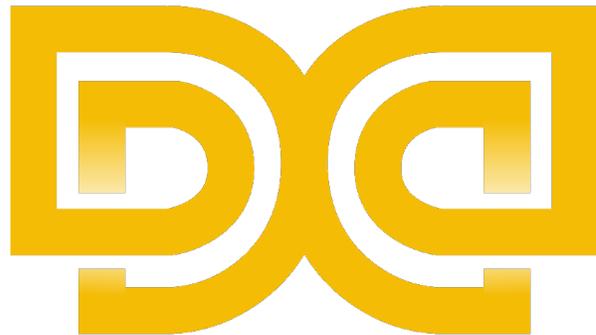


Love Data Week 2025

Université de Lorraine



Data & Corpus

La revue des données en SHS

Julien Muller

Chargé d'édition à la MSH Lorraine

Love Data Week – 14 mars 2025

1

Introduction sur les *data papers*

1

C'est quoi un *data paper* ?

Quelques définitions

« Un *data paper* est un document décrivant un jeu de données et publié dans une revue à comité de lecture. La préparation, la gestion et la description des données prennent du temps ; les *Data papers* permettent à cet effort d'être reconnu et valorisé par l'intermédiaire d'un article académique. »

([GBIF – Système Mondial d'Information sur la Biodiversité](#))

« Un *Data paper* est un article scientifique évalué par les pairs et citable. Il décrit un jeu de données (*data, data set*), la méthode ayant permis de l'obtenir et le potentiel de réutilisation de ce jeu. Le *Data paper* ne décrit pas de résultats de recherche et ne contient ni discussion, ni conclusion. »

(Dedieu, Laurence. 2022. Publier un Data paper, en 5 points. Montpellier (FRA) : CIRAD, 5 p. DOI: <https://doi.org/10.18167/coopist/0057>)

« Un *data paper* est un article sur les données de recherche qui permet de décrire les données produites au cours d'un processus de recherche (méthodologie, process, matériel utilisé, lien vers l'entrepôt où se trouve les données, etc.). Il est revu par les pairs et peut être publié soit dans une revue traditionnelle soit dans une revue dédiée que l'on appelle "data journal". Le *data paper* est citable au même titre qu'une publication classique. »

([Françoise Gouzi – Carnet Hypothèses TOUL'AO](#))

« Les articles de données sont des articles rédigés par un/des auteur(s), examinés par des pairs et citables, publiés dans une revue scientifique, dont le contenu principal est une description des jeux de données de recherche publiés, ainsi que des informations contextuelles sur la production et l'acquisition des jeux de données, dans le but de faciliter la recherche, l'accessibilité et la réutilisation des données de recherche. »

(Joachim Schöpfel – Présentation « Les articles de données (*data papers*) : Une introduction », Love Data Week 2024 – Université de Lorraine)



En bref, un *data paper*, c'est...

Une **description très précise** d'un jeu de données et de la méthodologie

qui favorise l'**accès aux données** et la **réutilisation de ces données**.

C'est un véritable **article scientifique**, soumis à une **validation par les pairs** et **citable**.

