

Diplôme national de master

Domaine - sciences humaines et sociales

Mention – sciences de l’information et des bibliothèques

Parcours – sciences de l’information et des bibliothèques et
information scientifique et technique

MEMOIRE

Apprendre à penser en « données de la recherche » : le cas de l’équipe de recherche en SHS, AAU-CRESSON

Louise-Anne Charles

Sous la direction de Joachim Schöpfel

Maître de conférences en sciences de l'information et de la documentation –
GERiiCO, Université de Lille

Remerciements

Je souhaite remercier Françoise Acquier, ma tutrice d'alternance, avec qui j'ai pu échanger de nombreuses fois sur le sujet de la gestion des données, de la science ouverte et de l'équipe CRESSON, qui m'a conseillée et relue.

Cette réflexion a été alimentée par un premier dossier sur le laboratoire AAU, pour l'UE4 - Communication et sociologie des organisations et des TIC, réalisé avec Morgane Peypoch et Ellen Simmons.

Ce mémoire n'aurait pas été possible sans la participation des membres de l'équipe CRESSON, dont je remercie à la fois la rapidité avec laquelle ils ont accepté de participer et le temps qu'ils m'ont accordé.

Je tiens également à remercier mon directeur de mémoire Joachim Schöpfel, qui a su me guider dans mes réflexions, mes lectures et la mise en place de l'enquête, ainsi que Susan Kovacs avec qui j'ai pu échanger à de nombreuses reprises sur le sujet.

Enfin, merci à ma famille et mes amis pour leur soutien, en particulier Emilie Salort pour sa relecture.

Résumé :

Le mouvement pour une science ouverte développe plusieurs axes dont un qui promeut une meilleure gestion des données de la recherche et leur ouverture. Différentes institutions, dont l'Agence Nationale de la Recherche, ont pris position en faveur de cette ouverture des données et ont parfois défini des obligations pour les chercheurs pour pouvoir atteindre ces objectifs.

Ce mémoire s'intéresse à la prise en compte de ces injonctions dans une équipe de recherche en SHS et des pratiques existantes autour de la gestion des données. Il souligne l'écart entre adhérence aux valeurs de la science ouverte et pratiques, en démontrant également les difficultés éprouvées par les chercheurs et la question primordiale des moyens à accorder. Par rapport à d'autres écrits qui se situent au niveau d'un établissement ou d'une discipline, ce travail souligne l'intérêt de l'échelle de l'équipe ou du laboratoire dans des questions qui sont, en fin de compte, bien collectives.

Descripteurs : données de la recherche, science ouverte, laboratoire

Abstract:

The open science movement works towards different goals such as the improvement of research data management and their opening. In the last few years, a number of institutions engaged themselves in favour of that goal, such as the ANR (Agence Nationale de la Recherche) and developed obligations for researchers to reach it.

This work describes how a humanities and social sciences research team takes into account those obligations and their practices regarding research data management. It highlights the gap between adhering to the values of open science and practices, demonstrating the difficulties felt by researchers and the crucial question of the means to attain this goal. Usually about a discipline or an establishment, this work puts in the spotlight the value of using the scale of a laboratory or a research team for questions that are, undoubtedly, collective.

Keywords: research data management, open science, laboratory



Cette création est mise à disposition selon le Contrat : « **Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France** » disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr> ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Sommaire

SIGLES ET ABREVIATIONS	7
INTRODUCTION.....	9
PARTIE 1 : LE DEVELOPPEMENT DE LA SCIENCE OUVERTE	11
I. Naissance et caractéristiques de la science ouverte	11
1) <i>Origines et définition.....</i>	<i>11</i>
2) <i>Science ouverte et numérique.....</i>	<i>14</i>
II. Présentation de l'étude.....	17
1) <i>Terrain choisi et particularités</i>	<i>17</i>
2) <i>Méthodologie choisie.....</i>	<i>18</i>
III. La science ouverte vue par l'équipe	18
1) <i>L'arrivée de la notion de science ouverte</i>	<i>18</i>
2) <i>Définir la science ouverte</i>	<i>19</i>
3) <i>Adhésion à la science ouverte et limites</i>	<i>20</i>
4) <i>Un sujet discuté dans l'équipe ?</i>	<i>22</i>
PARTIE 2 : DEFINIR LA DONNEE DE LA RECHERCHE : DE LA GESTION AU PARTAGE CONTROLE.....	23
I. Définir les données de recherche et ses enjeux	23
1) <i>Etat des lieux</i>	<i>23</i>
2) <i>La vision des enquêtés sur la notion de donnée</i>	<i>25</i>
II. La gestion des données de la recherche des enquêtés	27
III. Le partage des données	30
1) <i>Fonctionnement du monde scientifique.....</i>	<i>30</i>
2) <i>Des choix personnels.....</i>	<i>31</i>
3) <i>Le point de vue des chercheurs enquêtés</i>	<i>31</i>
PARTIE 3 : DES CHOIX PERSONNELS A UNE ORGANISATION COLLECTIVE	37
I. Un « nous » qui apparaît	37
1) <i>Les enseignants-chercheurs</i>	<i>37</i>
2) <i>Nous, l'équipe CRESSON</i>	<i>38</i>
II. Les expériences de CRESSOUND, Cartophonies et de la bibliothèque du centre de documentation de l'équipe AAU-CRESSON : acquérir une expertise sur le son	42
1) <i>Naissance du projet CRESSOUND : impulsion par des ingénieurs d'étude</i>	<i>43</i>
2) <i>Naissance du projet Cartophonies : impulsion par des chercheurs</i>	<i>45</i>
3) <i>Fonctionnement et pérennité.....</i>	<i>47</i>

4) Une donnée : des représentations, des projets	48
III. La question finale de qui va s'en occuper : le coût du partage (et de la gestion)	48
1) Développer une expertise qui dépasse les enregistrements sonores 48	
2) La sensibilisation et l'aide à apporter aux chercheurs.....	49
CONCLUSION	53
BIBLIOGRAPHIE.....	55
ANNEXES.....	59
TABLE DES MATIERES.....	65

Sigles et abréviations

AAU : laboratoire Ambiances, Architectures, Urbanités

ANR : Agence nationale de la recherche

CNRS : Centre national de la recherche scientifique

CRENAU : Centre de recherche nantais architectures urbanités

CRESSON : Centre de Recherche sur l'Espace Sonore et l'environnement urbain

DOI : Digital Object Identifier

DPD : délégué à la protection des données personnelles

ENSAG : Ecole nationale supérieure d'architecture de Grenoble

FAIR (les principes) : Findable, Accessible, Interoperable, Reproducible

ORCID: Open Researcher and Contributor ID

PGD : Plan de gestion de données (en anglais DMP pour Data Management Plan)

RGPD : le Règlement Général sur la Protection des Données

UMR : Unité mixte de recherche

INTRODUCTION

Ces dernières années ont vu accroître l'importance des données de la recherche dans le domaine scientifique. Cette importance se traduit par un investissement politique et financier forts, qui ont mené par exemple en juillet 2022 au lancement de l'écosystème Recherche Data Gouv (à la fois portail, entrepôt de données et certification de personnes ressources, formations sur le sujet).

L'ouverture des données de recherche fait également l'objet de nombreux plans pour leur valorisation, comme le sous-entend le site de la science ouverte : « Les données de la recherche constituent un très fort potentiel de connaissance et d'innovation, qui reste pour l'instant insuffisamment exploité »¹.

Différents articles et rapports ont tenté de comprendre les raisons du manque d'exploitation des données de la recherche par les chercheurs, qu'il s'agisse de leurs pratiques de sauvegarde, de partage ou de réutilisation. Par exemple, ils ont pu mettre en avant le rôle de la discipline², ou les pratiques de partage déjà existantes dans des cercles plus restreints et basées sur une relation de confiance³.

Cependant, dans le cadre de la science ouverte, mouvement qui cherche à mettre à disposition de manière libre et gratuite en ligne les productions scientifiques (dont les données), le partage doit se faire dans un cadre qui garantisse au mieux la visibilité et la réutilisation des données par le plus grand nombre d'acteurs possibles. Ce cadre, qui répond à des impératifs numériques, sera décrit avec plus de précision dans une première partie.

Par conséquent, des outils sont développés, comme l'entrepôt de données mentionné auparavant Recherche Data Gouv, et un besoin de formation se fait ressentir.

Le sujet de ce mémoire est né de l'écart apparent entre le discours politique et institutionnel sur l'importance des données et les premières réactions que j'ai observé sur mon lieu d'alternance, le centre de documentation du laboratoire AAU-CRESSON, qui pouvaient remettre en cause l'importance de l'ouverture des données en sciences humaines et sociales allant jusqu'à dire « Je crois que cela va être un grand cimetière de données »⁴.

Ce mémoire s'intéressera à la réception des discours sur l'ouverture des données de la recherche, et plus précisément à comment ceux-ci s'adaptent au cadre et au raisonnement au sein d'une équipe de laboratoire. Il s'agira également

¹ OUVRIR LA SCIENCE, 2022. La ministre de l'Enseignement supérieur et de la Recherche, Sylvie Retailleau, inaugure Recherche Data Gouv. *Ouvrir la science*. [en ligne]. [Consulté le 20 juillet 2022]. Disponible à l'adresse: <https://www.ouvrirlascience.fr/la-ministre-de-lenseignement-superieur-et-de-la-recherche-sylvie-retailleau-inaugure-recherche-data-gouv>

² SERRES, Alexandre, MALINGRE, Marie-Laure, MIGNON, Morgane, PIERRE, Cécile et COLLET, Didier, 2017. *Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l'Université Rennes 2* [en ligne]. Research Report. Université Rennes 2. [Consulté le 3 novembre 2021]. Disponible à l'adresse: <https://hal.archives-ouvertes.fr/hal-01635186>

³ REBOUILLAT, Violaine, 2021. Le partage des données vu par les chercheurs : une approche par la valeur. *Les Enjeux de l'Information et de la Communication*. [en ligne]. 2021. N° 22/1, pp. 35-53. [Consulté le 30 novembre 2021]. Disponible à l'adresse: <https://lesenjeux.univ-grenoble-alpes.fr/2021/varia/03-le-partage-des-donnees-vu-par-les-chercheurs-une-approche-par-la-valeur/>

⁴ E3_chercheur

d'observer l'influence des pratiques préexistantes de sensibilisation à la science ouverte et de partages de données dans cette équipe.

Par rapport à d'autres études effectuées sur le sujet qui pouvaient se faire à l'échelle d'un établissement ou d'une discipline, ce mémoire se concentre sur l'équipe CRESSON du laboratoire AAU qui a la particularité d'être interdisciplinaire.

Cette échelle réduite part de l'hypothèse suivante : le laboratoire est un lieu où les injonctions sont travaillées et transformées. Sur les questions de médiation et de sensibilisation, telles qu'elles ont été soulevées par Violaine Rebouillat dans sa thèse⁵, c'est une échelle qui peut se révéler intéressante puisqu'une politique ou des bonnes pratiques peuvent y être également formulées.

En effet, la deuxième hypothèse est que l'ensemble des pratiques autour des données ne relèveraient pas seulement de choix personnels, d'un chercheur face à son ordinateur, mais également de constructions collectives où plusieurs paramètres doivent être pris en compte et où le contexte a son importance. La dimension collective de la gestion des données de la recherche pourrait être ainsi soulignée.

Enfin, le manque d'exploitation des données de la recherche serait moins dû à un manque d'adhésion aux principes de la science ouverte, qu'à ces façons de procéder et au cadre auquel il faut adhérer.

Le mémoire se divise en trois parties. La première partie est dédiée à la science ouverte, sa définition et la perception des enquêtés sur le sujet. La deuxième s'interroge sur les pratiques actuelles des enquêtés autour des données de la recherche, après les avoir définies. La troisième s'interroge sur la dimension collective des données de la recherche, en étudiant deux projets de partage de données portés par l'équipe depuis plusieurs années sur des enregistrements sonores.

⁵ REBOUILLAT, Violaine, 2019. *Ouverture des données de la recherche: de la vision politique aux pratiques des chercheurs*. [en ligne]. Sciences de l'information et de la communication. Conservatoire national des arts et métiers. [Consulté le 6 juillet 2021]. Disponible à l'adresse: <https://tel.archives-ouvertes.fr/tel-02447653/document>

PARTIE 1 : LE DEVELOPPEMENT DE LA SCIENCE OUVERTE

I. NAISSANCE ET CARACTERISTIQUES DE LA SCIENCE OUVERTE

1) Origines et définition

L'introduction a été l'occasion de mettre à plat une première définition de la science ouverte. Cette partie s'attache à interroger de manière plus fine ses enjeux, sa mise en œuvre, et la vision qu'en ont des chercheurs dans une équipe. Plutôt que de partir sur une nouvelle définition, il semblait plus intéressant d'essayer de dater la science ouverte. D'où vient-elle exactement ?

La réponse ne peut pas être immédiate ni unique. Julien Roche, dans son entrée de dictionnaire « Science ouverte (origine) » date le mouvement du fin 16^{ème} - 17^{ème} siècle. En effet, la révolution scientifique a déjà pour enjeu de permettre une meilleure circulation des savoirs, ce qui va amener par exemple à l'apparition de deux journaux scientifiques⁶. Ici, ce n'est pas tant la notion de « science ouverte » qui existerait, puisqu'elle est assez récente, qu'une première expression de ses valeurs et de ses enjeux dans le système technique de l'époque.

Dans l'article « Bref historique de la science ouverte européenne »⁷, l'origine de la science ouverte prendrait plus précisément racine dans le mouvement du libre accès, qui daterait des années 1990-2000. Cette étape cruciale est également mentionnée par Julien Roche : « La science ouverte est assurément une forme de prolongement et d'amplification du mouvement d'accès ouvert, né dans les années 1980 et qui a pris son essor dans le cours des années 1990 »⁸.

Le libre accès, ou accès ouvert, est un mouvement qui vise à rendre plus accessible les publications scientifiques en en donnant un accès gratuit en ligne. Ce mouvement naît alors qu'Internet permet ce partage et que le coût des abonnements aux revues scientifiques ne cesse de grimper, devenant un poids dans le budget des bibliothèques⁹. Différentes voies pour publier les publications scientifiques ont été privilégiées au fur et à mesure des années :

- La voie dorée, où la publication se fait dans une revue en accès ouvert (Open Access), l'accès est alors gratuit pour tous et toutes en ligne. Cette voie a plusieurs modèles, dont celui diamant « lorsque les fonds sont fournis par une institution de recherche »¹⁰. Le prix n'est plus pris en charge par l'auteur

⁶ ROCHE, Julien, 2022. Science ouverte (origine). In : CAULI, Marie, FAVIER, Laurence et JEANNAS, Jean-Yves, *Dictionnaire du numérique*. Londres : ISTE Editions. pp. 287-290. Sciences, société et nouvelles technologies.

⁷ PASTEUR, Ceris-Institut, 2022. Bref historique de la science ouverte européenne. *Open science : évolutions, enjeux et pratiques*. [en ligne]. 7 avril 2022. [Consulté le 8 juin 2022]. Disponible à l'adresse: <https://openscience.pasteur.fr/2022/04/07/bref-historique-de-la-science-ouverte-europeenne/>

⁸ ROCHE, Julien, *op. cit.*, p. 289.

⁹ PASTEUR, Ceris-Institut, *op. cit.*

¹⁰ BLANC-FERACCI, Caroline, BRAUD, Marion, FORGET, Camille, MICHEL, Eugénie et TOURNIER, Floriane, 2022. « Bibliothèques, éditeurs, libraires, face au mouvement de l'open ». *Bulletin des bibliothèques de France*. [en ligne].

ou le lecteur mais par l'institution. Ce dernier modèle a été récemment privilégié par plusieurs institutions (Coalition S, Science Europe, OPERAS, ANR) avec le déploiement du plan d'action pour l'accès ouvert « diamant », annoncé le 4 février 2022, et par l'axe 1 du second plan national pour la science ouverte¹¹.

- Et la voie verte où, en parallèle d'une publication dans une revue, le chercheur dépose dans une archive ouverte : la version auteur de son texte (c'est-à-dire avant la mise en forme éditeur), son preprint, qui est le texte soumis au comité de relecture ou l'article selon l'accord de l'éditeur.

C'est dans ce contexte-là et dans l'objectif de cette deuxième voie que HAL, archive ouverte en ligne développé par le centre pour la communication scientifique directe (CCSD) en 2001 sur le modèle d'ArXiv (archive ouverte précurseur en physique et mathématiques), prend forme et devient de plus en plus important au fur et à mesure des années jusqu'à être soutenu dans l'axe 1 du premier plan national pour une science ouverte (2018-2021).

Le libre accès fait bien parti des axes de la science ouverte, mais ne le résume pas puisque d'autres enjeux sont également cités dans les deux plans nationaux pour une science ouverte et dans d'autres feuilles de route institutionnelles comme la « feuille de route du CNRS pour la science ouverte » de 2019 :

- L'accès aux données de la recherche, qui fait l'objet d'une description dans une partie dédiée et qui est au cœur de ce mémoire ;
- L'accès aux codes sources qui ont permis la production des données et qui renseignent sur la construction du programme informatique ;
- Le changement de système d'évaluation de la recherche.

Ce dernier point est souvent décrit dans les articles comme un point nécessaire pour permettre l'ouverture de la science¹², et notamment dans les articles que j'ai pu lire sur l'ouverture des données de la recherche (voir partie 2). Elle est aujourd'hui jugée trop portée sur la quantité de la production scientifique (notamment les articles de revues labellisés par les grandes bases de données) plutôt que sur la qualité, et ne prenant pas en compte la plus grande diversité de productions scientifiques.

La note du 23 mars 2022 du Secrétariat général du Conseil l'exprime ainsi : « RECONNAÎT que les systèmes d'évaluation de la recherche devraient être axés sur la qualité, l'excellence ainsi que l'impact, et promouvoir le partage des connaissances et la collaboration à un stade précoce, et RAPPELLE que les systèmes actuels d'évaluation de la recherche sont, dans l'ensemble, trop axés sur l'utilisation d'indicateurs quantitatifs, ce qui favorise les biais négatifs en termes de qualité, de reproductibilité et d'intégrité de la recherche, et est généralement incompatible avec

[Consulté le 8 juin 2022]. Disponible à l'adresse: https://bbf.enssib.fr/tour-d-horizon/bibliotheques-editeurs-libraires-face-au-mouvement-de-l-open_70298

¹¹ OUVRIR LA SCIENCE, 2022. Plan d'action pour les revues « diamant ». *Ouvrir la science*. [en ligne] [Consulté le 29 juin 2022]. Disponible à l'adresse: <https://www.ouvrirlascience.fr/plan-daction-pour-les-revues-diamant>

¹² CNRS, 2019, *Feuille de route du CNRS pour la science ouverte* [en ligne]. CNRS. [Consulté le 29 juin 2022]. Disponible à l'adresse: https://www.science-ouverte.cnrs.fr/wp-content/uploads/2019/11/Plaqueette_Science-Ouverte_18112019.pdf

la mise en œuvre des pratiques de la science ouverte »¹³. Les pratiques de bibliométrie sont remises en cause.

La science ouverte est alors un élargissement des objets de préoccupation du mouvement de libre accès et d'autres mouvements de partage, étendant l'enjeu d'accessibilités à d'autres productions scientifiques que les résultats et interrogeant le système même de la recherche. Elle semble également représenter une institutionnalisation du mouvement d'ouverture de la science, avec une approche qui serait plus « top-down » (décision prise en haut de la hiérarchie avant de descendre dans celle-ci), que le mouvement pour un libre accès.

En effet, deux plans pour une science ouverte ont été effectués au niveau national : le premier concerne 2018-2021 et le deuxième 2021-2024. Leur objectif, tel décrit lors du premier plan, est que « les résultats de la recherche scientifique soient ouverts à tous, chercheurs, entreprises et citoyens, sans entrave, sans délai, sans paiement »¹⁴. Au niveau français, d'autres acteurs de la recherche ont pris position pour la science ouverte comme le CNRS, l'Agence nationale de la recherche, et certaines universités et écoles à travers des engagements, incluant récemment l'Ecole Nationale supérieure des sciences de l'information et des bibliothèques.

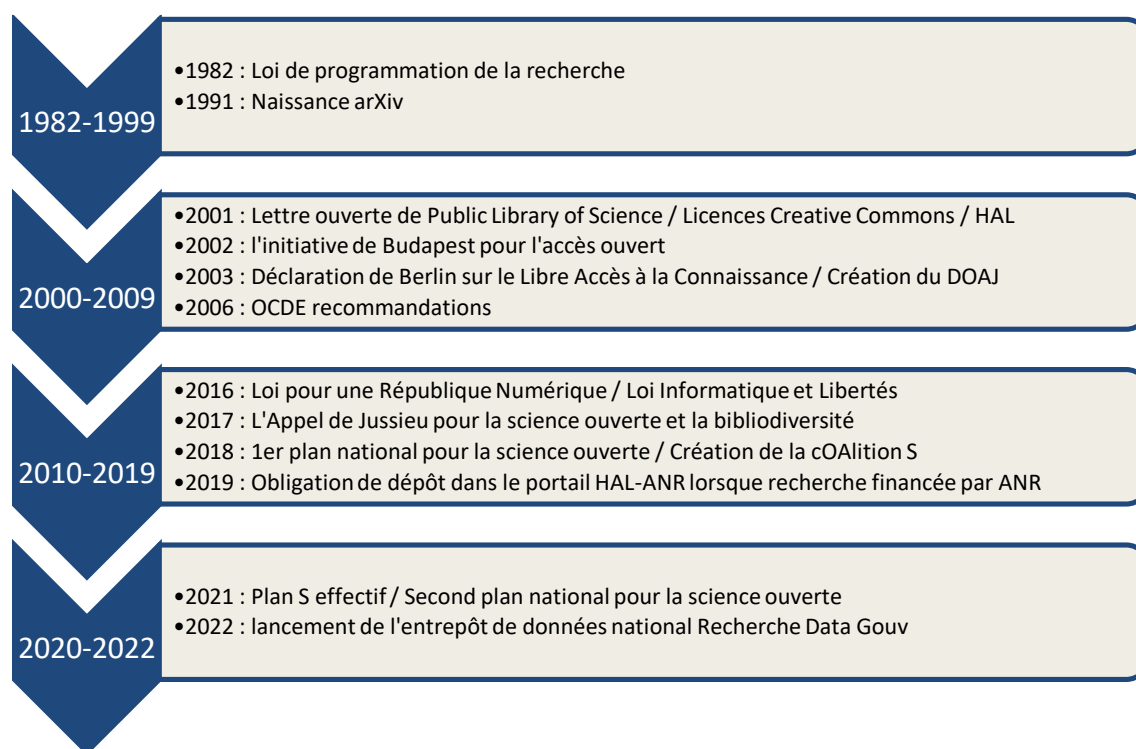
La difficulté de marquer le début de cette notion provient également du fait qu'il s'agit d'une construction progressive nationale, européenne et internationale (« open science »). De plus, cette construction s'appuie sur d'autres mouvements et actions, comme par exemple l'utilisation des licences Creative Commons¹⁵ par HAL.

