

couperin.org

Consortium Unifié des Établissements Universitaires et
de Recherche pour l'Accès aux Publications Numériques

Enquête sur l'appui à la gestion des données de la recherche en service de documentation et d'information scientifique et technique

Rapport produit par le Groupe Données du GTSO-Couperin

Juillet 2021

Version 1

DOI : [10.5281/zenodo.5078504](https://doi.org/10.5281/zenodo.5078504)

À propos des auteurs

Martine Augouvernaire est responsable du signalement et du contrôle qualité des données du SCD de l'Université de Tours. Elle participe au groupe de travail interne sur la gestion des données de la recherche et accompagne doctorants et enseignants-chercheurs sur les questions de Science Ouverte.

Laetitia Bracco (n°ORCID : [0000-0002-2939-9110](https://orcid.org/0000-0002-2939-9110)) est *data librarian* à l'Université de Lorraine. Elle accompagne doctorants et enseignants-chercheurs dans la gestion des données de la recherche et anime l'offre de services bibliométriques. Elle co-pilote le GTSO-Données de Couperin.

Romain Féret (n°ORCID : [0000-0002-1527-1482](https://orcid.org/0000-0002-1527-1482)) est directeur du CRFCB Média Normandie, à la ComUE Normandie Université. Il y propose notamment des formations à la science ouverte à destination de l'ensemble des personnels qui accompagnent les chercheurs.

Élise Lehoux (n°ORCID : [0000-0003-2929-6768](https://orcid.org/0000-0003-2929-6768)) est cheffe de projet Information Scientifique et Technique et référente pour les données de la recherche à Université de Paris. Elle accompagne doctorants et enseignants-chercheurs à la gestion des données de la recherche et participe aux projets dédiés à la Science Ouverte portés par le service des bibliothèques de l'université.

Avec la participation de **Maxence Larrieu** (Université de Paris).

Table des matières

Introduction.....	3
1. Méthodologie et analyse des établissements ayant répondu à l'enquête.....	4
2. Section 1 de l'enquête : Organisation du service.....	17
3. Section 2 de l'enquête : Signalement et archivage des données de recherche.....	26
4. Section 3 de l'enquête : Accompagnement à la rédaction des plans de gestion de données.....	28
5. Section 4 de l'enquête : Sensibilisation et formation.....	35
6. Section 5 de l'enquête : Services complémentaires.....	39
7. Section 6 de l'enquête : Publics.....	42
8. Section 7 de l'enquête : Indicateurs.....	43
9. Section 8 de l'enquête : Collaborations internes ou externes – Q34 et 35.....	48
10. Section 9 de l'enquête : Projets et besoins – Q36 à 39.....	52
Conclusion.....	54
Annexe – FAQ accompagnant l'enquête.....	59

Introduction

Le Groupe Données du GTSO (Groupe de Travail Science Ouverte) de Couperin est à l'initiative de cette enquête sur les services d'accompagnement à la gestion des données de la recherche, lancée en septembre 2020. Travaillant sur les modalités d'accompagnement que les services de documentation et d'information scientifique et technique présentent aux personnels de recherche en matière de données, le GTSO souhaitait disposer d'un état des lieux des services offerts. Cette enquête pourra être reproduite dans le temps afin de mesurer l'évolution de l'accompagnement proposé par ces services.

L'objectif était aussi de disposer d'un panorama le plus complet possible de ce que ces services proposent en matière de gestion des données, de leurs difficultés et de leurs projets pour les prochaines années, afin d'orienter la feuille de route et les réalisations du GTSO-Données. La communication autour de cette enquête a fortement insisté sur la nécessité que les services non positionnés sur ce sujet, ou démarrant à peine, y répondent aussi afin de disposer de résultats représentatifs et d'éviter un effet de halo sur les établissements les plus avancés. L'enquête porte sur l'ensemble des facettes de la gestion des données, à l'exception des services de fouille et de visualisation de textes et de données.

Cet état des lieux a paru d'autant plus nécessaire qu'aucun établissement français n'était représenté dans l'enquête menée par Tenopir sur l'appui à la gestion des données en 2017¹. Par ailleurs, des enquêtes du même type ont été réalisées à l'échelle d'un pays, notamment aux Etats-Unis (Murray 2019, Gowen 2020²) ou en Afrique du Sud (Chiwari 2018³), mais pas en France. Nous nous sommes inspirés de cette enquête afin de disposer d'éléments de comparaison et de ne pas négliger d'aspects de la gestion des données. Cette enquête montre que nombre de services était encore dans une phase de démarrage de leur service, notamment à travers le développement de politiques institutionnelles. Cette enquête insiste sur

¹ Tenopir, Carol, et al. "Research Data Services in European Academic Research Libraries." *LIBER QUARTERLY*, vol. 27, no. 1, 2017, pp. 23–44, doi:10.18352/lq.10180.

² Murray, Matthew, et al. "A Survey of the State of Research Data Services in 35 U.S. Academic Libraries, or 'Wow, What a Sweeping Question.'" *Research Ideas and Outcomes*, vol. 5, Dec. 2019, p. e48809, doi:10.3897/rio.5.e48809 et Gowen, Elise, and John J. Meier. "Research Data Management Services and Strategic Planning in Libraries Today: A Longitudinal Study." *Journal of Librarianship and Scholarly Communication*, vol. 8, no. 1, Apr. 2020, p. 2336, doi:10.7710/2162-3309.2336.

³ Chiwari, Elisha R.T. and Becker, Deborah Anne. "Research Data Management Services in Southern Africa: A Readiness Survey of Academic and Research Libraries." *African Journal of Library Archives and Information Science*, vol. 28, no. 1, Apr. 2018, pp.1-16.

les collaborations internes et externes mises en place sur ce sujet. Elle différencie aussi les services de sensibilisation (renseignement, formation) d'une part, et l'accompagnement technique d'autre part en montrant que les services de documentation ayant répondu à l'enquête se positionnaient nettement plus sur les premiers que sur les seconds en 2016.

Lors de la préparation de l'enquête, neuf hypothèses principales ont été identifiées pour guider l'analyse des résultats :

1. Une minorité non négligeable de services de documentation n'accompagnent pas du tout la gestion des données et n'ont pas de projets en la matière ;
2. Une majorité d'établissements sont au début de leur accompagnement en matière de GDR ou encore en phase de conception ;
3. Il existe un écart important entre les services d'appui à la gestion des données offerts par les organismes de recherche et les universités ;
4. Il existe un écart important entre les services offerts par les grandes universités et les petites et moyennes universités ;
5. Seule une minorité d'universités est en capacité de dégager des quotités de travail mesurables et significatives dédiées à la gestion des données ;
6. La situation est aussi très hétérogène entre les services de documentation des organismes de recherche, mais dans une moindre mesure ;
7. Les services offerts sont plutôt de type sensibilisation que de nature technique ;
8. Il y a une corrélation entre l'existence d'une politique institutionnelle en matière de GDR et les services offerts par les services documentaires ;
9. La gestion des données nécessite de nombreuses formes de coopérations internes et externes.

1. Méthodologie et analyse des établissements ayant répondu à l'enquête

1.1 Méthodologie et mise en œuvre de l'enquête

Questionnaire et administration de l'enquête

L'enquête est structurée en dix sections qui comportent chacune 2 à 6 questions, pour un total de 45 questions. Les questions 1 à 4 ont pour objet de disposer d'informations sur la structuration et les moyens du service de documentation répondant. Elles sont complétées par les questions 40 à 43 qui permettent

de connaître l'environnement dans lequel évoluent ces services et de s'assurer qu'il n'y ait pas de doublon dans les réponses obtenues. La question 44 a permis de recueillir un contact afin de consolider certaines données. La question 45 porte sur les modalités de confidentialité souhaitées par les répondants.

Le corps de l'enquête est donc composé des questions 5 à 39. Les questions 5 à 13 portent sur les moyens disponibles pour la mise en œuvre des services en matière de données et le positionnement des services de documentation au sein de leur établissement. Les questions 14 à 30 relèvent des services proposés par les services de documentation. Elles sont complétées par les questions 31 à 33 qui ont pour objet de mesurer les services effectivement rendus de manière quantitative. Les questions 34 et 35 sont relatives aux collaborations entretenues par les services de documentation dans leur établissement ou à l'extérieur. Enfin, les questions 36 à 39 sont des questions ouvertes qui interrogeaient les besoins, difficultés et projets des services de documentation dans la mise en œuvre d'un accompagnement à la gestion des données.

L'enquête a été conçue entre juin et septembre 2020. Elle a été ouverte le mardi 8 septembre et close le mardi 6 octobre. La communication sur l'enquête s'est faite par l'envoi de courriels sur des listes de diffusion professionnelles : gtso, agora, correspondants ao, datalibrarians, adbu-forum et sur Twitter. L'enquête au format pdf et une notice explicative sur l'enquête accompagnait ces courriels. L'enquête était à remplir en ligne sur l'outil Survey Monkey.

Traitement et protection des données

Différents traitements ont été appliqués aux données recueillies afin de s'assurer de leur fiabilité. L'ensemble de ces traitements sont précisés dans le jeu de données qui accompagne l'enquête. Une vérification sur les doublons a permis d'écarter 4 réponses sur les 93 obtenues. Par ailleurs, 7 réponses ont été obtenues de services sans fonction de documentation. Ces réponses ont fait l'objet d'une diffusion séparée et n'ont pas été analysées dans l'enquête. Cela a permis de travailler sur 82 réponses.

Les informations sur les établissements ont permis de les placer dans la catégorie qui leur correspondait en cas d'erreur des répondants. Les réponses à un certain nombre de questions ont fait l'objet d'une vérification ligne à ligne, parfois complétée par un échange avec le répondant pour confirmer ou corriger les informations recueillies. Enfin, la typologie d'établissements utilisée par Couperin a été partiellement modifiée pour affiner l'analyse par type d'établissements.

Les réponses aux questions 44 et 45 ont été supprimées avant diffusion. Les verbatims ont fait l'objet d'une vérification pour s'assurer que les informations données étaient pertinentes pour l'enquête et ne contenaient pas de données à caractère personnel. Lorsque des données identifiantes sur des établissements étaient fournies, elles ont fait l'objet d'une pseudonymisation dans les cas où les répondants avaient demandé à ce que les informations directement identifiantes sur leur établissement ne soient pas diffusées.

Disponibilité des données de l'enquête

Après la phase de traitement des données, un jeu de données a été déposé dans Zenodo avec l'ensemble de la documentation qui accompagne l'enquête le 9 décembre 2020. Le jeu de données était accessible uniquement sur demande jusqu'à la publication de cette analyse. Le jeu de données a été diffusé publiquement le 9 juillet 2021. Il est accessible par l'intermédiaire de son DOI : [10.5281/zenodo.4313147](https://doi.org/10.5281/zenodo.4313147)⁴.

1.2. Typologie des répondants

Typologie des établissements répondants

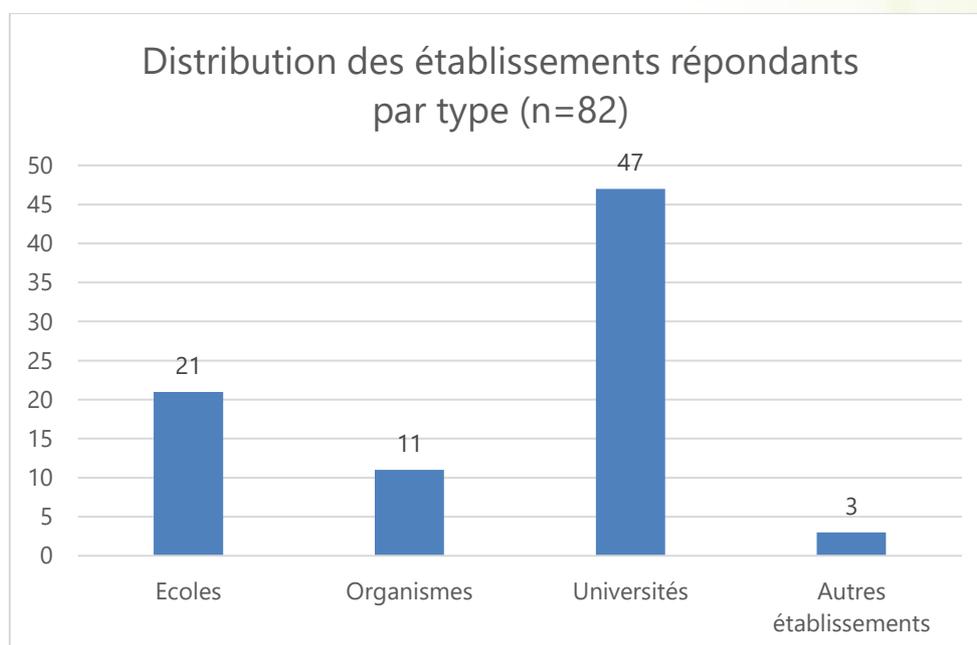
Le travail sur la typologie des établissements répondants s'est appuyé sur celle utilisée par Couperin pour déterminer les collèges électoraux, mais elle a été aménagée pour les besoins de l'enquête. La catégorie utilisée par Couperin "universités et assimilés" a été scindée pour ne garder que les universités d'une part et distribuer les établissements "assimilés universités" dans les autres catégories. Les principes suivants ont été retenus :

- Les écoles et institut de formation ont été placés dans la catégorie "Ecoles" ;
- Les établissements dont la mission principale est la recherche ont été placés dans la catégorie "Organismes de recherche".

Par ailleurs, les établissements dotés de la personnalité morale ont été placés dans la catégorie "Autres établissements". Comme le montre le graphique ci-dessous, les universités représentent plus de la moitié

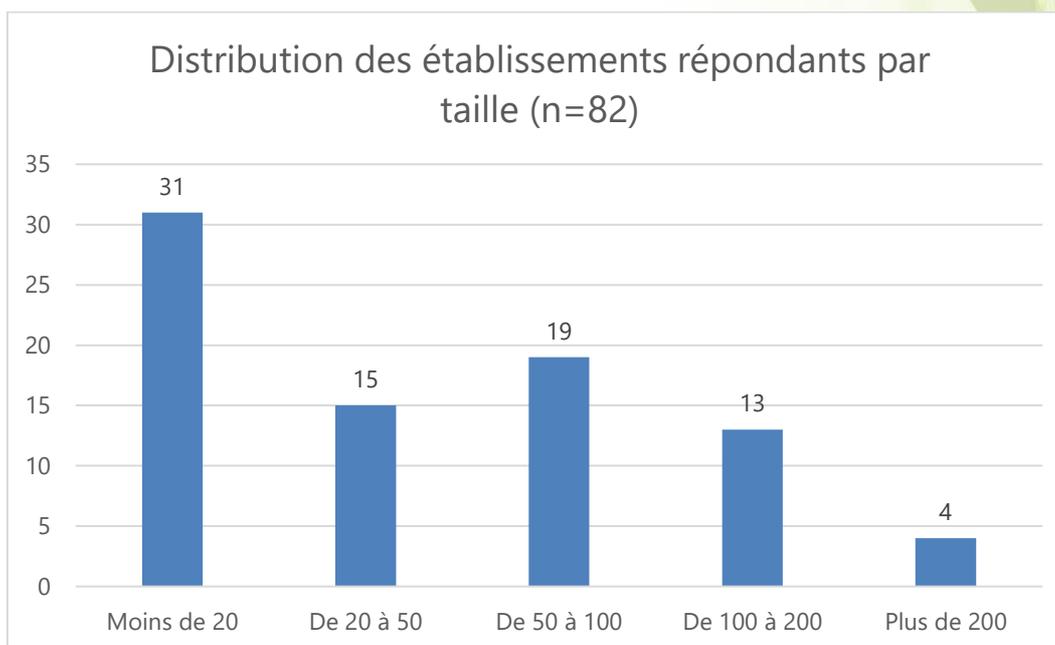
⁴ Féret, Romain, Augouvernaire, Martine, Bracco, Laetitia, Larrieu, Maxence, & Lehoux, Elise. (2020). Données de l'enquête Couperin sur l'accompagnement à la gestion des données de recherche par les services de documentation [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.4313147>

des répondants. Les écoles sont aussi particulièrement bien représentées de même que, dans une moindre mesure, les organismes de recherche.



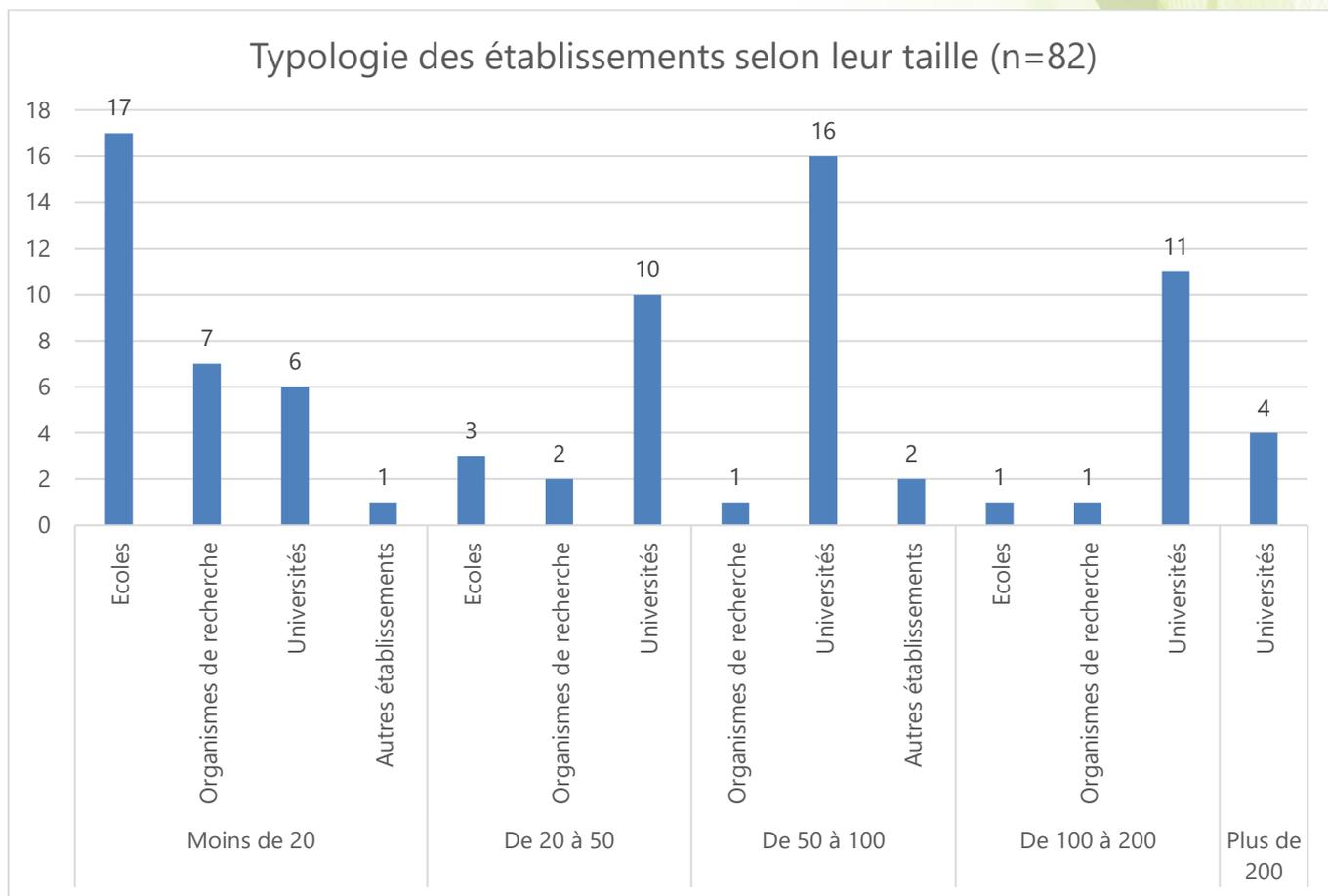
Taille des établissements répondants

Les services de documentation et d'information scientifique et technique qui ont répondu à l'enquête sont de tailles très variables, avec une **majorité d'établissements de petite ou moyenne taille**. Moins de 20 % des services regroupent plus de 100 personnes, alors que près de 40 % d'entre eux sont composés de moins de 20 personnes. Le reste des établissements, un peu plus de 40 %, sont compris entre 20 et 100 agents.



La distribution des services par taille n'est pas du tout homogène selon le type d'établissement. **Les écoles représentent plus de la moitié des services de moins de 20 agents (17/31)** et seules 4 écoles ont plus de 20 agents. **Les services de documentation des organismes de recherche sont de petite ou moyenne taille.** Ils sont répartis entre les différentes catégories à l'exception de la catégorie « plus de 200 agents ». Cependant, certains ont d'importants moyens dédiés à la Science Ouverte et à la gestion des données de recherche (Inist-CNRS, INRAE). Enfin, **les services de documentation des universités sont les mieux répartis entre les différentes catégories de taille**, et sont généralement de moyenne ou grande taille. Les deux tiers d'entre eux sont des services de plus de 50 agents (31/47).

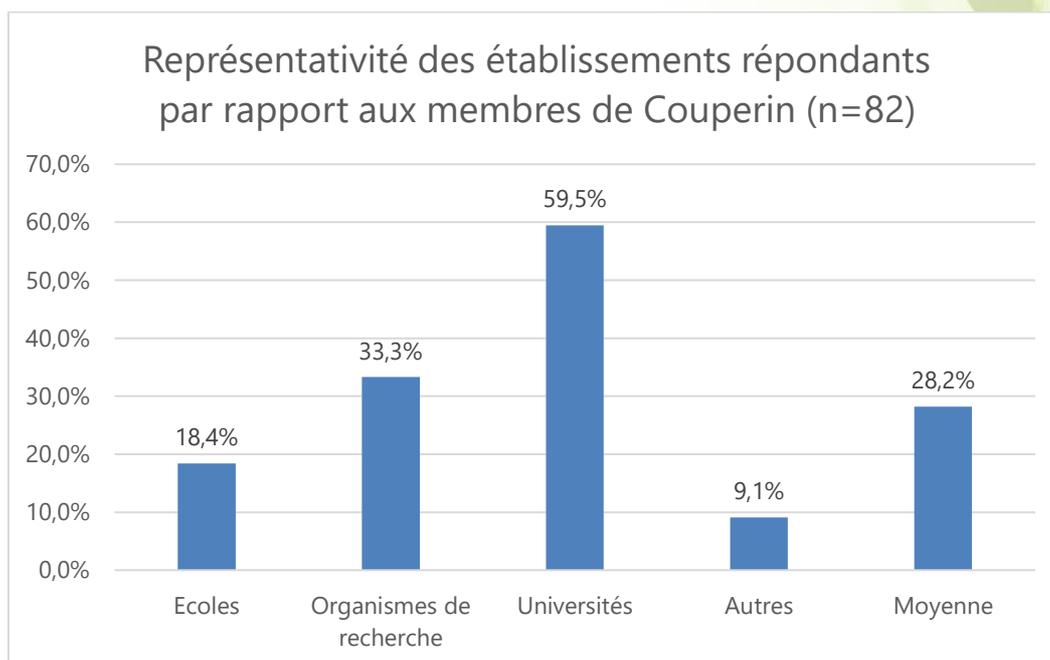
Comme il existe une corrélation partielle entre la taille des services de documentation et le type d'établissement, il n'est pas toujours aisé d'identifier laquelle de ces deux clés d'analyse est la plus pertinente. **Nous avons analysé l'enquête selon ces deux variables**, parfois en les croisant pour affiner au mieux la typologie des services de documentation. La taille des services semble la variable prépondérante, mais à taille égale les organismes de recherche offrent généralement des services plus étendus et structurés que les écoles et les universités.



