

Les articles de données (*data papers*) Une introduction

Joachim Schöpfel



Définition

Une proposition

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.*

Ils font partie de la gestion des données de recherche et sont liés aux entrepôts de données.

Un exemple

Data in Brief 54 (2024) 110278

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

ELSEVIER

Data Article

Using demographic, psychosocial, behavioural and safety-related factors to assess cyclists' behaviour: A comparative approach across 19 countries

Sergio A. Useche^{a,*}, Francisco Alonso^a, Aleksey Boyko^b, Polina Buyvol^b, Isaac Castañeda^c, Boris Cendales^d, Arturo Cervantes^c, Tomas Echiburu^e, Mireia Faus^a, Zuleide Feitosa^f, Cesareo Fernandez^g, Javier Gene-Morales^a, Bas de Geus^h, Jozef Gnapⁱ, Mohd K. Ibrahim^j, Kira H. Janstrup^k, Ignacio Lijarcio^a, Irina Makarova^b, Miroslava Mikusova^l, Mette Møller^k, Sylvain Ngueuteu-Fouaka^l, Steve O'Hern^m, Mauricio Orozco-Fontalvoⁿ, German Rojas^d, Ksenia Shubenkova^b, Felix Siebert^k, Jose Soto^e, Amanda N. Stephens^o, Yonggang Wang^p, Elias Willberg^q, Phillip Wintersberger^r, Yongzheng Yang^s, Linus Zeuwts^t, Zarir H. Zulkipli^b, Rich McIlroy^u

^aUniversity of Valencia, Valencia, Spain
^bKazan Federal University, Kazan, Russia
^cAaxhuac University, Mexico
^dEl Bosque University, Bogotá, Colombia
^eUniversidad Católica de Chile, Santiago, Chile
^fUniversidade de Brasília, Brasília, Brazil
^gJuane I University, Castellón, Spain
^hUniversity of Leuven, Belgium
ⁱUniversity of Žilina, Bratislava, Slovakia
^jMalaysian Institute of Road Safety Research, Kajang, Malaysia
^kTechnical University of Denmark, Copenhagen, Denmark
^lUniversité de Dschang, Dschang, Cameroun
^mUniversity of Leeds, United Kingdom
ⁿUniversidade de Lisboa, Lisbon, Portugal
^oMasashi University, Melbourne, Australia
^pChang'an University, Chang'an, China
^qUniversity of Helsinki, Helsinki, Finland

* Corresponding author.
E-mail address: sergio.useche@uv.es (S.A. Useche).

<https://doi.org/10.1016/j.dib.2024.110278>
2352-3409/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>)

- 9 pages
- Un article de revue
- Les données sont accessibles sur Harvard Dataverse (avec questionnaire et *codebook*)
- Ce matériel complémentaire se trouve également sur ScienceDirect
- L'article contient également des résultats d'analyse

Un autre exemple

Data paper en humanités numériques : Adressbuch 1854

The Third Life of the Historic Adressbuch of German Migrants in Paris in 1854: A Data paper

Mareike König, mkoenig@dhi-paris.fr (1), Gérald Kembellec, gerald.kembellec@lecnam.net (1,2), Evan Virevialle, evirevialle@dhi-paris.fr (1,3)

(1) Institut Historique Allemand (2) Laboratoire Dicen-IdF, Cnam (EA 7339) (3) Université Panthéon-Sorbonne

Correspondant : gerald.kembellec@lecnam.net

Mots clés : Base de données, Histoire numérique, Allemands, Paris, Migration, XIXe siècle

Keywords : Database, Digital History, Germans, Paris, Migration, 19th century

Mareike König est directrice adjointe et directrice du département Humanités Numériques à l'Institut historique allemand (IHA). Elle est à l'origine du projet « *Adressbuch der Deutschen in Paris von 1854* ». Historienne de l'histoire franco-allemande du XIX^e siècle, ses recherches portent également sur l'Histoire à l'ère numérique et sur la communication scientifique dans les médias sociaux.

Gérald Kembellec est maître de conférences en sciences de l'information et de la communication au Cnam, il anime la thématique « Data, médiation, valorisation » du laboratoire « Dispositifs d'Information et de Communication à l'Ère du Numérique ». Il était détaché au département des humanités numériques de l'IHA en 2020 et 2021.

Evan Virevialle est étudiant en 2^e année du Master Communication du savoir, technologies de la connaissance et management de l'information à l'université Paris 1 Panthéon-Sorbonne. Il est également assistant de recherche à l'IHA, il y collabore notamment au projet « *Adressbuch der Deutschen in Paris von 1854* » sur les aspects de qualification et d'enrichissement des données, d'interface et de cartographie.

Résumé

Ce chapitre présente un *data paper* qui porte sur les jeux de données liés au projet « *Adressbuch der Deutschen in Paris von 1854* » au long cours d'histoire sur l'immigration des Allemands à Paris au XIX^e siècle réalisé à l'Institut historique allemand de Paris. Le projet a pour objectif de mettre à la disposition des chercheuses et chercheurs et généalogistes, dans une interface, les informations sur les individus et entreprises allemands installés à Paris en 1854. Ces informations sont issues d'un document historique, une sorte de bottin de commerce ou pages jaunes des Allemands à Paris.