La science ouverte ne se délimite qu'avec difficulté et s'appuie sur des mouvements divers. En dessous, une frise chronologique retrace quelques dates marquantes des mouvements d'ouverture (sans exhaustivité).

¹³ SECRETARIAT GENERAL DU CONSEIL, 2022. *Projet de conclusions du Conseil sur l'accélération de la mise en œuvre et de l'impact des politiques et des pratiques de la science ouverte*. Note du 23 mars 2022 [consulté 23/03/2022]. Disponible à l'adresse : <https://data.consilium.europa.eu/doc/document/ST-7390-2022-INIT/fr/pdf>

¹⁴ MINISTERE DE L'ENSEIGNEMENT SUPERIEUR, DE LA RECHERCHE ET DE L'INNOVATION, 2018. *Plan national pour la science ouverte* (1). [en ligne] [Consulté le 29 juin 2022]. Disponible à l'adresse: https://www.ouvrirlascience.fr/wp-content/uploads/2018/08/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf

¹⁵ Licences créées pour les contenus publiés sur Internet. Elles permettent d'inciter au partage de ceux-ci tout en précisant les modalités de ce partage. La liste des licences disponibles est disponible sur le Web : <https://creativecommons.org/about/ccllicenses/>



Plusieurs de ces dates ou des organisations ont été discutées auparavant dans cette partie alors que d'autres reviendront lors de la partie sur les données de la recherche. La loi pour une République Numérique a notamment autorisé la diffusion ouverte des productions scientifiques, en créant un nouveau cadre juridique qui permet la voie verte, décrite plus en amont. La cOAlition S est un consortium composé d'agences de financement de la recherche de plusieurs pays européens. Elle est à l'origine du plan S, initiative en faveur du libre accès.

2) Science ouverte et numérique

Abordé que de manière transversale jusque-là dans le mémoire, le numérique, avec ses enjeux particuliers, joue un rôle important dans la mise en place de la science ouverte et sur ses conditions. Il est d'abord celui qui a rebattu les cartes dans l'édition scientifique et qui a mené au mouvement de l'accès libre, avec d'un côté l'enjeu du prix des abonnements et de l'autre les possibilités de partage d'Internet, comme discuté précédemment.

Le numérique est l'outil par lequel passe les travaux vers une science ouverte. Florence Millerand revient sur le discours entre sciences et numérique dans son article « les imaginaires de la « science 2.0 » ». Celui-ci relèverait : « de la promesse d'une « science meilleure », qui prendrait appui sur une pratique scientifique plus ouverte, plus collaborative, décloisonnée, et plus en phase avec la société »¹⁶.

Ce discours se retrouve dans la feuille de route pour une science ouverte de 2019 du CNRS : la naissance du mouvement international pour une science ouverte est datée de 30 ans, bien qu'il se développe de manière « sans précédent depuis que le web l'a rendu possible à une échelle globale avec des coûts raisonnables »¹⁷. Le

¹⁶ MILLERAND, Florence. Les imaginaires de la « science 2.0 ». *Communication. Information médias théories pratiques* [en ligne], no vol. 33/2, 2015 [consulté le 13 novembre 2021]. Disponible à l'adresse : <https://doi.org/10.4000/communication.6070>

¹⁷ CNRS, *op.cit.*

premier plan national pour une science ouverte met également en avant le numérique : « Elle [la science ouverte] s'appuie sur l'opportunité que représente la mutation numérique pour développer l'accès ouvert aux publications et – autant que possible – aux données de la recherche »¹⁸.

Néanmoins, ce qui est à noter et qui sera également crucial pour la gestion des données, c'est que la science ouverte entre en relation avec des enjeux propres au numérique. Une structuration forte tente d'émerger sur le Web, avec notamment le mouvement protéiforme en faveur du web de données.

Ce mouvement vise à structurer les données, à les relier les unes avec les autres pour ne plus créer de silos isolés. Par exemple, dans le cas de la transition bibliographique, cela est venu du constat que de nombreuses bibliothèques produisaient et avaient besoin de données identiques concernant le catalogage de documents. Au lieu de continuer à les produire tous, l'idée serait de pouvoir les produire à un endroit pour ensuite les communiquer.

Pour cela, ce mouvement cherche également à écrire de manière optimale pour les machines. Une différence se fait entre le web dit 1.0 qui se base sur des liens entre pages HTML et le web de données (ou encore Web sémantique ou Web 3.0) qui met en place des normes et des technologies pour faciliter l'exploitation des données par les machines :

- RDF, resource description framework, qui représente la grammaire avec laquelle on décrit, représente et relie les données à échanger sur le web, et qui permet la communication des bases de données entre elles (interopérabilité). Il part du postulat que tout peut se décomposer en une formule simple : sujet, verbe, complément.
- Vocabulaire et ontologies comme le FOAF (friend of a friend) ;
- URI (Uniform Resource Identifier), qui permet d'avoir un identifiant pérenne même si la ressource est déplacée ;
- Et SPARQL qui est un langage d'interrogation.

Nous soulignerons deux mots d'ordre dans ce contexte en particulier. Le premier est l'interopérabilité qui signifie la « capacité de matériels, de logiciels ou de protocoles différents à fonctionner ensemble et à partager des informations ».¹⁹ Le deuxième est pérennité. Le numérique, comme les formats, a connu de nombreuses évolutions en quelques dizaines d'années. L'objectif est de garantir un accès aux ressources, en évitant par exemple les erreurs 404.

Concrètement, dans le monde scientifique, cela s'est traduit par le développement et la mise en avant d'identifiants uniques : à la fois pour les chercheurs (IdHAL, ORCID, ...), les institutions (ROR), et les productions scientifiques avec le DOI (digital object identifier). L'objectif est également de faire communiquer ces identifiants entre eux, c'est ainsi qu'il est possible de lier l'IdHAL avec son ORCID et de réunir toutes ses productions dans un unique CV.

Cela impacte également les outils en soutien de la science ouverte : en janvier 2022, DMP OPIDoR, outil qui accompagne les utilisateurs dans la rédaction de plans de gestion de données, a annoncé le développement d'un nouveau modèle dit

¹⁸ MINISTERE DE L'ENSEIGNEMENT SUPERIEUR, DE LA RECHERCHE ET DE L'INNOVATION, *op. cit.*

¹⁹ EDITIONS LAROUSSE, [sans date]. Définitions : interopérabilité. Larousse. [en ligne]. [Consulté le 29 juin 2022]. Disponible à l'adresse : <https://www.larousse.fr/dictionnaires/francais/interop%C3%A9rabilit%C3%A9/43787>

structuré. Ce nouveau modèle, appelé également « machine actionnable », vise à faciliter l'exploitation de plans de gestion de données par des machines en proposant par exemple un « noyau commun » pour une plus grande interopérabilité et plus de référentiels que les autres modèles classiques²⁰.

En effet, les outils se dotent alors de référentiels avec lesquels ils sont en interface, à l'instar de DMP OPIDoR avec Wikidata²¹, base de connaissance avec des données structurées, ou la vidéothèque de l'enseignement supérieur et de la recherche Canal-U avec IDRef, une base d'identifiants pour l'enseignement supérieur et de la recherche basée sur autorités du Sudoc, de Calames et Star.

Par rapport à la pérennité, une attention est portée aux formats des données, avec une différence qui se fait entre format ouvert et propriétaire. Une liste de recommandations de formats a été établie par le Centre informatique national de l'enseignement supérieur (CINES) à l'adresse suivante : <https://facile.cines.fr/>. Ce sont les seuls formats acceptés en cas d'archivage numérique pérenne par le CINES.

Cette attention entre le privé ou le propriétaire en comparaison avec le public ou l'ouvert se ressent également sur les comparaisons qui sont faites entre ResearchGate, un réseau social académique ou d'autres réseaux du même type (services commerciaux), et HAL, une archive ouverte. La pérennité et la transparence sont particulièrement questionnées dans le cas des réseaux académiques²².

L'objectif de cette partie est de montrer comment la notion de « science ouverte » peut recouvrir une vision très formalisée de ce qu'est l'ouverture sur le Web et de ce qui est attendu. Ceci dépasse les seules valeurs d'accessibilité et de partage, ou plutôt elles sont redéfinies par ce contexte pour s'assurer de l'accessibilité. Des entrepôts de données, des identifiants uniques et des formats sont ainsi favorisés par rapport à d'autres.

Comme ce sont des outils qui sont en mutation avec un mouvement relativement récent, ils ne sont pas stabilisés et les intérêts de cette structuration sont souvent encore en cours. L'utilisation de référentiels a pu se révéler compliqué dans mon expérience d'alternance car ils ne sont pas forcément justes : des informations datées par exemple qu'il faut mettre à jour. Il faut alors savoir pouvoir modifier ces référentiels, puisqu'il n'est pas possible de les modifier à partir des outils qui en exploitent le contenu.

C'est en cela également que la science ouverte continue à se définir ; entre le mouvement du libre accès et l'institutionnalisation de la science ouverte, l'environnement du Web s'est profondément modifié. Et il ne fait aucun doute que dans quelques années, celui-ci soit encore différent.

²⁰ INIST-CNRS, 2022. DMP OPIDoR – Le modèle structuré. *DoRANum*. [en ligne]. [Consulté le 3 juillet 2022]. Disponible à l'adresse: https://doranum.fr/plan-gestion-donnees-dmp/dmp-opidor-le-modele-structure_10_13143_mzze-kz59/

²¹ Ce n'est pas le seul référentiel utilisé par DMP OPIDoR. La liste est accessible sur leur site : https://dmp.opidor.fr/static/about_registries.

²² BOUCHARD, Aline, 2016. *Éléments de comparaison : archives ouvertes (HAL) et réseaux sociaux académiques (Academia, Researchgate)* [en ligne]. URFIST de Paris. [Consulté le 29 juin 2022]. Disponible à l'adresse: <https://www.enssib.fr/bibliotheque-numerique/documents/67085-elements-de-comparaison-archives-ouvertes-hal-et-reseaux-sociaux-academiques-academia-researchgate.pdf>

II. PRESENTATION DE L'ETUDE

1) Terrain choisi et particularités

L'enquête concerne des chercheurs et des ingénieurs de l'équipe CRESSON, Le Centre de Recherche sur l'Espace Sonore et l'environnement urbain.

L'équipe CRESSON est l'équipe grenobloise du laboratoire AAU, Ambiances, Architectures, Urbanités créé en 1998. Ce dernier est une unité mixte de recherche du Centre National de la Recherche Scientifique (CNRS) et du Ministère de la Culture. La deuxième équipe, CRENAU (Le Centre de Recherche Nantais Architectures Urbanités), est située à Nantes.

Le laboratoire AAU a pour champs de recherche la ville, l'environnement, l'architecture, et a la particularité d'être pluridisciplinaire, à la croisée de l'urbanisme, de l'architecture, des sciences sociales et des sciences de l'ingénieur. Ses méthodologies sont diverses, allant de la réalité virtuelle à des enquêtes ethnographiques. Il est composé de 150 membres²³. Le laboratoire AAU appartient à la section 39 espace, territoire, et société de l'institut national des sciences humaines et sociales du CNRS.

Il a pour tutelles le CNRS, l'Ecole Nationale Supérieure d'Architecture de Grenoble (ENSAG) et de Nantes (Ministère de la Culture), l'Ecole Centrale de Nantes, et, récemment, deux établissements publics expérimentaux Nantes Université et Université Grenoble Alpes.

L'équipe CRESSON, composée à la fois de chercheurs permanents, contractuels et associés, de doctorants et d'un service de soutien et d'appui à la recherche, a été créée en 1979.

Sa création a fait suite à un cours commun sur le son à l'ENSAG avec deux des futurs créateurs de l'équipe : Jean-François Augoyard et Jean-Jacques Delétré. L'équipe s'est d'abord intéressé aux espaces sonores avant de s'étendre à d'autres perceptions sensibles in situ comme l'odorat, en développant la notion d'« ambiance ». Les recherches peuvent aborder à la fois les pratiques professionnelles, celles des habitants sur un territoire, comment capter ces sensations et ressentis et comment les restituer.

L'intérêt de cette équipe est d'avoir développé des projets depuis plusieurs années sur un type de donnée en particulier, le son, en ayant pour objectif son partage et son archivage. De plus, cette équipe est sensibilisée aux enjeux d'ouverture de la science depuis 2001 avec l'arrivée d'HAL et a pu observer des évolutions. Le travail sur la sensibilisation à la gestion des données, bien que plus récent, date tout de même de plusieurs années. Deux projets, financés par l'Agence Nationale de la Recherche, obligent également certains membres de l'équipe à se confronter à ses enjeux, puisque le financeur demande la rédaction de plusieurs plans de gestion de données.

²³ A propos du laboratoire, [sans date]. *Laboratoire AAU*. [en ligne]. [Consulté le 29 juin 2022]. Disponible à l'adresse: <https://aau.archi.fr/laboratoire-aau/>

2) Méthodologie choisie

La méthodologie privilégiée par ce mémoire est celle d'entretiens semi-directifs de membres de l'équipe du CRESSON. Ces entretiens visaient à aborder à la fois la vision de la science ouverte, leur gestion des données et leur compréhension ou utilisation des projets CRESSOUND et Cartophonies. Ils ont concerné 9 membres de l'équipe, dont 7 chercheurs et 2 ingénieurs de recherche. Ces entretiens ont duré entre 1h et 1h50 et la trame utilisée se trouve en Annexe 1 pour les chercheurs et Annexe 2 pour les ingénieurs de recherche.

Les chercheurs ont été sélectionnés sur deux critères majoritaires : leur antériorité dans l'équipe et leur proximité avec des projets d'ouverture de publications ou de données.

Les ingénieurs d'études sélectionnés sont la documentaliste Françoise Acquier et le médiaticien/ingénieur son juL McOisans, qui sont les moteurs du travail sur le partage et l'archivage des sons.

Ce mémoire ne pourra pas en prenant cette petite échelle faire de généralisation mais espère mettre en avant les particularités et les nuances sur un même sujet dans une équipe de recherche.

III. LA SCIENCE OUVERTE VUE PAR L'EQUIPE

1) L'arrivée de la notion de science ouverte

Le fait que le mouvement pour une science ouverte prenne racine à l'intérieur du mouvement du libre accès explique les réponses contrastées des chercheurs et des ingénieurs d'études lorsqu'ils ont été questionnés sur la première fois qu'ils avaient entendu parler de la « science ouverte ». Une grande majorité l'ont estimée à 4-5 ans, ce qui correspondrait à la date du premier plan pour une science ouverte. Il est à noter déjà ici que, contrairement à ce qui était l'hypothèse de départ, le rôle des tutelles et des financeurs n'ont été que très rarement discutés lors des entretiens.

En effet, l'introduction de cette notion dans le laboratoire s'est faite par le travail de la documentaliste Françoise Acquier, qui estime également en avoir entendu parler il y a 4-5 ans. Cependant, à l'instar d'autres, pour elle ce travail d'ouverture de la science (et non de la « science ouverte ») date des années 2000 avec le mouvement du libre accès « Parce qu'au départ ça passait plus par cette idée de libre accès des publications [...] depuis 2001 ça existe dans le travail. Ça a commencé autour des publications et puis ça s'est étendu à d'autres activités de la recherche »²⁴. D'autres enquêtés le date de l'introduction des licences Creative Commons « Il me semble qu'avec Françoise on a été un peu les premiers à l'entendre à l'avoir saisi, [...] quand il faut diffuser en open source et en creative commons, des choses comme ça, je crois que depuis les années 2000 – 2005 on a commencé à faire ce genre de choses »²⁵, l'introduction de HAL qui commence à changer les pratiques « Quand, je ne sais pas, une dizaine d'années, peut-être ? Et je pense que ça devait être HAL qui a déclenché mon attention à ça. »²⁶, ou la présence de HAL déjà dans les pratiques « Par contre, à partir du moment où je suis plus rentré dans

²⁴ E8_ITA

²⁵ E9_ITA

²⁶ E2_chercheur

la partie recherche [autour de 2013-2015], en fait il y avait déjà HAL et donc de fait... j'ai commencé la recherche et la science ouverte était déjà là »²⁷.

L'institutionnalisation du mouvement d'ouverture de la science qui ne concerne plus seulement les publications a été ressenti par les deux ingénieurs de recherche qui voyaient le mouvement du libre accès comme plus marginal au départ : « On pourrait pas penser autrement aujourd'hui mais il y a 10-15 ans on n'était pas les pionniers mais les premiers à se dire – ah apparemment il faut faire ça maintenant, et bien faisons le. Essayons ! »²⁸, « Ça s'institutionnalise. C'est moins des militants, des réseaux de bibliothèque ou d'informaticien ou disciplinaires, ça c'est institutionnalisé, ça s'appelle science ouverte, c'est un peu plus... ça prend une autre dimension et un autre vocable »²⁹. Ce premier mouvement qui est décrit comme plus militant a été amené dans le contexte de cette équipe par le réseau de documentalistes de laboratoires qui se réunissaient alors chaque année.

2) Définir la science ouverte

Lors de la question « pouvez-vous définir la science ouverte », les chercheurs ont généralement précisé que c'était leur vision de la définition et qu'ils n'étaient pas certains. D'un point de vue global, tous avaient une idée des valeurs et de l'objectif global du mouvement de la science ouverte, en accentuant l'aspect accessibilité (« rendre accessible »³⁰, « mettre à disposition »³¹ « partage »³²).

Alors que les questions étaient faites de façon à être ouvertes pour laisser plus de liberté à l'enquête et qu'aucune ne concernait directement HAL, l'outil a été à chaque fois cité montrant le lien fort reliant l'outil à la notion de science ouverte. Les connaissances semblaient plus complètes lorsqu'il s'agissait de la question de l'accès aux publications, tandis que les données de la recherche « ça c'est toute une autre discussion »³³, sans doute comme la sensibilisation au libre accès remonte à plus d'une quinzaine d'années dans l'équipe et a fait le fruit de différentes stratégies : « Avec des hauts et des bas, et énormément de réticences, et en ayant fait jouer énormément de stratégies. Pleins de choses ont été testées. J'ai fait moi-même, j'ai fait venir des gens du CCSD, j'ai fait venir des gens de l'UGA, il y a eu des phases comme ça, on avance, on descend »³⁴. Françoise Acquier estimait qu'il y avait dans l'équipe une certaine adhésion et que certaines postures opposées chez les chercheurs de l'équipe avaient disparu.

Les autres axes, l'ouverture du code source et l'évaluation de la recherche n'ont pas été mentionnées, sauf par la documentaliste Françoise Acquier qui a pu décrire chacun des axes, le contexte et les origines. Les causes étaient rarement mises en avant par les chercheurs (prix de la revue, etc.), sauf par un qui a également témoigné de l'opacité des sujets pour un chercheur (« C'est cette partie-là qui est invisible au chercheur c'est la valeur marchande, que paie les universités [...] Pour

²⁷ E7_chercheur

²⁸ E9_ITA

²⁹ E8_ITA

³⁰ E2_chercheur, E6_chercheur

³¹ E1_chercheur

³² E7_chercheur

³³ E5_chercheur

³⁴ Françoise Acquier

développer une conscience collective sur ces questions, on doit avoir l'information. A partir de là on peut se dire qu'il y a un problème »³⁵).

3) Adhésion à la science ouverte et limites

De ces entretiens, une globale adhésion aux valeurs et principes de la science ouverte se démarque, qui serait non seulement bien, mais nécessaire ou naturelle « Qu'est-ce qu'on pourrait faire sinon, faire nos recherches et les garder dans nos placards ? Ou alors la donner à nos commanditaires et ça sera dans leur placard à eux. Non tant qu'à faire on le met en partage »³⁶. Ce processus d'ouverture des publications sur HAL a même porté des fruits « J'ai reçu un mail l'autre jour ; je suis tombée sur HAL sur une super recherche que vous aviez fait, ça m'intéresse, et une chose qu'on avait fait, il y a 40 ans... »³⁷ ce qui a donné lieu à une collaboration pour une émission de radio en utilisant des sons précédemment produits.

Cependant cette adhésion est vite teintée dans le discours par la mise en place de limites à cette science ouverte, limite qui diffère d'un enquêté à l'autre et par une différence faite entre adhésion à des principes et pratiques de la science ouverte, qui est revenue particulièrement à la question de leur perception de l'avis de leurs collègues : « je ne dis pas que c'est la pratique qui est partagée, le sentiment [positif] oui »³⁸, « Je pense que la question c'est plutôt le décalage entre trouver ça super et le fait de pas prendre le temps ou de pas maîtriser les choses. Y a un retard à l'allumage dans la façon dont on peut s'impliquer, dont on peut rendre opératoire les outils »³⁹.

