

Wie verändert sich Religiosität? Religiöse Dynamiken im Spiegel eines multidimensionalen Modells der Religiosität

Data Management Plan (DMP)

1 Datenerhebung und -dokumentation

1.1 Welche Daten werden Sie erheben, beobachten, generieren oder wiederverwenden?

Das Projekt betreibt primär eine längsschnittlich angelegte, explorative Grundlagenforschung zu innerreligiösen Dynamiken, wobei quantitative und qualitative Daten erhoben und analysiert werden.

Das Projekt arbeitet mit drei verschiedenen Datenformaten:

Paneldaten: das Projekt greift auf bestehende und zukünftig erhobene Daten des Schweizerischen Haushaltpanels (SHP) zu. Das Datenmanagement im Zusammenhang mit SHP übernimmt der Projektpartner am FORS (Fondation suisse pour la recherche en sciences sociales) in Lausanne. Bisher liegen die Wellen 1 bis 20 in tabellarischer Form (Fälle x Variablen) vor. Im Laufe des Projektes kommen noch mindestens 2 Wellen hinzu. Der Datensatz bedarf in seiner derzeitigen Form unkomprimiert ca. 400MB an Speicherplatz. Den Fokus für das Projekt bildet das Rotationsmodul zu Glauben und Religion, das in 2012, 2015, 2018 Daten generiert hat und weiter alle 3 Jahre in die jährliche Befragung integriert wird. Die Daten werden von FORS in den gängigen Formaten für die Arbeit mit statistischer Software IBM SPSS® (.sav), STATA (.dta) und SAS (.sas7bdat) zur Verfügung gestellt.

Interviewdaten: hiermit sind Audioaufnahmen und Transskripte von Audioaufnahmen aus qualitativen Interviews gemeint. Bisher liegt dem Forschungsteam ein Datenumfang von 160 Interviewdatensätzen. Im Laufe des Projektes sollen diese Daten mit neu erhobenen Daten zu längsschnittlichen qualitativen Datensätzen zusammengefasst werden. Im Projekt werden zu 2 Zeitpunkten Interviews mit 120 Personen geführt. Pro Zeitpunkt und pro Fall entsteht dabei eine Datenmenge von durchschnittlich 200MB bei 90min Interviewzeit. Insgesamt ist dementsprechend bei der Zusammenführung der alten und neuen Daten mit einer Gesamtgrösse von $200\text{MB} \times (120 \times 2 + 160) = 80'000\text{MB}$, gerundet 80GB zu rechnen. Die erwarteten Dateiformate sind für die Audioaufnahmen ".mp3" und für die Transkripte ".docx", die Zwischenform wird in dem MAXQDA Format ".MX12" gespeichert.

Fragebogendaten: begleitend zu Interviews werden quantitative Erhebungen mithilfe von psychometrischen Kurzfragebögen erhoben. Erfasst werden demographische Angaben zur Person und religionspsychologische sowie -soziologische Variablen. Die Daten werden in Form von tabellarischen Datensätzen (Fälle x Variablen) zusammengeführt und verwaltet. Bei einer geplanten Fallzahl von 120 und zwei Messzeitpunkten ist mit einer Datei kleiner als 100MB zu rechnen. Das erwartete Rohformat sind ".csv"-Dateien, die später in IBM SPSS®-Datendateien (.sav) überführt werden.

Insgesamt ist mit einer Datenmenge von gerundet 80.5GB zu rechnen.

1.2 Wie werden die Daten erhoben, beobachtet oder generiert?

Grundsätzlich: um die Zusammenarbeit zu ermöglichen wird ein Cloud-Speicherdienst der Universität Bern für die Speicherung der Dateien verwendet. Nur Personen mit einer Anstellung im Projekt haben durch den Campus-Account einen Zugriff auf die Projektdaten. Die IT-Dienste der Universität Bern stellen innerhalb des Cloud-Dienstes Back-Up-Lösungen zur Verfügung und sorgen für die Datensicherheit. Innerhalb der Cloud gibt es ein Berechtigungssystem, d.h. dass der Projektverantwortliche je nach Aufgabengebiet Zugriff auf Ordner gewähren und zurücknehmen kann.

Paneldaten: für die Erhebung der Paneldaten ist der Projektpartner SHP-Arbeitsgruppe am FORS in Lausanne zuständig. Als von SNF mitgetragene Institution wendet FORS alle vom SNF als notwendig erachteten Qualitätssicherungsmaßnahmen an. Die längsschnittliche Analyse ist durch eine anonyme ID in den Daten möglich. Wiederholte Messungen lassen sich durch diese ID eindeutig identifizieren.