Outre la description de méthodes de collecte des données, ce *data paper* présente la manière dont les jeux de données sont enrichis et valorisés dans le dispositif de consultation. Nous portons une attention particulière sur le respect des valeurs des humanités numériques et de la

- 17 pages
- Un chapitre d'ouvrage (sur HAL)
- Les données se trouvent sur Zenodo
- Le code source est sur GitHub
- D'autres sites (projet, corpus)
- Avec quelques résultats de recherche

Un troisième exemple

scientific **data**

OPEN

DATA DESCRIPTOR

Check for updates

COVID Border Accountability Project, a hand-coded global database of border closures introduced during 2020

Mary A. Shiraef¹, Cora Hirst², Mark A. Weiss³, Sarah Naseer⁴, Nikolas Lazar^{5,6,7}, Elizabeth Beling⁸, Erin Straight⁹, Lukas Feddern⁹, Noah Rusk Taylor⁹, Cayleigh Jackson⁹, William Yu¹⁰, Aadya Bhaskaran¹¹, Layth Mattar¹², Matthew Amme¹³, Maggie Shum¹⁴, Mary Louise Mitsdarffer¹⁵, Johanna Sweere¹⁶, Susanna E. Brantley¹⁷, Luis L. Schenoni¹⁸, Colin Lewis-Beck¹⁹ & COBAP Team*

Quantifying the timing and content of policy changes affecting international travel and immigration is key to ongoing research on the spread of SARS-CoV-2 and the socioeconomic impacts of border closures. The COVID Border Accountability Project (COBAP) provides a hand-coded dataset of >1000 policies systematized to reflect a complete timeline of country-level restrictions on movement across international borders during 2020. Trained research assistants used pre-set definitions to source, categorize and verify for each new border policy: start and end dates, whether the closure is "complete" or "partial", which exceptions are made, which countries are banned, and which air/land/sea borders were closed. COBAP verified the database through internal and external audits from public health experts. For purposes of further verification and future data mining efforts of pandemic research, the full text of each policy was archived. The structure of the COBAP dataset is designed for use by social and biomedical scientists. For broad accessibility to policymakers and the public, our website depicts the data in an interactive, user-friendly, time-based map.

Background & Summary

Since the first reported case of the novel coronavirus (SARS-CoV-2) crossing international borders (into Thailand on 13-Jan 2020)¹⁻³, policymakers introduced sweeping public health measures (non-pharmaceutical interventions, or NPIs) at international borders—including restrictions on entry altogether. The first country-level restriction on movement across international borders was introduced on 22-Jan 2020 by North Korea. The number

¹Department of Political Science, University of Notre Dame, Notre Dame, Indiana, USA. ²Departments of Biology and Anthropology, Emory University, Atlanta, Georgia, USA. ³Departments of Political Science and Mathematics, Emory University, Atlanta, Georgia, USA. ⁴Department of Quantitative Theory and Methods, Emory University, Atlanta, Georgia, USA. ⁵Department of Economics, Brown University, Providence, Rhode Island, USA. ⁶Department of Sociology, Emory University, Atlanta, Georgia, USA. ⁷Center for Latin American, Caribbean, and Latinx Studies, Vanderbilt University, Nashville, Tennessee, USA. ⁸University Medical Center Hamburg-Eppendorf, Department of Health Economics and Health Services Research, Hamburg, Germany. ⁹Department of Philosophy, Emory University, Atlanta, Georgia, USA. ¹⁰Well-Woodruff School of Nursing, Emory University, Atlanta, Georgia, USA. ¹¹Department of Economics, Emory University, Atlanta, Georgia, USA. ¹²Departments of Biology and Middle Eastem and South Asian Studies, Emory University, Atlanta, Georgia, USA. ¹³COVID Border Accountability Project (COBAP), Austin, Texas, USA. ¹⁴University of Notre Dame, Keough School of Global Affairs, Notre Dame, Indiana, USA. ¹⁵Rutgers University, Department of Childhood Studies, Camden, New Jersey, USA. ¹⁶Charles River Associates, Boston, Massachusetts, USA. ¹⁷Department of Cell Biology, Duke University, Durham, North Carolina, USA. ¹⁸Universität Konstanz, Department of Politics and Public Administration, Baden, Württemberg, Germany. ¹⁹Department of Statistics & Actuarial Science, University of Iowa, Iowa City, Iowa, USA. *A list of authors and their affiliations appears at the end of the paper. [✉]e-mail: nikolas_lazar@brown.edu

