



CELLULE DATA STEWARDSHIP RAPPORT ENTRETIENS CHERCHEUR·EUSES DOMAINE TRAVAIL SOCIAL

Corinne Dallera, Data steward – Octobre 2024

TABLE DES MATIÈRES

| | |
|---|----|
| I. INTRODUCTION..... | 1 |
| II. PRATIQUES EN MATIÈRE DE GESTION DES DONNÉES | 3 |
| III. BESOINS ET SOLUTIONS | 10 |
| IV. CONCLUSION ET RECOMMANDATIONS | 11 |
| V. RÉFÉRENCES BIBLIOGRAPHIQUES..... | 12 |

I. INTRODUCTION

Grâce à un soutien de swissuniversities, le Dicastère Recherche et Innovation de la HES-SO (Dicastère R&I) a engagé en juin 2023 cinq data stewards, dont la principale mission était celle de mener des entretiens qualitatifs avec des chercheur·euses et des personnes de référence en matière d'Open Research Data (ORD). L'objectif était d'établir une première cartographie des pratiques ainsi que des enjeux/problématiques liées à ORD et plus généralement à la gestion des données, tout en identifiant les spécificités de chaque domaine¹. C'est à ce titre que j'ai été engagée pour le domaine Travail social (domaine TS)². Ces entretiens qualitatifs visaient également à compléter les résultats de l'enquête quantitative conduite en amont par le Dicastère Recherche et Innovation de la HES-SO (Delamadeleine, 2023).

Le rapport qui suit est centré sur les données recueillies lors de neuf entretiens menés avec des chercheur·euses des quatre Hautes écoles du domaine Travail social de la HES-SO (HETS), durant le premier semestre 2024 et qui avaient pour but de compléter les informations recueillies lors de

¹ Je remercie chaleureusement toutes les personnes interviewées pour le temps qu'elles m'ont accordé.

² En contrat de durée déterminée à 30% jusqu'à fin décembre 2024.





cinq entretiens réalisés au cours du deuxième semestre 2023 avec les membres du Groupe Open Data et éthique du domaine TS (Groupe OD et éthique)³.

Les entretiens ont duré entre 45 minutes et 1h30 et ont principalement eu lieu par Teams. Avec l'accord des participant·es, ils ont été enregistrés et partiellement retranscrits. Afin de garder une cohérence avec les entretiens menés par les autres data stewards, la grille d'entretien élaborée par Gregory Rauber⁴ et Constance Delamadeleine⁵ a guidé dans une large mesure ces entretiens, tout comme c'était le cas pour les entretiens réalisés avec les membres du Groupe OD et éthique. L'objectif étant également d'approfondir des problématiques particulières identifiées en amont, certains entretiens ont été axés sur des points particuliers de la grille. Cinq chercheur·euses interviewé·es ont été recruté·es à la suite d'une prise de contact par les membres du Groupe OD et éthique, quatre chercheur·euses l'ont été par moi-même dans le cadre de mon réseau professionnel personnel.

Ces entretiens visaient à recueillir de manière exploratoire des expériences diverses face aux enjeux liés à la gestion, la protection et au partage des données du point de vue des spécificités des méthodologies et des types de recherche menées au sein du domaine TS, mais aussi en termes de carrière et conditions de travail. Un·e doctorant·e, un·e collaborateur·trice scientifique, deux adjoint·es scientifiques, un·e professeur·e associé·e et quatre professeur·es ordinaires ont été interviewé·es⁶. Les méthodologies utilisées par sept personnes interviewées sont très majoritairement qualitatives, même si elles mobilisent également des données quantitatives dans leurs recherches. Deux personnes interviewées sont aguerries aux méthodes mixtes. Concernant les méthodologies qualitatives mobilisées par les chercheur·euses interviewé·es, elles couvrent une large palette de démarches, soit : des entretiens semi-directif, des entretiens en marchant⁷, des prises d'images et de sons (par les participant·es à l'enquête ou l'équipe de recherche), des focus groupes, des immersions participantes, des questionnaires, de l'analyse documentaire, etc. Ces démarches impliquent une multitude de matériaux : croquis, carnets de note, documents existants (format numérique ou papier) plus ou moins confidentiels⁸, images, sons, retranscriptions, statistiques, etc.

Ce document vise à nourrir les réflexions internes en vue de l'élaboration de la stratégie et du plan d'action ORD de la HES-SO et peut également servir de support pour alimenter d'autres processus de planification stratégique au sein des Hautes écoles et/ou du domaine.

³ Les membres du Groupe OD et éthique sont les personnes référentes au sein des quatre HETS concernant toutes les questions en lien avec le cycle de gestion des données ainsi que l'ORD. Le Groupe a été formellement constitué en 2022 par le Comité R&I du domaine TS dans le but de mieux coordonner les initiatives et la circulation des informations entre les quatre HETS au niveau du Domaine TS.