Interviewdaten: Interviews werden nur nach Zustimmung der Auskunftsperson durchgeführt. Die Interviewer unterliegen einer Verschwiegenheitsregelung, die sie schriftlich zu Projektbeginn erklären. Bei den Interviews werden Audioaufnahmen mit Diktafonen erstellt. Die Audioaufnahmen werden anschliessend auf Rechnern der Universität Bern gespeichert und verbleiben für die Transskription mit der Software MAXQDA lokal gespeichert. MAXQDA übernimmt die Versionierung von Transskripten und enthält zudem Werkzeuge zur Verwaltung von qualitativen Daten. Da MAXQDA eine proprietäre Software ist, ist der Zugriff nur für Lizenzinhaber innerhalb des Teams möglich. Lizenzen werden personengebunden ausgestellt. Vor der Analyse der Transskripte im Teamsitzungen erfolgt eine Pseudonymisierung der Daten im Text (z.B. für Namen, für Ortsbezeichnungen). Für die Rückverfolgung der Daten wird im Prozess der Pseudonymisierung eine Schlüsseldatei erstellt, diese enthält die Paarungen von Klarnamen und Pseudonymen und verbleibt bei dem Projektverantwortlichen für den Fall der Notwendigkeit der Rückverfolgung. Im Team kann dementsprechend nur der Interviewer selbst und der Projektverantwortliche wissen, um wenn es sich bei einem Pseudonym handelt. Die übrigen Teammitglieder verständigen sich bei den Analysen mithilfe der Pseudonyme. Bei der wiederholten Messung ist vorgesehen, dass der gleiche Interviewer dieses Interview durchführt, wodurch die Beziehung zur Auskunftsperson und die Pseudonymisierung bewahrt bleiben. Bei den qualitativen Datenanalysen entstehen sogenannte Codebücher. Die dort enthaltenen Codes sind das Produkt von zuerst individueller und anschliessend gemeinschaftlicher Arbeit, wobei das Prinzip des Peer-Reviews ins Spiel kommt. Auf diese Weise ist ein weiteres Prinzip der Qualitätssicherung involviert. Aus früheren Projekten hat sich eine bestimmte Ordnerstruktur für die Analysen etabliert, die sowohl die Chronologie der Erhebung und der Analyse als auch die Zusammenführung von Einzelfällen zu Datenbanken erleichtert. Wir werden diese Struktur für die Mitarbeiter im qualitativen Teil festlegen:

01_Fälle
02_Memos_Leitfadeninterview
03_Audiodateien_Notizen
04_Transkripte_original
05_Transkripte_anonymer_Datenname

Fragebogendaten: die Erhebung der Fragebogendaten erfolgt je nach Gegebenheit entweder im Paper-Pencil-Verfahren oder online. Auskunftspersonen füllen einen Fragebogen aus, bei dem sie zuerst die Zustimmung zur Verwendung in der Forschung geben. Fragebogendaten werden tabellarisch zusammengefasst und im Projektordner in der Team-Cloud gespeichert. Da eine Verknüpfung der Fragebogendaten mit den Interviewdaten vorgesehen ist, erfolgt diese über das Pseudonym. Die Versionierung der Fragebogendaten erfolgt im manuellen Modus über die Datumsangabe in der jeweiligen Dateiversion z.B.: JJJJMMTT_Dateiname_Namenskürzel.

1.3 Welche Dokumentationen und Metadaten sehen Sie im Zusammenhang mit den Daten vor?

Generell: zu Beginn des Projektes werden alle Mitarbeiter Namens Kürzel bekommen. Sollte eine Software keine Versionierung erlauben, werden die Mitarbeiter angewiesen die von ihnen bearbeiteten Versionen mit Namens Kürzeln und einem aktualisierten Datum zu versehen. Mitunter müssen Mitarbeiter auch singuläre Dateien erzeugen, die sie mit dem Namens Kürzel kennzeichnen siehe z.B. den Unterpunkt "Interviewdaten".

Paneldaten: da es sich beim SHP um ein Paneldatenformat handelt, welches von FORS in Lausanne gepflegt wird, existieren bereits ausführliche Metadaten wie z.B. Codebücher und kommentierte Fragenkataloge. Diese Arbeit wird vom Projektpartner weiterhin bewerkstelligt.

