[David A. Broniatowski and Christopher L. Magee]

Studying Group Behaviors

[A tutorial on text and network analysis methods]

Many important technical and policy decisions are made by small groups, especially by deliberative committees of technical experts. Such committees are charged with fairly combining information from multiple perspectives to reach a decision that one person could not make alone. Committees are social entities and are therefore affected by any number of mechanisms recorded in the social sciences. Our challenge is to determine which of these mechanisms are likely to be encountered in the deliberative process and to evaluate how they might impact upon decision outcomes. In particular, we examine the role of committee deliberations on the U.S. Food and Drug Administration's (FDA's) advisory panels.

Although the research on group decision making is vast, the empirical and quantitative analysis of committees of experts in a real-world setting has been relatively scarce. This may largely be ascribed to difficulty in gathering data (e.g., because it might be proprietary or simply not recorded) and the absence of a corresponding methodology. The advent of the Internet has made text data abundantly available. Furthermore, regulations such as the U.S. Federal Advisory Committee Act (FACA) of 1997 ensure that transcripts of real-world expert committee meetings are available online or upon request (e.g., through a Freedom of Information Act request). Finally, recent innovations in machine learning and computational linguistics have enabled the analysis of large sources of text data in a repeatable and consistent fashion. These methods have yet to be applied to the analysis of social data, and in particular, transcript data on a large scale. There is therefore an opportunity to apply some of these methods to enable a deeper empirical understanding of committee deliberation processes, a major component of decision making by committees of technical experts. Many of these methods use signal processing techniques, such as approaches to dimensionality reduction. Ultimately, these methods may help to generate quantitative insight into committee deliberation processes, perhaps enabling better decision outcomes.

This article provides a short tutorial of approaches in the literature that can be applied to the quantitative analysis of text data. We begin by reviewing selected literatures in the social sciences that are relevant to our inquiry. Next, we review existing attempts to extract social networks from text data. Limitations of these approaches are examined and addressed, culminating in an automated method for extracting social networks from transcripts of expert committee meetings based upon Bayesian clustering. This method also uses cross-correlation techniques to generate directed social networks illustrative of the flow of influence within these meetings. Throughout, we use data from the U.S. FDA's Circulatory Systems Devices Advisory Panel, with the goal of demonstrating the application of the method to real-world data. Findings include insights into the impact of professional specialty

Signal and Information Processing for Social Learning and Networking

©iSTOCKPHOTO.COM/ANDREY PROKHOROV

upon deliberation, identification of different leadership styles on these committees, and possible indications of panels in which voting members may have learned from one another to reach a consensus decision.

SOCIAL SCIENCE APPROACHES TO THE ANALYSIS OF EXPERT GROUPS

Work within sociology motivates approaches to the analysis of text data as a signal of social phenomena. Conversation analysis is perhaps most relevant to our inquiry. This tradition traces its roots to the ethnomethodology of Garfinkel (e.g., [1]), the observations of Goffman (e.g., [2]), and the work of Sacks, Schegloff, and Jefferson (e.g., [3]) and is focused on generating a qualitative understanding of how the unspoken rules of conversation signal social relations. Maynard [4] identifies topics as a key feature of conversations, arguing that changes in topic are nonrandom occurrences that can be related to the structure of the group that is discussing them. Gibson [5] notes that because only one person can generally speak at a time, external status characteristics manifest in conversation as participation-shifts and, often, as topic shifts. These results suggest that an analysis of the speech of committee member participants might provide insight into the dynamics of group decision making, including among committees of experts. Work within the anthropology and science, technology, and society (STS) literatures extends this notion to the realms of professional and institutional cultures. For example, Douglas emphasizes the extent to which group membership influences patterns of thought, and therefore affects perception of data [6]. Furthermore, each specialty possesses its own unique language and jargon, which carries with it an implicit scheme for categorizing perceived phenomena [7]. This is particularly true in technical disciplines, where conceptual precision is required. The STS literature extends this notion by noting that language is used as a cognitive mechanism to delineate professional boundaries, directing the attention of experts within a specialty toward a given interpretation of a problem that is consistent with that expert's training, while simultaneously directing that attention away from other possible interpretations (e.g., [8]). Formalizing these qualitative insights requires the use of signal processing techniques on text data.

QUANTITATIVE APPROACHES TO STUDYING TRANSCRIPTS

This section reviews empirical quantitative methodologies based upon a computational linguistic analysis of text data. For example, Sack uses conversation maps [9] to construct social network representations of very large-scale conversations over email. These visualizations are based on linguistic content as well as email characteristics (e.g., who sends/receives information), and require access to large-scale, annotated data.

NETWORK TEXT ANALYSIS

Perhaps most relevant is network text analysis (NTA) [10], a set of methodologies that attempts to extract social or cognitive structures from transcripts based upon conceptual linkages. One prominent example in this tradition is cognitive mapping [11]—a noncomputational approach for identifying key concepts and generating a network of causal relations from meeting transcripts. In this approach, an analyst identifies and handcodes major concepts and their interrelations. AutoMap (e.g., [12]) is a second, automated approach. AutoMap provides a general computational suite that performs word-frequency analysis, generates concept networks, and classifies these concepts into a preestablished ontology (e.g., entities and interconnections). Sentiment analyses are also performed. A third major approach, centering resonance analysis [13], is designed to identify and link important word phrases within a discourse. The resulting network is then analyzed to determine which of these word-phrases are most important. Each of these methods has been used to analyze group dynamics, where the dominant paradigm is the representation of relations among concepts (i.e., words) embedded within a text. The networks generated aggregate information from across multiple individuals and focus on relationships between concepts, as represented by word phrases. Consequently, these networks have not been used to analyze dyadic relationships between group members. In what follows, we describe a series of methods designed to measure the theoretical constructs outlined in the social science literature reviewed above. This series of methods builds toward an automated approach that does not require aggregation across individuals.

LATENT SEMANTIC ANALYSIS

Approaches to text analysis used by social scientists typically utilize "latent coding" schemes of varying complexity [14]. These tend to be labor intensive and resist unsupervised automation, with the exception of frequency counts of preidentified key words—an approach that does not capture the complexity of real-world group interactions. As suggested by the sociolinguistics literature, analyzing relations between systems of concepts (i.e., topics) is crucial. We therefore use latent semantic analysis (LSA) as a baseline for comparison against the methods reviewed in this article. LSA is a state-of-the-art natural language processing tool, which was developed for purposes of information retrieval and topic grouping [15]. In presenting LSA, we begin by considering a corpus of documents, $\mathcal{D}$, containing n documents $d_1 \ldots d_n$. Consider, as well, the union of all words over all documents, W. Suppose there are $m > n$ words, $w_1 \ldots w_m$. We construct an $m \times n$ "word-document matrix," $\mathbf{X}$, where each element in the matrix, x_{jk}, consists of a frequency count of the number of times each word (j) appears in each document (k). LSA performs a singular value decomposition on $\mathbf{X} = \mathbf{W} \, \mathbf{S} \, \mathbf{D}^{\mathbf{T}}$, where $\mathbf{W}$ is an $m \times l$

matrix of singular unit vectors of length l, each corresponding to a word. Similarly, $\mathbf{D}$ is an $n \times l$ matrix of mutually orthogonal singular unit vectors each corresponding to a document.

