



Retour d'expérience sur l'atelier “ Gestion des données de la recherche : réalités de terrain, profils métiers et pistes d'action ”

Diane Antoine, Sylvie Grésillaud, Claire Sowinski

► To cite this version:

Diane Antoine, Sylvie Grésillaud, Claire Sowinski. Retour d'expérience sur l'atelier “ Gestion des données de la recherche : réalités de terrain, profils métiers et pistes d'action ”. 2026. <hal-05742475>

HAL Id: hal-05742475

<https://hal.science/hal-05742475v1>

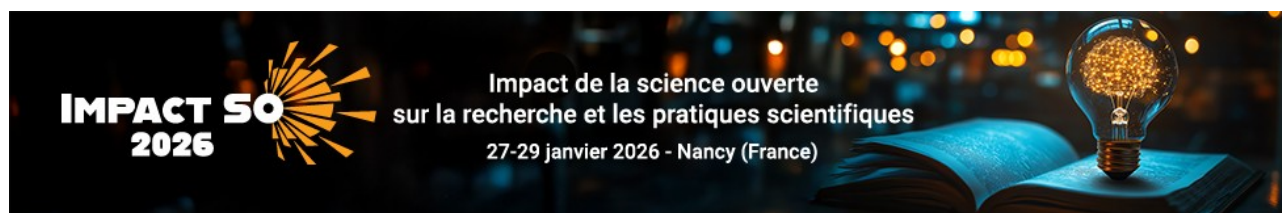
Preprint submitted on 8 Sep 2026

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

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



ETALAB - Open licence



Retour d'expérience sur l'atelier « Gestion des données de la recherche : réalités de terrain, profils métiers et pistes d'action »

Produit le 17 août 2026

Animatrices et autrices :

Diane Antoine (Inist-CNRS), Administratrice et rédactrice des plateformes e-learning DoRANum
(<https://orcid.org/0009-0007-0228-7980>)

Sylvie Gresillaud (Inist-CNRS), Chargée de missions transversales et coordinatrice du réseau Renatis (<https://orcid.org/0000-0001-6806-4893>)

Claire Sowinski (Inist-CNRS), Responsable du service formation - DoRANum
(<https://orcid.org/0000-0002-1048-8041>)

Sommaire

Résumé :.....	1
Abstract.....	2
1. Introduction et contexte.....	3
1.1. La gestion des données de la recherche : un enjeu structurant de la science ouverte.....	3
1.2. Des métiers en évolution et des contours encore en construction.....	3
2. Méthodologie.....	5
2.1. Objectifs et principes directeurs.....	5
2.2. Phase préparatoire.....	5
2.2.1. Sélection et analyse des ressources.....	5
2.2.2. Préparation des matériaux pédagogiques.....	6
2.3. Déroulé méthodologique de l'atelier.....	9
2.4. Modalités d'exploitation des résultats.....	9
3. Résultats et analyse.....	10
3.1. Profils et reconnaissance des métiers.....	10
3.3. Enjeux transversaux.....	15
4. Discussion.....	15
5. Conclusion et points de vigilance.....	16

Résumé :

La gestion des données de la recherche s'impose comme un enjeu central des politiques de science ouverte, porté par des initiatives européennes et nationales. Bien que les principes FAIR structurent désormais les pratiques, les métiers associés à la gestion des données restent en construction, avec des contours flous et une grande diversité de profils selon les disciplines et les établissements. Cet article présente un atelier participatif organisé lors des journées Impact SO 2026 à Nancy¹, visant à confronter des référentiels métiers théoriques (Skills4EOSC, Ministère de l'Enseignement Supérieur et de la Recherche) aux réalités professionnelles du terrain.

L'atelier a réuni des professionnel-le-s aux profils variés pour identifier les missions effectives, les compétences mobilisées et les besoins, notamment en formation. Les résultats révèlent un alignement limité entre les profils existants et les pratiques réelles : les participant-e-s cumulent des rôles multiples, avec une forte transversalité des activités. Les profils les plus représentés après normalisation sont ceux de documentaliste des données (MESR) et de data steward (Skills4EOSC). Les besoins exprimés portent sur des formations juridiques, éthiques et techniques, ainsi que sur une meilleure organisation des missions au sein des établissements.

L'article souligne un écart entre les modèles théoriques et les réalités du terrain, où les métiers de la gestion des données se construisent de manière pragmatique, en réponse aux besoins locaux. Les référentiels existants, bien qu'utiles, ne capturent pas cette diversité. Les perspectives incluent l'adaptation de profils aux spécificités locales, le développement de formations ciblées et la valorisation de ces métiers émergents.

L'atelier est disponible à la réutilisation pour compléter ces premiers résultats.

Abstract

Research data management has become a central issue in Open Science policies, driven by European and national initiatives. The FAIR principles now structure practices but the associated professions remain under construction, with unclear boundaries and significant diversity in profiles observed across disciplines and institutions. This article presents a participatory workshop that was held during the Impact SO conference in Nancy, France, with the aim of comparing theoretical frameworks (Skills4EOSC, French Ministry of Higher Education and Research) with professional realities.

The workshop brought together professionals with diverse backgrounds to identify actual missions, skills, and training needs. The results reveal a limited alignment between existing profiles and real-world practices, for example participants hold multiple roles, with strong cross-disciplinary activities. The most represented profiles after normalization are data documentation specialist (MESR) and data steward (Skills4EOSC). The needs expressed focus on legal, ethical, and technical training, as well as better organization of missions within institutions.

The article highlights a gap between theoretical models and field realities in which research data management professions are constructed pragmatically in response to local needs. While existing

¹ <https://www.science-ouverte.cnrs.fr/fr/actualite/journees-impact-so-2026-les-inscriptions-sont-ouvertes/>

frameworks are useful, they fail to capture this diversity. Future perspectives include adapting profiles to local specificities, developing targeted training, and valuing these emerging professions. The workshop is available for use to complement these initial findings.

