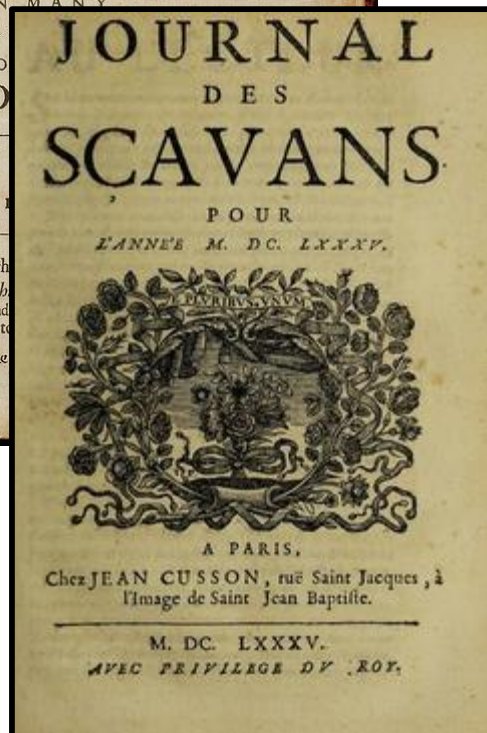
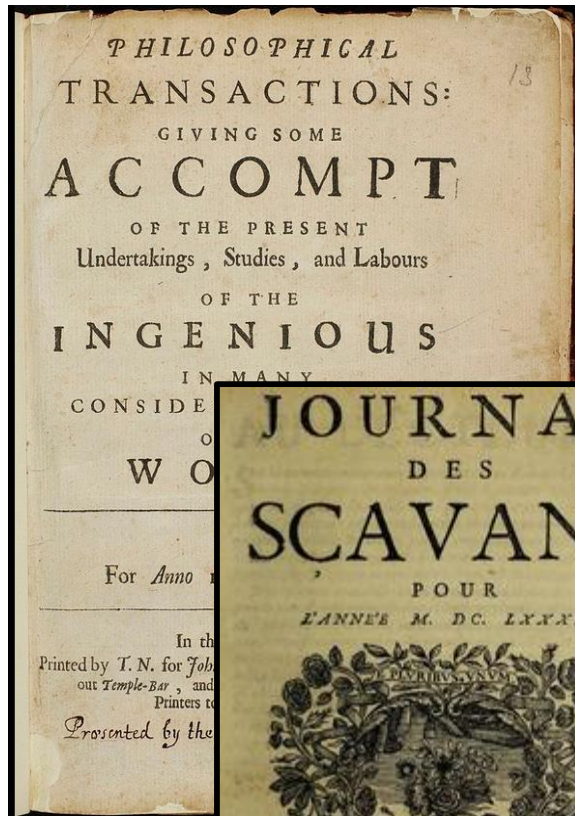


***Les data papers*, objets de négociation des régimes d'ouverture**

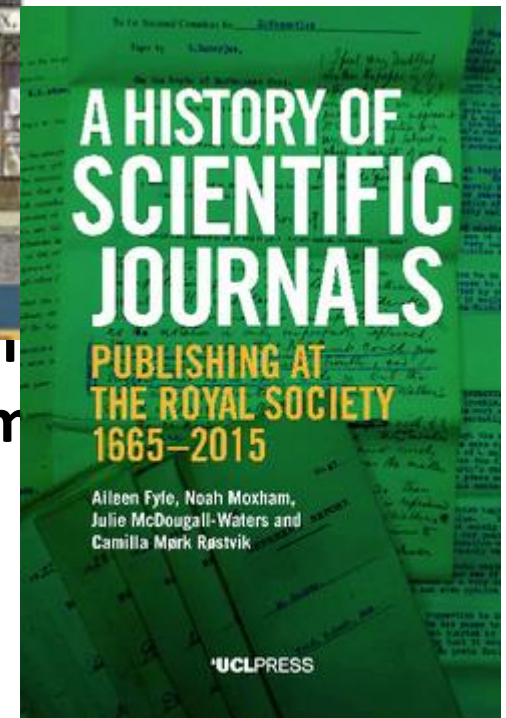
Chérifa BOUKACEM ZEGHMOURI
Professeure en SIC
Université Claude Bernard Lyon 1 (UCBL)

Love Data Week
Nancy - 14 mars 2024

La revue, un « artefact archéologique »



de la communauté
scientifique : m
macro



Fil d'Ariane

- Les données sont-elles publiables ?
- Quelle est la « véritable » place du Data Paper dans les formes de production des savoirs scientifiques ?
- Les motivations à « produire » un *Data Paper*
- Une entrée intéressante pour analyser les évolutions en cours : une nouvelle filière ?

Données de recherche : il faut "reconnaître à sa juste mesure le travail supplémentaire demandé aux chercheurs" (DGRI)

Une problématique de recherche

- Quelles sont les conditions socio-symboliques et socio-économiques dans lesquelles naissent les motivations (les stratégies ?) de publication de ce nouveau type d'articles ?
- Du côté des SHS et du côté des STM



Volume 35, Issue 1

Pages: 1-87

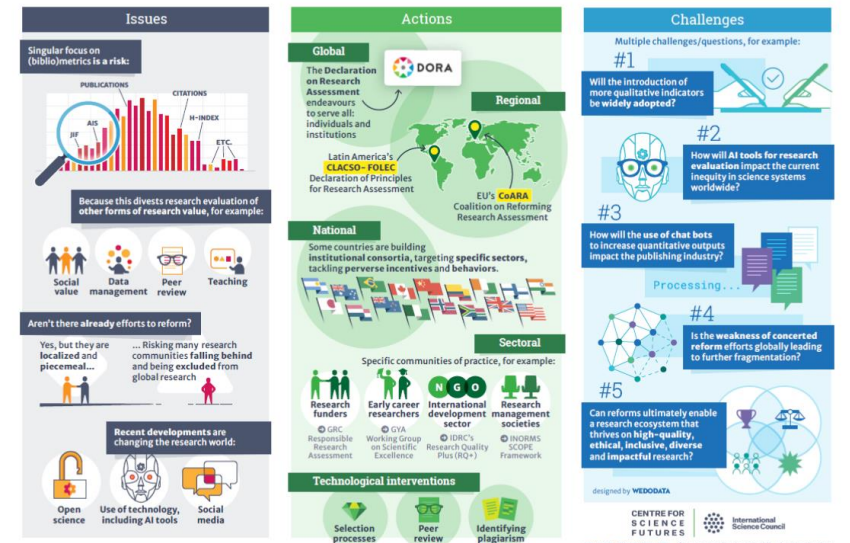
January 2022

Un contexte en tension

- Les éditeurs historiques
 - Elsevier, Springer, Wiley et les autres ...
- Les éditeurs Gold OA
 - Copernicus, MDPI et les autres ...



