



HAL
open science

Les données de la recherche AAU-CRESSON : résultats de l'enquête sur les usages des chercheurs, doctorants et ingénieurs en matière de gestion de données

Léa Mosnier, Françoise Acquier, Véronique Dom

► To cite this version:

Léa Mosnier, Françoise Acquier, Véronique Dom. Les données de la recherche AAU-CRESSON : résultats de l'enquête sur les usages des chercheurs, doctorants et ingénieurs en matière de gestion de données. [0] Centre de recherche sur l'espace sonore et l'environnement urbain. 2019. hal-02462300

HAL Id: hal-02462300

<https://hal.science/hal-02462300>

Submitted on 31 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

AAU cresson
ambiances
architectures
urbanités



Les données de la recherche AAU-CRESSON

Résultats de l'enquête sur les usages des
chercheurs, doctorants et ingénieurs
en matière de gestion de données

Léa Mosnier & Françoise Acquier - Mai 2019

Collaboration éditoriale : Véronique Dom

Table des matières

Contexte	2
La Synthèse	5
L'enquête statistique.....	6
Synthèse des entretiens semi-directifs.....	22
Conclusion.....	24
Tableau des préconisations.....	25
Annexe n°1 : enquête sur les pratiques des membres de l'équipe CRESSON en matière de données de la recherche	26
Annexe n°2 : questionnaire pour les entretiens semi-directifs.....	34
Bibliographie	33

Contexte

L'importance des données de la recherche dans plusieurs domaines est de plus en plus manifeste, notamment par les différentes évolutions engagées par l'État ces dernières années. En effet, dès octobre 2016, la loi pour la République Numérique a apporté des précisions sur le statut des données de la recherche et des données publiques.

Le 4 juillet 2018, la ministre Frédérique Vidal a annoncé le Plan National pour la Science Ouverte, celui-ci comportant un volet dédié aux données de la recherche. Ce plan national a été élaboré à l'automne 2018, il a donné naissance au **Comité pour la Science Ouverte** qui garantit la mise en place d'une politique de soutien à l'ouverture des publications et des données de la recherche¹. Depuis le début de l'année 2019, les principaux organismes financeurs de projets de recherche telle que l'ANR, rendent obligatoire la création de plans de gestion de données dans les projets déposés et financés. La communauté scientifique a donc aujourd'hui le devoir de se préoccuper de la gestion et de la diffusion des données produites au cours de ses travaux de recherche.

Les données de la recherche recouvrent des notions différentes selon les disciplines et sont un aspect du domaine scientifique qui ne cesse d'évoluer. Nous avons choisi de retenir la définition suivante, proposée par Duprat Julie² à l'occasion de l'enquête menée à l'Université Bordeaux-Montaigne auprès de chercheurs en SHS :

« les données de la recherche peuvent être définies simplement par toutes les productions, numériques ou non, collectées et réalisées par les chercheurs en amont de leur travail d'écriture proprement dit (...). Les données de la recherche regroupent donc un ensemble hétéroclite de sources et matériaux de recherche, aussi appelés données primaires, ainsi que toutes les formes de traitement et d'analyses desdites sources, que l'on appelle alors des données dérivées. »

Cette définition est plus proche des pratiques de recherche du CRESSON, elle peut être plus parlante pour les chercheurs que la définition plus générique établie par l'OCDE : *« 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. »*

Afin de respecter les engagements nationaux, l'équipe CRESSON de l'UMR AAU a décidé de se pencher sur la gestion de ses données de recherche et ainsi améliorer leur cycle de vie. La mise en place d'une enquête au sein de la structure a été optée afin de sensibiliser l'équipe pour préparer une future politique de gestion des données. Cette démarche d'enquête a été initiée conjointement au sein des deux équipes de l'UMR. Laurence Bizien, documentaliste de l'équipe CRENAU s'est occupée de formaliser une enquête spécifique aux chercheurs nantais tout en s'inspirant fortement des enquêtes proposées à Lille et Rennes. Afin d'étudier les pratiques initiales des chercheurs du CRESSON et de pouvoir soumettre une politique de gestion des données adaptée au contexte de l'équipe, nous avons ainsi décidé de mettre en place une enquête à la fois statistique et qualitative. Les enquêtes réalisées à Lille et à Rennes, essentiellement statistiques, ont montré certaines limites. En effet, elles mentionnent un fort taux d'abandon en cours de questionnaire par des chercheurs non-familiarisés à ce type de questions. Les entretiens qualitatifs réalisés par l'Université de Bordeaux Montaigne ont permis d'approfondir les questions sur les données et ainsi de prendre plus précisément conscience

¹ Selon le site internet <https://www.ouvrirlascience.fr/presentation-du-comite/> (Consulté le 30 avril 2019)

² J. Duprat, *Les données de la recherche à l'Université Bordeaux Montaigne*, Bordeaux, Université Bordeaux Montaigne, 2019

des enjeux et besoins de tous.

Nous avons donc décidé d'allier les deux types de méthodologie, c'est-à-dire, de proposer à la fois une enquête en ligne relativement statistique afin de dégager les grandes tendances parmi les membres du CRESSON, mais aussi de réaliser quelques entretiens semi-directifs afin d'approfondir certaines questions qui nous paraissaient importantes. Largement inspirée des études effectuées au sein des Universités de Lille, Rennes et Bordeaux Montaigne, cette enquête vise à établir un état des lieux le plus complet possible sur la nature et la place des données dans les travaux des membres de l'équipe, leurs pratiques d'archivage et de valorisation. Nous faisons également référence à l'enquête en cours d'analyse menée par une équipe au sein de l'Université Libre de Bruxelles qui a détaillé les pratiques des enseignants-chercheurs de l'École d'Architecture de la Cambre Horta³.

Les réponses nous permettront également de réfléchir à une offre de services pouvant répondre à leurs besoins énoncés. Cette enquête⁴ a été saisie sur GoogleForms, elle est composée de six rubriques et d'une trentaine de questions ouvertes et fermées. Le questionnaire a été diffusé par mail à l'ensemble de l'équipe CRESSON (59 membres) le 8 mars 2019 et est resté disponible jusqu'au 8 avril 2019.

³ Blondiau Geneviève, Bizien Laurence, Serlet Murielle, Fort Pascal, Acquier Françoise, Bsaiti Badia et Hattet Véronique, « Les outils de gestion des données de la recherche dans les universités belges », Lab & doc, 2019, consulté le 24 mai 2019, URL : <https://labedoc.hypotheses.org/5579>.

⁴ Voir Annexe n°1

Méthodologie

Avant de lancer l'enquête auprès de toute l'équipe nous avons décidé de la tester auprès d'un chercheur de l'équipe. Cet essai nous a permis de récolter un certain nombre de remarques pertinentes afin d'affiner et préciser les questions et choix proposés. Par exemple, l'emploi du terme de « données de recherche » a été trop compliqué à saisir pour le chercheur test. Ce questionnement nous a amené à clarifier les termes utilisés pour chaque question en les adaptant au contexte de l'équipe, à ses méthodes de recherche.

Afin de compléter les données recueillies par le questionnaire, nous avons mené une enquête qualitative auprès de quelques membres de l'équipe représentant chaque groupe : IT, doctorant et chercheur. Nous avons ainsi pu interroger une personne en charge des archives sonores, travaillant avec les chercheurs et sur le terrain, deux enseignants-chercheurs et deux doctorants afin de mener des entretiens semi-directifs⁵ pour préciser certaines questions et engager une réflexion. Nous avons rencontré les personnes pendant un mois pour des entretiens allant de 20 minutes à 1h30 selon l'engagement de la personne, sa motivation.

Nous nous sommes inspirés plus précisément des questions qui ont été posées lors des entretiens menés à l'Université de Bordeaux en ne sélectionnant que celles qui nous paraissaient les plus complémentaires en fonction des questions déjà posées par le questionnaire en ligne.

Nous avons rencontré les personnes pendant un mois pour des entretiens allant de 20 minutes à 1h30 selon l'engagement de la personne, sa motivation.

Nous nous sommes inspirés plus précisément des questions qui ont été posées lors des entretiens menés à l'Université de Bordeaux en ne sélectionnant que celles qui nous paraissaient les plus complémentaires en fonction des questions déjà posées par le questionnaire en ligne.

⁵ Voir Annexe n°2

La Synthèse

Un taux de participation de 30,5 % à l'enquête

Données sources	Les données sources utilisées : les données textuelles sont les plus importantes (94,4 %), puis les entretiens (88,9 %), les enregistrement sonores et images fixes de paysage, d'architecture ou de texte (55,6 %).
Types de données	Plusieurs types de données sont exploitées au CRESSON, ce qui peut complexifier la définition d'un plan de gestion et de conservation des données. On peut tout de même mettre en place des temps de formation afin d'établir une méthode commune de gestion et d'archivage des données de recherche .
Stockage et sauvegarde	Une grande majorité des chercheurs stockent leurs données sur un disque dur externe ou autre dispositif externe (clé USB, CD, DVD etc.) à 88,9%, 72,2 % sur un ordinateur personnel, 66,7 % sur le serveur Thor de l'équipe (un résultat assez encourageant), 61,1 % sur un ordinateur professionnel, 38,9 % sur le cloud privé . Le problème principal est le risque de perte, d'endommagement et du faible niveau de sécurité du dispositif utilisé notamment pour le disque dur externe. On constate plus globalement une absence de politique de conservation, 38,9 % des répondants déclarent sauvegarder quotidiennement leurs données , ou « irrégulièrement en fonction des besoins » ; 28 % toutes les semaines ; 22 % tous les mois.
Durée de conservation	Concernant la durée de conservation de leurs données, 88,9 % des répondants souhaitent qu'elle soit pérenne (pour valider les résultats de recherche, pour mener de nouvelles études ou pour l'enseignement).
Accès	72,2 % des participants déclare ne donner accès à personne à leurs données de recherche . Une majorité de répondant serait plutôt d'accord pour diffuser les matériaux de recherche en libre accès sur le web , 22,2 % ne savent pas, 16,7 % sont tout à fait enclin à les diffuser, mais 16,7 % ne sont pas vraiment d'accord. Aucune réponse formellement négative n'a été enregistrée. Parmi les raisons qui inciteraient à rendre les données de recherche accessibles en libre accès, le souhait de soutenir les valeurs de l'archive ouverte est la principale motivation (77,8 %) .
Types de dépôts	Le support à privilégier pour un dépôt en libre accès semble être celui des archives ouvertes de type HAL pour 77,8 % des répondants . 66,7 % souhaitent privilégier le site de l'UMR, 44,4 % un blog ou un site personnel et enfin 27,8 % souhaitent privilégier les serveurs institutionnels tel qu'Huma-Num. Nous devons donc penser à communiquer davantage autour des entrepôts existants et à suivre la mise en œuvre d'entrepôts universitaires. Parmi les raisons pour ne pas déposer, en tête nous retrouvons les freins juridiques liés à la nature des données (données personnelles, sous droit etc.) avec 61,1 % . Le manque de temps serait également une des raisons majeures empêchant les chercheurs de déposer (55,6 % au CRESSON), puis le manque de connaissance dans la description des données et les modalités de dépôt pour 44,4 % .
Services attendus	La plupart des services proposés sont demandés : en tête de liste des conseils d'ordre général pour gérer les données, des conseils sur des questions techniques, des conseils relatifs à la publication, des conseils d'ordre juridique, etc.