Représentativité des répondants par rapport à la population Couperin

A noter qu'aucune COMUE⁵, bibliothèque interuniversitaire ou bibliothèque de centre hospitalier n'ont répondu à l'enquête. Les établissements n'étant ni des écoles, des organismes de recherche ou des universités ont été classés dans la catégorie "Autres établissements". Le tableau ci-dessous présente la représentativité des répondants par rapport aux établissements adhérents de Couperin, selon le type d'établissement.

⁵ Les COMUE sont des communautés d'universités et établissements. Ce type de regroupement permet de regrouper des établissements d'[enseignement supérieur](#) et de [recherche](#).



A nouveau, les universités sont très bien représentées puisque 60 % d'entre elles ont répondu à cette enquête, ce qui montre un réel intérêt pour ce sujet. Cela peut aussi s'expliquer par les modalités de communication de l'enquête (sur le forum ADBU⁶ notamment), complémentaire de celle menée sur les listes de diffusion de Couperin.

A noter que la présence modeste en valeur absolue des organismes de recherche correspond en réalité à une représentativité intéressante puisqu'un tiers des organismes membres de Couperin ont répondu à l'enquête. A l'inverse, plus de 80 % des écoles n'ont pas répondu à l'enquête malgré une participation intéressante en valeur absolue. Cet écart est assez logique dans la mesure où la question des données de la recherche intéresse particulièrement les institutions pour lesquelles la recherche est une composante majeure de leurs missions, ce qui n'est pas toujours le cas des écoles.

La catégorie "Autres établissements" est de loin la catégorie la moins bien représentée. Dans le cas des COMUE, cela peut s'expliquer parfois par l'inexistence de services de documentation. Pour les centres hospitaliers, il serait nécessaire de mener une analyse complémentaire pour savoir pourquoi aucun d'entre eux n'a répondu. Une hypothèse est que les questions de données étant particulièrement prégnantes dans ces centres depuis plusieurs années, les services de documentation n'ont pas pu ou n'ont pas jugé utile de

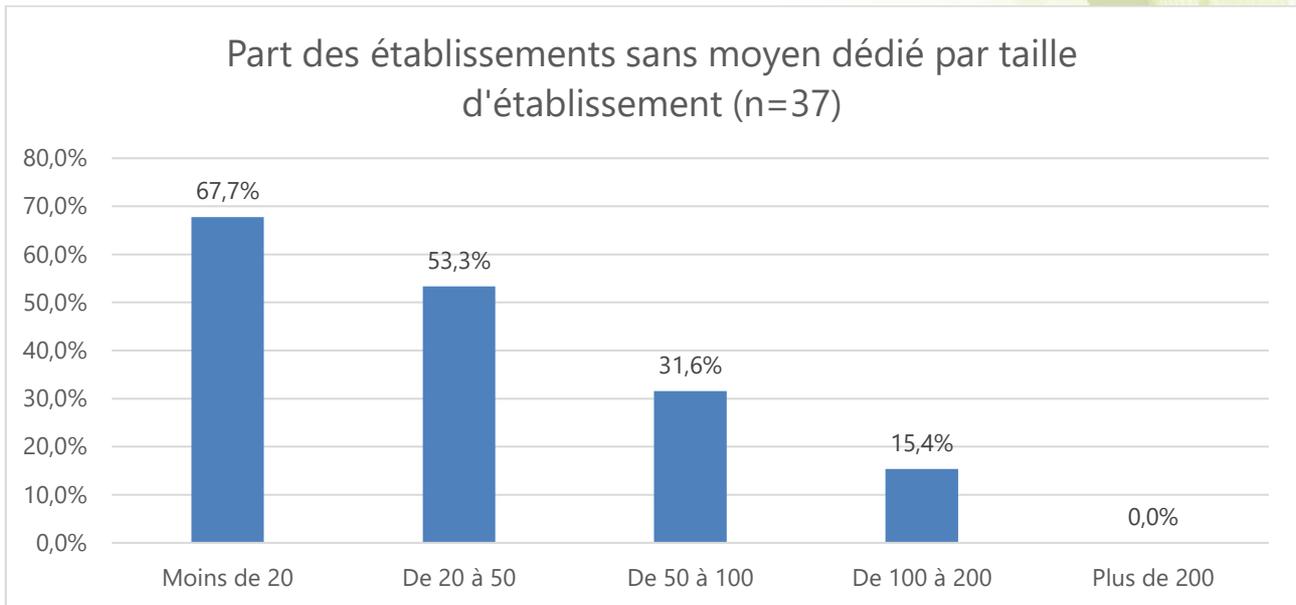
⁶ L'ADBU est l'Association française des directeurs et personnels de direction des bibliothèques universitaires et de la documentation.

se positionner sur ce sujet. En effet, du fait de la sensibilité des données patients, il existe une acculturation déjà forte des communautés professionnelles et de recherche à gérer ces données de manière sécurisée. Par ailleurs, il est rare que les données produites dans les hôpitaux soient diffusées librement alors que le rôle des services de documentation arrive souvent à cette étape du cycle de vie des données.

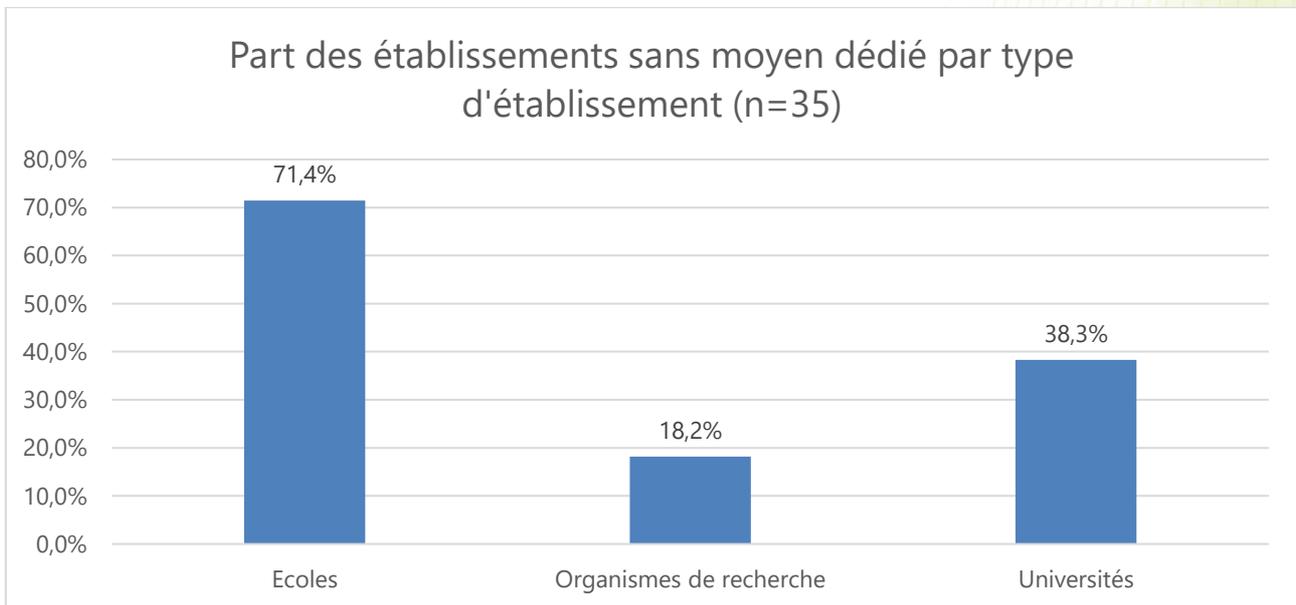
Dans leur ensemble, 30 % des établissements membres de Couperin ont répondu à l'enquête ce qui montre une certaine sensibilité à ce sujet avec de très grandes différences selon les types d'établissement comme expliqué précédemment. **Cela confère cependant une représentativité réelle aux résultats de cette enquête, en particulier pour les services de documentation des universités.**

1.3. Les établissements qui n'accompagnent pas la gestion des données de la recherche (GDR), quels sont-ils ?

37 établissements ont indiqué être dans ce cas, sans personnel dédié à ce sujet. A la lecture des résultats de l'enquête, l'existence de moyens humains dédiés à l'accompagnement de la gestion des données est un préalable à l'existence d'une offre de services. La question 3, et par corollaire la question 4, permettent de distinguer les établissements qui ont entamé une démarche pour accompagner la gestion des données, des établissements qui n'accompagnent pas du tout la gestion des données. **La taille des établissements est un facteur important puisque plus de la moitié des services qui n'accompagnent pas la gestion des données sont composés de moins de 20 personnes (21/37).** Et, seuls deux services de plus de 100 personnes n'accompagnent pas la gestion des données, les 14 autres comprenant entre 20 et 100 personnes. Il apparaît clairement que la part des services qui n'accompagnent pas la gestion des données est proportionnelle à leur taille.

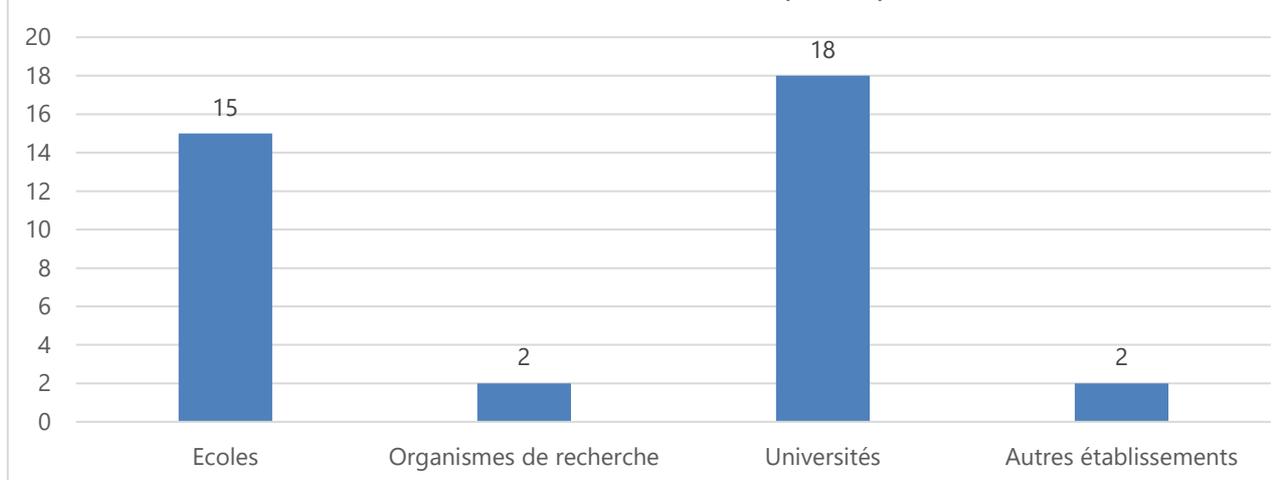


Dans la mesure où les écoles ont majoritairement de petits services de documentation, cette répartition des moyens d'accompagnement se retrouve logiquement dans une approche par type d'établissement puisque 70 % des écoles n'accompagnent pas la gestion des données contre seulement 20 % des organismes de recherche.



La grande majorité des services de documentation qui n'accompagnent pas la gestion des données relèvent des universités et des écoles (33 sur 37).

Distribution des établissements sans moyen dédié par type d'établissements (n=37)



En croisant les deux facteurs dans le tableau ci-dessous, on s'aperçoit que parmi les services de documentation de plus de 20 agents qui n'accompagnent pas la gestion des données, on retrouve 14 universités et une école mais aucun service IST d'organisme de recherche. Les seuls services IST des organismes de recherche qui ne disposent pas de moyens dédiés sont des services qui comprennent moins de 20 personnes, et il s'agit seulement d'une minorité d'entre eux (2/7).

A l'inverse, cela représente une grande majorité des écoles dont les services de documentation comprennent moins de 20 agents (14/17) et des universités (4/6). **Dans le cas des universités, la part des services qui n'accompagnent pas la gestion des données diminue à mesure que la taille du service augmente** : 7 sur 10 pour les services de 20 à 50 agents, 5 sur 16 pour ceux de 50 à 100 agents, et seulement 2 sur 15 pour les services de plus de 100 agents.

Taille des établissements	Ecoles	Organismes de recherche	Universités	Autres établissements
Moins de 20	14	2	4	1
De 20 à 50	1	0	7	0
De 50 à 100	0	0	5	1
De 100 à 200	0	0	2	0
Plus de 200	0	0	0	0

Les établissements où les services de documentation n'accompagnent pas la gestion des données actuellement correspondent donc aux caractéristiques suivantes :

- Une minorité des services de documentation de petite taille des organismes de recherche ;
- Une grande majorité des écoles dont les services de documentation incluent moins de 20 personnes ;
- Des universités dont les services de documentation sont de petite ou moyenne taille, généralement inférieure à 50 agents mais parfois comprise entre 50 et 100 agents.

A l'opposé, il est possible de dresser un profil type des services de documentation qui accompagnent la gestion des données :

- Les services de documentation des organismes de recherche, quelle que soit leur taille ;
- Les services de documentation des écoles lorsqu'ils excèdent 20 agents ;
- Une majorité des services de documentation des universités qui excèdent 50 agents, et presque tous ceux qui excèdent 100 agents.

Facteurs d'explications au non positionnement des établissements

L'analyse des verbatims issus des réponses de ces 37 établissements aux questions 36 à 39 du questionnaire permet de comprendre les raisons qui amènent ces établissements à ne pas avoir formalisé de services d'accompagnement à la gestion des données pour le moment. Il convient tout d'abord de préciser que **l'absence de moyens humains dédié ou de cellule structurée ne signifie pas que les services de documentation ne sont pas sollicités sur les questions relatives à l'appui à la gestion des données**. Certains services sont offerts ponctuellement aux chercheurs :

- Treize services ont déjà mené des actions de sensibilisation sous différentes formes : ateliers, formation doctorale, événements, parfois dans le cadre de dispositifs nationaux ou régionaux (réseaux des URFIST⁷, Opidor Tour⁸) ;

⁷ Les URFIST (Unité Régionale de Formation à l'Information Scientifique et Technique) ont pour missions la recherche et la formation des usagers universitaires, les pratiques et les évolutions de l'information scientifique. Voir <https://sygefor.reseau-urfist.fr/#/> ; URL consultée le 12/06/2021.

⁸ <https://tour.opidor.fr/> ; URL consultée le 12/06/2021.

- Trois ont mené une enquête sur les pratiques des chercheurs en matière de données ;
- Cinq ont été amenés à proposer un accompagnement sur des sujets tels que le choix des formats des fichiers de données ou les métadonnées ;
- Deux services ont déjà accompagné la rédaction de plan de gestion de données (PGD) ;
- Trois ont contribué à la mise en place d'un appui technique : mise en production d'une infrastructure liée aux données, attribution de DOI (*digital object identifier*).

Pour ce faire, ces services s'appuient notamment sur d'autres services de leur établissement et leur participation à des groupes de travail ou commissions professionnelles. La participation à des groupes de travail avec d'autres services de documentation, souvent à l'échelle régionale, semble un bon moyen de mutualiser leurs expériences pour les établissements qui ne se sont pas encore formellement positionnés sur ce sujet (par exemple : groupe Toul'AO en Occitanie⁹, gTIGRE¹⁰ dans la région Grand Est).

Parmi les facteurs qui expliquent que ces services de documentation n'accompagnent pas la GDR, le premier d'entre eux (15 occurrences) est **le manque de ressources humaines**, qui se caractérise à la fois par **l'absence de temps de travail mobilisable et par un manque de compétences identifiées au sein du service de documentation**. Les verbatims insistent particulièrement sur l'enjeu de formation des personnels : « nous avons besoin de plus de personnels formés et dédiés à la gestion des données de la recherche » ; « manque de personnel qualifié et formé aux problématiques de gestion des données de la recherche ».

Le 2^{ème} facteur (5 occurrences) recensé se caractérise par **l'absence d'impulsion politique de l'établissement et de coordination interservices sur le sujet**. La corrélation entre ces deux sujets est souvent mentionnée explicitement dans les verbatims : « pour avancer dans ce projet, nous avons besoin du portage politique par notre nouvelle présidente et par le VP Recherche. Nous avons besoin d'une collaboration étroite avec la Direction de la Recherche. On en revient donc à un soutien politique indispensable pour faire avancer ce dossier ». Plusieurs réponses font aussi apparaître les difficultés de

9 « Toul'AO » (groupe de travail « Toulouse Archives Ouvertes », SICD – réseau des bibliothèques de Toulouse, UFTMiP – Université Fédérale Toulouse Midi-Pyrénées). Voir le carnet Hypothèses : <https://openarchiv.hypotheses.org/> ; URL consultée le 12/06/2021.

¹⁰ Groupe de travail inter-établissements Grand Est sur les données de la recherche. Voir la communauté Zenodo : <https://zenodo.org/communities/gtigre> ; URL consultée le 12/06/2021.

collaboration avec les directions de la recherche : « accorder nos actions avec le département Recherche, préciser les périmètres » ; « manque de coordination avec la direction de la recherche et les équipes de recherche ». A l'inverse, les établissements où un lien existe avec les services centraux semblent favorisés pour qu'un travail sur ce sujet soit initié sans attendre qu'une offre de services ne soit formalisée : « Notre direction de la recherche et valorisation ne s'est pas encore saisi du sujet. Elle nous a demandé d'attendre. Toutefois, elle redirige certaines interrogations de chercheurs vers nous et nous les conseillons selon les préconisations d'OPIDoR».

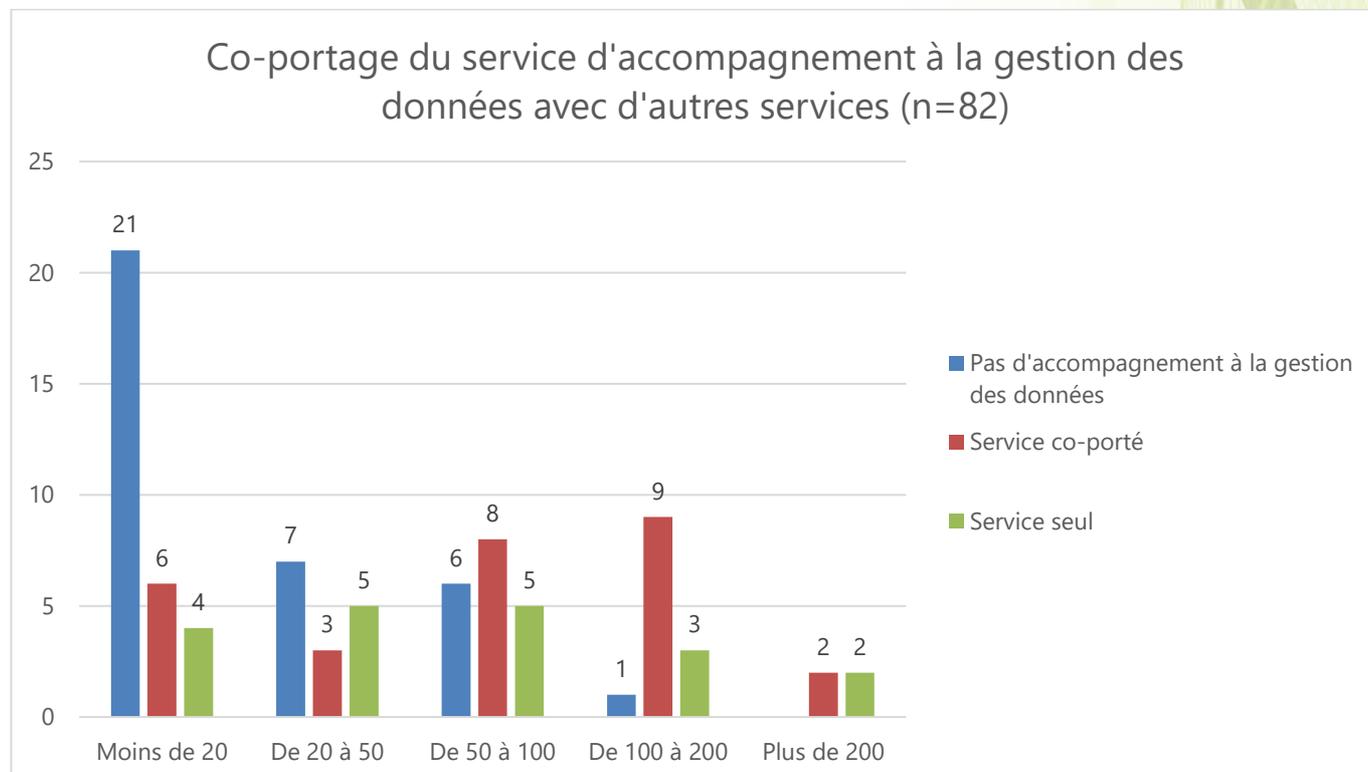
Enfin, le 3^{ème} facteur (6 occurrences) est le sentiment d'un **manque de légitimité de certains services de documentation**, éventuellement renforcé par le positionnement d'un autre service sur ce sujet au sein de leur établissement. Un des établissements indique par exemple : « la bibliothèque n'est pas aujourd'hui identifiée par l'établissement comme un acteur potentiel dans la gestion des données ou l'accompagnement des projets de recherche, ce qui explique qu'elle n'intervient pas dans ce domaine ». Parmi l'ensemble de ces services, certains n'envisagent pas de commencer à proposer un accompagnement à la gestion des données, mais la majorité souhaiterait pouvoir le faire, si les freins étaient levés. Les projets envisagés par ces services pour les deux prochaines années en vue d'initier véritablement une offre de services sont de 3 types :

- Structuration d'un service au niveau institutionnel avec d'autres services partenaires ;
- Contribution à la définition d'une politique institutionnelle ;
- Mise en place de services concrets : développement d'actions de sensibilisation, accompagnement à la rédaction de PGD et des projets, attribution de DOI.

En matière de développement de services, aucune tendance générale ne se dégage sur les actions prioritaires ou une logique de phasage.

