



HAL
open science

Laurence Bizien, Véronique Cohoner, Fiona Edmond, Arnaud Natal, Pierre Peraldi-Mittelette. Rapport d'Enquête Création d'une revue d'articles sur des jeux de données Data Journal SHS. MSH Lorraine. 2024, 54 p.

Marianne Duquenne

► **To cite this version:**

Marianne Duquenne. Laurence Bizien, Véronique Cohoner, Fiona Edmond, Arnaud Natal, Pierre Peraldi-Mittelette. Rapport d'Enquête Création d'une revue d'articles sur des jeux de données Data Journal SHS. MSH Lorraine. 2024, 54 p.. 2026. <hal-05575463>

HAL Id: hal-05575463

<https://hal.science/hal-05575463v1>

Preprint submitted on 1 Apr 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Laurence Bizien, Véronique Cohoner, Fiona Edmond, Arnaud Natal, Pierre Peraldi-Mittelette. *Rapport d'Enquête Création d'une revue d'articles sur des jeux de données Data Journal SHS*. MSH Lorraine. 2024, 54 p.

DUQUENNE, Marianne (1), Post-doctorante, <https://orcid.org/0000-0003-1493-7905>. Mail : marianne.duquenne@univ-lille.fr

(1) Université de Lille, ULR 4073 - GERiiCO - Groupe d'Études et de Recherche Interdisciplinaire en Information et Communication, F-59000 Lille, France

Situé à l'intersection de l'article scientifique, des métadonnées et du plan de gestion des données, le *data paper* vise avant tout à décrire un jeu de données afin d'en faciliter l'accès et la réutilisation par la communauté scientifique. Ce type de publication s'inscrit plus largement dans le paradigme de la science ouverte, en contribuant à l'ouverture des données de la recherche, à leur documentation, à leur circulation et à leur réutilisation par d'autres chercheurs (Schöpfel et al., 2019).

Toutefois, en sciences humaines et sociales (SHS), l'ouverture des données est marquée par de fortes disparités disciplinaires, tant au niveau des méthodologies de recherche que dans l'usage des infrastructures dédiées à la science ouverte (Larrieu et Schöpfel, 2022 ; Rousseau-Hans et al., 2020 ; Gruson-Daniel et al., 2022). Le *data paper* n'échappe pas à ces disparités, et ce, malgré l'évolution des pratiques de valorisation des données. Ce format demeure inégalement investi en SHS et encore insuffisamment reconnu dans le cadre plus général de l'évaluation de la recherche (Gay, 2025).

Ces constats sont confirmés par le rapport publié par le groupe de travail « Enquête » de Data Journal SHS et soulignent l'intérêt d'enquêtes empiriques pour mieux cerner les connaissances, les usages et les attentes associés aux articles de données en SHS. Au-delà d'un simple état des lieux, ses résultats renforcent la légitimité du projet éditorial de la revue scientifique *Data & Corpus*, consacrée aux articles de données en SHS et portée par la Maison des Sciences de l'Homme de Lorraine. Cette recension propose de revenir sur la démarche de cette enquête et d'en discuter les principaux résultats. Notre contribution s'inscrit dans une réflexion plus large sur la manière dont l'analyse des pratiques scientifiques peut accompagner et renforcer le déploiement de la science ouverte.

Fondée sur un questionnaire national diffusé entre novembre et décembre 2023, l'enquête avait pour double objectif d'identifier les besoins des communautés des SHS en matière d'articles de données et d'ajuster au plus près le projet éditorial de la future revue *Data & Corpus*. Le questionnaire comportait 36 questions. Organisées en plusieurs thématiques, elles visaient à saisir le profil des répondants, leurs connaissances relatives aux *data papers* et aux *data journals*, l'intérêt porté à ce type de publication, ainsi que leurs attentes à l'égard d'une revue consacrée aux jeux de données. La restitution des résultats présente l'intérêt d'offrir à la fois une lecture synthétique des principaux enseignements, organisés selon les grandes thématiques du questionnaire, et une présentation plus détaillée, structurée selon les questions posées.

Sur 915 réponses recueillies, 510 ont été retenues pour l'analyse. Le rapport l'analyse commet un résultat encourageant et un signe d'intérêt des communautés SHS pour ce format de publication, même s'il rappelle que le corpus de réponses ne permet pas de fonder des conclusions généralisables. Notons que l'enquête a enregistré un niveau de participation supérieur à celui de l'étude menée par l'équipe de Recherche Data Gouv portant sur l'évaluation d'un outil d'aide à la rédaction de *data papers* proposé sur l'entrepôt de données

national. Pourtant adressée à une communauté scientifique plus large, celle-ci n'a retenu que 226 réponses sur 427 (Gassama et al., 2024).

Concernant le profil des répondants en SHS, ces derniers appartiennent majoritairement aux milieux de l'enseignement supérieur et de la recherche, avec une forte représentation des enseignants-chercheurs et des personnels d'appui à la recherche. Le panel se distingue par la présence de nombreuses disciplines, particulièrement la sociologie-démographie, la géographie, la linguistique, la psychologie et l'anthropologie-ethnologie. La majorité des répondants déclarent avoir des pratiques de recherche interdisciplinaires, principalement au sein de leur équipe de laboratoire et dans le cadre des projets de recherche auxquels ils prennent part. Enfin, les affiliations institutionnelles relèvent surtout d'établissements d'enseignement supérieur et de recherche, tandis que les bibliothèques, les archives et les musées restent moins représentés.