L'enquête statistique

Le questionnaire

Nous avons ainsi formé notre questionnaire en six rubriques principales, chacune comportant des questions obligatoires et facultatives :

	Questions obligatoires	Questions facultatives	Total
1. Les champs disciplinaires de pratique des répondants et leur statut	4	0	4
2. La typologie des données (sources, matériaux de terrain collectés et issues des analyses)	3	1	4
3. Les pratiques de stockage et d'archivage des données	6	1	7
4. Les pratiques de partage et de diffusion des données	7	1	8
5. Les besoins et attentes pour les données de recherche	5	1	6
6. L'intérêt pour un entretien et les résultats de l'enquête	1	2	3
Total	26	6	32

Figure 2 : Structuration du questionnaire

Comme en témoigne le tableau ci-dessus, nous avons préféré proposer une majorité de questions obligatoires afin de récolter un maximum de réponses. Quelques questions sont restées facultatives car elles avaient pour but de préciser certaines réponses sans pour autant empêcher les répondants de poursuivre le questionnaire.

Sur les 32 questions, 12 d'entre elles permettent de répondre via une liste de choix et 3 questions proposent de répondre dans un champ en texte libre. Le temps nécessaire pour répondre au questionnaire est de 10 à 15 minutes environ.

L'échantillon collecté

En nous adressant à 59 membres de l'équipe nous avons pu observer un taux de réponse relativement bas avec **18 réponses** soit d'un taux de **30,5 %** après un mois d'ouverture au questionnaire et de relances régulières. Il est intéressant de constater que l'enquête, bien que présentée antérieurement à l'ensemble de l'équipe lors d'une assemblée générale n'a pas attiré l'attention de l'ensemble des chercheurs du CRESSON. Les membres de l'équipe CRESSON ne se sentent pas réellement impliqués par la question de la gestion des données. Il faut tout de même noter ici que la moyenne du taux de réponse des enquêtes en ligne est de 20 à 30 %, nous pouvons donc considérer que nous nous situons à peine au-dessus de la moyenne. Pour information, les taux de réponse recueillis à Rennes étaient de 28% et de 15 % pour Lille et l'Université Libre Bruxelles.

Les statuts

Les enseignants-chercheurs et les doctorants ont été les plus nombreux à répondre à l'enquête (7 personnes pour chaque corps), suivis des chercheurs CNRS (2 personnes).

Nous avons laissé la possibilité aux répondants d'ajouter le statut qui leur convenait avec la case « autre ». Une personne se définit comme artiste, docteur en architecture, professeur ; un ingénieur d'étude comme médiaticien (ITA).

L'âge des répondants

Trois tranches d'âge ont été proposées aux personnes interrogées. La majorité des répondants se situe entre 30 et 50 ans (**44,4%**), suivi par les moins de 30 ans (**38,9 %**) et 3 répondants ont plus de 50 ans (**16,7 %**).

Ces chiffres sont similaires à ceux du CRENAU.

Les responsabilités

Les répondants avaient la possibilité de cocher plusieurs responsabilités :

- 12 personnes enseignants ;
- 11 qui travaillent dans la recherche ;
- 4 exerçant un encadrement de thèse ;
- 3 directeurs de revues ou faisant partie d'un comité de rédaction ;
- 6 répondants en projet de thèse ;
- 1 personne déclarant exercer un développement de projet.

Les disciplines

6 disciplines étaient proposées selon le contexte des membres de l'équipe CRESSON :

- 10 personnes ont répondu exercer en architecture ;
- 9 en urbanisme ; 3 en géographie ;
- 6 en ingénierie (acoustique etc.) ;
- 5 en sociologie ;
- 2 en philosophie ;

- 2 répondants ont également ajouté exercer dans l'art notamment avec le « *son comme matériel plastique et de construction* » ou encore une autre discipline telle que la « *biophysique* » (thèse en cours).

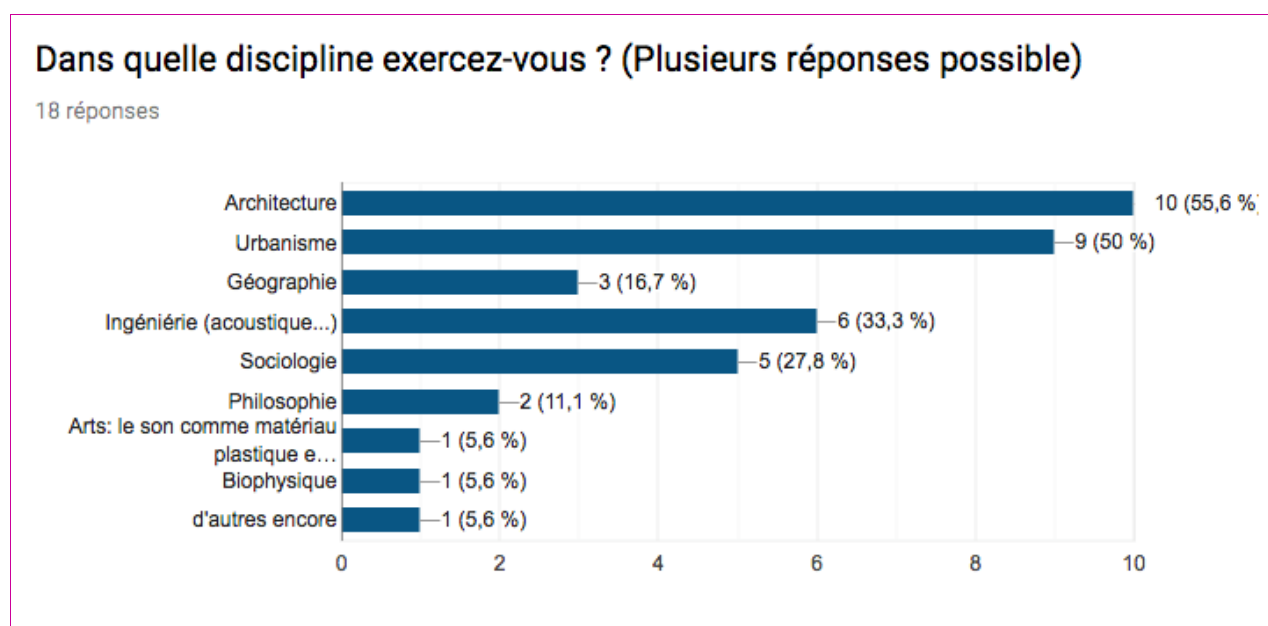


Figure 1 : Graphique des disciplines exercées par les répondants

La typologie des données (collectées et produites)

On peut noter deux grands groupes de données :

17 répondants déclarent réaliser leurs recherches à partir de textes (94,4 %) et 16 d'entre eux avec des entretiens (88,9 %).

Ces deux premiers chiffres sont à l'image du CRESSON puisque ses méthodologies de recherche s'appuient majoritairement sur ces deux types de matériaux.

Viennent ensuite pour 10 répondants les « *Vidéos* », « *Enregistrements sonores* » et « *Images fixes de paysage, d'architecture, de texte etc.* » (55,6 % pour chaque).

9 répondants travaillent à partir d'archives et 7 d'entre eux à partir de mesures (sonomètre, luxmètre etc.)

Enfin, une personne déclare travailler à partir d'exploration d'environnements tempérés, une autre à partir de la presse et de documents de communication de la ville et un dernier participant, à partir de la fabrication de prototypes.

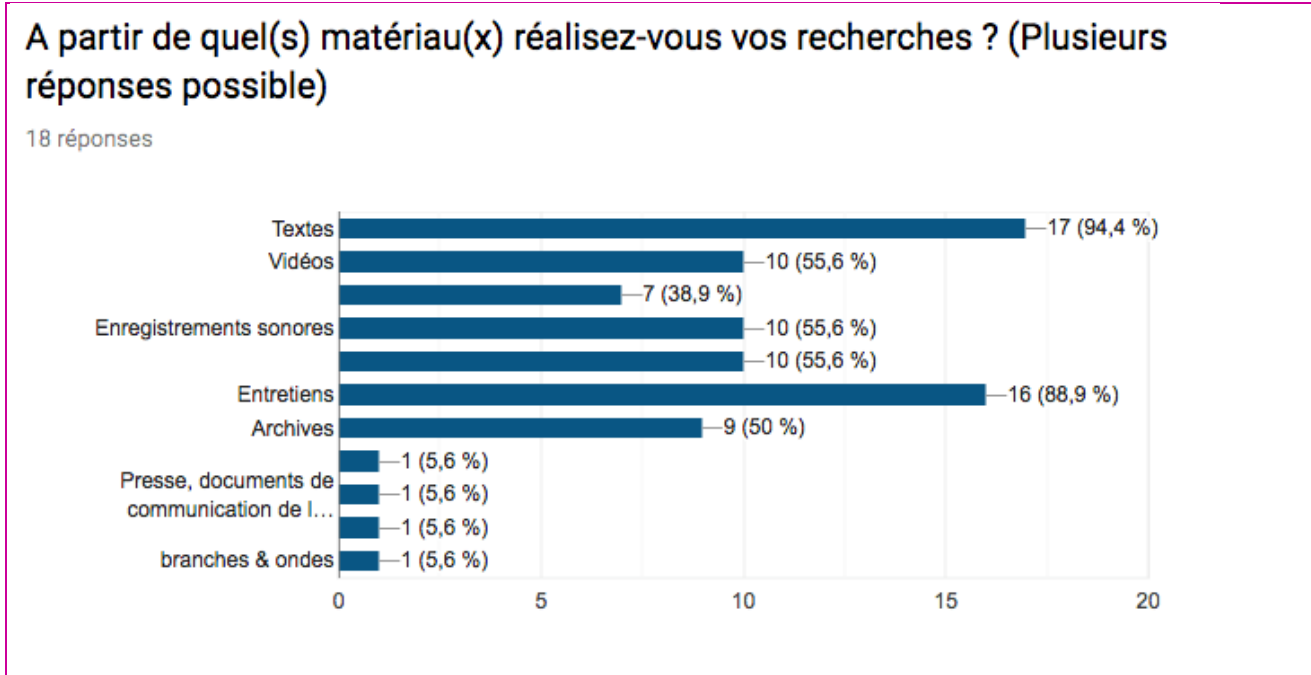
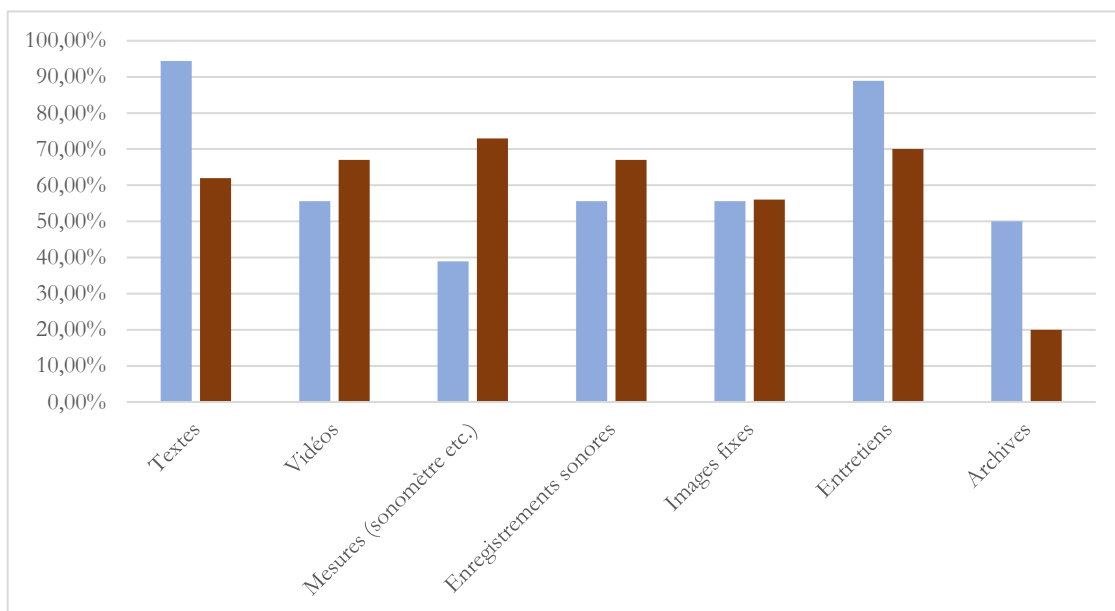


Figure 2 : Graphique des matériaux utilisés pour les recherches

Au CRENAU les **données d'observation** (remontées de terrain, données de capteur) représentent la source la plus importante de données (73 %), suivie par les **enquêtes** et **entretiens** (70%). Les données textuelles ne sont représentées qu'à 62 % et les images à 56 %.