2. Section 1 de l'enquête : Organisation du service

2.1. Le service d'accompagnement à la gestion des données est-il propre à votre service ou est-il partagé avec d'autres services ou directions ? (Plusieurs réponses possibles) – Q7

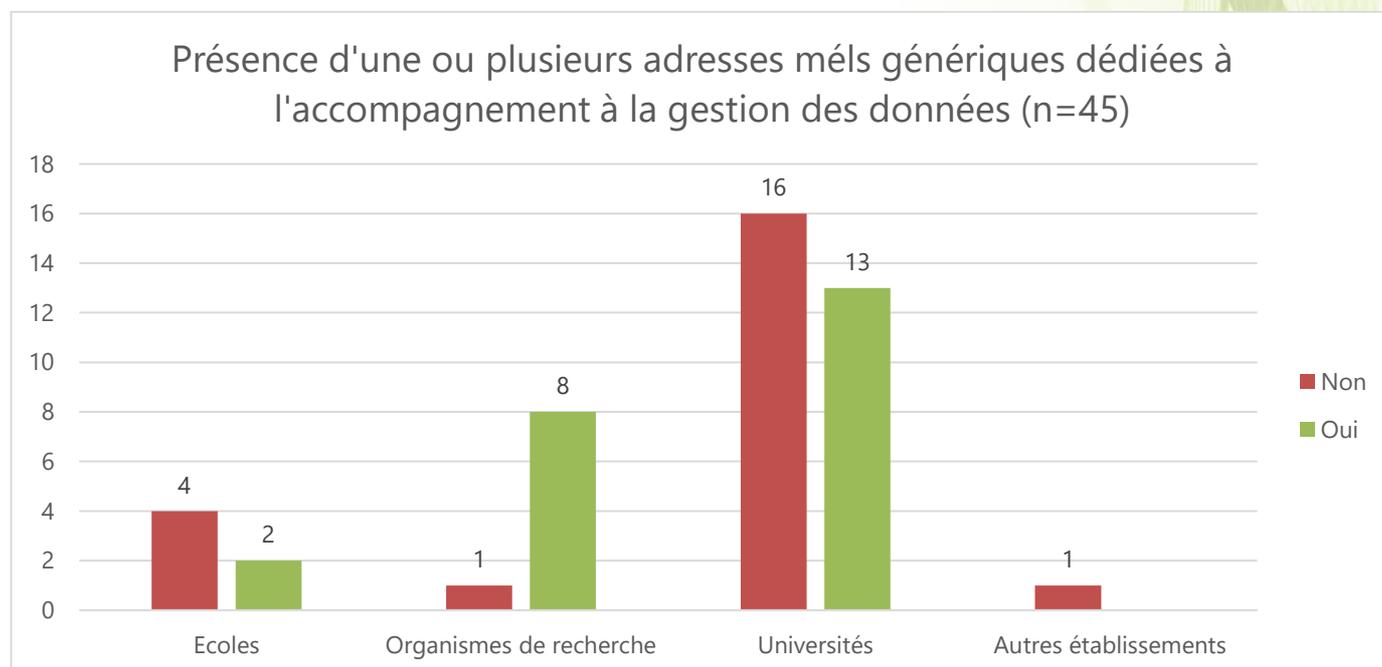


Cette question portait sur le **co-portage** de la gestion des données entre différents services. Il est possible qu'elle ait été interprétée différemment par les répondants, qui peuvent l'avoir également comprise en termes de « collaborations ». La collaboration entre les services sur la GDR étant abordée dans la suite de l'enquête, les résultats de cette question sont à envisager de manière nuancée.

Sur les 47 cellules d'accompagnement à la gestion des données, **28 (soit 60 %) sont co-portées**. Elles se situent majoritairement dans les établissements de grande taille (11 établissements de plus de 100 personnes) : plus la taille du service est importante (au-delà de 50 personnes), plus l'accompagnement à la GDR est partagé. Plus que la typologie d'établissement, c'est la taille qui semble un facteur déterminant pour expliquer ce phénomène.

Parmi les cellules co-partagées, on observe pour une majorité d'entre elles le co-partage avec les services informatiques, la direction Recherche ou les archivistes.

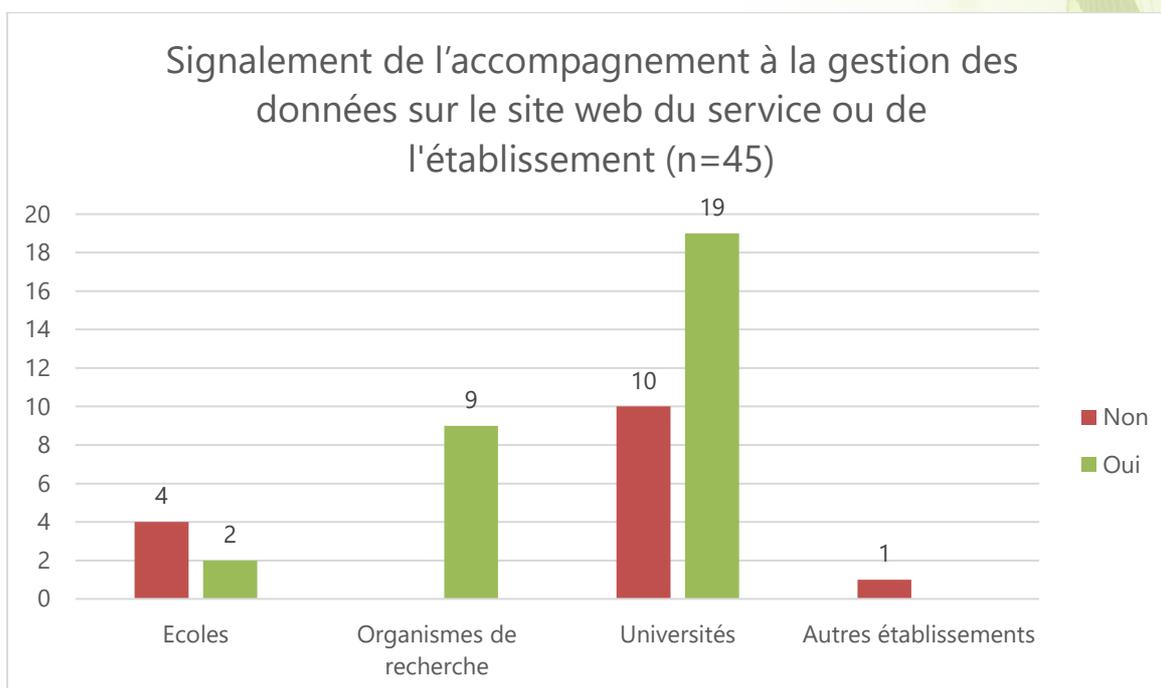
2.2. Une ou plusieurs adresses méls génériques sont-elles dédiées à l'accompagnement à la gestion des données ? – Q8



Le graphique représente la distribution des cellules d'accompagnement qui ont mis en place une adresse mél générique dédiée, selon le type d'établissement. **Près de la moitié n'ont pas mis en place d'adresse mél générique dédié (22/45)**. Cela représente une majorité des services de documentation des universités et des écoles (20/35), alors que les organismes de recherche disposent très majoritairement d'un ou plusieurs alias pour gérer leurs interactions au sujet des données de recherche (8/9).

Une représentation par taille d'établissement des services qui disposent d'un alias dédié à la gestion des données ne montre pas de corrélation probante.

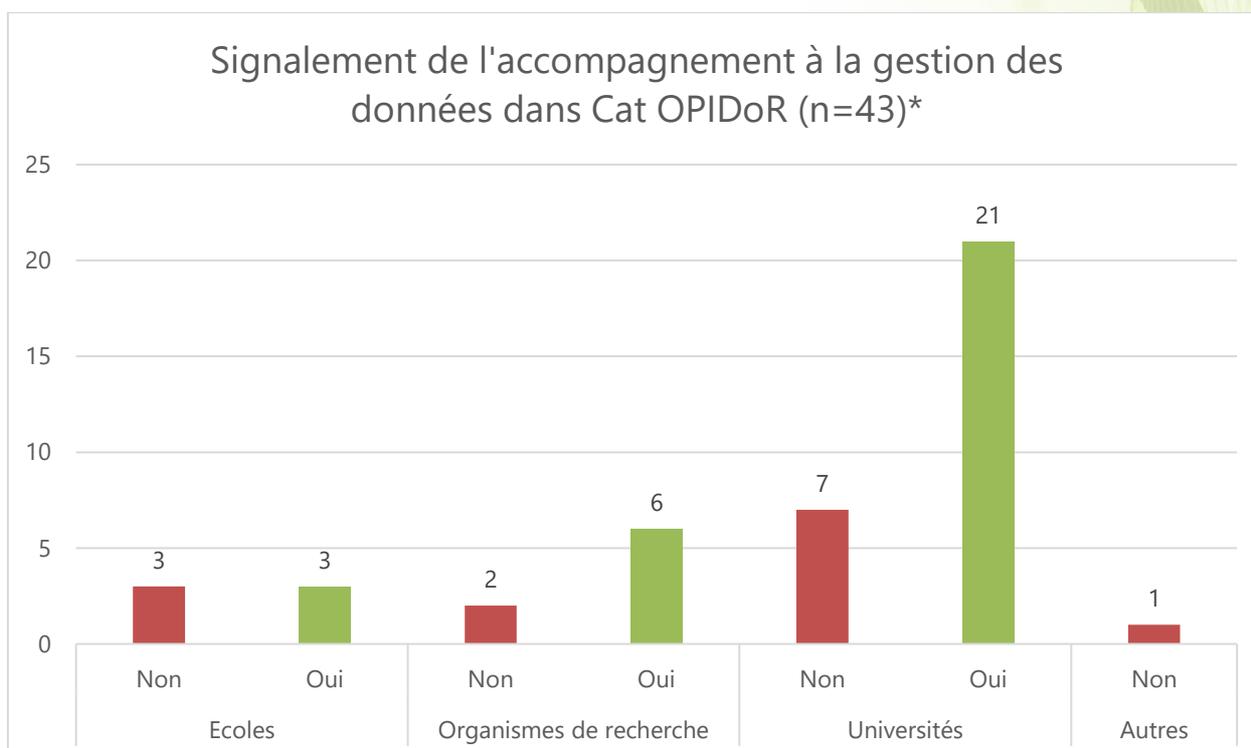
2.3. L'accompagnement à la gestion des données est-il signalé sur le site web de votre service ou de votre établissement ? – Q9



Une majorité d'établissements qui offrent des cellules d'accompagnement à la gestion des données signale ces services sur leur site web (30/45), mais un tiers des établissements ne le font pas. Les différences entre types d'établissements constatées dans les réponses à la question précédente se retrouvent dans celle-ci : les universités et les écoles représentent la grande majorité des établissements qui ne signalent pas leur service d'appui à la gestion des données sur leur site web (14/15), même si les deux tiers des universités le font. Tous les organismes de recherche signalent leur accompagnement sur leur site web (9/9).

La corrélation entre la taille des services et leur propension à signaler leurs services sur leur site web n'est pas franche entre les établissements de petite et de moyenne taille mais est claire par rapport à ceux de grande taille. En effet, **seul un service sur cinq de plus de 100 agents (3/15) ne signale pas son accompagnement sur son site web** tandis que la proportion est d'un sur deux et demi pour ceux inférieurs à 100 agents (12/30).

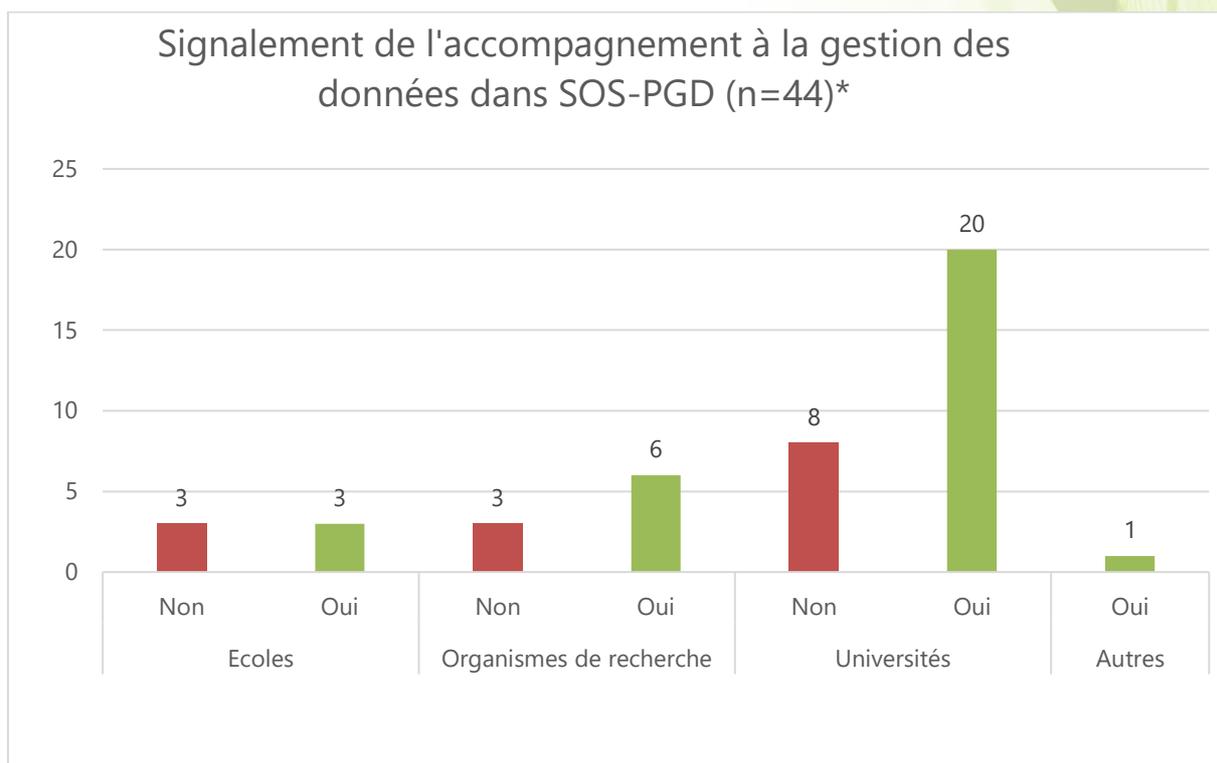
2.4. L'accompagnement à la gestion des données est-il signalé dans [Cat OPIDoR](#)¹¹ et/ou dans [SOS-PGD](#)¹² ? – Q10 et Q11



* Deux établissements ont répondu « ne sait pas » (une université et un organisme de recherche).

¹¹ Cat OPIDoR, wiki des services dédiés aux données de la recherche géré par l'Inist-CNRS. [https://cat.opidor.fr/index.php/Cat OPIDoR, wiki des services d %C3 %A9di %C3 %A9s aux donn %C3 %A9es de l a recherche](https://cat.opidor.fr/index.php/Cat_OPIDoR_wiki_des_services_d%C3%A9di%C3%A9s_aux_donn%C3%A9es_de_la_recherche) ; URL consultée le 12/06/2021.

¹² Le répertoire des Services Opérationnels de Soutien à la rédaction des Plans de Gestion des Données (SOS-PGD) recense les services accompagnant la rédaction des plans de gestion des données au sein des établissements d'enseignement supérieur et de la recherche. <https://scienceouverte.couperin.org/sos-pgd/>, URL consultée le 12/06/2021.



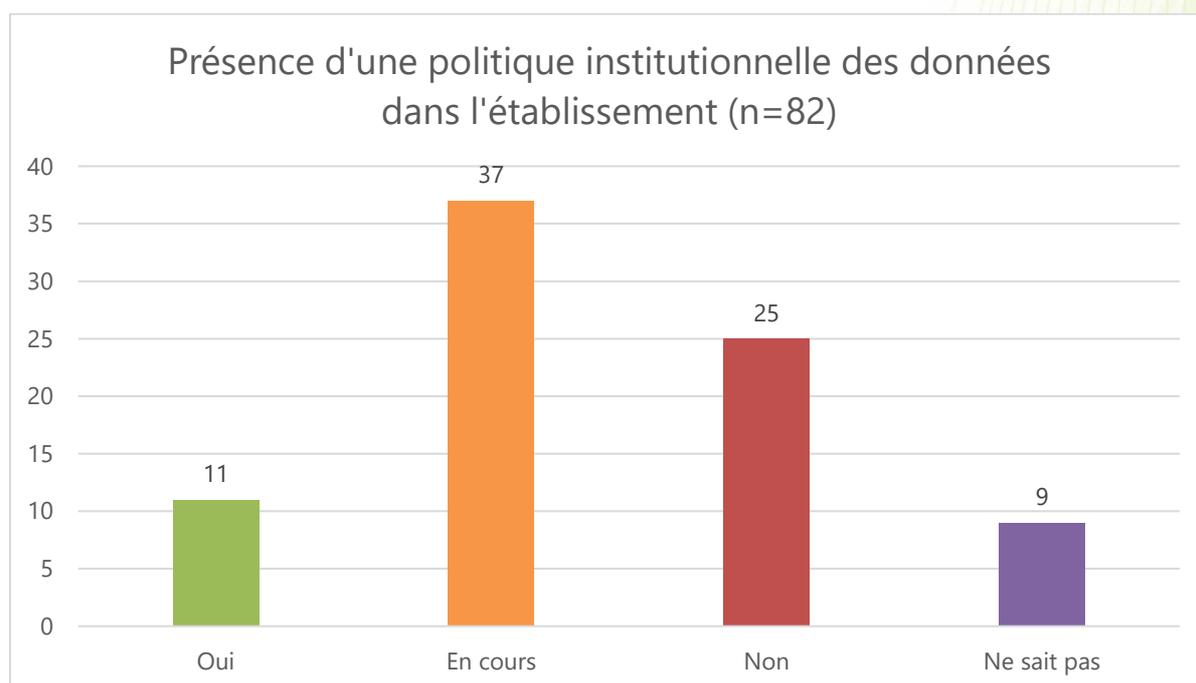
* Un établissement a répondu « ne sait pas » (un service de documentation de 50 à 100 agents, en université)

Les questions 10 et 11 portent sur le signalement des cellules d'appui dans des outils nationaux qui facilitent leur identification, respectivement Cat OPIDoR et SOS-PGD. **La proportion d'établissements qui signalent leur cellule dans Cat OPIDoR et dans SOS-PGD est d'environ deux tiers.** Elle est légèrement supérieure à la proportion de ceux qui signalent leur cellule sur leur site web, et la répartition par type d'établissement n'est pas identique. En effet, seuls 6 organismes de recherche sur 9 ont signalé leur cellule dans ces deux outils, alors qu'à l'inverse les services documentaires d'université sont plus nombreux à s'y être signalés, respectivement 21 dans Cat OPIDoR et 20 dans SOS-PGD contre 19 sur leur site web. Cela conduit les universités à être proportionnellement légèrement surreprésentées par rapport aux organismes de recherche pour leur présence sur Cat OPIDoR (20/27 contre 6/8) et sur SOS-PGD (21/28 contre 6/9). Le signalement dans SOS-PGD est facilement et rapidement réalisable et les services sont signalés directement par Cat OPIDoR, ce qui peut expliquer en partie ce constat.

La corrélation entre la taille de l'établissement et la communication sur les cellules d'accompagnement se maintient, avec une **large majorité des services de plus de 100 agents qui sont signalés sur ces deux outils** (13/15) tandis qu'environ **la moitié de ceux de moins de 50 agents y sont signalés** (respectivement 7/16 pour Cat OPIDoR et 9/17 pour SOS-PGD).

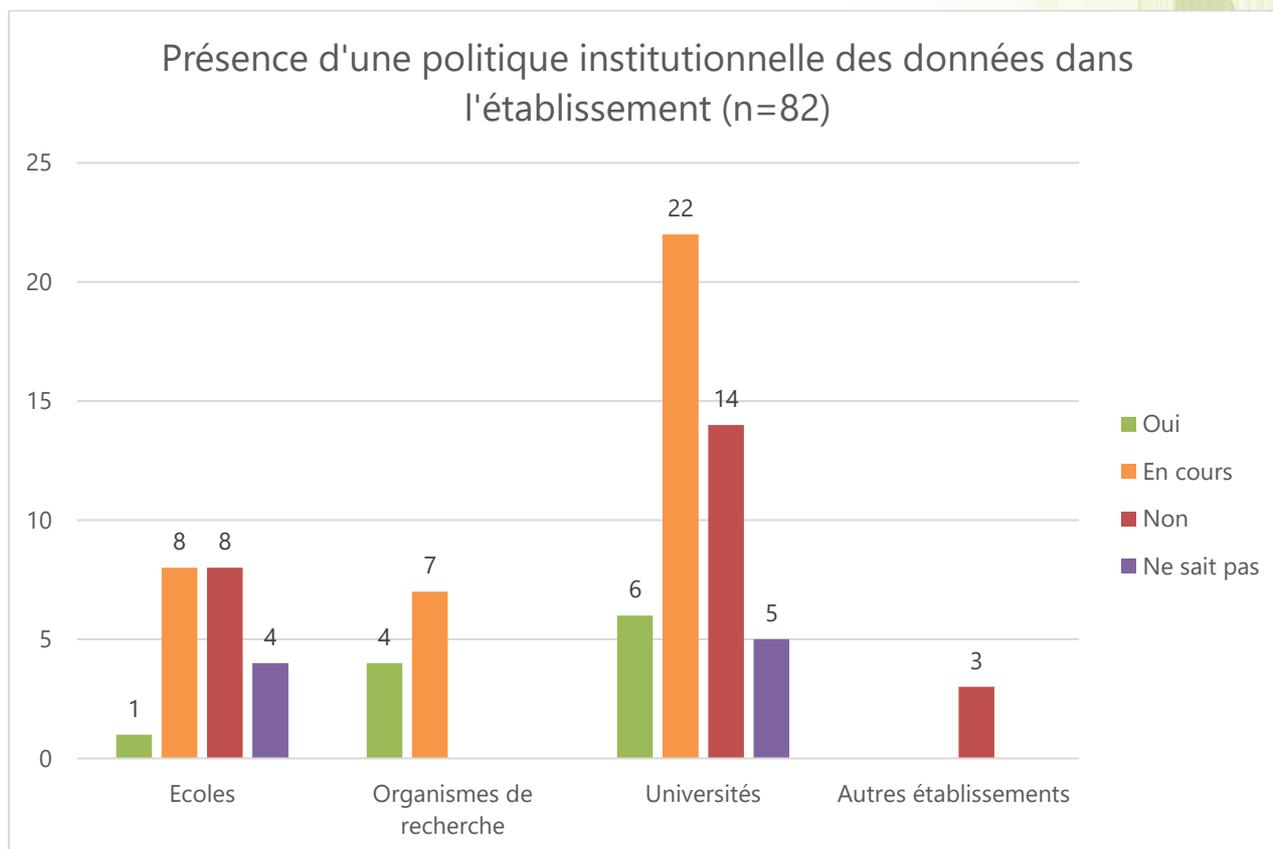
L'analyse conjointe de ces réponses montre que **les services documentaires des organismes de recherche affichent plus clairement leur soutien à la gestion des données que ceux des universités et surtout des écoles**. Les services de documentation d'universités sont plus présents dans les outils nationaux (Cat OPIDoR et SOS-PGD) que par le biais d'outils institutionnels (alias, site web). Il existe aussi une corrélation entre la taille des services de documentation et l'affirmation de leur positionnement. L'écart est particulièrement important entre les services de moins de 100 agents et ceux de plus de 100.