14/03/2024

SCIENTIFIC DATA | (2021) 8:253 | https://doi.org/10.1038/s41597-021-01031-5

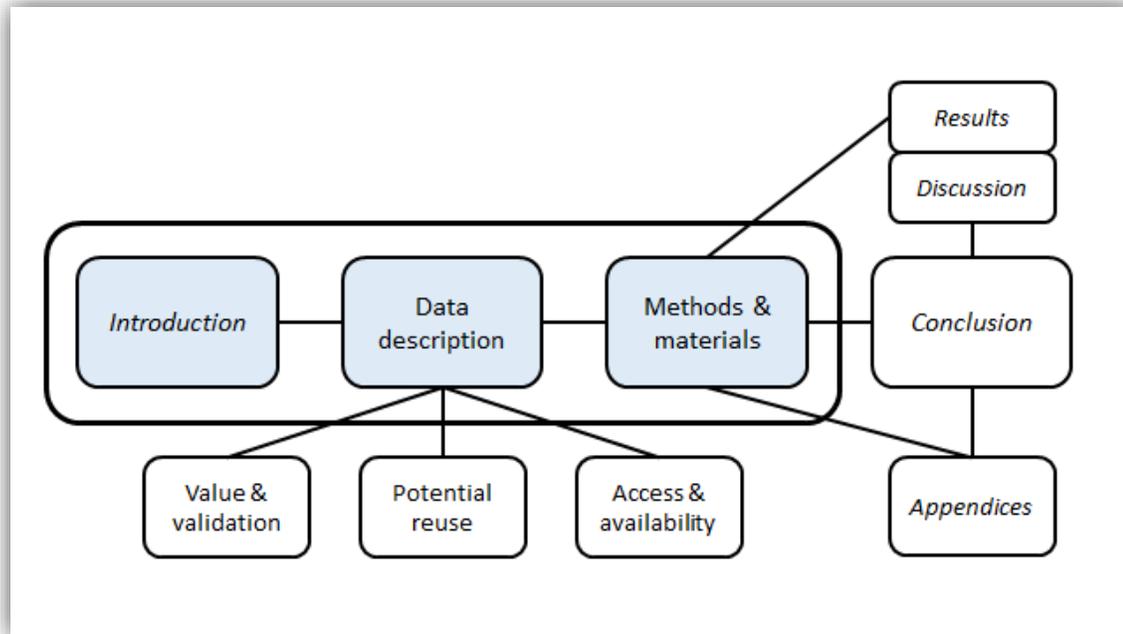
Love Data Week 2024 - Université de Lorraine

1

- 11 pages
- Un article de revue
- La base de données est accessible sur GitHub, avec du matériel complémentaire (enquête...)
- Pas d'analyse, pas de résultats
- Comparaison avec d'autres outils
- Intérêt potentiel pour d'autres

Structure

Structure type



En italique : les sections d'un article de recherche

- Des sections similaires aux métadonnées des données de recherche
 - Cf. DataCite Metadata Schema (ou Dublin Core)
- Intérêt potentiel (utilisateurs, utilisations)
- Conflit d'intérêts, éthique
- Lien vers données, accessibilité, conditions de réutilisation

Avec ou sans résultats ?

- **Pas de résultats, pas de discussion, pas de conclusion**
- Généralement, les directives des revues de données à l'intention des auteurs contiennent ces recommandations ou des recommandations similaires
- *Data in Brief* demande aux auteurs « d'éviter d'utiliser des mots tels que 'étude', 'résultats' et 'conclusions' ».
- **Mais...**
- *Atomic Data and Nuclear Data Tables* laisse aux auteurs le choix d'inclure oui ou non les résultats, la discussion et la conclusion dans la description des données.

Instructions aux auteurs

Instructions type

Format of Data Descriptors

Scientific Data publishes descriptions of datasets under its primary article-type, the Data Descriptor. Please see our [submission guidelines](#) to learn how to draft and format your Data Descriptor, and to download a manuscript template. The main elements of a Data Descriptor manuscript are:

- Title (recommended length 110 characters or fewer)
- Abstract (maximum 170 words)
- Background & Summary (unlimited length)
- Methods (unlimited length)
- Data Records (unlimited length)
- Technical Validation (unlimited length)
- Usage Notes (unlimited length)
- Code Availability
- References
- Figures
- Tables

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 (when applicable)
- 6 A complete description of the experimental design and materials and methods
- 7 References

© 2023 Published by Elsevier Inc.
 This is an open access article under the CC BY license
 (<http://creativecommons.org/licenses/by/4.0/>)

Specifications Table

Subject	Social Science, Environmental Science
Specific subject area	Geography, Planning and Development, Geographical Information System, Landscape Evolution, Spatial Humanities
Type of data	Geospatial dataset
How the data were acquired	The dataset layers were extracted by manual vectorization from the topographic map series K715 of 1969 (1:50,000). All spatial processes (georeferencing, digitization) were carried out using ArcGIS 10.6.1 and a file geodatabase to store the results. Census data were manually extracted from the published printed tables and organized in tabular form using .csv file format. Record matching with their corresponding locations digitized on the map, was performed manually.
Data format	Raw: ESRI shapefile
Description of data collection	Manual vectorization of the road network of Cyprus (22,968 km) categorized in six types, and land cover/use (9,251 km ²) categorized in thirty-three types as explained in the map's detailed legend. The data were collected from the detailed topographical map using ArcGIS 10.6.1, after georeferencing the 25 sheets of the map. Census data were manually digitized from the published tables and were geolocated using the coordinate system of the same map.
Data source location	Region: Eastern Mediterranean Country: Cyprus
Data accessibility	The data files are available in a public repository Repository name: Mendeley Data Data identification number: 836w392t82.1 Direct URL to data: https://data.mendeley.com/datasets/836w392t82/1