L'archive ouverte en ligne HAL a particulièrement été l'objet de différentes critiques, plus ou moins partagées par les enquêtés. Le sentiment qui est revenu plusieurs fois est que la prise en main de l'outil avait été difficile : « l'effort initial était un peu dur »⁴⁰, « il y a un côté un peu ingrat au début. Maintenant que c'est mis en marche, c'est hyper agréable de montrer ses travaux et de montrer les travaux des uns, des autres. Mais c'est vrai que la première marche était un peu raide »⁴¹. Le sentiment de confusion a persisté pour certains chercheurs que ce soit par rapport à l'outil en lui-même ou par rapport au statut de l'article à déposer et des délais réglementaires. Alors que certains chercheurs ont intégré HAL dans leur pratique, d'autres font les dépôts par obligation et doutent de l'intérêt de l'ouverture permise par l'archive.

Les notions d'« ouverture » et son pendant négatif de « fermeture » n'ont été finalement que très peu interrogés par les enquêtés directement. L'analyse menée par Stéphane Olives en 2006⁴² du terme « changement » montre comment par la caractérisation positive d'un terme, celui-ci devient un impératif auquel doit adhérer les membres d'une organisation. Il identifie trois fonctions pour ce terme, qui

³⁵ E1_chercheur

³⁶ E9_ITA

³⁷ E9_ITA

³⁸ E5_chercheur

³⁹ E2_chercheur

⁴⁰ E5_chercheur

⁴¹ E6_chercheur

⁴² OLIVESI, Stéphane, 2006. *La Communication au travail. Une critique des nouvelles formes de pouvoir dans les entreprises*. 2^{ème} éd. Grenoble : PUG. Communication en plus.

peuvent également correspondre à l'usage du terme « ouverture » bien qu'il ne s'agisse là ni d'une entreprise, ni de discours patronaux, et que ce terme se trouve dans l'univers scientifique ainsi que d'autres comme celui culturel ou administratif (open data).

La première fonction est celle de connotation. Le signifiant « ouverture » véhicule des valeurs positives relatives à la collaboration entre scientifiques et à une plus grande accessibilité du savoir vis-à-vis la société. Ainsi l'Agence nationale de la recherche parle d'« une science plus transparente, plus rapide et plus universelle »⁴³ tandis que le premier plan national parle de la science ouverte en tant que « progrès scientifique et un progrès de société »⁴⁴. Ce terme également s'oppose à celui de « fermeture », connoté négativement dans ce contexte.

La deuxième fonction est celle de nécessitation, soulignée par la feuille de route du CNRS pour la science ouverte : « Plusieurs raisons scientifiques, citoyennes et socio-économiques rendent aujourd'hui incontournable le développement de la science ouverte »⁴⁵.

La troisième et dernière fonction est celle d'euphémisation. Elle met en lumière que ce mot « ouverture » renvoie à des changements structurels importants du monde scientifique, notamment le fait de réfléchir au système de reconnaissance actuel, et nécessite une acculturation des chercheurs. C'est possible que ce soit cette euphémisation qui soit au cœur de la différence entre adhérence au principe de la science ouverte et mise en pratique.

C'est ainsi que HAL a été interrogé par certains chercheurs sur sa capacité d'ouverture, comme il s'agit d'un outil national et qu'il ne permet pas des échanges et des prises de contact a priori comme ResearchGate d'après un chercheur.

Pour un petit nombre d'enquêtés, c'est le concept de science qui serait pour l'instant fermée qui est interrogé. Ils ne le ressentent pas ainsi, notamment avec le rôle des bibliothèques et par rapport aux scientifiques du secteur privé qui ne peuvent pas être forcément trouvés. Pour l'un, si la science peut être fermée à des personnes extérieures, c'est plus relatif aux codes utilisés par les scientifiques d'où le besoin de médiation.

Dans les autres limites qui ont été soulevés par les enquêtés, est revenue cette idée que la première difficulté n'était pas d'ouvrir la science, mais de faire de la science tout simplement, par manque de temps qui serait pris par la pédagogie et, de manière croissante, par de l'administratif.

Par rapport à l'argument de la pérennité, il convainc peu les enquêtés qui estiment qu'un changement au niveau du système numérique se produit tous les cinq ans en moyenne ou alors qu'un outil soit public ne garantit pas plus sa pérennité que des outils privés, selon le changement de politique. Les ingénieurs interrogés ont ainsi déjà vu des projets apparaître et disparaître en une vingtaine d'années, qu'ils y aient participé comme Europeana Sound⁴⁶ (« On a travaillé 6 mois pour tout mettre

⁴³ ANR, [sans date]. La science ouverte. *anr.fr*. [en ligne]. [Consulté le 3 juillet 2022]. Disponible à l'adresse: <https://anr.fr/fr/lanr/engagements/la-science-ouverte/>

⁴⁴ 1 : *Plan national pour la science ouverte*, 2018. [en ligne]. Ministère de l'Enseignement supérieur, de la recherche et de l'innovation. [Consulté le 29 juin 2022]. Disponible à l'adresse: https://www.ouvrirelascience.fr/wp-content/uploads/2018/08/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf

⁴⁵ CNRS, *op. cit.*

⁴⁶ Projet qui a visé à valoriser des collections sonores à travers l'Europe. Les sons déposés sont toujours consultables mais il n'est pas possible d'en mettre d'autres.

aux normes et exporter, et puis au bout d'un an et demi le projet était fini. Donc on n'en parle plus (rire). C'était un programme de 2 ans européen et puis c'est fini plus personne s'en occupe »⁴⁷), ou non (« Parce qu'il y a eu des outils avant HAL du CNRS qui ont crashé. (...) Entre le minitel et HAL du coup j'ai pas fait. Y en a qui se sont tapés des boulots monstrueux qui n'ont jamais servis à rien »⁴⁸).

4) Un sujet discuté dans l'équipe ?

Par rapport à la question de savoir si c'était un sujet discuté entre collègues, deux réponses différentes se déploient. Premièrement, l'axe de sensibilisation au sujet de la science ouverte et de l'accès ouvert est porté par les documentalistes du laboratoire (Nantes et Grenoble) et est maintenant relayé par des incitations de la part de la direction comme sur HAL : « Oui, en tant que directeur de l'équipe, je mets un point d'honneur à d'abord mettre tout ce que je fais sur HAL en termes de publications ou même d'articles de statut différent. Mais surtout j'insiste énormément, je crois vraiment important que chacun fasse l'effort d'archiver ses données dans HAL, de les rendre publiques, quitte à mettre des embargos pour quelque mois, années s'il le faut »⁴⁹. De plus, des chercheurs ont également maîtrisé l'outil, devenant des « relais »⁵⁰.

Un engagement pour la science ouverte a été pris au niveau du laboratoire AAU, visible sur leur site :

« 1. AAU soutient l'archive ouverte HAL [...]

2. AAU favorise un accès ouvert aux publications scientifiques qui encourage la bibliodiversité et l'innovation [...]

3. AAU accompagne ses membres dans la rédaction de plans de gestion des données »⁵¹

Cet engagement n'est pas repris dans ces termes dans le projet scientifique du laboratoire tel qu'il a été décidé et voté en printemps 2021. Cet écart entre les deux projections a pu se retrouver dans le discours des chercheurs enquêtés. La « science ouverte » est très peu discutée en tant que telle entre chercheurs, il n'y a pas de débat de fond sur le sujet «Après le côté théorique, idéologique derrière, non, j'avoue n'en avoir jamais parlé avec personne »⁵². La seule fois où le sujet semble être abordé est lorsqu'il s'agit d'éléments pratiques : comment valoriser les résultats de la recherche, par exemple « Quand il faut le faire. Quand il faut organiser nos matériaux de données, c'est plus à ce moment-là qu'on en parle »⁵³. L'objectif premier reste de faire de la science.

⁴⁷ E9_ITA

⁴⁸ E8_ITA

⁴⁹ Nicolas Tixier, directeur de l'équipe CRESSON

⁵⁰ E8_ITA

⁵¹ Science ouverte, [sans date]. *Laboratoire AAU*. [en ligne]. [Consulté le 3 juillet 2022]. Disponible à l'adresse: <https://aau.archi.fr/laboratoire-aau/science-ouverte/>

⁵² E7_chercheur

⁵³ E6_chercheur

PARTIE 2 : DEFINIR LA DONNEE DE LA RECHERCHE : DE LA GESTION AU PARTAGE CONTROLE

I. DEFINIR LES DONNEES DE RECHERCHE ET SES ENJEUX

1) Etat des lieux

La définition la plus usuelle de donnée de la recherche, et qui a introduit la notion dans le débat public, est celle de l'Organisation de Coopération et de Développement Economiques (OCDE) en 2007 : « Les données de la recherche sont définies comme des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider des résultats de recherche. Un ensemble de données de recherche constitue une représentation systématique et partielle du sujet faisant l'objet de la recherche »⁵⁴. Cette définition formalise dans le sens qu'elle simplifie, réduit, et crée une forme pour ce qui n'en avait pas, les données variant de discipline à discipline. Dans le cas de l'équipe étudiée, les données mentionnées lors des entretiens sont à la fois des corpus d'entretiens, des enregistrements sonores, des archives, des textes, ou encore des modèles numériques pour n'en citer que quelques-uns.

La notion de « donnée de la recherche » a fait l'objet de nombreuses réflexions, car il s'agit d'un objet complexe, difficile à définir sans se cantonner à une définition qui reconnaît la diversité de l'existant et qui développe seulement quelques exemples. Dans l'article « « Pour commencer, pourriez-vous définir « données de la recherche » ? » Une tentative de réponse »⁵⁵, Joachim Schöpfel, Eric Kergosien et Hélène Prost mettent en avant quatre éléments qui doivent être pris en compte lors de cette tentative de définition : sa nature factuelle, la communauté, la finalité et l'enregistrement. Par cet article, la dimension contextuelle de la donnée est également mise en avant. La donnée n'existerait pas en soi, par elle-même, mais que par reconnaissance comme telle par un chercheur et une communauté, en rapport à un axe de recherche. Aucune liste prédéfinie de la forme que peut prendre une donnée ne peut exister.

De plus, il ne s'agit pas seulement de savoir définir l'objet, il importe également de formaliser et visualiser le parcours de ces données afin de mieux accompagner les chercheurs dans leur gestion des données. L'exemple ci-dessous provient du site sur la science ouverte de l'Université Grenoble Alpes.

⁵⁴ OCDE, 2007. *Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics*. [en ligne]. p.18. [Consulté le 6 juillet 2022]. Disponible à l'adresse: <https://www.oecd.org/fr/sti/inno/38500823.pdf>

⁵⁵ SCHÖPFEL, Joachim, KERGOSIEN, Eric et PROST, Hélène, 2017. `` Pour commencer, pourriez-vous définir « données de la recherche »? `` Une tentative de réponse. In : *Atelier VADOR : Valorisation et Analyse des Données de la Recherche; INFORSID 2017*. [en ligne]. Toulouse, France. mai 2017. pp. 15. [Consulté le 6 juillet 2022]. Disponible à l'adresse: <https://hal.univ-lille.fr/hal-01530937/document>

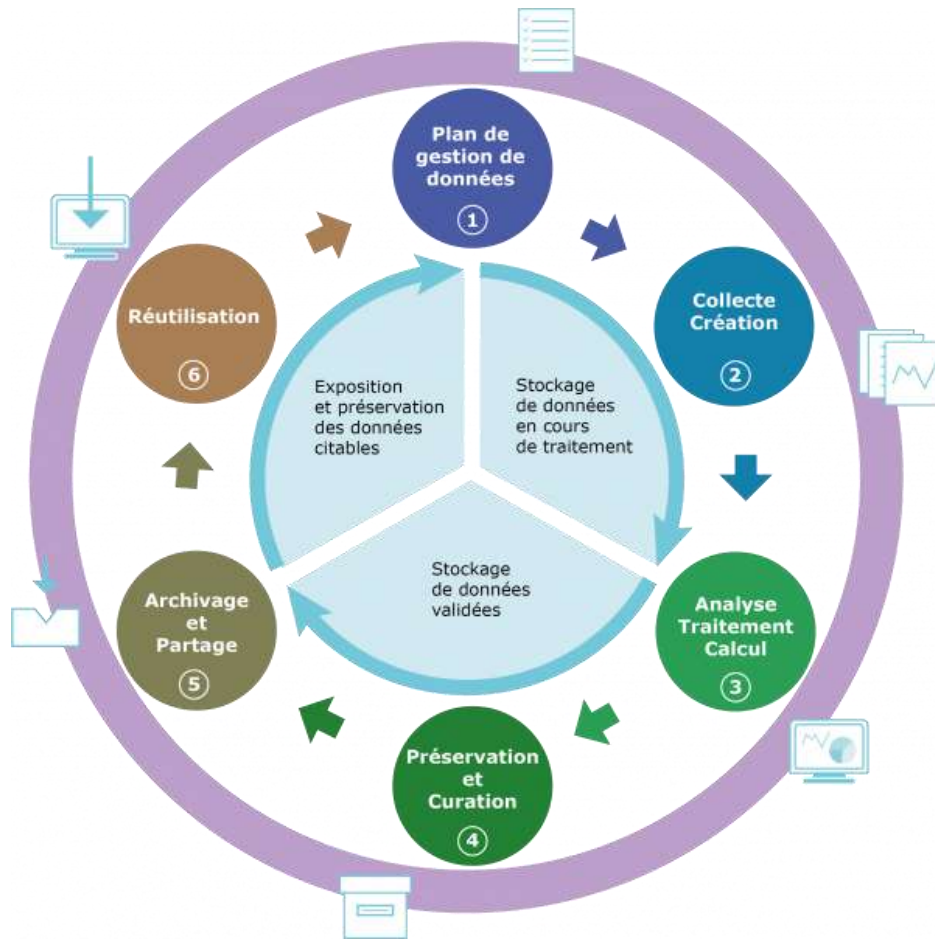


Figure 1 : Cycle de vie des données, schéma de Science ouverte Grenoble, 2022, <https://scienceouverte.univ-grenoble-alpes.fr/donnees/cycle-de-vie-principes-fair/>

Ce parcours est représenté comme un processus, sous la forme d'un cycle de vie qui se répète à l'infini, une représentation assez commune et partagée par les services d'aide à la gestion de données. Plusieurs étapes sont clairement distinguées dans un ordre précis : la collecte des données, leur analyse, leur stockage, leur partage et leur conservation, incluant également une étape de réutilisation.

Un autre standard régit également les enjeux de stockage et de partage des données : le principe FAIR (Findable, Accessible, Interoperable, Reproducible). Les données doivent pouvoir être trouvées facilement, téléchargeables, utilisables peu importe l'environnement informatique, et réutilisables dans d'autres fins de recherche que celles dans lesquelles elles ont été produites. Ce standard répondrait à l'objectif d'ouverture des données, en tant que « « lex technica », c'est-à-dire comme un ensemble de règles techniques à fort pouvoir normatif de fait »⁵⁶. C'est par le biais de ce standard que nous retrouvons la dimension normative de la science ouverte, en lien avec les enjeux numériques. De nouveau, ce serait pour permettre une lecture par les machines.

⁵⁶ AMIEL, Philippe, FRONTINI, Francesca, LACOUR, Pierre-Yves et ROBIN, Agnès, 2020. Pratiques de gestion des données de la recherche : une nécessaire acculturation des chercheurs aux enjeux de la science ouverte ? *Cahiers Droit, Sciences & Technologies*. [en ligne]. 11 mai 2020. N° 10, pp. 147-168. [Consulté le 31 octobre 2021]. DOI [10.4000/cdst.2061](https://doi.org/10.4000/cdst.2061).

Le cycle de vie de l'Université de Grenoble Alpes ne le met pas en valeur mais un autre enjeu est également primordial dans le stockage des données et leur ouverture et qui a mené à la maxime « aussi ouvert que possible, aussi fermé que nécessaire », repris par plusieurs institutions comme le CNRS et l'ANR. Cet enjeu concerne plus précisément la sécurité des données personnelles, qui sont récoltées et utilisées à des fins de recherche.

Une donnée personnelle est une donnée qui permet d'identifier directement ou indirectement une personne physique (un numéro de téléphone, une date de naissance par exemple). Certaines de ces données sont jugées sensibles lorsqu'elles peuvent donner lieu à une discrimination comme une opinion politique ou une orientation sexuelle.

Un cadre légal vise aujourd'hui à protéger les données personnelles. Cela passe par un contrôle des systèmes de traitement de données personnelles qui récoltent ce type de données, les traitent et les stockent. Ce cadre légal est posé par le règlement général sur la protection des données (RGPD), voté à l'échelle de l'Union Européenne, et mis en vigueur le 25 mai 2018 (avec une disposition nationale qui a modifié la loi de 1978 relative à l'informatique, aux fichiers et aux libertés).

Les chercheurs concernés par la récolte de données personnelles doivent également respecter ce cadre. Ceci passe par une déclaration à la CNIL sur le traitement effectué (qui passe pour l'équipe par une inscription au registre de la Déléguée à la Protection des Données DPD du CNRS) et le déploiement de formulaires d'autorisation (enregistrement de la voix et de l'image par exemple) et de notes d'information pour les enquêtés, expliquant le projet, quelle utilisation sera faite de leurs données, pendant combien de temps et pourquoi. Lorsqu'il s'agit de données sensibles, un formulaire de consentement est mis en place. Sans ces autorisations, le partage et l'ouverture des données ne sont pas possibles, ni même dans le cadre des publications et des communications scientifiques.

2) La vision des enquêtés sur la notion de donnée

A la question sur l'utilisation du terme de « donnée de la recherche », les réponses étaient contrastées. Nous pouvons noter une progression dans l'utilisation de ce terme, une possible acculturation, qui passe notamment par les obligations de l'Agence Nationale de la Recherche à produire des plans de gestion de données « je l'utilise de force [le terme de donnée], dans le sens dans le cadre de l'ANR il y a le plan de gestion de données »⁵⁷. A l'instar de l'image du point d'eau dans un désert, utilisé par Jens Klump⁵⁸, l'Agence Nationale de la Recherche, en intégrant dans son financement de projets de recherche des enjeux de la science ouverte, a mis en place un point de passe obligé pour les chercheurs, qui les sensibilisent à ces enjeux.

Cela inclut des chercheurs qui ne sont pas concernés directement pour l'instant par cette obligation et qui commencent à l'intégrer et à le prendre en compte dans le cas où un projet de recherche se monterait : « quand on a une réponse à faire, on est dans une course, et je préfère voir en amont comment ça se prépare plutôt qu'attendre le jour j, que ce soit obligé et qu'il faudra tout que je découvre. C'est pour ça que ça m'intéresse pas mal parce que je sais que c'est une étape. »⁵⁹

⁵⁷ E2_chercheur

⁵⁸ KLUMP, Jens, 2017. Data as social capital and the gift culture in research. *Data Science Journal*. [en ligne]. 2017. Vol. 16, pp. 1-8. [Consulté le 31 octobre 2021]. DOI [10.5334/dsj-2017-014](https://doi.org/10.5334/dsj-2017-014).

⁵⁹ E6_chercheur

« L'ANR le demande, je comprends pourquoi il demande à peu près donc si tu veux faire un projet ANR il faut le faire. C'est comme je dois faire mes rapports annuels sur mes activités, c'est une obligation qui vient avec mon travail et si je dois garder mon poste, il faut que je le fasse »⁶⁰.

Cependant, les pratiques seraient encore pour l'instant dans la désignation directe de l'objet concerné : des entretiens ou des vidéos par exemple, et non des données de la recherche « je parle beaucoup plus de la donnée en spécifique, j'analyse un entretien, j'analyse une observation que j'ai faite, des notes que j'ai prises. Mais quand je parle plus abstraitement j'utilise quand même le terme de donnée, oui. Tout dépend de l'échelle et du contexte... »⁶¹. Les termes de matériaux et de corpus empirique sont également revenus lorsqu'il s'agissait d'un ensemble de données. Ce terme de donnée n'apparaît qu'à une certaine échelle et dans un certain contexte notamment interdisciplinaire, quand il devient nécessaire pour communiquer.

Et à l'instar d'autres chercheurs en sciences humaines et sociales, le mot même de donnée est questionné par certains dans l'équipe qui relèvent le fait qu'il s'attacherait plus à une vision quantitative de la recherche « pour moi il est beaucoup plus attaché au quantitatif »⁶², association également vue par Violaine Rebouillat dans le cadre de sa thèse⁶³ et dans le rapport du projet « Partage et protection des données qualitatives à l'ère du numérique : expériences, enjeux, stratégies »⁶⁴. Ou alors qu'il cache une partie du travail de la prise de cette donnée qui est construite « C'est moi qui la prends, il ne m'est pas donné comme ça donc ce n'est pas vrai. On construit des données et après on appelle ça des données. Donc une donnée n'est pas donnée, elle est saisie. Ça change l'objet. En observant quelque chose, je le modifie »⁶⁵. Cette dernière citation permet de revenir sur l'étymologie du mot « donnée » qui vient du latin « Dare » qui signifie donnée. Le mot data, utilisé en anglais (et parfois en français, avec l'open data ou le big data), signifie « ce qui a été donné » en latin⁶⁶.

Il y a un aspect que je n'avais pas imaginé au premier abord car je l'avais considéré, à tort, comme acquis : la question de la sécurité des données et plus précisément du RGPD. Le RGPD semble également jouer un rôle de sensibilisation sur la gestion des données et cristallise des craintes et des critiques.

Il a été très rarement cité directement, sans que ce soit possible de savoir si c'était par manque de connaissance ou pour une autre raison. Il était souvent désigné indirectement par le « papier ». Ce papier représente les différents formulaires qui s'instaurent dans la recherche sur la gestion des données personnelles comme

⁶⁰ E3_chercheur

⁶¹ E3_chercheur

⁶² E3_chercheur

⁶³ REBOUILLAT, Violaine, 2019. *Ouverture des données de la recherche: de la vision politique aux pratiques des chercheurs*.

