

Data Documentation: List of Data Journals

Version 1.0 – November 03, 2022

Maxi Kindling^a, Dorothea Strecker^a

a) Berlin School of Library and Information Science, Humboldt-Universität zu Berlin

Abstract

This document describes a dataset that aggregates information about 135 data journals. Data journals focus on the publication of data papers – a specialized publication type describing datasets, their collection and reuse potential that is peer-reviewed, citable and indexed.

This dataset includes a comprehensive list of data journals that was compiled by aggregating existing sources, as well as an overview of these sources.

Citing the data set

Kindling, Maxi; Strecker, Dorothea (2022) List of Data Journals. 2022. Version 1.0: DOI: 10.5281/zenodo.7082126

Licenses



This documentation is licensed under Creative Commons Attribution 4.0 International terms.



The data files are licensed under Creative Commons CC0 1.0 Universal terms.

Motivation

Data citation and the adequate recognition of data publications as first-class research outputs still challenge scholarly communication. Data journals are presented as a solution to these issues. They focus on the publication of data papers – a specialized publication type describing datasets, their collection and reuse potential that is peer-reviewed, citable and indexed.

Currently, there is no consensus on a definition for the term *data journal*, for example with regards to how strong the focus on publishing data papers should be. As a result, data journals are difficult to find in bibliographic resources, and it is not possible to compile a definitive and complete list of data journals, for example because there are no facet filters for journal type. Several studies have already examined the role of data journals in scholarly communication, and named or compiled lists of the data journals under investigation. However, these lists of data journals vary significantly in extent and scope.

In the absence of suitable bibliographic data sources, this dataset combines information from existing sources and applies a set of inclusion criteria to create a comprehensive and coherent list of data journals.

Methodology

Figure 1 describes the approach to compiling the final list of data journals. The following sections outline the process in detail.

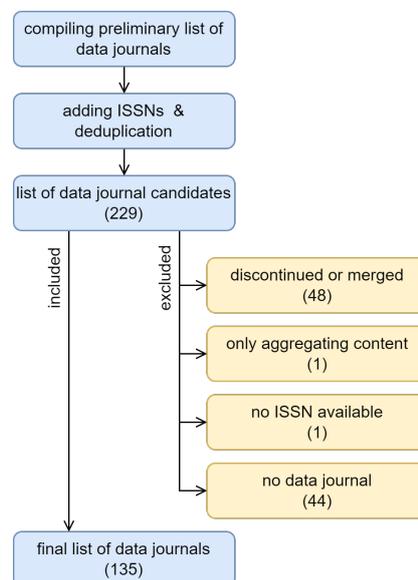


Figure 1: Systematic overview - compiling the final list of data journals

Compiling a preliminary list of data journals

First, a preliminary list of data journals was compiled. The discovery process was centered around 17 sources observing the data journal landscape published between 2013 and 2022 (see Table 1). Journals mentioned in these studies were added to the preliminary list.

Source	Number of data journals listed
Akers (2014)	73
Callaghan (2013)	28
Candela et al. (2015)	116
Finnish Committee for Research Data (2020)	18
Forschungsdaten-Wiki (2020)	32
García-García, López-Borrull, and Peset (2015)	20
Kim (2020)	24
Kong et al. (2019)	5
Li and Jiao (2022)	2
Li (2018)	15
Lindstädt and Rothe (2017)	12
Reymonet (2017)	6
Schöpfel (2019)	81
Seo and Kim (2020)	9
Thelwall (2020)	24
Walters (2020)	128
Ward (2016)	24

Table 1: Existing sources of data journals and the number of data journals listed

Next, websites of publishers represented in this list were searched for additional data journals. Data journals encountered during the discovery process accidentally were also added.

Data processing and application of inclusion criteria

After adding ISSNs, the preliminary list was deduplicated, resulting in a set of 229 data journal candidates. Of these, 223 journals were found via previous studies, an additional 3 were discovered via publisher websites, and 3 by coincidence.

Based on the websites of the 229 data journal candidates, the list was further narrowed down. Journals were removed from the list if they were discontinued or had merged with other journals (48), if they only aggregated content published by other journals (1) or if no ISSN was available (1).