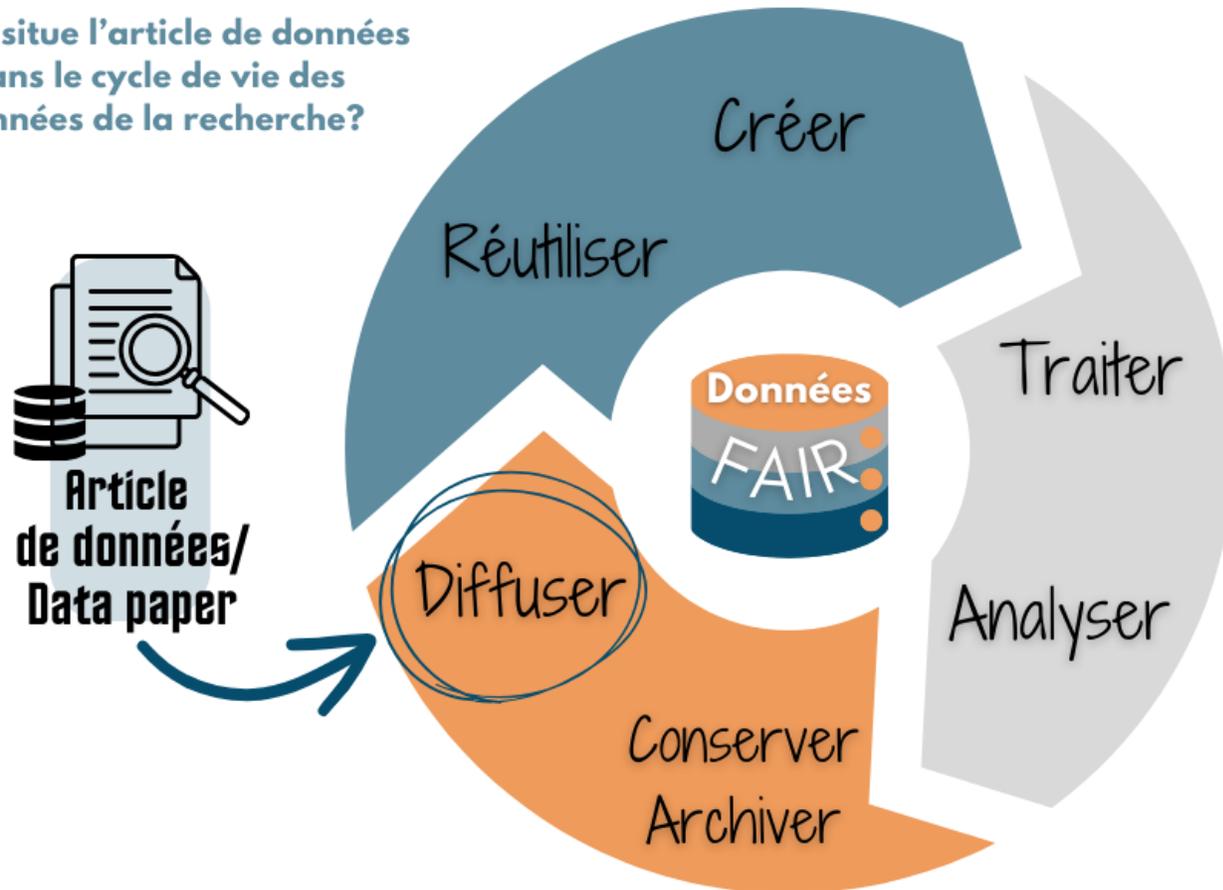
Peut être publié de manière **complémentaire à un article de recherche**.

C'est aussi un moyen de **valoriser le travail des producteurs de données**.

1

Quand publier un *data paper* ?

Où se situe l'article de données dans le cycle de vie des données de la recherche?



Quelques étapes avant de vous lancer dans la rédaction d'un *data paper* :

- Rédiger un plan de gestion de données (PGD ou DMP)
- Documenter finement les données et la méthode tout au long de la vie du projet (métadonnées)
- Déposer les données dans un entrepôt (liste des entrepôts de confiance établie par le CoSo : <https://recherche.data.gouv.fr/fr/entrepots>)

1

Par où commencer ?

Quelques questions importantes à se poser :

- ❑ Mes données peuvent-elles être partagées librement, ou certaines restrictions s'imposent-elles en fonction de la nature de ces données ? (données personnelles, données sensibles, données stratégiques, etc.)
- ❑ Mes données peuvent-elles être réutilisées ? (Sont-elles suffisamment documentées ? Présentent-elles un intérêt au-delà de mon projet de recherche ?)
- ❑ Selon la revue dans laquelle on souhaite publier, prendre connaissance des instructions fournies aux auteurs, et du modèle d'article.

Un outil de création de *data paper* sur recherche.data.gouv.fr : deux modèles fournis (Recherche Data Gouv et Data In Brief).

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

Formation sur ARCHE : *S'initier au processus de publication d'un article de données (data paper)*.

Crédits : Atelier de la donnée ADOC Lorraine. **Licence** : CC-BY 4.0



Création d'un Data Paper

Cet outil vous permet de créer une ébauche de data paper (publication scientifique décrivant un jeu de données) à partir du DOI d'un jeu de données déposé dans l'entrepôt/catalogue entrepot.recherche.data.gouv.fr

Modèle

DOI

2

Le projet de revue

2

À l'origine de *Data & Corpus...*

2022 – Une initiative d'abord locale, initialement circonscrite au territoire lorrain



10 mars 2023 → Journée d'étude : « [Un data journal interdisciplinaire pour les sciences humaines et sociales – Enjeux scientifiques et mise en œuvre pratique](#) »



Constitution d'un groupe de travail composé de chercheur·e·s, d'ingénieur·e·s des données, d'ingénieur·e·s documentaires, d'éditeur·rice·s, d'archivistes, etc.



