



Gestion et ouverture des données de recherche à la HES-SO

Enquête sur les pratiques
et besoins des chercheuses
et chercheurs



Constance Delamadeleine
Décembre 2023



Table des matières

	Les chiffres clés	3
	Liste des figures	4
1	Introduction	5
1.1	Objectifs	6
1.2	Structure et méthodologie	6
1.3	Limites	7
1.4	Remerciements	7
2	Les répondant-es	8
2.1	Taux de participation	8
2.2	Profil et affiliation	8
3	Le moment « données »	10
3.1	Caractéristiques des données	10
3.2	Des données hétérogènes : défis et opportunités	12
3.3	La propriété des données : la complexité de l'attribution	12
4	La gestion des données de recherche	14
4.1	La question de la responsabilité	14
4.2	Le Data Management Plan (DMP) : un standard à la HES-SO?	15
4.3	L'affaire des principes « FAIR »	16
4.4	Volumétrie, stockage et sauvegarde	17
4.5	Les données personnelles et leur sécurité	19
4.6	Les métadonnées et les standards : loin d'être une norme	20
5	Vers le partage des données	21
5.1	L'attitude face au partage	21
5.2	Les freins	22
5.3	L'expérience du partage	23
5.4	Focus sur la réutilisation	25
5.5	Connaissez-vous le Data paper?	25
6	Soutien et besoins	26
6.1	Etat des lieux	26
6.2	Les collègues comme premier point d'appui	26
6.3	Soutien financier	27
6.4	Les besoins en formation	27
6.5	Voix libre aux chercheur-euses	29
7	Conclusion et recommandations	30
7.1	Synthèse des résultats	30
7.2	Recommandations	31
	Annexe	
	Questionnaire	

Les chiffres clés

56%

ne connaissent pas
les principes « FAIR »

43%

ont déjà rempli un
Data Management
Plan

12%

reçoivent des
financements
pour traiter leurs
données en vue
du partage

223 répondant·es

58%

gèrent des données
personnelles

36%

ont déjà réutilisé
des données d'autres
chercheur·euses

65%

sont favorables
au partage
des données
de recherche

62%

n'ont jamais suivi
de formation
sur la gestion
des données
de recherche

67%

n'ont jamais partagé
leurs données
de recherche

Liste des figures

Fig. 1	Fonction principale
Fig. 2	Domaine de recherche
Fig. 3	Canton de rattachement
Fig. 4	Types de données collectées/générées (projet actuel)
Fig. 5	Types de données collectées/générées par domaine (projet actuel)
Fig. 6	Les données générées/collectées appartiennent à (projet actuel)
Fig. 7	Bailleur(s) de fonds (projet actuel)
Fig. 8	Responsable de la gestion des données de recherche
Fig. 9	Rédaction d'un DMP au cours de sa carrière
Fig. 10	Importance accordée à la planification de la gestion des données de recherche
Fig. 11	Connaissance des principes FAIR
Fig. 12	Volume de données (projet actuel)
Fig. 13	Volume de données par domaine (projet actuel)
Fig. 14	Support(s) de stockage utilisé(s) (projet actuel)
Fig. 15	Collecte et gestion de données à caractère confidentiel (e.g. données personnelles ou sensibles)
Fig. 16	Les données sont sécurisées (projet actuel)
Fig. 17	Utilisation de métadonnées pour documenter, décrire les données de recherche
Fig. 18	Utilisation de standard(s) de métadonnées
Fig. 19	Favorable au partage des données de recherche
Fig. 20	Raison(s) qui motive(nt) le choix de partager ses données de recherche
Fig. 21	Raison(s) qui freine(nt) le partage des données de recherche
Fig. 22	Projet actuel sous contrat (de confidentialité, ou formulaire de consentement, de non-divulgation, etc.)
Fig. 23	Expérience du partage des données de recherche
Fig. 24	Expérience du partage des données de recherche par domaine
Fig. 25	Intention de partager ses données de recherche (projet actuel)
Fig. 26	Connaissance du Data paper
Fig. 27	Existence de lignes directrices, recommandations internes pour la gestion des données de recherche
Fig. 28	Premier contact pour le soutien à la recherche
Fig. 29	Financement pour gérer les données de recherche (projet actuel)
Fig. 30	Formation(s) souhaitée(s)
Fig. 31	Formation(s) souhaitée(s) par domaine
Fig. 32	Format(s) de formation(s) souhaité(s)
Tableau 1	Nombre d'EPT Ra&D par domaine en 2022
Tableau 2	Freins au partage des données et exemples d'actions pour réduire voire éliminer ces freins
Tableau 3	Choix des dépôts de données par domaine
Tableau 4	Synthèse des résultats

1. Introduction

D'abord articulées autour de l'ouverture des publications, les politiques Open Science accordent désormais un intérêt croissant à l'Open Research Data (ORD). Ces dernières années, les gouvernements, les institutions académiques et les principaux organismes de financement en Suisse et au-delà, œuvrent en faveur de l'ORD. Comme le souligne swissuniversities, l'ouverture des données de recherche « favorise une recherche de meilleure qualité, plus efficace et ayant plus d'impact, au bénéfice de la société dans son ensemble ».¹ Afin d'assurer l'adoption de l'ORD, les grands bailleurs de fonds de la recherche demandent un engagement dans ce sens aux chercheuses et chercheurs. Par exemple, la politique ORD du Fonds national suisse (FNS) « s'attend [...] à ce que les données produites au cours des travaux de recherche soient archivées dans des dépôts de données numériques en libre accès, pour autant qu'aucune clause juridique, éthique, concernant le droit d'auteur ou autre ne s'y oppose ».²

La mise en œuvre réussie de l'ORD repose en grande partie sur l'adhésion aux quatre principes « FAIR » : *Findable, Accessible, Interoperable, Reusable*. Ces piliers directeurs de l'ORD s'appliquent dès la conception du projet de recherche et nécessitent une gestion rigoureuse des données. Cette gestion est « la partie du processus de recherche qui traite de l'organisation et de la manipulation des données de recherche, y compris la planification de la gestion des données, le stockage structuré, la description, la conservation, la préservation et la fourniture de métadonnées et d'algorithmes, de codes, de logiciels et de flux de travail complémentaires, ainsi que la conformité avec la législation interne, nationale et internationale sur la protection de la vie privée ».³ En d'autres termes, l'Open Research Data ou devrait-on dire le « FAIR Data », est loin de se limiter à un « drag and drop » d'un jeu de données sur Zenodo. En effet, le « FAIR Data » est exigeant. Il demande une approche systématique et continue de la gestion des données tout au

long de leur cycle de vie, engageant les chercheur·euses à connaître et à adopter des normes, des pratiques et des politiques qui favorisent la disponibilité, l'accessibilité, et la réutilisation des données de recherche.

Adoptée en 2018, la stratégie Open Science de la HES-SO qui intègre la mise en place des directives nationales concernant l'ouverture des données de recherche, met en lumière les multiples défis et enjeux relatifs à l'ORD et reconnaît la complexité de sa mise en œuvre.⁴ Depuis son adoption, diverses mesures ont été mises en place afin de guider les chercheur·euses vers cette nouvelle culture. Cela comprend entre autres, la mise en place de formations en ligne, d'ateliers de sensibilisation, un appel à projets en 2020, renouvelé en 2021 et l'élaboration de deux guides ([guide juridique](#) et [guide rapide sur la gestion des données de recherche](#)). Plus récemment, la cellule data stewardship a été mise en place en juin 2023 afin de renforcer le soutien aux chercheuses et chercheurs de la HES-SO. Des constats ont également été dressés. Parmi eux, et dans le sillage des exigences et recommandations formulées dans la stratégie nationale publiée en 2021, la nécessité de tenir compte des spécificités disciplinaires et des différentes pratiques de recherche de la HES-SO. Cette transition vers l'ouverture des données de recherche a mis en évidence une réalité supplémentaire : celle d'une approche ambitieuse qui exige de nouvelles pratiques, de nouvelles compétences et de nouvelles infrastructures. La stratégie Open Science 2025-2028 consolidera cette réalité en accordant une importance significative à l'ORD. Plusieurs années après la publication de la stratégie Open Science de la HES-SO et à l'aube de celle à venir, il s'agit de faire un bilan : dans quelle mesure cette nouvelle culture s'est-elle disséminée ? Quels sont les obstacles et les besoins des chercheuses et des chercheurs ? C'est dans cette perspective d'évaluation que cette enquête a été initiée. Ce rapport présente les résultats de cette enquête quantitative réalisée auprès de la communauté scientifique de la HES-SO au printemps 2023.

1 Swissuniversities (2021). Stratégie nationale Open Research Data, p.4. Traduit de l'anglais par l'auteure. Pour faire suite à la stratégie Open Access de 2017 visant le libre accès aux publications scientifiques, swissuniversities a adopté la stratégie ORD ainsi que son plan d'action fin 2021.

2 Fonds national suisse (2023). Open Research Data.

3 OCDE (2021). Recommandation du Conseil concernant l'accès aux données de la recherche financée sur fonds publics.

4 Voir : Pirinoli, C., Sauthier, G. (2018). Stratégie Open HES-SO.

1.1 Les objectifs de l'enquête

Plus concrètement, l'enquête a pour objectif premier de dresser un état des lieux, même partiel :

- des pratiques en matière de gestion des données de recherche
- des pratiques, connaissances et attitudes relatives à l'ouverture des données de recherche
- des besoins des communautés de recherche
- des tendances par domaine.⁵

Les résultats de cette enquête constituent une base de travail qui permettra d'élaborer et de soutenir la mise en œuvre de la stratégie Open Research Data de la HES-SO et de proposer un plan d'action au plus proche des besoins de la communauté scientifique. Ces résultats seront enrichis par les entretiens qualitatifs conduits par la Cellule data stewardship de la HES-SO qui visent entre autres, à identifier les spécificités par domaine.

1.2 Structure et méthodologie

Le questionnaire contient 54 questions divisées en 17 blocs visant à cerner le profil des répondant·es, les pratiques et besoins en matière de gestion des données (c.f. Annexe). Un champ libre est réservé aux commentaires et suggestions en fin de questionnaire. La durée moyenne est de 15 mn. Cette enquête s'inspire de celle menée par Carmen Jambé dans le cadre de son mémoire de Bachelor à la Haute école de gestion de Genève pour le Service des Ressources informationnelles et archives (UNIRIS) de l'Université de Lausanne.⁶ Les questions ont été adaptées au contexte de la HES-SO. En avril, la chargée de projet Open Research Data (Constance Delamadeleine) a approché la conseillère en évaluation de la HES-SO (Sophie Ruchat) pour la mise sur pied de l'enquête via le logiciel Limesurvey. Avant sa mise en ligne (du 22 mai au 23 juin 2023), des tests ont été réalisés en sollicitant la participation de chercheur·euses de chaque domaine et d'autres personnes concernées par la thématique. Cette étape préliminaire visait à évaluer la pertinence des questions, la clarté des instructions mais aussi à garantir la représentativité et l'adéquation

du questionnaire aux différents domaines de recherche. Les retours ont permis d'apporter des ajustements pour améliorer la qualité globale de l'enquête.

La communication de l'enquête s'est faite via différents canaux. Un email a été adressé aux responsables de recherche leur demandant de relayer le lien vers l'enquête auprès de leurs chercheuses et chercheurs suivi d'un rappel quinze jours après. Une communication a été faite également via la messagerie Teams de l'équipe de la Communauté Open Science suivie d'une publication sur LinkedIn via le service Communication de la HES-SO.

Dans le cadre de cet état des lieux des pratiques, des obstacles et des besoins des chercheur·euses, seuls les résultats les plus pertinents et exploitables sont présentés dans ce rapport. Le nombre de questionnaires traités s'élève à 223 (exprimé par n). Toutefois, étant donné que certaines questions étaient conditionnées par les réponses précédentes ou permettaient un choix multiple, les lecteur·rices trouveront dans certains cas, un nombre inférieur ou supérieur à 223. À l'issue de cette analyse, 32 graphiques ont été générés, couvrant de manière équilibrée les différentes rubriques du questionnaire. Pour des raisons de lisibilité et de représentation, les pourcentages présentés dans ces graphiques sont arrondis. De plus, certaines données annoncées dans le questionnaire sont normalisées. Par exemple, certain·es répondant·es ont indiqué être « doctorant·e » mais aussi « assistant·e ». Dans ce cas, la catégorie « doctorant·e » est retenue comme base pour le calcul pour des fins de représentation. Les catégories « Chargé·e de cours », « Maître d'enseignement » et « Professeur·e » sont regroupées sous « Corps professoral ».

Lors du traitement des réponses, une attention particulière a été accordée à la préservation de l'anonymat des répondant·es. Cela signifie que toutes les réponses fournies ont été traitées de manière confidentielle et sécurisée. Le questionnaire a été conçu de manière à

⁵ La comparaison entre les domaines a été réalisée uniquement lorsque cela s'avérait pertinent.

⁶ Jambé, C. (2015). La gestion des données de recherche à l'Université de Lausanne : enjeux transdisciplinaires (Haute école de gestion de Genève).

exclure toute question susceptible de générer des données personnelles à l'exception de l'adresse email demandée en fin de questionnaire pour celles et ceux souhaitant participer au tirage au sort. Toutefois ces données ont été collectées et conservées de manière distincte et isolées des réponses au questionnaire et sont à ce jour détruites.

1.3 **Les limites**

La difficulté de connaître le périmètre exact de la population cible à l'échelle des vingt-huit écoles constitue un biais sur la représentativité. Le questionnaire a suivi un protocole de communication et de diffusion qui repose sur les réseaux et connexions internes de la HES-SO. Les populations cibles n'ont donc pas été touchées de manière systématique.

De plus, la gestion et le partage des données englobent diverses activités et compétences toutes aussi cruciales les unes que les autres, ce qui explique la longueur du questionnaire. La durée peut être considérée comme l'un des principaux facteurs d'abandon. Au total, 33% ont abandonné en cours de route.

Enfin, au-delà des limites traditionnelles liées aux enquêtes quantitatives, l'échantillon constitue ici la limite principale de l'étude rendant impossible une analyse fine des pratiques disciplinaires.

1.4 **Remerciements**

La gestion des données et leur partage nécessitent des compétences transversales tout comme la mise en œuvre de cette enquête et de ses résultats. Nous remercions les différentes personnes qui ont contribué à ce travail et en particulier les chercheur·euses qui ont dégagé du temps pour répondre à ce questionnaire.

2. Les répondant·es

La HES-SO est composée de six domaines : Design et Arts visuels (DAV), Economie et Services (ES), Ingénierie et Architecture (IA), Musique et Arts de la scène (MAS), Santé (S) et Travail social (TS). Au total, le nombre de personnes affectées aux activités de recherche s'élève à 3264 chercheur·euses qui exercent leur fonction sur différents sites régionaux (BE-JU-NE, FR, GE, VD, VS).⁷

2.1 Taux de participation

Au total, 334 questionnaires ont été enregistrés. Comme indiqué plus haut, il n'a pas été possible de connaître le nombre de personnes ayant reçu l'enquête. Toutefois, le taux de participation s'élève à environ 10% (3264/334 participant·es), un chiffre plutôt encourageant si on considère le caractère récent de cette nouvelle culture. Par ailleurs, ce taux est relativement satisfaisant si on le compare à celui affiché dans des enquêtes comparables.⁸

Sur ces 334 questionnaires enregistrés, 223 sont entièrement complétés. L'échantillon analysé dans ce rapport se concentre uniquement sur les réponses de ces 223 questionnaires afin d'optimiser la cohérence des réponses, de renforcer la qualité des données analysées, et de faciliter une interprétation plus précise des tendances émergentes au sein de l'échantillon étudié.