⁶⁴ REVELIN, Florence, LEVAIN, Alix, MIGNON, Morgane, NOEL, Marianne, QUEFFELEC, Betty, RAUX, Pascal et SQUIVIDANT, Hervé, 2021. *L'ouverture des matériaux de recherche ethnographiques en question*. [en ligne]. Rapport de recherche. CNRS. [Consulté le 16 novembre 2021]. Rapport d'enquête du projet "Partage et protection des données qualitatives à l'ère du numérique : expériences, enjeux, stratégies". Disponible à l'adresse: <https://hal.archives-ouvertes.fr/hal-03238067/document>

⁶⁵ E4_chercheur

⁶⁶ REBOUILLAT, Violaine, *op. cit.*

l'autorisation d'enregistrement et de diffusion de la voix et de l'image et le consentement lorsqu'il s'agit de données sensibles.

II. LA GESTION DES DONNEES DE LA RECHERCHE DES ENQUETES

Une grande part de l'entretien portait sur les pratiques actuelles de gestion de données de la recherche. Chaque chercheur enquêté possédait un système pour cette gestion, notamment à travers la structuration des dossiers numériques. Des similitudes reviennent : une première distinction à faire entre pédagogie et recherche, puis dans recherche, une structuration par projet ou par événement (citation). Deux options se dessinaient ensuite. Pour une semaine sur le terrain par exemple, soit le tri se faisait de manière chronologique pour les données de la recherche (jour 1, jour 2...) soit de manière plus rare par type de données (regroupement des entretiens, des sons...). Deux chercheurs ont également détaillé des pratiques de sauvegarde mais cela n'a pas été le cas de tous.

Par rapport à cette question-là, ce qui revenait était le vocabulaire du manque et du « mauvais élève » (« j'ai un dispositif totalement médiocre mais qui marche très bien »⁶⁷, « Et puis j'ai une copie sur Google Drive, le truc qu'il ne faut pas faire [...] Donc là je ne suis pas un très bon élève »⁶⁸) avec une perception de ce qui était encore à faire pour améliorer cette gestion : un jugement de valeur ressort. Ce dernier est souvent en lien avec la possibilité de partager ces données avec des personnes tierces (« Ce qui fait qu'il n'y a que moi qui peut m'y repérer »⁶⁹, « c'est organisé selon une organisation propre à moi »⁷⁰). Les systèmes mis en place sont personnels et permettent au chercheur de se retrouver par des techniques propres « « Jusqu'à présent c'était plutôt on bidouille avec notre ordinateur, c'est assez intuitif et informel. Chacun a je pense sa technique de classement, de mémoire »⁷¹, mais ce ne serait pas partageables d'après eux et c'est en cela que le système trouve des limites.

A travers ces entretiens, ce qui se dessine est un écart entre une certaine connaissance des « bonnes pratiques », avec les pratiques réelles. Par exemple, un chercheur forme des étudiants sur le sujet mais se retrouve quand même en difficulté lorsqu'il s'agit de ses données « je suis conscient, je sais, je sais comment faire et je forme mes étudiants. Maintenant de l'autre côté dans ma vie de chercheur, mes données, je n'arrive pas à les faire comme ça tout le temps »⁷². La systématisation paraît difficile.

Dans le cadre de ce mémoire, il ne sera pas possible de définir précisément les raisons de cet écart, mais certaines causes sont revenues dans les entretiens et qui sont intéressantes à relever : la duplication des outils, la « masse » des données, les rapports multiples entre temps et donnée.

⁶⁷ E5_chercheur

⁶⁸ E7_chercheur

⁶⁹ E6_chercheur

⁷⁰ E1_chercheur

⁷¹ E2_chercheur

⁷² E1_chercheur

Pour la première cause, il s'agit des solutions de stockage déployées par les différentes institutions : les enseignants-chercheurs sont souvent à la croisée de plusieurs institutions (laboratoire, CNRS, école par exemple) et engagées dans des projets de recherche qui demandent un partage des données. La conséquence est le déploiement d'outils différents chacun dédié à un objectif précis « C'est un peu entre mon organisation perso, l'organisation des autres et puis qui est le chapeau institutionnel. Je n'ai pas la même stratégie. Parce que je ne vais pas utiliser le drive de l'école d'archi pour aller mettre mes trucs pro. Ou le drive de l'ENSI à Paris pour aller mettre les trucs de l'école d'archi. J'essaye d'être un peu clean un peu là-dessus et du coup ça rajoute du bordel dans le bordel »⁷³. Un autre chercheur de l'équipe estimait à une quinzaine le nombre de Dropbox en activité, causant parfois des problèmes techniques par rapport aux mots de passe et aux mails utilisés. Un éparpillement des données se fait alors même que de multiples structurations existent.

Cette idée de « masse de données » vient d'un discours d'un chercheur (« Et... en même temps tu peux avoir de plus en plus de données donc ... je ne sais pas si ça facilite [le numérique]. La masse. »⁷⁴) et est revenu dans plusieurs entretiens. Elle est en lien avec la facilité technologique aujourd'hui de produire des données dans leur recherche : photographies, prises de son... Un téléphone peut parfois suffire. La différence se creuse avec les premières années de l'équipe de recherche qui étudiait le son avec difficulté jonglant entre bandes magnétiques et coupes manuelles pour faire un montage (l'équipe CRESSON a eu 40 ans en 2019). Même si la gestion est censée être plus simple, le bénéfice est limité par la démultiplication des données.

Ce n'est pas la question du stockage qui semble se poser aujourd'hui parmi les enquêtés, qui est relativement évacuée (elle n'est pas problématique pour l'instant), mais celle de l'intérêt et de la légitimité des données produites parmi cette masse. Ces questions seront également vues à nouveau lors de la troisième partie lorsque nous discuterons du projet Cartophonies, porté par des membres de l'équipe. Un ingénieur d'études en parle plus particulièrement en relation avec un événement de la recherche, mettant en avant l'inutilité et l'aspect chronophage du tri a posteriori : « Sur le colloque à Volos, j'ai fait le tri un jour et ça m'a stupéfié. 5000 photos dont à peu près 2/3 étaient les mêmes. J'en ai gardé 500, ce qui est beaucoup. Et tu vois, c'était des gigas et des gigas, ok y a de la place mémoire, on les met et puis on appuie sur le bouton, ça fait des rafales [...] Il y aussi le côté politique écologique, si on veut faire fondre la banquise plus vite ok, on a qu'à faire des milliards de photos et des enregistrements tous les matins. On stocke ça sur des disques durs qui seront refroidis super [...] Si ce n'est pas ça l'objectif, on peut commencer à se dire, bon pourquoi je ferai une photo ici maintenant, pourquoi j'en ferai trois, pourquoi j'en enregistrerai une heure »⁷⁵.

L'enjeu derrière pour trois des enquêtés est écologique, puisque stocker ces données demande de l'énergie, ce qui n'est pas forcément un argument très développé dans les discours habituels sur la gestion des données.

En filigrane, ce sont donc les questions du tri et de la suppression des fichiers qui se posent, en plus de celles de production de données sur le terrain. Or, le tri des données de la recherche se fait habituellement en cours du projet, au fil des analyses

⁷³ E7_chercheur

⁷⁴ E4_chercheur

⁷⁵ E9_ITA

(« Plus ça va dans l'analyse, plus je récupère des documents qui me paraissent particulièrement intéressants et que je vais donc recopier ou transformer pour le mettre dans un autre dossier qui sera déjà pré-analysé, disons. Il y a aussi l'étape dans le déroulement de la recherche entre la production de données brutes sur le terrain et la phase d'analyse »⁷⁶). Même lors d'une multiplication du nombre de versions d'un même fichier jusqu'à ce qu'il convienne, deux enquêtés ne savaient pas quoi faire des fichiers intermédiaires.

Aucun enquêté ne semble effectuer un tri à la fin d'un projet de recherche, par manque de temps et nécessité de s'occuper d'autres projets ou de pédagogie. Le dossier conserve donc toutes les étapes d'analyse et de production de données. Tout est conservé tout le temps dans un principe d'accumulation (« Mais après ça s'entasse dans les disques durs. Par contre j'ai du mal à m'en séparer. Je ne supprime pas. »⁷⁷, « je travaille par addition, je ne travaille pas par remplacement [...] Tout s'accumule. Et je sauvegarde l'ensemble de mes dossiers, pas que ce qui est important pour la recherche »⁷⁸).

Ce n'est pas le seul moment où cette question du temps est ressortie. Même s'il est devenu plus simple de produire une donnée, le temps d'analyse de la donnée lui ne se modifie pas ce qui rend complexe le traitement des données et l'analyse (« La durée aussi entre le temps où tu obtiens tes choses et le temps où tu les traites. Ça c'est aussi un gros problème je trouve. On fait une chose puis on doit faire autre chose parce qu'on enseigne, parce qu'on fait ci, parce qu'on fait ça... puis après tu reviens dans ton matériau, il y a eu un délai »⁷⁹), ce qui aura même un impact sur le choix de réutiliser de données produites par d'autres : « Déjà il faut dire qu'on a déjà du mal à trouver le temps d'exploiter toutes les données qu'on produit nous même ! Donc il y a aussi une question très très concrète de disponibilité, tu vois ? [...] Les données qu'on produit, elles prennent énormément de temps, quand tu produis une vidéo ou un son, quand tu veux réécouter, ça prend beaucoup de temps. Et comme on a une telle facilité à produire des choses [...] qu'on a tendance à trop multiplier les choses, à être submergé par plein de données dont on exploite un dixième. Se dire je vais aller en chercher d'autres encore... ce n'est pas la priorité disons ! »⁸⁰. Pris dans des multiples projets et l'enseignement, ces étapes de traitement et d'analyse se font sur un temps relativement long.

De plus, certains enquêtés ont également exprimé des difficultés à se retrouver dans les données qu'ils produisent eux-mêmes à cause de la fracture entre la production d'une donnée produite dans un contexte précis et dont le nommage fait sens dans celui-ci, et le fait de la conserver sur un temps plus long cette donnée qui perdra alors son contexte de production (« Dans un an, je n'ai aucune idée de ce que c'est, même pas un an, dans une semaine »⁸¹). Ainsi un enquêté raconte comment il lui faut plusieurs heures pour retrouver la logique de l'organisation de ses données alors même qu'il en est le créateur : « Mon disque dur, même moi, si je reviens à des données qui datent de trois ans, je dois passer une matinée le temps pour me retrouver à la fin. [...] Je vais faire une nomination conjoncturelle, qui est pertinente

⁷⁶ E2_chercheur

⁷⁷ E6_chercheur

⁷⁸ E5_chercheur

⁷⁹ E4_chercheur

⁸⁰ E2_chercheur

⁸¹ E5_chercheur

à ce moment-là par rapport à ce que je fais. Il faut retrouver cette logique pour comprendre ». Il est possible de s'y retrouver mais après un effort. Ce qui ressort, c'est finalement de nouveau le fait que la donnée ne prend valeur qu'avec son contexte.

III. LE PARTAGE DES DONNEES

La question s'est souvent posée ainsi : pourquoi les chercheurs ne partagent-ils pas leurs données ?

Les réponses sont variées et nous pouvons distinguer deux différentes catégories de levier et de motivation d'après les articles scientifiques lus, avant de s'intéresser aux réponses des enquêtés.

1) Fonctionnement du monde scientifique

Tout d'abord il s'agit de souligner que le partage des données en tant que pratique culturelle est déjà vrai pour certaines disciplines scientifiques tandis que d'autres le font moins. Une dynamique de collaboration dans la recherche semblerait inciter à un partage des données⁸².

Le deuxième point porte sur l'écosystème du monde scientifique qui, d'après Jens Klump⁸³, se repose sur la « reputation economy » et le « gift culture », c'est-à-dire que publier un article reconnu par la suite permet d'obtenir des financements qui mènent à faire d'autres recherches et ainsi de suite. Les données ont également une valeur dans ce système. Or comme cette économie se base sur le potentiel de reconnaissance, celui-ci est fragile lorsque l'utilisateur d'un entrepôt de donnée est anonyme et que l'entrepôt permet un trajet de la donnée dans le temps et l'espace (utilisation dans un autre pays, dans 3 ans...). L'utilisateur peut alors être un chercheur ou ne pas l'être.

Ce constat est également partagé par l'article « Pratiques de gestion des données de la recherche : une nécessaire acculturation des chercheurs aux enjeux de la science ouverte ? » qui conclut ainsi : « La mise en œuvre de la politique de science ouverte suppose donc à la fois une prise de conscience des enjeux par les chercheurs et un changement de paradigme dans le fonctionnement de la recherche scientifique publique »⁸⁴, puisqu'il faudrait valoriser la pratique du partage de données ce qui changerait le système d'évaluation des chercheurs. Cette réflexion est maintenant menée comme le montre la note du 23 mars 2022 du Secrétariat général du Conseil, pour le groupe de travail de recherche du 31/03/2022, ayant pour objet « projet de conclusions du Conseil sur l'accélération de la mise en œuvre et de l'impact des politiques et des pratiques de la science ouverte ». Cette note par exemple « ESTIME que l'évolution des systèmes d'évaluation de la recherche en Europe devrait être fondée sur les principes suivants [...] : b. reconnaître tous les résultats de la recherche et innovation, y compris les ensembles de données, les logiciels, les codes, les brevets etc. et pas seulement les publications ; »⁸⁵.

⁸² AMIEL, Philippe, FRONTINI, Francesca, LACOUR, Pierre-Yves, et ROBIN Agnès, *op. cit.*

⁸³ KLUMP, Jens, *op. cit.*

⁸⁴ AMIEL, Philippe, FRONTINI, Francesca, LACOUR, Pierre-Yves, et ROBIN Agnès, *op. cit.*

⁸⁵ SECRÉTARIAT GÉNÉRAL DU CONSEIL, 2022. *Projet de conclusions du Conseil sur l'accélération de la mise en œuvre et de l'impact des politiques et des pratiques de la science ouverte* [en ligne]. Bruxelles : Conseil de l'Union

La question de la reconnaissance semble alors centrale, ainsi que les pratiques de collaboration préexistantes (à une discipline par exemple ou un projet comme Migreval⁸⁶). Ils soulignent le besoin d'avoir un changement structurel.

2) Des choix personnels

D'autres leviers et freins pour partager les données ont été relevés par le biais de partage d'expérience, de questionnaires administrés aux chercheurs. Ils sont directement liés à la partie précédente : peur de perdre le contrôle, partage encouragé lorsqu'il est valorisé économiquement ou faisant parti du contrat, le lien avec la collaboration, sentiment d'autorialité⁸⁷... La peur de perte de contrôle, alors qu'un lien de confiance se crée entre l'enquêteur et l'enquêté, a pu revenir plusieurs fois.

L'article de Violaine Rebouillat⁸⁸ articule ces deux catégories motivations personnelles et caractéristiques du monde scientifique en identifiant différents types de valeurs accordées aux données, à partir d'entretiens : la valeur pour soi, la valeur d'usage, la valeur d'échange. Ceci permet de regrouper les différents aspects vus auparavant : l'inscription des données dans une recherche scientifique, dans l'environnement du monde scientifique et relation entre données et producteur.

3) Le point de vue des chercheurs enquêtés

Un partage existant, partiel et local

Très rapidement est apparue la nécessité de définir ce que signifiait partager des données, comme un rappel qu'il n'y a pas d'un côté la science ouverte et celle fermée⁸⁹. Faut-il que cela rentre dans le cadre de la mise à disposition d'un ensemble de données sur un entrepôt de données certifié pour que ce soit science ouverte ? Car les chercheurs enquêtés partagent déjà des données de la recherche. Certains dans le cadre d'un projet de recherche en équipe (« Sur ShareDocs, on a un dossier commun où chacun met les données produites. Je peux aller voir ce qui a été produit par d'autres et d'autres vont voir ce que j'ai fait et du coup, ça alimente... ça permet de travailler collectivement, ça permet d'enrichir les données, il y en a plus et elles sont plus diverses, et puis ça permet aussi de travailler ensemble. Plus ça va, et plus on le fait ça »⁹⁰), d'autres le font de manière ponctuelle sur demande et cela passe par des relations interpersonnelles. Ces dernières ont parfois permis la réalisation d'une recherche qui n'aurait pas été possible par les canaux habituels : « J'avais un post-doc d'un an, pendant 3 mois, je ne voyais rien arriver et j'étais hyper inquiète de ça parce que j'avais rien sous la dent. Après je suis passée par tout mon réseau d'archis, d'amis, qui sont éparpillés dans tous les services et qui m'ont donné sous forme d'une clé USB au détour d'un repas de midi toutes les informations. Et ça, si je n'avais pas eu ce réseau-là, le transect qu'on expose un peu partout, toutes les

Européenne. [Consulté le 23 mars 2022]. Disponible à l'adresse: <https://data.consilium.europa.eu/doc/document/ST-7390-2022-INIT/fr/pdf>

⁸⁶ Plateforme collaborative sur « L'évaluation biographique des politiques par les migrants en Europe. » (<https://migreval.hypotheses.org/>)

⁸⁷ AMIEL, Philippe, FRONTINI, Francesca, LACOUR, Pierre-Yves, et ROBIN Agnès, *op. cit.*

⁸⁸ REBOUILLAT, Violaine, 2021. Le partage des données vu par les chercheurs : une approche par la valeur.

⁸⁹ REBOUILLAT, Violaine, 2019. *Ouverture des données de la recherche: de la vision politique aux pratiques des chercheurs.*

⁹⁰ E2_chercheur

coupes que j'ai pu faire là-dessus, elles n'auraient jamais existé »⁹¹. Ils sont à la fois prêteurs et demandeurs.

Cependant, dans le cas où ce n'est pas un projet en équipe, il s'agit généralement d'un fragment d'un ensemble de données de la recherche d'un projet qui fait l'objet d'un échange entre deux chercheurs dans un cadre informel. La raison invoquée par un chercheur est la faisabilité d'envoyer un jeu de données entier (« Mais des données partielles. J'ai besoin de ça, j'ai vu cette photo, rappelle-toi est-ce que tu peux me la donner ? Aussi parce que confier le jeu de données, on a tous plusieurs gigas de données quand on part »⁹²). Ce partage partiel concerne par exemple les sons, produits dans le cadre d'une recherche, qui sont mises à disposition sur le Web à travers deux projets portés par l'équipe, CRESSOUND et Cartophonies. Cela ne concerne pas toutes les données produites dans le cadre d'une recherche.

Ce côté partiel se retrouve également dans la réutilisation qui est faite de leurs données produites au cours du temps. Les enquêtés ont exprimé le fait d'y revenir pour des objectifs pédagogiques ou d'illustration dans le cadre de conférences ou d'articles, mais pas réétudier le corpus en son intégralité. Cela dépend également si le chercheur s'occupe toujours de la même thématique de recherche « C'est un changement thématique derrière et donc ça m'intéresse moins, je suis aussi intellectuellement trop détaché, ça me prendrait trop de temps de relire ça et de revoir ça. »⁹³, « Depuis trente ans, on travaille sur la question des ambiances et donc quand on a besoin à un certain moment d'expliquer ou de montrer des choses comme c'est un domaine compliqué, on a besoin de s'appuyer sur des choses assez factuelles et du coup, je peux revisiter certains travaux »⁹⁴.

Les « frictions » des échanges de données

Ce partage de données ne se fait pas sans difficulté entre chercheurs. Parmi les enquêtés plusieurs ont exprimé des difficultés à comprendre la donnée produite par un autre et la nécessité d'avoir de la documentation ou des échanges avec le producteur de la donnée. Une donnée qui aurait été accessible directement sans moyen d'échange ne serait pas forcément intéressante ni utilisable : « C'est-à-dire qu'il faut essayer de comprendre comment lui a produit ces matériaux, dans quel cadre et pourquoi. Il faut toujours faire un pas de côté, que ce soit du chiffre, de la mesure, du sonore... ce n'est pas ça le problème, c'est comprendre le cheminement et pourquoi il me dit que ça c'est important. Et là oui il faut faire un certain effort de comprendre sa posture, pourquoi il me présente ça, quel est l'intérêt, comment j'arrive à le saisir [...] ça passe par beaucoup d'échanges sur différents médiums. Mais c'est sûr que, par rapport à la science ouverte, si on m'avait donné les plans comme ça et on m'avait dit voilà, ça c'est les plans d'énergie de Grenoble, j'aurais dit ouais... c'est sympa mais j'en fais quoi, si je n'ai pas la personne pour m'aider à lire ? »⁹⁵.

D'autres ont exprimé ces mêmes difficultés à appréhender les données ouvertes par les métropoles « Dernier data que j'ai consulté c'était il y a quelques mois et

⁹¹ E6_chercheur

⁹² E5_chercheur

⁹³ E3_chercheur

⁹⁴ E2_chercheur

⁹⁵ E6_chercheur

c'était une machine à gaz. Ça peut être intéressant mais il faut accepter de passer déjà une semaine à savoir comment ça marche et extraire les informations »⁹⁶. S'approprier des données demande des compétences et des efforts.

L'idée remise en cause derrière par les enquêtés est celle de donnée brute, compréhensible facilement car se rapprochant d'une observation factuelle. Une nécessité se fait entendre : celle de comprendre le cheminement de la personne, de l'intérêt d'une donnée. Le besoin de documentation et de métadonnées est clairement énoncé par plusieurs enquêtés « il faut donner quelques éléments, des métadonnées comme dit Françoise. S'il n'y en a pas, on ne sait pas quoi en faire, s'il n'y a pas des données pour expliquer pourquoi il y a cette donnée »⁹⁷.

