Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2019 Proceedings | Digital Government and Smart Cities

Predicting success of online petitions from the perspective of agenda setting

Philip Tin Yun Lee
The University of Hong Kong, phil0127@hku.hk

Alvin Ying LU
National University of Singapore Singapore, luying@comp.nus.edu.sg

Feiyu E
The University of Hong Kong, u3005658@hku.hk

Michael Chau
The University of Hong Kong, mchau@business.hku.hk

Predicting Success of Online Petitions from the Perspective of Agenda Setting

Short Paper

Philip Tin Yun Lee
The University of Hong Kong
Pokfulam, Hong Kong
phil0127@connect.hku.hk

Alvin Ying Lu
National University of Singapore
Singapore
luying@comp.nus.edu.sg

Feiyu E
The University of Hong Kong
Pokfulam, Hong Kong
u3005658@connect.hku.hk

Michael Chau
The University of Hong Kong
Pokfulam, Hong Kong
mchau@business.hku.hk

Abstract

Existing predictive models of online petition popularity largely overlooked the literature of agenda setting. This study adheres to Cobb and Elder's (1972) issue expansion model and symbolism (Birkland, 2017) in the agenda-setting literature. Examining the literature, we identified features of popular petitions and examined the effects of these features on online petition success. Commonly used models were used to evaluate our proposed features and compare their performance with benchmark cases. The predictive model, i.e., the product of our study, is the combination of our proposed features and the best performing model. The contributions of the study are two-fold. This study demonstrates how we can translate the textual characteristics described by the literature of agenda setting into technical features that are comprehensible to machines. On practical implications, a better predictive model helps activists better utilize online platforms to secure support for their proposed policy changes.

Keywords: Online petitions, collective actions, text mining, agenda setting

Introduction

Agenda setting is the first stage of policy making cycle (Jann & Wegrich, 2017). Issues that successfully attract the public attention are put on the public agenda. This is the beginning of any possible changes of relevant policies (Birkland, 2017). When there is adequate public attention to the issues, policy decision makers will feel the pressure to put the issues on their formal agendas. Eventually, the decision makers decide whether and how to act on the issues (Cobb et al., 1976). Platforms of online petitions are important tools, since they can facilitate processes of individuals gathering support and attention from others for any proposed changes of existing policies in the online environment. Through online petition platforms, individuals can easily launch a petition campaign and gather support from other Internet users. Some popular examples of these online petition platforms include Avaaz and Change.org. The Internet inherently allows the petition to reach a large population with affordable resources.

In recent years, researchers have examined various types of collective actions through online platforms. Some popular topics include crowdfunding (e.g. Hong et al., 2018; Siering et al., 2016) and crowdsourcing (e.g. Mo et al., 2018). Online petitioning is also one of the popular topics. A number of studies on online petitions have explored how text-mining technologies can be used to understand a large volume of petitions proposed by the general public in a timely manner (e.g. Hagen, 2018; Suh et al., 2010). The results of these studies indicate a promising potential of the use of text-mining technologies in the context of online petitions. Yet, apparently, the existing studies on online petition platforms tended to focus more on viral effects of petitions through the social media (e.g. Jalali et al, 2016). Relatively few studies have investigated texts of the petitions and explored what textual characteristics can attract more people's attention. Specifically, existing predictive models of online petition popularity largely overlooked the literature of agenda setting. Unfortunately, the insights of previous scholars of agenda setting have been neglected. Although some of these scholars' observation and empirical studies were more relevant to the time before the age of the Internet, many of their insights shall still apply to the online environment.

This study adheres to Cobb and Elder's (1972) issue expansion model and symbolism (Birkland, 2017; Schattschneider, 1975; Stone, 2002), both of which are remarkable concepts in the literature of public policy and agenda setting. Examining these two concepts and other relevant literature, we identified features of popular petitions and will examine whether and to what extent these features improve prediction of online petition success. To develop a predictive model of the success, we will adopt Lash and Zhao's (2016) approach. Various commonly used models, including Decision Tree (DT), Naïve Bayesian (NB), Support Vector Machine (SVM), Random Forest (RF) and Neural Networks (NN), will be used to evaluate our proposed features with a real-life dataset of an online petition platform. We will then compare the performance of our features with benchmark cases and will discuss the potential of our proposed features. The predictive model, i.e. the product of our study, is the combination of our proposed features and the best performing model (i.e. DT /NB/SVM…).

