

Livrable 8 – Synthèse des réflexions méthodologiques

Objectifs du livrable 8 – Le partage des données de la recherche est de plus en plus encouragé, en raison d'objectifs scientifiques, méthodologiques, etc. (cf. phase 1, livrable 2) ; *a fortiori*, l'analyse secondaire des données de la recherche devrait se développer ; si c'est une démarche relativement pratiquée autour des données quantitatives, elle l'est moins pour les données qualitatives ; de fait, il existe relativement peu de retours d'expérience d'analyse secondaire de données qualitatives, notamment dans la littérature francophone (hormis le numéro spécial de Réanalyse, publié en 2017). Dans ce livrable 8, nous souhaitons donc faire un retour sur notre/nos propre(s) expérience(s) d'analyse secondaire, que nous analyserons d'un point de vue méthodologique.

1. Différentes démarches « d'analyse secondaire », de « retour dans l'enquête »

De manière générale, *l'analyse secondaire* concerne la « réutilisation de données (notamment qualitatives) préexistantes produites lors de recherches précédentes » (Heaton, 2008). S. Duchesne (2017), qui rejoint M. Hammersley (1997) et défend une position « entre positivisme et constructivisme »¹, considère néanmoins que les données ne peuvent être analysées séparément de leurs contextes (historique, scientifique) de production. Selon elle, une analyse secondaire ne consiste donc pas uniquement en une analyse secondaire de données existantes, mais bien de l'enquête dans sa globalité ; une analyse des données produites, mais également des hypothèses et problématiques de recherche, méthodes employées, grilles d'entretien, résultats, etc.

La *réanalyse* telle que la définit Duchesne (2017) consiste en une « analyse secondaire à laquelle participent les (ou certains des) chercheurs primaires » ; on parlera de réanalyse notamment lorsqu'un chercheur mène une deuxième analyse de ses matériaux, plusieurs années après l'analyse primaire, comme l'a fait Hamidi (2017). Selon Heaton (2008), sur les expériences d'analyse secondaire qu'elle avait répertoriées en 2004 (à partir du corpus Qualidata), un auteur ou plus avait été impliqué dans la recherche primaire dans 86% des cas ; autrement dit, 86% des expériences d'analyse secondaire étaient en réalité de la réanalyse. La réanalyse diffère de l'analyse secondaire notamment en ce qui concerne la connaissance des matériaux qu'aura le chercheur primaire : connaissance qui facilitera l'appropriation des données par les chercheurs secondaires et limitera les difficultés liées au « *not having been there* » (Heaton, 2008), en permettant les différentes recontextualisations nécessaires à la réanalyse (Belot & Van Ingelgom, 2017).

On distinguera par ailleurs les opérations de *mutualisation*, mise en commun de données produites séparément par les chercheurs participant à l'opération ; une situation de mutualisation peut par exemple s'observer lors de programmes de recherche impliquant différents chercheurs, qui collectent des données de leur côté et les mettent en commun au moment de l'analyse. Dans le cadre d'une mutualisation, tous les chercheurs sont à la fois primaires (ou « premiers »), fournisseurs de données, et secondaires, « receveurs » de données existantes : les chercheurs se trouvent donc dans une relation de réciprocité et de « bienveillance » vis-à-vis des données : elle permet notamment de limiter les craintes envers une « mauvaise utilisation » des données (même si la littérature sur le sujet – notamment Bishop, 2013 – illustre à partir de plusieurs exemples cette bienveillance entre chercheurs, y compris lorsqu'il y a discordance, voire divergence, entre analyses primaire et secondaire).

¹ A noter que les définitions de positivisme et constructivisme adoptées par Duchesne, que je relaye ici, ne correspondent pas à celles que nous avons retenues dans le livrable 2. Fondationalisme (nous) = positivisme (Duchesne); relationnelle (nous) = constructivisme (Duchesne) ; constructivisme (nous) = voie intermédiaire (Duchesne, Hammersley)

2. Le méta-projet, un contexte favorable à l'analyse secondaire

J. Heaton (2008) identifie deux freins à l'analyse secondaire : le problème du *data fit* (adéquation des données avec les nouvelles questions de recherche) et le problème du *not having been there* (absence des chercheurs secondaires lors de la première enquête). Au regard de ces deux freins, nous montrons ci-dessous que le méta-projet constitue un contexte favorable à l'analyse secondaire.

a. De la possibilité de former un méta-corpus et de comparer les données

Avant de traiter de la possibilité de réaliser une analyse secondaire des données du méta-projet, il nous semble important de revenir sur les données analysées.

Les données dont nous disposons sont issues de six programmes de recherche indépendants et ont donc été collectées et produites dans des contextes différents, scientifiques d'une part, historiques d'autre part (même si dans une moindre mesure, puisqu'à une dizaine d'années d'écart au maximum). Aussi, la question se pose d'une part de la possibilité de comparer ces données ; d'autre part, de mutualiser ces données et de former un « méta-corpus », ou de conserver les six corpus originels.

Belot & Van Ingelgom (2017) ont rencontré une situation similaire en voulant comparer des données produites à 10 ans d'écart, par des chercheurs différents, et autour de questionnements ressemblants mais pas identiques. Pour les chercheuses, certaines similarités entre les corpus rendent possible leur mutualisation et leur comparaison : les similarités liées à « *un point de départ commun, un corpus de littérature commun, une même approche méthodologique, un même rapport au terrain, ainsi qu'une même méconnaissance des enquêtés avant les entretiens* ».

Concernant les données qualitatives dont nous disposons pour le méta-projet (environ 250 entretiens semi-directifs), nous pouvons mentionner une cohérence d'ensemble qui a beaucoup compté dans la genèse du projet et nous est apparue comme une condition favorable pour la mutualisation et la réanalyse. En premier lieu, les données abordent des thématiques et répondent à des interrogations similaires, centrées sur l'analyse des modes d'habiter : parcours résidentiels, pratiques de mobilité et modes de vie au quotidien. En outre, elles s'inscrivent dans une démarche qualitative et ont été collectées *via* des méthodes comparables (entretiens semi-directifs), dans des territoires et avec des individus relativement similaires (ménages résidant dans des territoires périurbains ou ruraux d'Indre et Loire). Nous partageons également des références bibliographiques communes, nuancées toutefois par nos différentes appartenances disciplinaires (géographie, aménagement, sociologie) et nos approches (qualitatives pour la majorité d'entre nous, mais plutôt quantitatives pour d'autres). Enfin, tous les chercheurs appartiennent, ou ont appartenu, au même laboratoire de recherche et avaient déjà des expériences communes de recherche et/ou d'enseignement.

Il nous semble donc envisageable de mutualiser, partager entre nous ces données, et de réaliser une analyse secondaire commune : les similarités évoquées garantissent une compréhension des données par les chercheurs et des échanges possibles entre chercheurs qui se connaissaient bien en amont.

Si la réanalyse d'un ensemble de données mutualisées nous a semblé possible et riche de promesses scientifiques, nous n'envisageons pas pour autant de constituer un « méta-corpus », c'est-à-dire de mener une réanalyse globale sur l'ensemble des matériaux collectés (par exemple, en considérant les 250 entretiens comme une base de données cohérente à partir de laquelle il serait possible de mener une analyse globale). En effet, les échanges entre membres de l'équipe du méta-projet nous ont amenés au constat selon lequel il est impossible de séparer la donnée (le discours, la trace GPS) de son contexte de production spécifique (le terrain d'étude, la grille d'entretien, le contexte d'enregistrement) pour l'analyser « correctement ». Il est donc impossible de constituer un « méta-

corpus » regroupant l'ensemble des entretiens et données GPS, qui ferait disparaître les spécificités des contextes de production. Il est inenvisageable de mélanger les entretiens et de traiter (par quelque méthode que ce soit, papier-crayon ou CAQDAS) un méta-corpus de manière globale, ce qui nous ferait perdre leurs contextes spécifiques, nécessaires à leur bonne compréhension. Nous devons donc travailler à l'échelle de chaque corpus originel (ou de sous-corpus au sein des corpus originels).

