



Lier des données et des publications dans COMMONS

Le cas des *data papers*. Pratiques et usages

- Sandra Guigonis, *Directrice adjointe d'OpenEdition*
- Nicolas Larrousse, *Directeur adjoint d'Huma-Num*
- Dominique Roux, *Directeur de l'infrastructure Métopes*

Le projet COMMONS bénéficie d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du Programme d'Investissements d'Avenir portant la référence ANR-21-ESRE-0045





Sommaire

- Introduction & contexte
- Bilan des usages & pratiques observées sur OpenEdition, Sandra Guignonis
- Pratiques de liaisons de données, Nicolas Larrousse
- Pratiques de liaisons dans un contexte d'édition structurée, Dominique Roux





Introduction & contexte

COMMONS, a COnsortium of Mutualised Means for OpenN data & Services for SSH

- Réunit trois infrastructures nationales de recherche : OpenEdition, Métopes, Huma-Num
- Cinq établissements engagés dans la construction de l'équipement





Introduction & contexte

COMMONS, a COnsortium of Mutualised Means for Open data & Services for SSH

- Un environnement numérique unifié
- Accessible au travers d'une offre de services liant données et publications
- Principes de la science ouverte dans le domaine des SHS
- Appuyé sur les principes FAIR.



Lier des données et des publications

Données – Média – Documents – Dispositifs

Un axe de travail qui s'articule en 3 phases :

1. **Citer et lier.** Il s'agit de traiter toutes les modalités de citation et de crosslinking
2. **Représenter**, en traitant d'abord la représentation à plat, puis la représentation dynamique
3. **Use Cases.** Sur la base des avancées des phases 1 et 2, proposer des prototypes d'objets éditoriaux complexes pour dégager des modèles stabilisés de publication, par exemple pour l'édition savante ou les data papers.



Bilan des usages & pratiques sur OpenEdition

1. Place des Data papers dans le corpus d'OpenEdition
2. Cas d'usages & pratiques émergentes sur OpenEdition Journals
3. Outiller des Data papers dans COMMONS ?



Le corpus d'OpenEdition





- **Nature du corpus :** textuel
- **Volume :** 1 161 317 documents (31.12.2022)
- **Couverture spatiale :** monde
- **Couverture temporelle :** 1999 – 2023 (numérique) | 1900-2023 (imprimé)
- **Formats :** HTML (en ligne) | Dublin Core, METS, XML-TEI (dépôt OAI-PMH)
- **Langues :** français (69.5 %), anglais (11.3 %), allemand (10.3 %), espagnol (4 %), portugais (2.3 %), italien (1.9 %), autres langues (0.7 %)
- **Stockage :** BDD au Centre de Calcul de l'In2p3, dépôt OAI-PMH
- **Accès :** Navigation avec Search + métadonnées moissonnables
- **Licence :** CC0 pour les métadonnées + politiques de diffusion & droit d'auteur

Résultats d'une recherche

Une recherche dans Search sur l'expression « Data paper » fait remonter 252 résultats sur les quatre plateformes, que l'on peut affiner.

The screenshot shows the OpenEdition Search interface. At the top, the logo 'OpenEdition Search' is visible on the left, and 'Français' is selected on the right. Below the logo, there are links for 'RECHERCHE AVANCÉE' and 'AIDE À LA RECHERCHE'. The main heading is 'Résultats de la recherche'. Underneath, the search query 'data paper' is entered in a search bar, with a 'RECHERCHER' button and a 'MODIFIER LA RECHERCHE' link. A filter sidebar on the left allows selecting filters for 'SÉLECTIONNER DES FILTRES', 'PLATEFORMES', and 'LANGUES'. The 'PLATEFORMES' section shows: HYPOTHESES (108), OpenEdition JOURNALS (108), OpenEdition BOOKS (26), and CALENDIA (10). The main results area shows '252 DOCUMENTS TROUVÉS' and 'Afficher 12 documents par p...'. The first result is from 'HYPOTHESES' with the title 'Qu'est-ce qu'un "data paper" ?' by 'Françoise Gouzi' from 'TOULAO' (19-05-2017), marked as 'Accès ouvert'. The second result is from 'OpenEdition JOURNALS' with the title 'Cyberge Data Papers' from 'CYBERGEO : REVUE EUROPÉENNE DE GÉOGRAPHIE / EUROPEAN JOURNAL OF GEO...' (2020), marked as 'Freemium Accès ouvert'. The third result is also from 'OpenEdition JOURNALS' with the title 'Présentation de la rubrique Data Papers' from 'CYBERGEO : REVUE EUROPÉENNE DE GÉOGRAPHIE / EUROPEAN JOURNAL OF GEO...' (2017), marked as 'Freemium Accès ouvert'. Navigation controls at the bottom show 'PAGE 1 SUR 21'.

