



---

## CELLULE DATA STEWARDSHIP RAPPORT ENTRETIENS CHERCHEUR·EUSES DOMAINE INGÉNIERIE ET ARCHITECTURE

Xavier Brochet, Data steward – Octobre 2024

---

### TABLE DES MATIÈRES

I. INTRODUCTION.....	1
II. PRATIQUES EN MATIÈRE DE GESTION DES DONNÉES.....	2
III. PRINCIPAUX DÉFIS.....	7
IV. CONCLUSION ET RECOMMANDATIONS.....	9

### I. INTRODUCTION

---

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. » (OCDE, 2021). 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. Malgré les mesures de soutien instaurées dans certaines hautes écoles, un bilan ORD dressé pour les années 2019-2022 a pointé un obstacle majeur rencontré par les chercheur·euses, à savoir un important manque d'aide et d'information lié à leur domaine (Lucas, 2023). Afin de combler ce manque, dans le cadre du programme Open Science I Phase B – Open Research Data, lancé par swissuniversities,





Le Dicastère Recherche et Innovation a obtenu des fonds pour développer un projet pilote de Data Stewardship au sein de la HES-SO (période 2023-2024). Lancée en juin 2023, le projet rassemble 5 data stewards (1 par domaine à l'exception des domaines artistiques). L'une de leurs principales missions consiste à conduire des entretiens avec des chercheur·euses. Dans ce cadre, au niveau du domaine Ingénierie et Architecture, 22 entretiens semi-structurés ont été menés par Xavier Brochet, data steward du domaine IA, entre janvier et mai 2024 (voir liste des entretiens, tableau 1). Ces entretiens ont pour objectifs de recueillir des retours d'expérience, d'identifier les pratiques des chercheur·euses en matière de stockage, de partage et de diffusion des données, et de pointer les défis et spécificités disciplinaires.

Ce rapport synthétise les résultats de cette enquête. Anonymisées pour préserver la confidentialité des participant·es, les informations recueillies serviront à nourrir les réflexions internes en vue de l'élaboration de la stratégie et du plan d'action ORD de la HES-SO. Ces informations pourraient également servir de support pour alimenter d'autres processus de planification stratégique au sein des hautes écoles et/ou du domaine.

	Nom de la haute école	Discipline couverte (par ordre alphabétique)
Entretien 1	HES-SO Valais-Wallis - HEI	Architecture logiciel et application industrielle Biologie moléculaire et végétale Chimie industrielle, Automatisation, prototypage et conduite de procédé Conception mécanique et hydroélectricité Conception mécanique et simulation numérique Domaine d'application des sols et de l'environnement Énergie, Valorisation énergétique de la biomasse Foudre et champs électromagnétiques ; Génie électrique, électronique et communication Gestion de la nature, biodiversité spécialisée écosystème aquatique Informatique embarquée et pharmacologie chimique, approches mathématiques de la pharmacocinétique Ingénierie industrielle et informatique Internet des objets Mécanique des fluides et aérodynamique Nanotechnologie, Physique appliquée, Surfaces et matériaux fonctionnels Science du sol, pédologie et environnement Système embarqué et intelligence artificielle Technologie alimentaire et Biotechnologie Technologie de l'éducation UX, Conception centrée sur l'utilisateur
Entretien 2	HEIG-VD	
Entretien 3	HEIG-VD	
Entretien 4	HEIG-VD	
Entretien 5	HEIG-VD	
Entretien 6	HEIG-VD	
Entretien 7	HES-SO Valais-Wallis - HEI	
Entretien 8	Changins	
Entretien 9	Changins	
Entretien 10	HEIG-VD	
Entretien 11	HEIG-VD	
Entretien 12	HEIG-VD	
Entretien 13	HES-SO Valais-Wallis - HEI	
Entretien 14	HEIG-VD	
Entretien 15	HEIG-VD	
Entretien 16	HE-Arc	
Entretien 17	HEIA-FR	
Entretien 18	HEPIA	
Entretien 19	HEPIA	
Entretien 20	HEPIA	
Entretien 21	HE-Arc	
Entretien 22	HEIA-FR	

Tableau 1 - Liste des entretiens

## II. PRATIQUES EN MATIÈRE DE GESTION DES DONNÉES

### 2.1 Types de données

Les chercheur·euses du domaine IA manipulent une grande variété de données, principalement quantitatives et électroniques. La diversité des disciplines implique des outils et pratiques variés, rendant la transition vers des formats libres complexes pour certaines catégories de données (tableau 2).