Cette hétérogénéité du matériau nous amène à réfléchir, pour chaque axe d'analyse, à ce qui est comparable ou non, ou, dit autrement, aux conditions et possibilités de mise en comparaison et de traitement conjoint des données. Les traces GPS fournissent un bon exemple. Les traces GPS sont des enregistrements pendant plusieurs jours des déplacements quotidiens d'individus résidant dans un territoire périurbain ou rural. Ces enregistrements ont été réalisés lors de trois sessions d'enregistrement

(au cours de trois programmes de recherche), à l'aide de matériels similaires ; en revanche, deux différences importantes sont à noter : 1/ elles ont été recueillies sur des durées très différentes (de 4 à 14 jours) ; 2/ la fréquence d'enregistrement varie selon les traces, de 3 pt/s pour certaines à 1 pt/s pour d'autres. Il est donc a priori impossible de mener des analyses statistiques sur l'ensemble des enregistrements directement, sans traitement préalable : il a donc fallu réfléchir aux conditions de mise en comparaison de ces traces (cf livrable Axe 1).

A noter que la question se serait posée de la même manière si nous avions souhaité opérer une analyse quantitative des entretiens semi-directifs : un tel traitement nécessiterait une préparation en amont du matériau, et nécessiterait de descendre au niveau des contextes d'énonciation et de la signification des mots énoncés, ce qui ne peut se faire sans intégrer le contexte de production de la donnée.

b. Adéquation des données aux problématiques de recherche

Les données à notre disposition peuvent-elles être réutilisées pour mener une analyse secondaire ? Sont-elles pertinentes pour répondre aux problématiques de recherche que nous avons identifiées ? (Ces questions de recherche sont détaillées dans le livrable 6)

Pour qu'une donnée soit « adéquate », un premier critère concerne la « parenté » thématique qui doit exister entre les données et les questions de recherche. (S'il peut sembler évident, ce critère n'est néanmoins pas toujours vérifiable, notamment lorsqu'on ne connaît pas avec précision les données : c'est pourquoi Ziebland & Hunt (2004) proposent que, dans le cas où un chercheur demande à avoir accès à des données, il puisse avoir accès à quelques exemples pour juger de leur pertinence.)

Hinds et al (1997) et Long-Sutehall et al (2010) évoquent ensuite un critère de diversité, de richesse et de niveau de détail des données primaires : « *enough [has to be] said in the primary transcripts about the topic of interest so that it would be reasonable to assume that the secondary research questions can be answered* » (Long-Sutehall et al, 2010). Enfin, Long-Sutehall et al (2010) évoquent les similarités qui doivent exister entre les questions de recherche primaire et secondaire : « *it is recommended that the research questions for the secondary analysis be sufficiently close to those of the primary research* ».

Qu'en est-il des données du méta-projet et de leur adéquation avec nos questions de recherche ?

Concernant les entretiens semi-directifs, on note tout d'abord une parenté thématique évidente entre les données et les problématiques de recherche que nous souhaitons investiguer. Un objectif principal du méta-projet est de tester (conforter, nuancer, confronter) des résultats obtenus au cours des six programmes « primaires » en soumettant aux questions de recherche initiales un panel élargi, plus copieux ou situé dans différents contextes. D'autres questions de recherche du méta-projet constituent

des approfondissements de problématiques primaires, en continuité parfaite avec les questions initiales.

C'est par exemple le cas de l'axe 3 sur les changements de pratiques de mobilité, pour lequel les questions de recherche primaire et secondaire sont très proches, voire similaires. Ainsi, l'une des questions posées concerne la compréhension du déroulement du processus de changement, qui avait été posée aux individus enquêtés au cours d'une recherche (programme Modalter), et sert d'axe d'analyse pour les entretiens analysés dans le cadre de l'axe 3. Les résultats de l'analyse secondaire nous permettront donc de valider, invalider, aller plus loin dans les résultats primaires.

L'adéquation du matériau aux problématiques de recherche se pose par contre avec acuité pour les questions de recherche pour lesquelles le matériau n'a pas été produit initialement : dans le cas où on soumet un panel à de « nouvelles questions ». Dans ce cas, la réanalyse présente un caractère expérimental et incertain : impossible de savoir au commencement de l'analyse secondaire si les entretiens comporteraient des éléments susceptibles d'alimenter les nouveaux angles d'analyse. Il s'agit alors d'aller chercher dans les entretiens d'éventuels "angles dormants" pour étoffer le matériau disponible.

Dans l'axe 1 par exemple, l'hypothèse d'une adhérence de certaines pratiques (achats, loisirs, services) aux réseaux-flux domicile-travail est testée à travers l'analyse des traces GPS et des entretiens, sans avoir l'assurance *a priori* que le matériau permette de répondre convenablement à cette question de recherche. De même, dans l'axe 2, la question du vécu du déplacement était présente en creux dans les interrogations initiales de l'une des recherches, en tant que composante du processus du changement modal, mais elle n'en constituait pas la trame centrale ; elle n'avait pas été abordée directement dans les entretiens des autres recherches.

Les données dont nous disposons dans le méta-projet nous semblent donc pour l'essentiel adéquates pour répondre aux problématiques identifiées ; pour les nouvelles hypothèses cette adéquation (ou la pertinence) du matériau est sans cesse posée.

c. Recontextualisations des données

Considérant que les données « n'existent » pas en soi mais sont des « construits » dépendants de différents contextes, il est nécessaire de les recontextualiser, de documenter leurs contextes (historique, scientifique) de production et de réanalyse.

Concernant les contextes historiques de production, nous considérons qu'ils sont très similaires pour l'ensemble des six programmes de recherche, ceux-ci ayant été menés sur des territoires géographiquement très proches (en Indre-et-Loire), et surtout à des époques très rapprochées (le plus ancien a débuté en 2009 et le plus récent s'est achevé en 2016). Nous avons néanmoins conscience de la très rapide évolution des pratiques en matière de mobilité, notamment en matière d'altéromobilités, et de leur contingence, au contexte économique notamment mais pas seulement ; on peut supposer que le niveau de connaissance et de pratique du covoiturage ou de l'autopartage qu'ont les enquêtés ont fortement évolué entre les premiers programmes et les plus récents. Aussi, nous ne pourrions pas comparer rigoureusement la diffusion de l'altéromodalité (et ses facteurs) à l'échelle interindividuelle. Par contre, il semble tout à fait envisageable de travailler sur la reconstitution des trajectoires de mobilité ou la compréhension des représentations et vécus de la mobilité d'individus interviewés dans différentes recherches, pour lesquels on peut maîtriser davantage les effets de contextes.

Les contextes scientifiques de production des données concernent notamment la personne qui a assuré la passation des entretiens : parmi les 250 entretiens dont nous disposons pour le méta-projet, certains

ont été réalisés par le chercheur primaire, qui a écrit la grille d'entretien et les a ensuite analysés ; d'autres par des étudiants ou des chercheurs contractuels. A la lecture des entretiens, cette différence d'expertise et d'avancement dans la carrière ne semble néanmoins pas avoir d'impact, sur la qualité ou la richesse des interviews.

Les contextes scientifiques de production sont également liés au contexte de recherche, i.e. au laboratoire auxquels appartiennent/appartenaient les chercheurs-producteurs, mais surtout à la posture scientifique que revendiquent les chercheurs, à leur sensibilités (parfois liées à leur appartenance disciplinaire) et intentions. La lecture des entretiens a mis en évidence une diversité dans ces postures et sensibilités, qui peut par exemple s'observer dans le choix des questions posées (la thématique des modes d'habiter étant abordée par un prisme ou un autre) ou des personnes interviewées (la question de la représentativité du panel constituant, ou non, un prérequis). Il nous faut donc composer avec cette diversité, qui peut parfois constituer un frein à l'analyse.

