

Collection 2: U.S. News Media, c. 1989-2019

(articles mentioning "humanities" or "liberal arts")

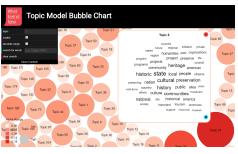
A collection of word-frequency and other data representing 94,816 unique articles mentioning "humanities" or "liberal arts" (no duplicate or close-variant documents) published mostly during 1989-2019 in 884 U.S. news sources and their associated blogs.¹ (5,492 articles originate from earlier years going back to 1977.) WE1S and other researchers use this data to look for broad patterns and to help guide closer study.

The top 20 news sources (ordered by number of articles in descending order) are: New York Times, Washington Post, Los Angeles Times, Chicago Tribune, Chronicle of Higher Education, News Day, Hartford Courant, Boston Globe, St Louis Post Dispatch, Oregonian, Birmingham News, Chicago Daily Herald, Dallas Morning News, Advocate (Baton Rouge), Star Ledger (Newark), Daily Oklahoman (Oklahoma City), Stanford Daily (Stanford University), Knoxville News Sentinel, Star Tribune (Minneapolis), Providence Journal. (Full sources and counts are available as a csv file in the Collection Dataset.)

In this collection, the word "humanities" occurs 134,948 times in 82,077 documents; and "liberal arts" 20,335 times in 14,621 documents.

Specification decision from the entire from the first better from the entire f





Topic model of this collection (250 topics visualized in three of the visualization interfaces in the WE1S <u>Topic Model Observatory</u>: <u>Dfr-browser</u>, <u>pyLDAvis</u>, and <u>TopicBubbles</u>)

WE1S Collection Registry ID:

20190621_0419_us-humanities-libarts-all-no-reddit

Data source: LexisNexis (via LN Web Services Kit)

Collection deposit (term frequencies, Mallet outputs, visualization files, and other non-consumptive data):

Zenodo (DOI: 10.5281/zenodo.TBD)

Topic models available (# of topics): 25, 50, 100,

150, 200, 250

Topic model visualizations: Start page

¹ WE1S makes available only "non-consumptive use" word frequency, topic model, and other datasets along with their visualizations. Datasets cannot be used to access, read, or reconstruct the original texts.