

SwissDS-ENV

Définition et profil de métiers *Data*

Auteur.e.s : Jean-Blaise Claivaz, Claire Wullemin (UNIGE), Gérard Bagnoud, Marielle Guirlet, Carmen Jambé, Sacha Auderset, Guillaume Beausire, Christelle Cocco, Zhargalma Dandarova, Auriane Marmier, Ellina Mourtazina (UNIL), Mathilde Panes (HEP Vaud), Vassilios Ioannidis, Patricia Palagi, Grégoire Rossier (SIB), Eloïse Morisod, René Schneider (HEG, HES-SO Genève), Martin Jaekel, Felix Moerman (ZHAW)

Formule de citation : Projet SwissDS-ENV, *Définition et profil de métiers Data*, Université de Lausanne, 2023.



Ce document est un livrable du projet Swiss Data Stewardship Environment : Profile - Training - Network (SwissDS-ENV), financé par swissuniversities dans le cadre de la ligne d'action B5.2 : Plans de mesures concernant le data stewardship.

La [licence CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) sous laquelle est publié ce document vous autorise à:

- **Partager** : copier, distribuer et communiquer le matériel par tous moyens et sous tous formats
- **Adapter : remixer, transformer et créer à partir du matériel pour toute utilisation, y compris commerciale**

à condition d'en créditer les auteur.e.s, d'intégrer un lien vers la licence, et d'indiquer si des modifications ont été faites.

Version	Date	Notes
v1	28.03.2023	Contenu section 1
v2	05.2023	Modifications texte
v2-1	24.05.2023	Sélection des métiers et représentations radar
v3-1	31.05.2023	Cycle de vie et fiches-métier
v4	07.07.2023	Ajout des figures 1, 2, 3 ; mise en page
v5	28.07.2023	Prise en compte des derniers commentaires et nouvelle version des représentations radar
VF_2023	02.08.2023 15.09.2023	Page de couverture Légende figure 3 (cycle de vie adapté de FORS 2023)
VF_2023_2	26.10.2023	Ajout d'une phrase de contexte et d'une conclusion Version en français : termes en français pour les métiers et les étapes du cycle de vie des données (liste des métiers, figures et fiches-métier)

Table des matières

1. Contexte	1
2. Présentation des métiers	2
2.1. Métiers retenus.....	2
2.2. Contributions aux étapes du cycle de vie des données	2
3. Description détaillée sous forme de fiches-métier	4
3.1. Coordinateur·rice des données/responsable gestion des données.....	4
3.2. Data Steward.....	5
3.3. Bibliothécaire de données.....	7
3.4. Scientifique de données	8
3.5. Curateur·rice de données	9
3.6. Archiviste de données.....	10
4. Conclusion	10
Bibliographie	11

Liste des figures

Figure 1: Cycle de vie des données (adapté de FORS 2023)	1
Figure 2: Contribution des métiers Data aux différentes étapes du cycle de vie des données (cycle de vie adapté de FORS 2023).....	3
Figure 3 : Liste de tâches effectuées par les Data Stewards à chaque étape du cycle de vie et niveau de contribution (« conseiller » : plus faible ; « soutenir » : plus fort). Cycle de vie adapté de FORS (2023)	6

Liste des tableaux

Tableau 1: Fiche-métier coordinateur·rice des données/responsable gestion des données.....	4
Tableau 2 : Fiche-métier Data Steward.....	5
Tableau 3: Fiche-métier bibliothécaire de données	7
Tableau 4 : Fiche-métier scientifique de données	8
Tableau 5: Fiche-métier curateur·rice	9
Tableau 6 : Fiche-métier archiviste de données.....	10

1. Contexte

Les résultats présentés ci-dessous ont été produits dans le cadre du projet Swiss Data Stewardship Environment (SwissDS-ENV), financé par swissuniversities, et qui a pour objectif principal de développer une formation certifiante en Data Stewardship.