Catégorie	Données	Outils propriétaires et Formats	Problèmes potentiels (Transition vers formats FAIR)	Complexité de la Transition
<b>Données quantitatives</b>	Mesures, Statistiques, Données tabulaires, Signaux de longueur d'onde de laser, Données de spectre	<b>Excel (XLSX), MATLAB,</b> logiciels de spectroscopie spécifiques	Conversion vers formats ouverts comme CSV, risque de perte de structure ou précision	<b>Faible</b> : Peu de perte de données avec CSV ou TXT
<b>Données qualitatives</b>	Entretiens, Observations, Description des terrains (structure, texture)	<b>Word (DOCX), PDF,</b> logiciels de traitement de texte propriétaires	Conversion facile vers des formats ouverts (TXT, JSON), perte de mise en page, mais peu d'impact sur les données	<b>Très faible</b> : Peu de risques
<b>Modèles et simulations</b>	Modèles numériques (CAO/DAO), Simulations et résultats d'expériences, Modèle physique	<b>Autocad (DWG), SolidWorks, Ansys</b>	Perte de détails ou fonctionnalités spécifiques lors de la conversion vers STEP, DXF	<b>Élevée</b> : Logiciels complexes, peu de solutions libres viables
<b>Données géospatiales</b>	Cartes, SIG (Systèmes d'Information Géographique)	<b>ArcGIS, Google Earth</b>	Difficulté à convertir des formats vers GeoJSON ou KML, perte de métadonnées spatiales	<b>Moyenne</b> : Solutions libres (QGIS) disponibles, mais conversion complexe
<b>Imagerie et échantillons</b>	Images en haute définition, Échantillons de sol, Scanner du sol pour obtenir des images 3D	<b>Photoshop (PSD), RAW, TIFF</b>	Conversion des formats propriétaires (RAW, PSD) vers PNG ou JPEG peut entraîner une perte de qualité	<b>Faible</b> : Perte de qualité possible, mais outils disponibles
<b>Données de conception et plans</b>	Données CAO, Plans mécaniques 3D, Fichiers d'impression 3D, Cartes électroniques, Listes de matériels, Analyse chimique	<b>Autocad (DWG), SolidWorks (SLDPRT), EagleCAD</b>	Conversion vers STEP ou STL entraîne souvent une perte de détails ou de fonctionnalités	<b>Élevée</b> : Logiciels complexes, peu d'alternatives libres
<b>Données biologiques</b>	Séquences biologiques	<b>Propriétaires des laboratoires de séquençage</b>	Conversion vers FASTA possible, mais risques de perte de métadonnées	<b>Moyenne</b> : Formats ouverts disponibles, mais conversion délicate
<b>Documents textuels</b>	Questionnaires, Rapports	<b>Word (DOCX), PDF, Excel (XLSX)</b>	Conversion facile vers des formats ouverts (TXT, CSV), risque minime pour les données	<b>Très faible</b> : Conversion simple et peu de perte
<b>Images</b>	Photos, Dessins, Scans	<b>Photoshop (PSD), RAW, TIFF, JPEG</b>	Compression et perte de qualité possibles lors de la conversion vers PNG, JPEG	<b>Faible</b> : Impact limité sur les données