Équipe CRESSON Équipe CRENAU

Figure 3 : Les données sources

Les entretiens issus de l'enquête qualitative ont permis de consolider les résultats pour cette question. En effet, nous constatons que les données collectées et produites sont assez diverses tels que les enregistrements sonores, les photographies, les archives ou encore les textes. Dans les deux cas, les données orales sont fortement représentées.

Le support matériel :

La majorité des matériaux exploités sont **nativement numériques** (photo, audio, vidéo, carte, plan etc.) à **55,6 %** contre **44,4 % non numériques** tel que les relevés de terrain, les notes manuscrites, les textes imprimés, les archives etc.). La différence entre les deux types de support n'est pas très importante.

La question suivante proposait donc de savoir si les matériaux non numériques utilisés étaient ensuite numérisés : **76,5 % ont répondu oui. (80 % pour le CRENAU avec 20 répondants)**

Les données produites :

A la dernière question de cette rubrique, il était proposé aux répondants de cocher plusieurs choix de données produites, en analysant leurs sources (données dérivées).

Une majorité, **13 d'entre eux, disent générer des prises de notes informatiques (72,2 %)**, 11 déclarent produire des **tableaux, schémas** etc. et également des **productions audio-visuelles** ce qui représente **61,1 % des réponses**. 10 répondants signalent produire des **notes manuscrites** (carnets de terrain etc.) (**55,6 %**) et 5 d'entre eux des **bases de données**. Enfin, 3 produisent des **logiciels (16,7 %)** ; 4 des **fabrications matérielles (22,2 %)** ; 1 personne déclare produire des « *œuvres (design d'espace, design d'expériences, visualisations sonores et/ou multi-sensorielles)* » et une autre des « *itinéraires et des entretiens retranscrits et analysés* ».

Pratiques de stockage et d'archivage des données

A la question « Où stockez-vous vos données ? » plusieurs réponses étaient possibles.

On constate que les répondants, pour une grande majorité d'entre eux (**88,9 %**), **stockent leurs données sur un disque dur externe ou sur d'autres dispositifs** (clé USB, CD, DVD etc.) puis, :

- **72,2 %** sur un **ordinateur personnel** ;
- **66,7 %** en réseau sur le **serveur du laboratoire** ;
- **61,1 %** sur un **ordinateur professionnel** ;
- **38,9 %** sur une **plateforme privée** dans le cloud (Google Drive, Dropbox etc.).
- **11,1 %** en réseau sur le **serveur de l'école d'architecture** de Grenoble,
- **11,1 %** sur un **serveur externe** d'une autre institution.

En revanche, une seule personne déclare stocker sur un serveur d'une TGIR et une autre sur un site internet.

Ces résultats sont relativement similaires à ceux de l'enquête réalisée par le CRENAU. En effet, les résultats collectés indiquent que 78 % des répondants utilisent un disque dur externe comme source principal de stockage. Une autre majorité privilégie le stockage local sur leur ordinateur personnel tout comme au CRESSON pour 56 % d'entre eux ou sur leur ordinateur professionnel à 54 %.

Au CRENAU, seuls 18 % des répondants utilisent le serveur de leur laboratoire. De plus, 51 % de l'équipe stocke dans le cloud et 13 % ont des données sur le serveur d'une autre institution.

Nous pouvons donc constater que l'usage de périphérique de stockage externe est majoritaire au sein de l'UMR alors qu'il est peu fiable.

Lors de l'enquête qualitative, la question « *Rencontrez-vous des difficultés particulières pour gérer vos données ? Est-ce un problème récurrent ? Pouvez-vous décrire la dernière fois où vous avez rencontré un problème dans la gestion de vos données ?* », certains répondants expliquent qu'ils rencontrent souvent des problèmes techniques (performance des PC pour traiter les fichiers très lourds telles que les photographies aériennes par exemple).

Le stockage sur un disque dur externe entraîne également des risques de perte ou d'endommagement des données. Cette dernière constatation s'appuie sur le témoignage de l'un des répondants qui s'est rendu compte récemment, de la perte d'un disque dur sur lequel était stocké beaucoup d'informations liées à une recherche ancienne. Cette dernière observation nous conforte donc dans l'idée que l'usage du disque dur externe n'est pas assez fiable comme espace de stockage et qu'il **pourrait être intéressant de sensibiliser les chercheurs à utiliser davantage le serveur du laboratoire mis à leur disposition.**

L'espace de stockage :

6 propositions étaient faites à propos de l'espace de stockage nécessaire pour les données :

- 33,3 % des répondants estiment générer entre **20 Go et 100 Go** de données ;
- 22,2 % estiment utiliser **100 Go à 1 To** d'espace ;
- 22,2 % à plus d'1 To.

La fourchette de 1 Go à 20 Go n'a pas été sélectionnée mais **16,7 % des répondants indiquent ne pas connaître l'espace de stockage nécessaire pour leurs données.**

Au CRENAU, nous pouvons remarquer que 45 % des répondants utilisent majoritairement 100 Go à 1 To et entre 1 et 20 Go pour 37 % d'entre eux. Nous pouvons également comparer ici les résultats avec ceux de la faculté d'architecture de La Cambre-Horta qui déclare utiliser entre 100 Go et 1 To pour 45 % d'entre eux et entre 50 Go et 100 Go pour 22 %.

A combien estimez-vous l'espace nécessaire au stockage de vos données ? (Une seule réponse)

18 réponses

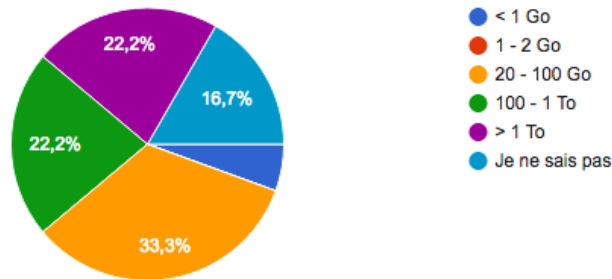


Figure 4 : Graphique secteur sur l'espace de stockage nécessaire estimé

Fréquence des copies de sauvegarde :

38,9 % des répondants déclarent **sauvegarder quotidiennement leurs données**, ou « *irrégulièrement en fonction des besoins* ».

27,8 % réalisent des **sauvegardes toutes les semaines** contre **22,2 % tous les mois**. Cependant, on ne trouve pas au CRESSON de pratique de sauvegarde tous les trimestres ou tous les ans.

Enfin, une personne déclare ne faire aucune copie de sauvegarde.

Comparatif entre le CRENAU et le CRESSON :

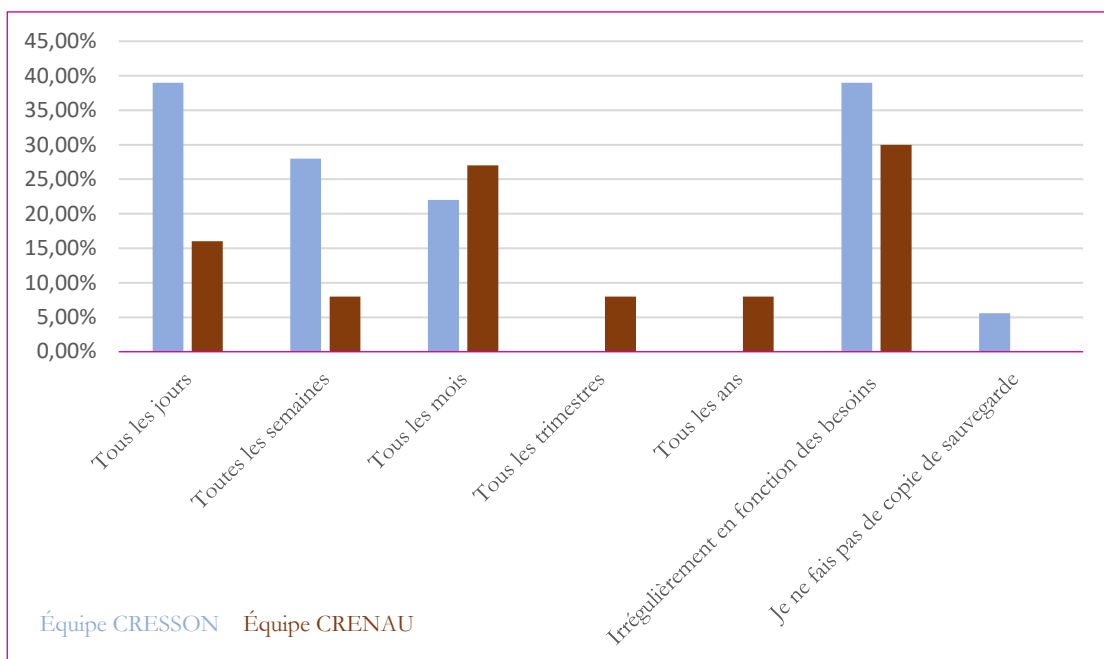


Figure 5 : La fréquence des copies de sauvegarde

Il est intéressant de constater qu'au CRENAU, seuls 16 % des répondants sauvegardent quotidiennement leurs données, et 30 % le font de manière irrégulière en fonction des besoins. Ils sont cependant 8 % à déclarer réaliser des copies tous les ans.

Comme indiqué dans l'enquête du CRENAU, le terme « sauvegarde » peut également être interprété différemment en fonction des répondants : enregistrements de mise à jour (sauvegarde quotidienne), à différencier des pratiques d'archivage des dossiers de manière plus espacée, etc.

Ces résultats mettent en lumière un besoin de clarification de la politique de conservation au sein de l'UMR.

Durée de conservation :

88,9 % des répondants **souhaitent que les données soient conservées et préservées à long terme au-delà du projet** (pour valider les résultats de recherche, pour mener de nouvelles études ou pour l'enseignement), contre 11,1 % qui souhaitent les conserver seulement pour la durée du projet.

Cette question soulève quelques interrogations : il serait intéressant de définir la durée que couvre le long terme. Comment définir cette durée de conservation ? Et ensuite, comment garantir l'accès aux données numériques notamment ? **S'engager dans un processus de conservation à long terme avec le CINES par exemple est vraiment complexe, il serait utile de clarifier le travail à mettre en oeuvre pour une procédure de conservation à long terme, notamment par le biais de formation collective.**

Combien de temps les données sont-elles/devraient être conservées et préservées ? (Une seule réponse possible)

18 réponses



Figure 6 : Graphique secteur du temps de conservation et de préservation estimé pour les données

Informations supplémentaires :

- **66,7 %** des répondants **ajoutent des informations scientifiques à leurs données** (l'origine des données, le projet de recherche, des mots-clés etc.) ;
- **61,1 %** rajoutent des **informations administratives** (auteur, date de création).