To ensure a consistent understanding of the term *data journal* throughout the final list, websites of all journals on the candidate list were checked manually to ensure they met the following inclusion criteria:

- The journal states that it is a data journal in its scope or description.
- The journal accepts data papers or comparable publication types.

Based on these criteria, 44 journals were not considered data journals and were therefore removed from the candidate list. The remaining 135 data journals comprise the final list of data journals.

Components of the dataset

Final list of data journals

The resulting data file *01_data_journals.csv* contains information on the 135 data journals identified. The following variables are available for all data journals:

- *issn*: ISSN of the data journal (manually added)
- *journal_title*: title of the data journal (retrieved from Crossref on 2022-10-11)
- *publisher*: publisher of the data journal (retrieved from Crossref on 2022-10-11 and harmonized)
- *data_journal_type*: type of the data journal (manually added); either *pure* (journal has a strong focus on publishing data papers) or *mixed* (journal also publishes other publication types)

Sources

The file *02_sources.csv* provides an overview of data journals listed in existing studies. Rows correspond to the 229 journals in the data journal correspondence list. For each journal, the following variables are available:

- *issn*: ISSN of the journal
- *journal_title*: title of the journal
- *reason_for_exclusion*: reason for excluding a journal from the final list of data journals; *NA* if a journal is included; last checked: 2020-10-07
- *akers_growing_2014* – *ward_sources_2016*: Boolean values indicating if a journal is listed in the respective source; *TRUE* if the journal is listed, otherwise *FALSE*; column names correspond to BibTeX keys of the sources in file *03_bibliographic_information.bib*
- *other*: Boolean values indicating if a journal if the data journal was identified on publisher websites or encountered accidentally (*TRUE*; otherwise *FALSE*)

This file documents the process of compiling the final list of data journals by listing all journals mentioned in the given sources and providing a reason if a journal was excluded. It can also be used to analyze the co-occurrence of journals in these sources, for example by generating a similarity matrix – as is visualized in Figure 2.

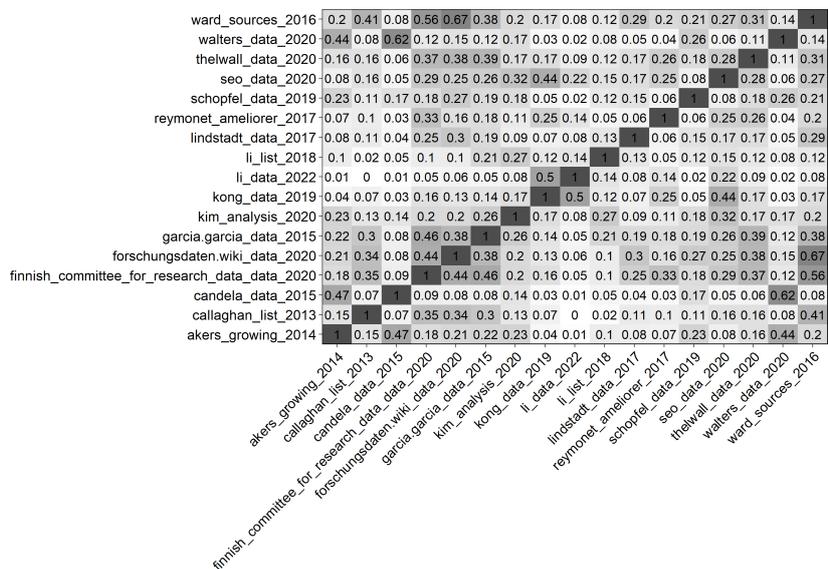


Figure 2: Similarity matrix of journals named in existing sources; similarity is measured using the Jaccard index

Bibliographic information

The file *03_bibliographic_information.bib* provides bibliographic information on the 17 studies on data journals at the center of the sampling process. Entries follow BibTeX specifications. The BibTeX keys are used as column names in the file *02_sources.csv*.

