

Rendre accessible et visibiliser ses données et ses codes

Retours sur une expérience d'entreposage

Plan

- Introduction: du cas particulier au regard sur les normes
- Des normes variables?
 - Deux types de données, deux temporalités
- Des défauts de normes?
 - Des zones troubles juridiques et scientifiques
- Des normes rationalisables?
 - Tenants et aboutissants d'un engagement atypique

Du cas particulier au regard sur les normes

- Retour sur une expérience d'auto-dépôt: une analyse sur N=1
- Les rouages d'une **expérience déroutante**
 - ❖ Un objectif politique de normalisation du partage des données (Galonnier, Le Courant, Pecqueux & Noûs, 2019)
 - ❖ Une « culture de la revisite » (Scot, 2006) dès les années 1990 et, au tournant du XXI^e siècle, la création de vastes entrepôts de données
 - Le projet Dataverse Network aux États-Unis (King, 2007)
 - La plateforme BeQuali en France (Duchesne & Garcia, 2014).

Du cas particulier au regard sur les normes

- Retour sur une expérience d'auto-dépôt: une analyse sur N=1
- Les rouages d'une **expérience déroutante**
 - ❖ En SHS, une démarche portée par les équipes de soutien à la recherche
 - ❖ Parmi les chercheur·es qualitativistes, tradition d'archivage (Mouton, 2001; Müller & Wolikow, 2006).
 - ❖ Le partage des données ethnographiques se heurte à des résistances (Revelin et al., 2021)

→ **Une confusion qui signifie quelque chose**

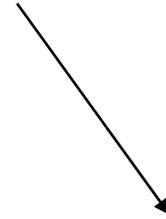
Que traduit la confusion d'une jeune chercheuse
qui ouvre certaines de ses données
sur les normes de cette activité?



Des normes
variables?



Des défauts de
normes?



Des normes
rationalisables?

Deux types de données, deux temporalités

- L'enquête
- Une recherche doctorale (CRIS-LIEPP, Sciences Po)
- L'empreinte du handicap sur les positions sociales à l'âge adulte
- Méthodes mixtes:
 - ❖ Exploitation de l'EEC et de son module *ad-hoc* 2011 (Insee/Dares)
 - ❖ Réalisation d'entretiens biographiques de décembre 2019 à juin 2020

Deux types de données, deux temporalités

- Le volet qualitatif et sa mise en ligne
- Des données primaires, à caractère personnel
- Forte anticipation et formalisation
 - ❖ Novembre 2019: mesures dès l'amont de la campagne d'entretien
 - ❖ Fiches d'information, recueil du consentement et trace matérielle

Deux types de données, deux temporalités

- Le volet quantitatif et sa mise en ligne
- Des données secondaires, déjà anonymisées
- Improvisation et flexibilité
 - ❖ Juillet 2021 : un support pour des excroissances de la thèse?
 - ❖ Printemps 2022 : délimitation post-soutenance des contenus versés

Des zones floues juridiques et scientifiques

- La protection de droits mal circonscrits
- Le paradigme d'ouverture des données de la recherche s'arrête là où commencent les droits de tiers (Stérin & Noûs, 2019)
 - ❖ Droit à la protection des données personnelles
 - ❖ Droit de propriété intellectuelle
- Les limites des normes formelles (RGPD, conventions d'utilisation)
- ... et informelles (déontologie, e. g. Stérin, 2018)

Des zones floues juridiques et scientifiques

- Les dilemmes des catégorisations
- La transparence scientifique plutôt que la « brutisation » (Denis & Goëta, 2017)
- Les nécessaires transformations et réductions d'information
 - ❖ Entretiens: risques de réidentification → l'anonymisation
 - ❖ Analyses statistiques: synthèse et priorisation → le codage
- Catégories d'analyse situées et marges de recatégorisation

Tenants et aboutissants d'un engagement atypique

- L'amorce de l'engagement
- Entre conditions de possibilité et injonctions
 - ❖ Début de thèse en 2018
 - une mise à l'agenda politique de la science ouverte
 - ❖ Laboratoire interdisciplinaire
 - une circulation des paradigmes propice à l'ouverture des données
 - ❖ Directrice de thèse sensible à ces enjeux
 - une relation hiérarchique qui prescrit et qui rend possible

Tenants et aboutissants d'un engagement atypique

- Le maintien de l'engagement
- Une rationalité en valeurs plutôt qu'en finalité?
 - ❖ Des centaines d'heures de travail gratuit // un débouché incertain
 - ❖ Des valeurs connues: « culture du libre », partage aux collègues...
 - ❖ ... et des enjeux de mise à l'agenda
 - 1/ valeurs scientifiques : développement d'un champ de recherche
 - 2/ valeurs déontologiques : un contre-don aux personnes rencontrées
 - 3/ valeurs politiques : *disability mainstreaming*

Tenants et aboutissants d'un engagement atypique

- La visibilisation de l'engagement
- De la valorisation des données à celle de leur mise en ligne
 - ❖ Le rôle proactif du soutien à la recherche
 - ❖ ... face à la tiédeur de la recherche



Jeux de données EHDS - usages et besoins



Questions

Réponses

6

Paramètres

- ❖ « Documente-le »: du *data paper* à l'article méthodologique

Merci beaucoup pour votre attention et merci
d'avance pour vos questions !

celia.bouchet@sciencespo.fr

Pour retrouver ce diaporama: <https://celiabouchet.hypotheses.org/files/2022/12/2022-Bouchet-DataSHS.pdf>

Pour les jeux de données: <https://data.sciencespo.fr/dataverse/ehds>

Bibliographie (1/2)

- Denis, J., & Goëta, S. (2017). Les facettes de l'Open Data : émergence, fondements et travail en coulisses. In *Big data et traçabilité numérique* (pp. 121–138). Collège de France.
- Duchesne, S., & Garcia, G. (2014). beQuali : une archive qualitative au service des sciences sociales. In *Archives de la recherche. Problèmes et enjeux de la construction du savoir scientifique* (pp. 35–56). Paris: L'Harmattan.
- Galonnier, J., Le Courant, S., Pecqueux, A., & Noûs, C. (2019). Ouvrir les données de la recherche ? *Tracés*, (19), 17–33.
- King, G. (2007). An Introduction to the Dataverse Network as an Infrastructure for Data Sharing. *Sociological Methods and Research*, 36, 173–199.
- Mouton, M.-D. (2001). Archiver la mémoire des ethnologues. *Gradhiva : Revue d'histoire et d'archives de l'anthropologie*, 30(1), 67–72.

Bibliographie (2/2)

- Müller, B., & Wolikow, S. (2006). Sciences sociales: Archives de la recherche. *Genèses*, 63(2), 2–3.
- Revelin, F., Levain, A., Mignon, M., Noel, M., Queffelec, B., Raux, P., & Squividant, H. (2021). *L'ouverture des matériaux de recherche ethnographiques en question*. Centre national de la recherche scientifique.
- Scot, M. (2006). Les archives britanniques des sciences sociales. Deux études de cas: UK data archive (UKDA) et qualidata. *Genèses*, 63(2), 46–65.
- Stérin, A.-L. (2018). Diffuser des données de la recherche dans le respect du droit et de l'éthique. In *La diffusion numérique des données en SHS - Guide de bonnes pratiques éthiques et juridiques* (pp. 19–29). Aix-en-Provence: Presses universitaires de Provence.
- Stérin, A.-L., & Noûs, C. (2019). Ouverture des données de la recherche : les mutations juridiques récentes. *Tracés*, (19), 37–50.