2.6. Votre établissement s'est-il doté d'une politique institutionnelle relative à la gestion des données de recherche ? – Q12

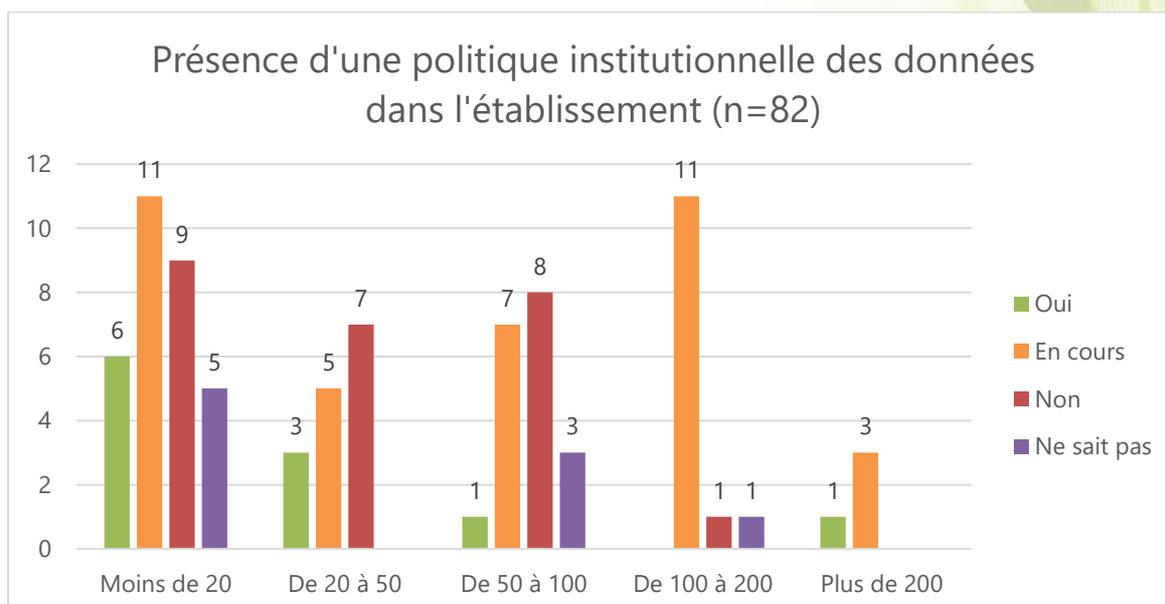


La mise en place d'une politique institutionnelle de données en est encore à ses prémices. Sur les 82 établissements ayant répondu à l'enquête, **seuls 11 disposent déjà d'une politique institutionnelle**,

tandis qu'elle est en cours de définition pour près de la moitié de l'échantillon (37). Il sera intéressant de suivre l'évolution de la formalisation de ces politiques et de leur mise en œuvre.

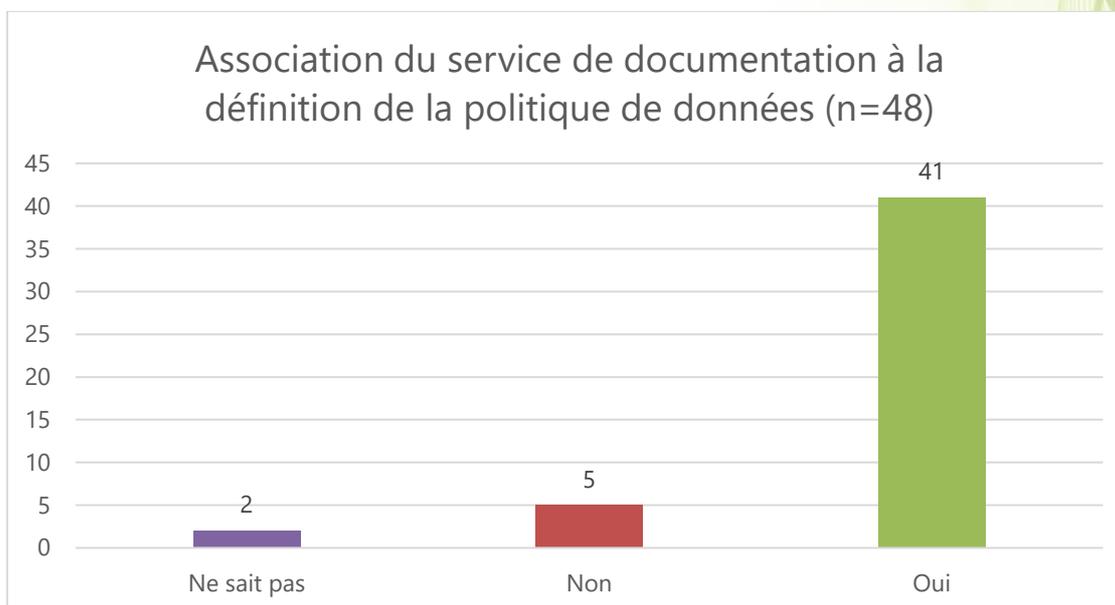


La différence entre types d'établissements est nette. D'un côté, plus de la moitié des écoles n'ont pas de politique (12/21), dont, pour 4 pour lesquelles les répondants ne savent pas si leur établissement s'est doté d'une politique. De l'autre, **l'ensemble des organismes de recherche ayant répondu à l'enquête ont une politique déjà en place ou en cours de définition**. Entre les deux, 60 % des universités ont déjà une politique ou se sont engagées dans ce processus (28/47).



Assez étonnement, **ce sont les établissements qui ont des services documentaires de moins de 20 personnes qui sont les plus nombreux à déclarer disposer d'une politique institutionnelle de données**. Les deux tiers des répondants de ces établissements indiquent disposer déjà d'une politique ou qu'elle est en cours de définition, tandis qu'ils sont seulement la moitié pour les établissements disposant de services de 20 à 100 agents. Seul un établissement de plus de 100 agents sur 17 indique avoir déjà une politique en place. Une explication pourrait être une plus grande difficulté des grandes universités à réunir l'ensemble des acteurs concernés et à définir une politique commune pour une recherche multidisciplinaire.

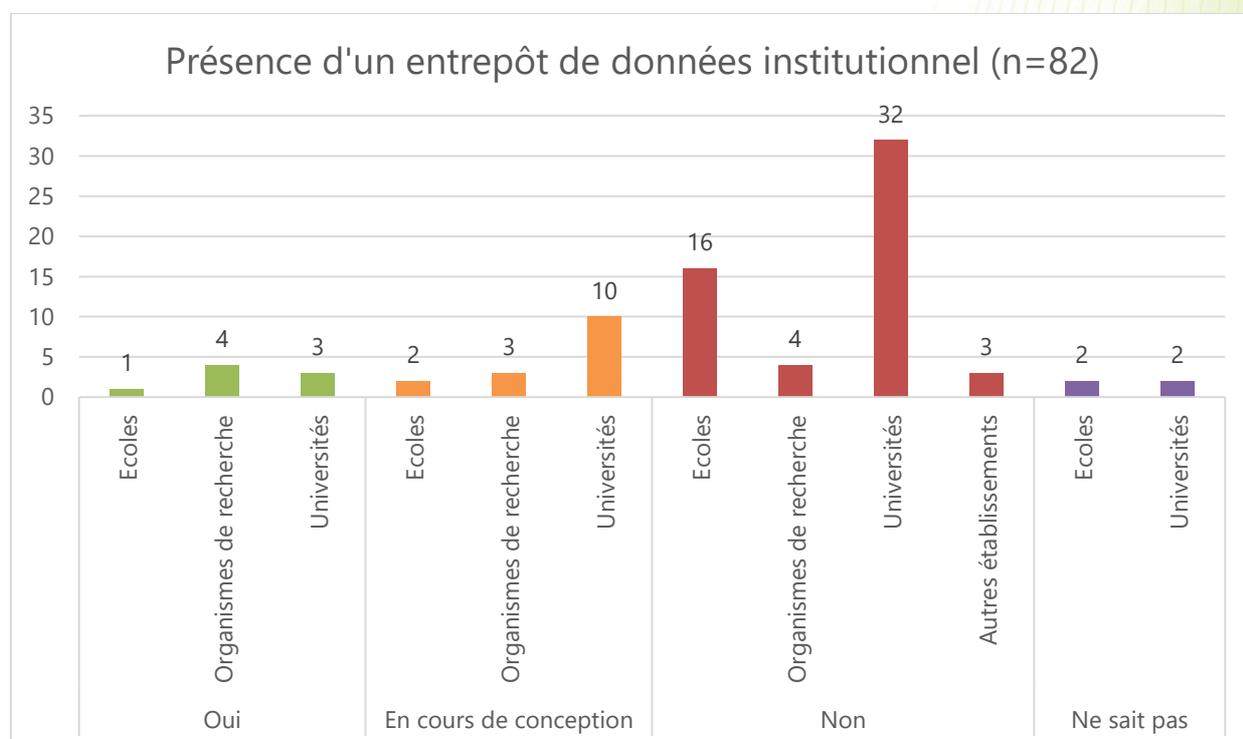
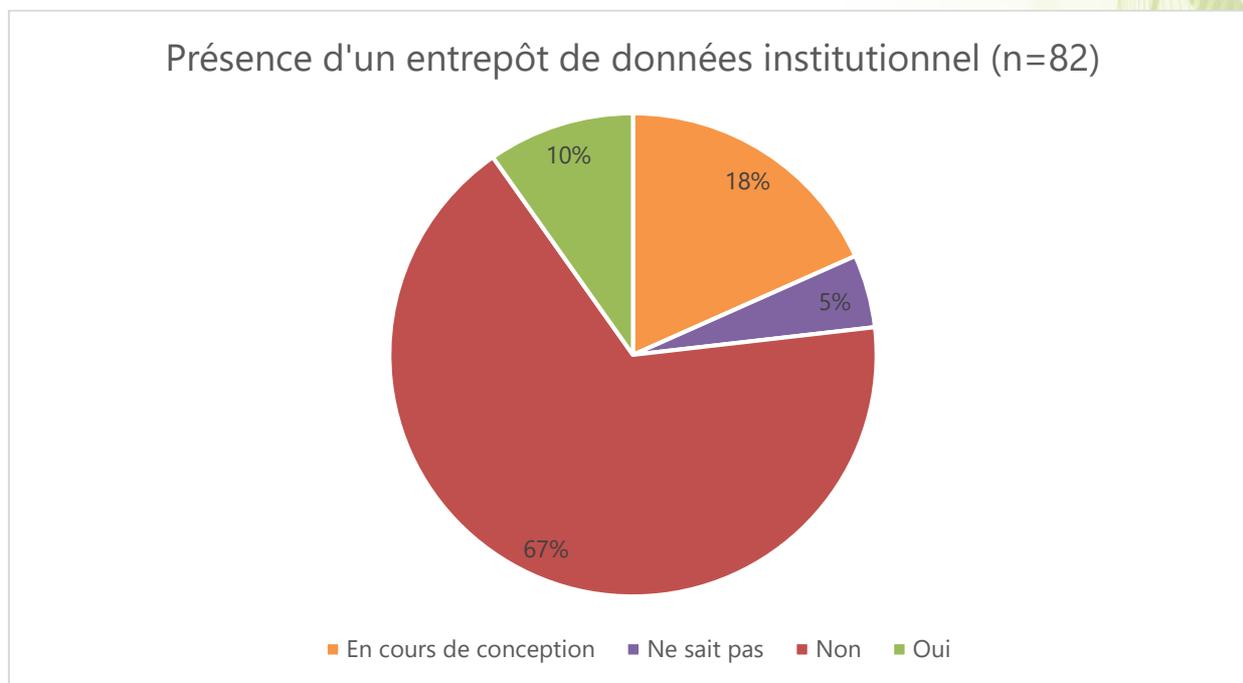
2.7. Votre service porte-t-il ou est-il associé à la définition et la mise en œuvre de la politique institutionnelle ou à la gouvernance de votre établissement en matière de gestion des données de recherche ? – Q13



Pour l'analyse des réponses apportées à cette question, nous avons retenu uniquement les 48 répondants ayant déclaré disposer déjà d'une politique institutionnelle de données ou en cours de rédaction. **Dans près de 90 % des cas, les services de documentation sont ou ont été associés à la définition de cette politique (41/46).** Deux des répondants ne connaissent pas le niveau d'association de leur service. Parmi les 5 établissements où le service de documentation n'a pas été associé, 4 sont des universités. Dans une majorité d'entre elles cependant, les services de documentation sont ou ont été associés (24/28).

3. Section 2 de l'enquête : Signalement et archivage des données de recherche

3.1. Disposez-vous d'un entrepôt institutionnel pour conserver des données de recherche ? – Q14



8 répondants (10 %) indiquent disposer d'un entrepôt institutionnel. Cependant, ce chiffre est à considérer avec précaution car il est possible que certains services aient répondu oui, dans le cas où leur archive ouverte permet de lier publications et données. Ce nombre est sans doute amené à évoluer puisque 15 répondants (18 %) indiquent qu'un entrepôt est en cours de conception dans leur établissement.

La répartition par type d'établissement montre que ces futurs entrepôts profiteront majoritairement aux universités (22 %). Parallèlement on constate que les établissements les moins équipés en entrepôts sont surtout les écoles et autres établissements.

3.2. Proposez-vous d'autres solutions institutionnelles pour signaler des jeux de données et/ou lier des publications et des jeux de données ? (ex. : possibilité de signaler un jeu de données dans votre archive ouverte, de manière autonome ou en lien avec une publication) – Q15

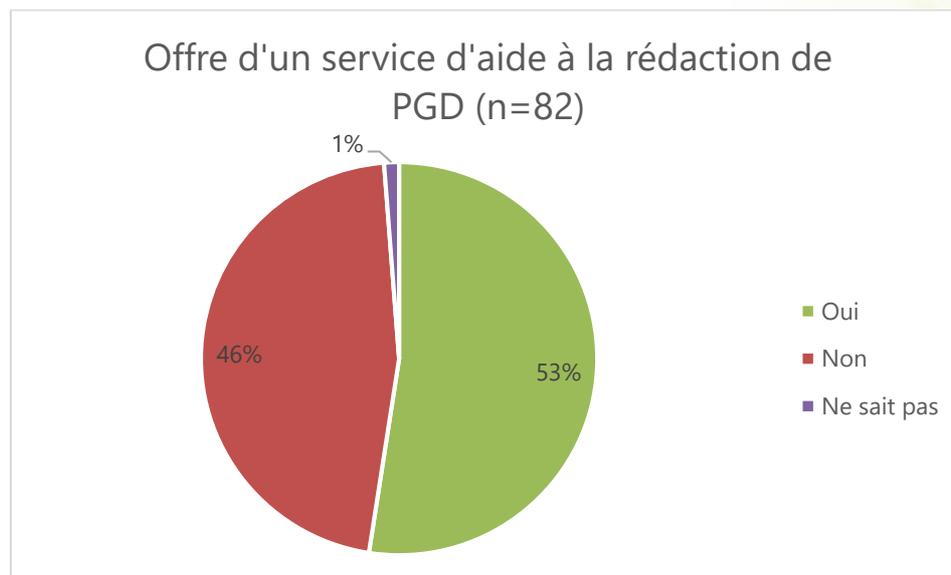
12 établissements (environ 15 %) ont indiqué des solutions institutionnelles pour signaler des jeux de données. Les solutions concernent essentiellement des possibilités de connexions entre archive ouverte et jeux de données et réciproquement. La présence dans l'archive ouverte peut se faire via une publication (comme c'est dans le cas dans HAL) ou bien indépendamment, sans publication, comme l'autorisent certaines archives ouvertes institutionnelles.

3.3. Disposez-vous d'un espace institutionnel sur des entrepôts de données généralistes ? (Plusieurs réponses possibles) – Q16

Seuls six répondants ont indiqué disposer d'une communauté institutionnelle sur Zenodo et aucun autre espace institutionnel sur des entrepôts généralistes (Figshare, Dryad) n'est mentionné. Cela semble assez étonnant car ces solutions semi-institutionnelles sont assez faciles à mettre en place et peuvent constituer une option intermédiaire avant la mise en place d'un entrepôt institutionnel de données. Le projet national de plateforme fédérée des données explique peut-être que certains établissements soient en attente sur ce sujet, la difficulté étant qu'aucune alternative ne semble émerger pour les chercheurs en attendant.

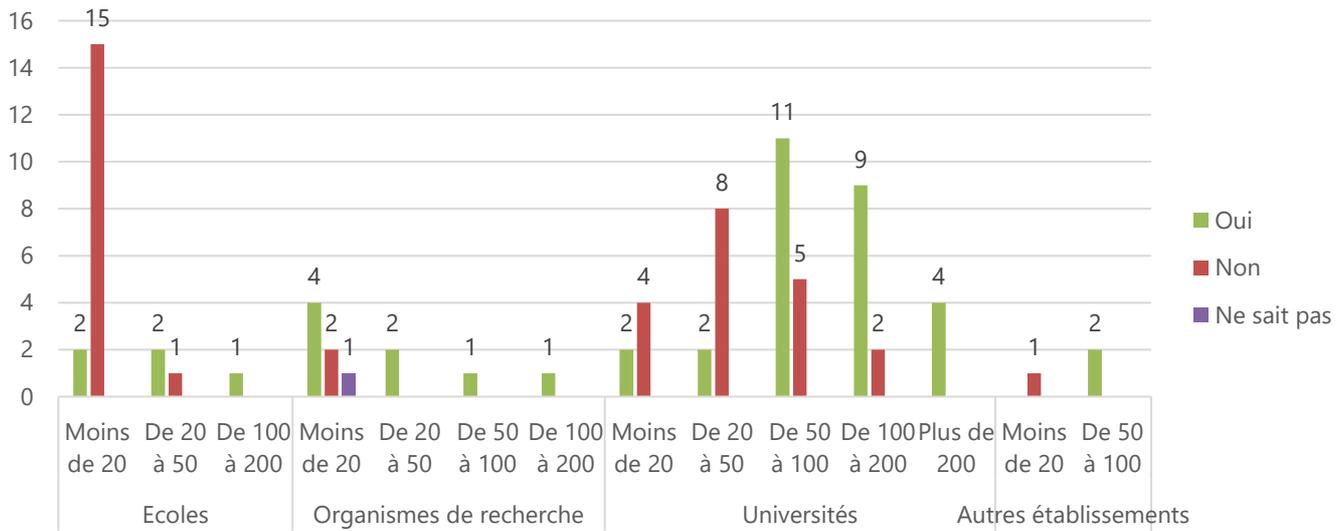
4. Section 3 de l'enquête : Accompagnement à la rédaction des plans de gestion de données

4.1. Votre service propose-t-il une aide à la rédaction des plans de gestion de données (explication des attendus du document, mise en relation avec d'autres services, mise à disposition de documents, conseils, etc.) ? – Q17



La **majorité des répondants (53 %)** a développé un service d'aide à la rédaction de plans de gestion de données. Il est présent dans 60 % des universités, 72 % des organismes de recherche et 24 % des écoles. Cette activité se développe surtout dans les **universités** disposant de services documentaires de plus de 50 agents et est très présente dans les organismes de recherche, indépendamment de leur taille, ce qui s'explique par leur expertise plus ancienne sur le sujet.

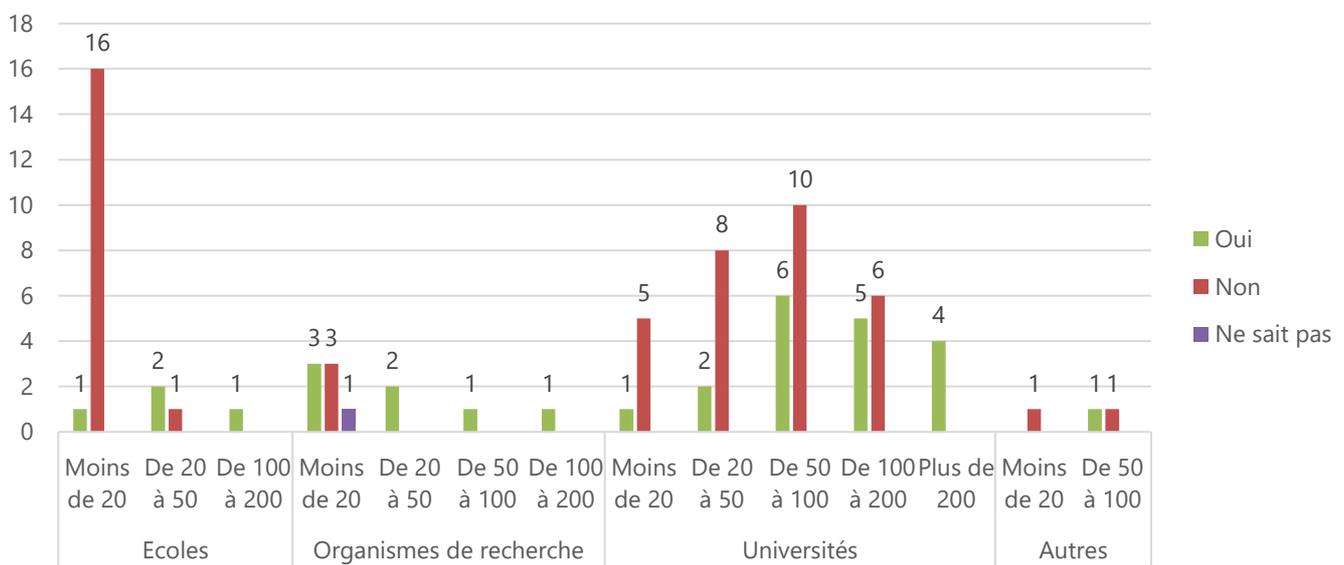
Offre d'un service d'aide à la rédaction de PGD (n=82)



Les questions suivantes permettent de préciser et quantifier le type d'accompagnement proposé, en dehors des formations et des conseils ponctuels.

4.2. Offrez-vous un service de relecture des plans de gestion de données ? – Q18

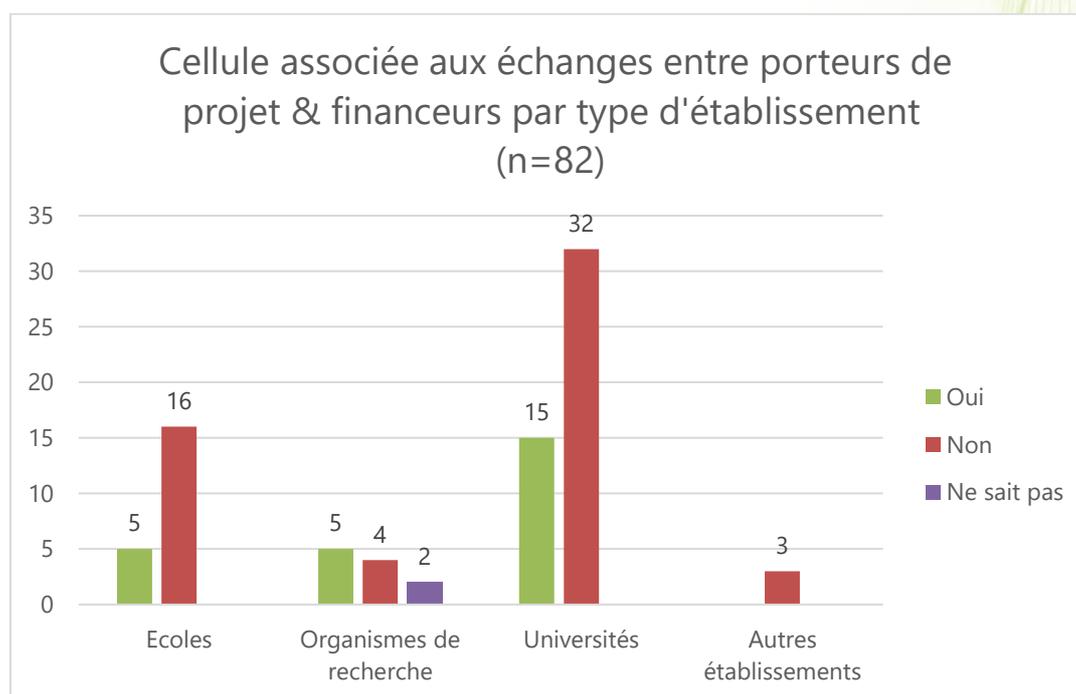
Offre d'un service de relecture de PGD (n=82)



Dans le cadre de l'accompagnement aux PGD, **37 % des établissements proposent un service de relecture**. On les retrouve surtout parmi les services de taille moyenne à grande, soit 63 % des organismes de recherche, 38 % des universités et 19 % des écoles. Cette différence, par rapport à un service plus global

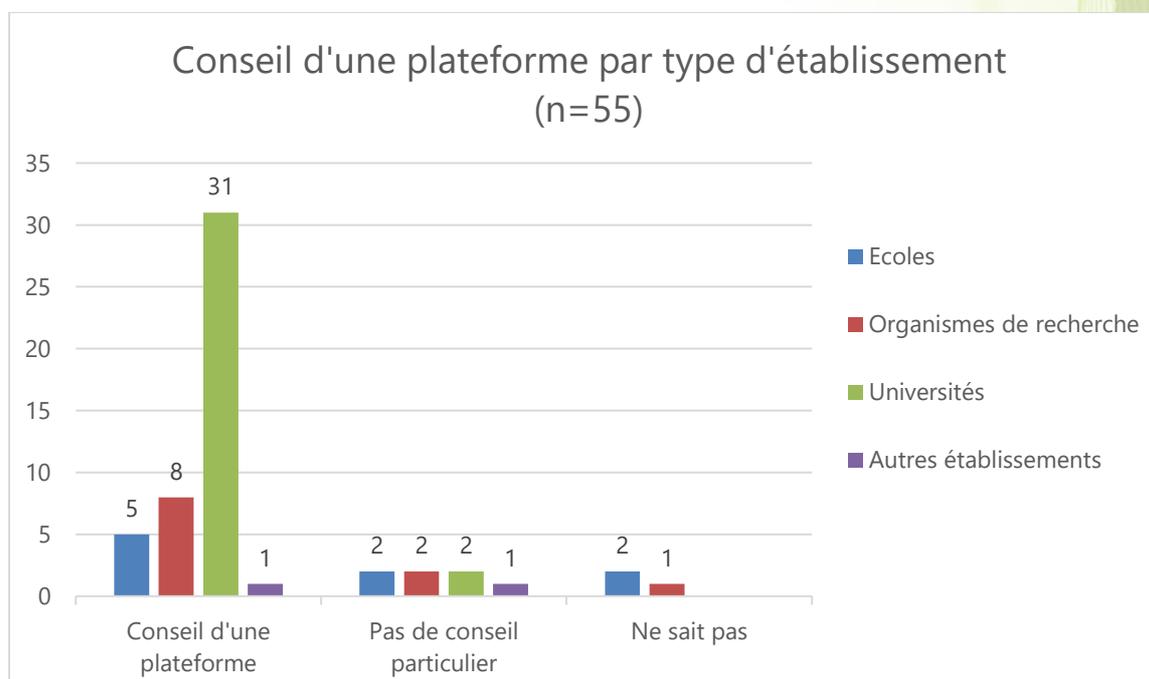
de conseils sur les PGD, s'explique probablement par le nombre de personnels nécessaire à déployer pour assurer cette expertise plus chronophage. Ces chiffres sont probablement appelés à augmenter dans les prochaines années.

