Supplementary Material for Water Narratives in Local Newspapers within the United States

1 SUPPLEMENTARY TABLES AND FIGURES

Table S1. Water resources-related topics identified in first STM implementation of the full corpus. PROB refers to words with the highest probability within each topic whereas FREX refers to words that were both FRequent and EXclusive for a given topic.

Topic Number	PROB Words	FREX Words
1	storm, flood, people, hurricane, emergency	fema, katrinas, nagin, femas, storm-related
6	virginia, west, county, charleston, richmond	virginians, manchin, gazette-mail, festivall, manchins
14	food, eat, foods, eating, milk	violationscomments, reach-, moderate-risk, inspectionsphdmcorg, re-inspection
23	water, system, plant, gallons, supply	boil-water, komiske, septic, pumping, filtration
32	santa, mexico, los, mexican, rio	pojoaque, fes, cerrillos, pueblos, agua
39	fire, said, firefighters, department, smoke	firefighters, two-alarm, water-dropping, lightning-caused, fireworks-related
44	river, lake, water, creek, dam	headgates, dams, low-head, downstream, silt
48	can, use, water, plastic, clean	dulley, fixitstartribunecom, royalgreen, thesuperhandymancom, wwwstartribunecomfixit
49	weather, ice, winter, snow, cold	heat-related, precipitation, snowfall, temperatures, freezing
50	news, show, television, media, video	moonves, gottschamer, advertisers, colmes, arkansasonline
52	colorado, trail, mountain, forest, denver	snowmaking, trailheads, fourteeners, crampons, thru-hikers
56	environmental, waste, health, pollution, state	epa, epas, deq, dioxin, pcbs
63	farm, farmers, drought, agriculture, farms	-till, silage, hundredweight, eqip, -cow
65	fishing, fish, bass, lake, trout	crappie, baits, minnows, largemouth, stripers
68	beach, pool, water, swimming, swim	beachgoers, nyad, belugas, beach-goers, nyads
70	energy, oil, gas, power, fuel	fracking, -megawatt, renewables, coal-bed, kilowatt-hour
75	city, property, development, land, project	rezone, rezoned, non-significant, developers, developer
76	fish, wildlife, hunting, fishing, salmon	antlerless, snakehead, wdfw, overfished, snakeheads
79	years, history, old, one, first	tribe, archaeologists, historians, archaeologist, indians
83	garden, plants, trees, plant, soil	annuals, potting, mulching, aphids, succulents
92	boat, island, boats, bay, harbor	pwcs, skippered, phrf, opsail, oneaustralia
97	state, public, environmental, plan, local	stakeholders, initiative, sustainability, implement, economic
98	state, law, court, federal, officials	injunction, lawsuit, sued, lawsuits, ruling

Table S2. Top twenty most frequently occurring named entities within each topic of interest. Drinking water pertains to topic 19, pollutants to topics 29 and 48, and energy production to topics 49 and 69.

Drinking Water			Pollutants			Energy Production		
Entity	Type	Number of Mentions	Entity	Type	Number of Mentions	Entity	Type	Number of Mentions
california	GPE	3,096	epa	ORG	38,367	u.s.	GPE	18,407
florida	GPE	2,035	u.s.	GPE	11,281	epa	ORG	9,039
u.s.	GPE	1,691	dep	ORG	10,691	west virginia	GPE	7,653
utah	GPE	1,665	west nile	LOC	4,950	dep	ORG	5,525
colorado	GPE	1,661	washington	GPE	4,095	colorado	GPE	5,040
denver	GPE	1,629	superfund	LAW	3,900	the united states	GPE	4,972
georgia	GPE	1,464	dupont	ORG	3,486	oklahoma	GPE	4,939
austin	GPE	1,342	utah	GPE	3,247	washington	GPE	4,891
santa fe	GPE	1,300	the united states	GPE	3,187	california	GPE	4,874
hillsborough	ORG	1,284	the u.s. environmental protection agency	ORG	3,146	congress	ORG	4,824
tampa bay	LOC	1,225	west virginia	GPE	3,145	texas	GPE	4,610
tampa	GPE	982	virginia	GPE	2,866	florida	GPE	4,271
texas	GPE	965	california	GPE	2,834	utah	GPE	4,114
oklahoma	GPE	679	congress	ORG	2,631	pennsylvania	GPE	4,011
american	NORP	664	colorado	GPE	2,554	bush	PERS	3,734
lcra	ORG	658	florida	GPE	2,512	american	NORP	3,733
nevada	GPE	657	maryland	GPE	2,374	virginia	GPE	3,511
oklahoma city	GPE	656	american	NORP	2,333	america	GPE	3,214
arizona	GPE	607	cdc	ORG	2,325	senate	ORG	2,863

Note: GPE = countries, cities, or states; LAW = named documents made into laws; LOC = non-GPE locations, mountain ranges, or bodies of water; NORP = nationalities, or religious/political groups; ORG = companies, agencies, institutions, etc.; and PERS = people, including fictional.
