

DIE VORBEREITUNG VON SOZIALWISSENSCHAFTLICHEN DATEN FÜR SWISSUBASE: IN KÜRZE

Um Ihre Daten in [SWISSUbase](#) zu hinterlegen und sie anderen Forschenden zugänglich zu machen, müssen Sie Ihr Forschungsprojekt beschreiben und einen (oder mehrere) Datensatz(-sätze) hinzufügen.

Darüber hinaus müssen die Daten bestimmten Anforderungen genügen. Insbesondere sollte ein Datensatz mindestens eine **bereinigte, anonymisierte Datendatei** in einem der weiter unten **aufgeführten Formate** (siehe S. 2), sowie die für eine sachgemässe Interpretation und Wiederverwendung der Daten erforderliche **Dokumentation** enthalten.

Datenbereinigung

Quantitative Daten:

- Beschriftungen für Variablen und Werte müssen in der Datendatei enthalten sein. Labels können die Form von verkürzten Umfrage-Fragen haben. Die Verbindungen zwischen den Variablen in der Datendatei und den entsprechenden Fragen im Fragebogen sollten eindeutig sein.
- Für fehlende Werte sollten nach Möglichkeit explizite Codes vergeben werden (z.B. weiss nicht, Verweigerung, Filter, Nichtbeantwortung).
- Die Häufigkeiten sollten überprüft und festgestellte Unstimmigkeiten oder Fehler korrigiert oder entfernt werden (z.B. ungültige oder sehr unwahrscheinliche Werte).

Qualitative Daten:

- Informationen über den Kontext der Datenerhebung – wie Datum, Ort, eindeutige Kennung der interviewenden und interviewten Person(en) – sollten enthalten sein; entweder in der Datei (z.B. Transkripte) oder in einem separaten Dokument.

SWISSUbase

[SWISSUbase](#) ist eine Online-Plattform für die Archivierung und Verbreitung von Forschungsdaten und deren Metadaten. Es werden zudem verschiedene Dienstleistungen für die Verwaltung von Forschungsprojekten angeboten.

SWISSUbase ist ein Gemeinschaftsprojekt mehrerer Institutionen und orientiert sich an internationalen Archivierungsstandards, um die langfristige Aufbewahrung und Zugänglichkeit der Daten zu gewährleisten und den Bedürfnissen von Universitäten und Forschungseinrichtungen gerecht zu werden.

SWISS  **base**

Anonymisierung

Es liegt in der Verantwortung der Forschenden, ihre Daten vor der Hinterlegung zu anonymisieren, so dass keine Person ohne “unverhältnismässigen Aufwand” identifiziert werden kann. Dies erfordert eine sorgfältige Behandlung von direkten und indirekten Identifikatoren:

- **Direkte Identifikatoren** sind Variablen oder sonstigen Elemente, die eine direkte Identifizierung von Personen ermöglichen (z.B. Namen, Adressen, Telefonnummern). Sie sollten entfernt oder durch Pseudonyme oder Kategorien ersetzt werden.
- **Indirekte Identifikatoren** sind Variablen oder sonstigen Elemente, welche, in Kombination, eine Person identifizieren können/zur Identifikation einer Person führen können (z.B. Alter, Beruf, Postleitzahl). Sie sollten ebenfalls überprüft und gegebenenfalls umkodiert, oder gelöscht, werden.

Falls eine Anonymisierung nicht möglich ist, ist die ausdrückliche Zustimmung der Teilnehmenden zur Archivierung und Weitergabe der Daten erforderlich. Eine Kopie der Einverständniserklärung sollte in diesem Fall miteingereicht werden.

SWISSUbase bietet verschiedene Optionen an, um die Kontrolle über den Datenzugriff zu behalten: Nutzungseinschränkungen, vorherige Zustimmung, Embargo.

Formate

Je nach Art der Daten werden die folgenden Formate in SWISSUbase angenommen:

- **Tabellarische Daten:** SPSS Data File & Portable Formate (.sav; .por), MS Excel (.xls; .xlsx), Comma-Separated value File (.csv).
Dateien wie Stata-Daten (.dta), Statistical Analysis System (.sas), R Dateiformat (.RData; .rds), OpenDocument spreadsheet (.ods) werden ebenfalls akzeptiert, sollten aber von Datenproduzierenden in SPSS, Excel oder CSV konvertiert werden, um die Datenqualität und -integrität nach der Konvertierung zu gewährleisten.
- **Textdaten:** Portable Document Format (.pdf), Rich Text Format (.rtf), ASCII (.txt), MS Word (.doc; .docx), OpenDocument Text (.odt).
- **Bilddaten:** JPEG (.jpeg; .jpg; .jp2), TIFF (.tif; .tiff), RAW-Bildformat (.raw), Bitmap-Bilddatei (.bmp), Portable Network Graphics (.png), Portable Document Format (.pdf).
- **Audiodaten:** MP-Format (.mp3), FLAC (.flac), Waveform Audio Format (.wav), Advanced Audio Coding (.aac).
- **Videodaten:** MP-Format (.mp4), QuickTime-Video (.mov), AVI-Datei (.avi), Motion JPEG 2000 (.mj2).

Ordner und Dateien in akzeptierten Formaten können in ZIP-Format (.zip; .rar; .7z) hinterlegt werden.

Dokumentation

Die Hinterlegung sollte alle Unterlagen enthalten, die für die sachgemässe Interpretation der Daten und die Gewährleistung ihrer Qualität erforderlich sind. Dies umfasst:

Ein Bericht, der die folgenden Aspekte beschreibt:

- Projekt: Kontext, Problematik, Ziele, Hypothesen.
- Methode: Grundgesamtheit, Stichprobe (Umfang, Rücklaufquote), Datenerhebungsmethode.
- Daten: Bereinigung, Anonymisierung, Kodierung, Validierung, bei Bedarf Beschreibung der umkodierten oder konstruierten Variablen, und Gewichtung.

Dokumente im Zusammenhang mit der Datenerhebung:

- Materialien, die den Befragten im Voraus zugesandt wurden (z.B. Unterlagen);
- Anweisungen oder Materialien für die Interviewenden;
- Materialien, die den Befragten während des Interviews vorgelegt wurden;
- Instrumente, die für die Datenerhebung verwendet wurden (Fragebögen, Raster, Protokolle).

Falls vorhanden, fügen Sie bitte auch Folgendes hinzu:

- Publikationen und Abschlussberichte;
- ein Codebuch;
- Kodierungsanweisungen;
- Syntaxen;
- und einen Data Management Plan.

Dokumentationsdateien sollten im PDF (mit Ausnahme der Syntaxen) und in den verfügbaren Sprachen eingereicht werden.

Weitere Informationen und Ressourcen

- [SWISSUbase Ressourcen](#)
- [Data Management](#) (FORS-Webseite)
- [FORS Data protection policy](#) (FORS-Webseite)
- [The informed consent as legal and ethical basis of research data production](#) (FORS Guide n°5)
- [How to draft a DMP from the perspective of the social sciences, using the SNSF template](#) (FORS Guide N°07)
- [Data anonymisation: legal, ethical, and strategic considerations](#) (FORS Guide n°11)
- [Data Management Expert Guide: Anonymisation](#) (CESSDA-Empfehlungen für die Anonymisierung von Daten)

Brauchen Sie Hilfe?

dataservice@fors.unil.ch