

ONLINE APPENDIX FOR:
RELATIO: TEXT SEMANTICS CAPTURE POLITICAL
AND ECONOMIC NARRATIVES

Elliott Ash¹, Germain Gauthier², Philine Widmer³

¹ETH Zürich, ashe@ethz.ch

²CREST, Ecole Polytechnique, germain.gauthier@polytechnique.edu

³University of St.Gallen, philine.widmer@unisg.ch

A Python Implementation with RELATIO

We detail the programmatic implementation of the procedure described in this paper, which is freely distributed as the Python package RELATIO. Our pipeline takes a collection of unstructured text documents as input and returns narrative statements (see Equation 2). Figure 2 presents a graphical outline of the different steps.

First, we split the raw text into sentences using spaCy (Honnibal et al., 2020). Second, we annotate the semantic roles (SRL). For this task, we rely on a state-of-the-art pre-trained model implemented by Gardner et al. (2017).⁴ SRL decomposes every sentence into one or multiple role sequences as defined in Equation 1. The current implementation covers agents (ARG0), verbs (V), patients (ARG1), attributes (ARG2), modals (ARGM-MOD) and negations (ARGM-NEG). An example of semantic role labeling annotations is provided in Figure 1.

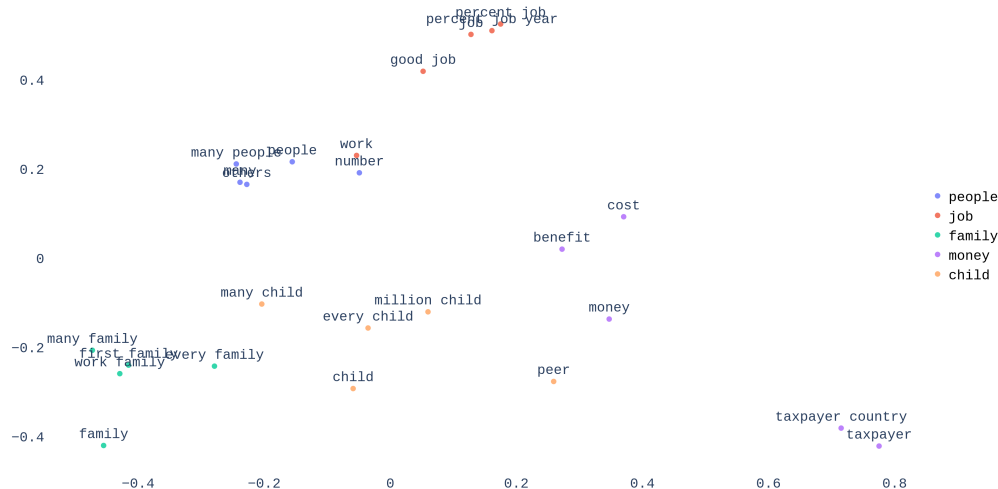
There are hundreds of thousands of unique semantic roles. We seek to regroup them into a smaller set of interpretable groups. An easy step in this direction is to “clean” all roles. We thus lowercase and lemmatize each word. We also drop punctuation, numbers and words from a list of pre-defined stopwords (for the complete list, see M). After these cleaning steps, we drop roles of more than four tokens as they are often uninterpretable.

We do not process the verbs (V) any further, but connect them to their negation (if any). For agents (ARG0), patients (ARG1) and attributes (ARG2), we rely on two methods: Named Entity Recognition (NER) and phrase embedding clustering. We first run spaCy’s Named Entity Recognizer (NER) to identify all named entities mentioned in the corpus. After careful inspection of this list, we keep the 1000 most frequent entities and assign to them their own cluster. Beyond this threshold, named entities are pronounced less than 40 times and are often idiosyncratic. If a role contains one of these named entities, then it is designated as that entity and not processed any further.

For all of the remaining roles that do not contain a frequent named entity, we produce text-embedding clusters as follows. First, we apply a phrase encoder to compress the plain text snippet to a low-dimensional dense vector. There are many encoders fit for this purpose (e.g. SentenceBERT (Reimers and Gurevych, 2019) and its variants). As a computationally efficient default, we use Arora, Liang and Ma’s “simple but tough-to-beat” sentence embeddings, which consist of the weighted average of the GloVe word vectors (Pennington, Socher and Manning, 2014) across each word in the role. The weights are chosen to be inversely proportional to the frequency of the word in the corpus. This means that relatively rare, more distinctive words are up-weighted in the phrase embedding. We use gensim’s pre-trained GloVe embeddings with 300 dimensions (Řehůřek and Sojka, 2010). The result of this procedure is a vector in the embedding space for each agent,

⁴See <https://demo.allennlp.org/semantic-role-labeling>

Figure A.1: Clustered Embeddings Recover Coherent Entities



Note: This figure illustrates the embedding-cluster process. We take the five most frequent entity clusters (produced using K -means applied to the phrase embeddings): people, job, family, money, and service. The embedding locations of the most frequent phrases in each cluster are illustrated here in a two-dimensional projection. The embedding space both co-locates and clusters similar phrases, providing dimension reduction while preserving semantically coherent entities. Embeddings are projected to two dimensions using PCA.

patient and attribute, where linguistically similar phrases (e.g. “higher taxes” or “high tax rates”) have geometrically proximate vector representations. A graphical depiction is presented in Figure A.1.

Exploiting this geometry, we further reduce the dimensionality of our role sets by unsupervised clustering applied to the associated matrix of encoded vectors. As with the phrase encoding step, there are a number of possible approaches for clustering (e.g., DBSCAN, agglomerative clustering). We obtain effective performance with `scikit-learn`’s implementation of the conventional K -means clustering algorithm, which is also computationally lightweight (Pedregosa et al., 2011). For K -means, the only hyperparameter is the number of clusters. Each remaining agent, patient and attribute in the corpus is assigned to a cluster. For interpretability, each cluster is labeled by the most frequent term within the cluster. Given the size of the corpus, we randomly select 50,000 speeches (roughly 5% of the entire corpus) to train the K -Means model, and predict out-of-sample for the remaining speeches. We cluster agents, patients and attributes together. In turn, the same entity can appear as an agent, a patient or an attribute. This joint clustering ensures that roles can be connected easily in a directed multigraph.

Finally, we focus on all narratives that have either an Agent-Verb-Patient or an Agent-Verb-Attribute structure. As the distribution of narrative frequencies is heavily skewed to the left, and since we are interested in recurring narratives, we restrict our attention to narratives pronounced at least 50 times in the corpus (i.e., at least twice a year on average).

B Human Validation

Tuning the number of clusters

The number of mined entities is the core hyperparameter in our analysis. It depends on the number of named entities recognized, as well as the number of clusters we specify for K-Means. To pick the correct number of named entities in our analysis, we manually inspect the most frequent named entities in the corpus. We fix a threshold at 1000 named entities, above which the named entities are pronounced fewer than 40 times and tend to become idiosyncratic. We keep this threshold fixed throughout the rest of the analysis.

Contrary to the number of named entities, the number of clusters to specify for K-Means is not trivial for at least two reasons. First, it is unclear how many unnamed entities are being referred to in the corpus. Second, there is a trade-off between the informativeness of clusters and the magnitude of the dimension reduction. If a small number of clusters is specified, then clusters are likely to become abstract and to encompass a broad variety of phrases. If a large number of clusters is specified, then clusters will be more coherent (and thus potentially more interpretable), but they can also become redundant, with multiple clusters referring to the same underlying entity. This trade-off suggests that we would like to maximize cluster coherence while minimizing cluster redundancy.

We investigate the optimal number of clusters with a simple human validation task. The task is performed by freelancers on Upwork who were based in the U.S. and whose primary language is English. First, we run four clustering scenarios and specify respectively 100, 500, 1000, and 2000 clusters. We then provide annotators with a random sample of semantic roles and their associated cluster labels (from different cluster numbers). Additionally, we add placebo observations, where we show again a random sample of semantic roles, but instead of the predicted cluster label, we combine every role with a random cluster label (drawn from the entire pool of available labels). Annotators are then asked to rank the similarity of the semantic role and the cluster label on a scale of 1 (not similar at all) to 10 (extremely similar/the same). Our random validation sample consists of 4000 unique role and label combinations (500 true and 500 placebo combinations for every clustering scenario). Every unique combination is, on average, annotated by 2.5 freelancers, resulting a total sample of 10000 annotated combinations. Annotators do not see the number of clusters that were used to generate the cluster label, nor whether the semantic role is combined with a predicted cluster label or a placebo label. This procedure jointly evaluates cluster coherence and redundancy (coherence is reflected in how similar the semantic roles are to their predicted label; redundancy is captured by similarities between random combinations of roles and clusters).

Annotator recruitment

In June 2021, we hired three different freelancers on Upwork to annotate 4000 combinations (applies to two freelancers) or 2000 combinations (applies to one freelancer) of roles and labels. The instructions as communicated in the job post are shown in Figure B.1.

Figure B.1: Wording of the validation job post on Upwork

“We need freelancers (U.S.-based, fluent in English) to check the similarity between two phrases. In the attached file, please compare the phrase in column 1 and 2. The question is: for every row, how similar are the entities described in column 1 and 2? Please indicate from 0 (not similar at all) to 10 (very similar/the same). Note that you can select values in between 0 and 10 (e.g., 5). This task is about evaluating a computer-based language model from a human perspective. It is about your “hunch” or intuition of how similar the phrases are. Do they refer to a similar entity?”

Upon posting the job, we sent a sample of 20 combinations to the first five freelancers who expressed their interest in the job. Out of these five, we hired the first three freelancers who submitted the annotated sample (all of these first three had annotated the sample correctly). In Figure B.2, we feature the form presented to the freelancers for annotation.

Figure B.2: Screenshot of Human Annotation File

ARG	ARG-RAW	similarity
water	soil water degradation	
tsa	voter supporter	
care	support treatment	
democrat	sherry life	
bad	product substance regulate	
constitution	constitution high law	
opportunity	many commercial advantage	
labor	nation overall growth	
objection	madison marketing	

Human validation Results

Table B.1 summarizes the key figures of the human validation across numbers of clusters. The average similarity of clustered roles is 4.49; for the placebos, this figure amounts to 2.57.

Figure B.3 decomposes the results by number of clusters: it shows the similarity between a random semantic role token and its cluster label for actual (role, cluster) tuples (yellow) and placebo tuples (blue; where roles are combined with a random cluster label), as annotated by human freelancers. Subfigure (1) shows the results for all arguments (that appear in narratives

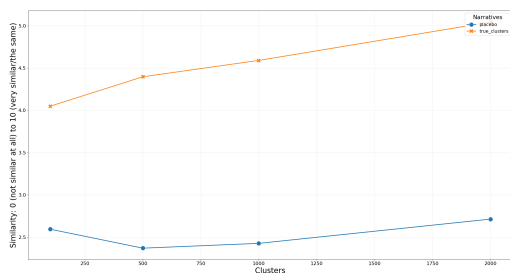
Table B.1: Summary Statistics on the Human Annotations

Total combinations annotated	10000
Unique combinations annotated	4000
Number of annotators	3
Average number of annotators per combination	2.5
Average number of annotations per annotator	3334
Average similarity (true combinations)	4.49
Average similarity (placebo combinations)	2.57
Average inter-annotator similarity standard deviation	2.42

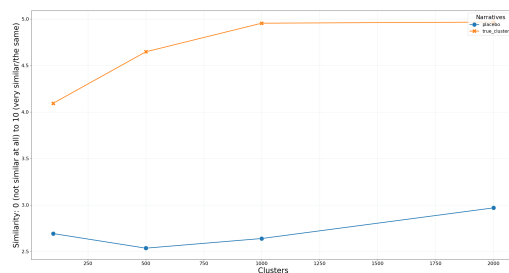
above our usual frequency threshold of 50, see Section 3.1). Here, the similarity increases for the clustering specifications studied here (100, 500, 1000, 2000). However, after 1000 clusters, the placebo similarity also increases.

Next, we seek to understand whether more clusters lead to precision gains for frequent or infrequent entities. Thus, in Subfigure (2), we show the subset of validated arguments that belong to the top 25% of entities (for the respective cluster numbers). Here, increasing the number of clusters seems to yield lower similarity gains after 1000 clusters. These results suggest that higher cluster numbers may be beneficial for less frequent entities. Again, however, the placebo similarity increases more after 1000 clusters. In our application, we focus on 1000 clusters.

Figure B.3: Similarity of semantic roles and cluster labels as reported by humans



All arguments
(1)



Top 25% arguments (by frequency)
(2)

Note: This figure shows the similarity between a random semantic role token and its cluster label for actual (role, cluster) tuples (yellow) and placebo tuples (blue; where roles are combined with a random cluster label), as annotated by human freelancers. Subfigure (1) shows the results for all arguments (that appear in narratives above our usual frequency threshold of 50). Here, the similarity increases for the clustering range studied here. However, after 1000 cluster, the placebo similarity also increases. Subfigure (2) shows the subset of validated arguments that belong to the top 25% of entities (for the respective cluster numbers). Here, increasing the number of clusters seems to yield lower similarity gains after 1000 clusters.