7 Pour la période 2022, on relève un total de 3264 chercheuses et chercheurs dont 1854 engagé·es à un taux minimum de 20% dans une haute école de la HES-SO.

8 Voir par exemple la récente enquête de l'EPFL qui affiche un taux de participation à 6,5%. Varrato, F. ; Blumer, E. ; Gabella, C. (2022). Quantitative assessment of research data management practice – 2021, p.2.

9 Dans la catégorie « Autre », sont indiquées d'autres fonctions : 1 « Chargé.e Ra&D », 1 Chef.fe de projet Ra&D » et 1 « Gestionnaire de projet Ra&D ».

10 En 2022, le domaine IA enregistre le plus grand nombre d'EPT Ra&D (589) suivi des domaines ES (218), Santé (102), Travail social (93), DAV (29) et MAS (12).

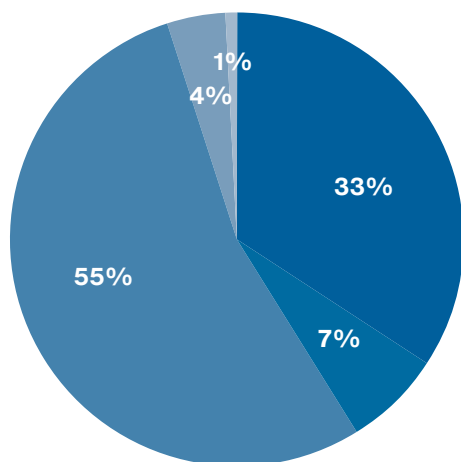
2.2 Profil et affiliation

Les catégories indiquées dans le graphique « Fonction principale » sont retenues comme base pour le calcul du taux d'affiliation (Fig.1). Avec 55%, le corps professoral constitue la catégorie la mieux représentée suivie des membres du corps intermédiaire.⁹ Si le taux de participation des doctorant·es demeure faible (7%) par rapport au corps professoral et corps intermédiaire (adjoint·e/assistant·e, etc.), il montre néanmoins une sensibilité et une implication pour la gestion et l'ouverture des données de recherche parmi les jeunes chercheur·euses.

L'enquête a suscité l'intérêt des chercheuses et chercheurs de tous les domaines ce qui montre un intérêt général pour la thématique de la gestion des données de recherche et de l'ORD. Les répondant·es du domaine Ingénierie et Architecture dominent largement l'échantillon (54%), contrairement aux domaines artistiques qui enregistrent un faible taux de participation soit 4% pour DAV et 2% pour MAS (Fig.2). Ce contraste n'est pas surprenant puisque le domaine IA enregistre un nombre élevé d'EPT Ra&D (589) contre 29 EPT pour le domaine DAV et 12 pour MAS (Tab.1).¹⁰ Toutefois, si on croise ces données, soit le nombre d'EPT par domaine/nombre de répondant·es à l'enquête pour les domaines IA et MAS, on obtient un ratio presque similaire (20% pour IA et 33% pour MAS). En d'autres termes, le taux de participation par domaine est équilibré.

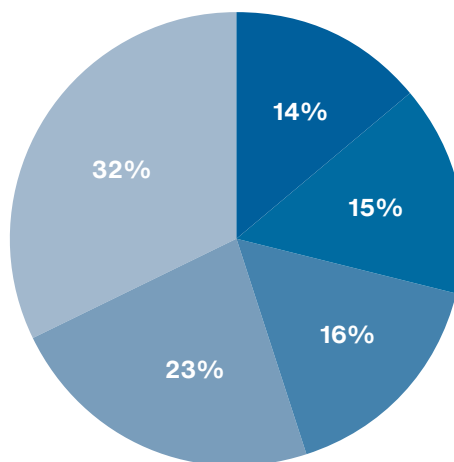
L'enquête couvre la majorité des écoles. Seulement trois (Changins, EDHEA et la HEM) n'ont pas été incluses dans cette représentation statistique, dont deux écoles des domaines artistiques. Cette absence peut être due à différents facteurs, tels que le faible nombre de chercheur·euses dans ces écoles (Changins comptabilise 10 EPT, et l'EDHEA et HEM, 4 chacun). Enfin, les résultats indiquent une participation significative des écoles des cantons du Valais (23%) et Vaud (32%), avec une participation plus ou moins à parts égales des cantons de Genève, Fribourg, Neuchâtel, Berne et Jura (ARC) (Fig.3).

Figure 1 - Fonction principale, n=223



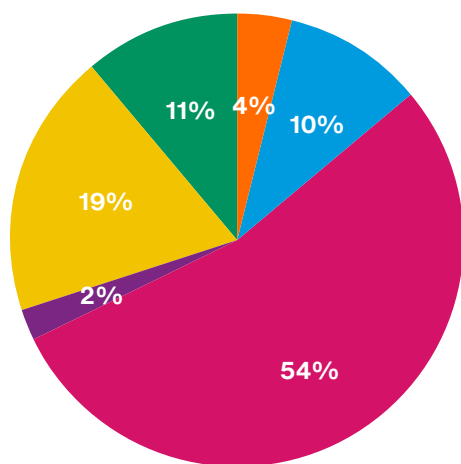
- Adjoint-e scientifique, Assistant-e HES, Collaborateur-riche scientifique
- Doctorant-e
- Corps professoral
- Responsable de recherche
- Autre

Figure 3 - Canton de rattachement, n=223



- ARC (Berne, Jura, Neuchâtel)
- Fribourg
- Genève
- Valais
- Vaud

Figure 2 - Domaine de recherche, n=223



- Design et Arts visuels
- Economie et Services
- Ingénierie et Architecture
- Musique et Arts de la scène
- Santé
- Travail social

Tableau 1 - Nombre d'EPT Ra&D par domaine en 2022 (chiffre arrondi)

Domaine/ Nombre d'EPT	DAV	ES	IA	MAS	S	TS
	29	218	589	12	102	93

3. Le moment « données »

3.1 Caractéristiques des données

Les réponses à la question « Quel(s) type(s) de données collectez-vous et/ou produisez-vous actuellement ? » permettent d’avoir une cartographie des différents types de données manipulés par les chercheuses et chercheurs de la HES-SO (Fig. 4). Deux groupes génériques de données dominent :¹¹

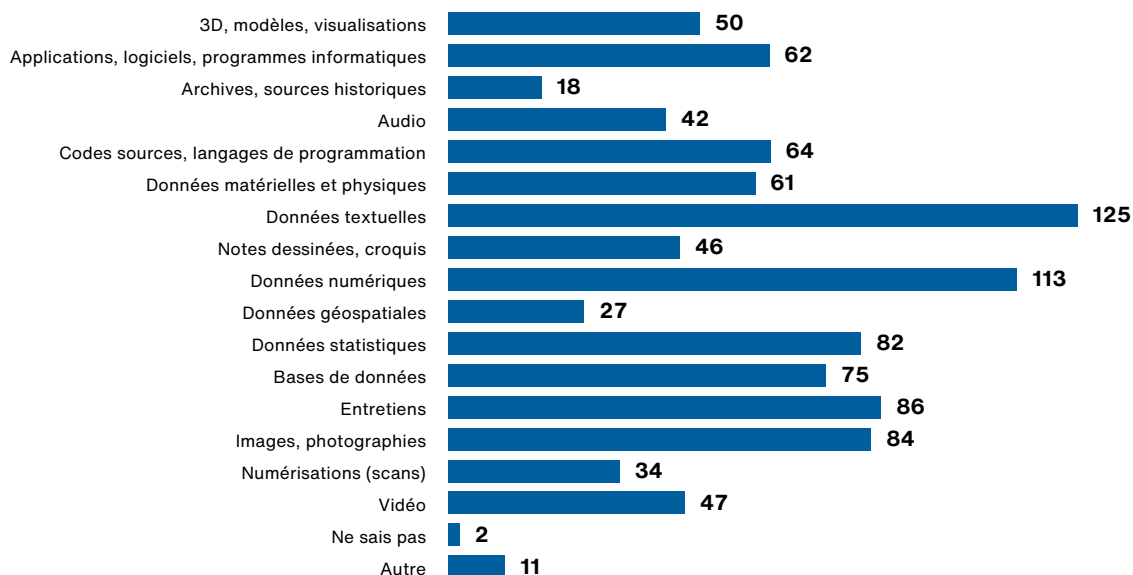
- Données textuelles : elles regroupent par exemple des corpus de texte, des notes manuscrites
- Données numériques : elles regroupent par exemple des tableurs, des feuilles de calcul.

En dehors de ces deux groupes génériques, les types de données les plus fréquemment cités sont :

- Les entretiens
- Les images, photographies
- Les codes sources, langages de programmation.

Parmi les différentes catégories de la classification canonique des types de données (données d’observation, expérimentales, computationnelles/simulation, dérivées ou de référence, informatique), les « données d’observation », (e.g. données d’enquête, images et photographies, notes dessinées, croquis) présentent un certains défi dans le contexte de l’ORD.¹² Ces données sont « collectées à l’instant T, nécessitent un appareil descriptif conséquent (conditions, méthodologie, équipement, etc.). Indissociables d’un contexte donné, et donc uniques et impossibles à reproduire ».¹³ La spécificité du moment de collecte, des circonstances contextuelles, et des méthodologies employées créent une singularité qui peut restreindre la généralisation ou la transposition directe de ces données à d’autres contextes. Les résultats révèlent que ce type de données est fréquemment collecté ou généré, représentant ainsi un défi majeur pour la mise en conformité avec les principes « FAIR ».

Figure 4 - Types de données collectés/générés (projet actuel), n=1029



11 Les catégories « Données numériques » et « Données textuelles » ne sont pas exclusives car elles peuvent regrouper d’autres catégories proposées dans l’enquête.

12 Voir par exemple: Institut de l’information scientifique et technique (2014). Une introduction à la gestion et au partage des données de recherche.

13 Gaillard, R. (2014). De l’Open data à l’Open research data : quelle(s) politique(s) pour les données de recherche ?, p.18.

Citées 61 fois, les données physiques et matérielles, les grandes oubliées de la classification canonique ou de la définition de l'OCDE, constituent un autre type de données qui requiert une attention particulière dans le cadre de l'ORD. Il pose notamment la question des bonnes pratiques pour convertir cette donnée matérielle en donnée numérique.

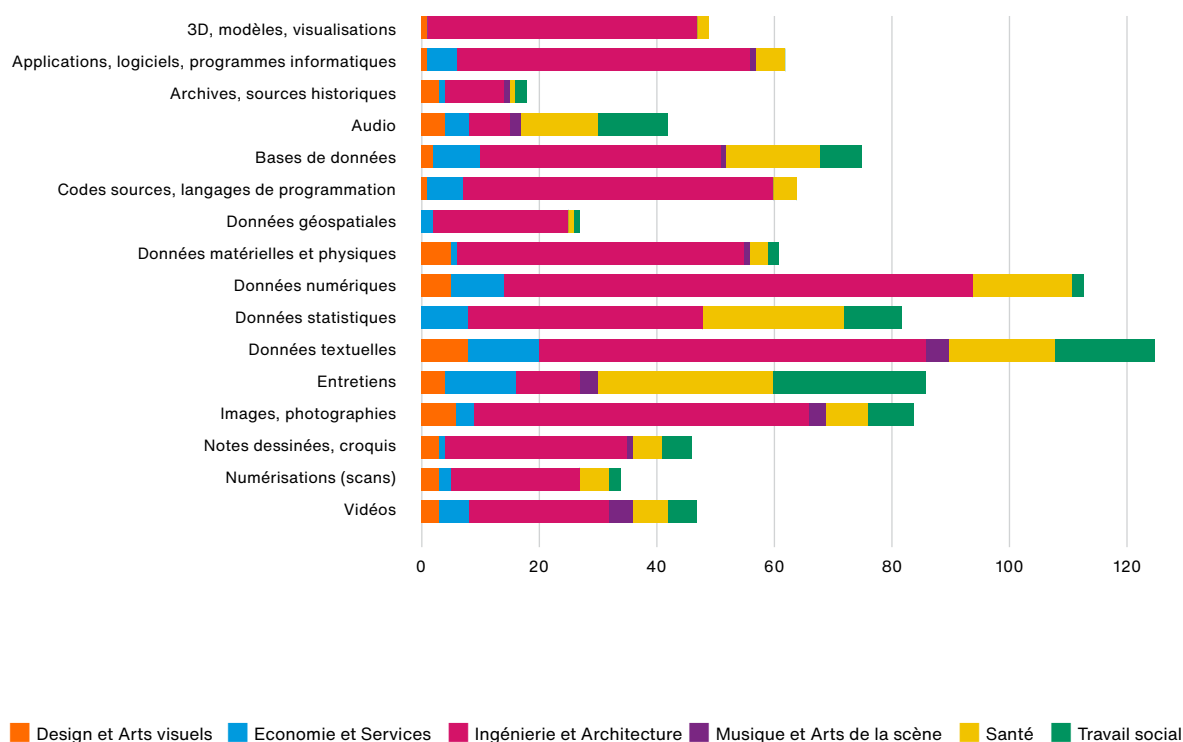
Au total, 1029 réponses ont été obtenues soit en moyenne 5 types de données par chercheur·euse. Certains types de données ne sont pas exclusifs car ils peuvent regrouper d'autres types proposés dans l'enquête (e.g. « Entretiens » peuvent intégrer la catégorie « Données textuelles »). Toutefois, ces résultats mettent en lumière une diversité qui n'est pas sans défis lorsqu'il s'agit de mettre en œuvre des protocoles de gestion standardisés en vue de rendre les données faciles à trouver, accessibles, interopérables et réutilisables. Par exemple, les vidéos peuvent nécessiter des métadonnées spécifiques mais aussi une standardisation des formats afin de faciliter l'interopérabilité et l'échange de données entre différents systèmes, logiciels ou utilisateur·ices. De plus, certains types de données peuvent nécessiter des solutions de stockage adaptées. L'option « Ne sais pas », qui ne collecte que 2 réponses, indique que la plupart ont une bonne

compréhension de ce qu'est une donnée de recherche. L'option « Autres » permet d'identifier d'autres types de données qui ne figurent pas parmi la liste fournies telles que les données physiologiques, spectrométriques, des relevés de mesure et des séquences ADN.

Une analyse plus fine permet de dégager des spécificités par domaine mais aussi des similarités entre eux. La vidéo, l'audio, les données matérielles et physiques, les notes dessinées, les bases de données, les images, archives/sources historiques et les entretiens sont des données communes aux six domaines (Fig.5). Il est intéressant de souligner que les domaines IA et S produisent tous les types de données proposées dans la liste. Les domaines DAV et ES, affichent la même tendance à quelques exceptions près.

