INFORMS Journal on Data Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Rhetoric Mining: A New Text-Analytics Approach for Quantifying Persuasion

Michelle M. H. Şeref, Onur Şeref, Alan S. Abrahams, Shawndra B. Hill, Quinn Warnick

informs
https://pubsonline.informs.org/journal/ijds

INFORMS JOURNAL ON DATA SCIENCE
Vol. 2, No. 1, April–June 2023, pp. 24–44
ISSN 2694-4022 (print), ISSN 2694-4030 (online)

Rhetoric Mining: A New Text-Analytics Approach for Quantifying Persuasion

Michelle M. H. Şeref,[a,*] Onur Şeref,[a] Alan S. Abrahams,[a] Shawndra B. Hill,[b] Quinn Warnick[a]

[a] Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061; [b] Columbia University, New York, New York 10027

*Corresponding author

Contact: mmhseref@vt.edu, https://orcid.org/0000-0002-3642-9875 (MMHŞ); seref@vt.edu, https://orcid.org/0000-0002-6918-1327 (OŞ); abra@vt.edu, https://orcid.org/0000-0002-4884-3192 (ASA); sbh2146@columbia.edu, https://orcid.org/0000-0003-1980-727X (SBH); qwarnick@vt.edu (QW)

Received: June 27, 2021
Revised: May 11, 2022; December 12, 2022
Accepted: December 19, 2022
Published Online in Articles in Advance: March 27, 2023

https://doi.org/10.1287/ijds.2022.0024

Abstract. *Rhetoric mining* is a novel text-analytics method for quantifying persuasion based on *rhetorical analysis theory*. Our mixed-methodology approach combines qualitative context analysis with automated tagging and quantification of rhetorical moves. Rhetorical moves are complex discursive patterns and, thus, require a sequence-based text-mining approach, rather than the simpler word-based frequency analyses. We apply a sequence-alignment method to detect *semantically equivalent* sequences with high precision and efficiency. We illustrate our method by analyzing arguments used to justify stock picks in an online investment community. For these data, we detect and quantify the rhetorical moves of *ethos* (personal versus cited expertise), *hedging* (confidence versus uncertainty), and *evidence type* (product-, company-, or stock-based evidence). We use rhetoric mining to identify argument styles of *persuasiveness* (pitches that receive community recommendations) and *trustworthiness* (pitches that are written by successful investors). Rhetoric mining provides a new analytic lens in Information Systems research to analyze the influence of persuasion in consumer decision making.

History: W. Nick Street served as the senior editor for this article.

Data Ethics & Reproducibility Note: The code capsule is available on Code Ocean at https://codeocean.com/capsule/9373643/tree/v1 and in the e-Companion to this article (available at https://doi.org/10.1287/ijds.2022.0024).

Keywords: rhetoric mining • sequence alignment • text analytics • persuasion • trustworthiness

1. Introduction and Motivation

Rhetorical approaches to studying language focus on analyzing the intentional use of language by speakers and writers to influence an audience to think or act in a particular way. The academic field of rhetorical studies has grown far beyond this simple description, but for our purposes, it will suffice to say that rhetoric is *any intentional language choice used to persuade others*. This type of strategic communication is pervasive in situations where language is used to influence decision making. For example, there is rhetoric in accounting reports, which may influence auditing or investor decisions; there is rhetoric in consumer reviews, which may influence purchase decisions; and there is rhetoric in healthcare documentation, which may influence patient procedure choices. Rhetoric enables us to understand language's relationship to decisions and actions, as well as to incorporate contextual and situational influences on decision making.

The academic study of rhetoric has its roots in public speaking and political persuasion, dating to the time of Greek and Roman thinkers such as Plato, Aristotle, Cicero, and Quintilian. A rhetorical analysis includes a critique of the *speaker* (or writer), the *audience* (or reader), the *situation* or context in which the speaker and audience interact, and the *goal* the speaker is trying to achieve by persuading the audience. Rhetorical criticism uses the initial contextual investigation to then identify the types of persuasive moves employed by the speaker, as well as the audience's reaction to the speaker's rhetoric. Rhetoric scholars study the variation of persuasive moves across speakers, contexts, and cultures and their influences on different audiences, particularly in decision-making situations. The academic disciplines of Communication Studies and English examine rhetoric in public relations, mass media, online communities, healthcare information, public policy, technical instructions, intercultural negotiations, and many other areas (McClosky 1985, Miller 1990, Simons 1990, Aristotle 1991, Kuypers 2009, Barrett et al. 2013). Traditionally, rhetorical analysis employs very small data sets—for example, a single political speech, a collection of corporate memos, or a sample of government documents. This research model is necessary because rhetorical analysis has always depended upon human scholars to collect, analyze, and interpret textual artifacts. However, if computational tools can be developed to perform some of the work traditionally undertaken by humans, researchers can begin studying rhetorical moves in massive data sets.

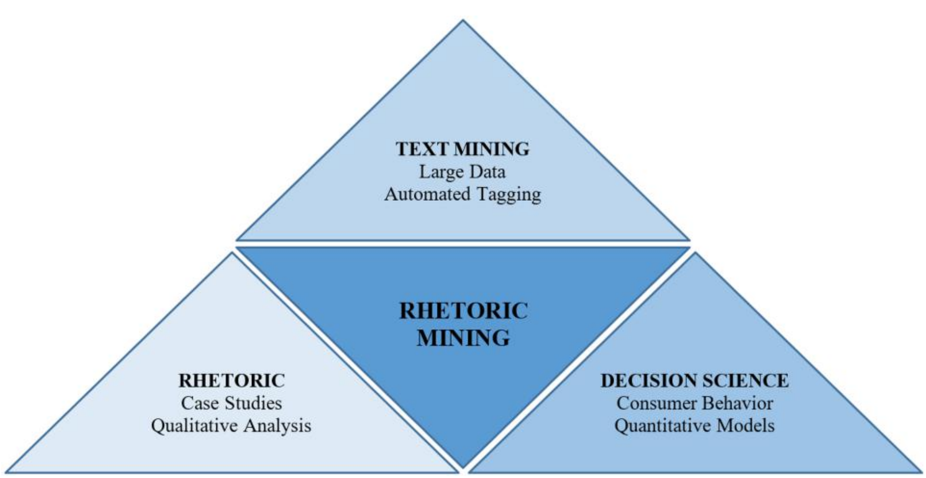
Rhetoric has not typically been used in analyzing business decision-making situations. There is an opportunity to extend the typically small-scale, in-depth analysis employed by rhetoric scholars in the humanities and social sciences to a larger-scale rhetorical analysis through the development of computational approaches in information systems and business analytics. Automated rhetoric tagging in large data sets could give management access to unique insights, such as: What types of persuasion are common in user communities? Which forms of persuasion has our company used? Which forms of persuasion do our competitors use? Which are successful? Which are contagious? Which forms of persuasion suit particular product types, customer types, or situations? Which forms of persuasion work with suppliers or business partners? Which forms of persuasion influence employees? Our work bridges a disciplinary gap between the related fields of rhetoric and decision science by offering a computational method for rhetorical analysis of large volumes of text data (Figure 1).

Automating rhetoric tagging, however, is not trivial. It requires an approach that combines human expertise, computational tagging, and machine learning, the relevance of which is increasing in the field of human machine learning (Zagalsky et al. 2021). In rhetoric mining, the role of the human expert is to identify different collections of sequences of words that represent different rhetorical moves using a manageable random sample from a corpus of text. The human expert uses context analysis and rhetorical analysis to understand the subtext and represent it as sequences of words. However, each sequence typically has multiple semantically equivalent variations in the corpus, many of which are not in the random sample. Moreover, a priori enumeration of such variations is inefficient and time-consuming.

To apply rhetoric mining to large-scale data, we utilize a sequence-alignment algorithm to efficiently and effectively identify sequences of words that represent particular rhetorical moves. Our approach is based on the Smith-Waterman local sequence-alignment algorithm (Smith and Waterman 1981) used in computational biology applications of genetic sequencing. We modify this dynamic programming method to identify semantically equivalent sequences of words, thus allowing us to use a relatively small training set to tag the larger text corpus. We derive these training sequences from an initial round of human coding (or manual highlighting), which allows us to define context-dependent rhetorical moves. The sequence-alignment algorithm is trained on these manual highlights to develop a computational highlighter, which can be applied to larger textual data with efficiency and accuracy. This computational highlighter outputs a vector enumerating instances for each identified rhetorical move in each text unit. Transforming this qualitative information into quantitative vectors of rhetoric instances allows us to perform a new level of text analytics on large data to address questions about persuasion.

In this paper, we employ rhetoric mining to study the argumentation styles used to justify investment decisions in the Motley Fool online community. In particular, we identify rhetorical moves used in stock pitches and characterize rhetorical styles (combinations of moves) used in persuasive pitches (receiving recommendations from the online community) and those used in trustworthy pitches (written by successful investors). Our paper is organized as follows. In Section 2, we review the literature related to rhetoric in text mining. In Section 3, we describe in detail our rhetoric mining methodology and

Figure 1. (Color online) Rhetoric Mining as an Interdisciplinary Approach



Note. This approach combines rhetoric theories and qualitative analysis methods with text-mining automated tagging approaches for large data to analyze the influence of persuasion in quantitative decision science models.

sequence-alignment algorithm. In Section 4, we illustrate an application of rhetoric mining via sequence alignment in analyzing stock-pick arguments in an online investment community. We provide performance results for our computational highlighter tool, as well as report descriptive analytics derived from the resulting vectors of rhetoric instances for this application. In Section 5, we discuss a generalizable approach for our rhetoric mining method, as well as potential future application areas.

2. Related Literature

In this section, we review research related to mixed methods, rhetorical analysis in decision making, sequence-tagging techniques, deep neural networks, and persuasion and trustworthiness in online communities. We review the current state of research in these areas and identify the research gap that our paper addresses—namely, the ability to automate the detection of sequences of words that define rhetorical moves of persuasion in large text data.

2.1. Mixed Methods

Many studies involve collection and analysis of both quantitative and qualitative data in the pursuit of the same research inquiry, where the research approach is a mixture and interplay of qualitative and quantitative methods (Creswell and Plano 2011). Mixed methods might help reveal insights broader and deeper than those that can be obtained by only qualitative or quantitative methods. There are methodological approaches in information systems research to bridge the gap between qualitative and quantitative methods (Venkatesh et al. 2013), and guidelines have been proposed for conducting mixed-methods research (Venkatesh et al. 2016). Studies that involve analysis of large volumes of text data have a significant overlap with mixed-methods research, as they often require qualitative coding and quantitative processing. Social media analysis is among the important domains of these studies, with applications spanning numerous fields from healthcare (Hamad et al. 2016) and production (Choi et al. 2016) to demographics (Chakrabarti and Fry 2022), behavioral research (Andreotta et al. 2019), big-data analytics (O'Halloran et al. 2016), and natural language processing (Parks and Peters 2022).

2.2. Rhetorical Analysis in Decision Making

Computational linguistics and natural language processing study the natural (or subconscious) use of language. Rhetoric, however, focuses on intentional language choices used to persuade an audience. Rhetorical analysis studies strategic language use, as well as the effect of persuasive moves or argumentation styles on a particular audience. Research in rhetoric points to the prospective application of rhetorical analysis to decision making (McClosky 1985, Miller 1990, Simons 1990). This points to a gap in both rhetoric research and business research in studying the effect of persuasion on decision-making situations in business: complementing current quantitative models of decision science with qualitative information now readily available in large text data.

Most persuasion studies in information systems research use psychology-based theories, such as the Elaboration Likelihood Model (Chaiken and Trope 1999, Cheung et al. 2012, Lowry et al. 2012), the Heuristic-Systematic Model (Chaiken 1980), Sensemaking (Weick 1995, Abbasi et al. 2018), Persuasive System Design (Weick 1995, Oinas-Kukkonen and Harjumaa 2009), and Intention Mining (Khodabandelou et al. 2014). These theories tend to focus on subjects' behaviors, cognitive processing, social structure, and motivation. There is still a large gap in information systems research related to the analysis of language theories in decision making (Lyytinen 1985, Pollach 2012). Some language theories, such as Speech-Act Theory (Searle 1969, Flores and Ludlow 1980) or Language Action Perspective (Habermas 1984, Johannesson 2001, Rittgen 2006), recognize the influence of language in decision making, but do not identify the same units of language as in rhetorical analysis. Rhetorical analysis identifies different persuasive moves, considers the context of the decision-making situation, and then seeks to understand the resulting influence on a particular audience.

