(WE1S core collection of articles mentioning “humanities”)

A collection of word-frequency and other data representing 82,324 unique articles mentioning “humanities” (no duplicate or close-variant documents) published mostly during 1989-2019 in 850 U.S. news sources and their associated blogs.¹ (About 5,000 articles originate from earlier in the 1980s.) The word “humanities” occurs 134,948 times in the collection. WE1S and other researchers use this data to look for broad patterns and to help guide closer study.


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.

Further Information

WE1S Collection Registry ID:
20190620_2238_us-humanities-all-no-reddit
Data source: LexisNexis (via LN Web Services Kit)
Collection dataset*: GitHub
Mallet topic model data files: GitHub

Topic models available (# of topics): 25, 50, 100, 150, 200, 250
Topic model visualizations: Start page

This card created 5 Jan. 2020; rev. 24 Feb. 2020 (Alan Liu)