



# SÉMINAIRE SO-MSH! 2023-2024

POURQUOI PUBLIER UN ARTICLE DE DONNÉES (DATA PAPER) ?

RETOURS D'EXPÉRIENCES DE CHERCHEURS

15 AVRIL 2024, LABORATOIRE AAU-CRENAU, NANTES

# PROGRAMME DE LA SÉANCE

Introduction et mise en contexte :

- enjeux et caractéristiques d'un data paper
- présentation des résultats de l'enquête nationale du projet « Data Journal SHS », Laurence Bizien (AAU-CRENAU), Véronique Cohoner (MSH Ange-Guépin).

# PROGRAMME DE LA SÉANCE

Exemples et partage d'expériences :

- Ni Zhu (GEOLOC, Université Gustave Eiffel)
- Stéphane Loret (ESO, Nantes Université)
- Gwendoline L'Her et Myriam Servières (AAU-CRENAU)

# DÉFINITION

Le data paper (ou data article) est une publication scientifique, soumise à la **validation des pairs**, dont le but principal est de décrire un ou plusieurs jeux de données, afin de permettre leur **réutilisation**. Regroupés **dans des data journals, ou publiés au sein de revues classiques** les data papers permettent de répondre à certaines exigences de la science ouverte et sont encouragés par le plan national pour la science ouverte (axe 2).

# POURQUOI PUBLIER UN DATA PAPER ?

- **Pour communiquer sur l'existence des données, et permettre de les trouver** : la publication des données se fait par l'intermédiaire des entrepôts de données et/ou de revues spécialisées ; dans ce cas les données sont intégrées dans l'article, en matériel supplémentaire, ou bien associées à un data paper.
- **Pour créditer les auteurs (reconnaissance, référence citable) et valoriser les données** : il permet aux producteurs de données de faire reconnaître leur contribution scientifique (par les comités d'évaluation et par les pairs) avec des données identifiables.
- **Pour faciliter la réutilisation des donnée**, en les décrivant avec des métadonnées conformes aux principes FAIR (Findable, Accessible, Interoperable, Reusable)

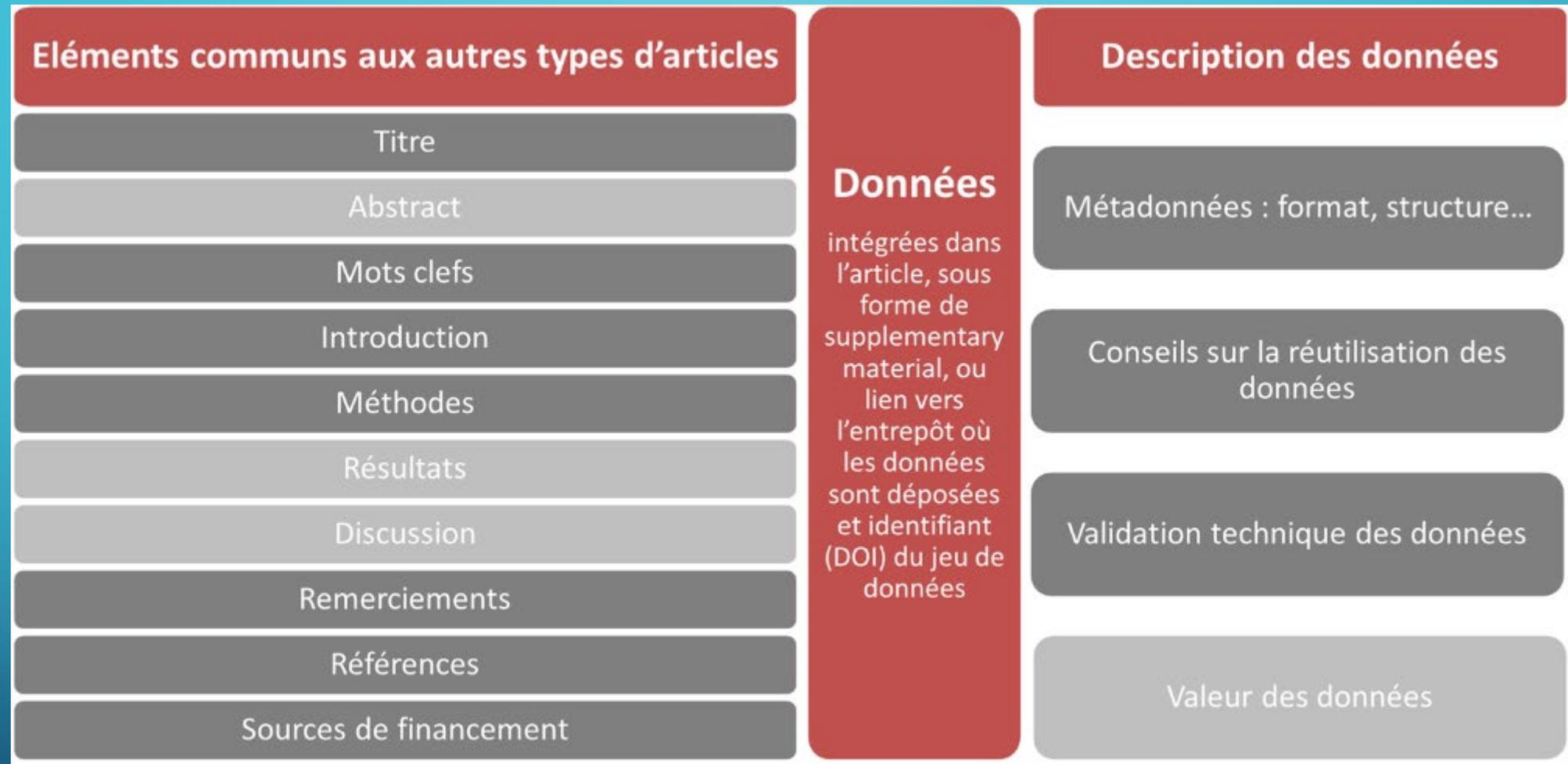
Source : [https://doranum.fr/data-paper-data-journal/data-papers-et-data-journals-fiche-synthetique\\_10\\_13143\\_2wcb-fw52/](https://doranum.fr/data-paper-data-journal/data-papers-et-data-journals-fiche-synthetique_10_13143_2wcb-fw52/)