Rhetorical analysis draws on several methods of rhetorical criticism established by seminal scholars in rhetoric: neo-Aristotelian criticism (examining tools of organization or invention, as well as persuasive appeals, such as ethos, pathos, and logos) (Archak et al. 2011), rhetorical situation (exigence, audience, and constraints) (Bitzer 1968), metaphor analysis (creating connections using comparisons) (Lakeoff and Johnson 1980), narrative critique (identifying people, places, events, actions, and plots) (Geertz 1973), pentadic analysis (examining ratios of act, agent, agency, scene, and purpose) (Burke 1969), ideological analysis (identifying underlying beliefs and values) (Althusser 1971), feminist criticism (seeking out marginalized perspectives) (Butler 1990), and genre analysis (critiquing conventional expectations of genre) (Measell 1976). The selection of a particular method of rhetorical criticism depends on the artifact of study, the perspective of the researcher, and the context of the rhetorical situation.

Rhetorical analysis methods typically follow a close-reading approach; in some cases, large-scale methods are used to aggregate close-reading analyses or examine a corpus at large. Digital humanities researchers have begun to employ computational methods to conduct large-scale textual analysis. Berry (2012) examined relationships among large-scale documents and related demographic or geographical data. Wojcik (2011, p. 20) develops heuristics from rhetorical theory in hopes of achieving a "higher level of abstraction than such linguistic fields as sociolinguistics or pragmatics." Ishizaki and Kaufer (2012) developed a tool, called DocuScope, that uses a standard home-grown dictionary classified into over 100 rhetorical functions. However, computational approaches to rhetorical analysis have still not been widely explored, evaluated, or accepted in the field of rhetoric, which still prefers in-depth research on a small data set to broader pattern analysis of a larger data set (Geisler 2016).

There has been some exploratory research on rhetoric in text mining that has examined sentence structure and argumentation (Petty et al. 1981, Baroni et al. 2005). Barrett et al. (2013) studied rhetorical features such as framing and ideology. They acknowledged rhetoric as a valuable, yet underdeveloped, method for examining IT diffusion. They argued that "by making persuasive arguments, actors justify and rationalize the adoption of a practice, thereby enabling its diffusion" (Barrett et al. 2013, p. 205). Zhang et al. (2014) studied how consumers perceive an online review to be informative or persuasive, claiming that argument quality is impacted by source credibility and perceived number of reviews. Hromada (2011) modeled rhetorical moves based on the repetition or reordering of text. He noted that "the study of persuasion remains understudied and underrepresented in current natural language systems" (Hromada 2011, p. 85). Grasso (2002) used various computer science techniques to tag inductive argumentation structures based on parts of speech and sentence-tree structures. Studies by Taboada and Mann (2006) and Markle-Hub et al. (2017) use Rhetorical Structure Theory to analyze arguments through a hierarchical, connected structure of texts. Ficcadenti et al. (2019) study the rhetoric dynamics of a large collection of U.S. presidents' speeches by implementing a rank-size procedure over word frequencies in individual speeches. Antons et al. (2019) use a dictionary-based automated content analysis to extract rhetorical signals in scientific texts. Ihlen et al. (2021) use rhetorical situations to address vaccine-skepticism challenges and introduce principles for content strategy for public health communicators. There have also been recent studies in information systems on user engagement in social media. Glass and Colbaugh (2012) examined rhetorical framing in social media. Other experimental methods for automatic tagging of rhetoric moves have used linguistic approaches, or language-independent approaches that rely on *n*-grams and statistical methods (Teufel and Moens 2002, Anthony and Lashkia 2006, Feltrim et al. 2006, Pendra and Cotos 2008). However, most of these methods still move toward developing a dictionary of moves for a particular language context. There is still much more to be discovered in terms of both the contribution of insights gained from rhetorical analytics and the development of methods for implementing computational approaches to identifying rhetorical moves.

2.3. Standard and State-of-the-Art Methods in Natural Language Processing and Understanding

Early text-mining methods involve standard natural language processing (NLP) tools, such as common expressions (Backus et al. 1963), part-of-speech tagging (Harris 1962, Kupiec 1992), stemming (Porter 1980), content analysis (Stone et al. 1966, Neuendorf 2002), information extraction (Cowie and Lehnert 1996, Grishman 1997, Gaizauskas and Wilks 1998, Craven et al. 2000), and semantic role determination (Gildea and Jurafsky 2002, Punyakanok et al. 2008, Wei et al. 2008). Many of these approaches typically depend on predefined language patterns (frames) or domain-specific dictionaries of terms with applications in medicine (Sager et al. 1987), finance (Antweiler and Frank 2005, Das and Chen 2007, Loughran and McDonald 2011, Lia et al. 2014, Mai et al. 2018), and the auto industry (Abrahams et al. 2012, 2015).

The vector space model (VSM) is frequently used in information retrieval, document classification, and document clustering. One of its most commonly used features is term frequency-inverse document frequency (Jones 1988). VSM vectors can be generated from semantic features, such as sentiment, opinion, and subjectivity (Choi et al. 2005, Wilson et al. 2005). Topic-modeling methods like Latent Dirichlet Allocation (Blei et al. 2003) produce a set of topics, in which each document is generally represented by a vector of weights over the set of topics. The methods in the generation of most of these representations, however, depend on individual terms or phrases and discard their sequence information. Within the last decade, distributed representations of words and phrases have become increasingly successful in NLP tasks, such as Word2Vec (Mikolov et al. 2013) and GloVe (Pennington et al. 2014). These methods embed a sparse, high-dimensional representation of words to a dense, low-dimensional vector representation, while implicitly preserving the semantic relationship among sequences of words. Extensions of these methods have been introduced that extend the embedded vector representation for longer sequences from sentences to documents (Le and Mikolov 2014). A sequence of these embedded vectors representing words are usually used for training recurrent neural network (RNN) architectures, such as long short-term memory (LSTM) (Józefowicz et al. 2016) and bidirectional LSTM (BiLSTM) (Zhou et al. 2016). Combining character-level tokens with BiLSTM, context-sensitive embeddings from the language model (ELMo) have been created (Peters et al. 2018).

The focus on contextual representations has recently led to transformer-based methods, such as bidirectional encoder representation from transformers (BERT), which have become the state-of-the-art (SOTA) in natural language understanding (NLU) (Devlin et al. 2018) and have been used widely (Kalyan et al. 2021). Transformer-based multilingual and language agnostic embeddings allow robust representation of text with superior performance in machine translation (Feng et al. 2020). Multitask deep neural networks featuring BERT provide SOTA results on multiple NLU tasks (Liu et al. 2019b). Knowledge-enabled language-representation models using BERT improve result in domain-specific tasks (Liu et al. 2019a). Transformer models work well with reinforcement learning models on a variety of tasks, from text classification (Chai et al. 2020) to text abstraction (Wang et al. 2019). Transformer models have also been advancing the natural language generation (NLG) field. Models like generative pretrained transformer 3 (GPT-3) (Brown et al. 2020) can generate human-like text from an initial text prompt. Google's breakthrough conversation technology, language model for dialogue applications (LaMDA) (Thoppilan et al. 2022), can engage in free-flowing, open-ended conversations and can be used for knowledge systems for human-machine interaction, challenging fundamental intelligence benchmarks like the Turing test. NLP, NLU, and NLG are rapidly changing fields, with transformer models leading the way.

We include RNN models such as LSTM and BiLSTM and transformer models such as BERT in our computational results in Section 4.4.

2.4. Sequence Alignment

Sequence alignment is a method mainly used in bioinformatics to identify regions of similarity among two or more sequences of DNA, RNA, or proteins (Mount 2004). Sequence-alignment methods can be categorized as global and local with respect to their scope. Global alignment methods seek to find the best possible alignment among the input sequences by performing editing operations for aligning the entirety of the sequences. In contrast, local alignment methods search for the best subsequence that is common among the input sequences, while ignoring remainder in each sequence. In terms of the number of input sequences, sequence-alignment methods can be categorized as pairwise alignment methods and multiple alignment methods. The pairwise alignment methods take two sequences for alignment. Multiple alignment methods consider more than two sequences for aligning them simultaneously, rather than aligning two sequences at a time.

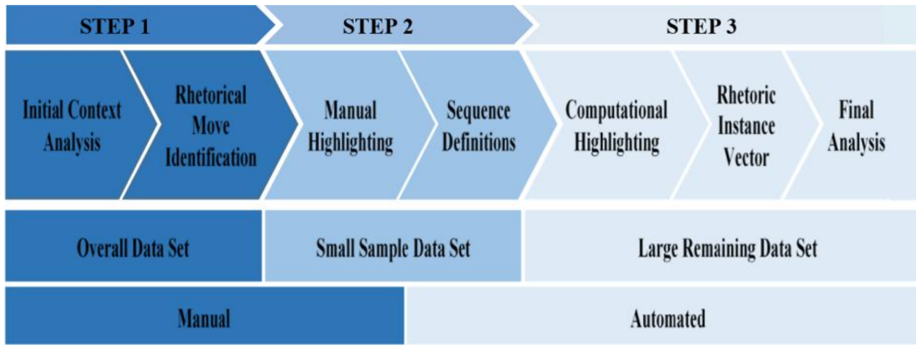
Dynamic programming (DP) is frequently used in both global and local alignment. For example, the Needleman-Wunsch algorithm (Needleman and Wunsch 1970) is a DP-based, pairwise sequence-alignment algorithm typically used for global alignment. However, there are other methods for multiple global alignment, like Kalign, which uses progressive alignment (Lassmann and Sonnhammer 2005), or SAGA, which uses genetic algorithms (Notredame and Higgins 1996). The Smith-Waterman (SW) algorithm is a DP-based, pairwise alignment algorithm, which is usually preferred for local alignment (Smith and Waterman 1981). The SW algorithm is easy to implement; however, there are other implementations of the SW algorithm like BLAST (Altschul et al. 1990), PSI-BLAST (Altschul et al. 1997), and ScalaBLAST (Oehmen and Nieplocha 2006) specifically optimized for a variety of problems in genomics research.

The Smith-Waterman local sequence-alignment algorithm has mostly been used as a similarity measure between two relatively large units of text (Irving 2004, Menon et al. 2018). However, we adopt SW to identify sequences of words that represent particular rhetorical moves. In our approach, sequences can be readily defined and generalized from examples. This means that rhetoricians do not require training in complex programming syntax (e.g., regular expression formats for pattern recognition), but can nevertheless rapidly configure an efficient and accurate automated rhetoric recognizer. We use a target sequence of few words as the first sequence and the entire corpus as the second sequence, in which we search for the target sequence and its variations. This task corresponds to a pairwise local alignment, for which the SW algorithm is the standard method with a linear time complexity. Global or multiple sequence-alignment methods are not relevant with this task, whereas our adoption of the SW algorithm can effectively identify all occurrences and variations of the target sequence in the corpus.

2.5. Persuasion and Trustworthiness in Online Communities

Several papers in Information Systems research have addressed questions of persuasion and trust in collaborative decision making, much more so recently in online communities. This topic is of particular interest to researchers studying online reviews and their influence on consumer purchasing decisions (Wang et al. 2015, Liu and Karahanna 2017). Hill and Ready-Campbell (2011) also examined the Motley Fool online investment community to evaluate the relative impact of "experts" in aggregate information collection, assigning experts more weight when measuring the "wisdom of crowds." We further evaluate the determination of experts or leaders by analyzing the arguments the authors made in the paper for "smart" investments. Johnson et al. (2015) found that community leaders write more concise posts with simple language. Kim and Benbasat (2006) used Toulmin's model of argumentation (Toulmin 1958) to examine the relative influence of claim only, claim plus data, and claim plus data and backing and found that more data and backing increase trust. Kim et al. (2009) examined the role of trust in reasoned action and expectation-confirmation theories of consumer behavior, noting that trust is evaluated over three stages—namely, prepurchase, purchase, and postpurchase experience. Robert et al. (2009) determined that initial

Figure 2. (Color online) Rhetoric Mining Methodology Diagram



assessment of a team member's characteristics, along with an individual's own disposition to trust, form first impressions of trustworthiness, which are later dominated by behavior assessment of perceived ability, integrity, and benevolence. Our research extends trust evaluation by challenging whether consumers are persuaded by particular statements or arguments. We address these questions by performing a rhetorical analysis of argument style for any unit of text.

