10. More Than Words: A Text Mining Approach to the Analysis of Topics and Skills in Technical Writing Job Ads

Bremen Vance
MERCER UNIVERSITY

Erica M. Stone
INDEPENDENT RESEARCHER

Abstract: Technical writing work can vary significantly in its content, skills, and goals. This study examines patterns in the language of recent technical writing job advertisements as a way of assisting efforts of job seekers and educators related to professional development and curricula. Job advertisements were collected from Indeed each month for a year across the ten most populated cities in the US. We created a corpus of 4,597 unique advertisements using the search terms *Technical Writing* and *Technical Writer* posted during this time frame. We analyzed this corpus using distant reading and text as data methodologies. We present the most frequently used words used in technical writing job ads, and we examine several topics of interest to get a sense of how frequently (and infrequently) particular words are used. We also identify several topics representing the industries and roles that technical writers work in. The patterns in the corpus have several implications for professional development, higher education, and job seekers. This study provides an entry point into examining the expectations and perceptions of work in the field of technical writing, and it highlights several opportunities for reflecting on and studying the practices, priorities, and opportunities in technical writing.

Keywords: Job Ads, Text mining, Topic Modeling, Professional Development, Technical Writing

For us, as practitioners and educators in technical and professional communication (TPC), articulating what we do has significant implications. Definitions, as turns out, matter. Because the work of a communicator is varied and context-specific, the boundaries of our work are far from definite. Concrete deliverables might appear deceptively simple when we do our jobs well, but the work and value of technical communication goes beyond the physical outputs. With every project, and in each context, the subject, form, and value of our work may look different. The situational differences between roles and projects can make it difficult to judge how the professional communicator fits in with project dynamics.

The problem is nothing new. Over three decades ago, Jo Allen (1990) described the attempts and failures of the Society for Technical Communication

(STC) to define technical writing. Allen's argument came after an incident in which a cookbook was disqualified from a technical-publications competition because the judges determined it did not meet their definition of *technical writing*. Allen also pointed out that the year before, the STC board had abandoned their attempt at developing an official definition of *technical writing* after members could not agree on a definition. In the end, Allen (1990) cautioned, "we should be careful, in our earnest desire to create a definition, not to exclude or disenfranchise writing that falls outside our strict categories" (p. 75). For Allen, the biggest problem is the way definitions draw firm lines that determine what is and isn't technical writing, and the firm boundaries lead to problems. Allen suggested that "we should reconsider works…that fall so naturally in line with almost all the criteria we claim for technical writing before we exclude these works for violating a single criterion" (p. 75). Allen's point is that definitions are a statement about what is valued in our field, and it can lead us to ignore or overlook otherwise important work.

In the years since Allen's argument, what it means to be a writer and a communicator has gone through significant changes. Perhaps unsurprisingly, scholars and practitioners find it challenging at times to describe the focus and value of the field (Batova et al., 2016; Henning & Bemer, 2016; Kimball, 2016; Lauer & Brumberger, 2016; Melonçon & Schreiber, 2022; Rice-Bailey, 2016; Shalamova et al., 2019; Swarts 2012). Each attempt by scholars and practitioners to articulate a vision of the field has its value, often providing readers the opportunity to consider new ways to apply their skills, opportunities to ask new questions, and arguments for the value of writing and communication work.

Keeping pace with the historical context, recent progress, and emerging trends of our field presents a formidable task. Perspective accumulated through years, or even decades, of experience in TPC can provide valuable insights, but rapid growth and change in the field necessitates continuous learning and adaptation. The challenge lies not only in monitoring new developments but also in integrating them beyond our areas of specialization within the vast scope of TPC — a feat often unrealistic. Derek Mueller (2017) has referred to the challenge of keeping pace with a field as the "reading problem," explaining, "this ongoing condition—the field's perpetually being written—means that more disciplinary material is generated than any one person reading by conventional strategies alone could reasonably handle" (p. 21). Of course, most members of our field specialize in a few areas, developing their own sense of what it means to be a technical communicator. In a sense, each of us will develop our own working understanding of what technical communicators have to offer.

Still, we often ask questions like: What kinds of skills, tools, and technologies are most desired? How do companies define the job and responsibilities of communicators in a job posting? How can educators train students to meet the expectations of an ever-changing job market? These types of questions have motivated several studies over the past few decades of TPC scholarship (Coffelt

et al., 2022; Lanier, 2009; Rainey et al., 2005; Rosselot-Merritt, 2020; Rude, 2009; Shalamova et al., 2019; Spyridakis, 2015; Stanton, 2017). These studies strive to use observations about the prevalence and importance of theories, competencies, skills, processes, tools, genres, and more to orient our attention towards ideas that are the most relevant to various goals. Studies that attempt to determine what is most common, most needed, and most valuable are useful because they examine and contextualize TPC work.

Ultimately, TPC is a dynamic and engaging field which led Eva Brumberger and Claire Lauer (2020) to observe that "technical and professional communicators must now acquire broader skill sets than in the past" (p. 1) an observation that motivated their work developing data-driven personas of TPC practitioner roles. The broad and evolving nature of TPC has also led to questions regarding its value and legitimacy.

The ways we explain our field to managers, colleagues, clients, students, and each other can have significant implications. Jeremy Rosselot-Merritt explained that "perceptions of TPC among non-TPC professionals reveal accurate conceptions of the field on a basic level; however, those conceptions lack the depth needed for an accurate understanding in a broader sense" (2020, p. 57). Rosselot-Merritt goes on to suggest that poor understanding of the value that TPC work brings is an existential problem based on two key realities: TPC is dependent on collaboration with other disciplines and decisions about technical communicator's roles and responsibilities are often made by stakeholders without a TPC background.

Like other scholars, our desire to orient ourselves to what matters most in our field has led us to an empirical study of job advertisements (ads) that would produce insights into some fundamental questions. TPC is a broad field and each facet is worth investigating, but in practice 'technical communicator' is a broad term. Very few job ads use the general term, offering more specific titles.

Why Technical Writing?

While the broader language of *technical and professional communication* or *technical communicator* are our preferred labels, the job titles and specializations that qualify as TPC work vary broadly. For example, Brumberger and Lauer (2020) used 55 job titles to collect job ads after reviewing the Society for Technical Communication (STC) job board and other sources, and they go on to remark after conducting interviews that job titles can be "fluid" and "arbitrary" (p. 7). Instead of attempting to study all TPC roles simultaneously based on the job title, we decided to collect job ads based on search results of a job board. We believe this helps ensure the collection represents the language job seekers would encounter when seeking employment as a technical writer.

For this study, we narrowed our focus to the way technical writing is represented in job ads. During a 12-month period, we developed a corpus of 4,597

unique job ads using the search terms *technical writer* and *technical writing*. As a genre that describes the work of our field, job ads themselves don't completely reflect what technical writers do, but rather what hiring managers *think* they can do. Our corpus reflects the way specific industries value technical writing, and it can potentially highlight the disconnects between what we actually do, what we think we do, and how we're being described.

We choose to analyze job ads because they, in part, directly define what it means to be a technical communicator. Individual job ads are limited, but the descriptions offered in them represent real attempts at explaining the necessary skills, experience, and responsibilities that are expected. This reveals how hiring managers (or sometimes, Human Resources) see technical writing, at least as it relates to their needs. Job ads also allow us to see TPC from the vantage point of a practitioner looking for work because we are analyzing the language they will encounter that most meaningfully explains what technical writers do.