Si le principe de l'article de données est théoriquement reconnu, sa consultation, comme sa production, reste en marge et varie selon les profils, les disciplines et les affiliations institutionnelles. Lorsqu'ils sont lus, les *data papers* sont avant tout appréhendés comme des outils de cadrage méthodologique, permettant de repérer de bonnes pratiques de gestion des données ou d'envisager la réutilisation de jeux de données. Les chercheurs sont motivés par la diffusion de leurs données et la reconnaissance scientifique associée à ce travail.

Toutefois, on observe des disparités disciplinaires dans la production d'articles de données. La géographie, portée par un pan quantitativiste bien développé, apparaît comme la discipline la plus engagée. Dans les faits, les pratiques se heurtent à des obstacles structurels tels que le manque de temps, le faible niveau de reconnaissance accordé à ce travail et la méconnaissance des normes de description des données. En effet, l'idée que les données en SHS seraient moins "normalisables" persiste, de même que les craintes associées à d'éventuelles réutilisations inappropriées de données sensibles. La faible lisibilité du paysage éditorial constitue un frein supplémentaire, tout comme la difficulté à identifier des revues adaptées pour certaines disciplines. Ces écarts mettent en évidence la nécessité de consolider une culture du partage des données (Gassama et al., 2024), en particulier dans les communautés scientifiques davantage amenées à produire et manipuler des données qualitatives.

L'ensemble de ces freins ne signifie en rien que les chercheurs se désintéressent complètement des articles de données, l'enquête met en évidence un accueil très favorable à la revue *Data & Corpus*. Les répondants s'y projettent en grande majorité, avant tout comme lecteurs, mais aussi comme auteurs potentiels, et, pour certains, comme membres du comité de rédaction.

Les réponses laissent apparaître de fortes attentes en matière d'accompagnement. Les chercheurs reconnaissent à la fois la nécessité d'avoir une revue d'articles de données dédiée en SHS et d'autre part, le besoin d'un modèle d'article détaillé pour les débutants. Pour accompagner les auteurs, la revue propose des modèles détaillés (*templates*), répondant ainsi à la forte demande d'un cadre pour soutenir les pratiques d'écriture. Afin de répondre à la volonté d'ouvrir la revue à d'autres types de productions, son périmètre a été élargi au-delà du seul *data paper*, de manière à intégrer des articles de méthodologie ainsi que des travaux portant sur les outils, les logiciels et les plateformes en lien avec les données de la recherche. Le multilinguisme est également perçu comme un atout supplémentaire, à condition de garantir la qualité des traductions et des métadonnées.

En définitive, l'enquête démontre que le *data paper* constitue un objet encore en cours de stabilisation, dont les contours et les usages continuent de se préciser au sein des communautés scientifiques. Si les articles de données demeurent encore peu présents dans

les pratiques effectives en SHS, ils suscitent néanmoins un intérêt réel, notamment en ce qu'ils répondent aux objectifs de diffusion, de réutilisation et de reconnaissance du travail sur les données. Le champ des SHS se caractérise par des pratiques hétérogènes en matière d'articles de données, ce qui confirme la pertinence de la démarche engagée pour définir la politique éditoriale d'une revue scientifique en accord avec la réalité du terrain et les attentes exprimées.

Bibliographie

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* [Rapport de recherche]. INRAE. <https://hal.science/hal-04690792>

Gay, V. (2025). Un data paper en SHS : pourquoi, pour qui, comment ? Dans C. Kosmopoulos & J. Schöpfel (dir.), *Publier, partager, réutiliser les données de la recherche : les data papers et leurs enjeux*. Presses des Mines. <https://hal.science/hal-03434216>

Gruson-Daniel, C. (2022). *Synthèse de l'étude « Décliner la science ouverte »* [Rapport de recherche]. Comité pour la science ouverte. <https://hal.science/hal-03798580>

Larrieu, M., et Schöpfel, J. (2022). Différences disciplinaires en contexte de Science ouverte. Étude avec les publications de l'archive ouverte HAL. Dans A. Mahé, I. Mayeur, E. Poupardin & C. Prime-Claverie (Dir.), *Communication scientifique et science ouverte: Opportunités, tensions et paradoxes - Actes du colloque « Document numérique et société »*, (pp. 221-230). DeBoeck Supérieur.

Rousseau-Hans, F., Ollendorff, C., et Harnais, V. (2020). *Les pratiques de publications et d'accès ouvert des chercheurs français en 2019*. [Rapport de recherche]. Consortium Couperin. <https://cea.hal.science/COUPERIN/cea-02450324v2>

Schöpfel, J., Farace, D., Prost, H., & Zane, A. (2019). Data papers as a new form of knowledge organization in the field of research data. *Knowledge Organization*, 46(8), 622–638. <https://doi.org/10.5771/0943-7444-2019-8-622>