



RÉPUBLIQUE
FRANÇAISE
*Liberté
Égalité
Fraternité*

recherche.data.gouv.fr

Résultats de l'enquête sur les pratiques de publication de data papers

Cathy Tang, Mouhamed Gassama (INRAE, DipSO)
Présentation C@fé Rénatis - Novembre 2024



Contexte

Enquête et méthodologie

Résultats de l'enquête

Conclusion

**Résultats de l'enquête sur
les pratiques de
publication de data
papers**





RÉPUBLIQUE
FRANÇAISE recherche.data.gouv.fr

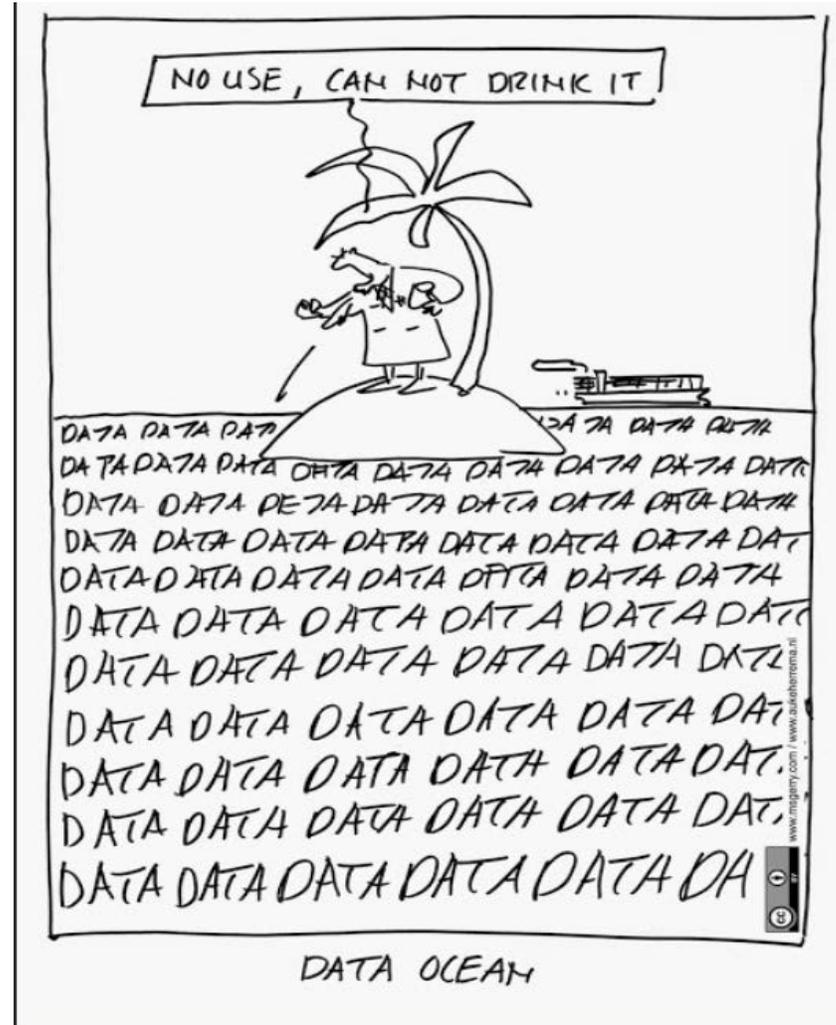
*Liberté
Égalité
Fraternité*

Contexte



Un peu de métaphore

- Impossibilité pour l'homme de boire l'eau de mer
- Océan de données
- Données de recherche nécessitent traitement pour être réutilisées



Les Data papers

Les data papers sont des articles rédigés, révisés par des pairs et cités dans des revues académiques ou savantes, dont le contenu principal est une **description des jeux de données de recherche** publiés, ainsi que des **renseignements contextuels sur la production et l'acquisition des données**, dans le but de faciliter l'accessibilité, la disponibilité et **la réutilisation des données de recherche (...)**

Joachim Schöpfel, Dominic J. Farace, Hélène Prost, Antonella Zane. Data papers as a new form of knowledge organization in the field of research data. Knowledge Organization, 2020, 46 (8), pp.622-638. (10.5771/0943-7444-2019-8-622). (hal-02994295)



Image: Herrema, A. (2014): FOSTER Cartoon: Data for Future Generations. FOSTER-Projekt (EU funding: 612425).

Disponible sur : <https://www.fosteropenscience.eu/foster-taxonomy/research-data-management>, 25.08.2015. (CC-BY)

Où publier des data papers ?

