

## 1. Allgemein

### 1.1. Thema

1.1.1. Wie lautet die primäre Forschungsfrage des Projektes?

1.1.2. Bitte geben Sie einige Schlagworte zur Forschungsfrage an.

1.1.3. Welcher Disziplin / welchen Disziplinen ist das Projekt zuzuordnen?

### 1.2. Projektablauf

1.2.1. Wann beginnt die Projektlaufzeit?

1.2.2. Wann endet die Projektlaufzeit?

### 1.3. Projektpartner

1.3.1. Welche Personen oder Institutionen sind verantwortlich für die Projektkoordination? 1.3.2.

Projektpartner

1.3.3. 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.

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

### 1.4. Förderung

1.4.1. Wer fördert das Projekt?

1.4.2. In welcher Förderlinie und/oder welchem Förderprogramm wird das Projekt gefördert? 1.4.3.

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.

### 1.5. Weitere Anforderungen

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

1.5.2. Welche Anforderungen an das Datenmanagement sind dies?

## 2. Inhaltliche Einordnung

### 2.1. Daten

2.1.1. Um was für einen Datensatz handelt es sich?

2.1.2. Wird der Datensatz selbst erzeugt oder nachgenutzt?

2.1.3. Wenn nachgenutzt, wer hat den Datensatz erzeugt?

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

2.1.5. Für welche Personen, Gruppen oder Institutionen könnte dieser Datensatz (für die Nachnutzung) von Interesse sein? Für welche Szenarien ist dies denkbar?

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

### 2.2. Technische Einordnung

#### 2.2.1. Daten

2.2.2. Wann beginnt die Erhebung bzw. Erstellung der Daten?

2.2.3. Wann endet die Erhebung bzw. Erstellung der Daten?

2.2.4. Wann beginnt die Datenbereinigung / -aufbereitung?

2.2.5. Wann endet die Datenbereinigung / -aufbereitung?

2.2.6. Wann beginnt die Datenanalyse?

2.2.7. Wann endet die Datenanalyse?

2.2.8. Was ist die tatsächliche oder erwartete Größe des Datensatzes?

2.2.9. Wie hoch ist die erwartete Erzeugungsrate der Daten pro Jahr?

2.2.10. In welchen Formaten liegen die Daten vor?

- 2.2.11. Welche Instrumente, Software, Technologien oder Verfahren werden zur Erzeugung oder Erfassung der Daten genutzt?
- 2.2.12. Welche Software, Verfahren oder Technologien sind notwendig, um die Daten zu nutzen?
- 2.2.13. Wird die Dokumentation von ggf. zur Nutzung notwendiger Software benötigt, um die Daten zu nutzen?
- 2.2.14. Werden verschiedene Versionen des Datensatzes erzeugt?
- 2.2.15. Welche Versionierungsstrategie wird für diesen Datensatz angewandt?
- 2.2.16. Welche Technologie bzw. welches Tool wird zur Versionierung verwendet?

### 3. Datennutzung

#### 3.1. Nutzungsszenarien

- 3.1.1. Wozu / wie wird dieser Datensatz während des Projektes genutzt?
- 3.1.2. Wie häufig wird dieser Datensatz genutzt?
- 3.1.3. In welchem Umfang werden Infrastrukturressourcen benötigt (CPU-Stunden, Bandbreite, Speicherplatz etc.)?
- 3.1.4. Gibt es beabsichtigte (ggf. auch potentielle) Nutzungsszenarien, für die die Unterstützung durch Datenmanagement- oder IT-ExpertInnen sinnvoll oder notwendig ist?

#### 3.2. Datenspeicherung und -sicherheit

- 3.2.1. Wo wird der Datensatz während des Projektes gespeichert?
- 3.2.2. Unter welcher URL kann der Datensatz während des Projektes abgerufen werden?
- 3.2.3. Gibt es projektinterne Richtlinien zur einheitlichen Organisation der Daten? Wenn ja, wo sind diese festgehalten?
- 3.2.4. Gibt es eine projektinterne Richtlinie zur Benennung der Daten? Wenn ja, bitte skizzieren Sie sie kurz und verlinken Sie ggf. zu einer ausführlicheren Dokumentation.
- 3.2.5. Wer darf auf den Datensatz zugreifen?
- 3.2.6. Wie und wie oft werden Backups der Daten erstellt?
- 3.2.7. Wer ist verantwortlich für die Erstellung der Backups?
- 3.2.8. Welche Maßnahmen zur Gewährleistung der Datensicherheit werden getroffen (z. B. Schutz vor unbefugtem Zugriff, Datenwiederherstellung, Übertragung sensibler Daten)?