1. Introduction et contexte

1.1. La gestion des données de la recherche : un enjeu structurant de la science ouverte

Au cours de la dernière décennie, la gestion des données de la recherche est devenue un pilier des politiques de science ouverte. Portée par les recommandations de la Commission européenne², les initiatives de l'European Open Science Cloud (EOSC)³, ainsi que par les plans nationaux pour la Science ouverte en France⁴, cette évolution traduit une volonté de rendre les résultats de la recherche plus transparents, reproductibles et réutilisables. Les données de la recherche sont désormais reconnues comme des productions scientifiques à part entière dont la qualité, la documentation, la conservation et le partage conditionnent leur potentiel de réutilisation et, plus largement, leur contribution à l'avancement des connaissances.

Les principes FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable), formulés par Wilkinson et al. en 2016⁵, constituent aujourd'hui un cadre de référence largement adopté pour améliorer la gestion et la diffusion des données de recherche. Leur mise en œuvre s'accompagne de nouvelles exigences : élaboration de plans de gestion de données, documentation des jeux de données, prise en compte des dimensions juridiques et éthiques, dépôt dans des entrepôts pérennes, utilisation de standards de métadonnées et d'identifiants persistants, ou encore accompagnement des chercheur·e·s dans l'ensemble du cycle de vie des données. Au-delà des aspects techniques, ces évolutions impliquent une transformation des pratiques professionnelles, des organisations et des modes de collaboration au sein des établissements de recherche.

La mise en œuvre de ces politiques repose sur une grande diversité d'acteur·rice·s. Chercheur·e·s, ingénieur·e·s, technicien·ne·s, personnels des bibliothèques, informaticien·ne·s, juristes, spécialistes de la qualité des données et responsables d'infrastructures contribuent collectivement à construire un écosystème favorable à une gestion responsable des données. Cette dimension collaborative est régulièrement soulignée dans la littérature⁶, qui insiste sur la nécessité de développer des compétences distribuées et des dispositifs d'accompagnement adaptés aux besoins des communautés scientifiques.

² <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018H0790>

³ <https://eosc.eu/>

⁴ <https://www.ouvrirlascience.fr/deuxieme-plan-national-pour-la-science-ouverte/>

⁵ Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

⁶ <https://arxiv.org/abs/2109.01694>

1.2. Des métiers en évolution et des contours encore en construction

L'importance croissante accordée à la gestion des données s'est traduite par l'émergence de nouvelles fonctions, ainsi que par une profonde évolution de métiers pré-existants. Les fonctions de data steward, de research data manager ou de data librarian se sont progressivement développées dans de nombreux pays, tandis que les personnels d'appui à la recherche et professionnel·le·s de l'information ont vu leurs missions s'élargir pour intégrer l'accompagnement à la gestion des données. Cette évolution répond autant aux attentes des politiques publiques qu'aux besoins exprimés par les communautés scientifiques.

Toutefois, malgré cette dynamique, les contours de ces métiers demeurent encore en construction. Plusieurs travaux soulignent l'hétérogénéité des intitulés de poste, la diversité des compétences mobilisées et les variations importantes selon les disciplines, les établissements ou les infrastructures de recherche^{7 8 9}. Les référentiels élaborés ces dernières années, notamment les Minimum Viable Skills (MVS) issus du projet Skills4EOSC¹⁰ et les fonctions clés liées à la gestion, l'exploitation, la diffusion et l'utilisation des données définies par le Ministère de l'Enseignement Supérieur et de la Recherche (MESR)¹¹, témoignent de la volonté de mieux caractériser ces professions. Néanmoins, ils mettent également en évidence le caractère évolutif de ces fonctions, situées à l'interface entre les activités scientifiques, les services d'appui, les infrastructures numériques et les enjeux de gouvernance des données.

Dans ce contexte, la professionnalisation des métiers de la gestion des données ne peut être appréhendée uniquement à travers les référentiels institutionnels. Elle se construit également au travers de pratiques quotidiennes, des coopérations entre métiers, des communautés professionnelles et des retours d'expérience des acteur·rice·s sur le terrain. Les dispositifs participatifs constituent un moyen privilégié pour faire émerger une compréhension collective des missions, des compétences et des difficultés rencontrées.

C'est dans cette perspective qu'a été conçu l'atelier participatif présenté dans cet article. Son objectif était de réunir des professionnel·le·s issus de différents horizons afin de recueillir leur perception de leur métier et de leurs missions. L'atelier « **Gestion des données de la recherche : réalités de terrain, profils métiers et pistes d'action** » s'est tenu dans le cadre des journées

⁷ https://rdmkit.elixir-europe.org/data_steward

⁸ Jetten, M., Grootveld, M., Mordant, A., Jansen, M., Bloemers, M., Miedema, M. & Van Gelder, C. W. G. (2021). Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship (Version 1.1). Zenodo. <https://doi.org/10.5281/zenodo.4623713>

⁹ Caldoni G, Iacovella V, Lazzeri E, Lehtsalu L, Marchegiani F, Paschetta M, Pasquale V, Pedonese G, Tarallo A. Mapping data stewardship in Italy: Findings from the first national survey. Open Res Eur. 2026 Feb 13;6:45. doi: 10.12688/openreseurope.22183.1. PMID: 41816254; PMCID: PMC12972714.