4.3. Au cours des 12 derniers mois, avez-vous été associé aux échanges entre un porteur de projet de votre institution et un financeur (ANR, H 2020) au sujet de leur plan de gestion de données ? – Q19



Une **minorité d'établissements (25 % des répondants) a été impliquée** dans des échanges entre le porteur de projet et le financeur au sujet du plan de gestion. Leur répartition au sein des différentes institutions est similaire au profil des cellules qui offrent une relecture de plan de gestion de données (Q. 18). Plus que la taille ou le type d'établissement, c'est probablement l'expertise acquise par les cellules qui explique le développement de ce type d'échanges avec les financeurs. Ces échanges témoignent du franchissement d'une étape supplémentaire dans la spécialisation des équipes qui portent cet accompagnement.

4.4. Conseillez-vous une plateforme spécifique aux chercheurs de votre établissement qui rédigent un plan de gestion de données ? (Plusieurs réponses possibles) – Q20

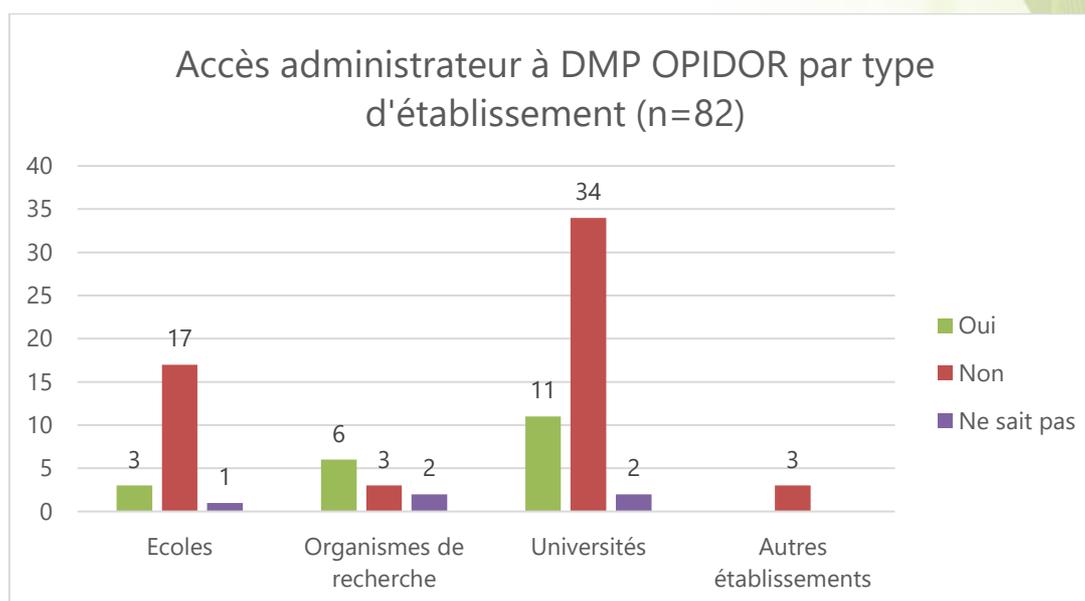


Parmi les établissements qui accompagnent à la gestion des données, **une grande majorité d'entre eux recommandent une plateforme pour la rédaction de PGD** (soit 82 % des cellules d'accompagnement). Quelques établissements ne recommandent pas de plateformes spécifiques.

Parmi les plateformes recommandées, **DMP OPIDoR¹³ est majoritairement conseillée** par 77 % des cellules d'accompagnement, ce qui témoigne de sa visibilité et de son appropriation par les institutions françaises. S'ajoute à cette référence, d'autres recommandations mais qui apparaissent comme relativement minoritaires (autre format et/ou autres plateformes conseillées), signe d'une probable spécialisation de ces répondants.

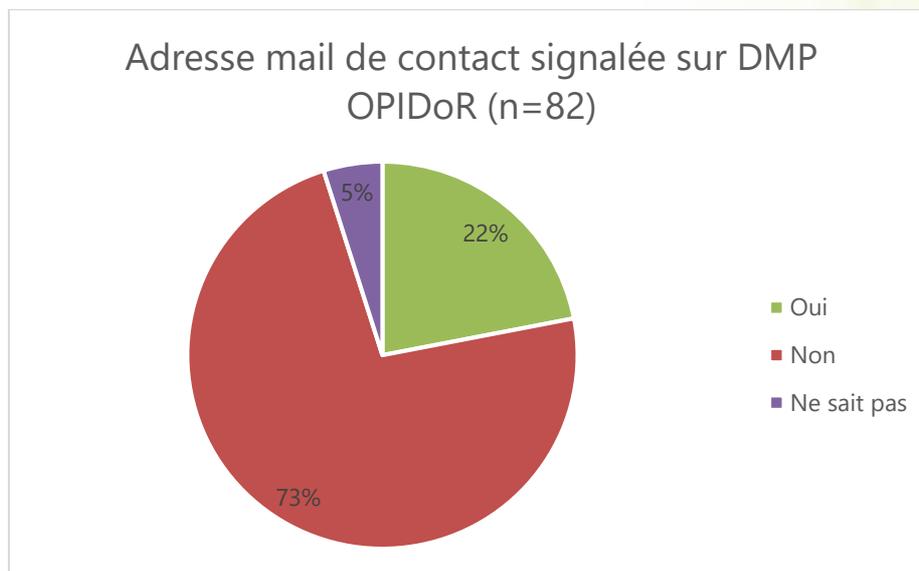
¹³ Outil de rédaction en ligne de plans de gestion de données. <https://dmp.opidor.fr/> ; URL consultée le 12/06/2021.

4.5. Disposez-vous d'un accès administrateur à DMP OPIDoR pour votre établissement ? – Q21



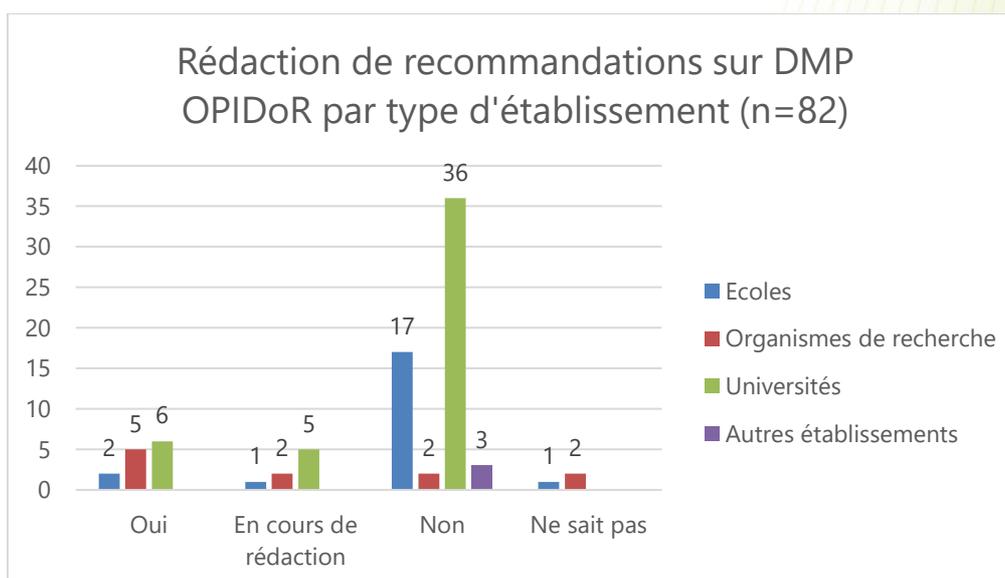
Toutefois, l'appropriation technique et/ou institutionnelle de la plateforme **DMP OPIDoR** est encore relativement récente, que ce soit du point de vue de son administration ou de la personnalisation institutionnelle. **24 % des cellules d'accompagnement disposent d'un accès administrateur à DMP OPIDoR**, essentiellement des organismes de recherche (6 sur 9 répondants) et des universités (11 sur 47). La très grande majorité des cellules d'accompagnement n'ont pas d'accès administrateur à la plateforme en 2020. Cette tendance sera certainement amenée à évoluer dans les prochaines années.

4.6. Avez-vous signalé une adresse mail de contact sur DMP OPIDoR permettant aux chercheurs de demander une assistance sur leur plan de gestion de données ? – Q22



Assez logiquement, une plus faible proportion des cellules d'accompagnement déclare avoir renseigné une adresse électronique de contact sur DMP OPIDoR, soit 22 % d'entre elles, essentiellement des organismes de recherche (6/9) et des universités (9/47). Les écoles ont répondu positivement pour 3 d'entre elles (sur 17).

4.7. Avez-vous ajouté des recommandations (aide contextualisée) propres à votre établissement sur DMP OPIDoR ? – Q23

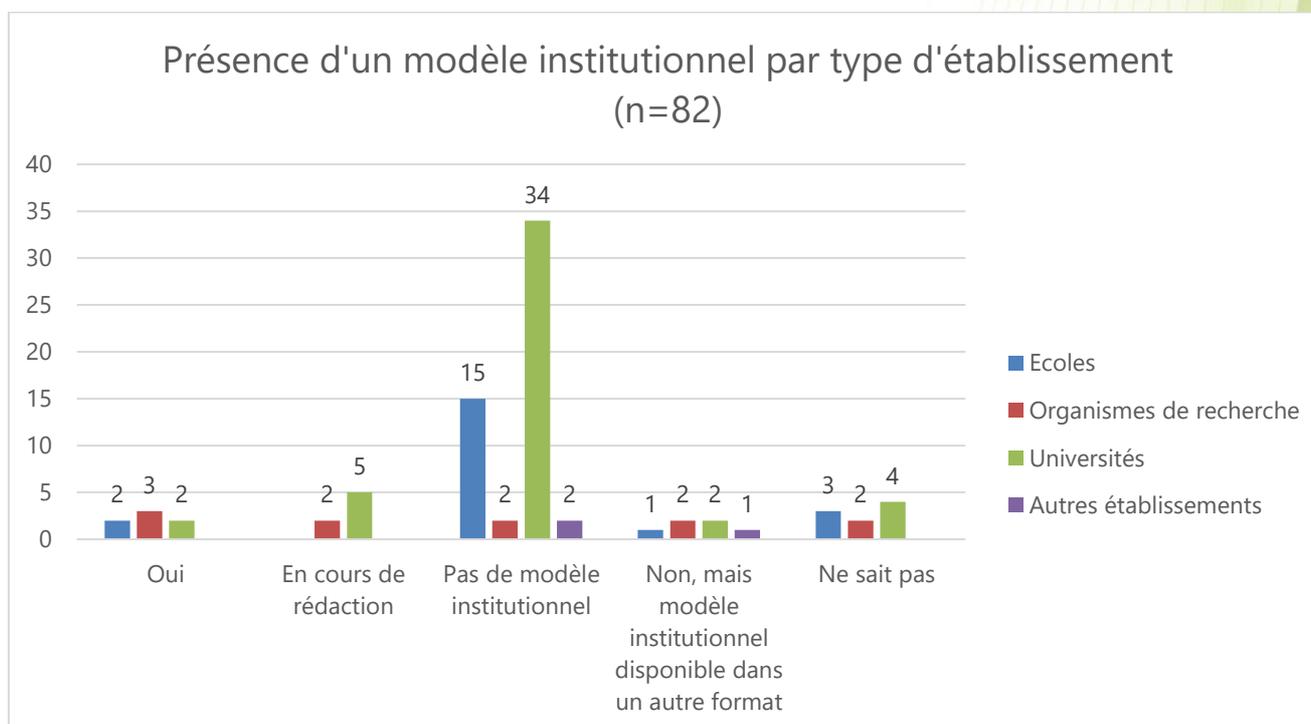


16 % des répondants ont ajouté des recommandations sur la plateforme DMP OPIDoR, essentiellement des universités et des organismes de recherche. Ce chiffre est amené à augmenter car parmi les répondants, **10 % ont signalé qu’elles étaient en cours de rédaction**, signe de la progressive spécialisation et montée en compétences des cellules d’accompagnement.

4.8. Disposez-vous d’un modèle institutionnel de plan de gestion de données sur DMP OPIDoR ? – Q24

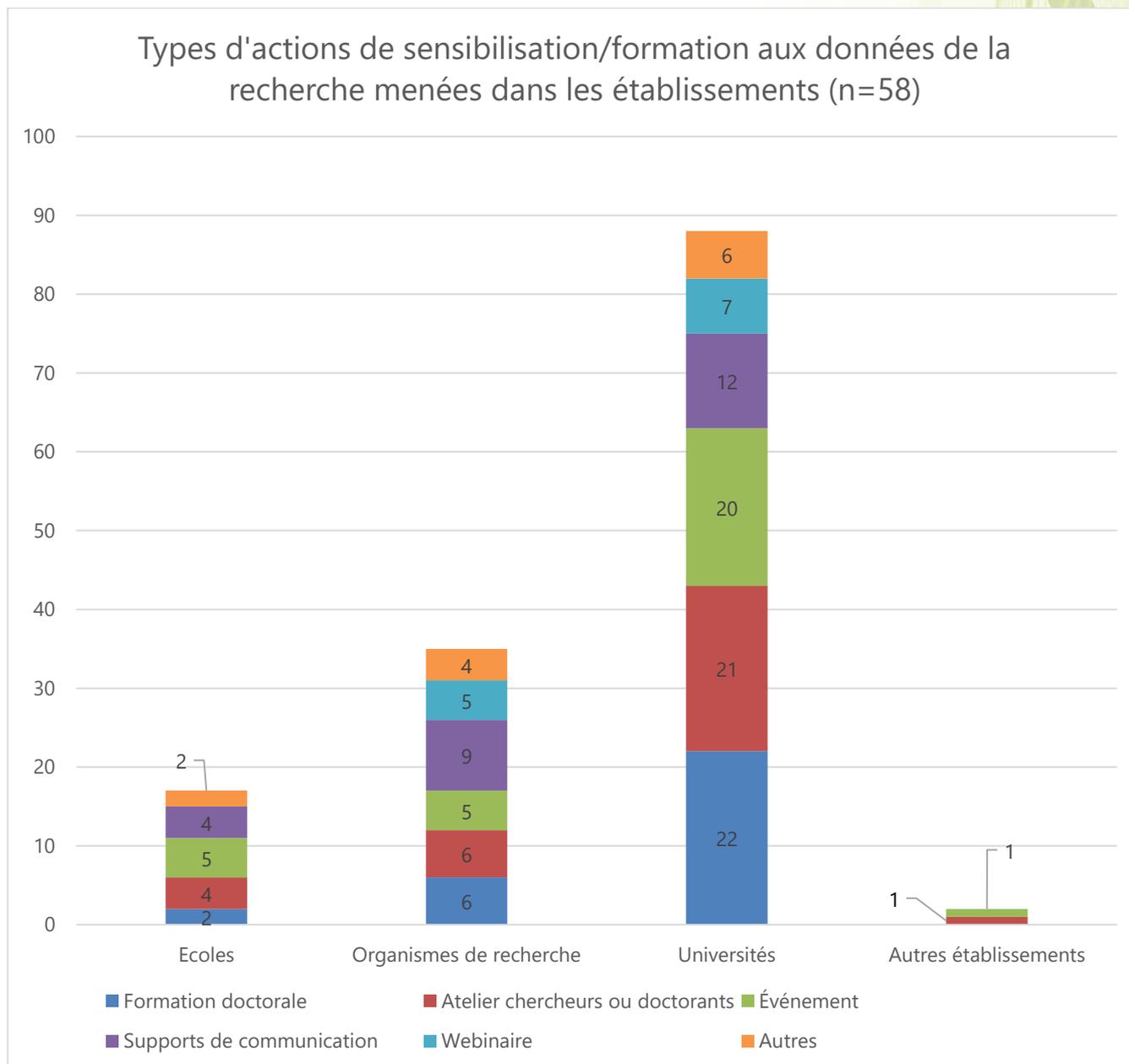
Oui	7
En cours de rédaction	7
Non, mais modèle institutionnel disponible dans un autre format	6
Pas de modèle institutionnel	53
Ne sait pas	9
Total général	82

Une très grande majorité des services (65 % des répondants) ne dispose pas de modèle institutionnel de PGD sur DMP OPIDoR. 7 déclarent en avoir un mais ce chiffre est en augmentation car 8,5 % des cellules ont déclaré que ce modèle est en cours de rédaction, majoritairement au sein des universités. 6 répondants déclarent disposer d’un modèle institutionnel de PGD dans un autre format (notamment pour les services de moins de 20 personnes).



5. Section 4 de l'enquête : Sensibilisation et formation

5.1. À quels types d'activités de sensibilisation et/ou de formation à la gestion des données votre service a-t-il contribué au cours des 12 derniers mois ? – Q25

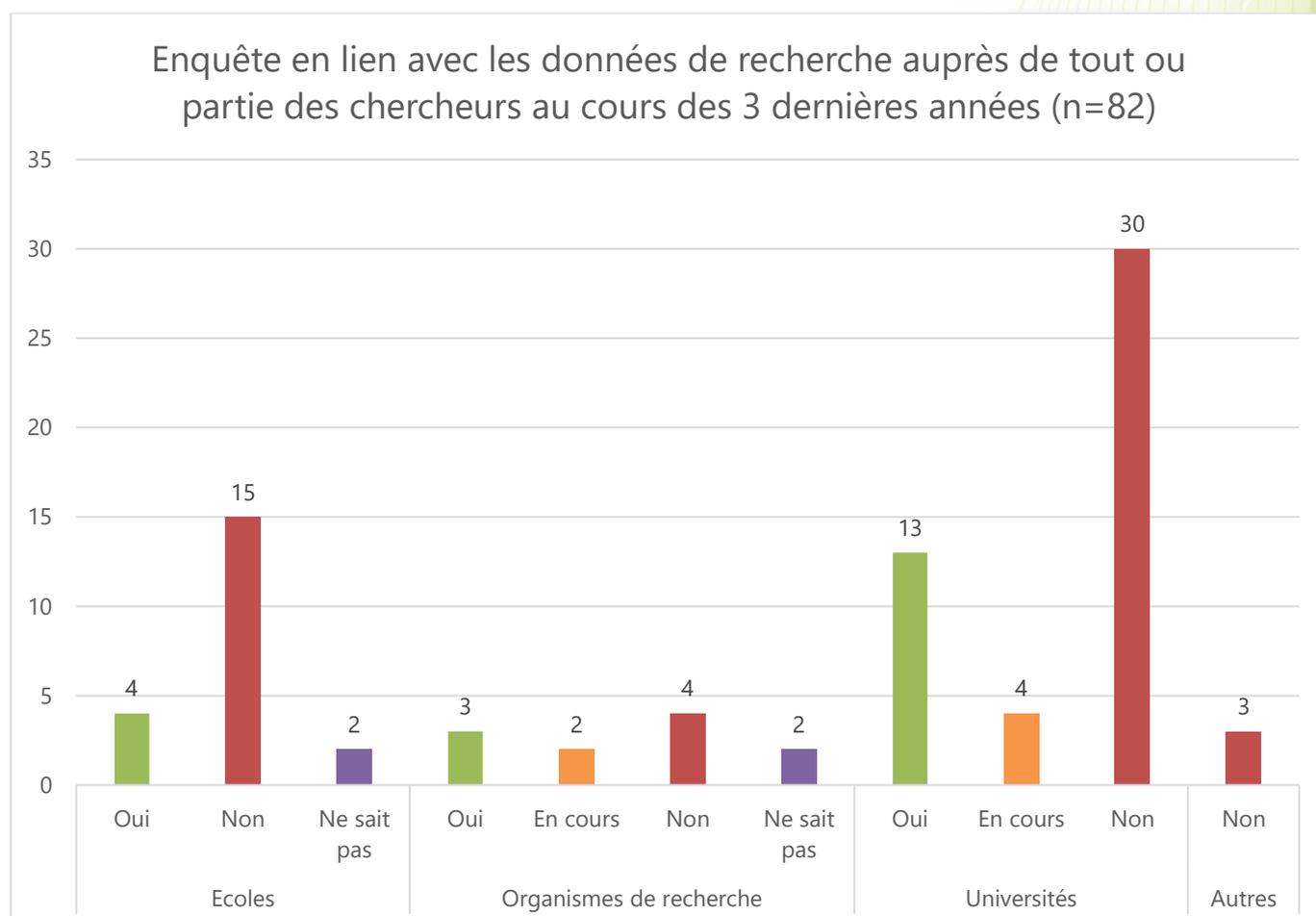


Le nombre indiqué dans chaque case correspond au nombre d'établissements ayant déclaré animer ce type d'action. 24 établissements sur 82 ont déclaré n'effectuer aucune action de formation ou de sensibilisation. Quel que soit le type d'établissement, la question de la sensibilisation aux données de la recherche est abordée sous des angles multiples : formations, ateliers, événements, communication, etc.