Dans des revues dédiées : Data journals

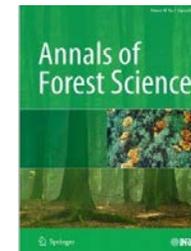
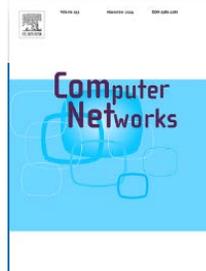


Où publier des data papers ?

Dans des revues dédiées : Data journals



Dans des revues dites traditionnelles



Modèles proposés par les revues

Templates for your manuscript file

Technical instructions for LaTeX ▶

Technical instructions for MS Word and compatible formats ▶

Technical instructions for R Markdown ▶

<https://www.earth-system-science-data.net/submission.html#templates>

Title of the manuscript

Angie A. Smith¹, B. Bradley Miller¹, Carl Carter²

¹Department example, University example, city, postal code, country

²Laboratory example, city, postal code, country

5 Correspondence to: Angie A. Smith (angie@email.com)

Abstract. Please use only the styles of this template (MS title, Authors, Affiliations, Correspondence, Normal for your text, and Headings 1–3). Figure 1 uses the style Caption. The abbreviation “Fig.” should be used when it appears in running text and should be followed by a number unless it comes at the beginning of a sentence, e.g.: “The results are depicted in Fig. 5.

10 Figure 9 reveals that.” Figures and tables as well as their captions must be inserted in the main text near the location of the first mention (not appended to the end of the manuscript). Regarding colour schemes, it is important that your maps and charts allow readers with colour vision deficiencies to correctly interpret your findings. Please check your figures using the [Coblis – Color Blindness Simulator](#) and revise the colour schemes accordingly.

1 Section (as Heading 1)

15 Suspendisse a elit ut leo pharetra cursus sed quis diam (Smith et al., 2014; Miller and Carter, 2015). Nullam dapibus, ante vitae congue egestas, sem ex semper orci, vel sodales sapien nibh sed lectus. Etiam vehicula lectus, quis orci ultricies dapibus. In sit amet lorem egestas, pretium sem sed, tempus lorem.

1.1 Subsection (as Heading 2)

Quisque cursus massa sed urna congue, ac convallis neque consectetur. Proin faucibus neque non metus mollis, suscipit pretium nisl blandit. In hac habitasse platea dictumst.

20 1.1.1 Subsubsection (as Heading 3)

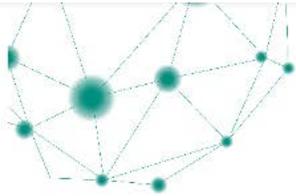
In placerat dictum urna ut interdum. Proin faucibus neque non metus mollis, suscipit pretium nisl blandit. In hac habitasse platea dictumst. Nam laoreet augue eu odio eleifend, non posuere quam pulvinar. Integer sit amet leo vitae nisl facilisis tristique calculated following Eq. (1):

$$Y = \frac{\Delta M_0}{\Delta[\text{isoprene}]}, \quad (1)$$

25 where ΔM_0 is Ut rutrum, sapien et vulputate molestie, augue velit consectetur lectus, bibendum porta justo odio lobortis ligula. In in urna nec arcu iaculis accumsan nec et quam. Integer ut orci mollis, varius justo vitae, pellentesque leo. Ut.



Outil d'ébauche d'un data paper



Recherche Data Gouv,
l'écosystème au service du partage et de l'ouverture
des données de recherche célèbre ses 1 an



Recherche Data Gouv

Génération datapaper

(Recherche Data gouv)

- Fonctionnalité spécifique à l'entrepôt *Recherche Data Gouv*
- 2 modèles disponibles
 - *Recherche Data Gouv*
 - **Data in Brief**