L'une des difficultés à la recontextualisation réside dans la mise à disposition, le porter à connaissance des éléments d'information. Différents « outils » de recontextualisations (mentionnés dans Belot & Van Ingelgom, 2017, notamment) peuvent être utilisés, telles que la lecture des grilles d'entretien, des rapports de recherche, des notes du chercheur primaire sur les entretiens ou d'autres documents de travail, et surtout d'échanges et discussions avec le chercheur primaire. Globalement, pus la donnée est « accompagnée » d'éléments d'explicitation du contexte d'enquête, y compris la façon dont l'enquêteur primaire l'analyse en premier intention, plus le travail de réanalyse s'avère facilité et pertinent. Or, lorsque le partage des données n'est pas anticipé ou considéré dès le départ comme un objectif, il est fréquent que les éléments de contexte (de production, d'analyse), qui permettent de comprendre et recontextualiser les données, ne soient pas détaillés ou transcrits, et restent connus du seul chercheur primaire.

A cet égard, le méta-projet a constitué un contexte particulièrement favorable : les chercheurs participant au projet étant également chercheurs primaires d'une ou plusieurs des six recherches, ils ont pu facilement apporter des éléments de recontextualisation aux chercheurs secondaires travaillant sur « leurs » données. Néanmoins, l'appartenance des chercheurs au même laboratoire, la récence des projets de recherches et les similarités dans les approches ont limité les besoins de recontextualisation : souvent, la simple lecture des entretiens a suffi pour pouvoir comprendre les discours. En outre, le temps de l'analyse (restreint, décalé suivant les axes, réalisé pendant les vacances, etc.) n'a pas toujours permis d'échanges préalables à l'analyse entre chercheurs primaire et secondaire.

Nous avons choisi de construire, au début du projet, un tableau de métadonnées qui comportait, pour l'ensemble des entretiens et traces GPS, le nom du programme de recherche ainsi que des éléments de caractérisation de l'enquêté (anonymisé) et de l'entretien². Ce tableau de métadonnées devait

² Indicateurs renseignés dans le tableau de métadonnées :

- Relatifs à l'enquêté : âge ; genre ; situation familiale ; nombre d'enfants à charge ; région de naissance ; niveau de scolarisation des enfants ; dernier diplôme obtenu ; situation professionnelle ; revenu mensuel du foyer ; patrimoine détenu ; PCS ; commune de résidence ; commune de travail ; distance à Tours ; distance au centre-bourg ; statut occupant ; type de logement ; cadre de vie ; taille du logement ; nombre de pièces ; espaces extérieurs et superficie ; ancienneté construction ; ancienneté emménagement ; mode de déplacement principal ; taux de motorisation ; abonnement aux transports en commun ; accès internet ; numéro d'identification (les enquêtés étaient anonymisés dans le tableau de métadonnées – un tableau de correspondance a également été construit)
- Relatifs à la recherche ou à l'entretien : nom du programme de recherche ; nom de l'enquêteur ; nom du retranscripteur ; lieu de l'entretien ; durée de l'entretien ; déroulé de l'entretien ; participation à d'autres moments de la recherche (expérimentation d'accompagnement personnalisé, entretien embarqué, recontact à long terme, enregistrement GPS) ; existence d'une grille d'observation

constituer un outil de recontextualisation ; divers objectifs étaient poursuivis au travers de sa construction :

- en termes de connaissance des entretiens : au commencement du projet, ce tableau devait permettre d'avoir une vue d'ensemble des données à notre disposition et de leur état (entretien retranscrit ou non, traces GPS à l'état brut ou nettoyée). Sa construction a notamment été l'occasion de collecter auprès des différents chercheurs, et stocker en un même endroit, la totalité des matériaux.
- en termes d'analyse : en prévision des analyses secondaires, ce tableau devait permettre de réaliser des sélections parmi les entretiens à notre disposition ; une autre ambition initiale était de pouvoir si besoin réaliser des statistiques à partir du « méta-corpus ».

En réalité, ce tableau de métadonnées a été peu utilisé par les chercheurs du méta-projet. D'une part, l'avancée de nos réflexions nous a convaincus de l'absence de pertinence de la notion de « méta-corpus » ; dès lors, quel aurait été le sens de réaliser des analyses statistiques sur ce panel ? quelle valeur auraient eu ces analyses ? En outre, seul un chercheur a eu recours au tableau pour sélectionner les entretiens qui l'intéressaient.

Finalement, l'intérêt principal de la construction de ce tableau a été d'être l'occasion de rassembler l'ensemble des matériaux ; cet exercice s'est révélé chronophage, mettant en évidence une difficulté au partage/à la réutilisation des données de la recherche : conserver, stocker, organiser les données.

3. Retour méthodologique sur des expériences d'analyse secondaire : une pluralité d'objectifs et de protocoles méthodologiques

a. Une pluralité d'objectifs poursuivis

La (maigre) littérature sur l'analyse secondaire de données qualitatives identifie une diversité dans les objectifs poursuivis lors d'analyses secondaires ou de réanalyses. Elle distingue cinq objectifs principaux (on retrouve certains d'entre eux dans les objectifs portés par l'ouverture et le partage des données de la recherche - cf. livrable 2).

1. Explorer un nouveau sujet ou préparer un nouveau projet de recherche ; pour Dargentas *et al* (2007), un tel travail préparatoire ne remplace cependant pas un travail de terrain et la production de données nouvelles. Concernant cet objectif, on peut néanmoins s'interroger sur l'utilité de « fouiller » aussi finement dans les données de la recherche : une méta-analyse (une recension ou une synthèse des résultats produits sur un sujet ou un terrain de recherche) ne suffirait-elle pas ?

2. Réutiliser les données à des fins de comparaisons, notamment diachronique (Belot & Van Ingelgom, 2017) ou géographique.

3. Renforcer la validité des analyses en travaillant sur des panels de données plus larges ou diversifiés (Heaton, 2008).

Pour les deuxième et troisième objectifs, les analyses secondaires se basent sur les données primaires. Pour les objectifs ci-après, l'analyse secondaire porte également sur les *résultats* de l'analyse primaire, voire plus largement sur les hypothèses de travail et les postures théoriques et épistémologiques initiales. Dans ces situations, la présence ou les échanges avec le chercheur primaire semblent d'autant plus indispensables.

4. Aller « plus loin » dans l'analyse primaire, produire une analyse « *plus complexe* » (Hamidi, 2017) : l'analyse secondaire doit alors être considérée comme une suite à l'analyse primaire, comme un moyen « *d'affiner les grilles d'interprétation issues de précédentes analyses* » (Belot & Van Ingelgom, 2017), « *d'approfondir la question de recherche initiale* » (Hamidi, 2017). Divers éléments permettent cet approfondissement de l'analyse : le temps supplémentaire accordé à l'analyse des données (les contraintes temporelles inhérentes à la recherche ne permettant pas toujours de consacrer

à l'analyse autant de temps que souhaité) ; les compétences et la maturité supplémentaires que le chercheur primaire aura acquises entre les deux analyses ; les compétences et le nouveau regard qu'apporte le chercheur secondaire ; possiblement, les technologies d'analyse qui auront été développées entre les deux analyses (Duchesne, 2017) (développement de logiciels d'analyse, etc.).