The contributions of the study are two-fold. First, we aim to demonstrate the use of the literature of political science and communication in the context of online petitions. Apparently, no researcher has considered this knowledge area when they develop predictive models of online petition success. We treat online petition platforms as a tool that can be used in setting the public agenda, but not simply a general platform of collective actions. This study demonstrates how we can translate the textual characteristics described by the literature of agenda setting into technical features that are comprehensible to machines. Second, on practical implications, a better predictive model can help activists to better utilize online platforms to secure support for their proposed policy changes, especially given that a real-life dataset is used in the evaluation. The improved model helps fulfill the promised potential of e-participation (Kim & Lee, 2012; Macintosh, 2004) of which the goal is to enable a larger population to engage in democratic debates.

Literature Review

Agenda-setting and Online Petition

Agenda setting, according to Birkland (2017), is "the process by which problems and alternative solutions gain or lose public and elite attention" (p. 63). It is the first stage of policy cycle, followed by policy formulation and decision making, implementation, and evaluation and termination (Jann & Wegrich, 2017). The public agenda is different from the agenda of political decision makers. The former agenda mostly determines what issues to be put on the latter agenda. Those issues on the agenda of political decision makers then pass through the policy cycle and possibly lead to changes of relevant policies. Thus, the mechanism of agenda setting acts like a filter of policies that will be focused by political decision makers in future. Since public attention is limited, issues have to compete among themselves for places in the public agenda. Through online petition platforms, individuals can propose petitions that express their views on different issues in the society. The petitions aim to gather wide attention of online users and hopefully lead to changes in the corresponding policies at last. Even if some petitions may target at bad business practices of some companies at first, the pressure will later be shifted to the government to implement new regulatory policies to rectify the wrong practices. As a result, policy changes are achieved.

Issue Expansion Model

Whether an issue in a petition can attract the public attention largely depends on how it is defined. According to Cobb and Elder (1972), the issue expansion depends on some defined characteristics of an issue. One of such characteristics is the concreteness of an issue: "The more ambiguously an issue is defined, the greater the likelihood that it will reach an expanded public" (p. 112). Specifically, Cobb and Elder (1972) identified the difference between goals and objectives which represents the two poles of concreteness-specificity continuum. Goals are vague terms and doctrines such as liberty and equality, whereas objectives are referred to as specific demands of political decision makers' actions.

Another characteristic is temporal relevance: "(T)he more an issue is defined as having extended temporal relevance, the greater the chance that it will be exposed to a larger audience" (p. 117). Actions for an issue can be described as the start of a long-lasting trend. Some people may not be affected by the issue, but they will be affected by the possible "spill-over" effect of the actions. Therefore, an issue which has strong temporal relevance can draw attention of this type of people.

The third characteristic is complexity: "(T)he more non-technical an issue is defined to be, the greater the likelihood that it will be expanded to a larger public" (p.120). The use of technical language prohibits the general public to participate in the issue discussion, and therefore reduces the chance of issue expansion.

The last characteristic is categorical precedence: "(T)he more an issue is defined as lacking a clear precedent, the greater the chance that it will be expanded to a larger population" (p. 122). If the issue has existed for long, then the public may have an impression that the issue was examined but could not be resolved. On the other hand, if an issue appears to be unprecedented, novelty of the issue can catch more people's eyeballs. This also concurs with Kingdon's (1995) streams metaphor of agenda change. According to Kingdon (1995), a change of perception of an issue opens a window of policy change, implying a possible increase of public attention to the issue and public pressure on political decision makers for problem solving.

Symbolism

Symbol, in agenda-setting literature, is referred to as "anything that stands for something else. Its meaning depends on how people interpret it, use it, or respond to it" (Stone, 2002, p. 137). Symbols can be used to elevate an issue to the public agenda by inducing empathy of the media and the general public (Birkland, 2017; Schattschneider, 1975). Cobb and Elder (1972) discussed several aspects of the use of symbols. They suggested that symbols with a long historical background are more likely to evoke reactions and attention of the public. Stronger symbols have usually been used in a large number of issues. Cobb and Elder (1972) also argued that symbols can be used together for better issue expansion.