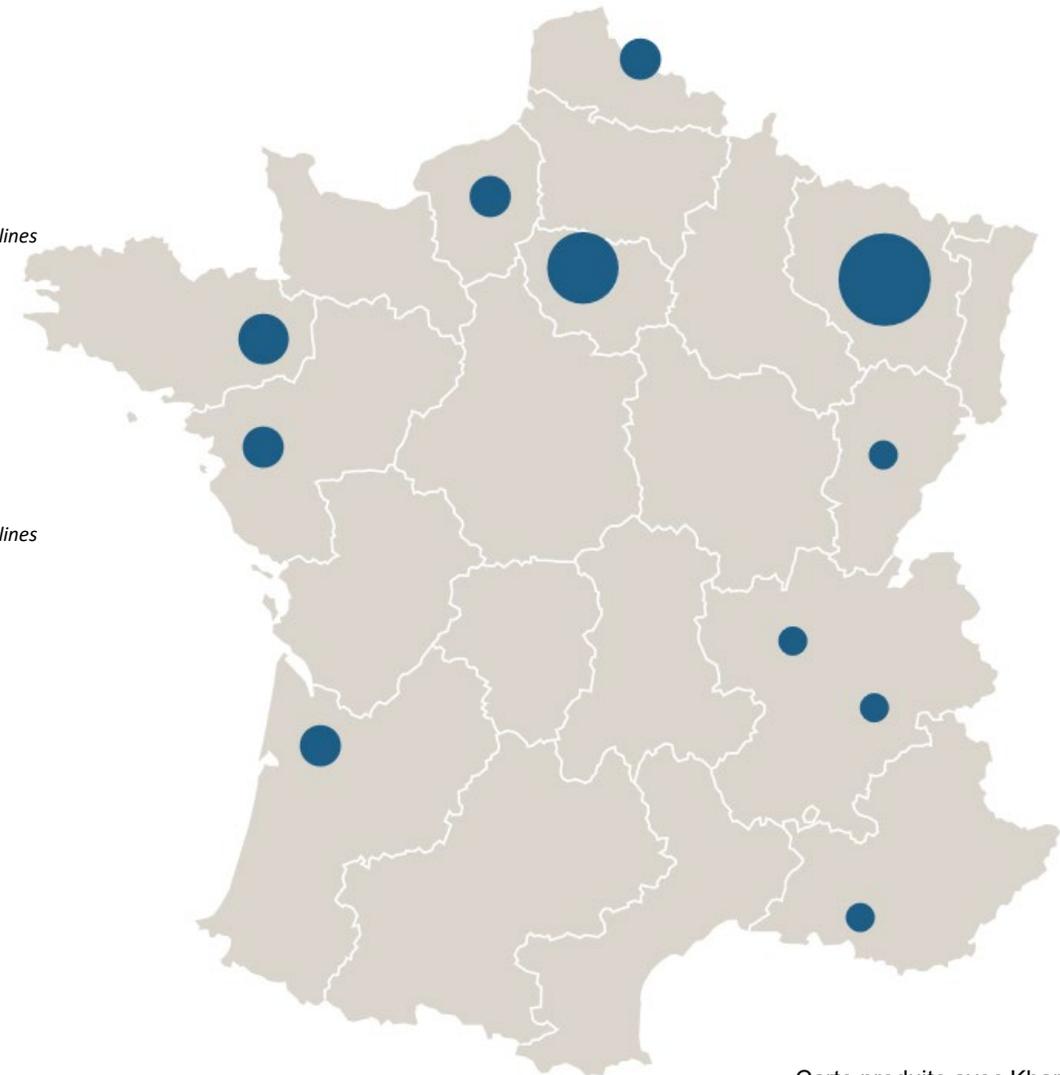
Principales réalisations du **groupe de travail (GT)** :

- ✓ choix de la plateforme de revues ;
- ✓ grande enquête nationale ;
- ✓ réflexion autour de l'identité de la revue : sa ligne éditoriale, son titre, son logo, son mode de fonctionnement, sa chaîne éditoriale, etc. ;
- ✓ recensement de ressources mobilisables pour accompagner les auteur·e·s ;
- ✓ constitution de l'équipe de la revue.

