

# The Media Frames Corpus: Annotations of Frames Across Issues: Supplementary Material

Dallas Card<sup>1</sup> Amber E. Boydston<sup>2</sup> Justin H. Gross<sup>3</sup> Philip Resnik<sup>4</sup> Noah A. Smith<sup>1</sup>

<sup>1</sup>School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA

<sup>2</sup>Department of Political Science, University of California, Davis, CA 95616, USA

<sup>3</sup>Department of Political Science, University of Massachusetts, Amherst, MA 01003, USA

<sup>4</sup>UMIACS, University of Maryland, College Park, MD 20742, USA

dcard@cmu.edu aboydstun@ucdavis.edu jhgross@polsci.umass.edu

resnik@umiacs.umd.edu nasmith@cs.cmu.edu

## A Additional details on preprocessing and annotation

The full set of news articles for this project was obtained using keyword searches on Lexis Nexis. The newspapers included in the search were: Atlanta Journal-Constitution, Daily News (New York), Denver Post, New York Times, Palm Beach Post, Philadelphia Inquirer, San Jose Mercury News, St. Louis Post-Dispatch, St. Paul Pioneer Press, Tampa Bay Times, The Herald-Sun (Durham), USA Today, and Washington Post.

De-duplication was performed by generating all the 4-grams from the text of each article and measuring the similarity between sets of 4-grams from each articles using the Jaccard coefficient. Articles published within two months of each other with a Jaccard coefficient  $> 0.2$  were considered to be the same, and each such group of articles was represented by one drawn at random.

Annotation was done using QDA Miner. The versions of articles shown to annotators had their source, author, and date removed, and were shortened to 225 words, rounded up to the nearest paragraph.

## B Reliability of pairwise arbitration of the primary frame

Beginning in stage 3, pairs of annotators who had been assigned the same articles in a given round met after completing their work independently to discuss cases where they disagreed on the primary frame and attempt to come to a consensus. To test the reliability of this process, we used three sets of 100 immigration articles each annotated by a separate pair of annotators in Stage 3. The cases where annotators disagreed on the primary

frame were then given to both the original annotators, as well as an additional pair of annotators, who saw the choices made by the original annotators, but were free to choose any primary frame. The inter-annotator agreement between pairs was then calculated for each batch of 100 articles, using the primary frame from the original annotators in cases where they initially agreed. The resulting average Krippendorff  $\alpha$  value was 0.78, and we thus concluded that this two-step process of annotating separately and then resolving disagreements was sufficiently reliable for determining an article's primary frame.

## C Patterns of Annotator Behavior

Ann.	Articles annotated	Av. frames per article	Av. spans per article	Av. chars. per span
1	5,751	2.8	6.6	146.0
2	3,576	2.8	6.1	115.4
3	3,324	3.7	10.0	72.5
4	2,957	2.6	5.4	113.2
5	2,950	2.4	5.6	214.5
6	2,850	3.4	6.4	191.7
7	2,451	2.8	5.4	161.9
8	2,399	3.5	8.5	134.1
9	2,351	2.9	6.0	200.5
10	1,550	3.2	6.4	160.3
11	1,300	3.6	8.8	169.2
12	1,200	2.6	5.6	126.3
13	1,200	3.3	7.3	167.2
14	1,200	2.7	6.3	180.6
15	900	3.4	6.9	168.1
16	650	2.0	4.1	191.0
17	500	2.9	5.4	241.6
18	300	2.3	4.8	192.3
19	125	2.5	3.4	73.7

Table 1: Summary statistics for the annotators who worked on this corpus, including average number of frames and spans selected per article, and average characters per span