About the project

Compiling a list of data journals was part of a PhD project on quality assurance of research data publications at the Berlin School of Library and Information Science, Humboldt-Universität zu Berlin and the project re3data COREF. re3data COREF is a joint project by DataCite, Helmholtz Open Science Office, Humboldt-Universität zu Berlin, and KIT Library.

The project is funded by the German Research Foundation (DFG) under the project number 422587133.

Sources of data journals

- Akers (2014). *A Growing List of Data Journals*. URL: <https://mlibrarydata.wordpress.com/2014/05/09/data-journals/> (visited on 09/12/2022).
- Callaghan, Sarah (2013). *A list of Data Journals (in no particular order)*. URL: https://www.atsr.rl.ac.uk/frozen_sites/preparde/preparde/blog/DataJournalsList.html (visited on 09/12/2022).
- Candela, Leonardo et al. (2015). "Data journals: A survey". In: *Journal of the Association for Information Science and Technology* 66.9, pp. 1747–1762. DOI: 10.1002/asi.23358.
- Finnish Committee for Research Data (2020). *Data Journal Directory - Data-asioiden kansalliskomitea*. URL: <http://web.archive.org/web/20200812040958/https://www.fcrd.fi/data-journal-directory/> (visited on 09/12/2022).
- Forschungsdaten-Wiki (2020). *Data Journals (as of 2020-01-24)*. URL: https://www.forschungsdaten.org/index.php/Data_Journals (visited on 09/12/2022).
- García-García, Alicia, Alexandre López-Borrull, and Fernanda Peset (2015). "Data journals: eclosión de nuevas revistas especializadas en datos". In: *El Profesional de la Información* 24.6, pp. 845–854. DOI: 10.3145/epi.2015.nov.17.
- Kim, Jihyun (2020). "An analysis of data paper templates and guidelines: types of contextual information described by data journals". In: *Science Editing* 7.1, pp. 16–23. DOI: 10.6087/kcse.185.
- Kong, Lihua et al. (2019). "A Data Quality Evaluation Index for Data Journals". en. In: *Big Scientific Data Management*. Ed. by Jianhui Li et al. Lecture Notes in Computer Science. Berlin: Springer, pp. 291–300. DOI: 10.1007/978-3-030-28061-1_29.
- Li, Kai (2018). *List of curated data papers by GBIF*. note: Dataset. DOI: 10.6084/m9.figshare.6307805.v1.
- Li, Kai and Chenyue Jiao (2022). "The data paper as a sociolinguistic epistemic object: A content analysis on the rhetorical moves used in data paper abstracts". In: *Journal of the Association for Information Science and Technology* 73.6, pp. 834–846. DOI: 10.1002/asi.24585.
- Lindstädt, Birte and Robin Rothe (2017). *Data Journals als Publikationsform in den Lebenswissenschaften*. URL: <http://zbmedblog.de/?p=482> (visited on 09/12/2022).
- Reymonet, Nathalie (2017). *Améliorer l'exposition des données de la recherche : la publication de data papers*. URL: https://archivesic.ccsd.cnrs.fr/sic_01427978 (visited on 09/12/2022).
- Schöpfel, Joachim (2019). *Data for: Data Papers as a New Form of Knowledge Organization in the Field of Research Data*. note: Dataset. DOI: 10.17026/dans-zk3-jkyb.
- Seo, SunKyung and Jihyun Kim (2020). "Data journals: types of peer review, review criteria, and editorial committee members' positions". In: *Science Editing* 7.2, pp. 130–135. DOI: 10.6087/kcse.207.
- Thelwall, Mike (2020). "Data in Brief: Can a mega-journal for data be useful?" In: *Scientometrics* 124.1, pp. 697–709. DOI: 10.1007/s11192-020-03437-1.

- Walters, William H. (2020). "Data journals: incentivizing data access and documentation within the scholarly communication system". In: *Insights* 33.1, p. 18. DOI: 10.1629/uksg.510.
- Ward, Pauline (2016). *Sources of dataset peer review (as of 2020-01-28)*. URL: <https://www.wiki.ed.ac.uk/display/datashare/Sources+of+dataset+peer+review> (visited on 09/12/2022).