2.6. Research Gap

The prior work discussed above validates the usefulness of rhetoric in text mining and points to the challenge of implementing various rhetorical features or argumentation moves in an efficient manner for large text-data applications. The methods currently used in text mining are typically word-based or dictionary-dependent. However, we believe we are the first to apply a sequence-alignment algorithm to identify sequences of words, with a particular associated word-sense that defines specific rhetorical moves. The underlying dynamic programming model of the sequence-alignment algorithm allows us to computationally highlight instances of rhetoric in an efficient manner with high accuracy. Computational linguistics has historically focused on information extraction: extracting facts, knowledge, or propositions from text. However, this approach inherently misses a primary purpose of many business communications, which is the intention of the sender to influence the thinking (through persuasive moves in text) of receivers.

This is a novel approach in the fields of both rhetoric and text mining. Rhetoric mining can be particularly useful in business-decision analytics in identifying types of persuasion used within a particular community in a specific decision-making context. The study of persuasion may be able to help explain or predict some social- or business-network behaviors, such that rhetoric mining could be complementary to network analytics for authority, hubs, and influencers (Cascavilla et al. 2015, Fang and Hu 2018). Additionally, studying the impact of rhetorical choices, in given contexts, has the potential to influence the cognitive psychology facet of Information Systems and Marketing, similar to the framing-effects literature. There is valuable information to be gained from a rhetorical analysis of large-scale text data, and, thus, there is a need to develop and validate a computational approach to studying how rhetoric is employed to influence business decision making.

3. Rhetoric Mining Methodology and Sequence-Alignment Algorithm

Our methodology involves three general steps: (1) context analysis of the overall data set to identify salient rhetorical moves, (2) manual highlighting of a sample of text to create sequence definitions of each rhetorical move, and (3) computational highlighting of the remaining data set to quantify and analyze vectors of instances of rhetorical moves (see Figure 2). In this section, we provide a general overview of this methodology, and in the following section (Section 4), we illustrate an exemplar application of our approach: analyzing stock-pick arguments in an online investment community.

3.1. Initial Context Analysis and Rhetorical Move Identification

We first perform a general context analysis and initial rhetorical analysis of the data to identify salient language features and rhetorical moves (Kuypers 2009). Both the context analysis and initial rhetorical analysis should be performed on the original data source with a large random sample to provide an overview impression. Context analysis includes understanding the decision-making environment, key decision makers, and language-based, as well as non-language-based, influences on decision makers. This also includes recognizing the "speakers" and "audience" in the data context and understanding why the speakers are motivated to persuade this particular audience.

We determine which rhetorical moves are most relevant to a particular text, given a particular context. This step requires *qualitative* analysis, resulting in identification of important textual and nontextual features, which will then inform *quantitative* analysis of the decision-

Table 1. Sequence Highlights with Dummy Markers and Lemmas

Original sequence highlight	Sequence with dummy markers	Sequence with lemmas
This is my *best* investment	This is my * investment	This be I * invest
This is my *worst* investment	This is my * investment	This be I * invest
This is my *best ever* investment	This is my * investment	This be I * invest
This is (*or was*) my *best* investment	This is my * investment	This be * I * invest

making process. The advantage of rhetoric mining is the ability to leverage the qualitative analysis to efficiently scan large amounts of text, converting qualitative information into quantitative vectors of tabulated instances of rhetorical moves. Rhetoric mining therefore extends a typically qualitative research approach to the computational generation of numerical data representing the rhetorical approaches used by the authors within a large collection of documents. This allows for a typically quantitative decision-science analysis to include richer information (Miller 1990, Lacity and Janson 1994).

3.2. Manual Highlighting and Sequence Definitions

Once a set of rhetorical moves has been identified, we perform manual highlighting on a sample set of the data to provide specific examples from the actual text source. This sample should be randomly drawn from several sources of relatively equal text length to ensure that one writer's style does not influence the examples of rhetorical moves. This qualitative process of content analysis should be performed by two or more human coders following an established protocol defining each rhetorical move (Krippendorf 2004). An initial small sample should be provided to both coders to compare manual highlights until a satisfactory interrater-reliability score is achieved. Then, coders can work independently on larger random samples.

We allow coders to denote any "dummy" words within rhetorical moves; the software does not require an exact match on these words to find a rhetorical move. For rhetorical moves, more importance is given to semantic equivalence rather than sentiment equivalence. For example, we may highlight the statement "this is my *best* investment" to denote a personal experience to gain trust (an ethos rhetorical move), but we do not necessarily care what sentiment that experience holds. That is, the statement "this is my *worst* investment" would be considered the same type of rhetorical move. Therefore, we would denote "best" as a dummy phrase and define an instance of this move as "this is my * investment." We collapse consecutive dummy words into one dummy marker, and we also allow for multiple nonconsecutive dummy markers in a sequence.

Once the sample instances have been highlighted for each rhetorical move, we extract the highlighted sequences to create a sequence bank for each rhetorical move definition. We run Stanford Natural Language Processing analysis on the sample data (Toutanova and Manning 2000) to identify sentence breaks and note word lemmas (stems or root words). In the previous examples of highlighted sequences, the lemma format would be transformed as in Table 1.

Longer sequences ensure high accuracy, as they provide context-based synonyms. Using sequences of lemmas helps the software control for synonym or part-of-speech issues when identifying the rhetorical moves. At the same time, shorter sequences may falsely highlight an excessive number of synonym equivalents. Thus, it is important that the manual highlights are accurate in their definition of each rhetorical move, keeping enough words to ensure context, but also being short enough to avoid overconstraining the sequence match. We have found that two- to five-word sequences work best. We review the details of our manual highlighting in the application of online stock pitches in Section 4.

3.3. Computational Highlighting via Sequence-Alignment Algorithm

The words in a sequence acquire their specific meaning (among multiple meanings in a reference dictionary) based on the context of the sentence in which the sequence is found. Moreover, variations of this sequence might be observed in the corpus where one or more words in the sequence can be replaced with their synonyms. Given a sequence of words as an example of a rhetorical move, our goal is to find all occurrences of this sequence and its semantically equivalent variations within the corpus in an efficient way. Our overall goal is to use this method for all examples in all rhetorical move types in order to quantify the prevalence of rhetorical moves in each pitch for further quantitative analyses.

Variations of a phrase can automatically be generated by replacing words with their synonyms. For example, if the first, second, and third words in a three-word phrase have three, five, and 10 different synonyms, then one can create $3 \times 5 \times 10 = 150$ different variations. However, only a few of these variations will turn out to be semantically equivalent to the original sequence, due to the diversity of different meanings each word in the sequence might have. It is difficult and time consuming to qualitatively evaluate these variations to find those sequences that are semantically equivalent to the original phrase. The potential dummy words between words in a sequence further complicate the variations.

Again, our goal is to locate all occurrences of a given sequence and its variations in the entire corpus. Although searching for a specific sequence in a corpus is a standard information-retrieval operation, the sheer volume of variations and other complications make it prohibitively expensive to apply a brute-force approach, which is the reason we resort to a more efficient and innovative method like sequence alignment to achieve our goal.

Our algorithm is based on the Smith-Waterman local sequence-alignment algorithm used for genetic sequencing in computational biology (Smith and Waterman 1981). The SW algorithm uses a dynamic programming model, which assigns rewards and penalties based on match locations when comparing two sequences. In our application of this algorithm, we consider a target sequence as the first sequence and the entire corpus as the second sequence in which we want to find local matches. We define a match to be an exact word match (lemma-based) or synonym match. We use *WordNet* to establish synonym-equivalence relationships. The sequence-alignment algorithm evaluates a matrix of rewards and penalties by comparing word-by-word matches (exact or synonym) to determine an alignment score (see Online Appendix A for details). Based on the sequence-alignment calculations, we can determine which sequence variations would indicate a computational highlight for any sentence.

We perform the sequence-alignment computational highlights using each sequence for each rhetorical move for the entire text corpus. Given that the number of words per sequence is typically less than some small constant n_{max} (n-gram size up to 10), the dynamic program for sequence alignment runs in linear time in the size of the entire text corpus of C words. Therefore, given S manually highlighted sequences, the total running time is $O(CS)$. Typically, S is substantially smaller than C, and our sequence-alignment method is an efficient approach to highlighting semantically equivalent phrases for a large text corpus using a relatively small number of manual highlights.

3.4. Rhetoric Instance Vector Output and Analysis

We run our sequence-alignment algorithm to generate computational highlights for the entire corpus for each rhetorical move. This produces rhetoric vectors with the number of instances of each rhetorical move for each unit of text in the decision-making context. A unit of text can be a sentence, a paragraph, a pitch, a forum entry, or a review posting. The rhetoric instance vector characterizes the text based on the rhetorical moves. We can then use this rhetoric mining output to perform a new level of text analytics by examining the relationship of rhetoric instance vectors to other quantitative attributes or decision parameters. We provide an example of this process in detail and then discuss further descriptive analysis of the rhetoric vector output in Section 4.

4. Application of Rhetoric Mining to an Online Investment Community

We apply our rhetoric mining method to analyze stock pitches in an online investment community. Stock pitches have been studied in several text-mining and Information Systems papers as an important application of the influence of online communities and social media in financial decision making (Antweiler and Frank 2005, Hill and Ready-Campbell 2011, Oh and Sheng 2011). We use our rhetoric mining approach to investigate intentional language choices of persuasion to justify and influence investment strategies. In this section, we describe in detail each step of our rhetoric mining method implemented for this particular data set.

4.1. Initial Context Analysis and Rhetorical Move Identification

Motley Fool CAPS is an online community of more than 28,000 investors, who learn investment strategies from each other by sharing their stock predictions. Players predict ("pick") whether a stock will outperform or underperform the market. Each pick has a score to denote whether the pick was correct (positive) or not (negative); the weight of the score is calculated as the difference between the stock value and the S&P market performance. For example, in Figure 3, we see that "Player 01" picked Facebook to underperform the market; however, Facebook outperformed the market by 89.05 points, so the pick score is −89.05. Players are assigned player ratings (on a 100-point scale) based on their overall stock-pick performance, calculated using stock-pick accuracy and overall stock-pick scores. In the example in Figure 3, we see that Player 01 has a player rating of 29.64 (substantially low).

Players can write a "pitch" to justify their stock picks by offering more detailed thoughts, analysis, concerns, or rationale behind their decision. In addition, players may recommend pitches that they agree with or find worthy for other investors to read (in the example in Figure 3, the Facebook pitch by Player 01 received 10 recommendations). Within the community, players with a high number of recommendations are labeled as *Pitch Writing Superstars*. These players are considered influential in the investment community; however, they are not all successful in their stock picks (see Figure 4). Out of the 1,000 Pitch Writing Superstar players, more than 500 had a total of 50 or more recommendations for their pitches, with the top Pitch Writing Superstar earning a cumulative 4,178 recommendations. However, the player ratings for these 1,000 Pitch Writing Superstars

Figure 3. (Color online) Example Stock Pick and Pitch Argument from Motley Fool

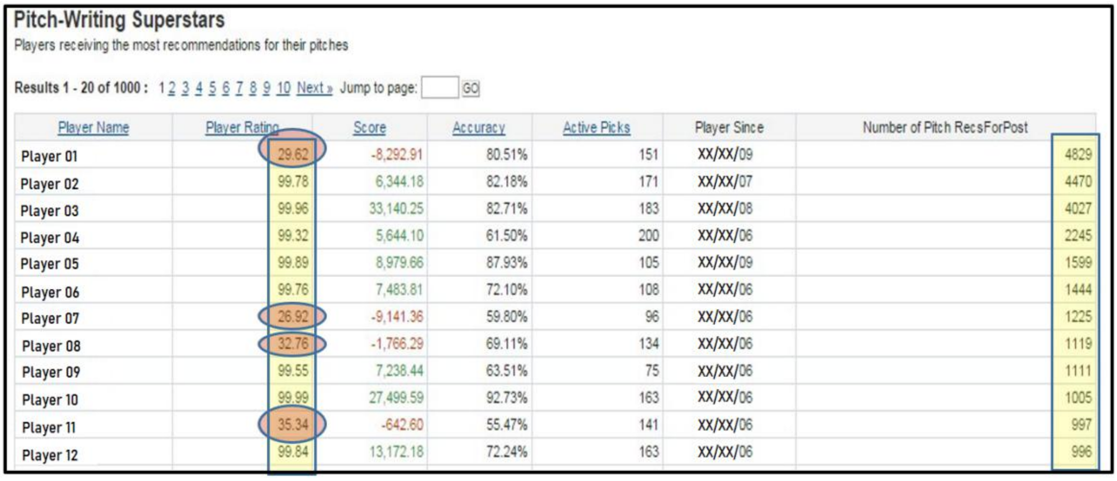
range from 0 to 100, with only 392 "good" players (having a player rating of 80 or above).