5. « Décaler le regard » : l'analyse secondaire est également le moyen de réinterroger la problématique ou les hypothèses de recherche de l'analyse primaire, d'« *apprendre à repérer ce qu'on avait appris à ne pas voir* » (Dupuy & Duchesne 2017). Ici encore, l'analyse secondaire doit être considérée dans son rapport à l'analyse primaire, non pas dans sa continuité (comme précédemment) mais dans le pas de côté qu'elle permet de réaliser. Là, le croisement des différents regards disciplinaires des chercheurs primaire et secondaire (de leurs cultures disciplinaires, de leurs backgrounds théoriques et épistémologiques) permet de réinterroger des présupposés qui auraient pu sembler aller de soi pour le chercheur primaire ou au moment de l'analyse primaire (en raison du contexte historique).

En participant au méta-projet, les chercheurs de l'équipe se sont tous engagés dans un travail « d'analyse secondaire » : ce travail recouvre en réalité (selon les axes investigués et les aspirations personnelles des chercheurs) différents objectifs, les principaux étant des objectifs de capitalisation, validation de résultats existants, et d'approfondissement de nouvelles questions. D'autres objectifs, propres à chaque axe, ont été poursuivis, comme le test ou la validation de protocoles méthodologiques.

En conséquence, la réanalyse s'est ensuite déclinée de manière variable selon les axes de recherche.

b. De l'utilisation de logiciels d'analyse textuelle

Une question s'est posée rapidement : comment analyser de manière intégrale 250 entretiens semi-directifs ? Comment analyser un panel « quantitatif » de données qualitatives ? Surtout, comment réaliser cette analyse dans le temps contraint qui nous est alloué ?

Les outils d'assistance automatisée à l'analyse, logiciels d'analyse textuelle (en particulier les CAQDAS³) nous sont au départ apparus comme pertinents pour mener l'analyse secondaire : la littérature mentionne en effet que « *la pratique de l'analyse secondaire est conditionnée par l'usage d'outils d'assistance au traitement des données, dans la mesure où elle impose au chercheur, dans la plupart des cas, d'analyser des corpus numériquement beaucoup plus importants que dans le cas d'une analyse primaire et qui ne pourraient donc que difficilement faire l'objet d'une analyse « manuelle* » » (Brugidou & Le Roux, 2005). Ces outils d'assistance constitueraient donc une « *opportunité pratique* » en « *facilitant l'exploration du matériau empirique* », mais ils permettraient également « *une objectivation de la démarche qualitative* » (Brugidou & Le Roux, 2005) en « *explicitant les procédures, permettant l'analyse à plusieurs ou un archivage de l'activité scientifique* » (Lejeune, 2010).

La littérature (en particulier le Hors-Série de Recherche Qualitatives) relaie l'utilisation de CAQDAS pour certaines expériences d'analyse secondaire. Ces CAQDAS (différents logiciels selon les recherches) sont utilisés pour coder les entretiens et repérer des premières catégories d'analyse ; ils sont reconnus comme obligeant les chercheurs à « *documenter et systématiser ses façons de faire* » (Hamidi, 2017), ce qui facilite notamment le travail à plusieurs et l'appropriation par des chercheurs secondaires de données primaires.

³ CAQDAS = famille d'outils développés dans les milieux de la recherche qualitative anglo-saxonne, là même que se sont développées les premières recherches d'analyse secondaire. Développés par et pour des chercheurs en SHS dans les années 1980. Les CAQDAS donnent à voir les manipulations concrètes des data, les opérations de surlignement, d'annotation et de codage, de rapprochement de d'opposition par lesquelles le chercheur prétend passer des percepts ou concepts »

Concernant le méta-projet, un seul chercheur s'est finalement orienté vers l'utilisation des CAQDAS. La réticence des autres chercheurs à avoir recours à ces outils peut être liée à leur méconnaissance des logiciels (Lejeune, 2010) ou à leur utilisation reconnue comme chronophage.

Le chercheur a eu recours au logiciel Tropes : parmi les différents logiciels d'analyse de texte, Tropes s'apparente à un dictionnaire informatique permettant de représenter des éléments d'interprétation (Lejeune, 2017) à travers la recherche de termes spécifiques et paramétrés dans le corpus analysé. Ce logiciel procède d'une démarche hypothético-déductive en classant les termes issus du corpus en fonction de leur présence dans le dictionnaire et de leur fréquence d'apparition. Le chercheur a commencé par appliquer à ses entretiens le vocabulaire en lien avec les termes « sentiment », « ressentiment », « sensation » ... (et aux déclinaisons sémantiques en lien avec la thématique de la perception et du vécu) contenu dans le dictionnaire pré-paramétré du logiciel. Cette première « fouille » des entretiens devait lui permettre d'appréhender les textes sans avoir besoin de les lire en intégralité.

Cette fouille a permis d'identifier des entretiens et parties d'entretiens potentiellement intéressants au regard de ses problématiques.

Néanmoins, des limites importantes sont apparues, non liées à la démarche de réanalyse mais bien à l'analyse systématique par logiciel CAQDAS de matériaux qualitatifs de type entretien semi-directif. D'une part, dans certains entretiens, il arrive que l'enquêteur « finisse » certaines phrases de l'enquêté ou traduise des discours « sous-entendus » : appliquée au discours du seul enquêté, la fouille automatisée invisibilise ces éléments de discours de l'enquêteur, pourtant riches en éléments d'analyse.

D'autre part, une fois les éléments du discours de l'enquêté identifiés grâce à la fouille automatisée, il était ensuite nécessaire de les remettre en contexte, de leur donner sens en relisant le passage en intégralité, puis des passages de plus en plus importants... puis l'intégralité de l'entretien. Rapidement, la capacité du chercheur à appréhender et mémoriser les entretiens dans leur globalité a été dépassée. Le temps « gagné » grâce à la fouille automatique était donc finalement peu utile, puisqu'il fallait de toute manière lire la totalité de l'entretien.

Plus important, la fouille automatique des entretiens n'a permis d'identifier que peu de vocabulaire en lien avec la perception et le vécu, ce qui est lié d'une part aux programmes de recherche et à la construction des grilles d'entretien, d'autre part (Lejeune, 2017) aux dictionnaires pré-paramétrés dans le logiciel, peu pertinents pour identifier l'expression des sentiments ou des ressentis personnels.

Pourtant, la lecture intégrale de certains entretiens permettait d'identifier des éléments « entre les lignes », intéressants pour l'analyse, des éléments implicites, non explicités verbalement mais sous-entendus dans les propos de l'enquêté. Une lecture attentive révélait donc des entretiens plus riches que la fouille ne le laissait supposer.

Globalement, l'utilisation d'un logiciel automatisé et la fouille automatisée se sont révélés peu adaptés et peu satisfaisants pour appréhender le matériau qualitatif à notre disposition par le prisme de la thématique investiguée.

Le chercheur a donc fait évoluer sa méthodologie, et a opéré l'analyse sur une sélection d'entretiens plutôt que sur la globalité des entretiens.

L'utilisation de logiciels d'analyse textuelle aurait-elle néanmoins pu être pertinente ?

Peut-être, à condition que les entretiens aient été « préparés » pour être analysés par ce biais, i.e. retranscrits (en distinguant les locuteurs) et catégorisés (par thématique) ; à condition également que les chercheurs aient été suffisamment familiers de ce type de logiciel pour qu'au temps de l'analyse ne s'ajoute pas un temps de familiarisation avec l'outil ; à condition, enfin, que les entretiens analysés soient suffisamment riches et explicites sur la thématique investiguée, pour que la « fouille » soit

efficace. Alors, le recours à un logiciel d'analyse textuelle aurait effectivement pu permettre d'engager l'analyse secondaire.

Aurait-on pu travailler sur la totalité des entretiens, réaliser une analyse secondaire exhaustive à partir d'une ou l'autre thématique ?

Nos expériences d'analyse secondaire semblent justifier les doutes que nous ressentions au départ du méta-projet sur notre capacité à analyser par une approche qualitative une telle masse de données. Avec ou sans recours à un logiciel d'analyse textuelle, une approche « traditionnelle » et un passage par la lecture des entretiens semblent nécessaires pour comprendre et pouvoir produire une analyse des données. Or, appréhender (mémoriser) et être capable de comparer un échantillon de 250 entretiens semble totalement hors de portée de nos cerveaux, et incompatible avec le temps à notre disposition.