A Text Mining Approach

In this article, we analyze the language employers use in communication job ads to identify useful patterns. We use text mining and analysis methods that are common in computer science, machine learning, and natural language processing. Text mining refers to computer-aided methods that analyze large collections of text. More specifically, we treat text as data following approaches described by Justin Grimmer et al. (2022). Grimmer et al. explained the reason, writing, "forcing researchers to use data to test theories that were developed before the data arrived also has substantial weakness. Scholars in the social sciences have acknowledged the importance of more inductive forms of analysis in qualitative research including full cycle research design, grounded theory, and nested analysis" (2022, p. 14). The goals of text mining are to identify patterns of facts through systematized observations, and the careful analysis of the patterns will lead to interpretations that can be further tested.

These methods resonate with the distant reading techniques introduced by Franco Moretti (2013) as an alternative to close reading and further elaborated by Mueller (2017) for investigating disciplines. Distant reading prioritizes the analysis of patterns across a body of literature instead of focusing on an individual passage, allowing for conversations about the features and relationships across a network of texts. Distant reading encourages a relationship with text that is critical and analytical, but it recognizes that viewing text in context of related materials can provide useful context.

A text mining approach is valuable when researchers have a broad interest in an area of inquiry but do not want to apply pre-determined assumptions. The methods we are working from are exploratory and iterative in nature, using abductive reasoning which is a logical approach that "involves an inference in which the conclusion is a hypothesis that can be tested with a new or modified research

design" (Ignatow & Mihalcea, 2018, p. 51). While many data-driven studies often start with specific research questions, text mining research often avoids predetermined assumptions about what appears in the collection. Researchers work, instead, to carefully report the data, observing patterns of interest, attempting to explain and formulate new questions. The findings and interpretation often result in new, testable hypotheses that are grounded in the data. Our goal was to uncover patterns and features within technical writing descriptions that represent the field systematically and in a way that can be reproduced and expanded on. The identified patterns and insights reflect the needs and expectations of managers and organizations hiring technical writers.

To help situate this research, we first discuss the complications of explaining the work of technical writers. The history of scholars discussing TPC's boundaries, key topics, and questions illustrates why distant reading and text as data methods can provide valuable perspective. We take a look at this history to provide context and to suggest that scholarship that maps the field in TPC helps the field reflect and grow. We argue that text mining compliments these efforts. Then, we outline our method and present comprehensive data sets from technical writing job ads. We identify several patterns in the data that add to previous scholars' efforts to review the field. Through sharing our findings and discussing various interpretations, we suggest several perspectives and questions that merit further exploration.

▮ Literature Review

Demonstrating the value, variability, and complexity of TPC work in a way that is both illustrative and productive is no simple task. In a faculty meeting, a senior administrator recently remarked that technical communicators are 'the ones who write instructions for how to use the toaster.' This accurate, but diminutive, explanation illustrates the problem of explaining TPC. Similarly, a former graduate student recently remarked, "I thought technical writing was just making things clearer and easier to understand, but now that I'm actually doing the work, I see that it's so much more complex."

At times we must contend with popular, reductive, descriptions of the work we do, which as Rosselot-Merritt demonstrated, may "reveal accurate conceptions of the field on a basic level…' (2020, p. 57). A layman understanding of technical communication is quick to grasp, but it hides the difficulty, diversity, and impact of our field. Easy definitions like "writing simple instructions" or "translating complicated topics" may be accurate, but they can also significantly reduce the sense of complexity and value that our work involves, a problem that is not unique to TPC.

We have often used the definition of *technical communication* offered by STC, which includes three categories: 1) "Communicating about technical or specialized topics," 2) "Communicating by using technology," and 3) "Providing

instructions about how to do something" (n.d.). While this definition is both more accurate and adds clarity, the variety of work that is covered by these areas is substantial and often dependent on industry-specific norms. Within each of the three categories, we could quickly add a variety of roles and skill sets that are distinct while clearly falling within the field of TPC.

Scholars within TPC often publish work with the primary goal of making sense of the field, publishing about the field's history, methods, and research questions. Scholarship about the history and current state of TPC is filled with discussions about what subjects are most important, salient, or suitable for the profession (Brumberger & Lauer, 2020; Clegg et al., 2021; Coffelt et al., 2022; Melonçon & St.Amant, 2019; Robinson et al., 2019; Roundtree, 2018b; Spyridakis, 2015; Tham et al., 2022; Zheng et al., 2016). Questions about the status of the field are not new, and many past articles lay out the topics and perspectives that the authors view as growing in significance and essential to the future. TPC scholars have built entire careers by describing and redefining the field. Ongoing scholarly self-reflections have played an important role in developing a clear understanding of the field and the value of our work.

◼ A Dynamic Landscape

TPC is a continuously evolving domain, shaped by diverse professional expectations and industry trends. A mature field is not a monolith, which is why scholars must engage with this type of disciplinary maintenance. The complexity of a field can be seen as a sign of its strength and maturity, and as a complex and growing field, TPC is a perfect example. TPC is a complex field, and it often faces rapid change due to the evolution of tools, technologies, and industry practices. Jason Tham et al. (2022) recently explored the implications of integrating design thinking, content strategy, and artificial intelligence into TPC programs, insightfully pointing to the need to adapt and suggesting several strategies for doing so.

In the face of such a field, the argument put forward by Allen (1990) seems as relevant today as it did three decades ago. A strict definition of technical writing can arbitrarily exclude important work. Of course, definitions are needed at times, but we also need ways to investigate the broader workings of our field.

Still, establishing the focus, boundaries, and directions of TPC is a frequent goal of scholars, and there are several methods that have been developed. The implications of the definitions and descriptions can be significant. Two decades ago, Barbara Giammona (2004) conducted a survey and suggested that the future of technical communicators would shift focus toward information management and the change would require that the field adapt to technological and societal changes. Miles A. Kimball's (2017) analysis of the field leads him to argue for academic programs in technical communication to broaden their appeal by adjusting the scope and definition beyond a field of study and a profession "to include the vast, unrecognized bulk of technical communication performed every day" (p. 346).

In the effort to define what we do, Giammona and Kimball both offer engaging analyses and arguments about the potential influences and directions that could shape the fields related to technical communication at the time, and they both offer suggestions based on their observations. The predictions of Giammona and Kimball are engaging and valuable, but their efforts are not unique in the field and speculating about the future of the field is an imprecise art. David Wright et al. (2011) showed the robust history of predicting the future within technical communication with a detailed annotated bibliography covering six decades, and making the observation that "Many predictions do seem to be motivated by fear" (p. 448). While many scholars discuss the many changing facets of the field to consider its future, other scholars do so to create useful frameworks that invite us to notice the many corners and interconnected elements of our field.

■ Mapping the Field

Attempts to define the field or predict its future are essential efforts for determining how scholars, students, and practitioners should apply themselves. Discussions about what matters in TPC help inform our decisions when we design courses and programs, decide on academic paths, seek professional development opportunities, plan future career goals, select research topics, or reflect on our own skill sets as we draft a resume. As researchers approach the field from different perspectives, they help show the many sided, thriving complexity of technical and professional communication. Each study adds a new vision, and in so doing we may find at times that there are significant differences. For example, there is a disconnect between TPC articles written by practitioners and researchers in focus and motivation with practitioner articles focusing on process and profession and academic researchers more often focusing on writing products and educational matters (Boettger & Friess, 2016).

Studies that examine features of a field are valuable, especially during periods of change, because they can help clarify and strengthen connections in the field. Mueller (2017) has called such studies forms of "discipliniography," which he defines as "a genre that writes the field and is written by scholars in the field, and as such is a genre that is responsive to the growth of the field and its changing, contested state(s)" (p. 13). This type of scholarship helps to mediate discussions within the field, and it facilitates the process of familiarizing newcomers and outsiders by surfacing patterns and histories that would take years to learn about through first-hand study.

