen zu "Identifying data and researchers").

Fragen

- **Text:** Sollen für diesen Datensatz persistente Identifikatoren (PIDs) genutzt werden?
- **Text:** Welches System von persistenten Identifikatoren soll genutzt werden?

Optionen:

- Handle / DOI
 - PURL
 - ARK
 - URN
 - ISLRN
 - Anderes: _____
- **Text:** Welche (Sub-)Entitäten / Untereinheiten sollten sinnvollerweise eigene Identifikatoren erhalten? Welche dieser Identifikatoren sollten dauerhaft und zitierfähig sein?
 - **Text:** 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)?

Hilfe: Die Voraussetzung dafür, dass persistente Identifikatoren auch wirklich halten, was sie versprechen, ist, dass sie - wie auch die Objekte, auf die referenziert wird - dauerhaft und verlässlich gepflegt werden. D.h. beispielsweise, dass bei einer Änderung des Speicherortes des Objekts die Information über den neuen Speicherort entsprechend aktualisiert werden muss. Werden die Daten in einem Datenzentrum oder -repository gespeichert, wird diese Verantwortung i.d.R. von diesem übernommen. Um sicher zu sein, sollten die Verantwortlichkeiten jedoch in jedem Fall geprüft bzw. geklärt werden.

Fragenset

Fragen

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

Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.

Bereich:

- Minimum: 0,0
- Maximum: 12,0
- Schrittgröße: 0,1

- **Text:** Welche Sachkosten entstehen im Zusammenhang mit persistenten Identifikatoren im Projekt?
Hilfe: Bitte schätzen sie die Kosten in Euro.

Abschnitt Rechtliche und ethische Fragen

Unterabschnitt Recht allgemein

Frage

Text Muss die rechtliche Situation verschiedener Länder berücksichtigt werden?

Hilfe Falls Sie diese Frage mit “Ja” beantworten, setzen Sie sich bitte mit der Rechtsabteilung bzw. einem/r entsprechenden Ansprechpartner/in an Ihrer Institution in Verbindung, um zu klären, ob sich daraus Konsequenzen für Ihr Projekt ergeben und wenn ja, welche dies sind.

Unterabschnitt Sensible Daten

Fragenset

Fragen

- **Text:** Enthält dieser Datensatz personenbezogene Daten?

Hilfe: Europäisches Recht kann Geltung haben. Das Bundesdatenschutzgesetz definiert personenbezogene Daten als „Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbar natürlichen Person“ (https://www.gesetze-im-internet.de/bdsg_1990/___3.html) (BDSG, §3 Abs. 1) Dabei wird von einer „bestimmten Person“ gesprochen, wenn eindeutig erkennbar ist, zu welcher Person die Daten gehören. Um eine „bestimmbare Person“ handelt es sich, wenn mittels Zusatzinformationen ermittelt werden kann, um welche Person es geht. (Vgl. Häder 2009, S. 7). Der Umgang mit solchen Daten ist gesetzlich geregelt. Mehr Informationen (zwar auf bestimmte Disziplinen / Datentypen fokussiert, aber übertragbar) zum Thema bieten bspw.:
* http://www.ratswd.de/download/RatSWD_WP_2009/RatSWD_WP_90.pdf target=_blank">Michael Häder (2009): Der Datenschutz in den Sozialwissenschaften. RatSWD Working Paper No. 90.
* http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_107.pdf target=_blank">Uwe Jensen (2012): Leitlinien zum Management von Forschungsdaten. Sozialwissenschaftliche Umfragedaten. GESIS Technical Report 2012|07. (v.a. S. 13ff.)
* http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/Sprachkorpora. (v.a. Teil 1.1 und Teil 2.2)

Frage

Text Welches Gesetz ist bezüglich der Fragen des Datenschutzes für das Projekt maßgeblich?