⁴ Data Steward pour le domaine Musique & Arts de la scène et le domaine Design & Arts visuels de la HES-SO

⁵ Chargée de projet Open Data à la HES-SO.

⁶ Pour plus de détails sur les implications en termes de cahiers des charges formels des interviewé·es, voir : *Règles communes pour le personnel d'enseignement et de recherche. Typologie des fonctions*. Adoptée par le Comité gouvernemental le 20.11.2014.

⁷ « Walking interviews ».

⁸ Soit des lois, des directives internes, des statistiques administratives (déjà produites ou que l'équipe de recherche demande de produire).





II. PRATIQUES EN MATIÈRE DE GESTION DES DONNÉES

2.1 Propriété des données

Les données des projets évoqués au cours des entretiens appartiennent à trois types de régime de propriété des données⁹ :

- la propriété des données appartient à l'institution ;
- la propriété des données est partagée entre deux institutions académiques de deux pays différents ;
- la propriété des données appartient au mandant.

La question de la propriété des données n'a toutefois pas été systématiquement abordée et approfondie dans les entretiens.

Les exemples ci-dessous illustrent les flous actuels entre propriété des données, propriété intellectuelle et droits d'utilisation avec lesquels les chercheur·euses sont amené·es à devoir jongler.

Les contrats institutionnels avec les mandants d'une HETS indiquent que les données appartiennent de plein droit au mandant et que le mandataire peut utiliser les analyses dans l'enseignement, les publications ou à des fins de recherche sur la base d'un accord du mandant. La question de la propriété intellectuelle des données des chercheur·euses n'étant pas clairement stipulée dans le contrat écrit, cela pourrait constituer un obstacle de réutilisation des données par les équipes menant des recherches mandatées au sein de cette HETS, par exemple pour des publications ou pour le partage sur un dépôt de données.

L'expérience de la personne en parcours de thèse interviewée illustre les enjeux liés à l'auteurat en termes de rapports de pouvoir et de pratiques disciplinaires différentes. Dans le cadre du projet de doctorat évoqué et dont le terrain est hospitalier, un·e médecin a exigé de pouvoir cosigner tous les articles en lien avec le projet de doctorat, une pratique qu'une Commission cantonale d'éthique a considérée comme inacceptable. Un·e autre médecin qui avait la même exigence est revenu·e en arrière sur sa position après négociation avec la direction de thèse, mais par contre ne s'est pas investi·e dans le recrutement de participant·es à la recherche.

Un entretien souligne que les recherches participatives dans lesquelles les participant·es sont amené·es à créer des matériaux ou des produits artistiques soulèvent des questions importantes concernant la répartition des droits d'auteurs entre les chercheur·euses et les participant·es.

2.2 Gestion des données de recherche

Sensibilité(s) des données

Les neuf chercheur·euses interviewé·es collectent ou produisent des données personnelles et dans la majorité des cas sensibles. Telle qu'elle est décrite lors des entretiens, cette sensibilité va au-delà

⁹ Ces régimes juridiques ne concernent pas la propriété intellectuelle des données.





de la définition d'une donnée sensible selon la Loi sur la protection des données (LPD)¹⁰, comme l'illustrent les extraits d'entretien suivants :

« On a des données qui permettent de reconnaître la personne. Sensibles des fois au sens de la loi, mais aussi sensibles au sens sociologique du terme. C'est donc très sensible parce que parfois au sein d'une institution, on travaille avec différents groupes : Ne pas monter les uns contre les autres. [Des données] sensibles au sens où ça peut porter préjudice à quelqu'un, du coup il y en a tout le temps dans le travail social. » (E11)

« C'est surtout sensible en termes de thèmes qui portent sur l'intimité. On n'a pas de données juridiques qui mettraient des personnes en danger. C'est plutôt qu'elles nous font confiance sur le fait qu'on ne va pas divulguer leur vie privée. [...] Ça peut être des données de famille. Elles n'aimeraient pas que leurs enfants sachent certaines choses. Rien ne pourrait les mettre dans une situation illégale, mais dans une mauvaise posture par rapport à des travailleurs sociaux. Les personnes ne voudraient pas que l'on divulgue leur point de vue personnel sur les travailleurs sociaux. » (E8)

Anonymiser au risque d'une « compréhension désincarnée »

La sensibilité des données soulève des défis importants pour rendre les matériaux les plus anonymes possibles sans trop désincarner les situations :

« Donc le défi est que l'on ne puisse pas reconnaître les gens parce que l'on fera tout ce qui est en notre pouvoir pour qu'ils ne soient pas reconnus sauf si on cherche vraiment. Mais on a remarqué que les gens cherchent. On doit essayer de noyer l'information. Et parfois, on ne peut pas utiliser des informations très intéressantes parce que cela permettrait trop facilement de remonter aux personnes. » (E11)

C'est notamment le cas dans le cadre d'évaluations d'institutions ou de focus groupes où la démarche consiste à discuter dans chaque groupe de ce qui a été dit dans l'autre groupe et lors de terrains de recherche où l'interconnaissance est très forte.