The Future of Research Evaluation

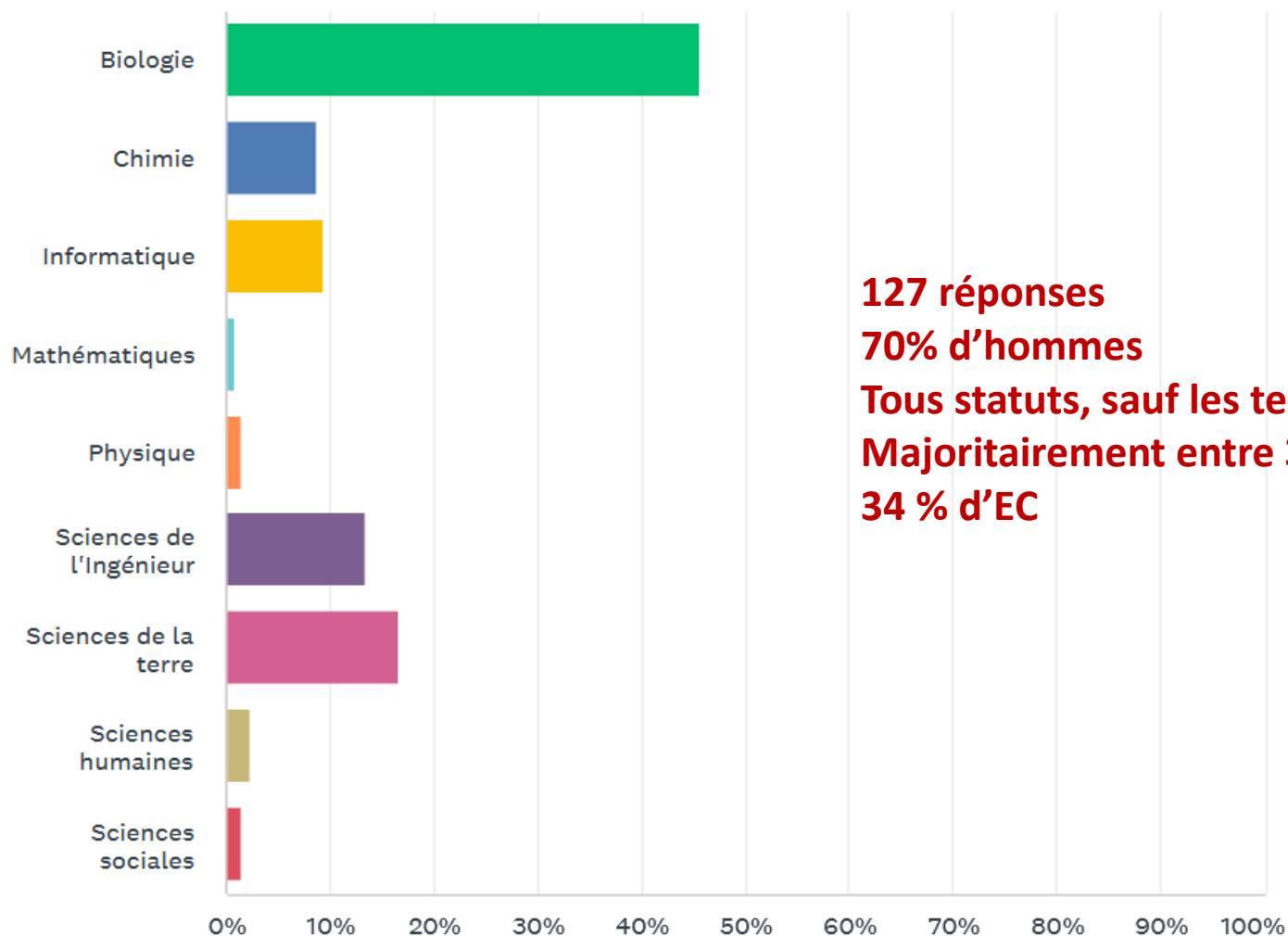


Choix méthodologiques

- Questionner les auteurs « français » identifiés
- Une première phase exploratoire
- Un pendant à des travaux antérieurs (étude de cas de SHS)
- Un questionnaire : questions fermées et ouvertes
- Se donner la possibilité d'une seconde phase : entretiens semi-directifs ou bien Focus Group
- Tous domaines confondus
- 6 mois ...de relances
- Participation des étudiants
- Un partage de réflexion avec Violaine REBOUILLAT

RÉSULTATS ... NON PUBLIÉS !

Qui a répondu ?