¹⁰ Whyte, A., Green, D., Evangelinou, B., Gingold, A., Di Giorgio, S., Koteska, B., Linés, C., Rauste, P., Sokolova, E., Sowinski, C., Torres Ramos, G., van Leersum, N., & Lines Diaz, C. (2025). D2.6 Catalogue of OS career profiles and MVS - update. Zenodo. <https://doi.org/10.5281/zenodo.15731870>

¹¹ Isabelle Blanc, Anne Laurent. Données de recherche : caractérisation des fonctions clés liées à leur gestion, leur exploitation, leur diffusion et leur utilisation. Ministère de l'enseignement supérieur et de la recherche. 2025. (hal-05265596)

Impact SO 2026¹², qui ont été consacrées à l'impact des politiques en faveur de la science ouverte sur les pratiques et les résultats scientifiques. Elles se sont tenues à Nancy du 27 au 29 janvier 2026. Il a été fait le choix d'un atelier participatif pour évaluer les similitudes et potentielles différences entre les profils dressés par les travaux pré-identifiés et ceux effectivement existants sur le terrain. L'atelier avait également pour objectifs de faire émerger les différentes missions existantes et les compétences liées, ainsi que de potentiels besoins, notamment en formations, ou difficultés. L'objectif n'était pas de produire un référentiel normatif mais de faire émerger une vision collective issue d'échanges entre professionnel-le-s de la donnée.

Cet article présente la démarche mise en œuvre lors de cet atelier participatif, les principaux enseignements issus des échanges entre participant-e-s et les perspectives qu'ils ouvrent pour mieux comprendre les évolutions des métiers de la gestion des données de la recherche et mieux répondre aux besoins identifiés sur le terrain.

2. Méthodologie

2.1. Objectifs et principes directeurs

L'atelier visait à analyser les écarts entre les cadres théoriques décrivant les métiers de la gestion des données de la recherche et les réalités professionnelles observées sur le terrain.

Il reposait sur trois principes méthodologiques :

- **Approche participative et collaborative** : valorisation des retours d'expérience des participant-e-s.
- **Ancrage dans des travaux récents** : mobilisation d'enquêtes et livrables nationaux et européens relatifs aux métiers de la donnée.
- **Orientation vers l'action** : production de pistes concrètes en matière de reconnaissance, structuration et montée en compétences.

2.2. Phase préparatoire

2.2.1. Sélection et analyse des ressources

Les animatrices ont sélectionné plusieurs livrables récents traitant du sujet des compétences et métiers en gestion des données :

- MESR, Données de recherche : caractérisation des fonctions clés liées à leur gestion, leur exploitation, leur diffusion et leur utilisation¹³
- Skills4EOSC, D2.6 Catalogue of OS career profiles and MVS – update¹⁴

¹² <https://impactso2026.sciencesconf.org/>

¹³ Isabelle Blanc, Anne Laurent. Données de recherche : caractérisation des fonctions clés liées à leur gestion, leur exploitation, leur diffusion et leur utilisation. Ministère de l'enseignement supérieur et de la recherche. 2025. (hal-05265596)

¹⁴ Whyte, A., Green, D., Evangelinou, B., Gingold, A., Di Giorgio, S., Koteska, B., Linés, C., Rauste, P., Sokolova, E., Sowinski, C., Torres Ramos, G., van Leersum, N., & Lines Diaz, C. (2025). D2.6 Catalogue of OS career profiles and MVS - update. Zenodo. <https://doi.org/10.5281/zenodo.15731870>

Après analyse des 14 profils MVS (Minimum Viable Skills) du catalogue Skills4EOSC, et compte tenu des contraintes organisationnelles de l'atelier (durée limitée à 1 h 30), il a été décidé de retenir les 7 profils les plus directement liés à la gestion opérationnelle des données et d'exclure les profils liés au pilotage. Ces 7 profils correspondaient d'autant plus aux fonctions occupées par les participant·e·s. Ces profils ont été traduits en français. Les activités associées ont également été synthétisées et simplifiées, dans un souci d'efficacité pédagogique.

Les profils Skills4EOSC utilisés pour l'atelier sont les suivants :

1. Data Steward coordinateur·rice / Data librarian (Data Steward Coordinator)
2. Data Steward intégré·e (Data Steward Embedded)
3. Curateur·rice de collections numériques (Digital Collections Curator)
4. Conseiller·ère éthique (Ethics Advisor)
5. Expert·e légal (Legal Expert)
6. Chercheur·e - début de carrière (Early Career Researcher)
7. Chercheur·e - senior (Senior Researcher)

En complément, 7 fonctions clés liées à la gestion des données présentées dans le rapport du MESR ont été sélectionnées. Dans un souci d'harmonisation, ces fonctions ont été transposées sous la forme de profils métier. Les fonctions (profils) de l'étude du MESR utilisées pour l'atelier sont les suivantes :

1. Intendant·e des données (Intendance des données)
2. Curateur·rice de données (Curation des données ou research data curation)
3. Ingénieur·e des données (Ingénierie des données)
4. Analyste des données (Analyse des données)
5. Scientifique des données (Science des données)
6. Documentaliste des données (Documentation des données)
7. Curateur·rice des métadonnées documentaires (Curation des métadonnées documentaires)

Par ailleurs, deux autres livrables ont été pris en compte dans cette phase préparatoire afin d'alimenter la réflexion :

- L'enquête sur l'implication des structures documentaires dans l'accompagnement sur les données de recherche de l'ADBU¹⁵ a permis d'élargir le périmètre de réflexion aux services en gestion de donnée mis en place au sein de structures documentaires.
- L'étude sur les « acteurs de l'IST » dans l'enseignement supérieur et la recherche menée par le Réseau RENATIS¹⁶, et plus particulièrement les travaux en cours pour l'analyse des

¹⁵ ADBU (2025), Enquête sur l'implication des structures documentaires dans l'accompagnement sur les données de recherche <https://hal.science/view/index/docid/5221870>

¹⁶ Réseau RENATIS, Étude sur les « acteurs de l'IST » dans l'enseignement supérieur et la recherche

besoins en formation des répondant·es, a permis d'obtenir un certain nombre d'éléments de réflexions concernant ce sujet.

2.2.2. Préparation des matériaux pédagogiques

Les 14 profils métiers et activités dédiées ont été transformés en **14 fiches métier** (disponibles en annexes) destinées à être imprimées et affichées dans la salle de formation. Chaque fiche profil s'est vu attribuer une lettre de A à N.



