

Lana Burke

LIS 545 - Data Curation

Term Project - #1 (Data and Metadata)

Word count (not including headers): 750

Project link: [Humour practices and European attitudes towards democracy - data documentation](#)

Data

What are the data?

A curated list of 315 studies related to contemporary, 21st-century humour practices, attitudes toward democracy, and modes of civic engagement in the six Democratic Literacy and Humour (DELIAH) consortium countries (Belgium, Estonia, Germany, the Netherlands, Slovakia, and Spain). The dataset includes studies published in English or in the official language of the country the study refers to.

Each record in the dataset includes basic bibliographic information (author(s), title, and where and when it was published) as well as classification fields such as “medium of humour practices,” “attitudes toward democracy,” “modes of civic engagement,” “topic of humour,” and “groups.”

Where do the data come from, i.e. what is their origin?

The data originate from a two-step collection process conducted by DELIAH. First, the team carried out a literature review using a combination of narrative and systematic search strategies to account for differences in language and scholarly practices across the six countries. This step used predefined search strings to identify relevant studies. In the second step, DELIAH conducted country-specific, ad hoc searches to capture additional relevant publications (Engelken-Jorge, Moreno, & Castañeda-Zumeta, 2025, pp. 3–4).

Who are the key stakeholders for the data? Stakeholders could include the creators, users, managers, funders, or other relevant individuals/organizations.

DELIAH is the creator of this dataset and the primary stakeholder. DELIAH plans to share its findings through toolkits, academic resources, and policy recommendations on deliah.eu.

Additional stakeholders include academics and researchers interested in humour, civic engagement, and democracy, especially those studying these topics within the European Union, and particularly within the six DELIAH focus countries.

How many data files are there? What do they contain?

Two files: one data file (an Excel spreadsheet with 315 coded bibliographic records) and one documentation file.

What file format are these files?

An Excel file (.xlsx) for the data file and a PDF (.pdf) for the documentation file.

Is there any specific software required to open or analyze the data files?

No specialized software is required. The dataset is provided as an .xlsx spreadsheet, so it can be opened and reviewed in common spreadsheet programs.

Does the data set come with any usage restrictions? If so, what are they?

Yes. According to the documentation, the dataset is licensed under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which allows reuse and adaptation (including commercial use) as long as you provide appropriate attribution (Engelken-Jorge, Moreno, & Castañeda-Zumeta, 2025).

Metadata

What metadata does the data come with?

The dataset is accompanied by a PDF that documents the data sources used to identify studies for each consortium country and defines the controlled vocabulary terms used in the dataset's classification fields. (Engelken-Jorge, Moreno, & Castañeda-Zumeta, 2025, p. 6).

Where is the metadata? E.g. is it in the data files, included in separate files, or both?

The metadata is in a separate PDF file and not within the dataset file itself.

How comprehensive is the metadata provided?

The metadata is fairly comprehensive. It defines each controlled vocabulary term and includes brief narrative sections describing the methodology used to gather and curate the dataset, as well as documenting key data decisions.

Is the metadata structured according to any metadata standard?

The dataset documentation functions as a project-specific codebook and does not seem to be following any metadata standard.

How could the current data and/or metadata be enriched? In particular, what additional information could be provided to: improve users' ability to discover the data set in the repository environment?

Although the project provides a strong "About" section in the Zenodo record, I think the dataset would be more discoverable if the record included additional keyword tags to improve visibility through Zenodo's search and browsing features.

Currently, the Zenodo record is tagged only with the keyword "humor" under Zenodo's keyword/subject tags, and it includes a single tag—"democracy"—under the MeSH, GEMET, and EuroSciVoc keyword sections.

How could the current data and/or metadata be enriched? In particular, what additional information could be provided to: assist somebody unfamiliar with the data to make use of the dataset for new purposes?

The project provides a substantial amount of supporting information; the PDF combines the codebook with methodology and scope notes, but is titled *Database of humour practices assembled and catalogued*, which may cause users searching specifically for controlled vocabulary terms or methodology documents to overlook it. One helpful approach would be to split this PDF into two separate documents and give each a clear, descriptive title indicating its contents.

It would also be useful to include the controlled vocabulary terms within the dataset file itself (for example, as an additional tab within the spreadsheet) to make them easier to find and reference.

What publications have been written (if any) based on this dataset?

Reviewing the Zenodo deposit page, there are no citations listed. I also entered the title of the dataset into Web of Science and also found no results.

Are any publications listed/provided with the data set?

[DELIAH's own substack newsletter](#) archive is linked on the Zenodo deposit, as well as a link to [DELIAH's LinkedIn](#) and the [DELIAH website](#) for the research project.

What publications can you find using a literature database or general web search engine that cite or otherwise reference the use of this dataset? How did you search for these publications (if any)?

Interestingly, when I searched Google using the dataset title, the seventh result was a Zenodo deposit for a DELIAH paper that cites this dataset within the paper (though it is not explicitly cited or listed on the Zenodo deposit page). In that paper, the dataset is described as one part of the project's research output, alongside a narrative meta-analysis as the complementary second component.

I was surprised that these two deposits are not formally linked on Zenodo: neither Zenodo record explicitly references the other, even though both appear in the DELIAH Zenodo Community (and the meta-analysis is also additionally listed in the EU Open Research Repository community).

Reference list

Engelken-Jorge, M., Moreno, C., & Castañeda-Zumeta, A. (2025). *Humour practices and European attitudes toward democracy: Data documentation* (Version 1; M1/MS1.1—Database of humour practices assembled and catalogued). DELIAH.
<https://doi.org/10.5281/zenodo.15756409>