Notre expérience ne remet pas du tout en question l'intérêt d'une analyse secondaire, qui peut se révéler (on le verra ensuite) parfaitement intéressante. En revanche, elle interroge l'intérêt et la possibilité de travailler, par une approche qualitative, sur des panels d'entretiens de cette taille.

Par ailleurs, il semble tout à fait possible de mener une analyse secondaire sur un panel quantitativement important de données : c'est notamment ce qui a été fait dans l'axe 1, à partir de traces GPS.

c. De la délimitation de sous-corpus d'entretiens

Un point commun entre les méthodologies mises en œuvre dans les trois axes de recherche concerne la délimitation de sous-corpus d'entretiens au sein du panel. Différentes raisons ont justifié ces sélections : d'une part, l'impossibilité d'adopter une approche qualitative sur la totalité des 250 entretiens ; d'autre part, l'envie de travailler sur certains profils d'individus : individus résidant en périurbain (axe 1), individus ayant « fait évoluer leurs pratiques de mobilité » (axe 3) ; enfin, l'impossibilité temporelle d'analyser la totalité des matériaux à notre disposition.

Comment opérer cette sélection ? Quels entretiens sélectionner ?

Ces contraintes temporelles, associées à une incertitude quant au contenu des entretiens (cf. [5. Difficultés rencontrées](#)), a conduit les chercheurs des différents axes à vouloir identifier parmi le panel de 250 entretiens les entretiens « les plus riches » sur leur thématique, les entretiens comprenant les données les plus pertinentes pour répondre aux problématiques. Mais comment les identifier ?

D'une part, en répondant à des « commandes », passées par les chercheurs secondaires aux chercheurs primaires. Ainsi, le chercheur secondaire soumettait aux chercheurs primaires sa problématique de recherche, afin qu'il identifie les (parties d') entretiens les plus intéressant(e)s ou riches à analyser : de fait, le chercheur primaire « connaît » ses entretiens, pour les avoir menés et/ analysés lors d'une recherche primaire. En revanche, sa connaissance est nécessairement imparfaite et partielle : d'une part, il a pu oublier certains entretiens ou parties d'entretiens, moins marquants ou menés il y a trop longtemps ; d'autre part, il connaît surtout les parties d'entretiens, éléments d'analyse et angles d'approche qui l'avaient intéressé au moment de l'analyse primaire. En réponse à la problématique proposée par le chercheur secondaire, le chercheur primaire va sélectionner certains entretiens et en éliminer d'autres, moins ou pas intéressants : un biais est alors induit.

Notre expérience d'analyse secondaire a en effet montré que cette méthode de sélection d'entretiens « intéressants » peut se révéler problématique : pour l'axe 2 par exemple, une sélection d'entretiens et passages ont été soumis au chercheur secondaire, censés être les plus intéressants ; d'autres ont été écartés de la sélection, le chercheur primaire considérant « qu'il ne devait rien y avoir d'intéressant pour l'axe 2 dans ses entretiens ». Cette sélection a orienté le chercheur secondaire, l'amenant à lire certains entretiens ou ne pas s'y pencher du tout.

Or, le chercheur secondaire a identifié des éléments très intéressants dans des entretiens non sélectionnés par le chercheur primaire, mais qu'il avait quand même choisi de lire. Ceci met en évidence un élément plusieurs fois vérifié au cours de nos expériences d'analyse secondaire : la différence dans les lectures et l'analyse que font les chercheurs primaire et secondaire d'un même entretien. Le chercheur secondaire, qui maîtrise sa question de recherche et « sait » ce qu'il cherche dans les entretiens, a su identifier des éléments qui n'avaient pas paru intéressants au chercheur primaire ; son ignorance de l'entretien lui a fait porter un regard nouveau sur les éléments.

Ainsi, la pré-sélection par les chercheurs primaires des « entretiens les plus intéressants », si elle peut permettre de gagner du temps, induit certainement aussi une perte d'éléments intéressants. Elle peut notamment sembler incohérente avec un objectif que nous poursuivions initialement, i.e. la valorisation de « parties dormantes » des entretiens. Se fier uniquement au chercheur primaire pour sélectionner les matériaux à analyser pourrait ainsi se révéler néfaste pour la qualité de l'analyse.

Alors, comment faire ? Hormis les entretiens présélectionnés par les chercheurs primaires, les chercheurs secondaires ont choisi de lire des entretiens au hasard... lectures qui se sont donc révélées plus ou moins riches.

d. Des échanges entre chercheurs primaire et secondaire

Le bénéfice pour la qualité de l'analyse secondaire de liens entre chercheurs primaire et secondaire est largement reconnu, que ces échanges se fassent sous la forme d'entretiens (Laferté, 2016), de « *collaboration étendue* » (Dargentas et al, 2007) ou de « *dispositifs collaboratifs* » (Irwin, Bornat & Winterton, 2014). Dans le cadre de nos réflexions sur le partage des données de la recherche (Huyghe et al, 2018), nous avons d'ailleurs préconisé que le chercheur primaire soit impliqué au maximum tout au long du partage des données, et que des échanges soient prévus entre chercheur primaire et chercheurs secondaires.

Dans le cadre du méta-projet, ces échanges entre chercheurs primaire et secondaire ont été plus ou moins développés suivant les étapes de l'analyse secondaire.

1/ Lancement du projet : on l'a dit précédemment, l'interconnaissance et la confiance entre les sept chercheurs participant au projet (qui sont tous à la fois chercheurs primaire et secondaire) ont largement participé au lancement du projet ; sans qu'elles aient constitué une condition à la mutualisation de « nos » données, elles ont permis de réduire les réticences que nous pouvions ressentir, au début du projet, à partager « nos » entretiens et traces GPS.

2/ Recontextualisation : on l'a également mentionné précédemment, la possibilité d'échanges entre les chercheurs secondaire et primaire participait à ce que le contexte du méta-projet soit favorable à une analyse secondaire : elle devait faciliter la nécessaire re-contextualisation des recherches primaires (entretiens, métadonnées, hypothèses, méthodologies, résultats).

Globalement, nous avons peu échangé à ce sujet, peu discuté des contextes de production des données primaires ; deux raisons principales peuvent être avancées : d'une part, un manque de temps ou d'opportunités (peu de réunions collectives par exemple) ; d'autre part, un manque de besoin (en raison du contexte très spécifique du méta-projet se basant sur une mutualisation de projets récents). Exception notable néanmoins : l'axe 3, mené par un binôme de chercheurs, a débuté puis donné lieu à de nombreuses discussions *au cours* du travail, en particulier sur les cadres théoriques et analytiques utilisés dans les programmes de recherche antérieurs.

3/ Méthodologie d'analyse : en revanche, nous avons eu de nombreux échanges autour de la pré-sélection d'entretiens « riches » à analyser (voir [De la délimitation de sous-corpus d'entretiens](#)), ainsi que sur la/les méthodes à mettre en œuvre pour l'analyse secondaire en fonction des matériaux à analyser et des objectifs poursuivis par axe.

4/ Résultats des analyses : nous avons également échangé autour des résultats des analyses (en cours ou finalisées), au cours de l'analyse, à partir de nos écrits ou au cours de réunions ou discussions informelles. Ces échanges nous ont permis, comme dans toute recherche collective, d'avancer dans nos réflexions, de décaler nos regards, d'enrichir nos analyses. Dans le cadre spécifique d'une analyse secondaire, les échanges avec les chercheurs primaires ont également permis de valider les résultats produits, qu'ils soient en accord ou non avec les résultats primaires.