En dehors des données numériques et textuelles, le domaine IA mentionne en première position les images/photographies suivies des logiciels et du code (et d'une manière plus large les programmes informatiques). Si il n'est pas surprenant de voir certains types de données émerger en prédominance dans certains domaines, d'autres résultats sont plutôt surprenants. Par exemple, les archives/sources historiques, un type de données habituellement

Figure 5 - Types de données collectées/générées (projet actuel), n=1029



manipulé par les chercheuses et chercheurs en sciences humaines et sociales, sont observées, et en majorité, dans le domaine IA. Le domaine DAV produit davantage de données textuelles suivies d'images/photographies et de données matérielles et physiques. Le domaine MAS présente une tendance comparable, à l'exception des données matérielles qui sont générées en quantité limitée. Proportionnellement au nombre de répondant·es, l'utilisation de l'audio est identique dans les domaines MAS et TS. Dans le domaine ES sont mentionnés en priorité, les entretiens suivis à parts égales les bases de données et données statistiques. Les domaines S et TS affichent plus ou moins les mêmes types de données avec en première position les entretiens.

3.2 Des données hétérogènes : défis et opportunités

Comme souligné plus haut, l'hétérogénéité des types de données par domaine mais aussi par chercheur·euse présente certains défis en matière de gestion, de standardisation, de confidentialité, de sécurité des données et nécessite de mettre en place des mesures organisationnelles et structurelles adaptées en vue de rendre les données « FAIR ».

Toutefois, cette diversité et les similarités observées entre les domaines, offrent des opportunités de collaborations entre chercheur·euses et le personnel de soutien qui peuvent bénéficier des connaissances et expertises en matière de gestion et partage des données au-delà de leur propre communauté scientifique ou institution.

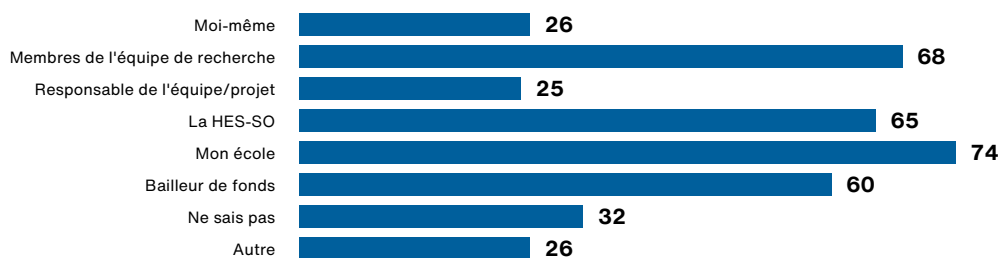
3.3 La propriété des données : la complexité de l'attribution

Selon la convention intercantonale sur la Haute École Spécialisée de Suisse Occidentale (HES-SO), « [l]es hautes écoles sont titulaires des droits de propriété intellectuelle portant sur toutes les créations intellectuelles ainsi que les résultats de recherches obtenus dans l'exercice de leurs fonctions par les personnes ayant une relation contractuelle de travail avec ces dernières ».¹⁴

A première lecture du graphique ci-dessous, cette réglementation institutionnelle est majoritairement connue des chercheuses et chercheurs puisque « Mon école » apparaît en première position, collectant 74 réponses (Fig.6). Cependant, une analyse plus fine révèle une situation plus nuancée. En effet, parmi les 376 réponses recueillies, le cumul des choix « Moi-même », « Membres de l'équipe de recherche », « Responsable de l'équipe/projet », « La HES-SO » et « Ne sais pas » représente un total de 216 réponses soit 57% des répondant·es qui semble ne pas connaître cette réglementation.

La titularité de la propriété des données n'est en réalité pas exclusive et peut-être complexe. Le guide intitulé *Les grands principes juridiques et les recommandations de la HES-SO* met en lumière cette complexité. Par exemple, lorsque la chercheuse ou le chercheur collabore avec des tiers, des accords sont fixés en amont du projet entre les parties prenantes définissant l'attribution de la propriété.¹⁵ Autre exemple, si la ou le partenaire fournit les données pour le projet, la propriété revient généralement au partenaire et non à l'institution de rattachement.

Figure 6 - Les données collectées/générées (projet actuel) appartiennent à, n=376:



14 Voir l' Art. 15 de la Convention intercantonale sur la Haute École Spécialisée de Suisse Occidentale (HES-SO), p.5.

15 Cruchon, B.; Lucas, I. (2023). Ouverture des données de la recherche. Les grands principes juridiques et les recommandations de la HES-SO, p. 15.

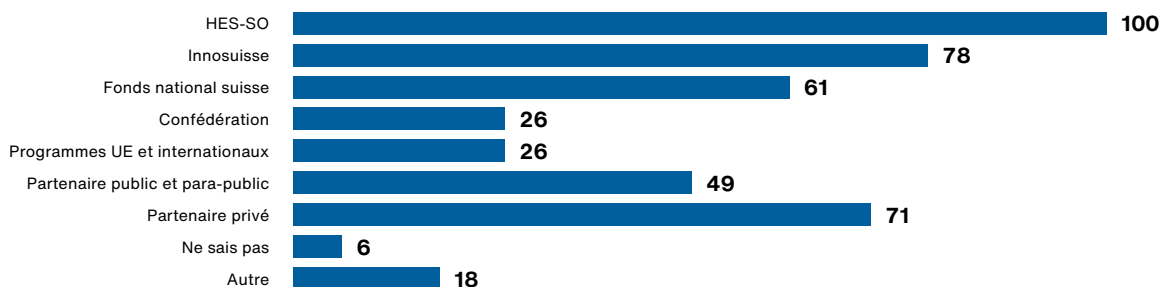
Les bailleurs de fonds (option citée 60 fois), comme le Fonds national suisse, ne possèdent pas les droits de propriété sur les données. En revanche, cette situation diffère pour d'autres sources de financement, notamment les partenaires privés qui sont les principaux contributeurs financiers des participant·es à l'enquête (Fig.7). Il n'est donc pas étonnant d'observer un nombre significatif de réponses provenant de cette option.

Le champ de réponse libre « Autre » met en évidence d'autres cas. Les chercheur·es d'autres institutions et les partenaires publics peuvent également détenir la propriété en fonction des accords fixés en amont.

Un autre cas qui émerge du champ libre, est celui des participant·es non-scientifiques au projet de recherche. Si il n'y a que deux occurrences, la question de l'attribution de la propriété est pertinente notamment dans le cadre de la recherche participative. Comme le souligne un récent rapport, la propriété des données dans ce contexte est complexe et doit être clairement définie et communiquée aux participant·es en amont du projet.¹⁶

Enfin, les multiples réponses indiquent qu'une grande partie identifie plusieurs propriétaires des données. Cette observation peut être interprétée, en fonction de l'analyse précédente, de deux manières : soit une méconnaissance du cadre réglementaire, soit une forte tendance au partage de la propriété des données.

Figure 7 - Bailleur(s) de fonds (projet actuel), n=435



16. Maussang, K; Jouguet, H; Jouneau, T.; Martin, J-F; Larroussen, N. (2023). Données et recherches participatives: Enjeux et recommandations issues d'exemples de projets de recherches participatives. Comité pour la science ouverte, p.50.

4. La gestion des données de recherche

4.1 La question de la responsabilité

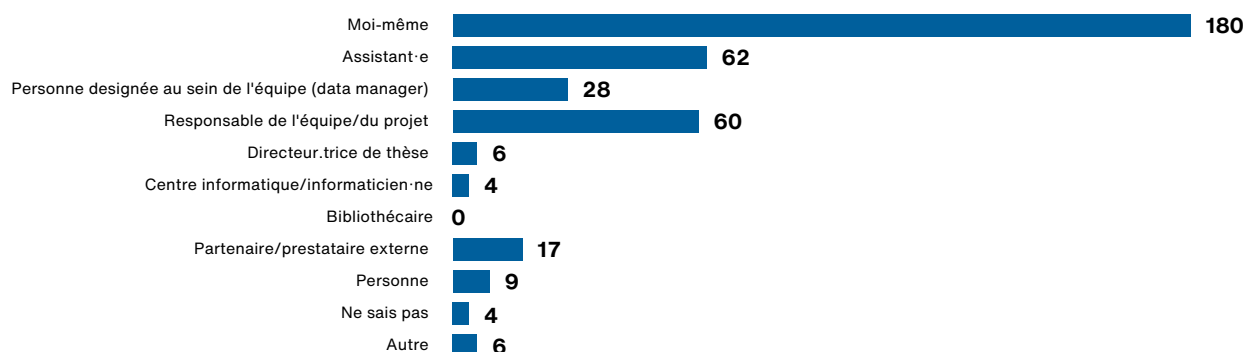
Comme l'illustre le graphique ci-dessous, la gestion des données de recherche est dans la plupart des cas, assumée par les chercheur·euses (Fig.8). Toutefois, les multiples réponses collectées (376) indiquent que cette responsabilité est souvent partagée. Plus rarement, on observe que cette charge est confiée au « Centre informatique/informaticien·ne », une option qui recueille 4 réponses, avec un résultat identique pour l'option « Ne sais pas ». 9 ont indiqué que « Personne » n'est désigné à cette tâche. Dans 17 cas, la responsabilité est déléguée à un·e « Partenaire/prestataire externe ».

La gestion des données de recherche concerne l'ensemble du cycle de vie des données, c'est-à-dire des opérations de collecte, production, description, stockage, traitement, analyse, diffusion et archivage des données de recherche. A l'ère de l'ouverture des données de recherche, la question de la responsabilité nécessite d'être clarifiée dès la phase de conception du projet et surtout lorsque des données personnelles sont traitées. Comme le précise le guide *Ouverture des données de la recherche. Les grands principes juridiques et les recommandations de la HES-SO*, « le·la responsable de traitement est généralement

le·la responsable du projet ».¹⁷ Cette responsabilité entraîne l'obligation de respecter la législation actuelle, les obligations contractuelles ainsi que les conditions des bailleurs de fonds. Cette responsabilité peut-être déléguée, comme le suggère certaines réponses, toutefois sous certaines conditions : « [I]orsque le traitement de données personnelles (collecte des données, analyse, etc.) est confié à un tiers, la HES-SO recommande que ce traitement soit défini contractuellement entre l'équipe de recherche et le sous-traitant ».¹⁸

Si la ou le responsable du projet a la charge de faire en sorte que les données de recherche soient gérées tout au long de leur cycle de vie et de respecter la législation actuelle, elle ou il n'en a pas pour autant la responsabilité juridique. La notion de « responsable de traitement » est définie par les différentes lois cantonales comme : « l'entité qui, seule ou conjointement avec d'autres, détermine les finalités et les moyens du traitement de données ».¹⁹ En d'autres termes, l'institution d'affiliation est la responsable de traitement. Toutefois, cette question de la responsabilité doit être définie et réglementée par chaque institution.

Figure 8 - Responsable de la gestion des données de recherche, n=376



17 Cruchon, B.; Lucas, I. (2023). Ouverture des données de la recherche. Les grands principes juridiques et les recommandations de la HES-SO, p. 7.

18 Ibid., 15.

19 Art. 14 let. f. Convention JUNE sur la protection des données et la transparence des 8 et 9 mai 2012. Se référer également à la LIPDA (Art. 3 al. 6) qui précise que la ou le responsable du traitement « est l'autorité, le service ou tout autre organisme public ou privé qui, seul ou conjointement avec d'autres, dans l'accomplissement de ses tâches légales, détermine les finalités et les moyens du traitement de données personnelles ».

4.2 Le Data Management Plan (DMP) : un standard à la HES-SO?

Le DMP est un outil qui permet de planifier la gestion des données. Il joue un rôle essentiel dans la mise en œuvre des principes de « FAIRification » de la recherche, favorisant la découverte, l'accessibilité, l'interopérabilité et la réutilisation des données. 96 répondant·es soit 43% se sont déjà livrés·es à l'exercice de la rédaction d'un DMP (Fig.9). Si ce n'est pas encore devenu un standard dans le paysage de la recherche HES-SO, ce chiffre est satisfaisant par rapport aux proportions présentées dans des enquêtes similaires.²⁰ Sur ces 96 répondant·es, 85 indiquent que la rédaction d'un DMP découle d'une exigence de leur organisme de financement. Ce résultat souligne l'influence et l'importance des bailleurs de fonds dans la promotion de bonnes pratiques de

gestion des données au sein de la communauté scientifique. Toutefois, il est important de noter que seul un nombre restreint de bailleurs de fonds, parmi lesquels le FNS et Horizon Europe, imposent actuellement cette exigence.

Dans la majorité des cas, le DMP est considéré comme étant plutôt utile (46 répondant·es) à tout à fait utile (17). Si les répondant·es n'ont pas toutes et tous l'expérience du DMP, d'une manière générale, elles ou ils s'accordent sur le fait que la planification de la gestion des données est importante (Fig.10).

Figure 9 - Rédaction d'un DMP au cours de sa carrière, n=223

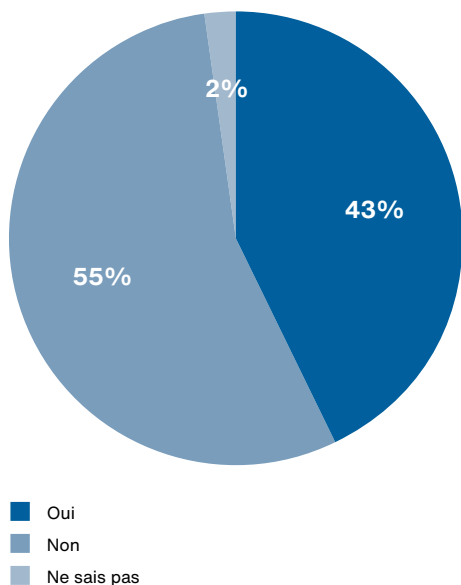
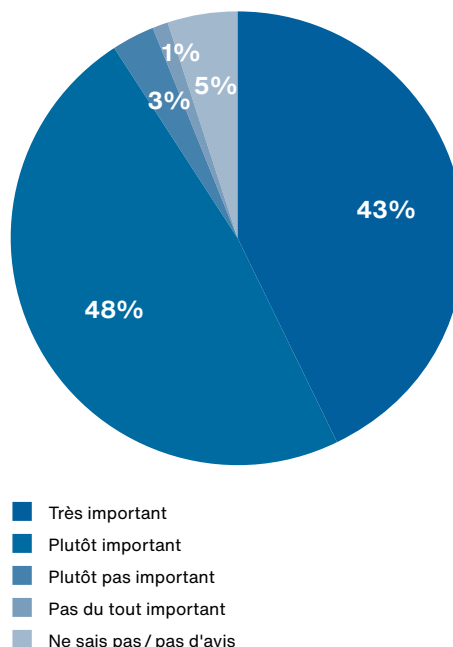


Figure 10 - Importance accordée à la planification de la gestion des données de recherche, n=223

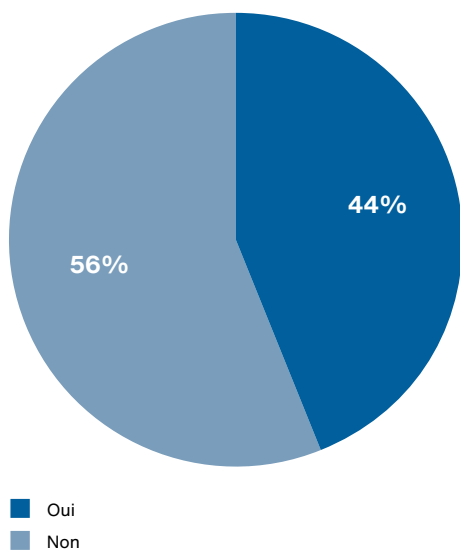


20 Voir par exemples : Herbet, M-E. ; Dugué, A ; Billy, R. ; Maire, M. et al. (2023). Gestion et ouverture des données de recherche en SHS : une enquête auprès des communautés de Lyon 2, Lyon 3 et la MSH Lyon St-Étienne, p. 18 ; Cellule Data (2022). Rapport sur l'enquête sur les usages et les besoins pour la gestion des données de la recherche sur le site de l'université Grenoble Alpes, p. 50.