127 réponses

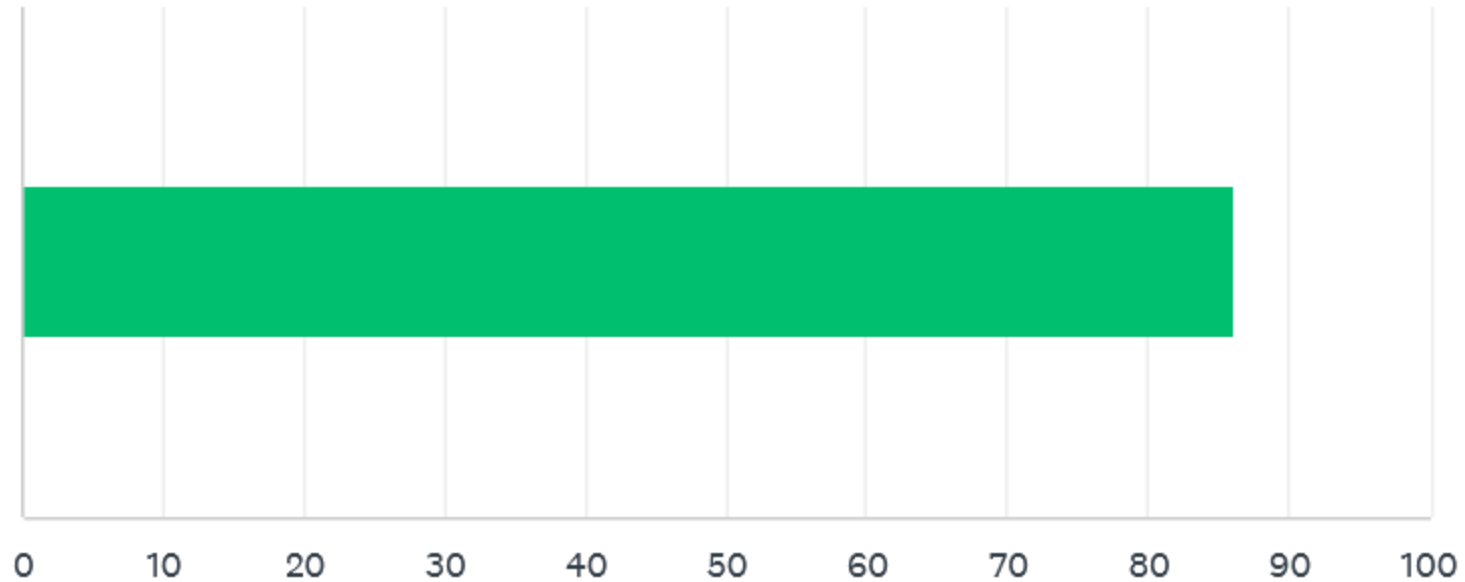
70% d'hommes

Tous statuts, sauf les techniciens

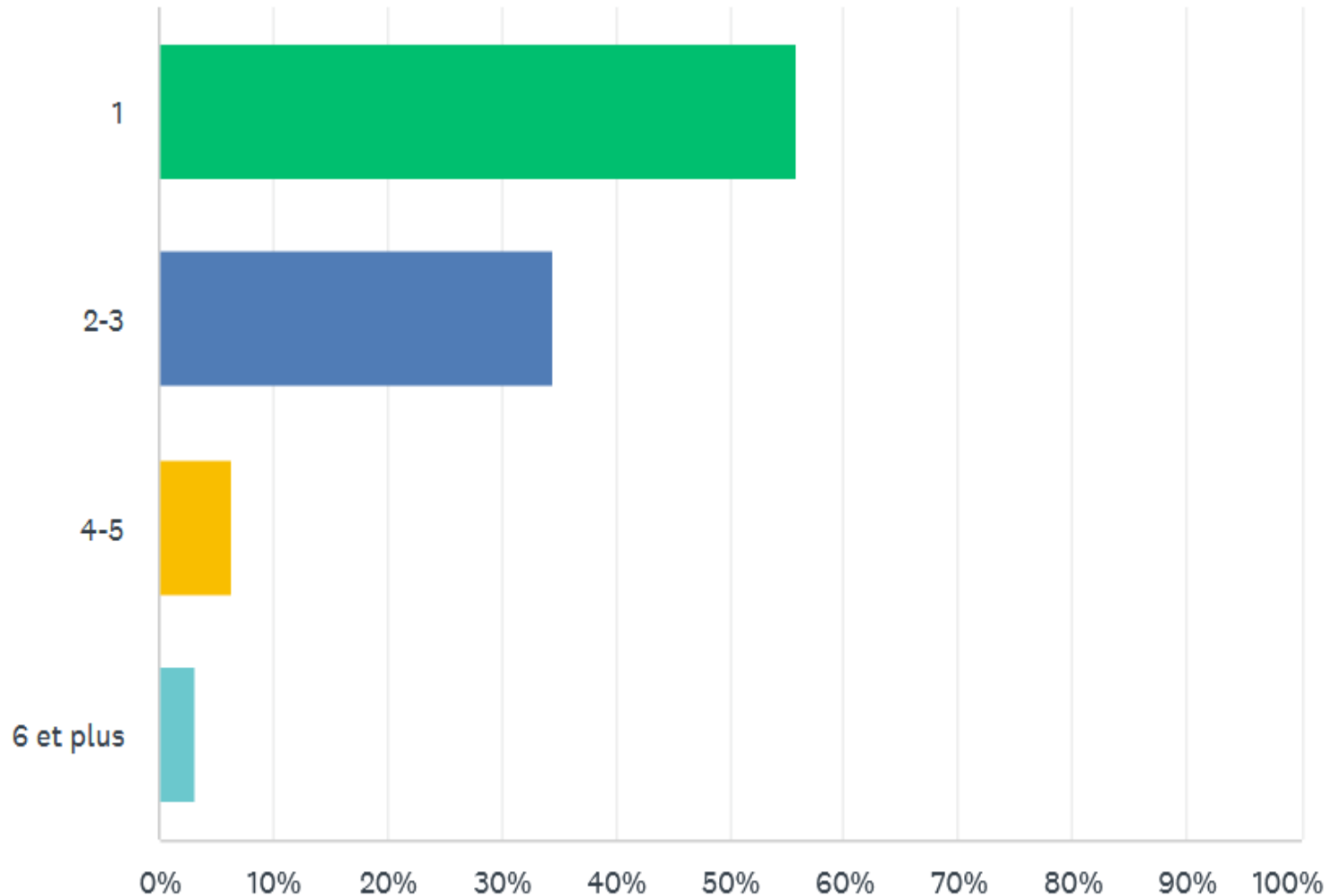
Majoritairement entre 31 et 65 ans

34 % d'EC

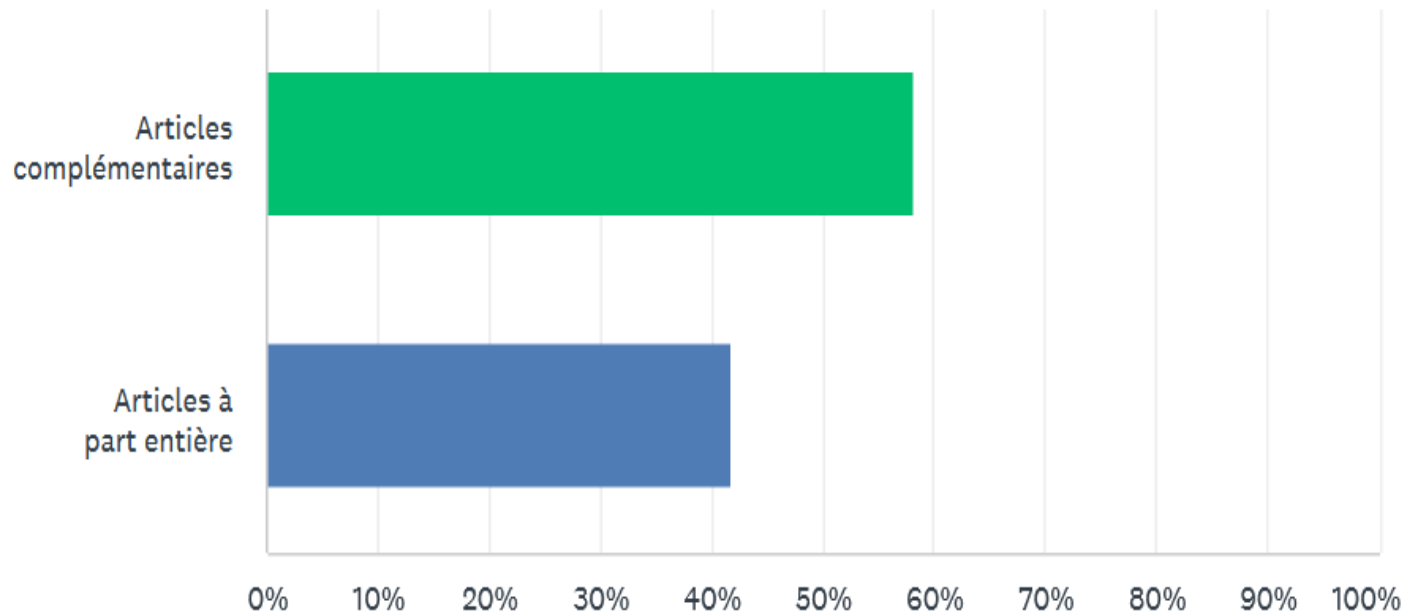