La plupart des répondants ont déclaré animer plusieurs actions différentes. On constate que les **ateliers à destination des chercheurs** sont plébiscités (en rouge, 32 des 58 établissements (55%) ayant répondu positivement) ainsi que les **événements** (en vert, 31) et la **formation doctorale** (en bleu, 30). Cette variété d'actions menées indépendamment du type d'établissement est également constatée en analysant les réponses au prisme de la taille de l'établissement, puisque tous les types d'actions sont présents dans toutes les catégories.

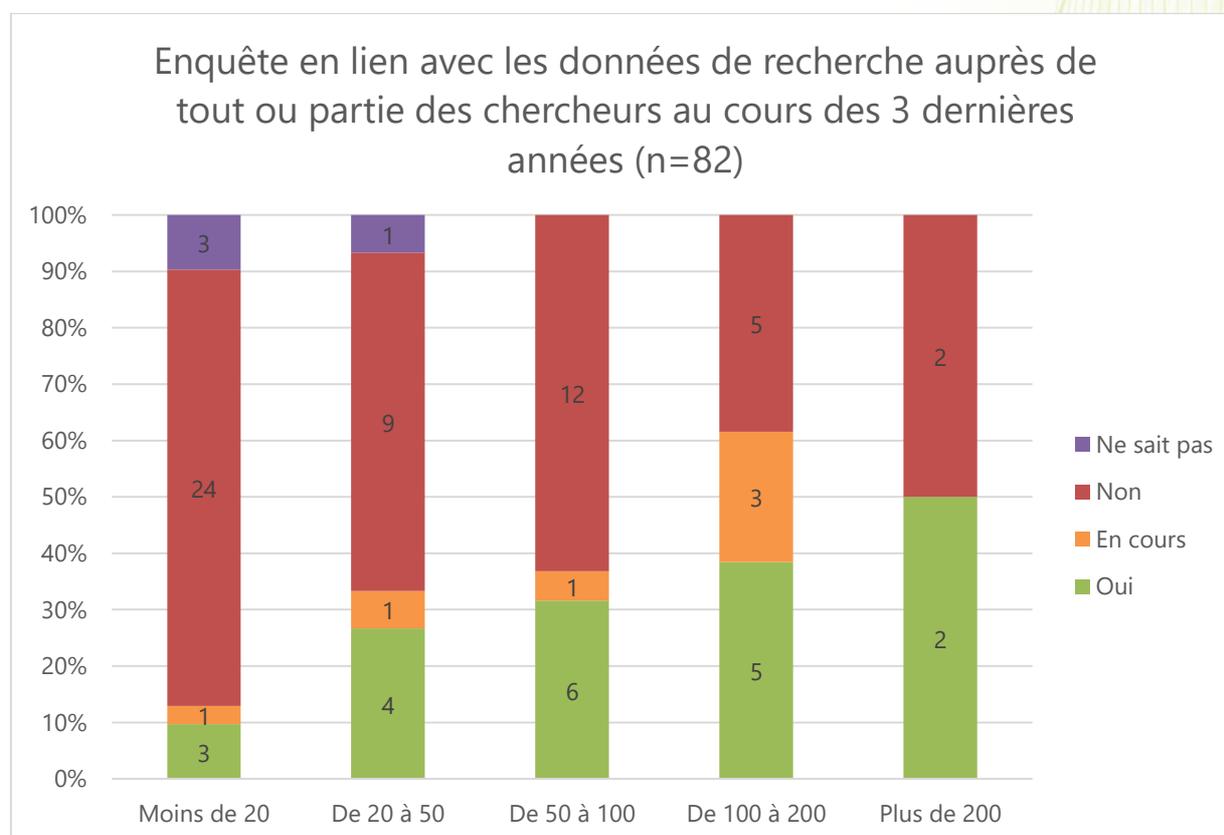
Les autres actions non listées dans les questions signalées par les établissements sont minoritaires, mais méritent d'être signalées : hackathon, datathon, ou encore interventions au sein de commissions Recherche.

5.2. Avez-vous mené une enquête en lien avec les données de recherche auprès de tout ou partie des chercheurs de votre établissement au cours des 3 dernières années ? – Q26



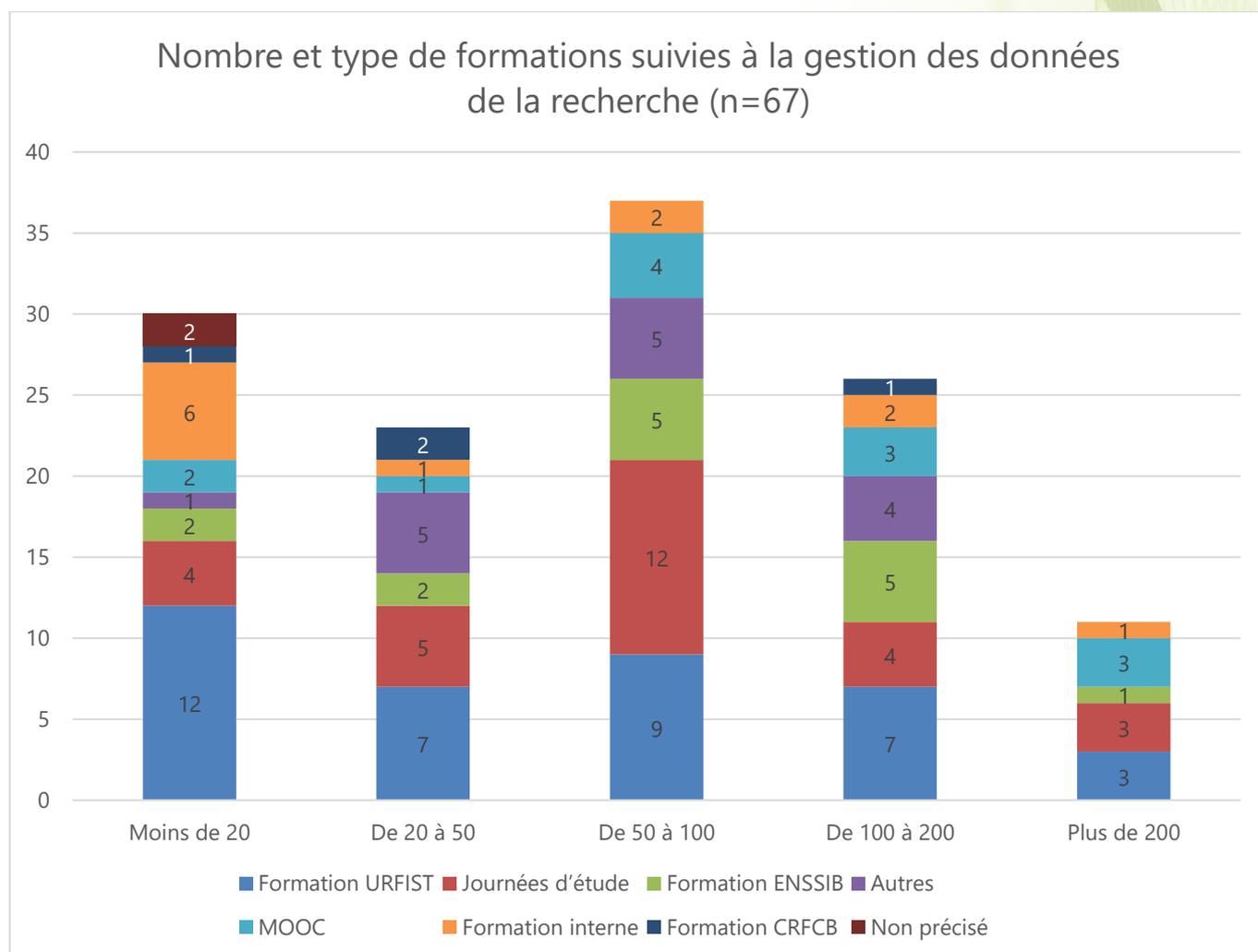
Les enquêtes sur les données de la recherche sont loin d'être majoritaires au sein des établissements répondants ; cela ne semble donc pas présider à la mise en place d'une offre de service d'accompagnement aux données. En effet, les enquêtes de ce type nécessitent de **disposer d'un nombre suffisamment élevé d'agents** à mobiliser pendant plusieurs mois, tant pour l'élaboration du questionnaire et sa diffusion que pour l'analyse des réponses. Le graphique suivant témoigne de la corrélation directe entre taille de l'établissement et organisation d'une enquête de cette envergure.

Par ailleurs, plusieurs établissements ont signalé les résultats de leurs enquêtes sur le site Science Ouverte de Couperin¹⁴. Chacune de ces enquêtes a mobilisé les services pendant plusieurs mois.



¹⁴ <https://scienceouverte.couperin.org/donnees-recherche-enquetes/> ; URL consultée le 22/06/21.

5.3. Au cours des 12 derniers mois, est-ce qu'au moins un agent de votre service a suivi une formation sur la gestion des données de la recherche ? – Q27



Le nombre indiqué dans chaque case correspond au nombre d'établissements ayant déclaré que leurs agents avaient suivi ce type de formation. Seuls 15 services déclarent que leurs agents n'ont suivi aucune formation. La plupart des répondants ont déclaré avoir suivi plusieurs formations différentes. Les formations dans les URFIST sont les plus souvent citées (38 services), ainsi que les journées d'étude (28), ces dernières étant particulièrement adaptées au domaine des données de la recherche qui est en constante mutation. Les CRFCB¹⁵ sont moins identifiés sur cette question à ce stade (4 services). Parmi les

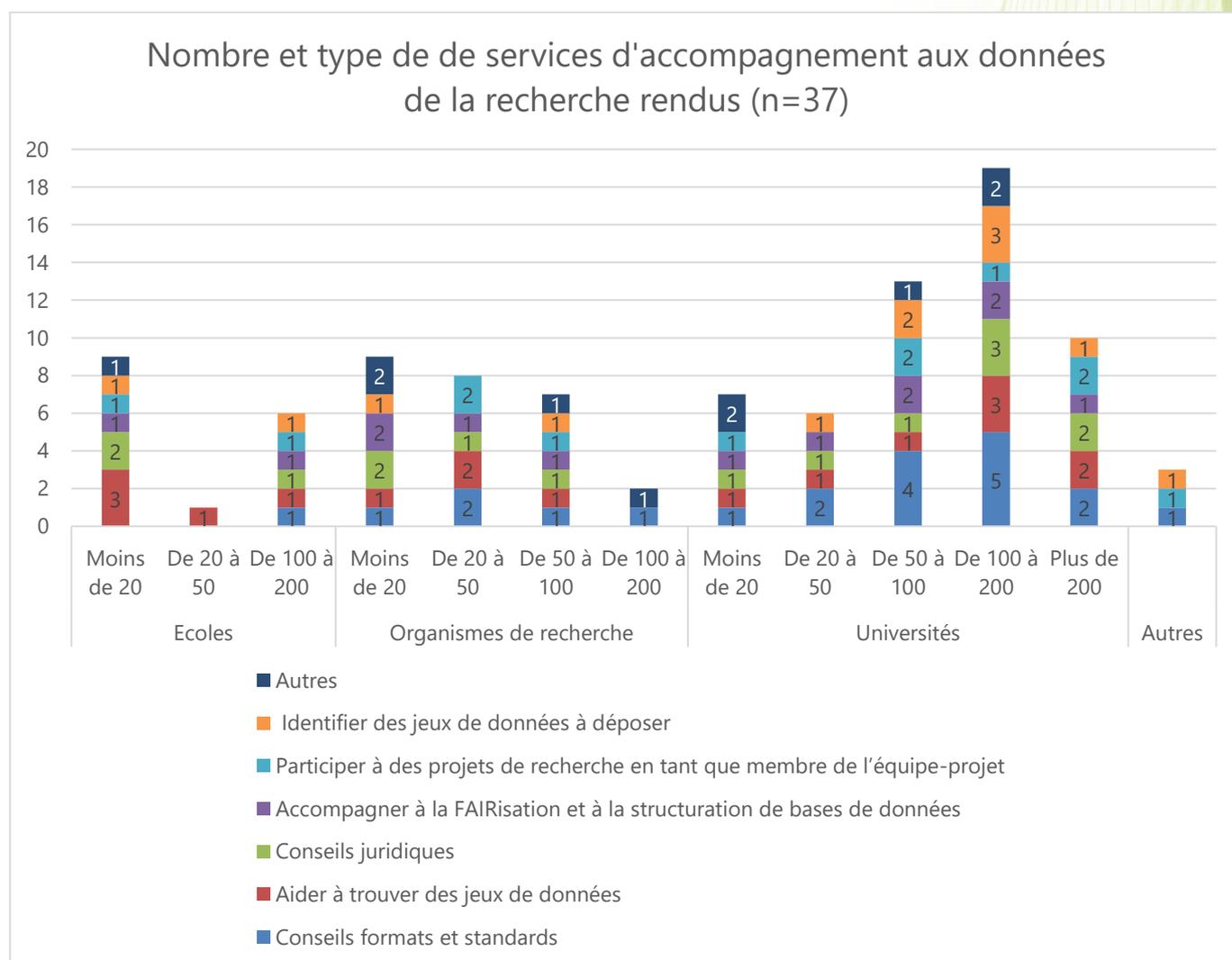
¹⁵ Centres régionaux de formation aux carrières des bibliothèques.

autres canaux de formation (en violet), on retrouve l'Opidor Tour et les actions de formation menées par la Research Data Alliance¹⁶.

Quelle que soit la taille des établissements, la palette de formations suivies est large et témoigne d'une montée en compétences globale sur la question. On constate que les plus petits services font davantage appel à la formation interne, probablement dans une logique de mutualisation des compétences plus systématique que dans les établissements plus importants.

6. Section 5 de l'enquête : Services complémentaires

6.1. Parmi ces services d'accompagnement, lesquels avez-vous déjà rendus ? – Q28

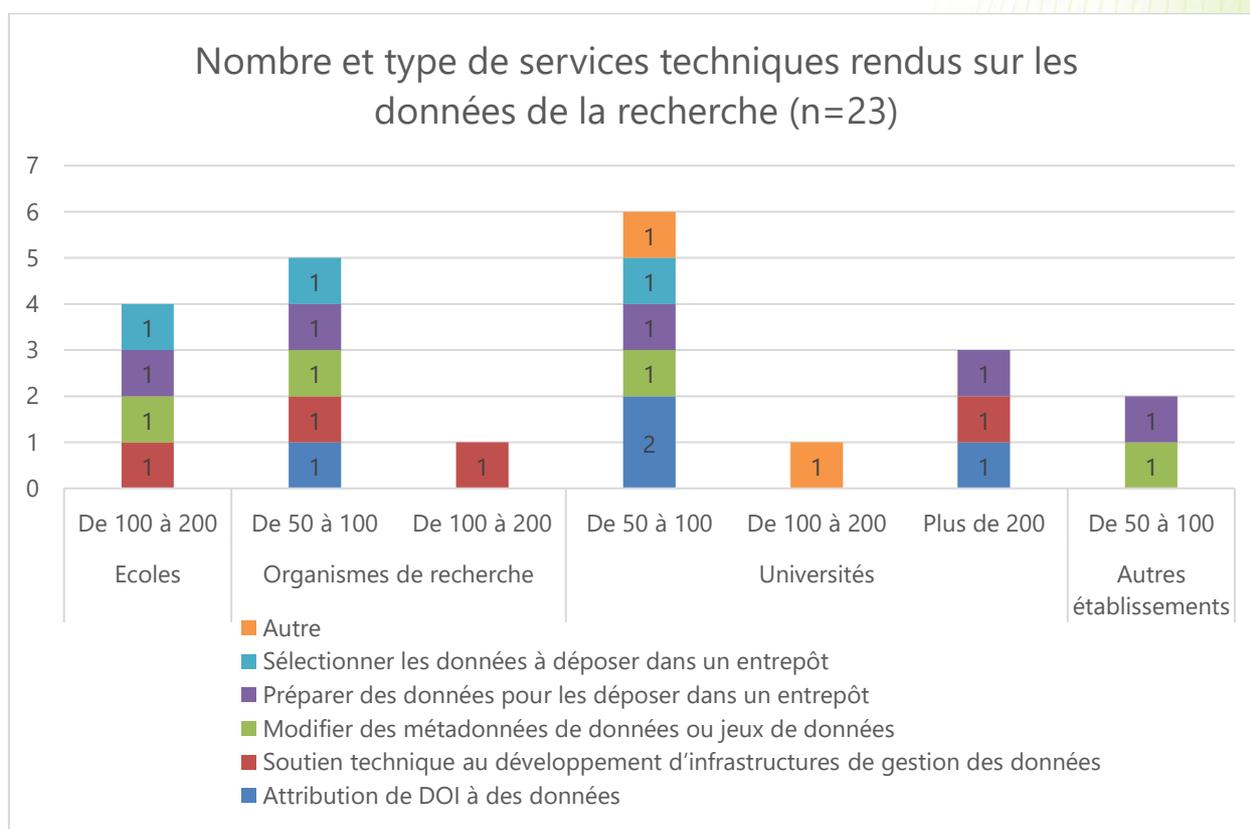


¹⁶ Voir : <https://rd-alliance.org/> ; URL consultée le 22/06/21.

45 établissements, soit près de la moitié des répondants, déclarent n'avoir encore rendu aucun service d'accompagnement relatif aux données de la recherche. La montée en compétences évoquée dans la question précédente laisse supposer que le sujet, encore relativement nouveau pour les établissements, va prendre peu à peu de l'importance et se matérialiser dans les années à venir par des actions concrètes.

Les services d'accompagnement les plus rendus sont les **conseils sur les formats et standards de métadonnées** (21/37 déclarent rendre ce service), la **recherche de jeux de données** (17 établissements) et, plus étonnamment car cela entre moins dans la catégorie des services standards rendus traditionnellement par les services de documentation, les **conseils juridiques** (15 établissements). On observe ainsi l'existence d'une palette très variée de services d'accompagnement, même dans les établissements de taille modeste. Les autres services rendus et non détaillés dans le graphique sont le conseil sur un entrepôt pour publier ses données ou encore la rédaction d'un paragraphe concernant le plan de gestion de données dans le cadre d'une réponse à un appel à projet.

6.2. Parmi ces services techniques, lesquels avez-vous déjà rendus ? – Q29



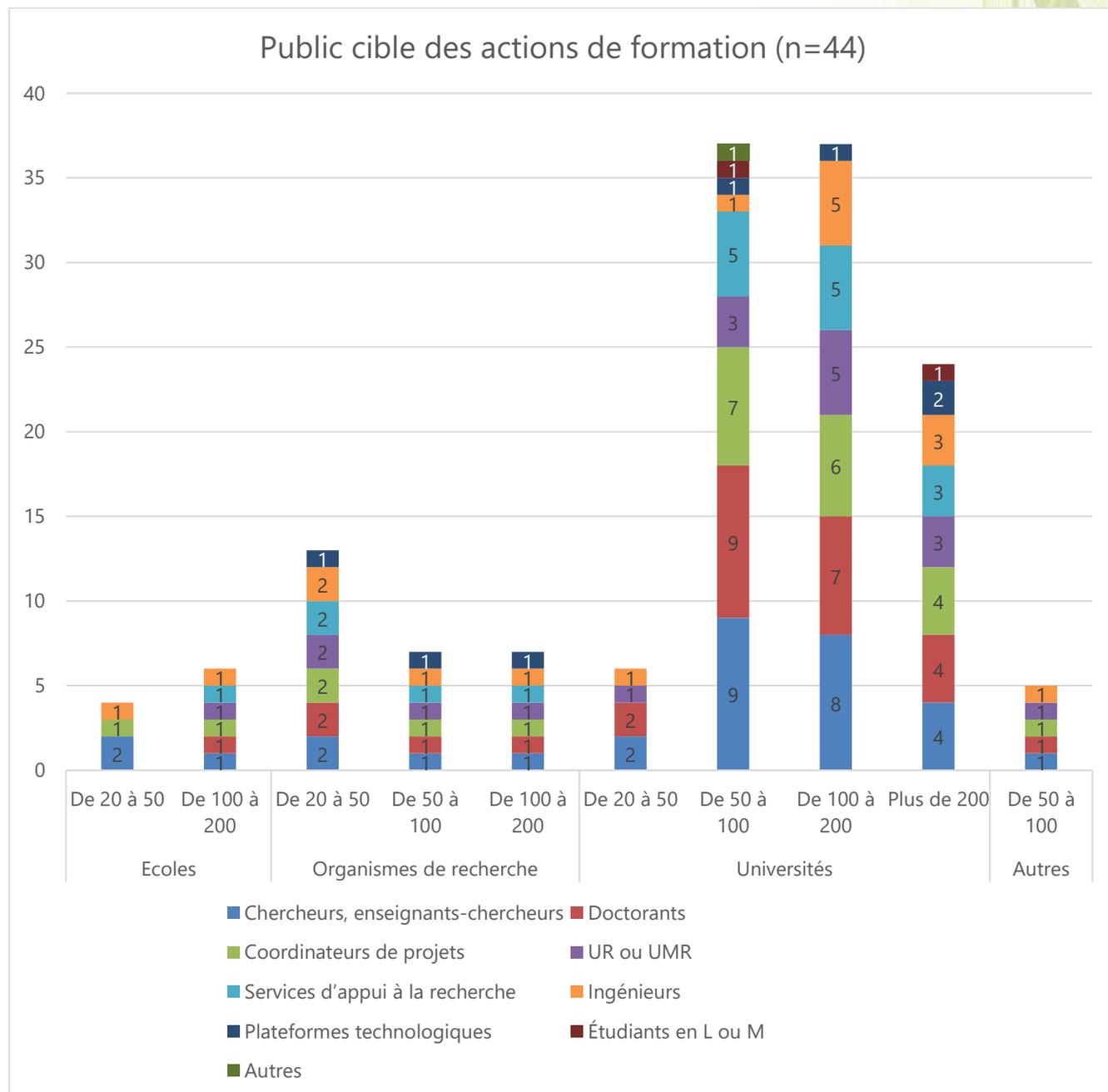
La grande majorité des répondants (59/82) déclare ne rendre aucun service technique autour des données de la recherche. Ici, la taille de l'établissement est clairement un critère, puisqu'aucun service de moins de 50 agents n'est présent dans ce graphique. Les services techniques sont donc encore assez limités, et consistent le plus souvent à la préparation de données pour les déposer dans un entrepôt (5 établissements), l'attribution de DOI à des jeux de données, le soutien technique au développement d'infrastructures de gestion des données et la modification de métadonnées (4 établissements rendent ces trois services).

Le travail de manipulation directe des données, type de mission confié notamment à des *data scientists*, est donc minoritaire au sein des services d'accompagnement. Si le service d'attribution de DOI a vocation à être centralisé au niveau de l'établissement (notamment pour les raisons de suivi budgétaire) et peut être mis en place relativement facilement, les autres services techniques nécessitent des compétences allant au-delà de la simple sensibilisation et sont fortement corrélés à la présence, ou non, d'infrastructures numériques adaptées.

Avec l'émergence des entrepôts de données institutionnels et la mise en ligne prochaine de la plateforme nationale fédérée de données Recherche Data Gouv, il est probable que ces types de services soient de plus en plus développés au sein des services d'appui à la recherche, notamment par la mise en place des ateliers de la donnée et des centres de compétences.