#### 3.3. Weitergabe und Veröffentlichung

- 3.3.1. Ist der Datensatz interoperabel, d.h. geeignet für den Datenaustausch und die Nachnutzung zwischen bzw. von unterschiedlichen Forschenden, Institutionen, Organisationen und Ländern?
- 3.3.2. Soll dieser Datensatz veröffentlicht oder geteilt werden?
- 3.3.3. Wenn nicht, begründen Sie dies bitte und unterscheiden Sie dabei zwischen rechtlichen und/oder vertraglichen Gründen und freiwilligen Einschränkungen.
- 3.3.4. Wenn ja, unter welchen Nutzungsbedingungen oder welcher Lizenz sollen die Daten veröffentlicht bzw. geteilt werden?
- 3.3.5. Sollte die Nachnutzung dieses Datensatzes Einschränkungen unterliegen, erläutern Sie bitte die Gründe.
- 3.3.6. Wann werden die Daten veröffentlicht?

#### 3.4. Kollaboratives Arbeiten

- 3.4.1. Werden die Daten kollaborativ genutzt?
- 3.4.2. Welche Plattform, welche Werkzeuge werden zum kollaborativen Arbeiten an Daten und Publikationen genutzt?
- 3.4.3. Wie ist das kollaborative Arbeiten an denselben Dateien geregelt?

#### 3.5 Qualitätssicherung

- 3.5.1. Welche Maßnahmen zur Qualitätssicherung werden für diesen Datensatz ergriffen?
- 3.5.2. Wird die Integration zwischen nachgenutzten und erzeugten Daten gewährleistet? Wenn ja, wie?

### 3.6. Kosten

- 3.6.1. Welcher Personalaufwand für das Datenmanagement entsteht im Rahmen der Erhebung, Erstellung oder Akquise der Daten im Projekt?
- 3.6.2. Welche Sachkosten für das Datenmanagement entstehen im Rahmen der Erhebung, Erstellung oder Akquise der Daten im Projekt?
- 3.6.3. Welcher Personalaufwand für das Datenmanagement entsteht im Zusammenhang mit der Nutzung der Daten im Projekt?
- 3.6.4. Welche Sachkosten für das Datenmanagement entstehen im Zusammenhang mit der Nutzung der Daten im Projekt?
- 3.6.5. Welcher Personalaufwand entsteht im Zusammenhang mit der Speicherung der Daten und Maßnahmen zur Datensicherheit während des Projektes?
- 3.6.6. Welche Sachkosten entstehen im Zusammenhang mit der Speicherung der Datensätze während des Projektes?

### 4. Metadaten und Referenzierung

#### 4.1. Metadaten

- 4.1.1. 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?
- 4.1.2. Welche Standards, Ontologien, Klassifikationen etc. werden zur Beschreibung der Daten und Kontextinformation genutzt?
- 4.1.3. Welche Metadaten werden automatisch erhoben?
- 4.1.4. Sollte es unvermeidbar sein, projektspezifische oder seltene Ontologien, Metadatenschemata oder Vokabulare zu nutzen, werden Mappings zu gängigen Ontologien etc. erstellt?
- 4.1.5. Welche Metadaten werden semi-automatisch erhoben?
- 4.1.6. Welche Metadaten werden manuell erhoben?
- 4.1.7. Werden Metadaten und Kontextinformation auf Korrektheit und Vollständigkeit geprüft?
- 4.1.8. Wer ist verantwortlich für die Dokumentation und Prüfung der Metadaten und Kontextinformationen auf Richtigkeit und Vollständigkeit?
- 4.1.9. Welcher Personalaufwand entsteht im Zusammenhang mit der Erstellung von Metadaten und Kontextinformation im Projekt?
- 4.1.10. Welche Sachkosten entstehen im Zusammenhang mit der Erstellung von Metadaten und Kontextinformation im Projekt?

