The Twitter Network of Think Tank Experts in D.C.

Chen Liang

Master's in Computational Social Science

University of Chicago

ABSTRACT

The influence of a political idea rarely relies only on a few scholars and politicians. Instead, it emerges, evolves, and transforms into feasible policies when policy experts of various backgrounds and reputations cooperate and communicate through organizations and their personal networks. Previous research on the roles of policy experts in the policy making process, however, focus primarily on celebrities whose personal connections with other political actors are more publicly visible and politically influential. Two narratives are thus missing: First, what are the roles of the majority of lesser-known experts? Second, how much are experts' cooperation and communication shaped by organizations such as think tanks? To provide a more complete picture, I construct a Twitter follow graph which covers 609 experts with valid Twitter accounts, which is about 70% of all listed experts on the websites of the Brookings Institution, the American Enterprise Institute, the Center for American Progress, and the Heritage Foundation. Then, I analyze the structure of this Twitter follow graph and explore the implication and limits of Twitter-based connections. I argue that, because the Twitter connections among experts are highly reciprocal and organizational based, the decision to follow someone appears to be more of a meaningful consideration than a whimsical click. Third, with a DeepWalk Algorithm, I show that experts' online and informal social media connections are strongly bound to their offline and formal organizational affiliations. That is, social media usage seems ineffective in exposing experts to opinions from those beyond the reach defined by their offline constraints. In all, I construct a framework to study the expert network through Twitter data and provide a complementary narrative to the existing qualitative discussions on how opinions and information may circulate among policy expert networks.

KEYWORDS

Think Tanks, Political Polarization, Interest Groups, Social Network Analysis, Amazon Web Services, Natural Language Processing, Topic Modeling, DeepWalk

INTRODUCTION

Think tanks have been studied by political and historical scholars as research institutes, lobby organizations, and even business entities which produce and market research products for political capital. However defined, they are expected to compete to provide policy advice as detailed and feasible as possible so that elected officials do not have to be professionally trained to comprehend and interpret sophisticated economic and social situations.

This process is essential to a well-functioning democratic system. Like Max Weber writes in his famous essay, *Politics as a Vocation*, amateur political leaders could better break the bureaucratic inertia and prioritize the greater good. However, his assumptions regarding the sense of responsibility and honor of the political leaders and the moral discipline of the professional policy experts to "execute conscientiously the order of the superior authorities" even when it means denying their own beliefs, are not always easily confirmed in reality (Weber, 1965). When Weber called for these ethical values, he already deeply recognized how they could be undermined by the demagogues emerging from party politics. His thoughts, of course, have been well-researched by scholars in the past century and cannot possibly be exhaustively covered within this paper. Yet, this expectation of policy experts to be value-neutral to the extent of self-denial and the disappointment when they fail to do so have been common scenes in today's public discussion.

In the early 20th Century, think tank experts generally insisted on this principle of value neutrality despite unavoidable ideological subtleties and gained credibility through rigorous and unbiased research. The expansion and complication of government responsibility following the World Wars and the expansion of social security programs brought forth demand for professional social research expertise. Social scientists, mostly university professors, were interested particularly in advancing policy efficiency through quantitative measures such as surveys and established significant credibility among policy makers in various levels of government. While some joined administrations as policy advisors, many of these experts, concerned with the potential political distortion in government bureaucracy, assembled in privately endowed foundations such as the Russell Sage Foundation and the Institute for Government Research—later renamed the Brookings Institution (Smith, 1993).

The scholarly principles face two major challenges. First, while committed to ordinary citizen education in their mission statements, these first generation think tanks often found it difficult to translate the complexities of social phenomena to the busy policymakers without risking the methodological rigor of their research. Kent Weaver mentioned a running joke among Brookings scholars that their “books [we]re written for policy-makers and read by college students.” (Rich, 2010) Facing this tradeoff, the scholars believed that credibility relies on an image of political detachment and that the “policy community” and the “political people” (Brodkin, 1985) should be mutually exclusive. The former equips the latter with theoretical frameworks and historical references, but it should be the elected officials instead of certain technocrat elites who take the ultimate responsibility in making political decisions in a democratic society (Dewey, 1927). “By the method of converting general principles into specific indices of action, the policy sciences provide criteria by which to test the applicability of general principles in specific situations,” writes Charles Rothwell, a drafter of the United Nations Charter (Lerner, 2002). But if so, how could politicians choose among thousands of active and well-trained policy experts and trust these experts’ ability to solve the problem without suspecting their overconfidence (Mckenzie, 2008) or being confused by their often-conflicting predictions (Zickfeld, 2010)?

Second, the movement of fixing “social illness” through scientific research triggered resentment from conservatives who found their values and power ramshackle in radical social reforms. The conservative intellectuals around the 1950s started to challenge not only the progressivists but the old-school, “ivory-towerish”, economically focused conservatives, following the motto “Ideas have consequences” from Richard Weaver. Weaver urges reforms to restore traditional values on piety, private property, and social hierarchy while claiming that sensational and egalitarian ideas have led to the decline of the Western civilization. Inspired by Weaver and frustrated by the “continued ineffectiveness of conservatives” in Washington, Edwin Feulner and his friends established a think tank, later named the Heritage Foundation, which actively engages in advocacy and explicitly demonstrates their ideological intentions. He translated the traditionalist philosophical concepts into a practical instruction for conservative experts: First, “mobilizing data in support of ideas changes minds and helps set the public policy agenda.” Second, “marketing ideas is as important as producing the research.” (Edwards, 2013)

Both challenges lead to a growing gap between credibility and influence. As Andrew Rich’s 1997 survey shows, 90% of the Republican and 65% of the Democratic Congressional staff considered Heritage as most influential, while only 39% of the Republican and 58% of the Democratic Congressional staff named Brookings as most influential. On the contrary, Brookings was perceived as the most credible organization, well above Heritage which scored as only the ninth most credible think tank. Admittedly, surveys on Congressional staff working for a Republican-controlled Congress could have strong partisan biases and the political environment has also changed significantly since then. In the next section, I use the personal descriptions of experts on think tank websites to show that think tanks have various percentages of doctoral and professional degrees. As I will elaborate there, think tanks do have different preference rankings based on academic training, media appearance, and political experience.