Nous pouvons remarquer que **55,6 %** des répondants n'attribuent **aucune information** supplémentaire et une minorité de **22,2 %** ajoutent des **informations techniques** (formats des fichiers, taille, logiciels utilisés etc.).

La question suivante, en lien avec cette dernière, n'était pas obligatoire, 10 personnes sur 18 ont apporté une réponse. Nous pouvons noter qu'une majorité attribue les informations supplémentaires sous forme de **fiches**, d'autres **dans le nom du fichier** ou sous la forme de **tableur**.

Le serveur Thor :

66,7 % des répondants **connaissent le serveur « Thor »** mis à disposition par le CRESSON et les possibilités de sauvegarde de cet outil. 6 personnes sur 18 ont tout de même répondu ne pas le connaître (33,3 %).

Cet écart est peut-être dû au manque d'information sur l'utilisation de cet outil. Les personnes ayant répondu non, sont très rarement dans les locaux du CRESSON. Dans tous les cas, nous pouvons réfléchir à mettre ou remettre en place de courtes réunions pour informer les membres du CRESSON sur la présence de ce serveur, ou souligner particulièrement sa présence dans le livret d'accueil diffusé auprès des nouveaux entrants.

Les pratiques de partage et de diffusion des données (sources, matériaux de terrain collectés et produits)

Cette rubrique a pour but de comprendre le point de vue des membres de l'équipe sur le partage des données notamment en Open Access, une notion inscrite dans le statut des données publiques mais limitée par les questions juridiques et éthiques.

Accéder aux données :

A la question « qui, en dehors de vous, peut accéder à vos données ? »

- **72,2 %** des participants ont répondu « *personne* », une majorité que l'on retrouve dans les résultats au CRENAU à 70 % ;
- L'accès est autorisé **pour les chercheurs du même groupe de recherche à 50 %** ;
- **22,2 %** indiquent **donner l'accès à l'équipe** et à « *tout le monde* » ;
- **11,1 %** (2 répondants) donnent accès à leurs données aux **chercheurs de la discipline** ;
- **5,6 %** à l'**ENSA Grenoble**.

Comme au CRENAU, les données au CRESSON sont **majoritairement inaccessibles**.

Partager les données :

Nous avons voulu connaître par le biais de cette question le(s) moyen(s) utilisé(s) par les chercheurs pour partager leurs données avec d'autres collègues et/ou leur groupe de recherche.

- Une majorité à **88,9 %** annonce **partager leurs données par le biais de mails**
- **66,7 %** indiquent les transmettre en **présentiel**
- **61,1 %** par les **plateformes privées** en cloud tel que Dropbox, GoogleDrive etc.
- **22,2 %** par les **plateformes institutionnelles** (ShareDocs, Thor etc.)
- **16,7 %** par les **réseaux sociaux** tel que Facebook par exemple
- **5,6 %** par le biais de **Workshops, conférences et séminaires**

Nous pouvons donc en déduire qu'il serait intéressant et important, de présenter la plateforme Hum-Num aux chercheurs, afin de leur faire découvrir l'existence de cet outil, tout comme les autres entrepôts de données.

Publier les données en libre accès sur le web au moins une fois :

55,6 % des répondants ont déjà mis en ligne des vidéos, des sons, des photographies en dehors du texte de l'article. 16,7 % ne l'ont pas encore fait mais en ont l'intention ou, ne connaissent pas cette possibilité. Enfin, 11,1 % n'ont pas publié leurs données en libre accès et déclarent ne pas souhaiter le faire.

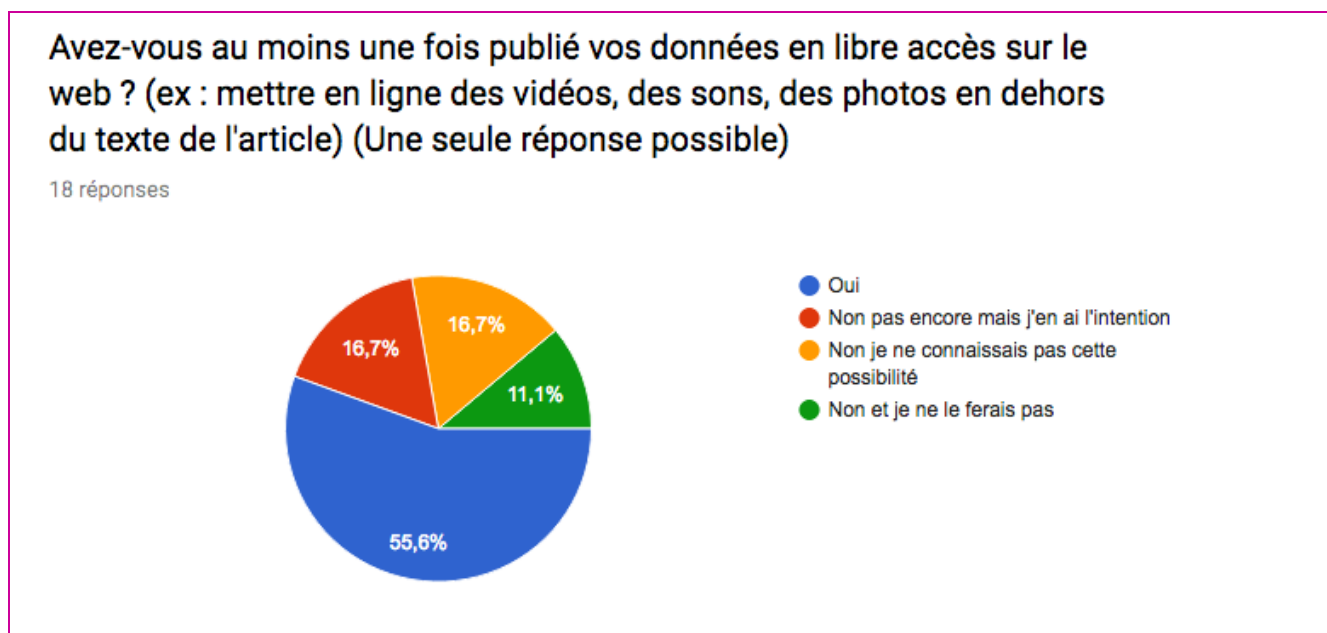


Figure 7 : Graphique secteur pour la publication des données en libre accès sur le web au moins une fois

L'utilisation des matériaux provenant d'autres chercheurs :

44,4 % des répondants ont déjà utilisé au moins une fois des matériaux provenant d'autre(s) chercheur(s). Mais 27,8 % indiquent ne pas l'avoir encore fait mais en avoir l'intention ou encore. Le même pourcentage déclare ne pas l'avoir fait et ne pas souhaiter le faire. Les entretiens semi-directifs ont permis de montrer que l'utilisation des matériaux d'autres chercheurs sont le plus souvent des ouvrages et des articles publiés.

Diffuser les matériaux en libre accès sur le web :

Cette question cherche à définir la volonté des membres de l'équipe à diffuser leurs données en libre accès.

- 38,9 % des répondants sont **plutôt d'accord** ;
- 22,2 % **ne savent pas** ;
- 16,7 % sont **tout à fait enclin** à diffuser leurs données ;
- 16,7 % ne sont **pas vraiment d'accord** pour diffuser ;
- 5,6 % sont **plutôt d'accord, mais seulement après un certain délai**.

L'étude au CRENAU a révélé une position majoritairement positive à l'idée de diffuser les données de recherche en libre accès avec 29,7 % pour la réponse « oui après un certain délai » et 27 % à « plutôt oui ». Nous pouvons noter que 16,7 % de CRESSON ne sont pas vraiment d'accord.

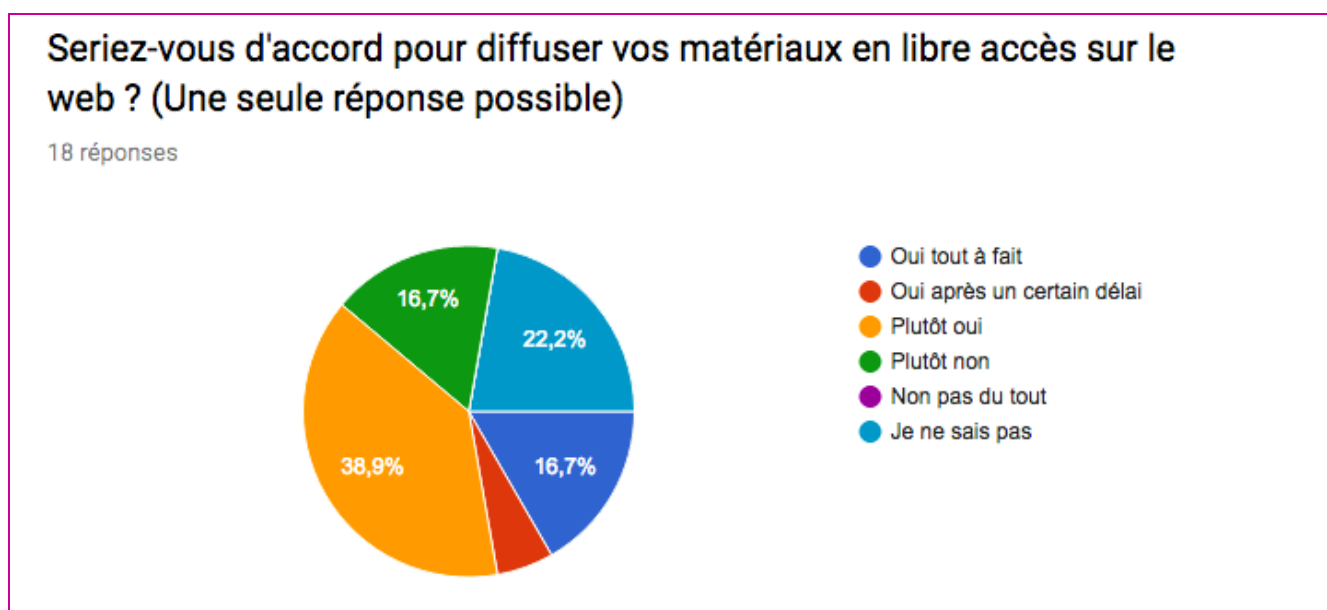


Figure 8 : Graphique secteur pour la diffusion en libre accès des données

Motivation à rendre les matériaux de recherche accessibles en libre accès :

Le souhait de soutenir **les valeurs de l'archive ouverte est la principale motivation des répondants (77,8 %)**. Puis, suivent avec **50 %** chacun, l'occasion de créer de nouveaux contacts, de nouvelles coopérations et la motivation d'avoir une meilleure visibilité de leurs travaux. Enfin, la reconnaissance de la communauté scientifique se place à **38,9 %** et l'obligation faite par le financeur de la recherche (Horizon 2020, ANR etc.) a été indiquée par **22,2 %** des répondants.

Au CRENAU, la motivation principale reste le fait de voir **se développer d'autres recherches** à partir des jeux de données à 63 % ; l'occasion d'établir de nouveaux contacts à 54 % tandis que l'adhésion aux valeurs des archives ouvertes n'atteint que 27 %. « *La reconnaissance de la communauté scientifique* » est un enjeu commun pour les deux équipes (45 % pour le CRENAU et 38,9 % pour le CRESSON.)