4.3 L'affaire des principes « FAIR »

Comme souligné en introduction, les principes « FAIR » représentent un ensemble de lignes directrices pour gérer les données de recherche, avec pour objectif de rendre les données de recherche faciles à trouver, accessibles, interopérables, et réutilisables tant par les êtres humains que par les machines. Les principes « FAIR » sont les piliers de l'ouverture des données. Si certains de ces principes s'appliquent lors du partage, d'autres interviennent dès le début du projet. Une majorité de répondantes et répondants (56%) au questionnaire ne sont pas familiar·ères avec ces principes (Fig.11). Si on rapporte de chiffre aux 3264 chercheuses et chercheurs de la HES-SO, il y a environ 1815 personnes à former ou à sensibiliser sur ces incontournables principes. Notons par ailleurs, que sur les 7% des doctorant·es qui ont répondu à l'enquête, environ la moitié connaissent ces principes, un résultat plutôt encourageant mais qui invite à renforcer la sensibilisation sur les avantages d'implémenter ces principes tout au long de la thèse de doctorat.

Figure 11 - Connaissance des principes FAIR, n=223



4.4 Volumétrie, stockage et sauvegarde

4.4.1 Volumétrie

La plupart des répondantes et répondants produisent moins de 50 Go de données, ce qui représente un volume soutenable (Fig.12). Les volumes les plus importants de données, supérieurs à 1 To, concernent 17 répondant·es sur les 223 ayant répondu à la question, soit environ 7%. Un quart environ se situe dans une zone mitoyenne supérieure à 100 Go mais inférieure à 1 To. Les besoins les plus importants de stockage, lorsqu'ils existent, sont surtout exprimés dans le domaine IA et dans une moindre mesure dans les domaines ES, S et TS (Fig.13).

Figure 12 - Volume de données (projet actuel), n=223

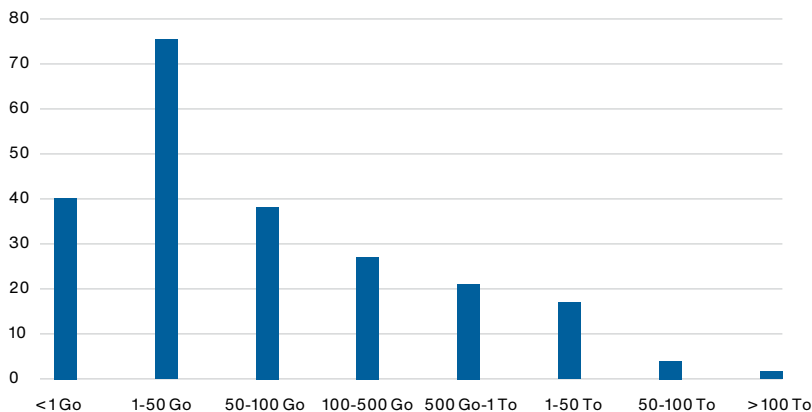
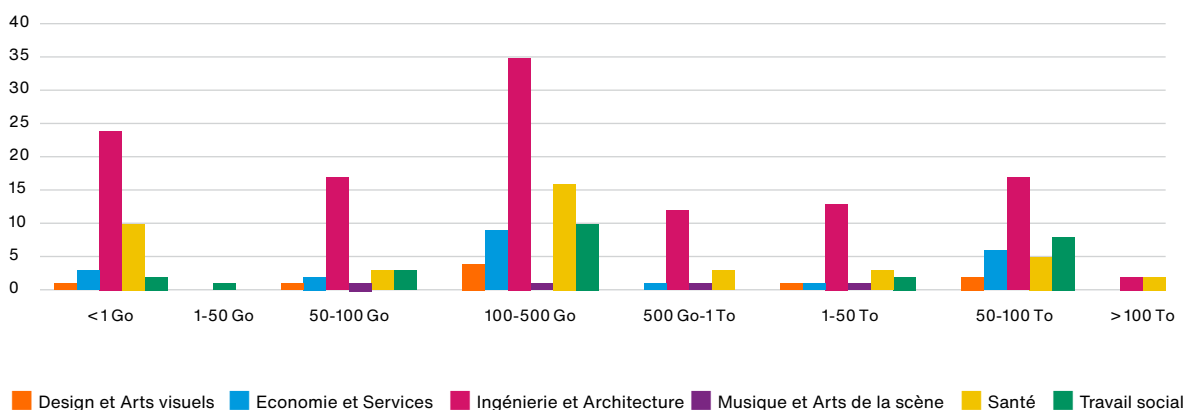


Figure 13 - Volume de données par domaine (projet actuel), n=223



4.4.2 Stockage

La question « Où stockez-vous vos données de recherche pour le projet actuel ? » obtient 584 réponses ce qui indique une tendance vers l'utilisation de plusieurs solutions de stockage (Fig.14). Le disque dur de l'ordinateur professionnel est le plus fréquemment cité. Si cette solution paraît fragile, elle est dans la majorité des cas combinée à d'autres solutions plus sécurisées. Les solutions sécurisées telles que le serveur institutionnel et Switchdrive figurent en deuxième et troisième position. Les espaces de stockage de Microsoft (OneDrive, Sharepoint et Teams) sont eux aussi souvent mentionnés avec une préférence pour Teams (59). Les solutions fragiles sont peu employées. Si le disque dur de l'ordinateur personnel apparaît 31 fois, seules 2 répondant·es l'utilisent comme unique solution de stockage. Les clouds privés sont également peu cités. Dropbox revient seulement 11 fois et parmi les réponses non affichées, Google Drive apparaît moins de cinq fois. Cela indique cependant des pratiques qui ne sont pas optimales. Dans l'option « Autres », des solutions telles que GitLab et REDCap sont citées 6 fois pour la première et 2 fois pour la seconde.

De façon générale, les répondant·es privilégient des solutions sécurisées et multiplient les supports de stockage ce qui montre une intégration globale des bonnes pratiques. Toutefois, les résultats mettent en lumière une grande diversité dans les choix de solutions de stockage ce qui souligne la nécessité d'harmoniser et de normaliser les pratiques.

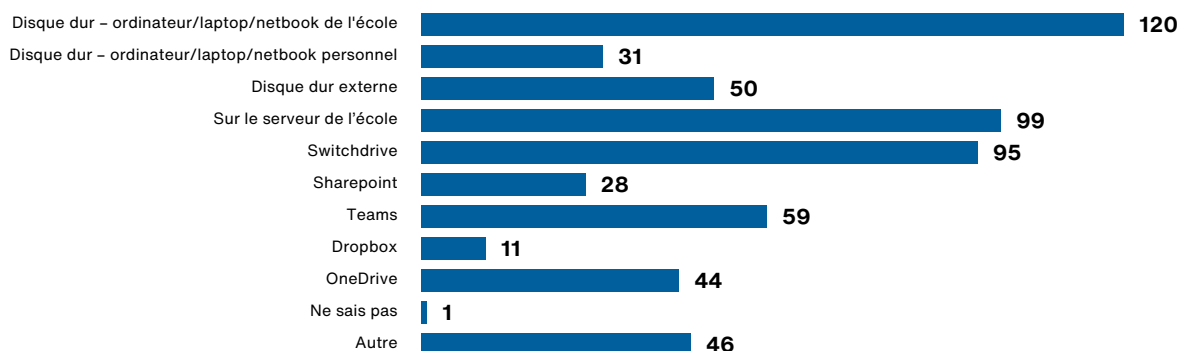
4.4.3 Sauvegarde

La majorité des répondant·es (56%) effectue des sauvegardes de toutes leurs données de recherche, et 26% pour une partie de leurs données ce qui montre que les pratiques de sauvegarde sont bien intégrées. Le pourcentage de réponses entre celles et ceux qui ne sauvegardent pas (ou qui n'ont pas l'intention de le faire) ou qui ne savent pas atteint un total de 18%.

En matière de fréquence, 32% indiquent sauvegarder tous les jours. En deuxième position, on observe un pourcentage significatif de participant·es (14%) qui ne savent pas exactement à quelle fréquence elles ou ils sauvegardent leurs données suivi de 10% indiquant une fréquence irrégulière de sauvegarde, 5% tous les 6 mois et 1%, seulement une fois par an soit 25% des répondant·es qui manquent de routine ou de système en place pour sauvegarder leurs données de façon optimale. Cette variation dans les pratiques de sauvegarde souligne un besoin potentiel de sensibilisation et d'amélioration des stratégies de sécurité des données.

Les « lieux » de sauvegarde sont aussi multiples totalisant 306 réponses. Le serveur institutionnel est cité 87 fois suivi en deuxième et troisième positions par l'ordinateur personnel/professionnel (74) et du disque dur externe (54). Les clouds privés tels que Dropbox sont fréquemment cités (46) ce qui souligne la nécessité de communiquer sur la sécurité des données sauvegardées de cette manière.

Figure 14 - Supports de stockage utilisés (projet actuel), n=584



4.4.4 Conservation des données de recherche

La tendance générale révélée par les réponses à la question « Conservez-vous vos données de recherche à la fin d'un projet? » met en lumière une forte propension à la conservation des données. En effet, 75% ont affirmé conserver l'intégralité de leurs données, suggérant l'intégration de protocole rigoureux pour assurer la préservation et l'accessibilité des données de recherche au sein des hautes écoles. Cependant, en vue de garantir la durabilité et d'optimiser les coûts, il est légitime de se questionner sur les meilleures pratiques à adopter pour la conservation des données.

4.5 Les données personnelles et leur sécurité

Le graphique ci-dessous indique qu'une majorité (58%) gère, traite ou collecte des données personnelles ou sensibles (Fig.15). Ce pourcentage significatif inclut tous les domaines: 1 répondant·e du côté des domaines DAV et MAS, 12 pour le domaine ES, 56 pour le domaine IA, 36 pour le domaine S et 24 du côté du domaine TS. Ce résultat met en lumière d'une part une prise de conscience significative quant à la nature des informations manipulées et d'autre part, l'importance et le besoin d'une gestion appropriée de ces données, incluant des mesures techniques et organisationnelles de protection pour assurer leur sécurité et être conforme aux exigences légales. Toutefois, si une majorité des répondant·es indiquent sécuriser leurs données, le pourcentage de celles et ceux qui ne le font pas (5% et 21%) ou qui ne savent pas (11%), est significatif: au total 37% soit 84 répondant·es (Fig.16). En d'autres termes, ces résultats suggèrent une lacune dans la compréhension ou la mise en œuvre et la communication sur les mesures de sécurité au sein des institutions. Parmi les mesures de sécurité proposées, le mot de passe arrive en première position (107), suivi de l'utilisation de serveur sécurisé (91). A parts plus ou moins égales, on retrouve la mise en place de contrôle d'accès (46) et de logiciel de sécurité anti-virus (44). Le cryptage des données est cité 19 fois, ce qui est faible par rapport au 129 répondant·es (58%) qui ont indiqué collecter/traiter des données à caractère confidentiel.

Figure 15 - Collecte et gestion de données à caractère confidentiel (e.g. données personnelles ou sensibles), (projet actuel), n=223

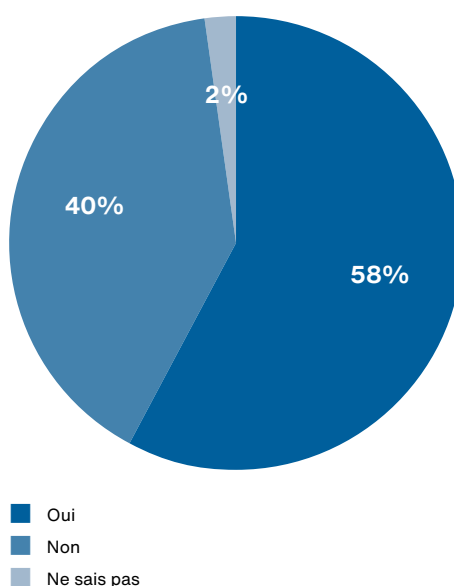
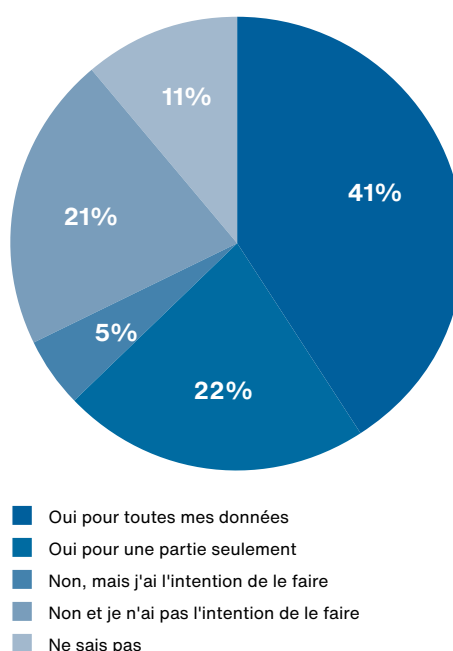


Figure 16 - Les données sont sécurisées (contrôle d'accès, cryptage, etc.), (projet actuel), n=223



4.6 Les métadonnées et les standards : loin d'être une norme

Les métadonnées soutiennent les principes « FAIR ». Elles jouent un rôle clé dans l'accès aux données et à leur réutilisation. Toutefois la description et la documentation des données ne sont pas encore des pratiques courantes (Fig.17). Parmi les 34% (76 réponses) qui décrivent et documentent leurs données de recherche, une grande majorité n'utilisent pas de standards de métadonnées (Fig.18). Ce résultat est à corréler avec celui obtenu sur la méconnaissance des principes « FAIR ». Il reflète un besoin de formation sur l'importance des standards de métadonnées et sur la manière de les mettre en œuvre. Ces standards permettent de décrire les données de façon riche et précise, et assurent l'interopérabilité sémantique. L'absence d'utilisation de ces standards peut entraîner des problèmes d'interopérabilité et de compréhension des jeux de données, limitant ainsi leur potentiel de réutilisation.

Figure 17 - Utilisation de métadonnées pour documenter, décrire les données, n=223

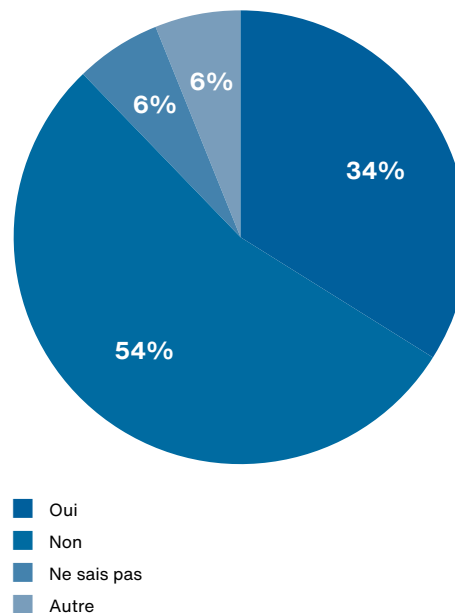


Figure 18 - Utilisation de standard(s) de métadonnées (si utilisation de métadonnées), n=76



5. Vers le partage

5.1 L'attitude face au partage

Dans l'ensemble, une grande majorité des répondant·es sont tout à fait (31%) ou plutôt favorable (34%) au partage des données de recherche soit au total 65% (Fig.19).