While the dilemma of credibility and influence seems synchronous with an ideological spectrum, attributing it purely to a conflict between fact-driven progressivists and value-driven conservatives could be dangerous. Instead, both sides successively transform and advance their value-based policy agenda in response to the challenges from each other and to the dramatic societal changes from the Great Depression and the two World Wars, to the civil movements especially following Vietnam and Iraq wars. Today, liberal think tanks rarely disguise their values such as egalitarianism and social justice. Economic conservatives, most notably the Austrian School economists, also have an academically respected tradition of fact-based research with convincing and rigorous models.

In fact, the dilemma might show more of the absence of an accountability system among experts. Policy researchers often deal with sensitive governmental data, work on specific cases that cannot be translated into repeatable conclusions and focus more on drafting confidential memos than publishing academic articles. Thus, without a feasible fact-checking system to verify credibility, policy advising may rely more on personal connections, colleague reputation, prior projects, and media coverage. Given that think tanks as professional organizations help experts build connections, fund research projects, and advertise results, to which extent do they determine who can be heard by policymakers? What does it mean to the policy-making process?

Rarely have the political roles of policy experts been studied within the context of think tanks. Previous research on policy experts mostly comes from individualistic narratives and disproportionately favors famous people and historical events. To this extent, Walter Heller tutoring John F. Kennedy in Keynesian economics or William Kristol and Carnes Lord tutoring Vice President J. Danforth Quayle in works of history are not so different from how Machiavelli advised the prince or how Hobbes educated the young Charles II. Otherwise, think tanks are modeled like business entities, which maximize their political influence and public appearance and market their research products within the constraints of political and financial resources. This type of organizational approach is represented by the annual think tank index reported by the University of Pennsylvania (McGann, 2020). While these indices tend to be straightforward and often backed by established economic theories, the individual names of policy experts are hidden behind think tanks just like employers for companies in business research.

In response to these difficulties, I propose to use the Twitter follow graph as a proxy for the communication networks for policy experts. A Twitter follow graph is defined here as a directed, unweighted graph where a vertex A is directed to vertex B if expert A follows expert B on Twitter. In the following sections, I start by explaining the data collection and preparation process and briefly summarizing the key findings prior to the network analysis. Second, I overview the previous literature which uses social media as the proxy for social interactions and compare this expert graph with an average Twitter follow graph of millions of users. I argue that, while one may suspect that experts either randomly or exhaustively follow others since the following action takes almost no costs, experts do tend to limit their news inbox to selected account sources with clear patterns. Third, I apply a DeepWalk algorithm to interpret the clustering pattern of experts in the network and tentatively argue that organizational factors differentiate experts even with similar ideological stances and specialties. Finally, I show that experts with the most influence over the network tend to engage more in think tank affairs but do not tend to have higher education degrees.

As I elaborate later, it is hard to conclude what a Twitter follow relationship represents, whether these relationships are strong or weak, or whether it represents relationships between friends or enemies. Nevertheless, I argue that the Twitter follow graph has an undeniable advantage by weighting down significant names and focusing more on the incremental interactions among "ordinary" policy researchers that may accumulate to political ideas and policy agenda. This approach can serve as a complementary narrative for the discussion on policy experts by raising the possibility that think tanks may possess implicit yet powerful influence in organizing and mobilizing, if not limiting, the process of idea formation. The call for independent, unbiased, and non-partisan research is therefore not only a matter of personal research ethics but also demands institutional arrangements and deliberation.

DATA COLLECTION

While I would like to include as many and as diverse of think tanks as possible, manually matching experts with their Twitter accounts is time consuming and scraping Twitter is computationally expensive. Thus, I focus only on four major think tanks in D.C., namely **the Brookings Institution (Brookings), the American Enterprise Institute (AEI), the Center for American Progress (CAP), and the Heritage Foundation (Heritage).** All the data used by this paper is obtained by web-scraping think tank websites and Twitter.

1. Think Tank Selection and Overview

First, I limit the scope to D.C. to minimize the influence of physical proximity and regional differences in explaining cross-organizational interactions. Second, among leading D.C. think tanks, I exclude those with particular policy specialties such as the Center for Strategic and International Studies

(CSIS) and the Peterson Institute for International Economics (PIIE), both which emphasize international development, diplomatic strategies, and national security. Third, I hope to reach a balance between liberals and conservatives and between research- and advocacy-oriented organizations. For this selection, I appreciate the help from Dr. Andrew Rich, dean of the Colin Powell School for Civic and Global Leadership at The City College of New York, and Dr. James Smith, former vice president of the Rockefeller Archive Center, who both wrote extensively on think tank history.

Here, I argue that Brookings and AEI are research-based organizations while CAP and Heritage are advocacy-based. Also, to project the multi-dimensional policy positions onto a one-dimensional ideological spectrum, I argue for an ideology spectrum of CAP - Brookings - AEI - Heritage, with the leftmost being the most liberal to the rightmost being the most conservative. Both Brookings and AEI describe themselves as non-partisan and neutral. Brookings is widely considered as "promoting a pragmatic center position between those who wished to extend and deepen the Great Society" and "those who saw almost any federal intervention as likely to make things worse." (Peschek, 1987) It is often considered to hold a center-left position, which is also supported by Andrew Rich's 1997 survey on congressional staff and journalists. Similar to Brookings, AEI claims to operate "independently of any political party and has no institutional positions" and to stick to "rigorous, data-driven research and broad-ranging evidence." (AEI, 2019) But its history as a leading opponent to the New Deal and its tradition of advocating for free enterprise and strong national security mark it as right-leaning.