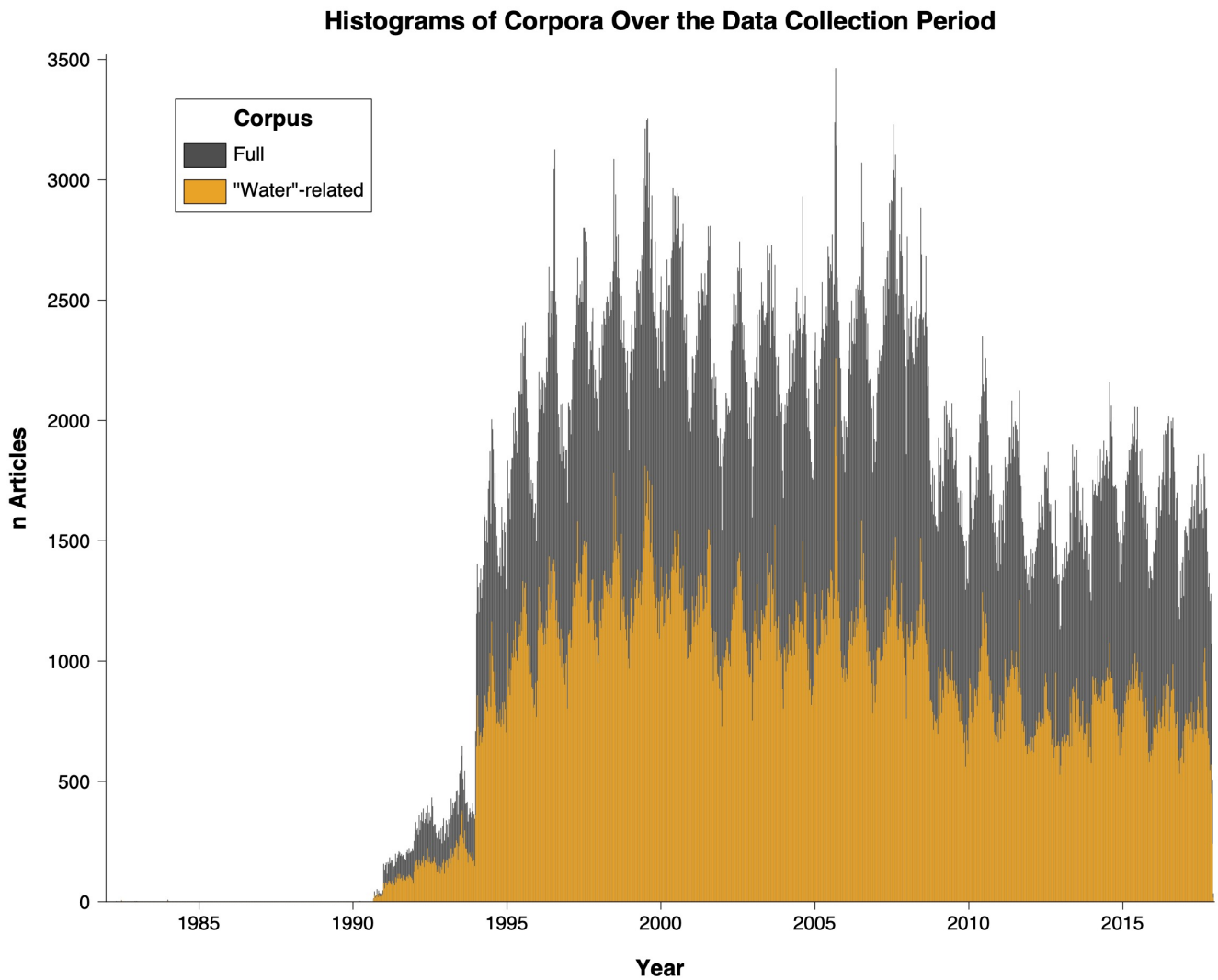
Figure S1. Histogram of publication dates across the data collection period. Articles were grouped into 1,000 evenly-sized bins. Bar color indicates whether the total number of articles published within a bin was summed across the whole corpus (gray) or the "water"-related subset.