La fonction d'un article de données

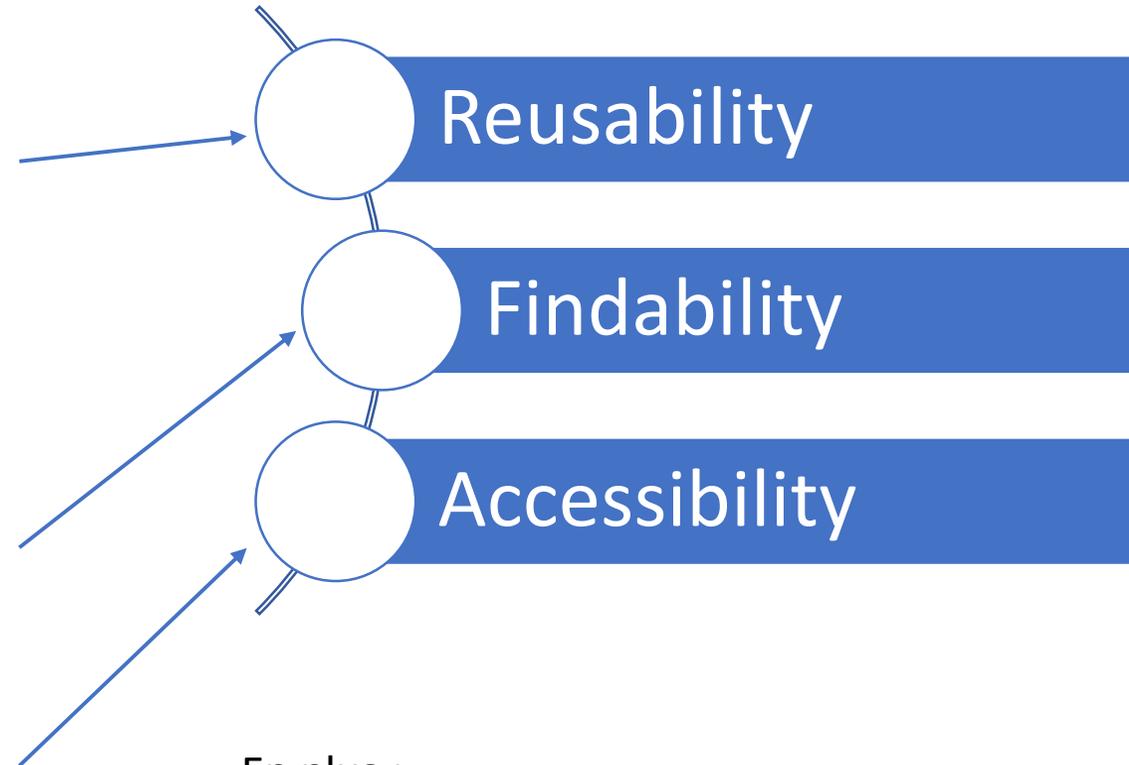
- “*information on the what, where, why, how and who of the data*”
- Description des données
 - D'abord dépôt (conservation) et publication (partage) des données
 - Ensuite publication d'un data paper

Les *data papers* contribuent au respect des principes FAIR

- Findability : ils aident les personnes (et les machines) à trouver les données
- Reusability : ils informent sur la provenance, sur l'intérêt potentiel et sur les droits de réutilisation des données

Scientific Data is founded on six key principles

<p>Credit</p> 	<p>Scientists who share their data in a FAIR manner deserve appropriate credit and recognition. Publishing at Scientific Data:</p> <ul style="list-style-type: none"> • Provides citable, peer-reviewed credit for dataset creation. • Grants recognition to researchers who may not qualify for authorship on traditional articles. • Allows publication of valuable datasets that may not be well-suited for traditional research journals.
<p>Reuse</p> 	<p>Standardized and detailed descriptions make research data easier to find and reuse. Data Descriptors:</p> <ul style="list-style-type: none"> • Provide the information needed to interpret, reuse and reproduce data. • Ensure linking to one or more trusted data repositories where data files, code and/or workflows are stored. • Fulfill a significant part of funders' data-management requirements, particularly by demonstrating and promoting the reuse potential of research data.
<p>Quality</p> 	<p>If released data are to be truly reusable, critical evaluation is needed to verify experimental rigour and the completeness of their description.</p> <ul style="list-style-type: none"> • Focused peer-review evaluates the technical quality and completeness of each Data Descriptor and associated datasets. • Standards are upheld by an academic Editorial Board of recognized experts from a broad range of fields. • Editors and referees ensure alignment with community standards.
<p>Discovery</p> 	<p>Scientists should be able to easily find datasets that are relevant to their research. Content at Scientific Data:</p> <ul style="list-style-type: none"> • Is uniformly searchable and discoverable. • Provides validated links to related data-repository records. • Accelerates integrative analyses by helping authors find relevant datasets across a wide range of different data-types.
<p>Open</p> 	<p>We believe scientists work best when they can easily connect and collaborate with their peers, so Scientific Data aims to:</p> <ul style="list-style-type: none"> • Offer transparency in experimental methodology, observation and collection of data. • Use open licenses that allow for modifications and derivative works. • Break down barriers to interdisciplinary research — facilitating understanding, connectivity and collaboration. • Ensure all interested parties — scientists, policy-makers, NGOs, companies, funders and the public — can find, access, understand and reuse the data they need.
<p>Service</p> 	<p>Scientific Data is committed to providing excellent service to both authors and readers.</p> <ul style="list-style-type: none"> • Authors can deposit datasets in figshare or Dryad during submission, ensuring that datasets can be rapidly peer-reviewed, even when repositories do not exist for the authors' specific data-type(s). • The technology and experience of the Nature Research provides powerful searching of, linking to and visualization of content.

