

# DH 2019 technical notes

## Creating the subset of hearings used in this paper

The exact search was

```
'\Wlabor union|\Wlabor organization|\Worganized labor|\Wtrade union|industrial relation|labor relation|labor management|\Wworker\W|\Wworkers\W|\Wworking people\W|\Wworking women\W|\Wjobs\W|\Wjob\W|\Wwage\W|\Wwages\W|employment'
```

over a string consisting of the title of the hearing, the title of the committee or subcommittee, and the list of subjects the metadata contained for that hearing. Committee and subcommittee were included because sometimes the Congress creates special committees or commissions whose title is revealing. Note that the search does not capture merely "Labor" in the committee name, as in the House Committee on Labor or the Senate Education and Labor committee; those were only included if the title or one of the subjects included keywords related specifically to labor unions, industrial relations, workers, or jobs.

## Assigning witnesses into broader categories

The procedure for doing this was iterative and is not error-free. A number of regular-expression searches were performed on the full list of organizations to try to extract organizations belonging to a particular category. For example, a series of 'grep' commands for strings (or sometimes regular expressions) like "company," "railroad", "bank" was used to extract organizations from the full list of organizations; these were written into a file that served as the basis of a "company" classification. That preliminary categorization was then manually inspected and re-processed to remove false hits like e.g. "railroad brotherhoods" which should be in the labor list instead. The resulting final lists were then used for assigning a category (that is, if an organization is in the labor list, it gets categorized as labor, and so on.) The process is not perfect and errors remain; at present I have not yet attempted to classify all witnesses but have focused on a few categories like labor, companies, political parties, etc.