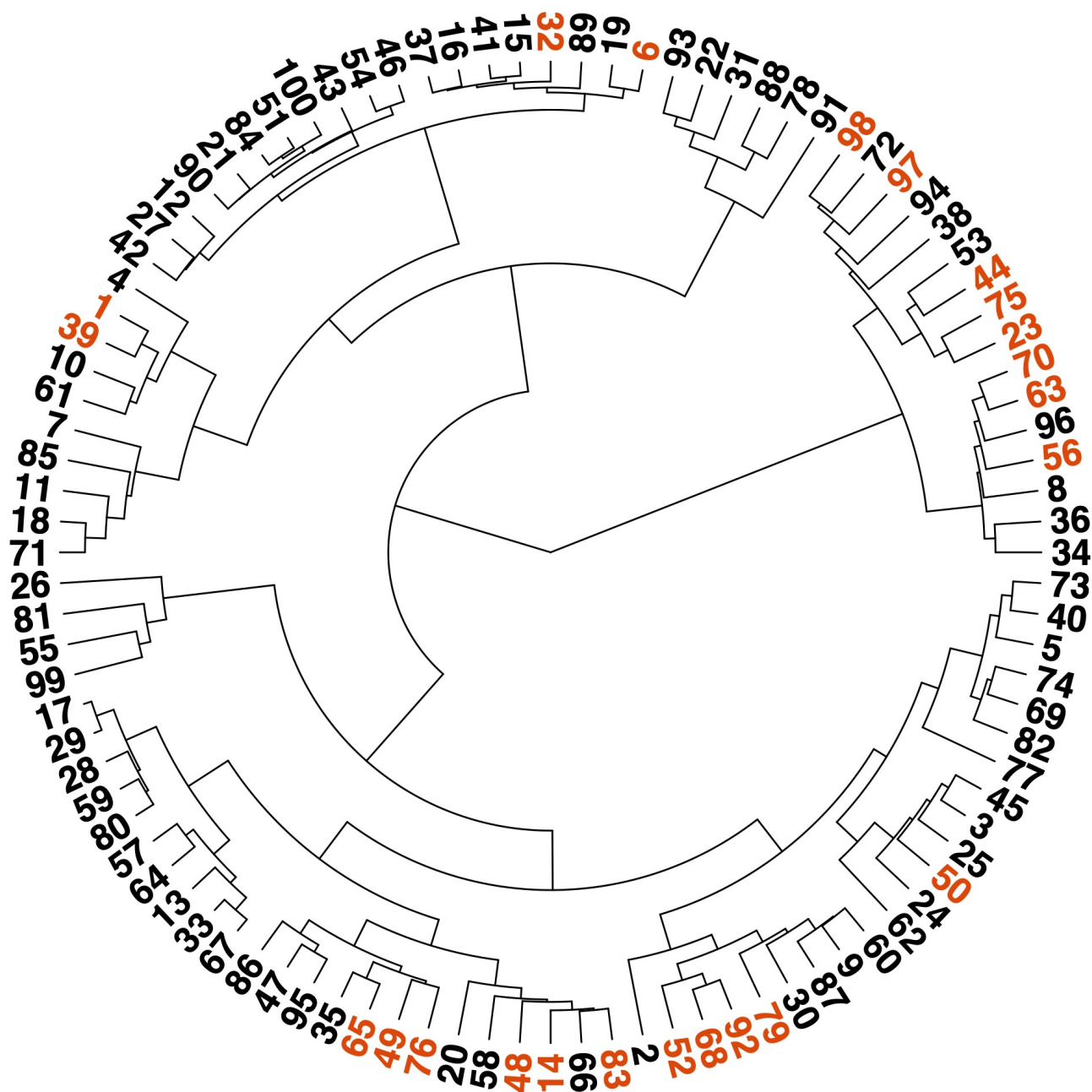
Figure S2. Dendrogram of topic correlations for the full corpus 1.8M articles. Numbers indicate associated topic numbers from the first STM implementation for $k = 100$. Topic numbers in red were identified as those closely related to water resources based on review of PROB and FREX words (see Table S1).