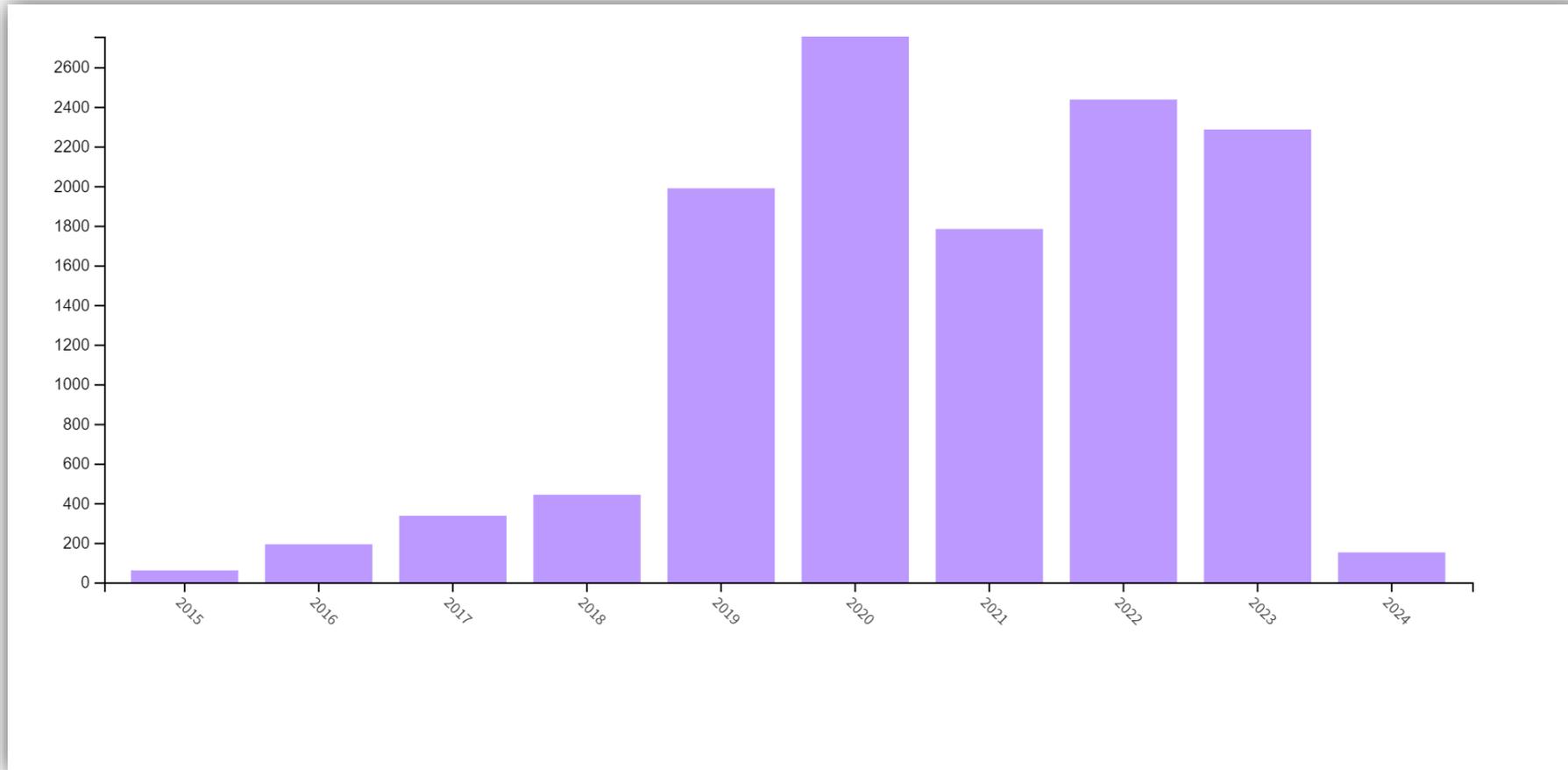


En plus :

- Reconnaissance des auteurs
- Peer review des données associées
- Environnement de service (interopérabilité avec d'autres plateformes)

Revue

Combien d'articles de données ?



N = 12 496 articles

+ 4% depuis un an

Stagnation ?

< 0,1% des articles

Surtout en anglais

Source : Web of Science, 10 mars 2024

Dans quelles revues ?