2

Le groupe de travail

Marine Amouroux	<i>CNRS / Université de Lorraine</i>
Fanny Arnaud	<i>CNRS / École Normale Supérieure de Lyon</i>
Nathalie Barré-Lemaire	<i>Université de Lille</i>
Aricia Bassinet	<i>Université de Lorraine</i>
Laurence Bizien	<i>CNRS / École Nationale Supérieure d'Architecture de Nantes / Nantes Université / École Centrale Nantes</i>
Florence Bouchet Moneret	<i>CNRS / Université de Lorraine</i>
Lison Burlat	<i>CNRS / Université Paris-Saclay / École Normale Supérieure de Paris-Saclay / Université Versailles Saint Quentin en Yvelines</i>
Pierre Chiron	<i>CNRS / Université Grenoble Alpes / Sciences Po Grenoble</i>
Véronique Cohoner	<i>CNRS / Nantes Université / Université d'Angers / Le Mans Université</i>
Christine Da Silva-Genest	<i>CNRS / Université de Lorraine</i>
Fiona Edmond	<i>Université Rennes 2</i>
Cyril Heude	<i>Sciences Po Paris</i>
Audrey Knauf	<i>Université de Lorraine</i>
Raphaëlle Krummeich	<i>Université de Rouen Normandie / CNRS</i>
Nathalie Lac	<i>Université Paris Saclay</i>
Flavie Lavallée	<i>CNRS / Université Paris-Saclay / École Normale Supérieure de Paris-Saclay / Université Versailles Saint Quentin en Yvelines</i>
Annaïg Mahé	<i>Ecole Nationale des Chartes</i>
Aurore Montébran	<i>CNRS / École Normale Supérieure PSL Université / Université Sorbonne Nouvelle</i>
Julien Muller	<i>CNRS / Université de Lorraine</i>
Arnaud Natal	<i>Université de Bordeaux</i>
Karine Onfroy	<i>CNRS / Université de Bordeaux / INRAE</i>
Élodie Papin	<i>Université de Lorraine</i>
Pierre Peraldi-Mittelette	<i>Université de Lorraine</i>
Aemil Querre	<i>Université de Lorraine</i>
Stéphane Renault	<i>CNRS / Aix Marseille Université / Ministère de la Culture</i>
François-Claude Rey	<i>Université de Franche-Comté / Université Bourgogne Franche-Comté</i>
Sébastien Rey-Coyrehourcq	<i>CNRS / Université de Rouen Normandie</i>
Johanna Sauvage	<i>Université Rennes 2</i>
Joachim Schöpffel	<i>Université de Lille</i>
Stela Suhan	<i>Université Paris Saclay</i>
Florence Thiault	<i>Université Rennes 2</i>
Frédéric Vergnaud	<i>CNRS / MinesParis-PSL</i>



3

L'enquête

Rapport d'Enquête

Création d'une revue d'articles
sur des jeux de données
Data Journal SHS

Laurence BIZIEN
Responsable du centre documentaire, Référente Science Ouverte, Coordinatrice du groupe de
documentalistes Lab&doc
Nantes Université, ENSA Nantes, École Centrale Nantes, CNRS, AAU-CRENAU, UMR 1563, F-44000
Nantes, France

Véronique COHONER
Responsable de la Cellule d'édition électronique de revues, publications en libre accès - IR CNRS, Pôle revues
grand ouest, MSH Ange-Guépin, Nantes Université, CNRS

Fiona EDMOND
Responsable Données de la recherche, Co-responsable de l'atelier de la donnée ARDoISE, Département
Recherche et science ouverte
SCD, Université Rennes 2

Arnaud NATAL
Economiste
Bordeaux Sciences Économiques - UMR-CNRS 6060, UMR-INRAE 1441, Université de Bordeaux

Pierre PERALDI-MITTELETTE
Chargé d'édition de corpus numérique et non numérique,
Institut François Génys (UR 7301) - Université de Lorraine
Chercheur en ethnologie affilié à l'Institut Convergences Migrations (2021-2025), et membre associé du
LACNAD (UR4092)

Le rapport d'enquête complet est accessible sur HAL :
<https://hal.univ-lorraine.fr/hal-04541094>

Un travail réalisé par le sous-GT Enquête :

- Laurence Bizien,
- Véronique Cohoner,
- Fiona Edmond,
- Arnaud Natal,
- Pierre Peraldi-Mittelette

L'ensemble des analyses
présentés dans cette partie sont
extraites du rapport produit par
le sous-GT Enquête.

3

Objectifs de l'enquête et calendrier

- Affiner le projet
- Proposer une revue de données qui réponde au plus près des besoins des communautés de recherche en SHS
- Qualifier et quantifier les connaissances de ces communautés en matière d'articles de données (*data papers*)
- Recueillir le ressenti et l'intérêt des répondants pour notre projet
- Faire connaître le projet

Sept-Oct 2023

Élaboration du questionnaire

Oct 2023

Implémentation dans *LimeSurvey*

Nov-Déc 2023

Diffusion de l'enquête

Janvier-Mars 2024

Dépouillement et analyse des résultats

3

Le questionnaire

Questionnaire assez court :

- 36 questions (dont 11 optionnelles)
- temps de réponse moyen estimé à 15 minutes

Quatre grandes parties

Vous êtes ?