Compared to Brookings and AEI, CAP and Heritage state their missions explicitly with ideological inclinations. CAP describes itself as "dedicated to improving the lives of all Americans, through bold, progressive ideas as well as strong leadership and concerted action" and seeks to "challenge conservative misinformation with the facts." While it is an organization 501(c)(3) of Internal Revenue code with donations tax-deductible and lobbying activity limited, its advocacy branch, the CAP Action Fund, is under section 501(c)(4). A 501(c)(4) organization is allowed to engage more in advocacy and receive unlimited money without having to disclose the donors (CAP, 2013). Heritage, as mentioned in the last section, also actively engages in lobbying for "free enterprise, limited government, individual freedom, traditional American values, and a strong national defense (Heritage, n.d.).” It initiated a multimedia news platform, *The Daily Signal*, to broadcast--in its own words--“conservative commentary and policy analysis (Daily Signal, 2020).” The two think tanks’ action-oriented strategy can be further confirmed by their annual financial report. In 2019, Brookings spent 14% of their expenses on Brookings Press, communications, publications, and fundraising (Brookings, 2019). AEI spent 18% of their expenses on fundraising and communications (AEI, 2020). As a comparison, Heritage spent 30% in media, government relations, and fundraising and an additional 36% in “education” (Heritage, 2020a) --which might refer to its mentorship programs such as the Young Leaders Program and Phillip N. Truluck Center for Leadership Development (Heritage, 2020b). CAP does not report many details in its financial statements and its lobbying expenses spread across its affiliated organizations such as the CAP Action Fund, ThinkProgress, and Generation Progress, so it is hard to conclude with an exact percentage. However, its dedication to political action beyond policy research is clear with these advocacy affiliates.

2. Profile Data from Think Tank Websites

I scraped a full list of each think tank’s experts and each expert’s profile link from the staff pages on think tank websites on September 18th, 2020. The specific links are provided in Table 1. I did not include communication, campaign, and technical support staff and excluded them by departments. For example, CAP has a “Art, Editorial, and Video” department and Heritage lists non-expert employers as “other staff.” (CAP, n.d.; Heritage, n.d.) I then scraped each expert’s profile information, including his or her full name, job title, Twitter account (if any), and personal description through the profile links. The personal descriptions include education level and past working experience. The websites of Brookings and AEI list education and working experience in different sections on the same page and I saved them together to match the format with the rest two. I then searched automatically the name of each expert in the corresponding think tank websites and recorded the URLs for all articles under the expert's name. The scraping scripts, along with all other codes mentioned below, can be found at my GitHub[1].

Table 1: Website Links

Brookings	https://www.brookings.edu/experts/
CAP	https://www.americanprogress.org/staff/
AEI	https://www.aei.org/our-scholars/
Heritage	https://www.heritage.org/about-heritage/staff/experts

3. Twitter Data

Some experts do not report their Twitter accounts on think tank pages, so I matched them by searching their names on Twitter. Only 1) when a Twitter account explicitly mentions the expert's affiliated think tank or refers to the articles or events of this think tank, or 2) when the account's profile photo is exactly the same as the expert's photo on think tank websites or LinkedIn should the account be matched to the expert. LinkedIn is an online professional networking platform where some experts would report more complete personal backgrounds. To ensure exactness of match, even if a Twitter account and an expert share the same full name and similar policy expertise, they are not matched without one of the two above criteria checked.

I implemented a Python-based Twitter scraper with the help of Pywren (Jonas et al, 2017), a package that helps simultaneously run multiple functions on Amazon Web Services (AWS). Given that this paper is still at its exploratory stage, I did not use Twitter API but circumvented it with a technique similar to that used by Twint. The scraper can parallelly extract an exhaustive list of following and follower accounts for any user, even popular ones like Ben Bernanke who has more than 100K followers. It is also possible to obtain friendship connections through API, but with Twitter's API updates in August 2020 (Cairns, 2020) and the general limits of API-based scrapers, I did not use it for now.

[1] Please see https://github.com/bjcliang-uchi/sample_code/tree/main/Expert_Follow_Graph

DATA PREPROCESSING AND EXPLORATION

Here I discuss how I label the experts based on their profile information. I show that think tank strategies and agenda can lead to significantly different personnel and thus distinct office contexts for social interactions. I show that the experts who use and don't use Twitter are not significantly different in job titles and education levels, but the Twitter data is indeed slightly biased for resident and leadership experts and those with law degrees.

1. Expert Demographics

Experts' profile information extracted from think tank websites provides interesting observations which indicate think tanks' operating strategies and show that the organizational heterogeneity among policy experts is salient. I labeled an expert as a resident scholar, a senior scholar, and/or a leader, if his or her job title include the keywords specified in Table 2. I labeled an expert as holding a law, medical, or PhD degree if his or her personal description includes the corresponding keywords. Here, being a resident scholar or not reflects an expert's physical involvement. Nonresident scholars often cooperate with a think tank remotely as university scholars but do not physically have an office in this organization. Temporary scholars such as visiting scholars, research assistants, and interns tend to stay at the think tank for only a few months to 1-2 years, a time which might be too short for them to build solid social connections. Also, I did not differentiate between research degrees (PhD) from applied professional degrees (Doctor of Medicine, for example), because they both indicate extensive schooling in their corresponding specialty. The goal is to differentiate them from those who gained expertise mainly through work experience and personal exploration as politicians, journalists, and commentators. Finally, I manually confirmed the accuracy of these labels by checking, for example, whether an expert labeled as a "doctor" may have undertaken doctoral studies but never finished (N=2), or an expert labeled as "Law" may not actually have a professional law degree (N = 3).