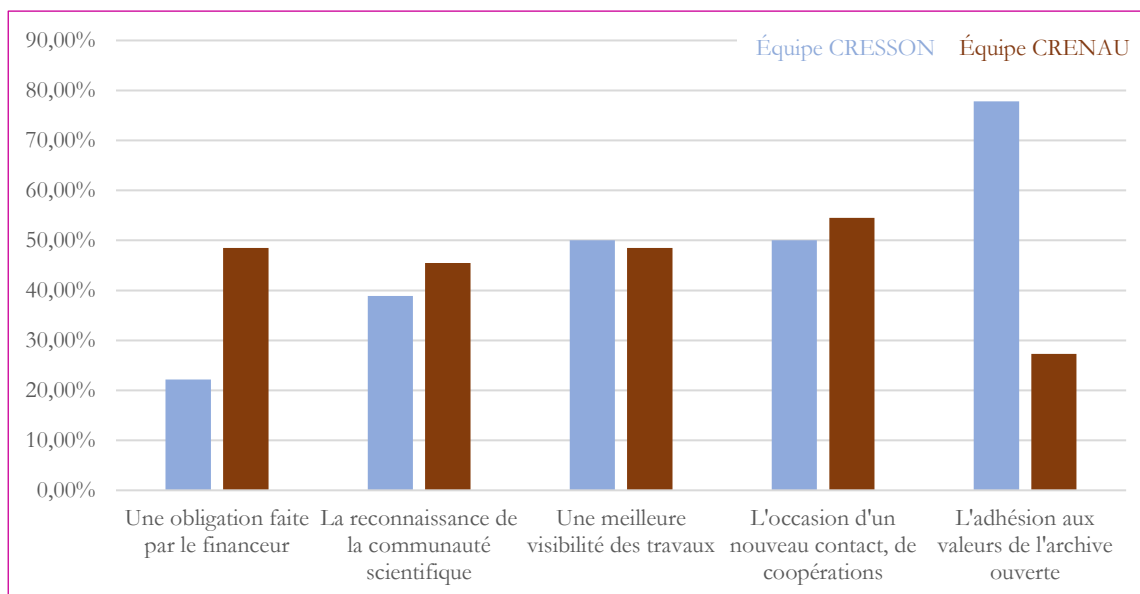


Figure 9 : Les raisons qui inciteraient à publier les données en Open Access

Les freins à rendre accessible les données :

Parmi les raisons pour ne pas déposer, nous retrouvons en tête les **freins juridiques liés à la nature des données** (données personnelles, sous droit etc.) avec **61,1 %**. Une majorité qui se démarque également au CRENAU avec 64 % de répondants pour cette raison.

Ensuite, le **manque de temps** serait également une des raisons majeures empêchant les chercheurs à déposer (55,6 % CRESSON, 51% CRENAU), puis le **manque de connaissance** dans la description des données et leurs **modalités de dépôt** pour **44,4 %**. Enfin, nous retrouvons les causes telles que :

- Le manque de compétence dans le dépôt à **38,9 %**
- Pour des raisons scientifiques pour **33,3 %**
- La crainte du plagiat à **27,8 %**
- Pour des raisons personnelles (mes données m'appartiennent) à **11,1 %**
- Aucune des raisons présentes pour une personne (**5,6 %**).

Une personne a rajouté dans la case « Autre » « cela dépend du type de donnée » (5,6 %) car effectivement les données peuvent être partageables à la condition de respecter le cadre juridique lié aux données personnelles et sensibles.

Le pourcentage concernant la **crainte du plagiat** est similaire à celui du CRENAU (27 %) toute comme l'invocation des raisons plus personnelles avec 5,4 %.

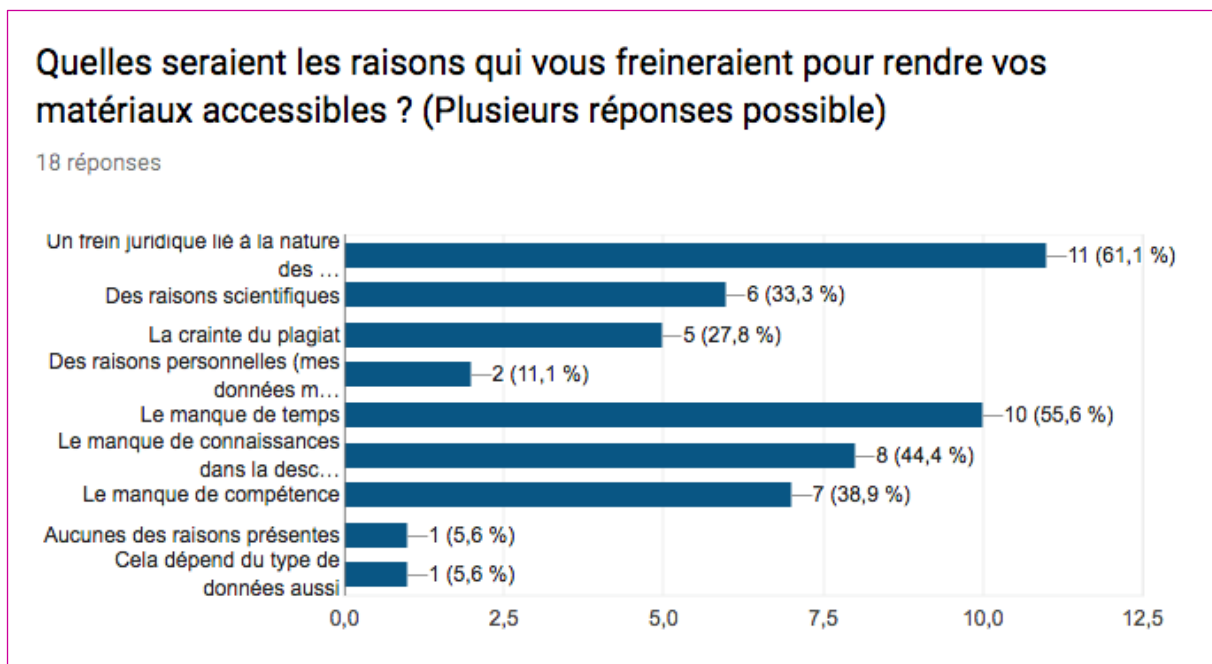


Figure 10 : Les raisons qui freineraient les chercheurs du CRESSON à rendre leurs données publiques

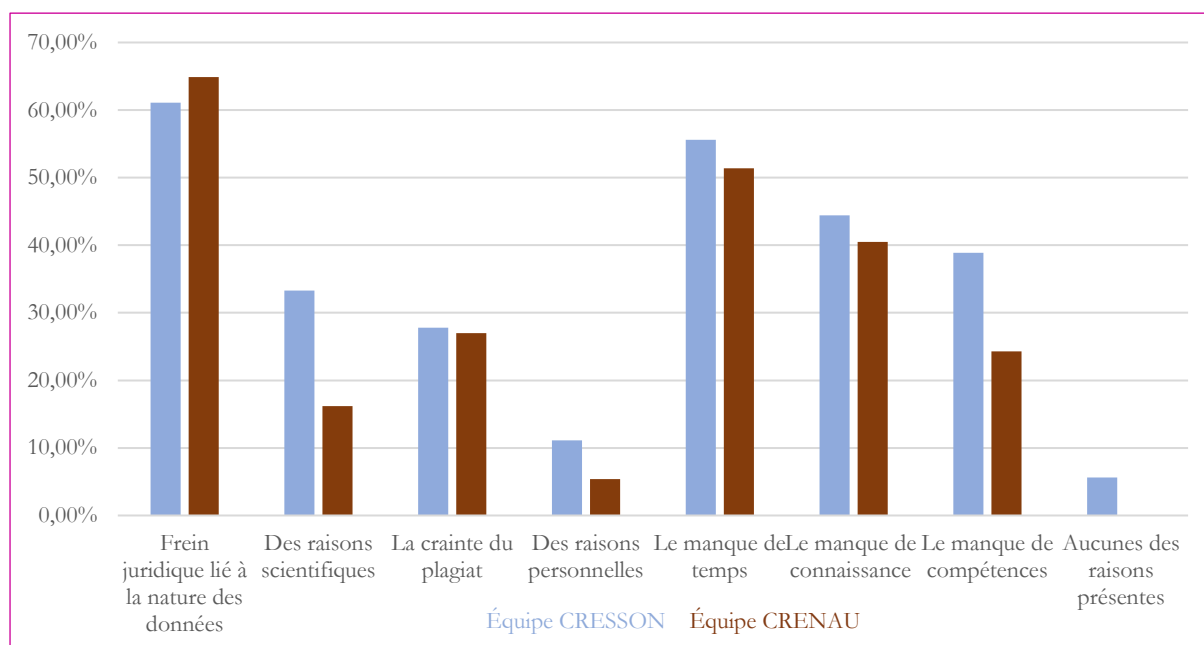


Figure 11 : Les freins empêchant les chercheurs de rendre accessible leurs données

Préférences par rapport au dispositif :

Le support à privilégier pour un dépôt en libre accès semble être celui des **archives ouvertes de type HAL** pour **77,8 %** des répondants, tout comme au CRENAU qui indique 57 % de répondants pour cette proposition ; ensuite le **site de l'UMR** semble être l'un des supports à privilégier avec **66,7%** contre 24 % au CRENAU. De plus, l'option « **un blog ou un site personnel** » collecte **44,4 %** de répondants. Un **serveur institutionnel** tel qu'Huma-Num ou du laboratoire pourrait être un support à privilégier pour **27,8 %** des membres du CRESSON.

- **11,1 %** indiquent vouloir **favoriser des plateformes spécifiques** dans l'archivage et la diffusion des données de recherche (Zenodo etc.), alors que le CRENAU enregistre 36 % des répondants pour cette option ;
- **11,1 %** pour un **réseau social académique** tels que Academia, ResearchData ... ;
- **5,6 %** indiquent le **data paper**.

Nous pouvons donc penser, qu'il est nécessaire de communiquer autour des entrepôts de données existants, sur leurs modalités d'accès et les évolutions des entrepôts universitaires naissant.

Les besoins et attentes pour les données de recherche

Cette dernière rubrique permet de faire ressortir les connaissances et besoins des membres du CRESSON en matière de gestion de données.

Le PGD :

A la question : « *Savez-vous ce qu'est un Plan de Gestion de Données (PGD) ou dit Data Management Plan (DMP) ?* », une grande majorité **a répondu non (77,8 %)**, (oui 22,2 %).

Cette même proportion se retrouve au CRENAU avec 70 % de non pour le DMP.

La question suivante permettait de savoir si les membres de l'équipe connaissaient les enjeux derrière ce document administratif, devenu obligatoire pour le dépôt d'un projet ANR ou H2020. Cette question collecte une **majorité de réponses négatives pour 61,1 %**. On peut néanmoins relever que 38,9 % avaient connaissance de ces évolutions.

