Having A Way with Words: Innovations and Improvements in Text Analysis Methods

Submission Deadline: August 31, 2024

Guest Editors

Jason Kiley (Clemson University)
Aaron McKenny (Indiana University)
Jeremy Short (University of North Texas)
Anne Smith (University of Tennessee)

In recent decades, the increased availability of digital texts and other narratives, coupled with programming tools such as web scraping and APIs, has enabled scholars to study organizational phenomena in new ways. As a result, techniques such as content analysis, computer-aided text analysis, topic modeling, document classification, and automated aspects in computer-assisted qualitative data analysis have emerged as valuable tools to make sense of such data.

For quantitative researchers, the ability to measure organizational constructs from text data opens new doors for unobtrusive measurement. For qualitative researchers, text analysis offers a potentially valuable tool for exploring and understanding text data to complement the suite of interpretive tools already used. With the help of these new tools, researchers can now analyze the large volumes of organizational text data in ways that were not possible in the past.

The goals of this feature topic are four-fold. First, we aim to encourage the introduction and dissemination of new and innovative techniques that have the potential to inform organizational research, enabling researchers to provide stronger tests of existing research questions and opening the door to research questions that are difficult to test with current tools. For example, previous work has introduced innovations in text analysis such as topic modeling (e.g., Schmiedel, Müller, & Vom Brocke, 2019), transformer-based text classification (e.g., Fyffe, Lee, & Kaplan, 2023), and CATA dictionary development (e.g., Short et al., 2010) to organizational research, demonstrating how each improves upon current practices in text analysis.

Second, we aim to better understand the limitations of current text analytic approaches and critically evaluate the application of, tools employed in, and key decisions affecting the appropriateness and rigor of text analyses in management research. We also need to evaluate the extent to which researcher decisions can ameliorate these limitations. For example, previous studies have identified measurement error as a threat to research using dictionaries in computer-aided text analysis and identified procedures for estimating and reducing such error (e.g., McKenny et al., 2016). Similarly, whereas topic modeling is growing in popularity in management research, key researcher decisions affect the interpretability of the uncovered topics. Schmiedel and colleagues (2019) outline these important decisions and provide researchers with guidance regarding how to make them. While text analysis algorithms and tools are frequently discussed, it is also important to understand the appropriateness of other inputs to text analysis. For instance, the collection and transformation of text data for analysis is a key input into text analysis (Hickman et al., 2022). How should researchers select, source, and preprocess these texts to enhance the quality of insights drawn from their analysis?

Third, we welcome articles that take innovations in text analysis from other domains, adapt them for
organizational research, and evaluate their promise and limitations for developing and testing management theory. For example, Arseniev-Koehler (2022) documents how word embeddings developed by researchers in computational linguistics can be used in cultural sociology to model the meaning of words in text, highlighting both the opportunities for and the limitations of these embeddings for understanding meaning from a sociological lens. Because text analysis techniques employed from other fields are often evaluated using criteria not catered to organizational research, we need similarly evaluate the applicability of novel text analytic techniques for organizational research as well. For instance, how can organizational researchers use word embeddings or other text analysis innovations from other fields to improve our research? What adaptations must be made to improve their relevance to our context? And what limitations must we attend to when employing these innovative approaches?

Fourth, this feature topic offers an opportunity to consider how text analysis might be used by an interpretive researcher, not to turn text into numbers, but as an option to assist in the intensive work of qualitative data analysis. Interpretive research has a rich tradition of drawing insights about organizations and their stakeholders from text data such as interviews, observation notes, documents, and other secondary sources. Text analysis techniques need not diminish the critical role of the interpretive researcher in coding, pattern identification, and insight creation. In an early use of text analysis in management research labeled “interpretive text analysis”, Geiphart (1997: 583) undertook a detailed analysis of the differences in how managers and regulators made sense of risk, using their testimony in a public inquiry of a sour gas blow up. O’Kane, Smith, and Lerman (2021) also noted how some existing dictionaries might be used to replicate patterns from open coding, as a supplemental analysis during qualitative data analysis. Finally, topic modeling can generate patterns from textual data, perhaps to see how topics shift over time or across industries (e.g., Hannigan et al., 2019). This inductive topic modeling approach will not “spit out” an answer but relies on researchers to make sense of the topics and their relation to the overall research question. We welcome submissions that carefully examine how innovations in text analysis can complement the researcher’s primary interpretive role in making sense of qualitative data. Text analysis tools are just that: tools in the service of furthering potential insights into qualitative data. Submissions in this fourth approach should also attend to the limitations of these techniques and discuss contexts in which computer-assisted text analyses are inappropriate.

Text analytic techniques have been used in a variety of contexts including entrepreneurship (Anglin et al., 2023), strategic management (Harrison et al., 2019), and organizational psychology and organizational behavior (Short et al., 2018). As such, we welcome submissions across macro, meso, and micro domains.

To help methods scholars respond to our call for research, we offer a few possible topics and potential questions. These include, but are not limited to the following suggestions:

- Text analysis in computational linguistics are often evaluated using metrics of predictive accuracy and perplexity. Organizational research often considers a broader range of validity and reliability considerations. How do text analytic techniques perform on these validity and reliability metrics and how can text analyses be improved to meet the validity and reliability needs of organizational research?
- Can new machine learning/AI techniques help test novel hypotheses? If so, how should organizational researchers address the ‘black box’ problem commonly associated with these techniques?
- What techniques common in other fields (economics, psychology, communications, linguistics) might be adapted to provide new insights for organizational scholars in entrepreneurship, strategic management, organizational behavior, and other organizational fields of study?
- Are traditionally inductive text analytic techniques such as topic modeling and deductive text analytic techniques such as text classification potentially useful but underutilized in deductive/inductive research? If so, how must they be adapted, how should their outputs be interpreted differently, and what are the limitations to their application in this new context?
• How might recent innovations in text analysis be of use to qualitative researchers without violating the primary role of the researcher in data sensemaking? What are pragmatic ways to consider text analysis approaches without violating the researcher’s primacy in data analysis or interpretive stance? facilitate interpretive research? What are the limits of their applicability and in what contexts should these techniques be avoided?

• How can text analysis be integrated with other research techniques (e.g., surveys, experiments, grounded theory, case analysis) to offer a new synthetic method with promise to uncover new insights/stronger tests of theory that would not be possible by either method in isolation?

• What are the limitations of the tools and techniques currently used by organizational scholars to collect and process text data and how do these limitations affect insights derived from text data? How can these tools and techniques be adapted to improve the quality of organizational research?

• In what ways is text analysis currently being applied uncritically, inappropriately, or incorrectly? How do these applications of text analysis affect our understanding of organizational phenomena and what should organizational scholars do to avoid these mistakes?

• What text analysis practices are currently common but no longer apply, need revision, or even lead to problematic insights in organizational research? What should organizational scholars do differently and what are the implications of these changes for the development and testing of organizational theory?

References


Schmiedel, T., Müller, O., & Vom Brocke, J. (2019). Topic modeling as a strategy of inquiry in organizational
research: A tutorial with an application example on organizational culture. Organizational Research Methods, 22: 941-968.


