

Title of your data paper

First Author Name^{a*}, Second Author Name^{bc}, etc.^{ac}

^aDepartment, University, City, Country

^bDepartment, University, City, Country

^cDepartment, University, City, Country

*Corresponding author: first name, initials, surname; `email.address`

Abstract

A short (ca. 100 word) summary of the dataset being described: what the data covers, how it was collected, how it is stored, and a short description of its reuse potential.

Keywords: keyword 1; keyword 2; lower case except names, max 6

Author roles: For determining author roles, please use following taxonomy: <https://casrai.org/credit/>. Please list the roles for each author.

1 Overview

Repository location Indicate where the dataset repository is located.

Context Was this data produced as part of a research project, thesis, course work, or is this data used in a paper(s)? If so, please list the appropriate bibliographic information here. Note: This journal uses a style based on the APA system (see [here](#)).

The following are some basic citation commands in L^AT_EX:

`\citet` → [Jenset and McGillivray \(2017\)](#)

`\citet` → [Shree, Liu, Gordon, and Hobbs \(2019\)](#)

`\citep` → [\(Fabricius-Hansen & Haug, 2012\)](#)

`\citealp` → [\(Eckhoff et al., 2018; Fabricius-Hansen & Haug, 2012; Shree et al., 2019\)](#)

The following are other basic functions you might find useful:

Bullet points:

- Some point
- Another point

Numbered points:

1. Some numbered point
2. Another numbered point

Adding a footnote¹.

A simple table:

Table 1: A caption.

1	2	3	4
a	b	c	d
e	f	g	h

Please refer to your table as: Table 1.

To add a figure, upload the figure into the `images` folder, and then embed it:



Figure 1: JOHD's logo.

To resize the figure:



Figure 2: JOHD's logo.



Figure 3: JOHD's logo.

Please refer to your figures as: Figure 1, Figure 2, etc.

2 Method

Describe the methods used to create the dataset, including the following sub-headings:

Steps The series of procedures followed to produce the dataset. This should include any source data used, as well as software and instrumentation involved.

Sampling strategy If relevant, please outline the sampling strategy used to produce the data.

Quality control If applicable. Please list the methods used for quality control in the production of the data (i.e., steps taken to normalize the data).

3 Dataset Description

Object name Typically the name of the file or file set in the repository.

Format names and versions E.g., ASCII, CSV, Autocad, EPS, JPEG, Excel, SQL, etc.

Creation dates The start and end dates of when the data was created (YYYY-MM-DD).

Dataset creators Please list anyone who helped to create the dataset (who may or may not be an author of the data paper), including their roles (using and affiliations).

Language Languages used in the dataset (i.e., for variable names etc.).

License The open license under which the data has been deposited (e.g., CC0).

Repository name The name of the repository to which the data is uploaded. E.g., Figshare, Dataverse, etc.

Publication date If already known, the date in which the dataset was published in the repository (YYYY-MM-DD).

4 Reuse Potential

Please describe the ways in which your data could be reused by other researchers both within and outside of your field. For example, this might include aggregation, further analysis, reference, validation, teaching or collaboration. This section should also include limitations to, or potential barriers for reuse.

Acknowledgements

Please add any relevant acknowledgements to anyone else that assisted with the project in which the data was created but did not work directly on the data itself.

Funding Statement

If the research resulted from funded research please list the funder and grant number here.

Competing interests

If any of the authors have any competing interests then these must be declared. If there are no competing interests to declare then the following statement should be present: The author(s) has/have no competing interests to declare.

Notes

¹This is a footnote

References

- Eckhoff, H. M., Bech, K., Bouma, G., Eide, K., Haug, D., Haugen, O. E., & Jøhndal, M. (2018). The PROIEL treebank family: A standard for early attestations of Indo-European languages. *Language Resources and Evaluation*, 52(1), 29–65. DOI: 10.1007/s10579-017-9388-5
- Fabricius-Hansen, C., & Haug, D. T. T. (2012). Co-eventive adjuncts: Main issues and clarifications. In C. Fabricius-Hansen & D. T. T. Haug (Eds.), *Big events, small clauses* (pp. 21–54). Berlin-Boston: De Gruyter.
- Jenset, G. B., & McGillivray, B. (2017). *Quantitative historical linguistics: A corpus framework*. Oxford: Oxford University Press. DOI: 10.1093/lc/fqy046
- Shree, J., Liu, E., Gordon, A., & Hobbs, J. (2019). Deep natural language understanding of news text. In *Proceedings of the first workshop on narrative understanding* (pp. 19–27). Minneapolis, Minnesota: Association for Computational Linguistics. DOI: 10.18653/v1/W19-2403

Supplementary Files (optional)

Any supplementary/additional files that should link to the main publication must be listed, with a corresponding number, title and option description. Ideally the supplementary files are also cited in the main text. Note: supplementary files will not be typeset so they must be provided in their final form. They will be assigned a DOI and linked to from the publication.