<b>Données 3D</b>	Schémas, Conception assistée par ordinateur	<b>Autocad (DWG), SolidWorks (SLDPRT)</b>	Conversion vers STL ou DXF peut entraîner une perte de détails et de précision	<b>Élevée</b> : Transition difficile avec peu d'outils libres
<b>Audio</b>	Enregistrements sonores	<b>Audacity (AUP), MP3, WAV</b>	Conversion des fichiers MP3 en formats libres peut réduire la qualité sonore	<b>Faible</b> : Risque de qualité sonore, mais outils disponibles
<b>Vidéo</b>	Enregistrements vidéo	<b>Adobe Premiere (PRPROJ), MP4</b>	Conversion vers des formats ouverts entraîne parfois des pertes de qualité vidéo	<b>Faible</b> : Problèmes minimes avec des outils standards
<b>Mesures</b>	Microscopes, spectromètre, capteur de mouvement, température, hydrométrie	<b>LabVIEW, MATLAB, logiciels spécifiques propriétaires</b>	Conversion vers des formats ouverts est possible, mais perte de précision et métadonnées	<b>Moyenne</b> : Conversion faisable, mais perte de détails
<b>Code informatique et algorithmes</b>	Scripts, Algorithmes	<b>MATLAB, RStudio, Eclipse (JAVA)</b>	Conversion possible vers des formats libres comme Python ou R, mais difficulté de compatibilité	<b>Faible à Moyenne</b> : Outils libres disponibles, mais compatibilité limitée

Tableau 2 - Types de données ; Outils & Formats

Les types de données et les outils utilisés varient grandement en termes de facilité de conversion vers des formats libres conformes aux principes FAIR. Les outils de traitement de texte, comme **Word (DOCX)** ou **Excel (XLSX)**, sont faciles à convertir en formats libres tels que **TXT** ou **CSV** avec peu de risque de perte significative des données. La principale perte est la mise en page ou le formatage, ce qui n'impacte généralement pas la précision des données elles-mêmes. En revanche, pour des logiciels plus complexes comme **Autocad (DWG)**, **SolidWorks** ou **Ansys**, la conversion pose des défis majeurs. Ces outils propriétaires gèrent des données très spécifiques, comme des modèles 3D ou des simulations numériques, qui peuvent inclure des couches, des annotations ou des détails précis difficilement convertibles dans des formats libres comme **STEP** ou **STL** sans perte de précision. De plus, les solutions libres équivalentes à ces logiciels propriétaires sont limitées et souvent moins performantes, ce qui rend la transition beaucoup plus difficile pour des projets complexes. Pour des données comme celles gérées par des systèmes **SIG (ArcGIS)**, des solutions libres existent (**QGIS** par exemple), mais la conversion entre formats peut entraîner la perte de couches géospatiales ou de métadonnées. Ce qui rend ces transitions plus compliquées que des simples documents textuels. Les données issues des logiciels de CAO ou de simulation nécessitent une attention particulière dans la transition vers des formats libres, car les pertes de détails, de calibrage ou de fonctionnalités peuvent affecter les résultats finaux. En revanche, les données textuelles ou basiques peuvent être converties avec peu de contraintes, et les outils libres de traitement de texte sont largement accessibles.

## 2.2 Stockage des données

Les volumes de données manipulés par les chercheur·euses peuvent atteindre des tailles très importantes, allant jusqu'à plusieurs dizaines de To. Lorsque ces grands volumes dépassent les capacités des infrastructures institutionnelles, les chercheur·euses se tournent souvent vers des





solutions personnelles. Celles-ci impliquent fréquemment des pratiques non standardisées et non supervisées ou évaluées par le service informatique de l'institution, ce qui augmente les risques pour l'intégrité des données (voir tableau 3).