Résultats affinés

	1900-2012	1900-2017	1900-Déc 2022	1900-2023
 Hypothèses	0	22	104	108
 OE Journals - article	0	5	39	42
 OE Books - chapitre	2	4	8	8
 Calenda	0	3	10	10
Total « data paper »	2	34	161	168



Ordres de grandeur



Recherches sur OpenEdition, expression « data paper » : **168** 1 161 317 documents



Recherches sur HAL, type « data paper » : 723 dont **126 SHS** 3 246 036 documents



Recherches dans Nakala, type « data paper » : **19** 542 668 données



Échantillon de data papers en SHS

Sur les 42 résultats dans OpenEdition Journals, 14 sont des formes de data papers parus dans

- *Cybergeo*
- *la Revue française des sciences de l'information et de la communication*
- *Préhistoires méditerranéennes*
- *Humanités numériques*
- *Questions de communication*

Consultation des parutions entre 2015 et 2023 : 21 data papers

(fin février 2023)



Échantillon de data papers en SHS

Titre URL	Domaines disciplinaires	Page présentant la ligne éditoriale	Total	depuis
Cybergeog	Géographie	Présentation de la rubrique Data Papers	11	2017
RFSIC	SIC	Une nouvelle rubrique pour la RFSIC : Le Data Paper	7	2019
Humanités numériques	DH	Éditorial. Décrire des données, publier une éthique	1	2021
Préhistoires méditerranéennes	Préhistoire, archéologie	Recommandations aux auteurs	1	2021
Questions de communication	SIC		1	2017



Mise en œuvre #1

Bideau G., Ysebaert R. (2022). Data paper : Les communes nouvelles françaises (2012-2022) : une méthode pour l'analyse de données à l'échelon municipal selon des limites évolutives. *Cybergeo: European Journal of Geography, Data papers*, document 1027. <https://doi.org/10.4000/cybergeo.39387>

Liaisons | Citation

- [Lien vers Cybergeo Dataverse](#)
- Article cité dans Dataverse

Structure de l'article

Introduction
Objectifs généraux : gérer des structures évolutives
Présentation des différentes sources utilisées
Méthodes d'agrégation des données dans le cas des communes nouvelles
Présentation de quelques traitements possibles
Conclusion
Remerciements

Présentation des données

- [Matériel supplémentaire](#)
- 2 paragraphes pour renvoyer vers le Dataverse et la mention de licence



Mise en œuvre #2

Pernin J.-L. (2022). Intention d'abandonner ses études chez les étudiants et Covid-19 : une application de la théorie du comportement planifié, *Revue française des sciences de l'information et de la communication*. (24).

<https://doi.org/10.4000/rfsic.12259>

Liaisons | Citation

- [Lien vers Zenodo](#)
- Pas de référence à l'article dans Zenodo

Structure de l'article

Phase qualitative
La mesure des concepts
Particularités de l'enquête
Contenu de la base de données
Méthodologie
Résultats exploratoires
Disponibilité et formats de la BDD

Présentation des données

- [Disponibilité et formats de la BDD](#)
- 1 encadré présente les informations principales, dont la licence, et renvoie vers Zenodo



Mise en œuvre #3

Ouvrard L., Degioanni A. (2021). Modélisation de l'extinction de Néandertal, *Préhistoires méditerranéennes*. 9(1)
<https://doi.org/10.4000/pm.3312>

Liaisons | Citation

- [Lien vers Nakala](#)
- Article cité dans Nakala avec lien dans les métadonnées