Finally, $\mathbf{S}$ is an $l \times l$ diagonal matrix of decreasing, nonnegative singular values, with each element corresponding to a linear combination of weights associated with each singular vector. LSA truncates $\mathbf{S}$ to generate $\mathbf{S}'$, a reduced number of putative latent "semantic dimensions" within D. The resulting matrix, $\mathbf{X}' = \mathbf{W}\,\mathbf{S}'\,\mathbf{D}^{\mathbf{T}}$, may be thought of as a Euclidean space, such that the normalized inner product of two word-vectors (represented as rows of the matrix $\mathbf{W}\,\mathbf{S}'$) can be thought of as the projection of each word upon a set of axes, each of which corresponds to a latent concept. Two speakers who share similar concepts might therefore be related. Dong [16] used this approach to study group decision making in design team contexts. Broniatowski et al. then applied Dong's method in a pilot study, with the goal of exploring the applicability of LSA to the transcript data from the FDA's Circulatory Systems Devices Panel meetings [17]. Five limitations of LSA were identified and must be addressed:

1) LSA assumes the existence of a set of latent, mutually orthogonal, dimensions underlying a Euclidean semantic space. Furthermore, each word's location in the Euclidean space is linearly distributed, an assumption that introduces increasingly more distortion into the analysis as a given speaker uses fewer words. These limitations make it difficult to resolve the linguistic attributes of individual speakers, particularly in the absence of extensive speaker data within a given meeting [18].

2) LSA was unable to account for polysemy—the existence of words with multiple meanings (e.g., a baseball bat versus the flying mammal called a bat).

3) LSA was unable to distinguish procedural language from jargon in FDA panel meetings.

4) LSA defines relationships between speakers in terms of dimensions in a Euclidean feature space. These dimensions nominally correspond to latent concepts within a discourse, but often lack an intuitive interpretation. As a result, relationships between speakers are correspondingly difficult to interpret.

5) LSA was unable to account for additional dimensions, especially the passage of time, within a meeting.

LSA is widely used, and later variants of the technique have addressed some of the issues raised above. For example, tensor representations might be used to account for the passage of time [19]–[21]. Nevertheless, they are still sensitive to the Euclidean space assumptions underlying all LSA-based approaches and therefore introduce significant noise into the analysis. In general, none of these approaches have addressed all of the issues outlined above. We therefore use these five limitations to guide the remainder of our review.

BAYESIAN TOPIC MODELS
The leading alternative to LSA is latent Dirichlet allocation (LDA), a Bayesian "topic model" [22]. Griffiths et al. provide an excellent comparison of LSA to Bayesian models of text analysis [23]. Consistent with the literature on sociolingusitics, topic models aim to extract a set of general topics from specific texts. Approaches based on Bayesian inference, such as LDA, provide a platform that may be used to avoid many of the limitations noted above. Of particular interest are topic-modeling approaches to studying social phenomena in various contexts, including studies of the evolution of specialized corpora [24], analysis of structure in scientific journals [25], finding author trends over time in scientific journals [26], topic and role discovery in e-mail and other text networks [27], analysis of agenda-setting in the U.S. Congress [28], and group discovery in sociometric data [29].

TOPIC MODELS ADDRESS THE LIMITATIONS OF LSA
Unlike LSA, which assumes a continuous Euclidean metric "semantic space" representation of a corpus, LDA assumes probabilistic assignment of each word to a discrete topic. Rather than modeling a word as an average between two locations in a latent Euclidean space, a word is instead modeled as having been drawn from a discrete probability distribution over topics. This provides a natural solution to the polysemy problem outlined above [23]. The basic structure of an LDA model is shown in Figure 1.

Like LSA, LDA stores corpus data in a word-document matrix, $\mathbf{X}$. Each word (w) is assumed to be drawn from a topic (z). A topic is accordingly defined as a multinomial distribution (ϕ) over words (i.e., a word is chosen at random by rolling a weighted w-sided die, where w is the total number of words in the corpus).
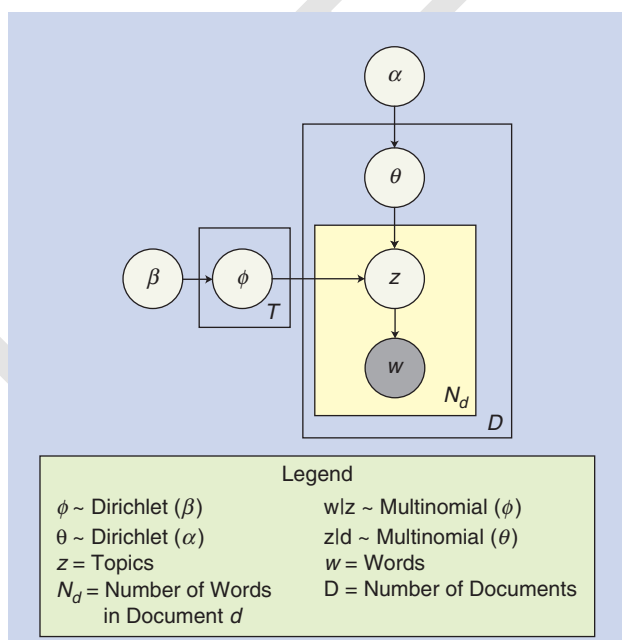
[FIG1] A plate-notation representation of the LDA algorithm [22].

Legend

$\phi \sim$ Dirichlet (β)	w|z $\sim$ Multinomial (ϕ)
$\theta \sim$ Dirichlet (α)	z|d $\sim$ Multinomial (θ)
z = Topics	w = Words
N_d = Number of Words in Document d	D = Number of Documents

Each document is similarly modeled as a multinomial distribution (θ) over topics. The parameters (i.e., the die-weights) for each multinomial distribution are drawn from a symmetric Dirichlet prior distribution—a multivariate distribution that is the conjugate prior to the multinomial distribution. Each Dirichlet distribution has a number of "hyperparameters" equal to the number of parameters of its corresponding multinomial distribution. Nevertheless, early LDA models all assume that the Dirichlet priors are symmetric, i.e., all of the hyperparameter values are the same. The effectiveness of nonsymmetric Dirichlet priors has only recently been tested [30].

Gibbs sampling, a Markov-chain Monte Carlo (MCMC) method adopted from statistical physics, although potentially slower than other approaches, is currently in widespread use as an inference tool among topic modelers [25]. Evaluating Markov chain convergence is an open area of research. It is therefore standard practice for a Markov chain to be run for multiple iterations to ensure convergence. These initial iterations are known as a "burn-in" period. Throughout this article, burn-in length is set to 1,000 iterations—a value used in the topic modeling literature because it ensures sufficient empirical evidence of convergence without undue computational cost [25].

The ability to fit the probability distribution underlying the LDA model to a specific corpus neatly solves the problem inherent in the statistical distribution assumptions resulting from the Euclidean space formulation of LSA. Once the LDA model has been defined, variants may be utilized, given the nature of the problem being solved. In particular, the LDA model as outlined above is still sensitive to the arbitrary document boundaries imposed by the court recorder. Furthermore, documents vary significantly in length—some might only be two words (e.g., "thank you") whereas others might be significant monologues.

ADDRESSING SPEAKER IDENTITY: THE AUTHOR-TOPIC MODEL

The LDA model possesses significant flexibility in that additional independent variables may be added to guide probabilistic inference. Given that the literature suggests that each speaker possesses an experiential, educational, institutional, and/or role-based signature in his or her word choice, we would like to have the identity of the speaker inform the selection of topics. We therefore use a variant of LDA, known as the author-topic (AT) model [26], which adds probabilistic pressure to assign each author to a specific topic. Shared topics are therefore more likely to represent common jargon. The AT model provides an analysis that is guided by the authorship data of the documents in the corpus, in addition to the word cooccurrence data used by LSA and LDA. Each author (in this case, a speaker in the discourse) is modeled as a multinomial distribution over a fixed number of topics that is preset by the modeler. Each topic is, in turn, modeled as a multinomial distribution over words. A plate-notation representation of the generative process

> **CONSISTENT WITH THE LITERATURE ON SOCIOLINGUISTICS, TOPIC MODELS AIM TO EXTRACT A SET OF GENERAL TOPICS FROM SPECIFIC TEXTS.**

underlying the AT model is shown in Figure 2.