We perform an initial context analysis with a random sample of pitches and identify the following seven salient rhetorical moves: *ethos* (either using personal experience (ethos 1) or referencing others' expertise (ethos 2)), *hedging* (expressing either certainty (hedging 1) or uncertainty (hedging 2)), and different types of *evidence* (either product-based (evidence 1), company-based (evidence 2), or stock-based (evidence 3) evidence). These moves were identified, discussed, and agreed upon by two human coders with advanced degrees in rhetoric to develop a protocol with definitions and examples of each move.

4.2. Manual Highlighting and Rhetorical Move Sequence Definition

We used the Motley Fool Application Programming Interface to extract all pitches from all 1,000 players denoted as Pitch Writing Superstars to create an initial data set of 186,043 pitches. The data were collected in June 2015, which includes pitches from April 18, 2006, to June 11, 2015. We randomly selected 200 pitches with 5–15 sentences (for uniform text length). We then randomly selected 1,000 sentences out of all the pitch sentences in this subset. We applied manual highlighting for rhetorical moves on these randomly selected 1,000 sentences of training data. Two human coders initially reviewed a random selection of 50 sentences from this sample to identify salient rhetorical moves and develop a coding protocol. The protocol provides definitions and examples of each move, as well as guidelines for how to record the highlighted phrases. The two coders agreed that the most prominent rhetorical moves were *ethos* moves (drawing on one's personal experience versus referencing the expertise of other CAPS members or outside experts), *hedging* moves (expressing certainty versus admitting uncertainty), and different types of argument *evidence* (product-based, company-based, or stock-based evidence). These moves were grouped into mutually exclusive categories to create seven unique rhetorical moves (shown in Table 2). The coders then compared highlighted sequences from a new set of 50 randomly selected sentences to attain an average Cohen's kappa

Figure 4. (Color online) Pitch Writing Superstars


Pitch-Writing Superstars
Players receiving the most recommendations for their pitches

Results 1 - 20 of 1000 : 1 2 3 4 5 6 7 8 9 10 Next » Jump to page: GO

Player Name	Player Rating	Score	Accuracy	Active Picks	Player Since	Number of Pitch RecsForPost
Player 01	29.62	-8,292.91	80.51%	151	XX/XX/09	4829
Player 02	99.78	6,344.18	82.18%	171	XX/XX/07	4470
Player 03	99.96	33,140.25	82.71%	183	XX/XX/08	4027
Player 04	99.32	5,644.10	61.50%	200	XX/XX/06	2245
Player 05	99.89	8,979.66	87.93%	105	XX/XX/09	1599
Player 06	99.76	7,483.81	72.10%	108	XX/XX/06	1444
Player 07	26.92	-9,141.36	59.80%	96	XX/XX/06	1225
Player 08	32.76	-1,766.29	69.11%	134	XX/XX/06	1119
Player 09	99.55	7,238.44	63.51%	75	XX/XX/06	1111
Player 10	99.99	27,499.59	92.73%	163	XX/XX/06	1005
Player 11	35.34	-642.60	55.47%	141	XX/XX/06	997
Player 12	99.84	13,172.18	72.24%	163	XX/XX/06	996

Table 2. Rhetorical Moves by Category

Rhetoric category	Rhetorical move	Definition
Ethos	*Ethos 1*	Personal expertise
	Ethos 2	Referencing others' expertise
Hedging	*Hedging 1*	Certainty
	Hedging 2	Uncertainty
Evidence	*Evidence 1*	Product-based evidence
	Evidence 2	Company-based evidence
	Evidence 3	Stock-based evidence

score of 83.87% over all moves, indicating strong agreement (Fleiss 1981). Having established interrater reliability, each rater then worked independently to code 450 additional sentences, so that the entire set of 1,000 randomly selected sentences had complete manual highlights identifying examples of all seven rhetorical moves. Our coding scheme is in agreement with recently suggested procedures for intercoder-reliability assessment in the literature (O'Connor and Joffee 2020).

Table 3 provides examples of each of these moves from our sample of 1,000 sentences. These examples also illustrate how the sequences are generated for each move using dummy word notations and word lemmas. Lemmas capture some of the variation created by word inflections: For example, "owned" is an inflection of the lemma "own." The lemma "own" is the root form of the word "owned." For the 1,000 sample sentences with manual highlighting, we recorded 1,375 unique sequences with a total of 3,105 sequence instances (see the online appendix for details).

4.3. Computational Highlighting via Sequence-Alignment Algorithm

To evaluate our sequence-alignment performance, we use leave-one-out (LOO) cross-validation. For the 1,000 sentences in the sample set, we calculate the precision, recall, and F-scores to measure the performance of the computational highlights using sequence alignment compared with the original manual highlights (see the online appendix for details). We achieve very good performance results, with an average F-score greater than 70% for all rhetorical moves (see Table 4 for precision, recall, and F-score results).

We use n-grams as a baseline to compare the performance of our sequence-alignment algorithms. We assume that n-grams require n exact word matches, and we allow up to k nonmatches between any words within an n-gram. For example, the 3-gram "my best investment" can have the following variations with two nonmatches: "my best investment", "my * best investment", "my best * investment", "my * * best investment", "my best * * investment", "my * best * investment", "my * * best * investment", "my * best * * investment", or "my * * best * * investment". Using our original manual highlights from the training set, we repeat our LOO cross-validation for each rhetorical move for n-grams of size n from one to six. Among all variations, our sequence-alignment F-scores outperform even the best n-gram with $k = 1, 2, 3$ (see Table 5).

We further explore how the size of our training set influences sequence-alignment performance. We use bootstrapping to randomly select from 5% to 95% of the 1,000 sentences to serve as the training set. We reserve 5% as the test set and calculate the average F-scores over 20 iterations. We generate learning curves for each rhetorical move in Figure 5, which shows the average F-score for each training-data threshold percentage.

For each move, we calculate its learning-growth rate using a log-linear learning curve (Buntine 1996, Krajewski and Ritzman 2004). The learning rate is calculated as $r = 2^b$, where b is the linear regression slope of the $\log(x)$ and $\log(y)$ values; here, x is the training-data threshold, and y is the average F-score. We observe the training-data threshold point at which the average F-score is consistently greater than 80, which is our baseline for good F-score performance (see the online appendix for more details). From these results, we see that several moves require a relatively small training-data percentage (less than 50%). This implies little variation in the types of sequences used to identify this move. That is, from a small sample of manual highlights, we can accurately tag semantically equivalent phrases. However, other moves with a slower growth rate have significant variation, or several unique phrases, in the sample sequences.

Although F-score performance continues to increase for some moves, most moves achieve an average F-score

Table 3. Rhetorical Move Examples

Rhetorical move	Manual highlights	Sequences generated
Ethos 1	I HAVE OWNED this stock for over a year.	*I * own*
Ethos 2	ACCORDING TO MF, Wall Street ANALYSTS PROJECT about 10% growth for 2010.	*accord to analyst project*
Hedging 1	MS WILL CONTINUE to underperform for AT LEAST the next year.	*be continue at least*
Hedging 2	I''M NOT SURE of the longevity and the profitability of the LCD industry.	*I be not sure*
Evidence 1	Great DEMAND FOR NEW/IMPROVED PRODUCTS.	*demand for * products*
Evidence 2	I think it's a SOLID COMPANY with a good history and seems to BE MANAGED very well.	*solid company be managed*
Evidence 3	On 10/28/2010, GSIT reported 2nd QUARTER 2011 EARNINGS of $0.18 PER SHARE.	*quarter * earnings per share*

Table 4. Precision, Recall, and F-scores from Sequence Alignment

Move	Precision (%)	Recall (%)	F-score (%)
Ethos 1	100.00	76.31	86.56
Ethos 2	97.56	56.74	71.75
Hedging 1	99.65	76.13	86.32
Hedging 2	99.11	77.97	87.28
Evidence 1	100.00	84.65	91.69
Evidence 2	97.38	94.97	96.16
Evidence 3	99.85	94.80	97.26

greater than 80 with a less than 75% training threshold, which represent less than 750 sample sentences from our manually highlighted data. This demonstrates the benefit of our sequence-alignment approach in achieving good performance with a relatively small training set. Unlike n-gram approaches, which depend on dictionary size, sequence alignment requires less training by allowing semantically equivalent highlights through synonym matches. Therefore, the qualitative "set-up cost" of manual highlighting is minimal. For this particular research, it took each human coder an average of 5 seconds per sentence to perform highlights for one rhetoric move. Each coder reviewed a total of 500 sentences for seven moves, with a total time approximation of 17,500 seconds (around 4.8 hours), working in parallel to create all manual highlights. The sequence-alignment algorithm was able to evaluate a single sequence against the entire corpus (5.8 million words) in less than 2 seconds. For a total of 1,375 sequences, it took less than 1 hour to perform all computational highlights. Our computational highlighting software was implemented in Visual Studio 2015 on a machine with four 3.4-GHz Intel Core i7-2600 processors and 8-GB memory running on the Windows 10 operating system.

4.4. Association Analysis of Rhetorical Styles on Persuasiveness and Trustworthiness

After highlighting all moves in all sentences, we created a rhetoric instance vector for each pitch. That is, for each pitch, we recorded the number of instances of each rhetorical move. This allows us to characterize the rhetoric used in each pitch for each stock-pick justification. For example, we might have the rhetoric instance vector for a particular pitch shown in Table 6. This example indicates that personal experience and product-based evidence were used to justify the stock pick. At this point, we have converted the qualitative rhetorical analysis from our initial manual highlighting of 1,000 sentences to a quantitative representation of rhetoric instance vectors for each pitch as a unit of text.

We normalize the rhetorical move frequencies within a pitch with respect to the pitch length to remove length-based bias. The normalized frequencies are typically exponentially distributed over pitches; we therefore use a log transformation to create log-normalized scores for each move in the rhetoric vector. The distribution of the nonzero log-normalized scores follow a bell-curve shape (see Online Appendix D). We analyze the log-normalized rhetoric vectors to define styles based on move prevalence, where each move may be prevalent or not in a given pitch. We therefore convert our log-normalized rhetoric vectors into binary vectors with $0, 1$ values for each move: This creates a total of $2^7 = 128$ possible rhetorical styles.

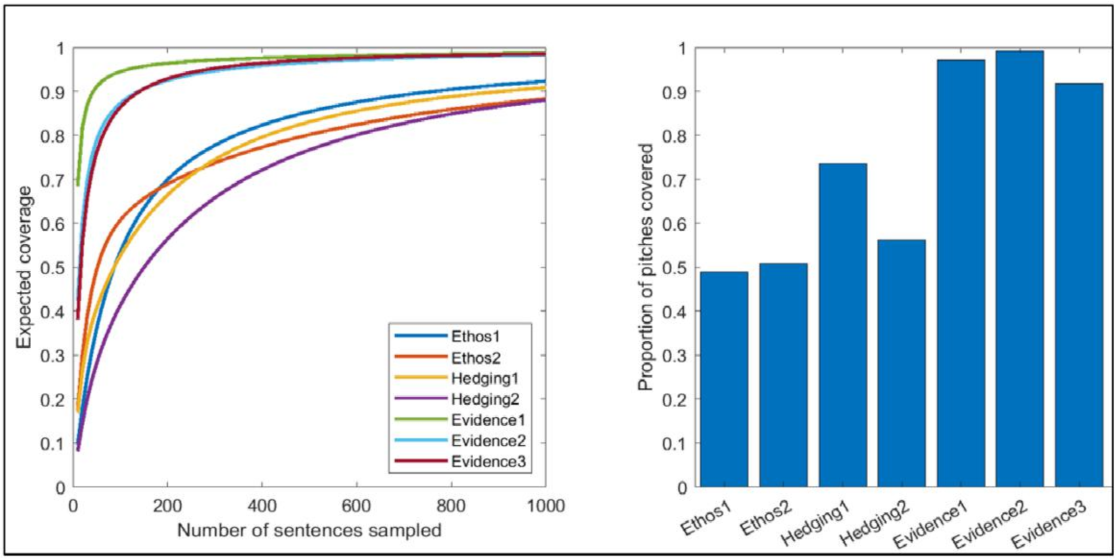
We define *persuasive* pitches to be any pitch that receives at least one recommendation; this results in a subset of 38,894 persuasive pitches. From these, we randomly select 20,000 pitches to represent the persuasive class P and randomly select 20,000 pitches from the remaining data set of pitches with zero recommendations to represent the nonpersuasive class P'. We refer to this as the persuasiveness data set. We define *trustworthy* pitches to be pitches that are written by investors who have a Motley Fool player rating greater than the median rating (good investors). We prepare a data set with 20,000 random pitches from trustworthy players representing the trustworthy class T and 20,000 random pitches from the remaining players representing the nontrustworthy class T'. We refer to this as the trustworthiness data set.