Code	Activité
D-13	Analyses statistiques des données
D-14	Visualisation et communication des résultats
D-15	Gestion des données et veille technologique
D-16	Accompagnement des équipes dans l'utilisation des outils d'analyse et l'interprétation des données

Figure 1 : Exemple de fiche profil : analyste des données. La lettre de ce profil est D. A ce profil sont liées 4 activités, numérotées de 13 à 16 (D-13 à D-16).

Chaque profil a été subdivisé en activités répertoriées selon la lettre du profil et un numéro (A1, A2, B1, B2...) (figure 1).

L'ensemble des activités de tous les profils ont été listés de façon aléatoire sur une fiche (disponible en annexe), destinée à être utilisée individuellement (figure 2). Les codes des fiches profils et activités n'étaient pas connus des participant·e·s.

Chacun des participant·e· était ensuite amené à cocher sur sa fiche personnelle les activités qu'il/elle exerçait.

Atelier

Profils et activités liés à la gestion des données de la recherche

Profil de poste du répondant :

Activités	Code	Je fais	Je ne fais pas
Coordination du cycle de vie des données (mobilisation des différents acteurs pour assurer un enchaînement fluide entre les phases)	A-1	☐	☐
Orientation des chercheurs vers les installations et services liés à la gestion des données de recherche	H-30	☐	☐
Promotion et soutien du libre accès et des principes juridiques et éthiques liés à la science ouverte	L-53	☐	☐
Soutien et conseil aux chercheurs sur les meilleures pratiques en matière de gestion des données (éthique, conformité)	I-39	☐	☐
Renseignement, mise à jour et contrôle des métadonnées documentaires	G-24	☐	☐
Établissement d'un réseau de collaboration	N-71	☐	☐
Promotion et soutien de la science ouverte Adhésion à ses initiatives	M-59	☐	☐
Contribution à la mise en œuvre des politiques institutionnelles, des outils et des services destinés à la gestion des données de recherche issues des collections numériques muséales, selon les principes FAIR et CARE	J-41	☐	☐

Figure 2 : Extrait de la fiche destinée à être remplie par les participant·e·s à l'atelier. Les activités sont réparties aléatoirement dans la fiche et les participant·e·s déterminent pour chacune d'elles s'ils l'exercent ou non.

2.3. Déroulé méthodologique de l'atelier

Le scénario pédagogique a été structuré en trois séquences :

- Une séquence d'introduction visant à présenter les objectifs et le déroulé de l'atelier ;
- Une séquence de travail sur les activités ;
- Une séquence de restitution en plénière.

Durant la séquence de travail sur les activités, les participant·e·s étaient réparti·es en sous-groupes de cinq à six personnes. Dans un premier temps, chaque participant·e renseignait une fiche individuelle recensant les différentes activités proposées, en cochant celles qu'il/elle exerçait dans le cadre de ses fonctions.

Dans un second temps, les membres de chaque sous-groupe échangeaient autour de leurs perceptions des activités, de leurs pratiques professionnelles, des réalités de terrain auxquelles ils étaient confrontés ainsi que des difficultés rencontrées. Ces discussions constituaient un objectif central de l'atelier, en favorisant le partage d'expériences et la mise en perspective des pratiques.

Enfin, les participant·e·s étaient invité·es à consulter les profils métiers affichés dans la salle. En comparant leurs fiches individuelles aux référentiels présentés, ils/elles pouvaient identifier le ou les profils auxquels leur activité correspondait le mieux. Cette étape permettait également de questionner l'adéquation entre les profils théoriques et la réalité des missions exercées.

2.4. Modalités d'exploitation des résultats

Lors de l'atelier, chaque participant·e a été invité·e à partager ses propres résultats. Les animatrices ont alors globalement comptabilisé les profils sélectionnés.

A l'issue de l'atelier, toutes les fiches remplies par les participant·e·s des 2 sessions ont été compilées dans un tableau : n° fiche / activités regroupées par profils (figure 3).

Source	Profil	Fonction	1	2	3	5	6	7	8
MESRI	Intendant·e des données (A1/A2/A3/A4/A5)	A1			1	1	1		
MESRI	Intendant·e des données (A1/A2/A3/A4/A5)	A2		1	1		1		
MESRI	Intendant·e des données (A1/A2/A3/A4/A5)	A3					1		
MESRI	Intendant·e des données (A1/A2/A3/A4/A5)	A4				1			
MESRI	Intendant·e des données (A1/A2/A3/A4/A5)	A5		1	1		1		
MESRI	Curateur·rice de données (B)	B6							
MESRI	Curateur·rice de données (B)	B7					1		
MESRI	Curateur·rice de données (B)	B8		1			1		
MESRI	Ingénieur·e des données (C9/C10/C11/C12)	C9			1				
MESRI	Ingénieur·e des données (C9/C10/C11/C12)	C10							
MESRI	Ingénieur·e des données (C9/C10/C11/C12)	C11							
MESRI	Ingénieur·e des données (C9/C10/C11/C12)	C12							
MESRI	Analyste des données (D13/D14/D15/D16)	D13							
MESRI	Analyste des données (D13/D14/D15/D16)	D14					1		
MESRI	Analyste des données (D13/D14/D15/D16)	D15					1	1	
MESRI	Analyste des données (D13/D14/D15/D16)	D16							
MESRI	Scientifique des données (E17/E18/E19/E20)	E17							
MESRI	Scientifique des données (E17/E18/E19/E20)	E18							
MESRI	Scientifique des données (E17/E18/E19/E20)	E19							
MESRI	Scientifique des données (E17/E18/E19/E20)	E20							
MESRI	Documentaliste des données (F21/F22)	F21		1	1		1	1	
MESRI	Documentaliste des données (F21/F22)	F22	1	1	1		1	1	

Figure 3 : Extrait du tableau regroupant les données anonymisées. Chaque ligne représente une activité d'un profil (activité A1, A2...). Les colonnes numérotées représentent chacune un participant·e.

Différentes analyses croisées (activités cochées/activités proposées/regroupement des activités par profil) ont été directement réalisées dans le tableau. Les graphiques qui en résultent sont intégrés ci-après.