Autour de la problématique de plus en plus prégnante de la gestion des données de recherche, de nouvelles professions apparaissent tandis que d'anciennes évoluent et se recomposent pour répondre aux demandes des différentes parties prenantes. Étant donné la pluralité de domaines où ces professions interviennent, leurs titres, leurs définitions, leurs périmètres d'action se chevauchent, et plus généralement leurs caractéristiques sont intimement liées aux contextes dans lesquels elles s'inscrivent (milieu, institutions, équipes, approche généraliste versus spécialisée, etc.).

Afin de clarifier la situation et d'utiliser un langage commun, il est apparu nécessaire de préciser les différents termes et concepts qui seront utilisés dans le cadre du projet SwissDS-ENV. Pour ce faire, nous nous appuyons sur le cycle de vie des données de recherche en six étapes de FORS (2023, p.9) (Figure 1) et nous ne retiendrons que six professions que nous pensons représentatives et nécessaires pour couvrir le spectre des besoins des différentes situations et institutions de recherche.



Figure 1: Cycle de vie des données (adapté de FORS 2023)

2. Présentation des métiers

2.1. Métiers retenus

La liste des six métiers retenus a été réalisée à la suite d'une revue de la littérature et des projets existants. Elle se base sur le paysage des métiers liés à la gestion des données présenté aux pages 10 à 13 par Wuillemin (2022). Le premier métier (*Data Coordinator/Data Manager Officer*) est un peu différent des autres car il est lié à la gouvernance et à la supervision de la gestion des données de recherche au niveau institutionnel, tandis que les cinq autres métiers sont plus directement en prise avec le soutien direct aux chercheur·euse·s lors des étapes de création, utilisation, analyse, stockage et préservation des données de recherche. D'autres terminologies existent que celles que nous avons retenues, et nous en mentionnons certaines dans les fiches descriptives de chacun de ces métiers.

Liste des métiers

1. Coordinateur·rice des données/Responsable gestion des données (*Data Coordinator/Data Manager Officer*)
2. Data Steward
3. Bibliothécaire de données (*Data Librarian*)
4. Scientifique de données (*Data Scientist*)
5. Curateur·rice de données (*Data Curator*)
6. Archiviste de données (*Data Archivist*)

2.2. Contributions aux étapes du cycle de vie des données

L'interaction entre les étapes du cycle de vie des données et les métiers permet d'illustrer, de manière très schématique, les principales différences et le périmètre d'action des professions retenues. À titre d'exemple, dans le graphique ci-dessous (Figure 2) nous voyons que le·la Data Steward est impliqué·e à tous les échelons du cycle de vie, tandis que l'archiviste de données n'est sollicité·e que pendant la deuxième partie du cycle de vie. Ce qui n'apparaît pas dans cette représentation, c'est l'intensité avec laquelle un métier intervient dans une étape. Cette information doit être cherchée dans la fiche descriptive du métier, et plus spécifiquement dans la représentation sous forme de radar (voir section 3). Pour le·la Data Steward, cette représentation est complétée par une autre figure (Figure 3) détaillant les tâches effectuées par étape du cycle de vie. On distingue ici deux niveaux de contribution : conseiller (contribution plus faible) et soutenir (contribution plus forte). D'un contexte à l'autre, en fonction de l'environnement institutionnel et des ressources, ce niveau de contribution pourra varier légèrement de celui représenté sur cette figure.