<https://entrepot.recherche.data.gouv.fr/datapartage-datapapers-web>

Exemple d'ébauche de data paper



Métadonnées générales ^

Identifiant pérenne	doi:10.57745/TODBTM
Date de publication	2024-09-13
Titre	Données de l'enquête sur les pratiques des scientifiques en matière de publication de data paper
Autre titre	Data from the survey on scientists' practices in terms of data paper publication
Point de contact	Utiliser le bouton de courriel ci-dessus pour joindre la personne-contact. Gassama, Mouhamed (DipSO ; INRAE) Szabo, Dimitri (DipSO ; INRAE) Tang, Cathy (DipSO ; INRAE)
Auteur	Gassama, Mouhamed (DipSO ; INRAE) - ORCID: 0009-0002-2977-749X Tang, Cathy (DipSO ; INRAE) - ORCID: 0000-0003-4787-1690 Szabo, Dimitri (DipSO ; INRAE) - ORCID: 0000-0002-1524-437X
Contributeur	Other : Fournier, Dominique (SDAR Montpellier ; INRAE) - ORCID: 0000-0001-6150-9079 Other : Grenier, Anne-Sophie (IGEPP ; INRAE) Other : Chemla, Sandrine (DipSO ; INRAE) Other : Bravo, Sonia (DipSO ; INRAE) - ORCID: 0000-0002-0640-5733
Distributeur	Entrepôt-Catalogue Recherche Data Gouv https://entrepot.recherche.data.gouv.fr
Description	<p>Ce jeu de données analyse les résultats d'une enquête réalisée entre septembre et novembre 2023, sur les pratiques des scientifiques en matière de publication de data papers. Cette enquête a été menée auprès de 226 scientifiques de l'enseignement supérieur et de la recherche française produisant des données de recherche. Elle est composée de 3 parties : le contexte professionnel des répondants, leur pratique en matière de publication de data paper (revue, difficultés rencontrées, publications associées) et leur retour sur l'utilisation de l'outil permettant de générer une ébauche de data paper à partir d'un jeu de données déposé dans l'entrepôt Recherche Data Gouv. French</p> <p>This dataset analyses the results of a survey carried out between September and November 2023 on scientists' practices in terms of publishing data papers. The survey was carried out among 226 scientists in higher education and research in France who produce research data. It consists of 3 parts: the professional context of the respondents, their practice in terms of publishing data papers (reviews, difficulties encountered, associated publications) and their feedback on the use of the tool for generating a draft data paper from a dataset deposited in Recherche Data Gouv repository. English</p>
Langue	French
Sujet	Computer and Information Science; Other
Mot-clé	data paper http://opendata.inrae.fr/thesaurus/INRAE/c_20880 (INRAE Thesaurus) https://thesaurus.inrae.fr/thesaurus-inrae/fr/données-ouvertes http://opendata.inrae.fr/thesaurus/INRAE/c_15504 (INRAE Thesaurus) https://thesaurus.inrae.fr/thesaurus-inrae/fr/enquête-auprès-des-utilisateurs http://opendata.inrae.fr/thesaurus/INRAE/c_16064 (INRAE Thesaurus) https://thesaurus.inrae.fr/thesaurus-inrae/fr/



Données de l'enquête sur les pratiques des scientifiques en matière de publication de data paper.
Data from the survey on scientists' practices in terms of data paper publication

Gassama, Mouhamed [1], Tang, Cathy [2], Szabo, Dimitri [3]

[1] DipSO ; INRAE. [2] DipSO ; INRAE. [3] DipSO ; INRAE

Author's identifier(s) : ORCID:0009-0002-2977-749X (Gassama, Mouhamed); ORCID:0000-0003-4787-1690 (Tang, Cathy); ORCID:0000-0002-1524-437X (Szabo, Dimitri)

Corresponding author(s): Gassama, Mouhamed (mouhamed.gassama@inrae.fr) DipSO ; INRAE; Szabo, Dimitri (dimitri.szabo@inrae.fr) DipSO ; INRAE; Tang, Cathy (cathy.tang@inrae.fr) DipSO ; INRAE

Citation: Gassama, Mouhamed, Tang, Cathy, Szabo, Dimitri... (year) Données de l'enquête sur les pratiques des scientifiques en matière de publication de data paper. . *Journal Name, Volume, (Issue number), doi of the data paper.*

Abstract

Ce jeu de données analyse les résultats d'une enquête réalisée entre septembre et novembre 2023, sur les pratiques des scientifiques en matière de publication de data papers. Cette enquête a été menée auprès de 226 scientifiques de l'enseignement supérieur et de la recherche française produisant des données de recherche. Elle est composée de 3 parties : le contexte professionnel des répondants, leur pratique en matière de publication de data paper (revue, difficultés rencontrées, publications associées) et leur retour sur l'utilisation de l'outil permettant de générer une ébauche de data paper à partir d'un jeu de données déposé dans [l'entrepôt Recherche Data Gouv.](https://entrepot.recherche.data.gouv.fr/)

This dataset analyses the results of a survey carried out between September and November 2023 on scientists' practices in terms of publishing data papers. The survey was carried out among 226 scientists in higher education and research in France who produce research data. It consists of 3 parts: the professional context of the respondents, their practice in terms of publishing data papers (reviews, difficulties encountered, associated publications) and their feedback on the use of the tool for generating a draft data paper from a dataset deposited in [Recherche Data Gouv repository.](https://entrepot.recherche.data.gouv.fr/)

Kind of Data: Dataset, Work flow,



RÉPUBLIQUE
FRANÇAISE recherche.data.gouv.fr

*Liberté
Égalité
Fraternité*

Enquête et méthodologie



Pourquoi une enquête ?