Deux traitements supplémentaires ont été apportés. Tout d'abord, afin de se délester du biais lié au nombre différent d'activités listées par profil (de 2 à 8), il a été fait un ratio entre le nombre de fois où les activités de chaque profil ont été cochées et le nombre d'activités proposées dans chaque profil. Par exemple, les activités du profil « scientifique des données » ont été cochées 11 fois. Il y a 4 activités dans ce profil. Le ratio est de $11/4 = 2,8$. Ensuite et indépendamment, il a été choisi de ne retenir dans le tableau, pour chacune des fiches et par profil, que les données pour lesquelles le total des activités sélectionnées par le/la participant·e est égal au moins à la moitié des activités du profil. Par exemple, le profil « Intendant·e des données » regroupe 5 activités. Si le/la participant·e n'en a coché que 2, sa fonction n'est pas alors prise en compte.

3. Résultats et analyse

Les deux groupes ont été subdivisés en sous-groupes de 5 personnes maximum. Les profils présents étaient diversifiés : les personnels d'appui à la recherche étaient majoritaires, avec une représentation non négligeable de chercheur·e·s.

3.1. Profils et reconnaissance des métiers

Concernant l'alignement des compétences des participant·e·s avec les profils présents, ils se sont peu reconnus dans les profils pré-établis (ex. : data stewards, data librarians, chercheur·e·s). Leur très grande majorité a des missions éclatées et transversales, ce qui rend difficile l'identification à un seul profil. Certains profils sont particulièrement représentés dans les résultats des participant·e·s : c'est le cas notamment des profils de data steward coordinateur·rice et de chercheur·e·s, issus des profils métiers créés dans le cadre du projet européen Skills4EOSC. Plusieurs groupes soulignent que les activités se recoupent entre profils. Par exemple, les data stewards et chercheur·e·s font des tâches similaires en gestion de la donnée, accompagnement et promotion d'une bonne gestion des données sans toujours le savoir, ce qui explique que des personnels d'appui à la recherche se retrouvent dans le profil de chercheur·e·s. Le profil d'expert·e légal, également issu de Skills4EOSC, a également été représenté dans les résultats malgré l'absence de participant·e·s spécialistes légaux. Ceci est à nouveau expliqué par la multiplicité des activités effectuées en pratique par les personnels qui traitent les données ou accompagnent à le faire (« il faut que quelqu'un le fasse donc on le fait »). Ce sont des postes souvent occupés par défaut, faute de mieux.

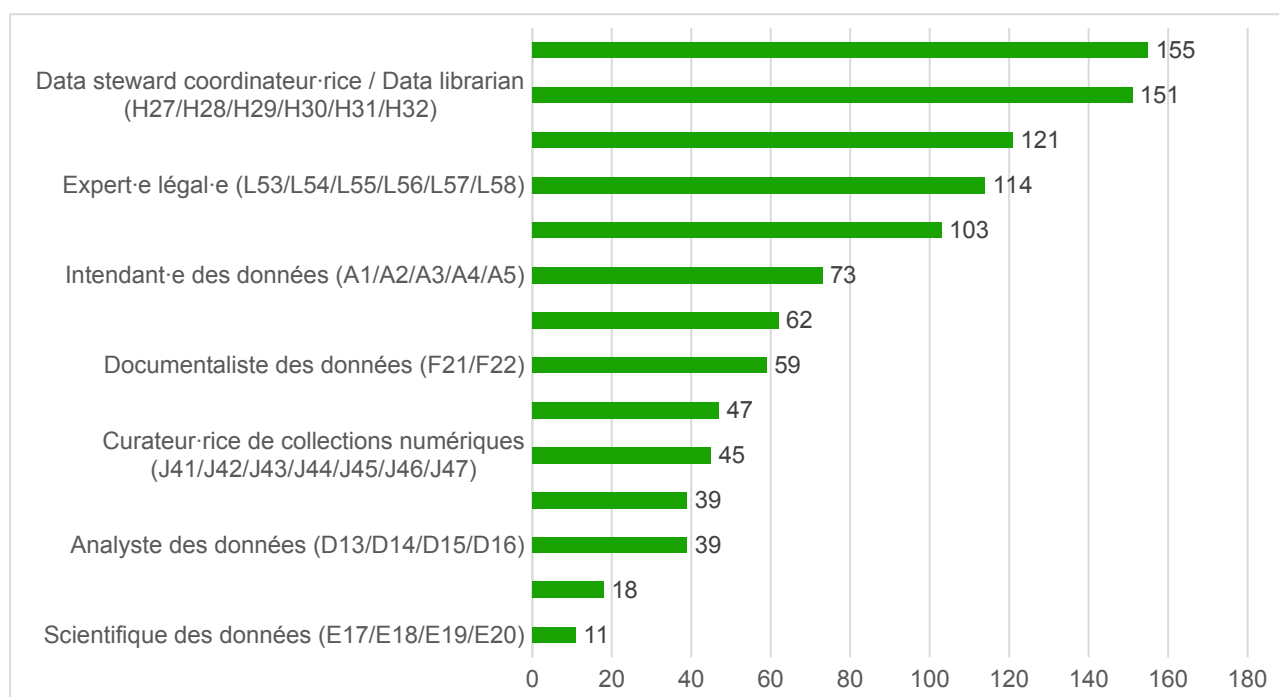


Figure 4 : Histogramme représentant le nombre d'activités cochées par profil. Les activités sont représentées sous le format LN, où L représente la lettre correspondant au profil, et N le chiffre correspondant à l'activité nommée.

En termes de nombre d'activités exercées, ce sont les activités des profils de data steward intégré·e et data steward coordinateur·rice / data librarian (tous les deux Skills4EOSC) qui sont les plus représentées (figure 4). Ces premiers résultats sont à prendre avec précautions car il s'agit de valeurs absolues. Ce sont tous les deux des profils avec un nombre importants d'activités : respectivement 8 et 6. Ils font ainsi partie des profils avec le plus d'activités.