Hilfe Welches Gesetz gilt, hängt von der Art der Einrichtung ab. Für öffentliche Stellen sowie private Unternehmen ist das Bundesdatenschutzgesetz (BDSG) maßgeblich, für öffentlichen Stellen der Länder (z.B. Universitäten) das entsprechende Landesdatenschutzgesetz. (Vgl. http://www.ratswd.de/download/RatSWD_WP_2009/RatSWD_WP_90.pdf target=_blank">Michael Häder (2009): Der Datenschutz in den Sozialwissenschaften. RatSWD Working Paper No. 90, S. 6 f.). In bestimmten Bereichen gelten bereichsspezifische Gesetze, die über dem Bundesdatenschutzgesetz bzw. den Landesdatenschutzgesetze stehen. Für medizinischen Daten

ist bspw. das Zehnte Sozialgesetzbuch (SGB X) anwendbar, für Daten des Mikrozensus das Bundesstatistikgesetz (BStatG). (Vgl. http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_method07.pdf Uwe Jensen (2012): Leitlinien zum Management von Forschungsdaten. Sozialwissenschaftliche Umfragedaten. 2012. S. 14.)

Optionen

- Bundesdatenschutzgesetz (BDSG)
- Landesdatenschutzgesetz Baden-Württemberg
- Landesdatenschutzgesetz Bayern
- Landesdatenschutzgesetz Berlin
- Landesdatenschutzgesetz Bremen
- Landesdatenschutzgesetz Brandenburg
- Landesdatenschutzgesetz Hamburg
- Landesdatenschutzgesetz Mecklenburg-Vorpommern
- Landesdatenschutzgesetz Hessen
- Landesdatenschutzgesetz Nordrhein-Westfalen
- Landesdatenschutzgesetz Rheinland-Pfalz
- Landesdatenschutzgesetz Niedersachsen
- Landesdatenschutzgesetz Saarland
- Landesdatenschutzgesetz Sachsen
- Landesdatenschutzgesetz Sachsen-Anhalt
- Landesdatenschutzgesetz Schleswig-Holstein
- Landesdatenschutzgesetz Thüringen
- Sozialgesetzbuch X (z.B. für medizinische Daten)
- Bundesstatistikgesetz (z.B. für Zensusdaten)
- Sonstiges: _____

Fragenset

Fragen

- **Text:** Enthält der Datensatz “Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit, Gesundheit oder Sexualleben” (BDSG §3, Abs.9)?

Hilfe: Solche Daten gelten als besonders sensibel und erfordern noch strengere Schutzmaßnahmen, als bei personenbezogenen Daten generell bereits erforderlich sind. Wenn Sie diese Frage mit “Ja” beantworten, informieren Sie sich bei der/dem Datenschutzbeauftragten Ihrer Institution, welche zusätzlichen Schutzmaßnahmen notwendig sind.

- **Text:** Werden die Daten anonymisiert oder pseudonymisiert?

Optionen:

- Ja, während der Erhebung
 - Ja, vor / zu Beginn der Datenanalyse
 - Ja, nach der Datenanalyse / vor der Publikation
 - Nein
- **Text:** In welchem Umfang wird die “informierte Einwilligung” der Betroffenen eingeholt?
- Hilfe:** Grundsätzlich gilt, dass eine Erhebung, Verarbeitung, Archivierung und Veröffentlichung personenbezogener Daten nur zulässig ist, wenn eine entsprechende „informierte Einwilligung“ der Betroffenen vorliegt. Nur in ganz wenigen, jeweils in den Datenschutzgesetzen

definierten Ausnahmefällen wird diese nicht benötigt (vgl., auch für mehr Information, http://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/target=_blank)>Informationen zu rechtlichen Aspekten bei der Handhabung von Sprachkorpora, S. 6).

Optionen:

- Zur Analyse/Nutzung der Daten im Rahmen des Projektes sowie zur Nachnutzung
 - Nur zur Analyse/Nutzung der Daten im Rahmen des Projektes
 - Es wird keine “informierte Einwilligung” eingeholt
- **Text:** Wenn keine “informierte Einwilligung” eingeholt wird, begründen Sie dies bitte.
 - **Text:** Wo und wie sind die “informierten Einwilligungen” abgelegt?
 - **Text:** Bis wann werden die (unanonymisierten bzw. unpseudonymisierten) Originaldaten spätestens sicher vernichtet?

Fragenset

Fragen

- **Text:** Enthält dieser Datensatz nicht-personenbezogene sensible Daten?
Hilfe: Beispiele hierfür sind etwa Daten, die Betriebs- oder Geschäftsgeheimnisse oder Ortsangaben zu bedrohten Tier- oder Pflanzenarten enthalten.
- **Text:** Wenn ja, um welche nicht personenbezogenen sensiblen Daten handelt es sich?