Ces échanges aux différentes étapes de l'analyse ont été principalement permis par la proximité géographique des chercheurs ; pour le seul chercheur ne travaillant plus au laboratoire CITERES, les échanges « en cours d'analyse » ont d'ailleurs été moins fréquents, plus compliqués à avoir. Plus largement, ce constat nous interroge donc sur la pertinence/l'opportunité de mener des analyses secondaires dans le cas où les échanges avec les chercheurs primaires seraient difficiles ou impossibles : lorsque les données sont anciennes, lorsque le chercheur est géographiquement éloigné, a quitté le monde académique, ne travaille plus sur les mêmes thématiques, etc. En particulier, comment recontextualiser les données ?

4. Plus-value de l'analyse secondaire ?

Au terme de nos expériences d'analyses secondaires, quel regard portons-nous sur l'exercice ? Est-ce finalement un exercice utile, pertinent ?

Nous identifions deux types de plus-value à l'analyse secondaire :

a. Des plus-value en matière d'analyse scientifique

L'analyse secondaire est notamment préconisée en ce qu'elle permettrait d'optimiser des données de la recherche existantes, d'exploiter des matériaux non exploités. Les expériences d'analyse secondaire menées dans le cadre du méta-projet nous ont en effet permis de proposer des résultats scientifiques parfaitement intéressants dans le champ des mobilités spatiales : de valider des analyses antérieures (obtenues au cours des programmes de recherche primaires, ou au cours d'autres recherches) mais également de proposer de nouveaux résultats.

Par ces diverses analyses secondaires, les données de recherche existantes ont donc bien été « optimisées », ont permis de produire plus de résultats.

Quid de la valorisation de données « dormantes », non encore exploitées ? Si c'était un objectif initial du méta-projet, les méthodologies que nous avons mises en œuvre pour analyser les corpus (en particulier, la sélection des entretiens « les plus riches ») n'ont certainement pas réellement permis de l'atteindre. Ces données dormantes, peu ou non connues des chercheurs primaires car non exploitées lors de l'analyse primaire, n'ont sans doute pas (toutes) été identifiées comme matériau de recherche intéressant pour les nouvelles problématiques... et resteront à l'état de données « dormantes ».

Aussi, comme mentionné précédemment, d'autres méthodologies d'analyse secondaire permettraient certainement d'aller encore plus loin dans la valorisation des matériaux. De même que, interrogés par le prisme d'autres problématiques, ces matériaux pourraient certainement avoir encore beaucoup à dire.

Nous avons également identifié une autre plus-value scientifique de l'analyse secondaire, non mentionnée dans la littérature : l'analyse secondaire constitue une aubaine, puisqu'elle met à disposition des chercheurs des données « prêtes à exploiter », qui incitent à creuser ou expérimenter de nouvelles thématiques ou méthodes. Une aubaine puisque, si ces données n'existaient pas, les chercheurs ne les auraient sans doute jamais collectées ou produites.

Ainsi, l'existence de traces GPS facilement exploitables a permis à un chercheur de l'axe 1 qui travaillait jusque-là exclusivement sur des entretiens semi-directifs, de travailler à partir de ce nouveau type de matériau ; aubaine pour le chercheur (cf. [Des plus-values pour les chercheurs participant à l'analyse secondaire](#)), c'est également une aubaine pour la qualité et la richesse des analyses : le chercheur a ainsi pu tester ses hypothèses, valoriser son expertise et son approche de qualitatif sur un matériau « quanti », et ainsi produire des analyses plus complètes et plus riches.

De même, l'existence de traces GPS « à exploiter » a permis à d'autres chercheurs de l'axe 1 de tester de nouvelles méthodes d'analyse : le temps « économisé » à ne pas produire les matériaux de recherche a pu être investi dans des recherches méthodologiques⁴.

b. Des plus-values pour les chercheurs participant à l'analyse secondaire

L'analyse secondaire constitue une aubaine pour les chercheurs qui y participent, en leur permettant de gagner en expertise sur leur thématique en travaillant sur de nouvelles données/des données d'un nouveau type, qu'ils n'auraient pas produites eux-mêmes. Dans l'exemple mentionné ci-dessus, l'existence des données GPS a permis à un chercheur de se familiariser avec ce matériau et de monter en expertise, alors qu'il est probable qu'il n'aurait pas cherché à les produire lui-même.

En outre, participer à une expérience d'analyse secondaire (en encore plus à une expérience de mutualisation) favorise, voire impose, des échanges et des réflexions autour des ancrages théoriques, postures, méthodes mises en œuvres par les différents chercheurs.

Deux étapes invitent particulièrement à cette « mise à nu » : d'une part, la recontextualisation, les échanges autour des contextes de production des données, qui demandent de rentrer dans le détail du mode de faire de chaque chercheur. D'autre part, la lecture des entretiens semi-directifs : on en apprend énormément sur les méthodes respectives des chercheurs-enquêteur, sur leurs manières de mener leur entretien, de « donner » plus ou moins d'eux aux enquêtés. De même, la lecture des rapports de recherche, dont les parties méthodologiques sont généralement relativement détaillées, sont particulièrement formatrices. Ces échanges ou lectures peuvent amener les chercheurs secondaires à faire évoluer leurs postures, leurs systèmes interprétatifs.

Outre par ces expériences d'analyse secondaire ou des recherches collectives qui imposent de mettre en commun données, méthodes et résultats, un chercheur a peu d'occasions d'avoir de tels échanges sur ses manières de faire et de se positionner ; en particulier, les articles scientifiques (qui constituent nos ressources de base quand on se lance dans un nouveau projet de recherche) sont souvent très pauvres en matière de méthodologie ou de posture scientifique... et invitent peu à une remise en question de nos manières de faire.

Les échanges et réflexions sur les méthodes et résultats sont donc une opportunité pour les chercheurs de progresser dans leurs réflexions et leur manière de faire. Pour les chercheurs secondaires, mais également pour les **chercheurs primaires** : les échanges menés avec les chercheurs secondaires au cours du processus d'analyse, en particulier sur les méthodes développées et les résultats proposés, permettent au chercheur primaire d'avoir un « retour » critique, un regard extérieur, sur sa recherche.

5. Difficultés rencontrées

Le processus de mise en commun et de réanalyse des entretiens pose un certain nombre de problèmes, en particulier son caractère chronophage. En effet, la principale difficulté à laquelle nous avons été

⁴ Cette question du temps « économisé » n'est peut-être valable que pour les traces GPS mais pas pour les entretiens semi-directifs, qui nécessitent des temps de recontextualisation ou d'appropriation élevés qui « compensent » ce temps gagné.

confrontés, déjà bien documentée dans la littérature (Duchesne, 2017) est celle du temps nécessaire à l'analyse secondaire. Si une des motivations du méta-projet était d'optimiser des données de la recherche, d'exploiter et de valoriser des matériaux relativement peu explorés, il convient de garder en tête que l'analyse secondaire prend quasiment autant de temps (donc d'argent) qu'une analyse primaire avec production de ses propres données.

Ce caractère chronophage se heurte au temps, contraint, de l'analyse (et, plus globalement, de la recherche). Le temps limité dont nous disposons pour la réanalyse nous a poussés à faire des choix dans les problématiques et les matériaux investigués (sélection de sous-corpus), et ont nécessairement eu un impact sur la richesse, la profondeur et la qualité de la réanalyse.

Pourquoi cette analyse secondaire a-t-elle été si chronophage ? Il faut considérer que le traitement de matériaux constitués par d'autres comporte un "coût temporel" lié à la réappropriation des données.

D'une part en raison d'une nécessaire remise en forme des données en amont de la réanalyse, données qui sont rarement exploitables immédiatement : ainsi, certains entretiens n'avaient pas été retranscrits. De même, la ré-exploitation des traces GPS a nécessité tout un travail préalable (par exemple d'identification des points d'arrêt), qui n'avait pas forcément été anticipé au départ.