Catégorie	Solution	Responsabilité/Utilisation	Remarques
Gestion par équipes ou SI	NAS (Network Attached Storage)	Géré par les équipes de recherche	Partagé entre plusieurs utilisateur·rices pour stockage localisé
	Serveur de données	Géré par les équipes ou le SI	Utilisé pour héberger des données critiques
	Serveur de calcul	Géré par les instituts ou le SI des écoles	Haute performance pour calcul et stockage temporaire
	Serveurs de stockage et d'archivage de l'école	Géré par l'école	Solution centralisée pour conservation à long terme
Solutions de stockage en ligne	OneDrive, SwitchDrive, Google Drive, Dropbox	Partagé entre les utilisateur·rices	Accessible depuis n'importe quel appareil connecté
	REDCap	Utilisé pour les données médicales	Conforme aux normes de sécurité médicale
	Git (GitLab, GitHub)	Pour le versioning et la gestion de code	Utilisé principalement par les développeurs
	Google Docs	Collaboratif	Pour l'édition en temps réel
	Teams	Collaboration et stockage de fichiers	Intégré aux outils Microsoft Office
Stockage local	Ordinateur de bureau	Personnel ou institutionnel	Risque de perte si non sauvegardé
	Ordinateur portable	Personnel ou institutionnel	Mobilité, mais plus fragile
	Disque dur externe	Personnel ou institutionnel	Utile pour les sauvegardes, mais risque de panne matérielle
	Site web	Pour le stockage en ligne accessible	Visibilité externe potentielle
	Base de données	Gestion efficace des données	Adapté pour les grandes quantités de données structurées
Autres	Boîtes e-mail	Communication	Risque de surcharge et perte de traçabilité

Tableau 3 - Cartographie des Différentes Options de Stockage de Données

## 2.3 Archivage à long terme

Les entretiens ont révélé des pratiques d'archivage peu systématiques. Il n'existe peu, voire pas de lignes directrices pour l'archivage des données de recherche. Ces lignes directrices permettraient de renseigner les chercheur·euses sur les données à conserver, celles à détruire, mais aussi à identifier les dépôts de données qui sont conformes aux normes d'archivage (OAIS). Si Zenodo est conforme aux principes FAIR, il n'a pourtant pas la mission de préserver les données à long terme selon les normes OAIS. De plus certaines données doivent être conservées ou supprimées pour des raisons contractuelles, légales ou réglementaires. Ces lignes directrices sont essentielles pour identifier quelles données peuvent être partagées en libre accès.

## 2.4 Partage des données

Sur les 22 chercheurs·euses interrogés.es, 9 ont partagé des données sur des dépôts FAIR non commerciaux. Cependant, ce chiffre ne reflète qu'une partie des pratiques au sein de la HES-SO, l'échantillon étant limité. Les partenariats avec des entreprises constituent un frein majeur au partage des données. Les partenaires privés ayant souvent la titularité des droits d'utilisation. Le faible taux de partage des données ne s'explique pas uniquement par des contraintes techniques, contractuelles ou juridiques. Les chercheur·euses évoquent également un manque de motivation.





Beaucoup estiment que le partage des données représente un coût en temps et en ressources, sans réel bénéfice immédiat. Ils·elles ont des difficultés à percevoir les avantages, en particulier lorsque le champ d'application de leurs données est limité à une communauté restreinte.

Le manque de reconnaissance pour les efforts fournis contribue à freiner le partage des données. De plus, les chercheur·euses interrogé·es perçoivent souvent leurs données comme trop contextuelles ou spécialisées pour être utiles à d'autres.

D'autres freins au partage ont été pointés :

- **Connaissance des dépôts** : La plupart des chercheur·euses connaissent peu les dépôts de données disponibles, qu'ils soient spécialisés ou généralistes.
- **Sélection des données à partager** : Les chercheur·euses hésitent souvent sur le type de données à partager (données brutes, nettoyées, ou seulement les métadonnées).
- **Compréhension des données partagées** : Il est crucial pour eux de s'assurer que les données partagées soient compréhensibles et utilisables par d'autres, ce qui nécessite un effort de documentation et de contextualisation (métadonnées).