Afin de tenir compte de cette différence en termes de nombre d'activités par profil, le ratio entre nombre d'activités cochées par les participant·e·s et nombre d'activité par profil a été fait (figure 5). Les profils les plus représentés sont alors documentaliste des données (MESR), data steward coordinateur·rice / data librarian (Skills4EOSC), data steward intégré·e (Skills4EOSC) et expert·e légal (Skills4EOSC).

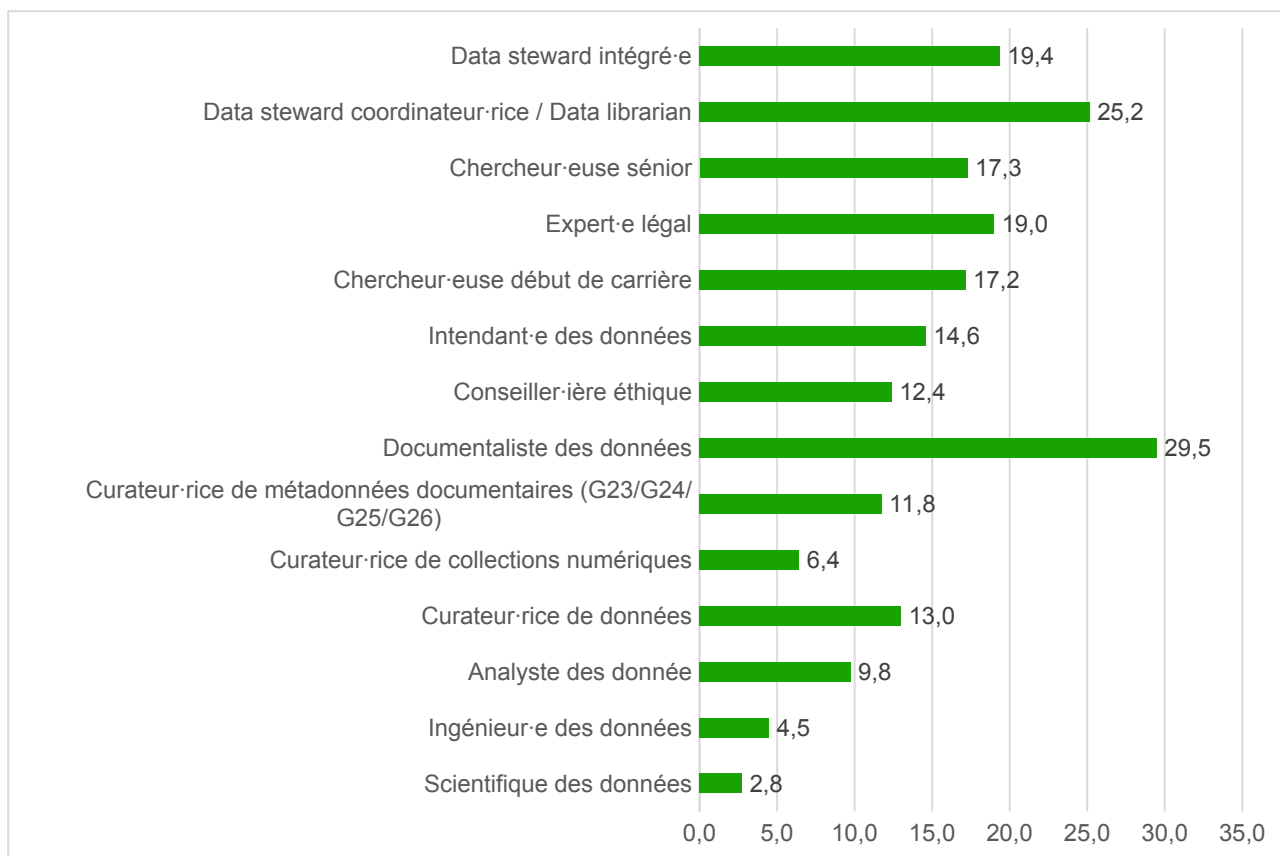


Figure 5 : Histogramme représentant les ratios nombre d'activités cochées par profil par rapport au nombre d'activités.

En tenant uniquement compte des profils sélectionnés à au moins 50% de leurs activités par les participant·e·s, l'ordre évolue également (figure 11). La fonction principale de documentaliste des données (MESR) arrive bien en première position devant le profil de data steward coordinateur·rice / data librarian (Skills4EOSC). Le classement des profils les plus représentés se rapproche de celui obtenu avec normalisation en fonction du nombre d'activités (figure 5), notamment avec le profil de documentaliste des données (MESR) qui arrive en tête. Les deux profils de data stewards (coordinateur·rice et intégré·e, tous les deux de Skills4EOSC), ainsi que les profils d'expert·e légal et de chercheur·e·s en début de carrière et senior (tous trois de Skills4EOSC) restent très représentés.



Figure 6 : Histogramme représentant le nombre de profils dont les activités ont été choisies à au moins 50 % parmi les participant·e·s.

Dans le détail des fonctions (figure 2), les activités « Veille sur le paysage réglementaire et les politiques nationales concernant la science ouverte » (L54), « Promotion et communication sur l'importance de la science ouverte et des principes FAIR » (H28), « Promotion et soutien du libre accès et des principes juridiques et éthiques liés à la science ouverte » (L53), « Orientation des chercheur·e·s vers les installations et services liés à la gestion des données de recherche » (H30) sont exercées par la quasi-totalité des participant·e·s. Ces activités sont transversales et reflètent bien les missions premières des personnels d'appui à la recherche impliqués dans la Science ouverte la gestion des données de la recherche.

À l'inverse, les activités « Aide aux chercheurs et aux utilisateurs des collections muséales pour appliquer les meilleures pratiques en matière de gestion des données et des objets numériques dans les domaines concernés » (J43) et « Industrialisation et mise en production des traitements et modèles » (C12) ne sont exercées par aucun des participant·e·s. Ces activités correspondent à des personnels précis, travaillant respectivement dans les musées ou dans un service informatique.