Fragenset

Fragen

- **Text:** Welcher Personalaufwand entsteht für die Anonymisierung von sensiblen Daten im Projekt?
Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.
- **Text:** Welche Sachkosten entstehen im Zusammenhang mit der Anonymisierung von sensiblen Daten im Projekt?
Hilfe: Bitte schätzen sie die Kosten in **Euro**.
Bereich:
 - Minimum: 0,0
 - Maximum: 12,0
 - Schrittgröße: 0,1
- **Text:** Welcher Personalaufwand entsteht im Zusammenhang mit weiteren (nicht-technischen) Sicherheitsmaßnahmen für sensible Daten im Projekt?
Hilfe: Bitte schätzen Sie den Aufwand in Personenmonaten.
- **Text:** Wie hoch sind die weiteren Sachkosten für weitere (nicht-technische) Sicherheitsmaßnahmen für sensible Daten im Projekt?
Hilfe: Bitte schätzen sie die Kosten in **Euro**.
Bereich:
 - Minimum: 0,0
 - Maximum: 12,0
 - Schrittgröße: 0,1

Fragenset

Fragen

- **Text:** Wurde das Forschungsvorhaben von einer Ethikkommission begutachtet?

Optionen:

- Ja, positiv begutachtet von folgender Kommission: _____
- Ja, begutachtet mit Auflagen, die folgendermaßen erfüllt werden: _____
- Noch nicht, es befindet sich aber bereits in Begutachtung
- Noch nicht, es wird zur Begutachtung eingereicht bis spätestens: _____
- Nein, eine Begutachtung ist nicht notwendig, weil: _____

- **Text:** Wird für das Forschungsvorhaben eine offizielle Genehmigung benötigt?

Optionen:

- Nein
- Ja. Die Genehmigung liegt bereits vor.
- Ja. Die Genehmigung wurde beantragt am/im: _____
- Ja. Die Genehmigung wird beantragt bis spätestens: _____

- **Text:** Wenn ja, um welche Genehmigung handelt es sich?

- **Text:** Wenn ja, welche ist die ausstellende Behörde?

- **Text:** Wird eine Datenzugangs-Kommission benötigt, die bei Zugriffsanfragen auf vom Projekt veröffentlichte Daten entscheidet, ob Zugang gewährt wird oder nicht?

Unterabschnitt Urheber- oder verwandte Schutzrechte

Frage

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

Hilfe Daten oder Software können Urheber- oder anderen Schutzrechten unterliegen. Die Rechtslage kann selbst in der EU von Land zu Land erheblich abweichen. In Deutschland sind nach dem Urheberrechtsgesetz (UrhG) Werke der Literatur, Wissenschaft und Kunst, die eine „persönliche geistige Schöpfung“ darstellen, urheberrechtlich geschützt. Der urheberrechtliche Schutz erlischt 70 Jahre nach dem Tod der bzw. des Urheberin/s. Reine Daten, z.B. Messdaten oder Surveydaten, aber auch Metadaten (bis auf ggf. „beschreibende“ Metadaten) sind hingegen nicht schutzfähig. In § 2 nennt das UrhG folgende geschützte Werkarten, wobei die Aufzählung nicht abschließend ist: * Sprachwerke, wie Schriftwerke, Reden und Computerprogramme * Werke der Musik * pantomimische Werke einschließlich Werke der Tanzkunst * Werke der bildenden Künste einschließlich der Werke der Baukunst und der angewandten Kunst und Entwürfe solcher Werke * Lichtbildwerke einschließlich der Werke, die ähnlich wie Lichtbildwerke geschaffen werden * Darstellungen wissenschaftlicher oder technischer Art wie Zeichnungen, Pläne, Karten, Skizzen, Tabellen und plastische Darstellungen. Nach § 3 sind auch „Übersetzungen und andere Bearbeitungen“ von Werken geschützt, die persönliche geistige Schöpfungen des Bearbeiters sind“. Schließlich sind nach § 4 auch Sammelwerke und Datenbankwerke geschützt, was im Bereich Forschungsdaten durchaus relevant sein kann. Datenbankwerke werden dabei definiert als „Sammlungen von Werken, Daten oder anderen unabhängigen Elementen, die aufgrund der Auswahl oder Anordnung der Elemente eine persönliche geistige Schöpfung sind“. Bei einem „Datenbankwerk im Sinne des Gesetzes“ handelt es sich um ein „Sammelwerk, dessen Elemente systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind“. Weitere relevante Schutzrechte können gewerbliche Schutzrechte wie Patente, Gebrauchsmuster, Sortenschutz [bei Pflanzenzüchtungen], Halbleiterschutz, Marken, geographische Herkunftsangaben, eingetragene Designs oder geschäftliche Bezeichnungen sein.