As in LDA, the hyperparameters defining each Dirichlet prior (α and β) are chosen by the modeler and are the primary means, along with choosing the number of topics, by which the form of the model might be controlled. Therefore, two authors' likelihoods of discussing the same topic will depend on the hyperparameters chosen. In general, larger values of α will lead to more topic overlap for any given corpus, motivating the use of a consistent hyperparameter selection algorithm across all corpora analyzed. All hyperparameter settings used for the analyses presented in this article follow the guidelines derived empirically by Griffiths and Steyvers [25]. In particular, $\alpha = 50/(\# \text{ topics})$, inducing topics that are mildly smoothed across authors, and $\beta = 200/(\# \text{ words})$, inducing topics that are specific to small numbers of words.

Like LDA, the AT model is fit using an MCMC approach. Information about individual authors is included in the Bayesian inference mechanism, such that each word is assigned to a topic in proportion to the number of words by that author already in that topic, and in proportion to the number of times that specific word appears in that topic. Thus, if two authors use the same word in two different senses, the AT model will account for this polysemy. Details of the MCMC algorithm derivation are given in the paper by Rosen-Zvi et al. [26].

FILTERING PROCEDURAL LANGUAGE

The AT model was initially developed to analyze corpora of academic papers, each of which may have had multiple authors.
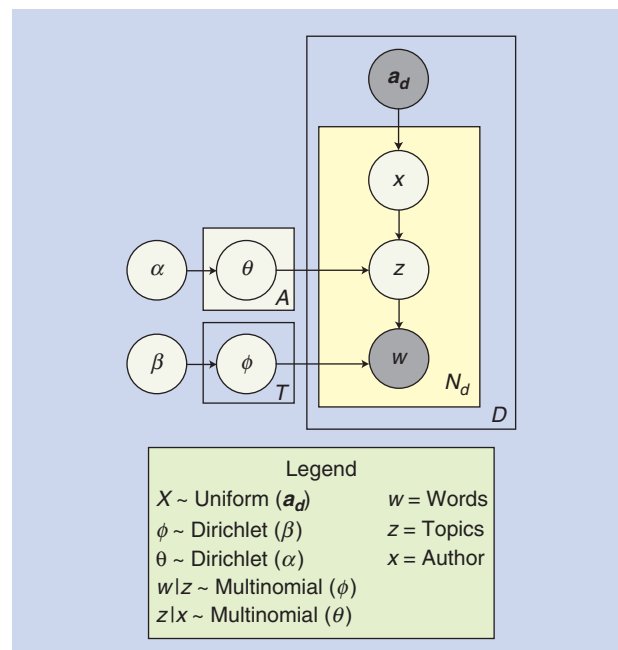
[FIG2] A plate-notation representation of the AT model from [26]. Authors are represented by a multinomial distribution over topics, which are in turn represented by a multinomial distribution over all words in the corpus.

Conversely, when applying the AT model to committee transcripts, each utterance has only one speaker. In such a situation it is often difficult to differentiate between panel members. This is especially true since the majority of the speech during an FDA panel meeting is occupied by presentations from the device sponsor and the FDA. A given voting member might speak relatively rarely. Furthermore, panel members share certain language in common including procedural words and domain-specific words that are sufficiently frequent to prevent good topic identification. As a result, a large proportion of the words spoken by each committee member may be assigned to the same topic, preventing the AT model from identifying important differences between speakers. Ideally, we would be able to separate the unique language used by committee members from that language that they all use as a matter of course. The AT model provides a natural mechanism to perform this separation. Rosen-Zvi et al. [32] used the AT model to create one "fictitious author" for each document in a corpus to remove document-specific words. We use the inverse of this approach by creating one fictitious author named "committee" that is assigned to all utterances spoken by a committee voting members (the authors would like to thank Dr. Mark Dredze for suggesting this approach). Prior to running the AT model's algorithm, all committee voting members' statements are labeled with two possible authors—the actual speaker and "committee." Since the AT model's MCMC algorithm randomizes over all possible authors, words that are held in common to all committee members are assigned to "committee," whereas words that are unique to each speaker are assigned to that speaker. In practice, this allows individual committee members' unique language to be identified. In the limiting case where all committee members' language is common, half of all words would be assigned to "committee" and the other half would be assigned at random to the individual speakers in such a way as to preserve the initial

distribution of that author's words over topics. Sample topics generated using this approach are presented in Table 1.

CONSTRUCTING NETWORKS FROM TOPICS

Consistent with the NTA tradition, Broniatowski and Magee [32] developed a method for extracting social networks from meeting transcripts that is based on the AT model. When applied to a transcript, each utterance is treated as a document. Thus, the meeting transcript may be viewed as a corpus. When the AT model is applied to this corpus, words within each utterance are grouped into topics with probability proportional to the number of times that word has been previously used in that topic, and the number of times that word's "author" (i.e., speaker) has previously used that topic. As discussed above, hyperparameter selection dictates the extent to which topics might overlap. This emphasizes the need for a consistent approach to topic selection that is nevertheless tailored to a specific transcript.

The number of topics for each model is chosen in an automated fashion, using a method developed by Broniatowski and Magee [32]. Briefly, the number of topics is chosen by fitting 35 AT models to the transcript for $T = 1 \ldots 35$ topics, a range that was established empirically. Twenty samples were generated from each of these models, defining a distribution. Each sample's goodness-of-fit is evaluated using a cross-entropy metric based on the AT model structure. The smallest number of topics, t_0, was assigned such that the 95th percentile of the cross-entropy distribution for all larger values of T is greater than the fifth percentile of the cross-entropy distribution associated with t_0. A value of $T = t_0 + 1$ is used to ensure that the model fit is no worse than any defined by a smaller number of topics. Once the number of topics has been chosen, a T-topic AT Model is again fit to the transcript. Ten samples are taken from 20 randomly initialized Markov chains, such that there are 200 samples in total, forming the basis for subsequent analysis. Once a topic model has been fit, constructing social networks requires developing principles regarding what constitutes a link within a given model iteration. As noted above, one would like to link together speakers who commonly use the same topics of discourse. This is accomplished by examining each author-pair's joint probability of speaking about the same topic

$$P(X_1 \cap X_2) = \sum_{i}^{T} P(Z = z_i | X_1) * P(Z = z_i | X_2).$$

Speakers who frequently shared the same topics, i.e., their joint probability exceeded 1/T, were considered to be linked. A square author-author matrix with entries equal to one for each linked author pair, and entries equal to zero otherwise, was constructed. This procedure was then repeated several times for each transcript to average across whatever probabilistic noise might exist in the AT model fit. Speakers who linked across multiple AT model fits more often than would be expected due to chance (i.e., more than 125 times out of 200 samples, conservatively assuming a binomial

[TABLE 1] THE TOP FIVE WORD STEMS FOR ONE TEN-TOPIC RUN OF THE AT MODEL WITH A FICTITIOUS AUTHOR. THE DATA SOURCE IS THE TRANSCRIPT OF THE FDA CIRCULATORY SYSTEMS DEVICES ADVISORY PANEL MEETING HELD ON 4 MARCH 2002. THESE RESULTS WERE GENERATED USING THE FILTERING METHOD OUTLINED ABOVE.

