

DATA & CORPUS, intégrer les métiers de la donnée au cœur de l'édition scientifique

Le projet : lancer une revue Sciences Humaines et Sociales (SHS), nativement numérique et dédiée aux *data papers* en accès ouvert Diamant.



Initié en 2022 par un collectif interdisciplinaire, ce projet vise à combler un « chaînon manquant » éditorial : transformer la donnée de recherche en une contribution scientifique citable et validée.

LE CONSTAT : UN VIDE ÉDITORIAL EN SHS

La Science ouverte a fait de l'ouverture des données un principe central. Pourtant, en SHS, les chercheurs disposent encore de peu de revues adaptées pour publier des *data papers*. Les titres existants sont souvent disciplinaires ou reposent sur des frais de publication, et offrent peu d'espace pour décrire des données hétérogènes : archives électorales, corpus sonores, bases historiques ou collections d'images.

Résultat : faute de contextualisation des données produites, celles-ci restent peu visibles ou difficilement réutilisables dans les circuits traditionnels.

LA MÉTHODE : UN PROCESSUS ÉDITORIAL INVERSÉ

La revue **Data&Corpus** inverse la logique classique de publication : le processus ne débute pas par l'écriture d'un article, mais par le traitement et la mise à disposition de la donnée. Trois étapes structurent ce dispositif.

1. Dépôt et structuration. Les données doivent être déposées dans un entrepôt de confiance (Nakala, Ortolang, Recherche Data Gouv, etc.). L'article ne les héberge pas mais y renvoie via un identifiant pérenne (DOI). L'usage de la taxonomie CRediT permet d'identifier les contributions et d'associer ingénieurs et spécialistes de la donnée à la signature des articles.

2. Soumission via HAL. Avant évaluation, le texte est déposé comme préprint dans l'archive ouverte HAL. Le comité éditorial vérifie sa conformité avant d'engager l'expertise.

3. Evaluation croisée. La spécificité centrale de la revue réside dans une double expertise, réalisée en simple aveugle sous 6 mois :

- un expert scientifique valide la cohérence disciplinaire et la méthodologie
- un professionnel de l'information scientifique et technique analyse la structuration des données, les métadonnées et la conformité aux principes FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable).

GOVERNANCE ET INFRASTRUCTURE

Le projet repose sur une gouvernance inédite où chercheurs, ingénieurs et documentalistes participent ensemble au comité éditorial. La revue est hébergée sur la plateforme Episciences¹, qui permet de développer des « épi-revues » (*overlay journals*) greffées sur les archives ouvertes. Son modèle économique est celui de l'accès ouvert « Diamant » (gratuit pour les auteurs comme pour les lecteurs), un choix éthique soutenu par l'Université de Lorraine et la MSH Lorraine.

BILAN ET PERSPECTIVES

Le premier numéro, publié en décembre 2025, montre la diversité des données pouvant être valorisées : corpus de trans-

criptions d'enfants en psycholinguistique, photographies aériennes anciennes du Rhône pour l'étude des dynamiques fluviales, archives électorales du Web ou encore bases historiques réévaluées pour en analyser les biais méthodologiques.

Lancée officiellement en mars 2025, la revue se veut à la fois espace éditorial et outil de diffusion des bonnes pratiques de gestion des données. Le modèle Diamant limite les coûts financiers mais repose fortement sur l'investissement en temps des équipes.

L'enjeu reste désormais institutionnel : faire reconnaître le *data paper* comme une production scientifique à part entière dans l'évaluation de la recherche. Une évolution que défend notamment la coalition CoARA² en faveur d'une reconnaissance plus large de la diversité des productions scientifiques.

FLORENCE THIAULT
rédactrice en chef de la revue
Data & Corpus
Co-responsable de l'URFIST
Bretagne Pays de la Loire
Université Rennes 2
florence.thiault@univ-rennes2.fr

[1] Consulter : <https://dc.episciences.org>

[2] Consulter : <https://coara.fr>