Structure de l'article

Introduction
Matériel et méthodes
Résultats
Discussion et Conclusion
Données supplémentaires

Présentation des données

- [Données supplémentaires](#)
- Reprise des principales informations et métadonnées : Nom, Emplacement du dépôt, Lien permanent, Date du dépôt, Format, Version du logiciel, Licence

Data in Brief

Open access

Mise en œuvre sur *Data In Brief*

Jiang X.-F., Bai L., Zhao N., Xiong L. (2023). A dataset of dynamical social map in ancient China: 618–1644. *Data in Brief* (46). <https://doi.org/10.1016/j.dib.2022.108804>

Liaisons | Citation

- [Lien vers Mendeley Data](#)
- [Lien vers article de recherche](#)
- Lien depuis l'article dans *Physica*
- Lien des données vers l'article dans *Physica*

Structure de l'article

Value of the Data
1. Objective ; 2. Data Description
3. Experimental Design, Materials and Methods
Ethics, CRediT, Competing Interest
Acknowledgments
Appendix. Supplementary materials
Data Availability

Présentation des données

- [Data availability](#)
- *Specification Table* normée après l'Abstract et les Keywords



Quels outils pour les data papers ?

- La revue présente-t-elle sa ligne éditoriale ?
- Propose-t-elle une définition des data papers ?
- L'enrichit-elle de spécificités disciplinaires ?

« A data paper is a searchable metadata document, describing a particular dataset or a group of datasets, published in the form of a peer-reviewed article in a scholarly journal. Unlike a conventional research article, the primary purpose of a data paper is to describe data and the circumstances of their collection, rather than to report hypotheses and conclusions. »

GBIF – the Global Biodiversity Information Facility



Quels outils pour les data papers ?

- Y a-t-il des recommandations aux auteurs et autrices ?
- La revue fournit-elle un template ?
- Quels jeux de métadonnées met-elle en œuvre pour les data papers ?

Data in Brief only publishes data articles that describe research data you've made publicly available through a repository. These articles ensure that your data, which is often buried in supplementary material, is actively peer-reviewed, curated, formatted, indexed, given a DOI and publicly available to all upon publication. All Data in Brief submissions must contain data produced and owned by either the author or the author's institution. Data in Brief will not accept any submissions that contain datasets with scant variables or samples.

Article Structure

We ask all authors to use our Data in Brief template when preparing their manuscript. All Data in Brief articles must contain the following sections:

- 1 Title, authors and author affiliations
- 2 Abstract
- 3 A completed Specifications Table
- 4 3-5 bullet points that explain why the data is valuable to other researchers
- 5 A direct link to the deposited data
- 6 A complete description of the experimental design and materials and methods
- 7 References

For a quick, easy guide see our ↓ [Data in Brief at a Glance](#)



Quels outils pour les data papers ?

- Y a-t-il des procédures particulières pour l'évaluation ?
- Y a-t-il des recommandations aux évaluateurs et évaluatrices ?
- Leur fournit-on une grille ?

The *Data In Brief* peer review process

You can find out more about how reviewers assess your Data in Brief manuscripts [here](#) ↗ .

Data articles

Because Data in Brief articles are pure descriptions of data they are reviewed differently than a typical research article. The Data in Brief peer review process focuses on data transparency.

Reviewers review manuscripts based on the following criteria:

- Do the description and data make sense?
- Do the authors adequately explain its utility to the community?
- Are the protocol/references for generating data adequate?
- Data format (is it standard? potentially re-usable?)
- Does the article follow the Data in Brief template?
- Is the data well documented?



Quels outils pour les data papers ?

- Quel processus éditorial met-elle en œuvre pour la publication ?
- Comment est assurée la citation réciproque des données et de la publication ?
- Qui assure cette citation ?



Merci !

- Sandra Guigonis, *Directrice adjointe d'OpenEdition*

sandra.guigonis[at]openedition.org

 [0000-0001-7976-1408](https://orcid.org/0000-0001-7976-1408)



Le projet COMMONS bénéficie d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du Programme d'Investissements d'Avenir portant la référence ANR-21-ESRE-0045