C Inspecting the Quality of Mined Entities

Table C.1: Most Frequent Clustered Entities (i)

Label	Most frequent phrases	Frequency
people	people - 182086 — many - 28671 — others - 17168 — many people - 7433 — number - 6196 — have - 5907 — several - 2299 — many others - 1413 — include - 883 — many country - 684	195976
money	money - 30048 — taxpayer - 15388 — benefit - 9898 — cost - 9534 — pay - 2496 — expense - 1298 — amount money - 831 — cash - 744 — taxpayer money - 657 — earn - 529	70503
family	family - 37326 — life - 25338 — work family - 2310 — great - 1981 — many family - 1444 — real problem - 585 — family business - 537 — well life - 524 — great country - 438 — real people - 430	67615
someone	someone - 17689 — anyone - 13401 — everyone - 12217 — everybody - 10819 — somebody - 9924 — anybody - 8159 — nobody - 7671 — whoever - 1030 — someone else - 833 — else - 739	66452
person	person - 30540 — woman - 22444 — man - 10881 — every person - 755 — many woman - 582 — woman child - 545 — first woman - 454 — another person - 404 — man woman - 336 — million woman - 266	58586
job	job - 40808 — work - 16726 — good job - 1357 — job do - 648 — create job - 584 — great job - 548 — work people - 510 — well job - 405 — good pay job - 356 — do job - 353	56871
country	country - 53529 — people country - 2428 — around - 725 — along - 626 — people across country - 388 — across country - 354 — across - 350 — people around world - 311 — around world - 255 — country around world - 229	54044
child	child - 40835 — million child - 1278 — every child - 1224 — peer - 779 — many child - 714 — child family - 593 — child care - 508 — child support - 440 — care need - 439 — family child - 289	43809
business	business - 20863 — industry - 8018 — many company - 719 — management - 719 — enterprise - 691 — foreign company - 628 — financial - 433 — business community - 397 — many business - 386 — management plan - 297	40514
employee	employer - 20943 — employee - 16335 — contractor - 3404 — small employer - 475 — employer employee - 416 — many employer - 393 — private contractor - 334 — percent employer - 232 — many employee - 182 — large employer - 168	39434
law	law - 41489 — public law - 1375 — tax law - 488 — legislative branch - 384 — legal system - 361 — civil action - 327 — legal status - 307 — judicial review - 246 — judicial system - 223 — judicial branch - 172	38472

Table C.2: Most Frequent Clustered Entities (ii)

Label	Most frequent phrases	Frequency
plan	plan - 19094 — proposal - 13111 — initiative - 3198 — propose - 2008 — agenda - 1479 — president plan - 573 — president proposal - 422 — legislative proposal - 312 — be propose - 247 — blueprint - 213	37723
individual	individual - 33657 — account - 3647 — individual family - 643 — many individual - 463 — personal information - 283 — personal responsibility - 235 — individual business - 207 — every individual - 207 — million individual - 179 — family individual - 172	36330
community	community - 13902 — organization - 8763 — local community - 2305 — international community - 2100 — youth - 1386 — membership - 564 — behalf - 487 — many community - 487 — nonprofit organization - 462 — participation - 405	34277
worker	worker - 16944 — million worker - 554 — many worker - 399 — foreign worker - 277 — social worker - 213 — worker right - 133 — every worker - 127 — worker country - 126 — worker family - 126 — percent worker - 123	33601
tax	tax - 15994 — tax cut - 5609 — spending - 4017 — cut - 3774 — tax increase - 1480 — raise - 648 — raise tax - 399 — spending cut - 318 — billion tax cut - 229 — percent cut - 221	32317
company	company - 34769 — consortium - 288 — subsidiary - 222 — affiliate - 196 — ag - 178 — venture - 147 — ge - 129 — foreign subsidiary - 63 — startup - 63 — conglomerate - 62	32205
school	school - 16925 — college - 2924 — local school - 1676 — high school - 1305 — public school - 1178 — college student - 519 — academy - 460 — school system - 417 — many school - 403 — community college - 335	27607
student	student - 18503 — teacher - 6398 — many student - 542 — student family - 407 — million student - 393 — every student - 268 — percent student - 190 — student teacher - 181 — parent teacher - 142 — qualify teacher - 131	26535
fund	fund - 17564 — investment - 3647 — investor - 2817 — foreign investor - 293 — invest - 228 — fund provide - 208 — hedge fund - 192 — foreign investment - 182 — private investment - 175 — mutual fund - 158	26216

Note: This table presents the 20 most frequent clustered entities identified in the U.S. Congress. The column “Label” refers to the final label assigned to the entity after manual inspection. The column “Most frequent phrases” lists the ten most frequent phrases which were assigned to these clustered entities, as well as their frequency of occurrence (i.e., phrase - frequency). The column “Frequency” is the total number of mentions of this clustered entity in the corpus.

Table C.3: Most Frequent Named Entities (i)

Label	Most frequent phrases	Frequency
american	american - 49293 — american people - 35168 — million american - 8209 — american public - 3212 — american taxpayer - 3187 — many american - 3118 — american family - 3074 — every american - 3031 — american worker - 2682 — american citizen - 1956	149225
program	program - 44125 — pilot program - 1125 — federal program - 958 — important program - 617 — government program - 580 — medicaid program - 448 — cop program - 403 — education program - 314 — many program - 293 — entitlement program - 280	104383
government	government - 62923 — local government - 5866 — foreign government - 1085 — government official - 553 — government shutdown - 437 — russian government - 383 — tribal government - 357 — government spending - 344 — big government - 303 — branch government - 285	90865
service	service - 14224 — arm service - 3166 — forest service - 2589 — postal service - 2282 — internal revenue service - 896 — wildlife service - 862 — good service - 812 — custom service - 690 — public service - 637 — military service - 607	61409
nation	nation - 34474 — great nation - 918 — nation world - 445 — people nation - 351 — many nation - 280 — entire nation - 277 — nation economy - 272 — rogue nation - 239 — nation child - 231 — foreign nation - 226	57571
republican	republican - 51157 — republican leadership - 6332 — republican majority - 3678 — republican friend - 1246 — republican plan - 1169 — republican side - 614 — republican president - 486 — many republican - 478 — republican proposal - 471 — republican control - 410	56582
administration	administration - 56284 — president administration - 606 — administration official - 484 — previous administration - 476 — administration policy - 268 — administration proposal - 217 — administration budget - 200 — republican administration - 149 — administration president - 125 — last administration - 121	53429
agency	agency - 23605 — federal agency - 6048 — local educational agency - 1605 — government agency - 942 — educational agency - 859 — law enforcement agency - 723 — head agency - 592 — intelligence agency - 530 — public housing agency - 514 — department agency - 389	47452
budget	budget - 28857 — republican budget - 2091 — president budget - 1780 — budget deficit - 663 — democratic budget - 446 — budget plan - 427 — budget agreement - 392 — budget proposal - 345 — budget process - 319 — budget surplus - 298	46434
america	america - 24636 — people america - 971 — rural america - 481 — corporate america - 306 — america senior - 258 — job america - 246 — america family - 209 — america child - 202 — bless america - 197 — everybody america - 191	46053

Table C.4: Most Frequent Named Entities (ii)

Label	Most frequent phrases	Frequency
national	national security - 2623 — national - 1426 — national guard - 1159 — national debt - 1153 — national academy science - 672 — national interest - 563 — national labor relation - 471 — national intelligence director - 422 — director national intelligence - 407 — national science foundation - 358	39427
energy	energy - 5592 — department energy - 2259 — energy policy - 1119 — energy commerce - 652 — high energy - 327 — energy company - 312 — renewable energy - 305 — energy efficiency - 294 — energy independence - 258 — energy crisis - 246	28045
democrat	democrat - 25073 — democrat republican - 2377 — republican democrat - 2183 — democrat majority - 459 — democrat leadership - 418 — president democrat - 406 — republican democrat alike - 297 — many democrat - 280 — democrat republican alike - 260 — liberal democrat - 220	28012
group	group - 8480 — bipartisan group - 464 — special interest group - 458 — work group - 413 — terrorist group - 389 — group people - 256 — environmental group - 252 — interest group - 250 — group health plan - 246 — outside group - 231	25784
project	project - 9652 — demonstration project - 610 — pilot project - 310 — construction project - 241 — program project - 104 — important project - 100 — research project - 100 — military construction project - 98 — project authorize - 93 — water project - 92	23013
war	war - 6299 — war iraq - 831 — world war ii - 819 — war terrorism - 482 — civil war - 476 — cold war - 408 — war terror - 338 — war power - 217 — war drug - 201 — world war - 196	22704
small business	small business - 14213 — small business owner - 1592 — small business administration - 538 — small business concern - 517 — many small business - 439 — small business people - 183 — percent small business - 156 — small business community - 131 — individual small business - 128 — family small business - 124	22379
medicare	medicare - 10868 — medicare beneficiary - 1220 — medicare patient - 808 — medicare system - 309 — medicare trust fund - 215 — medicare benefit - 199 — medicare cut - 150 — medicare recipient - 149 — cut medicare - 138 — traditional medicare - 127	22348
veteran	veteran - 11324 — veteran affair - 1688 — many veteran - 412 — disabled veteran - 377 — nation veteran - 328 — veteran family - 189 — homeless veteran - 188 — million veteran - 181 — veteran benefit - 156 — america veteran - 156	20792
authority	authority - 7485 — local authority - 527 — palestinian authority - 508 — housing authority - 310 — administer authority - 299 — federal authority - 256 — public housing authority - 236 — budget authority - 212 — trade promotion authority - 170 — exist authority - 158	19468

Note: This table presents the 20 most frequent named entities identified in the U.S. Congress. The column “Label” refers to the final label assigned to the entity after manual inspection. The column “Most frequent phrases” lists the ten most frequent phrases which were assigned to these named entities, as well as their frequency of occurrence (i.e., phrase - frequency). The column “Frequency” is the total number of mentions of this named entity in the corpus.

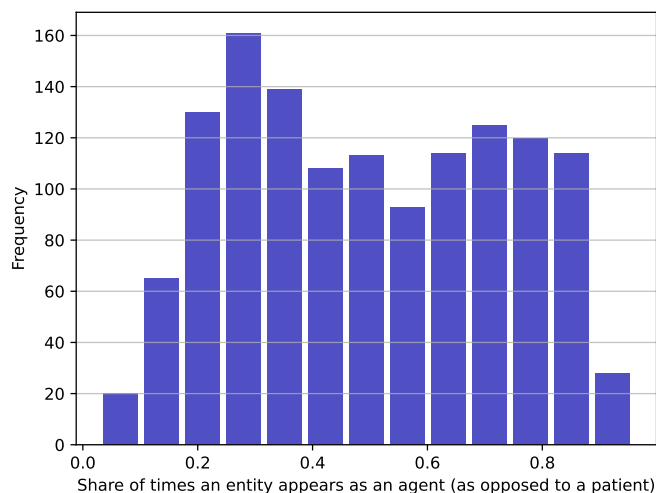
D Summary Statistics

Table D.1: Summary Statistics on the Congressional Record Corpus

Variable	All Narratives	Complete Narratives	Frequent Narratives	Relevant Narratives
Speeches	1190969	550561	187379	105514
Sentences	17312358	4196393	279044	160827
Statements	76395962	5774295	346685	180964
Narratives, unique	5963192	3570119	2838	1638
Agents raw	76395962	5774288	346685	180964
Agents raw, unique	1604754	945412	21041	15035
Agents clustered, unique	1884	1879	415	245
Patients raw	76395962	5774293	346685	180964
Patients raw, unique	8322792	1794942	33774	21411
Patients clustered, unique	1884	1877	499	353
Verbs raw	76395962	5774295	346685	180964
Verbs raw, unique	44889	18642	391	276
Verbs cleaned, unique	34780	11807	183	140

Note: This table shows descriptive statistics on the U.S. Congressional Record at different steps of the pipeline. “All narratives” are those with at least an agent or a patient. “Complete narratives” are those with both an agent and a patient. Among those, “frequent narratives” were pronounced at least 50 times. Finally, we focus on politically or economically relevant entities (“relevant narratives”) – as opposed to procedural/uninterpretable entities.

Figure D.1: Share of Times An Entity appears as an Agent (as opposed to a Patient)



Note: This figure presents, for every entity, the share of times it appears as an agent (as opposed to a patient). For example, if an entity occurred as an agent 8 times and as a patient 2 times, the corresponding share would be 0.8. Most entities appear both as an agent and a patient. However, entities also tend to be “specialized”: Most of them appear as either an agent or a patient more frequently.

E Examples of Narratives and Sentences

This appendix shows examples of random sentences and their associated narratives as well as the most frequent narratives and random examples of sentences they appear in. In both cases, we focus on narratives that contain (at least) an agent, a verb, and a patient (i.e., “complete” narratives).

1. Sampled sentences and their associated narratives

The following list shows a random sample of sentences, along with the complete narrative(s) we mine for each sentence.

As the ranking member of the Transportation and Infrastructure Subcommittee during the 111th Congress, he helped us ensure ongoing investment in Federal highway, transit and highway safety programs and the jobs that those investments create.

fund create job

.....
Many of our Senators during campaigns promise to vote for a constitutional amendment to balance the budget, in the heat of the campaign.

constitutional (thing) balance budget

.....
Companies doing business online want to reassure consumers and potential customers that their interests will be protected online, not heighten their concern about electronic commerce.

company do business

.....
The Progressive Caucus Budget raises taxes by almost \$950 billion.

budget raise tax

.....
Many are asking what happened to the promise Republicans made in 2010 when, in their pledge to America, they wrote—and again I quote—in their pledge to America:

republican make promise

.....
Faced with soaring gasoline prices, agencies around the country that provide services to the elderly say they are having to cut back on programs such as Meals on Wheels, transportation assistance, and home care, especially in rural areas that depend on volunteers to provide their own gas.

agency provide service

.....

And we are doing so at a time when our budget, the budget put through by the Republicans who control the Congress, is slashing funds for education, cancer research, health care, other essential needs at home.

republican put budget

.....
Subject to the direction of the Secretary, a person charged with law enforcement responsibilities by the Secretary who is performing a duty related to enforcement of a law regarding fisheries or other marine resources may make an arrest without a warrant for an offense against the United States committed in his presence, or for a felony cognizable under the laws of the United States, if he has reasonable grounds to believe that the person to be arrested has committed or is committing a felony.

person commit (legal) charge

.....
We desperately need the largest volunteer effort in our history to mobilize our citizens to act as role models and mentors for young people who have no real family encouragement.

people have appreciation

.....
What I don't understand is why our Democratic friends want to ruin a good thing that 80 percent of seniors who have access to this prescription drug plan say they like and 90 percent of seniors eligible have signed up for, saving on average \$1,200 a year.

senior have prescription drug

.....
President, I ask unanimous consent that the Committee on the Judiciary be authorized to meet to conduct a markup on Tuesday, October 5, 2004, at 10 a.m., in SD-226.

judiciary conduct markup

.....
Recognizing the public health threat caused by antibiotic resistance, Congress took several steps to curb antibiotic overuse in human medicine through amendments to the Public Health Service Act made by section 102 of the Public Health Threats and Emergencies Act , but has not yet addressed antibiotic overuse in agriculture.

healthcare make healthcare

.....
As the crisis approached, Democrats did nothing.

democrat do nothing

.....
I join with my friend in asking everyone to support this measure to name this post office after

Sergeant Shawn T. Hannon and Master Sergeant Jeffrey J. Rieck, and I align myself with all of the comments that my friend and colleague has made on behalf of these two gentlemen, these two heroes.

friend make comment

.....
History has shown that, without strict limits on spending, tax increases will simply go to new spending rather than to reduce the deficit.

tax reduce deficit

.....
Provided further, That funds available under this heading for opportunity scholarships, including from prior-year appropriations acts, may be made available for scholarships to students who received scholarships in the 2009-2010 school year:

student receive degree (education)

.....
Well, it is important because people who have made America, who have made America strong, ought to in fact be remembered.

people make america

.....
President, earlier this week I placed on my Web site– sanders.senate.gov–a letter to Secretary Paulson, and I asked people who shared the sentiments of that letter to sign a petition.

people share concern

people sign petition

.....
It seeks to solve the problem of soft money , not just with the political parties, but with the special interest groups who run attack ads, who are completely unregulated by the system, and whose contributors are undisclosed.

group run ad

.....
I raise these points because when someone stands up and says, “We are the ones who want the taxpayers to keep their money and you on this side of the aisle, you are the ones who want to take it from them.”

money keep money

.....

2. Most frequent narratives and examples of sentences they appear in

The following list shows the most frequent narratives in our corpus (ordered by frequency), along with a random sample of sentences in which they appear.

people lose job

they knew people would lose their jobs.

