



CELLULE DATA STEWARDSHIP DE LA HES-SO ENTRETIENS AUPRÈS DES CHERCHEUR.EUSES/DOMAINE

Constance Delamadeleine – Octobre 2024

TABLE DES MATIERES

I. CONTEXTE	1
II. OBJECTIFS DES ENTRETIENS.....	2
III. CADRE MÉTHODOLOGIQUE ET DIFFUSION.....	3
ANNEXE QUESTIONNAIRE.....	5

I. CONTEXTE

L'Open Research Data (ORD) vise à rendre les données de recherches financées par des fonds publics accessibles de façon permanente et gratuite. Le but étant leur réutilisation par d'autres afin d'améliorer la qualité et l'intégrité du travail scientifique. C'est dans ce cadre que des actrices et acteurs du monde académique et politique, mais aussi des éditeurs, ont développé les principes FAIR pour rendre ces données Findable, Accessible, Interoperable and Reusable. Désormais, les grands bailleurs de fonds de la recherche et la HES-SO demandent un engagement dans ce sens à leurs chercheuses et chercheurs.

Cette nouvelle culture et ces nouvelles pratiques impliquent une gestion efficace des données de la recherche. 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 ».¹ En d'autres termes, l'ORD ou devrait-on dire le « FAIR Data », est loin de se limiter à un « drag and drop » d'un ensemble de fichiers sur Zenodo. Ce nouveau paradigme requiert un accompagnement continu des chercheur.euses. C'est précisément ce que vise la stratégie [Open Science de la HES-SO](#), lancée en 2018 : apporter un soutien à la communauté scientifique pour faciliter l'adoption de cette nouvelle culture. Dans cette perspective, et dans le cadre du programme Open Science I Phase B – Open

¹ OCDE (2021). [Recommandation du Conseil concernant l'accès aux données de la recherche financée sur fonds publics](#) (consulté le 5 mars 2024).



Research Data initié par swissuniversities, le Dicastère Recherche et Innovation a obtenu un financement afin de mener un projet pilote de Data Stewardship au sein de la HES-SO pour la période 2023-2024. Lancée en juin 2023, la [Cellule Data Stewardship](#) se compose de 5 data stewards pour la plupart employés à 30% (un·e data steward par domaine à l'exception d'un data steward pour les domaines DAV/MAS) et d'une coordinatrice.

Cette cellule poursuit les objectifs suivants :

- Renforcer le soutien aux chercheur·euses en matière de gestion et d'ouverture des données de recherche en identifiant les besoins spécifiques à leur domaine de recherche
- Coordonner les initiatives menées au sein des hautes écoles
- Contribuer à la mise en œuvre de la stratégie ORD de la HES-SO ainsi que son plan d'action.

Afin d'atteindre ces objectifs, les data stewards ont mené différentes missions dont la conduite d'entretiens avec des chercheur·euses des différents domaines.

II. OBJECTIFS DES ENTRETIENS

La récente enquête quantitative menée en 2023 « [Gestion et ouverture des données de recherche à la HES-SO. Enquête sur les pratiques et besoins des chercheuses et chercheurs](#) », a permis de dresser un état des lieux partiel des pratiques en matière de gestion des données de recherche, des connaissances et attitudes relatives à l'ouverture des données de recherche. 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 nécessitent d'être enrichis par des entretiens qualitatifs afin d'obtenir une analyse plus fine des besoins et spécificités disciplinaires que ne couvre pas l'enquête quantitative. Cette approche, orientée vers les spécificités disciplinaires, s'appuie sur l'expérience de l'équipe de data stewards de l'Université technique de Delft. Active depuis 2017, cette dernière a accumulé une expertise précieuse en Data Stewardship, mettant en évidence une réalité incontournable : les préoccupations, les méthodologies et les flux de travail diffèrent considérablement selon les domaines et disciplines, rendant difficile l'adoption d'une solution unique.² Forte de ces enseignements, l'équipe souligne l'importance d'adapter les conseils en gestion des données aux besoins spécifiques de chaque discipline pour garantir un soutien efficace et pertinent. Cette nécessité est également confirmée par un bilan ORD couvrant les années 2019-2022, qui a révélé un obstacle majeur pour les chercheur·euses de la HES-SO : un manque notable d'aide et d'information spécifique à leur domaine.³

Outre l'identification des besoins et des particularités disciplinaires des communautés scientifiques, notamment des domaines moins représentés dans l'enquête quantitative, ces entretiens ont présenté différentes opportunités : sensibiliser les chercheur·euses aux enjeux et avantages de l'ORD ainsi que l'importance de la gestion des données (pour celles et ceux qui n'ont pas encore

² Université technique de Delft (2017). Addressing disciplinary data management needs. <https://openworking.wordpress.com/2017/08/29/data-stewardship-addressing-disciplinary-data-management-needs/> (consulté le 5 mars 2024).