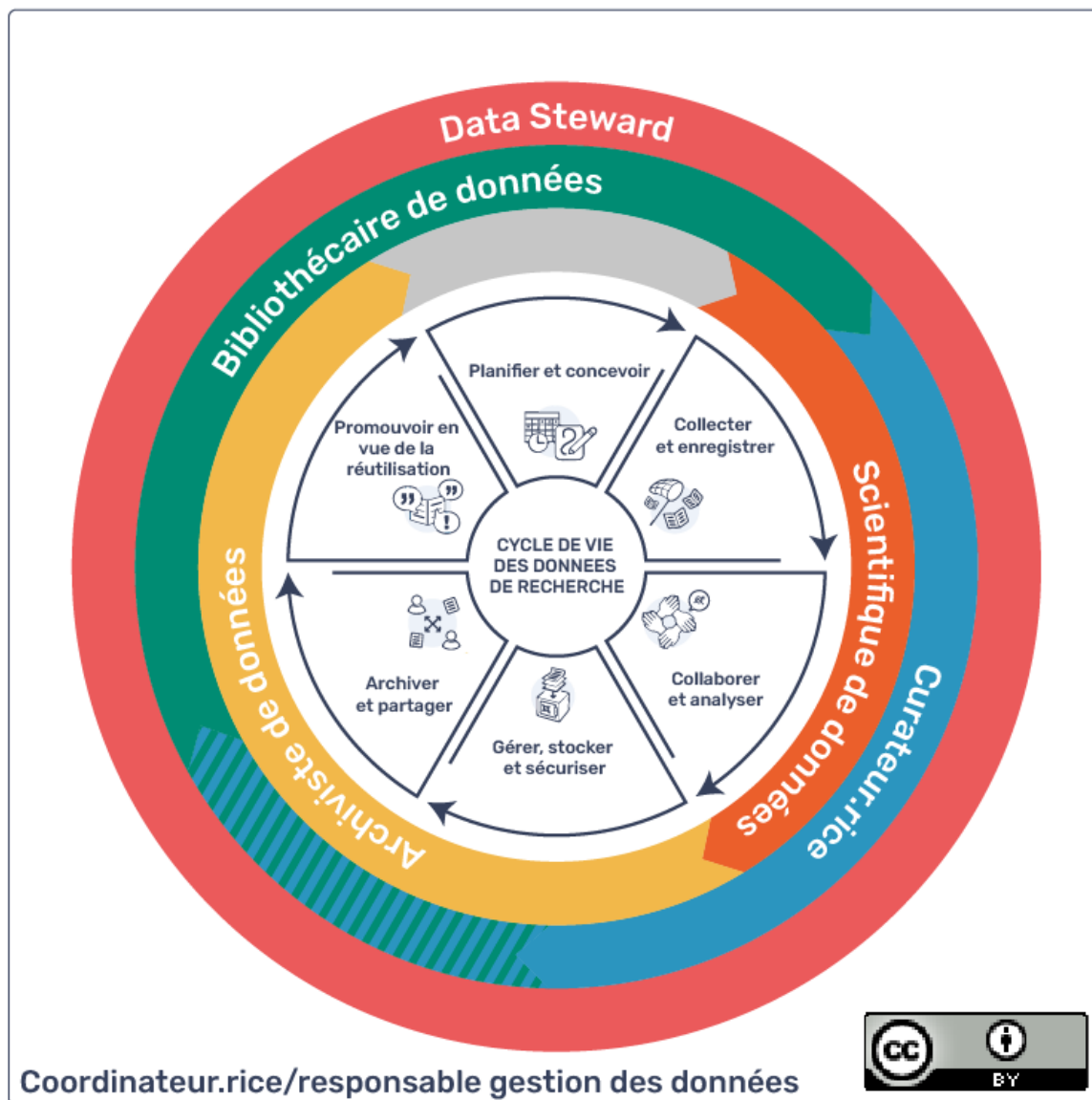


Figure 2: Contribution des métiers Data aux différentes étapes du cycle de vie des données (cycle de vie adapté de FORS 2023)