Next, we perform an association analysis between different rhetorical styles and a given class Y of pitches. Based on the bell-shaped distribution of the nonzero log-normalized scores, we consider three thresholds for each move k to determine move prevalence: one at the mean μ_k and two others at the lower and upper tails at a distance of $c\sigma_k$, where σ_k is the standard deviation of

Table 5. Comparison of F-Scores from Sequence Alignment (SA) to n-Gram with up to k Nonmatches

Move	SA (%)	n-gram ($k = 0$) (%)	n-gram ($k = 1$) (%)	n-gram ($k = 2$) (%)	n-gram ($k = 3$) (%)
Ethos 1	86.56	60.44	67.05	68.59	70.69
Ethos 2	71.75	52.91	55.67	55.56	55.28
Hedging 1	86.32	76.32	79.20	78.63	78.25
Hedging 2	87.28	80.44	81.58	81.48	81.13
Evidence 1	91.96	88.67	89.76	89.00	88.42
Evidence 2	96.16	94.74	94.83	94.76	94.58
Evidence 3	97.26	95.51	95.67	95.68	95.37

Figure 5. (Color online) Average F-Score for Increasing Training Data Threshold



Notes. Left: Expected coverage for each move as a function of the number of sentences sampled. Right: Proportion of pitches covered by each move and at least one move.

the scores for move k and c is a *tuning parameter*. We denote these three thresholds to determine prevalence as low ($\mathrm{L} = \mu_k - c\sigma_k$), medium ($\mathrm{M} = \mu_k$), and high ($\mathrm{H} = \mu_k + c\sigma_k$). For a given value of the parameter c, let $\mathcal{T}_c$ be the set of threshold combinations induced by c, where $|\mathcal{T}_c| = 3^7 = 2187$. Let $\mathbf{t}_c \in \mathcal{T}_c$ denote a threshold combination. Let $\mathcal{I}$ denote the set of 128 binary vectors representing the rhetorical styles, $I \in \mathcal{I}$ denote a binary vector for a specific rhetorical style, and $I(\mathbf{t}_c)$ denote the subset of instances whose binary vectors match I with respect to the threshold combination $\mathbf{t}_c$. We denote *support* for $I(\mathbf{t}_c)$ as $\mathrm{Supp}(I(\mathbf{t}_c)) = |I(\mathbf{t}_c)|/N$, where N is the number of all instances. For a target class Y (e.g., persuasiveness), we denote the label of an instance i as $y_i \in \{0,1\}$, where $y_i = 1$ if the instance belongs to that class (e.g., a persuasive pitch), and $y_i = 0$ otherwise. We define confidence for $I(\mathbf{t}_c)$ with respect to the target class Y as

$$\mathrm{Conf}(I(\mathbf{t}_c), Y) = |\{i : i \in I(\mathbf{t}_c), y_i = 1\}|/|I(\mathbf{t}_c)|.$$

Then, using a linear set values $c \in \mathcal{C}$, where $\mathcal{C} = \{0.50, 0.51, \ldots, 2.50\}$, we define the maximum confidence for a value of c with respect to class Y for subsets of $I(\mathbf{t}_c)$ with a minimum support of 1% as

$$\mathrm{MaxConf}(c, Y) = \max_{\mathbf{t}_c \in \mathcal{T}_c, I \in \mathcal{I}} \{\mathrm{Conf}(I(\mathbf{t}_c), Y) : \mathrm{Supp}(I(\mathbf{t}_c)) \geq 0.01\}.$$

Finally, we define the maximum confidence for class Y using a set of values $c \in \mathcal{C}$, where $\mathcal{C} = \{0.50, 0.51, \ldots, 2.50\}$, as

$$\mathrm{MaxConf}(Y) = \max_{c \in \mathcal{C}} \{\mathrm{MaxConf}(c, Y)\},$$

and denote $c^* = \arg\max_{c \in \mathcal{C}} \{\mathrm{MaxConf}(c, Y)\}$ as the optimal tuning parameter for class Y. We refer to the maximum confidence for a specific rhetorical style I with respect to class Y with minimum support of 1% as

$$\mathrm{MaxConf}(I, Y) = \max_{\mathbf{t}_{c^*} \in \mathcal{T}_{c^*}} \{\mathrm{Conf}(I(\mathbf{t}_{c^*}), Y) : \mathrm{Supp}(I(\mathbf{t}_{c^*})) \geq 0.01\},$$

and denote $\mathbf{t}^*_{c^*} = \arg\max_{\mathbf{t}_{c^*} \in \mathcal{T}_{c^*}} \{\mathrm{Conf}(I(\mathbf{t}_{c^*}), Y) : \mathrm{Supp}(I(\mathbf{t}_{c^*})) \geq 0.01\}$ as the optimal threshold combination for style I.

We apply association analysis to the persuasiveness and trustworthiness data sets. Using the persuasiveness data set, we first set the class as $Y = P$ to analyze trustworthy rhetorical styles and, then, $Y = P'$ to analyze nontrustworthy styles. We repeat the same analyses for $Y = T$, the trustworthy, and $Y = T'$, nontrustworthy, styles using the trustworthiness data set. In Tables 7, 8, 9, and 10, we present rhetorical styles with high maximum confidence values and their corresponding optimal threshold combinations for persuasive, nonpersuasive, trustworthy, and nontrustworthy pitches, respectively.

Table 6. Rhetoric Instance Vector Example

Ethos 1	Ethos 2	Hedging 1	Hedging 2	Evidence 1	Evidence 2	Evidence 3
3	0	0	0	0	7	0

Table 7. Rhetorical Styles for Persuasive Pitches

Persuasive														
Style								Threshold						
Ethos 1	Ethos 2	***Hedging 1***	Hedging 2	Evidence 1	Evidence 2	***Evidence 3***	MaxConf (%)	Ethos 1	Ethos 2	Hedging 1	Hedging 2	Evidence 1	Evidence 2	Evidence 3
0	0	1	0	0	0	1	84.1	L	L	M	M	L	L	L
0	0	1	0	0	0	0	81.1	L	L	M	L	L	L	H
0	0	0	0	0	0	1	72.8	L	L	H	L	L	L	L

Note. The bold headings signify columns with non-zero values.

In Table 7, for example, pitches that use a style combining medium levels of Hedging 1 (confidence) with low levels of Evidence 3 (stock-based evidence) are highly likely to receive recommendations. We find similarly that pitches using this same style (with same threshold levels) is highly likely to be written by a good (trustworthy) investor (Table 8). We find even higher confidence levels in detecting nonpersuasive and nontrustworthy pitches containing the styles described in Tables 9 and 10. We have thus used rhetoric mining to identify qualitative rhetorical moves in this online investment community, quantify these moves in each pitch, and analyze rhetorical move vectors over all pitches to identify rhetorical styles indicative of persuasive or trustworthy pitches with very high confidence.

4.5. Other Approaches on Persuasiveness and Trustworthiness

We consider three different approaches from standard text-classification methods to the state-of-the-art methods in natural language processing as a comparison of methods that do not include any rhetorical analysis: (1) frequency-based n-gram features combined with support vector machine (SVM) classifiers, (2) embedded features with recurrent neural networks, and (3) bidirectional encoder representations from transformers. That is, we want to see whether standard machine learning methods, as well as the state-of-the-art deep learning methods for text-mining analysis, can be used to classify persuasive or trustworthy pitches using the persuasiveness and trustworthiness data sets from Section 4.4. For each data set, we first randomly split the set of instance indices into 10 equal-sized partitions $\Pi = \{\Pi_1, \ldots, \Pi_{10}\}$ for cross-validation, where $\Pi_i \cap \Pi_j = \emptyset$ for $1 \le i < j \le 10$ and $I = \cup_{i=1}^{10} \Pi_i$. For fair comparison, for each iteration i of the cross-validation, we use set Π_i as the test set and the remaining instances $\Pi \setminus \Pi_i$ as the training set for all experiment setups in all three approaches.

In approach (1), we use four different types of n-grams sets that gradually include larger n-grams up to 4-grams (i.e., $N = \{\{1\}, \{1,2\}, \{1,2,3\}, \{1,2,3,4\}\}$). For each set, we use four different bounds $B = \{50, 100, 150, 200\}$ for minimum n-gram frequencies to eliminate less frequently used words from the features, which still produces thousands to tens of thousands of features. The resulting combinations create $|N \times B| = 16$ experiment sets, each of which uses the same partition Π for cross-validation for the same data set. We use a linear SVM classifier with lasso regularization as a classifier in these 16 experiment sets. In Table 11, we show the best results for classification accuracy, precision, recall, and F-score for persuasiveness and trustworthiness, each for all n-gram settings in N along with the parameter in B that produced the best result within that n-gram setting for that performance measure.

Table 8. Rhetorical Styles for Nonpersuasive Pitches

Nonpersuasive														
Style								Threshold						
Ethos 1	***Ethos 2***	Hedging 1	Hedging 2	***Evidence 1***	***Evidence 2***	Evidence 3	MaxConf (%)	Ethos 1	Ethos 2	Hedging 1	Hedging 2	Evidence 1	Evidence 2	Evidence 3
1	1	0	0	1	1	0	87.0	L	M	M	L	H	M	H
1	1	0	0	1	1	1	85.4	M	M	L	H	H	M	L
1	1	0	0	1	0	0	84.8	L	M	L	M	H	H	H
1	1	0	0	1	0	1	84.4	L	M	L	L	H	H	L
1	0	0	0	1	1	0	82.5	L	H	L	L	H	M	M

Note. The bold headings signify columns with non-zero values.

Table 9. Rhetorical Styles for Trustworthy Pitches

Trustworthy														
Style								Threshold						
Ethos 1	Ethos 2	***Hedging 1***	Hedging 2	Evidence 1	Evidence 2	***Evidence 3***	MaxConf (%)	Ethos 1	Ethos 2	Hedging 1	Hedging 2	Evidence 1	Evidence 2	Evidence 3
0	0	1	0	0	0	1	87.5	L	L	M	M	L	L	L
0	0	1	0	0	0	0	82.5	L	L	M	M	L	L	M
0	0	0	0	0	0	1	77.5	L	L	H	M	L	L	L
1	0	0	1	0	0	1	75.9	L	M	L	L	M	H	M
1	0	0	1	0	1	1	75.6	L	M	L	L	M	L	M

Note. The bold headings signify columns with non-zero values.

Table 10. Rhetorical Styles for Nontrustworthy Pitches

Nontrustworthy														
Style								Threshold						
Ethos 1	***Ethos 2***	Hedging 1	Hedging 2	***Evidence 1***	Evidence 2	Evidence 3	MaxConf (%)	Ethos 1	Ethos 2	Hedging 1	Hedging 2	Evidence 1	Evidence 2	Evidence 3
1	1	0	0	1	1	0	94.0	L	M	L	L	H	L	H
1	1	0	0	1	0	0	93.7	M	M	L	H	H	H	H
0	1	0	0	1	0	0	93.0	H	M	L	L	H	H	M
1	1	0	0	1	1	1	92.3	L	M	L	L	H	L	L
1	1	0	0	1	0	1	92.3	L	M	L	L	H	H	L

Note. The bold headings signify columns with non-zero values.