Une étude plus détaillée serait appropriée mais est dépendante des établissements et/ou services d'appartenance des participant·e·s présents.

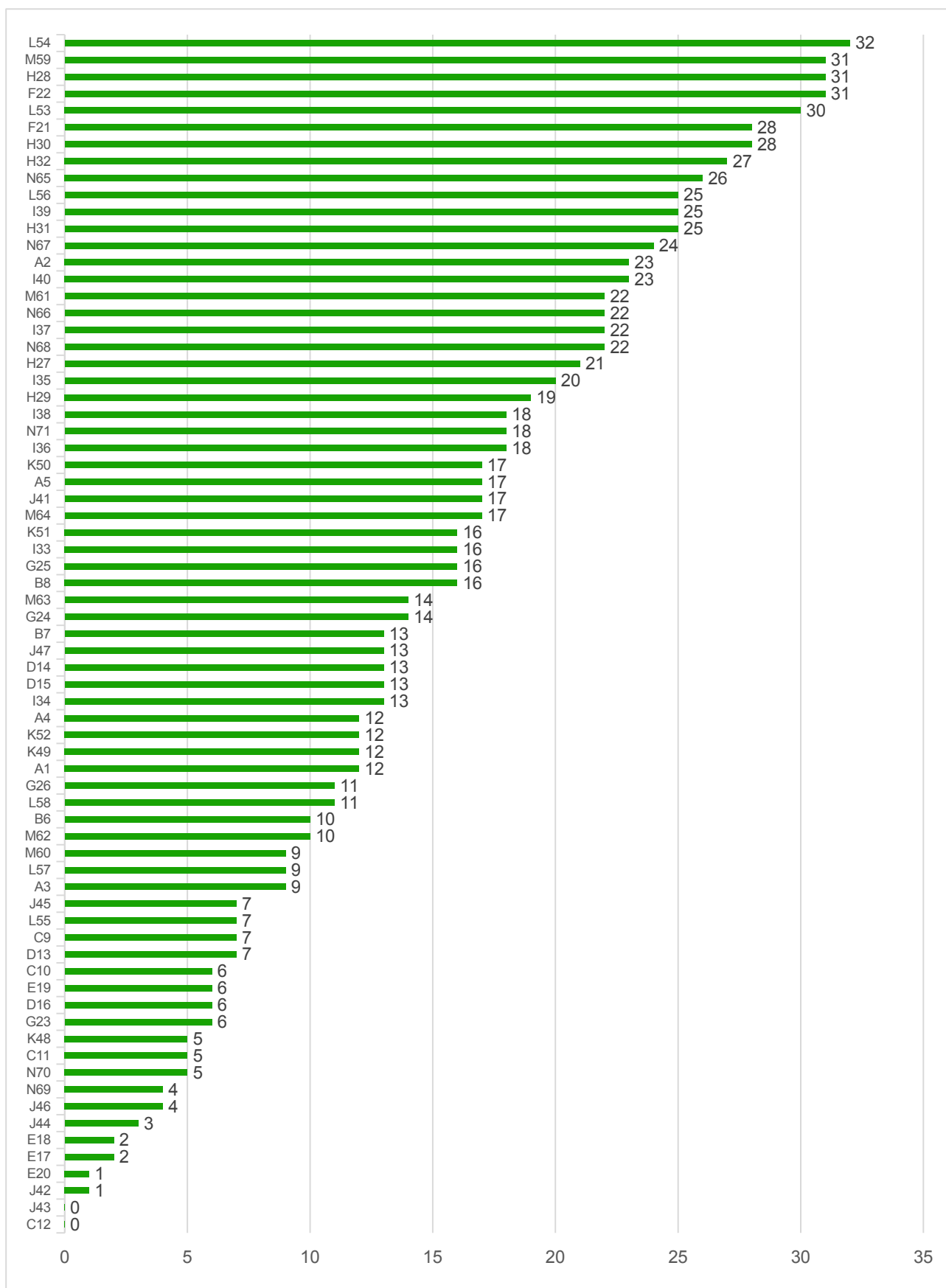


Figure 7 : Histogramme représentant la répartition des activités sélectionnées par les participant·e·s, classées par ordre décroissant.

3.2. Besoins en compétences et formations

À travers les activités effectives et en faisant un alignement avec des profils déterminés, un second objectif de ces ateliers était de faire ressortir les potentiels besoins en formations.

Un besoin de formations juridiques et éthiques a été unanimement exprimé. Les accompagnements légaux offerts par des personnels non spécialisés sont à un premier niveau, nécessaire mais non suffisant. Les appuis légaux des institutions sont rarement spécialisés dans les données de la recherche et la plupart des établissements manque de personnes-relais pour ces questions. Il existe donc un gap qu'une formation plus poussée des personnels d'accompagnement pourrait offrir. Dans le même sens, le besoin que les juristes soient eux-mêmes formés aux enjeux de la donnée a été exprimé. Il a également été exprimé un manque de lien, à travers une personne relai, entre les comités d'éthique et les collectes de données.

Avec l'émergence continue de nouveaux outils et l'arrivée de l'intelligence artificielle, un besoin de formations a été exprimé.

Un besoin en formations de formateurs a également été relevé : il faut pouvoir former les personnels à la transmission de compétences.

3.3. Enjeux transversaux

Les participant·e·s cumulent de 5 à 14 rôles différents, ce qui nuit à leur efficacité et à la clarté de leur poste. Avec la surreprésentation de multitudes d'activités éclatées dans différents profils, les participant·e·s ne se sentent ni experts ni généralistes. Il est suggéré de mieux répartir les missions et de créer des liens entre les acteur·rice·s.