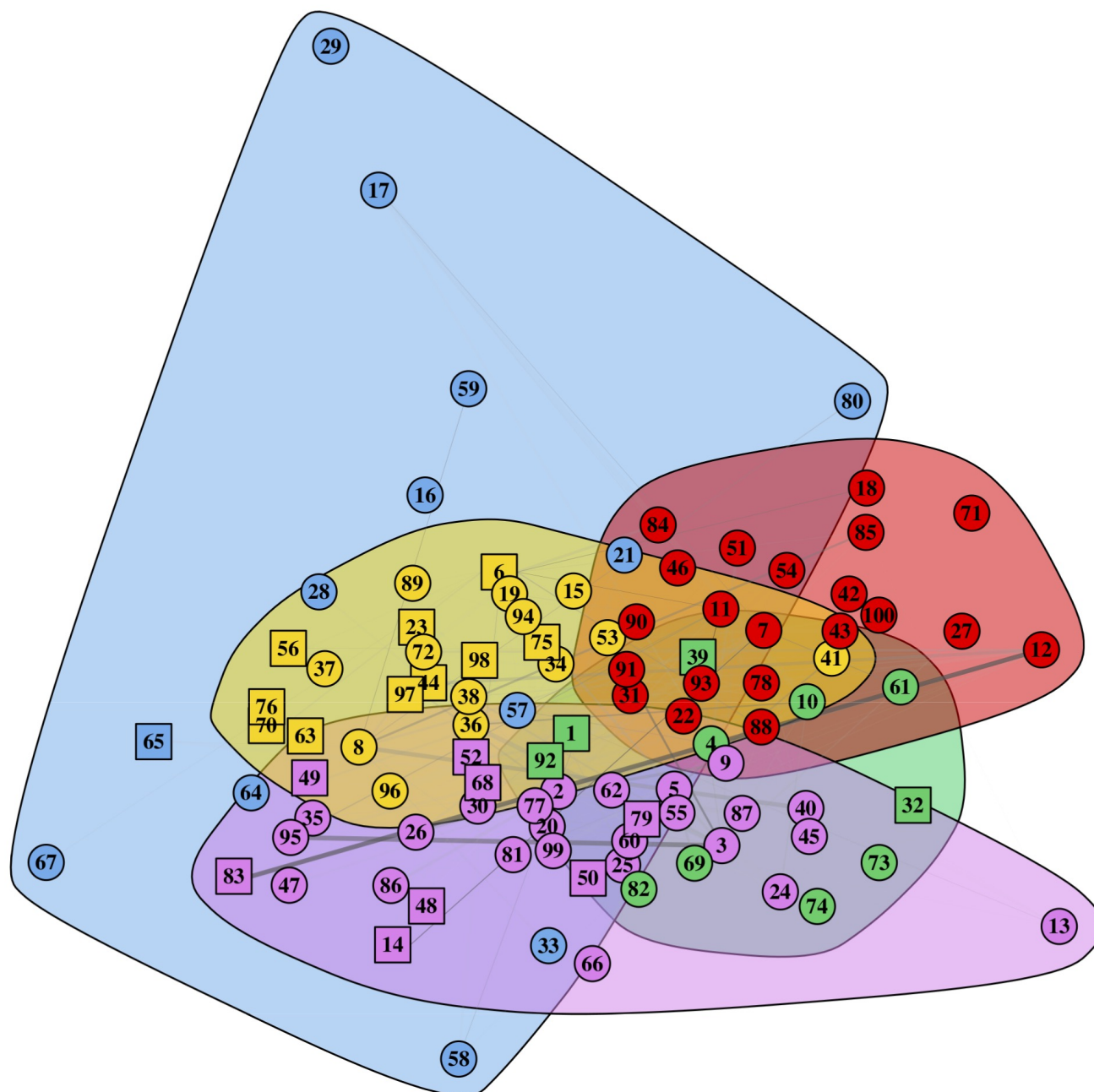
Figure S3. Network graph of topics in the first STM ($k = 100$, $n = 1,819,490$ articles). The thickness of edges corresponds to the strength of correlation between topics across documents. Node colors and shaded areas correspond to the community assignment of topics using the spinglass algorithm. Square-shaped nodes are those topics which were determined to be related to water resources, while circular nodes were considered extraneous.