In approach (2), we use two different RNN architectures: long short-term memory and bidirectional LSTM as classifiers (i.e., $A = \{\text{LSTM}, \text{BiLSTM}\}$) and two different word-embedding methods: Word2Vect (Abbasi and Chen 2009) and global vectors for word representation (GloVe) (Abbasi et al. 2008) (i.e., $E = \{\text{W2V}, \text{GloVe}\}$), which produce $|A \times E| = 4$ architecture and embedding combinations. For each combination, we use three embedding dimension parameters, $ED = \{100, 200, 300\}$, and the maximum number of epoch parameters, $Ep = \{10, 20, 30\}$, which results in a three-by-three grid of $|ED \times Ep| = 9$ possible settings. Overall, these options result in $|A \times E \times ED \times Ep| = 36$ experiment sets. Because of computation-platform-related limitations, we use pretrained GloVe word embeddings of 100, 200, and 300 dimensions from the Stanford NLP GloVe site (Abbasi and Chen 2009), which are generated by using the Wikipedia2014 and Gigaword5 corpora that have a total of 6 billion tokens and a vocabulary size of 400,000.

For comparability of results, we note that for each of the 36 experiment sets in approach (2), we use the same partition Π as in approach (1) for the same data set. In Table 12, we show the best results for classification accuracy, precision, recall, and F-score for persuasiveness and trustworthiness, each for all architecture and embedding combination in $A \times E$ along with the parameter combination in $ED \times Ep$ that produced the best results within that architecture and embedding combination.

In approach (3), we use BERT, which uses a pretrained, 12-layer model with hidden size of 768 elements to extract feature vectors. These vectors are then used as inputs to train a deep neural network for classification. Again, we use the same partition Π as in approaches (1) and (2) for the same data set for the comparability of results. In Table 13, we show the best results for classification accuracy, precision, recall, and F-score for persuasiveness and trustworthiness using BERT.

5. Conclusions and Future Potential Applications

In this paper, we describe our rhetoric mining methodology via sequence alignment and illustrate its high-performance results and valuable insights through an application of analyzing stock-pick arguments in the Motley Fool online investment community. Our rhetoric mining results show that we can identify rhetorical styles for a given pitch, which, in turn, has very promising results and very high confidence in identifying persuasion or trustworthiness where even the state-of-the-art methods fail to produce good results.

The analysis results illustrate the valuable insight that can be achieved through rhetoric mining. In our application to stock-pitch arguments, we are able to

Table 11. Classification Performance for Best *n*-Gram Parameters

Pitch class	*n*-gram	Accuracy	Min Cnt	Precision	Min Cnt	Recall	Min Cnt	F-score	Min Cnt
Persuasive	1	**0.5844**	100	**0.5948**	200	0.8747	50	0.6619	50
	1,2	0.5641	100	0.5444	100	0.8761	50	0.6642	50
	1,2,3	0.5611	100	0.5402	100	0.9008	200	0.6689	200
	1,2,3,4	0.5484	200	0.5301	300	**0.9474**	100	**0.6728**	100
Trustworthy	1	0.6329	100	0.6412	100	0.7112	50	0.6595	50
	1,2	**0.6378**	150	**0.6533**	50	0.7734	100	0.6770	100
	1,2,3	0.6340	100	0.6156	100	**0.8160**	50	**0.6826**	200
	1,2,3,4	0.6334	200	0.6233	100	0.7509	50	0.6795	150

Notes. The bold values indicate best results. Min Cnt, minimum count.

characterize rhetorical styles of persuasive moves used in investment arguments. Persuasion, which is typically studied qualitatively in small data samples, is quantified and evaluated for a large corpus of text. Considering the rhetorical features of text offers an additional perspective to sentiment, or word or topic, frequency analysis, allowing researchers in Information Systems to answer new social behavior questions. Rather than addressing frequently used words or phrases, common topics, or basic sentiment categories, rhetoric mining allows a researcher to capture more complex language usage in relation to how it influences decision making. Rhetoric mining provides an efficient way to include rhetorical theories (such as argument theory, narrative theory, and persuasion in decision making) in automated large-scale text analysis to achieve truly interdisciplinary research insights.

We believe rhetoric mining can be applied to other business-decision situations in which language is a clear factor in the decision-making process. Rhetoric mining offers a new lens for analyzing persuasion and argumentation in decision making. Rhetoric mining can be used to study business strategies used to influence consumers or other stakeholders, as well as consumer or other stakeholder reactions to decision-making situations.

We believe that the rhetoric mining methodology we outline in Section 3 can be generalized to different text corpora. Our sequence-alignment algorithm is dependent on the sequences generated from the initial manual highlighting. Thus, manual highlighting is key for achieving accurate and consistent results in computational highlighting. However, as we see from our expected coverage (Figure 5), only a relatively small set of text examples is needed for qualitative sequence identification, as our implementation of the sequence-alignment method quickly learns semantically equivalent phrases without extensive tedious training.

Although large amounts of digital text lend themselves well to computational approaches, the use of rhetorical analysis requires attention to context. Fairclough (2003, p. 16) cautioned that "we cannot assume that a text in its full actuality can be made transparent through applying the categories of a pre-existing analytical framework." Most text-mining approaches apply a general tool or set of tools to quickly generate output. However, these approaches require a large cost in interpreting these outputs; that is, once a general tool or set of tools is applied to a new data set, the analysts must spend some time interpreting the output of these tools, so they can understand the potential insight of their analysis results. We believe that this can cause problems in interpretation accuracy and may also result in missing important aspects of the data that may have been ignored in an initial context analysis. In our approach, we argue that it is important to customize the analysis to each discourse community, rather than applying one general set of rhetorical moves to every situation. Even though there is more upfront cost in our approach, we have found that the initial context analysis allows for more easily extractable interpretations of the final results.

One future potential application of rhetoric mining may be in the analysis of online reviews. Such analysis may reveal whether there is a relationship between different rhetorical moves and a positive or negative critique, the price of the product, the average product

Table 12. Classification Performance for Best RNN Parameters

Pitch class	Arch+Emb	Accuracy	(ED, Ep)	Precision	(ED, Ep)	Recall	(ED, Ep)	F-score	(ED, Ep)
Persuasive	LSTM+W2V	0.5659	(100, 30)	0.5701	(100, 30)	0.6602	(200, 20)	0.6017	(200, 20)
	LSTM+GloVe	0.5536	(100, 10)	0.5642	(200, 10)	0.6423	(300, 20)	0.5891	(300, 20)
	BiLSTM+W2V	**0.5664**	(100, 30)	0.5675	(100, 30)	**0.6707**	(300, 20)	**0.6058**	(300, 20)
	BiLSTM+GloVe	0.5631	(300, 20)	**0.5803**	(200, 20)	0.6523	(200, 30)	0.5958	(300, 30)
Trustworthy	LSTM+W2V	0.6680	(300, 10)	0.6763	(100, 20)	0.7326	(100, 30)	0.6853	(100, 30)
	LSTM+GloVe	0.6574	(200, 20)	0.6796	(300, 20)	**0.7706**	(200,10)	0.6859	(200, 10)
	BiLSTM+W2V	**0.6713**	(100, 10)	**0.7345**	(200, 10)	0.7704	(300, 10)	**0.6963**	(300, 10)
	BiLSTM+GloVe	0.6582	(100, 30)	0.6740	(100, 10)	0.7037	(200, 20)	0.6645	(200, 20)

Notes. The bold values indicate best results. Arch, architecture; ED, embedding dimension; Emb, embedding; Ep, epoch.

Table 13. Classification Performance for BERT

Pitch class	Accuracy	Precision	Recall	F-score
Persuasive	0.6049	0.6232	0.5644	0.5924
Trustworthy	0.6788	0.6570	0.5959	0.6249

rating, and so forth. Future work in this area could extend work in consumer-review segment modeling (Archak et al. 2011), opinion-credibility analysis (Mackiewicz 2010, Wang et al. 2011), and user-generated content analysis in social media (Goh et al. 2013, Johnston et al. 2015, Xu and Guo 2018). Rhetorical analysis can also offer new insights into social network analysis of online communities by examining argument styles of leader personas and possible argument mimicking. This type of work could also contribute to research on trust in online communities (Lu et al. 2013).

We also see potential applications of rhetoric mining in comparing persuasive moves used between competitors by measuring their relative influence on consumers. Rhetoric mining could also be applied to the analysis of financial news articles to study their influence on the stock market. Further, rhetoric could contribute to current work examining the relationship between news sentiment and stock market behavior (Xiong and Bharadwaj 2013, Lash and Zhao 2016, Mai et al. 2018) by identifying specific types of intentional language choices made by journalists in framing their news reports. Such research could lead to more studies on the predictive potential of rhetoric mining, adding to work such as that by Oh and Sheng (2011) on predicting stock movement based on stock blog sentiments.

Another application of rhetoric mining may be in the detection of potential auditing issues in managerial accounting or financial reports. Using rhetoric mining to analyze financial reports may reveal a set of patterns and relationships that has not yet been discovered or appropriately emphasized in the business-research literature. In this context, rhetorical analysis may identify markers of power, blame, agency, or obligation, perhaps through genre analysis or ideological analysis. This work could add to research on risk perception based on financial report disclosures (Bao and Datta 2014). Rhetoric mining could be useful in studying fraud detection in insurance claims or tax documentation, possibly extending the work of Abbasi et al. (2008), because false claimants may use distinctive rhetoric. A rhetorical analysis of websites or authors to detect inflammatory speech could add to work such as that by Abbasi and Chen (2009), with future applications in terrorism monitoring to identify unique rhetoric styles to determine authorship or tracing cascades of influence (potentially following replicated rhetoric patterns to their origin or source). Likewise, rhetoric mining can add to current research on deception theory (Glancy and Yadav 2011, Li et al. 2016, Ludwig et al. 2016, Siering et al. 2016) by identifying moves used in deception, rather than just key terms.

Other potential future applications of rhetoric mining in business-decision analysis include assessing what types of rhetorical moves are successful marketing products or acquiring new customers. This analysis could examine online chats or consumer rhetoric in inbound consumer messages (e.g., via email messages or online contact forms) to determine the expected value of the prospect when prioritizing sales leads (Monteserin and Amandi 2015, Han et al. 2018, Ordenes et al. 2019). We could use rhetoric mining to study which communication techniques among members of project teams (e.g., emails from a project manager to team members) are associated with project success or failure; similarly, we could assess which types of private and/or public rhetoric from corporate CEOs are indicative of corporate success. Some of these applications of rhetoric could be mapped to Porter's Five Forces Framework (Porter 2008), to Ives and Learmonth's Customer Resource Life Cycle (Ives and Learmonth 1984), or to other popular business frameworks to identify further gaps and potential areas for application of rhetoric.

We can also use rhetoric mining to study changes in persuasion over time. For example, we could extend the work of Aggarwal and Singh (2013) to examine rhetorical moves used over different stages in a decision-making situation. We can also combine rhetorical analytics with other numerical values related to the decision-making context to develop a complete decision-analytics report. For example, we could consider the relationship between persuasive moves used in marketing materials, total demand, and product pricing to develop a multivariate understanding of both qualitative and quantitative decision variables.

In this paper, we have presented a new analytics perspective of rhetoric mining to recognize intentional language choices used to influence a particular audience. We developed a sequence-alignment method to successfully detect sequences of words that represent different rhetorical moves. We illustrated highly precise performance of our sequence-alignment algorithm with efficient implementation. Rhetoric mining is particularly useful in business-decision analytics to identify types of persuasion used by and on a particular community in a particular decision-making context. Our transformation of the qualitative analysis of persuasion into quantitative instances of rhetorical moves provides a new lens in business-decision analytics. Rhetoric mining has many potential applications in Information Systems research, and our sequence-alignment approach opens the door for radical new studies in rhetoric mining and beyond.

References

Abbasi A, Chen H (2009) A comparison of fraud cues and classification methods for fake escrow website detection. *Inform. Tech. Management* 10(2):83–101.

Abbasi A, Chen H, Writeprints H (2008) A stylometric approach to identity-level identification and similarity detection in cyberspace. *ACM Trans. Inform. Systems* 26(2):7.

Abbasi A, Zhou Y, Deng S, Zhang P (2018) Test analytics to support sense-making in social media: A language-action perspective. *MIS Quart.* 42(2):427–464.

Abrahams A, Jiao J, Wang GA, Fan W (2012) Vehicle defect discovery from social media. *Decision Support Systems* 54:87–97.

Abrahams A, Fan W, Wang GA, Zhang Z, Jiao J (2015) An integrated text analytic framework for product defect discovery. *Production Oper. Management.*26(4):975–990.