Questions d'identification (statut, discipline, établissement)

La découverte de l'article de données

Évaluation du niveau de connaissance des répondants concernant les articles de données

Intérêt de la publication d'articles de données

Évaluation de l'intérêt des répondants pour les articles de données

Projet de revue sur les (jeux de) données

Évaluation de l'intérêt des répondants pour le projet de création d'une revue publiant des articles de données

La population en quelques chiffres

- 510 répondants
- Une forte majorité d'enseignants-chercheurs et chercheurs (42,7 %)
- 25,3 % d'ingénieurs de recherche et ingénieurs d'études
- 19,6 % de doctorants

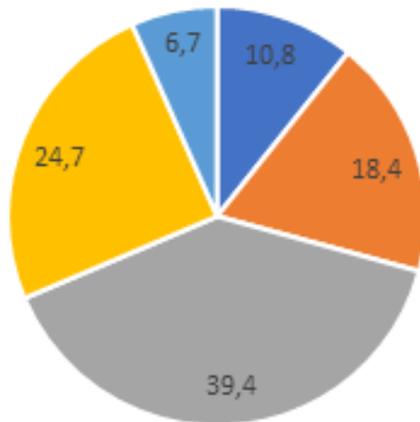
Certaines disciplines sont davantage représentées :

- 72 sociologues et démographes (14 %)
- 56 géographes (11 %)
- 31 psychologues (6,1 %)
- 30 anthropologues et ethnologues (5,9 %)

On trouve aussi parmi les répondants des spécialistes en archéologie, en architecture (urbanisme et aménagement de l'espace), en art, en histoire de l'art, en droit, en histoire, en sciences de l'éducation, en humanités numériques, en langues étrangères, en littérature et en science politique.

Plus de 80 % des répondants sont impliqués dans des pratiques interdisciplinaires, principalement entre disciplines des SHS (pour 348 répondants), et plus rarement avec des disciplines des STM (pour 143 répondants).

Savez-vous ce qu'est un article de données ?



- Je n'en ai jamais entendu parler
- J'en ai entendu parler, mais je ne sais pas ce que c'est
- J'en ai entendu parler et je vois à peu près de quoi il s'agit
- Oui je sais parfaitement de quoi il s'agit
- Manquants

62 % des répondants, dont une part importante de doctorants, manifestent un intérêt pour des formations et de l'accompagnement.

Des disparités en fonction des statuts, disciplines, établissements d'appartenance :

- Des doctorants moins sensibilisés
- Une moindre connaissance en anthropologie, en psychologie et en langues étrangères
- Une meilleure connaissance de la part des membres d'établissements publics à caractère scientifique et technologique

Parmi les 327 répondants ayant connaissance des articles de données, ils sont 130 (soit 40 %) à en consulter plus ou moins régulièrement.

Sur l'ensemble du panel, près de 20 % (101 répondants sur 510) en consultent au moins une fois par mois.

Motivations des lecteurs

Parmi les 130 répondants qui consultent des articles de données, les principales motivations sont :

- Étudier les bonnes pratiques en matière de gestion de données (70 %)
- Réutiliser des données (58 %)
- Reproduire des résultats de recherche (22 %)
- Autres motivations (11 %)

3

Partie 3 - Intérêt de la publication d'articles de données

9 % des enquêtés ont déjà publié un article de données. Forte représentation des géographes.

46 % envisagent la rédaction d'un article de données à plus ou moins long terme (14,3 % dans un délai de moins d'un an).

Motivations à publier	Freins à la publication
Diffuser et valoriser les données de la recherche	Manque de temps
Favoriser la reproductibilité de la recherche	Méconnaissance du format
Rendre les données citables et interopérables	Contraintes techniques
S'approprier les standards de la science ouverte	Travail important pour obtenir des données partageables (problème des données personnelles et des données sensibles notamment)
Faire reconnaître une expertise méthodologique	
Favoriser l'interdisciplinarité	
Produire une nouvelle forme de connaissance	Reconnaissance insuffisante
Optimiser son travail de recherche (évolution de carrière, opportunités de collaboration, etc.)	Un format encore très associé aux « sciences dures »
Diversifier les supports de publication et ouvrir son audience	Crainte du « vol » des données et de leur mauvais usage
Répondre aux exigences de l'évaluation de la recherche, à la demande de financeurs ou de son laboratoire	Un type de publication parfois perçu comme davantage destiné aux « ingénieurs » qu'aux chercheurs

3

Partie 4 - Projet de revue sur les (jeux de) données

81 % des répondants sont intéressés par le projet de création d'une revue interdisciplinaire sur les données en SHS, en libre accès et sans frais pour les auteurs et les lecteurs (modèle diamant).

- 77 % des répondants intéressés en tant que lecteurs
- 57 % des répondants intéressés en tant qu'auteurs potentiels

63 % des répondants souhaitent une diversité linguistique au sein de la revue.