7. Section 6 de l'enquête : Publics

7.1. À quels types de publics destinez-vous actuellement vos services en matière de gestion des données ? – Q30

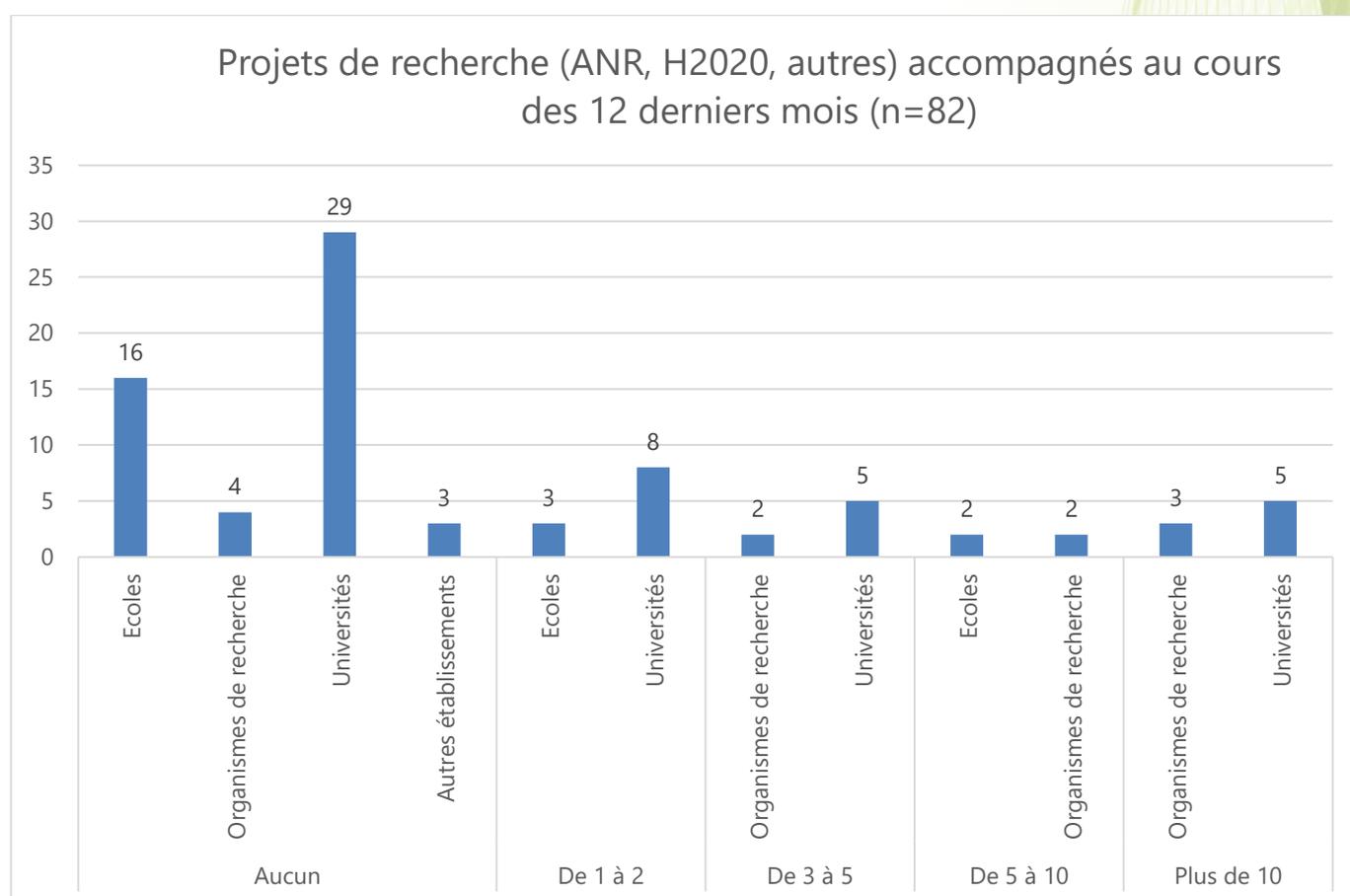


38 établissements ont déclaré ne pas proposer de service d'accompagnement spécifiquement dédié à un public. **Le public cible des formations est majoritairement constitué de chercheurs (31/44), de doctorants (28 établissements) et de coordinateurs de projets (24 établissements).** La question des données de la recherche est très rarement abordée au niveau Licence ou Master (2 établissements

seulement). La formation des personnels d'appui à la recherche est également répandue (18 établissements). Quelle que soit la taille de l'établissement considéré, on constate une certaine homogénéité des approches au sein des différents types d'établissements, qui adressent leurs services à un public varié.

8. Section 7 de l'enquête : Indicateurs

8.1. Au cours des 12 derniers mois, combien de projets de recherche (ANR, H2020, autres) avez-vous accompagné ? – Q31

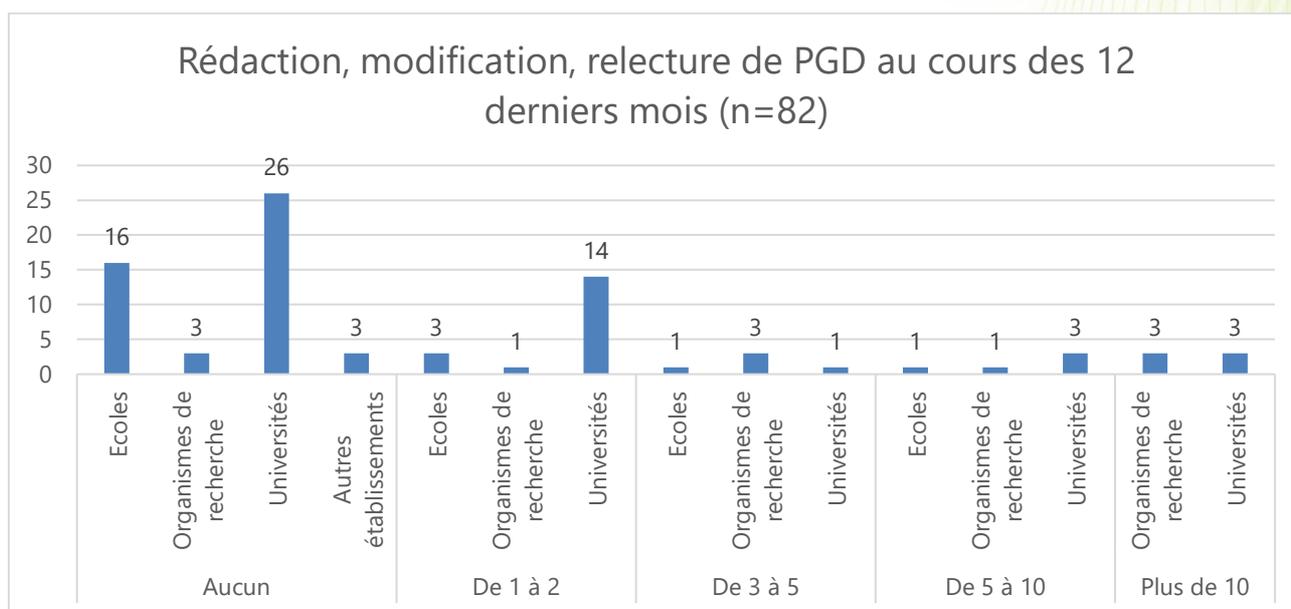


L'accompagnement des projets de recherche est offert par 30 services de documentation, soit les deux tiers de ceux qui accompagnent la gestion des données. Le graphique ci-dessus montre que cette activité ne s'est pas encore généralisée puisque seulement 8 services sur 30 ont accompagné plus de 10 projets, 5 universités et 3 organismes de recherche. Seule une minorité de services de documentation accompagne

donc régulièrement des projets. 52 services répondent qu'ils n'ont accompagné **aucun projet au cours des 12 derniers mois**.

37/82 services ont déclaré dans cette enquête ne pas accompagner à la gestion des données de la recherche. Si l'on regarde uniquement ces établissements, **deux répondants ont quand même accompagné des projets de recherche** : 1 à 2 (une université) pour l'un et 5 à 10 pour l'autre (un organisme de recherche). La répartition par taille d'établissement n'est pas significative, on peut noter cependant que près de la moitié des établissements qui n'ont accompagné aucun projet de recherche sont des établissements de moins de 20 personnes.

8.2. Au cours des 12 derniers mois, sur combien de PGD avez-vous travaillé (rédaction, modification, relecture) ? – Q32



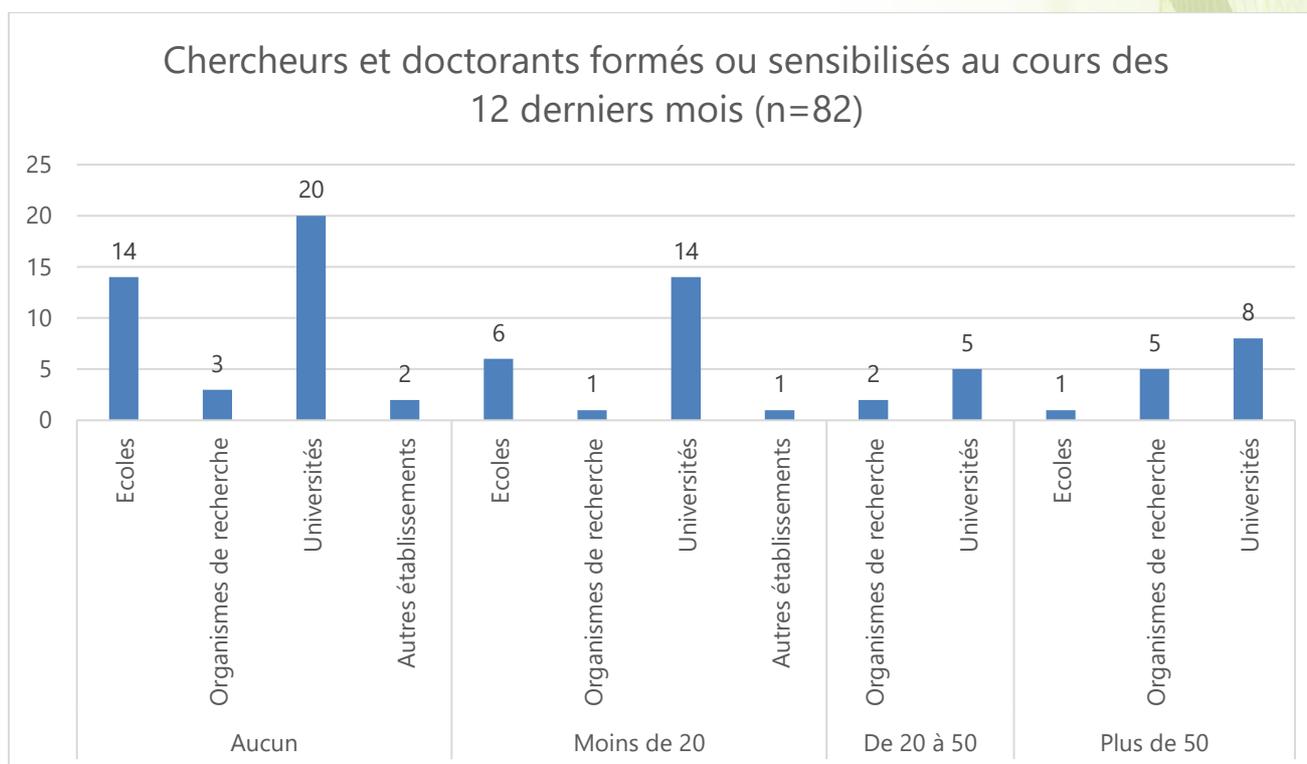
Comme pour la question précédente, **la majorité des établissements (48/82, soit 58.5 %) n'a travaillé sur aucun PGD** au cours des 12 derniers mois.

Le travail sur les PGD a cependant concerné un peu plus d'établissements, puisqu'ils sont 34/82 à déclarer avoir travaillé sur au moins un PGD au cours des 12 derniers mois, 6 établissements ayant travaillé sur plus de 10 PGD (3 organismes de recherche et 3 universités). Si l'on s'intéresse aux 37 établissements ne

proposant pas d'accompagnement à la gestion des données (Q.3), on note quand même qu'une université et un organisme de recherche déclarent pourtant avoir travaillé sur 1 à 2 PGD.

Enfin, la répartition par taille d'établissement permet là aussi de remarquer que la majorité des services de documentation de moins de 20 personnes (24/32) n'a accompagné aucun PGD.

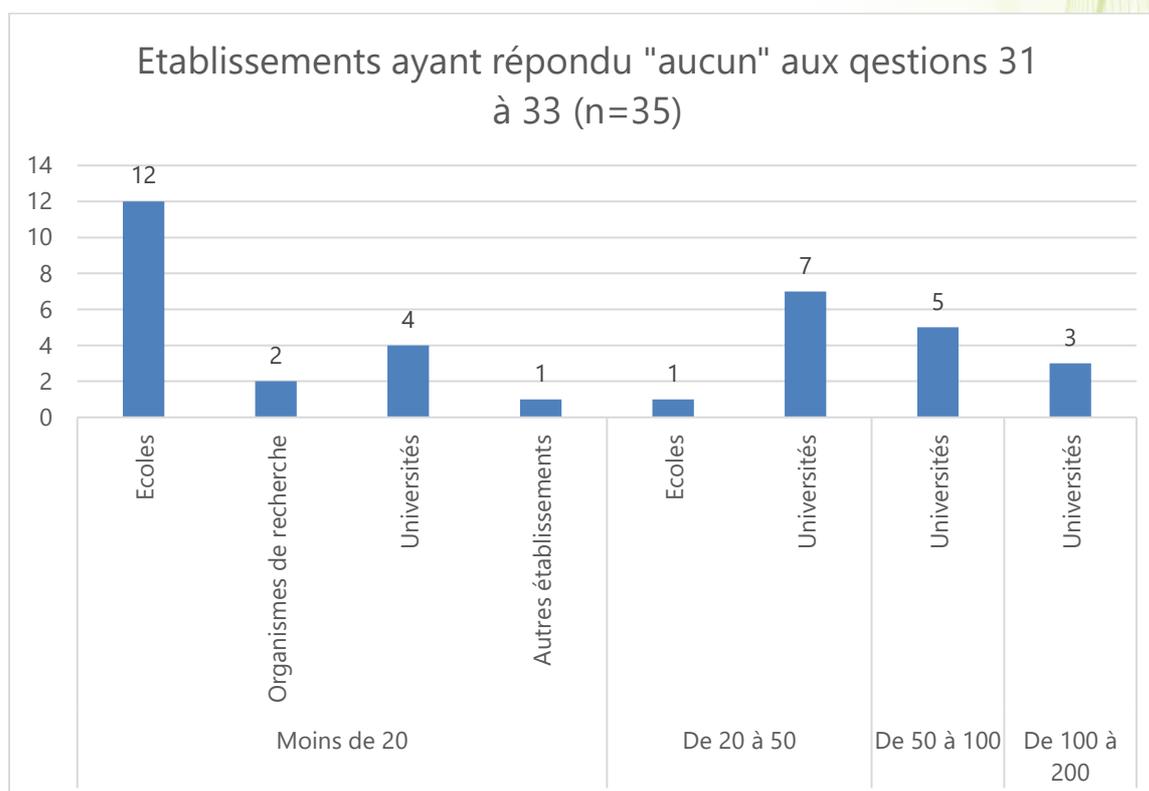
8.3. Au cours des 12 derniers mois, combien de chercheurs et doctorants avez-vous formés ou sensibilisés ? – Q33



La formation et la sensibilisation des doctorants semblent être un domaine mieux maîtrisé par les répondants, **plus de la moitié des services de documentation** (43/82) assure des formations ou sensibilise les chercheurs et les doctorants à la gestion des données de la recherche. Parmi ceux n'ayant formé aucun chercheur à ces questions se trouvent 20/46 universités et 14/21 écoles. Là encore, ce sont **majoritairement des services documentaires de moins de 20 agents** (12 écoles, 4 universités et 3 organismes de recherche, 1 établissement « autre ») qui ne proposent aucune formation. En revanche le type des établissements ne semble pas être un facteur discriminant parmi les 14 établissements ayant formé plus de 50 personnes (8 universités, 5 organismes de recherche et 1 école).

Du côté des 37 services de documentation sans cellule d'accompagnement à la gestion des données de la recherche, 5 ont formé moins de 20 étudiants ou doctorants (2 universités, 2 écoles et 1 autre établissement), 1 organisme de recherche a formé 20 à 50 personnes.

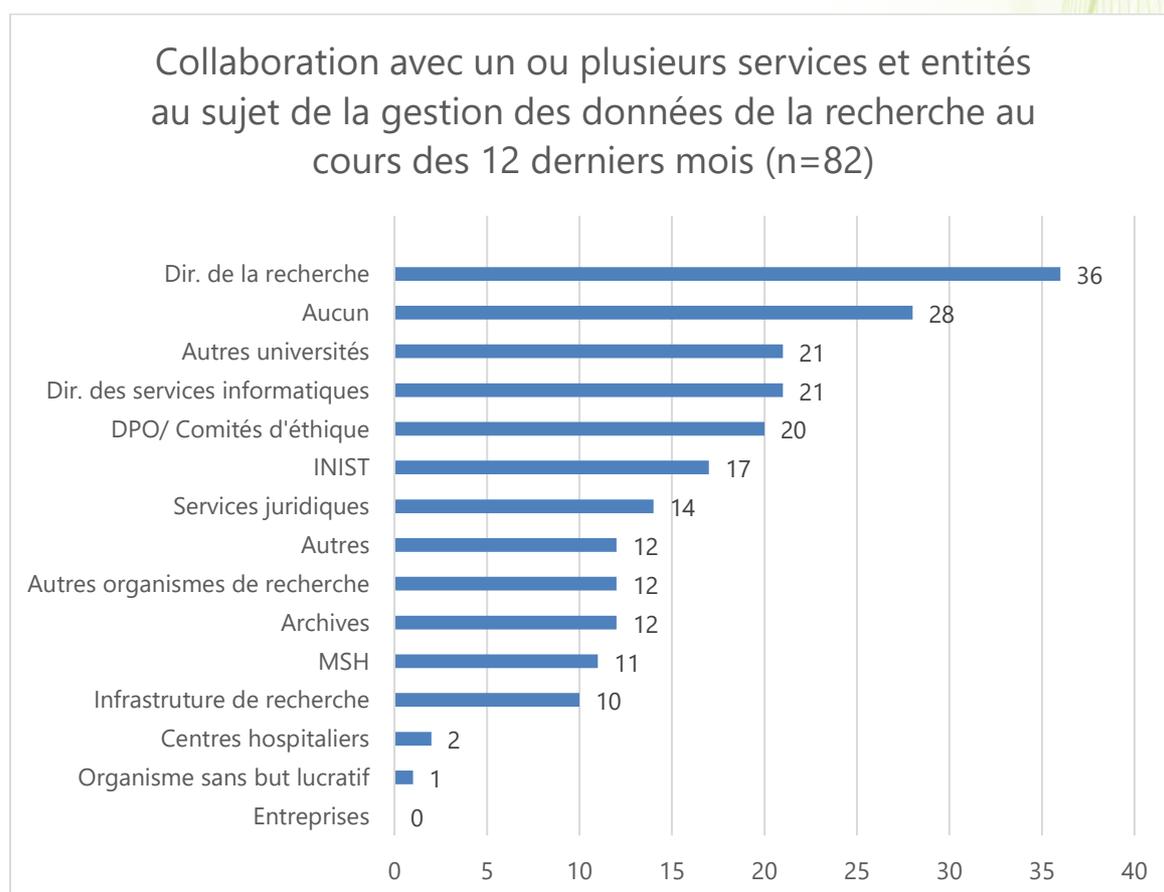
Un grand nombre de services (35/82) ne propose encore aucun des trois services examinés ici et a donc répondu « aucun » aux trois questions. Il s'agit de 13 écoles, 2 organismes de recherche, 19 universités et 1 établissement de la catégorie « autres ». On retrouve ici aussi majoritairement des services de moins de 20 personnes (19/35).



Parmi les services ayant répondu « aucun » aux trois questions, 6 universités et 2 organismes ont pourtant indiqué avoir entre 2 et 4 personnes travaillant à l'accompagnement à la gestion et la diffusion des données de la recherche. Ils apportent probablement d'autres services ou peuvent avoir assimilé l'accompagnement à la gestion des données de la recherche au service plus large que de nombreux établissements rendent déjà, à savoir l'appui à la recherche, qui concerne plus la publication des résultats de la recherche que la gestion des données elles-mêmes.

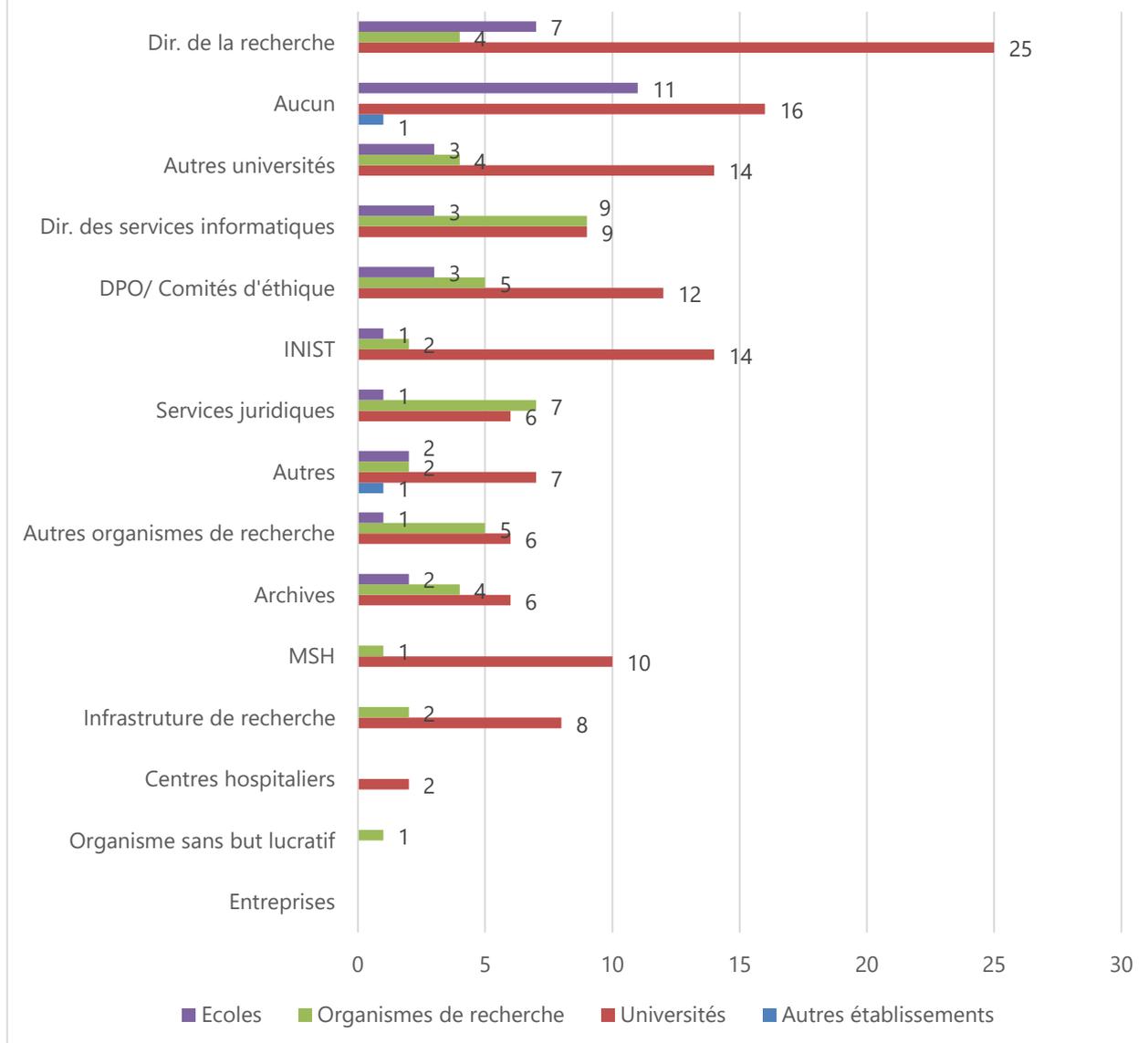
9. Section 8 de l'enquête : Collaborations internes ou externes – Q34 et 35

9.1. Au cours des 12 derniers mois, avez-vous collaboré avec un ou plusieurs des services et entités suivants au sujet de la gestion des données de la recherche ? (Plusieurs réponses possibles) – Q34



Seulement 6 services sur les 45 qui accompagnent la gestion des données n'ont collaboré sur cette question avec aucun autre service. A l'inverse, 13 services ne proposant pas d'accompagnement collaborent néanmoins sur la question de la gestion des données avec d'autres services, principalement la direction de la recherche (5/13) ou d'autres universités (3/13). Sans surprise, donc **c'est avec la direction de la recherche que les répondants collaborent le plus**, (36 sur 45). Viennent ensuite la collaboration avec la DSI ou une autre université pour 21 répondants, et les DPO ou les comités d'éthique pour 20 répondants. La répartition par type d'établissement fait apparaître des partenaires plus spécifiques aux universités : l'Inist, les MSH ou les infrastructures de recherche.