Carolyn D. Rude (2009) is an example of a study that has done this type of disciplinary work by documenting the research questions in our field — what to study (disciplinarity), how to teach (pedagogy), what we do (practice), and how to engage in rhetorical change work (social action). Rude referred to her work as 'mapping,' which is an apt metaphor for identifying a field's features and relationships. Rude's approach to explaining the field focused on research questions,

and the result is a useful entry point into a complex field of academic study and industry practice. Her map helps newcomers to the field understand the conceptual landscape, and it helps established members reflect and discuss the features and connections.

Several studies have documented the perspectives and experiences of practitioners in technical writing through surveys (Rainey et al., 2005; Shalamova et al., 2019). The findings of surveys, on their own, provide valuable insights; however, triangulating the findings from surveys with analyses of technical writing job ads generates a richer sense of the field that is grounded in the most current data.

Another approach to clarifying the focal points and systematizing the values of the field is to identify core concepts and keywords. Several books have been published in recent decades dedicated to explaining the values of TPC and our related fields through a close examination of our keywords (Adler-Kassner & Wardle, 2015; Heilker & Vandenberg, 2015; Tham, 2022; Yu & Buehl 2023). These types of collections that draw on the deep expertise and experience of established members of the field can be authoritative and useful, and they help the field examine its own complexity.

In the introduction to a recent collection, Han Yu & Jonathan Buehl write, "Keyword essays attempt to open up the meanings of words, to emphasize that meanings are always in flux, and to celebrate the different (but also overlapping) meanings of words as they are used in varied social, cultural, and disciplinary circles" (2023, p. 9). They go on to invite scholars to use the keyword essays in their collection to help examine the "evolving, divergent, and contested meanings" (p. 23). The process of identifying important terms and examining the full, complex usage of the terms in our field reflects that type of disciplinary maintenance described by Mueller, albeit through other means. Yu and Buehl's (2023) collection offers the field an insightful glimpse into the complexity of many topics in our field, and what makes their collection particularly valuable is how the terms were selected. They explain that they used a data-driven process that began with an analysis of the field's academic publications (p. 11). Scholarship that engages in the careful identification of key features provide valuable entry points. Tracking the field's boundaries, concepts, and language is essential for working through questions like those regarding TPC's evolving scope, core concepts and terms, interdisciplinary nature, and ethical commitments through social and technological changes. These questions are not only of interest, but they should regularly be reexamined so we can confidently make sense of the field itself, what we offer as members of a shared discipline, and what avenues of professionalization are needed.

Researchers will often review scholarly and practitioner publications to create their own conceptual maps to highlight the various domains of the field, just as Yu and Buehl (2023) did to identify keywords for their collection. For example, Lisa Melonçon and Kirk St. Amant (2019) and Chris Lam and Ryan Boettger (2017) analyzed the research methods in top TPC journals. Rude (2009) investigated the research questions in technical communication by analyzing books

published in our field, an approach that is similar to Kimball's (2013) approach to gathering information about design principles. Books or articles are scholarly genres that represent the interests, questions, topics, perspectives, and values of our field, but scholarly perspectives may not provide a complete picture of how industries are (re)shaping the field.

Careful analysis of our most representative texts, such as job ads and associated documents, is an avenue that can help us develop a grounded understanding of the many facets of our field. Several scholars in TPC have involved technical writing job ads as data for this purpose (Brumberger & Lauer, 2015; Lanier, 2009; Stanton, 2017). While job ads themselves are imperfect, they yield insights into current and future expectations for technical writing positions. More recently, Brumberger and Lauer (2020) were able to identify key competencies, personal characteristics, and common information products in their grounded theory research. Because job ads are a meta-genre (Giltrow, 2001) of technical communication that are often composed by non-technical writers within human resources, their patterns—both how they define technical writing skills and who they claim to be seeking to fill technical writing roles—a large (and ever-expanding) corpus of technical writing job ads can help us begin to move from definitional work to a more in-depth understanding of the everyday knowledge work of technical writers across industries.

For our purposes, we define meta-genre as situated, self-referential language that characterizes who technical writers are and what they do. Job ads are a meta-genre of a primary set of TPC genres: resumes, cover letters, applicant instructions, and even LinkedIn profiles. They contain information commonly found in these primary TPC genres, such as lists of required skill sets, descriptions of experience, or desired certifications and credentials. They are, as Janet Giltrow (2001) describes them, "wordings and activities, demonstrated precedents or sequestered expectations" that surround a genre and indicate how readers and writers should appropriately take it up" (p. 195). As such, a corpus of job ads from a variety of locations and industries collected over a period of time can provide the TPC field with a better understanding of "how texts and related communication practices mediate knowledge, values, and action in a variety of social and professional contexts" (Rude, 2009, p. 176).

■ Available, Scalable, and Repeatable

Other researchers have studied patterns in the field using job ads, but typically at a smaller scale. Text mining lets us zoom out and view things at a different scale. We approach this study with the understanding that text-mining allows for a more comprehensive analysis of the field than is typical in other methods. By treating advertisements as machine-readable data that can be aggregated, we can observe language patterns that are difficult to identify by close reading. Automated frequency counts, collocation searches, topic modeling, and related methods

for computationally analyzing collections of text create statistical representations of the text that help us understand the disciplinary landscape at a wider scale. Computer-aided textual analysis is a useful way to conduct a distant-reading study due to the availability of text, the scalability of the analysis, and the repeatability of the methods. A thoughtful and consistent analysis provides the opportunity to zoom in or zoom out to understand patterns in context and from afar.

In each job ad study we reviewed, researchers relied primarily on human readers to identify the features for analysis. While the methods and findings of previous studies on job ads are effective and valuable, the choice of textual features are predetermined or determined by human readers as they review the ads, which means the identification of patterns in the text is driven by the research team. The study that follows this review, in contrast, employs a different approach, viewing the language that describes technical writing in aggregate before human analysis. The full text in each ad across the entire collection is counted and considered for potential patterns by the computer during the analysis. The computer does not begin with preconceptions about which language is most important, so when the human reader reviews the patterns, they are faced with a different sense of what is important.

Large-scale analyses of job ads can provide grounded descriptions about the current state of the field for stakeholders to engage more confidently and comfortably with the most relevant and current concerns of the field. For example, Rhonda Stanton (2017) has argued that understanding what employers want will help university faculty and administrators (p. 224), a point that had been made in an earlier study (Rainey et al., 2005, p. 335). We ourselves have had many discussions with colleagues speculating about what employers are seeking from applicants in technical writing and the need for reliable data about various paths that are available. Studying job ads and employment trends that move quickly can help us trace new subfields and specialized skill sets as they gain importance (Shalamova et al., 2019).

To establish a more holistic picture, to strive for a sense of the field with all of its complexity, we can use computational methods of textual analysis based on consistent features across the texts. Using computational approaches to analyzing text offers a reflective analysis of TPC, and there are several ways this type of work has been described including text mining (Fan et al., 2006; Roundtree, 2018a), machine learning (Lindstedt, 2019; Murakami et al., 2017), and text as data (Grimmer et al. 2022). Text mining may include a wider array of activities that include acquiring and parsing text as well as various forms of analysis, including machine learning.