TOPIC NUMBER	TOP FIVE WORD STEMS
1	"CLINIC ENDPOINT EFFICACI COMMENT BASE"
2	"TRIAL INSYNC ICD STUDI WAS"
3	"WAS WERE SPONSOR JUST QUESTION"
4	"PATIENT HEART GROUP WERE FAILUR"
5	"DEVIC PANEL PLEAS APPROV RECOMMEND"
6	"THINK WOULD PATIENT QUESTION DON"
7	"DR CONDIT VOTE DATA PANEL"
8	"EFFECT JUST TRIAL LOOK WOULD"
9	"LEAD IMPLANT COMPLIC VENTRICULAR EVENT"
10	"PATIENT PACE LEAD WERE DEVIC"

distribution with ~15 voting members and applying a Bonferroni correction for a family-wise error-rate of 5%) were considered to be linked in the social network associated with that transcript. Although the structure of these networks is insensitive to small fluctuations in hyperparameter values (e.g., those of the same rough order of magnitude), large changes in hyperparameter values have a strong effect. This is because topic overlap is more likely for larger hyperparameter values, underscoring the importance of using a consistent scheme across meeting transcripts. Indeed, one may interpret hyperparameter values as calibration settings; one must choose hyperparameter values that are likely to provide meaningful results in the same way that one must appropriately focus a microscope to see a clear picture. The steps for constructing this network are found in "Steps Required to Generate Social Networks from Meeting Transcripts," whereas full details of the methodology used to generate these social networks may be found in the article by Broniatowski and Magee [32].

A sample network is presented in Figure 3.

In general, application of this method to the set of 37 FDA advisory panel meetings tested found that voting members tend to link to members of the same medical specialty more often than one would expect due to chance (one-sided Kolmogorov-Smirnov test; $p = 0.0045; n = 37$). Furthermore, results for the subset of meetings in which there was a voting minority of at least two

members provide support for the notion that panel members who vote similarly tend to be linked (one-sided Kolmogorov-Smirnov test; $p = 0.02; n = 11$). These findings are consistent with the literatures reviewed above, suggesting that the method is
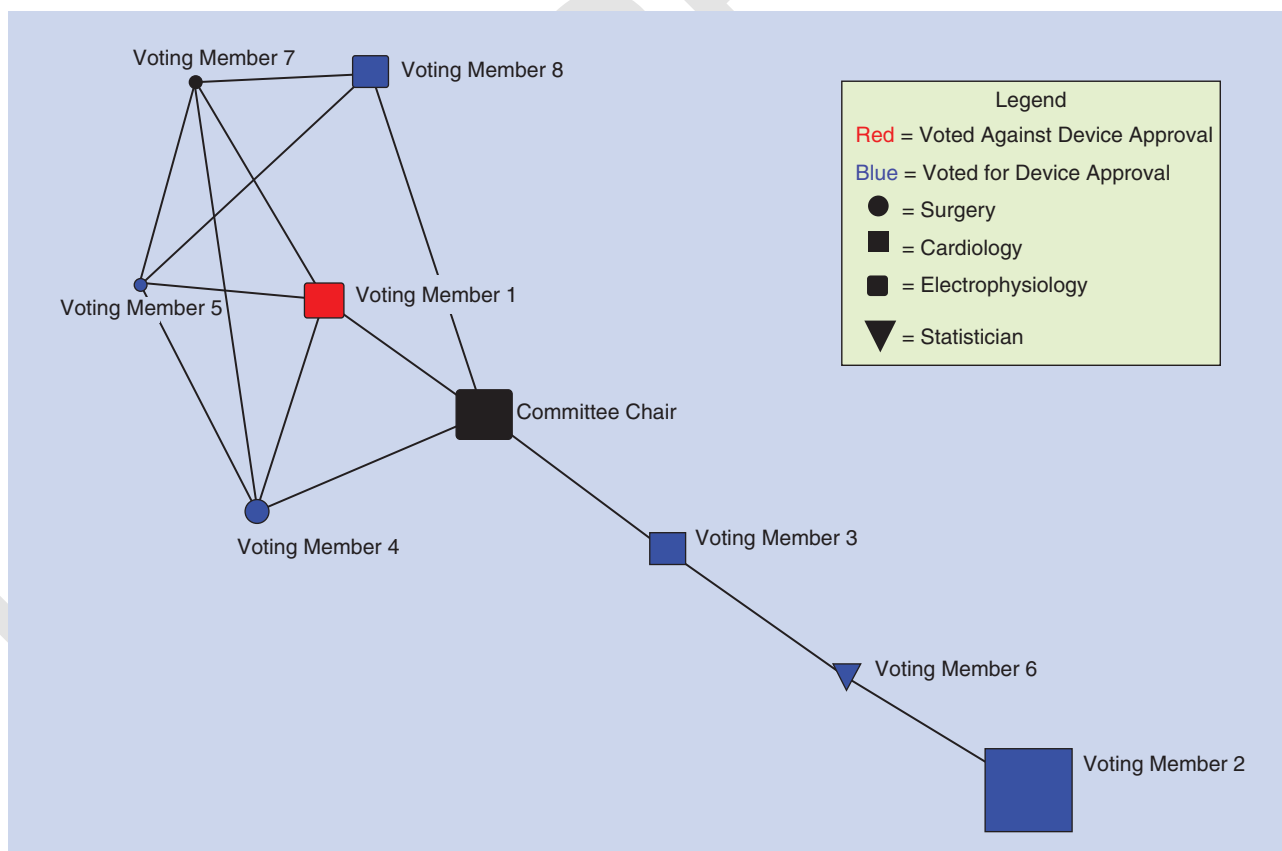
[FIG3] A sample social network generated from the transcript of the FDA Circulatory Systems Advisory Panel meeting held on 27 October 1998. Node shape represents medical specialty, node size is proportional to number of words spoken, and node color represents vote (blue is approval, red is nonapproval, and black is abstention). A directed version of this social network is shown in Figure 8.
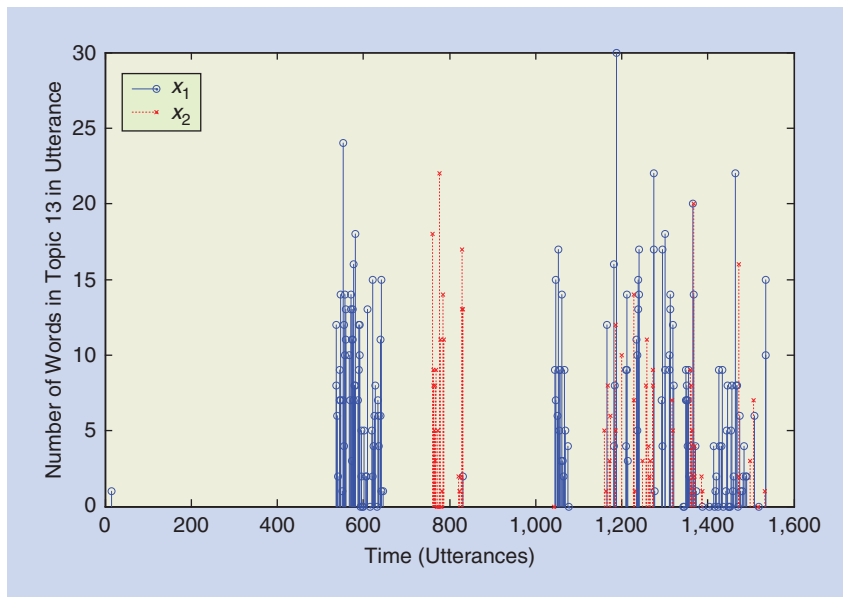
[FIG4] Time series for two speakers on topic 13 during the meeting held on 13 January 2005.

generating valid results. In addition, meetings in which members of the same medical specialty were tightly linked were the same meetings in which members who voted the same way were tightly linked (Spearman rho $= 0.79$, $p = 0.0061$; $n = 11$). This suggests that medical specialty may have served to guide voting behavior in meetings where there was no consensus.