Table 2: Labels and Keywords

Label	Keywords
Resident	'non-resident', 'nonresident', 'visiting', 'research associate', 'intern', 'research assistant', 'part-time'
Senior	'senior', 'distinguished'
Leadership	'president', 'director', 'chair', 'chief'
Doctor	'phd', 'ph.d', 'dr. ', 'doctora', 'ph. d.'
Law	'law school', 'law degree', 'jd', 'j.d', 'school of law', 'juris doctor'

As shown in Table 3, Brookings seems to have a strong preference for established scholars in academia over experts from other career backgrounds with its unusually high percentage of nonresident experts and experts with Ph.D. degrees and senior titles. AEI has a similar taste but a much higher percentage of resident scholars. It perhaps also has a stronger sense of inner-organizational power hierarchy by granting senior positions only to a limited number of experts. It seems that AEI focuses more on building an expert community within the organization itself rather than directly cooperating with academia like Brookings. This difference might correspond to the observation confirmed by multiple surveys that left-wing liberals outnumber conservatives among university professors in social science disciplines (Gross 2013; Langbert, Quain, & Klein, 2016). This conclusion has registered among high-ranking conservatives. Education Secretary Betsy DeVos, for example, accused left-wing faculty members of trying to indoctrinate students in a speech at the Conservative Political Action Conference in 2017 (Jaschik, 2017).

CAP and Heritage, the two think tanks which both stress advocacy and political action, seem to have much larger bureaucratic bodies with more resident scholars and leadership positions and less non-law doctoral degrees. This is because they invest more in non-academic affairs related to government relations, communications, fundraising, and voter mobilization. Moreover, CAP seems to have a specific focus on legislative issues and the least attention on its experts' academic background. This observation may correspond to CAP's history of favoring former presidential advisors who worked for the Obama and Clinton administrations.

Table 3: Think Tank Demographics

	Brookings	AEI	CAP	Heritage
Resident	37%	62%	89%	82%
Senior	62%	8%	41%	31%
Leadership	12%	15%	37%	47%
PhD	67%	57%	11%	25%
Law	14%	12%	21%	14%

2. Twitter Coverage

Excluding communication, campaign, and technical support staff, out of the 865 policy experts, 609 (70%) are matched to valid and non-private Twitter accounts. The matching process is stated in the previous section. The percentage is consistent across all four think tanks except for Heritage, which has a slightly higher percentage of Twitter users at 78%. As a comparison, only about one-in-five U.S. adults use Twitter in 2019 (Perrin & Anderson, 2019). Without enough data, we cannot conclude whether our research subjects tend to use Twitter more because of their deeper political involvement or simply due to their unique social demographics such as higher level of education (more than half of the experts have doctoral or professional degrees) and income (Perrin & Anderson, 2019).

Table 4. Comparisons between Twitter and non-Twitter user coverage [2]

		Resident		**Senior**		**Leadership**		**PhD**		**Law**		**Total**
		Count	% Total	Count	%	Count	%	Count	%	Count	%	
Brookings	Twitter	114	39%	172	60%	36	12%	195	67%	49	17%	289
	Other	41	33%	80	64%	13	10%	84	67%	10	8%	125
AEI	Twitter	58	67%	3	3%	13	15%	47	55%	10	12%	86
	Other	20	50%	3	8%	6	15%	25	63%	5	13%	40
CAP	Twitter	141	89%	67	42%	64	41%	15	9%	39	25%	158
	Other	60	87%	26	38%	19	28%	9	13%	8	12%	69
Heritage	Twitter	59	78%	30	39%	18	24%	24	32%	24	32%	76
	Other	15	68%	10	45%	5	23%	5	23%	2	9%	22
Total	Twitter	372	61%	272	45%	131	22%	281	46%	113	19%	609
	Other	136	53%	119	46%	43	17%	123	48%	25	10%	256

[2] "Total" refers to the total number of experts who use/ do not use Twitter in each think tank. The percentage is calculated by using the "Total" to divide the "Count" at the left.

The demographics for experts with and without Twitter accounts are similar as presented in Table 4. But the Twitter data slightly disproportionally cover more resident and leadership experts as well as those with law degrees, and less senior and doctoral experts. For example, in AEI, 67% of Twitter users are resident scholars, but only 50% of non-Twitter users are resident scholars. I suspect that the Twitter accounts for resident and leadership experts are more likely to be matched to their expert profiles. With more frequent engagement in logistic affairs such as organizing events and conferences, these full-time employees tend to mention more about their think tanks on social media. Also, experts with law degrees tend to be former presidential advisors or legislative staff who might, when in office, need to publicize their opinions and gain electorate support through social media.

TWITTER: A SOCIAL & INFORMATION NETWORK

A wide range of studies have explored the offline implications of social media networks. First, active online connections can create social capital and stimulate feelings of social connectedness (See Verduyn et al., 2017 for an overview). Social media help establish friendships in the digital space (Lenhart, Smith, Anderson, Duggan, & Perrin, 2015), facilitate personal ties of friendship and intimacy through informal interactions, and create a sense of "community (Chambers, 2013)." This sense of community sometimes may evolve into "echo chambers" or "engineered sociality" (Dijck, 2011) if users and algorithms consistently personalize the networks. Second, social media are also effective in disseminating information (Bakshy, Rosenn, & Marlow, 2012). For example, data-mining studies show that the abundant casual interactions--"weak ties" --on Twitter help spread both important real-time health information and potential misunderstandings of antibiotics (Scanfeld, Scanfeld, & Larson, 2010). But either way, these technologically mediated networks clearly influence users' perceptions, social relationships, and in-person behaviors.

Regarding the political implication in particular, most studies focus on social media's function as an information and opinion disseminating mediator. They predominantly analyze tweet contents, hashtags, and retweeting behaviors rather than the network. Twitter, specifically, has been extensively studied as a mechanism which facilitates political polarization and disinformation. Data of various magnitude show that political retweets exhibit a highly segregated partisan structure (Conover et al, 2011) and that politicians with more extreme ideological positions tend to have more followers (Hong & Kim, 2016; Abramowitz & Saunders, 2008). Even for those embedded in ideologically diverse networks, Bail et al (2018) found that simply being exposed to opposing partisan views online may increase polarization. In all, while social media might also enhance the quality of democratic governance by encouraging "greater citizen engagement," their potential to undermine the credibility of the democratic system is widely recognized. (See Tucker, 2018 for an overview)

Two research gaps, however, remain unsolved. First, previous studies gather tweet data from Twitter primarily by keyword searching and thus assume that the specific wordings of a political opinion should be fixed within the research scope. This assumption on symbolic stability does not often hold true. Thus, it is helpful to examine an existing network which has the potential of disseminating and reframing any undefined opinions and ideas. Second, it is not clear that the conclusions of the previous research are robust for political "elites," mostly of whom well educated and experienced enough to comprehend opposite opinions and fully aware of the impact of social media on political polarization. There are indeed a few studies on how politicians intentionally utilize social media to mobilize political actions, disseminate (mis)information (Farhall et al, 2019; Gaxie, 2018; Aelst & Walgrave, 2016) and, as Hetherington (2001) argues, to "clarify public perceptions of the parties' ideological differences." But rarely are the roles of policy experts, those responsible to collect and analyze information for policymakers, studied in a similar context. Does Twitter expose them to more diverse information and connect them to people they cannot physically interact with easily, or does it reinforce the existing institutional structures in which they are embedded?