Interviewdaten: direkt nach einem Interview ist der Interviewer verpflichtet ein sogenanntes Memo zu erstellen. Dabei handelt es sich um eine am unseren Institut standardisierte Form von Metadaten über ein Interview in Form einer formatierten Textdatei. Es enthält Angaben über die Kontaktaufnahme, Kontaktdaten und kurze Zusammenfassung der demographischen Daten der Zielperson, Beschreibungen der Interviewumgebung (Datum, Ort, Dauer) und der interpersonellen Atmosphäre, Situation vor dem Beginn der Aufnahme, erste Eindrücke des Interviewer nach dem Interview, kurze Zusammenfassung der Schwerpunkte des Gesprächs sowie die ersten Hypothesen über die Äusserungen im Zusammenhang mit den Forschungsfragen, "off-the-record-Situationen", Verbesserungsvorschläge, offene Fragen, Anmerkungen zum Projekt als Gesamtes. Das Memo ist eine zwei bis drei Seiten lange Datei, die die Audioaufnahmen kontextuell einordnet und den Zugang für andere erleichtert. Memodateien werden nach einem festen Prinzip benannt: JJMMTT_Interviewerkürzel_Name, Vorname der Auskunftsperson.

Fragebogendaten: die Fragebogendaten werden tabellarisch als numerischer Code (Zahlen anstatt Zeichenketten) zusammengefasst. Das von uns verwendete Softwarepaket IBM SPSS® erlaubt das Labeln von numerischen Codes. Aus Labels erstellen wir in der Software automatisiert ein Codebuch,

welches während des Projektes und nach dem Abschluss als Metadatei vorliegen wird.

2 Ethische, rechtliche und Sicherheitsfragen

2.1 Wie gestalten sich der Umgang mit und die Behandlung von ethischen Fragen?

Wir sehen vor, bei einem positiven Entscheid für das Projekt einen Ethikantrag bei der zuständigen lokalen Ethikkommission zu stellen. Ein für die Postdoc-Stelle vorgesehener Projektmitarbeiter ist aktuell Mitglied in der Ethikkommission seiner Universität und wird auf die Einhaltung der Ethikrichtlinien bei der Ausarbeitung der konkreten Arbeitsschritte Acht geben. Insbesondere sind wir uns folgender Punkte bewusst:

- wir erheben und verarbeiten sensible personenbezogene Daten, die besonderer Schutzmassnahmen bedürfen
- eine Anonymisierung ist nicht mit allen erhobenen Daten möglich (z.B. bei Interviewdaten), wir müssen an manchen Stellen zu Pseudonymisierung greifen, da sonst Analysen nicht möglich sind. Wir unterscheiden klar zwischen Anonymisierung und Pseudonymisierung und wenden, wo möglich, die striktere Schutzmassnahme der Anonymisierung an
- die Datei mit den Pseudonymisierungsschlüsseln wird verschlüsselt und der Zugriff auf den Projektverantwortlichen beschränkt
- beim gemeinsamen Arbeiten mit den Daten im Cloud-Speicher der Universität werden die Zugriffsberechtigungen entsprechend der Aufgaben zugeteilt, nur der Projektverantwortliche hat den Zugriff auf der obersten Berechtigungsebene und autorisiert die Projektmitarbeiter zum Datenzugriff entsprechend der Aufgaben
- sollte die Ethikkommission eine Ende-zu-Ende-Verschlüsselung für notwendig erachten, werden wir diese technische Massnahme in Kooperation mit den IT-Diensten der Universität Bern umsetzen
- von FORS haben wir die notwendige Lizenz, die Daten zu erhalten, zu verarbeiten und aufzubewahren, bereits bekommen

2.2 Wie werden der Zugriff auf Daten und die Datensicherheit verwaltet?

In Fragen des Zugriffs auf die Daten und die Datensicherheit lehnen wir uns an die Empfehlungen der Universität Bern zum Umgang mit Forschungsdaten an und benutzen eine Cloudspeicherlösung, die von der Universität Bern den Mitarbeitern zur Verfügung gestellt wird. Es handelt sich dabei um Speicher, der sich auf lokalen Servern befinden, die von der Universität Bern selbst verwaltet werden. Es existieren tägliche Back-Ups. Der Datenzugriff wird durch den Projektverantwortlichen in der Ordnerstruktur verwaltet, entsprechende Werkzeuge liegen vor und wurden bereits erfolgreich ausprobiert. Projektmitarbeiter werden Zugriff auf gemeinsame Ordner haben und auch einen eigenen Ordner besitzen, auf denen nur sie selbst Zugriff haben. Alle benutzten Computer werden mit Passwörtern geschützt, die Mitarbeiter werden angewiesen nach Empfehlungen der IT-Dienste die Passwörter auf den von der Universität zur Verfügung gestellten Rechnern halbjährlich zu wechseln.