People will lose their jobs, people will lose their savings, people will find that they cannot get the credit necessary to keep their businesses open or to function at a reasonable level.

But last week the Associated Press reported that an even greater number of people, 108,000, lost their jobs, as U.S. companies dealt with the battered economy right here at home.

This bill allows for clear actions, up-or-down votes on the conference report to prevent Wall Street from melting down like it did 2 years ago and a bill to provide unemployment compensation to people who have lost their jobs who cannot find work in this economy.

On the one hand we have people who lose their jobs after working years and years and are temporarily in need of assistance and should have that assistance.

.....

citizen abide law

By forcing the IRS, and not the taxpayer, to carry the burden of proof in disputes, we protect legal, law-abiding citizens and end harassing and frivolous claims by maverick agents.

A bill to increase and enhance law enforcement resources committed to investigation and prosecution of violent gangs, to deter and punish violent gang crime, to protect law-abiding citizens and communities from violent criminals, to revise and enhance criminal penalties for violent crimes, to expand and improve gang prevention programs, and for other purposes; to the Committee on the Judiciary.

Senators from every State of the Union, from every political persuasion, agreed to a version of the PATRIOT Act that would reauthorize the provisions that were set to expire and which provided the Government with the tools to aggressively pursue the war on terror, while protecting the rights of law-abiding citizens.

This is a blatant attack on our second amendment rights and on our law-abiding citizens.

Furthermore, it is my belief that this basic right should include law-abiding citizens.

.....

american lose job

And now this bill that was cooked up in secret between the White House and Congressional Republicans without any input from Congressional Democrats will bring our country further into debt, lead to more hard- working Americans losing their jobs, and put a greater share of the tax receipts in the pockets of the Nation's most privileged.

But they are not cheaper because they pay 33 cents an hour; it is just that Americans lost their jobs.

How can we stand by as millions of American workers lose their jobs, as thousands of American companies cannot compete fairly, as our country as a whole has wealth drained from it?

For all the time that we do not pass it, each month 500,000 Americans will lose their jobs.

All these benefits, permanent tax benefits for middle-class families making one-quarter of a million dollars or less, the opportunity to create jobs, the opportunity to take care of a couple million Americans who lost their jobs, the opportunity to bring the private sector back again, the opportunity to give the private sector certainty, none of that is good enough for them.

.....

government run healthcare

Secretary Mike Leavitt, who is the former Secretary of the Department of Health and Human Services, said: Advocates for a public health-care plan continue to look for a way to give political cover to moderates while advancing their goal of implementing a government-run health-care system.

This is a debate about getting us on that path toward government-run health care.

We know that some here in Washington have wanted government-run health care for many years.

Now they have a government-run health care full of tax hikes and mandates on small businesses, which the NFIB estimates will cost 1.6 million more jobs lost.

We didn't see it materializing into a government-run health care program, as many have alleged here today.

.....

american have healthcare

As a nation we have 50 million Americans today who have no health insurance and that number has increased.

Speaker, as millions go without health insurance and costs skyrocket, I strongly believe we need commonsense reforms so that all American families can have access to affordable health coverage. The bill would create affordability credits to ensure that all Americans have more affordable health care coverage.

What we need to do is subject any health proposal, Democrat or Republican, to one simple test: Does it guarantee that all Americans will have health care coverage?

It is time for us to be a part of that representative democracy and forge a bill together that assures all Americans have access to health care reform.

.....

god bless america

The President said that we ought to pray that God bless America.

When in Namibia, where we were, I suggest you play the Namibian national anthem and leave it to the local mayor to say, as he said to us “God bless America.”

Mr. Speaker, God bless all listening, and God bless the United States of America.

Mr. Speaker, with that, before I yield back the balance of my time, I will ask God to please bless our men and women in uniform, both in Afghanistan and in Iraq and throughout the world, and I will ask God to please bless America.

I pray that God will always continue to bless America.

.....

people need help

According to the executive director of the Community Action Agency in Kentucky, Kip Bowmar: February of 2008 marked the first time in the program’s history that all 120 Counties in Kentucky ran out of LIHEAP funds, forcing us to close our doors as fuel prices were soaring and people needed help.

Look, there are lots of different things, and I will not go through the litany in terms of unemployment provisions and in terms of helping small businesses, because the thing I would like to do is just say thank the gentleman to the gentleman on behalf of New York, on behalf of the liberty zone, on behalf of all those people who need your help.

Why then should we be concerned about the increase in numbers of people who are condemned to a career as a temporary worker without benefits or minimum wage workers, people who work every day and still need public help?

Those young people need more than help from their families in order to be able to go to college.

There are people who need help who don’t have help, and we need to find a way to do that, but it is not this 2,000-page bill.

.....

god bless troop

In conclusion, God bless our troops, and we will never forget September the 11th and the global war on terrorism.

In conclusion, God bless our troops, and we will never forget September 11

In conclusion, God bless the memory of Congressman Sonny Montgomery, God bless our troops, and we will never forget September 11.

In conclusion, God bless our troops, and we will never forget September the 11th in the global war of terrorism.

In conclusion, God bless our troops—and we will never forget September the 11th and the global war on terrorism.

.....

worker lose job

Workers are losing their jobs, their homes, and their hopes.

This would create a system that is more focused on punishing employers than truly helping workers who lose their jobs.

There are truckers who are losing their jobs.

the impact of not doing something here is huge, on workers losing their jobs.

Speaker, Sara Lee is closing their Virginia apparel factory; 42 workers lose their jobs.

.....

small business create job

The No. 1 issue facing our country is jobs, and the No. 1 goal of Republicans in 2012 is to continue to make it easier for American small business to create jobs.

Two out of every three new jobs in the United States is created by a small business, but the shutdown is hitting them hard.

Small business owners will be able to spend less time searching for affordable health plans, and more time creating jobs.

President, our Nation's small businesses have created 64 percent of all new jobs in the last 15 years, yet in the last year nearly 85 percent of the jobs lost have come from small businesses.

That is why the President said let us focus on doing things that will help these families and equally, if not more importantly, help small businesses create jobs.

.....

service connect disability

A bill to increase, effective as of December 1, 1996, the rates of disability compensation for veterans with service-connected disabilities and the rates of dependency and indemnity compensation for survivors of such veterans, and for other purposes; to the Committee on Veterans Affairs.

I thank the Senator from Maine, Ms. Snowe, the chairman of the Small Business Committee, for making the important point that Federal law requires that a certain percentage of contracts be awarded to firms owned by veterans with service-connected disabilities.

As we reported in VA Health Care: Issues Affecting Eligibility Reform , VA uses a complex priority system—based on such factors as the presence and extent of any service-connected disability, the incomes of veterans with nonservice-connected disabilities, and the type and purpose of care needed—to determine which eligible veterans receive care within available resources.

Under the law that existed and controlled at that time, only those who had been granted a claim for a service-connected disability or demonstrated financial need could turn to the VA for health care services.

Most military retirees who have a service-connected disability are not permitted to collect both their retirement and disability benefits concurrently.

.....

constitutional (thing) balance budget

I have always wanted to be on the side of those supporting a constitutional amendment to balance the budget.

Let me say that I am someone who will support a constitutional amendment to balance the budget.

Let me, to the extent that I can, say to the seniors who are listening to those who would like to make you believe that they are really here arguing to save Social Security, suggest to you that what they are really arguing about is that they don't want a constitutional amendment to balance the budget and they have now hit on what I perceive to be a risky gimmick in an effort to frighten seniors and by that approach defeat a constitutional balanced budget amendment.

They say if we pass this constitutional amendment to balance the budget, the budget will be balanced.

One of the reasons the polls show so many people want a constitutional amendment to balance the budget is two things.

.....

people have healthcare

But more importantly than that, we have got people who have health care emergencies because of the earthquake, business emergencies, lives literally at stake.

The same on health insurance: If you are trying to increase the number of people who have health insurance, you have to do more than what is in this tax package.

Some writers have called that group the "contentedly covered," the people who already have health care coverage in America today.

We generally think of most people having private health insurance coverage.

People are going to have health insurance they do not now have.

.....

public held debt

446 511 available to repay debt held by the public.

It has brought with it, as Chairman Greenspan noted, a significant increase in revenues to the Federal Government, which are allowing us to pay down even more rapidly than previously thought the debt now held by the public.

Change in debt held by the public.

The best way to gauge the level of debt held by the public is to compute that debt as a percentage of our national income, our Gross Domestic Product.

202 Resolved, That upon the adoption of this resolution it shall be in order to consider in the House the bill to require that the Government prioritize all obligations on the debt held by the public in the event that the debt limit is reached.

.....

people pay tax

Real jobs in the U.S. economy for people who will be paying taxes, not being paid unemployment compensation for not working.

It is the people who are going to pay these taxes.

That is why the increase in seniors is going up so dramatically, and the increase in the people working and paying their taxes is going up modestly.

We are talking about giving the money back to the people who earn it, and the first place we ought to look is to people who are married who pay more taxes just because they are married.

This entire debate is about whether or not people who pay no taxes can get a tax cut.

.....

people lose family

But I think it is also so important for us to look back, figure out what went wrong, and move forward and fix the problems; not to create a political blame game, not to blame the Clinton administration or the Bush administration, but to fix what went wrong yesterday and make it right for tomorrow, so that we do not have 2,800 people lose their lives again.

Now is the time to speak up, not after the people have lost their lives, not after the American people—the American people know what is going to happen here.

We have had a number of people—maybe a dozen or so—who have lost their lives prior to this month in Iraq and Afghanistan.

Let’s recognize, whether they were in the World Trade Center, at the Pentagon, or on airplanes that were commandeered, people were in dangerous situations, and many lost their lives.

Yet the U.S. Government has chosen to discriminate against these people who lost their lives at Fort Hood 4 years ago at the hands of a terrorist, who tragically happened to be a member of the uniformed military of the United States, MAJ Nidal Hasan, who has subsequently been convicted of these crimes.

.....

american pay tax

Unless there is some extraneous circumstance, I think most Americans voluntarily pay their taxes. I think that Members of Congress, members of the administration need to take their heads out of the sand, need to start dealing with really some of the very tough issues of Medicare, of Social Security, of annual government overspending, and I would just ask an American that pays taxes to spend a few moments thinking about the absurdity of our tax code in this country.

However, it is not small to Dave and Joann and to the millions of young Americans who pay more in taxes simply because they have formed the basic unit of society—a family.

Why should middle-class Americans, who work hard, pay their taxes, and send their kids to school, not have that kind of security, too?

This bill takes away that freedom, requiring every American to purchase a government-approved health plan, pay a tax, or even go to jail.

.....
employee provide healthcare

The vast majority of employers who provide health care merely pay for the benefit. We urge the Senate to support your amendment to ensure that employers will not be sued for voluntarily providing health coverage to 172 million workers. As chairman of the Subcommittee on Employer-Employee Relations with jurisdiction over employer-provided health care, the issue of mental health parity falls within the purview of this subcommittee. According to a Gallup poll conducted last September, the vast majority of Americans, 70 percent, are satisfied with their health insurance provided by their employer. I don't think it's right that you mandate an employer to provide health care insurance or

.....
people do job

We have people doing this work for free in California, not costing any Federal taxpayer dollars, and the amendment that I offered with the gentleman from California that was passed by, I believe, 380 to 0, allows our people to have continued access. If you are giving the money to the people doing the work, they will work wonderfully, and things will happen beautifully. And now you've got people who once had three people doing a job, now there's one. My amendment to the Department of Defense authorization bill is meant to strengthen our efforts to verify if people in the United States are here legally to do their work. The people in Mexico who have the jobs that the people in Ohio used to do do not make enough money to buy the very products they make.

.....
men women serve country

So I have had the privilege of coming from a State that has sent to Washington, DC, men of great honor, great integrity, men and women who have served the country so well. We are talking about young men and women in our Armed Forces returning home after serving their country in Iraq. Throughout my years of service to Minnesotans in the Senate, Frank provided wise counsel—especially on behalf of the men and women who, like himself, had served their country in time of war. ' To the smart young man who wrote me that letter and to all of America's veterans, this bill builds on efforts to meet our country's moral obligations to the men and women who so bravely served our country.

I encourage my colleagues to honor the men and women who serve this country and pass this important legislation.

.....

F Comparison to Alternative Text Features

To identify latent topics in political discourse, we run Latent Dirichlet Allocation (LDA) on the U.S. Congressional Record. We then compare our method to topics and n-grams by presenting the ten most representative words as well as the ten most representative narratives associated to each topic share in Tables F.1, F.2, and F.3. We restrict ourselves to narratives that were at least pronounced 50 times in the corpus.

Table F.1: Topics, Words and Narratives (i)

Topic	Top Words	Top Narratives
banking	bank, financi, compani, credit, consum, drug, card, institut, account, market	bank issue card, company hold business, money fund outlay, measure provide authority, company hold loan, asset back military, american want reform, mortgage back military, action have effect, consumer make instruct
communications	page, line, internet, servic, amount, communic, comput, increas, access, broadcast	program serve child, provider offer service, tax benefit american, reconciliation carry recommendation, budget receive recommendation, facility provide service, company provide service, network provide service, provider provide service, citizen have healthcare
crime	crime, law, enforc, victim, crimin, violenc, justic, state, prison, act	law enforcement add category, crime add category, injustice motivate crime, juvenile commit crime, violent crime serve (legal) sentence, program meet regulation , juvenile commit violent crime, offender serve (legal) sentence, child witness violence, offender commit crime
culture	art, smith, museum, delawar, idaho, cultur, music, endow, recognit, artist	american have service, university conduct study, american do job, parent care child, someone seek recognition, program help community, federally fund project, program have (positive/negative) effect, applicant receive grant, people leave country
economy	busi, small, compani, capit, invest, contract, administr, job, creat, econom	person control small business, eligible participate program, program provide fund, small business employ workforce, agency give priority, agency require information , veteran control small business, service control small business, individual receive funding, founder start business
education	educ, school, student, teacher, program, colleg, year, high, state, higher	student use technology, student leave discipline, school reduce class size, school hire student, agency serve school, school use fund, student receive degree (education), agency receives grant, agency require information , student have education
employment and labor	worker, employ, employe, work, labor, job, wage, union, pay, hour	employee pay overtime, employee hire permanent, person earn dollar, men women earn dollar, worker join union, worker form union, person make dollar, market take healthcare, person make workforce, commonwealth perform duty
environment	environment, wast, water, regul, epa, facil, state, clean, protect, administr	individual take opportunity, child have disability, toxic (thing) cause cancer, national preserve social security, program need reform, environmental protection agency do job, supreme declare unconstitutional, republican want repeal (something), money justify money, republican want medicare
family	children, famili, child, parent, welfare, care, young, work, support, kid	girl have baby (boomer), action have effect, family adopt child, good start provide service, parent care child, good start serve child, family leave welfare, child witness violence, child have child, community participate program
farming	food, product, farm, agricultur, farmer, tobacco, produc, price, fda, crop	kid graduate school, F.D.A. regulate tobacco, tobacco cause cancer, child start smoking, food drug administration regulate tobacco, agency require information , nothing limit authority, law provide remedy, F.D.A. have authority, uniform serve country

