



Le partage de données de recherche en Sciences Humaines: exemples pratiques

Philippe Schneider, PhD

Haute école de santé Vaud HESAV – HES-SO

8/09/2026 - Open Science Days - UNIL

Pour me présenter

- Bachelor en psychologie – Université de Genève, Suisse (2013)
- Master en neuropsychologie – Université Libre de Bruxelles, Belgique (2017)
- Doctorat en psychologie cognitive – Université de Fribourg, Suisse (2022)

- Actuellement:
Chargé RaD à la Haute école de santé Vaud (HESAV – HES-SO) (2026)

- Anciennement:
 - Chargé de recherche, Haute école de de travail social et de la santé de Lausanne (HETSL- HES-SO) (2023 – 2026)
 - Chargé de recherche, Ecole La Source (HES-SO) (2022 – 2024)

Expérience personnelle

- Toutes mes données de thèse déposées sur OSF
- Participation au Swiss Reproducibility Network
- Dépôt de données en sciences infirmières sur SwissUbase
- Travail actuellement sur un projet de dépôt de données en science occupationnelle (ergothérapie)
- → Principalement avec données expérimentales et quantitatives, facilement anonymisable

Pourquoi déposer ses données de recherches?

- Scientifiques :
 - Meta analyses futurs simplifiées
 - Transparence pour la publication scientifique
 - Analyses secondaires par d'autres chercheurs
- Ethiques et matériels :
 - Données produites par financement publique → les données doivent être accessible
 - Les données stockées en local peuvent se perdre (changement d'ordinateur, matériel perdu ou cassé, etc.)
 - Requis par bailleurs de fonds (FNS, HORIZON) et par de nombreux journaux scientifiques

Limites au dépôt de données

- Données sensibles
- Balance entre pertinence scientifique et anonymisation
- Accord des participants pour le dépôt de données

FAIR research data

➤ Findable

- Données liées à un DOI; accompagnées de métadonnées complètes; déposé sur une plateforme indexée

➤ Accessible

- Peuvent être retrouvées grâce à leur identifiant

➤ Interoperable

- Données en format accessible, ouvert si possible; vocabulaire claire

➤ Reusable

- Données décrites de façon complète; publiées avec une licence; répondent aux normes pertinentes pour le domaine

Données de recherche

- Dans mon cas: format .csv
- Car:
 - Petits datasets (max 10'000 lignes, 100 colonnes)
 - Format ouvert, facilement lisible par tous les tableurs et logiciel de traitement statistique (R, python, excel, openTable, etc.)
- Importance d'utiliser un format « long » ou « large » de façon consistante
- Préférence pour stocker les données brutes plutôt que transformées
 - MAIS: Attention aux données sensibles
- Limitations: grand datasets, peu flexible, formatage automatique par Excel

Exemple dataset

Fichier

Accueil

Insertion

Dessin

Mise en page

Formules

Données

Révision

Affichage

Automatiser

Aide

Acrobat

Coller

Presse-papiers

Aptos Narrow

11

G

I

S

A

A

Police

Alignement

Nombre

Mise en forme conditionnelle

Mettre sous forme de tableau

Styles de cellules

Cellules

Édition

Sensibilité

Compléments

Créer un PDF et partager le lien

Créer un PDF et le partager via Outlook

Adobe Acrobat

A1

:

<

Metadonnées

- Document(s) associé(s) au dataset
- Contient une description générale de chaque dataset, de la méthodologie, puis une description claire du contenu
- Description du type de donnée
 - Quantitative continue: portée (minimum – maximum)
 - Quantitative discrète: toutes les valeurs possibles
 - Qualitative nominale: liste des valeurs possibles
 - Qualitative ordinale: liste des valeurs possible et leur ordre

Exemple codebook

Dictionary explaining the meaning of each column in the output of experiments 1A, 1B, 2A and 2B

Index: trial number (starting from 0)

NumSujet: Subject number

Age: age (in years)

Lateralite: Hand laterality (D=right hand; G=left hand, Ambi: both hands)