Whereas communication within a specialty should be relatively easy because of shared ways of interpreting data, communication across specialties requires learning and translation, and would consequently be more difficult. Consensus meetings might therefore suggest the possibility of communication across medical spe-

[FIG5] Cross-correlation of the time series shown in Figure 4. Here, the maximum value of the cross-correlation function is less than zero. This is a quantitative indication that x_2 lags x_1.

cialty boundaries, e.g., due to panel members learning to see the same data from multiple, potentially convergent, perspectives. In-depth analyses of these phenomena in terms of psychological and sociological theories of group decision making are left to future work.

DIRECTED GRAPHS

The final limitation of LSA concerns the incorporation of temporal information into the analysis. The conversation analysis literature in sociology (e.g., [5]) notes that, within small groups, influence is often linked to capacity to affect a topic shift. Broniatowski and Magee [33] examined topic overlap among speakers to take advantage of the temporal aspect of our data to develop insights about topic changing. The method, outlined below, considers a sample, s, from the posterior distribution of the AT model. Within this sample, choose a pair of speakers, x_1 and x_2, and a topic z. Given that utterances are temporally ordered, this defines two separate time series (see Figure 4). If x_1 tends to speak about topic z before x_2 does, we can say that x_1 leads x_2. These time series can be compared to generate the topic-specific cross-correlation for speakers x_1 and x_2, in topic z

$$(f_{i,z}^s * f_{j,z}^s)[\delta] = \sum_{d=-\infty}^{\infty} f_{i,z}^s * [d]f_{j,z}^s[\delta + d],$$

where $f_{i,t}^s(d)$ is the number of words spoken by author i and assigned to topic z in document d, in sample s (see Figure 5).

For each sample, s, from the AT Model's posterior distribution, we examine the cross-correlation function for each author pair, $\{x_i, x_j\}$, in topic z. Let there be K peaks in the cross-correlation function described by the expression $m_k = \arg\max_\delta(f_{i,z}^s * f_{j,z}^s)[\delta]$, where $k \in 1 \ldots K$. For each peak, k, if $m_k > 0$, we say that author i lags author j in topic z, at point m_k (i.e., $l_{i,j,z,m_k}^s = 1$). Similarly, we say that author i leads author j in topic z at point m_k (i.e., $l_{i,j,z,m_k}^s = -1$) if $m_k < 0$. Otherwise, $l_{i,j,z,m_k}^s = 0$. For each sample, s, we define the polarity of authors i and j in topic z to be: $p_{i,j,t}^s = \text{median}(l_{i,j,z}^s)$. If most of the peaks in the cross-correlation function are greater than zero, then the polarity $= 1$; if most of the peaks are less than zero, then the polarity $= -1$; otherwise, the polarity $= 0$.

Of particular interest are the topic polarities for author-pairs who are linked
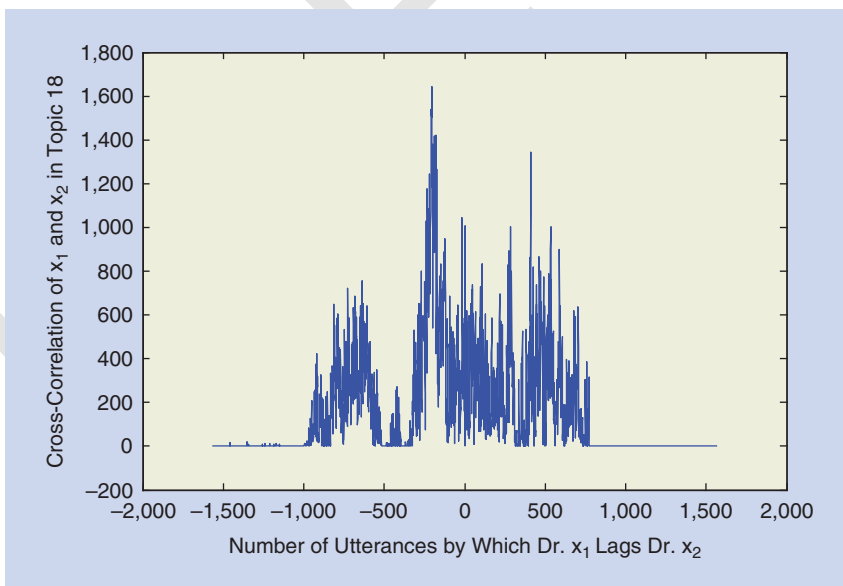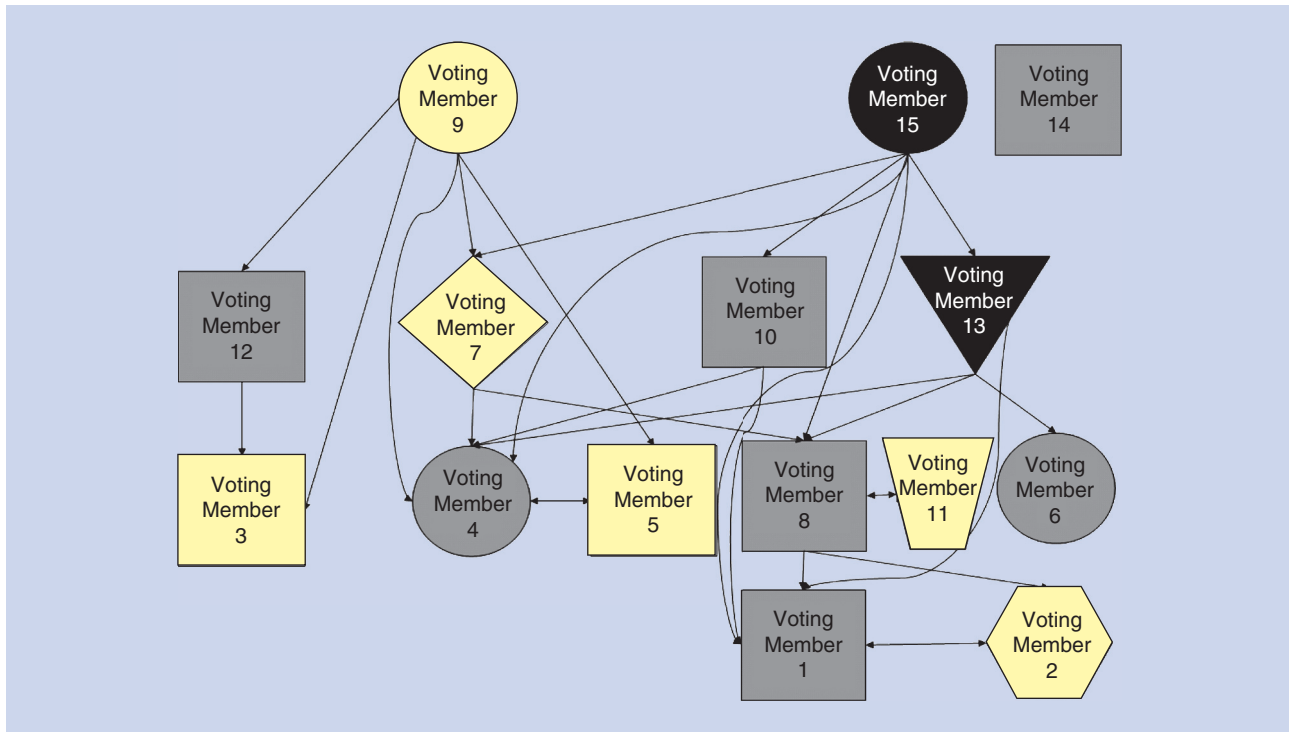
[FIG6] Directed graph representation of meeting held on 23 June 2005 [33]. Hierarchy metric = 0.35 [35]. Node shape represents medical specialty and node color represents vote (yellow is approval, grey is nonapproval, black is abstention). (Figure used with permission.)