In this section below, I compare the policy experts' Twitter network with a general Twitter network of 175 million active users and their twenty billion edges in 2012 as presented by Myers et al (2014) from Twitter. I argue that the expert network is not a whimsical network despite the minimal cost of following someone online and is informative of the process of how political opinions are formed in this age of social media. It is important to note, however, that this Twitter graph should be interpreted only in a digital context. It represents offline relationships only to the extent that social media interactions, as shown by previous research, may erode, enhance, redefine experts' perceptions about their real-life situations. Of course, given how much social media platforms like Twitter are politicalized in today's U.S. social context, a clear distinction between online and offline networks might not be possible.

Unfortunately, existing scraping techniques provide no information on when and how a Twitter user follows another user or is followed. In my future research, I plan to maintain a panel dataset that updates quarterly, such that network evolution can be analyzed as well. As for now, the research focuses only on a Twitter follow graph extracted on September 19th-20th, 2020.

A Highly Reciprocal Celebrity Network

To start, the expert network is a "celebrity" network with a broad distribution of popularity. Opposite to Myers' observation of a typical Twitter user, experts follow less people than they have followers at the 25^{th}, 50^{th}, and 75^{th} percentiles. This asymmetric relationship between experts and non-expert users suggests that experts appear to be more like celebrities who, loosely speaking, more often inform people than are informed themselves. The distribution of popularity also spreads widely. About 2% of the experts have less than 10 followers. Influential experts like Julia Gillard, the 27th Prime Minister of Australia, and Dr. Lawrence Summers, the former Vice President of Development Economics and Chief Economist of the World Bank, have about 812 and 158 thousand followers respectively.

In addition to informing others, experts also clearly use Twitter for non-social purposes such as accessing information. Goncalves et al (2011) suggests that the maximum number of stable social relationships an individual can maintain on Twitter at a time is only about 150. But similar to a typical Twitter user reported by Myers, an expert on average follows over 900 users including non-experts. I argue that while following over 900 users sounds unbelievable, it does not mean that experts do not curate their Twitter accounts at all: about 87% of these expert-users tweeted, retweeted, or replied to tweets more than 10 times in 2020. Although more research is needed, I propose two possible explanations. First, Twitter users may not often clean up their past and often inactive following accounts and thus the follow graph should be considered a stock network with both past and current relationships. Second, if most of the followed accounts update only on a weekly-to-monthly basis and if those who update tend to forward similar materials rather than posting new information that needs to be read carefully, then actively following about a thousand accounts is not impossible.

While disseminating and accessing information is clearly one of the reasons why experts use social media, an expert-only subgraph also exhibits characteristics that are consistent with an acquaintance network. A follow relationship is often asymmetric: users often follow those who are more popular than themselves and do not expect to be followed back (Shi, Rui, & Whinston, 2014). To this extent, a reciprocal follow relationship is likely to indicate that the two users mutually know each other and/or believe that each other's tweet content could be valuable either as opposite opinions or as informative sources. This level of mutual recognition often implies repeated personal communications, especially given that the D.C. policy research community is a small social circle. In the expert follow graph, about 58% of the directed edges between experts are mutual, much higher than the 42% average for the general Twitter network. The clustering coefficients are also high for the expert follow graph. The average cluster coefficient is 0.51 for degree = 5, 0.47 for degree = 20, and 0.30 for degree = 100. As a comparison, the average cluster coefficient for the Facebook graph is 0.4 for degree = 5, 0.3 for degree = 20, and 0.14 for degree=100 (Ugander, Karrer, Backstrom, & Marlow, 2011). The general Twitter graph has even lower cluster coefficients. Admittedly, while these characteristics suggest that the network connections may imply social relationships among individual experts, they may also simply represent a highly homogeneous social context. Given that all four think tanks are located in D.C., acquaintance might be attributed more to frequent conference encounters and Friday happy hours for policy enthusiasts who tend to share similar social status and intellectual interests.

In all, based on multiple graph characteristics, I argue that while the Twitter follow graph is not a perfect proxy given individual variations on Twitter usage, it can inform us of the structure of a broadly defined expert network both to access and disseminate information and to establish and maintain personal relationships.

HOW DO THEY CLUSTER: A DEEPWALK APPROACH

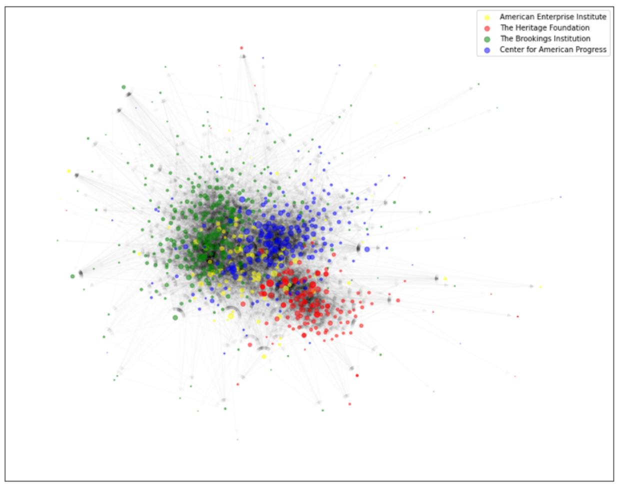
In this section, I discuss how to measure the clustering structure of the Twitter graph with a DeepWalk approach. I show how this social media network--through which experts socialize, access information, share opinions, and perhaps attack opponents on almost a daily basis--could be deeply embedded in the organizational structures of their affiliated think tanks. In another word, regardless of their professional reputation and ideological stances, policy experts seem to be engaged members of their think tanks instead of mercenaries simply funded by these organizations to work for their independent interests and communicate within their own communities.