Ce sont un ensemble de « frictions », notion d'Edwards⁹⁸ mise en avant dans la thèse de Violaine Rebouillat, causés par la réutilisation de données existantes comme « d'importantes transformations doivent être opérées pour rendre les données intelligibles à leurs réutilisateurs »⁹⁹. Ces difficultés sont une des raisons utilisées par un enquêté pour expliquer le fait que le partage se fait plutôt dans des cercles proches : « Mais après pour rendre explicite des données à une autre équipe avec qui je ne travaille jamais... je crois pas, non ça ne m'est pas arrivé je pense. Est-ce que c'est aussi dû à ce que je disais tout à l'heure ; nos données sont tellement liées aux situation et contexte, c'est assez difficile de les rendre communicables, accessibles oui, mais exploitables après pour quelqu'un qui ne connaît pas le terrain, qui ne connaît pas les conditions d'enquête etc. et qui se retrouve avec des enregistrements, des vidéos ou même d'entretiens sans connaître le contexte, c'est pas si facile que ça »¹⁰⁰.

Un intérêt pour les méthodes et la représentation des données

Ce qui est revenu chez plusieurs enquêtés est le désintérêt à analyser un corpus de données produites et déjà analysées par d'autres. Néanmoins, il existe tout de même un intérêt pour ces données si elles renseignent sur le processus particulier de production. L'intérêt porte plus sur la méthode, qui pourrait être appliquée à d'autres terrains, que la donnée elle-même jugée trop contextuelle « S'il a des données qui montrent l'enquête en train de se faire, quel type de protocole, quel type de technique, de dispositif d'enquête, c'est là où j'aurais le plus d'envie. Exploiter des données produites par d'autres, non je n'ai jamais eu cette envie vraiment »¹⁰¹, « L'autre jour j'ai voulu me replonger dans les parcours commentés d'une thèse faite au sein du CRESSON pourtant par A., mais je n'ai pas réussi parce qu'elle n'a retranscrit que la parole de la personne, elle n'a pas mis les questions. C'était insaisissable [...] la recherche avec le CSTB et AAU porte sur Paris Rive Gauche. A. avait travaillé sur Paris Rive Gauche et j'avais voulu voir comment elle avait fait ses parcours commentés et je n'ai pas réussi à m'en saisir »¹⁰².

⁹⁶ E1_chercheur

⁹⁷ E4_chercheur

⁹⁸ EDWARDS, Paul N., MAYERNIK, Matthew S., BATCHELLER, Archer L., BOWKER, Geoffrey C. et BORGMAN, Christine L., 2011. Science friction: Data, metadata, and collaboration. *Social Studies of Science*. [en ligne]. Vol. 41, n° 5, pp. 667-690. [Consulté le 6 juillet 2022]. DOI [10.1177/0306312711413314](https://doi.org/10.1177/0306312711413314).

⁹⁹ REBOUILLAT, Violaine, *op. cit.*, p. 258

¹⁰⁰ E2_chercheur

¹⁰¹ E2_chercheur

¹⁰² E6_chercheur

Cet intérêt pour la méthode s'exprime également par l'importance donnée à la représentation de la donnée produite, aussi importante que sa méthode de production : « Après je ne sais pas sous quelle forme ces données vont être représentées [...] Est-ce que c'est un tableau de chiffres ou est-ce que c'est une image sur laquelle je vais voir des lignes qui bougent ? C'est aussi important que la donnée elle-même. Elle est construite 2 fois, la manière dont je la saisis, et la manière dont je la représente. Au moins »¹⁰³.

Une étape pourrait alors manquer dans le cycle de vie de la donnée tel vu au début de cette partie. Il faudrait y inclure la mise en forme de données qui parfois mène à la création de nouvelles données qui n'ont pas exactement le même statut que les premières, ou comment l'analyse produit des données également. Le cycle de vie, reproduit dans l'article « `` Pour commencer, pourriez-vous définir 'données de la recherche'? " Une tentative de réponse »¹⁰⁴, prend en compte ces différentes catégories des données de la recherche et provient du mémoire de Marilou Pain : « les données préliminaires ou données préparatoires (exclues dans un contexte de diffusion) ; les données brutes (données acquises lors du processus de recherche et déjà potentiellement traitées) ; et enfin les données traitées et analysées : c'est-à-dire ayant subi une transformation telle qu'il n'est plus possible d'accéder aux données brutes, un graphique par exemple »¹⁰⁵. C'est également une fonction discutée par Violaine Rebouillat : « Les données sont collectées par les chercheurs dans la perspective d'être utilisées comme preuve d'une théorie scientifique au sujet d'un phénomène particulier. Afin d'assurer cette fonction auprès des pairs, elles sont rendues « portables », c'est-à-dire présentées sous une forme adéquate à leur communication »¹⁰⁶.

Il semble alors que, dans un projet de recherche, différentes couches de données existent. Dans ces couches, certaines pourraient faire l'objet d'un partage et d'autres non selon leur statut et le contexte de production. Une dimension de présentabilité apparaît également que nous reverrons dans la dernière partie : « Surtout que c'était un peu limite mes entretiens donc oui. Après si maintenant je fais quelque chose de plus organisé cela ne me dérange pas. Comme ODE, les petits carnets qu'on a faits, bah là ça a été plus que balisé, donc là ça ne me dérange pas que ce soit en accès libre sur HAL etc. ça dépend des matériaux demandés et comment j'ai pu les produire », « il y a toute une ressource disponible, mais il y a des pans de la recherche, des couches de la donnée ou du travail qui ne sont pas diffusés, donnés que j'aurais pu mettre [...] Mais ça je ne pourrai pas le diffuser quoi [un type de données de sa recherche]... par éthique vis-à-vis des participants et puis parce que je ne l'ai pas anticipé, ce n'est pas hyper propre ». Il y a cette idée que cela doit être anticipé en début de recherche et que cela peut modifier les méthodes de récolte ou de création de données.

En prenant en compte tous les éléments décrits précédemment, à la question de savoir s'ils seraient d'accord de partager leurs données dans des entrepôts de données, les réponses ont été multiples : oui dès que le temps d'analyse est passé,

¹⁰³ E4_chercheur

¹⁰⁴ SCHÖPFEL, Joachim, KERGOSIEN, Eric et PROST, Hélène, *op. cit.*

¹⁰⁵ PAIN, Marilou, 2016. *Les données de la recherche et leurs entrepôts, de la documentation à la réutilisation : étude de cas pour l'archive HAL*. [en ligne]. Enssib. P.18. [Consulté le 6 juillet 2022]. Disponible à l'adresse: https://memsic.ccsd.cnrs.fr/mem_01374509

¹⁰⁶ REBOUILLAT, Violaine, *op. cit.*, p.46

Partie 2 : Définir la donnée de la recherche : de la gestion au partage contrôlé

non car il n'y a pas le temps, oui si connaissance du cadre de recherche, dans l'idée
oui mais dans la pratique comment faire, oui si on me contacte...

PARTIE 3 : DES CHOIX PERSONNELS A UNE ORGANISATION COLLECTIVE

I. UN « NOUS » QUI APPARAÎT

Dans la partie de mon mémoire dédiée à l'état de l'art sur la question du partage des données, j'ai intitulé une partie « des choix personnels ». Or, il me semble que les entretiens révèlent : ce n'est pas entièrement personnel et ce n'est pas non plus que relatif à la discipline ou au fonctionnement du monde scientifique et de son système de reconnaissance. Le cadre de l'équipe semble jouer un rôle dans les questions de gestion des données et de partage, avec des pratiques qui s'instaurent en communauté. Ceci peut également s'expliquer par le lien étroit entre protocoles partagés et gestion des données.

Ceci reprend la première hypothèse de la thèse de Violaine Rebouillat qui s'appuie sur la notion de « cadre épistémique » de Karin KnorrCetina¹⁰⁷. A ceci près que, bien qu'il semble s'agir d'une culture scientifique partagée, elle ne repose pas sur une discipline en particulier comme l'équipe est pluridisciplinaire.

En effet, ce qui m'a interpellé lors de ces entretiens, c'est la facilité avec laquelle tous les enquêtés passaient du « je » à « nous » sur des questions qui auraient pu paraître très personnelles : celles d'un chercheur face à ses données sur un ordinateur ou d'un chercheur sur son terrain. Il n'y avait aucune question sur le rôle du collectif dans le questionnaire et c'est naturellement qu'ils passaient d'un pronom à l'autre.

Premièrement, ce « nous » n'a pas toujours été le même au long de l'entretien. Il a été tantôt relatif aux enseignants-chercheurs, tantôt relatif à l'équipe AAU-CRESSON, à ses objets de recherche, et à ses protocoles.

1) Les enseignants-chercheurs

Par rapport au premier, la difficulté de porter ces deux casquettes a été relevé, à la fois de faire de la science et de la pédagogie et le temps que prend la pédagogie : « Ça devient vite compliqué, ouais [d'avoir autant de données à analyser]. Surtout quand on n'est pas chercheur à plein temps et qu'on est enseignant »¹⁰⁸.

Cette question, en plus de la dimension temporelle que nous retrouvons, est également en lien avec la (re)définition des missions du chercheur, puisque ce n'est pas seulement la pédagogie qui a été citée mais également :

- la prise en charge de plus en plus importante des côtés administratifs de la recherche : « c'est le fait que je vois dans l'évolution de la recherche, j'ai l'impression d'être de plus en plus sollicité par des choses très diverses qui sont périphériques, disons qui se font en plus des recherches à proprement parler. Une tendance à une démultiplication des demandes de contribution de participation à certains outils, etc. etc. Donc c'est le contrepoids de l'intérêt

¹⁰⁷ REBOUILLAT, Violaine, 2019. *Ouverture des données de la recherche: de la vision politique aux pratiques des chercheurs*.

¹⁰⁸ E4_chercheur

que peut avoir ce type de choses, c'est-à-dire que ça suppose que le chercheur s'investisse là-dedans en plus et c'est au détriment du temps qu'on dispose qui est de plus en plus réduit car la partie administrative est de plus en plus importante »¹⁰⁹ ;

- et la demande à communiquer et maintenant à archiver « C'est infini quoi. Ça veut dire que dès que tu produis quelque chose, tu dois en produire son archive. Et c'est un des problèmes de la recherche actuelle, où on te dit qu'il faut communiquer dès qu'on fait quelque chose. Mais en fait on n'a déjà pas le temps de faire la chose qu'on fait. Si en plus il faut prendre le temps de l'archiver... [...] Et nous... on n'est pas solides. Et il y a aussi la peur de la critique quand on communique trop vite »¹¹⁰.

Un autre enquêté a ainsi souligné comment d'autres métiers jouaient le rôle de relai de la science avec la société, comme les journalistes ou les bibliothécaires (« Il y a des intermédiaires qui sont... qui doivent pouvoir guider les personnes [...] elle peut très bien aller dans une bibliothèque et la bibliothécaire la guide. Pour moi ça serait ça un chemin [...] J'ai fait une petite expertise pour une ONG, je crois qu'il fallait partir à trois conférences, l'expertise était tout de suite mise en ligne etc. Et là c'est vraiment de la science ouverte. C'est un peu leur tâche, et c'est eux qui s'en occupent, c'est eux qui réclament et c'est super »¹¹¹), mais que c'était maintenant attendu également de la part des chercheurs de jouer ce rôle directement.

L'ensemble de ces remarques nous amène à observer deux tendances ressenties parmi les enquêtés : les missions du chercheur évoluent et il est difficile de faire de la science parmi les multiples sollicitations vécues par les chercheurs « On a déjà beaucoup de mal à faire ce qu'on a à faire »¹¹². Ces remarques proviennent d'enquêtés qui sont dans la recherche depuis de nombreuses années. La science ouverte est alors vécue comme une nouvelle sollicitation qui prendra du temps, ce qui inclut l'ouverture des données.

2) Nous, l'équipe CRESSON

La particularité de ce « nous » qui va nous occuper dans cette partie est d'avoir été souvent formé en opposition à d'autres systèmes, à un « eux ». Les relever permet d'avoir une compréhension des travaux rencontrés lors de l'enquête.

L'intérêt pour l'« écriture » des données

Bien que certains objets de recherches et méthodes puissent différer d'un chercheur à l'autre, une certaine culture scientifique semble être partagée par l'équipe, par des positions, des méthodes et des objets de recherche. C'est cette culture scientifique, ces pratiques partagés par un certain nombre, qui permet l'émergence d'un « nous ». Par exemple, avec ce premier enquêté, qui sur une expérience avec une équipe différente, parle de ces difficultés à s'adapter : « On a expérimenté quelque chose qui n'est pas commun au CRESSON puisqu'on a travaillé avec un labo de psychologie expérimentale qui travaille beaucoup avec

¹⁰⁹ E2_chercheur

¹¹⁰ E5_chercheur

¹¹¹ E3_chercheur

¹¹² E4_chercheur

tableaux Excel par exemple. Chose dont on n'a pas l'habitude au CRESSON [...] nous on travaille plutôt de façon synthétique et non analytique. Deux modes de pensée, de travail et d'outils. Ça nous a pas mal déstabilisés parce que ce n'est pas du tout notre culture scientifique »¹¹³.

Cela s'exprime également par des réflexions communes comme sur la représentation des données à partir de la notion d'écriture qui est l'axe suivant du projet scientifique du laboratoire AAU (2022-2026) « EXPRIMER les mondes perçus »¹¹⁴. Les particularités de ces recherches sur l'ambiance dans ce laboratoire sont de s'intéresser :

- Aux sensations dans l'espace et aux différents moyens de les capter, par des protocoles ou techniques particuliers (comme pour l'odorat, le son et autres sensibilités dites minoritaires) ;
- Aux moyens de représentation de ces captations et perceptions de l'espace ;
- Et aux partages de ces représentations.

Ces trois points qui proviennent de l'axe du projet scientifique ne mentionnent pas le mot « donnée », lui préférant celui de « captation », mais cet axe rend tout de même visible une réflexion sur la production de données, leur qualité, leur partage. Ces « objets intermédiaires », les représentations des données, sont ces données « portables » ou ces données analysées et transformées mentionnées dans la partie 2 « il n'y avait aucune photo, aucune bande sonore, rien attaché au texte, c'est pour ça que le texte est à retravailler parce que c'est incompréhensible et elle a tellement de richesses de matériaux d'ambiances qu'elle devrait nous permettre de nous plonger dedans. C'est ça de mémo que j'ai écrit. Pour moi c'est la tâche du scientifique de mettre en valeur ses données pour la compréhension »¹¹⁵.

La particularité est que ces représentations sont elles-mêmes des objets de recherche mis en débat ainsi que les processus employés, dans cette recherche d'écriture des ambiances : « Une chose qui devient aussi de plus en plus émergeant dans les travaux du CRESSON [...] c'est la question des formes et des formats de la recherche. Ces données brutes, enregistrements sonores, vidéos etc., il y a aussi une façon de les travailler, de les monter, de les transformer pour les faire parler d'une certaine manière, les restituer d'une certaine manière. Toute cette question-là des données brutes, sonores entre autres, mais pas seulement, comment est-ce qu'on les analyse et comment on les remet en chantier dans l'usage presque ... du montage. C'est quelque chose qui apparaît de plus en plus. C'est-à-dire entre les données brutes et les résultats finaux, il peut y avoir des processus de transformation des données pour les monter d'une certaine manière »¹¹⁶.

Ces données « portables » qui ont été travaillées d'après un critère de présentabilité sont plus simple à partager, comme pour le projet les Ondes de l'Eau¹¹⁷ où sur l'entrepôt de données SHS d'Huma-Num Nakala ont été publiées leurs

¹¹³ E2_chercheur

¹¹⁴ Projet scientifique AAU 2022-2026, [sans date]. *Laboratoire AAU*. [en ligne]. [Consulté le 13 juillet 2022]. Disponible à l'adresse: <https://aau.archi.fr/laboratoire-aau/prochain-projet-scientifique-aau-2022-2026/>

¹¹⁵ E3_chercheur

¹¹⁶ E2_chercheur

¹¹⁷ projet de recherche-cr  ation Les Ondes de l'Eau : M  moires des lieux et du travail dans la vall  e de la Romanche AAU-CRESSON (Laure Brayer, direction scientifique) - Regards des Lieux (Laure Nicoladz  , direction culturelle)

matériaux de terrain éditorialisés, mais non les matériaux bruts. Par exemple, la « Carte des mémoires sensibles de la Romanche » replace sur une carte des paroles des habitants, mais les entretiens eux-mêmes qui ont servi à cette carte ne sont pas accessibles. De plus, ce sont également ce type de données qui sont restituées aux enquêtés sous la forme d'expositions, de « parole donnée, parole rendues »¹¹⁸ ... Ce sont des données qui possèdent un double statut, à la fois qui ne sont pas « brutes » et en même temps qui restent des objets de recherche, auxquels des publications peuvent être dédiés, et qui sont préférées lors d'un partage ou d'une ouverture (en ligne ou non).

Cette idée d'écriture renforce l'idée d'auteur de données, même si une ligne est faite avec l'auteur de fiction, puisqu'il y a tout de même un certain acte de création mis en exergue : « Par exemple, faire un entretien c'est une dimension quand même de ... pas de création mais presque, faire des photos, faire des films, ce n'est pas ... ce n'est pas des données neutres, c'est quand même un acte créatif à un moment donné de faire ça. [...] il y a une dimension d'auteur, non pas d'auteur comme un auteur artistique au sens mais... il y a une forme d'auteur »¹¹⁹. Le chercheur est crédité comme auteur lors de prises de photo ou la réalisation de films, sauf si par exemple l'appareil photo est posé et se déclenche automatiquement selon un protocole. Au contraire, de par leur statut administratif, les ingénieurs du laboratoire ne sont pas auteurs.

Sur le terrain : la peur de la surformalisation

Une crainte a été soulevée par plusieurs enquêtés : celle de la « surformalisation » par la gestion des données qui ne correspondrait pas forcément à leurs façons de procéder. Cette surformalisation est ressentie ainsi parce qu'elle entre dans une recherche qualifiée de bricolage : « Le terme bricolage : on a des hypothèses qu'on essaye de sonder, d'évaluer sur le terrain, on a des hypothèses, des pressentis et on n'essaye pas de l'argumenter mais de l'évaluer et donc on y va... étape par étape, pas à pas, et puis aussi ce sujet évolue énormément en fonction de l'enquête de terrain, des personnes rencontrées, de la richesse de l'environnement exploité, et ... on n'est pas non plus des romanciers, mais ce n'est pas comme une validation. Je vois mes collègues ingé, aucun point de vue hautain, ils ont une démarche complètement antagoniste qui va valider une posture de débat, de base. Alors que nous on déduit, eux c'est toujours dans le sens inverse. Et ... donc ça ne fait pas la même chose quoi. D'où ces méandres, ces retours qui se font au gré des situations des rencontres... »¹²⁰.

Ce terme de « bricolage » utilisé par l'enquêté peut être mis en relation avec celui d'« artisanal » développé par Joachim Schöpfel¹²¹ puis par Violaine Rebouillat : « Les termes d'« artisanat » et de « bricolage » renvoient à une méthodologie individuelle, unique car propre à chaque chercheur. Celle-ci n'aboutit pas au recueil de données standardisées et homogènes. C'est ce que véhicule la comparaison sous-jacente entre artisanat et industrie : l'objet artisanal a un caractère

¹¹⁸ ROUX, Jean-Michel et TIXIER, Nicolas, 2011. Paroles données, paroles rendues : la fabrique de la ville à l'épreuve des usages. *Vies des villes, trimestriel de l'architecture et de la ville*. [en ligne]. 2011. N° 16, pp. 84. [Consulté le 13 juillet 2022]. Disponible à l'adresse: <https://hal.univ-grenoble-alpes.fr/hal-00995534>

¹¹⁹ E5_chercheur

¹²⁰ E6_chercheur

¹²¹ SCHÖPFEL, Joachim, 2018. *Vers une culture de la donnée en SHS* [en ligne]. Rapport de recherche. Université de Lille. [Consulté le 13 juillet 2022]. Disponible à l'adresse: <https://hal.archives-ouvertes.fr/hal-01846849>

unique car il procède d'ajustements, qui sont fonction du matériau travaillé ; dans l'industrie, des protocoles normalisés et des gestes répétés permettent de produire des objets chaque fois identiques. C'est cette nuance qui distingue la théologie ou l'anthropologie des autres disciplines scientifiques, dans lesquelles le terme de « donnée » est courant »¹²².

Selon l'espace étudié, selon l'objet de recherche, selon les habitants rencontrés, le projet de recherche évolue en allant sur le terrain, qui, sans remettre en cause le plan de gestion de données, rend difficile son côté prévisionnel et formel. Il est alors encore plus nécessaire de le faire évoluer avec le temps et rend intéressant ceux qui sont rétrospectifs. Ici, ce n'est pas qu'une méthodologie individuelle mais également une méthodologie qui se construit collectivement.

Cette surformalisation a également été décrite comme problématique vis-à-vis de leurs enquêtés, avec les nouvelles pratiques introduites par le RGPD. Celles-ci ont la capacité de modifier profondément les relations avec les enquêtés. Des pratiques d'autorisation, informelles, existent depuis de nombreuses années et leur rôle en tant que chercheur, d'après eux, est d'instaurer un climat de confiance. Or, d'après plusieurs chercheurs, les formulaires n'aident pas à renforcer ce lien de confiance et éveillent souvent la méfiance des enquêtés : « Non [le formulaire ne permet pas d'instaurer une bulle de confiance]. Moi je trouve que c'est même l'inverse. Parce que ça formalise. Alors qu'au début avec la parole on part sur une confiance. On part plutôt sur quelque chose d'intime que je te dévoile quelque part »¹²³. En effet, paradoxalement, ce qui est censé redonner du pouvoir à l'enquêté renvoie plutôt l'image d'une autorisation absolue pour le chercheur d'utiliser leurs données personnelles (« Je rentre dans une phase comme si il m'autorisait à faire ce que je veux, il me libère, je ne sais que ce n'est qu'un papier mais quand on signe un truc, on s'engage à quelque chose, on se désengage de quelque chose »¹²⁴). Enfin, cela participe à une formalisation, à une entrée dans le domaine juridique qui change le rapport.

