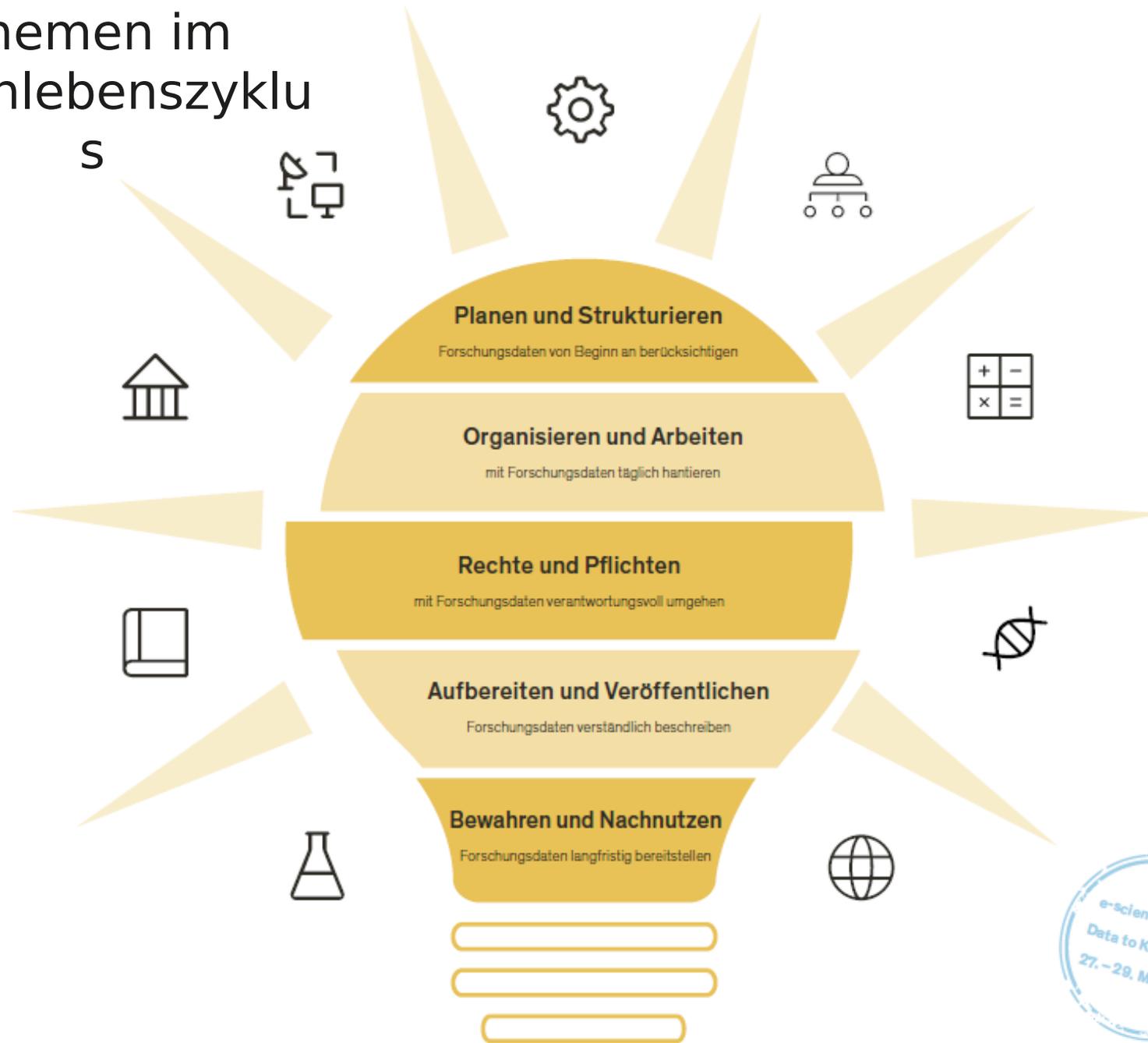


Themen im Datenlebenszyklus



-  Bauwesen und Architektur
-  Chemie
-  Geisteswissenschaften
-  Geowissenschaften
-  Informatik, System- und Elektrotechnik
-  Klassisches Ingenieurwesen
-  Lebenswissenschaften
-  Physik und Mathematik
-  Verhaltens- und Gesellschaftswissenschaften

Die RDMO-Instanz auf rdmo.forschungsdaten.info



- Speziell auf forschungsdaten.info
 - Shibboleth-Zugang
 - Mitarbeit an der Vorlage für Datenschutz und Nutzungsbedingungen
 - Werbung und Koordination in Baden-Württemberg: Im AK FDM wollen die meisten Mitglieder RDMO nutzen
 - Erstellung angepasster Fragenkataloge in Kooperation mit potentiellen Nutzern (z.B. Forscher aus einem SFB)
 - Ein sehr vereinfachter Fragenkatalog für den User als Einstieg gedacht
 - Mehrere modulare Fragenkataloge, angepasst an die Disziplin bzw. Art der Forschungsdaten
 - Kurzen DMP-Abriss erstellt für den Bedarf unterhalb eines Tools