Table F.2: Topics, Words and Narratives (ii)

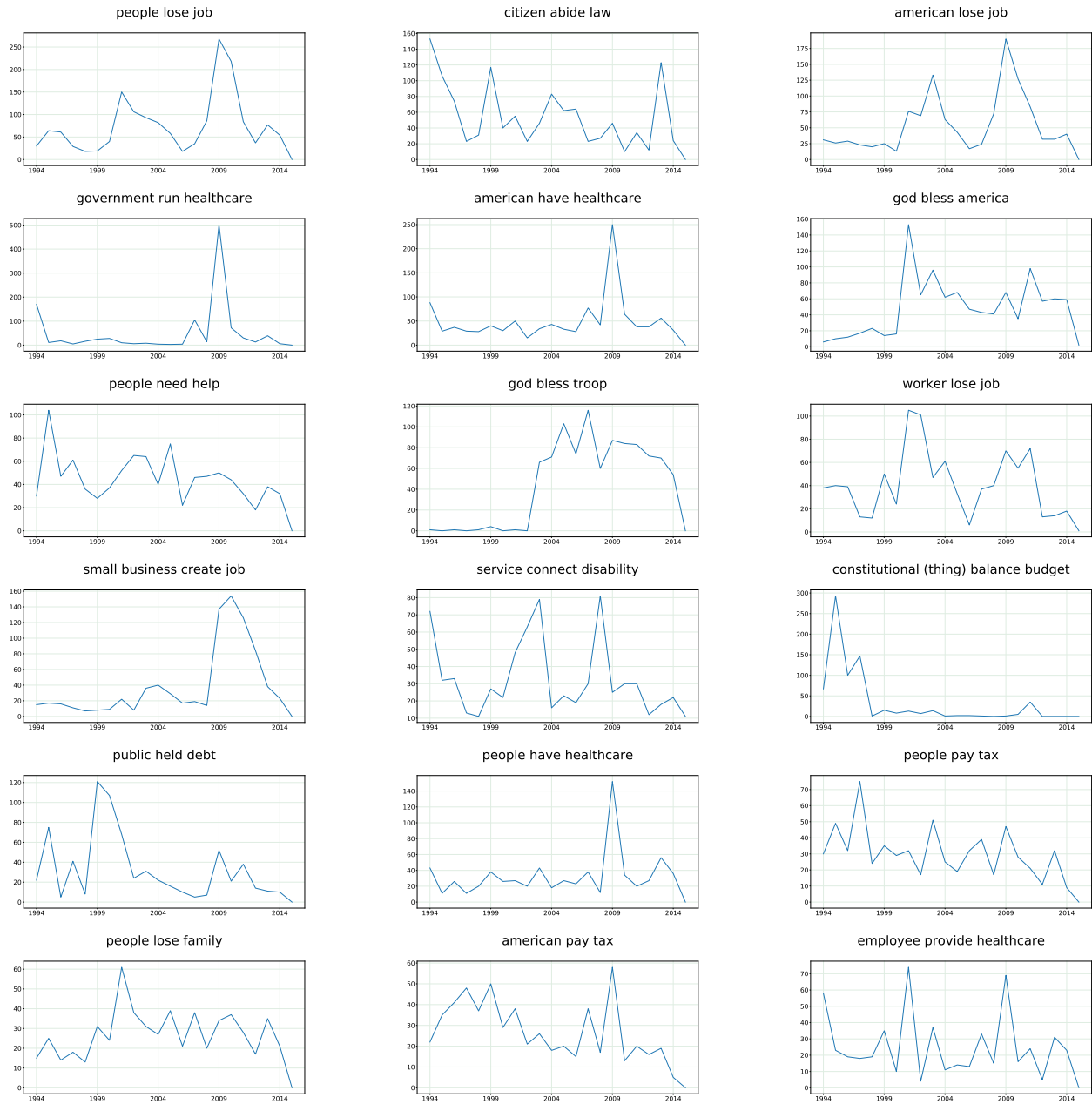
Topic	Top Words	Top Narratives
fiscal policy and taxation	tax, percent, american, incom, famili, pay, job, increas, year, would	measure breach budget, measure would budget, measure provide authority, program require appropriation, budget receive recommendation, reconciliation carry recommendation, measure not-increase deficit, public held debt, fact impose tax, budget add trillion
foreign policy	state, unit, nuclear, treati, weapon, nation, presid, peac, intern, iran	wealthy pay fair, iran obtain nuclear (weapon), iran acquire nuclear (weapon), iran develop nuclear (weapon), iran have nuclear (weapon), iran pose threat, nuclear (weapon) pose threat, bank do business, constitution prohibit america, people help people
gun violence	china, gun, chines, law, weapon, firearm, taiwan, kill, year, state	weapon pierce tank, national provide service, institute conduct study, customer sell gun, veteran need help, soldier pierce tank, gun abide law, crime get gun, witness give witness, american scratch staff
healthcare	hospit, rural, nurs, johnson, home, physician, mitchel, area, facil, lungren	person perform abortion, pregnancy feel pain, healthcare deny coverage, person have corrective action, physician deliver baby (boomer), insurance drop coverage, insurance deny coverage, physician perform abortion, doctor acquire disease, patient get child
housing	loan, hous, home, mortgag, bankruptci, rate, debt, famili, credit, interest	american make fund, homeowner face mortgage, soldier serve country, family face mortgage, law provide protection, federally assist housing, trade create job, people face mortgage, program provide housing, authority provide funding
human rights	unit, govern, state, wherea, right, peopl, human, freedom, intern, religi	program help community, message transmit concurrent, attorney (judge) represent client, people save family, american travel cuba, money amount taxable, burmese contain export, small business provide service, government make payment, business pay tax
immigration	secur, immigr, border, homeland, state, law, illeg, unit, alien, countri	bible fulfil self-employed, american hold market, country participate program, terrorist enter country, service administer program, employee hire illegal immigrant, people cross border, people enter country, employee hire immigrant, immigrant enter country
judiciary	fee, claim, patent, liabil, properti, damag, file, compani, lawsuit, case	party seek remedy, supreme have (legal) statute, association give grade, bush nominate circuit (courts), attorney (judge) follow law, attorney (judge) apply law, bush nominate attorney (judge), law confer power, agreement enters force, attorney (judge) interpret law
legislation	amend, motion, bill, ask, resolut, consent, unanim, read, question, presid	(sick/paid/unpaid) leave be columbia, (sick/paid/unpaid) leave be guam, (sick/paid/unpaid) leave be virgin, money amount taxable, message transmit concurrent, burmese contain export, corporation make charity, natural resource recommend substitute, service recommend substitute, individual make legislative (election)
military	veteran, servic, disabl, militari, care, member, benefit, mental, health, provid	air force request application, force base defense, service connect condition, veteran suffer illness, veteran earn money, veteran suffer diabetes, veteran receive compensation, defense submit service, service connect veteran, defense submit congressional
names	vote, smith, johnson, lewi, brown, taylor, davi, miller, weldon, peterson	application decrease appropriation, law authorize study, brave men women give family, government make payment, corporation make charity, veteran control small business, veteran administer law, nothing limit authority, person commit (legal) charge, congressional budget define money
national pride	american, nation, year, serv, great, honor, mani, today, life, one	people remember courage, god wipe shot, god grant strong, american hold market, word do justice, commonwealth perform duty, african make contribution, church offer prayer, hero give family, god bless family
native americans	indian, tribe, nativ, alaska, tribal, land, mine, trust, reserv, state	decision have effect, budget do job, someone get money, program carry program, employee perform service, healthcare save money, good start provide service, institute conduct study, someone know nothing, raw material meet road

Table F.3: Topics, Words and Narratives (iii)

Topic	Top Words	Top Narratives
natural resources / energy	energi, oil, gas, price, fuel, electr, product, use, produc, natur	someone buy gun, school use fund, ethanol blend gasoline, nation need energy, production shut supply and demand, production use american, oil account america, american rely oil, O.P.E.C. increase production, america need energy
nature	land, nation, area, park, forest, manag, conserv, resourc, state, acr	county mean county, city mean city , authority provide funding, worker earn minimum wage, grant receive grant, agency provide technical assistance, family leave welfare, someone buy gun, nothing affect authority, service do job
no label	presid, veto, congress, aid, reagan, item, bush, earmark, clinton, power	company bring job, story detail fight, energy affect family, people not-want job, american make facility, people send kid, people buy car, country lose job, small business hire people, people try job
political debate	speaker, time, gentleman, chairman, member, hous, would, vote, bill, colleagu	(sick/paid/unpaid) leave be guam, (sick/paid/unpaid) leave be columbia, (sick/paid/unpaid) leave be virgin , someone seek recognition, period offer appropriate, republican refuse action, people have commitment, democratic offer alternative, law authorize appointment, friend make comment
politics	senat, clerk, follow, pro, read, tempor, amend, presid, legisl, rule	commonwealth perform duty, judiciary conduct promulgate regulation, air force indicate appointment, self-employed executes budget, commitment have witness, news sign text, party raise soft money, witness limit medicine, group run ad, witness give witness
public services	shall, secretari, fund, provid, program, act, section, state, appropri, avail	law authorize uniform, U.S.C. authorize service, application decrease appropriation, fund subsidize deficit, U.S.C. authorize uniform, fund make funding, congressional budget define money, fund carry funding, person perform abortion, government estimate contract
science	nation, develop, research, technolog, communiti, program, support, sci-enc, center, improv	commonwealth perform duty, student pursue degree (education), director award grant, agency participate program, (sick/paid/unpaid) leave be columbia, federally fund study, fund pay profit, company participate program, program award grant, program achieve goal
social welfare	drug, medicar, senior, prescript, benefit, cost, plan, program, beneficiari, year	oil account america, production shut supply and demand, medicare operate prescription drug, american rely oil, production use american, senior join H.M.O., senior have coverage, medicare have prescription drug, senior need prescription drug, senior pay premium
sports	univers, team, year, colleg, game, state, nation, wherea, award, first	people have exhaust, parent take child, fund provide service, small business employ workforce, individual perform service, defense prescribed regulation , people sign petition, budget cut education , american obtain healthcare, individual participate program
states	water, project, river, flood, corp, construct, lake, feder, cost, louisiana	water resource development authorize river, person commit offense, university conduct study, law authorize project, project provide water, F.D.A. approve medicine, program have effect, person take healthcare, worker earn minimum wage, agency carry program
trade	trade, u.s, countri, agreement, unit, state, foreign, export, market, american	administration negotiate trade, china join wto, export create job, trade create job, china manipulate dollar, american make life quality, agreement enters force, country have deficit, constitution give authority, bank provide loan
transportation	transport, highway, safeti, airport, system, state, passeng, vehicl, air, rail	grant fund program, amtrak provide service, good start provide service, project meet regulation , railroad provide service, railroad provide transportation, project receive funding, association give grade, carrier provide service, government receive funding
war	iraq, war, troop, terrorist, american, attack, presid, militari, terror, iraqi	iraqi take responsibility, iraqi take power, iraq pose jeopardy, iraq have weapon mass destruction, disability receive education , saddam hussein pose threat, iraq pose threat, saddam hussein use weapon mass destruction, saddam hussein use chemical (weapon), saddam hussein have weapon mass destruction

G Additional Results

Figure G.1: Most Frequent Narratives Over Time



Note: In this figure, we plot the occurrences of the 18 most frequent narratives in the U.S. Congressional Record over the period 1994 to 2015. We focus on narratives pronounced at least 50 times and remove entities labeled as procedural or noise.

Table G.1: Most Frequent Narratives Per Entity

family	money	someone	agency
people lose family	american keep money	someone have opportunity	agency provide service
american lose family	people keep money	someone pay attention	agency administer program
family lose love	people make money	someone have healthcare	agency furnish information
men women put family	people get money	someone do nothing	agency do job
american give family	government take money	someone have idea	agency receive grant
men women risk family	people have money	someone take healthcare	agency take action
people give family	family keep money	someone pay tax	agency provide funding
men women give family	people earn money	someone do job	agency submit application
god bless family	people put money	someone lose job	agency receive fund
healthcare save family	hard earn money	someone make money	agency participate program
program	american	child	service
program provide service	american lose job	child have healthcare	service connect disability
program provide funding	american have healthcare	child attend school	program provide service
program provide healthcare	american pay tax	child get education	healthcare provide service
agency administer program	american lose healthcare	child receive education	U.S.C. authorize service
program provide grant	american keep money	child have education	agency provide service
program make difference	american lose family	child receive healthcare	government provide service
program create job	american look job	child have opportunity	community provide service
student participate program	american abide law	person have child	service provide service
federally fund program	american not-have healthcare	child need help	company provide service
government run program	american want job	child not-have healthcare	provider provide service
person	budget	law	republican
person have abortion	constitutional (thing) balance budget	citizen abide law	republican take power
person serve country	budget balance budget	american abide law	republican join democrat
person have child	republican balance budget	law respect faith	republican do nothing
person make difference	budget raise tax	people violate law	republican balance budget
person do job	budget increase tax	bush sign law	republican shut government
person earn dollar	budget do nothing	gun abide law	democrat join republican
person make decision	plan balance budget	people break law	republican cut medicare
person have healthcare	budget cut program	law prescribed witness	republican put together
person provide service	government balance budget	veteran administer law	republican cut program
person make disaster	budget reduce deficit	law make appropriation	republican give tax break
employee	healthcare	business	people
employee provide healthcare	government run healthcare	business create job	people lose job
employee sponsor healthcare	american have healthcare	company do business	people need help
employee offer healthcare	people have healthcare	government sponsor business	people have healthcare
employee drop coverage	employee provide healthcare	company hold business	people pay tax
employee provide coverage	american lose healthcare	american do business	people lose family
employee take (paid/unpaid) leave	healthcare provide healthcare	corporation do business	people do job
employee sponsor plan	healthcare provide service	government do business	people take opportunity
employee sponsor coverage	child have healthcare	people do business	people make decision
employee provide service	people lose healthcare	business do business	people look job
employee hire worker	people get healthcare	business hire worker	people have concern
government	job	tax	country
government run healthcare	people lose job	people pay tax	men women serve country
government take healthcare	american lose job	american pay tax	person serve country
government provide service	worker lose job	budget raise tax	veteran serve country
government make decision	small business create job	capital gain tax	problem face country
government take money	people do job	democrat raise tax	people serve country
republican shut government	people look job	budget increase tax	challenge face country
government sponsor business	business create job	tax kill job	people enter country
government do nothing	american look job	family pay tax	country face challenge
government take property	people have job	someone pay tax	american serve country
government run program	people want job	people not-pay tax	country produce oil

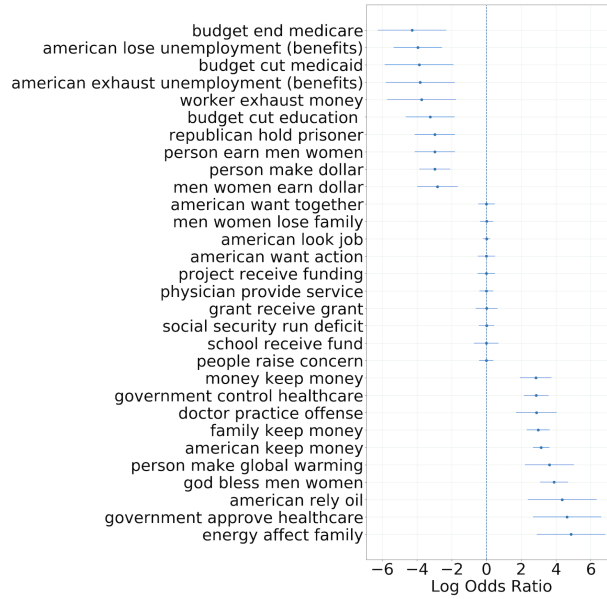
Note: In this table, we plot the ten most frequent narratives for the 20 most frequent entities in the U.S. Congressional Record (1994-2015).