The occurrence of focusing events is one of the main triggers of public attention to a specific issue (Baumgartner and Jones, 1993; Birkland, 1998; Cobb & Elder, 1983; Kingdon, 1995). A focusing event is defined as "an event that is sudden; relatively uncommon; can be reasonably defined as harmful or revealing the possibility of potentially greater future harms; has harms that are concentrated in a particular geographical area or community of interest; and that is known to policy makers and the public simultaneously" (Birkland, 1998, p. 54). The media reporting the events will leave readers a memory of "casual stories" of the event (Stone, 1989, p. 281). There are four different types of causes of an issue: mechanical cause, accidental cause, intentional cause, and inadvertent cause (Stone, 2002). The media reporting generates symbols of the issue, and the symbols facilitate the issue expansion (Birkland, 1998).

Predictive Models on Online Petitions

Several studies have explored how predictions of online petition popularity can be improved. Hagen et al. (2016) examined how various linguistic and semantic factors respectively influence the popularity of online petitions. They made three concluding remarks. First, extreme language inhibits the success of petitions. Following Craig and Blankenship's (2011) approach, Hagen et al. (2016) found that petitions which mention words "much more", "extremely", "very", and "wonderful" are less popular among the online audience. Second, names in petitions are not appealing to the online population. Using the StandfordCoreNLP NER tagger (Finkel et al., 2005), Hagen et al. (2016) extracted named entities from the petitions. These named entities included persons, locations and organizations. Hagen et al. (2016) showed that only names of persons are negatively correlated with the popularity of the petitions. Third, petitions which mentioned well-known topics or important events are more popular. The authors adopted a semi-automatic approach with the use of LDA to produce a list of topics. They then qualitatively analyzed and identified topics that were significantly correlated with the popularity of petitions.

Associating the words in petitions with the relevant categories of General Inquirer (Stone et al., 1966), Chen et al. (2019), on the basis of the dual-process theory of persuasion (Petty & Brinol, 2015), found that petitions with positive emotions and enlightening information are more appealing to online users. Furthermore, they showed that online population is not interested in moral and cognitive reasoning.

Methods

Feature Engineering

Features are derived from the aforementioned literature of agenda setting. These features are divided into two types, namely linguistic features and semantic features.

Linguistic Features

Concreteness of a petition is represented by the frequency of abstract words appeared in a petition. More abstract words used indicate a lower level of concreteness of an issue and a higher chance to attract public attention, based on Cobb and Elder (1972). The list of abstract words is derived from General Inquirer (GI), a commonly used linguistic and content analytical tool in academic studies (Chen et al., 2019). Specifically, the list includes words that are tagged abstract in GI.

Temporal relevance is represented by the frequency of words related to "future" in a petition. The repository of words related to "future" will be prepared with the use of WordNet, which is a large lexical database of English words and all the words are organized in synonym sets. Each set expresses a distinct concept, in which items share similar conceptual-semantic meaning and are lexically related (Miller, 1998). Besides, we will use the online Oxford Dictionary[1] to complement the synonym sets. A petition with stronger temporal relevance is more likely to attract people's attention (Cobb & Elder, 1972).

Complexity is represented by the frequency of technical words in a petition. By technical words, we mean those words that are highly relevant to specific domain knowledge. Understanding the technical words requires some previous exposure to the domains. The less technical words are therefore more appealing to the general public, since the exposure to the domains is not necessary for clear understanding of this kind of words. The repository of technical words will be prepared with the use of online Oxford Dictionary. Those technical words were tagged with a "domain" label in the Online Oxford Dictionary. We will calculate the number of words with the domain tag in each petition. More words of this kind indicate a higher level of complexity of a petition.