3. Description détaillée sous forme de fiches-métier

3.1. Coordinateur·rice des données/responsable gestion des données

Autre(s) appellation(s) / Recoupement avec	<i>Data Coordinator, Data Manager Officer (Chief) Data Manager, (Chief) Data Officer</i>
Mots-clés (tags)	Gouvernance, stratégies, politiques, conformité, bonnes pratiques, infrastructures, pilotage, supervision, coordination, soutien, relations internes et interinstitutionnelles
Définition	
<p>Agissant au niveau institutionnel et transversal, le.la Coordinateur·rice des données/Responsable gestion des données occupe un poste stratégique et politique au sein de l'institution. Responsable de la gouvernance des données de recherche tout au long de leur cycle de vie, il-elle contribue à l'établissement de stratégies et politiques conformes aux diverses exigences – légales, réglementaires ; nationales et internationales – en matière de gestion (GDR) et ouverture (ORD) des données de recherche. De même, le.la Data Coordinator/Data Manager Officer participe au développement d'infrastructures et outils ; à l'établissement de bonnes pratiques, recommandations, (in)formations destinées aux chercheur·euse·s. En outre, il-elle coordonne les différents métiers impliqués dans le soutien à la gestion des données de recherche.</p>	
Mission(s)	<ul style="list-style-type: none"> • Superviser, à l'échelle institutionnelle, la gouvernance des données de recherche, ainsi que le soutien à leur gestion et ouverture, tout au long du cycle de vie • Assurer la conformité avec les exigences – nationales et internationales – et les normes et standards en matière de gestion et ouverture des données de recherche
Exemples de tâches	<ul style="list-style-type: none"> • Développer des stratégies et politiques institutionnelles de gestion et ouverture des données de recherche • Instaurer un cadre de bonnes pratiques de gestion et ouverture des données de recherche • Développer des infrastructures/outils de gestion et d'ouverture des données de recherche ; • Développer du matériel de communication et (in)formation destiné aux chercheur·euse·s • Coordonner les différentes parties prenantes impliquées dans le soutien à la gestion des données de recherche, notamment les Data Stewards. • Représenter l'institution dans diverses instances nationales ou internationales
Étapes du cycle de vie <ul style="list-style-type: none"> • <i>Planifier et concevoir</i> • <i>Collecter et enregistrer</i> • <i>Collaborer et analyser</i> • <i>Gérer, stocker et sécuriser</i> • <i>Archiver et partager</i> • <i>Promouvoir en vue de la réutilisation</i> 	<p>Le.la coordinateur·rice des données/responsable gestion des données intervient à toutes les étapes du cycle de vie de façon transversale</p>

Tableau 1: Fiche-métier coordinateur·rice des données/responsable gestion des données

3.2.Data Steward

Autre(s) appellation(s)	<i>Data Consultant, (Research) Data Manager, Research Data (Management) Specialist</i> Consultant-e pour les données, gestionnaire de données (de recherche), spécialiste en gestion des données
Mots-clés (tags)	Soutien aux chercheur.euse.s, bonnes pratiques de gestion et ouverture des données de recherche, conformité aux exigences
Définition	
<p>De par leur expertise consultative, les Data Stewards soutiennent et facilitent la gestion par les chercheur.euse.s des données (et métadonnées) de recherche tout au long de leur cycle de vie. Il-elle-s sont souvent au bénéfice d'une formation et d'une expérience professionnelle dans la recherche.</p> <p>Les Data Stewards peuvent être embarqué.e-s (<i>embedded</i>) ou génériques (<i>generic</i>) (Verheul et al. 2017, p.7): les <i>Embedded Data Stewards</i> sont souvent rattachés à une unité/département de recherche spécifique et directement impliqués dans les recherches menées. Il-elle-s possèdent de solides connaissances dans les méthodes de travail propres à leur entité de rattachement ce qui leur permet de soutenir opérationnellement les chercheur.euse.s par exemple dans la création de code ou de scripts pour l'analyse des données, mais aussi de traduire les politiques institutionnelles en mesures opérationnelles en adéquation avec les réalités du domaine.</p> <p>Pour leur part, les <i>Generic Data Stewards</i> sont situés au sein d'un service d'assistance ou au niveau d'une faculté et possèdent des connaissances générales sur la gestion des données de recherche ainsi qu'une vision globale de la situation institutionnelle qui leur permettent de répondre à toutes sortes de questions et de les réorienter si besoin vers d'autres services institutionnels compétents.</p>	
Mission(s)	<ul style="list-style-type: none"> • Promouvoir les bonnes pratiques de gestion des données de recherche au sein de l'institution et auprès des chercheur.euse.s ; leur apporter conseil, soutien et expertise en la matière • Participer au développement de projets institutionnels sur la gestion et l'ouverture des données de recherche
Exemples de tâches	<ul style="list-style-type: none"> • Promouvoir les stratégies, politiques institutionnelles et bonnes pratiques de gestion et ouverture des données de recherche • Apporter un soutien aux chercheur.euse.s pour la gestion des données de recherche tout au long de leur cycle de vie • Informer les chercheur.euse.s sur les infrastructures/outils existants pour la gestion et l'ouverture des données de recherche • Co-développer une offre de formations aux compétences de gestion des données de recherche • Relayer les chercheur.euse.s vers des expert.e.s (techniques, DPO, éthiques, légaux, ...) si besoin
Étapes du cycle de vie	<ul style="list-style-type: none"> • Planifier et concevoir • Collecter et enregistrer • Collaborer et analyser • Gérer, stocker et sécuriser • Archiver et partager • Promouvoir en vue de la réutilisation