Rendre plus visible l'outil d'ébauche de data
paper

Utilisation
Connaissance
Difficultés

Pourquoi une enquête ?



Rendre plus visible l'outil d'ébauche de data paper

Utilisation
Connaissance
Difficultés



Comprendre les pratiques en matière de data paper

Difficultés
Ordre
Publication
Qui
Revue



Pourquoi une enquête ?



Rendre plus visible l'outil d'ébauche de data paper

Utilisation
Connaissance
Difficultés



Comprendre les pratiques en matière de data paper

Difficultés
Ordre
Publication
Qui
Revue



Élargir les usages sur les data papers

Accessibilité
Partage
Emergence
Données
Réutilisation



Public cible de l'enquête

- Adressée principalement aux **utilisateurs de l'entrepôt Recherche Data Gouv**
- Pouvant intéresser toute personne de **l'enseignement supérieur et de la recherche française** intéressée dans une démarche d'ouverture des données



Nous lançons une enquête sur les **pratiques des scientifiques en matière de publication de data papers**.

Cette enquête s'adresse à tout scientifique de l'enseignement supérieur et de la recherche française qui produit des données de recherche et qui souhaite les valoriser, y compris à ceux qui n'ont jamais (ou pas encore) publié de data papers.

Les réponses nous permettront de mieux connaître vos pratiques, vos besoins, les difficultés éventuelles rencontrées lors de la rédaction ou la soumission du data paper, etc. Ceci afin notamment d'améliorer [l'outil de génération d'ébauche de data paper](#) actuellement proposé dans [l'entrepôt Recherche Data Gouv](#), ainsi que les formations sur le sujet.



Structure du questionnaire

- Caractéristiques des répondants
 - Statut, établissement de rattachement, domaine scientifique
- Pratique de publication en matière de data paper
 - Qui a déjà publié ?, A quel moment ?, Dans quelle revue ?, Difficultés rencontrées ?
- L'outil data paper proposé par Recherche Data Gouv
 - Connaissance, Utilisation, Nouveaux modèles à proposer