# QUE CONTIENT UN DATA PAPER ?

- Les data papers ont une structure très variable selon les revues. Ils ont cependant des composantes communes :
- **Une partie descriptive avec des éléments communs aux articles classiques** (titre, résumé, mots clefs, etc.) **et des éléments spécifiques liés aux données** (métadonnées, réutilisation, etc.)
- **L'accès aux données** : elles peuvent être intégrées ou jointes à l'article, ou accessibles après avoir été déposées dans un entrepôt (recherche.data.gouv, Nakala, Zenodo ..) via l'identifiant des données (exemple : le DOI) permettant d'établir le lien du data paper vers les données.

Voir :

- [https://doranum.fr/data-paper-data-journal/contenu-data-paper\\_10\\_13143\\_8r3d-k505/](https://doranum.fr/data-paper-data-journal/contenu-data-paper_10_13143_8r3d-k505/)
- [https://www.ouvrirlascience.fr/wp-content/uploads/2022/04/Guide\\_Partager\\_les\\_donnees\\_web.pdf](https://www.ouvrirlascience.fr/wp-content/uploads/2022/04/Guide_Partager_les_donnees_web.pdf)
- <https://openscience.pasteur.fr/2022/09/14/comment-partager-des-donnees-associees-a-une-publication/>



Source : Infodoc INRA - Data Papers - Octobre 2016

# COMMENT RÉDIGER UN DATA PAPER ?

Selon la revue et ses exigences, le contenu et la taille du data paper peuvent varier considérablement d'une forme très synthétique à un article très complet. Pour savoir comment les rédiger, vous pouvez :

- Suivre les instructions aux auteurs, utiliser les outils de rédaction s'ils existent (certaines revues autorisent la soumission à partir de plateformes externes)
- Vous appuyer sur les articles déjà publiés dans la revue ciblée
- Utiliser les modèles (ou templates) propres aux revues, quand il en existe (par ex *Data in Brief* ).

**Recherche Data Gouv propose un tutoriel permettant de générer une ébauche d'article :**

<https://recherche.data.gouv.fr/fr/page/tutoriel-generer-une-ebauche-de-data-paper>

Fovet-Robot, C. 2015. *Rédiger la lettre d'accompagnement (cover letter) de l'article scientifique en 5 points*. Montpellier (FRA) : CIRAD, 4 p. <http://url.cirad.fr/ist/lettre-accompagnement>

\* Facultatif si données intégrées dans l'article

## La minute "Publier un Data paper"

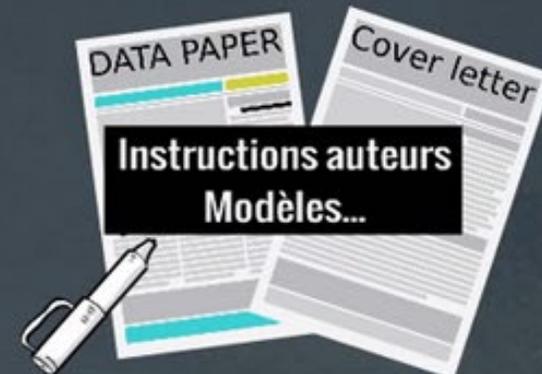
1 Choisir une revue



2 Choisir un entrepôt & déposer les données



3 Rédiger le data paper & la cover letter



4 Lier le data paper aux données déposées



5 Soumettre le data paper



6 Attribuer un identifiant au data paper



# OÙ PUBLIER UN DATA PAPER ?

- Dans un data journal, revue dédiée à ce type de publication : il existe désormais plusieurs data journals en SHS.
- Dans une revue classique qui publie des data papers en plus des articles : plusieurs revues en SHS intègrent déjà, ou souhaitent intégrer, des data papers à leurs formats de publications scientifiques.