Semantic Features

A code book is developed to determine an issue's policy category. Two faculty members of political science are invited to formulate the code book. For each policy category, the code book includes a list of keywords for each policy category. The policy category which has the highest frequency of keywords appears in a petition is treated as the petition's policy category. The level of categorical precedence is represented by how a petition is different from previous petitions in the same policy category. We will apply the average term frequency (TF value) of all the words in each petition to reflect the degree to which the petition is different to other petitions under the same category. We assume that the more common words one petition owns, which means a higher average TF value in mathematics, the less likely this petition is different from others.

Symbolic topics are obtained by training LDA model. LDA is a generative probabilistic model which expresses document via a distribution of topics, and each topic is further represented via a probabilistic distribution of words (Blei et al., 2013). We not only extract topics using petition contents as written by Hagen et al. (2016), but also use the corpus from New York Times. The corpus that was written by professional journalists is believed to reflect issues' symbols more consistently and succinctly than user-generated online petitions. Through unsupervised training process, topics generated by LDA are assumed to be symbols with different influential power to petition popularity. Direct application of New York

[1]We access Oxford Dictionary via its API (https://developer.oxforddictionaries.com/).

Times enables us to make a clearer expression of how one petition contains implicit meanings that are shared among different policy categories. To measure the influential impacts of the symbolic topics, we firstly use LDA model that is trained by datasets with more professional petition information to statistically express each petition in a probability distribution of topics, and then we evaluate the helpfulness of importing such features into various classification models.

Benchmark features

Few studies have examined the textual characteristics that lead to higher popularity. Hagen et al. (2016) found that linguistic *extremity* negatively affects the popularity. This *extremity* linguistic feature was measured by whether a petition contained any of the following words: “much more”, “extremely”, “very”, and “wonderful”. Besides, the study showed that linguistic features *repetition* of words and *internet activity* (i.e., whether a petition mentions words including “http”, “www”, “html”, and “Youtube”) have negative effects on the popularity. The features *urgency* (i.e. whether a petition mentions words such as “immediately”, “immediate”, “urgent”, and their synonyms) and *sentiment* positively affect the popularity. However, except *extremity*, the effects of the remaining linguistic features disappeared when topic variables were incorporated into the predictive model. Hagen et al. (2016) also discovered that petition with more names of *persons* are likely to be unpopular. The named entities were identified using Stanford CoreNLP NER tagger (Finkel et al., 2005). Furthermore, the authors derived topics from their training data set using an approach which involves LDA, post-hoc human annotations and calculated statistics on a dataset different from the training dataset. They eventually identified *15 topics* and suggested that topics related to people’s daily life, such as veteran, children and investigation, or important public events, such as study-visa, military and China, have positive relationships with petition popularity.

Chen et al. (2019) developed a multi-appeal model which was composed of cognitive appeals, moral appeals and emotional appeals to predict online petitions’ popularities and successes. They used categories in GI that are relevant to these appeals to measure the petitions’ strength of these appeals. Specifically, they found four significant appeal factors, namely negative emotion (an emotional appeal factor), linguistic modality (a moral appeal factor), enlightenment and understatement (cognitive appeals).

Experiment Design

To demonstrate practical usefulness of our proposed features, we evaluate the features with a dataset of Avaaz[2], one of the most influential online petition websites. The website allows individuals to advocate their political views and to acquire support from others. The dataset consists of around fifty thousand English petitions across six years span. We tag a petition with “successful” as long as the numbers of signatures reached the targets set by the petition authors. Otherwise, a petition is labeled with “failed”. Thus, the evaluation of proposed features generates binary classification results.

Following Lash and Zhao’s (2016) approach, we first use 10-fold cross validation to prove the effectiveness of the full feature set which contains linguistic features, semantic features and benchmark features. In addition, we compare the performance of the classification model fed by the full feature set with the performance of the model fed only by benchmark features. All classification experiments are conducted with exploration of a number of commonly used models, including Decision Tree (DT), k-Nearest Neighbor(KNN), Naïve Bayesian (NB), Support Vector Machine (SVM), Random Forest (RF) and Neural Networks (NN). This set of models was also used in Lash and Zhao (2016). The predictive model, i.e. the product of our study, is the combination of our proposed features and the best performing model (i.e. DT/NB/SVM…). Figure 1 shows a data-oriented illustration of our overall experimental set-up.

[2]The data were collected from https://avaaz.org/.