Machine learning employs algorithms to improve upon tasks by learning from data and can facilitate the analysis of extensive textual data in ways that human reading and interpretation alone cannot accomplish. These methods have received limited attention in the field of TPC. After many months of attempting to code our data set with more conventional methods of text and rhetorical analysis, we began to see the need for another approach. There are several methods in

natural language processing and machine learning that are appropriate for making sense of large collections of texts, including text clustering, text classification, information retrieval, summarization, opinion mining, and sentiment analysis (Aggarwal, 2018). Grimmer et al. (2022) use the term 'text as data' to capture many of these methods, and they provide several useful guidelines for applying machine learning, most notably they argue that text methods are useful because they support discovery and can lead to new insights and questions. They also clarify that "Text analysis does not replace humans–it augments them" (p. 24).

Changing the analytic scope from a hand-read sample to a comprehensive corpus has several important implications for research. Close-reading supports deep critique, but cannot claim representativeness. A wider, more comprehensive scope builds what Mueller calls *network sense*, which "mitigates the negative consequences of excessive specialization" (2017, p. 164). Distant reading and corpus methods operationalizes that network sense by focusing on surface-level, countable features, what Mueller calls *thin description*. Reducing and simplifying texts can take many forms, as Mueller writes, "everyday examples where distant reading and thin description do their thin-distant work…table of contents, indexes and the notes on a book jacket" (2017, p. 6). In short, thin description and distant reading are the tools that Mueller identifies for adjusting our scope of inquiry. Thin description makes it possible to engage in distant reading by treating content as data that can be tracked across thousands of texts, after which researchers can then analyze and discuss the significance of those patterns based on the goals and theories of the field.

In our study, we work to better understand technical writing using job ads, a genre of communication explicitly designed to describe the work expected and the skills involved. In particular, we have examined language patterns and the presence of communication forms and tools that fall outside of traditional notions of writing. By identifying common features and their relationships, the findings of our study function as a conceptual map of the field, providing data about the topics and their relationships in context. While job advertisements cannot reveal everything about the work of technical writers, the language in the advertisements can act as an indicator of what working professionals face as they navigate the profession. Understanding technical writing from multiple perspectives is important for getting an overall sense of our complex field.

◼ Methods

The project has gone through several phases, and the texts have been analyzed in multiple ways as described below. The findings are based on 4,597 job advertisements available on Indeed, a popular job board, that result from the search terms *technical writer* or *technical writing*. Throughout the study, we treat text as data (Grimmer et al., 2022), with the understanding that it is through certain texts that we define, describe, and create the field.

Data Collection and Data Cleaning

The first phase of the study was developing a principled corpus of texts. A corpus is principled if the collection and inclusion of items in a corpus follow specific and predefined rules. The process involved collecting, organizing, and cleaning the texts. To ensure repeatability and to clearly explain what the data represents, we carefully managed the collection process and parameters. During the first week of the month, we identified and collected job ads by searching Indeed using the target search phrases. Collecting advertisements took place over a 12-month period, November 2021-October 2022. We searched for *technical writer* and *technical writing*, which yielded job postings from a wide variety of organizations and with a wide variety of job titles. We repeated the search using the same search terms for the 10 most populous cities in the United States according to the U.S. Census: New York, Los Angeles, Chicago, Houston, Phoenix, Philadelphia, San Antonio, San Diego, Dallas, and San Jose.

Understanding the Corpus

The tables in this section describe the corpus makeup: the distribution of the ads across months (Table 10.1), the number of word tokens and word types (Table 10.2), and the job titles found in the collection (Table 10.3 and Table 10.4).

Table 10.1. Ads Collected and Ads Identified as
Unique Within Each Collection Cycle

Month	Ads Collected	Unique Ads
Nov -2021	1,187	670
Dec-2021	1,195	632
Jan-2022	1,307	657
Feb-2022	1,273	609
Mar-2022	1,212	650
Apr-2022	1,202	637
May-2022	654	459
June-2022	1,142	642
July-2022	967	530
Aug-2022	1,140	606
Sept-2022	534	370
Oct-2022	1,387	577
Total	13,200	7,039

Table 10.2. Final Corpus Construction by Texts, Word Tokens, and Word Types

Collection Period	Total Unique Ads	Word Tokens	Word Types
Nov 2021–Oct 2022	4,597	2,449,139	20,399

The final number reflects the corpus after exact duplicates were removed across the whole collection period.

Table 10.3. Job Titles Containing Target Features

Title Characteristics	Total
"tech*" and "writ*"	3,224
"sr." or "senior" or "manag*" or "supervis*"	750
Contains "edit*"	303

Table 10.4. Job Titles that Occur 10 or More Times in the Corpus

Title	Count
Technical Writer	1,197
Senior Technical Writer	168
Proposal Writer	123
UX Writer	56
Technical Editor	41
Technical Writer II	35
Technical Writer/Editor	31
Sr. Technical Writer	31
Technical Content Writer	26
Technical and Engineering Writer	24
Senior UX Writer	24
Specifications Writer	22
Technical Writer I	20
Tech Writer	19
Content Writer	19
Technical Writer III	16
Technical Report Writer	15
Specification Writer	15
Document Control Specialist	15
Grant Writer	14
Associate Technical Writer	13
Technical Writer (Contract)	12
Technical Proposal Writer	12

Title	Count
Technical Proposal Manager	12
Technical Writing Manager	11
Proposal Manager	11
IT Technical Writer	11
Impact Report Writer	11
Writer	10
Technical Volume Writer	10
Sr Technical Writer	10

During the collection process, each job ad was saved as an individual file that included the job title, company, location, and the description of the position. Tags were used to designate each content type so that computer-aided searches could more easily recognize these pieces of information with each position. The development of a data set, or a corpus, involves careful planning to ensure that the computer can reliably assist in the analysis, but also to ensure that the contents are consistent and principled. The choice of the composition of a corpus can have significant impacts on how it is interpreted.

The process of preparing the files for analysis involves multiple steps of organizing and cleaning. After the initial collection process, the job advertisements are then organized and stored based on month and search term. While there are many ways the corpus could be structured and analyzed, for this study we consolidated job ads for the entire collection period into a single corpus for analysis. In future studies, a comparison of the advertisements could be made based on location, month, or search term.

Using a program named SearchMyFiles, we removed duplicated postings from the collection. The program compares files to remove identical information. Companies often reuse substantial portions of their postings, so many advertisements share large portions of text. However, small differences in language are considered a different position for the sake of this study. Date and location were disregarded when identifying duplicate ads, so ads that remained across collection periods or ads that were reposted regularly were identified and exact duplicates were removed. The removal of duplicates was completed using automated methods for exact matches only, so some redundant postings with only small differences are included in the final corpus. Differences, no matter how small, were assumed to be an intentional distinction between different positions or updates intended to improve the advertisement in some other way.

Additional cleaning was needed for the frequency analysis, which is described below, to move past words with little relevance to the study. To ensure the analysis would generate the most useful data, a list of words was generated for AntConc, specialized software for analyzing text, to ignore when generating frequencies.

The list included words with primarily grammatical function and words that are generalizable to job advertisements. The grammatical words include pronouns, prepositions, articles, conjunctions, helping verbs, demonstratives, and some vague adverbs. Additional words were filtered because they appear frequently in job advertisements because they refer to the position, applying, or legal disclaimers but do not provide insight into the position (e.g. words related to locations, pay or benefits, and legal protections). Examples of these types of words include *job*, *apply*, *applicant*, *work*, *race*, *sex*, *gender*, etc. Words were only ignored when creating frequency lists, not during the topic modeling.