D'accord avec la définition de Joachim Schöpfel ?



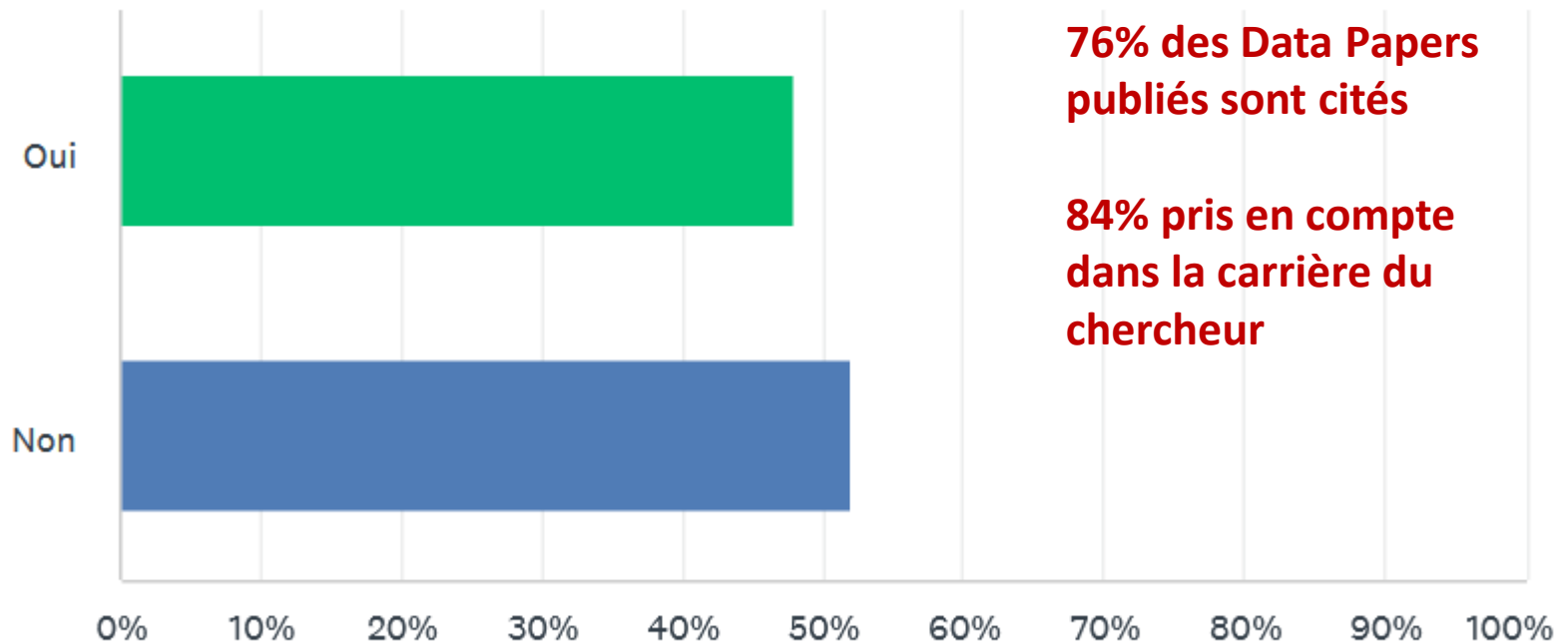
Nombre de *data papers* publiés sur les 5 dernières années



Des articles complémentaires ou bien à part entière ?