H Additional Material on Divisive Entities

Partisan and Neutral Narratives

Figure H.1 shows again the most Republican and Democrat narratives, now along with the most neutral narratives.

Figure H.1: Partisan Narratives in U.S. Congress



Note: This figure presents the most partisan narratives in the U.S. Congressional Record, as well as the least partisan narratives (i.e., with log-odds ratio closest to 0). For each narrative, we present the associated log-odds ratio and 95% confidence interval. A high log-odds ratio reflects narratives pronounced more by Republicans (relative to Democrats), and vice versa for a low log-odds ratio. One of the least partisan narratives, “men women lose family”, was pronounced most often in 2001 – suggesting that his narrative captures a reaction to the 09/11 terrorist attacks.

Table H.1: Longer, Unfiltered List of Divisive Entities

Rank	Entity	Score	Log-Odds Ratio		# Narratives		
			Narratives	Entity	Total	Dem	Repub
1	filibuster	1.93	2.10	0.16	2	1	1
2	oil	1.12	1.39	0.27	8	4	4
3	H.M.O.	0.96	1.05	0.10	2	1	1
4	remedy	0.91	0.94	0.02	2	1	1
5	contract	0.87	1.18	0.31	4	2	2
6	mortgage	0.79	1.03	0.24	5	3	2
7	budget	0.76	0.82	0.06	42	17	25
8	men women*	0.75	0.86	0.10	21	10	11
9	workforce	0.75	1.26	0.51	2	1	1
10	economy	0.74	0.78	0.04	4	2	2
11	pregnancy	0.73	1.07	0.34	3	2	1
12	medicaid	0.73	2.43	1.71	2	2	0
13	facility	0.71	0.92	0.21	3	2	1
14	policy	0.70	0.83	0.14	2	1	1
15	gun	0.69	0.85	0.16	6	4	2
16	healthcare	0.68	0.74	0.06	116	76	40
17	environmental protection agency	0.68	1.13	0.46	2	1	1
18	interest/rate	0.65	0.97	0.32	8	5	3
19	loan	0.65	0.72	0.07	12	6	6
20	center	0.64	0.79	0.15	2	1	1
21	democrat*	0.64	1.03	0.40	11	4	7
22	constitution*	0.62	0.63	0.01	8	5	3
23	worker	0.61	1.23	0.62	16	11	5
24	american*	0.60	0.61	0.01	158	86	72
25	plant	0.59	0.97	0.38	3	2	1
26	insurance	0.58	0.67	0.09	16	7	9
27	trade	0.57	0.74	0.17	2	1	1
28	action	0.57	0.61	0.05	14	6	8
29	resource	0.56	0.79	0.23	4	3	1
30	medicine	0.55	0.90	0.35	3	1	2
31	priority	0.54	0.93	0.39	3	2	1
32	community*	0.53	0.60	0.07	14	10	4
33	bank	0.52	0.88	0.35	6	4	2
34	abortion	0.52	0.92	0.39	10	2	8
35	republican*	0.52	1.12	0.60	40	31	9
36	medicare	0.52	0.83	0.31	18	12	6
37	energy	0.52	1.30	0.78	7	3	4
38	job	0.51	0.65	0.14	117	73	44
39	nation*	0.50	0.54	0.04	28	15	13
40	fund	0.49	0.49	0.00	19	8	11
41	plan	0.48	0.53	0.05	20	9	11
42	troop	0.48	0.64	0.16	3	1	2
43	missile	0.48	0.53	0.05	3	1	2
44	weapon mass destruction	0.47	0.64	0.16	3	2	1
45	doctor	0.47	0.79	0.31	16	6	10
46	company	0.46	0.87	0.41	17	11	6
47	people*	0.46	0.49	0.04	240	135	105
48	F.D.A.	0.45	0.68	0.23	8	4	4
49	family*	0.45	0.55	0.10	90	63	27
50	bad (problem)	0.45	0.46	0.01	2	1	1

Note: This table shows the full list of most divisive narratives, as measured by the score in column 3, given by average absolute log-odds ratio of the narratives where the entity appears (column 4), minus the absolute log odds ratio of the entity itself (column 5). The other columns give the number of unique narratives (total, Democrat-slanted, and Republican-slanted) where the entity appears. Bold type indicates a policy narrative listed in Table 2. Asterisks indicate an identity/symbolic narrative listed in Table 2.

Table H.2: List of Entities with Least Divisive Narratives

Rank	Entity	Score	Log-Odds Ratio		# Narratives		
			Narratives	Entity	Total	Dem	Repub
416	population	-0.04	0.06	0.10	1	0	1
417	client	-0.04	0.06	0.10	1	0	1
418	(positive/negative) effect	-0.04	0.06	0.10	1	0	1
419	strong	-0.04	0.34	0.38	1	0	1
420	union	-0.04	0.02	0.06	1	0	1
421	natural resource	-0.04	0.02	0.06	1	0	1
422	shot	-0.04	0.06	0.10	1	0	1
423	patient	-0.04	0.23	0.27	7	0	7
424	reform	-0.04	0.81	0.85	2	0	2
425	opinion	-0.05	0.10	0.14	3	0	3
426	truth	-0.05	0.18	0.22	2	0	2
427	buy	-0.05	0.07	0.11	2	0	2
428	county	-0.05	0.17	0.22	3	0	3
429	consequence	-0.06	0.75	0.81	2	0	2
430	railroad	-0.06	0.46	0.52	2	2	0
431	air force	-0.06	0.73	0.79	2	1	1
432	founder	-0.07	0.52	0.59	3	1	2
433	god	-0.08	1.04	1.12	9	1	8
434	border	-0.08	0.66	0.73	2	1	1
435	profit	-0.10	1.10	1.20	8	8	0
436	adult	-0.12	0.37	0.49	4	4	0
437	force	-0.15	0.27	0.42	2	0	2
438	congressional	-0.15	0.75	0.90	2	0	2
439	(sick/paid/unpaid) leave	-0.17	1.32	1.49	2	0	2
440	U.S.C.	-0.18	1.37	1.55	2	0	2
441	freedom	-0.19	0.58	0.77	4	0	4
442	legislative (election)	-0.19	0.58	0.77	2	0	2
443	word	-0.24	0.53	0.77	3	2	1
444	winners and losers	-0.53	0.81	1.33	2	0	2
445	appointment	-0.57	1.44	2.01	2	1	1

Note: This table shows the least divisive narratives, as measured by the score in column 3, given by average absolute log-odds ratio of the narratives where the entity appears (column 4), minus the absolute log odds ratio of the entity itself (column 5). The other columns give the number of unique narratives (total, Democrat-slanted, and Republican-slanted) where the entity appears.

I Which narratives co-occur together?

In many cases, agent-verb-patient tuples are building blocks of a broader political story. In Figure 3a, for example, we show that two prevalent narratives related to the war in Iraq are: “Saddam Hussein has weapons of mass destruction” and “Saddam Hussein poses a threat”. A more comprehensive story might read: “Saddam Hussein has weapons of mass destruction and poses a threat to national security. With a military intervention, the United States can make a difference, but at the cost of American lives.” Similarly, “people lose job” is the most common narrative in the corpus, but to which political stories does it belong to? Can we identify narratives that co-occur together to eventually reconstruct these broader political stories? As a first step in this direction, we build on two common approaches in the text-as-data literature below.

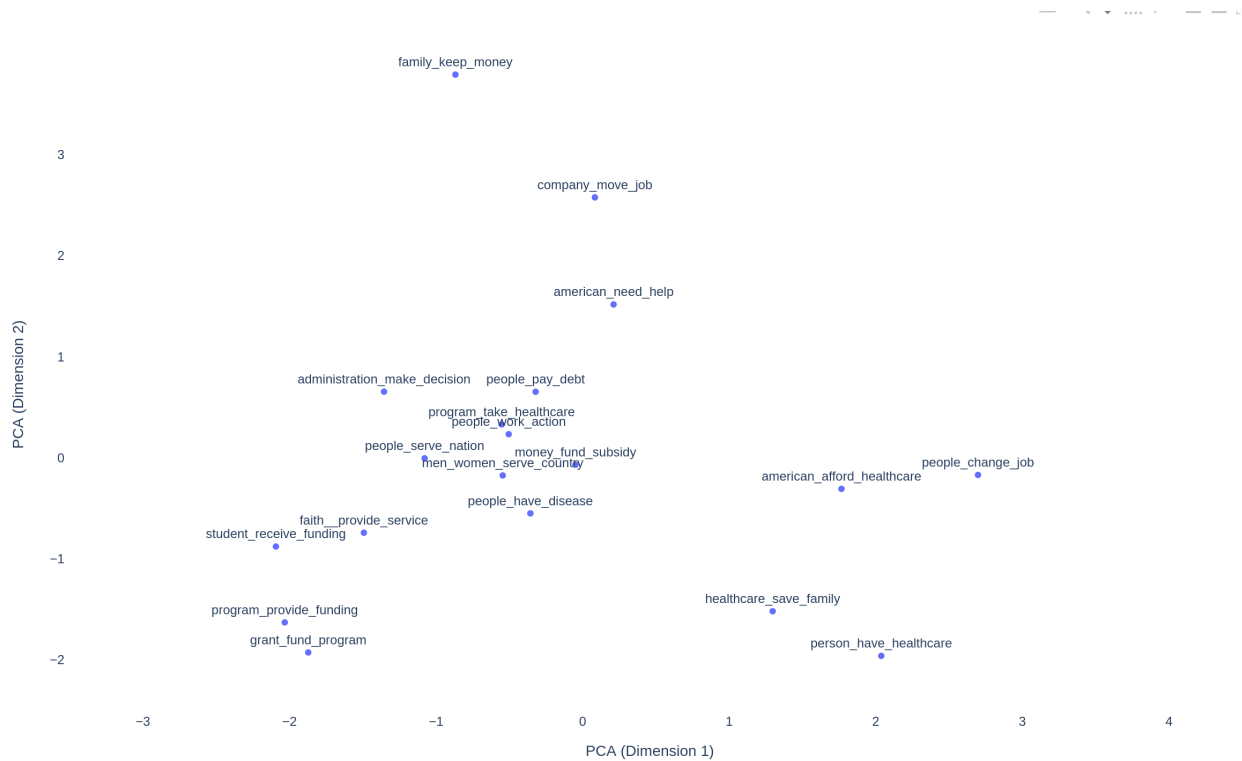
Pointwise Mutual Information. A simple approach to recover co-occurring narratives is to compute the pointwise mutual information (PMI) between pairs of narratives. For two narratives n_i and n_j , the pointwise mutual information is defined as:

$$PMI(n_i, n_j) = \log \left(\frac{P(n_i \cap n_j)}{P(n_i)P(n_j)} \right) \quad (3)$$

If two narratives are independent from one another, then their PMI is equal to zero. The larger the PMI, the more likely two narratives are to co-occur together, adjusting for their overall frequency. To analyze the odds of two narratives co-occurring in the same speech, we compute PMIs for pairs of narratives in the U.S. Congressional Record.

Narrative Embedding. A second approach is to train skip-gram embeddings for narratives. As done with word embeddings (Mikolov et al., 2013), this approach locates narratives that appear in similar contexts close to one another in a low-dimensional, dense vectorial space. For the U.S. Congressional Record, we define speeches as documents, narratives as tokens, and then train a Word2Vec model (skip gram with negative sampling) over ten epochs. The window size is set to an arbitrarily large number so that two narratives co-occurring in a speech are always in their respective windows. The dimension of vectors is set to 50. To graphically represent the resulting embedding, Figure I.1 plots a random sample of 20 narratives along the two first dimensions of a Principal Component Analysis (PCA), showing how the geometry represents intuitive topical information. Hence, the relationship between two narratives can be computed as the cosine similarity between their respective vectors.

Figure I.1: Geometry of Embedded Narratives



Note: This figure shows 20 random narratives in the embedding, projected to two dimensions using PCA.

Results. Table ?? applies these methods to the narrative “Saddam Hussein poses a threat”. We provide the 20 most associated narratives with our PMI metric as well as the cosine similarity of the narrative embedding vectors. Overall, both methods paint a similar picture. Saddam Hussein posed a threat to national security because of (supposedly possessed) nuclear weapons, the United States won the war but American troops had to make a sacrifice. In fact, Saddam Hussein did not possess nuclear weapons. The administration made a mistake and misled Americans.

Table I.1: Co-occurring Narratives – War on Terror

Pointwise Mutual Information	Narrative Embedding
iraq pose threat	iraq pose threat
iraq pose jeopardy	nation love peace
saddam hussein use chemical (weapon)	people love freedom
iraq have weapon mass destruction	american lose job
saddam hussein have weapon mass destruction	god bless troop
saddam hussein invade kuwait	people do job
saddam hussein use weapon mass destruction	saddam hussein have weapon mass destruction
administration mislead american	god bless america
freedom take tree	american lose family
iraqi take responsibility	iraq pose jeopardy
american risk family	men women put family
nation love peace	saddam hussein use chemical (weapon)
men women wear uniform	men women wear uniform
people love freedom	administration make mistake
administration make mistake	person serve country
iraqi take power	terrorist kill american
nation love freedom	economy create job
war torn country	democrat join republican
terrorist pose threat	american do job
military do job	people have problem

Note: This table presents narratives that tend to co-occur with “Saddam Hussein poses a threat”. Column (1) lists the 20 narratives with the highest PMI associated to this narrative. Column (2) lists the 20 narratives that are the closest to this narrative in the embedding (as measured by cosine similarity).