Le nombre substantiel de réponses recueillies à la question « Quelle(s) est (sont) la(es) raison(s) qui motive(nt) votre choix de partager vos données de recherche ? » soit 659 réponses, reflète une conscience des avantages multiples associés au partage de données chez celles et ceux qui y sont favorables (Fig.20). Le partage comme opportunité de prouver l'intégrité et la qualité de la recherche arrive en tête de classement (89) suivie de près par la possibilité d'augmenter la visibilité et l'impact de ses résultats (83). Une autre raison fréquemment cochée est celle du partage des données comme pratique courante au sein de la communauté scientifique, citée 80 fois ce qui souligne l'influence de la communauté scientifique sur l'adhésion des chercheur·euses au partage des données. D'une manière globale, ce résultat reflète l'acceptation généralisée du partage en tant que norme établie.

Le champ de réponse libre offre la possibilité d'identifier une raison qui motive le partage et qui n'a pas été proposée dans le questionnaire, à savoir celle de réutiliser les données de participant·es aux enquêtes dans le but de réduire la fréquence des sollicitations des participant·es.

Figure 19 - Favorable au partage des données de recherche, n=223

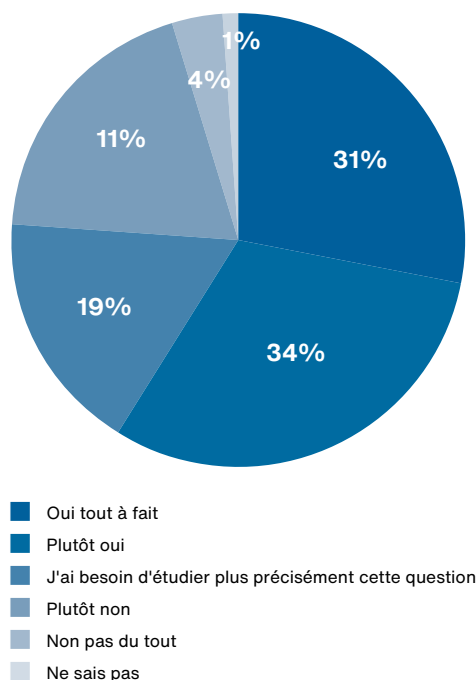


Figure 20 - Raison(s) qui motive(nt) le choix de partager ses données de recherche, n=659



5.2 Les freins

Parmi les 35% qui ne sont pas favorables au partage ou qui sont indécis-es, plusieurs raisons sont citées, affichant un total de 93 réponses provenant de tous les domaines, à l'exception du domaine DAV qui n'enregistre aucune réponse.

Les questions éthiques et de confidentialité arrivent en première position, raison majoritairement cochée par les domaines IA, S et TS (Fig.21). Ce résultat indique que les chercheur·euses accordent une priorité élevée aux enjeux éthiques dans leurs décisions individuelles et met en lumière l'importance de trouver un équilibre entre partage et éthique des données. D'autres raisons citées à parts égales et qui arrivent en deuxième position, concernent les exigences des bailleurs de fonds/éditeur/institution, le risque de fausse interprétation des données et sans surprise les raisons juridiques et contractuelles. Comme le montre le graphique ci-dessous, 66% des projets en cours sont régis par des contrats (e.g. de confidentialité, de consentements), ce qui constitue un obstacle majeur au partage des données de recherche (Fig.22). D'une manière générale, les motifs de non-partage sont indépendants de la volonté des chercheur·euses et reflètent dans certains cas, une méconnaissance des conditions qui encadrent le partage des données.

Figure 22 - Projet actuel sous contrat (de confidentialité, ou formulaire de consentement, de non-divulgateion, etc.), n=223

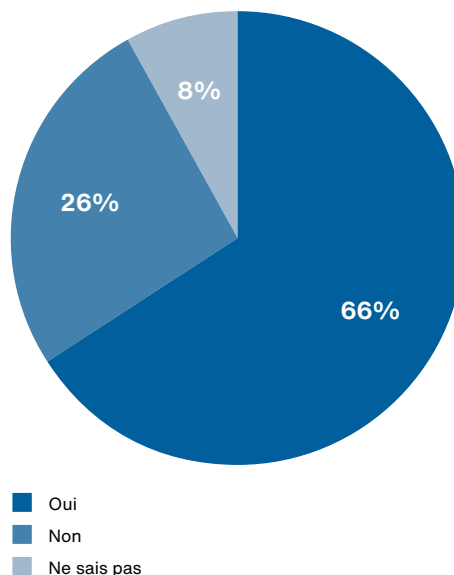
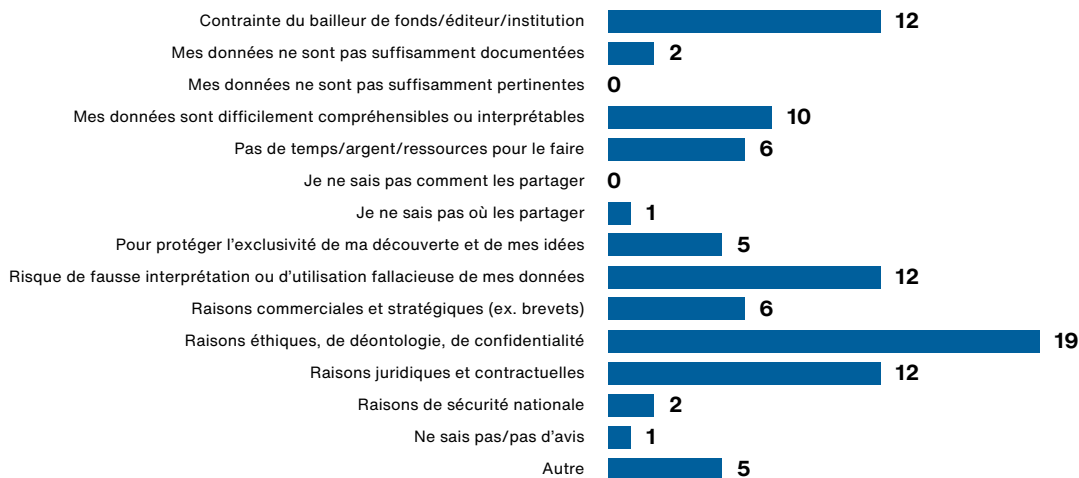


Figure 21 - Raison(s) qui freine(nt) le partage des données de recherche, n=93



Parmi l'éventail de raisons proposées, certaines représentent des obstacles substantiels au partage des données et sont difficilement contournables (e.g. raisons juridiques, éthiques et contractuelles). Toutefois, certaines (listées dans le tableau ci-dessous), peuvent être dépassées par la mise en œuvre de certaines mesures (Tab.2).

Tableau 2 - Freins au partage des données et exemples d'actions pour réduire voire éliminer ces freins.

Freins au partage	Exemples d'actions
Mes données ne sont pas assez documentées	Proposer des formations sur la documentation et les métadonnées
Mes données sont difficilement compréhensibles ou interprétables	Fournir des conseils personnalisés + Proposer des formations sur la documentation et les métadonnées
Pas de temps/argent/ressources pour le faire	Déblocage de fonds et renforcer la communication sur les fonds alloués par le FNS et d'autres bailleurs de fonds tels qu'Horizon-Europe pour la gestion des données
Pour protéger l'exclusivité de ma découverte et de mes idées Risque de fausse interprétation ou d'utilisation fallacieuse de mes données	Renforcer les formations sur les licences et de manière plus large sur la propriété intellectuelle
Je ne sais pas où les partager	Renforcer la communication sur les inventaires disponibles en ligne de dépôts de données « FAIR » : liste HES-SO , re3data.org , etc.

5.3 L'expérience du partage

Le partage des données n'est pas du tout intégré dans la routine des chercheur·euses (Fig.23). Une grande majorité d'entre elles et eux (67%) ne s'y sont jamais livré·es. 12 % indiquent ne pas savoir répondre à cette question ce qui souligne une certaine méconnaissance de cette pratique. Parmi les 21% qui ont déjà l'expérience du partage, sont représentés tous les domaines avec une prédominance du côté IA suivi du domaine S puis du TS (Fig.24).

Parmi celles et ceux qui se sont déjà livré·es à l'exercice du partage (47), on relève une forte tendance à se tourner vers des dépôts généralistes et nationaux : Zenodo (15), SWISSUbase (8), Yareta (5). Moins fréquemment cités, Figshare, GitLab, GitHub ou encore OSF (Tab.3). Certain·es ont indiqué avoir partagé plusieurs fois. Tous ces dépôts sont référencés sur re3data (un registre mondial des dépôts de données de recherche) et sont recommandés par le FNS, ce qui montre une bonne connaissance des infrastructures conformes aux principes « FAIR ».²¹ Cependant, la mention de solutions telles que « Disque réseau HES-SO (1); Serveur de l'école (1); REDCap (1) et Switchdrive (2) » indique une certaine confusion entre des solutions de stockage et d'archivage et des dépôts de données.

21 Voir la liste des dépôts recommandés par le FNS et le registre en ligne re3data.

Figure 23 - Expérience du partage des données de recherche, n=223

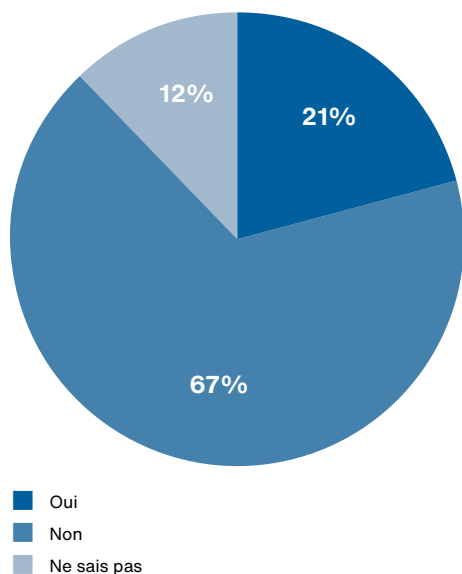
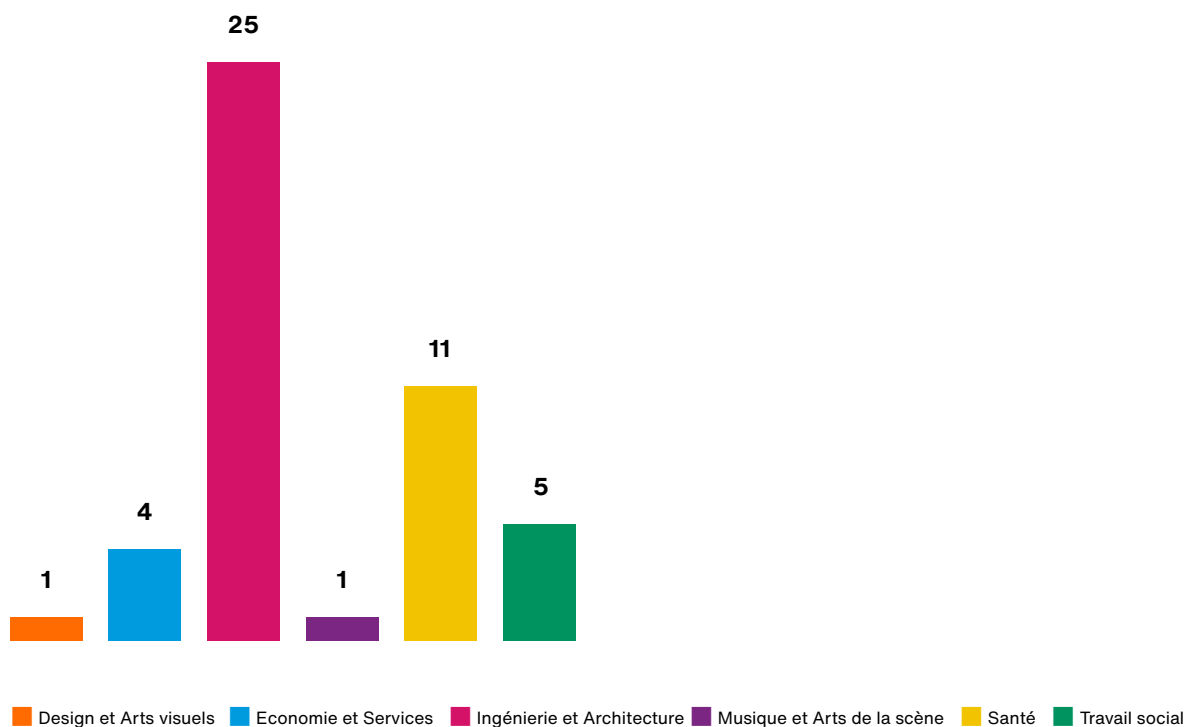


Tableau 3 - Choix des dépôts de données par domaine

Domaine/ Nom du dépôt	DAV	ES	IA	MAS	S	TS
Dryad			1			
Figshare		1			2	
GitHub	1		3			
GitLab			4			
Harvard Dataverse			1			
NCBI			1			
OSF		1	1			
SWISSUbase				1	2	5
Yareta			3		2	
Zenodo		4	9		2	

DAV = Design et Arts visuels
 ES = Economie et Services
 IA = Ingénierie et Architecture
 MAS = Musique et Arts de la scène
 S = Santé
 TS = Travail social

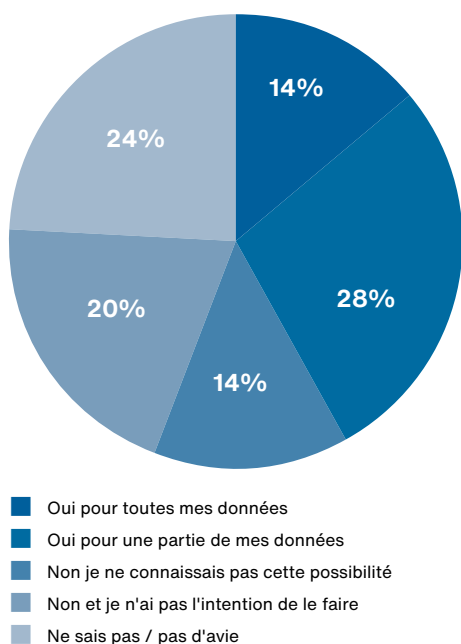
Figure 24 - Expérience du partage des données de recherche par domaine, n=47



Le partage des données est une pratique émergente qui n'est exigée que par une minorité de bailleurs de fonds, ce qui pourrait expliquer le pourcentage d'expérience relativement faible dans ce domaine. Cependant, il est pertinent de se questionner sur les intentions à venir. A la question « Avez-vous l'intention de déposer vos données de recherche sur un dépôt pour le projet en cours ? », 24% ne savent pas, 20% n'en ont pas l'intention et 14% ne connaissent pas cette possibilité (Fig.25). En d'autres termes, une majorité (58%) des répondant·es n'intègrent pas le partage dans leur road map actuelle.

Pour résumer, un état contradictoire se dégage de l'enquête : les répondant·es dépeignent un climat globalement favorable à l'ouverture des données, mais se montreraient plutôt réticent·es dans la pratique d'où la nécessité de renforcer la communication sur les avantages du partage mais aussi de continuer et/ou d'intégrer dans les politiques de financement des incitations au partage en cas d'absence d'exigences de certains bailleurs de fonds.