Ce défi est également important lors de la restitution des résultats dans des valorisations, mais aussi dans l'enseignement ou à des partenaires de terrain comme le souligne cet extrait d'entretien :

« Par exemple quand on présente notre recherche, les gens qui connaissent un peu la montagne aimeraient avoir un peu plus de précisions : Est-ce que c'est un village à 1'500 mètres ? Combien il y a d'habitants pour pouvoir imaginer un peu la situation ? Ce qui est légitime, mais nous on ne peut pas faire. Il y a toujours cette demande de justification et moi je ne sais jamais trop quoi dire sans dévoiler. Et si tu décontextualises, tu ne comprends pas grand-chose. » (E8)

¹⁰ LPP Art. 5 al. c. entend par « données personnelles sensibles (données sensibles) : les données sur les opinions ou les activités religieuses, philosophiques, politiques ou syndicales ; les données sur la santé, la sphère intime ou l'origine raciale ou ethnique ; les données génétiques ; les données biométriques identifiant une personne physique de manière univoque ; les données sur des poursuites ou sanctions pénales et administratives, les données sur des mesures d'aide sociale. »





La confrontation avec les protocoles éthiques standardisés

De manière plus générale, un enjeu important souligné par les entretiens est que les protocoles éthiques standardisés visant la protection des sujets de recherche et qui sont construits sur des approches hypothético-déductives sont incompatibles avec la grande majorité des recherches menées au sein du domaine TS, soit des recherches majoritairement qualitatives, dont les questionnements sont construits au fur et à mesure de l'avancement de la recherche de manière itérative. Les exigences d'une Commission d'éthique concernant le traitement d'enregistrements d'entretiens et de leur retranscription décrites par une personne interviewée témoignent également de potentielles contradictions entre éthique des Commissions d'éthique et éthique des chercheur·euses en sciences humaines et sociales (SHS).

« Et puis [la Commission d'éthique] a demandé par rapport aux retranscriptions d'entretien que si j'allais modifier a posteriori mes retranscriptions ou mes observations, il fallait que je mette tout en mode 'suivi-révision' et qu'à côté je tiens un journal qui documente toutes les modifications que j'aurais faites sur mon journal de terrain ou sur mes retranscriptions en argumentant pourquoi je l'ai fait, pour vraiment garantir la fiabilité des données etc., en vue de garantir la reproductibilité des résultats. Pour la Commission d'éthique, il faut garder une trace de toutes les modifications pour garder un certain contrôle sur les échantillons. Du coup, j'ai mis ça en place, mais ça n'a aucun sens pour moi et cela va même à l'encontre de ma conception de l'éthique. Par exemple, il y a des professionnel·les que j'interroge où il y a des gros problèmes au niveau de l'institution et [les personnes interviewées] m'ont dit : "OK, tu peux enregistrer ça". Mais après coup, je me dis : "non j'ai plus envie que ça apparaisse quelque part". Mais théoriquement, je devrais quand même garder une trace. » (E12)

Stockage et sécurisation des données numériques et numérisées

Les entretiens avec les chercheur·euses confirment que les formats de données numériques ou numérisées des recherches du domaine TS ne soulèvent généralement pas d'enjeux particuliers d'un point de vue du stockage IT, à l'exception du poids de certains fichiers vidéo ou de bases de données importantes. Les entretiens réalisés montrent cependant qu'en termes de volume et de « puissance de calcul pour l'analyse » des solutions ont pu finalement être trouvées avec les services informatiques des HETS.

Partager les données/matériaux et échanger avec les membres de l'équipe de recherche

Le partage des données/matériaux avec les membres de l'équipe de recherche peuvent s'avérer complexes, notamment dans des projets comprenant des terrains situés dans plusieurs pays et impliquant des personnes de l'équipe de recherche sous contrats vacataires, comme l'illustre l'extrait ci-dessous :

« On a plusieurs systèmes : en termes de lieu protégé, on utilise Switchdrive et il y a un backup sur un dossier protégé à la HETS où ils ont fait des dossiers, sur-sur-protégés. Je fais un backup régulièrement, vu qu'il n'y a pas de backup sur Switchdrive. Et sinon, on utilise Google Drive pour tout ce qui n'est pas dangereux, comme si on utilisait Teams, car comme certains membres de l'équipe sont sous contrats vacataires, ils n'ont pas accès à Teams. Mais c'est vraiment pour les données qu'on pourrait publier, on va dire. On a encore un autre endroit pour le chat de tous les jours, mais dans lequel on n'écrit rien de sensible, on a SLACK. C'est un tchat gratuit et tous les 30 jours, 60 jours, ça efface tout. » (E8)