Intérêt pour différents types d'articles

Articles de données (<i>data papers</i>)	72,4 %	Articles sur les articles de données	37,1 %
Articles de recherche portant sur des outils, des méthodologies, etc.	60,6 %	Annonces (appels, recherches de coauteurs, etc.)	28,5 %
Articles de veille (articles de formation, comptes rendus d'événements, notes de lecture)	47,3 %	Articles exécutables	27,4 %
Articles de corpus, décrivant des collections d'objets physiques	46,1 %	Lexique	27,1 %

3

En conclusion

Un intérêt qui paraît plus marqué au sein de certaines disciplines acculturées à la pratique du partage de données et au travail en équipe (ex. : géographie et archéologie).

Les freins identifiés pour la publication d'articles de données sont généralement les mêmes que ceux qui existent en SHS concernant le partage des données (notamment manque de temps, de compétences et de reconnaissance académique).

Distinction SHS / STM concernant le travail à fournir pour rendre les données partageables et réutilisables. D'où la nécessité d'un accompagnement des auteurs/producteurs de données, non seulement dans la rédaction d'articles de données (formations, modèles fournis, etc.), mais également, plus en amont, dans le traitement de leurs données (FAIRisation, aide au dépôt, conseils relatif aux droits des données, etc.).

Quelques précautions d'usage concernant l'interprétation des résultats

- Inégale représentation des différentes disciplines SHS : biais relatif aux voies par lesquelles a été diffusé le questionnaire (listes de diffusion mobilisées par les membres du GT).
- Échantillon trop restreint pour tirer des conclusions d'ordre général sur des différences observées entre disciplines, types d'établissements, etc.
- Biais d'autosélection : il y a certainement une surreprésentation dans notre panel d'individus sensibilisés et intéressés par les articles de données.

4

La revue *Data & Corpus*

4

La revue

- Une revue de données qui s'adresse à l'**ensemble de la communauté SHS**
- Promotion de l'**interdisciplinarité**
- Philosophie de promotion de la **science ouverte** portée par l'Université de Lorraine (modèle diamant, ouverture des données et réutilisation)
- **Le cœur de la revue** : des articles de données (*data papers*), mais aussi des articles de recherche autour de l'innovation en termes d'outils et de méthodologie
- **Langues de publication** : français, anglais
- Une revue portée par l'**Université de Lorraine** et accessible sur la plateforme de revues **Episciences**

4

L'équipe de la revue : Comité de rédaction

MEMBRES	AFFILIATIONS	PRINCIPALES MISSIONS
<p>Directrice Florence Thiault</p>	<p>Université Rennes 2</p>	<p>Coordonne la revue Pilote le comité de rédaction</p>
<p>Directrice adjointe Audrey Knauf</p>	<p>CREM – Université de Lorraine</p>	
<p>Équipe de rédaction Cyril Heude Pierre Peraldi-Mittelette François-Claude Rey Christine da Silva-Genest</p>	<p>Sciences Po Paris IFG – Université de Lorraine Université Bourgogne Franche-Comté UMR ATILF – Université de Lorraine / CNRS</p>	<p>Définit la ligne éditoriale de la revue Détermine les thématiques des numéros Assure une première lecture des articles avant la phase d'expertise Prend la décision finale sur la publication des articles (au regard des évaluations rendues) Recherche des fonds et promeut la revue Assure le référencement de la revue dans les BDD bibliographiques</p>
<p>Chargé d'édition Julien Muller</p>	<p>MSH Lorraine – Université de Lorraine / CNRS</p>	<p>Publie les appels à contribution Assure le suivi éditorial des articles publiés dans la revue Gère les contrats auteurs</p>

4

L'équipe de la revue : Comité scientifique 32 expert·e·s scientifiques

Anthropologie, ethnologie, sociologie

Noël Jouenne
Sébastien Plutniak
Florence Weber

Archéologie

Vincenzo Capozzoli
Loïc Jeanson

Démographie

Xavier Thierry

Droit

Christophe Quézel-Ambrunaz
Vincent Rivollier

Économie

Victor Gay
Stéphane Jugnot

Études littéraires

Pascal Brissette

Études urbaines

Anne-Lise Humain-Lamoure

Géographie

Sébastien Gadat
Zakariyao Koumou
Francesca Di Pietro

Histoire

Christelle Loubet

Humanités numériques

Frédérique Bordignon
Jovan Kostov
Florent Laroche

Linguistique

Aurore Montébran
Frédéric Pugniere-Saavedra

Philologie

Gianluca Valenti

Psychologie

Melissa Arneton

Science des données

David Reymond

Sciences de gestion

Antoine Chollet

Sciences de l'éducation

Amélie Duguet
Muriel Epstein

Sciences de l'information et de la communication

Fanny Georges
Joachim Schöpfel

Science politique

Fabienne Greffet

Traitement automatique des langues

Delphine Bernhard
Juan-Manuel Torres

PRINCIPALES MISSIONS

Évalue les articles : **évaluation en simple aveugle**
Selon les articles proposés, recherche des évaluateurs à l'extérieur de la revue

