

PRÉPARER VOS DONNÉES POUR LEUR DÉPÔT DANS SWISSUBASE

Pour déposer vos données dans [SWISSUbase](#) et les rendre accessibles à d'autres chercheurs-euses, vous devez décrire votre projet de recherche et y ajouter un (ou plusieurs) jeu(x) de données.

Par ailleurs, les données doivent répondre à certaines exigences. En effet, un jeu de données doit être composé au minimum d'un fichier de **données nettoyées, anonymisées** et dans un des **formats listés plus bas**, ainsi que de la **documentation** nécessaire à la bonne interprétation et à la réutilisation des données.

Nettoyage des données

Données quantitatives:

- Les labels de variables et de valeurs doivent être inclus dans le fichier de données. Les labels peuvent prendre la forme des questions de l'enquête raccourcies. Les liens entre les variables dans le fichier de données et les questions correspondantes du questionnaire doivent être clairs.
- Des codes explicites doivent être attribués aux valeurs manquantes (par ex. ne sait pas, refus, filtre, non réponse), lorsque cela est possible.
- Les fréquences doivent être contrôlées et les inconsistances ou erreurs détectées doivent être corrigées ou supprimées (ex. valeurs invalides ou hautement improbables).

Données qualitatives:

- Des informations sur la récolte des données - telles que la date, le lieu, l'identifiant unique de l'interviewer et de la personne interviewée - doivent être disponibles soit dans le fichier (par ex. transcriptions), soit dans un document séparé.

SWISSUbase

[SWISSUbase](#) est une plateforme qui permet l'accès aux données de recherche et offre une solution et des services pour la gestion de projets de recherche ainsi que pour l'archivage, la diffusion et la promotion des données de recherche et leurs métadonnées.

Issue de la collaboration de plusieurs institutions, SWISSUbase s'appuie sur des normes d'archivage internationales pour garantir la préservation et l'accessibilité des données à long terme et répondre aux besoins des universités et autres organismes de recherches.

Anonymisation

Il est de la responsabilité des chercheurs-euses d'anonymiser leurs données avant le dépôt, de manière à ce qu'aucun individu ne puisse être identifié sans "efforts disproportionnés". Cela passe par un traitement rigoureux des identifiants directs et indirects:

- Les **identifiants directs**, c'est-à-dire, toutes les variables ou autres éléments qui permettraient d'identifier directement des individus (par ex. noms, adresses, numéros de téléphone) doivent être supprimés ou remplacés par des pseudonymes ou des catégories.
- Les **identifiants indirects**, qui, combinés, pourraient permettre d'identifier un individu (par ex. âge, profession, code postal), doivent également être contrôlés et recodés ou supprimés si nécessaire.

Dans le cas où il n'est pas possible d'anonymiser les données, un **consentement** explicite des participant-e-s est nécessaire pour l'archivage et le partage des données. Un exemplaire du formulaire de consentement doit alors être fourni.

SWISSUbase permet également de restreindre l'accès aux données (par ex. exclure l'enseignement) et/ou d'ajouter un accord préalable de la part de l'auteur-e.

Formats

En fonction du type de données, les formats suivants sont à privilégier dans SWISSUbase:

- **Données en tableaux:** SPSS Data File & Portable Data file (.sav ; .por), MS Excel (.xls ; .xlsx), Comma-Separated value (.csv).
Les fichiers tels que Stata data file (.dta), Statistical Analysis System (.sas), OpenDocument spreadsheet (.ods) sont également acceptés mais devraient être convertis en SPSS, Excel ou CSV par les producteurs-trices des données ceci en vue d'assurer la qualité des données et leur intégrité suite à leur conversion.
- **Données textuelles:** Portable Document Format (.pdf), Rich Text Format (.rtf), ASCII (.txt), MS Word (.doc ; .docx), OpenDocument text (.odt)
- **Données constituées d'image:** JPEG (.jpeg ; .jpg ; .jp2), TIFF (.tif ; .tiff), RAW image format (.raw), Bitmap Image File (.bmp), Portable Network Graphics (.png), Portable Document Format (.pdf)
- **Données audio:** MP Format (.mp3), FLAC (.flac), Waveform Audio Format (.wav), Advanced Audio Coding (.aac)
- **Données vidéo:** MP Format (.mp4), QuickTime video (.mov), Fichier AVI (.avi), Motion JPEG 2000 (.mj2)

Les dossiers et fichiers dans des formats acceptés peuvent être déposés en zip (.zip, .7z, .rar).

Documentation

Le dépôt devrait inclure toute documentation nécessaire à la bonne interprétation des données et permettant d'assurer leur qualité, à savoir:

Un rapport décrivant les aspects suivants:

- Le projet: contexte, problématique, objectifs, hypothèses
- La méthode: population, échantillonnage (taille, taux de réponse), mode de collecte des données
- Les données: nettoyage, anonymisation, codage, validation, le cas échéant description des variables recodées ou construites, pondération

Les documents relatifs à la récolte des données:

- Le matériel envoyé à l'avance aux répondant·e·s (par ex. courriers préliminaires)
- Les instruments utilisés pour la récolte des données (questionnaires, grilles, protocoles)
- Le matériel présenté aux répondant·e·s durant les entretiens
- Les instructions ou matériel pour les enquêtrices et enquêteurs

Si disponibles, veuillez aussi ajouter:

- Les publications et rapports finaux
- Le codebook
- Les instructions de codage
- Les syntaxes
- Le Data Management Plan

Les fichiers de documentation doivent être soumis **au format PDF** (à l'exception des syntaxes) et dans les différentes langues disponibles.

Informations et ressources

- [SWISSUbase](#)
- SWISSUbase user guide
- [Data Management](#)
(site internet de FORS)
- [Politique FORS en matière de protection des données](#)
(site internet de FORS)
- [The informed consent as legal and ethical basis of research data production](#)
(publication FORS Guide)
- [Data anonymisation: legal, ethical, and strategic considerations](#)
(publication FORS Guide)
- [Data Management Expert Guide: Anonymisation](#)
(Recommandations CESSDA pour l'anonymisation des données)

Besoin d'aide?

Contactez-nous!

dataservice@fors.unil.ch