Retranscriptions

La pratique des personnes interviewées est de faire retranscrire leurs entretiens par des personnes qui sont soit engagées au sein de la HETS comme vacataires, soit externes sous contrats de confidentialité. Les retranscriptions effectuées par des logiciels utilisant des clouds soulèvent par contre des questions éthiques, notamment lorsqu'il a été annoncé aux personnes interviewées que seul·es les membres de l'équipe auront accès aux données. Dans plusieurs entretiens, il est mentionné que l'utilisation de ces logiciels existe au sein des HETS et que les personnes qui recourent à ces services ne sont pas toujours conscientes des risques en termes de confidentialité des données de cette pratique.

Quelques enjeux liés à la gestion des données quantitatives

Un entretien montre la complexité de la gestion des données quantitatives, une complexité probablement sous-estimée au niveau du domaine TS, dont la majorité des chercheur·euses s'inscrivent dans des démarches qualitatives :

« Parce qu'on a demandé un appariement de six jeux de données, dont quatre appartiennent à l'OFS, mais relèvent de trois départements différents, de trois enquêtes et deux viennent de la Centrale de compensation. Mais là, il y a un enjeu. Il a fallu déjà définir dans chaque jeu de données, quelles variables on voulait obtenir. Donc cela implique de devoir avoir des échanges avec chacune des personnes qui s'occupent de ces bases de données. Et puis après faire une demande spéciale d'appariement parce que la Centrale de compensation envoie les données à l'OFS, puis l'OFS s'occupe d'appareiller tous les jeux de données à partir du numéro AVS. Ensuite l'OFS réanonymise le tout pour que l'équipe de recherche ne puisse pas identifier les personnes. Et l'OFS nous livre chaque jeu de données avec un identifiant anonymisé, mais en s'assurant que la personne soit bien la même et apparaisse bien dans chaque jeu de données. Et après nous on reçoit tout ça et on doit remettre ensemble les fichiers. [...] Donc ça, c'est un temps assez effrayant pour réunir tout ça. » (E14)

2.3 Partage des données

Le partage des données n'est pas une pratique courante des chercheur·euses du domaine. Seules deux personnes interviewées ont déposé des données et une personne a tenté de le faire. La démarche n'a toutefois pas abouti en raison de la complexité de la désidentification des données, comme relaté dans l'entretien :

« Dans ce gros projet avec les collègues [pays de l'UE], on a voulu aussi faire de l'open data. On a préparé, on a payé une personne qui nous a aidé à tout mettre en forme, mais en fait c'était beaucoup trop compliqué. Pour l'instant c'est toujours dans le tiroir et c'était un gros investissement en temps et en argent qui pour l'instant n'a pas vraiment abouti. Donc on en est un petit peu là dans cette expérience : classification des données, des fiches, des partages, classifications pour qu'on retrouve les chemins pour les personnes qui sont dans la recherche et puis d'être conforme à une arborescence. Cette arborescence permet ensuite de créer un corpus de données avec des entretiens en bout de chaîne. Mais les entretiens en bout de chaîne, quand on les écoute, si on les écoute en fait, c'est beaucoup trop identifiable quand c'est une donnée brute, donc voilà, il y a la voix déjà, donc on oublie l'audio. Après, on peut retomber sur la retranscription. Et puis la retranscription, il faudrait caviarder beaucoup de choses. Ça rejoint ce que je disais tout à l'heure : jusqu'où on assume le fait que ce soit partagé. Le rôle du chercheur, pour moi, c'est de transformer la donnée d'une certaine manière pour que ce soit restituable mais suffisamment ancré, sans qu'on puisse forcément reconnaître où identifier directement la personne. Alors que si on partage la donnée brute, on oublie tout ça, en fait. Et, j'ai eu des échanges avec [un spécialiste d'un dépôt de données spécialisé SHS] à





un moment donné. Et puis effectivement, vu les obstacles, ça paraissait malheureusement très difficile de partager ces données. » (E13)

L'impossibilité d'anonymiser totalement les données sensibles

La plupart des personnes interviewées mentionnent l'impossibilité d'anonymiser totalement les données sensibles dans l'objectif de les partager.