Une différence dans la culture managériale a été relevée au cours des échanges : les bibliothèques ont un cadre plus rigide que les laboratoires et les attentes sont variées en termes de coordination et de management opérationnel. L'absence d'activités strictement liées au pilotage ou au management dans les profils exposés a également été relevée.

Un point essentiel discuté a également été la nécessité de sensibiliser les établissements à l'utilité des plans de gestion de données et à l'importance des métiers de la donnée, au-delà des obligations réglementaires.

4. Discussion

Globalement, les profils métiers proposés par Skills4EOSC et par l'étude du Ministère de l'Enseignement Supérieur et de la Recherche constituent des modèles de référence utiles pour décrire les compétences attendues, mais ils ne reflètent qu'imparfaitement la réalité des établissements. Les professionnel·le·s rencontré·e·s occupent des fonctions hybrides, construites progressivement en réponse aux besoins locaux davantage qu'à partir de fiches de postes stabilisées. Cette situation traduit une phase de transition entre l'émergence de nouveaux métiers et leur institutionnalisation.

La transversalité apparaît comme une caractéristique constitutive des métiers de la donnée en France. Si elle favorise les collaborations entre communautés professionnelles, elle conduit

également à une dilution des responsabilités et à un sentiment de ne maîtriser qu'imparfaitement un ensemble très vaste de compétences. Les professionnel·le·s occupent ainsi une position intermédiaire, ni spécialistes d'un domaine unique, ni généralistes pleinement reconnus.

Les profils ainsi utilisés dans l'atelier constituent des repères mais ne reflètent que partiellement les réalités professionnelles. La polyvalence des professionnel·le·s est aujourd'hui constitutive des métiers mais atteint ses limites : les besoins identifiés concernent autant l'organisation des établissements que la formation des individus sur des points émergents et / ou spécialisés, comme l'intelligence artificielle ou les expertises éthique ou juridique. Les référentiels français (MESR, CNRS et autres établissements) ou européens (Skills4EOSC, EOSC, ELIXIR) peuvent servir d'outils d'accompagnement de cette professionnalisation plutôt que de modèles prescriptifs. Les résultats de cet atelier peuvent aussi alimenter une réflexion sur l'évolution de ces référentiels.

La discussion qu'a permis cet atelier montre que les métiers de la gestion des données de la recherche ne sont pas seulement des métiers émergents : ce sont encore des métiers en construction. En effet, les professionnel·le·s construisent leurs périmètres d'activité au quotidien, en fonction des besoins de leurs laboratoires, des compétences disponibles et des ressources institutionnelles. Les référentiels français et européens décrivent un horizon de professionnalisation, tandis que les pratiques observées dans le paysage français témoignent d'une phase de transition où les frontières entre métiers restent poreuses. Cette tension entre modèles théoriques et réalités de terrain paraissent être l'un des principaux enseignements de cet atelier et constitue une contribution à la réflexion sur la structuration des métiers de la gestion des données de la recherche.

Par ailleurs, les participant·e·s aux deux sessions de cet atelier ont fait part d'une appréciation positive de l'atelier et de son format ludique et interactif, tout en soulignant sa pertinence et son côté engageant. Plusieurs d'entre eux ont d'ailleurs également suggéré de renouveler l'atelier dans d'autres contextes afin d'enrichir les retours d'expérience, et ont demandé de disposer des supports ici en annexes.

5. Conclusion et points de vigilance

La création de profils métiers structurés est une étape essentielle à la reconnaissance de ces métiers nouveaux. Cependant, cet atelier joué à deux reprises permet de mettre en lumière la multiplicité des activités encore concrètement expérimentée. Les besoins du terrain engendrent des métiers disparates très, voire trop, diversifiés et construits « à la carte » par celles et ceux qui les pratiquent. Ces métiers nouveaux mériteraient à être complétés, notamment par de réels spécialistes juridiques et éthiques en gestion des données de la recherche. Ces états confirment les résultats observés lors de l'enquête RENATIS sur les acteurs et actrices de l'IST¹⁷ : diversification des activités, parfois par défaut ou bien encore exercice de ces activités par des non ou « peu » professionnels de la donnée.

Le développement de formations ciblées, telles que celles du projet Skills4EOSC¹⁸, qui permettra

¹⁷ Réseau RENATIS, Étude sur les « acteurs de l'IST » dans l'enseignement supérieur et la recherche
<https://cnrs.hal.science/hal-05128131v1>

¹⁸ <https://www.skills4eosc.eu/participate/skills4eosc-training-courses>

une montée en compétences des profils pourra aiguïser les réponses aux besoins remontés du terrain : besoins juridiques, éthiques, besoins de compétences en outils techniques, et besoins en formations de formateurs.

La transversalité est également un atout à favoriser : il est souhaitable d'encourager les échanges entre laboratoires et bibliothèques pour harmoniser les pratiques. Les professionnel·le·s compensent aujourd'hui les manquent organisationnels des établissements en développant eux-mêmes des compétences multiples.

En mettant en œuvre ces pistes d'action, il est également préconisé d'éviter de créer des profils trop rigides ou trop éclatés. Les profils proposés pendant cet atelier sont une excellente base pour une adaptation aux besoins du terrain, mais également aux spécificités locales : il faut prendre en compte les particularités des laboratoires et des établissements, mais également du paysage français. Les profils proposés dans le cadre du projet Skills4EOSC sont des profils européens : ils pourraient être adaptés au terrain français et à ses spécificités.

Toutes ces pistes d'action, qui font partie intégrante de l'évolution des profils utilisés dans cet atelier, doivent être discutées dans chaque contexte institutionnel. Les établissements doivent avoir les moyens humains et financiers pour mettre en œuvre ces changements. La valorisation de ces rôles et métiers reste également essentielle dans leur évolution (recrutement, progression de carrière).

Cet atelier a été pensé comme le premier d'une réflexion collective et progressive. Les annexes mises à disposition de cette publication offrent à chacun la possibilité de le rejouer et de l'adapter pour compléter ces premiers résultats et la discussion qui en résulte.