N = 237 revues

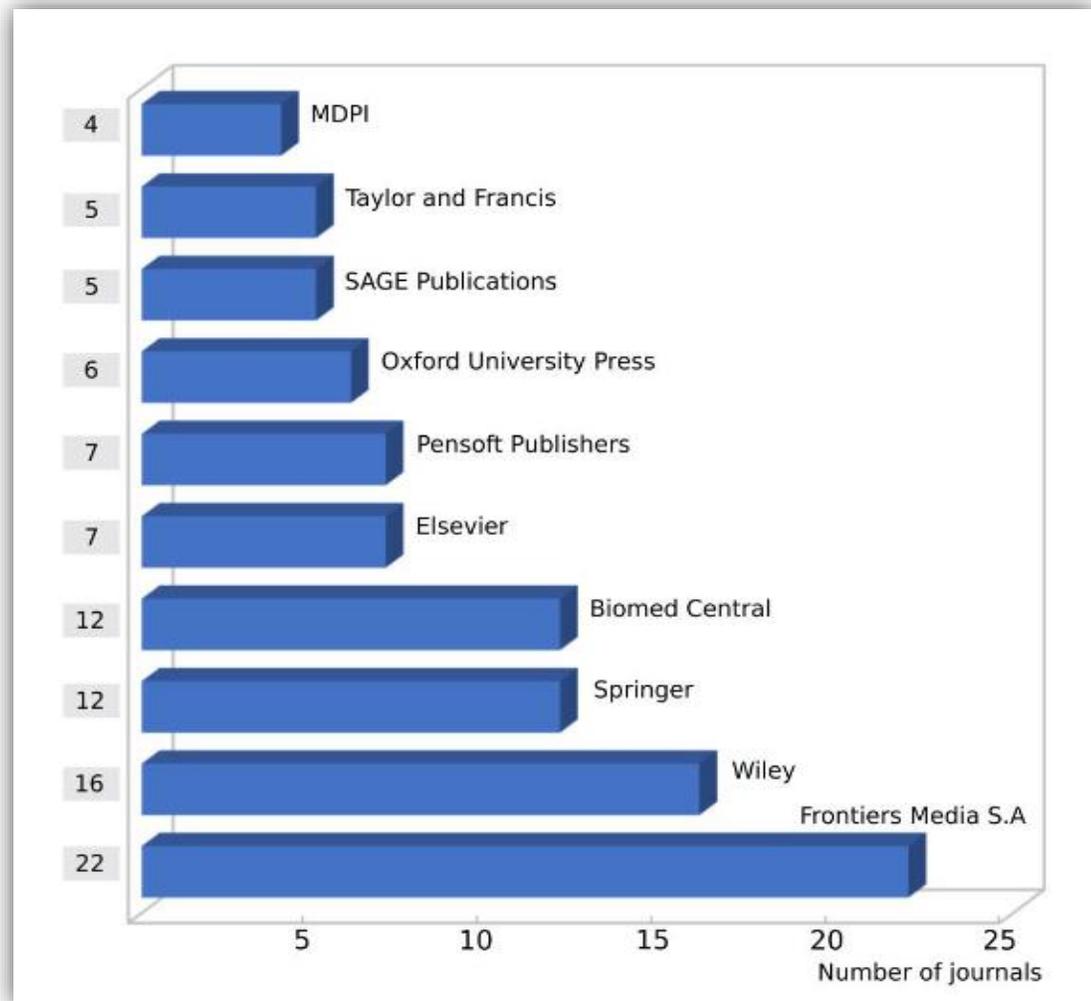
10 revues = 89% des articles

Elsevier & Springer
Nature (plus les autres)

Surtout STM

Des revues de données
Des revues de recherche

Environ 100 revues de données



- Les cinq grands éditeurs commerciaux
- Les éditeurs « open access »

- Modèle d'affaire : Gold avec APC
 - 91% des articles
 - Mais : 84% dans une archives ouverte

Des projets français en SHS

DEMC Journal

- Multidisciplinaire SHS
- Lancement en 2023 (4 articles)
- Sciences Po Paris
- NumeRev
- Modèle diamant
- Peer review
- Quatre rubriques : données, expériences, méthodes, codes
- Pas de lien avec un entrepôt

Data Journal SHS (titre provisoire)

- Multidisciplinaire SHS
- Lancement en 2024 ou 2025
- Université de Lorraine
- Episciences
- Modèle diamant
- Peer review
- Plusieurs types d'articles
- Partenariat avec RDG et/ou Nakala ?

Rédaction

Rédiger un article de données



The screenshot shows a website header for 'factu' with navigation tabs: 'NOTRE ÉTABLISSEMENT', 'NOS FORMATIONS', and 'NOS LABORAT'. Below the header are category buttons: 'À LAUNE', 'Journée internationale des droits des femmes', 'Agenda culturel', 'MT180', and 'A votre santé'. The main content area is titled 'NOTRE ETABLISSEMENT | Publier un article de données ou data paper : retour d'expérience de deux ambassadrices des données'. It includes a publication date 'Publié le 20/02/2024 - Mis à jour le 21/02/2024' and an illustration of a person with a laptop and data charts. The text describes an atelier at ADOC Lorraine and lists two data papers: AK (a corpus of 12 years of publications in R2IE) and MA (optical spectra of skin cancer patients).

factu
l'info de l'université de lorraine

NOTRE ÉTABLISSEMENT | NOS FORMATIONS | NOS LABORAT

À LAUNE | Journée internationale des droits des femmes | Agenda culturel | MT180 | A votre santé

NOTRE ETABLISSEMENT |
Publier un article de données ou data paper : retour d'expérience de deux ambassadrices des données

Publié le 20/02/2024 - Mis à jour le 21/02/2024

L'atelier de la donnée [ADOC Lorraine](#) a recueilli les retours d'expérience d'Audrey Knauf (maitresse de conférences au [CREM](#)) et de Marine Amouroux (ingénieure de recherche au [CRAN](#)), [ambassadrices des données](#) pour leur laboratoire, suite à la publication de leur *data paper* respectif. Envie d'en savoir plus ? Rendez-vous les 11 et 15 mars pendant la [Love Data Week de l'Université de Lorraine](#), dates auxquelles sont organisés des événements sur les *data papers* (sur inscription).

Quelles données avez-vous publiées ?

- AK : Un corpus de données, représentant douze années de publications au sein de la revue internationale d'intelligence économique (R2IE).
- MA : Le jeu de données publié correspond à des spectres lumineux (optiques) acquis sur des patients porteurs de cancers de la peau. Concrètement ces "spectres" prennent la

- Rédaction très standardisée
 - “follow the structure and filling the gap”
 - Utilisation des métadonnées des données
 - Utilisation des fichiers *readme*
- Moins de temps
 - Pas d'état de l'art
 - Pas d'analyse
 - Pas de présentation ou discussion des résultats
- Diversité des formats
- Particularité des SHS

La génération assistée ou automatique



- Les *data papers* peuvent être générés automatiquement
- Les *data papers* sont potentiellement lisibles par une machine
 - Condition : un haut degré de normalisation et d'interopérabilité entre les entrepôts de données, les outils de traitement de texte et les plateformes de revues (format de métadonnées, terminologie, PIDs)
- L'intelligence artificielle facilitera l'automatisation
- Ingestion automatique : cf. Pensoft

Evaluation

Évaluer, oui, mais quoi et comment ?