« Donc j'ai toujours rempli des DMP en disant qu'il y aurait peut-être des résumés, mais qu'en tout cas, on ne va jamais déposer des enregistrements, ça c'est sûr. Peut-être déposer des résumés anonymisés. Mais en même temps c'est jamais totalement anonymisable. A la fin, on finissait toujours par dire : "on ne dépose pas". » (E14)

« Mais après si tu floutes tout le temps, c'est une catastrophe, ça ne donne plus rien » (E6)

Le temps important et les compétences que nécessitent la préparation des données en vue du partage

Les personnes ayant partagé leurs données ou une partie de leurs données évoquent toutes le temps de travail très important que cela a engendré. La démarche implique également des compétences particulières que les membres de l'équipe de recherche n'ont peut-être pas et le fait que la tâche ne peut pas être entièrement externalisée dans la mesure où il faut également connaître les enjeux de la recherche.

« Ça demande un travail stratosphérique. Il faut tout automatiser pour garantir la sécurité des données et pour les mettre dans le bon format, il faut faire des scripts. Si tu fais ça à la main, t'es sûr que tu te plantes. Il faut tout documenter et faire des métadonnées et si tu le fais tout à la main, tu es sûr que tu dois tout refaire. » (E11)

« Donc quand on a retranscrit les entretiens pour les déposer sur [dépôt de données spécialisé en SHS] où il fallait encore plus anonymiser, on a anonymisé le nom des personnes et on a aussi anonymisé le nom des entreprises sociales. On a anonymisé aussi les cantons pour pas qu'on puisse faire des recoupements, parce que dans certains cantons, il y a peu d'entreprises sociales. Pour que le canton ne puisse pas être reconnu, quand la mesure était trop spécifique au canton, on a aussi anonymisé. [Les retranscriptions ayant été faites par deux personnes qui ont été engagées sur des contrats vacataires], on a dû tout reprendre derrière [...]. C'est la collaboratrice scientifique qui a relu les entretiens et qui a ajouté cette confidentialité sur le lieu, sur le canton, sur l'entreprise sociale. Et aussi du côté de [dépôt de données spécialisé en SHS], la collaboratrice a aussi vérifié que la façon d'anonymiser soit OK. » (E10)

Le manque de moyens financiers pour effectuer ce travail

Les entretiens soulignent le manque de moyens financiers pour la préparation des données en vue du partage. C'est particulièrement le cas pour les recherches mandatées car il n'est pas possible de faire porter ces coûts à certains mandants qui n'en n'ont pas les moyens (p ex de petites associations) et pour celles qui en auraient les moyens (p. ex des administrations), cela pourrait les encourager à donner le mandat à une autre institution proposant une offre moins coûteuse.





Le partage des données, notamment « brutes » n'est pas adapté aux méthodologies des sciences humaines et sociales

De manière plus générale les personnes interviewées considèrent le partage de données brutes comme une injonction inappropriée aux méthodologies de récolte et de construction des matériaux propres à leur champ de recherche.

« C'est aussi qu'on a un peu l'impression que ça nous tombe dessus et que ça vient des sciences médicales. [...] Et comme notre manière de faire de la recherche n'est pas la même, on n'est pas tout à fait préparé-e. [...] En tant qu'ethnologue, sociologue, on a une sensibilité sur le fait que ces données, elles valent quelque chose et qu'on ne peut pas faire n'importe quoi, notamment parce qu'on a vu des gens et ce n'est quand même pas la même chose que quand tu calcules des cellules ou autres. » (E9)

« Si on est dans le quali, c'est aussi de se dire : qu'est ce qui s'est joué autour de ces entretiens, qu'est-ce qu'ils m'ont dit implicitement. Dans la sociologie que je fais, on est vraiment dans du quali et pas dans une logique d'accumulation. [...] Avoir deux-cents entretiens en reprenant ceux des autres, ce n'est pas ça qui se joue. C'est ça, je m'intéresse au vivant et ce vivant il ne peut pas être capté dans une retranscription. » (E7)

« Au fond dans une équipe de recherche on travaille sur des données qui ont été collectées par d'autres. La grande différence c'est qu'on collectivise. On échange sur les contenus, la dynamique, sur les interprétations. Donc on travaille vraiment en équipe pendant un bon moment là-autour. Mais ça, c'est absent de la réutilisation de données brutes. Tout ce travail d'équipe dans une temporalité donnée, ce n'est pas juste quelques séances, c'est sur deux ans, trois ans, tout ce travail-là ne peut pas être substitué par la métadonnée. » (E13)

ORD et durabilité

La pertinence de l'ORD est également questionnée par rapport au contexte climatique actuel et à la condition de chercheur·euse :

« Pour moi cela ferait plus sens dans une perspective écologique de se donner les moyens de valoriser plutôt que de stocker des données à tout prix. [...] Je n'adhère pas à ce mouvement de tout stocker sans une réflexion de fonds plus macro sur le pourquoi on fait tout ça ; [...] avoir du recul sur le prendre soins du vivant actuellement, le prendre soins des chercheur·euses, des chercheur·euses précaires. Il y aurait de l'énergie à mettre plus sérieusement sur nos conditions de vie actuelles, plutôt que de se soucier des futur·es chercheur·euses dans cinquante ans. Quel est le sens de stocker, sachant que moi-même mes anciennes données, je ne les exploite pas parce que je n'ai pas le temps. » (E7)

En raison des difficultés évoquées ci-dessus, la majorité des chercheur·euses interviewé·es n'ont pas partagé ou prévu de partager leurs données sur des dépôts de données.