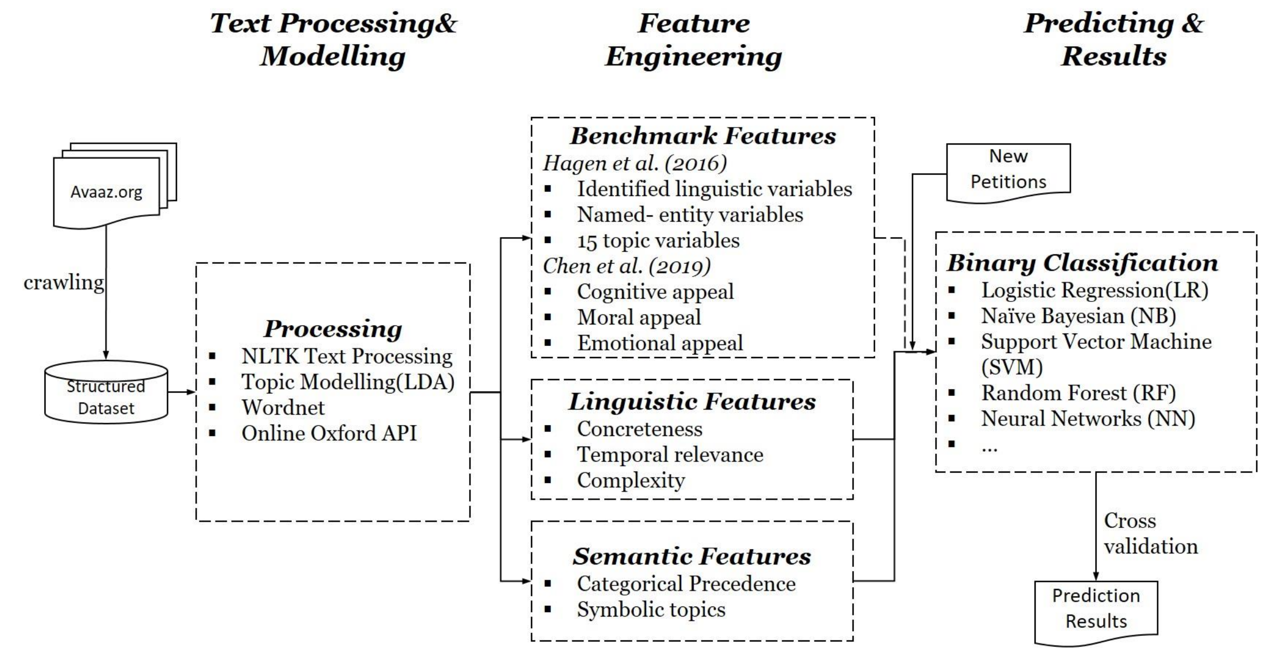

Figure 1. Data-oriented Illustration of the Overall Experimental Set-up

Preliminary Results

Five thousand pieces of petitions were selected randomly from the successful and unsuccessful group respectively to preliminarily explore the performance of the proposed semantic features and the linguistic features. In the preliminary data analysis, the benchmark features include Hagen et al.'s (2016) identified linguistic variables and named-entity variables as well as Chen et al.'s (2019) features of cognitive appeal, moral appeal and emotional appeal. All our proposed linguistic and semantic features were included in the analysis. An initial version of the code book was used in our preliminary data analysis. Three commonly used performance metrics of classification, namely Recall, Precision and F1, were used to evaluate the performance. Six classification models were used in the experiment: Support Vector Machine (SVM), Neural Network (NN), K-Nearest Neighbor (KNN), Naïve Bayesian (NB), Decision Tree (DT) and Random Forest (RF).

Table 1 shows the average scores of the three performance measures (Precision, Recall and F1) achieved by importing the different types of features into the classification models. As shown in Table 1, the combination of the proposed semantic features, the linguistic features, and the basic features had the best performance. The encouraging preliminary results support our exploration of the proposed features in future.