Community Detection

To start, the organizational influence on expert clustering in the network can be inferred in Figure 1. Nodes are sized by their degree centrality and positioned using a Fruchterman-Reingold (FR) force-directed algorithm (Fruchterman & Reingold, 1991) with 50 iterations. Simulating an electrical field, the FR-algorithm adds repulsive and attractive forces between each pair of nodes and tries to minimize the energy of the system by moving the nodes and adjusting forces in each iteration. However, while the resulting visualized distances do represent nodes' topological similarities, the FR-algorithm works better with undirected graphs and its results tend to be sensitive to the change of parameters such as the number of iterations.

A better measure of node clustering is thus needed. Specifically, a graph partition is a family of pairwise disjoint sets of nodes whose union includes all nodes in the graph. Community quality indices such as modularity and coverage are commonly used to measure the performance of a graph partition, namely, the strength of division of a graph into clusters. Modularity calculates the difference between a given and a random edge distribution (Newman, 2006) and coverage is the percentage of total edges explained by intra-community relationship (Fortunato, 2010). While mostly used in undirected graphs, they are still applicable to directed cases (Nicosia et al., 2009). In our case, the graph modularity with a partition based on organizational affiliation is 0.51, and the coverage of such partition is 0.79. However, these absolute values are meaningless unless analyzed with comparisons such as improving a community detection method within a graph and describing how a graph evolves over time.

Figure 1: FR-Algorithm Visualization of Expert Follow Graph



For decades, algorithms have been developed to optimize community detection problems such as how to partition a graph to maximize inter-community differences and minimize intra-community differences. They seek to maximize community performance by cutting or adding components to the graph, classify nodes based on the vote of other local nodes (weighted-vote relational neighbor (Macskassy & Provost, 2003)), or reduce the dimensions of adjacency or similarity matrixes for vector clustering (spectral clustering, edge clustering, etc.). However, in our case, we hope to describe the performance of a given partition and specify which community clusters most closely instead of looking for the most accurate partition method. That is, we need a node representation technique, which, independent to its classification process, can summarize a node's local information in the graph.

DeepWalk Algorithm

The DeepWalk algorithm (Perozzi 2014) is a community detection method which represents nodes based on truncated random walks and then classifies the corresponding vectors. It treats a sequence of nodes (path) as a sequence of words (sentence). A Skip-Gram unsupervised neural language model is then trained to learn latent representations of nodes and project nodes to a vector space. The Skip-Gram algorithm is originally used for word representation, which predicts context words for a given word based on co-occurrence frequency. The intuition is that words used in similar contexts should have similar semantic meanings and thus closer distance in the vector space. For example, the addition and subtraction between word vectors have semantic meanings. For example, this concept is reapplied to network analysis, representing the relative differences between the local information of nodes with the spatial distance of their corresponding vectors. The specific Skip-Gram algorithm I used is Word2Vec (Mikolov 2013) and implemented in a Python package named Generate Similar (Gensim) (Řehůřek & Sojka 2010).

For the expert Twitter graph, a random walk is a sequence of experts in which each successor is a random choice from all the users the previous expert follows. Hypothetically, each random walk simulates a possible path of Twitter-based information flow among experts. An important sidenote for potential future research is the difference between principal component analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) in reducing the dimensions of vectorized word/node representations. Mathematically, PCA projects from a high dimensional space into a very low dimension space such that the first dimension of the low dimension space captures the largest amount of variance in the original space, the second the largest amount remaining after the first, and so on. It assumes that interesting patterns exist in variance and seeks to find orthogonal project vectors that can maximize the variance. Because the only zero-mean probability distributions that can be fully described by the variance is Gaussian (Shlens, 2014), PCA tends to work less effectively on highly skewed data. In our case specifically, for models with different combinations of parameters, the top five principal components can only explain about 24-30% of the total variances. Therefore, I choose t-SNE instead. It is a nonlinear dimensionality reduction technique which minimizes the Kullback-Leibler Divergence (KL divergence) between the original data distribution and the mapped distribution for visualization, and therefore is less sensitive to data skewness.

Figure 2: DeepWalk Algorithm

Observations

Simply eyeballing the FR-algorithm visualization or looking at individual experts, it seems that, although organizational affiliations matter, experts from different think tanks do follow each other frequently. In fact, on average, in AEI, Brookings, CAP, and Heritage correspondingly, 37%, 19%, 18%, and 20% of the experts that an expert follows, are not their think tank colleagues. But these methods fail to tell whether or how exactly the outward connections matter to an expert's relative position in the graph.

A DeepWalk-based visualization can provide a new perspective to this question. The figures below show the distribution of node vectors, with different combinations of training parameters trained. Thus intuitively, nodes should stay closer if they share more contextual nodes. Outliers—data points lingering relatively at the margin—are mostly users who don't follow many other people or get followed. These reclusive nodes create repeated short circuits in their local walks, similar to the concept of sinks in a PageRank algorithm.

In all, the figures show that experts from the same organization tend to have more similar network contexts. In another word, formal organizational affiliation does seem to be strongly related to experts' informal and online communications. Heritage, in particular, is more and more isolated with the increase in path length even though Heritage experts appear to follow slightly more outside experts than those in Brookings and CAP. This observation supports my previous argument that the decision for experts to follow each other on Twitter is more deliberate than whimsical. But more importantly, it pushes forward the question of what the organizational affiliation to a think tank really means to an expert.