« Au sein de notre école on a été parmi la première équipe [soumise au DMP]. J'ai trouvé très intéressant quand il y a eu cette première table ronde. Parce qu'il y avait [X, spécialiste engagé par un dépôt de données institutionnel SHS reconnu par le FNS] qui nous expliquait les enjeux. Cela nous interrogeait sur nos rapports aux données. Il y avait beaucoup de crispations autour du partage, de résistance. Moi je m'étais dit "pourquoi pas", mais par rapport au contenu de nos données, c'était juste inapproprié. » (E7)





« Alors dans le cadre des mandats, absolument pas, ça s'est sûr. Et dans le cadre du projet, on a clairement dit dans le DMP que les données ne seront pas partagées parce que, simplement, le contrat avec l'OFS contredit ce principe. » (E14)

D'autres n'ont partagé ou n'envisagent de partager qu'une partie des données de leur recherche.

« J'ai fait un DMP sur le modèle qu'avait fait X. Et donc j'ai mis que les données ne seraient pas totalement accessibles à tous les chercheur·euses parce qu'on est dans un milieu d'interconnaissance où les lieux sont proches et que les gens ont forcément des liens communautaires, familiaux et sociaux. » (E9)

Il est intéressant de noter que ce n'est pas le partage « en soi » des données/matériaux/résultats de leurs recherches qui pose problème aux chercheur·euses du domaine TS interviewé·es. D'autres modalités de partage ont en effet été évoquées, notamment à des stades plus avancés de l'analyse.

Par exemple, dans un projet quantitatif, il est prévu et annoncé au FNS – qui l'a accepté – que les syntaxes qui serviront aux analyses seront transmises aux administrations concernées afin qu'elles-mêmes puissent reproduire des analyses sur la base de ces syntaxes dans l'objectif de faire un monitoring.

D'autres formes de transmissions qui sont considérées comme plus ouvertes par rapport à la société et plus démocratique que l'ORD ont également été mentionnées :

« Cela fait réfléchir au statut de cet open data. Moi, je suis plus favorable à soutenir des outils de médiation scientifique, plutôt que l'open data, en fait. C'est-à-dire mettre des moyens pour que ces données brutes se déclinent selon différentes modalités pour aller à la rencontre d'un public plus large. Parce que dans le fonds, qui va s'intéresser à une donnée brute à part un chercheur ? donc on pourrait chercher à diversifier, par exemple à faire des textes qui soient plus accessibles, [...] cela pourrait être sur des réseaux sociaux, dans des éditions intermédiaires pour les écoles, etc. » (E13)

Par ailleurs, les chercheur·euses ne remettent pas en cause le partage de données sur des dépôts de données pour certains corpus qui s'y prêtent :

« Mais attention, je mets des précautions aussi. Il y a peut-être des projets de recherche qui se prêtent très bien, avec un périmètre très clair : il y a le questionnaire et il y a tout qui est en place. A ça, je suis tout à fait favorable, par contre. » (E13)

2.4 Réutilisation des données

Aucune personne interviewée n'a réutilisé des données d'autres chercheur·euses, notamment parce qu'elleux ne verraient pas très bien l'utilité qu'elleux feraient de données brutes récoltées par d'autres équipes de recherche dans leurs projets de recherche :

« J'ai fait des entretiens, mais ce qui m'a aidée à chaque fois pour comprendre les enjeux, c'est toute l'agitation qu'il y a eu autour. [...] La créativité se fait par le frottement avec le terrain, donc je n'aurais pas un très grand élan [à consulter et réutiliser les données des autres]. La créativité scientifique, elle se joue sur des trucs, les gens qui se lâchent un peu à la fin d'un entretien, des choses comme ça. Donc non, pour ces raisons qui sont scientifiques, j'aurais plutôt envie de me donner les moyens de ça avec mes propres entretiens. » (E7)





La réutilisation de données dans d'autres types de démarches que les leurs n'est toutefois pas exclue :