Performance Metrics	Features	Average across Classification Models
Precision	BF	0.62
	BF & SF	0.64
	BF & LF	0.62
	BF & SF & LF	**0.65**
Recall	BF	0.73
	BF & SF	0.72
	BF & LF	0.68
	BF & SF & LF	**0.74**
F1-score	BF	0.66
	BF & SF	0.67
	BF & LF	0.64
	BF & SF & LF	**0.69**
SVM: Support Vector Machine; NN: Neural Network; KNN: K-nearby Neighbors; NB: Naïve Bayes; DT: Decision Tree; RF: Random Forest BF: Baseline Features SF: Semantic Features LF: Linguistic Features		

Table 1. Preliminary Results

Discussions and Future Work

Our preliminary results show that our proposed new features perform better than the benchmark features. Decades ago, only the more educated people used the Internet. Nowadays, many people in developed countries have good access to the Internet. The online population becomes more comparable to the offline population. Although the agenda setting literature was contextualized by the offline environment, we believe that the literature shall still provide us with insights and ideas about how to improve the predictive model of petition success in the online environment. The results of the predictive model will offer a better picture of the dynamics in the online petition platforms. The weights of the variables will be found to help us identify important variables for online petition successes. The results will help interest groups, activists or actually any individuals to make better use of petition platforms to gather support for their proposed changes in public policies. Specifically, for those interest groups who work on issues that are less popular, they shall spend extra efforts on those variables that can be manipulated.

We are currently incorporating the concept of participatory journalism into our proposed model. Previous researchers in agenda setting have widely discussed the important role of the mass media in setting the public agenda, gaining public attention and subsequently influencing government actions on certain issues (e.g. Downs, 1996; Lodge and Hood, 2002). The mass media was able to direct readers' thinking, highlighting certain issues' salience and implying real reasons for issues (McCombs & Shaw, 1972). However, recent studies have shown that the traditional mass media has rather weak influence in the online environment (e.g. Meraz, 2009 and Wu et al., 2013). In Bowman and Willis' (2003) book *We Media, How Audiences Are Shaping the Future of News and Information,* they defined participatory journalism as "the act of a citizen, or group of citizens, playing an active role in the process of collecting, reporting, analyzing and disseminating news and information" (p.9). Social media users can now play the roles of reporters in the online environment. We shall examine the interaction between features of petition content and responses of users on different social media, and possibly incorporate the interaction into our models for better predictions.

References

Bamgartner, F., and Bryan D. J. 1993. *Agenda and Instability in American Politics*. Chicago: University of Chicago Press.

Birkland, T. A. 1998. "Focusing events, mobilization, and agenda setting," *Journal of Public Policy* (18:1), pp. 53-74.

Birkland, T. A. 2017. "Agenda setting in public policy," in *Handbook of Public Policy Analysis: Theory, politics, and methods,* F. Frank, J. M. Gerald, and S. S. Mara (eds.), CRC Press, pp. 63-78.

Blei, D. M., Ng, A. Y. and Jordan, M. I. 2003. "Latent dirichlet allocation," *Journal of Machine Learning Research*, pp. 993-1022.

Bowman, S., and Willis, C. 2003. *We Media. How Audiences Are Shaping the Future of News and Information*. American Press Institute.

Chen, Y., Deng, S., Kwak, D., Elnoshokaty, A. and Wu, J. 2019. "A Multi-Appeal Model of Persuasion for Online Petition Success: A Linguistic Cue-Based Approach," *Journal of the Association for Information Systems* (20:2), pp. 105-131.

Cobb, R. W., and Elder, C. D. 1972. *Participation in American Politics: The Dynamics of Agenda-Building*. Johns Hopkins University Press.

Cobb, R., Ross, J. K., and Ross, M. H. 1976. "Agenda building as a comparative political process," *American Political Science Review* (70:1), pp. 126-138.

Craig, T. Y., and Blankenship, K. L. 2011. "Language and persuasion: Linguistic extremity influences message processing and behavioral intentions," *Journal of Language and Social Psychology* (30:3), pp. 290-310.

Downs, A. 1996. "Up and Down with Ecology: The "Issue-Attention Cycle"," in *The politics of American Economic Policy Making*, P. Peretz (ed.), M. E. Sharpe, Inc, pp. 48-59.