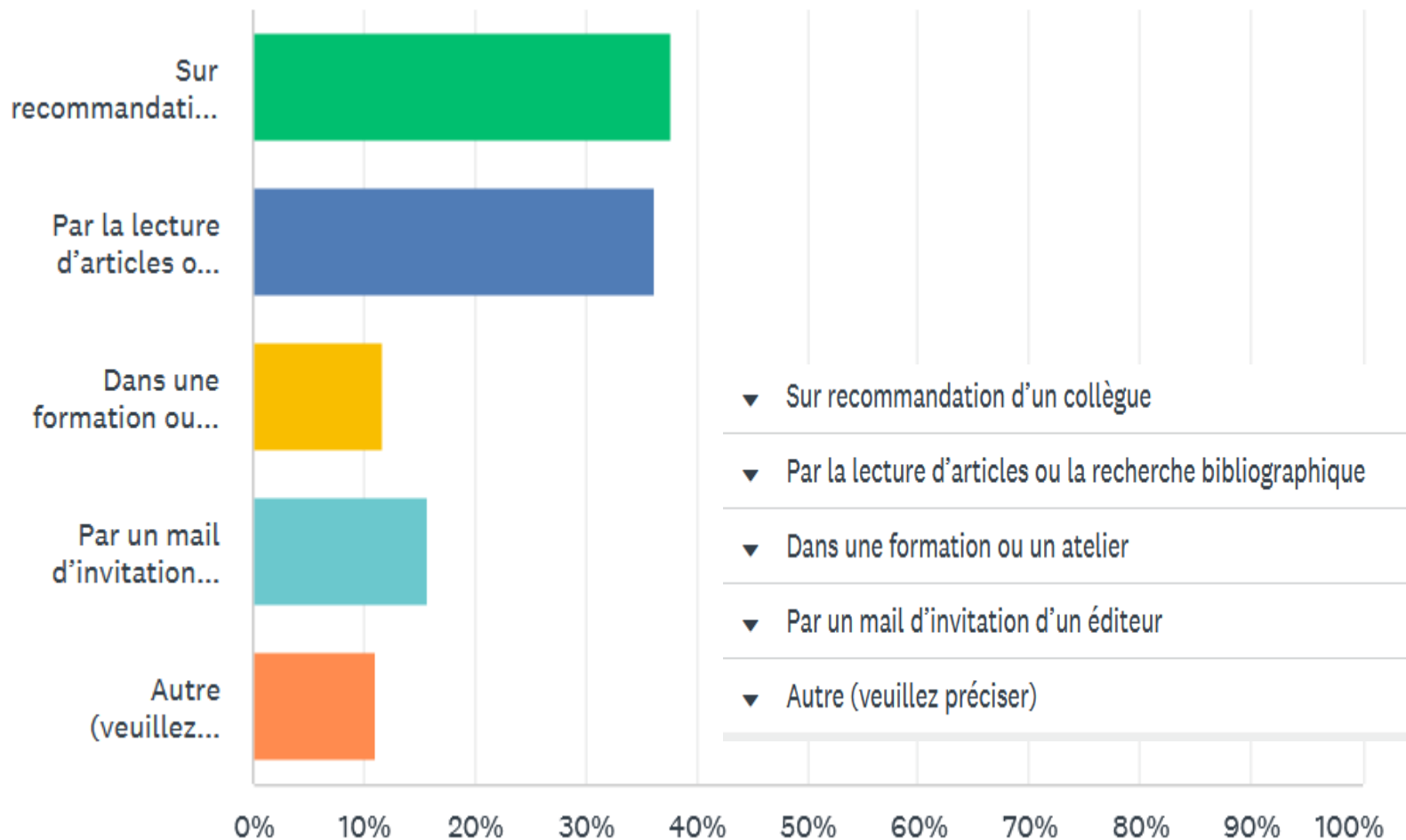
Avez-vous noté un impact ? (visibilité, collaboration)



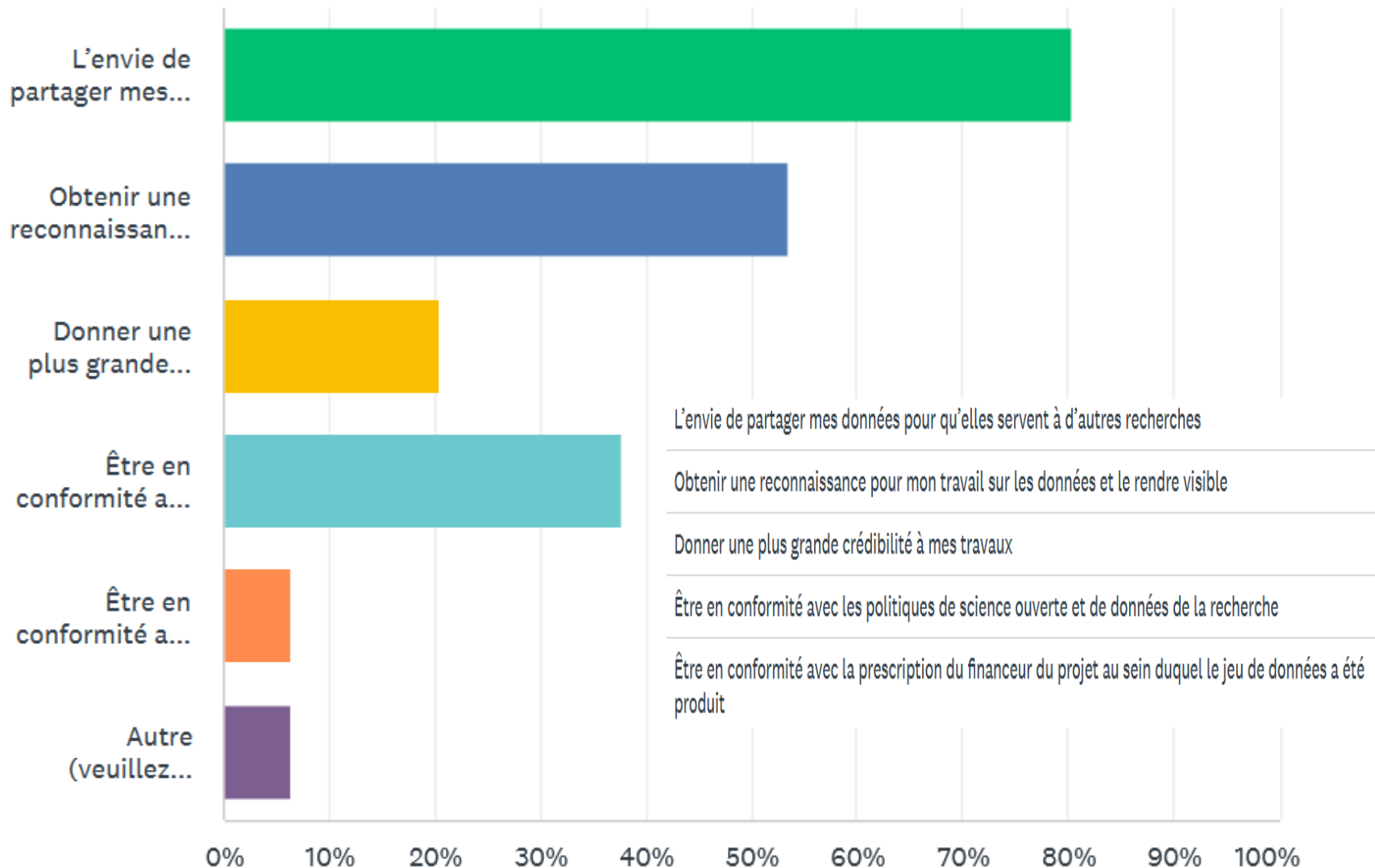
**76% des Data Papers
publiés sont cités**

**84% pris en compte
dans la carrière du
chercheur**

Comme avez-vous découvert le data paper ?