**RÉPUBLIQUE
FRANÇAISE** recherche.data.gouv.fr

*Liberté
Égalité
Fraternité*

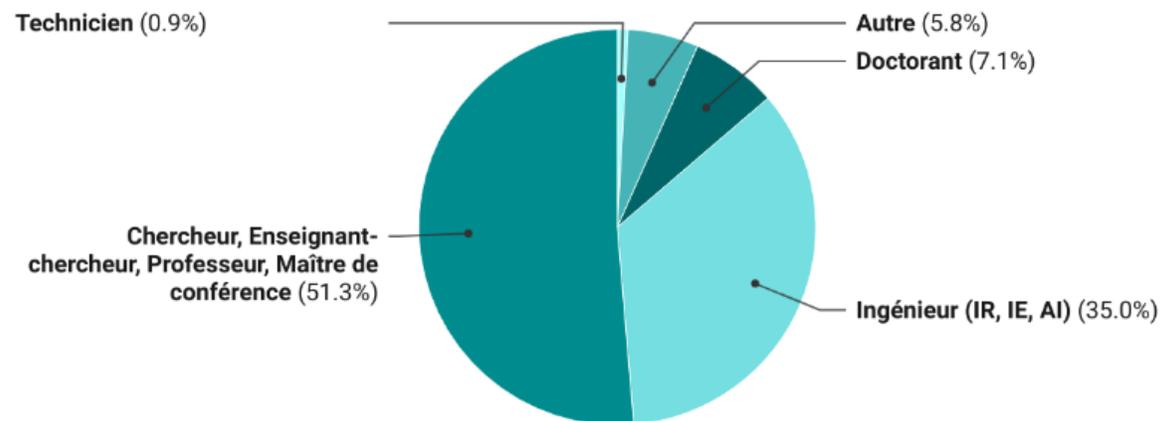
Résultats de l'enquête



Typologie des répondants

Statut des répondants

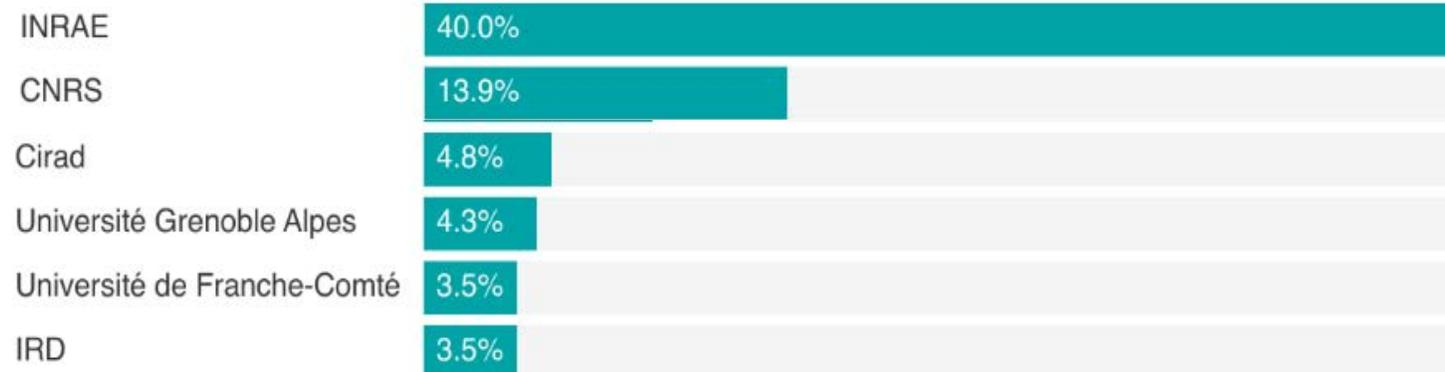
226 réponses retenues
dont



Typologie des répondants

Principaux établissements d'appartenance des répondants

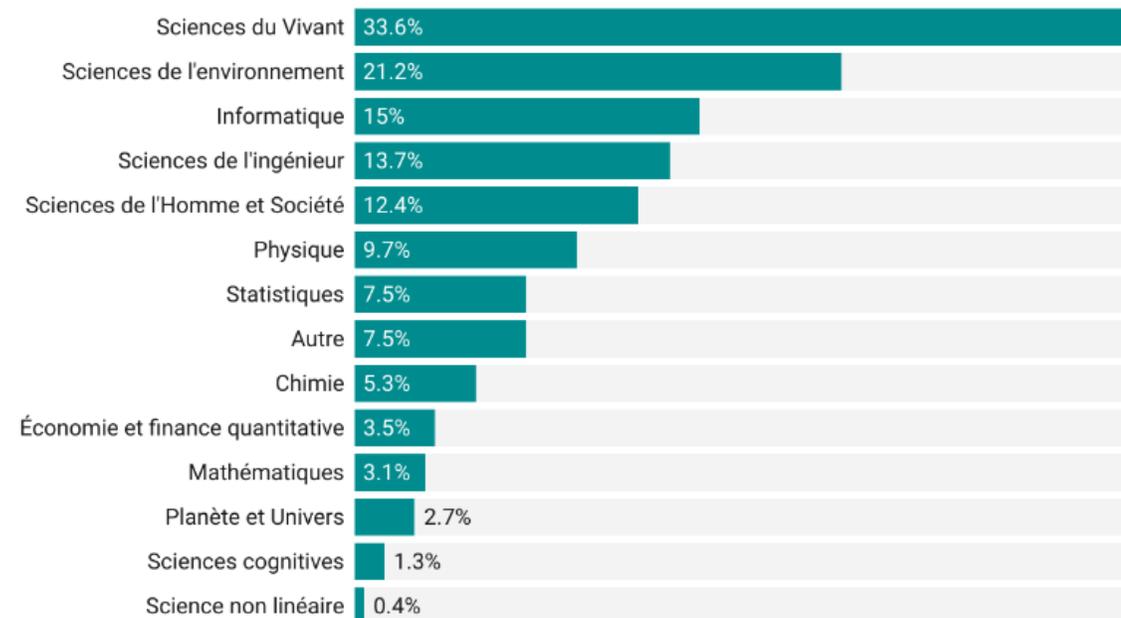
42 établissements représentés dont



Typologie des répondants

Domaine scientifique de recherche

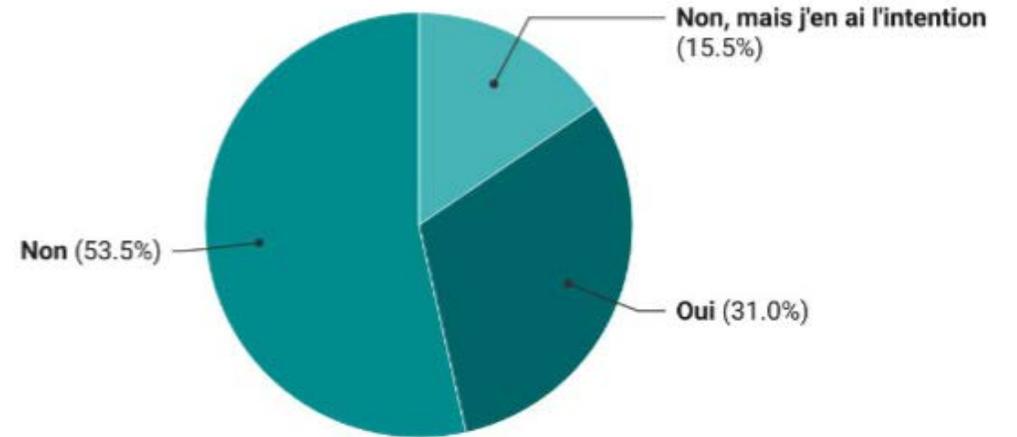
Surreprésentation des domaines des **Sciences du vivant (33.6%)** et de **l'environnement (21.2%)**.