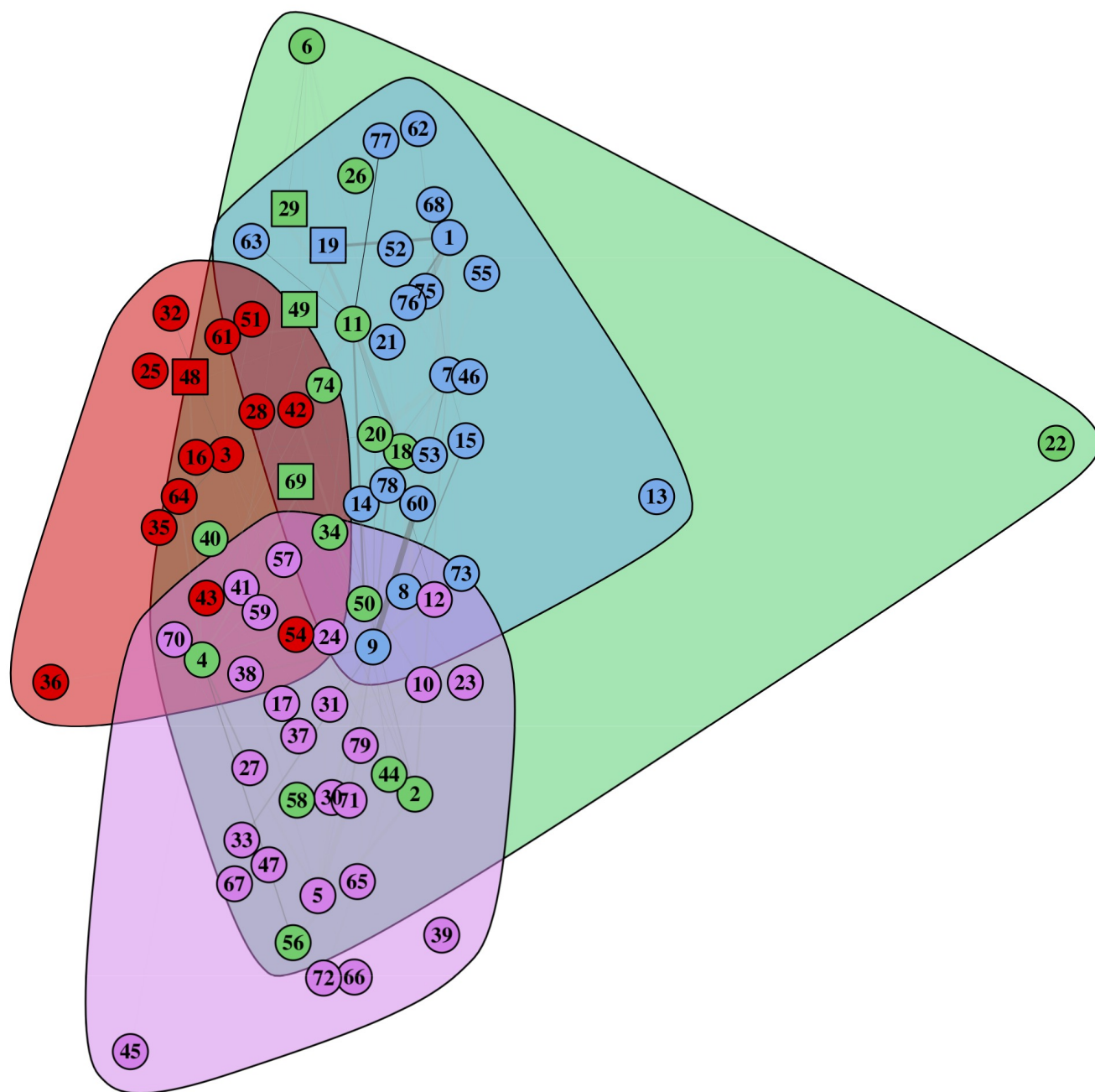
Figure S4. Network graph of topics from the second STM ($k = 79$, $n = 911,132$ articles). Edge thickness and shading (grey lines) corresponds to the strength of correlation between topics across documents. Node colors and shaded areas correspond to the four community assignments of topics using the spinglass algorithm. Square-shaped nodes are subselected topics of interest that were examined in greater detail in this study, while circular nodes reflect topics that can be studied in greater detail in future studies.