▮ Results

In the sections that follow, we explain the method for each phase of analysis and present the findings. We begin by sharing information about the job titles to provide additional context for the corpus and to provide insight into the roles that job seekers would likely encounter when conducting a generic search for technical writing jobs. We then provide an overview of frequently occurring language organized into general frequency and then frequency by selected topics. Finally, we explain how we applied topic modeling to further investigate the corpus.

▮ Job Titles

Examining the job titles in the corpus helps to clarify the range of work represented in the collection. Table 10.4 shows that just 70% of the job titles represented in the corpus match for both *tech** and *writ**. These titles include many variations, though, with 1,035 unique job titles. 31 out of 1,698 job titles in the corpus appear 10 or more times. 1,149 job titles in the corpus appear in only a single job ad. It is worth noting that even though *technical communicator* is often the preferred term by practitioners and scholars, *communicator* only appears twice in job titles we collected.

The review of job titles also shows that the corpus includes a small number of positions explicitly identified as senior and editor roles. Technical writer levels are distinguished by a variety of labels and levels. The collection includes technical writer levels 1–5, often represented with roman numerals. Titles also include distinctions such as *senior*, *sr*, *junior*, *manager*, *lead*, *principal*, and *associate*.

▮ Exploring Frequent Language

As an initial point of analysis, a general frequency analysis helps provide additional context for the makeup of the corpus of job ads. The frequency list (Table 10.5) shows the most frequently occurring words, along with their total frequencies (number of occurrences) and ranges (number of job ads containing the word at least once). Additional frequency lists were generated using premade lists of

terms associated with several topics of interest in technical writing classes. Table 10.6 shows both the frequency and range for selected word types. The frequency lists help add context for the analysis, and creates the opportunity for some initial observations.

Table 10.5. Words that Appear in 50% or More of the Ads, Ordered by Range

Type	Freq	Range	Type	Freq	Range
technical	21569	4240	development	6161	2725
writing	11090	3903	management	5986	2685
skills	9775	3703	communication	3845	2621
team	9847	3367	knowledge	5157	2612
writer	6274	3288	opportunity	4204	2602
ability	9098	3275	support	5144	2542
degree	3966	3050	new	4790	2520
information	7029	3008	environment	4427	2507
time	5464	2877	company	5182	2455
content	11349	2859	related	4039	2422
documentation	11083	2824	business	5567	2409
required	6779	2822	strong	4057	2340
requirements	5587	2814	written	3210	2328
responsibilities	3557	2764	preferred	4084	2302

Table 10.6. Frequency of Words Indicative of Modes of Communication and Communication Technologies

Category	Term (freq, range)
Writing, Designing, and Making	writing (11090,3903), writer (6274,3288), design (3624,1722), written (3210,2328), write (2775,1866), make (2154,1474), writers (1529,926), making (796,680), designers (762,537), writes (438,344), makes (371,322), designed (308,266), designing (222,202), designs (220,205), designer (109,86), makers (72,71), designated (71,68), maker (31,27)
Oral Communication	meet (1822, 1362), meetings (871, 669), presentations (775, 614), interpersonal (774, 712), presentation (546, 450), meeting (513, 399), present (506, 451), meets (477, 414), speak (159, 142), presenting (126, 124), speaking (108, 92), presented (84, 63), presents (38, 38), speaks (28, 28), speakers (14, 11), presenter (11, 11), presently (11, 11), meetups (11, 10), speaker (8, 8), presentational (4, 4), presenters (4, 2), presentable (2, 2), presentations (1, 1), presentation (1, 1)

Category	Term (freq, range)
Visual Communication	visual (435, 379), visuals (103, 99), display (90, 87), slides (61, 60), visualization (57, 50), visually (40, 37), slide (32, 27), visualizations (30, 30), displays (30, 27), visualize (14, 14), displaying (8, 8), slideshow (4, 4), displayed (4, 4)
Multimedia	video (718,490), interface (525,417), videos (338,274), multimedia (166,142), interfaces (163,145), podcasts (37,34), interfacing (34,32), podcast (30,10)
Authoring and Design Tools	microsoft (2067,1592), adobe (1285,883), google (630,280), acrobat (540,415), confluence (417,336), visio (363,332), illustrator (368,312), jira (346,310), photoshop (329,298), indesign (305,257), framemaker (302,249), madcap (232,182), robohelp (69,65), figma (51,44), lucidchart (18,18), xd (7,6)
Coding and Programming	code (662, 476), html (525, 451), xml (495, 389), dita (296, 214), markdown (199, 185), git (198, 184), css (206, 182), javascript (130, 118)
CMSes	aws (198,68), wiki (67,57), squarespace (62,6), wordpress (50,44), drupal (9,9), joomla (1,1), wix (1,1), blogger (1,1)
Social Media	linkedin (161,150), facebook (153,126), twitter (125,118), instagram (49,45), youtube (44,41), tiktok (79,14), snapchat (10,8), whatsapp (1,1)

◼ Topic Model

Topic modeling is an iterative process of identifying groups of words that commonly occur together in documents. There are a few methods for identifying topics and themes within a body of text, and topic modeling is one of the most well-established uses of machine learning when working with language. Latent Dirichlet Allocation (LDA) is one of the processes for identifying topics (Grimmer et al., 2022; Lindstedt, 2019). For this study, an LDA model was created using Gensim, a python library for topic modeling.

The outputs of an LDA model are groups of words that frequently appear together in the same texts throughout the corpus (see Table 10.7). It's important to note that the LDA model outputs topics as numbered groups. The actual naming of these topics is done post hoc by the researcher based on the words in each group and their perceived thematic relevance. These groups of words are called topics, and researchers rely on an iterative process to identify a useful model, which includes labeling the numbered topics in a way that is meaningful for their analysis.

Researchers must ultimately determine, based on a review of the word groupings and the texts that contain them, how many topics to include in the model and what each 'topic' is about. An analysis of the topics and their presence in various texts can then be used to characterize the overall themes that are discussed in the texts. As a process for finding hidden patterns in text, LDA is an effective tool for

finding hidden (latent) patterns, and by analyzing these patterns researchers can discover relationships that may not be immediately clear through other means.

To analyze the topics further, we created visualizations. We used the pyLDAvis, a Python package that implements the visualization described by Carson Sievert and Kenneth Shirley (2014). This type of visualization, shown in Figures 10.1, 10.2, and 10.3, is a tool that helps researchers interpret a topic model.
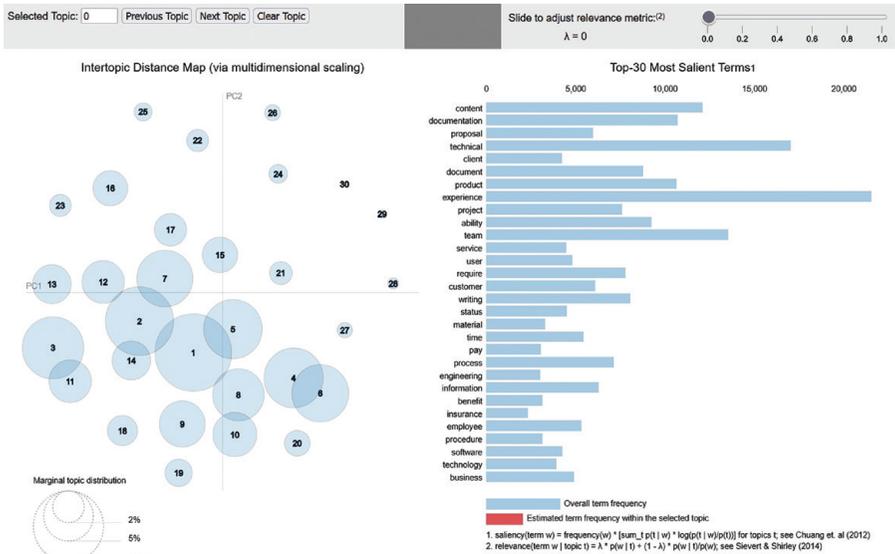