in the graph methodology outlined above. Using the polarity values defined above, it is useful to determine directionality for each link. For each sample, s, we define the direction of $e_{i,j}$ in sample s as

$$d^s(e_{i,j}) = \sum_{t=1}^{T} (p_{i,j,t}^s * P^s(Z = z_i | x_i) * P^s(Z = z_i | x_j)).$$

This expression weights each topic polarity by its importance in the joint probability distribution between x_i and x_j, and is constrained to be between -1 and 1 by definition. The set of 200 $d^s(e_{i,j})$ defines a distribution. Next, the net edge direction, $d(e_{i,j})$ is determined by partition of the unit interval into three equal segments. In particular, we examine the proportion of the $d^s(e_{i,j})$ that are greater than zero. If more than 66% of the $d^s(e_{i,j}) > 0$ then $d(e_{i,j}) = 1$ (the arrow points from j to i). If less than 33% of $d^s(e_{i,j}) > 0$ then $d(e_{i,j}) = -1$ (the arrow points from i to j). Otherwise, $d(e_{i,j}) = 0$ (the arrow is bidirectional). The result is a directed network, examples of which are seen in Figures 6–8. Although we did not conduct rigorous tests of sensitivity to these thresholds, the results of preliminary sensitivity testing are qualitatively similar; indeed, with few exceptions, most topic polarity distributions are either sharply skewed toward one end of the distribution or largely symmetric.

SAMPLE DIRECTED GRAPH RESULTS: THE EFFECTS OF PANEL MEMBER SPEAKING ORDER
A sample application of this approach to the same FDA advisory committees studied by Broniatowski and Magee [32], [33] shows that influential panel members, who initiate topics that

others follow, are more likely to be near the "top" of the graph (i.e., a low indegree) whereas panel members who are not followed tend to be near the "bottom" (i.e., a low outdegree). Broniatowski and Magee [34] found a tendency for members of the voting minority to speak later in the meeting, compared to members of the voting majority. Consistent with this finding, members of the voting minority tend to have a lower graph outdegree than do members of the voting majority (Mann Whitney U; $p = 0.045$; $n = 17$). In addition, as meetings get longer, the maximum outdegree of a member of the voting minority increases, i.e., a voting minority member is more likely to reach the "top" of the graph (Spearman rho = 0.50; $p = 0.04$). Under such conditions, voting minorities also become larger—indeed, the maximum outdegree of a voting minority member is strongly associated with the proportion of voting members in the minority (Spearman rho = 0.62, $p = 0.0082$).

LEADERSHIP DYNAMICS ON FDA PANELS
Broniatowski and Magee [27] also explored how these graphs may change their topology on the basis of which nodes are included. In particular, inclusion of the committee chair may or may not "flatten" the graph by connecting the sink nodes (i.e., nodes on the bottom of the graph) to higher nodes. In other meetings, the chair may act to connect otherwise disparate groups of voters or serve to pass information from one set of nodes to another. Using a metric of the hierarchy in the graph [35], one may quantify the impact that the committee chair has upon the meeting. To do so, we first determine, for each graph, which proportion of edges is part of a cycle

$$\frac{\#links\_in\_cycle}{\#total\_links}.$$

We then add the node representing the committee chair to the graph and calculate the hierarchy for this updated graph.

The difference in this hierarchy metric between graphs with and without the chair therefore quantifies impact of the chair on the meeting

$$Chair\ Impact = \left(\frac{\#links\_in\_cycle}{\#total\_links}\right)_{no\_chair}$$
$$- \left(\frac{\#links\_in\_cycle}{\#total\_links}\right)_{chair}.$$

For example, in a directed acyclic graph, the value of this metric would be zero. If, upon including the committee chair, 50% of the links in the graph take part in a cycle, then the value of this metric would increase to 0.5. We display this metric for the graph without the chair (Figure 6) and with the chair (Figure 7). For the meeting held on 23 June 2005 (Figures 6 and 7), this value is $0.78-0.35 = 0.43$. This suggests that the chair is significantly changing the topology of the meeting structure – in particular, it seems that the chair is serving a mediation role by connecting members at the "bottom" of the graph to those at the "top." Among the many roles of the committee chair is to serve as a facilitator, ensuring that all of the panel members present are able to express their views. Other meetings display different behavior by the

chair. Consider the meeting held on 27 October 1998 (in Figure 8).

In this meeting, the committee chair combined the viewpoints of two disparate clusters on the panel. When the chair does not act to significantly change the graph structure, he or she may be taking on the role of a synthesizer—a hands-off approach that allows panel members more freedom to respond to FDA questions, perhaps because additional mediation is deemed unnecessary.

DISCUSSION

The initial findings presented here provide an empirical formalization of, and insight into, the impact of social structure on committee deliberation and voting. The effects of medical specialty on voting behavior have been mentioned above. Additionally, the results of the directed graph analysis are particularly suggestive in that they provide an empirical verification and visualization of paths of influence within a committee meeting. Additionally, these graphs offer quantifiable insights into how different committee chairs might approach running meetings. In our view, the methods outlined in this article hold significant promise for application across a range of social phenomena. Although a full social-scientific treatment of the findings presented here is outside of the scope of this article, these methods provide an emerging toolkit that may be used to analyze group decision making in committees of experts.
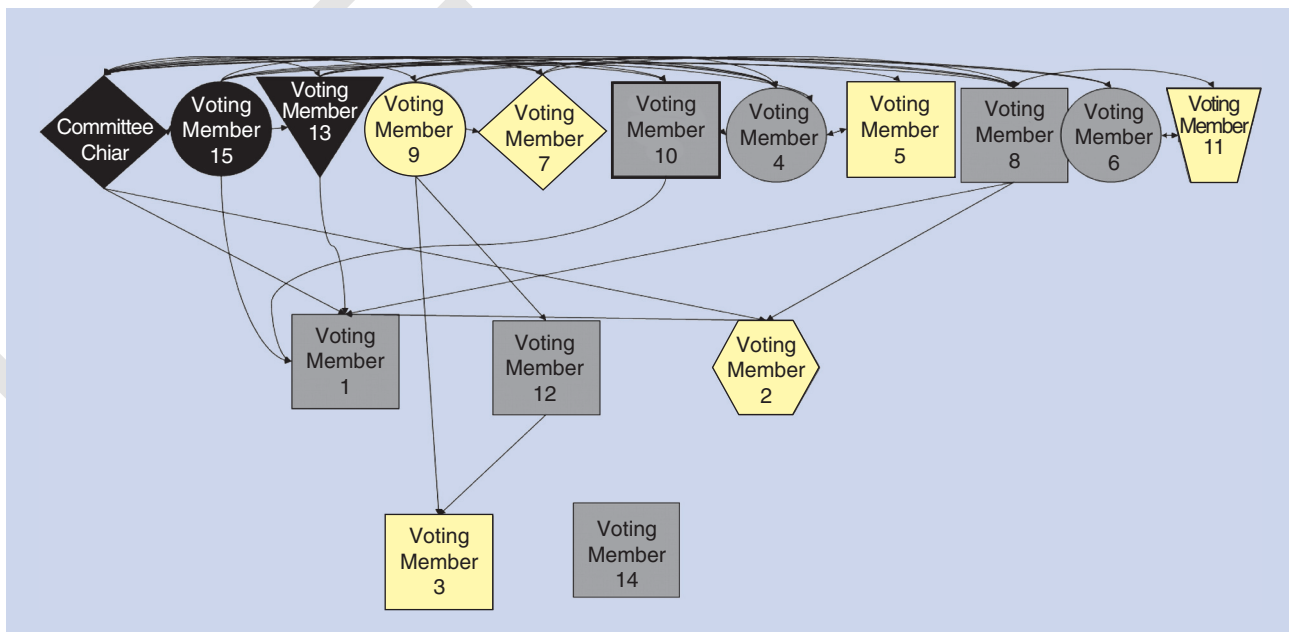
[FIG7] Directed graph representation of meeting held on 23 June 2005, with the committee chair included [33]. Hierarchy metric = 0.78 [35]. Node shape represents medical specialty and node color represents vote (yellow is approval, grey is nonapproval, black is abstention). (Figure used with permission).