## 2.5 Réutilisation des données

La réutilisation des jeux de données déposés sur des plateformes FAIR reste une pratique peu courante. Aucun exemple concret de réutilisation de données provenant de ces dépôts n'a été identifié lors de cette enquête. Bien que certains chercheur·euses consultent des informations dans des bases de données publiques, ils·elles ne considèrent généralement pas cette démarche comme une réutilisation de données, révélant une certaine confusion ou une incompréhension du terme « réutilisation de données ». Il semble également exister une incompréhension concernant la distinction entre les bases de données traditionnelles, reconnues et bien établies au sein de la communauté scientifique, et les dépôts de données qui prennent de plus en plus d'importance avec l'adoption des procédures liées à l'ORD. Alors que les premières sont perçues comme des sources classiques et fiables, les dépôts de données ORD sont encore mal compris, et leur rôle dans la réutilisation des données est souvent sous-estimé. En théorie, les chercheur·euses sont ouvert·es à l'idée d'utiliser des données externes, à condition qu'elles soient pertinentes, fiables et bien documentées. Toutefois, ils·elles se montrent réticents à réutiliser des données sans en évaluer rigoureusement la qualité. Cette méfiance s'explique par divers facteurs: la fiabilité des sources, des méthodologies de collecte potentiellement biaisées, le manque de standardisation, des erreurs humaines possibles, ou encore l'obsolescence des données.

L'absence de transparence sur les processus de collecte et de traitement renforce également ces doutes, amenant les chercheur·euses à remettre en question la pertinence des données partagées pour leur propre recherche. Souvent, ils·elles préfèrent obtenir directement les données auprès de leurs pairs, via leurs réseaux de connaissances. Une exception notable concerne les codes informatiques, qui sont fréquemment partagés et réutilisés via des plateformes de gestion de versions comme GitHub ou GitLab, ainsi que par des Gits institutionnels. Contrairement à Zenodo, Git n'est pas un dépôt de données, mais un système de contrôle de version utilisé pour la gestion des logiciels. La réutilisation des données varie également selon les domaines d'application :

- **Biologie** : Utilisation fréquente de bases de données en ligne telles que NCBI et UniProt.
- **Data Science** : Utilisation de jeux de données disponibles sur des plateformes comme Kaggle.





- **Imagerie** : Usage d'images et de vidéos provenant de sources telles que la NASA ou l'ESA.
- **Environnement** : Exploitation de données provenant de MétéoSuisse, InfoSpécies, SITG (Système d'Information du Territoire Genevois) ou de l'Office fédéral de la topographie.

Les chercheur·euses peuvent également contribuer à enrichir ces bases de données en y ajoutant les données collectées lors de leurs recherches. Cependant, la réutilisation des jeux de données FAIR reste, pour l'instant, limitée et dépend fortement du domaine d'application et des besoins spécifiques des chercheur·euses.

### III. PRINCIPAUX DÉFIS

---

#### 3.1 La « propriété » des données

La question de la propriété des données est un enjeu central dans l'ORD. La majorité des personnes interrogées sont conscientes que les données collectées et/ou produites appartiennent à la haute école qui les emploie. Toutefois, une confusion s'observe notamment sur les droits de propriété intellectuelle. Comme le stipule l'art. 15 de la [Convention intercantonale sur la haute école spécialisée de Suisse occidentale](#) (HES-SO) « les 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 ».<sup>1</sup> Toutefois, les chercheur·euses considèrent dans certains cas que les données produites sont les résultats d'un travail intellectuel, et par conséquent que les données leur appartiennent. Cette confusion peut, dans certains cas, se révéler problématique, notamment dans les deux cas observés lors des entretiens :

- **Absence de contrat spécifique** : Contrairement aux projets avec des partenaires industriels, les projets impliquant plusieurs chercheur·euses ou institutions ne font pas toujours l'objet de contrats spécifiques, laissant floue la question du partage des droits de propriété intellectuelle entre les parties.
- **Projets menés par des étudiant·es** : Les personnes interrogées s'interrogent sur la titularité des droits de propriété intellectuelle pour le corps étudiant·e qui n'est pas contractuellement lié à une haute école.