Figure 10.1. Interactive visual of the topic model generated using Gensim and pyLDAvis. The visualization helps with the interpretation of topics.
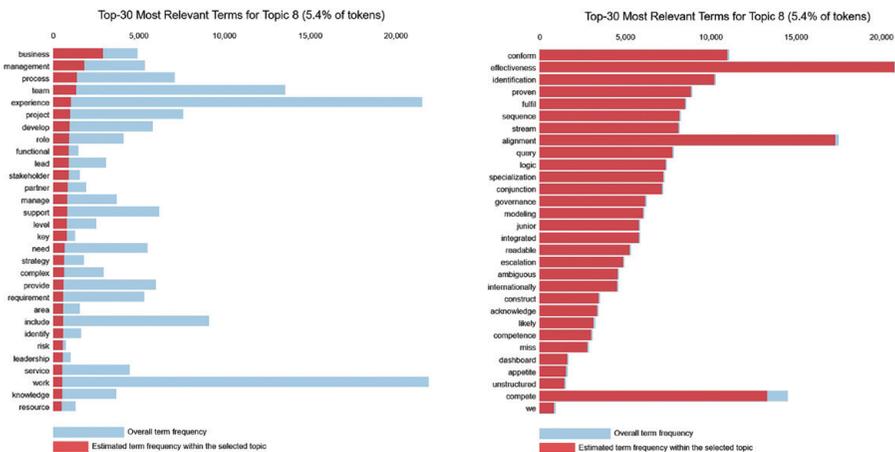

Figure 10.2. A visual comparison of the words in Topic 8, Analysis and Risk. The most probable words are shown on the left, and the words with the highest lift (indicating less likely to appear in other topics) is on the right.
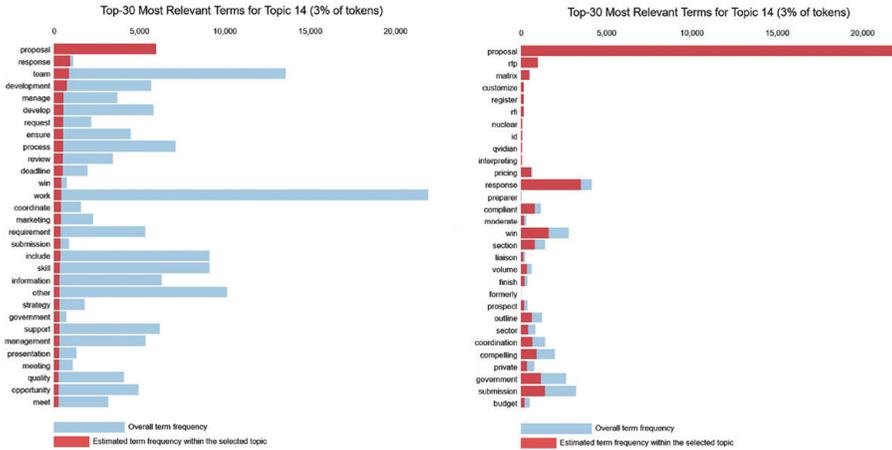
Figure 10.3. A visual comparison of the words in Topic 14, Proposals. The most probable words are shown on the left, and the words with the highest lift (indicating less likely to appear in other topics) is on the right.

Figure 10.4. Network diagram of strongest topic correlations. Lines connect nodes with a tie strength > .1, indicating a potential relationship between the topics. Correlations were calculated between each topic based on the topic–document matrix using Pearson correlation. Numbers in parentheses reflect the topic number in the topic model.

Figure 10.4 shows the correlation between the topics (Lindstedt, 2019). According to Sievert and Shirley (2014), the interactive visualization "enables deep inspection of topic-term relationships in an LDA model, while simultan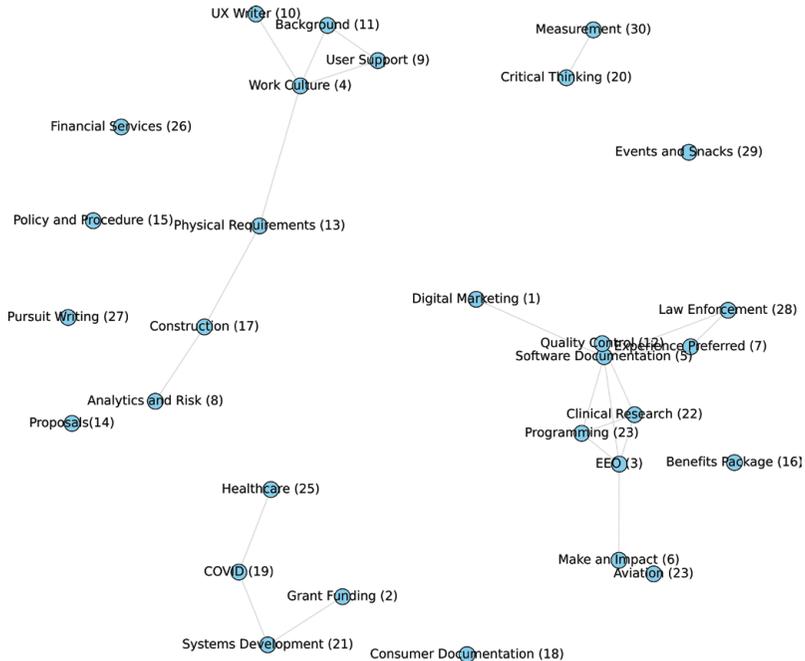eously providing a global view of the topics" (p. 69). The key functionality in the LDA visualization is the relevance slider which allows users to adjust the way 'relevance' is calculated by balancing two measurements: its *probability* and its *lift*, and the top terms by both are presented in Table 10.7. The probability of a word occurring in a topic is higher for more common words, while the lift score of each word is determined by calculating a ratio of the probability within a topic to its probability across the corpus. Understanding topics becomes easier when looking at both frequently occurring words in a topic and words that are less frequent but unique to the topic. The lift score often reveals words that occur infrequently overall, but are much more common with a particular topic.

Table 10.7. Word Lists Representing Each Topic
Generated in the LDA Topic Model*

	1 - Digital Marketing		2 - Grant Funding	
	Probable	High Lift	Probable	High Lift
1.	content	image	include	student
2.	skill	audio	program	grant
3.	ability	seo	experience	administrative
4.	write	original	provide	funding
5.	experience	white	other	validation
6.	writing	threat	position	letter
7.	work	proofreading	support	behavior
8.	strong	succinct	report	dependent
9.	communication	minority	work	brochure
10.	technical	constructive	require	classification
	3 - EEO		4 - Work Culture	
	Probable	High Lift	Probable	High Lift
1.	status	creed	team	even
2.	gender	discrimination	work	cell
3.	disability	genetic	look	much
4.	protect	prohibit	role	choose
5.	work	harassment	experience	air
6.	national	permit	new	raise
7.	race	grade	base	eager
8.	origin	tough	benefit	excited
9.	orientation	political	company	happy
10.	sexual	affiliation	learn	entertainment