**Il est important de bien choisir sa revue** : les revues peuvent imposer une structure spécifique au data paper, un standard de métadonnées particulier, recommander certains entrepôts, imposer une licence de diffusion pour les données, autoriser ou non un embargo sur l'accès aux données... Par ailleurs, certains data journals ne sont disponibles que sur abonnement et d'autres requièrent des APC élevés. (Dedieu, L. 2022. *Publier un Data paper, en 5 points*. Montpellier (FRA) : CIRAD, 5 p. <https://doi.org/10.18167/coopist/0057>)

# ENQUÊTE NATIONALE : UN INTÉRÊT DES COMMUNAUTÉS SHS POUR LE FORMAT



- Une enquête nationale a été réalisée dans le cadre du projet de création d'une revue de données interdisciplinaire en SHS à l'horizon 2025 (cf. journée d'étude MSH Lorraine, mars 2023 : <https://msh-lorraine.fr/nos-services/data-papers/>), en vue d'identifier les besoins des communautés SHS en matière d'articles de données.

<https://msh-lorraine.fr/nos-services/data-shs-la-revue-des-donnees-en-shs/>

- Ses résultats permettent d'établir **le constat d'un intérêt**, avec cependant des disparités significatives selon la discipline et l'institution d'appartenance. Les différentes réponses laissent apparaître **une pratique de l'écriture d'articles relativement faible**, dans des journaux assez variés. Une proportion significative des répondants exprime **un besoin de formations ou un accompagnement supplémentaire** dans la rédaction d'articles de données.

[Rapport de l'enquête :](#)

<https://hal.univ-lorraine.fr/CRENAU/hal-04541094v1>

# ENQUÊTE NATIONALE : FREINS ET MOTIVATIONS

- **Freins** : la **quantification du temps** à y consacrer (manque de temps, méconnaissance des procédures et contraintes techniques du format, un travail important pour rendre les données partageables), alors que la **perception et la reconnaissance** de ce type de publication par les pairs ou l'institution **demeure insuffisante** par rapport aux articles classiques de recherche.
- **Motivations** : **diffuser et valoriser ses données**, puis la **reproductibilité** et le gain de **citation**.

# FOCUS SUR 2 REVUES



- **Cybergegeo** : revue de géographie en libre accès, sans APC, qui publie des data papers

<https://journals.openedition.org/cybergegeo/23412#tocto1n2>

- **Data in Brief** : revue de données multidisciplinaire, en libre accès, avec APC (840\$)

<https://www.sciencedirect.com/journal/data-in-brief/publish/guide-for-authors>

# POURQUOI AVEZ-VOUS PUBLIE UN ARTICLE DE DONNÉES ?



- Pourquoi avoir choisi d'écrire un data paper ? Est-il le complément d'une publication classique ou peut-il la remplacer ? Avez-vous accordé de l'importance à l'impact potentiel du data paper dans l'évaluation de votre recherche ou cherchez-vous surtout à valoriser les données décrites par l'article ? Vous êtes-vous posé la question d'identifier les usages et profils d'utilisateurs potentiels : la communauté que vous connaissez, un public plus large, voire inconnu !
- Quelles données avez-vous publiées ? Dans quel but ? Avez-vous décrit un contenant (une base de données, une collection), un contenu (le jeu de données) ? Comment avez-vous sélectionné ce jeu de données et à quel niveau de granularité ? Pour vous, tous les jeux de données ont-ils vocation à figurer dans un data paper ?
- Avez-vous du réaliser un travail important de curation ? Avez-vous employé des outils pour gérer les données (curation) que vous vouliez publier, selon les principes FAIR ? Des recommandations et spécifications strictes et précises énoncées par certaines revues comme Cybergeio, ou un modèle, comme proposé par Data in Brief, sont-ils des facilitateurs ?
- Comment s'est déroulé le processus de rédaction et sa publication ? Quels problèmes avez-vous rencontrés ? Qui assure la rédaction de l'article ? Y'a-t-il eu un travail conjoint de rédaction avec des ingénieurs et techniciens dans la rédaction des data papers ? Les frais de publication ont-ils joué un rôle ? Quels conseils donneriez-vous pour se lancer dans la rédaction d'un article ? Quelle méthodologie ?
- Pouvez-vous déjà mesurer un impact sur la réutilisation des données mise à disposition ? Vos données sont-elles plus citées ?

<https://doi.org/10.1371/journal.pbio.1002295.g001>