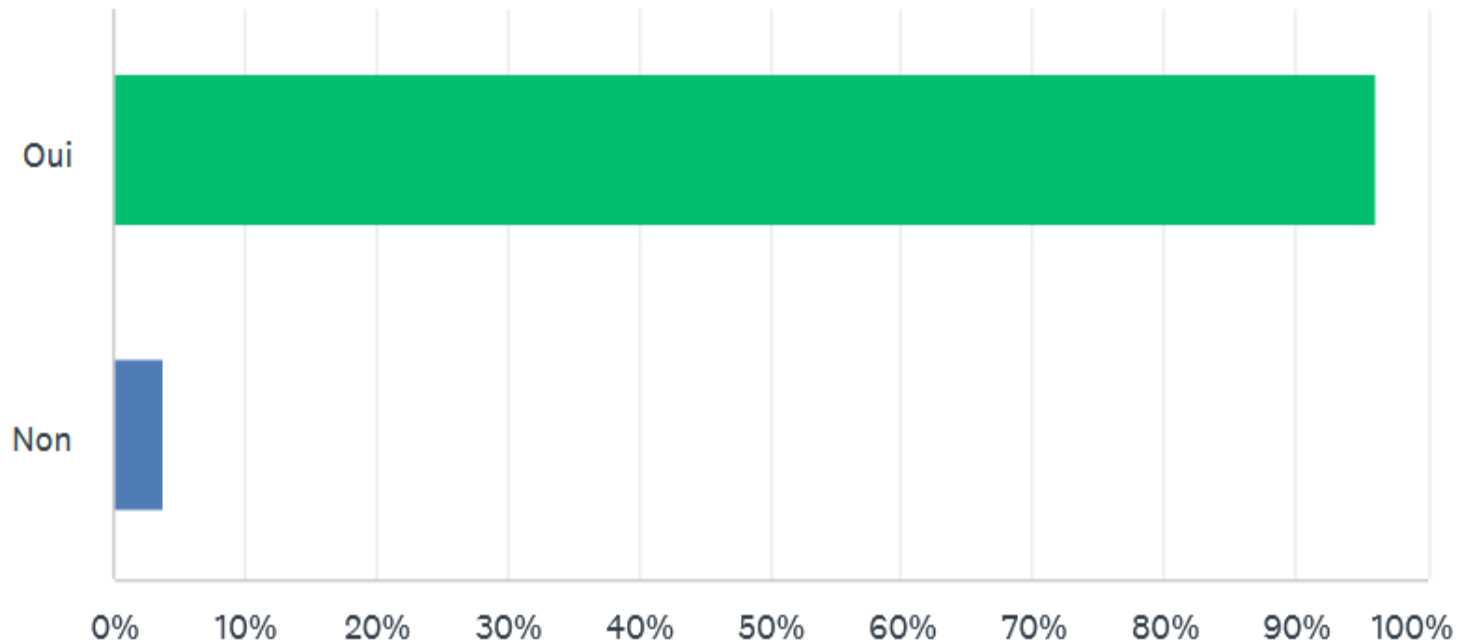
Soumis sur quelles motivations ?



D'autres arguments

- « Une publication gratuite et qui sert à étoffer le CV de mon étudiant en thèse »
- « Augmenter le nombre de mes publications »
- « Avoir un article en plus facilement, car c'est ce que m'avait vendu l'éditeur »
- « La plus grande facilité à faire publier un article »

Mettez vous à disposition les jeux de données associés ?



Sur quels critères la revue a-t-elle été choisie ?

▼ La visibilité auprès de ma communauté	22.05%	28
▼ La politique open access	24.41%	31
▼ La qualité de son système d'évaluation par les pairs (peer review)	7.87%	10
▼ Le lectorat de la revue	9.45%	12
▼ L'indexation de la revue dans les bases de données de référence	7.09%	9
▼ Le facteur d'impact de la revue	8.66%	11
▼ Les frais de publication en open access (APC)	7.09%	9
▼ Autre (veuillez préciser)	13.39%	17

Responses

13.39%

17



elico

Équipe de recherche de Lyon en sciences
de l'information et de la communication

Data in Brief

Data in Brief



Impact. In more ways than one
Open Access

Abstracting

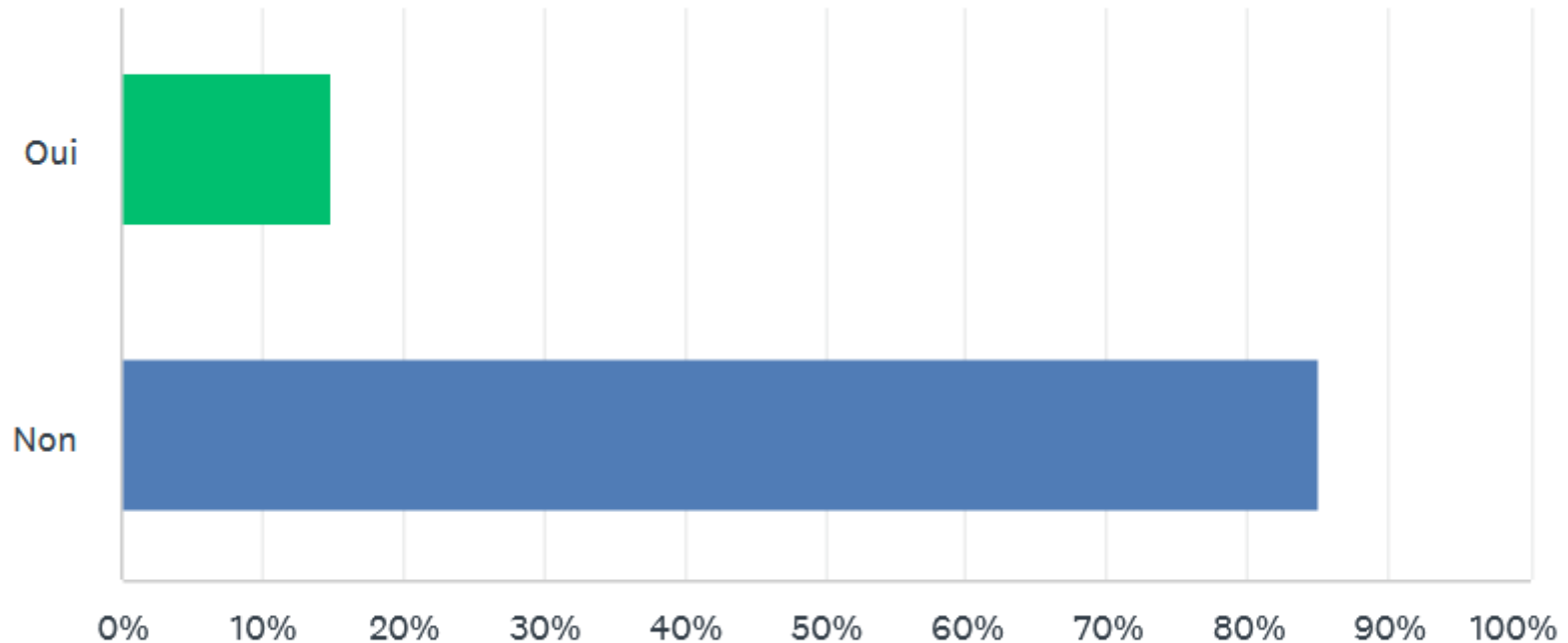
- PubMed Central
- PubMed/Medline
- Scopus
- Emerging Sources Citation Index (ESCI)
- Directory of Open Access Journals (DOAJ)