Pratique de publication des data papers

Avez-vous déjà publié des data papers ?

46.5% ont publié un data paper (ou l'envisagent)



Pratique de publication des data papers

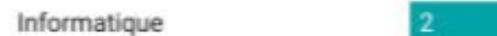
Principales revues utilisées : Data In Brief, Scientific Data, Annals of Forest Science

■ Nombre de répondants

Data in Brief



Language Resource and Evaluation Conference (LREC)



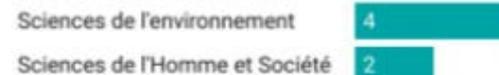
PLOS ONE



Scientific Data



Earth System Science Data



Journal of Open Archaeology Data (JOAD)



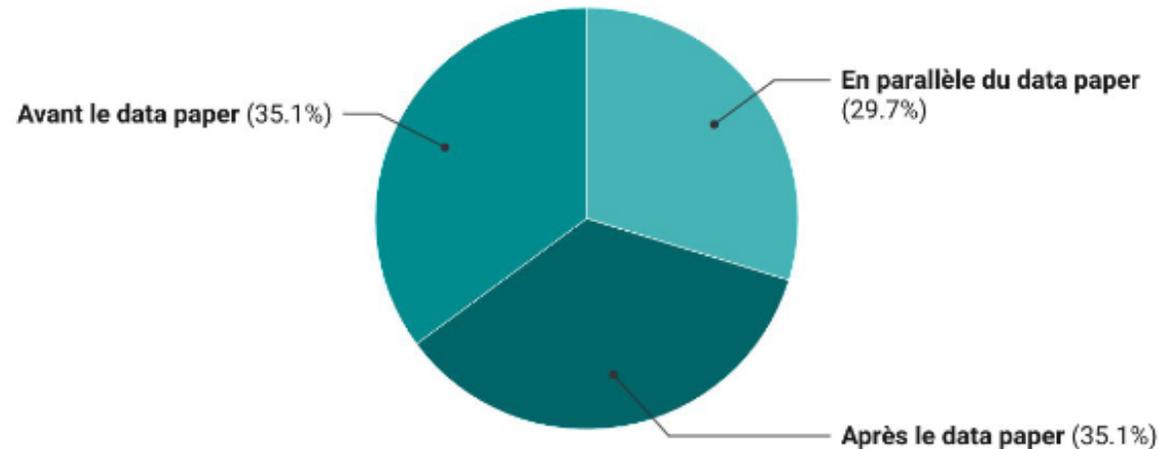
Annals of Forest Science



Sur le même sujet, voir : Rémy Decoupes. Analyse de l'impact des data papers de l'UMR TETIS. <https://hal.science/hal-04428092v2>