	5 - Software Documentation		6 - Make an Impact	
	Probable	**High Lift**	**Probable**	**High Lift**
1.	documentation	api	work	difference
2.	product	programmer	people	embrace
3.	technical	fix	world	too
4.	software	scratch	make	boundary
5.	experience	cyber	help	dream
6.	user	screenshot	life	belong
7.	team	alone	well	authentic
8.	create	facing	opportunity	redefine
9.	work	distinguish	diverse	forefront
10.	use	instrumental	way	ahead
	7 - Experience Preferred		**8 - Analytics and Risk**	
	Probable	**High Lift**	**Probable**	**High Lift**
1.	experience	commute	business	conform
2.	technical	relocate	management	effectiveness
3.	work	mechanical	process	identification
4.	year	reliably	team	proven
5.	location	desktop	experience	fulfill
6.	time	podcast	project	sequence
7.	require	shift	develop	stream
8.	hour	registration	role	alignment
9.	writing	historical	functional	query
10.	document	raw	lead	logic
	9 - User Support		**10 - UX Writer**	
	Probable	**High Lift**	**Probable**	**High Lift**
1.	customer	edge	content	unlock
2.	technology	campus	product	localization
3.	technical	cut	experience	failure
4.	solution	authoring	user	strategist
5.	experience	computing	customer	encouraging
6.	service	listening	work	landing
7.	cloud	proprietary	design	evolution
8.	provide	tier	create	obsession
9.	datum	gas	strategy	iterate
10.	company	unparalleled	team	uncover

	11 - Background		12 - Quality Control	
	Probable	High Lift	Probable	High Lift
1.	applicant	criminal	document	monitoring
2.	employment	affirmative	system	deviation
3.	accommodation	hiring	ensure	inspection
4.	opportunity	recruitment	procedure	batch
5.	application	charge	control	instructor
6.	position	recruiting	requirement	root
7.	disability	conviction	review	conclusion
8.	process	placement	change	classroom
9.	provide	disclose	record	capa
10.	consider	disabled	compliance	corrective
	13 - Physical Requirements		**14 - Proposals**	
	Probable	High Lift	Probable	High Lift
1.	perform	representative	proposal	proposal
2.	require	site	win	rfp
3.	duty	laboratory	requirement	matrix
4.	job	occasionally	submission	customize
5.	ability	lift	include	register
6.	employee	walk	skill	rfi
7.	other	pound	information	nuclear
8.	assign	contain	other	id
9.	function	object	strategy	qvidian
10.	essential	incumbent	government	interpreting
	15 - Policy and Procedure		**16 - Benefits Package**	
	Probable	High Lift	Probable	High Lift
1.	document	continuity	insurance	dental
2.	procedure	confidentiality	pay	savings
3.	documentation	mortgage	benefit	spending
4.	process	suit	time	checklist
5.	technical	moderately	dental	retirement
6.	policy	css	company	discount
7.	information	lending	employee	immediately
8.	standard	intranet	plan	sick
9.	maintain	simplifie	offer	tuition
10.	template	bond	match	technician

	17 - Construction		18 - Consumer Documentation	
	Probable	High Lift	Probable	High Lift
1.	project	construction	content	teaching
2.	design	selection	training	maximum
3.	engineering	sustainable	create	centralized
4.	technical	architectural	material	informative
5.	specification	abreast	information	groundbreaking
6.	construction	committee	write	complement
7.	firm	profession	need	outdoor
8.	coordinate	forum	understand	distill
9.	provide	notch	audience	financially
10	development	attach	instructional	confluence
	19 - COVID		20 - Critical Thinking	
	Probable	High Lift	Probable	High Lift
1.	covid	faculty	ability	judgment
2.	vaccination	vaccine	knowledge	comment
3.	vaccinate	light	information	lesson
4.	faculty	card	problem	diploma
5.	fully	mathematic	understand	exchange
6.	health	attribute	require	certificate
7.	include	emergency	concept	concurrently
8.	safety	chemistry	method	fashion
9.	requirement	mask	basic	delegate
10	vaccine	pandemic	technique	cognizant
	21 - Systems Development		22 - Clinical Research	
	Probable	High Lift	Probable	High Lift
1.	experience	assembly	medical	clinical
2.	engineering	lab	clinical	scientific
3.	product	carrier	scientific	phd
4.	development	ease	regulatory	manuscript
5.	technical	smartphone	science	dose
6.	demonstrate	actalent	device	systematic
7.	support	established	document	statistical
8.	drive	aviation	submission	collaborator
9.	process	transparent	datum	communicating
10	good	discrepancy	review	endeavor

	23 - Programming		24 - Aviation	
	Probable	High Lift	Probable	High Lift
1.	programming	programming	production	operator
2.	application	multimedia	material	http
3.	technical	electrical	engineering	sketch
4.	developer	networking	drawing	flight
5.	language	server	specification	slide
6.	system	troubleshoot	technical	aircraft
7.	engineering	licensing	equipment	bulletin
8.	xml	scam	chart	schematic
9.	technology	setup	interview	revolutionary
10	multimedia	demonstration	illustration	usd
	25 - Healthcare		26 - Financial Services	
	Probable	High Lift	Probable	High Lift
1.	healthcare	healthcare	client	employ
2.	patient	patient	service	capital
3.	university	university	financial	catalog
4.	medicine	medicine	firm	side
5.	school	school	employ	expense
6.	departmental	packaging	capital	upcoming
7.	college	ideally	investment	legacy
8.	education	accredited	credit	client
9.	care	artwork	market	corporation
10.	show	initial	catalog	investment
	27 - Pursuit Writing		28 - Law Enforcement	
	Probable	High Lift	Probable	High Lift
1.	bid	bid	productivity	productivity
2.	executive	extract	citizen	citizen
3.	pursuit	pose	mark	mark
4.	persuasive	humanity	sensitive	externally
5.	messaging	grammatically	client	rank
6.	extract	tune	externally	recognition
7.	pose	jargon	rank	touch
8.	team	journalistic	recognition	residence
9.	humanity	liberal	touch	autonomy
10.	business	crisp	internally	justice

	29 - Events and Snacks		30 - Measurement	
	Probable	High Lift	Probable	High Lift
1.	event	frequent	store	store
2.	frequent	sport	sometimes	sometimes
3.	investor	lunch	retail	measurement
4.	space	snack	measurement	stereotype
5.	sport	publicly	stereotype	fantastic
6.	lunch	ramp	fantastic	retail
7.	snack	off	abroad	abroad
8.	regulate	breakfast	cellphone	cellphone
9.	publicly	premier	filesystem	filesystem
10.	ramp	kitchen	git	git

* Each topic includes two word lists. The most probable words (most frequently occurring) for the topic and words with a high lift (which is calculated as a ratio with the whole corpus indicating these words are more exclusive to the topic).

▌ Discussion

While there are several limitations that are worth acknowledging with this study, we believe that the topic and study design ensure that the limitations point to the potential for additional research to map the field through job advertisements and related texts. The methods are extendable by design, and the available content for investigation pose both challenges and opportunities. In the discussion that follows, we begin to lay out our interpretation of the results, but time and the constraints of this article make it impossible to fully investigate all of the patterns that emerge from a large corpus. It is worth noting that the corpus itself likely favors entry-level positions for several reasons. First, as a ratio, there are fewer senior-level or managerial positions, so the corpus represents more entry or mid-level positions. Additionally, senior positions may not be advertised as frequently as companies seek candidates through other strategies such as internal promotion or targeted recruiting. We have identified one avenue for investigating senior positions through advertisements, though, by isolating positions based on the job title. However, the analysis here primarily focuses on the corpus as a whole.