Évaluer un data paper
retour d'expérience de la revue

cybergeog european journal of geography
revue européenne de géographie

Clémentine Cottineau, CNRS
Christine Kosmopoulos, CNRS
Denise Pumain, Paris 1

Webinaire du groupe de travail inter-reseaux
Atelier Données de la MITI (CNRS)
Jeudi 5 novembre 2020

- Évaluer quoi :
 - le *data paper* ?
 - les données (méthodologie, qualité ?
 - le dépôt des données (FAIRness) ?
- Évaluer comment :
 - en simple aveugle ?
 - par qui ?
- Évaluer pourquoi :
 - pour la qualité de la revue ?
 - pour l'impact des données ?
 - pour la crédibilité des entrepôts ?

Open peer review ?

The screenshot shows the F1000Research interface for an article titled "University students dataset related to achievement, classroom practices, perceptions and attitudes of multimedia-based learning quantum physics". The article is marked as "REVISED" and has 445 views and 40 downloads. It features two invited reviewers: Peter C. Samuels (Birmingham City University, UK) and Nicolas Labrosse (University of Glasgow, UK). The article is currently in "Version 2 (revision)" status, dated 12 Feb 24, with a "read" button. "Version 1" (dated 03 Jan 23) also has a "read" button. The "Open Peer Review" section shows a "Reviewer Status" of "Approved" (green checkmark) and "Reviewer Reports" for both reviewers, with "Invited Reviewers" 1 and 2. The article is available in PDF, XML, and Cite formats, and can be exported, tracked, or shared.

F1000Research de Taylor & Francis

Une revue OA avec des *data notes*

Des relecteurs invités

Commentaires

Impact

Quel impact ?

Biodiversity Data Journal 8: e50451
<https://doi.org/10.3897/BDJ.8.e50451> (27 Mar 2020)

Transdisciplinary Bioblitz: Rapid biotic and abiotic inventory allows studying environmental changes over 60 years at the Biological Field Station of Paimpont (Brittany, France) and opens new interdisciplinary research opportunities

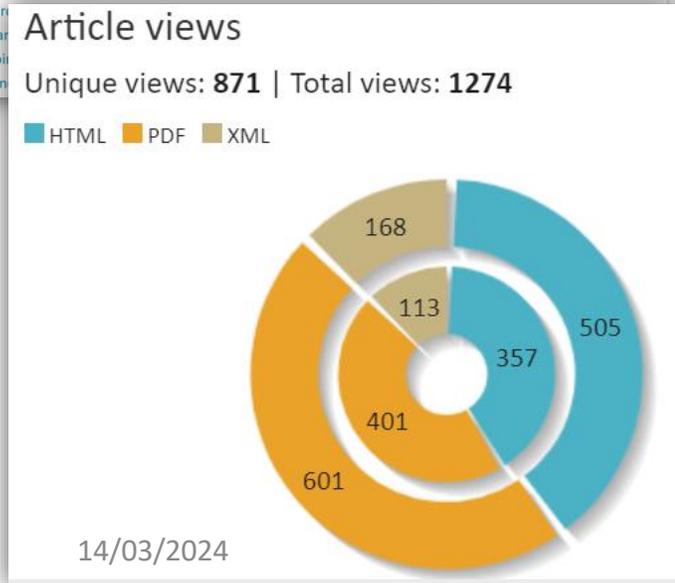
Annegret Nicolaj, Muriel Guernion, Sarah Guillocheau, Kevin Hoeffner, Pascaline Le Gouar, Nelly Ménard, Cheddar, Grégoire, Maximilien

Article views
Unique views: 871 | Total views: 1274

Altmetric
5

Tweeted by 9
17 readers on Mendeley

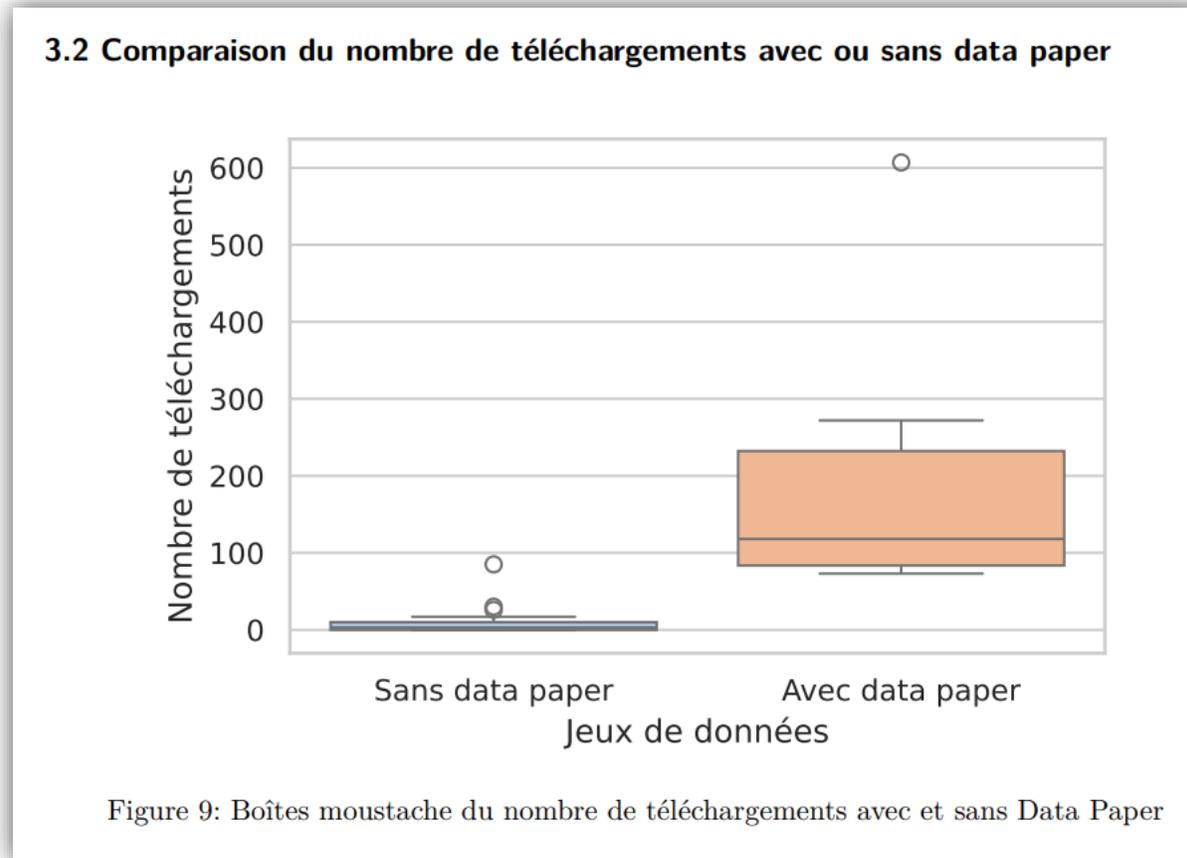
Email alerts for articles from similar subjects
Email