elico

Équipe de recherche de Lyon en sciences
de l'information et de la communication

Réticences ? Appréhensions ?



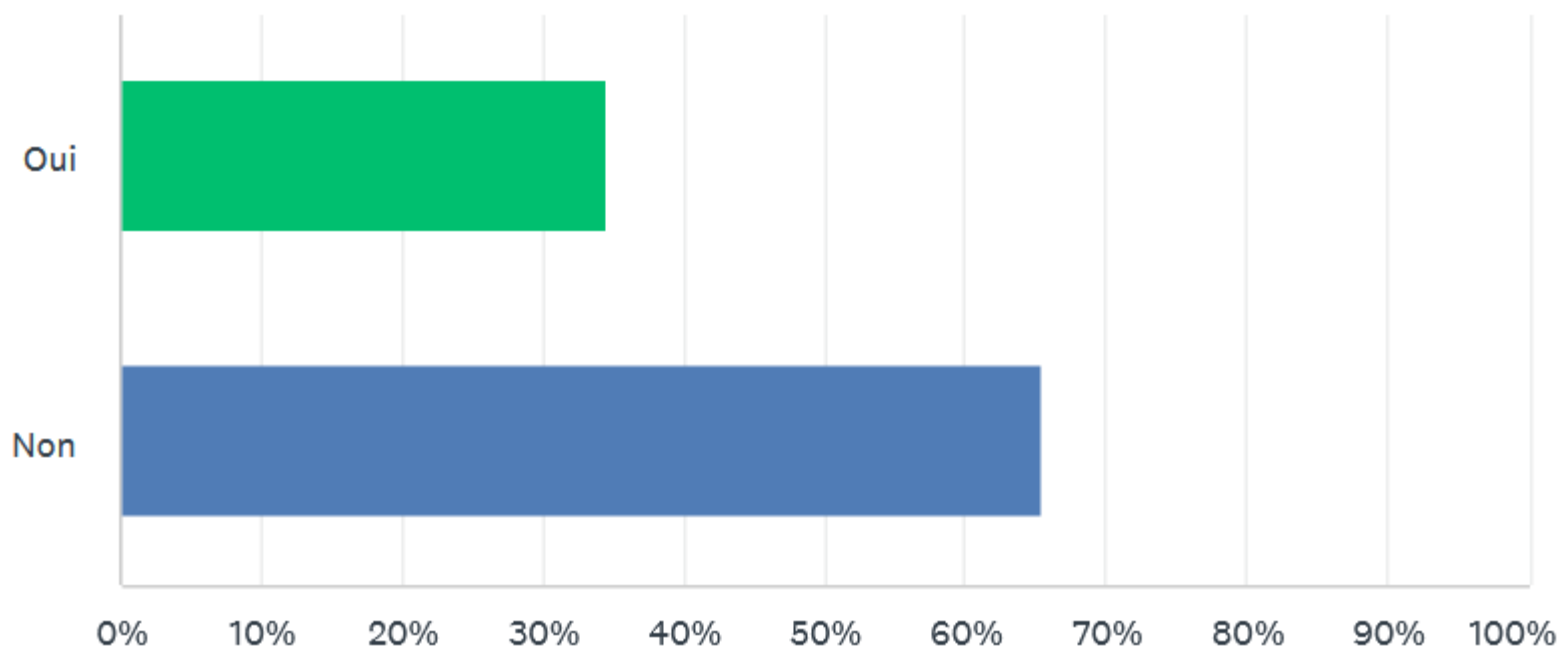
Quelques verbatims éclairants

- Temps, argent, valeur, visibilité
- « La peur d'être restreint à une description basique des données. Ce qui n'est pas du tout le cas pour ESSD, mais le cas pour Data In Brief. Ce dernier journal (DIB) ne bénéficie d'ailleurs pas d'une bonne qualité de relecture ce qui est dommage »
- « Moi non. Ce sont mes collègues des autres instituts de recherche qui lors de ma proposition ont eu des réticences. Méconnaissance du data paper, ne voulaient pas faire le data paper tant que leur articles scientifiques n'étaient pas fini de publier »