Propose des numéros thématiques
Peut proposer des évolutions de la revue au comité de rédaction

4

L'équipe de la revue : Comité scientifique 13 expert·e·s des données

Lorraine Adam

PROGEDO

Fanny Arnaud

École normale supérieure de Lyon

Flora Badin

UMR Laboratoire Ligérien de Linguistique – CNRS / Université de Tours / Université d'Orléans

Myriam Carrère

INRAE

Jean-Damien Généro

UMR Centre de recherches historiques – CNRS / EHESS

Justine Lascar

UMR Interactions, Corpus, Apprentissages, Représentations – CNRS / ENS de Lyon / Université Lumière Lyon 2

Delphine Montagne

UMR TRansitions Energétiques et Environnementales – Université de Pau et des Pays de l'Adour

Margaux Nguyen Ngoc Minh

MSH Paris Saclay

Sandrine Ollinger

UMR Analyse et Traitement Informatique de la Langue Française – CNRS / Université de Lorraine

Karine Onfroy

Université de Bordeaux

Frédérique Quidu

École des hautes études en santé publique

Marie-Laure Trémélo

CNRS Sciences Humaines et Sociales

Pierre Willaime

UMR Archives Henri-Poincaré – Université de Lorraine / Université de Strasbourg

PRINCIPALES MISSIONS

Évalue les articles : **évaluation en simple aveugle**

Selon les articles proposés, recherche des évaluateurs à l'extérieur de la revue

Propose des numéros thématiques

Peut proposer des évolutions de la revue au comité de rédaction

4

L'équipe de la revue : Réseau partenaire

Aricia Bassinet	Université de Lorraine
Laurence Bizien	CNRS / École Nationale Supérieure d'Architecture de Nantes / Nantes Université / École Centrale Nantes
Florence Bouchet Moneret	Université de Lorraine
Pierre Chiron	Université Grenoble Alpes
Véronique Cohoner	CNRS / Nantes Université / Université d'Angers / Le Mans Université
Fiona Edmond	Université Rennes 2
Arnaud Natal	Université de Bordeaux
Élodie Papin	Université de Lorraine
Aemil Querre	Université de Lorraine
Stéphane Renault	CNRS / Aix-Marseille Université
Johanna Sauvage	Université Rennes 2
Stela Suhan	Université Paris-Saclay

PRINCIPALES MISSIONS

Joue un rôle de **relai d'informations** pour la revue
Peut faire remonter des propositions au comité de rédaction

Le comité de rédaction tient informé le réseau partenaire concernant la vie de la revue



En 2025, et au-delà...

Journée de lancement – 31 mars 2025 à Nancy

*« Les articles de données en SHS : enjeux sociétaux
et innovations méthodologiques »*