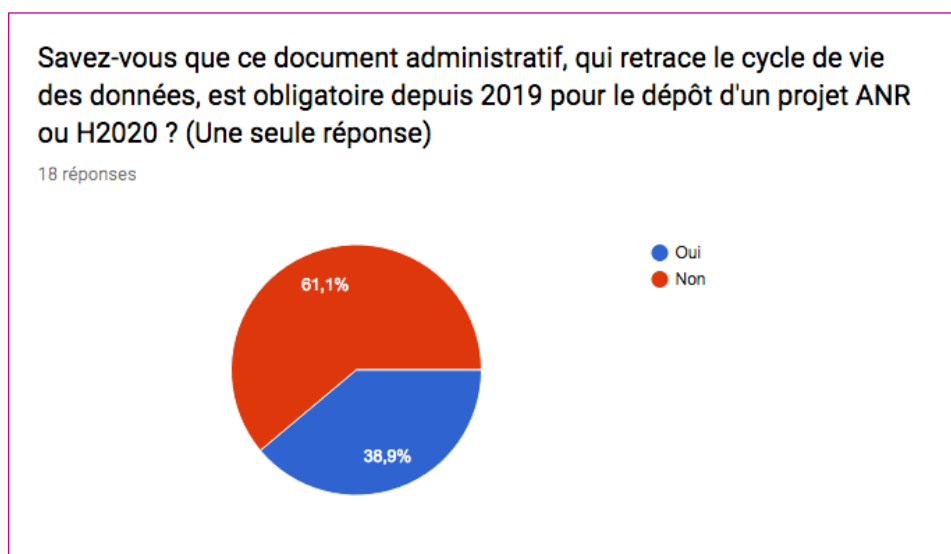


Figure 12 : Graphique secteur à propos du DMP obligatoire pour les projets ANR et H2020

Ces derniers chiffres sont assez représentatifs de la méconnaissance générale des législations en vigueur au sein des laboratoires et les entretiens menés soulignent cette observation. En effet, la quasi-totalité des répondants a indiqué ne pas connaître les nouveautés sur les données de la recherche et ne pas vraiment s'y intéresser. La phase administrative que requiert un PGD serait, selon eux, trop longue et laborieuse, en un mot : décourageante. Il est vrai que l'instauration d'un PGD implique plusieurs modifications d'usage et ce sont des modifications que beaucoup d'entre eux ne souhaitent pas réaliser. Cependant, plusieurs personnes interrogées ont émis l'hypothèse que s'ils embauchent une personne compétente dans le domaine pour s'occuper de cette charge administrative, alors la conception d'un PGD leur semblerait plus réalisable. Enfin, une personne a expliqué qu'il pourrait être intéressant de mettre en place un PGD pour les thèses afin d'exposer précisément la méthodologie d'enquête élaborée et utilisée sur le terrain. Nous pouvons noter ici que les doctorants en Angleterre, à Édimbourg⁶ notamment, sont fortement encouragés à rédiger un PGD pour partager les données collectées, exploitées et produites dans le cadre de leur thèse.

Il pourrait ainsi être intéressant de s'en inspirer afin de rendre les thèses plus transparentes dans leur rapport aux données. Cela pourrait enrichir le travail de réflexion et d'observation du doctorant et aider les futurs doctorants à penser leur méthodologie.

Bénéficiaire d'une aide :

Selon les réponses, la plupart des services proposés ont été sélectionnés par les répondants :

- **33,3 %** souhaitent recevoir des **conseils d'ordre général** pour la gestion des données de la recherche ;
- **16,7 %** souhaitent des **conseils sur des questions techniques** (métadonnées, normes, archives etc.) ;
- **11,1 %** veulent des **conseils relatifs à la publication** et la citation des données de la recherche ;
- **11,1 %** des **conseils juridiques** (restrictions d'accès, traitement des données sensibles, licences etc.) ;
- **11,1 %** d'une **assistance technique** (publications des données dans un article par exemple).

Les propositions sur les offres de formations ou encore l'aide pour préparer un DMP n'ont pas été sélectionnées par les répondants. Cependant, une personne indique avoir besoin « d'un peu de tout », une autre a besoin de conseils juridiques et techniques.

Du côté du CRENAU, une majorité demande également des conseils d'ordre général pour la gestion des données à 60 %, puis les conseils techniques et ensuite les conseils juridiques (57 %).

Quels projets collectifs pour le laboratoire ? :

L'une des dernières questions du questionnaire proposait 5 choix pour les projets collectifs à moyen terme au sein de l'équipe :

- **55,6 %** **Définir un modèle** pour nommer les fichiers de données ;
- **50 %** **Formations** pour bien utiliser le serveur du laboratoire Thor ;
- **44,4 %** Présentation des **serveurs d'archivages** d'HumNum ;
- **61,1 %** **Élaborer un DMP** au sein de l'équipe ;
- Autre.

⁶ <https://www.ed.ac.uk/>

A la proposition « *autre* », nous avons pu collecter des réponses telles que « *Peut-être quelque chose sur les moyens de partager les données, de ceux qui sont sur les mêmes terrains, ou les mêmes sujets* » ou encore « *Adapter ces plateformes aux besoins de l'équipe* ».

Nous pouvons donc voir que deux majorités se distinguent parmi les besoins de l'équipe :

- L'élaboration d'un DMP au sein de l'équipe et,
- Définir un modèle pour nommer les fichiers des données collectées.

Nous pouvons ainsi réfléchir à la mise en place de ces deux éléments en faisant participer les chercheurs sous la forme d'atelier par exemple, afin d'être au plus près des besoins et des attentes des membres du CRESSON. Nous pouvons remarquer que la création d'un DMP est aussi une demande majoritaire au sein du CRENAU avec 90 %.

Accompagnement dans la gestion :

« Avez-vous un besoin personnel d'accompagnement dans la gestion de vos données de recherche ? »

De façon assez surprenante, une majorité de répondants a indiqué « *non* » à cette question avec 61,1 % d'entre eux. 38,9 % ont répondu « *oui* ». Nous avons demandé aux répondants de préciser, en cas de réponse positive, le type d'accompagnement qu'il souhaitait : cette question n'étant pas obligatoire nous n'avons pu récolter que 4 réponses.

Une personne a indiqué « *gestion des données qui sont à la fois artistiques et scientifiques ; dans un souci d'ouverture, d'exploitabilité par l'un ou l'autre champ :*

1. la pluri - et transdisciplinarité mérite des formes originales d'écriture, de classement, de référencement, ainsi que des formes originales d'accès aux données ;

2. l'hybridation des données pose non seulement des questions de référencement et de protection des données, mais de valorisation : ce qui est fait la valeur pour l'un n'est pas/ ne fait pas valeur pour l'autre : c'est pourtant dans cet "entre" que se trouve souvent la plus-value du travail pluri, inter, transdisciplinaires). Un répondant a indiqué souhaiter avoir plus d'accompagnement à propos des outils pour la gestion des données photographiques, comment les trier, les annoter, leur relier des métadonnées de manière pérenne etc. Il est donc nécessaire d'approfondir la typologie des données produites pour mieux en comprendre la valeur et leur potentielle valorisation.

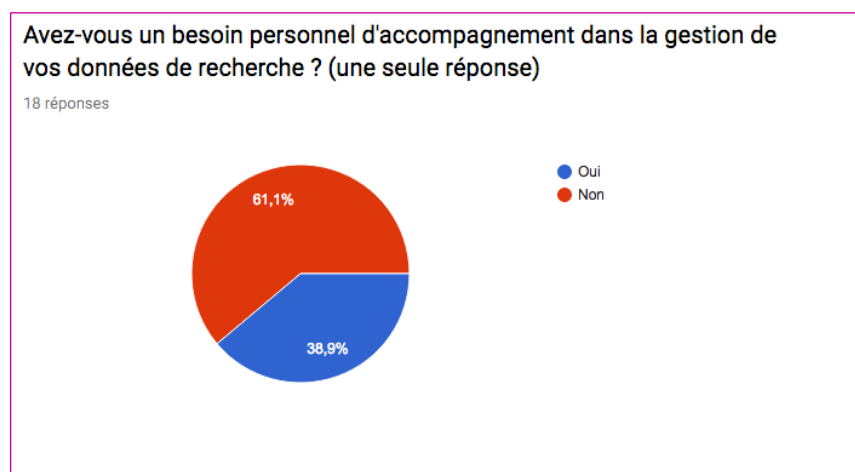


Figure 13 : Graphique secteur du besoin d'accompagnement personnel des chercheurs

Synthèse des entretiens semi-directifs

Pour les données collectées et produites au CRESSON, les entretiens ont permis d'enrichir les résultats de l'enquête. Les données sont des typologies assez diverses, tels que les enregistrements sonores, les photographies, les archives ou encore les textes. Dans tous les cas, la présence de données orales est fortement représentée, ce qui est cohérent puisque c'est une base forte de la méthodologie de recherche de l'équipe.

A propos du stockage de ces données :

Il apparaît que beaucoup de participants rencontrent des soucis de stockage et de gestion. Le manque de place est un problème récurrent, la taille des fichiers est trop importante pour les sauvegarder et les traiter sur un ordinateur. Certains expliquent qu'ils rencontrent souvent des problèmes d'ordre technique, les machines ne sont pas assez performantes pour traiter les fichiers très lourds telles que les photographies aériennes. Le stockage sur un disque dur externe entraîne également des risques de perte ou d'endommagement du dispositif.

L'usage du disque dur externe n'est pas assez fiable et il pourrait être intéressant de sensibiliser les chercheurs à utiliser davantage le serveur du laboratoire mis à leur disposition.

Qui en dehors de vous peut accéder à vos données ?

En dehors de l'équipe scientifique, au sein de l'unité ou dans le cadre d'un projet, nous pouvons constater que **les données sont donc majoritairement inaccessibles**, autant pour le CRESSON que pour le CRENAU. Nous pouvons expliquer ce résultat en partie par la réponse de l'une des personnes interviewées qui nous indique ne pas être encline à partager ses données car les matériaux récoltés sont en général, très difficiles à anonymiser. Les entretiens réalisés parlent très souvent de l'intimité des personnes etc. ce qui en fait des données personnelles⁷, elles sont donc protégées. Le partage de ce type de données prendrait trop de temps et il serait trop complexe de les rendre publiques.

Les raisons qui inciteraient les chercheurs à publier les données en Open Access :

Certaines personnes interrogées ont précisé être d'accord pour diffuser des données en Open Access seulement quand les données sont stables, afin de ne pas divulguer des informations erronées au grand public. L'adhésion aux valeurs de l'archive ouverte se justifie en partie, par le fait que toutes les personnes interrogées ont indiqué être d'accord pour l'Open Access pour les publications. Elles soutiennent les valeurs de la Science Ouverte par ce biais. Plusieurs personnes ont évoqué l'importance de rendre la science transparente ou encore de partager les connaissances afin de pouvoir les réutiliser.

Les freins principaux :

Les résultats concernant la **crainte du plagiat** sont similaires avec ceux du CRENAU (27 %), ainsi que ceux pour des **raisons personnelles** (5,4 %). Nous pouvons donc compléter ce constat en ajoutant un point déjà évoqué plus haut : beaucoup de données contiennent des données à caractère personnel, il est donc difficile de les partager car elles sont protégées, mais aussi par respect pour les personnes enregistrées : une relation de confiance s'est construite avec le chercheur qui ne souhaite pas porter préjudice aux personnes qui acceptent ces entretiens.

Savez-vous ce qu'est un Plan de Gestion de Données ?

⁷ Toute information identifiante directement ou indirectement une personne physique (ex. nom, no d'immatriculation, no de téléphone, photographie, date de naissance, commune de résidence, empreinte digitale...) selon C. Delplanque *et al.*, *Guide RGPD 2019 - fiches pratiques à destination des chercheurs*, Université Paris Nanterre, 2018

La méconnaissance des PDG et leur nécessaire élaboration pour les ANR en cours et futures appellent une acculturation collective, par la création de supports partagés au sein de l'UMR.