2.3 Wie lösen Sie Urheberrechtsfragen und Fragen des geistigen Eigentums?

Gemäß dem Vertrag der Universität Bern gehören alle Daten, die im Laufe dieses Projekts erstellt oder geändert wurden, der Universität Bern, mit Ausnahme der Urheberrechte des Erstellers. Dies bedeutet, dass für jede andere Verwendung als wissenschaftliche und akademische Zwecke eine Genehmigung der Universität Bern erforderlich ist (z. B. für Patentierung und Kommerzialisierung). Unsere Studie sieht keine Handels- und Patentfragen auf sich kommen. Sollte dies der Fall werden, werden wir uns an das institutionelle Amt für Technologietransfer wenden.

Die unterstützenden Daten werden unter einer offenen Creative Commons Attribution Lizenz veröffentlicht, speziell unter CC BY - Attribution (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Ausnahme: Die Rechte für die SHP-Daten liegen bei FORS in Lausanne.

3 Datenspeicherung und Datenerhalt

3.1 Wie werden Ihre Daten während der Forschungsarbeiten gespeichert und gesichert?

Die Daten des Projektes werden auf einem von den IT-Diensten der Universität Bern gestellten Cloudspeicher abgelegt und verwaltet. Wir richten uns in dieser Frage nach den Empfehlungen der Universität Bern, hier ein Ausschnitt aus der Beschreibung des Dienstes:
"Campus Cloud Storage ist eine von der Uni Bern betriebene Filesharing Lösung ähnlich wie Dropbox, Google Drive etc. Die Daten werden dabei ausschliesslich innerhalb der Uni Bern gespeichert. Mit Campus Cloud Storage gewährt die Universität Bern ihren Mitarbeitenden sicheren und einfachen Zugriff auf Dateien. Es bietet sich die Möglichkeit, gemeinsam vom Arbeitsplatz oder unterwegs an Dokumenten zu arbeiten und Kommentare auszutauschen. Die Dateien verbleiben stets auf den Servern der Universität Bern und werden täglich gesichert (Backup)."

Jedem Mitarbeiter stehen dabei 100GB Speicherplatz zur Verfügung, die nach Bedarf bis zu 1TB erweitert werden können.

3.2 Wie gestaltet sich Ihre Planung für den Datenerhalt?

Generell:

Veröffentlichte Daten: Depositions- und Open-Data-Daten über fachspezifische Repositories sorgen langfristig für Langlebigkeit der Daten. Hier bietet z.B. FORS eine Plattform namens FORSBASE an: <https://forsbase.unil.ch>. Das Projekt soll dabei schon in der Anfangsphase registriert und die Daten nach Abschluss der Publikationen der Fachöffentlichkeit zur Verfügung gestellt werden. Die einem Artikel zugrunde liegenden Teildaten werden zeitnah an die Veröffentlichung des Artikels öffentlich zugänglich publiziert werden.

Unveröffentlichte Daten: Wir speichern unveröffentlichte, hochwertige Enddaten, die während dieses Projekts generiert wurden, zur Wiederverwendung in unseren zukünftigen Projekten. Wir werden einige Daten für die Verwendung durch Forscher in zukünftigen Kollaborationen zur Verfügung stellen, wenn unser Institut sie nicht mehr verwendet. Daten schlechter Qualität werden am Ende des Projekts dauerhaft verworfen.

Offene Dateiformate, wie z.B. "comma separates values" oder "rich text format", können von jedem verwendet werden, da die Dateispezifikationen öffentlich verfügbar sind. Im Gegenteil, proprietäre Dateiformate, wie z.B. Microsoft Word ".docx" oder SPSS Data-Files ".sav" funktionieren nur mit Software, die vom Anbieter bereitgestellt wird, und wenn die Software nicht mehr unterstützt wird, sind Dateien in diesem Format in der Regel nicht lesbar. Daher werden wir, soweit möglich, universelle oder offene Dateiformate bevorzugen, um Daten zu teilen und zu archivieren. Das Überführen von proprietären Dateiformaten in offene Dateiformate gestaltet sich in der Regel beim Speichern als eine leicht zu bewerkstellende Prozedur.