METHODOLOGICAL LIMITATIONS

The method outlined in this article relies on meeting transcripts to generate empirical findings regarding committee decision making. It is seemingly sensitive to the limitation that not all committee members might express their views truthfully. Nevertheless, it is very difficult for individuals to avoid using jargon with which they are familiar. A more significant challenge to the use of linguistic data for the analysis of social behavior on expert committees stems from the assumption that such dynamics are entirely reflected in language. Another similar concern is absence of data that might result if a particular voting member of the committee remains silent, says little, or strategically uses speech to hide preferences. Indeed, work by Pentland [36] has shown that much social signaling occurs through body language and vocal dynamics that are not able to be captured in a transcript. In addition, procedural elements of decision making, including the impact of commit-
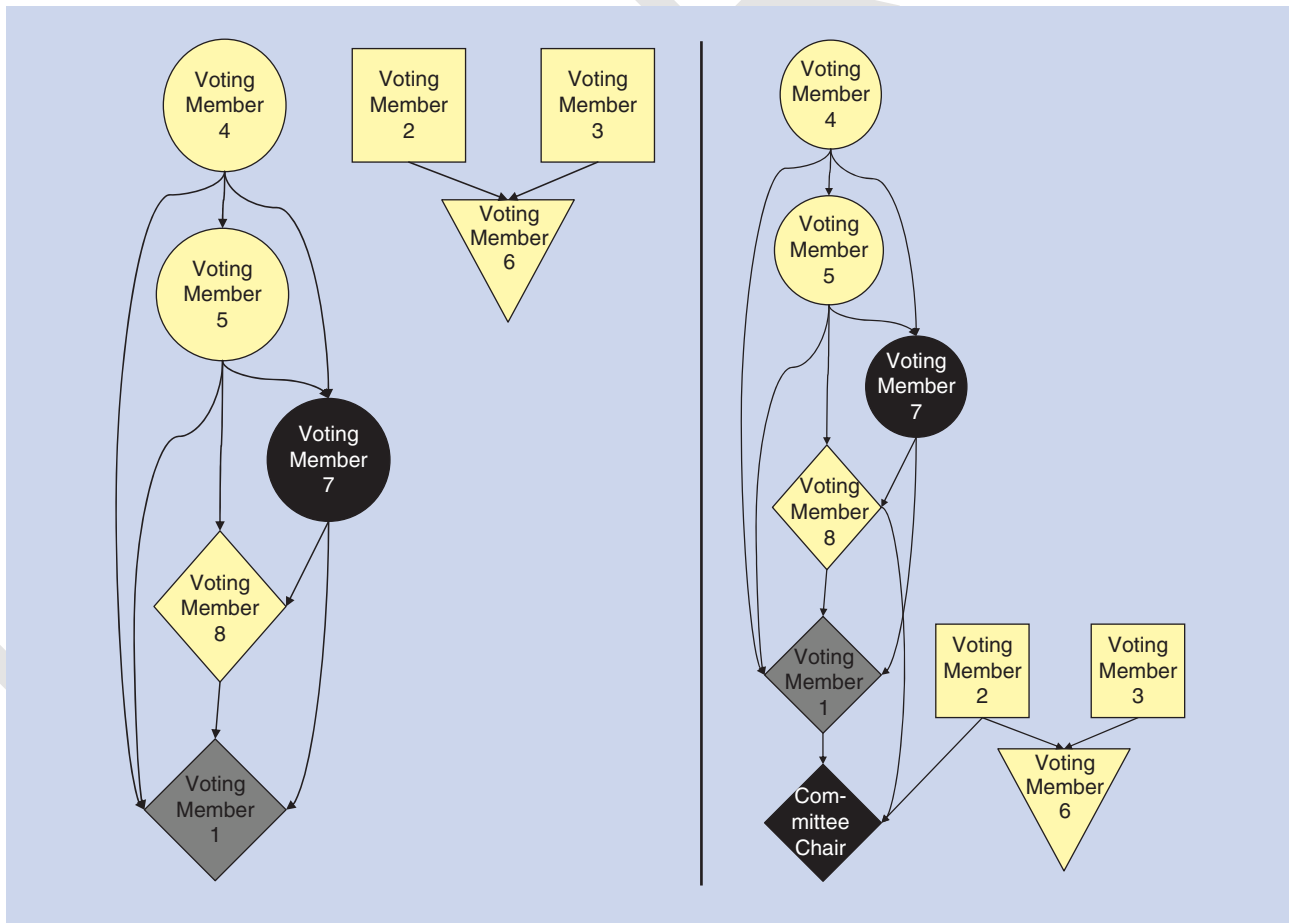
> **WE BELIEVE THAT METHODOLOGICAL INNOVATIONS WILL ENABLE THE FURTHER APPLICATION OF SIGNAL PROCESSING TECHNIQUES TO TEXT, PARTICULARLY AS TEXT DATA FROM REAL-WORLD SETTINGS BECOMES INCREASINGLY AVAILABLE.**

tee member preferences, ordering of alternatives, and voting rule, are important [37]. For a more detailed analysis of these factors in the FDA context, see the article by Broniatowski and Magee [34]. It should therefore be clarified that this article does not claim that all social dynamics are manifest in language, rather, word choice provides one source of insight into a complex, multimodal process.

CONCLUSIONS

Many of the approaches presented in this article use signal processing techniques to analyze text data. In particular, we analyze the flow of information between individuals on technical expert committees using words as representative of the signals in question. Dimensionality-reduction methods, such as LSA and LDA, are common to signal processing literature but have only just started to be applied to the social sciences. A cross-correlation analysis of the resulting topics allows the further

[FIG8] Directed graph representation of meeting held on 27 October 1998 with the committee chair [33]. Hierarchy metric = 0 [35]. Node shape represents medical specialty and node color represents vote (yellow is approval, grey is nonapproval, black is abstention). (Figure used with permission.)

incorporation of temporal information. In general, we believe that methodological innovations will enable the further application of signal processing techniques to text, particularly as text data from real-world settings becomes increasingly available. Policy-relevant findings are sure to follow, enabling the creation of new theory, and opening an exciting new quantitative frontier in computational social science.

ACKNOWLEDGMENT

AUTHORS

David A. Broniatowski (david@mit.edu) is a research affiliate at MIT and senior research scientist at SYNEXXUS, Inc. He holds a Ph.D. degree from MIT's engineering systems division, with research experiences cutting across engineering, management, the social sciences, and aerospace, defense, and health care domains. His research interests focus on computational social science from both empirical and theoretical perspectives. He earned S.B. and S.M. degrees in aeronautics and astronautics, respectively, as well as an S.M. degree in technology and policy, all from MIT. He has industry and research experience in the engineering, public policy, foreign policy, and entrepreneurial sectors.