Conclusion

Par l'ensemble de ces résultats, nous pouvons mettre en évidence les points communs et les différences entre les deux équipes CRENAU et CRESSON sur le sujet des données de recherche. Tout d'abord, le CRENAU a répondu de façon plus massive à l'enquête proposée par le Centre de documentation.

Nous pouvons remarquer que les chercheurs de CRENAU sont sans doute plus engagés sur certaines questions inhérentes à la gestion des données. La MSH Angé-Guépin a un rôle moteur dans la gestion des données et entraîne ainsi les différentes structures à se préoccuper de la question.

Dans l'environnement du CRESSON, la MSH Alpes commence à se positionner sur ces questions et a mis en place un cycle de conférences à l'attention des chercheurs, afin de vulgariser ces notions difficiles à appréhender.

Outre le taux de réponses, nous avons pu collecter certaines informations grâce aux différentes questions et rubriques nous permettant ainsi de connaître dans les grandes lignes les pratiques et habitudes des chercheurs du CRESSON. Parmi les chiffres les plus marquants nous retrouvons par exemple 94,4 % de personnes utilisant des textes comme matériel source de recherche et 88,9 % utilisent les entretiens.

De plus, nous avons pu constater que 88,9 % de personnes utilisent un disque dur externe pour stocker leurs données et travaux de recherche, un chiffre similaire à celui de l'équipe CRENAU. Cette dernière pratique de stockage soulève des questions sur la pérennité et la sécurité des informations conservées sur ce type de support.

Cependant un disque dur est simple d'utilisation et facilement transportable pour les chercheurs ; des conditions avantageuses pour sauvegarder ses données.

Faut-il davantage pousser les membres de l'équipe à utiliser les serveurs ?

La mise en place d'un groupe de travail à l'échelle de l'UMR est une des conséquences de ce travail d'enquête qui a sensibilisé les équipes aux enjeux des données de recherche. C'est à cette échelle que nous allons définir une feuille de route précisant éléments de politique et conseils pratiques afin d'accompagner les chercheurs dans leur besoin.

Tableau des préconisations

Espace de stockage : nous pouvons constater que l'usage de périphérique de stockage externe est majoritaire. Il pourrait être intéressant de sensibiliser les chercheurs à utiliser davantage le serveur du laboratoire mis à leur disposition

Fréquence des copies de sauvegarde : nous pouvons soulever ce sujet pour clarifier la politique de conservation au sein de la structure. (38,9 % réalisent des copies irrégulièrement selon les besoins).

Durée de conservation des données : 8,9 % des répondants souhaitent que les données soient ou devraient être conservées et préservées à long terme au-delà du projet. S'engager dans un processus de conservation à long terme avec le CINES est complexe, il serait donc utile de clarifier le travail pour une procédure de conservation à long terme.

Utilisation du serveur Thor : nous pouvons réfléchir à mettre ou remettre en place de courtes réunions pour informer les membres du CRESSON sur la présence de ce serveur ou souligner particulièrement sa présence dans le livret d'accueil.

Utilisation des plateformes institutionnelles pour partager leurs données : Il pourrait être intéressant et important de présenter la plateforme Huma-Num afin de faire connaître aux chercheurs la présence de cet outil et les autres entrepôts de données.

Préférence pour déposer : il est nécessaire de communiquer autour des entrepôts de données existants, leurs modalités et prendre connaissance des entrepôts universitaires naissant.

A propos du Plan de Gestion des données : il pourrait ainsi être intéressant de s'inspirer des universités du Royaume-Uni afin de rendre les thèses plus transparentes et se former collectivement afin de répondre aux demandes de l'ANR.

Les projets collectifs de l'équipe : réfléchir à la mise en place de l'élaboration d'un DMP au sein du CRESSON et définir un modèle pour nommer les fichiers des données collectées, en faisant participer les chercheurs sous la forme d'ateliers par exemple, afin d'être au plus près des besoins et attentes des membres de l'équipe.

Annexe n°1 : Enquête sur les pratiques des membres de l'équipe CRESSON du laboratoire AAU en matière de données de la recherche

Bonjour,

Les données de la recherche sont un enjeu majeur du Plan National pour la Science Ouverte mis en place depuis novembre 2018.

Ces dernières, dont la définition varie en fonction des disciplines, peuvent être définies par toutes les productions, numériques ou non, collectées et réalisées par les chercheurs en amont de leur travail d'écriture (sources, matériaux - mesures, images, sons...) ainsi que toutes les formes de traitement et d'analyses des sources, que l'on appelle alors des données dérivées.

Cette enquête vise donc à établir un état des lieux le plus complet possible sur :

- La nature et la place des données dans vos travaux,
- Vos pratiques d'archivage et de valorisation,
- Pour réfléchir à une offre de services pouvant répondre aux besoins des membres.

Merci de bien vouloir remplir ce questionnaire qui vous demandera 10 à 15 minutes.

La participation est anonyme.

Ce questionnaire est fortement inspiré d'une enquête réalisée par l'URFIST de Rennes, la MSHB et le SCD Rennes 2 et d'une enquête de l'Université de Lille 3.

Contact : mosnier.l@grenoble.archi.fr

Vos champs disciplinaires et votre statut

Vous êtes ? (Plusieurs réponses possible)

- Chercheur.e CNRS
- Enseignant.e et chercheur.e
- Doctorante
- Post doctorante
- Autre :

Veillez cocher votre tranche d'âge (Une seule réponse)

- < 30 ans
- 30-50 ans
- > 50 ans

Quelle(s) activité(s) exercez-vous actuellement ? (Plusieurs réponses possible)

- Enseignement
- Recherche
- Encadrement de thèse
- Direction d'une revue, membre d'un comité de rédaction ...
- Projet de thèse
- Autre :

Dans quelle discipline exercez-vous ? (Plusieurs réponses possible)

- Architecture
- Urbanisme
- Géographie
- Ingénierie (acoustique etc.)
- Sociologie
- Philosophie
- Autre :

La typologie des données (sources, matériaux de terrain collectés et produits)

A partir de quel(s) matériau(x) réalisez-vous vos recherches ? (Plusieurs réponses possible)

- Textes
- Vidéos
- Mesures (sonomètre, luxmètre etc.)
- Enregistrements sonores
- Images fixes de paysages, d'architectures, de textes etc.
- Entretiens
- Archives
- Autre :

Ces matériaux sont-ils en majorité ... (Une seule réponse possible)

- Non numériques (relevés de terrain, notes manuscrites, textes imprimés, archives etc.)
- Nativement numériques (photo, audio, vidéo, carte, plan etc.)

Pour vos matériaux non numériques, les numérisez-vous ensuite ? (Scan, retranscription etc.) (Une seule réponse possible)

- Oui
- Non

Quelles ont les données produites en analysant vos sources ? (Données dérivées) (Plusieurs réponses possibles)

- Prises de notes informatiques
- Notes manuscrites (carnets de terrain etc.)
- Tableurs, schémas etc.
- Bases de données
- Logiciels
- Cartes
- Fabrications matérielles
- Productions audio-visuelles
- Autre

Pratiques de stockage et d'archivage des données (sources, matériaux de terrain collectés et produits)

Où stockez-vous vos données ? (Plusieurs réponses possible)

- Sur un ordinateur personnel
- Sur un ordinateur professionnel
- Sur un disque dur externe et autres dispositifs (Clé USB, CD, DVD etc.)
- Sur une plateforme privée dans le cloud (Google Drive, Dropbox etc.)
- En réseau sur le serveur du laboratoire
- En réseau sur le serveur de l'école d'architecture
- Sur un serveur externe d'une autre institution
- Sur un serveur d'une TGR (Très Grandes Infrastructures de Recherche par ex. Huma-Num)
- Autre :

A combien estimez-vous l'espace nécessaire au stockage de vos données ? (Une seule réponse)

- < 1 Go
- 1 – 20 Go
- 20 – 100 Go
- 100 – 1 To
- > 1 To
- Je ne sais pas

A quelle fréquence faites-vous des copies de sauvegarde ? (Plusieurs réponses possible)

- Tous les jours

- Toutes les semaines
- Tous les mois
- Tous les trimestres
- Tous les ans
- Irrégulièrement en fonction des besoins
- En fonction du prestataire de service ou du gestionnaire du site
- Je ne fais pas de copie de sauvegarde

Combien de temps les données sont-elles/devraient être conservées et préservées ? (Une seule réponse possible)

- La durée du projet
- A long terme au-delà du projet (pour valider vos résultats de recherche, pour mener de nouvelles études ou pour l'enseignement)

Ajoutez-vous des informations supplémentaires à vos données ? (Plusieurs réponses possible)

- Des informations administratives (auteur, date de création etc.)
- Des informations techniques (formats des fichiers, taille, logiciels utilisés etc.)
- Des informations scientifiques (l'origine des données, le projet de recherche, des mots-clés etc.)
- Je n'en n'attribue pas

Si vous en attribuez, sous quelle forme le faites-vous ? (Tableur, fiche...)

- Réponse courte :

Connaissez-vous le serveur du laboratoire "Thor" et les possibilités qui vous sont proposées pour sauvegarder vos données ? (Une seule réponse possible)

- Oui
- Non

Pratiques de partage et diffusion des données (sources, matériaux de terrain collectés et produits)

Qui, en dehors de vous, peut accéder à vos données ? (Plusieurs réponses possible)

- Personne
- Le groupe de recherche
- Le laboratoire
- L'école d'architecture
- Les chercheurs de la discipline
- Tout le monde

Par quel(s) moyen(s) partagez-vous les données avec d'autres chercheur.e.s ?

- Par mails
- Par réseaux sociaux (Facebook etc.)
- Par plateforme privées en cloud (Dropbox, GoogleDrive etc.)
- Par plateformes institutionnelles (ShareDocs, Thor etc.)
- Par présentiel
- Autre

Avez-vous au moins une fois publié vos données en libre accès sur le web ? (Ex : mettre en ligne des vidéos, des sons, des photos en dehors du texte de l'article) (Une seule réponse possible)

- Oui
- Non pas encore mais j'en ai l'intention
- Non je ne connaissais pas cette possibilité
- Non et je ne le ferais pas

Avez-vous déjà au moins une fois utilisé des matériaux provenant d'autres chercheur.e.s ? (Une seule réponse possible)

- Oui
- Non pas encore mais j'en ai l'intention
- Non et je ne le ferais pas

Seriez-vous d'accord pour diffuser vos matériaux en libre accès sur le web ? (Une seule réponse possible)

- Oui tout à fait
- Oui après un certain délai
- Plutôt oui
- Plutôt non
- Non pas du tout
- Je ne sais pas

Qu'est-ce qui vous inciterait à rendre vos matériaux de recherche accessibles en libre accès ? (Plusieurs réponses possible)

- Une obligation faite par le financeur de la recherche (Horizon 2020, ANR)
- La reconnaissance de la communauté scientifique
- Une meilleure visibilité de vos travaux
- L'occasion d'un nouveau contact, de coopérations
- L'adhésion aux valeurs de l'archive ouverte

Quelles seraient les raisons qui vous freineraient pour rendre vos matériaux accessibles ? (Plusieurs réponses possible)

- Un frein juridique lié à la nature des données (données personnelles, sous droit etc.)
- Des raisons scientifiques
- La crainte du plagiat
- Des raisons personnelles (mes données m'appartiennent)
- Le manque de temps
- Le manque de connaissance dans la description des données et les modalités de dépôt
- Le manque de compétence
- Aucune des raisons présentes
- Autre

Si vous souhaitez les diffuser, quel est ou serait votre support principal ? (Plusieurs réponses possible)

