

## Collection 1: U.S. News Media, c. 1989-2019

(WE1S core collection of articles mentioning “humanities”)

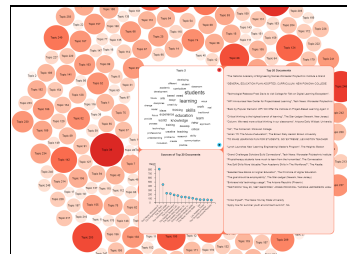
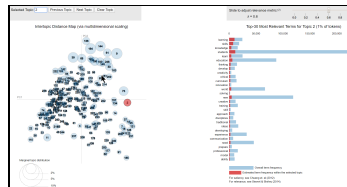
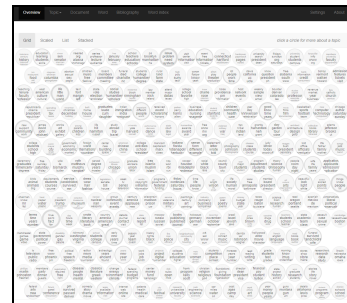
A collection of word-frequency and other data representing 82,324 unique articles mentioning “humanities” (no duplicate or close-variant documents) published mostly during 1989-2019 in 850 U.S. news sources and their associated blogs.<sup>1</sup> (About 5,000 articles originate from earlier in the 1980s.) 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.

The top 20 news sources in Collection 1 (ordered by number of articles in descending order) are: *New York Times*, *Washington Post*, *Los Angeles Times*, *Chronicle of Higher Education*, *Chicago Tribune*, *News Day*, *Boston Globe*, *Hartford Courant*, *St Louis Post Dispatch*, *Oregonian*, *Dallas Morning News*, *Stanford Daily* (Stanford University), *Advocate* (Baton Rouge), *Birmingham News*, *Star Ledger* (Newark), *Knoxville News Sentinel*, *Chicago Daily Herald*, *Providence Journal*, *Daily Oklahoman* (Oklahoma City), *Charlotte Observer*. (Full sources and counts are available as a csv file in the Collection Dataset.)

Sources in Collection 1 are tagged (through metadata) as associated with categories that include those listed on the [Metadata Tags for WE1S Document Sources page](#). (Sources are assigned to categories based solely on explicit publication information and/or self-identification.)

<sup>1</sup> 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.

(See also a [visualization](#) in an experimental “GeoD” interface created by WE1S researchers of a version of Collection 1 enhanced through “wikification” to create a topic model that includes geotagged and named entity tokens. See [GeoD](#) in the *WE1S Topic Model Observatory Guide*.)



Topic model of C-1 (250 topics visualized in three of the visualization interfaces in the WE1S [Topic Model Observatory](#): [Dfr-browser](#), [pyLDAvis](#), & [TopicBubbles](#))

### WE1S Collection Registry ID:

20190620\_2238\_us-humanities-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.4902187](https://doi.org/10.5281/zenodo.4902187))

**Topic models available (# of topics):** 25, 50, 100, 150, 200, 250

**Topic model visualizations:** [Start page](#)