J Narratives on Social Media Data

Though we illustrate RELATIO’s workings on the U.S. Congressional Record, the package is not designed to be corpus-specific. In this section, we provide a short analysis of a Twitter corpus.⁵ We focus on 32,323 tweets from the [Trump Tweet Archive](#). The tweets were written by Donald Trump between August 2011 and October 2020 and amount to 68,616 sentences.

We clean the text the same way as done for the congressional speeches, with the exception of stopwords, taken directly from the spaCy english model “en_web_core_md”. We keep the 100 most frequent named entities and specify 100 additional clusters for the embeddings. We then count the resulting narrative tuples.

Figure J.1 represents Trump’s 50 most frequent narratives in the form of a narrative graph. The former president notably claims that “The Democrats hurt the country”, “steal the election” and “raise taxes”. “China manipulates its currency” and “steals from the people”, while “the establishment and special interests kill the country”.

The corpus is much smaller than the U.S. Congressional Record. It takes just a few minutes to run the full RELATIO pipeline on it. Interested readers will find a step-by-step tutorial to extract narratives from this corpus on our [GitHub repository](#).

⁵Researchers not involved in the development of RELATIO have applied it to newspaper articles ([Otonello and Song, 2022](#)) as well as social media data from different platforms ([Sipka, Hannak and Urman, 2021](#)).

K List of Analyzed Narratives

Here we list all narratives mined from the Congressional Record corpus. The list includes dimension-reduced narratives appearing at least 50 times in the period 1994-2015. Narratives are sorted by most to least frequent.

people lose job, citizen abide law, american lose job, government run healthcare, american have healthcare, god bless america, people need help, god bless troop, worker lose job, small business create job, service connect disability, constitutional (thing) balance budget, people have healthcare, public held debt, people pay tax, people lose family, american pay tax, employee provide healthcare, people do job, men women serve country, people take opportunity, mortgage back military, american lose healthcare, american keep money, american lose family, program provide service, people make decision, people look job, healthcare provide healthcare, business create job, people have concern, american look job, family lose love, program provide funding, healthcare provide service, people have job, child have healthcare, people lose healthcare, company do business, people have opportunity, people want job, people keep money, american abide law, people get healthcare, U.S.C. authorize service, F.D.A. approve drug, people have insurance, agency provide service, american not-have healthcare, child attend school, period offer appropriate, program provide healthcare, government take healthcare, person have abortion, student attend school, republican take power, people get job, amount bear ratio, people not-have healthcare, defense submit congressional, people love freedom, people take healthcare, american want job, person serve country, american work hard, someone have opportunity, veteran receive healthcare, men women put family, people create job, someone pay attention, someone have healthcare, men women serve nation, people make money, government provide service, veteran serve country, god bless men women, someone do nothing, doctor make decision, republican join democrat, problem face country, agency administer program, agency furnish information , people find job, oil make profit, american give family, people commit crime, doctor take healthcare, law respect faith , child get education , program provide grant, republican do nothing, child receive education , community provide service, budget balance budget, american get healthcare, democrat take power, government make decision, american have concern, employee sponsor healthcare, administration take action, american pay attention, child have education , project create job, challenge face nation, people pay attention, men women risk family, service provide service, people have idea, people serve country, constitutional (thing) require balance budget, american need help, iraq pose threat, agency do job, american do job, program make difference, insurance deny coverage, american find job, challenge face country, economy create job, coal fire plant, program create job, student participate program, people get money, people have problem, government take money, people need

healthcare, people give family, men women give family, republican balance budget, republican shut government, bank make loan, god bless family, law enforcement add category, budget raise tax, government sponsor business, american take opportunity, problem face nation, federally fund program, federally fund study, federally recognize tribe, child receive healthcare, saddam hussein pose threat, healthcare save family, small business provide healthcare, people have money, family make sacrifice, people make choice, capital gain tax, people cross border, democrat raise tax, america have healthcare, child have opportunity, employee offer healthcare, commonwealth perform duty, american make sacrifice, director mean director, god continue america, american have opportunity, men women wear uniform, government do nothing, democrat join republican, company provide service, company hold business, person have child, family need help, budget increase tax, missile propel bomb, someone have idea, agency receive grant, people enter country, american want healthcare, family keep money, doctor provide healthcare, individual not-attain adult, legislative (election) have consequence, child need help, public create job, budget do nothing, country face challenge, government take property, judiciary conduct markup, agency take action, healthcare offer healthcare, family receive funding, administration do nothing, american create job, veteran make sacrifice, american do business, republican cut medicare, child not-have healthcare, people risk family, person make difference, american serve country, someone take healthcare, men women make sacrifice, tax kill job, tide lift ship, provider provide service, plan balance budget, america lead world, people have choice, american receive healthcare, family pay tax, government run program, american have job, people buy healthcare, program award grant, people earn money, people buy insurance, government take action, someone pay tax, people not-pay tax, company make profit, kid get education , people put family, senior have healthcare, corporation pay tax, attorney (judge) make decision, corporation do business, center provide service, person do job, student graduate school, nation face challenge, senior have prescription drug, people violate law, government do everything, american lack healthcare, agreement enters force, school provide education , government do business, government control healthcare, doctor provide service, person earn dollar, raw material meet road, energy create job, people not-have insurance, people express concern, lender make loan, job pay life quality, employee drop coverage, brave men women give family, company create job, employee provide coverage, healthcare offering healthcare, people have family, parent raise child, people buy car, parent send child, employee sponsor plan, person make decision, employee take (sick/paid/unpaid) leave, service connect veteran, program save family, child need healthcare, agency provide funding, people do business, american make choice, iraqi take responsibility, founder create job, people have child, american lose unemployment (benefits), staff do job, people put money, american need job, nation produce oil, director considers appropriate, saddam hussein have weapon mass destruction, bush sign law, crime add category, person have healthcare, budget cut program, family have healthcare, company move job, business do business, physician

provide service, people work job, program receive funding, government do job, people take risk, people lose insurance, stimulus package create job, country produce oil, american know truth, defense waive limit, people not-have job, american make energy, someone do job, hard earn money, people receive money, people need job, mother have child, program save money, program help american, gun abide law, men women serve military, people get education , senior pay premium, (sick/paid/unpaid) leave be guam, doctor perform abortion, american risk family, men women lose family, republican put together, defense notify congressional, individual receive service, terrorist kill american, child have diabetes, people provide service, outlay not-exceed payment, people have coverage, american make decision, american suffer diabetes, plan do nothing, american lose coverage, program provide education , people break law, veteran need healthcare, child review study, agency submit application, senior have choice, person provide service, american not-do job, people make sacrifice, people put together, american have choice, people have skill, individual purchase healthcare, iran develop nuclear (weapon), people need funding, employee sponsor coverage, god wipe shot, person make disaster, american need healthcare, child participate program, law prescribed witness, people participate program, people graduate school, people not-find job, government pick winners and losers, nation owes appreciation, democrat want tax, god grant strong, american make consumer, F.D.A. approve prescription drug, senior get prescription drug, republican cut program, people save money, energy affect family, person give family, government balance budget, fund create job, program help people, american need understand, brave men women risk family, republican give tax break, country support terrorist, people make mistake, person make dollar, agency receive fund, republican cut tax, country face problem, mother raise child, child reach future, family work hard, doctor practice healthcare, family send child, american sent message, program provide food, voter cast ballot, family take healthcare, individual have healthcare, people have income, employee provide service, american travel cuba, someone lose job, program provide money, person lose family, church offer prayer, people lose money, person dedicate family, american need (tax) relief, school hire student, people take money, people need service, american pay income, employee hire worker, administration make decision, people give money, employee drop healthcare, fuel drive economy, budget reduce deficit, grant submit application, people make living, individual participate program, veteran administer law, people make contribution, employee lose job, supreme make decision, story detail fight, company take opportunity, carrier provide service, people raise concern, constitution require balance budget, individual perform service, someone make money, individual provide service, american have diabetes, budget increase deficit, school participate program, vacancy not-affect power, employee pay minimum wage, people make difference, people drive car, program meet regulation , men women do job, person seek abortion, child receive service, customer sell gun, senior need help, people work minimum wage, people pay money, people get insurance, american have insurance, regulation kill job, family have child,

people suffer illness, people need understand, agency participate program, business hire worker, child get healthcare, action have consequence, student need help, veteran receive service, someone get money, drug save family, veteran receive money, person give baby (boomer), agency use fund, child lose parent, citizen pay tax, administration request level (spending), employee hire immigrant, someone seek recognition, republican play politics, small business offer healthcare, government approve healthcare, person have corrective action, american put food, storm cause damage, child witness violence, people do nothing, doctor deliver baby (boomer), american hold market, people call tax, student reduce class size, individual receive money, person have cancer, iran have nuclear (weapon), people work living, medicare provide healthcare, america face challenge, student receive grant, american pay money, american balance budget, life quality pay job, money pick deductible, expenditure not-exceed payment, mother take healthcare, car get gasoline, plan create job, american suffer illness, republican put plan, word have word, grant award grant, iran acquire nuclear (weapon), american want balance budget, budget cut deficit, policy create job, government fund program, car carry passenger, individual receive funding, senior have coverage, republican have idea, illness threaten family, person make contribution, F.D.A. regulate tobacco, american pay gasoline, people have diabetes, economy lose job, someone know nothing, person serve nation, country receive funding, government raise tax, people pay premium, people take grant, attorney (judge) do job, employee provide insurance, neighbor help neighbor, american make contribution, employee participate program, republican have plan, agency issue regulation, mother have diabetes, administration make mistake, american get tax, agency receive funding, firefighter lose family, medicare operate prescription drug, person have pregnancy, people have experience, people work family, pregnancy feel pain, plan meet regulation, people have education, american have government, government borrow money, judiciary recommend substitute, money pay tax, republican propose tax, physician take healthcare, freedom take tree, agency have authority, republican propose budget, attorney (judge) have (legal) statute, federally fund center, word do justice, student take loan, american buy healthcare, people work government, government solve problem, program serve child, people not-have opportunity, family give afghanistan, american want government, money save money, budget cut tax, people make dollar, student receive education, individual attain adult, terrorist pose threat, consumer have choice, grant make grant, brave men women put family, person serve military, american not-find job, program help child, challenge face america, school provide service, employee hire illegal immigrant, family raise child, country need energy, family send kid, american tighten belt, group run ad, program require appropriation, veteran control small business, people start business, company make money, people make fund, someone pay fair, child have future, tax create job, american lose money, friend make comment, F.D.A. approve consumer, government not-create job, law make appropriation, people work hard, people earn minimum wage, senior pay pocket, people need money, people re-

member courage, child suffer diabetes, someone take opportunity, american get job, people feel life quality, attorney (judge) follow law, appropriation do job, democrat do nothing, corporation make profit, someone make mistake, someone have insurance, hard earn dollar, agency have (legal) statute, bank issue card, government provide healthcare, money keep money, people want government, american make facility, people have illness, veteran get healthcare, employee do job, veteran have service, people make minimum wage, homeowner face mortgage, brave men women serve country, person perform abortion, program provide loan, program improve healthcare, parent make decision, nation face threat, agency receives grant, people lose coverage, people have program, student have education , school need help, american afford healthcare, defense prescribed regulation , republican play game, employee make contribution, disease threaten family, troop do job, american have coverage, crisis face country, constitution give authority, fund subsidize deficit, american take healthcare, american have voice, witness give witness, budget end medicare, people lose unemployment (benefits), person commit crime, people take job, worker join union, government create job, senior need prescription drug, authority provide funding, kid have healthcare, republican want medicare, H.M.O. make decision, program have effect, farmer need help, people love peace, mother have cancer, grant provide funding, people suffer diabetes, director award grant, tax benefit wealthy, people invest money, word give education , republican do job, country face crisis, people abide law, decision affect family, brave men women lose family, someone want job, appropriation make grant, government run deficit, business employ people, agency provide information , budget have deficit, someone have problem, person commit (legal) charge, agency serve school, people change job, american save money, person love family, worker form union, association give grade, community base healthcare, someone make decision, child enter school, men women earn dollar, animal lay milk, veteran suffer diabetes, immigrant enter country, people raise family, child graduate school, people have voice, corporation not-pay tax, child have disability, social security balance budget, american have money, someone get job, administration impose sanction, people filing bankruptcy, nothing create job, american deserve truth, agency provide technical assistance, injury fire missile, soldier lose family, democrat put together, government provide funding, hard work american, constitution give power, social security run deficit, fund provide service, student take opportunity, grant fund project, law authorize compensation, law enforcement do job, service administer program, father serve war, food drug administration regulate tobacco, problem face american, iraq have weapon mass destruction, employee take opportunity, action have effect, people make future, grant use grant, people want healthcare, attorney (judge) interpret law, community promote abortion, child have family, family lose job, crime get gun, law authorize uniform, worker find job, doctor treat medicare, insurance raise interest/rate, obligation bearing interest/rate, agency enter contract, person get corrective action, people afford healthcare, doctor take medicare, people get help, employee pay premium, nation face crisis, people serve

nation, doctor treat patient, money justify money, people have faith , business provide service, small business take opportunity, american earn money, company hold loan, project receive funding, american take job, senior take opportunity, military do job, someone loses job, corrective action save family, american get money, american put family, business hire people, people do best, service recommend substitute, people serve war, republican hold prisoner, attorney (judge) apply law, physician perform abortion, small business hire worker, america create job, american lose faith , american purchase healthcare, asset back military, self-employed executes budget, soldier give family, individual meet regulation , supreme rule unconstitutional, iran pose threat, people have difficulty, fund make funding, money pay money, program provide fund, community receive grant, mother give baby (boomer), bank provide loan, veteran serve nation, parent lose child, democratic offer alternative, physician assist suicide, money fund outlay, administration propose budget, institute conduct study, F.D.A. have authority, fed raise interest/rate, program provide resource, outlay exceed payment, program benefit american, agency share information , employee perform job, family put food, application meet regulation , reserve raise interest/rate, student receive degree (education), threat face nation, company employ people, people want buy, men women defend nation, electronic device save family, american hold love, people lose everything, appropriation fund program, budget reduces deficit, student get education , employee make decision, agency require information , service control small business, employee hire people, person take healthcare, american want together, crime commit crime, american keep healthcare, china join wto, person lose job, defense submit service, people lose love, republican want tax, tax benefit american, kid need help, physician deliver baby (boomer), news sign text, appropriation report appropriation, attorney (judge) enforce law, people do harm, program help poor, employee provide institution, god bless nation, people borrow money, american paid tax, crime use gun, project meet criterion, american deserve (tax) relief, fact impose tax, business make profit, small business hire people, people keep staff, network provide service, american face problem, american face challenge, people want program, measure breach budget, people want help, government have power, country have deficit, american have enough, administration submit budget, business provide healthcare, child lose healthcare, people solve problem, tax stimulate economy, employee pay overtime, company sell consumer, public provide service, american want action, people have exhaust, employee hire permanent, director determines appropriate, person have baby (boomer), american have faith , worker earn minimum wage, student pursue education , agency receives fund, environmental protection agency regulate emission (pollution), american serve nation, government estimate contract, indian provide service, american send money, challenge face american, people take responsibility, attorney (judge) make determination, physician practice healthcare, justice write opinion, democrat have plan, country need help, project meet regulation , country make progress, nothing limit authority, government not-do nothing, student afford school, american rely oil, production shut