Des moyens d'adaptation sont alors mis en place, en faisant signer les formulaires à la fin plutôt qu'au début de l'entretien lorsqu'il s'agit de rencontres dans la rue avec des habitants. Ou alors des autorisations qui se font à l'oral (enregistré) plutôt que sur papier, notamment lorsque l'enquête ne se passe pas en France et qu'une méfiance existe.

Il semble qu'il y a également le sentiment que cela va trop loin, qu'il y a un décalage entre leurs besoins de sécurité et ce qui est demandé par le RGPD. C'était comme si les conséquences d'une faille de sécurité ne les concernaient pas, comme la vie numérique de la science n'est que peu visible dans le reste de la société.

Des données contextuelles et situées

Une différence fondamentale a été faite à plusieurs reprises sur la question du partage des données, en opposition au « big data » ou aux sciences fondamentales. Les données produites seraient plus contextuelles et situées, plus sujet aux frictions lors de partage de données : « ce n'est pas comme si nous étions dans des méthodologies big data où les choses sont extrêmement formalisés, analytiques, très

¹²² REBOUILLAT, Violaine, *op. cit.*, p. 166

¹²³ E6_chercheur

¹²⁴ E5_chercheur

comptable et statistiques. Ce sont des corpus qu'on peut plus facilement communiquer, on peut triturer les chiffres dans tous les sens. [...] c'est important de connaître le contexte de production des big data bien sûr, mais je pense que c'est plus facilement possible de travailler avec sans avoir été dans le cadre dans lequel ils ont été produits que dans le cas d'enquête qualitative »¹²⁵.

L'ouverture à la société des données dans l'équipe est fortement compromise par la réglementation RGPD comme il s'agit de données personnelles, voire sensibles. L'intérêt dépend de la nature des données qui seraient ouvertes, selon la discipline, et ceci est perçu par un enquêté : « S'il y avait un site d'agrégation de toutes les données de mesure acoustique, par type, par mesure de réverbération, que tu puisses trier... On pourrait faire des trucs incroyables, on pourrait comparer des données, des trucs qui n'ont jamais été fait, à des échelles où ça prend tout de suite de la pertinence, puisqu'on voit tout de suite les dynamiques. Pour ces choses-là ça serait super d'avoir de l'ouvert, surtout qu'il n'y a pas trop de donnée sensible, il n'y a pas trop d'humain. Voilà il ne faut pas systématiser, selon les types de données, les types de recherche, selon les disciplines »¹²⁶.

L'influence de l'environnement dans lequel s'inscrit la recherche

L'influence de l'environnement a été visible lorsqu'il a été discuté avec deux enquêtés d'un logiciel produit dans le cadre d'une recherche, qui est libre et gratuit. Plusieurs raisons ont été mises en avant pour que ce soit libre, qui dépassent la demande du financeur, comme son objectif premier d'être enrichi par d'autres « il y a une partie logicielle et celle-là elle est complètement ouverte. Elle est sur GitHub, le code source est ouvert, tout le monde peut le télécharger comme il veut. Toute la partie logicielle est disponible et en fait c'est le but : c'est que les gens se l'approprient, le prennent, le transforment... »¹²⁷. De plus, cela paraissait naturel puisqu'ils ont bénéficié d'outils ouverts dans le domaine de l'architecture et du design et qu'ils souhaitaient s'inscrire dans le même mouvement « c'est que nous on utilise tous les autres et on ne voit pas pourquoi on ne proposerait pas un truc que les autres peuvent utiliser. C'est beaucoup plus intéressant, ça s'enrichit »¹²⁸.

II. LES EXPERIENCES DE CRESSOUND, CARTOPHONIES ET DE LA BIBLIOTHEQUE DU CENTRE DE DOCUMENTATION DE L'EQUIPE AAU-CRESSON : ACQUERIR UNE EXPERTISE SUR LE SON

Il a été mentionné plus tôt l'attention donnée par l'équipe dans le cadre de recherches aux sensations. Il y en a une plus particulièrement qui a été développée depuis la naissance de l'équipe : l'ouïe. L'équipe a développé rapidement une expertise en enregistrements sonores et en analyse, avec le développement d'

¹²⁵ E2_chercheur

¹²⁶ E7_chercheur

¹²⁷ E7_chercheur

¹²⁸ E4_chercheur

« effets sonores »¹²⁹, suivie par une expertise en documentation de ces données sonores depuis les années 2000 pour pouvoir les conserver et les partager. Deux projets de partage en particulier seront discutés dans le cadre de ce mémoire : CRESSOUND et Cartophonies.

1) Naissance du projet CRESSOUND : impulsion par des ingénieurs d'étude

Le projet CRESSOUND (CRESSON et sound, son en anglais) est une bibliothèque sonore composée d'enregistrements sonores, de montages sonores produits lors de recherches et d'enregistrements de séminaires de l'équipe AAU-CRESSON. Ce projet a débuté avec l'idée de conserver et de valoriser les archives sonores du CRESSON qui étaient alors sur bandes magnétiques : « C'est aussi une manière de centraliser. C'est comme quand tu as un bouquin, ce n'est pas la même chose quand tu as une bibliothèque, [...] ce ne sont pas juste des objets éparpillés, mis en relation les uns avec les autres, ça dessine un certain ... champ de la connaissance on va dire. Qui serait lié au champ de la connaissance du CRESSON autour du sonore »¹³⁰.

C'était d'abord une collection de CD avant d'être maintenant accessible en ligne sur le site du laboratoire AAU : <https://aau.archi.fr/CRESSON/cres-s-o-u-n-d/>. Ce travail s'accompagne d'un catalogage sur le SIGB du centre de documentation CRESSON qui relie les sons à la publication, que ce soit une thèse, un rapport de recherche, un article...



Figure 2 : Capture d'écran de la page dédiée au projet CRESSOUND sur le site du laboratoire AAU, 2022, <https://aau.archi.fr/cresson/cres-s-o-u-n-d/>

¹²⁹ AUGOYARD, Jean François et TORGUE, Henry, 1995. *A l'écoute de l'environnement: répertoire des effets sonores*. Marseille : Parenthèses. Collection Habitat/ressources. ISBN 978-2-86364-078-4.

¹³⁰ E9_ITA



Figure 3 : Capture d'écran de la notice « Marcher à l'écoute de la médina à Alegma : Nefta », enregistrement sonore fait dans le cadre de « Résonances Oasiennes », sur le catalogue en ligne du centre de documentation du CRESSON, 2022
http://doc.cresson.grenoble.archi.fr/index.php?lvl=notice_display&id=8603

Ce projet est né suite à des échanges avec Véronique Ginouvès, responsable de la Phonothèque de la Maison méditerranéenne des sciences de l'homme, et de son guide sur le catalogage de sons. Celui-ci, dédié aux sons inédits, a été adapté pour correspondre aux besoins des enregistrements sonores produits par l'équipe, provenant d'environnements sonores urbains en majorité.

Ce travail d'adaptation s'est fait également à partir du travail déjà réalisé par les chercheurs dans les rapports de recherche. En effet, à la fin de chaque rapport de recherche, se trouvaient des fiches décrivant les données : « et après il y a eu ce projet de Cartophonies qui s'est mis en place dans l'idée de valoriser le travail descriptif qu'il y avait dans catalogue PMB, qui était lui-même issu de toutes les fiches descriptives de sons. Il faut imaginer que si on transposait ça aujourd'hui, il y aurait des rapports de recherche, ce qu'il n'y a plus. A la fin du rapport de recherche il y aurait des fiches descriptives sur toutes les données. C'est ça qu'il y a depuis le début. Il y aurait ça sur les cartes, sur les photos, les itinéraires... Des métadonnées qui sont issues de fiches cartonnées que faisaient faire Jean-François Augoyard quand ils sortaient leur magnétophone à bande. Lui il partait sur cette idée de descriptif. Après ils ont petit à petit travaillé sur des métadonnées de leurs sons ». Ce travail de documentation dans le cadre de la recherche prend du temps : « c'était un boulot énorme, c'était un boulot gigantesque, il faut être ... pour informer les données comme ça, il faut énormément de données »¹³¹.

Ce travail a été à la fois utile pour les chercheurs, puisqu'ainsi ils développaient le répertoire des effets sonores, et pour le travail de documentation qui a suivi, en stabilisant ce qui était attendu et nécessaire (matériel, lieu, descriptif avec intérêt ...), en complétant avec les normes alors en vigueur et les pratiques d'autres institutions comme la Phonothèque de la Maison méditerranéenne des sciences de l'homme.

¹³¹ E4_chercheur

Les deux membres de l'équipe ITA qui ont pris en charge ce projet, Françoise Acquier et juL McOisans, avaient d'abord répondu à un appel à projet pour numériser la collection mais celui-ci n'a pas abouti. La numérisation a alors été prise en charge par ce dernier. Ils ont au fur et à mesure des années développé le système, avec parfois l'obligation de tout recommencer à zéro « on a essayé de stabiliser un truc (entre normes et pratiques chercheurs) et ça nous a pris vachement de temps. Puis quand on est arrivé à stabiliser un truc, on a tout catalogué et on s'est rendu compte qu'un truc n'allait pas, on a tout repris. Avec juL, avec des stages... à chaque fois on progressait et on se disait : merde 500 [notices] à reprendre »¹³².

2) Naissance du projet Cartophonies : impulsion par des chercheurs

Le projet Cartophonies a sur son site des enregistrements sonores en commun avec CRESSOUND, mais son origine et ses buts diffèrent de ce dernier. Il a été imaginé suite au projet de recherche sur l'Isle d'Abeau, qui avait pour objectif de valoriser son patrimoine sonore. Or, les chercheurs étaient confrontés à un terrain important. Ils ont eu alors l'idée de géolocaliser les sons, d'abord avec un processus assez long et difficile « Et comme c'est un territoire très grand de 20 km de long sur 10 de large ou 5, on s'est dit en fait on ne peut pas c'est inappréhensible, c'est énorme. Il nous faut une carte [...] on va localiser les sons qu'on prend et les mettre sur la carte. On va appeler ça une cartophonie »¹³³, « c'était hyper stimulant mais on l'avait fait sur format flash, c'était très épuisant à coder »¹³⁴.

De cette expérience réussie est venue l'envie de créer un site où les sons seraient géolocalisés, idée financée par le laboratoire pour permettre son développement. Les archives sonores du CRESSON ont été reprises pour ajouter cette métadonnée lorsque cela était possible avant de l'importer sur le site. D'autres projets de recherche ont suivi pour l'alimenter de leurs enregistrements sonores.



Figure 4 : Capture d'écran du site du projet Cartophonies, 2022, <https://www.cartophonies.fr/>

¹³² E8_ITA

¹³³ E4_chercheur

¹³⁴ E6_chercheur

Le choix a été fait de créer un site participatif, qui donne la possibilité à des personnes externes à l'équipe de proposer des sons. Par rapport à d'autres sites qui ont le même concept de sons géolocalisés, les enquêtés ont souligné la particularité de celui-ci (« c'est orienté CRESSON, j'ai un espace, j'ai des pratiques, j'ai des données physiques et j'essaye d'expliquer pourquoi le son est comme ça là »¹³⁵) et l'importance d'avoir des enregistrements sonores qui ont de l'intérêt (« Je n'ai jamais déposé sur Cartophonies des prises de son que j'ai faites dans le vide, comme ça, parce que je pense que c'est intéressant que ça reste un outil, qu'il y ait une problématique, un objet d'étude clair etc. pour que ce soit de la donnée de la recherche »¹³⁶, « La spécificité de Cartophonies, c'est qu'ils sont là parce que. Ça intéresse telle problématique, c'est par rapport à tel mot clé, tel intérêt d'écouter ce son »¹³⁷). C'est alors une plateforme qui peut intéresser les autres communautés (qui ne sont pas forcément scientifiques mais artistes) qui s'intéressent au sonore « La communauté sur les ambiances sonores n'est pas non plus très grande, donc si on peut croiser nos différentes approches tant mieux »¹³⁸. Cela a par exemple suscité l'intérêt d'enseignants, de chercheurs en acoustique et géographie.

Les enregistrements proviennent régulièrement de recherches et une attention forte est donnée aux métadonnées, par un résumé qui présente le son, des photos... Cette plateforme a également été mobilisée lors d'un appel à enregistrer son environnement sonore lors du confinement. Le résultat est à demi-teinte pour l'instant : « C'était une manière d'ouvrir un processus qui me paraissait intéressant, on va faire de la prise de son alors qu'on est enfermé derrière nos fenêtres ou sur nos balcons [...] le problème c'est que les gens font n'importe quoi. Ils ne vérifient même pas si ça marche. Y en a qui sont loupés et y en a qui sont bien [...] Mais ça pose quand même la question que tu retrouves avec une avalanche de trucs qui sont pas forcément prévus, consignés. Et finalement est-ce que tu peux vraiment les utiliser ? Un peu mais pas autant que je l'imaginais »¹³⁹. Cependant, ce côté participatif permettrait quand même de donner la possibilité aux étudiants de l'enrichir « tu fais un atelier tu peux dire à tes étudiants vous les mettez directement. Il y a un côté où c'est eux qui sont acteurs qui est vraiment valorisant »¹⁴⁰.

Deux enquêtés sont revenus sur les intérêts qu'ont eu ce projet ou qu'il pourrait y avoir. Comme Cartophonies a récupéré de manière rétrospective des sons produits par des chercheurs de l'équipe, le projet a été l'occasion de se replonger dans les travaux portés par celle-ci : « ce qui était hyper stimulant pour moi, c'était de replonger dans l'héritage de tous les premiers rapports d'étude faits au CRESSON sur la sonorité. Ce qu'avait fait Pascal Amphoux, Jean-François Augoyard et autres... Et comment ils l'avaient abordé ce fragment sonore. C'était très riche ; pouvoir localiser, rendre des métadonnées autour... c'était extra »¹⁴¹. Et c'est également une proposition pour le futur, sans savoir ce que cela peut devenir « Tu ne sais pas à quoi peut servir quelque chose que tu proposes et que tu ouvres, effectivement, que tu donnes en pâture, même si toi tu l'as fait dans une autre intention. Ce qui compte, ce n'est pas tellement l'intention que nous on avait au

¹³⁵ E4_chercheur

¹³⁶ E7_chercheur

¹³⁷ E9_ITA

¹³⁸ E6_chercheur

¹³⁹ E4_chercheur

¹⁴⁰ E6_chercheur

¹⁴¹ E6_chercheur

départ, c'est l'intention de celui qui trouve la chose. Ça c'est de la science ouverte. Peut-être que ça servira à d'autres intérêts qui ne sont pas spécifiquement les nôtres aujourd'hui »¹⁴².

3) Fonctionnement et pérennité

Ces deux sous parties dédiées aux projets nous amènent à une figure importante, celle du médiateur. Bien que certains chercheurs soient relativement autonomes, d'autres passent encore par les ITA qui font également attention aux projets en cours et qui pourraient contenir des sons intéressants. Ils sont également là pour vérifier la qualité des métadonnées des sons déposés, que ce soit sur Cartophonies (ce qui inclut des personnes externes, et des chercheurs sont également administrateurs) et sur CRESSOUND et la bibliothèque.

Même pour ceux et celles qui font attention à la documentation de leurs données, il peut y avoir des manques « Même par exemple les sons de ma thèse, où j'avais toutes les métadonnées. C'était ma thèse donc du coup les sons étaient hyper calibrés : sur chaque son, j'avais un fichier texte avec toutes les métadonnées possibles et imaginables. J'ai tout envoyé à juL, écoute voilà il y a ça qui est en lien avec ma thèse et on peut mettre en ligne, et il m'a dit c'est super par contre il manque ça, ça, ça. Je me dis ce n'est pas mal de passer par eux, du coup c'est vérifié, ils savent exactement ce qui est pertinent de mettre, pertinent de ne pas mettre parce qu'en fait c'est leur écosystème »¹⁴³. Cela demande un temps d'échanges pour pouvoir déposer les sons « Tous ces fragments-là, une vingtaine par recherche, il faut passer après un par un, et tu t'aperçois qu'il y a toujours des erreurs, toujours des fautes. Il y a un côté vampirisant »¹⁴⁴.

Ce sont également des personnes qui portent ces projets et qui font attention à ce qu'ils continuent à vivre (« Ce n'est pas un truc automatique qui se génère tout seul, il faut qu'il y ait quelqu'un derrière »¹⁴⁵, « Soit c'est porté par une équipe chercheurs et IT, soit c'est pas grave, ça s'arrêtera là »¹⁴⁶), ce qui pose par conséquent la question de la pérennité de ceux-ci. Il y a un besoin constant de faire de la veille et de faire évoluer pour correspondre à l'environnement (« cette veille-là technique, organisationnelle, ces chercheurs ne le voient pas, ne veulent pas dialoguer... [...] Il faut nous donner du temps pour travailler sur les données du labo, leur structuration, leur archivage, sinon tu perds le fil, c'est irrattrapable quoi. Ce n'était déjà pas parfait, mais plus tu perds du temps et plus tu t'éloignes des évolutions »¹⁴⁷). Il y a à la fois le désir de faire vivre ces projets aussi longtemps que possibles (« ad vitam aeternam tant qu'il y aura des serveurs. C'est plutôt le principe, c'est comme un... pas un musée. C'est un truc immatériel qui peut être consulté. Je pense que cette dimension historique n'est pas inintéressante »¹⁴⁸) et en même temps de savoir que tout peut s'arrêter un jour.

¹⁴² E4_chercheur

¹⁴³ E7_chercheur

¹⁴⁴ E6_chercheur

¹⁴⁵ E9_ITA

¹⁴⁶ E8_ITA

¹⁴⁷ E8_ITA

¹⁴⁸ E4_chercheur

Par ailleurs, c'est peut-être par ces expériences que les résultats de l'enquête ont été fortement marqué par une sorte d'incrédulité sur l'argument de la pérennité numérique. De voir l'évolution du support pour le son en une quarantaine d'années ou même vingtaine d'années, de devoir faire évoluer les systèmes de dépôt, se poser au moment de ce mémoire la question des serveurs d'Huma-Num, une très grande infrastructure de recherche, et des DOI (digital object identifier, identifiant unique).

4) Une donnée : des représentations, des projets

Avec ces deux projets, avec parfois les mêmes sons, deux propositions sont faites qui n'ont pas les mêmes objectifs, ni les mêmes visualisations ou rapports, mais sur un même type de données.

CRESSOUND est une collection et la mémoire de travaux de l'équipe. Il s'appuie également sur un travail d'archivage sur le serveur local et d'une signalisation sur le catalogue du centre de documentation. Cartophonies est un site Web qui a la particularité de mettre en avant la localisation de la prise de sons par une apparence de carte, il s'ouvre également à l'extérieur en proposant de contribuer. La particularité du sonore est d'être à la croisée de plusieurs communautés intéressées par le même objet (« On est dans un micromilieu de recherche sur l'environnement, le territoire, le son, mais comme je disais tout à l'heure, il y a des tas d'autres gens qui s'intéressent à ces matériaux et qui ne sont pas dans la recherche architecturale... paysagistes sonores, artistes... »¹⁴⁹), et ces projets permettent un contact avec celles-ci.

Dans la sous-partie précédente, nous avons discuté de l'argument de données « teintées » par des êtres humains, contrairement au big data. Par l'histoire de ces deux projets, nous voyons l'importance de nouveau de la représentation des données : comment celles-ci sont teintées dans ces projets par des besoins et des méthodes particulières comme l'étude du rapport entre le son et l'espace ou le déploiement des effets sonores. Avec d'autres besoins, nous pourrions imaginer d'autres formes et d'autres types de documentation nécessaire. Comment prendre en compte tous les besoins possibles sur une plateforme externe à tous, dont (ou pour ?) des personnes avec d'autres méthodes ? Comment construire à partir du local quelque chose qui le dépassera ?

III. LA QUESTION FINALE DE QUI VA S'EN OCCUPER : LE COUT DU PARTAGE (ET DE LA GESTION)

1) Développer une expertise qui dépasse les enregistrements sonores

La partie précédente met en avant le processus qui a mené à une réflexion sur la documentation de données, leur stockage et leur partage. Cela a été un processus lent, plusieurs fois renouvelés et qui se doit de suivre des évolutions. Il prend racine dans une reconnaissance dans l'importance des enregistrements sonores, dans le cadre de recherches et dans le cadre de la mémoire du laboratoire, ce qui réunit à la fois les chercheurs, le service de soutien et d'appui à la recherche, et la direction.

¹⁴⁹ E9_ITA

En effet, cela a été une réflexion collective, à une échelle locale, qui a permis de systématiser.

A titre de comparaison, la gestionnaire de l'information et de la documentation Audrey Carbonnelle du laboratoire de recherche Architecture, Environnement & Cultures Constructives qui est aussi à l'ENSAG, accompagne également les chercheurs sur la question de la gestion des données. Elle a développé de son côté un thésaurus et un accompagnement sur un outil de photothèque qui correspondait plus aux besoins de la recherche de ce laboratoire, dont une partie des preuves repose sur les photographies. Il est nécessaire de pouvoir les retrouver et les partager plus facilement.

Dans chacun de ces laboratoires, un travail sur la documentation de données a été réalisé, ainsi que son stockage et son partage, en partant de besoins communs à plusieurs recherches. Cependant, il s'est toujours agi d'un type de donnée en particulier. Que ce soit les enregistrements sonores ou les photographies, les outils à utiliser et la documentation à réaliser changent sensiblement (ainsi que les aspects juridiques), et cela sans doute également d'un laboratoire à un autre.