Paneldaten: hier übernimmt der Projektpartner FORS alle Schritte im Zusammenhang mit dem Datenerhalt.

Interviewtranskripte: diese werden im Rich Text Format gespeichert und möglicherweise nach dem Projektende und nach erfolgreicher Anonymisierung zur Verfügung gestellt. Siehe 4.2. für die detaillierte Erklärung.

Fragebogendaten: diese werden im Comma Separated Value Format gespeichert und möglicherweise nach dem Projektende zur Verfügung gestellt. Siehe 4.2. für die detaillierte Erklärung.

4 Datenaustausch und Weiterverwendung von Daten

4.1 Wie und wo werden die Daten zugänglich gemacht?

Wir stellen wichtige Datensätze und ergänzende Dateien, die Publikationen begleiten, in geeigneten digitalen Daten-Repositories offen zur Verfügung z.B. bei BORIS (Bern Open Repository and Information System), die den Fair Data-Prinzipien entsprechen und von einer Non-Profit-Organisation verwaltet werden. Wir teilen spezifische Datensätze über domänenspezifische öffentliche Repositories. Solch eine Plattform bietet mit dem FORSBASE das FORS in Lausanne an.

Paneldaten sind bereits auf FORSBASE hinterlegt und werden jährlich aktualisiert.

Interviewtransskripte werden in FORSBASE hinterlegt werden.

Die Daten von Fragebogendaten werden ebenfalls in FORSBASE hinterlegt.

Datasets erhalten nach Möglichkeit einen Digital Object Identifier (DOI) und zugehörige Metadaten. Der DOI, der den Datasets im Repository entspricht, wird in die Referenzliste des Artikels aufgenommen, sodass alle Datasets in einer Publikation identifiziert und darauf zugreifen können.

Sichtbarkeit und Valorisierung von Datasets: Wir verknüpfen auch DOIs mit entsprechenden Datensätzen im Publikationsrepository der Universität Bern (BORIS), um die Sichtbarkeit des Datasets zu verbessern, sofern dem keine rechtlichen, ethischen, urheberrechtlichen oder andere Klauseln entgegenstehen. Metadaten zu Datasets sind öffentlich durchsuchbar und auffindbar und geben an, wie und zu welchen Bedingungen auf die Datasets zugegriffen werden kann. Wir zeigen Informationen über Datensätze auf der Webseite des Instituts, auf den Profelseiten von beteiligten Forschern sowie über ORCID-IDs der Forscher an, die die Sichtbarkeit der Datensätze erhöhen.

Die Daten in den Repositories werden in Übereinstimmung mit den Datenrichtlinien des Geldgebers gespeichert. Der Aufbewahrungszeitplan für Daten beträgt 10 Jahre ab dem Datum der Deposition in erster Instanz, wobei Erweiterungen auf Datensätze angewendet werden, auf die anschließend zugegriffen wird.

4.2 Sind bestimmte Einschränkungen erforderlich, um sensible Daten zu schützen?

Die Datei, die die Pseudonomisierung entschlüsselt, bleibt unter Verschluss bei dem Projektverantwortlichen für die Dauer der Existenz der Daten, mindestens jedoch 10 Jahre.

Bei Interviewtranskripten und Fragebogendaten handelt es sich um personenbezogene sensible Daten. Mit zunehmenden Möglichkeiten der Verlinkung von Daten, kann mittlerweile sogar von anonymisierten Daten mithilfe von computerisierten Mustererkennungen auf Personen geschlossen werden. Wir beobachten die technische Entwicklung und behalten uns vor, die Veröffentlichung entsprechend der Lage zum Schutz der Auskunftspersonen in ihrer Gänze abzulehnen und nur Teildaten zu veröffentlichen, die keinen Rückschluss auf Personen zulassen. Wir sind uns bewusst, dass die Güterabwägung der transparenten Forschung und des Personenschutzes nach den FAIR Data Principles erfolgen soll. Dies werden wir im Projekt gewährleisten.

4.3 Ich werde ausschliesslich digitale Archive wählen, die den FAIR Data Principles entsprechen

Ja

4.4 Ich werde digitale Archive wählen, die von einer gemeinnützigen Organisation verwaltet werden.

Ja