Aggarwal R, Singh H (2013) Differential influence of blogs across different stages of decision making: The case of venture capitalists. *MIS Quart.* 37(4):1093–1112.

Althusser L (1971) Ideology and ideological state apparatuses. Althusser L, ed. *Lenin and Philosophy and Other Essays* (Monthly Review Press, New York), 100–140.

Altschul S, Madden T, Schaffer A (1997) Gapped BLAST and PSI-BLAST: A new generation of protain database search programs. *Nucleic Acids Res.* 25(17):3389–3402.

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic molecular alignment search tool. *J. Mol. Biol.* 215(3):403–410.

Andreotta M, Nugroho R, Hurlstone M, Boschetti F, Farrell S, Walker I, Paris C (2019) Analyzing social media data: A mixed-methods framework combining computational and qualitative text analysis. *Behav. Res. Methods* 51(4):1766–1781.

Anthony L, Lashkia G (2006) Mover: A machine learning tool to assist in the reading and writing of technical papers. *IEEE Trans. Professional Comm.* 46(3):185–193.

Antons D, Joshi A, Salge T (2019) Content, contribution, and knowledge consumption: Uncovering hidden topic structure and rhetorical signals in scientific texts. *J. Management* 45(7):3035–3076.

Antweiler W, Frank M (2005) Is all that talk just noise? The information content of Internet stock message boards. *J. Finance* 59(3): 1259–1294.

Archak N, Ghose N, Ghose A, Ipeirotis P (2011) Deriving the pricing power of product features by mining consumer reviews. *Management Sci.* 57(8):1485–1509.

Aristotle (1991) *On Rhetoric. A Theory of Civic Discourse.* Translated by Kennedy G (Oxford University Press, New York).

Backus J, Bauer J, Bauer F, Green F, Green J, Katz J, Katz C et al. (1963) Revised report on the algorithmic language Algol 60. *Comm. ACM* 6:1–17.

Bao Y, Datta A (2014) Simultaneously discovering and quantifying risk types for textual risk disclosures. *Management Sci.* 60(6):1371–1391.

Baroni P, Giacomin P, Giacomin M, Guida G (2005) SCC-recursiveness: A general schema for argumentation semantics. *Artificial Intelligence* 168:(1–2):162–210.

Barrett M, Heracleous L, Walsham G (2013) A rhetorical approach to IT diffusion: Reconceptualizing the ideology-framing relationship in computerization movements. *MIS Quart.* 37(1):201–220.

Berry D (2012) *Understanding Digital Humanities* (Palgrave Macmillan, London).

Bitzer L (1968) The rhetorical situation. *Philosophy Rhetoric* 1(1):1–14.

Blei D, Ng A, Jordan M (2003) Latent Dirichlet allocation. *J. Machine Learning Res.* 3(0):993–1022.

Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, Neelakantan A, et al (2020) Language models are few-shot learners. Preprint, submitted May 28, https://arxiv.org/abs/2005.14165.

Buntine W (1996) A guide to the literature on learning probabilistic networks from data. *IEEE Trans. Knowledge Data Engrg.* 8(2):195–210.

Burke K (1969) *A Grammar of Motives* (University of California Press, Berkeley, CA).

Butler J (1990) *Gender Trouble: Feminism and the Subversion of Identity* (Routledge, London).

Cascavilla G, Conti M, Schwartz D, Yahav I (2015) Revealing censored information through comments and commenters in online social networks. Pei J, Silvestri F, Tang J, eds. *ASONAM'15 Proc. 2015 IEEE/ACM Internat. Conf. Adv. Soc. Networks Anal. Mining* (Association for Computing Machinery, New York), 675–680.

Chai D, Wu W, Han Q, Wu F, Li J (2020) Description based text classification with reinforcement learning. Preprint, submitted February 8, https://arxiv.org/abs/2002.03067.

Chaiken S (1980) Heuristic vs. systematic information processing and the use of source vs. message cues in persuasion. *J. Personality Soc. Psych.* 39(5):752–766.

Chaiken S, Trope Y (1999) *Dual-Process Theories in Social Psychology* (Guilford Press, New York).

Chakrabarti P, Fry M (2022) A mixed-methods framework for analyzing text data: Integrating computational techniques with qualitative methods in demography. *Demographic Res.* 37:1351–1381.

Cheung C, Sia C, Kuan K (2012) Is this review believable? a study of factors affecting the credibility of online consumer reviews from an ELM perspective. *J. Assoc. Inform. Systems* 13(8):2.

Choi J, Cardie C, Riloff E, Patwardhan S (2005) Identifying sources of opinions with conditional random fields and extraction patterns. *HLT'05 Proc. Conf. Human Language Tech. Empirical Methods Natural Language Processing* (Association for Computational Linguistics, Stroudsburg, PA), 355–362.

Choi T, Cheng T, Zhao X, Chan H, Wnag X, Lacka E, Zhang M (2016) A mixed-method approach to extracting the value of social media data. *Production Oper. Management* 25(3):568–583.

Cowie J, Lehnert W (1996) Information extraction. *Comm. ACM* 39(1):80–91.

Craven M, DiPasquo D, Freitag D, McCallum A, Mitchell T, Nigam K, Slattery S (2000) Learning to construct knowledge bases from the World Wide Web. *Artificial Intelligence* 118(1):69–113.

Creswell J, Plano V (2011) *Designing and Conducting Mixed Methods Research*, 2nd ed. (Sage Publications, Los Angeles).

Das S, Chen M (2007) Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Sci.* 53(9):1375–1388.

Devlin J, Chang M, Lee K, Toutanova K (2018) BERT: Pre-training of deep bidirectional transformers for language understanding. Preprint, submitted October 11, http://arxiv.org/abs/1810.04805.

Fairclough N (2003) *Analysing Discourse: Textual Analysis for Social Research* (Routledge, New York).

Fang X, Hu P (2018) Top persuader prediction for social networks. *MIS Quart.* 42(2):1–59.

Feltrim V, Teufel S, Nunes M, Alurisio S (2006) Argumentative zoning applied to critiquing novices' scientific abstracts. Shanahan J, Qu Y, Wiebe J, eds. *Computing Attitude and Affect in Text: Theory and Applications* (Springer, Dordrecht, Netherlands), 233–247.

Feng F, Yang Y, Cer D, Arivazhagan N, Wang W (2020) Language-agnostic BERT sentence embedding. Preprint, submitted July 3, https://arxiv.org/abs/2007.01852.

Ficcadenti V, Cerqueti R, Ausloos M (2019) A joint text mining-rank size investigation of the rhetoric structures of the US presidents' speeches. *Expert Systems Appl.* 123:127–142.

Fleiss J (1981) *Statistical Methods for Rates and Proportions*, 2nd ed. (John Wiley, New York).

Flores F, Ludlow J (1980) Doing and speaking in the office. Fick G, Sprague R, eds. *Decision Support Systems: Issues and Challenges* (Pergamon Press, New York), 95–118.

Gaizauskas R, Wilks Y (1998) Information extraction: Beyond document retrieval. *J. Documentation* 54(1):70–105.

Geertz C (1973) *The Interpretation of Cultures*, 2nd ed. (Basic, New York).

Geisler C (2016) Current and emerging methods in the rhetorical analysis of texts—introduction: Toward an integrated approach. *J. Writing Res.* 7(3):417–424.

Gildea D, Jurafsky D (2002) Automatic labeling of semantic roles. *Comput. Linguistics* 28(3):245–288.

Glancy F, Yadav S (2011) A computational model for financial reporting fraud detection. *Decision Support Systems.* 50(0):595–601.

Glass K, Colbaugh R (2012) Agile detection of framing rhetoric in social media. *2012 IEEE Internat. Conf. Intelligence Security Informatics* (IEEE, Piscataway, NJ), 120–122.

Goh KY, Heng CS, Lin Z (2013) Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content. *Inform. Systems Res.* 24(1):88–107.

Grasso F (2002) Toward computational rhetoric. *Informal Logic* 22(3): 195–229.

Grishman R (1997) Information extraction: Techniques and challenges. Pazienza MT, ed. *Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology* (Springer, Berlin), 10–27.

Habermas J (1984) *Theory of Communicative Action, Volume One: Reason and the Rationalization of Society.* Translated by McCarthy TA (Beacon Press, Boston).

Hamad E, Savundranayagam M, Holmes J, Kinsella E, Johnson A (2016) Toward a mixed-methods research approach to content analysis in the digital age: The combined content-analysis model and its applications to healthcare Twitter feeds. *J. Medical Internet Res.* 18(3):e60.

Han J, Chen Q, Liu J, Luo X, Fan W (2018) The persuasion of borrowers' voluntary information in peer to peer lending: An empirical study based on elaboration likelihood model. *Comput. Human Behav.* 78(January):200–214.

Harris Z (1962) *String Analysis of Sentence Structure* (Mouton, The Hague).

Hill S, Ready-Campbell N (2011) Expert stock picker: The wisdom of (experts in) crowds. *Internat. J. Electronic Commerce* 15(3):73–102.

Hromada D (2011) Initial experiments with multilingual extraction of rhetoric figures by means of PERL-compatible regular expressions. *Proc. Second Student Res. Workshop Associated RANLP 2011* (Association for Computational Linguistics, Stroudsburg, PA), 85–90.

Ihlen Ø, Toledano M, Just S (2021) Using rhetorical situations to examine and improve vaccination communication. *Frontiers Comm.* 6:697383.

Irving R (2004) Plagiarism and collusion detection using the Smith-Waterman algorithm. Technical Report TR-2004-164, University of Glasgow, Glasgow, UK.

Ishizaki S, Kaufer D (2012) Computer-aided rhetorical analysis. McCarthy PM, Boonthum-Denecke C, eds. *Applied Natural Language Processing: Identification Investigation and Resolution* (IGI Global, Hershey, PA).

Ives B, Learmonth G (1984) The information system as a competitive weapon. *Comm. ACM* 27(12):1193–1201.

Johannesson P (2001) A language/action based approach to information modeling. Rossi M, Siau K, eds. *Information Modeling in the New Millennium* (IGI Global, Hershey, PA), 94–109.

Johnson S, Safadi H, Faraj S (2015) The emergence of online community leadership. *Inform. Systems Res.* 26(1):165–187.

Johnston A, Warkentin M, Siponen M (2015) An enhanced fear appeal rhetorical framework: Leveraging threats to the human asset through sanctioning rhetoric. *MIS Quart.* 39(1):113–134.

Jones K (1988) A statistical interpretation of term specificity and its application in retrieval. Willett P, ed. *Document Retrieval Systems Taylor Graham Series in Foundations of Information Science* (Taylor Graham Publishing, London), 132–142.

Józefowicz R, Vinyals O, Schuster M, Shazeer N, Wu Y (2016) Exploring the limits of language modeling. Preprint, submitted February 7, http://arxiv.org/abs/1602.02410.

Kalyan K, Rajasekharan A, Sangeetha S (2021) AMMUS: A survey of transformer-based pretrained models in natural language processing. Preprint, submitted August 12, https://arxiv.org/abs/2108.05542.

Khodabandelou G, Hug C, Salinesi C (2014) A novel approach to process mining: Intentional process models discovery. *2014 IEEE Eighth Internat. Conf. Res. Challenges Inform. Sci. RCIS* (IEEE, Piscataway, NJ), 1–12.

Kim D, Benbasat I (2006) The effects of trust-assuring arguments on consumer trust in Internet stores: Application of Toulmin's model of argumentation. *Inform. Systems Res.* 17(3):286–300.

Kim D, Ferrin D, Rao H (2009) Trust and satisfaction, two stepping stones for successful ecommerce relationships: A longitudinal exploration. *Inform. Systems Res.* 20(2):237–257.

Krajewski L, Ritzman L (2004) *Operations Management: Strategy and Analysis* (Addison-Wesley Publishing Company, Boston).

Krippendorf K (2004) *Content Analysis* (SAGE Publications, Thousand Oaks, CA).

Kupiec J (1992) Robust part-of-speech tagging using a hidden Markov model. *Comput. Speech Language* 6:225–242.

Kuypers J, ed. (2009) *Rhetorical Criticism: Perspectives in Action* (Lexington Books, Lanham, MD).

Lacity M, Janson M (1994) Understanding qualitative data: A framework of text analysis methods. *J. Management Inform. Systems* 11(2):137–155.

Lakeoff G, Johnson M (1980) *Metaphors We Live By* (University of Chicago Press, Chicago).