#### 3.2 Les aspects légaux et contractuels

- **Utilisation des logiciels sous licence** : des questions se posent sur l'utilisation de logiciels sous licence universitaire pour des projets facturés (mandats) ou de recherche (InnoSuisse). Si l'utilisation est restreinte à l'enseignement, il s'agirait de proposer des alternatives pour des logiciels non libres comme Matlab ou des logiciels de CAO (exemples : Ansys, COMSOL, AutoCAD, SolidWorks).
- **Problèmes juridiques avec des partenaires étrangers** : lors de projets internationaux, certain·es chercheur·es s'interrogent sur la législation applicable (suisse ou du pays tiers).

---

<sup>1</sup> Voir l' Art. 15 de la Convention intercantonale sur la haute école spécialisée de suisse occidentale, HES-SO, 2011, p.5.





### 3.3 Transfert des données

Le transfert de données est une pratique courante dans les projets collaboratifs, notamment lorsque plusieurs partenaires sont impliqués. Cependant, ce processus présente des défis importants, en particulier lorsque les données sont sensibles ou contiennent des informations personnelles. Bien que des mesures de sécurité soient souvent mises en place, l'absence de protocoles standardisés et de bonnes pratiques partagées poussent fréquemment les chercheur·euses consulté·es à développer des solutions « internes ». Ces pratiques soulèvent des questions en termes de sécurité et de protection des données.

### 3.4 Partage des données

Plusieurs obstacles rendent le partage des données difficile :

- **Temps et coût** : préparer des données pour le partage est chronophage et coûteux.
- **Projets avec partenaires industriels** : La propriété intellectuelle et les données générées sont souvent protégées par des contrats, limitant ainsi la possibilité de partager ces données.
- **Manque d'intérêt** : Certains chercheur·euses estiment que le partage de leurs données n'apporte pas de réelle valeur, surtout si leur domaine de recherche est très spécialisé.
- **Risque de concurrence** : La crainte de perdre un avantage concurrentiel en partageant des données est un frein important.
- **Propriété des données** : Il y a une confusion récurrente sur la « propriété des données » qui sont partagées.
- **Confusion conceptuelle** : Les chercheur·euses mélangent souvent la valorisation et le partage des données, et d'une manière plus large l'ORD avec l'Open Access.

### 3.5 La réutilisation des données existantes

**Utilisation des données issues de forums ou outils numériques** : Les chercheur·euses se posent des questions sur les régulations légales entourant l'extraction et l'utilisation de données provenant de forums en ligne ou d'autres sources numériques. Il est essentiel de connaître les lois applicables pour garantir une utilisation éthique et légale de ces données.

### 3.6 Soutien

**Le rôle des bibliothécaires** : le rôle des bibliothécaires dans le soutien à l'Open Science, en particulier pour la relecture des DMP, est apprécié. Toutefois, il a été mentionné à plusieurs reprises, qu'un appui supplémentaire de personnes ayant une expérience en recherche, renforcerait la pertinence de cet accompagnement. En d'autres termes, une collaboration entre bibliothécaires et chercheur·euses permettrait de mieux répondre aux besoins spécifiques des différents types de recherche.

### 3.7 Hétérogénéité des données et des formats

Comme souligné au point 2.1, les chercheur·euses produisent ou collectent différents types de données. Cette hétérogénéité pose plusieurs défis liés à la gestion, la standardisation, la confidentialité et la sécurité des données, nécessitant l'adoption de mesures organisationnelles et structurelles appropriées pour rendre les données « FAIR ».





### 3.8 Divers

**Impact environnemental** : le stockage de certaines données est considéré comme inutile, et apparaît comme un point important de préoccupation sur l'impact environnemental d'une telle pratique.