Sexe: sexe of the participant (M=male, F=female)

Span: Number of images presented in the trial

ImgType (Exp. 1B, 2A, 2B): Type of image used in the trial (NR=Non-Real, Real=Real)

Block (Exp. 1A only): Type of image used in the block (NR=Non-Real, Real=Real)

Pace (Exp. 1A and 1B only): Cognitive load condition ("LENT" =low cognitive load, "RAPIDE" = High Cognitive load condition)

OrdreBlock (Exp. 1A and 1B only): Block order (1= Real then Non-Real; 2= Non-Real then Real)

OrdreList (Exp. 1B only): Which list was attributed to which cognitive load condition (1: List1= low cognitive load, List2= high cognitive load; 2: List1= high cognitive load, list2=Low cognitive load)

Resp[J]: Participant answer to the Jth image of the trial

RTresp[J]: RT to the Jth response (since the start of the response phase in exp. 1A and 1B, since the last answer in exp. 2A and 2B)

RepCorr[J]: The correct answer for the Jth image of the current trial

RepCorrTrial: 1 if every image in the current trial recalled in the correct position, 0 otherwise

Où déposer ses données?

- ▀ Osf.io (deprecated)
 - ▀ Problème: arrêt du dépôt de données à partir de novembre 2026, toujours utilisable pour pré-registrations
- ▀ SwissUbase.ch
 - ▀ Gratuit pour chercheur affilié à une institution Suisse
 - ▀ Équipe interne de soutien au dépôt
 - ▀ Permet le dépôt de données quantitatives et qualitatives
- ▀ Zenodo.org
- ▀ Hors plateforme
 - ▀ Sur son propre site, mais demande infrastructure spécifique et indexation
 - ▀ Risque si site disparaît

Exemple de SwissUbase

≡
Projet
Aperçu
Description générale
Financement
Références bibliographiques
Méthodes
Historique des versions
Jeux de données
Fichiers

Fichiers

↓ Réf.	Titre	Type	Lié à	Type de documentation
32125	Belgium - T0: Codebook (FR)	Documentation	Dataset 2665	Codebook
32124	Belgium - T0: Cross-sectional data	Données	Dataset 2665	
32123	Canada - T0-T1-T2: Codebook (FR)	Documentation	Dataset 2664	Codebook
32122	Canada - T0-T1-T2: Longitudinal data	Données	Dataset 2664	
32121	Portugal - T0-T1-T2: Codebook (PT)	Documentation	Dataset 2663	Codebook
32120	Portugal - T0-T1-T2: Longitudinal data	Données	Dataset 2663	
32119	Switzerland German-speaking - T0-T1-T2: Codebook (DE)	Documentation	Dataset 2662	Codebook
32118	Switzerland German-speaking - T0-T1-T2: Longitudinal data	Données	Dataset 2662	
32117	France - T0-T1-T2-T3: Codebook (FR)	Documentation	Dataset 2661	Codebook
32116	France - T0-T1-T2-T3: Longitudinal data	Données	Dataset 2661	

Éléments par page 10 1 - 10 de 12 < > >|

Bon à savoir

- Plus simple de directement intégrer le dépôt de données dès la création d'un projet, plutôt que de le faire à la fin
- Attention au type de données:
 - Les données « sensibles » ou protégées par la loi ne peuvent pas être déposées en accès libre. D'autres solutions existent pour le partage et le dépôt de ce type de données, mais demande une infrastructure légale lourde
- Les «data management plan» sont maintenant souvent demandés par les bailleurs de fonds, il est important d'y inscrire un plan sur le partage de données
- Accord à-priori avec participants pour que les données soient stockées en ligne (inclus dans consentement éclairé)

Ressources

- Portail « Open (FAIR) research data » de la HES-SO
 - <https://hes-so.libguides.com/open-research-data-software/introduction>
- Swiss Reproducibility Network
 - <https://www.swissrn.org/>
- Data.org
 - <https://data.org/resources/the-fair-data-principles/>
- Center for Open Science
 - <https://www.cos.io/>