D'autre part, mener une analyse secondaire suppose un temps nécessaire pour se familiariser ou se re-familiariser avec les données (que le chercheur les ait, ou non, produites) : un temps pour recontextualiser les entretiens (lecture des grilles d'entretien, des rapports produits, etc.), un temps pour (re)lire les entretiens de manière minutieuse.

En outre, lorsqu'il n'est pas le producteur primaire des données, le chercheur peut être amené à travailler sur des données avec lesquelles il n'est pas familier, comme cela a été le cas pour l'un d'entre nous sur les traces GPS. Cette méconnaissance initiale entraîne des interrogations du chercheur sur le caractère exploitable ou non des données et la manière adéquate de les analyser, ce qui encore une fois prend du temps.

Enfin, l'utilisation de logiciels d'assistance automatisée à l'analyse (mise en œuvre pour un axe) s'est également révélée chronophage, en raison notamment du matériau à analyser, qui n'avait pas été préparé pour être analysé de cette manière (i.e. retranscrit puis catégorisé *via* ce logiciel).

Une autre difficulté est liée au caractère incertain de l'analyse secondaire. On ne sait pas vraiment, au début de la réanalyse, "ce qu'on va trouver dans les entretiens" et si l'analyse va se révéler véritablement opérante. Les données vont-elles permettre de répondre aux nouvelles problématiques ? Les résultats produits vont-ils apporter une réelle plus-value vis-à-vis des analyses primaires ?

La démarche de réanalyse n'est ainsi pas exempte de déceptions. Dans le cadre de l'axe 3 (focalisé sur le processus de changement) nous souhaitions par exemple approfondir la compréhension du processus de changement *via* l'analyse du déroulement du changement, à partir d'entretiens de Modalter qui n'avaient initialement pas été orientés par cette problématique. La réanalyse de ces entretiens s'est révélée globalement décevante et n'a pas permis d'apporter de nouveaux éléments à cette question.

Le caractère incertain de la réanalyse est également lié au matériau investigué. Ainsi, si nous savions au début de l'analyse que nous disposions d'environ 250 entretiens et d'une centaine de traces GPS (grâce à une revue des parties « méthodologique » des rapports finaux de chaque programme), nous ne connaissions pas l'état de ce matériau et ne savions pas s'il était réellement exploitable. Ainsi, nous nous sommes aperçus au cours de la réanalyse qu'une partie des entretiens n'était pas retranscrite et que nous disposions uniquement d'un résumé de l'entretien. De même, nous ne savions pas réellement

dans quel état étaient les traces GPS : enregistrements bruts, nettoyés, segmentés ? (Depeau et alii, 2018) ; ni le temps de préparation qui pourrait être nécessaire avant de pouvoir les exploiter.

Une dernière difficulté, dans le cadre d'une réanalyse, consiste à lire « autrement » ses propres données : des données qu'on a produites, déjà analysées (selon un certain angle d'analyse), et qu'il s'agit au cours de la réanalyse de lire sous un autre angle afin d'en tirer de nouveaux résultats. Les souvenirs qu'on a des entretiens/parties d'entretiens peuvent « masquer » les éléments (voire, les entretiens complets) qu'on n'avait pas retenus ou jugés intéressants lors de l'analyse primaire ; peuvent nous faire considérer que « tel entretien » n'a rien à apporter sur telle ou telle question.

Cette difficulté s'est observée au cours du méta-projet : ainsi, certains entretiens de Modalter, qui avaient été écartés par le chercheur primaire car jugés inintéressants, ont finalement retenu l'attention du chercheur secondaire en charge de l'axe.

Ainsi, si la présence des chercheurs primaires facilite l'appropriation des données et la prise de connaissance des analyses primaires, n'amène-t-elle pas à minimiser ou masquer certains aspects utiles à l'analyse secondaire ?

Retour critique sur notre choix de ne pas anonymiser les entretiens

Nous avons fait le choix, justifié dans Huyghe et al (2018) de travailler sur des entretiens non anonymisés : un choix qui se justifie principalement par le temps qu'aurait demandé l'anonymisation du matériau. En revanche, nous avons produit un tableau de métadonnées (cf. [supra](#)) anonymisé, pour l'ensemble des entretiens et traces à notre disposition, accompagné d'un tableau de réidentification permettant de faire le lien entre le tableau et les entretiens.

Cette démarche « entre anonymisation et non-anonymisation » (peu cohérente en réalité) ne nous a pas convaincus.

La lecture d'entretiens non-anonymisés nous a permis de réidentifier l'enquêteur et, parfois, l'enquêté (connu personnellement, ou de nom) : nous avons ainsi vérifié le risque, mentionné dans Huyghe et al, (2018 : 15), que faisait courir la réalisation d'enquêtes sur les « *terrains de proximité, où la distance entre les enquêtés et les enquêteurs est faible* ». Reconnaître les enquêtés n'est pas « grave » ou « dangereux » puisque les données obtenues à la lecture des entretiens ne seront pas utilisées à mauvais escient, par aucun d'entre nous. Néanmoins, la lecture de retranscriptions pour lesquelles l'enquêteur primaire s'engage, au début de l'entretien, auprès de l'enquêté, à ce que ses propos soient anonymisés et confidentiels, pose de véritables questions éthique et déontologique. Aussi, une anonymisation aurait été souhaitable, même très légère (suppression du nom de famille ou du prénom de l'enquêté), qui aurait grandement limité la possibilité de reconnaître les enquêtés.

Par ailleurs, concernant la réidentification de l'enquêteur. D'une part, il peut être utile de connaître l'enquêteur pour pouvoir l'interroger et obtenir des éléments de contextualisation des entretiens. D'un autre, la lecture d'un entretien permet d'en apprendre parfois beaucoup sur la vie privée de l'enquêteur : il est en effet fréquent qu'au cours de l'entretien on en « donne » un peu à l'enquêté, pour créer notamment une ambiance de discussion plutôt que d'entretien trop formel. Ces lectures mettent en évidence différentes façons de mener des entretiens et une diversité dans les investissements des chercheurs dans la construction de la réalité sociale. Le décalage entre les choix et positions épistémologiques de chaque enquêteur peut alors provoquer un sentiment de gêne, qui apparaît à la lecture des entretiens.

Conclusion : Analyse secondaire - clés de réussite

L'incitation au partage des données de la recherche (relayée ces dernières années par les principaux organismes de recherche publique - CNRS, INRA, CIRAD) devrait induire un développement de la

pratique de l'analyse secondaire parmi les chercheurs. Pour le moment, on identifie relativement peu d'expériences de ce type dans la littérature scientifique ; *a fortiori*, on ne trouve que peu d'articles consacrés aux aspects méthodologiques de l'analyse secondaire : quelles données peut-on mutualiser ? Comment recontextualiser les données ? Comment analyser par une approche qualitative un panel quantitativement important de données ? Quid des CAQDAS, de leur utilité/caractère nécessaire dans une analyse secondaire ?

C'est à ces différentes questions que cette synthèse entendait apporter des réponses, sur la base de notre/nos expérience(s) d'analyse secondaire. De fait, les sept membres de l'équipe du méta-projet ont vécu des expériences d'analyse secondaire relativement différentes, selon les problématiques qu'ils ont investiguées, selon les aspirations individuelles qu'ils poursuivaient, selon les méthodologies qu'ils ont mises en œuvre. Des expériences différentes, qui laissent néanmoins apparaître un certain nombre de similarités et nous permettent de proposer ci-après quelques clés de réussite.