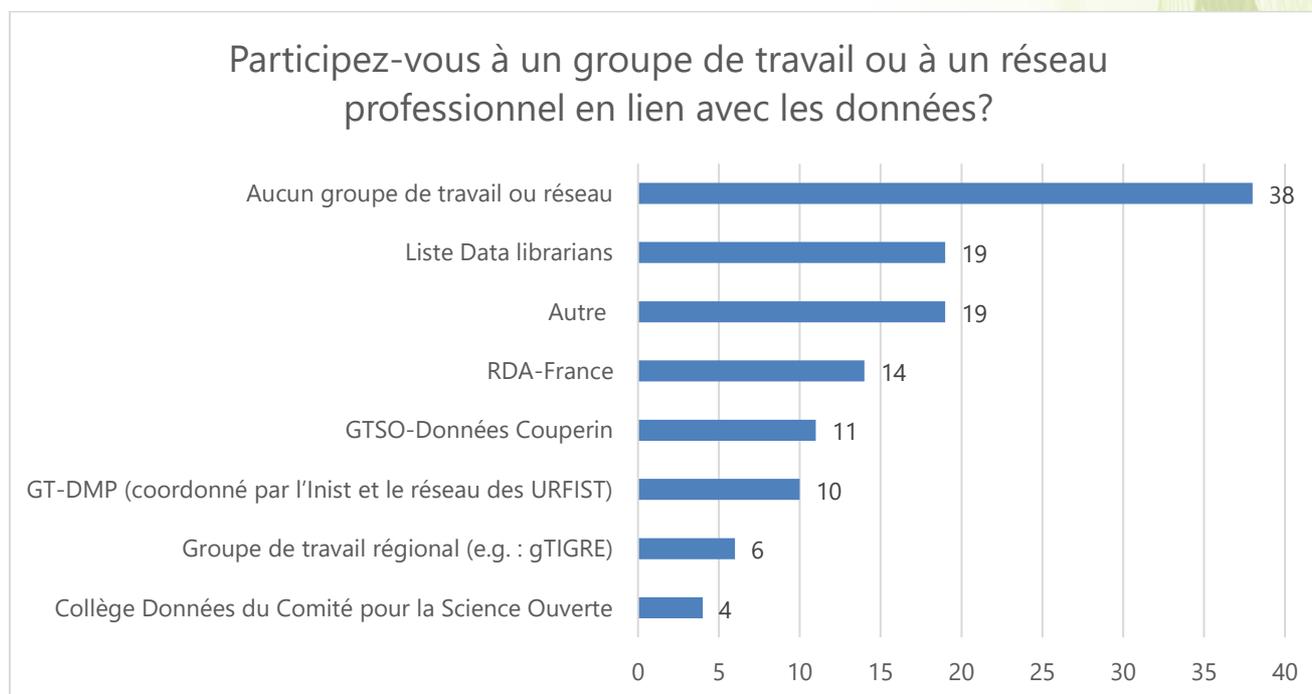
Au cours des 12 derniers mois avez-vous collaboré avec un ou plusieurs services et entités au sujet de la gestion des données de la recherche?(Par type d'établissement)



9.2. Participez-vous à un groupe de travail ou à un réseau professionnel en lien avec les données ? (Plusieurs réponses possibles) – Q35

Près de la moitié des répondants (38 sur 82) ne participe à aucun groupe de travail ou réseau sur la gestion des données. La liste Data librarians est la plus citée (19 fois). Dans la catégorie « Autre », les répondants ont cité 5 fois des groupes de travail internes, trois fois l'ANRT

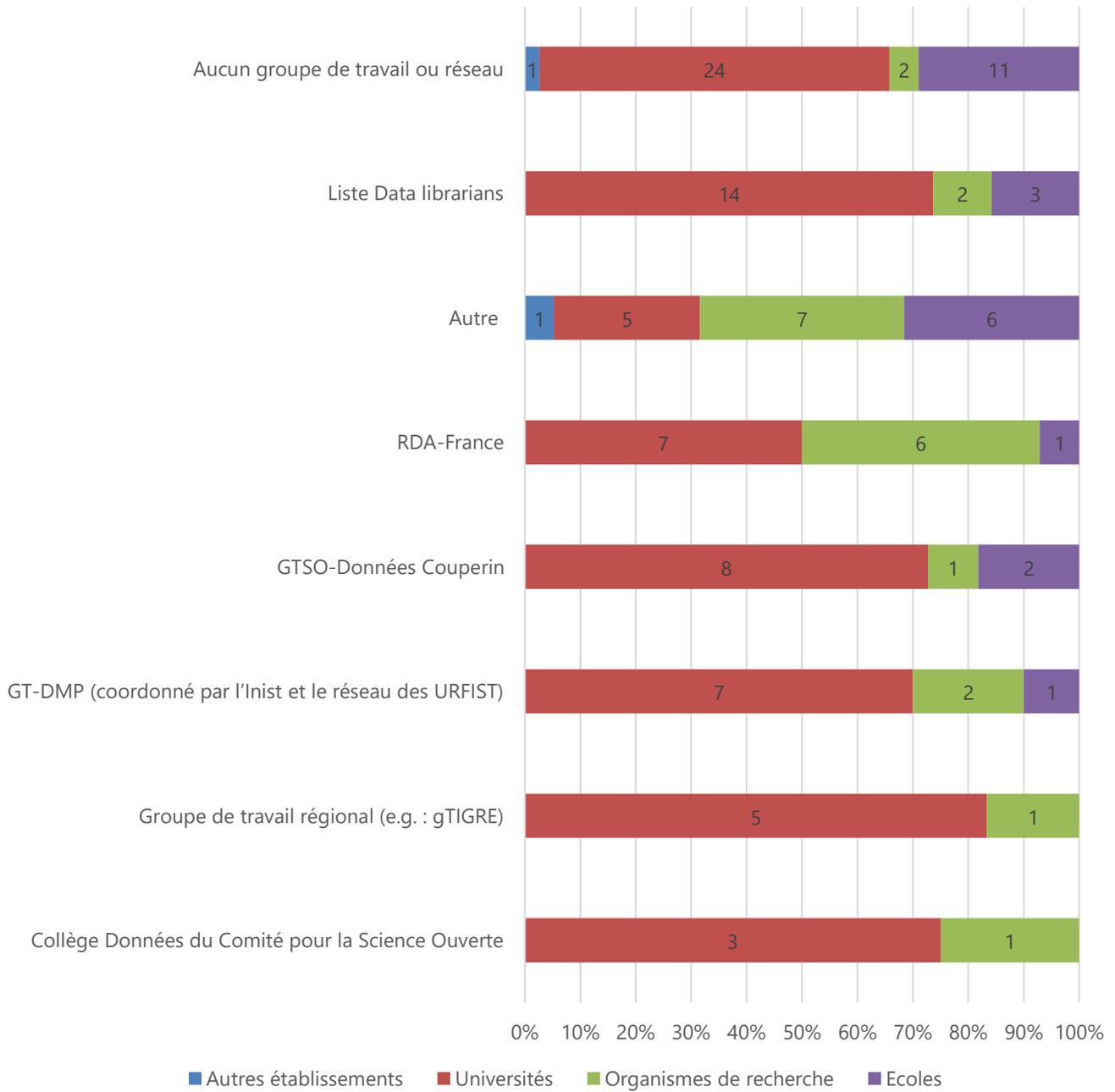
Data group (Association Nationale Recherche Technologie). Sont aussi cités le réseau MITI du CNRS, EOSC, la section Aurore de l'AAF, Toul'AO, Casuhal,...



La répartition par type d'établissements montre que les écoles participent à moins de groupes que les organismes de recherche ou les universités. Les écoles ne citent que Data librarians, RDA-France, le GTSO et le GT-DMP.

Si l'on considère la taille des établissements, on note que ce sont surtout les petits établissements qui ne participent à aucun groupe de travail. La taille de l'établissement est déterminante pour ce type de participation. Les établissements de plus de 200 personnes qui ont répondu à l'enquête participent tous à au moins un groupe de travail national.

Participez-vous à un groupe de travail ou à un réseau professionnel en lien avec les données ? (par type d'établissement)



10. Section 9 de l'enquête : Projets et besoins – Q36 à 39

Cette dernière section présente une synthèse des projets, difficultés et besoins des établissements **disposant déjà d'un service de gestion des données**. La première question portait sur les projets des services dans les deux prochaines années. Plusieurs dynamiques se dégagent des réponses, parmi lesquelles **la poursuite de l'élaboration et de la structuration de l'offre de service** (guichet unique, site web) en interne et au sein de l'établissement (mise en place de groupes de travail, développement de partenariats), **la sensibilisation des communautés** (pouvant passer par une phase de recueil de besoins) et de **la gouvernance** (charte, politique institutionnelle).

Un deuxième axe vise à développer et intensifier les services proposés : **développer l'expertise autour de l'accompagnement à la rédaction de PGD** (notamment via la rédaction de modèle) ; **intensifier – voire massifier – la formation aux doctorants**, élargir aux étudiants de master ou former d'autres publics ; **développer ou systématiser l'accompagnement des porteurs de projet** et enfin, **proposer une offre plus technique** (expertise et recommandations autour des données associées aux publications, cahiers de laboratoire, offre logicielle).

Une troisième dynamique identifiée porte sur le développement d'outils, qui passe notamment par **l'offre d'une solution institutionnelle d'archivage et de diffusion des données** (entrepôt de données).

Pour développer ces projets, les établissements ayant mis en œuvre la GDR rencontrent diverses difficultés, structurelles, organisationnelles ou politiques. Les difficultés structurelles s'expriment par le **manque de temps pour développer des collaborations, d'expertise ou de visibilité des cellules qui débutent, le manque de moyens** ou de cadrage politique clair (politique institutionnelle, stabilité de la gouvernance). On peut aussi voir apparaître des difficultés d'ordre plus organisationnelle : **multiplicité des acteurs, difficulté à coordonner la cellule et les services de proximité existants** ou la **difficulté à passer d'un service d'accompagnement générique à un service d'accompagnement disciplinaire**.

Ce dernier aspect s'accompagne de la difficulté à pouvoir rendre un « service sur mesure » ou à gagner en efficacité au vu de la très grande hétérogénéité des besoins ou des demandes. La **faible sensibilisation**

ou le manque d'intérêt de certains établissements aux enjeux des données de la recherche et/ou des laboratoires est également mentionnée.

Enfin, la **Science Ouverte** peut également être ressentie comme une imposition « *top down* » par certaines composantes, voire même rejetée, le plan de gestion de données considéré comme une charge administrative lourde et la gestion des données chronophage, voire même inutiles car « les données expérimentales ne seront pas réutilisées par d'autres » selon les chercheurs.

En miroir, les **principaux besoins des cellules d'accompagnement** sont des besoins en **formation permettant une montée en compétence des équipes** (notamment techniques et juridiques), **une structuration de l'offre et une mutualisation des efforts au sein des établissements au-delà des aspects structurels** (comme bénéficier de l'appui d'une politique d'établissement autour de la GDR par exemple).

Conclusion

Cette enquête confirme l'important travail mené par les services de documentation pour contribuer à l'accompagnement des personnels de la recherche en matière de gestion des données. La majorité des hypothèses de travail initiales ont été validées, mais certains points mériteraient une analyse plus approfondie pour en tirer des conclusions, probablement sous forme d'une enquête qualitative.

1. Une minorité non négligeable de services de documentation n'accompagnent pas du tout la gestion des données et n'ont pas de projets en la matière ;

En effet, près de la moitié des services ayant répondu à l'enquête n'ont pas formalisé de services d'appui à la GDR, même si cela n'empêche pas certains d'entre eux d'accompagner ponctuellement leurs chercheurs sur ce sujet. Il est à noter que plus de 80% des services de documentation des écoles et 40% de ceux des universités membres de Couperin n'ont pas répondu à cette enquête. Il est possible de supposer que la part de ces services qui ne disposent pas de service formalisé de gestion des données est au moins équivalente à celle des services qui ont répondu à l'enquête.

2. Une majorité d'établissements sont au début de leur accompagnement en matière de GDR ou encore en phase de conception ;

Au-delà des intentions et de l'offre de services affichés à l'intention des chercheurs, les réponses aux questions 31 à 33 montrent qu'un nombre important de cellules d'appui n'accompagnent les chercheurs que de manière relativement ponctuelle voire anecdotique et ne sont pas identifiées au sein de leur établissement.

3. Il existe un écart important entre les services d'appui à la gestion des données offerts par les organismes de recherche et les universités ;

A taille égale, on constate en effet que les services de documentation des organismes de recherche proposent une offre d'accompagnement à la gestion des données plus importante que ceux des universités, mais le facteur de taille est prépondérant puisque les services de documentation de plus de 100 agents en universités accompagnent largement à la gestion des données, ce que ne font pas les services de documentation de petite taille des organismes de recherche.

4. Il existe un écart important entre les services offerts par les grandes universités et les petites et moyennes universités

Cette hypothèse paraît assez nettement vérifiée, avec un gradient assez clair selon la taille des services de documentation et un effet de seuil entre les services de moins de 50 agents et ceux de plus de 100. Il serait nécessaire d'affiner la typologie par taille pour savoir s'il est possible d'identifier un point de bascule. Par ailleurs, renouveler cette enquête d'ici 3 à 4 ans permettra peut-être de constater que des services de documentation d'une taille de plus en plus petite s'emparent de ces sujets. Dans l'ensemble, le soutien aux services de documentation de petite et moyenne taille devrait être une priorité pour les initiatives de coopération professionnelle afin de réduire les inégalités d'accompagnement des chercheurs selon la taille de leur établissement et ce, quel que soit le type d'établissement (école, université ou organisme de recherche).

5. Seules une minorité d'universités est en capacité de dégager des quotités de travail mesurables et significatives dédiées à la gestion des données ;

Cette hypothèse n'est pas facile à vérifier car les réponses apportées à l'enquête ont montré une grande variété d'interprétation de cette question. Dans certains établissements, la gestion des données est visiblement intégrée à un item général sur la science ouverte et les services aux chercheurs, ce qui a conduit certains services à indiquer un nombre important d'agents travaillant sur ce sujet, avec parfois des quotités associées très faibles. Vu les enjeux de montée en compétences et de maintien de ces compétences, il est probable que 10% d'un ETP consacré à la gestion des données est le strict minimum en dessous duquel il n'est pas possible de considérer qu'un agent est véritablement en mesure de proposer des services sur ce sujet. Ceci mis à part, un nombre plus important que prévu de services de documentation d'université ont un ou des agents identifiés sur ce sujet, ce qui montre une prise de conscience intéressante de l'enjeu de ressources humaines pour développer des services en matière de données. Un travail plus fin sur les missions et les compétences attendues de ces agents permettrait de mieux cerner cette dimension.

6. La situation est aussi très hétérogène entre les services de documentation des organismes de recherche, mais dans une moindre mesure ;

L'effectif concerné (11 organismes de recherche) est trop faible pour en tirer des conclusions générales. Il existe tout de même des écarts importants dans les réponses apportées par les services de documentation

des organismes de recherche et certains ne se sont pas encore emparé de ce sujet du tout. Il serait souhaitable que les logiques de coopération documentaire inter établissements ne négligent pas ce type d'établissement même si certains d'entre eux sont particulièrement en avance par l'étendue des services qu'ils offrent.

7. Les services offerts sont plutôt de type sensibilisation que de nature technique ;

Cette hypothèse est globalement vérifiée. On peut distinguer les actions de sensibilisation et formation qui sont les plus communes pour les cellules d'appui, suivies par des services de conseil et enfin des services plus techniques. Plus d'une trentaine de cellules d'appui proposent de la formation doctorale (30), des ateliers sur la gestion des données (32) ou accompagnent à la rédaction d'un plan de gestion des données (43), tandis que seulement une dizaine apportent des conseils juridiques (15), des conseils sur la citation et la recherche de jeux de données (17) ou sur les formats et les métadonnées (20). Enfin, à l'exception de l'attribution de DOIs (12), moins de dix cellules proposent des services d'appui technique tels que la préparation de jeux de données en vue de leur diffusion (7) ou la création de métadonnées (8). Cependant, ces chiffres ne sont pas anecdotiques pour autant et il sera intéressant de voir si ce type de services se développent à l'avenir.

8. Il y a une corrélation entre l'existence d'une politique institutionnelle en matière de GDR et les services offerts par les services documentaires ;

Etonnamment, cette hypothèse ne se vérifie pas. Au contraire, les services de documentation des établissements qui ont mis en place des politiques de données sont plutôt de petite taille et correspondent donc aux services qui offrent le moins de service d'appui à la GDR. A l'inverse, les établissements de rattachement de nombreux services de documentation de grande taille (plus de 100 agents) n'ont pas encore mis en place de politique de données, ce qui n'empêche pas leur cellule d'appui à la GDR d'avoir développé des offres de service ambitieuses. Ce sujet mériterait d'être creusé, mais une hypothèse peut être que l'émergence d'un consensus sur une politique de données est plus longue dans des établissements de grande taille, surtout s'ils sont multidisciplinaires.

9. La gestion des données nécessite de nombreuses formes de coopération internes et externes.

Sans surprise, cette hypothèse a été confirmée, aussi bien pour les coopérations internes qu'externes. Dans les établissements où les services de documentation peinent à élaborer une offre de services d'appui à la gestion des données, les difficultés de coopération interservices sont un des freins principaux. Il y a donc un enjeu particulier à mettre en place des espaces d'échange interprofessionnels au sein des établissements pour favoriser des coopérations entre les différentes parties prenantes de la gestion des données.

En guise de conclusion, il est possible de retenir qu'il existe aujourd'hui un fort enjeu de ressources humaines pour accompagner la gestion des données, qui passe par un développement des compétences et l'augmentation des quotités de travail dédiées à la gestion des données. Il est aussi nécessaire que les gouvernances s'emparent de ces sujets dans les établissements et impulsent des dynamiques de coopération interservices. Sur ce point, il serait intéressant de développer les partages d'expérience d'établissements ayant mis en place ce type d'approche. D'une manière générale, mutualiser les expériences et les efforts semble être un levier essentiel pour accélérer le développement de l'accompagnement des chercheurs, en particulier dans les établissements de petite taille.

Pour le GTSO-Données, cette enquête a été l'occasion d'ouvrir un nouvel axe de travail, à savoir la mise à disposition de fiches pratiques à destination de la communauté professionnelle. En effet, les difficultés exprimées par les répondants, notamment par le biais des verbatims, ont été la source d'inspiration principale pour ces fiches, dont les statistiques de téléchargements (entre 320 et 760 par fiche à ce jour) soulignent la portée auprès des établissements. Ce sont donc 4 fiches pratiques¹⁷ qui ont déjà été produites pour répondre aux questionnements signalés dans l'enquête :

- Comment sensibiliser les chercheur.e.s en SHS à la gestion des données ?
<https://doi.org/10.5281/zenodo.4627817>
- Comment former les doctorant.e.s aux données de la recherche ?
<https://doi.org/10.5281/zenodo.4627784>

¹⁷ L'ensemble des productions du GTSO-Données est à retrouver sur la communauté Zenodo Couperin : <https://zenodo.org/communities/couperin>

- Comment animer un atelier plan de gestion de données ?
<https://doi.org/10.5281/zenodo.4719277>
- Comment monter un service d'accompagnement aux données ?
<https://doi.org/10.5281/zenodo.4890409>

D'autres fiches pratiques seront élaborées pour poursuivre ce travail. Par ailleurs, le GTSO-Données continuera à organiser des cycles de visioconférences¹⁸ pour accompagner au mieux les établissements.

¹⁸ Pour visionner les visioconférences passées, rendez-vous sur la chaîne Youtube du GTSO-Données : <https://www.youtube.com/channel/UCswnouUXeMhHSeZNpYGhMA>

Annexe – FAQ accompagnant l'enquête

Qui mène cette enquête ?

Le Groupe Données du GTSO (Groupe de Travail Science Ouverte) de Couperin est à l'initiative de cette enquête. Le GTSO-Données travaille sur les modalités d'accompagnement que les services de documentation et d'information scientifique et technique proposent aux personnels de recherche en matière de données. Pour toute question : gsto_donnees@listes.couperin.org

A quoi sert cette enquête ?

L'enquête a pour objectif de disposer d'un état des lieux des services en matière de gestion des données de recherche proposés par les services de documentation et d'information scientifique et technique. Elle pourra être reproduite dans le temps afin de mesurer l'évolution de ces services. C'est aussi un moyen pour le groupe Données du GTSO d'identifier des axes de travail pour appuyer les initiatives menées par les établissements membres de Couperin.

A qui est destinée cette enquête ? Qui peut y répondre ?

Tous les établissements d'enseignement supérieur et de recherche sont invités à y répondre. Une seule réponse sera acceptée par établissement. Les réponses aux questions pourront être plus facilement apportées par les responsables des services liés aux données de recherche ou par les responsables des services aux chercheurs et à la recherche.

Les établissements qui commencent à travailler sur les données de recherche ou qui n'ont pas encore engagés de démarche de ce type sont fortement encouragés à répondre pour contribuer à la représentativité de l'enquête et faire part de leurs difficultés, réticences et besoins sur ce sujet.

Comment répondre à l'enquête ?

L'enquête est mise à disposition sur SurveyMonkey : <https://fr.surveymonkey.com/r/MVPHZYZ>. Elle est ouverte du mardi 8 septembre au dimanche 27 septembre 2020. Cette période pourra être étendue si nécessaire. L'enquête est structurée en dix sections (2 à 5 questions par section) et peut être remplie en 10 à 15 minutes.

Comment seront diffusés les résultats ?

Les données obtenues dans le cadre de l'enquête et les analyses complémentaires seront diffusées en libre accès. En fonction des réponses obtenues, cette enquête pourra donner lieu à quelques entretiens complémentaires dont les comptes rendus seront diffusés en cas d'accord des participants.