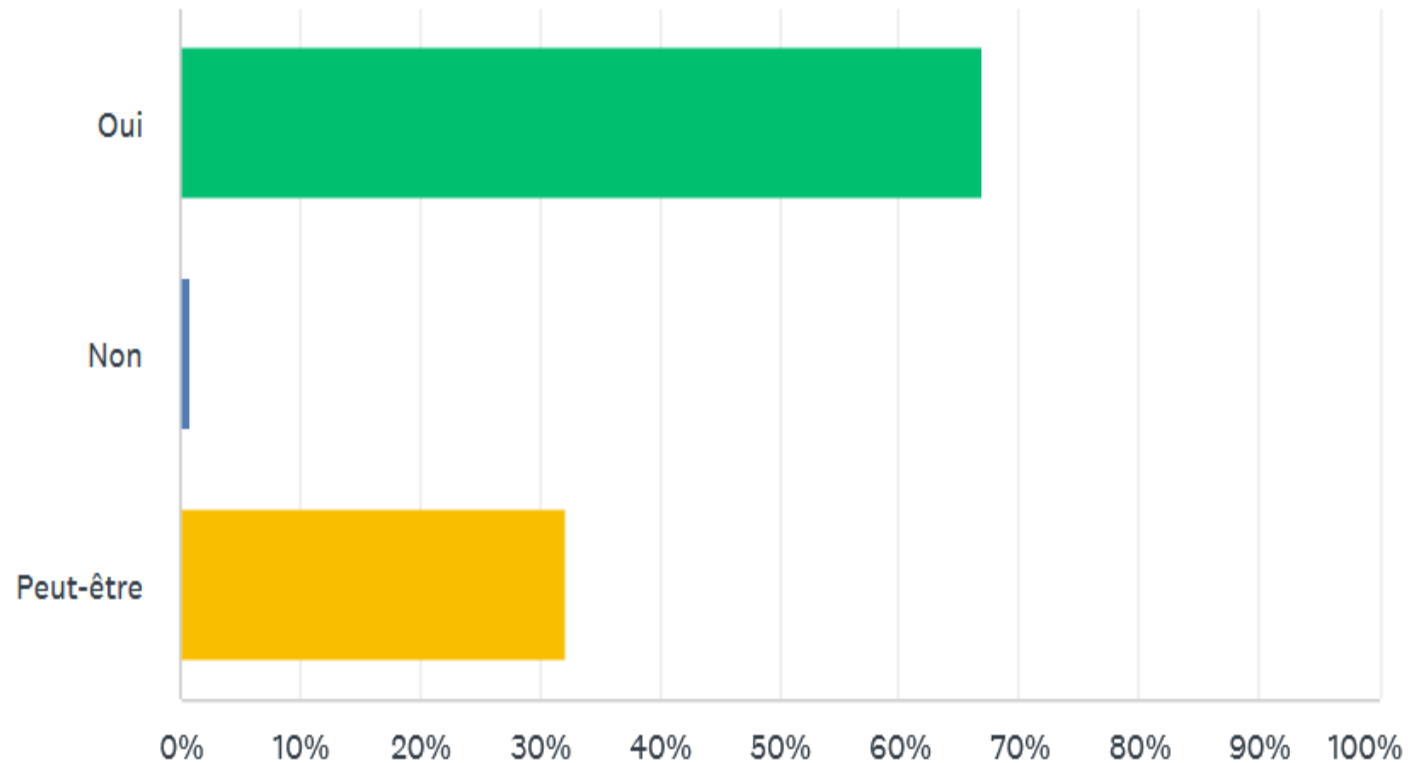
Quelle place pour le *Data Paper* dans les stratégies de publication ?

▼ Une publication régulière en même temps que les résultats de recherche	35.43%	45
▼ Une publication ponctuelle en attendant d'obtenir des résultats de recherche	14.17%	18
▼ Une publication systématique valorisant tous les jeux de données produits	13.39%	17
▼ Une publication exceptionnelle à la faveur d'une opportunité	37.01%	47

Le lecteur du data paper réutilise-t-il les données ?



De futurs « récidivistes » ?



Un argumentaire svp ?

- ***Partage, reproductibilité, ROI, visibilité (DOI), collaboration***
- **« Assurer la pérennité des jeux de données valoriser des collectes de données qui ne donnent pas suite à une analyse (sciences participatives, projets d'études réalisées par des étudiants, recherche non-concluante et "non-publiable", etc)**
- **« Vas-y, fonce. Ça prends moins de temps qu'un article conventionnel et est considéré à la même valeur. »**
- **« Publier autrement ses papiers, en mode allégé sur la partie Material, et se concentrer sur Method »**
- **« Si tu as des données insuffisantes pour faire un article scientifique, tu peux les publier dans un data paper rapidement et avoir un papier supplémentaire à moindre coût sur ces travaux ».**

- **« Ca valorise les techniciens et ingénieurs qui ont produits les données et ça les aide dans leur évolution de carrière Ca facilite l'écriture des articles scientifiques complexes (plusieurs jeux de données, nombreuses techniques et protocoles) »**
- **« Ne pas publier un data paper augmente le risque de garder des résultats pour toujours dans un tiroir. C'est à dire, un travail qui n'a jamais existé »**
- **« Faire le ménage » dans les données (curation)**

Les +

- « ça permet d'avoir un papier issue d'un processus "peer reviewed" associé aux données et de citer des données ».
- « Rendre les données publiques et faire connaître le projet. Il y a aussi la bénéfice de prendre le temps pour faire un résumé de l'ensemble des données collectées pour mieux réfléchir sur la valorisation possible (notamment dans le cas de grands projets impliquant de nombreuses équipes) ».
- « Cela peut donner plus de visibilité et citations à l'article principal auquel se rattache le data paper mais également valoriser le jeu de données et de le pérenniser »
- « Plus de personnes demandent accès à nos données, les réutilisent. Cela donne de la visibilité à nos recherches, et crée de nouvelles collaborations ».
- « On passe du temps à travailler sur les data, les publier permet de valoriser cela ».
- « Extension du lectorat à d'autres thématiques (interdisciplinarité) »

Les -

- **L'aspect chronophage du travail**
- **Convaincre des co-auteurs réticents (manque de reconnaissance)**
- **Le coût**
- **« L'impact facteur est très faible pour la quantité de travail »**
- **« Risque d'avoir une concurrence sur nos propres données »**
- **« Les articles data ne sont pas encore reconnus correctement par les comités scientifiques traditionnels ».**
- **« A mes yeux, un 'data paper' est surtout un argument habile qu'un éditeur utilisera pour faire publier un article payant supplémentaire (tout est dans "payant") ; en pratique, il existe un grand nombre de 'repositories' dédiés aux données (e.g. <https://zenodo.org/>) qui sont gratuits et robustes (e.g. association de chaque 'data' avec un doi), et offrent des avantages très similaires aux 'data papers'**

Des résultats intermédiaires

- **Data Papers =**
 - Le CV et la liste de publication
 - Les opportunités de collaboration (Rebouillat, 2019)
 - L'impact et la visibilité
 - L'invisibilité du travail des données
 - L'ouverture : des marqueurs discursifs
 - Un mode de management
- **Le rôle des éditeurs**
 - Incitation et accompagnement

Perspectives

- **Analyses à partir de Focus Group**
- **Approfondir la question de la diversification des registres de la reconnaissance**
 - Diversifier les registres de la reconnaissance
 - Production de la relation, de l'inter-reconnaissance (Bourdieu, 1980)
 - Les formes de collaborations, fondées sur les données, sur le long terme.
 - L'interdisciplinarité
- **Le craft, le travail de la donnée ouverte**
 - Les métaphores et la réalité
 - Une réponse à l'accélération
 - La réutilisation des données

EN ATTENTE DE QUESTIONS ...