Finkel, J. R., Grenager, T., and Manning, C. 2005. "Incorporating non-local information into information extraction systems by Gibbs sampling," in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, K. Knight (ed.), pp. 363–370.

Hagen, L., Teresa, M. H., Uzuner, O., May, W., Fake, T. and Katragadda, S. 2016. "E-petition Popularity: Do Linguistic and Semantic Factors Matter?" *Government Information Quarterly* (33:4), pp. 783-795.

Hagan, L. 2018. "Content analysis of e-petition with topic modeling: How to train and evaluate LDA models?," *Information Processing & Management*, (54:6), pp. 1292-1307.

Hong, Y., Hu, Y. and Burtch, G. 2018. "Embeddedness, Prosociality, and Social Influence: Evidence from Online Crowdfunding," *MIS Quarterly* (42:4), pp. 1211-1224.

Jalali, M. S., Ashouri, A., Herrera-Restrepo, O., and Zhang, H. 2016. "Information diffusion through social networks: The case of an online petition," *Expert Systems with Applications,* pp. 187-197.

Jann, W., and Wegrich, K. 2017. "Agenda setting in public policy," in *Handbook of Public Policy Analysis: Theory, Politics, and Methods,* F. Frank, J. M. Gerald, and S. S. Mara (eds.), pp. 43-62.

Kim, S., and Lee, J. 2012. "E-participation, transparency, and trust in local government," *Public Administration Review* (76:2), pp. 819-828.

Kingdon, J. W. 1995. *Agenda, Alternatives and Public Policies. 2nd ed.* Pearson.

Lash, M. T. and Zhao, K. 2016. "Early predictions of movie success: The who, what, and when of profitability," *Journal of Management Information Systems* (33:3), pp. 874-903.

Lodge, M., and Hood, C. 2002. "Pavlovian policy responses to media feeding frenzies? Dangerous dogs regulation in comparative perspective," *Journal of Contingencies and Crisis Management* (10:1), pp. 1-13.

Macintosh, A. 2004. "Characterizing e-participation in policy-making," in *Proceedings of the 37th Annual Hawaii International Conference on System Sciences,* IEEE, pp. 117-127.

McCombs, M. E., and Shaw, D. L. 1972. "The agenda-setting function of mass media," *Public Opinion Quarterly* (36:2), pp. 176-187.

Meraz, S. (2009). "Is there an elite hold? Traditional media to social media agenda setting influence in blog networks." *Journal of Computer-Mediated Communication* (14:3), pp. 682-707.

Miller, G. 1998. *WordNet: An Electronic Lexical Database*. MIT press.

Mo, J., Sarkar, S. and Menon, S. 2018. "Know When to Run: Recommendations in Crowdsourcing Contests," *MIS Quarterly* (42:3), pp. 919-943.

Petty, R. E. and Briñol, P. 2015. "Emotion and persuasion: Cognitive and meta-cognitive processes impact attitudes," *Cognition and Emotion* (29:1), pp. 1-26.

Stone, D. A. 1989. "Causal stories and the formation of policy agendas," *Political Science Quarterly* (104:2), pp. 281-300.

Stone, D. 2002. *Policy Paradox: The Art of Political Decision Making*. Norton.

Stone, P. J., Dunphy, D. C. and Smith, M. S. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. Oxford, U.K.: MIT Press.

Schattschneider, E. E. 1975. *The Semi-Sovereign People: A Realist's View of Democracy in America*. Cengage Learning.

Siering, M., Koch, J. and Deokar, A. 2016. "Detecting Fraudulent Behavior on Crowdfunding Platforms the Role of Linguisitic and Content-based Cues in Static and Dynamic Context," *Journal of Management Information Systems* (33:2), pp. 421-55.

Suh, J.H., Park, C.H. and Jeon, S. H. 2010. "Applying text and data mining techniques to forecasting the trend of petitions filed to e-People," *Expert Systems with Applications* (37:10), pp. 7255-7268.

Wu, Y., Atkin, D., Lau, T. Y., Lin, C., and Mou, Y. 2013. "Agenda setting and micro-blog use: An analysis of the relationship between Sina Weibo and newspaper agendas in China," *The Journal of Social Media in Society* (2:2), pp. 8-25.