- Les archives ouvertes (comme HAL par exemple)
- Des plateformes spécifiques dans l'archivage et la diffusion des données de recherche (Zenodo etc.)
- Un réseau social académique (Academia, ResearchGate etc.)
- Un blog ou site personnel
- Le site du laboratoire
- Un serveur institutionnel (HumaNum) ou du laboratoire
- Un data paper (Un data paper est une publication décrivant un ou plusieurs corpus scientifiques avec des métadonnées/informations précises de type : obtention, description des données, etc.)
- Autre

Vos besoins et attentes pour vos données de recherche

Savez-vous ce qu'est un Plan de Gestion de Données (PGD) ou dit Data Management Plan (DMP) ? (Une seule réponse)

- Oui
- Non

Savez-vous que ce document administratif qui retrace le cycle de vie des données est obligatoire depuis 2019 pour le dépôt d'un projet ANR ou H2020 ? (Une seule réponse)

- Oui
- Non

Sur quels points souhaiteriez-vous bénéficier d'une aide en priorité ? (Une seule réponse)

- Des conseils d'ordre général pour la gestion des données de la recherche
- Des conseils relatifs à la publication et la citation des données de la recherche

- Des conseils sur des questions techniques (métadonnées, normes, archives etc.)
- Des conseils juridiques (restrictions d'accès, traitement des données sensibles, licences etc.)
- Sur les offres de formations sur la question des données de la recherche
- Une assistance technique (publications des données dans un article par exemple)
- Une aide pour préparer un DMP
- Autre

Quels projets collectifs voyez-vous à moyen terme au sein du laboratoire ? (Plusieurs réponses possible)

- Définir un modèle pour nommer les fichiers de données
- Formations pour bien utiliser le serveur du laboratoire "Thor"
- Définir un ensemble de métadonnées pour décrire les données
- Élaborer un DMP au sein du laboratoire

Avez-vous un besoin personnel d'accompagnement dans la gestion de vos données de recherche ? (Une seule réponse)

- Oui
- Non

Si oui, pouvez-vous préciser lequel ?

- Réponse courte :
-

Intérêt pour un entretien et les résultats de l'enquête

Seriez-vous prêt à participer à un entretien sur votre expérience et vos attentes dans le domaine des données de recherche ?

- Oui
- Non

Si oui, veuillez-nous laisser vos coordonnées

- Réponse courte :

Des remarques particulières ?

Annexe 2 : questionnaire pour les entretiens semi-directifs

Nous avons proposé les questions suivantes abordant différents aspects de la gestion des données :

Avez-vous déjà entendu le terme « Données de recherche » ?

Comment le définiriez-vous ?

Faites-vous une différence entre les données collectées et les données produites ?

Qu'est-ce qui pose problème pour définir le terme ?

Pouvez-vous préciser quels types de données vous collectez et produisez ?

Comment les collectez-vous ?

Sous quel(s) format(s) sont ces données sources ?

Quelles sont les types de données provenant d'autres chercheurs que vous êtes amenés à utiliser ?

Avez-vous perçu une évolution des sources collectées dans votre discipline ?

Rencontrez-vous des difficultés particulières pour gérer vos données ?

Est-ce un problème récurrent ?

Pouvez-vous décrire la dernière fois où vous avez rencontré un problème dans la gestion de vos données ?

Quelle est votre approche personnelle de la question du partage des données ?

Êtes-vous plutôt favorable ou défavorable à ce partage ?

Connaissez-vous les obligations liées aux données ?

Avez-vous déjà entendu parler des obligations liées au programme H2020 ?

Trouvez-vous que cette question d'obligation influence votre travail quotidien ?

Bibliographie

- AGENCE NATIONALE DE LA RECHERCHE, « Plan d'action 2019 », ANR, 2018.
- BERT Hélène, « Bibliothèques et Open data : retour d'expérience de la Bibliothèque départementale des Côtes d'Armor », sur *DLIS*, <https://dlis.hypotheses.org/985>, sans date, consulté le 3 mai 2019.
- BIZIEN Laurence, « Questionnaire : enquête sur les pratiques en matière de données de la recherche du Crenau- mars 2019 », 2019.
- CALIMAQ, « Affaire DisinfoLab : quelles retombées potentielles sur la recherche publique et la science ouverte ? », sur - *S.Lex* -, <https://scinfolex.com/2018/08/21/affaire-disinfo-lab-queelles-retombees-potentielles-sur-la-recherche-publique-et-la-science-ouverte/>, 21 août 2018, consulté le 29 janvier 2019.
- COMMISSION EUROPÉENNE et DIRECTION GENERALE DE LA RECHERCHE ET DE L'INNOVATION, « Programme H2020 : Lignes directrices sur les règles pour le libre accès aux publications scientifiques et pour le libre accès aux données de la recherche dans Horizon 2020 », 2017.
- CONTAT Odile, « Gestion des données de la recherche : quelques pistes pour démarrer, comprendre et se former | CORIST-SHS », sur *ISTSHS Correspondants*, http://corist-shs.cnrs.fr/gestion_donnees_recherche_guideetformation, 18 août 2014, consulté le 14 février 2019.
- CORVOL Pierre, *Bilan et propositions de mise en oeuvre de la charte nationale d'intégrité scientifique*, 2016.
- DATA.GOUV.FR, « Découvrir l'Open Data », sur *Institutionnel*, <https://doc.data.gouv.fr/>, sans date, consulté le 21 février 2019.
- DIRECTION GENERALE DE LA RECHERCHE ET DE L'INNOVATION, « HORIZON 2020 en bref - Le programme-cadre de l'UE pour la recherche et l'innovation », sans date, p. 40.
- DORANUM, *La gestion des données de la recherche : retour d'expérience*, sans date, 319 seconds.
- DUPRAT Julie, *Les données de la recherche à l'Université Bordeaux Montaigne*, Bordeaux, Université Bordeaux Montaigne, 2019.
- FACHINOTTI Elena, Eva GOZZELINO, Sara LONATI et René (Dir) SCHNEIDER, *Les bibliothèques scientifiques et les données de la recherche: défis et enjeux*, sans lieu, coll. « Travail de recherche réalisé dans le cadre du Master of Science en Information documentaire HES à la Haute école de gestion de Genève (HEG-GE), Filière information documentaire, 2016 », sans date.
- GAILLARD Rémi, *De l'Open data à l'Open research data : quelle(s) politique(s) pour les données de recherche ?*, ENSSIB, Lyon, 2014.
- GBONGI, « Les métadonnées, ou l'art de la description », sur *Recherche d'ID*, <https://recherchemid.wordpress.com/2017/10/02/les-metadonnees-ou-lart-de-la-description/>, 2 octobre 2017, consulté le 29 janvier 2019.
- GENVRESS Sophie, *Données personnelles et sensibles : quels enjeux pour la recherche en SHS ?*, Université Paris Nanterre, Institut des Sciences sociales du Politique, 2017, 25:39.
- GINOUVES Véronique et Isabelle GRAS, « Aix-Marseille Université - La diffusion numérique des données en SHS - Guide de bonnes pratiques éthiques et juridiques », sur *Institutionnel*, <https://hal-amu.archives-ouvertes.fr/page/guide-de-bonnes-pratiques>, sans date, consulté le 8 mars 2019.

- INGRAM Caroline, « Managing research data in your institution », sur *Jisc*, <https://www.jisc.ac.uk/guides/research-data-management>, 2 octobre 2017, consulté le 29 janvier 2019.
- IST, « Datapartage - Plan National pour la Science Ouverte », <https://www6.inra.fr/datapartage/Actualites/Plan-National-pour-la-Science-Ouverte>, 14 août 2018, consulté le 14 février 2019.
- MINISTERE DE L'ENSEIGNEMENT SUPERIEUR, DE LA RECHERCHE ET DE L'INNOVATION, « Ouverture des données de recherche : guide d'analyse du cadre juridique en France », Ministère de l'Enseignement supérieur, de la recherche et de l'innovation, 2017.
- PROST Hélène et Joachim SCHÖPFEL, *Les données de la recherche en SHS. Une enquête à l'Université de Lille 3*, 2015.
- ROBIN Agnès, *Les données scientifiques et le droit*, 2018.
- ROYER Chantal, François GUILLEMETTE et Jocelyne MOREAU, « L'instrumentation dans la collecte des données », *Recherches qualitatives*, n° 2, 2005, p. 114.
- SCHÖPFEL Joachim, *Les données de la recherche : politiques et enjeux*, 2018.
- SERRES Alexandre, « Partage des données de recherche en SHS : quelles pratiques des (...) », sur *MSHSUD.TV*, <http://www.mshsud.tv/spip.php?article872>, 16 octobre 2018, consulté le 14 février 2019.
- SERRES Alexandre, MALINGRE Marie-Laure, MIGNON Morgane, PIERRE Cécile et COLLET Didier, *Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l'Université Rennes 2*, 2019.
- SEVERO Marta, « Open-access à l'université Paris Nanterre - Les données de la recherche », sur *Open-access à l'université Paris Nanterre*, <https://openaccess.parisnanterre.fr/open-data/les-donnees-de-la-recherche/les-donnees-de-la-recherche-826916.kjsp?RH=1518708014719>, sans date, consulté le 14 février 2019.
- STERIN Anne-Laure, « Le point sur les données à caractère personnel », sur *Ethique et droit*, <https://ethiquedroit.hypotheses.org/1717>, 9 novembre 2018, consulté le 21 février 2019.
- STERIN Anne-Laure, « Le chercheur « fouille » dans des textes ou dans des données », sur *Ethique et droit*, <https://ethiquedroit.hypotheses.org/1528>, 27 juillet 2017, consulté le 21 février 2019.
- SUBER Peter, *Qu'est-ce que l'accès ouvert ?*, Marie Lebert (trad.), Marseille, OpenEdition Press, coll. « Encyclopédie numérique », 2016.
- UNKNOWN, « Blog UniVersDoc: Interessez-vous au Data Management: ce métier qui s'impose aux gestionnaires de l'information et aux organisations », sur *Blog UniVersDoc*, <http://bloguniversdoc.blogspot.com/2015/03/interessez-vous-au-data-management-ce.html>, 5 mars 2015, consulté le 3 mai 2019.
- VERONIQUE Ginouvès et Gras ISABELLE (éd.), *La diffusion numérique des données en SHS : guide des bonnes pratiques éthiques et juridiques*, Presses Universitaires de Provence, Aix Marseille Université, coll. « Digitales », 2018.
- « Enquête sur les bibliothèques universitaires et de recherche », sur *Archimag*, <https://www.archimag.com/bibliotheque-edition/2019/01/14/enquete-bibliotheques-universitaires-recherche>, 14 janvier 2019, consulté le 7 février 2019.
- « Pour une action publique transparente et collaborative : plan d'action national pour la France 2018-2020 », 2018.
- *Les clés de la Science Ouverte*, sans lieu, coll. « DidakTIC : Convergence des cultures de l'information », n° 2, 2018.
- « Réaliser un plan de gestion de données "FAIR" : guide de rédaction », Université Paris Diderot Paris 7, 2018.

- « Les données, quel enjeu pour la recherche et les bibliothèques? », sur *HTTBU*, <http://bibliotheque-blogs.unice.fr/httbu/2017/01/06/les-donnees-quel-enjeu-pour-la-recherche-et-les-bibliotheques/>, 6 janvier 2017, consulté le 3 mai 2019.
- « Budapest Open Access Initiative », <https://www.budapestopenaccessinitiative.org/read>, 14 février 2002, consulté le 21 février 2019.