³ Lucas, I. (2023). Plan de mesures. Cellule Data Stewardship, p.4.

été accompagné·es) et d'engager une réflexion sur la définition des données, sur la planification, la collecte, l'analyse, la curation, le dépôt, la publication et la réutilisation de données.

Comme indiqué plus haut, les résultats de ces entretiens combinés à ceux de l'enquête quantitative serviront de support pour alimenter la stratégie ORD de la HES-SO et proposer un plan d'action au plus proche des besoins des domaines.

III. CADRE MÉTHODOLOGIQUE ET DIFFUSION

Entre septembre 2023 et juin 2024, les data stewards ont mené au total 69 entretiens avec des chercheur·euses des différents domaines (en moyenne une dizaine par domaine). D'un point de vue méthodologique, il semblait judicieux d'effectuer des entretiens de type semi-structuré afin de permettre aux participant·es de se sentir libres d'exprimer leurs attentes et leurs difficultés concernant la gestion et l'ouverture des données de recherche. De plus, l'enquête quantitative a accusé un certain taux d'abandon en cours de route du questionnaire (ca. 30%) face à des chercheuses et chercheurs probablement peu ou pas familiarisé·es à ces questions, ce qui a souligné la nécessité de recourir à des entretiens qualitatifs. Enfin, certaines hautes écoles ne sont pas représentées dans cette enquête, ces entretiens ont permis de combler ce manque.

Une grille de questionnaire générique a été élaborée par les membres de la cellule sur la base du questionnaire de l'enquête quantitative susmentionnée, elle-même inspirée de celle établie par Carmen Jambé pour le Service des Ressources informationnelles et archives (UNIRIS) de l'Université de Lausanne (voir Annexe).⁴ Cette grille a été adaptée pour certains domaines. Composée d'une vingtaine de questions, elle se divise en trois sections (Description du projet de recherche ; Données de recherche/Gestion des données ; Open Research Data) et balaie les thèmes suivants : la connaissance du contexte général des données et plus largement de l'ORD, la typologie des données produites par les répondant·es, leurs pratiques en termes de stockage/archivage, partage et réutilisation ainsi que les besoins. Afin d'aider les participant·es à réfléchir plus facilement aux questions posées, le questionnaire les invitait à s'appuyer sur un projet en cours ou récemment terminé. L'échantillon des participant·es visait à inclure des personnes ayant peu ou pas d'expérience de partage des données, mais ayant potentiellement des habitudes en matière de gestion des données, ce qui a conduit à concevoir une seconde trame de questionnaire adapté à ce profil.

Outre la prise en compte des diverses expériences en matière de gestion et d'ouverture des données, la constitution de chaque échantillon a suivi d'autres logiques afin de refléter, dans la mesure du possible, différents types de recherche, ainsi que des disciplines variées au sein d'un même domaine. L'objectif était également d'inclure dans la mesure du possible, toutes les hautes écoles et de couvrir l'ensemble des profils du personnel d'enseignement et de recherche, y compris les doctorant·es. Les porteurs de projets ont été directement contactés par les data stewards.

⁴ Jambé, C. (2015). La gestion des données de recherche à l'Université de Lausanne : enjeux transdisciplinaires (Haute école de gestion de Genève).



Toutefois, il convient de noter que le nombre de chercheur.euses interrogé.es ne reflète pas entièrement la proportion des effectifs respectifs des domaines.

La plupart des entretiens, d'une durée moyenne de 1h30, ont été enregistrés. Il n'a pas été question de réaliser des transcriptions intégrales d'entretiens. Néanmoins, sur la base des enregistrements et prise de notes, chaque data steward était chargé·e de transcrire les extraits significatifs et de les catégoriser en fonction des axes suivants : les pratiques en matière de gestion et d'ouverture ; les besoins et les défis/problématiques rencontrés. Certains rapports incluent des verbatims des entretiens.

Afin de préserver l'anonymat des personnes interrogées, leurs noms ne sont pas mentionnés dans les rapports ni les fonctions afin d'éviter toute réidentification. Nous souhaitons les remercier pour le temps consacré à répondre aux entretiens.