Tableau 2 : Fiche-métier Data Steward

Définition et profil de métiers Data

Projet SwissDS-ENV – Action 1 : définition et profil des Data Stewards

02.08.2023

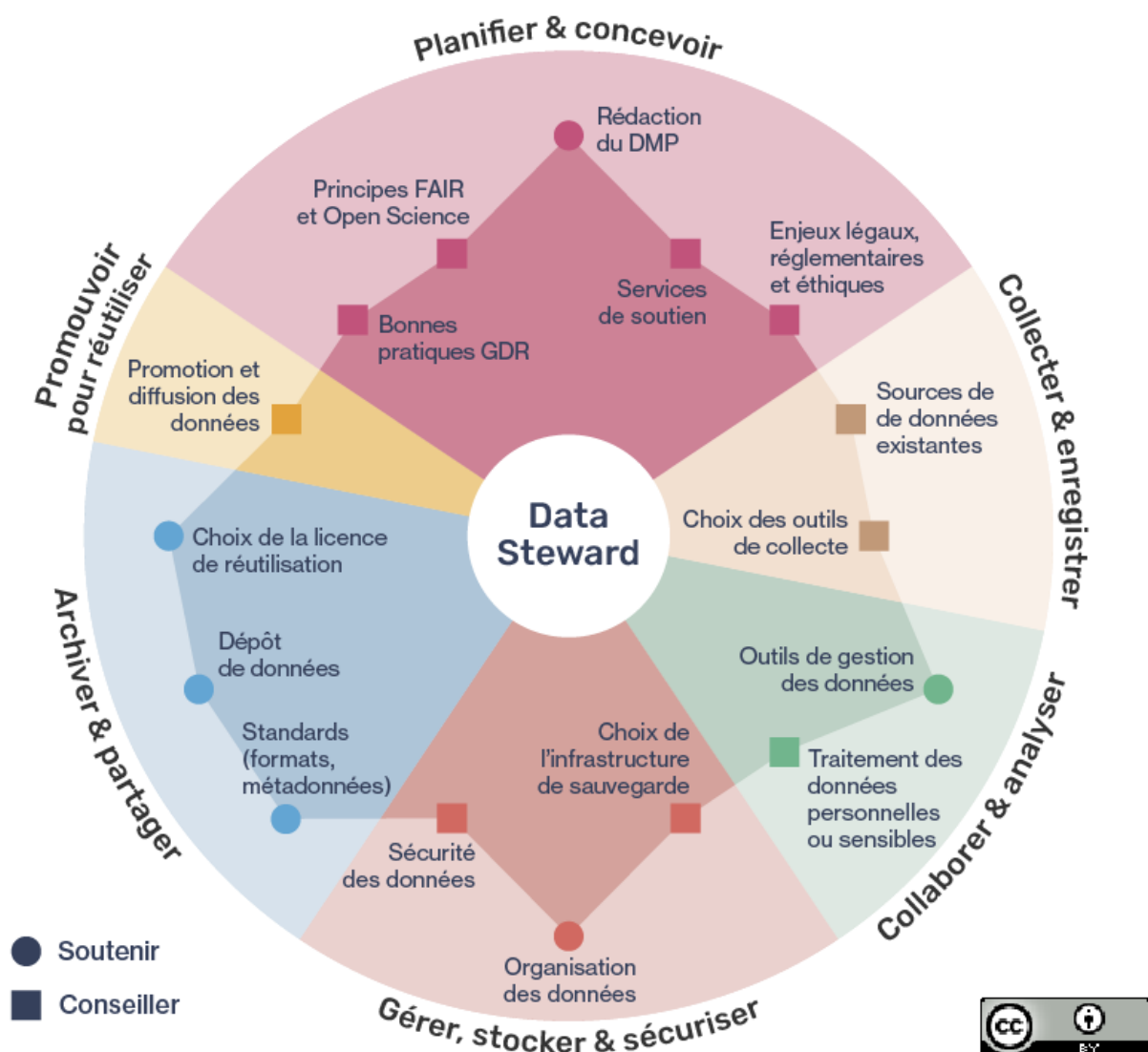


Figure 3 : Liste de tâches effectuées par les Data Stewards à chaque étape du cycle de vie et niveau de contribution (« conseiller » : plus faible ; « soutenir » : plus fort). Cycle de vie adapté de FORS (2023)

3.3. Bibliothécaire de données

Autre(s) appellation(s)	<i>Data Services Librarian, Research Librarian</i> Bibliothécaire spécialisé-e en services de gestion des données, bibliothécaire de recherche
Mots-clés (tags)	Spécialiste bibliothéconomie et gestion des données de recherche, soutien aux chercheur·euse·s
Définition	
Professionnel·le·s de l'information spécialisé·e·s dans la gestion de recherche, les bibliothécaires de données sont responsables de mettre en œuvre des services et des solutions de soutien et de formation pour la gestion des données de recherche et la diffusion de ses bonnes pratiques auprès des chercheur·euse·s et du personnel institutionnel.	
Mission(s)	<ul style="list-style-type: none"> • Mettre les compétences bibliothéconomiques au service de la gestion des données de recherche (copyright, préservation, ...) • Faciliter le travail de gestion des données de recherche par les chercheur·euse·s