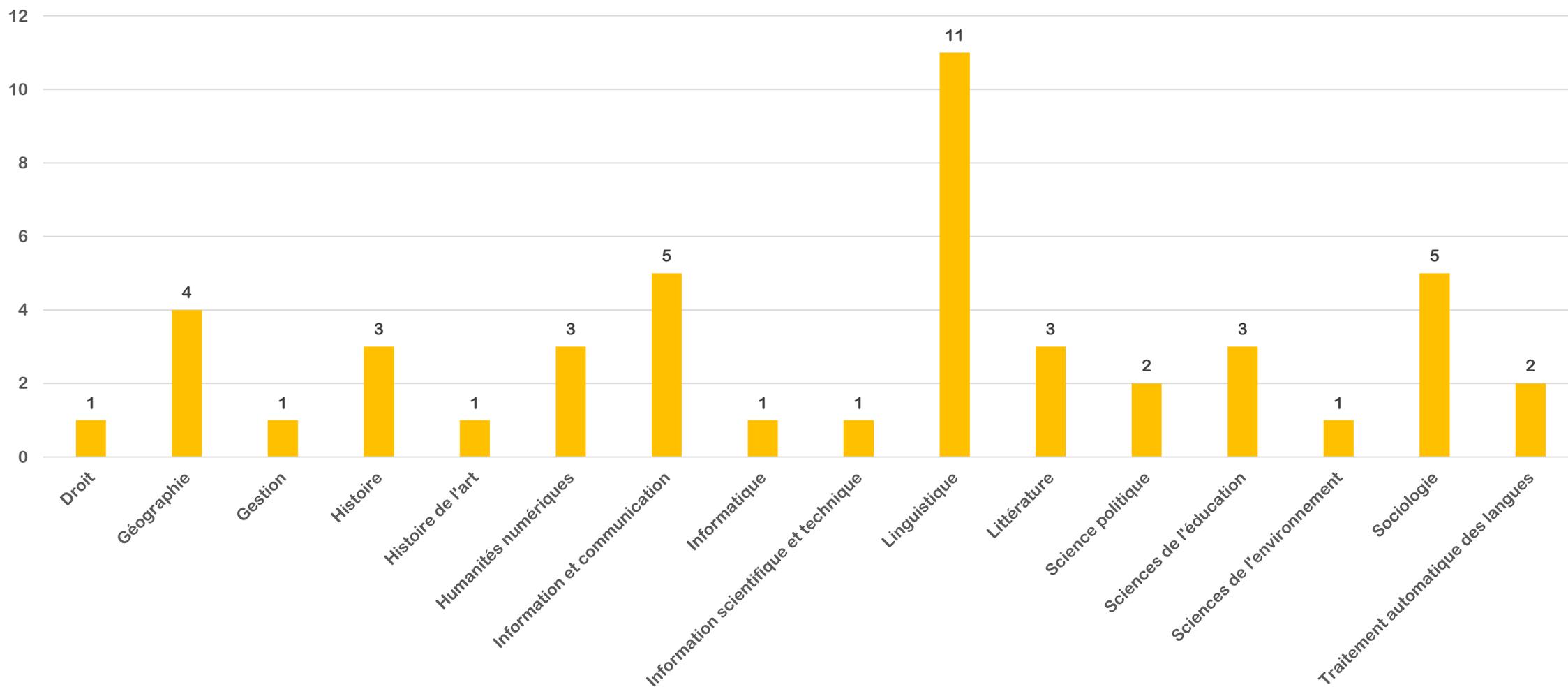
Premier numéro : fin 2025

Second numéro prévu au printemps 2026

4

Aperçu du premier appel à contributions

47 propositions reçues



Quelques ressources

- ❑ Bordignon, F. (2022). *Rédaction, soumission, peer-review : Un retour d'expérience multi-perspectives de la publication d'un data paper*. <https://enpc.hal.science/hal-03714491>
- ❑ *Data papers et data journals*. (s. d.). DoRANum. <https://doranum.fr/data-paper-data-journal/>
- ❑ Dedieu, L. (2022a). *Publier un data paper, en 5 points* (Version 2, p. 5). CIRAD. <https://doi.org/10.18167/COOPIST/0057>
- ❑ Dedieu, L. (2022b). *Publier un Data Paper : Enjeux de bonnes pratiques, d'intégrité scientifique et de science ouverte*. <https://doi.org/10.5281/ZENODO.6817003>
- ❑ Dillaerts, H. (2019, novembre 18). Les data Papers : Quels périmètres, quelles définitions ? Entretien avec Joachim Schöpfel [Billet]. *DLIS*. <https://doi.org/10.58079/nt2i>
- ❑ Gassama, M., Szabo, D., Tang, C., & Bravo, S. (2024). *Analyse de l'enquête sur les pratiques des scientifiques en matière de publication de data paper* (p. 28). INRAE. <https://doi.org/10.17180/vrh7-r606>
- ❑ Gay, V. (2021, novembre). Un data paper en SHS : Pourquoi, pour qui, comment ? #dhnord2021 - *Publier, partager, réutiliser les données de la recherche : les data papers et leurs enjeux*. <https://hal.science/hal-03434216>
- ❑ Kembellec, G., & Le Deuff, O. (2022). Poétique et ingénierie des data papers. *Revue française des sciences de l'information et de la communication*, 24. <https://doi.org/10.4000/rfsic.12938>
- ❑ Le Béhec, M. (2020). Le data paper, un nouvel outil de communication scientifique ? [Text/html]. *Territoires numériques de marques*. <https://doi.org/10.58079/UPEE>
- ❑ Reymonet, N. (2017). *Améliorer l'exposition des données de la recherche : La publication de data papers*. https://archivesic.ccsd.cnrs.fr/sic_01427978v1
- ❑ Schöpfel, J. et Kosmopoulos, C. (dir.). (2025). *Publier, partager, réutiliser les données de la recherche : les data papers et leurs enjeux*. Presses Universitaires du Septentrion. <https://www.septentrion.com/fr/book/?GCOI=27574100316700>
- ❑ Thiault, F. (2024, mars 7). *Comment publier un data paper ?* Stage Urfist / 7 mars 2024, Rennes. <https://docs.google.com/presentation/d/1JaLudSitlME88t4Gbrf6T6zcc9yngM5p>