supply and demand, company pay tax, worker support retiree, self-employed deduct healthcare, child have special (need/rule), school do job, water resource development authorize river, patient receive healthcare, service do job, people need (tax) relief, people do wrong, healthcare deny coverage, healthcare provide coverage, american keep job, measure would budget, program mean program, country sign treaty, bush nominate attorney (judge), community need help, people paid tax, business take opportunity, people wear clothes, people work money, government receive funding, people have government, democrat balance budget, people keep healthcare, F.D.A. approve facility, american pay fair, republican put budget, people serve military, someone have fact, iran obtain nuclear (weapon), physician make decision, mother pay tax, person make global warming, veteran need help, administration do everything, country participate program, ship carry passenger, person risk family, nation take action, veteran put family, appropriate rely payment, people live family, person earn men women, republican create job, country face threat, kid have opportunity, republican use filibuster, administration request amount, family have income, money fund abortion, threat face country, administration cut program, people lose saving, administration do job, people take action, offender commit crime, coal fire electricity, faith provide service, country have problem, law require budget, parent make choice, policy held debt, people make profit, defense transfer fund, manufacturer collect sale, government make promise, farmer plant crop, small business purchase healthcare, law authorize program, nation face problem, tax have effect, agency make decision, community make decision, grant give priority, budget cut education , people not-afford healthcare, administration threaten veto, immigrant pay tax, program benefit people, people provide healthcare, founder start business, american have housing, american have service, national preserve social security, money fund program, american provide family, federally fund project, patient save family, education award grant, parent take child, someone have job, interest/rate authorize agency, innocent lose family, ship launch missile, individual pay tax, oil account america, parent have child, law authorize appointment, parent protect child, nothing do nothing, insurance drop coverage, country take opportunity, mother lose job, hero give family, people try help, people have power, legislative (election) call word, community provide funding, people pay debt, budget add trillion, veteran find job, american want reform, government run insurance, child have child, law prohibit practice, people have life quality, person make workforce, force base defense, someone raise staff, law authorize project, people do everything, staff put together, war torn country, plan provide (tax) relief, country lose job, budget reach together, budget report budget, american make ultimate sacrifice, agency award contract, people have opinion, kid graduate school, girl attend school, flood cause damage, senior have medicare, someone need help, men women fight war, american make difference, republican give tax, agency carry program, country have healthcare, physician provide healthcare, american deserve nothing, people receive funding, people make family, employee use (sick/paid/unpaid) leave, terrorist enter country, funding orphan child, juve-

nile commit violent crime, family adopt child, administration take position, job require skill, news face nation, people have cosponsor, republican end medicare, applicant submit application, program provide help, money fund subsidy, student leave school, story wrench heart, message transmit concurrent, news face country, plan have effect, person terminate pregnancy, patient need healthcare, medicare have prescription drug, people work small business, decision have effect, job pay minimum wage, people put food, child lose family, doctor make healthcare, person choose abortion, worker pay tax, terrorist attack country, american hold accountable, soldier pierce tank, small business pay tax, doctor practice offense, people support family, community participate program, someone get tax, institution provide service, people have confusion, american access healthcare, american have ballot, family make decision, raw material hit road, fund pay profit, people leave welfare, iraq invade kuwait, government have authority, student enter school, U.S.C. authorize uniform, someone commit crime, family balance budget, someone have experience, law provide authority, child eat lunch, medicaid provide healthcare, national conduct study, government issue identification, veteran receive compensation, law provide remedy, city mean city , grant provide service, administration propose tax, american do nothing, military make sacrifice, supreme declare unconstitutional, government take opportunity, america need help, small business create america, senior join H.M.O., family save money, people keep job, firefighter open fire, government issue picture, people move job, air force request application, internal revenue code strike period, mother have baby (boomer), employee sponsor insurance, corporation pay fair, bank lend money, government pay money, kid start smoking, someone do business, program outlive effectiveness, program provide profit, someone have heart, nation sign treaty, god rest love, company have employee, individual lose job, worker do job, student make difference, government save money, agency conduct analysis, budget cut medicaid, people go job, parent care child, administration provide information , small business get loan, american do best, money outweigh money, physician treat patient, unemployment (benefits) look job, attorney (judge) make law, people get coverage, people balance budget, individual make decision, employee create job, program serve american, medicare cover service, american work money, railroad provide transportation, money get money, program need reform, filibuster raise tax, america need energy, agency make determination, wealthy pay fair, people help people, student receive funding, party reach agreement, program help family, veteran take opportunity, natural resource recommend substitute, people kill people, nuclear (weapon) pose threat, people have common sense, child lack healthcare, program take healthcare, people wear uniform, O.P.E.C. increase production, person have diabetes, market take healthcare, people have disease, senior pay deductible, republican cut education , program have success, police give family, consumer cause harm, child start smoking, american use site, person make choice, oil hit gasoline, republican take action, parent take healthcare, individual lose family, small business pool resource, child get good start, someone paid attention, police do job, child reach adult, government

provide information , school use fund, healthcare improve healthcare, people preexist condition, employee not-provide healthcare, person make payment, program provide technical assistance, person commit offense, american deserve healthcare, american receive money, employee offering healthcare, republican want repeal (something), people want tax, people share concern, veteran receive recognition, senior make choise, people want america, senior afford prescription drug, job require education , nation love freedom, people not-get job, family leave welfare, parent send kid, appropriation provide funding, (sick/paid/unpaid) leave be columbia, father have diabetes, family face mortgage, government run plan, law authorize activity, capital gain interest/rate, american put money, veteran have healthcare, insurance provide coverage, america win war, money make difference, program provide essential , veteran participate program, university conduct study, authority make decision, measure not-increase deficit, doctor acquire disease, student use technology, bank provide service, american work government, problem face america, veteran serve war, veteran seek healthcare, person need help, american obtain healthcare, senior depend medicare, people have plan, county mean county, winners and losers appropriate fund, family take opportunity, employee provide money, family make choise, facility provide service, founder endow men women, person have heart, men women serve arm force, people have appreciation, tobacco cause cancer, county lose population, people make america, juvenile commit crime, amtrak provide service, saddam hussein use chemical (weapon), person lay family, individual take opportunity, fund carry funding, saddam hussein use weapon mass destruction, government have responsibility, community use fund, party raise soft money, nation lead world, small business provide service, agency use information , veteran earn money, bureaucrat make decision, american not-afford healthcare, employee have healthcare, application decrease appropriation, party seek remedy, american feel pain, agency promulgate regulation , oil pay fee, student leave discipline, individual buy healthcare, people have feeling, major issue face country, commitment have witness, budget provide (tax) relief, soldier serve country, grant fund program, saddam hussein invade kuwait, person control small business, country lead world, medicare run money, worker exhaust money, people pay income, country make fund, family receive money, agency receive amount, american make fund, job support family, agency collect information , plan amends plan, disability receive education , african make contribution, judiciary conduct promulgate regulation, someone have diabetes, child need service, provider offer service, employee perform function, father write constitution, money hold bag (trash), person enter contract, medicare manage plan, people promote welfare, person graduate school, university do study, school receive funding, bible fulfil self-employed, student afford education , individual commit crime, medicare provide prescription drug, corporation take opportunity, program have (positive/negative) effect, student pay interest/rate, individual make legislative (election), people save family, family lose healthcare, people get loan, grant provide grant, government receive grant, idea have merit, program do nothing, store collect sale, weapon pierce tank,

person receive amount, american lose insurance, company ship job, worker have exhaust, people buy consumer, fund provide funding, someone get healthcare, republican do everything, mother lose mother, mother graduate school, american sell consumer, american work minimum wage, reconciliation carry recommendation, supreme have (legal) statute, prescription drug save family, republican raise tax, school educate child, country head direction, community perform abortion, worker exhaust unemployment (benefits), people kill american, treasury strikingfor purpose, of-fender serve (legal) sentence, railroad provide service, healthcare take healthcare, patient receive patient, people want spend, american work family, doctor see medicare, agency desire grant, american lose concern, republican take money, individual bearing interest/rate, government give money, healthcare work american, republican gain power, government incur money, employee do business, transportation issue certification, business have employee, people have ability, plan cut tax, senior get healthcare, american owe appreciation, people not-pay attention, company participate program, faith base community, service connect condition, program have bipartisan (support), people drive mile , american pay dollar, O.P.E.C. cut production, american have life quality, air force indicate appointment, government have money, people work american, people fight war, people leave country, american support action, law authorize study, attorney (judge) represent client, program reduce disease, administration enforce law, consumer make instruct, american fight war, project provide water, applicant receive grant, grant receive grant, government make payment, program provide community, baby (boomer) reach adult, people try job, american have idea, someone know someone, child have food, democrat have idea, small business need help, people face mortgage, government protect citizen, people sign petition, republican provide (tax) relief, school make decision, program help nation, people have trouble, money paid amount, federally assist housing, servicemembers lose family, budget receive recommendation, republican propose medicare, senior rely medicare, person obtain abortion, disaster affected area, republican have opportunity, transaction lack abuse, environmental protection agency do job, regulation have effect, budget pay debt, F.D.A. approve plant, people use money, veteran suffer illness, senior not-afford prescription drug, troop make sacrifice, trade create job, people not-have money, constitution vested power, administration make bad (problem), police lose family, patient get child, american rely program, oil reach gasoline, nation make progress, good start serve child, people use drug, american have confusion, men women defend country, people work action, american deserve government, people have courage, people bear staff, program help community, people establish justice, republican make promise, healthcare make healthcare, budget do job, american make life quality, american buy car, program benefit child, injustice motivate crime, people make comment, veteran need service, citizen do extraordinary, corporation make charity, illegal immigrant commit crime, small business afford healthcare, judiciary do job, plan have merit, social security pay money, person receive money, F.D.A. approve medicine, people provide family, student pay loan, people have

freedom, national submit intelligence, democrat have power, program provide housing, business pay tax, violent crime serve (legal) sentence, healthcare save money, people pay mortgage, law confer power, people not-want job, medicare pay doctor, people have resource, resource do job, national make recommendation, employee provide notice, program help farmer, people commit suicide, student bring gun, business employ american, measure provide authority, people make judgment, bush threaten veto, child receive food, american want soluton, medicine save family, doctor save family, oil create job, unemployment (benefits) find job, america pay tax, child brought gift, child have succeed, american exhaust unemployment (benefits), someone have concern, girl have baby (boomer), government pay debt, worker change job, tax reduce deficit, water provide water, school reduce class size, family attend school, person perform service, program serve people, company provide healthcare, constitution prohibit america, county provide service, people take loan, people establish constitution, people watch network, employee perform service, good start provide service, government work american, airline lose job, employee provide education , pesticide cause disease, child watch network, someone drive car, victim receive compensation, people have nothing, people have commitment, disability limit major issue, iraqi take power, iraq pose jeopardy, someone want tax, defendant award punitive damage, employee maintain plan, america have opportunity, parent have choise, appropriation provide fund, judiciary take action, budget take healthcare, eligible participate program, school receive fund, bank do business, worker lose healthcare, refinery make profit, american take grant, money amount taxable, program achieve goal, nation face fight, budget make bad (problem), director issue information , (sick/paid/unpaid) leave be virgin , ethanol blend gasoline, senior receive healthcare, kid not-have healthcare, farmer lose crop, law provide protection, business provide job, republican offer budget, attorney (judge) promulgate regulation , nation love peace, men women write constitution, social security strike period, worker look job, someone serve country, diabetes affect american, program help the needy, plan reduce deficit, american send child, congressional budget define money, export create job, people attend school, china manipulate dollar, program do job, farmer have crop, irs target group, production use american, american provide service, budget set priority, folk do job, citizen have healthcare, nation need energy, administration mislead american, homeland attack unit, business estimate tax, people get tax, lender have lender, child have illness, someone get everything, opinion show american, government not-do job, small business employ workforce, child have cancer, student pursue degree (education), witness limit medicine, small business provide job, crime use weapon, group provide service, disaster cause damage, program carry program, american buy insurance, employee sponsor saving, employee offer coverage, toxic (thing) cause cancer, employee provide employee, project provide service, people face problem, application contain information , men women do nothing, nation make commitment, someone want buy, people stop smoking, american do everything, government make commitment, guard save family, company bring job,

administration have plan, people broken law, agency conduct program, self-employed employ individual, director take action, plan save money, plant lose job, program receive grant, employee put money, people loan money, american drive car, car cross border, american produce energy, small business receive loan, katrina hit ocean, burmese contain export, someone have child, federally fund healthcare, republican refuse action, lender participate program, people have disability, budget increase debt, people lose concern, national provide service, government make healthcare, employee have employee, police put family, person make healthcare, student get loan, american earn minimum wage, administration negotiate trade, soldier make sacrifice, american scratch staff, agency give priority, brave men women serve nation, nothing affect authority, program receives funding, people get service, brave men women make sacrifice, men women lose job, child receive funding, family afford healthcare, people send kid, community base community, someone do everything, bush nominate circuit (courts), student earn degree (education), someone buy gun, unemployment (benefits) exhaust money, uniform make sacrifice, student teach child, person get healthcare, uniform serve country,

L List of entities labeled as procedural or noise terms

We present here the entities that we consider to be procedural or not interpretable after dimension reduction. We name them by the most frequent term in their cluster. These entities are excluded in our main analysis. The list below shows all such excluded entities.