Exemples de tâches	<ul style="list-style-type: none"> • Préparer et animer des formations relatives à la gestion des données de recherche • Conduire des actions de communication, de sensibilisation et d'implication de la communauté • Soutenir la rédaction de demandes de subventions (<i>grants</i>), par exemple en tant que spécialistes Open Science (voir Féret et Cros 2019) • Soutenir la gestion des données de recherche, notamment : <ul style="list-style-type: none"> ○ Aider à la rédaction et/ou relecture et évaluation de DMP ○ Participer à la gestion d'un <i>data repository</i> institutionnel ○ Identifier/localiser des données réutilisables ○ Orienter le choix vers une licence appropriée pour les jeux de données
Étapes du cycle de vie <ul style="list-style-type: none"> • Planifier et concevoir • Collecter et enregistrer • Gérer, stocker et sécuriser • Archiver et partager • Promouvoir en vue de la réutilisation 	<p>Bibliothécaire de données et cycle de vie des données</p> <p>Le diagramme est un hexagone à six axes. Les axes sont étiquetés : 'Planifier et concevoir' (en haut), 'Collecter et enregistrer' (en haut-droite), 'Collaborer et analyser' (en bas-droite), 'Gérer, stocker et sécuriser' (en bas), 'Archiver et partager' (en bas-gauche), et 'Promouvoir en vue de la réutilisation' (en haut-gauche). À l'intérieur de l'hexagone, il y a cinq lignes concentriques numérotées de 0 à 5. Une ligne verte relie les points de performance sur chaque axe, montrant des scores élevés (proches de 5) pour 'Planifier et concevoir', 'Collecter et enregistrer' et 'Collaborer et analyser', et des scores plus faibles (autour de 2-3) pour 'Gérer, stocker et sécuriser', 'Archiver et partager' et 'Promouvoir en vue de la réutilisation'.</p>