Fragenset

Fragen

- **Text:** Be- oder entstehen an diesem Datensatz Urheberrechte?

Optionen:

- Werk der Literatur, Wissenschaft oder Kunst
- Übersetzung oder andere Bearbeitung eines Werkes
- Sammelwerk oder Datenbankwerk
- Andere: _____
- Nein

- **Text:** Be- oder entstehen an diesem Datensatz andere Schutzrechte?

Optionen:

- Patent
- Gebrauchsmuster
- Marke
- Sortenschutz
- Halbleiterschutz
- geographische Herkunftsangabe
- eingetragenes Design
- Andere: _____
- Nein

- **Text:** Wurde der/die Rechteinhaber/in recherchiert?

Optionen:

- Ja: _____
- Nein

Fragenset

Fragen

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

Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.

Bereich:

- Minimum: 0,0
- Maximum: 12,0
- Schrittgröße: 0,1

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

Hilfe: Bitte schätzen sie die Kosten in ****Euro****.

Abschnitt Speicherung und Langzeitarchivierung

Unterabschnitt Auswahl

Fragenset

Fragen

- **Text:** Auf Basis welcher Kriterien / Regeln werden die Daten zur Archivierung (nach Projektende) ausgesucht?
- **Text:** Durch wen erfolgt die Auswahl?

Unterabschnitt Langzeitarchivierung

Fragenset

Fragen

- **Text:** Muss dieser Datensatz langfristig aufbewahrt werden?
- **Text:** Aus welchen Gründen müssen die Daten langfristig aufbewahrt werden?
Optionen:
 - Grundlage einer Publikation / Nachweis guter wissenschaftlicher Praxis
 - Nachnutzung in Folgeprojekten oder durch andere
 - Aufgrund gesetzlicher Bestimmungen
 - Dokumentation aufgrund gesellschaftlicher Relevanz
 - Selbstverpflichtung
 - Andere: _____
- **Text:** Wie lange müssen die Daten aufbewahrt werden?
- **Text:** Wie lang sollen die Daten nach Projektende (nach)nutzbar sein?
Kommentar: Original H2020 question
- **Text:** Wo werden die Daten (einschließlich Metadaten, Dokumentation und ggf. relevantem Code bzw. relevanter Software) nach Projektende gespeichert bzw. archiviert?
Optionen:
 - Eigene Institution
 - Disziplinspezifisches Datenzentrum: _____
 - Generisches Datenzentrum: _____
 - Wurde noch nicht entschieden
 - Sonstiges: _____
- **Text:** Handelt es sich dabei um ein zertifiziertes Repositorium oder Datenzentrum (z.B. durch das Data Seal of Approval, nestor-Siegel oder ISO 16363)? (Wurden mehrere Langzeitarchivierungsoptionen ausgewählt, kann die Frage bejaht werden, wenn dies auf mindestens eine der Optionen zutrifft).
- **Text:** Wurde mit dem Repositorium oder Datenzentrum bereits angemessene Archivierungslösungen besprochen?
- **Text:** Sollen die Daten erst nach Ablauf einer Sperrfrist zugänglich gemacht werden?

- **Text:** Wie wird die Identität von Personen, die auf die Daten zugreifen, verifiziert?
Kommentar: Original from H2020 FAIR Data Management Plan
- **Text:** Wann werden die Daten archiviert?

Fragenset

Fragen

- **Text:** Welcher Personalaufwand entsteht im Zusammenhang mit Langzeitarchivierung für dieses Projekt?
Hilfe: Bitte schätzen sie den Aufwand in Personenmonaten.
Bereich:
 - Minimum: 0,0
 - Maximum: 12,0
 - Schrittgröße: 0,1
- **Text:** Welche Sachkosten entstehen im Zusammenhang mit Langzeitarchivierung für dieses Projekt?
Hilfe: Bitte schätzen sie die Kosten in ****Euro****.
- **Text:** Wie werden die Kosten für das Datenmanagement im Projekt aufgebracht?