« Cela pose la question de la réutilisation de ces corpus. Il faut donc une géométrie variable. Il y a des corpus qui sont très très bien définis. Peut-être assez petits, précis, pointus. Et on peut plonger dedans et on peut réutiliser les données. Pourquoi pas, je n'exclus pas ça, mais en tant qu'anthropologue avec une approche plutôt réflexive, située et contextualisée, j'ai un peu de peine à me dire que quelqu'un d'autre va se projeter dans le travail que j'ai fait. » (E13)

Les conditions dans lesquels travaillent les chercheur·euses actuellement sont également évoquées comme un obstacle majeur à la réutilisation des données :

« Je n'ai pas été consulter des données. Mais peut-être que cela pourrait avoir un intérêt si je pouvais consulter des données récoltées dans un autre endroit [avec des analogies du terrain du/de la chercheur·euse interviewé·e]. Je pourrais peut-être les mettre en discussion avec les miennes. Mais pour ça, on n'a pas d'argent, on n'est pas payé·e pour ça. On ne peut pas déposer un projet autour de ça. Dans la réalité de la recherche ce n'est pas possible car ce serait un nouveau projet de recherche, ce ne serait pas pour compléter ma recherche. » (E9)

Une personne interviewée a reçu – à sa plus grande surprise – une demande de pouvoir consulter les données qu'elle a déposées :

« Une personne qui m'a demandé des données et finalement c'était une question de recherche assez éloignée de moi. Elle s'intéressait à comment différents professionnels – qui peuvent être du travail social ou non – parlent de leur travail d'insertion. C'était pas du tout ma question. Moi ma question, c'est comment ils placent, où ils placent, qu'est-ce qu'ils font pendant ce placement. Ce n'est pas comment ils parlent de leur travail. Mais finalement avec sa question de recherche, je pense qu'elle peut choper des données qui n'ont rien à voir avec mon objet initial et qu'elle peut trouver des choses qui peuvent lui être utiles. Je n'avais pas du tout perçu une utilisation comme ça des données. Ça ne porte pas forcément sur les mêmes fins, mais sur d'autres trucs qui peuvent émerger. [...] Ce n'était pas ma question, mais la personne, elle va lire mes entretiens avec cette grille-là. » (E10)

Il s'agit de la seule demande d'accès à des données déposées par des chercheur·euses de domaine TS à ma connaissance.

III. BESOINS ET SOLUTIONS

Un aspect frappant concernant les entretiens menés est qu'à la question « *De quels types de services ou soutiens auriez-vous besoin concernant la gestion des données de vos projets de recherche ?* », très peu de demandes ont été formulées, notamment par les chercheur·euses expérimenté·es.

Les jeunes chercheur·euses tendent à verbaliser davantage de besoins en termes de soutien. Par exemple, au niveau technique, le souhait est exprimé d'avoir d'avantage d'information sur les outils, logiciels et instruments disponibles au sein de leur HETS et d'avoir un appui plus important au niveau informatique. Il est également mentionné que de disposer d'une documentation sur les enjeux juridiques liées aux images serait très utile. Dans un entretien, l'opacité de l'information institutionnelle est soulignée et la personne souhaiterait pouvoir disposer d'un document synthétisant quel service pourrait apporter de l'aide et à quelle étape.





La personne menant un projet de doctorat interviewée a souligné avoir été particulièrement démunie face aux exigences des commissions d'éthiques auxquelles elle a dû soumettre son projet :

« Je me rends compte que dans mon parcours académique, j'avais été très peu formé·e sur ces questions éthiques. On est formé à la question de se situer en tant que chercheur·euses. À la subjectivité, mais la thématique de l'éthique en tant que telle, elle est peu problématisée. Et je me suis senti·e très peu outillé·e et il y a eu un bon moment où j'avais envie de chercher un emploi ailleurs tellement ça m'impactait au niveau de la motivation, où j'avais l'impression que je ne comprenais plus ce que je faisais. [...] J'aurai eu besoin d'être coaché·e. [...] D'être coaché·e, mais par des personnes qui viennent des sciences sociales. [...] Il faudrait une prise de position institutionnelle par rapport à ces commissions d'éthique et aussi par rapport à ces questions du co-auteurat. Il faut des discussions institutionnelles, pas juste portées par des individus. » (E12)

Parmi les autres souhaits de soutien exprimés, notons encore celui de pouvoir avoir un soutien juridique spécialisé dans le domaine de la recherche en SHS.