## IV. CONCLUSION ET RECOMMANDATIONS

---

### 4.1 Conclusion

Les pratiques liées à l'ORD sont encore relativement nouvelles et doivent être progressivement intégrées dans les processus de travail des chercheur·euses. Ces changements peuvent être perçus comme complexes, voire perturbants, mais il est essentiel de préparer les équipes de recherche à adopter ces pratiques, qui deviendront progressivement la norme. L'enquête révèle que les pratiques en matière de gestion des données de recherche ne sont pas toujours optimales. Les équipes définissent leur propre méthode, mais peu d'entre elles suivent des procédures formelles, notamment pour la réutilisation et la protection et la sécurité des données. Bien que la nécessité d'une bonne gestion soit reconnue, elle n'est pas systématiquement mise en œuvre, et les pratiques varient d'une équipe à l'autre. Pour encourager une gestion efficace des données de recherche selon les principes FAIR, un soutien renforcé est nécessaire sous forme de formations, de ressources humaines dédiées, d'infrastructures adaptées et de guides pratiques. Bien que certaines initiatives soient déjà en place au sein des hautes écoles et du Rectorat, il est nécessaire d'intensifier les efforts pour sensibiliser les chercheur·euses aux bonnes pratiques de gestion des données et aux avantages du partage des données.

### 4.2 Recommandations

Dans le cadre de l'amélioration de la gestion des données de recherche et de l'adoption des principes FAIR, nous proposons cinq axes de recommandations à court et moyen terme.

#### Guides et outils

- Créer une check-list pour la gestion des données de recherche adaptée aux spécificités du domaine IA
- Créer un guide sur les bonnes pratiques pour l'ouverture des logiciels et codes sources selon les principes FAIR
- Promouvoir et encourager l'utilisation d'outils dédiés pour aider les chercheur·euses à sélectionner la licence la plus appropriée pour leurs codes et logiciels, garantissant une gestion claire des droits d'auteur, des conditions de partage et d'utilisation
- Produire des documents-cadres pour la gestion et l'ouverture des données de recherche, précisant clairement les processus de stockage, de transfert, et d'archivage des données, ainsi que les rôles et responsabilités
- Produire un guide de bonnes pratiques pour une gestion des données qui tient compte des enjeux de durabilité





## Ressources humaines et réseau

- Renforcer et valoriser le réseau des référents en Open Research Data (ORD) en favorisant les collaborations intercantionales pour faciliter le partage des bonnes pratiques et améliorer l'accès aux ressources, notamment dans les cantons où ces services sont un peu moins développés.
- Soutenir l'interaction entre chercheur·euses et bibliothécaires spécialisés en Open Science, en particulier sur les questions d'ORD, pour enrichir et renforcer l'aide existante.
- Renforcer la collaboration entre les équipes de recherche des hautes écoles, les services informatiques et juridiques pour mieux répondre aux exigences de sécurité et de protection des données de recherche.

## Information, formation et sensibilisation

- Proposer un coaching sur la gestion des données de recherche pour tout nouveau projet
- Renforcer les ateliers de sensibilisation sur l'ORD au sein des hautes écoles.
- Améliorer la communication au sein des hautes écoles et du Rectorat pour mieux promouvoir les ressources en ligne et les services de soutien disponibles, tels que les outils et documentations offerts par la HEIG-VD, la HES-SO Valais-Wallis et les pages ORD du Rectorat.
- Favoriser l'interconnexion des sites web des hautes écoles et du Rectorat afin de rendre les ressources plus accessibles et complémentaires. Cela permettrait de capitaliser sur l'expertise des établissements plus avancés et de renforcer l'accès aux services de soutien à travers un réseau intégré.

## Incitations et soutien financier

- **Aide financière** : Allouer un budget spécifique pour soutenir l'ouverture des données de recherche dans les projets non financés par le FNS ou Horizon-Europe, qui incluent déjà des fonds pour l'ORD.
- **Prix ORD du domaine IA** : Créer un fonds dédié à récompenser les chercheur·euses qui partagent activement leurs données et s'engagent dans la promotion de l'Open Research Data.
- **Témoignages et webinaires** : Organiser des présentations de *Success stories* pour démontrer concrètement les avantages du partage des données, notamment en termes de collaborations et de visibilité accrue.

## Valorisation

- **Affichage des DOI** : Intégrer les DOI des jeux de données dans les systèmes internes (comme ARODES) pour donner visibilité et reconnaissance aux chercheur·euses.