Christopher L. Magee (cmagee@mit.edu) is a professor in the engineering systems division and mechanical engineering department, MIT, and codirector, International Design Center, associated with the Singapore University of Technology and Design being codeveloped by MIT and Singapore. His recent research has emphasized innovation and technology development in complex systems. He was elected to the National Academy of Engineering (1997) while with the Ford Motor Company for contributions to advanced vehicle development. He was a Ford technical fellow (1996), and is a fellow of the American Society for Metals. He has a B.S. degree and M.S. and Ph.D. degrees in metallurgy and materials science, respectively, from Carnegie Institute of Technology (now Carnegie Mellon University), and an M.B.A. degree from Michigan State University.

REFERENCES

[1] H. Garfinkel, *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall, 1967.

[2] E. Goffman, *Forms of Talk*. Oxford, U.K.: Blackwell, 1981.

[3] H. Sacks, E. A. Schegloff, and G. Jefferson, "A simplest systematics for the organization of turn-taking for conversation," *Language*, vol. 50, no. 4, pp. 696–735, Dec. 1974.

[4] D. W. Maynard, "Placement of topic changes in conversation," *Semiot. La Haye*, vol. 30, no. 3, pp. 263–290, 1980.

[5] D. R. Gibson, "Participation shifts: Order and differentiation in group conversation," *Social Forces*, vol. 81, no. 4, pp. 1335–1380, June 2003.

[6] M. Douglas, *How Institutions Think*. Syracuse, NY: Syracuse Univ. Press, 1986.

[7] J. A. Brown, "Professional language: Words that succeed," *Radical Hist. Rev.*, vol. 34, no. 34, pp. 33–51, 1986.

[8] M. Mulkay, T. Pinch, and M. Ashmore, "Colonizing the mind: Dilemmas in the application of social science," *Social Stud. Sci.*, vol. 17, no. 2, p. 231, 1987.

[9] W. Sack. "Conversation map: An interface for very large-scale conversations," *J. Manage. Inform. Syst.*, vol. 17, no. 3, pp. 73–92, 2000.

[10] C. W. Roberts, *Text Analysis for the Social Sciences*. Evanston, IL: Routledge, 1997.

[11] R. M. Axelrod, *Structure of Decision: The Cognitive Maps of Political Elites*. Princeton, NJ: Princeton Univ. Press, 1976.

[12] K. Carley, D. Columbus, M. Bigrigg, and F. Kunkel, "AutoMap user's guide 2010," Carnegie Mellon Univ., Pittsburgh, PA, Inst. Software Res., Tech. Rep. CMU-ISR-10-121, 2010.

[13] S. R. Corman, T. Kuhn, R. D. Mcphee, and K. J. Dooley, "Studying complex discursive systems," *Human Commun. Res.*, vol. 28, no. 2, pp. 157–206, 2002.

[14] W. L. Neuman, *Social Research Methods: Quantitative and Qualitative Approaches*, 6th ed. Boston, MA: Allyn & Bacon, 2005.

[15] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *J. Amer. Soc. Inform. Sci.*, vol. 41, no. 6, pp. 391–407, 1990.

[16] A. Dong, "The latent semantic approach to studying design team communication," *Des. Stud.*, vol. 26, no. 5, pp. 445–461, 2005.

[17] D. A. Broniatowski, C. L. Magee, J. F. Coughlin, and M. Yang, "The influence of institutional background on the approval of engineered systems," in *Proc. Conf. Systems Engineering Research*, Redondo Beach, CA, 2008, pp. 1–11.

[18] C. H. Papadimitriou, P. Raghavan, H. Tamaki, and S. Vempala, "Latent semantic indexing: A probabilistic analysis," *J. Comput. Syst. Sci.*, vol. 61, no. 2, pp. 217–235, 2000.

[19] B. Bader, R. Harshman, and T. Kolda, "Temporal analysis of semantic graphs using ASALSAN," in *Proc. 7th IEEE Int. Conf. Data Mining, ICDM*, 2007, pp. 33–42.

[20] L. de Lathauwer, B. de Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Applicat.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[21] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Trans. Knowledge Data Eng.*, vol. 21, no. 1, pp. 6–20, 2009.

[22] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Machine Learn. Res.*, vol. 3, no. 4–5, pp. 993–1022, 2003.

[23] T. L. Griffiths, M. Steyvers, and J. B. Tenenbaum, "Topics in semantic representation," *Psychol. Rev.*, vol. 114, no. 2, pp. 211–244, 2007.

[24] D. Hall, D. Jurafsky, and C. D. Manning, "Studying the history of ideas using topic models," in *Proc. Conf. Empirical Methods in Natural Language Processing*, 2008, pp. 363–371.

[25] T. L. Griffiths and M. Steyvers, "Finding scientific topics," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 1, pp. 5228–5235, Apr. 2004.

[26] M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth, "The author-topic model for authors and documents," in *Proc. 20th Conf. Uncertainty in Artificial Intelligence*, 2004, pp. 487–494.

[27] A. McCallum, X. Wang, and A. Corrada-Emmanuel, "Topic and role discovery in social networks with experiments on Enron and academic email," *J. Artif. Intell. Res.*, vol. 30, no. 1, pp. 249–272, 2007.

[28] K. M. Quinn, B. L. Monroe, M. Colaresi, M. H. Crespin, and D. R. Radev, "How to analyze political attention with minimal assumptions and costs," *Amer. J. Polit. Sci.*, vol. 54, pp. 209–228, Jan. 2010.

[29] X. Wang, N. Mohanty, and A. McCallum, "Group and topic discovery from relations and their attributes," *Adv. Neural Inform. Process. Syst.*, no. 18, 1449–1456, 2006.

[30] H. M. Wallach and A. McCallum, "Rethinking LDA: Why priors matter," presented at the NIPS Topic Modeling Workshop, Whistler, BC, 2009, pp. 1973–1981.

[31] M. Rosen-Zvi, C. Chemudugunta, T. Griffiths, P. Smyth, and M. Steyvers, "Learning author-topic models from text corpora," *ACM Trans. Inform. Syst.*, vol. 28, no. 1, pp. 1–38, 2010.

[32] D. A. Broniatowski and C. L. Magee, "Analysis of social dynamics on FDA panels using social networks extracted from meeting transcripts," in *Proc. 2nd IEEE Social Computing Conf. and 2nd Symp. Social Intelligence and Networking*, 2010, pp. 329–334.

[33] D. A. Broniatowski and C. L. Magee, "Towards a computational analysis of status and leadership styles on FDA panels," in *Proc. 4th Int. Conf. Social Computing, Behavioral-Cultural Modeling and Prediction (SBP)*, J. Salerno, S. J. Yang, D. Nau, and S.-K. Chai, Eds. Berlin: Springer-Verlag, 2011, pp. 212–218.

[34] D. A. Broniatowski and C. L. Magee, "Does seating location impact voting behavior on FDA advisory committees?" *Amer. J. Therapeut.* [Online]. Available: http://journals.lww.com/americantherapeutics/Abstract/publishahead/Does_Seating_Location_Impact_Voting_Behavior_on.99593.aspx

[35] J. Luo and C. L. Magee, "Detecting evolving patterns of self-organizing networks by flow hierarchy measurement," *Complexity*, vol. 16, no. 6, pp. 53–61, 2011.

[36] A. S. Pentland, *Honest Signals: How They Shape Our World*. Cambridge, MA: MIT Press, 2008.

[37] R. K. Sah and J. E. Stiglitz, "Committees, hierarchies and polyarchies," *Econ. J.*, vol. 98, no. 391, pp. 451–470, 1988.

[SP]