La nécessité est aujourd'hui d'appréhender l'ensemble des types de données pour les documenter, et comme vu dans la partie 2, ceux-ci sont multiples. Parfois, ils peuvent se délimiter par leur type (les photographies, les sons, les cartes etc.). Parfois, une donnée est hybride et rassemble plusieurs types. C'est le cas avec le protocole du « baluchon multisensoriel » qui vise à appréhender les expériences sensibles des habitants sur un territoire en leur donnant différents moyens d'expression sur une semaine¹⁵⁰. C'est à la fois un objet matériel, des images de tout type (photographies, dessins), des enregistrements sonores, un carnet. Il s'agit aujourd'hui de savoir documenter aujourd'hui toutes ces données « la pluralité des supports des données aujourd'hui ... c'est plus difficile d'arriver à cet état de propreté, de perfection qu'on a eu pour le son, mais c'est mobilisateur pour les chercheurs, pour les IT »¹⁵¹.

2) La sensibilisation et l'aide à apporter aux chercheurs

Dans le domaine de la science ouverte, il est attendu que les professionnels de l'information scientifique et technique puissent appréhender ces nouveaux outils (qui semblent se multiplier de jour en jour), ces enjeux et puissent ainsi accompagner les chercheurs dans les changements à venir¹⁵². C'est une demande qui a, par exemple, été prise en charge par la chargée des ressources documentaires Françoise Acquier « Je pense quand même que c'est ce que nous demande le CNRS, d'une certaine manière on est ce personnel d'accompagnement [...] Je pense que l'institution comptait sur nous, sur le relai que l'on pouvait être. Donc on se formait entre nous, on s'épaulait entre documentalistes et puis on adaptait nos stratégies dans les labos pour porter ça ».

¹⁵⁰ MANOLA, Théa, 2013. Rapports multisensoriels des habitants à leurs territoires de vie. Retours critiques sur une démarche méthodologique. *Noroi. Environnement, aménagement, société*. [en ligne]. N° 227, pp. 25-42. [Consulté le 20 juillet 2022]. DOI : [10.4000/noroi.4649](https://doi.org/10.4000/noroi.4649).

¹⁵¹ E8_ITA

¹⁵² DASSA, Michèle et KOSMOPOULOS, Christine, 2021. Les évolutions transversales de l'information scientifique et technique avec l'apparition du numérique. *Histoire de la recherche contemporaine* [En ligne], Tome X - n°2 [Consulté le 16 mars 2022.] DOI : <https://doi.org/10.4000/hrc.6097>

Ce travail de sensibilisation se fait alors au niveau local, en accompagnant les projets de recherche et leurs besoins spécifiques. Il diffère à celui sur HAL et les publications car il est plus récent et s'appuie sur des injonctions de financeurs « après les commanditeurs font beaucoup. Quand c'est l'ANR qui demande et que, pour eux, c'est le financeur de leur travail et bien c'est un appui énorme. Donc c'était évidemment super bien vu. Stratégiquement c'était bien joué ». Un autre enquête met également en avant que le RGPD rend plus fort le besoin du plan de gestion de données.

Le travail de sensibilisation sur HAL, lui, a demandé de nombreuses années et a vu la mise en place de modèles différents dans l'accompagnement des chercheurs « Avec des hauts et des bas, et énormément de réticences, et en ayant fait jouer énormément de stratégies. Pleins de choses ont été testées. J'ai fait moi-même, j'ai fait venir des gens du CCSD, j'ai fait venir des gens de l'UGA [...] On a dit qu'on allait déposer pour les chercheurs, qu'on les formait mais qu'on était quand même beaucoup là. Après 2007, on importe des notices comme ça ils vont ajouter des fichiers [...] on vous met tout et vous complétez mais ça n'a pas super bien marché. 2014, l'université de Grenoble met en place un portail, là les autres labos sont sensibilisés aux questions d'archives ouvertes, on était quand même plutôt précurseurs, donc là émulation entre équipes au niveau local [...] à force d'élan de travail qu'on donne, qu'il leur en demande mais on fournit beaucoup aussi, et du coup il y a une forme de reconnaissance de tout cet effort. Et certaines personnes qui ont vraiment pris le pli maintenant, comme Nicolas, Rachel, il y a des chercheurs qui y pensent vraiment maintenant. Ils sont un relai pour eux et pour les autres »¹⁵³.

Un enquête, en contact avec plusieurs institutions pour des recherches, a observé plusieurs modèles qui se mettent en place : un où un service est dédié à la gestion des données de la recherche « le service informatique et de gestion de données où s'est complètement externalisé mais avec aucun contact, c'est un peu pour eux comme le service administratif, on va valider notre budget, on va valider notre plan de gestion de données », un où la question n'avait pas été prise en charge à temps « quand c'est arrivé par exemple, ma collègue était en panique, parce qu'on lui a demandé de faire un plan de gestion de données et là c'était grosse panique à bord parce qu'elle ne savait pas du tout ce que c'était », et enfin le CRESSON avec le travail d'accompagnement proposé « nous on fait plutôt l'opposé je pense [...] sur le lien comme peut le faire, j'ai l'impression, Françoise, et toi, de qu'est-ce que ça apporte à ces chercheurs d'intégrer ces plans de gestion de données [...] il n'y a pas d'appui à la recherche comme on a avec vous. Et donc tout ça, nos journées en tant que chercheur sont déjà assez pleines, on n'a pas forcément le moment de se dire tiens si on allait voir le traitement des données »¹⁵⁴.

Cependant, ce travail d'accompagnement trouve ses limites dans les ressources qu'il peut apporter en termes humain et de temps. Rappelé lors de la présentation du décret n°2021-1572 de décembre dernier¹⁵⁵, la gestion des données mobilise une série de professionnels qui ont un champ de compétences différents des uns des autres, incitant notamment à la collaboration entre professionnels.

La question se pose à la fois au niveau de la sensibilisation à effectuer et à la gestion concrète des données : quel modèle mettre en place ? Cela a été la question

¹⁵³ Françoise Acquier

¹⁵⁴ E6_chercheur

¹⁵⁵ DATACTIVIST et COUPERIN, 2022. *De la science ouverte à l'intégrité scientifique : comment le décret n°2021-1572 bouleverse la conservation et la mise à disposition des données*. [Webinaire en ligne] [notes personnelles].

principale d'un enquêté qui a résumé ainsi sa pensée « Le problème de fond, ce n'est pas la volonté de oui ou non, c'est la mise en œuvre derrière. Et ce que ça implique »¹⁵⁶.

Voyant la veille technique à effectuer, les compétences numériques (sur les outils, les langages), et nous pourrions ajouter le travail d'adaptation à des référentiels et les dimensions juridiques, cet enquêté se demandait les ressources qui seraient fournies aux chercheurs et si celles-ci se feraient par projet ou par laboratoire, avec du personnel dédié ou non « Ce n'est pas facile comme question. Il y a des gens qui se forment pour ça. Dans le projet qu'on a déposé, ça va marcher parce qu'ils ont prévu des centaines de milliers d'euros, donc il y a des compétences qui seront recrutés pour ce projet-là. Mais après ce projet-là, comment ça se passe ? [...] Dans les budgets de labo, il faut que les cotations soient 20 % de plus pour financer ça. Ou alors que des fonctions supports soient attribués mais ça m'étonnerait parce qu'ils sont en train de couper tout. Soit de dire que tous les projets de recherche doivent prévoir un volet pour financer ça. L'idée est noble, mais la mise en pratique... »¹⁵⁷. L'ouverture des données demande une préparation particulière, du travail et du temps supplémentaires : « après je ne sais pas dans quelle mesure il faut tout donner ... si tu l'anticipes et tu te dis, ok dans cette recherche tout va être donné, c'est vachement bien mais par contre ça demande beaucoup plus de travail »¹⁵⁸.

Il est difficile aujourd'hui de répondre à la question du modèle qui sera mis en place, car nous sommes au début de ce mouvement, qui nécessitera sans doute de nombreuses années avant que cela aille jusqu'à la formation des étudiants en recherche, comme projette le billet de Claudine Dellode¹⁵⁹ ou comme c'est le cas dans certaines formations « avant [elle] faisait de la recherche en psychologie. Et elle depuis son master, ça a toujours été là : comment on gère, alors ils n'appelaient pas ça donnée de recherche, mais voilà comme on gère toutes les données qu'on accumule que ce soit sur un enregistreur, un papier, comment on les stocke, quelles sont les règles de stockage. Quand elle m'avait expliqué ça en début de ma thèse, je m'étais dit mais jamais je ne vais faire ça. Qu'il fallait dire aux gens quand tu commençais l'entretien comment tu allais stocker les données, où tu allais les stocker, combien de temps tu allais les stocker. Des trucs ... mais elle, elle fait ça depuis son master, son master c'était en 2009, 2008. Et j'imagine qu'ils faisaient ça depuis bien longtemps comme ils l'enseignaient aux étudiants. C'était une méthode de travail obligatoire, il n'y avait pas le choix, il fallait être très transparent sur comment était gérée la donnée. Il y avait tout un truc sur la sécurité aussi, que maintenant nous on doit faire à cause, à cause ou grâce, au RGPD, mais elle, elle le gérait déjà quoi [...] en fait ce n'est vraiment pas récent mais nous on en parle depuis peu de temps »¹⁶⁰.

Des solutions pour sensibiliser et accompagner ont été mis en place au niveau national, avec l'inauguration de l'entrepôt national Recherche Data Gouv le 8 juillet 2022. L'entrepôt de données représente une solution dite « souveraine », qui

¹⁵⁶ E1_chercheur

¹⁵⁷ E1_chercheur

¹⁵⁸ E7_chercheur

¹⁵⁹ DELODDE, Claudine, 2022. Les bibliothèques universitaires : un acteur clef dans la sensibilisation et la formation des enseignants-chercheurs à la science ouverte. *DLIS*. [en ligne]. [Consulté le 20 juillet 2022]. Disponible à l'adresse: <https://dlis.hypotheses.org/6064>

¹⁶⁰ E7_chercheur

participerait au rayonnement de la recherche en France¹⁶¹. C'est une solution pour les communautés scientifiques qui n'ont pas d'entrepôt de données institutionnels ou thématiques dédiées.

L'ouverture de cet entrepôt s'accompagne avec la mise en place d'ateliers de la donnée (des services de proximité), des centres de référence thématique et des centres de ressource. Les services communs de la documentation développent également de plus en plus au niveau des universités des formations et des propositions d'accompagnement sur des enjeux de science ouverte¹⁶².

Dans le cas de l'équipe CRESSON, ce genre d'offre n'est pas porté par le centre de documentation de l'école d'architecture de Grenoble. Cependant, la Cellule Data de l'Université Grenoble Alpes a été certifiée « atelier de la donnée » qui en ferait le service de proximité. La cellule repose également sur des référents donnés en laboratoire, dont fait partie Françoise Acquier.

¹⁶¹ OUVRIER LA SCIENCE, 2022. La ministre de l'Enseignement supérieur et de la Recherche, Sylvie Retailleau, inaugure Recherche Data Gouv.

¹⁶² DELODDE, Claudine, *op. cit.*

CONCLUSION

Il paraît difficile au premier abord de conclure sur le sujet, tellement celui-ci apparaît comme un processus dont les contours restent toujours à définir. Après avoir été sensibilisée de nombreuses années à l'accès ouvert, l'équipe est maintenant sensibilisée à l'ouverture des données de la recherche. Cette sensibilisation se caractérise, par rapport à la première, par une institutionnalisation et la mise en place de points de passage obligés. C'est par l'influence de ceux-ci, par exemple l'obligation de rédaction d'un plan de gestion de données, que la notion de « donnée de la recherche » est de plus en plus utilisée et que la question de la gestion se pose clairement.

Il s'agit également de souligner que cette question intéresse de nombreux enquêtés qui parfois ont des difficultés à assurer cette gestion. En effet, la notion qui a surgi tout au long du mémoire est celle de la systématisation. Les enquêtés avaient souvent un système personnel de gestion et, même s'ils avaient conscience des pratiques à effectuer, ils n'avaient parfois ni le temps ni les moyens ni les conditions nécessaires pour les respecter.

Alors même qu'un certain nombre d'enquêtés ont souligné l'aspect personnel de leur système et donc non partageable, un certain nombre de questionnements est commun à l'ensemble de l'équipe, que ce soit au niveau des considérations scientifiques ou des pratiques de recherche en lien avec la gestion des données. Cependant, le lien entre science ouverte et considérations scientifiques ne se fait pas toujours, la science ouverte n'étant que rarement discutée entre les enquêtés et encore moins sur ses considérations idéologiques.

Les questions de systématisation de gestion et d'ouverture des données sont également collectives, et elles l'ont déjà été pour les enregistrements sonores produits lors de recherche dans le cadre de Cartophonies et CRESSOUND. L'étude de ces deux projets ont montré l'importance de personnes médiatrices qui assurent la dynamique des projets et vérifient la qualité des données et de leurs métadonnées. Le degré d'autonomie des chercheurs dans la publication des données continue également de varier selon le profil. De plus, ces deux projets reposent sur la pertinence des données et de leurs métadonnées à un niveau local, tout en donnant deux propositions très différentes.

La question de systématiser pour l'ensemble des données produites, très variées dans l'équipe, se pose, ainsi que les moyens qui seront mis en place. Quelles pratiques collectives ces pratiques personnelles devront-elles être intégrées ? Celles du laboratoire ? Ou par projet, par établissement ? Qui seront les personnes relais ?

L'ouverture des données, par une méthode de qualité qui permet leur réutilisation, demande des moyens importants, sans compter le coût énergétique et les problèmes de stockage. A l'instar de ce que met en avant le rapport « Usage et gouvernance des données »¹⁶³ du comité pour la science ouverte de 2019 et certains enquêtés, la problématique de la pertinence se pose : « **Il y aura nécessairement une allocation des moyens dans la mesure où tout ne pourra pas être bien stocké,**

¹⁶³ CHAMBEFORT, Hélène, HUEBER, Juliette, LEMERCIER, Claire, MAUSSANG, Kenneth et VANET, Anne, 2019. *Usage et gouvernance des données*. [en ligne]. Rapport. Comité pour la science ouverte. [Consulté le 11 août 2022]. Disponible à l'adresse : <https://www.ouvrir.la-science.fr/usage-et-gouvernance-des-donnees/>

bien décrit et bien communiqué : il faut réfléchir à la manière de la réaliser. On gagnerait dès lors à **dire que la science ouverte implique des choix** ».

BIBLIOGRAPHIE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR, DE LA RECHERCHE ET DE L'INNOVATION, 2018. *Plan national pour la science ouverte* (1). [en ligne] [Consulté le 29 juin 2022]. Disponible à l'adresse: https://www.ouvrirlascience.fr/wp-content/uploads/2018/08/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf

A propos du laboratoire, [sans date]. *Laboratoire AAU*. [en ligne]. [Consulté le 29 juin 2022]. Disponible à l'adresse: <https://aau.archi.fr/laboratoire-aau/>

AMIEL, Philippe, FRONTINI, Francesca, LACOUR, Pierre-Yves et ROBIN, Agnès, 2020. Pratiques de gestion des données de la recherche : une nécessaire acculturation des chercheurs aux enjeux de la science ouverte ? *Cahiers Droit, Sciences & Technologies*. [en ligne]. 11 mai 2020. N° 10, pp. 147-168. [Consulté le 31 octobre 2021]. DOI [10.4000/cdst.2061](https://doi.org/10.4000/cdst.2061).

ANR, [sans date]. La science ouverte. *anr.fr*. [en ligne]. [Consulté le 3 juillet 2022]. Disponible à l'adresse: <https://anr.fr/fr/lanr/engagements/la-science-ouverte/>

AUGOYARD, Jean François et TORGUE, Henry, 1995. *A l'écoute de l'environnement: répertoire des effets sonores*. Marseille : Parenthèses. Collection Habitat/ressources. ISBN 978-2-86364-078-4.

BLANC-FERACCI, Caroline, BRAUD, Marion, FORGET, Camille, MICHEL, Eugénie et TOURNIER, Floriane, 2022. « Bibliothèques, éditeurs, libraires, face au mouvement de l'open ». *Bulletin des bibliothèques de France*. [en ligne]. [Consulté le 8 juin 2022]. Disponible à l'adresse: https://bbf.enssib.fr/tour-d-horizon/bibliothèques-editeurs-libraires-face-au-mouvement-de-l-open_70298

BOUCHARD, Aline, 2016. *Éléments de comparaison : archives ouvertes (HAL) et réseaux sociaux académiques (Academia, Researchgate)* [en ligne]. URFIST de Paris. [Consulté le 29 juin 2022]. Disponible à l'adresse: <https://www.enssib.fr/bibliotheque-numerique/documents/67085-elements-de-comparaison-archives-ouvertes-hal-et-reseaux-sociaux-academiques-academia-researchgate.pdf>

RENATIS, 2021. *C@fé Renatis - Un retour d'expérience sur le dépôt de données d'enquête qualitative en sciences sociales sur un entrepôt de données, Datasciencespo (Dataverse)*. [Vidéo]. Canal-U. [en ligne]. [Consulté le 31 octobre 2021]. Disponible à l'adresse: <https://www.canal-u.tv/107849>

CHAMBEFORT, Hélène, HUEBER, Juliette, LEMERCIER, Claire, MAUSSANG, Kenneth et VANET, Anne, 2019. *Usage et gouvernance des données*. [en ligne]. Rapport. Comité pour la science ouverte. [Consulté le 11 août 2022]. Disponible à l'adresse : <https://www.ouvrirlascience.fr/usage-et-gouvernance-des-donnees/>

CNRS, 2019. *Feuille de route du CNRS pour la science ouverte* [en ligne]. CNRS. [Consulté le 29 juin 2022]. Disponible à l'adresse: https://www.science-ouverte.cnrs.fr/wp-content/uploads/2019/11/Plaquette_Science-Ouverte_18112019.pdf

DASSA, Michèle et KOSMOPOULOS, Christine, 2021. Les évolutions transversales de l'information scientifique et technique avec l'apparition du

numérique. *Histoire de la recherche contemporaine* [En ligne], Tome X - n°2 [Consulté le 16 mars 2022.] DOI : <https://doi.org/10.4000/hrc.6097>

DATACTIVIST et COUPERIN, 2022. *De la science ouverte à l'intégrité scientifique : comment le décret n°2021 -1572 bouleverse la conservation et la mise à disposition des données*. [Webinaire en ligne] [notes personnelles].

DELODDE, Claudine, 2022. Les bibliothèques universitaires : un acteur clef dans la sensibilisation et la formation des enseignants-chercheurs à la science ouverte. *DLIS*. [en ligne]. 2022. [Consulté le 20 juillet 2022]. Disponible à l'adresse: <https://dlis.hypotheses.org/6064>

EDWARDS, Paul N., MAYERNIK, Matthew S., BATCHELLER, Archer L., BOWKER, Geoffrey C. et BORGMAN, Christine L., 2011. Science friction: Data, metadata, and collaboration. *Social Studies of Science*. [en ligne]. Vol. 41, n° 5, pp. 667-690. [Consulté le 6 juillet 2022]. DOI [10.1177/0306312711413314](https://doi.org/10.1177/0306312711413314).

INIST-CNRS, 2022. DMP OPIDoR – Le modèle structuré. *DoRANum*. [en ligne] [Consulté le 3 juillet 2022]. Disponible à l'adresse: https://doranum.fr/plan-gestion-donnees-dmp/dmp-opidor-le-modele-structure_10_13143_mzze-kz59/

JACQUEMIN, Bernard, SCHÖPFEL, Joachim, CHAUDIRON, Stéphane et KERGOSE, Eric, 2018. *L'éthique des données de la recherche en sciences humaines et sociales*. [en ligne]. De Boeck Supérieur. [Consulté le 30 octobre 2021]. ISBN 978-2-8073-1577-8. Disponible à l'adresse: <https://www.cairn.info/l-ethique-en-contexte-info-communicationnel--9782807315778-page-71.htm?contenu=article>

JACQUEMIN, Bernard, SCHÖPFEL, Joachim et FABRE, Renaud, 2019. Libre accès et données de recherche. De l'utopie à l'idéal réaliste. *Études de communication. Langages, information, médiations*. [en ligne]. N° 52, pp. 11-26. [Consulté le 30 octobre 2021]. DOI [10.4000/edc.8468](https://doi.org/10.4000/edc.8468).

KLUMP, Jens, 2017. Data as social capital and the gift culture in research. *Data Science Journal*. [en ligne]. 2017. Vol. 16, pp. 1-8. [Consulté le 31 octobre 2021]. DOI [10.5334/dsj-2017-014](https://doi.org/10.5334/dsj-2017-014).

EDITIONS LAROUSSE. Définitions : interopérabilité. *Larousse*. [en ligne]. [Consulté le 29 juin 2022]. Disponible à l'adresse: <https://www.larousse.fr/dictionnaires/francais/interop%C3%A9rabilit%C3%A9/43787>

MANOLA, Théa, 2013. Rapports multisensoriels des habitants à leurs territoires de vie. Retours critiques sur une démarche méthodologique. *Noroi. Environnement, aménagement, société*. [en ligne]. 30 juin 2013. N° 227, pp. 25-42. [Consulté le 20 juillet 2022]. DOI [10.4000/noroi.4649](https://doi.org/10.4000/noroi.4649).

MILLERAND, Florence, 2015. Les imaginaires de la « science 2.0 ». *Communication. Information médias théories pratiques*. [en ligne]. 11 décembre 2015. N° vol. 33/2. [Consulté le 3 novembre 2021]. DOI [10.4000/communication.6070](https://doi.org/10.4000/communication.6070).

OCDE, 2007. *Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics*. [en ligne]. [Consulté le 6 juillet 2022]. Disponible à l'adresse: <https://www.oecd.org/fr/sti/inno/38500823.pdf>

OLIVESI, Stéphane, 2006. *La Communication au travail. Une critique des nouvelles formes de pouvoir dans les entreprises*. 2ème éd. Grenoble : PUG. Communication en plus.