Entstehung der Daten

Die aus dem Projekt bwFDM-Communities stammenden Interviewdaten erfassen Verbesserungsanliegen der Forschenden in Baden-Württemberg für ihren Umgang mit Forschungsdaten. Die Interviews an allen neun Landesuniversitäten werden von einem dedizierten Projektmitarbeiter an jedem Standort persönlich anhand eines Interviewleitfadens geführt und decken alle Fachdisziplinen ab.

Die Daten dienen zur Bedarfsquantifizierung gegenüber dem MWK, um gegebenenfalls Folgeprojekte im Rahmen einer Digitalisierungsstrategie anstoßen zu können.

Beschreibung der Daten

Die Interviewdaten werden mit Treesheets oder von Hand erhoben und in ein gemeinsames Leitfadensformat überführt. Aus den Interviews werden User-Stories der Form „Als [Rolle] wünsche ich mir [Mittel], um [Zweck]“ gewonnen. Die Zusammenführung aller Daten erfolgt in einer SQL-Datenbank, in der zusätzliche Metadaten (Themen-Kategorien, DFG-Facheinordnung etc.) abgelegt werden.

Die Verständlichkeit der User-Stories und der Bezug zum Originalinterview werden durch intensiven gegenseitigen, kritischen und doppelten Review unter den Projektmitarbeitern sichergestellt, sodass die Kernaussagen auch ohne Informationen über das zugrundeliegende Interview verständlich sind.

Verwendbarkeit der Daten

Da es die größte Studie dieser Art ist, sind die Daten für Studien und Dienstentwickler interessant, die FDM-Services mit Nutzerwünschen vergleichen oder Forschungsbedarfe im IT-Bereich generell untersuchen.

Die User-Stories werden als .csv Datei veröffentlicht. Der Bericht des Projekts analysiert die User-Stories gruppiert anhand der zugeordneten Themen-Kategorien, wird als pdf veröffentlicht und enthält auch den Interview-Fragebogen. User-Stories und Bericht werden über KITopen dauerhaft referenzierbar unter CC BY-NC-ND veröffentlicht und archiviert, wobei auf Wunsch auch Speziallizenzen verfügbar sein sollen.

Bewahrung der Daten

Für den geeigneten Umgang mit den Interviewdaten sind die jeweiligen Erhebungsstandorte verantwortlich. Die Projektleitung verfolgt die Aufbewahrung, Vollständigkeit und Zugänglichmachung der zur Veröffentlichung angedachten Daten unter Einbeziehung aller Standorte an zentraler Stelle.

Die anfallenden Daten sind von geringer Größe (<50 MB), da sie von Hand erhoben werden. Die Interviewdaten werden nach Projektende gelöscht und nicht zwischen den Projektstandorten geteilt.

Die hochwertige Aufbereitung, Anonymisierung und Qualitätskontrolle der User-Stories aus den Interviews ist als Kernaufgabe im Projekt mit 36 Personennomonaten bedacht, sodass tatsächlich alle Anliegen aller Forschenden langfristig strukturiert verfügbar bleiben. Aufgrund der geringen Datenmenge wird die IT-Infrastruktur zur Speicherung und Präsentation der Daten als Eigenleistung durch Grundmittel abgedeckt.

1-2 Sätze, die die grundlegende Herkunft der Daten allgemeinverständlich erklären.

Ein Satz, der klar herausstellt, welche Frage die Daten im Forschungskontext beantworten sollen.

3-4 Sätze des genaueren Hintergrunds was für Informationen enthalten sind und welche Formate / Metadaten / Standards benutzt werden.

1-2 Sätze zum Qualitätssicherungskonzept der konkreten Datenauswertung.

Bei Mehrbedarf einen eigenen Abschnitt „Beschreibung der Metadaten“ formulieren.

1 Satz der erklärt, für welche Art von Forschung die Daten generell interessant sein könnten.

3-4 Sätze die erklären, wie die Daten strukturiert sind oder andere die Daten in die eigene Forschung integrieren können oder wie die Daten mit anderen Daten zusammen abgelegt werden.

Bei Mehrbedarf beispielsweise für das Datenschutskonzept einen eigenen Abschnitt vorschlagen.