1. De l'importance des liens entre chercheurs primaire et secondaire

Déjà mis en exergue dans la littérature, c'est en effet un point qui nous semble essentiel : les échanges entre chercheurs primaire et secondaire permettent d'échanger sur le fond (sur les résultats de l'analyse primaire, sur les données et leurs contextes de production ; sur les résultats de l'analyse secondaire... qui peuvent entraîner des controverses scientifiques qui enrichissent singulièrement les analyses) ; ils permettent au chercheur secondaire de se positionner, d'adopter un regard qui prolonge ou au contraire se décale par rapport au regard primaire ; ils permettent au chercheur primaire d'avoir un retour sur ses analyses primaires, qui peut lui permettre de progresser dans son analyse. Bref, ces échanges entre chercheurs primaire et secondaire sont bénéfiques tant à l'analyse en elle-même qu'aux chercheurs participant au processus.

L'absence de telles relations n'empêcherait pas l'analyse secondaire, des outils de recontextualisation existant (notes du chercheur primaire, documents de travail, tableaux de présentation des données, rapports et synthèses des résultats) qui permettent d'en apprendre suffisamment sur les données pour pouvoir les analyser « correctement ». Néanmoins, l'absence de relations nuirait certainement à la richesse et la profondeur de l'analyse secondaire. Ce constat nous interroge donc sur l'intérêt de réaliser une analyse secondaire dans le cas où le chercheur primaire serait injoignable.

2. De l'importance du temps

Mentionnée plusieurs fois dans cette synthèse, la question du temps est au cœur de cette expérience d'analyse secondaire.

On pense d'une part au temps consacré à la réanalyse, considéré comme toujours trop court, qui limite la richesse et la profondeur des résultats avancés ; qui impose parfois de faire des choix sur la quantité de matériau réanalysée. Aussi, lors du montage de projets reposant sur la réanalyse de matériaux de recherche, il est essentiel de prévoir au moins autant de temps d'analyse que pour des projets demandant une analyse primaire, et d'anticiper notamment le temps dévolu aux échanges entre chercheurs.

On pense d'autre part au temps entre les analyses primaire et secondaire qui, dans le cas où un chercheur réanalyse ses propres données, lui permet de gagner de l'expérience, d'acquérir de nouvelles connaissances et compétences... et de décaler son regard. Ce temps entre les deux analyses peut ainsi donner une nouvelle valeur aux données, faire apparaître des éléments non « vus » et non analysés initialement.

3. De l'importance de (se) former à l'analyse secondaire

Pour que l'analyse secondaire se développe en tant que pratique au même type que l'analyse primaire, une étape essentielle est celle de la formation des chercheurs, en particulier des doctorants.

Formation aux enjeux, avantages et limites d'une telle pratique : si cette pratique est encouragée (notamment dans un objectif d'optimisation de données de la recherche collectées sur fonds publics) et possède un vrai intérêt scientifique, elle a nécessairement des limites (en termes de chronophage, de richesse des données et des résultats produits, etc.), à connaître.

Pour que la pratique de l'analyse secondaire se développe, il faut également que les données de la recherche soient d'une part plus partagées (ce qui pourrait passer par la multiplication des structures de stockage et leur publicisation), soient d'autre part mieux préparées à la réanalyse. Cela passe peut-être par une certaine « normalisation » des données : par une retranscription, codage, thématisation, mises en forme dans des logiciels d'analyse automatisée. Cette normalisation faciliterait et accélérerait l'appropriation des données par le chercheur secondaire et le travail de réanalyse, voire rendrait possibles certains types d'analyse secondaire (l'analyse statistique de données textuelles par exemple). Un tel travail autour de la mise en partage des données de la recherche est mené depuis quelques années dans le monde de la recherche en archéologie, notamment au sein du [consortium MASA](#) (Mémoires des Archéologues et des Sites Archéologiques) qui travaille au développement de Plans de Gestion de Données et à la normalisation des données mises en partage. Des retours d'expérience présentés lors d'une [journée d'étude](#) sur le « *Partage des données qualitatives et ses enjeux* » laissent penser que ce travail participe réellement à la diffusion de la pratique et d'une culture d'analyse secondaire et aux échanges entre chercheurs.

6. Bibliographie

- Beaucher V. & Jutras F., 2007. « Etude comparative de la méta-synthèse et de la méta-analyse qualitative », *Recherches qualitatives* Vol. 27 (2) : 58-77
- Belot C. & Van Ingelgom V., 2017. « La réanalyse à l'épreuve d'une comparaison qualitative diachronique : stabilités et évolutions dans les perceptions citoyennes de l'Europe en 1995-1996 et 2005-2006 », *Recherches qualitatives* Hors-Série (21) : 99-121
- Bishop L., 2013. "Reply to: Natasha Mauthner and Odette Parry's "Open acces digital data sharing: principles, policies and practices"", *Social epistemology review and reply collective* 2 (8): 71-78
- Brugidou M. & Le Roux D., 2005. *L'usage des CAQDAS et des logiciels d'analyse textuelle dans l'analyse secondaire*, Colloque « L'analyse secondaire en recherche qualitative », Grenoble 2005
- Dargentas M., Le Roux D., Salomon A.C. & Brugidou M., 2007. « Sur les prospectives de la recherche qualitative en France : capitalisation et réutilisation d'entretiens de recherche », *Recherches qualitatives* Hors Série (3)
- Depeau & Collectif Mobikids, 2018. La trace : une notion partagée dans une recherche interdisciplinaire sur la mobilité et les rapports à l'espace des enfants, Communication MSFS 2018
- Duchesne S., 2017. « De l'analyse secondaire à la réanalyse. Une innovation méthodologique en débats », *Recherches qualitatives* Hors-Série (21) : 7-28
- Dupuy C. & Duchesne S., 2017. « La réanalyse au service de l'interdisciplinarité ? », *Recherches qualitatives* Hors-Série (21) : 76-98
- Hamidi C., 2017. « Saisir les processus de catégorisation ethniques, sociales et territoriales dans les rapports ordinaires au politique : le prisme de la réanalyse », *Recherches qualitatives* Hors-Série (21) : 54-75
- Hammersley M., 1997. "Qualitative data archiving: Some reflections on its prospects and problems". *Sociology*, 31(1): 131-142
- Heaton J., 2008. "Secondary analysis of qualitative data: an overview", *Historical social research* 33 (3): 33-45
- Hinds PS, Vogel RJ & Clarke-Steffen L., 1997. "The possibilities and pitfalls of doing a secondary analysis of a qualitative data set", *Qual Health*, 7: 408-424.
- Huyghe M., Cailly L., Oppenheim N., 2018. Partage des données d'enquêtes en recherche qualitative : approche juridique, éthique, déontologique et épistémologique. Rapport de recherche, CITERES. 2018
- Irwin, S., Bornat, J., & Winterton, M., 2012. "Timescapes secondary analysis: Comparison, context and working across data sets", *Qualitative Research*, 12(1) : 66-80.
- Laferté G., 2016. « Retours d'expériences. Plaidoyer pour l'ethnographie collective », *ethnographiques.org* (32)
- Lejeune C., 2010. « Texte de la conférence d'ouverture – Montrer, calculer, explorer, analyser. Ce que l'informatique fait (faire) à l'analyse qualitative », *Recherches qualitatives* Hors-Série (9) : 15-32
- Lejeune C., 2017. « Analyser les contenus, les discours ou les vécus ? À chaque méthode ses logiciels ! », Santiago-Delefosse M. & del Rio Carral M. (dir.) *Les méthodes qualitatives en psychologie et sciences humaines de la santé*, Malakoff, Dunod, p. 203-224.
- Long Sutehall T., Sque M. & Addington-Hall J., 2010. "Secondary analysis of qualitative data: a valuable method for exploring sensitive issues with an elusive population?", *Journal of research in nursing*, 16 (4): 335-344
- Maison P., 2010. « La méta-analyse sur données résumées », *Recherche en soins infirmiers* (101)
- Ziebland S. & Hunt K., 2014. "Using secondary analysis of qualitative data of patient experiences of health care to inform health services research and policy", *Journal of Health services research & Policy* 0 (0)