IV. CONCLUSION ET RECOMMANDATIONS

Les entretiens réalisés complètent les informations récoltées dans le cadre des entretiens menés avec les membres du Groupe OD et éthique et confirment les difficultés et défis soulevés par les types de démarches, les types de participant·es aux enquêtes et les types de données/matériaux produits dans le cadre de la majorité des recherches menées au sein du domaine TS. La très grande majorité des recherches menées par les chercheur·euses interviewé·es se heurte en effet à des injonctions contradictoires, d'une part, entre le respect de normes éthiques standardisées ayant été définies à partir de démarches exogènes à celles des recherches en SHS et, d'autre part, à des incitations à partager des « données brutes » interopérables et reproductibles, des incitations qui ne sont pas compatibles avec les démarches réflexives, itératives et processuelles adoptées dans la majeure partie des recherches menées au sein du domaine TS. Les personnes interviewées expriment également de sérieux doutes quant au sens de partager des données/matériaux décontextualisés du cadre interprétatif *in situ* de la production de ces données/matériaux et de leur(s) interprétation(s). Elleux-mêmes ne réutiliseraient pas les données d'autres chercheur·euses pour les mêmes raisons. Concernant le partage et la réutilisation des données, un bilan quantitatif du nombre de sets de données qualitatives en sciences humaines réutilisés de part du FNS est attendu. Les entretiens réalisés font ainsi largement écho à différentes prises de positions (SAA, 2021) et réflexions critiques sur l'ORD menées par des chercheur·euses en SHS (p.ex : Burton-Jeangros, 2017 ; Roca i Escoda, 2020 ; Perrin & al., 2020 ; Amiotte-Suchet, 2022 ; Siméant-Germanos, 2022).

Cela dit, dans le cas de certain corpus bien définis, déposer ses données de recherche est pertinent aux yeux des personnes interviewées. La diffusion d'exemples concrets de réutilisations de sets de données pourrait inciter certain·es chercheur·euses à partager leurs données via des dépôts reconnus comme « FAIR » pour les SHS. Il est également apparu lors des entretiens que peu de chercheur·euses étaient informé·es de la possibilité offerte par certains dépôts de données d'opter pour un « partage restreint ». Une meilleure communication à ce propos pourrait écarter certaines appréhensions quant à un mauvais usage des données déposées. Au niveau « pratico-pratique », des documents courts et centrés sur des questions précises pourraient être utiles, par exemple, un document synthétisant les principales bases juridiques sur l'utilisation des images. Concernant la sécurité informatique, les chercheur·euses du domaine ne semblent pas être bien informé·es des risques en termes de protection des données, mais aussi de perte de leurs données. Des recommandations sous forme de guideline pourraient être utiles. Aucune des HETS ne dispose de





services juridiques spécialisés sur des questions de recherche, notamment en SHS. Un soutien juridique au niveau du Rectorat pourrait s'avérer précieux. Enfin, lors des deux entretiens réalisés avec des jeunes chercheur·euses, il est apparu qu'elleux n'ont pas toujours trouvé des interlocuteur·trices en mesure de répondre à leurs questionnements concernant des questions juridique en lien avec des recherches mobilisant des démarches innovantes ou face aux instances de régulation éthique de la recherche (commission d'éthique, p. ex). Il pourrait être intéressant de réfléchir avec l'équipe « Relève et carrières académiques » du Rectorat au soutien qui pourrait être apporté aux jeunes chercheur·euses par rapport aux défis que soulèvent la gestion et le partage des données/matériaux de recherche.

V. RÉFÉRENCES BIBLIOGRAPHIQUES

Amiotte-Suchet Laurent (2022). « Les données qualitatives, que peut-on en FAIR ? : les recherches qualitatives au défi de l'open data ». *Traverse*. 1, pp. 63-68.

Burton-Jeangros Claudine (éd.) (2017). *L'éthique (en) pratique : la recherche en sciences sociales*. Genève: Université de Genève (Sociograph - Sociological Research Studies, 34).

Delamadeleine Constance (2023). *Gestion et ouverture des données de recherche à la HES-SO Enquête sur les pratiques et besoins des chercheuses et chercheurs*. Delémont : HES-SO.

Perrin Julie, Bühler Nolwenn, Berthod Marc-Antoine, Forney Jérémie, Kradolfer Sabine, Ossipow Laurence (2020). « En quête d'éthique. Dispositions légales et enjeux empiriques pour l'anthropologie ». *Tsantsa : revue de la Société suisse d'ethnologie*. 25. p. 225-242.

Roca i Escoda Marta, Burton-Jeangros Claudine, Diaz Pablo et Rossi Ilario (éds.) (2020). *Enjeux éthiques dans l'enquête en sciences sociales*. Genève : Université de Genève (Sociograph - Sociological Research Studies, 45).

Siméant-Germanos Johanna (2022). « Qui protéger, consentir à quoi, enquêter comment ? ». *Genèses. Sciences sociales et histoire*. 129 (4), pp. 66-87.

Swiss Anthropological Association (SAA) (2021). *Open Science and Data Management in Anthropological Research Position Paper of the Swiss Anthropological Association (SAA)*. Adopted at the SAA Annual Meeting, held on-line on November 4, 2021.