Pratique de publication des data papers

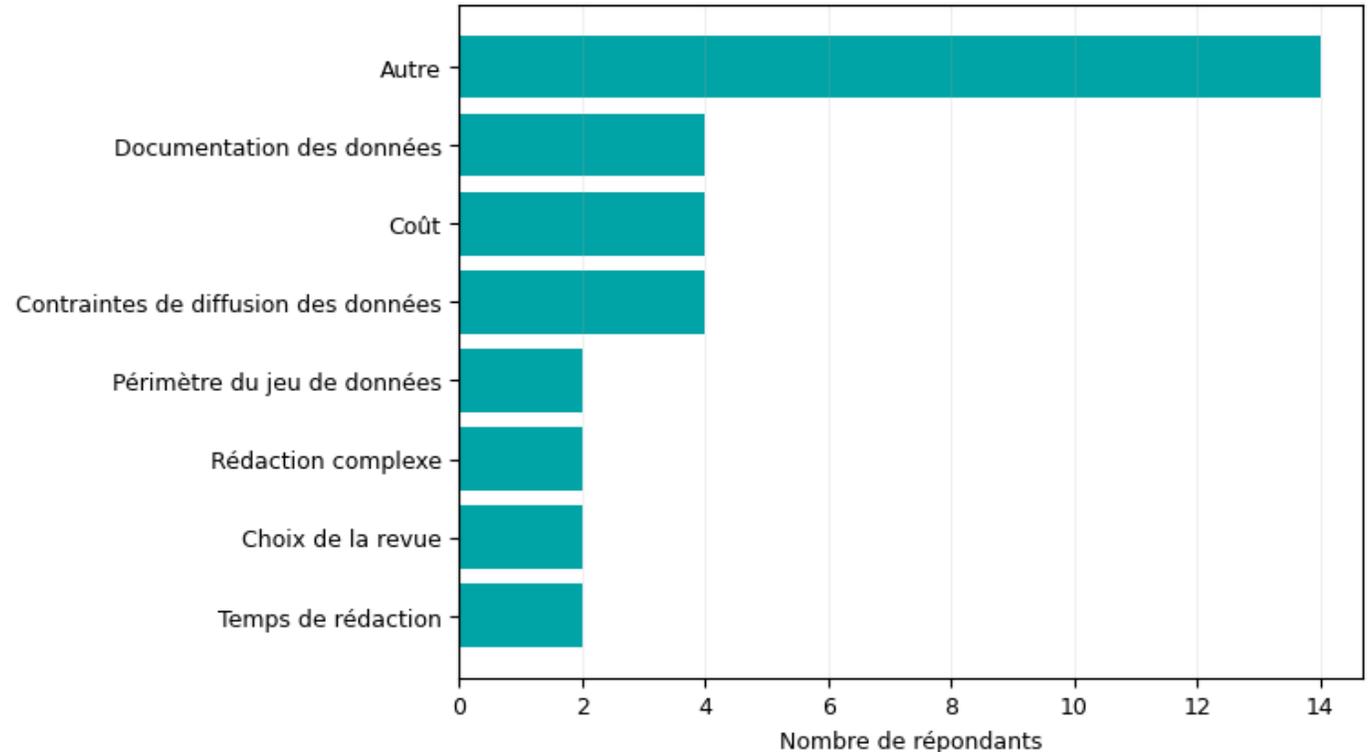
Période de publication de l'article de recherche en lien avec le data paper



Pratique de publication des data papers

Principales difficultés rencontrées par les répondants lors de la rédaction d'un data paper

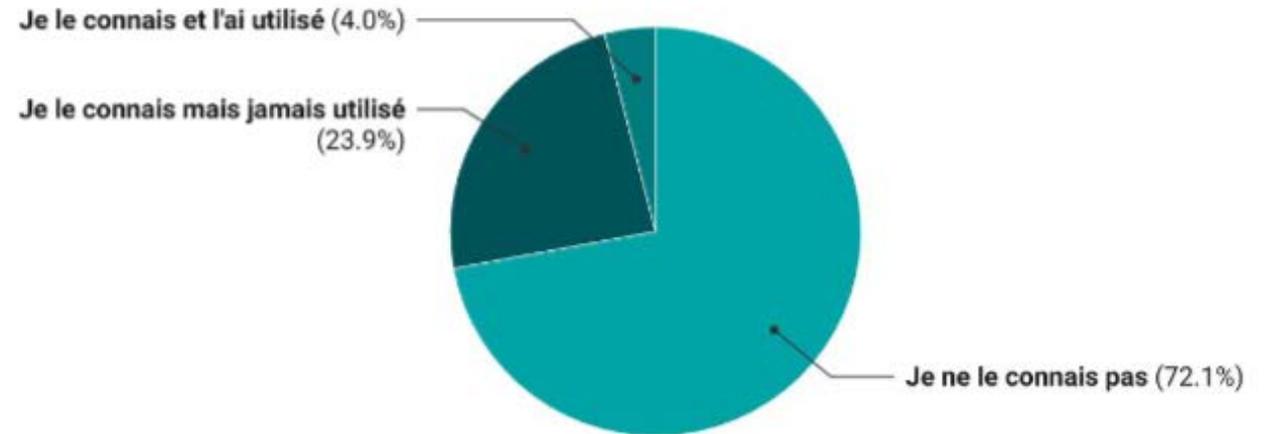
- **Documentation des données**
- **Coût**
- **Contraintes de diffusion des données**



Connaissance de l'outil de génération de data paper

Avez-vous connaissance de l'outil et l'avez-vous déjà utilisé ?