Figure 25 - Intention de partager ses données de recherche (projet actuel), n=223



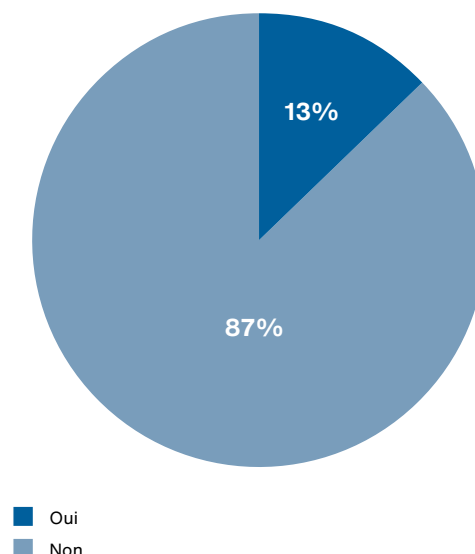
5.4 Focus sur la réutilisation

Si le partage des données n'est pas encore une pratique courante, le pourcentage de celles et ceux qui ont déjà réutilisé des données de recherche est encourageant. En effet 36% des répondant·es ont eu recours à la réutilisation de données de recherche et 22% ont indiqué leur intention de le faire, ce qui suggère une évolution positive vers une culture de partage et d'utilisation collaborative des informations scientifiques. Cette tendance laisse entrevoir un potentiel prometteur pour l'accroissement futur du partage de données au sein de la communauté scientifique.

5.5 Connaissez-vous le Data paper?

Aussi appelé Data article ou Data descriptor, le Data paper est un article scientifique évalué par les pairs et citable qui permet de valoriser les jeux de données et par extension les résultats de recherche. Une large majorité des répondant·es (87 %), ne connaissent pas le « Data paper », un résultat qui concorde avec l'expérience limitée du partage (Fig.26). Cela traduit une réalité où la formalisation et la reconnaissance institutionnelle ou communautaire des efforts liés au partage des données de recherche ne sont pas (encore) intégrés dans la culture et les pratiques actuelles.

Figure 26 - Connaissance du Data paper, n=223



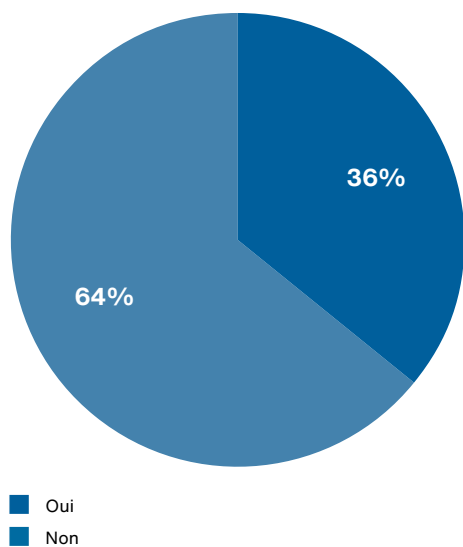
6. Soutien et besoins

6.1 Etat des lieux

Le graphique ci-dessous montre une absence de recommandations ou de lignes directrices internes pour la gestion des données (Fig.27). Toutefois, le champ libre réservé aux commentaires qui permet d'identifier quelques outils et mesures déjà mis en place à savoir des guides de bonnes pratiques. Ces commentaires précisent dans certains cas que ces outils sont peut-être existants mais pas ou peu communiqués.

De manière générale, une minorité (37%) des participant·es disent être informé·es de temps en temps sur la question des données de recherche et 27% de façon régulière. Ces pourcentages sont suivis par celles et ceux qui considèrent être rarement (21%) ou jamais informé·es (10%).

Figure 27 - Existence de lignes directrices, recommandations internes pour la gestion des données de recherche, n=223

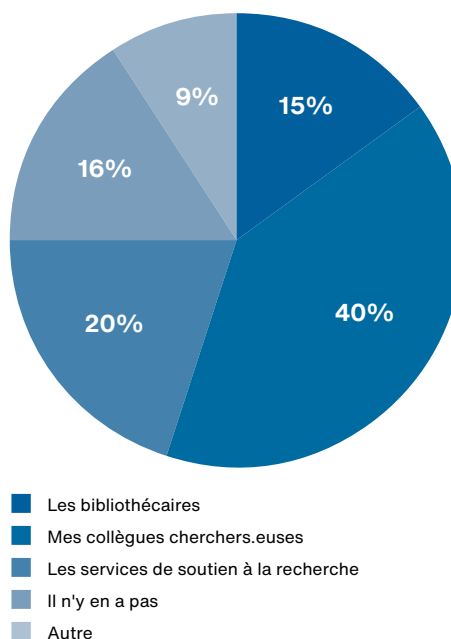


22 Pour plus d'information sur ce projet, consulter le site de la Médiathèque de la HES-SO Valais.

6.2 Les collègues comme premier point d'appui

De l'aide est recherchée auprès de différentes actant·es comme l'illustre le graphique ci-dessous (Fig.28). En première position, on retrouve les collègues chercheur·euses (40%) suivi des services de soutien à la recherche (20%). Une minorité significative (16%) indique une absence de soutien ou de services ce qui n'est pas surprenant puisque l'ORD est une nouvelle culture qui nécessite de mettre en place des structures et des services adaptés aux besoins des chercheur·euses. Ce pourcentage est suivi de près par les bibliothécaires qui ont jusqu'ici, un rôle clé dans le soutien à l'ouverture des publications (Open Access). Dans l'option « Autre », les « Data managers » sont mentionnés à plusieurs reprises, un des nouveaux profils qui émergent avec l'ORD. Cette réponse fait probablement référence au projet Research Data Service Network (RDSN) développé par la HES-SO Valais dans lequel un pool de Data managers intervient en soutien aux chercheur·euses par domaine depuis 2021.²²

Figure 28 - Premier contact pour le soutien à la recherche, n=223

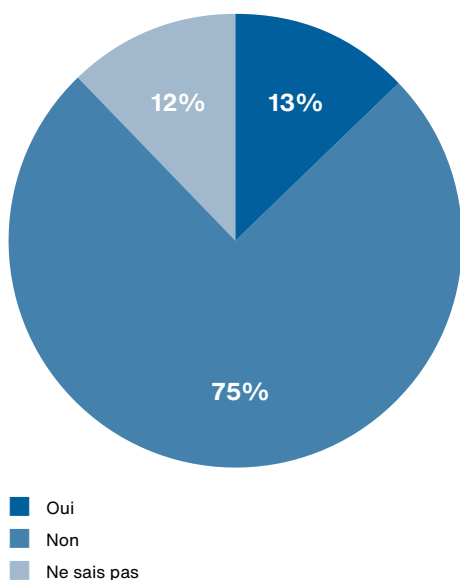


Comme mentionné, les chercheur·euses font appel à divers intervenant·es lors du processus de gestion des données en vue de leur partage et ne savent pas dans certains cas vers qui se tourner. L'ORD requiert une orientation vers des compétences spécifiques qui sont transversales. Il s'agit dans un futur proche de clarifier les rôles et communiquer davantage sur les services de soutien existants afin de mieux les soutenir dans ce processus exigeant.

6.3 Soutien financier

Parmi les répondantes et répondants, seulement 12% ont reçu un financement dans le cadre de leur projet en cours pour gérer, traiter, et/ou partager leurs données de recherche. 13% déclarent ne pas avoir connaissance de cette information (Fig.29). Il semble nécessaire de communiquer davantage sur les fonds alloués tels que le FNS qui prend en charge jusqu'à CHF 10'000.- pour les activités relatives à la préparation des données en vue de l'archivage ainsi que les coûts éventuels d'archivage.²³

Figure 29 - Financement pour gérer les données de recherche (project actuel), n=223



23 Fonds national suisse, « Coûts éligibles » .

6.4 Les besoins en formation

62% des répondantes et répondants ont indiqué ne jamais avoir suivi de formations sur la gestion des données de recherche. En termes de besoins, les formations sur la rédaction du DMP, la documentation et la protection des données se hissent en tête du classement suivies de près par l'archivage et les basiques de la gestion des données (Fig.30). La diffusion des données, encore peu répandue comme nous avons pu le voir, ferme la marche. Les multiples réponses indiquent que des formations sont souhaitées pour l'ensemble des tâches du cycle de la gestion des données et par tous les domaines (Fig.31). Toutefois si les formations DMP sont largement convoitées par le domaine IA, ce sont d'autres formations qui sont sollicitées dans les autres domaines :

- DAV et S sur les normes et standards (formats de fichiers, métadonnées)
- ES sur l'anonymisation, la protection des données
- MAS et TS sur les questions de stockage, d'archivage et préservation à long terme.

Les formats souhaités sont quant à eux variables, avec une préférence pour la documentation en ligne, format suivi de près par les ateliers pratiques (Fig.32).

Du côté des doctorant·es, chaque formation listée est citée entre 5 à 8 fois (sur 15 répondant·es), le DMP arrive en tête de position avec 10 réponses. Ces résultats suggèrent une reconnaissance et une adhésion élevées des doctorant·es à la gestion des données de recherche.

Figure 30 - Formation(s) souhaité(s), n=944

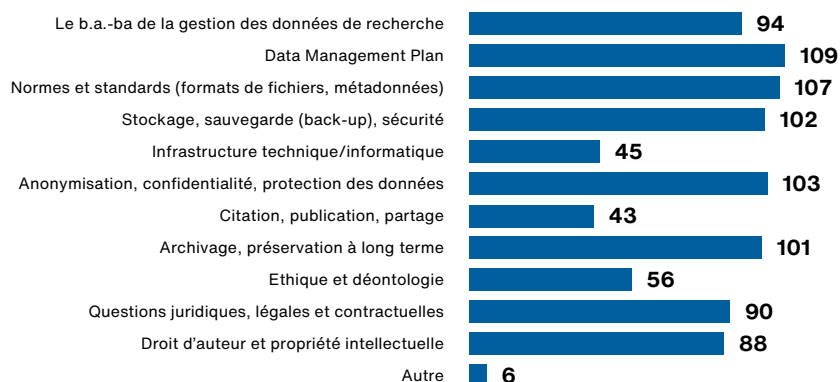


Figure 31 - Formation(s) souhaité(s) par domaine, n=944

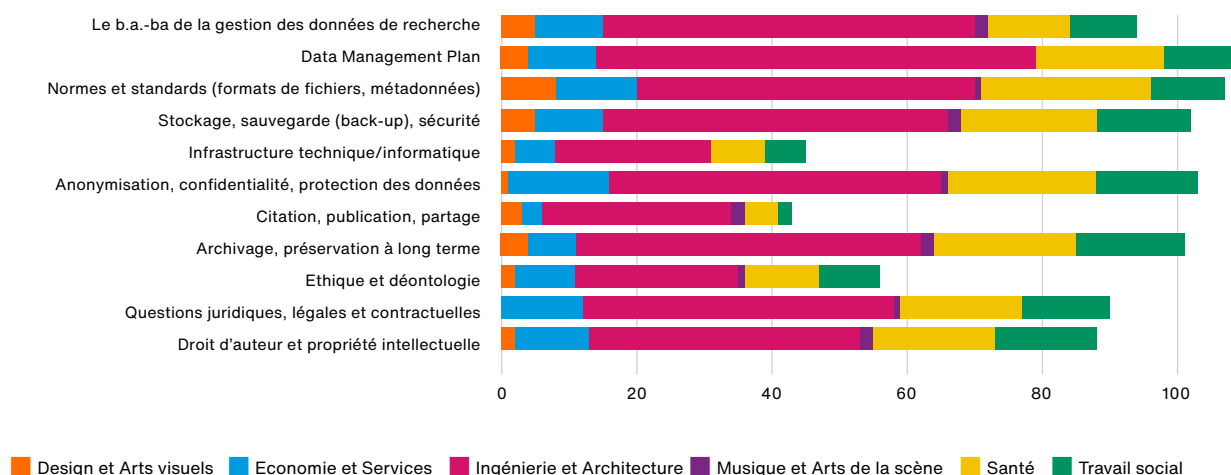
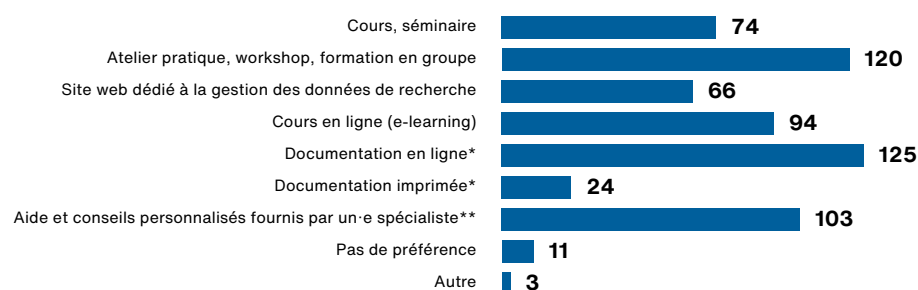


Figure 32 - Format(s) de formation(s) souhaité(s), n=620



* (fiches techniques, guides, modes d'emploi, check-lists, tutoriels, etc.)

** (archiviste, informaticien, juriste, etc.)

6.5 Voix libre aux chercheur·euses

La partie finale du questionnaire donnait l'occasion aux chercheur·euses de faire part des obstacles rencontrés, des attentes et besoins spécifiques en matière de gestion des données de recherche, de leurs suggestions et propositions d'amélioration. Au total, soixante commentaires ont été recueillis. Ils sont ici classés en 3 catégories afin de dégager les tendances et les points saillants, offrant une vue d'ensemble des perceptions et des opinions exprimées par les répondant·es : 1/Besoins ; 2/Obstacles et 3/Expériences positives (réactions positives ou réussites liées à la gestion et à l'ouverture des données).

1/ Besoins

Les besoins exprimés s'accompagnent dans la plupart des cas, d'une attitude axée sur la solution. En effet, une grande majorité des répondant·es ont formulé plusieurs propositions afin de répondre à ces besoins :

- Création d'une guideline spécifique pour les collaborations avec les partenaires privés
- Elaboration de lignes directrices fournies par les départements Ra&D des écoles afin d'éviter de chercher les informations sur de multiples sites
- Création d'une guideline et directive pour la sécurité des données : stockage, sauvegarde et archivage
- Développement d'un help-desk composé d'expert·es aux compétences transversales
- Engagement de data stewards pour la gestion des données de recherche selon les domaines disciplinaires, mais aussi selon le type de méthodologie et le type de données récoltées
- Développement de formations thématiques
- Nécessité d'une flexibilité dans l'application des principes « FAIR » en fonction des caractéristiques spécifiques des projets
- Soutien pour la gestion des données sensibles.

2/ Obstacles

- Le partenariat avec les entreprises est un frein au partage des données
- Certain·es client·es privé·es ne veulent pas partager les données de recherche générées
- Difficulté à saisir la valeur ajoutée de l'ORD
- Manque de compétences et d'expertises dans les équipes de recherche
- Principe de l'ORD pas applicable à certaines disciplines
- Les recommandations ou procédures à suivre sont fréquemment modelées sur les sciences dures et s'appliquent de manière limitée, et à d'autres types de recherche
- Pas de soutien financier pour appliquer ces nouvelles pratiques FAIR
- L'objectif principal de la recherche scientifique semble relégué au second plan avec l'ORD.