- Consultations
- Téléchargements
 - Mais réutilisation ?
- Citations
- *Dataset ou data paper ?*
- Lien avec richesse des métadonnées ?

Jiao, C., & Darch, P. T. (2020). The role of the data paper in scholarly communication. *Proceedings of the Association for Information Science and Technology*, 57(1), 2–7.
<https://doi.org/10.1002/pr2.316>

De l'évidence empirique (1)



Source : Rémy Decoupes. *Analyse de l'impact des data papers de l'UMR TETIS*. UMR TETIS, 500 rue Jean-François Breton, 34000 Montpellier. 2024. (hal-04428092v2)

De l'évidence empirique (2)



Data Papers provide an Innovative Tool for Information and Data Management

This study seeks to demonstrate how the data paper provides an innovative tool for information and data management, as part of an "ecosystem" of conference proceedings, journal articles, research data and open repositories. It relies upon GreyNet's current collection of 46 published datasets and 16 data papers. The study

highlights the importance of the human contribution for the writing of data papers and the enrichment of their metadata. To this end, key shared components of GreyNet's collection of data papers are discussed, namely the stakeholders, linked metadata, open data archiving, preservation, and issues of quality and information rights. The study concludes from a user perspective by addressing the value of data papers drawn from available statistics. The results are expected to move beyond a simple case study to a use case in which the key components of data papers can be implemented in other communities of practice dealing with grey literature.

Data Papers Extend GreyNet's Document Trail

Document Types	Proposal / Abstract	Conference Poster	Conference Slides	Video Presentation	Conference Paper	Published Dataset	Data Paper Preprint	Data Paper Article
----------------	---------------------	-------------------	-------------------	--------------------	------------------	-------------------	---------------------	--------------------



← Data Papers are a multifaceted Tool

↓ Data Papers increase file Downloads

As of August 31, 2020	With Data Papers	Without Data Papers
GreyNet's Datasets in DANS		
46 Datasets	16 (34.8%)	30 (65.2%)
899 Downloads	423(47.1%)	476 (52.9%)
Average Downloads per Dataset	26.4	15.9

Source : Dominic J. Farace, Joachim Schöpfel. *Data Papers provide an Innovative Tool for Information and Data Management*. GL 2020 Twenty-Second International Conference on Grey Literature "Applications of Grey Literature for Science and Society", Nov 2020, Rome, Italy. (hal-03825967)

Dominic Farace, GreyNet International



Joachim Schöpfel, France



Questions annexes

“Getting merits for creating FAIR research data”

- Curating data = research work ?
- *Data paper* = research article ?
 - En fait, qui rédige un *data paper* ?
 - Travail d’équipe ? Rôle des ingénieurs et bibliothécaires de données ?
 - Même reconnaissance (Hcéres, bilans...) ?
- Process or output ?
 - Evaluer une gestion FAIR ou le résultat ?

(EOSC bootcamp 2021)

Lien avec processus

- Quelles données :
 - données « froides » ?
 - données « chaudes (tièdes) » ?
- Lien avec plan de gestion ?
- Lien avec publication ?
 - Avant, après, à la place ?
- Quels métiers ?
- Quelles fonctions ?
- Quelles activités ?
- Quelles compétences ?



Merci !

joachim.schopfel@univ-lille.fr

Joachim Schöpfel, Dominic J. Farace, Hélène Prost, Antonella Zane. Data papers as a new form of knowledge organization in the field of research data. *12ème Colloque international d'ISKO-France : Données et mégadonnées ouvertes en SHS : de nouveaux enjeux pour l'état et l'organisation des connaissances ?*, ISKO France, Oct 2019, Montpellier, France. ⟨halshs-02284548⟩

Christine Kosmopoulos, Joachim Schöpfel (dir.). *Publier, partager, réutiliser les données de la recherche: les data papers et leurs enjeux*. Presses Universitaires du Septentrion, inPress. ⟨hal-03669965⟩