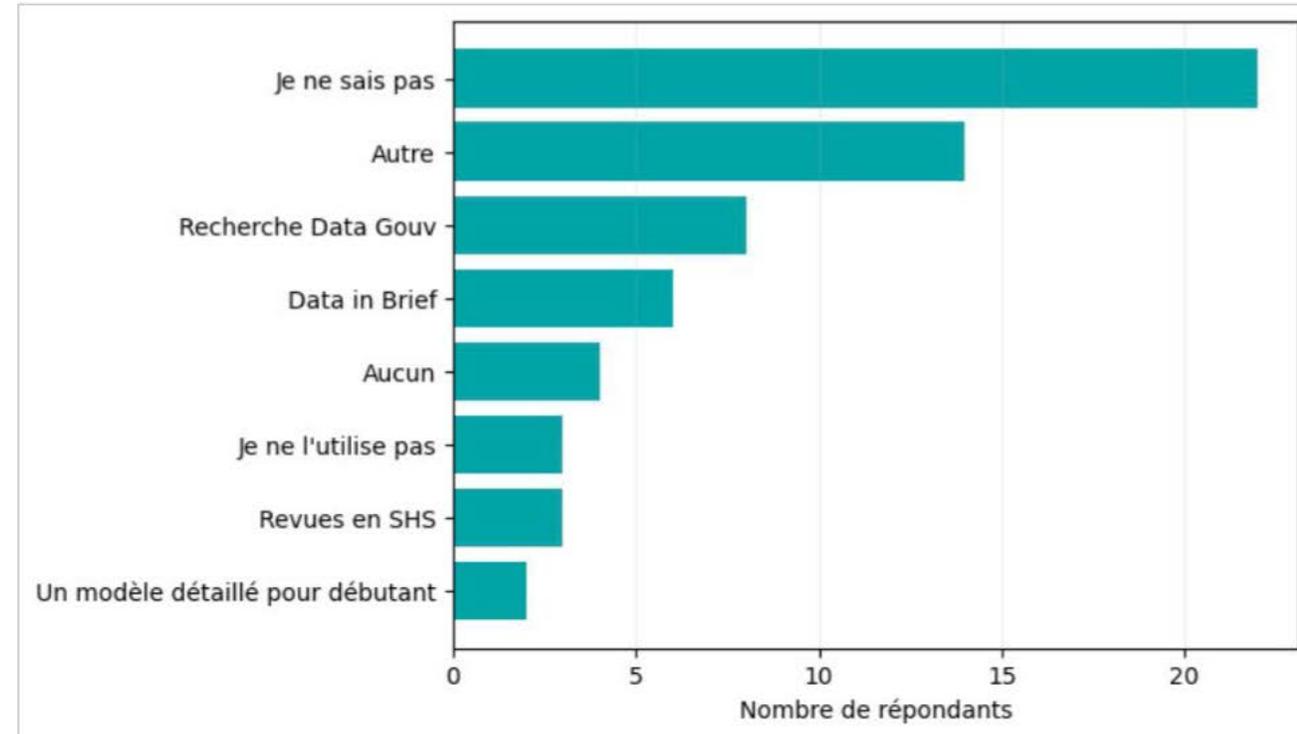
Méconnaissance de l'outil :
72.1% disent ne pas le connaître



Connaissance de l'outil de génération de data paper

Quels modèles vous seraient nécessaires pour utiliser cet outil ?

- **Revue en SHS**
- **Un modèle détaillé pour débutant**





RÉPUBLIQUE
FRANÇAISE recherche.data.gouv.fr

*Liberté
Égalité
Fraternité*

CONCLUSION





Partage des résultats

- Diffusion des données via un jeu de données

DOI : <https://doi.org/10.57745/TODBTM>

- Analyse statistiques de chaque colonne, recoupement des colonnes (fichier)
- Jeu de données final

- Code publié sur la forge logicielle institutionnelle

<https://forgemia.inra.fr/rdg/entrepot/datapapers/enquete-datapapers>



- Rédaction d'un rapport d'analyse publié dans HAL : <https://hal.science/hal-04690792>

Bibliographie / Webographie

Joachim Schöpfel, Dominic J. Farace, Hélène Prost, Antonella Zane. Data papers as a new form of knowledge organization in the field of research data. *Knowledge Organization*, 2020, 46 (8), pp.622-638.

Rémy Decoupes. Analyse de l'impact des data papers de l'UMR TETIS. <https://hal.science/hal-04428092v2>

<https://www.fosteropenscience.eu/content/cartoon-data>

<https://data.ird.fr/datapapers/>

<https://github.com/LimeSurvey/LimeSurvey>

<https://entrepot.recherche.data.gouv.fr/>

<https://entrepot.recherche.data.gouv.fr/datapartage-datapapers-web/>

https://gitlab.com/users/sign_in

https://cartong.pages.gitlab.cartong.org/learning-corner/fr/3_usable_data_analysis/3_1_first_steps



**Merci
pour votre écoute**

Vos Questions ?