absolutely right, accept, accomplish, accurate, adam, address, adjustment, administrative office, advisory, afford, agreeable, akin, alan, allegiance, amaze, anderson, andrew, angel, approach, argument, aristide, arlen, available, ayes, baca, ball, barbara, bartlett, basic, be, beautiful, behind, ben, benjamin, berkley, berman, biggert, billy, bipartisanship, blm, bob, bobby, body, bolton, bonner, bonnie, boswell, bottom, brian, briefing, brook, brownfields, buck, buckley, building, bumper, bureau, burn, burton, byron, cabinet, cafta, candidate, cardoza, carl, carolyn, certain, chamber, chandler, chapter, charles, charlie, cheap, cheney, chenoweth, chris, christopher, chuck, clark, classify annex, clear, close, cloture, coat, code, colin powell, colloquy, columbus, commenters, complicate, compromise, comptroller, conferee, conference, confirm, congressional caucus, congressional gold medal, congressman, consent, conservation, consideration, continue, corp, corrine, costello, counsel, couple, cranston, damato, dan, daniel, danny, date, dave, david, day, de la garza, debate, debbie, demint, dennis, department, designate, desjarlais, determine, dialogue, dianne, diaz balart, dice, dick, dick cheney, different, dirksen office building, disabled, disagree, disappoint, discuss, distinguish, division, do, doable, document, don, double, downstream, draconian, draft, dreier, due, ed, eddie bernice, edward, elizabeth, emerge, enforce, engage, enter, entity, eric, erisa, essence, establish, estimate, estrada, exactly, exceed, excite, executive, exist, extension, extent, extraneous material, extremely important, eye, factor, familiar, fannie, far, far proceeding, federal, filner, final, fine, finish, first, first responder, fit, follow, foot, formal, forward, frank, frank frank, frank gallegly, fred, freddie, fy, ga, gallegly, gary, gate, generalparagraph, george, gibbon, ginny, glosson, go, government accountability office, government printing office, governor, graf, gram, grisham, half, ham, happen, happy, harry, hearing, helm, helpful, henry, herger, herseth, high, hinchey, hj re, holden, hollen, homeownership, hour equally divide, hour workweek, howard, hr, hussein, ii, iii, implement, implication, important, inc, inspire, interior, introduce, ironic, irresponsible, item, jack, james, james madison, jeff, jerry, jesse, jim, jimmy, jo ann, joe, johnny, join, joint, joint chief staff, jon, joseph, journal last day, jr, judicial conference, kansa, keith, kent, kucinich, lahood, larry, late, latourette, ledbetter, letter, linda, list, lng, local coordinate entity, loretta, lot, lungren, ma, madam, madison, mail, main street, majority, manner, manzullo, mark, martha, martin, mary, max, mccotter, mcdonald, mcmorris, md, medically necessary, mel, michael, miguel estrada, mike, milosevic, minute, minute remain, mitch, mn, modify, moment, multiple, myrick, na, naive, nancy, nc, nd, net neutrality, newark, newt, nez, nick, nj, nomination, non, norton, np, nt do, numbered, numerous, objection, observation, offer, office, ok, olver, open, oppose, orrin,

outside, overall, oversight, overwhelm, owen, page, park, parliamentarian, pas, pat, patrick, patten, patty, paulson, percent, perez, pete, peter, petitioner, petraeus, phil, pitt, plate, point, point order, politically correct, polluter, popular, post, powell, preside officer, president, president tempore, procedure, process, projection, proponent, provision, qualify, quayle, question, quick, randy, rank, rapidly, re, read, ready, recommit, reconsider, redesignate, redesignating, reduction, regular order, rehberg, rehnquist, rein, remain available expend, remark, reno, repeat, report, representative, requisite number word, revolve fund, reyes, richard, rick, riggs, robert, rodney, roger, ron, rothman, rule, rule administration, rulemaking, rumsfeld, russell, russell office building, sam, sander, save, schedule, schmidt, sean, secure, select, select intelligence, select reserve, sen, sense, separate, session, sex offender, shay, shuler, side, sj re, skelton, solomon, somalia, sorry, sotomayor, southern, southwick, specie, speech, spouse, st, stack, standard, starr, statement, station, status quo, stearns, stennis, steve, straight, string, structure, stump, subparagraph, subtitle, sufficient second, sullivan, sunset, supplemental, support, susan, sutton, table, ted, territory, th, thanks, thing, thornton, tim, tim myrick, time, tn, todd, tom, tomorrow, tonight, tradition, tsa, tx, unacceptable, underfunded, underlie, unique, unprecedented, upon, va, valid, ve get, version, view, vitally important, wait, walter, wasserman, way mean, wayne, whatever, whichever, whole, whose, william, willing do, withdraw, woolsey, worthwhile, xiv, xx, xxi, xxii, year, year old, zero, zoe,

M List of Stopwords

We remove common words in English, procedural terms specific to the U.S. Congressional Record, numbers, as well as all the names of Congress members and U.S. States. The complete list is below:

i me my myself we us our ours ourselves you your yours yourself yourselves he him his himself she her hers herself it its itself they them their theirs themselves what which who whom this that these those a an the and but if or because as until while of at by for with about against between into through during before after above below to from up down in out on off over under again further then once here there when where why how all any both each few more most other some such no nor not only own same so than too very much actual abraham aderholt adler akaka alexander allard allen amash amodei andrews applegate archer arcuri arney ashcroft ayotte bacchus bachmann bachus baesler baird baker baldacci baldwin ballance ballenger barber barca barcia barkley barletta barlow barr barrasso barrett barrow barton bass bateman baucus bayh bean beatty beauprez becerra begich beilenson bell benishek bennet bennett bentivolio bentley bentsen bera bereuter berry bevill biden bilbray bilirakis bingaman bishop black blackburn blackwell blagojevich bliley blumenauer blumenthal blunt blute bocchieri boehlert boehner bonamici bond bonilla bonior bono boozman boren borski boucher boustany boxer boyd boyda bradley brady braley breaux brewster bridenstine bright brooks broun browder brown brownback brownley bryan bryant buchanan bucshon bumpers bunn bunning burgess burns burr burris bustos butterfield buyer byrd byrne callahan calvert camp campbell canady cannon cantor cantwell cao capito capps capuano cardenas cardin carnahan carney carper carson carter cartwright case casey cassidy castle castor castro cazayoux chabot chafee chaffetz chambliss chapman chiesa childers chocola christensen chrysler chu cicilline clarke clay clayton cleaver cleland clement clinger clinton clyburn coats coble coburn cochran coffman cohen cole coleman collins combest conaway condit connolly conrad conyers cook cooksey cooley coons cooper coppersmith corker cornyn corzine costa cotton courtney coverdell cox coyne craig cramer crane crapo crawford cremeans crenshaw crowley cruz cubin cuellar culberson cummings cunningham dahlkemper daines danforth danner darden daschle davis dayton deal deconcini defazio degette delahunt delaney delauro delay dellums denham dent derrick deutch deutsch dewine dickey dingell dixon djou dodd doggett dole domenici donnelly dooley doolittle dorgan doyle drake driehaus duckworth duffy duncan dunn durbin durenberger edwards ehlers ehrlich ellison ellmers ellsworth emanuel emerson engel english ensign enyart enzi eshoo esty etheridge evans everett ewing exon faircloth fallin farenthold farr fattah fawell fazio feeney feingold feinstein ferguson fields fincher fingerhut fischer fish fitzgerald fitzpatrick flake flanagan fleischmann fleming fletcher flores foley forbes ford fortentberry fossella foster fowler fox foxx frahm frankel franken franks frelinghuysen frisa frist frost fudge funderburk furse gabbard

gallego gallo ganske garamendi garcia gardner garrett gejdenson gekas gephardt geren gerlach gibbons gibbs gibson giffords gilcrest gillibrand gillmor gilman gingrey gingrich glenn glickman gohmert gonzalez goode goodlatte goodling goodwin gordon gorton gosar goss gowdy graham gramm grams grandy granger grassley graves grayson green greene greenwood gregg griffin griffith grijalva grimm grucci gunderson guthrie gutierrez gutknecht hagan hagel hahn hall halverson hamburg hamilton hanabusa hancock hanna hansen hare harkin harman harper harris hart hartzler hastert hastings hatch hatfield hayes hayworth heck hefley heflin hefner heineman heinrich heitkamp heller helms hensarling herrera higgins hill hilleary hilliard himes hinojosa hirono hoagland hobson hochbrueckner hodes hoeffel hoekstra hoeven hoke holding hollings holt honda hooley horn horsford hostettler houghton hoyer hudson huelskamp huffington huffman hughes huizenga hulshof hultgren hunter hurt hutchinson hutchison hutto hyde inglis inhofe inouye inslee isakson israel issa istook jackson janklow jefferson jeffords jeffries jenkins jindal johanns john johnson johnston jones jordan joyce kagen kaine kanjorski kaptur kasich kassebaum kaufman keating keller kelly kempthorne kennedy kennelly kerns kerrey kerry kildee kilmer kilpatrick kilroy kim kind king kingston kinzinger kirk kirkpatrick kleczka klein kline klink klobuchar klug knollenberg kohl kolbe kopetski kosmas kratovil kreidler kuhl kuster kuykendall kyl labrador lafalce lamborn lampson lancaster lance landrieu langevin lankford lantos largent larsen larson latham latta laughlin lautenberg lazio leach Leahy lee lehman levin levy lewis lieberman lightfoot lincoln linder lipinski livingston lloyd lobiondo loeb sack lofgren long longley lott lowenthal lowey lucas luetkemeyer lugar lujan lummis luther lynch machtley mack maffei mahoney majette maloney manchin mann manton marchant marino markey marshall martinez martini mascara massa massie matheson mathews matsui mazzoli mccain mccandless mccarthy mccaskill mccauley mcclintock mccloskey mccollum mcconnell mccrery mcurdy mcdermott mcgovern mchale mchenry mchugh mcinnis mcintosh mcintyre mckean mckinley mckinney mcLeod mcMahon mcMillan mcnerney mcNulty meadows meehan meek meeks melancon menendez meng merkley messer metcalf meyers mfume mica michaud michel mikulski miller mineta minge minnick mitchell moakley molinari mollohan montgomery moore moorhead moran morella moynihan mullin mulvaney murkowski murphy murray murtha musgrave myers nadler napolitano natcher neal nelson nethercutt neugebauer neumann ney nickles noem nolan northup norwood nugent nunes nunn nunnelee nussle nye obama oberstar obey olson ortiz orton osborne ose otter owens oxley packard packwood palazzo pallone pappas parker pascrell pastor paul paulsen paxon payne pearce pease pell pelosi penny perlmutter perriello perry peters peterson petri phelps pickering pickett pickle pingree pittenger pitts pocan poe polis pombo pomeroy pompeo porter portman posey poshard pressler price pryce pryor putnam quigley quillen quinn radanovich radel rahall ramstad rangel ravenel redmond reed regula reichert reid renacci renzi reynolds ribble rice richardson richmond ridge riegle rigell riley risch rivers robb roberts roby rockefeller rogers rodriguez roe roemer rogan rogers rohrbacher

rokita rooney ros rose roskam ross rostenkowski roth rothfus roukema rowland royce rubio ruiz runyan ruppertsberger rush ryan ryun sabo salazar sali salmon sanchez sanders sandlin sanford sangmeister santorum sarbanes sarpalius sasser sawyer saxton scalise scarborough schaefer schaffer schakowsky schatz schauer schenk schiff schneider schock schrader schrock schroeder schultz schumer schwartz schwarz schweikert scott seastrand sekula sensenbrenner serrano sessions sestak sewell shadegg shaheen sharp shaw shays shelby shepherd sherman sherwood shimkus shows shuster simmons simon simpson sinema sires sisisky skaggs skeen slattery slaughter smith snowbarger snowe snyder sodrel solis souder southerland space specter speier spence spratt stabenow stenholtz stevens stewart stivers stockman stokes strickland studds stupak stutzman sundquist sununu swalwell sweeney swett swift synar takano talent tancredo tanner tate tauscher tauzin taylor tejeda terry tester thomas thompson thornberry thune thurman thurmond tiaht tiberi tierney tipton titus tonko toomey torkildsen torres torricelli traficant tsongas tucker turner udall unsoeld upton valadao valentine van vargas veasey vela velazquez vento visclosky vitter voinovich volkmer vucanovich wagner walberg walden walker wallop walorski walsh walz wamp ward warner warren waters watson watt watts waxman webb weber webster weiner welch weldon weller wellstone wenstrup westmoreland wexler weygand wheat white whitehouse whitfield whitten wicker williams wilson wise wittman wofford wolf womack woodall wu wyden wynn yarmuth yoder yoho young zeliff zimmer aye section subsection sec act roll yea yes nai no none nay mr mrs senator gentleman gentlemen gentlewoman gentlewomen gentelady genteladies sir committee speaker chairman ask yield will shall can say let adjourn speak etc thank absent tell per part now month today con pro thereabout thereafter thereagainst thereat therebefore therebeforn thereby therefor therefore therefrom therein thereinafter thereof thereon thereto theretofore thereunder thereunto thereupon therewith therewithal hereabout hereafter hereat hereby herein hereinafter hereinbefore hereinto hereof hereon hereto heretofore hereunto hereunder hereupon herewith whereabouts whereas whereafter whereat whereby wherefore wherefrom wherein whereinto whereof whereon whereto whereunder whereupon wherever wherewith wherewithal bill house member senate amendment section subsection committee subcommittee line article vote insert clerk speaker district state adjourn monday tuesday wednesday thursday friday saturday sunday january february march april may june july august september october november december adjourn board general legislature legislation parliament amend reject appeal appear answer appoint ask author calendar chair chairman colleague commission congress court file motion floor leader object paragraph unanimous pending revise quorum permission secretary resolution pass record print current one two three four five six seven eight nine ten eleven twelve thirteen fourteen fifteen sixteen seventeen eighteen nineteen twenty thirty forty fifty sixty seventy eighty ninety hundred hundreds thousand thousands alabama alaska arizona arkansas california colorado connecticut delaware florida georgia hawaii idaho illinois indiana iowa kansas kentucky louisiana maine maryland massachussets michigan minnesota missis-

sippi missouri montana nebraska nevada hampshire new jersey mexico york carolina north south
east west dakota ohio oklahoma oregon pennsylvania island rhode tennessee texas utah vermont
virginia washington wisconsin wyoming united

References

- Arora, Sanjeev, Yingyu Liang, and Tengyu Ma.** 2016. “A Simple but Tough-to-beat Baseline for Sentence Embeddings.”
- Gardner, Matt, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, Nelson F. Liu, Matthew Peters, Michael Schmitz, and Luke S. Zettlemoyer.** 2017. “AllenNLP: A Deep Semantic Natural Language Processing Platform.”
- Honnibal, Matthew, Ines Montani, Sofie Van Landeghem, and Adriane Boyd.** 2020. “spaCy: Industrial-strength Natural Language Processing in Python.”
- Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean.** 2013. “Efficient Estimation of Word Representations in Vector Space.” *arXiv preprint 1301.3781*.
- Ottonello, Pablo, and Wenting Song.** 2022. “Financial Intermediaries and the Macroeconomy: Evidence from a High-Frequency Identification.” National Bureau of Economic Research.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay.** 2011. “Scikit-learn: Machine Learning in Python.” *Journal of Machine Learning Research*, 12: 2825–2830.
- Pennington, Jeffrey, Richard Socher, and Christopher Manning.** 2014. “Glove: Global Vectors for Word Representation.” 1532–1543. Doha, Qatar: Association for Computational Linguistics.
- Řehůřek, Radim, and Petr Sojka.** 2010. “Software Framework for Topic Modelling with Large Corpora.” 45–50. Valletta, Malta: ELRA. <http://is.muni.cz/publication/884893/en>.
- Reimers, Nils, and Iryna Gurevych.** 2019. “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.” *arXiv preprint arXiv:1908.10084*.
- Sipka, Andrea, Aniko Hannak, and Aleksandra Urman.** 2021. “Comparing the Language of QAnon-related content on Parler, Gab, and Twitter.” *arXiv preprint arXiv:2111.11118*.