Admittedly, experts in the same think tank may also have relatively similar ideological stances, opinions, working and social backgrounds. They also have more opportunity to get to know each other, not only through social events and conferences but also through daily encounters in the elevator or dining hall. Thus, the fact that experts cluster by organization in their networks, does not necessarily suggest that experts intentionally avoid exposing themselves to more opinions from other organizations. Therefore, we need more research to conclude whether this pattern suggests an echo chamber effect.

Interestingly, while organizational affiliations and ideological stances are closely related as discussed before, even two organizations with similar ideological preferences have segregated networks. While all four think tanks do not limit their focus to particular policy topics, they do have different emphases due to their history, sources of funding, and leading schools of thoughts. For example, Heritage was founded in response to the "failure of the existing conservative think tanks" such as AEI. The historical factions between economic conservatives as represented by AEI and social conservatives as represented by Heritage date back to decades of American political history and might have, under the

Trump Administration, evolved into a rupture between establishment and populists. Therefore, while existing computational research on political polarization focuses mostly on how people and texts cluster based on partisanship or particular policy issues, we need to recognize and incorporate the hidden mediating effects of organizational and institutional arrangements.

DISCUSSION & FUTURE RESEARCH

In conclusion, I constructed a Twitter follow graph for over six hundred experts from four leading think tanks in D.C.: Brookings, AEI, CAP, and Heritage. Through the experts' profile information scraped from their think tank websites, I show that the experts who use and don't use Twitter are not significantly different in job titles and education levels. Think tank strategies and agenda, on the other hand, correspond to distinct personnel and thus social contexts for personal interactions.

The major theme of this primary paper is to discuss the implications of a Twitter follow graph for policy experts. Specifically, I compared the expert Twitter graph with an overall Twitter graph and show that the former is a highly reciprocal celebrity network. While more research is needed, these graph characteristics suggest that experts may use social media both to disseminate and access information within and outside of the policy research community, and to maintain acquaintance relationships particularly within the community.

I also used a DeepWalk algorithm to analyze the graph structure and the clustering patterns of the nodes and visualized the network-based contextual similarities using t-SNE. I show that this approach, compared to traditional methods, focuses more on multi-degree connections and thus better represents the structure of each expert's local Twitter network. The results show that first, normal organizational affiliation does seem to be strongly related to experts' informal and online communications and second, even two organizations with similar ideological preferences have segregated networks. Therefore, I argue that the mediating effects of organizational structures in the procedure of political discussion and ideological polarization should not be ignored.

In all, I hope to construct a theoretical and methodological framework to study the expert network through Twitter data and provide a complementary narrative to the existing qualitative discussions on how opinions and information may circulate among policy expert networks.

In my future research, I hope to focus on further interpreting this network from the following perspectives: First, in my current analysis, I assume Brookings and CAP to be relatively liberal and Heritage and AEI to be relatively conservative. But I would try to measure the ideological identity of each expert instead of assuming an organizational one. I will also measure the content similarity between the articles of each pair of experts. I hope that the ideological and content similarity can help clarify the question of whether experts follow each other because of organizational affiliation, or because experts with similar ideological and topic interests tend to follow each other and work in the same organization. Second, I hope to expand the data scope to more think tanks. With more organizations involved, we would be able to see more cross-organizational patterns and better interpret the influence of organizations on expert networks. As a further plan, I hope to study how a particular idea such as wearing masks during the coronavirus pandemic has been circulating around the policy expert networks and evolved into a political narrative.

REFERENCES

Abramowitz, A. I., & Saunders, K. L. (2008). Is Polarization a Myth? *The Journal of Politics*, *70*(2), 542–555. https://doi.org/10.1017/s0022381608080493

Aelst, P. V., & Walgrave, S. (2016). Information and Arena: The Dual Function of the News Media for Political Elites. *Journal of Communication*, *66*(3), 496–518. https://doi.org/10.1111/jcom.12229

American Enterprise Institute. (2019). *AEI's Organization and Purposes*. https://www.aei.org/about/aeis-organization-and-purposes/.

American Enterprise Institute. (2020, January). *American Enterprise Institute 2019 Annual Report*. https://www.aei.org/wp-content/uploads/2020/01/2019-AEI-Annual-Report-LR-copy.pdf.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., … Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

Bakshy, E., Rosenn, I., & Marlow, C. (2012). The role of social networks in information diffusion. *Proceedings of the 21st International Conference on World Wide Web*, 519–528. https://doi.org/10.1145/2187836.2187907

Brodkin, E., & Kingdon, J. W. (1985). Agendas, Alternatives, and Public Policy. *Political Science Quarterly*, *100*(1), 165. https://doi.org/10.2307/2150882

Brookings Institution. (2019, November). *Brookings 2019 Annual Report*. https://www.brookings.edu/wp-content/uploads/2019/11/2019-annual-report.pdf.

Cairns, I. (2020, July 16). *Introducing a new and improved Twitter API*. https://blog.twitter.com/developer/en_us/topics/tools/2020/introducing_new_twitter_api.html.

Center for American Progress. (2013). *State Notices*. https://www.americanprogress.org/state-notices/.

Center for American Progress (n.d.). *Staff*. https://www.americanprogress.org/staff/.

Chambers, D. (2013). Technologically Mediated Personal Relationships. *Social Media and Personal Relationships*, 21–39. https://doi.org/10.1057/9781137314444_2

Conover, M. D., Francisco, M., Gonc¸alves, B., Flammini, A., & Menczer, F. (2011). Political Polarization on Twitter. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*.

Daily Signal. (2020, June 17). *About The Daily Signal*. https://www.dailysignal.com/daily-signal/.

Dijck, J. V. (2011). Facebook as a Tool for Producing Sociality and Connectivity. *Television & New Media*, *13*(2), 160–176. https://doi.org/10.1177/1527476411415291

Dewey, John. (1927). *The Public and its Problems*. New York: Holt.

Edwards, L. (2013). Chapter 12. In *Leading the way: the story of Ed Feulner and the Heritage Foundation*. Crown Forum.

Erdős, P., & Rényi, A. (1961). On the evolution of random graphs. *Bull. Inst. Internat. Statist*, *38*, 343–347.