Tableau 3: Fiche-métier bibliothécaire de données

3.4. Scientifique de données

Autre(s) appellation(s)	/
Mots-clés (tags)	Programmation, statistiques, analyse de données
Définition	
<p>Les scientifiques de données conseillent et soutiennent les chercheur·euse·s pour l'analyse et la modélisation de leurs données, pour répondre à des questions scientifiques précises ou pour développer des algorithmes qui exécutent automatiquement certaines tâches difficiles. Il·elle·s sont expert·e·s en <i>machine learning</i>, statistiques, <i>computer vision</i>, <i>natural language processing</i>, ou autres disciplines et compétences connexes. Les <i>Data Scientists</i> peuvent évoluer autant dans les milieux académiques qu'en entreprise ou dans l'industrie. Leurs fonctions sont hautement dépendantes de leur contexte de travail.</p>	
Mission(s)	<ul style="list-style-type: none"> • Conseiller et soutenir les chercheur·euse·s dans la collecte, le traitement et l'exploitation et l'analyse de leurs données • Développer et mettre en œuvre les outils adéquats
Exemples de tâches	<ul style="list-style-type: none"> • Développer des modèles et/ou algorithmes/scripts • Produire et appliquer des outils adaptés, par exemple : <ul style="list-style-type: none"> ◦ Formulaire de collecte de données ◦ Scripts de traitement ◦ Outils de visualisation et d'analyse ◦ Outils de contrôle qualité et validation formelle
Étapes du cycle de vie <ul style="list-style-type: none"> • Collecter et enregistrer • Collaborer et analyser 	<p>Scientifique de données et cycle de vie des données</p>

Tableau 4 : Fiche-métier scientifique de données

3.5. Curateur·rice de données

Autre(s) appellation(s)	/
Mots-clés (tags)	Collecte de données, sélection, traitement, enrichissement, documentation et archivage
Définition	
Les curateur·rice·s assurent les tâches de traitement des données en vue d'une préservation à long terme et d'une réutilisation. Leurs activités permettent la découverte et l'accessibilité aux données, le maintien de leur qualité, l'ajout de valeur et leur réutilisation dans le temps (Cragin sans date). Ces activités comprennent la collecte des données, leur sélection, leur enrichissement, leur préservation, leur maintenance et leur archivage. Ces activités s'effectuent souvent dans le contexte d'une archive de données par les curateurs de cette infrastructure (Johnston 2018).	
Mission(s)	<ul style="list-style-type: none"> • Soutenir les chercheur·euse·s pour la collecte, l'évaluation et l'enrichissement des données • Aider à la sélection des données à forte valeur ajoutée et les aider pour leur préparation avant versement dans le système d'archivage pour préservation et partage
Exemples de tâches	<ul style="list-style-type: none"> • Visualiser les données, aider à les évaluer, les enrichir (y.c les métadonnées) • Conseiller les chercheur·euse·s dans la sélection des données en vue de leur préservation • Conseiller pour le choix des formats les plus conformes pour la préservation • Compléter les métadonnées • Valider la qualité des données avant archivage et publication
Étapes du cycle de vie <ul style="list-style-type: none"> • Collecter et enregistrer • Collaborer et analyser • Gérer, stocker et sécuriser • Archiver et partager 	<p>Curateur·rice et cycle de vie des données</p>