3/ Expériences positives

- Sujet d'importance, le soutien apporté aux chercheur·euses jusqu'ici est apprécié.

Pour résumer, ces commentaires reflètent les principales tendances observées dans ce rapport. Ils mettent en lumière les défis associés à l'ouverture des données, avec une prédominance marquée pour les aspects juridiques et de sécurité et le besoin d'être accompagné·e et soutenu·e dans ce processus non seulement au niveau des ressources mais aussi sur le plan financier. Le temps et les compétences spécifiques nécessaires pour gérer et partager les données de recherche sont également signalés comme primordiales pour garantir l'adoption réussie de l'ouverture des données selon les principes « FAIR ». Si les expériences positives sont rares, les bénéfices potentiels de l'ORD pourraient se révéler au fil de l'évolution des pratiques.

7. Conclusion et recommandations

Les résultats de cette enquête, enrichis par les commentaires libres exprimés par les répondantes et répondants au questionnaire, permettent de dégager plusieurs enseignements et d'émettre une série de recommandations pour accompagner au mieux la transition vers l'ouverture des données et d'en faire une norme.

7.1 Synthèse des résultats

Cet état des lieux se base sur 223 répondant·es, il est loin d'être représentatif des pratiques et des besoins en matière de gestion et d'ouverture des données de toute la communauté scientifique de la HES-SO. Toutefois, l'échantillon couvre l'ensemble des domaines, des cantons et des fonctions et permet de dresser des tendances tout à fait pertinentes pour une première analyse. Le tableau ci-dessous propose une synthèse des résultats (Tab.4) :

Tableau 4 - Synthèse des résultats

Gestion des données	Partage des données	Soutien et formation
Un flou autour de la propriété des données de recherche et de la responsabilité du traitement des données	Une majorité favorable au partage des données de recherche	Une absence de centralisation des outils et un manque de visibilité des services de soutien existants
Une gestion complexe des données : diversité des types de données manipulés par chercheur·euse > une tendance commune à tous les domaines	Une pratique de réutilisation des données de recherche encourageante	Un besoin en formation sur les différentes étapes de la gestion et du partage des données
Une pratique pas systématique du DMP	Si une majorité est favorable, on relève une faible pratique du partage des données de recherche et une faible intention de le faire dans le futur	Un besoin d'expert·es et de compétences spécifiques
Une méconnaissance des principes FAIR. Comme mentionné plus haut, il y aurait environ 1815 personnes à former ou sensibiliser sur ces incontournables principes au sein de la HES-SO (cf. 4.3)	Une série d'obstacles qui freinent le partage des données de recherche notamment d'ordre juridique, éthique et contractuel	Un manque de soutien financier pour traiter les données en vue du partage
Une adoption des bonnes pratiques en matière de sécurité des données à améliorer		
Une faible adoption des pratiques de documentation et de description des données de recherche		

7.2 Recommandations

En fonction des éléments listés ci-dessus, détaillés au fil de ce rapport, et des commentaires laissés par les répondant·es à la fin du questionnaire, cette section propose une série de recommandations.

Sensibilisation et information

Si une majorité des répondant·es sont en faveur de l'ouverture et d'une gestion planifiée des données, un travail important reste à réaliser pour informer et sensibiliser la communauté scientifique sur les avantages et les enjeux de la gestion et de l'ouverture des données de recherche.

- Augmenter la fréquence des présentations, expositions et/ou ateliers de sensibilisation sur le « FAIR data » et les bonnes pratiques en matière de gestion des données au sein des écoles
- Renforcer la communication au sein des hautes écoles et du Rectorat sur les outils disponibles en ligne et les services de soutien existants.²⁴

Formations

- Organiser des ateliers thématiques adaptés à un large éventail de types de données/ méthodologies au sein des hautes écoles
- Poursuivre l'organisation de webinaires « Open research data lunch » > les répondant·es ont formulé à plusieurs reprises le besoin d'obtenir des conseils personnalisés. Depuis l'automne 2023, un cycle de webinaires est organisé sur différentes thématiques ORD. Ces webinaires sont conçus à partir des questions des chercheur·euses afin de fournir des réponses individualisées, adaptées à la réalité du terrain
- Organiser des formations en ligne pour les jeunes chercheur·euses et doctorant·es sur les bases de la gestion des données de recherche. Informer également les doctorant·es des formations sur la gestion des données organisées par la CUSO.

Incitations, outils et services

- Produire des documents cadres sur la gestion et l'ouverture des données de recherche au sein des hautes écoles qui précisent entre autres, les rôles et les responsabilités
- Identifier les compétences et les ressources nécessaires au sein des hautes écoles afin de soutenir au mieux les chercheur·euses
- Demander le DMP en l'absence d'exigences externes
- Encourager les doctorant·es à rédiger un DMP en début de thèse, sous l'impulsion des encadrant·es de thèse
- Pérenniser la cellule Data stewardship : un·e data steward par domaine. Plusieurs commentaires laissés par les répondant·es soulignent la nécessité de centraliser les services et le besoin d'expert·es. Les data stewards sont actuellement formé·es pour devenir des expert·es de la gestion des données de recherche dans leur domaine respectif. Elles et ils œuvrent à la mise en commun des connaissances et des compétences, ainsi qu'à la formalisation de pratiques spécifiques à l'ORD par domaine, en collaboration avec le personnel de soutien
- Créer des incitations au sein des domaines et des hautes écoles sur le partage des données de recherche
- Créer un guide sur les bonnes pratiques en matière de sécurité, stockage et archivage des données de recherche
- Créer/mutualiser un poste de DPO par canton (pour la recherche spécifiquement).

Soutien financier

- Prévoir un budget pour la mise en ouverture des données de recherche au sein des hautes écoles et/ou des domaines > pour les projets non financés par le FNS ou Horizon-Europe qui allouent des fonds pour l'ORD.

²⁴ Depuis début 2023, deux guides et autres ressources sont à disposition sur les pages ORD du site de la HES-SO, onglet « Ressources et formations ».

Annexe



ENQUÊTE PRATIQUES ET BESOINS DES CHERCHEUR·EUSES DE LA HES-SO EN MATIÈRE DE GESTION ET D'OUVERTURE DES DONNÉES DE RECHERCHE

Question avec un * = obligatoire

PROFIL

1. *Merci d'indiquer à quelle Haute école vous êtes rattaché.e (si vous êtes rattaché.e à plusieurs hautes écoles, veuillez indiquer celle à laquelle vous consacrez actuellement le plus de temps pour vos activités de recherche):

Veuillez choisir la réponse qui convient :

- a) HE-Arc Conservation-restauration
- b) HE-Arc Ingénierie
- c) HE-Arc Santé
- d) HE-Arc Gestion (HEG Arc)
- e) Haute école d'ingénierie et d'architecture Fribourg - HEIA-FR Hochschule für Technik und Architektur Freiburg - HTA-FR
- f) Haute école de gestion Fribourg - HEG-FR Hochschule für Wirtschaft Freiburg - HSW-FR
- g) Haute école de santé Fribourg - Hochschule für Gesundheit Freiburg - HEdS-FR
- h) Haute école de travail social Fribourg - HETS-FR Hochschule für Soziale Arbeit Freiburg - HSA-FR
- i) Haute école d'art et de design – Genève (HEAD – Genève)
- j) Haute école de gestion de Genève (HEG-Genève)
- k) HEPIA - Haute école du paysage, d'ingénierie et d'architecture de Genève
- l) Haute école de musique de Genève - HEM
- m) Haute école de santé de Genève (HEdS-Genève)
- n) Haute école de travail social de Genève (HETS-Genève)
- o) HES-SO Valais-Wallis - Haute Ecole de Gestion - HEG
- p) HES-SO Valais-Wallis - Haute Ecole et Ecole Supérieure de Travail Social - HESTS
- q) HES-SO Valais-Wallis - Haute Ecole de Santé - HEdS
- r) HES-SO Valais-Wallis - Haute Ecole d'Ingénierie - HEI
- s) HES-SO Valais-Wallis - Ecole de design et haute école d'art – EDHEA
- t) ECAL/Ecole cantonale d'art de Lausanne
- u) Haute Ecole d'Ingénierie et de Gestion du Canton de Vaud - HEIG-VD
- v) Institut et Haute Ecole de la Santé La Source
- w) HEMU – Haute Ecole de Musique
- x) Haute école de travail social et de la santé Lausanne - HETSL
- y) HESAV - Haute Ecole de Santé Vaud
- z) EHL- Hospitality Business School





- aa) CHANGINS – Haute école de viticulture et œnologie
- bb) La Manufacture - Haute école des arts de la scène
- cc) HES-SO Master
- dd) Autre:

2. *Quelle(s) est/sont votre/vos fonction(s) ?

Veillez choisir les réponses qui conviennent :

- a) Adjoint.e scientifique
- b) Assistant.e
- c) Collaborateur·trice scientifique ou artistique
- d) Doctorant.e
- e) Chargé.e de cours
- f) Maître d'enseignement
- g) Professeur.e
- h) Responsable Ra&D
- i) Autre:

3. *Quel est votre principal domaine de recherche ?

Veillez choisir la réponse qui convient :

- a) Design et Arts visuels
- b) Economie et Services
- c) Ingénierie et architecture
- d) Musique et Arts de la scène
- e) Santé
- f) Travail social

4. *Quelle est la discipline dans laquelle votre projet de recherche actuel s'inscrit (si plusieurs disciplines, et/ou plusieurs projets en cours veuillez préciser) ?

Commentaire :

RECHERCHE ACTUELLE

5. *Quel(s) est (sont) le(s) bailleurs de fonds qui finance(nt) votre recherche actuelle?

Veillez choisir toutes les réponses qui conviennent

- a) HES.SO
- b) INNOSuisse
- c) Recherche Fonds national suisse (FNS)
- d) Confédération
- e) Programmes UE et internationaux
- f) Partenaires publics et para-publics
- g) Partenaires privés
- h) Ne sais pas
- i) Autre:





6. * Recevez-vous des financements particuliers pour gérer vos données de recherche dans le cadre du projet actuel (ex. les sauvegarder, les publier, les traiter, les conserver, les partager...)?

Veillez choisir la réponse qui convient :

- a) Oui par mes bailleurs de fonds
- b) Non
- c) Ne sais pas
- d) Autre:

DONNÉES DE RECHERCHE

7. * Quel(s) type(s) de données collectez-vous et produisez-vous actuellement ?

Veillez choisir toutes les réponses qui conviennent :

- a) 3D, modèles, visualisations
- b) Applications, logiciels, programmes informatiques
- c) Archives, sources historiques
- d) Audio
- e) Codes sources, langages de programmation
- f) Données matérielles et physiques
- g) Données textuelles (ex. documents, textes, notes manuscrites)
- h) Données numériques (ex. tableurs)
- i) Données géospatiales
- j) Données statistiques
- k) Données qualitatives
- l) Données quantitatives
- m) Bases de données
- n) Entretiens
- o) Images, photographies
- p) Numérisations (scans)
- q) Vidéo
- r) Ne sais pas
- s) Autre:

8. * D'une manière générale, qui est responsable de la gestion des données collectées et produites durant votre recherche ?

Veillez choisir toutes les réponses qui conviennent :

- a) Moi-même
- b) Assistant.e
- c) Personne désignée au sein de l'équipe (data manager)
- d) Responsable de l'équipe/du projet
- e) Directeur.trice de thèse
- f) Centre informatique/informaticiens
- g) Bibliothèques
- h) Partenaire/prestataire externe
- i) Personne
- j) Ne sais pas
- k) Autre:





9. * Laquelle de ces affirmations reflète au mieux l'importance que vous accordez à vos données de recherche actuelles ?
Veuillez choisir la réponse qui convient :
- a) Elles sont importantes pour la société en général
 - b) Elles sont importantes pour l'ensemble de la communauté scientifique
 - c) Elles sont surtout importantes dans mon domaine spécifique de recherche
 - d) Elles sont surtout importantes pour ma carrière professionnelle
 - e) Elles sont surtout importantes pour mon projet de recherche
 - f) Elles sont importantes pour tous ces aspects cités
 - g) Ne sais pas/pas d'avis

PLAN DE GESTION DES DONNÉES (DATA MANAGEMENT PLAN)

10. * Un « Plan de gestion des données » (Data Management Plan - DMP) est un document qui décrit quelles actions seront réalisées, et avec quels moyens, pour gérer et sécuriser les données de recherche. Au cours de vos différentes recherches, avez-vous déjà réalisé un « Plan de gestion des données » ?
Veuillez choisir la réponse qui convient :
- a) Oui
 - b) Non
 - c) Ne sais pas

11. Était-ce alors une exigence de votre bailleur de fonds ?

Répondre à cette question seulement si les conditions suivantes sont réunies
- La réponse à la question '10 était égale à a)

- Veuillez choisir la réponse qui convient :
- a) Oui
 - b) Non
 - c) Ne sais pas

12. Ce DMP vous a-t-il été utile et pour quelles raisons ? Si oui/non, veuillez préciser les raisons.

Répondre à cette question seulement si les conditions suivantes sont réunies
- La réponse à la question '10 était égale à a)

- Veuillez choisir la réponse qui convient :
- a) Oui tout à fait
 - b) Plutôt oui
 - c) Plutôt non
 - d) Non pas du tout
 - e) Ne sais pas/pas d'avis
- Commentaire :

13. Selon vous et de manière générale, planifier la gestion de ses données de recherche est...

- Veuillez choisir la réponse qui convient :
- a) Très important
 - b) Plutôt important
 - c) Plutôt pas important
 - d) Pas du tout important
 - e) Ne sais pas/pas d'avis





DÉCRIRE ET DOCUMENTER VOTRE PROJET ET VOS DONNÉES DE RECHERCHE

14. Avez-vous créé de la documentation (ex. codebook, readme file, dictionnaire de variables, rapport) pour vos données de recherche ?

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non
- c) Ne sais pas

15. Quelle(s) principale(s) information(s) cette documentation contient-elle ?

Répondre à cette question seulement si les conditions suivantes sont réunies

- La réponse à la question '14 était égale à a)

Veillez choisir toutes les réponses qui conviennent :

- a) Contexte de la recherche
- b) Méthodologie et processus de recherche
- c) Récolte des données (ex. leur type, nature, contenu, structure,...)
- d) Instruments et manipulations
- e) Variables (ex. (noms, questions, descriptions, algorithmes, syntaxes,...)
- f) Ne sais pas
- g) Autre:

UTILISER LES MÉTADONNÉES

Les « métadonnées », généralement structurées et standardisées, sont des « données sur les données » qui permettent de définir et décrire certaines caractéristiques des données, comme :

- Leurs aspects administratifs (ex. droits d'accès et d'utilisation)
- Leurs aspects techniques (ex. date)
- Leur contenant (ex. format de fichier)
- Leur contenu (ex. titre, sujet)
- Leur contexte de création (ex. auteur, institution)
- Leur structure logique et physique (ex. liens entre les chapitres/fichiers)

16. Utilisez-vous des métadonnées pour documenter et décrire vos données de recherche actuellement ?

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non
- c) Ne sais pas

17. Si vous utilisez des métadonnées, sont-elles normalisées, structurées et standardisées (ex. Dublin Core, DDI, METS, ISO,...) ?