Farhall, K., Carson, A., Wright, S., Gibbons, A., & Lukamto, W. (2019). Political Elites' Use of Fake News Discourse Across Communications Platforms. *International Journal of Communication*, *13*, 4353–4375.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3-5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11). https://doi.org/10.1002/spe.4380211102

Gaxie, D. (2018). When political elites select political elites. *The Contested Status of Political Elites*, 108–123. https://doi.org/10.4324/9781315211916-7

Gonçalves, B., Perra, N., & Vespignani, A. (2011). Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0022656

Gross, N. (2013). *Why Are Professors Liberal and Why Do Conservatives Care?* Harvard University Press.

Heritage Foundation. (2020). *2019 Annual Report of The Heritage Foundation*. https://www.heritage.org/annual-report-2019/policy-promotion/.

Heritage Foundation. (2020). *2019 Financial Statements*. https://thf-reports.s3.amazonaws.com/Financial/2019_AnnualReport_Financials.pdf.

Heritage Foundation (n.d.). *About Heritage*. https://www.heritage.org/about-heritage/mission.

Hetherington, M. J. (2001). Resurgent Mass Partisanship: The Role of Elite Polarization. *American Political Science Review*, *95*(3), 619–631. https://doi.org/10.1017/s0003055401003045

Hong, S., & Kim, S. H. (2016). Political polarization on twitter: Implications for the use of social media in digital governments. *Government Information Quarterly*, *33*(4), 777–782. https://doi.org/10.1016/j.giq.2016.04.007

Jaschik, S. (2017, February 24). *Education secretary criticizes professors as telling students 'what to think'*. https://www.insidehighered.com/news/2017/02/24/education-secretary-criticizes-professors-telling-students-what-think.

Jonas, E., Pu, Q., Venkataraman, S., Stoica, I., & Recht, B. (2017). Occupy the Cloud: Distributed computing for the 99%. *Proceedings of the 2017 Symposium on Cloud Computing*, 445–451. https://doi.org/https://doi.org/10.1145/3127479.3128601

Langbert, M., Quain, A. J., & Klein, D. B. (2016). Faculty Voter Registration in Economics, History, Journalism, Law, and Psychology. *Econ Journal Watch*, *13*(3), 422–451.

Lenhart, A., Smith, A., Anderson, M., Duggan, M., & Perrin, A. (2015, August). *Teens, Technology and Friendships*. http://www.pewinternet.org/2015/08/06/teens-technology-and-friendships/.

Lerner, D., Lasswell, H. D., & Fisher, H. H. (2002). *The policy sciences: recent developments in scope and method*. Ann Arbor, Mi.

Macskassy, S. A., & Provost, F. (2003). A Simple Relational Classifier. *Proceedings of the MultiRelational Data Mining Workshop (MRDM) at the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.21236/ada452802

Mckenzie, C. R. M., Liersch, M. J., & Yaniv, I. (2008). Overconfidence in interval estimates: What does expertise buy you? *Organizational Behavior and Human Decision Processes*, *107*(2), 179–191. https://doi.org/10.1016/j.obhdp.2008.02.007

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. arXiv:1310.4546.

Myers, S. A., Sharma, A., Gupta, P., & Lin, J. (2014). Information network or social network? *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion*. https://doi.org/10.1145/2567948.2576939

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Nicosia, V., Mangioni, G., Carchiolo, V., & Malgeri, M. (2009). Extending the definition of modularity to directed graphs with overlapping communities. *Journal of Statistical Mechanics: Theory and Experiment*, *2009*(03). https://doi.org/10.1088/1742-5468/2009/03/p03024

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank Citation Ranking: Bringing Order to the Web. *Stanford InfoLab*. http://ilpubs.stanford.edu:8090/422/.

Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). DeepWalk. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '14*. https://doi.org/10.1145/2623330.2623732

Perrin, A., & Anderson, M. (2019, April 10). *Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018*. https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/.

Peschek, J. G. (1987). *Policy-planning organizations: elite agendas and America's rightward turn*. Temple University Press.

Raven, B. H., & French, J. R. (1958). Legitimate Power, Coercive Power, and Observability in Social Influence. *Sociometry*, 21(2), 83. doi:10.2307/2785895

Rich, A. (2010). *Think tanks, public policy, and the politics of expertise*. Cambridge University Press.

Scanfeld, D., Scanfeld, V., & Larson, E. L. (2010). Dissemination of health information through social networks: Twitter and antibiotics. *American Journal of Infection Control*, *38*(3), 182–188. https://doi.org/10.1016/j.ajic.2009.11.004

Shi, Z., Rui, H., & Whinston, A. B. (2014). Content Sharing in a Social Broadcasting Environment: Evidence from Twitter. *MIS Quarterly*, *38*, 123–142. https://doi.org/10.2139/ssrn.2341243

Shlens, J. (2014, April 3). *A Tutorial on Principal Component Analysis*. arXiv:1404.1100 [cs.LG].

Smith, J. A. (1993). *The Idea brokers think tanks and the rise of the new policy elite*. The Free Press.

Tucker, J., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., … Nyhan, B. (2018). Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. https://doi.org/10.2139/ssrn.3144139

Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The Anatomy of the Facebook Social Graph. *Social and Information Networks (Cs.SI)*. arXiv:1111.4503.

Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do Social Network Sites Enhance or Undermine Subjective Well-Being? A Critical Review. *Social Issues and Policy Review*, *11*(1), 274–302. https://doi.org/10.1111/sipr.12033

Weber, M. (1965). *Politics as a vocation*. Fortress Press.

Zickfeld, K., Morgan, M. G., Frame, D. J., & Keith, D. W. (2010). Expert judgments about transient climate response to alternative future trajectories of radiative forcing. *Proceedings of the National Academy of Sciences*, *107*(28), 12451–12456. https://doi.org/10.1073/pnas.0908906107

Řehůřek, R., & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora. *Proceedings of LREC 2010 Workshop New Challenges for NLP Frameworks*, 46–50.