1-2 Sätze zur Verantwortlichkeit

1 Satz zur erwartenden Datengröße und zur geplanten Aufbewahrungszeit. Ein Satz mehr, falls Rechtsfragen erwähnt werden

2-3 Sätze zu benötigten Ressourcen

Kurzer DMP-Abriss: Das Minimalbeispiel eines DMP

Projektbeschreibung

1-2 Sätze, die die grundlegende Herkunft der Daten allgemeinverständlich erklären.

Ein Satz, der klar herausstellt, welche Frage die Daten im Forschungskontext beantworten sollen.

Die aus dem Projekt bwFDM-Communities stammenden Interviewdaten beschreiben die Unterstützungsanliegen der Forschenden in Baden-Württemberg beim Umgang mit Forschungsdaten. Die Interviews an allen neun Landesuniversitäten werden von einem dezidierten Projektmitarbeiter an jedem Standort persönlich anhand eines Interviewleitfadens geführt und decken alle Fachdisziplinen ab.

Die Daten dienen zur Bedarfsquantifizierung gegenüber dem MWK, um gegebenenfalls Folgeprojekte im Rahmen einer Digitalisierungsstrategie anstoßen zu können.

Beschreibung der Daten

3-4 Sätze des genaueren Hintergrunds was für Informationen enthalten sind und welche Formate / Metadaten / Standards benutzt werden.

[Bei Mehrbedarf einen eigenen Abschnitt „Beschreibung der Metadaten einschieben“.]

1-2 Sätze zum Qualitätssicherungskonzept der Datenauswertung.

Die Interviewdaten werden mit Treesheets oder von Hand erhoben und in ein gemeinsames Leitfadensformat überführt. Aus den Interviews werden User Storys der Form „Als [Rolle] wünsche ich mir [Mittel], um [Zweck]“ gewonnen. Die Zusammenführung aller Daten erfolgt in einer SQL-Datenbank, in der zusätzliche Metadaten (Themen-Kategorien, DFG-Facheinordnung etc.) abgelegt werden.

Die Verständlichkeit der User-Storys und der Bezug zum Originalinterview werden durch intensiven gegenseitigen, kritischen und doppelten Review unter den Projektmitarbeitern sichergestellt, sodass auch Personen die ein Interview nicht geführt haben die Kernaussagen nachvollziehen.

Verwendbarkeit der Daten

Ein Satz der erklärt, für welche Art von Forschung die Daten generell interessant sein könnten.

3-4 Sätze die erklären, wie die Daten strukturiert sind oder andere die Daten in die eigene Forschung integrieren können oder wie die Daten mit anderen Daten zusammen abgelegt werden.

Da es die größte Studie dieser Art ist, sind die Daten für Studien und Dienstentwickler interessant, die FDM-Services mit Nutzerwünschen vergleichen oder Forschungsbedarfe im IT-Bereich generell untersuchen.

Die User-Storys werden als .csv Datei veröffentlicht. Der Bericht des Projekts analysiert die User-Storys gruppiert anhand der zugeordneten Themen-Kategorien, wird als pdf veröffentlicht und enthält auch den Interview-Fragebogen. User Storys und Bericht werden über KITopen dauerhaft referenzierbar unter CC BY-NC-ND veröffentlicht und archiviert, wobei auf Wunsch auch Speziallizenzen verfügbar sein sollen.

Datenspeicherung und Datenhaltung

1-2 Sätze zur erwartenden Datengröße und zur geplanten Aufbewahrungszeit. Ein Satz mehr, falls Rechtsfragen erwähnt werden

[sonst ggf. für Datenschutz o.ä. ein extra Block zwischen Verfügbarmachung und Datenhaltung]

1-2 Sätze zur Verantwortlichkeit

2-3 Sätze zu benötigten Ressourcen

Die anfallenden Daten sind von geringer Größe (<50 MB), da sie von Hand erhoben werden. Die Interviewdaten werden nach Projektende gelöscht und nicht zwischen den Projektstandorten geteilt.

Für den geeigneten Umgang mit den Interviewdaten sind die jeweiligen Erhebungsstandorte verantwortlich. Die Projektleitung verfolgt die Aufbewahrung, Vollständigkeit und Zugänglichmachung der zur Veröffentlichung angedachten Daten unter Einbeziehung aller Standorte an zentraler Stelle.

Die hochwertige Aufbereitung, Anonymisierung und Qualitätskontrolle der User-Storys aus den Interviews ist als Kernaufgabe im Projekt mit 36 Personenmonaten bedacht, sodass tatsächlich alle Anliegen aller Forschenden langfristig strukturiert verfügbar bleiben. Aufgrund der geringen Datenmenge wird die IT-Infrastruktur zur Speicherung und Präsentation der Daten als Eigenleistung durch Grundmittel abgedeckt.