Répondre à cette question seulement si les conditions suivantes sont réunies

La réponse à la question '16 était égale à a)

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non
- c) Ne sais pas

18. Quel(s) standard(s) utilisez-vous pour ces métadonnées ?





Répondre à cette question seulement si les conditions suivantes sont réunies
- La réponse à la question '17 était égale à a)

Veillez choisir toutes les réponses qui conviennent :

- a) Dublin Core
- b) DDI
- c) METS
- d) ISO
- e) Ne sais pas
- f) Autre:

STOCKER VOS DONNÉES DE RECHERCHE

19. Où stockez-vous vos données de recherche pour le projet actuel ?

Veillez choisir toutes les réponses qui conviennent :

- a) Disque dur – ordinateur/laptop/netbook de la HES-SO
- b) Disque dur – ordinateur/laptop/netbook personnel
- c) Disque dur externe
- d) Sur le serveur de l'école
- e) Switchdrive
- f) Sharepoint
- g) Teams
- h) Dropbox
- i) OneDrive
- j) Ne sais pas
- k) Autre:

20. À combien estimez-vous le volume de données générées de votre projet de recherche actuel? Veuillez choisir la réponse qui convient :

- a) < 1 Go
- b) 1-50 Go
- c) 50-100 Go
- d) 100-500 Go
- e) 500 Go-1 To
- f) 1-50 To
- g) 50-100 To
- h) > 100 To
- i) Ne sais pas

SAUVEGARDER VOS DONNÉES DE RECHERCHE

21. Existe-t-il des sauvegardes (back-ups, copies de secours) de vos données de recherche actuelles?

Veillez choisir la réponse qui convient :

- a) Oui pour toutes mes données
- b) Oui pour une partie seulement
- c) Non, mais j'ai l'intention de le faire
- d) Non et je n'ai pas l'intention de le faire
- e) Ne sais pas





22. À quelle fréquence vos données sont-elles actuellement sauvegardées ?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '21 était égale à a) et à b)

Veillez choisir la réponse qui convient :

- a) Tous les jours
- b) 1x/semaine
- c) Plusieurs fois par semaine
- d) 1x/mois
- e) Plusieurs fois par mois
- f) Tous les 6 mois
- g) 1x/an
- h) Fréquence irrégulière
- i) Ne sais pas
- j) Autre:

23. Où vos données actuelles sont-elles sauvegardées?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '21 était égale à a) et à b)

Veillez choisir toutes les réponses qui conviennent :

- a) Disque dur externe
- b) Ordinateur personnel ou HES-SO
- c) Serveur de votre école
- d) USB/Flash
- e) Serveur externe
- f) Services Cloud (ex. Dropbox, OpenDrive, CrashPlan,...)
- g) Ne sais pas
- h) Autre:

SÉCURISER VOS DONNÉES DE RECHERCHE

24. Vos données de recherche sont-elles actuellement sécurisées (ex. contrôle des accès, cryptage,...) ?

Veillez choisir la réponse qui convient :

- a) Oui pour toutes mes données
- b) Oui pour une partie seulement
- c) Non, mais j'ai l'intention de le faire
- d) Non et je n'ai pas l'intention de le faire
- e) Ne sais pas

25. Comment vos données sont-elles actuellement sécurisées ?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '24 était égale à a) et à b)

Veillez choisir toutes les réponses qui conviennent :

- a) Contrôle des accès aux labos/locaux/salles
- b) Utilisation d'un coffre-fort
- c) Pare-feu, logiciels de sécurité antivirus et contre le piratage informatique
- d) Serveurs sécurisés
- e) Cryptage des données
- f) Protocoles de transfert sécurisés
- g) Login/mots de passe pour l'accès
- h) Destruction des données
- i) Ne sais pas





j) Autre:

PARTAGER ET DONNER ACCÈS À VOS DONNÉES DE RECHERCHE

Les principes FAIR « permettent d'assurer qu'un jeu de données puisse être trouvé (« Findable »), qu'il soit accessible (« Accessible »), interopérable (« Interoperable ») et réutilisable (« Reusable »)». ([Lien](#) pour en savoir plus)

26. *Etes-vous familier.ère avec les principes FAIR ?

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non

27. *De manière générale, êtes-vous favorable au partage de vos données de recherche ?

Veillez choisir la réponse qui convient :

- a) Oui tout à fait
- b) Plutôt oui
- c) J'ai besoin d'étudier plus précisément cette question
- d) Plutôt non
- e) Non pas du tout
- f) Ne sais pas/pas d'avis

28. *Quelle(s) raison(s) motivent votre choix de partager vos données ?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '27 était égale à a) et b)

Veillez choisir toutes les réponses qui conviennent

- a) Exigence de mon bailleur de fonds/éditeur/institution
- b) Pratique courante au sein de la communauté scientifique
- c) Réseau, relations et collaborations professionnelles
- d) Augmente la visibilité et l'impact de mes recherches
- e) Prouve que la recherche est intègre, responsable et de qualité
- f) Facilite la vérification, l'évaluation (peer-review) et la reproductibilité de la recherche
- g) Renforce la validité des méthodes et des résultats
- h) Augmente le potentiel d'innovation
- i) Responsabilité du chercheur envers la société
- j) Ne sais pas
- k) Autre:

29. *Quelle(s) raison(s) motivent votre choix de ne pas partager vos données en général ?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '27 était égale à d) ou e)

Veillez choisir toutes les réponses qui conviennent :

- a) Contrainte du bailleur de fonds/éditeur/institution
- b) Mes données ne sont pas suffisamment documentées
- c) Mes données ne sont pas suffisamment pertinentes
- d) Mes données sont difficilement compréhensibles ou interprétables
- e) Pas de temps/argent/ressources pour le faire
- f) Je ne sais pas comment les partager
- g) Je ne sais pas où les partager
- h) Pour protéger l'exclusivité de ma découverte et de mes idées
- i) Risque de fausse interprétation ou d'utilisation fallacieuse de mes données





- j) Raisons commerciales et stratégiques (ex. brevets)
- k) Raisons éthiques, de déontologie, de confidentialité
- l) Raisons juridiques et contractuelles
- m) Raisons de sécurité nationale
- n) Raisons personnelles
- o) Ne sais pas/pas d'avis
- p) Autre:

ARCHIVER ET PRÉSERVER VOS DONNÉES DE RECHERCHE

30. *En général, conservez-vous vos données de recherche à la fin d'un projet ?
Veuillez choisir la réponse qui convient :

- a) Oui toutes
- b) Oui env. 75%
- c) Oui env. 50%
- d) Oui env. 25%
- e) Non, mais j'ai l'intention de le faire
- f) Non et je n'ai pas l'intention de le faire
- g) Ne sais pas

31. *Avez-vous déjà déposé vos données sur un dépôt/répertoire de données ?
Veuillez choisir la réponse qui convient :

- a) Oui
- b) Non
- c) Ne sais pas /pas d'avis

32. Si vous avez déjà déposé vos données sur un dépôt/répertoire de données, veuillez préciser le(s)quel(s) :

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '31 était égale à a)

Commentaire :

33. *Avez-vous l'intention de déposer vos données de recherche dans un dépôt/répertoire de données ou service d'archives pour le projet en cours ?
Veuillez choisir la réponse qui convient :

- a) Oui pour toutes mes données
- b) Oui pour une partie de mes données
- c) Non, je ne connaissais pas cette possibilité
- d) Non et je n'ai pas l'intention de le faire
- e) Ne sais pas / pas d'avis

34. *Si vous avez déposé ou avez l'intention de déposer vos données de recherche pour le projet en cours, à quel(s) endroit(s) souhaiteriez-vous les déposer ?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '33 était égale à a) et b)

Veuillez choisir toutes les réponses qui conviennent :

- a) Dépôt généraliste (ex : Olos, SWISSUbase, Zenodo, etc.)
- b) Dépôt disciplinaire (ex : Dasch, Material Cloud, etc.)





- c) Dépôt défini par le bailleur de fonds ou partenaire institutionnel
- d) Dépôt commercial, privé
- e) Ne sais pas/pas d'avis
- f) Autre:

35. *Si vous avez déposé ou avez l'intention de déposer vos données de recherche pour le projet en cours dans un répertoire/dépôt institutionnel ou service d'archives, quelle(s) raison(s) ont motivé ou motivent votre choix ?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '33 était égale à a) et b)

Veuillez choisir toutes les réponses qui conviennent :

- a) Recommandé par mes collègues/spécialistes du domaine
- b) Renommée du dépôt
- c) Orientation disciplinaire
- d) Nombre de données déposées sur le dépôt/répertoire
- e) Conditions générales d'utilisation
- f) Gestion des accès et des droits
- g) Gratuité du service
- h) Capacité du volume de stockage
- i) Sécurité des données
- j) Préservation et archivage à long terme
- k) Impact, citation, visibilité des données déposées
- l) Réseau, relations et collaborations professionnelles
- m) Avis/choix personnel
- n) Interface ergonomique
- o) Ne sais pas/pas d'avis
- p) Autre:

RÉUTILISER LES DONNÉES DE RECHERCHE

36. *De manière générale, acceptez-vous que d'autres chercheurs.euses réutilisent vos données dans le cadre de leur propre recherche ?

Veuillez choisir la réponse qui convient :

- a) Oui tout à fait
- b) Plutôt oui
- c) Plutôt non
- d) Non pas du tout
- e) Ne sais pas/pas d'avis

37. *Avez-vous déjà réutilisé les données de recherche d'autres chercheurs.euses dans le cadre de votre propre recherche en général ?

Veuillez choisir la réponse qui convient :

- a) Oui
- b) Non, je ne connaissais pas cette possibilité
- c) Non, mais j'ai l'intention de le faire
- d) Non et je n'ai pas l'intention de le faire
- e) Ne sais pas/pas d'avis





38. *Si vous avez réutilisé les données de recherche d'autres chercheurs.euses, pour quelle(s) raison(s) les avez-vous réutilisées ?

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '37 était égale à a)

Veillez choisir toutes les réponses qui conviennent :

- a) Pour les comparer à mes propres données
- b) Pour développer une nouvelle recherche
- c) Pour évaluer la recherche d'autrui (peer review)
- d) Pour éviter de refaire à double l'expérience
- e) Pour gagner du temps
- f) Pour économiser de l'argent
- g) Ne sais pas/pas d'avis
- h) Autre:

QUESTIONS JURIDIQUES, ÉTHIQUES ET DÉONTOLOGIQUES

39. *À qui appartiennent les données que vous créez lors de votre activité de recherche actuelle ?

Veillez choisir toutes les réponses qui conviennent :

- a) Moi-même
- b) Tous les membres de l'équipe de recherche
- c) Responsable de l'équipe/du projet
- d) La HES-SO
- e) Mon école
- f) Bailleurs de fonds
- g) Ne sais pas
- h) Autre:

40. *Des contrats (ex. de secret, de confidentialité ou de non-divulgence, de transfert de matériel biologique, formulaire de consentement,...) régissent-ils votre projet de recherche actuel ?

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non
- c) Ne sais pas

41. *Pour ce projet de recherche, collectez-vous et gérez-vous des données à caractère confidentiel (ex. données personnelles ou sensibles) ?

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non
- c) Ne sais pas

42. *Utilisez-vous des outils pour gérer et sécuriser vos données de recherche (outil de gestion, outil d'anonymisation, ...) ?

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non

43. *Si vous utilisez des outils pour gérer et sécuriser vos données de recherche, le(s)quel(s) utilisez-vous ?

Répondre à cette question seulement si les conditions suivantes sont réunies





La réponse à la question '42 était égale à a)

Commentaire :

VALORISATION

44. *Connaissez-vous ORCID ?
Veuillez choisir la réponse qui convient :
- a) Oui
 - b) Non
45. *Avez-vous un identifiant ORCID ?
Veuillez choisir la réponse qui convient :
- a) Oui
 - b) Non
 - c) Ne sais pas
46. *Connaissez-vous le Data paper ?
Veuillez choisir la réponse qui convient :
- a) Oui
 - b) Non
47. *Avez-vous déjà publié un Data paper ?
Veuillez choisir la réponse qui convient :
- a) Oui
 - b) Non

AIDE, SOUTIEN ET FORMATION À LA GESTION DES DONNÉES DE RECHERCHE

48. *Avez-vous déjà suivi des formations sur la gestion des données de recherche ?
Veuillez choisir la réponse qui convient :
- a) Oui
 - b) Non
49. *Parmi les différents types de formation ci-dessous, lesquels préférez-vous ?
Veuillez choisir toutes les réponses qui conviennent :
- a) Cours, séminaires
 - b) Ateliers pratiques, workshops, formations en groupe
 - c) Site web dédié à la gestion des données de recherche
 - d) Cours en ligne (e-learning)
 - e) Documentation en ligne (fiches techniques, guides, modes d'emploi, check-lists, tutoriels...)
 - f) Documentation imprimée (fiches techniques, guides, modes d'emploi, check-lists, tutoriels...)
 - g) Aide et conseils personnalisés fournis par un-e spécialiste (ex. archiviste, informaticien, juriste,...)
 - h) Pas de préférence
 - i) Autre:
50. *Si vous deviez participer à une formation (ex. cours, séminaires, workshops,...), quel(s) sujet(s) vous intéresseraient le plus ?
Veuillez choisir toutes les réponses qui conviennent :





- a) Le b.a.-ba de la gestion des données de recherche
- b) Data Management Plan
- c) Normes et standards (formats de fichiers, métadonnées)
- d) Stockage, sauvegardes (back-up), sécurité
- e) Infrastructure technique/informatique
- f) Anonymisation, confidentialité, protection des données
- g) Citations, publications, partage
- h) Archivage, préservation à long terme
- i) Ethique et déontologie
- j) Questions juridiques, légales et contractuelles
- k) Droit d'auteur et propriété intellectuelle
- l) Autre:

51. *Quel est votre premier point de contact pour vous aider dans vos activités liées à la gestion des données ?

Veillez choisir la réponse qui convient :

- a) Les bibliothécaires
- b) Mes collègues chercheurs.euses
- c) Services de soutien à la recherche
- d) Il n'y en a pas
- e) Autre:

INFORMATION SUR LA GESTION DES DONNÉES DE RECHERCHE

52. *De manière générale, êtes-vous informé.e (via avis officiels, communiqués internes, séances,...) sur la question des données de recherche ?

Veillez choisir la réponse qui convient :

- a) Oui régulièrement
- b) De temps en temps
- c) Rarement
- d) Jamais
- e) Ne sais pas

53. *Souhaiteriez-vous recevoir plus d'informations sur cette question ?

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non
- c) Cela m'est égal

54. *Existe-t-il au sein de votre unité de recherche des directives, lignes directrices, recommandations internes pour la gestion des données de recherche ? Si oui, veuillez préciser.

Veillez choisir la réponse qui convient :

- a) Oui
- b) Non

Commentaire :





COMMENTAIRE, SUGGESTION

55. Ici vous pouvez nous faire part de vos attentes et besoins spécifiques en matière de gestion des données de recherche, de vos commentaires, suggestions et propositions d'amélioration :

A préciser

TIRAGE AU SORT

56. *Je souhaite participer au tirage au sort (5 x bons de 100.- à la Librairie Payot à gagner).
Veuillez choisir la réponse qui convient :

- a) Oui
- b) Non

57. *Mes coordonnées mails pour le tirage au sort :

Répondre à cette question seulement si les conditions suivantes sont réunies
La réponse à la question '56 était égale à a)

Commentaire : à préciser

Merci beaucoup d'avoir pris le temps de compléter ce questionnaire.