Lash M, Zhao K (2016) Early predictions of movie success: The who, what, and when of profitability. *J. Management Inform. Systems* 33(3):874–903.

Lassmann T, Sonnhammer E (2005) Kalign—An accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics* 6(1):1–9.

Le Q, Mikolov T (2014) Distributed representations of sentences and documents. Preprint, submitted May 16, http://arxiv.org/abs/1405.4053.

Li W, Chen H, Nunamaker J (2016) Identifying and profiling key sellers in cyber carding community: AZSecure text mining system. *J. Management Inform. Systems* 33(4):1059–1086.

Lia Q, Wanga T, Gonga Q, Chenb Y, Linc Z, Song S (2014) Media-aware quantitative trading based on public web information. *Decision Support Systems* 61(0):93–105.

Liu Q, Karahanna E (2017) The dark side of reviews: The swaying effects of online product reviews on attribute preferences construction. *MIS Quart.* 41(2):427–448.

Liu W, Zhou P, Zhao Z, Wang Z, Ju Q, Deng H, Wang P (2019a) K-BERT: Enabling language representation with knowledge graph. Preprint, submitted September 17, http://arxiv.org/abs/1909.07606.

Liu X, He P, Chen W, Gao J (2019b) Multi-task deep neural networks for natural language understanding. Preprint, submitted January 31, http://arxiv.org/abs/1901.11504.

Loughran T, McDonald B (2011) When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *J. Finance* 66(1):35–65.

Lowry P, Moody G, Vance A, Jensen M, Jenkins J, Wells T (2012) Using an elaboration likelihood approach to better understand the persuasiveness of website privacy assurance cues for online consumers. *J. Assoc. Inform. Sci. Tech.* 63(4):755–766.

Lu Y, Jerath K, Singh P (2013) The emergence of opinion leaders in a networked online community: A dyadic model with time dynamics and a heuristic for fast estimation. *Management Sci.* 59(8):1783–1799.

Ludwig S, Laer TV, Ruyter K, Friedman M (2016) Untangling a web of lies: Exploring automated detection of deception in computer-mediated communication. *J. Management Inform. Systems* 33(2): 511–541.

Lyytinen K (1985) Implications of theories of language for information systems. *MIS Quart.* 9(1):61–74.

Mackiewicz J (2010) The co-construction of credibility in online product reviews. *Tech. Comm. Quart.* 19(0):403–426.

Mai F, Shan Z, Bai Q, Wang X, Chiang R (2018) How does social media impact Bitcoin value? A test of the silent majority hypothesis. *J. Management Inform. Systems* 35(1):19–52.

Markle-Hub J, Feueeriegel S, Prendinger H (2017) Improving sentiment analysis with document-level semantic relationships from rhetoric discourse structures. *Proc. 50th Hawaii Internat. Conf. System Sci.* (University of Hawai'i at Mānoa, Honolulu, HI), 1142–1151. https://www.research-collection.ethz.ch/handle/20.500.11850/182422.

McClosky D (1985) *The Rhetoric of Economics.* (University of Wisconsin Press, Madison, WI).

Measell JS (1976) Rhetorical criticism: Genre criticism—counter statements: Whither genre? (or, genre withered?). *Rhetoric Soc. Quart.* 6(1):1–4.

Menon V, Sabdhi M, Harikumar K, Soman K (2018) Semantic analysis using pairwise sentence comparison with word embeddings. Thampi S, Mitra S, Mukhopadhyay J, Li KC, James A, Berretti S, eds. *Intelligent Systems Technologies and Applications* (Springer, Cham, Switzerland), 268–278.

Mikolov T, Sutskever I, Chen K, Corrado G, Dean J (2013) Distributed representations of words and phrases and their compositionality. Burges CJ, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, eds. *NIPS'13 Proc. 26th Internat. Conf. Neural Inform. Processing Systems*, vol 2 (Curran Associates Inc., Red Hook, NY), 3111–3119.

Miller C (1990) The rhetoric of decision science, or Herbert A. Simon says. Simon H, ed. *The Rhetorical Turn* (University of Chicago Press, Chicago), 162–184.

Monteserin A, Amandi A (2015) Whom should I persuade during a negotiation? An approach based on social influence maximization. *Decision Support Systems* 77(0):1–20.

Mount D (2004) *Bioinformatics: Sequence and Genome Analysis*, 2nd ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Needleman S, Wunsch C (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48(3):443–453.

Neuendorf K (2002) *The Content Analysis Guidebook* (Sage Publications, Thousand Oaks, CA).

Notredame C, Higgins D (1996) SAGA: Sequence alignment by generic algorithm. *Nucleic Acids Res.* 24(8):1515–1524.

O'Connor C, Joffee H (2020) Intercoder reliability in qualitative research: Debates and practical guidelines. *Internat. J. Qualitative Methods* 19:1–13.

Oehmen C, Nieplocha J (2006) ScalaBLAST: A scalable implementation of BLAST for high-performance data-intensive bioinformatics analysis. *IEEE Trans. Parallel Distrib. Syst.* 17(8):740–749.

Oh C, Sheng O (2011) Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement. *Proc. 32nd Internat. Conf. Inform. Systems*. https://www.researchgate.net/publication/221599803_Investigating_Predictive_Power_of_Stock_Micro_Blog_Sentiment_in_Forecasting_Future_Stock_Price_Directional_Movement.

O'Halloran K, Tan S, Pham DS, Bateman J, Moere A (2016) A digital mixed methods research design: Integrating multimodal analysis with data mining and information visualization for big data analytics. *J. Mixed Methods Res.* 12(1):11–30.

Oinas-Kukkonen H, Harjumaa M (2009) Persuasive systems design: Key issues, process model, and system features. *Comm. Assoc. Inform. Systems* 24(1):485–500.

Ordenes F, Grewal D, Ludwig S, Ruyter K, Mahr D, Wetzels M (2019) Cutting through content clutter: How speech and image acts drive consumer sharing of social media brand messages. *J. Consumer Res.* 45(5):988–1012.

Parks L, Peters W (2022) Natural language processing in mixed-methods text analysis: A workflow approach. *Internat. J. Soc. Res. Methodology*, ePub ahead of print January 12, https://doi.org/10.1080/13645579.2021.2018905.

Pendra N, Cotos E (2008) Automatic identification of discourse moves in scientific article introductions. *Proc. Third Workshop Innovative Use NLP Building Ed. Appl. Assoc. Comput. Linguistics* (Association for Computational Linguistics, Stroudsburg, PA), 62–70.

Pennington J, Socher R, Manning C (2014) GloVe: Global vectors for word representation. *Proc. 2014 Conf. Empirical Methods Natural Language Processing EMNLP* (Association for Computational Linguistics, Stroudsburg, PA), 1532–1543.

Peters M, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, Zettlemoyer L (2018) Deep contextualized word representations. Preprint, submitted February 15, http://arxiv.org/abs/1802.05365.

Petty R, Cacioppo J, Goldman R (1981) Personal involvement as a determinant of argument-based persuasion. *J. Personality Soc. Psych.* 41(5):847–855.

Pollach I (2012) Taming textual data: The contribution of corpus linguistics to computer-aided text analysis. *Organ. Res. Methods* 15(2):263–287.

Porter M (1980) An algorithm for suffix stripping. *Program* 14(3):130–137.

Porter M (2008) The five competitive forces that shape strategy. *Harvard Bus. Rev.* 86:79–93.

Punyakanok V, Roth D, Yih W (2008) The importance of syntactic parsing and inference in semantic role labeling. *Comput. Linguistics* 34(2):257–287.

Rittgen P (2006) A language-mapping approach to action-oriented development of information systems. *Eur. J. Inform. Systems* 15(1):70–81.

Robert L, Dennis A, Hung Y (2009) Individual swift trust and knowledge-based trust in face-to-face and virtual team members. *J. Management Inform. Systems* 26(2):241–279.

Sager N, Friedman C, Lyman M (1987) *Medical Language Processing: Computer Management of Narrative Data* (Addison-Wesley, Boston).

Searle J (1969) *Speech Acts: An Essay in the Philosophy of Language* (Cambridge University Press, Cambridge, UK).

Siering M, Koch J, Deokar A (2016) Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts. *J. Management Inform. Systems* 33(2):421–455.

Simons H, ed. (1990) *The Rhetorical Turn: Invention and Persuasion in the Conduct of Inquiry* (University of Chicago Press, Chicago).

Smith T, Waterman M (1981) Identification of common molecular subsequences. *J. Mol. Biol.* 147:195–197.

Stone P, Dunphy D, Ogilvie D (1966) *The General Inquirer: A Computer Approach to Content Analysis* (MIT Press, Cambridge, MA).

Taboada M, Mann W (2006) Applications of rhetorical structure theory. *Discourse Stud.* 8(4):567–588.

Teufel S, Moens M (2002) Summarizing scientific articles: Experiments with relevance and rhetorical status. *Comput. Linguistics* 28(4):409–445.

Thoppilan R, Freitas DD, Hall J, Shazeer N, Kulshreshtha A, Cheng H, Jin A, et al. (2022) LaMDA: Language models for dialog applications. Preprint, submitted January 20, https://arxiv.org/abs/2201.08239.

Toulmin S (1958) *The Uses of Argument* (Cambridge University Press, Cambridge, UK).

Toutanova K, Manning C (2000) Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. *Proc. Joint SIGDAT Conf. Empirical Methods Natural Language Processing Very Large Corpora* (Association for Computational Linguistics, Stroudsburg, PA), 63–70.

Venkatesh V, Brown S, Bala H (2013) Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Quart.* 37(1):21–54.

Venkatesh V, Brown S, Sullivan Y (2016) Guidelines for conducting mixed-methods research: An extension and illustration. *J. Assoc. Inform. Systems* 17(7):435–495.

Wang G, Liu X, Fan W (2011) A knowledge adoption model based framework for finding helpful user-generated contents in online communities. *Internat. Conf. Inform. Systems*, 1–11. https://www.researchgate.net/publication/221599363_A_Knowledge_Adoption_Model_Based_Framework_for_Finding_Helpful_User-Generated_Contents_in_Online_Communities.

Wang GA, Wang HJ, Li J, Abrahams AS, Fan W (2015) An analytical framework for understanding knowledge-sharing processes in online Q&A communities. *ACM Trans. Management Inform. Systems* 5(4):18.

Wang Q, Liu P, Zhu Z, Yin H, Zhang Q, Zhang L (2019) A text abstraction summary model based on BERT word embedding and reinforcement learning. *Appl. Sci.* 9(21):4701.

Wei C, Hu P, Tai C, Huang C, Yang C (2008) Managing word mismatch problems in information retrieval: A topic-based query expansion approach. *J. Management Inform. Systems* 24(3):269–295.

Weick K (1995) *Sensemaking in Organisations* (Sage, London).

Wilson T, Hoffmann P, Somasundaran S, Kessler J, Wiebe J, Choi Y, Cardie C, Riloff E, Patwardhan S (2005) Opinionfinder: A system for subjectivity analysis. *Proc. HLT/EMNLP 2005 Interactive Demonstrations* (Association for Computational Linguistics, Stroudsburg, PA), 34–35.

Wojcik M (2011) Inventing computational rhetoric. PhD thesis, Michigan State University.

Xiong G, Bharadwaj S (2013) Asymmetric roles of advertising and marketing capability in financial returns to news: Turning bad into good and into great. *Marketing Sci.* 75(5):88–104.

Xu Z, Guo H (2018) Using text mining to compare online pro- and anti-vaccine headlines: Word usage, sentiments, and online popularity. *Comm. Stud.* 69(1):103–122.

Zagalsky A, Te'eni D, Yahav I, Schwartz DG, Silverman G, Cohen D, Mann Y, Lewinsky D (2021) The design of reciprocal learning between human and artificial intelligence. *Proc. ACM Human-Computer Interaction* 5(CSCW2):443.

Zhang K, Zhao S, Cheung C, Lee M (2014) Examining the influence of online reviews on consumers' decision-making: A heuristic-systematic model. *Decision Support Systems* 67:78–89.

Zhou P, Qi Z, Zheng S, Xu J, Bao H, Xu B (2016) Text classification improved by integrating bidirectional LSTM with two-dimensional max pooling. *Proc. COLING 2016 26th Internat. Conf. Comput. Linguistics Tech. Papers* (The COLING 2016 Organizing Committee, Osaka, Japan), 3485–3495.