- OUVRIR LA SCIENCE, 2022a. La ministre de l'Enseignement supérieur et de la Recherche, Sylvie Retailleau, inaugure Recherche Data Gouv. *Ouvrir la science*. [en ligne]. [Consulté le 20 juillet 2022]. Disponible à l'adresse: <https://www.ouvrirlascience.fr/la-ministre-de-lenseignement-superieur-et-de-la-recherche-sylvie-retailleau-inaugure-recherche-data-gouv>
- OUVRIR LA SCIENCE, 2022b. Plan d'action pour les revues « diamant ». *Ouvrir la science*. [en ligne]. [Consulté le 29 juin 2022]. Disponible à l'adresse: <https://www.ouvrirlascience.fr/plan-daction-pour-les-revues-diamant>
- PAIN, Marilou, 2016. *Les données de la recherche et leurs entrepôts, de la documentation à la réutilisation : étude de cas pour l'archive HAL*. [en ligne]. Enssib. [Consulté le 6 juillet 2022]. Disponible à l'adresse: https://memsic.ccsd.cnrs.fr/mem_01374509
- PASTEUR, Ceris-Institut, 2022. Bref historique de la science ouverte européenne. *Open science : évolutions, enjeux et pratiques*. [en ligne]. 7 avril 2022. [Consulté le 8 juin 2022]. Disponible à l'adresse: <https://openscience.pasteur.fr/2022/04/07/bref-historique-de-la-science-ouverte-europeenne/>
- Projet scientifique AAU 2022-2026, [sans date]. *Laboratoire AAU*. [en ligne]. [Consulté le 13 juillet 2022]. Disponible à l'adresse: <https://aau.archi.fr/laboratoire-aau/prochain-projet-scientifique-aau-2022-2026/>
- REBOUILLAT, Violaine, 2019. *Ouverture des données de la recherche: de la vision politique aux pratiques des chercheurs*. [en ligne]. Sciences de l'information et de la communication. Conservatoire national des arts et métiers. [Consulté le 6 juillet 2021]. Disponible à l'adresse: <https://tel.archives-ouvertes.fr/tel-02447653/document>
- REBOUILLAT, Violaine, 2021. Le partage des données vu par les chercheurs : une approche par la valeur. *Les Enjeux de l'Information et de la Communication*. [en ligne]. 2021. N° 22/1, pp. 35-53. [Consulté le 30 novembre 2021]. Disponible à l'adresse: <https://lesenjeux.univ-grenoble-alpes.fr/2021/varia/03-le-partage-des-donnees-vu-par-les-chercheurs-une-approche-par-la-valeur/>
- REVELIN, Florence, LEVAIN, Alix, MIGNON, Morgane, NOEL, Marianne, QUEFFELEC, Betty, RAUX, Pascal et SQUIVIDANT, Hervé, 2021. *L'ouverture des matériaux de recherche ethnographiques en question*. [en ligne]. Rapport de recherche. CNRS. [Consulté le 16 novembre 2021]. Rapport d'enquête du projet « Partage et protection des données qualitatives à l'ère du numérique : expériences, enjeux, stratégies ». Disponible à l'adresse: <https://hal.archives-ouvertes.fr/hal-03238067/document>
- ROCHE, Julien, 2022. Science ouverte (origine). In : CAULI, Marie, FAVIER, Laurence et JEANNAS, Jean-Yves, *Dictionnaire du numérique*. Londres : ISTE Editions. pp. 287-290. Sciences, société et nouvelles technologies.
- ROUX, Jean-Michel et TIXIER, Nicolas, 2011. Paroles données, paroles rendues : la fabrique de la ville à l'épreuve des usages. *Vies des villes, trimestriel de l'architecture et de la ville*. [en ligne]. 2011. N° 16, pp. 84. [Consulté le 13 juillet 2022]. Disponible à l'adresse: <https://hal.univ-grenoble-alpes.fr/hal-00995534>
- SCHÖPFEL, Joachim, 2018. *Vers une culture de la donnée en SHS* [en ligne]. Rapport de recherche. Université de Lille. [Consulté le 13 juillet 2022]. Disponible à l'adresse: <https://hal.archives-ouvertes.fr/hal-01846849>

SCHÖPFEL, Joachim, KERGOSIEN, Eric et PROST, Hélène, 2017. `` Pour commencer, pourriez-vous définir « données de la recherche »? ’ ’ Une tentative de réponse. In : *Atelier VADOR : Valorisation et Analyse des Données de la Recherche; INFORSID 2017*. [en ligne]. Toulouse, France. pp. 15. [Consulté le 6 juillet 2022]. Disponible à l’adresse: <https://hal.univ-lille.fr/hal-01530937/document>

Science ouverte, [sans date]. *Laboratoire AAU*. [en ligne]. [Consulté le 3 juillet 2022]. Disponible à l’adresse: <https://aau.archi.fr/laboratoire-aau/science-ouverte/>

SECRÉTARIAT GÉNÉRAL DU CONSEIL, 2022. *Projet de conclusions du Conseil sur l’accélération de la mise en œuvre et de l’impact des politiques et des pratiques de la science ouverte* [en ligne]. Bruxelles : Conseil de l’Union Européenne. [Consulté le 23 mars 2022]. Disponible à l’adresse: <https://data.consilium.europa.eu/doc/document/ST-7390-2022-INIT/fr/pdf>

SERRES, Alexandre, MALINGRE, Marie-Laure, MIGNON, Morgane, PIERRE, Cécile et COLLET, Didier, 2017. *Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l’Université Rennes 2* [en ligne]. Rapport de recherche. Université Rennes 2. [Consulté le 3 novembre 2021]. Disponible à l’adresse: <https://hal.archives-ouvertes.fr/hal-01635186>

Bibliographie réalisée en suivant la norme ISO-690.

ANNEXES

Table des annexes

ANNEXE 1 : GRILLE D’ENTRETIEN POUR LES CHERCHEURS	61
ANNEXE 2 : GRILLE D’ENTRETIEN POUR LES INGENIEURS	62
ANNEXE 3 : ORGANIGRAMME DU LABORATOIRE AMBIANCES, ARCHITECTURES, URBANITES, JUIN 2022	63

ANNEXE 1 : GRILLE D'ENTRETIEN POUR LES CHERCHEURS

PARTIE 1 : PRESENTATION

Pouvez-vous présenter ? Depuis combien de temps travaillez-vous au CRESSON ?

PARTIE 2 : SCIENCE OUVERTE

Avez-vous entendu parler de la science ouverte ? (quand/comment)

Qu'en pensez-vous ?

Qu'en pensent vos collègues ?

Quel impact sur votre travail ?

PARTIE 3 : LES DONNEES DE LA RECHERCHE : GESTION ET OUVERTURE

Utilisez-vous le terme donnée de la recherche ?

Quelles méthodes utilisez-vous dans vos projets de recherche ?

Comment stockez-vous vos données ? Selon une organisation précise ? Les communiquez-vous pendant un projet ou après ? Que faites-vous des données à la fin d'un projet ?

Seriez-vous prêt à partager les résultats sur une plateforme en ligne ?

Avez-vous déjà utilisé les résultats (données, questionnaires...) de quelqu'un d'autre ? (Si oui : comment, où ...)

Avez-vous déjà essayé d'avoir accès aux résultats d'un autre projet ? (ou bien voulu avoir accès)

Est-ce que les enquêtés demandent à connaître les résultats d'une recherche ?

PARTIE 4 : CRESSOUND ET CARTOPHONIES

Avez-vous participé au projet CRESSOUND ou au projet Cartophonies ?

Quels sont leurs objectifs selon vous ?

Avez-vous déjà déposé des enregistrements sonores ? Lesquels et comment ?

ANNEXE 2 : GRILLE D'ENTRETIEN POUR LES INGENIEURS

PARTIE 1 : PRESENTATION

Pouvez-vous présenter ? Depuis combien de temps travaillez-vous au CRESSON ?

PARTIE 2 : SCIENCE OUVERTE

Avez-vous entendu parler de la science ouverte ? (quand/comment)

Qu'en pensez-vous ?

Qu'en pensent vos collègues ?

Quel impact sur votre travail ?

PARTIE 3 : CRESSOUND ET CARTOPHONIES

Avez-vous participé au projet CRESSOUND ?

Pouvez-vous me présenter le projet CRESSOUND ; comment est-il né ?

Comment se passe la publication d'enregistrements sonores ?

Quelles sont les réactions externes à ce projet ?

Avez-vous participé au projet Cartophonie ?

Pouvez-vous me le présenter, sa naissance et sa relation avec le premier projet ?

Quelle pérennité pour ces deux projets ?

ANNEXE 3 : ORGANIGRAMME DU LABORATOIRE AMBIANCES, ARCHITECTURES, URBANITES, JUIN 2022



Unité mixte de recherche 1563

CNRS - Univ. Grenoble Alpes - ENSA Nantes - ENSA Grenoble - ECN

Directrice : Rachel Thomas

Directeurs adjoints : Myriam Servières et Nicolas Tixier

ORGANIGRAMME

01.06.2022

148 Membres

Recherche et enseignement

92 membres

(20 HDR)

46 Permanents

20 ENSAN
7 ENSAG
5 Autres ENSA
5 CNRS
5 ECN
1 ESBANM
1 HETS
1 UGA
1 Conservatoire Cuneo

11 Contractuels

3 ENSAN
3 ENSAG
4 CNRS
1 Nantes Université

35 Associés

CNRS - Centre National de la Recherche Scientifique

ECN - Ecole Centrale de Nantes

ENSAG - Ecole Nationale Supérieure d'Architecture de Grenoble

ENSAN - Ecole Nationale Supérieure d'Architecture de Nantes

ESBANM - Ecole Supérieure des Beaux-Arts de Nantes Métropole

HETS - Haute école de travail social Genève

MC - Ministère de la Culture

UGA - Université Grenoble Alpes

CRENAU Nantes

84 Membres - Directrice : Myriam Servières

Permanents

BOSSE Anne - MCF ENSA Nantes
CHÉREL Emmanuelle - Enseignante ESBANM, HDR
DE LA FOYE Alexandre - PR ENSA Nantes
DEVISME Laurent - PR ENSA Nantes, HDR
DROZD Céline - IR ENSA Nantes
FORT-JACQUES Théo - MCF ENSA Lyon
FRIBOURG Rebecca - MCF Ecole Centrale Nantes
HALGAND Marie-Paule - PR ENSA Nantes
HORSCH Bettina - MCF ENSA Nantes
JOANNE Pascal - MCF ENSA Nantes
LEDUC Thomas - IR CNRS
LESCOP Laurent - PR ENSA Nantes, HDR
MARENNE Christian - PR ENSA Nantes
MARTIN Jean-Yves - MCF Ecole Centrale Nantes
MEUNIER Virginie - MCF ENSA Nantes
MIGUET Francis - PR ENSA Nantes
MOSCONI Léa - MCF ENSA Nantes
NICOLAS Amélie - MCF ENSA Nantes
NORMAND Jean-Marie - MCF Ecole Centrale Nantes, HDR
OLIVARD Pauline - MCF ENSA Nantes
PHILIPPE Anne - MCF ENSA Nantes
REQUENA-RUIZ Ignacio - MCF ENSA Nantes
ROY Elise - MCF ENSA Nantes
SERVIÈRES Myriam - MCF Ecole Centrale Nantes, HDR
SIRET Daniel - IR ENSA Nantes, HDR
SUNER Bruno - MCF ENSA Nantes
SUTTER Yannick - MCF ENSA Montpellier
TOURRE Vincent - MCF Ecole Centrale Nantes
VIOLEAU Jean-Louis - PR ENSA Nantes, HDR

Contractuels

CHESNEL Kevin - Post-doctorant ENSAN
GANGNEUX-KEBE Julie - MCFA ENSA Nantes
L'HER Gwendoline - Post-doctorante ENSAN
SOPHER Hadas - Post-doctorante Nantes Université Projet IDE

Associés

BARBE Frédéric - Chercheur indépendant
BATAILLE Nicolas - Post-doctorant UGA PACTE
BELGACEM Houada - Chercheuse indépendante
BIENVENU Gilles - Chercheur indépendant
BRENU Nathan - Chercheur indépendant
CHRISTOPHE Sidonie - DR IGN, HDR
COLIN Florie - UBO Géoarchitecture
EN-NEJARI Maroua - Chercheuse indépendante
FINGEROVA Markéta - Chercheuse indépendante
GANGLOFF Emmanuelle - UGA PACTE
GIRALDO Gabriel - Chercheur indépendant
GOUSSANOU Rossila - Chercheuse indépendante
HOUEL Nicolas - Chercheur indépendant
JAHANGEER Farrah - ENSA Nantes Mauritius
LAFFAILLE Xenia - Chercheuse indépendante
MARCHADIER Côme - Chercheur indépendant
MOREAU Guillaume - PR IMT Atlantique, HDR
MOULEVRIER Pascale - PR Nantes Université, HDR
SPUR Maxim - Chercheur indépendant
TOUYA Guillaume - DR IGN, HDR
VIGNE Margaux - Chercheuse indépendante

CRESSON Grenoble

64 Membres - Directeur : Nicolas Tixier

Permanents

AUDAS Nathalie - MCF UGA (délégation CNRS)
BALAY Olivier - PR ENSA Lyon, HDR
BALEZ Suzel - MCF ENSA Paris-la-Villette, HDR
BEN SACI Abdelkader - MCF ENSA Grenoble
BONICCO-DONATO Céline - MCF ENSA Grenoble, HDR
BRAYER Laure - MCF ENSA Grenoble
BREVIGLIERI Marc - PR HETS, Genève (Suisse), HDR
CHELKOFF Grégoire - PR ENSA Grenoble, HDR
GAVAZZA Giuseppe - PR Conservatoire de Cuneo (Italie)
KAZIG Rainer - CR CNRS
LIVENEAU Philippe - PR ENSA Grenoble
MANOLA Thia - MCF ENSA Grenoble (délégation CNRS)
PARIS Magali - MCF ENSAP Versailles
POUPIN Perrine - CR CNRS
THIBAUD Jean-Paul - DR CNRS, HDR
THOMAS Rachel - DR CNRS, HDR
TIXIER Nicolas - PR ENSA Grenoble, HDR

Contractuels

CHARLES Louise Anne - Apprentie documentation CNRS
GHELLI Roberta - MCFA ENSA Grenoble
GIRAUD Mathilde - Post-doctorante CNRS Projet PROSECO
HIGGIN Marc - Post-doctorant CNRS Projet SENSIBILIA
MARCHAL Théo - MCFA ENSA Grenoble
MORTELETTE Camille - Post-doctorante CNRS Projet PROSECO
THOLLIERE Pascaline - MCFA ENSA Grenoble

Associés

AMPHOUX Pascal - Chercheur indépendant
BAK Eleonore - ESAL Metz
BOUMOU Abdelhakim - Chercheur indépendant
GAMAL SAID Noha - Université Ain Shams, Le Caire (Egypte)
KIMMEL Laurence - University of New South Wales (Australie), HDR
LAROUCHE Sylvie - CSTB
MASSIANI Lena - Chercheuse indépendante
MASSON Damien - Université de Cergy-Pontoise
MOUSSAWI Nour - Chercheuse indépendante
PAXINOU Evangelia - Chercheuse indépendante
RÉMY Nicolas - Université de Thessalie, Volos (Grèce)
SÉRAUDIE Christophe - Chercheur indépendant
STASSI Vanessa - Ecole Spéciale d'Architecture
VAN HOLLEBEKE Sarah - Chercheuse indépendante

Appui et soutien à la recherche

9 membres

9 Permanents

6 CNRS 2 ENSAN 1 ENSAG

Permanents

BIZIEN Laurence - Chargée d'études documentaires, MC ENSA Nantes
BOUQUET Marine - Gestionnaire, SA MC ENSA Nantes
CHARRIEAU Laurent - Administrateur systèmes, TC CNRS
DOM Véronique - Chargée de communication (Assist. Pré), IE CNRS

Permanents

ACQUIER Françoise - Chargée de ressources documentaires, IE CNRS
ARGOUD David - Administrateur systèmes (Assist. Pré), AI CNRS
DI CHIARO Stéphane - Gestionnaire (Corres. Formation), TC CNRS
MCOISANS jul - Médiateur, IE MC ENSA Grenoble
PICHAT Cédric - Métrologie et expérimentations, IE CNRS

Doctorants

47 membres

(Dont 43 avec financement)

21 ED SHPT Grenoble
12 ED SPI Nantes
13 ED STT Nantes
1 ED DSP Rennes

ATIE Michele - ED SPI Nantes
AYMARD Diane - ED STT Nantes
AYOUBI Rima - ED SPI Nantes
BANIADAM Pooria - ED SPI Nantes
BEAUCAMP Benjamin - ED SPI Nantes
BRILLET Denis - ED STT Nantes
COURNOYER-GENDRON Maude - ED STT Nantes
EL-MIR EL-AYOUBI Rima - ED SPI Nantes
FAYREAU Jean - ED STT Nantes
FOURRIER Nicolas - ED SPI Nantes
GARNIER Lucile - ED STT Nantes
GRISOT Sylvain - ED STT Nantes
GUY Martin - ED SPI Nantes
LAMBERT-AOUZERAT Solai - ED STT Nantes
LU Yang - ED SPI Nantes
MARGUO Petra - ED STT Nantes
MIRANDA SANTOS Marta - ED SPI Nantes
PATRON Lise - ED STT Nantes
PILLOT Angèle - ED SPI Nantes
PLANCHOT Mathilde - ED SPI Nantes
RABIN Thibault - ED STT Nantes
RACHET Clémentin - ED STT Nantes
REDIER Antoine - ED DSP Rennes
ROCHER Yann - ED SPI Nantes
SETEN Sevil - ED STT Nantes
SHANKLAND Stefan - ED STT Nantes

ABOU MRAD Georges - ED SHPT Grenoble
ABU-DAYA Amal - ED SHPT Grenoble
ALVAREZ COLL Nuria - ED SHPT Grenoble
CAMUS Alexia - ED SHPT Grenoble
DE PERTAT Sébastien - ED SHPT Grenoble
DIAB Jinane - ED SHPT Grenoble
ELNESR Maya - ED SHPT Grenoble
GOMEZ Lauren - ED SHPT Grenoble
HADBI Ryma - ED SHPT Grenoble
HAMROUNI Anwar - ED SHPT Grenoble
LAFFONT Charlotte - ED SHPT Grenoble
LOISELLE Mathilde - ED SHPT Grenoble
MRAD Fatma - ED SHPT Grenoble
OUNI Ines - ED SHPT Grenoble
OZTURK Ummuhan - ED SHPT Grenoble
PAVERO Axelle - ED SHPT Grenoble
ROSILIS Elai - ED SHPT Grenoble
SANCHEZ ARIAS Andres - ED SHPT Grenoble
SAULNIER-SIHAN Steven - ED SHPT Grenoble
SCHUMACHER Noa - ED SHPT Grenoble
TOUSSAINT Malys - ED SHPT Grenoble



TABLE DES MATIERES

SIGLES ET ABBREVIATIONS	7
INTRODUCTION.....	9
PARTIE 1 : LE DEVELOPPEMENT DE LA SCIENCE OUVERTE	11
I. Naissance et caractéristiques de la science ouverte	11
1) <i>Origines et définition.....</i>	<i>11</i>
2) <i>Science ouverte et numérique.....</i>	<i>14</i>
II. Présentation de l'étude.....	17
1) <i>Terrain choisi et particularités</i>	<i>17</i>
2) <i>Méthodologie choisie.....</i>	<i>18</i>
III. La science ouverte vue par l'équipe	18
1) <i>L'arrivée de la notion de science ouverte</i>	<i>18</i>
2) <i>Définir la science ouverte</i>	<i>19</i>
3) <i>Adhésion à la science ouverte et limites</i>	<i>20</i>
4) <i>Un sujet discuté dans l'équipe ?</i>	<i>22</i>
PARTIE 2 : DEFINIR LA DONNEE DE LA RECHERCHE : DE LA GESTION AU PARTAGE CONTROLE.....	23
I. Définir les données de recherche et ses enjeux	23
1) <i>Etat des lieux</i>	<i>23</i>
2) <i>La vision des enquêtés sur la notion de donnée</i>	<i>25</i>
II. La gestion des données de la recherche des enquêtés	27
III. Le partage des données	30
1) <i>Fonctionnement du monde scientifique.....</i>	<i>30</i>
2) <i>Des choix personnels.....</i>	<i>31</i>
3) <i>Le point de vue des chercheurs enquêtés</i>	<i>31</i>
<i>Un partage existant, partiel et local.....</i>	<i>31</i>
<i>Les « frictions » des échanges de données.....</i>	<i>32</i>
<i>Un intérêt pour les méthodes et la représentation des données</i>	<i>33</i>
PARTIE 3 : DES CHOIX PERSONNELS A UNE ORGANISATION COLLECTIVE	37
I. Un « nous » qui apparaît	37
1) <i>Les enseignants-chercheurs</i>	<i>37</i>
2) <i>Nous, l'équipe CRESSON</i>	<i>38</i>
<i>L'intérêt pour l'« écriture » des données</i>	<i>38</i>
<i>Sur le terrain : la peur de la surformalisation.....</i>	<i>40</i>
<i>Des données contextuelles et situées</i>	<i>41</i>

L'influence de l'environnement dans lequel s'inscrit la recherche ...	42
II. Les expériences de CRESSOUND, Cartophonies et de la bibliothèque du centre de documentation de l'équipe AAU-CRESSON : acquérir une expertise sur le son	42
1) <i>Naissance du projet CRESSOUND : impulsion par des ingénieurs d'étude</i>	43
2) <i>Naissance du projet Cartophonies : impulsion par des chercheurs</i>	45
3) <i>Fonctionnement et pérennité.....</i>	47
4) <i>Une donnée : des représentations, des projets.....</i>	48
III. La question finale de qui va s'en occuper : le coût du partage (et de la gestion)	48
1) <i>Développer une expertise qui dépasse les enregistrements sonores</i>	48
2) <i>La sensibilisation et l'aide à apporter aux chercheurs.....</i>	49
CONCLUSION	53
BIBLIOGRAPHIE.....	55
ANNEXES.....	59
TABLE DES MATIERES.....	65