Les rapports synthétisent les principaux besoins et défis liés à la gestion et à l'ouverture des données de recherche, tels qu'identifiés par les data stewards lors de leurs entretiens et proposent des recommandations. Ils sont principalement destinés à être diffusés en interne, notamment aux responsables de recherche, aux membres de la Communauté Open Science et à toutes personnes contactées lors de ces entretiens (chercheur.euses, personnel de soutien ORD). Toutefois, dans une logique d'ouverture et de favoriser le partage des connaissances, les rapports sont disponibles sur les pages ORD du site internet de la HES-SO ([Cellule Data Stewardship](#)).



ANNEXE QUESTIONNAIRE

Pour définir le groupe de questions A ou B : avez-vous déjà déposé vos données sur un data repository (dépôt de données de recherche) ?

Si non, veuillez vous reporter au groupe A (aucune ou peu d'expérience ORD)

Si oui, veuillez vous reporter au groupe B (expérience ORD)

QUESTIONS GROUPE A

I. Description du projet

1. Pouvez-vous décrire brièvement votre projet de recherche et son objectif principal ?
2. Dans quel type/approche de recherche, ce projet s'inscrit-il ?

II. Données de recherche et gestion des données

3. Dans le cadre de votre projet actuel ou récemment terminé, quel(s) type(s) de données avez-vous collecté et/ou produit ?
4. Pourriez-vous préciser les formats de fichiers que vous utilisez ?
5. A qui appartiennent ces données selon vous ?
6. Ces données sont-elles à caractère confidentiel (personnelle, sensible, etc.) ? Si oui, avez-vous eu recours à un soutien particulier pour gérer/traiter ces données ?
7. Connaissez-vous l'identité de la personne en charge des questions juridiques de votre haute école ?
8. Avez-vous reçu des recommandations concernant la sécurité des données de recherche au début de votre projet (sauvegarde, stockage, etc.) ?
9. Où stockez-vous vos données de recherche pendant le projet ?

III. Open Research Data

10. Avez-vous déjà entendu parler du terme « Open Research Data » (ORD) ?
11. Quelle est la différence selon vous entre ORD et gestion des données de recherche ?
12. Recevez-vous des informations ou un soutien particulier en matière d'ORD dans votre haute école ?
13. Connaissez-vous le DMP ? Si oui, avez-vous rédigé un DMP dans le cadre de votre projet ?
14. Connaissez-vous des dépôts de données ? Si oui, veuillez préciser le(s)quel(s).
15. Avez-vous déjà consulté des sets de données sur des dépôts ?
16. Connaissez-vous les principes FAIR ?
17. Si vous deviez partager vos données sur un dépôt FAIR, quels seraient, selon vous, les principaux obstacles ?
18. De quels types de services ou outils auriez-vous besoin concernant la gestion et l'ouverture de vos données de recherche ?

QUESTIONS GROUPE B

I. Description du projet

1. Pouvez-vous décrire brièvement votre projet de recherche et son objectif principal ?
2. Dans quel type/approche de recherche, ce projet s'inscrit-il ?



II. Données de recherche et gestion des données

3. Dans le cadre de votre projet actuel ou récemment terminé, quel(s) type(s) de données avez-vous collecté et/ou produit ?
4. Pourriez-vous préciser les formats de fichiers que vous utilisez ?
5. A qui appartiennent ces données selon vous ?
6. Ces données sont-elles à caractère confidentiel (personnelle, sensible, etc.) ? Si oui, avez-vous eu recours à un soutien particulier pour gérer/traiter ces données ?
7. Connaissez-vous l'identité de la personne en charge des questions juridiques de votre haute école ?
8. Avez-vous reçu des recommandations concernant la sécurité des données de recherche au début de votre projet (sauvegarde, stockage, etc.) ?
9. Où stockez-vous vos données de recherche pendant le projet ?

III. Open Research Data

10. Avez-vous rédigé un DMP dans le cadre de votre projet ? Si oui, veuillez préciser le contexte.
11. Avez-vous consulté et/ou eu réutilisé des sets de données ? Si oui, quel dépôt de données avez-vous consulté ?
12. De manière générale, avez-vous bénéficié d'un soutien pour gérer vos données de recherche ?
13. Où avez-vous déposé vos données ?
14. De quels types de services ou outils auriez-vous besoin concernant la gestion et l'ouverture de vos données de recherche ?