Tableau 5: Fiche-métier curateur·rice

3.6. Archiviste de données

Autre(s) appellation(s)	/
Mots-clés (tags)	Archivage et préservation des données
Définition	
Les archivistes de données assurent les tâches requises pour la préservation au long terme des données de recherche. Les archivistes de données sont responsables de l'archivage de données de recherche sans distinction : il·elle·s n'effectuent pas de sélection délibérée de données, mais traitent celles-ci comme un patrimoine à part entière qu'il est important de sauvegarder.	
Mission(s)	<ul style="list-style-type: none"> • Mettre les compétences archivistiques au service de la gestion des données de recherche (<i>copyright</i>, préservation, ...) • Faciliter la préservation des données par les chercheur·euse·s
Exemples de tâches	<ul style="list-style-type: none"> • Soutenir les chercheur·euse·s pour la préparation des données en vue de leur archivage • Mettre en œuvre des normes institutionnelles et/ou nationales de conservation des données de recherche comme biens informationnels de l'institution, en collaboration avec les autorités de l'institution et son administration. • Orienter le choix vers une licence appropriée pour les jeux de données
Étapes du cycle de vie	<ul style="list-style-type: none"> • Gérer, stocker et sécuriser • Archiver et partager • Promouvoir en vue de la réutilisation
	<p>Archiviste de données et cycle de vie des données</p>

Tableau 6 : Fiche-métier archiviste de données

4. Conclusion

Les résultats présentés ci-dessus permettent de mieux cerner les missions et les responsabilités typiques des principaux métiers Data, ainsi que les étapes du cycle de vie des données pendant lesquelles ces professions interviennent. Nous avons également caractérisé plus précisément le profil de Data Steward, en identifiant les tâches spécifiques à ce profil à chaque étape du cycle de vie des données, ainsi que le niveau de sa contribution. Dans la suite du projet SwissDS-ENV, une grille de compétences pour les Data Stewards, la plus exhaustive possible, sera développée en se basant sur les premiers résultats présentés dans ce document, ceci dans le but d'identifier le plus finement possible les besoins et de formuler au mieux les objectifs de la formation certifiante en Data Stewardship.

Définition et profil de métiers Data

Projet SwissDS-ENV – Action 1 : définition et profil des Data Stewards

02.08.2023

Bibliographie

CRAGIN, Melissa H, HEIDORN, P Bryan, PALMER, Carole L and SMITH, Linda C, [sans date]. An Educational Program on Data Curation. p. 1.

FERET, Romain, et CROS, Marie, 2019. « The Embedded Research Librarian: A Project Partner ». *LIBER Quarterly: The Journal of the Association of European Research Libraries* 29 (1): 1-20. <https://doi.org/10.18352/lq.10304>.

FORS, 2023. Introduction à SWISSUBase. Présentation à l'Université de Lausanne. 6 juin 2023.

JOHNSTON, Lisa R, CARLSON, Jacob, HUDSON-VITALE, Cynthia, IMKER, Heidi, OLENDORF, Robert and STEWART, Claire, 2018. How Important Are Data Curation Activities to Researchers? Gaps and Opportunities for Academic Libraries. *Journal of Librarianship and Scholarly Communication*. 2018. Vol. 6, p. 25.

VERHEUL, Ingeborg, IMMING, Melanie, RINGERMA, Jacquelyn, MORDANT, Annemie, van der PLOEG, Jan-Lucas, et PRONK, Martine. 2019. « Data Stewardship on the map: A study of tasks and roles in Dutch research institutes ». Zenodo. <https://doi.org/10.5281/zenodo.2669150>.

WUILLEMIN, Claire. *État des lieux des services de soutien à la gestion des données de recherche à l'Université de Genève*. 2022 <https://archive-ouverte.unige.ch/unige:162963>