#### 4.2. Objektstruktur, Granularität und Referenzierung

- 4.2.1. 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?
- 4.2.2. Sollen für diesen Datensatz persistente Identifikatoren (PIDs) genutzt werden?
- 4.2.3. Welches System von persistenten Identifikatoren soll genutzt werden?
- 4.2.4. Welche (Sub-)Entitäten / Untereinheiten sollten sinnvollerweise eigene Identifikatoren erhalten? Welche dieser Identifikatoren sollten dauerhaft und zitierfähig sein?
- 4.2.5. Wer ist verantwortlich für die Pflege der PIDs und die Objektpflege (d. h. die Langzeitarchivierung des Objekts und somit dafür, dem PID-Service einen Objektzug und die neue Adresse mitzuteilen)?
- 4.2.6. Welcher Personalaufwand entsteht im Zusammenhang mit der Vergabe von persistenten Identifikatoren im Projekt?
- 4.2.7. Welche Sachkosten entstehen im Zusammenhang mit persistenten Identifikatoren im Projekt?

### 5. Rechtliche und ethische Fragen

#### 5.1. Recht allgemein

5.1.1. Muss die rechtliche Situation verschiedener Länder berücksichtigt werden?

## 5.2. Sensible Daten

5.2.1. Enthält dieser Datensatz personenbezogene Daten?

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

5.2.3. Enthält der Datensatz "Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit, Gesundheit oder Sexualleben" (BDSG §3, Abs.9)?

5.2.4. Werden die Daten anonymisiert oder pseudonymisiert?

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

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

5.2.7. Wo und wie sind die "informierten Einwilligungen" abgelegt?

5.2.8. Bis wann werden die (un anonymisierten bzw. unpseudonymisierten) Originaldaten spätestens sicher vernichtet?

5.2.9. Enthält dieser Datensatz nicht-personenbezogene sensible Daten?

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

5.2.11. Welcher Personalaufwand entsteht für die Anonymisierung von sensiblen Daten im Projekt?

5.2.12. Welche Sachkosten entstehen im Zusammenhang mit der Anonymisierung von sensiblen Daten im Projekt?

5.2.13. Welcher Personalaufwand entsteht im Zusammenhang mit weiteren (nicht-technischen) Sicherheitsmaßnahmen für sensible Daten im Projekt?

5.2.14. Wie hoch sind die weiteren Sachkosten für weitere (nicht-technische) Sicherheitsmaßnahmen für sensible Daten im Projekt?

5.2.15. Wurde das Forschungsvorhaben von einer Ethikkommission begutachtet?

5.2.16. Wird für das Forschungsvorhaben eine offizielle Genehmigung benötigt?

5.2.17. Wenn ja, um welche Genehmigung handelt es sich?

5.2.18. Wenn ja, welche ist die ausstellende Behörde?

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

## 5.3. Urheber- oder verwandte Schutzrechte

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

5.3.2. Be- oder entstehen an diesem Datensatz Urheberrechte?

5.3.3. Be- oder entstehen an diesem Datensatz andere Schutzrechte?

5.3.4. Wurde der/die Rechteinhaber/in recherchiert?

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

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

## 6. Speicherung und Langzeitarchivierung

### 6.1. Auswahl

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

6.1.2. Durch wen erfolgt die Auswahl?

### 6.2. Langzeitarchivierung

6.2.1. Muss dieser Datensatz langfristig aufbewahrt werden?

6.2.2. Aus welchen Gründen müssen die Daten langfristig aufbewahrt werden?

6.2.3. Wie lange müssen die Daten aufbewahrt werden?

6.2.4. Wie lang sollen die Daten nach Projektende (nach)nutzbar sein?

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

- 6.2.6. 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).
- 6.2.7. Wurde mit dem Repositorium oder Datenzentrum bereits angemessene Archivierungslösungen besprochen?
- 6.2.8. Sollen die Daten erst nach Ablauf einer Sperrfrist zugänglich gemacht werden?
- 6.2.9. Wie wird die Identität von Personen, die auf die Daten zugreifen, verifiziert?
- 6.2.10. Wann werden die Daten archiviert?
- 6.2.11. Welcher Personalaufwand entsteht im Zusammenhang mit Langzeitarchivierung für dieses Projekt?
- 6.2.12. Wie werden die Kosten für das Datenmanagement im Projekt aufgebracht?
- 6.2.13. Welche Sachkosten entstehen im Zusammenhang mit Langzeitarchivierung für dieses Projekt?