Using text mining as a distant reading technique requires judgment and interpretation. While the findings from this study are driven by the data and the process could be repeated, the choices made during the analysis (which topics to investigate when generating frequency lists, the parameters for the LDA model, which topics to analyze from the topic model) are driven by researcher judgment. Topic modeling is a form of unsupervised machine learning. Small adjustments in the parameters can lead to variations in the topics. Determining an appropriate number of topics can be a challenge, and in this study the final model was

selected after several iterations and review topic distribution using pyLDAvis. The topic model discussed in this study is one model among many, so this should not be understood as a definitive list.

Additionally, while the corpus analyzed in this study was systematically constructed, the data cannot fully represent the field of technical writing for several reasons. First, the collection strategy focuses on high-population locations in the US, and further investigation should be done in rural and international locations. Additionally, the search terms of *technical writer* and *technical writing* limit what is included. Technical writers have many responsibilities and job titles that were not the focus of the collection process, and it is our hope that a more complete picture can be developed with follow-up studies. Finally, job advertisements are imperfect representations of the jobs they describe as they are at times incomplete or overwritten to include more than necessary.

As an imperfect reflection of what it means to be a technical writer, though, the patterns and language within the job advertisements may have several implications. In this section, we discuss some of the patterns that we find most notable based on the data presented above. We offer these observations, hypotheses, and questions as an invitation for further inquiry about how the meta-genre of job ads shapes the field.

■ Industry Implications

Domain-specific topics present opportunities for understanding when and where technical writers can make an impact. These topics reflect industry-based domain knowledge and specialized skills that may be most beneficial when seeking career opportunities. The topic model, presented in Table 10.7, highlights many specialized industries and roles for technical writers. Working in each domain could require additional expertise beyond the scope of the core technical writing and communication skills. Topic 17, construction, may include several industries related to design, engineering, and architecture.

The knowledge and interests needed for working in a construction-focused role will differ from those needed to write in a user support (Topic 9) or clinical research position (Topic 22). For example, analytics and risk, Topic 8, includes several positions that focus on business, data, and risk. To work in this capacity, the ads ask technical writers to be comfortable with "modeling process flows, data relationship flows and event sequencing" and "Advanced business and data analysis, business process and application knowledge." The ads also discuss a type of technical document that we have not encountered in our own training or in our teaching materials called "model documentation," a formal and essential type of documentation for communicating risk. Several of the positions appear to require a combination of business knowledge, data analytics, and communication skills that make for an exciting and unique combination of skills. Certain aspects of the work within each of these fields will be unique. The

goals, genres, domain knowledge, and ancillary skills needed will vary for technical writers documenting software or drafting policies for corporations. If we consider job descriptions as collections of "wordings and activities, demonstrated precedents or sequestered expectations" about how technical writers should work (Giltrow, 2002, p.195), we can better understand how our discipline is being shaped by these external, but related documents. Future studies could map these domains more precisely.

Many of the topics–including digital marketing, grant writing, software documentation, analytics and risk, user support, UX writer, quality control, policy and procedure, construction, consumer documentation, clinical research, aviation, healthcare, and financial services–provide some insight into different professional avenues that technical writers could pursue. These topics also appear to identify the industries and roles that have been recently hiring technical writers, and investigations into the similarities and unique skills that are required between these topics could be valuable. Like our study, some of the most insightful studies about communication skills in the workplace (Coffelt et al., 2022; Shalamova et al., 2019) provide insights about communication skills by looking across industries and roles. Further analysis that compares the required skills and career options for technical writers in each industry could provide additional insights. An analysis that provides a detailed comparison of the experiences, skills, and expectations for technical writers in several industries could add additional valuable insights for educators, researchers, practitioners, and hiring managers.

■ Digital Media and Writing Technologies

A comparison of the topic model and frequency data for language related to technology illustrates one of the challenges of our field. Many writing technologies and their related skills appear in the data underscoring the value or necessity of writing technologies. In the topic model in Table 10.7, Topic 1 shows language associated with digital media and the web, and Topic 23 contains language related to programming and code. The topic model indicates that there are patterns in which digital media and programming are discussed in the job ads. We reviewed the frequency data for the whole corpus, and the data shows that specific tools are rarely mentioned in the job ads. Table 10.6 reveals that in the entire corpus of job ads, there are many tools and technologies discussed, but specific technologies (which we categorized as authoring tools, programming languages, content management systems, or social media platforms) are each only rarely mentioned. These observable patterns have several potential implications and raise some important questions about the practical relationship between technical writing and the technologies of writing. For example, do the practices of technical writers govern the technologies used for documentation, or are new technologies further expanding the definition of *technical writing*? Is the practice of technical writing too closely linked to the work of translation within specific industries? Are the

tools for technical writers evolving so quickly that it's difficult to define technical communication in succinct terms?

Broad familiarity with tools appears most useful.

Generic mentions of authoring and design tools include Microsoft (35%), Adobe (19%) and Google (6.1%) which may function as a need for broad familiarity with common writing applications. The most frequent specific applications identified include infrequent mentions of tools for collaboration, visual design, structured authoring, and prototyping. Further research and analysis could be conducted to understand the perspectives of hiring managers and applicants about the transferability of skills between similar tool types to better understand these findings. For example, do employers see comfort with Microsoft Office or Google Workspace as interchangeable? Is familiarity with JIRA enough for teams that use Confluence? Are Adobe XD skills sufficiently valuable to join workflows that rely on Figma? Understanding the assumptions of both employers and job seekers on these questions would be valuable.

Minimal coding is requested.

Words related to coding, programming, and structured authoring appear infrequently. The most frequent type of coding involves markup languages with HTML appearing in 451 (9.8%), XML in 389 (8.5%), DITA in 214 (4.7%), and Markdown in 185 (4%). We take these observations to indicate that a strong awareness of and comfort writing in coding environments is often important, but the ability to code or program is rarely requested. Further research should be done to investigate the roles and industries in which coding is most and least valued. Additional research could also be conducted to determine the ways career paths evolve in terms of opportunities and promotions based on various combinations of skill sets.

Content systems are rarely mentioned.

Content management and social media are both significant cornerstones of TPC curricula and are frequently discussed in our scholarship (Andersen, 2014; Bowdon, 2014; Bridgeford, 2020; Clark, 2008; Roundtree, 2018b; Shen et al., 2022; Wang & Gu, 2016). Given the significance of both, we anticipated technical writing job ads to discuss Content Management Systems (CMSes) and Social Media Platforms; however, in our study, both are rarely mentioned as shown in Table 10.6. There appears to be a gap between the significance these systems have in the work of communicators and their inclusion in job ads. One explanation could be that writers in these positions will be more focused on producing content that can be shared in digital spaces than on managing the digital spaces. It is likely that more senior roles will need to focus on managing and distributing content, which entails a higher level of responsibility and specialization.

Other potential reasons for these findings should be considered and investigated as well. It is plausible that businesses don't want to share information about which tools they use. As a component of business operations, protecting this information may be viewed as necessary to maintain a competitive advantage. Alternatively, job advertisements are an initial call that requires candidates to self-select into applying, and naming the technologies may dissuade appealing candidates. Hiring managers may be more comfortable training or retraining good candidates on the necessary technologies.

In regards to technological skills and writing tools, this study aligns with previous studies that have suggested that a broad aptitude for using writing technologies is needed by technical writers, but that deep knowledge of any specific tools set may not be needed. Scholars have established that for technical writers, a broad comfort with technology and an ability to work with content across systems is necessary, while hiring managers likely expect new hires to quickly adapt. The expectation at the point of hiring appears to be that technical writers can adapt and adopt the necessary technologies. Avoiding the names of specific technologies that a writer will engage with in the job advertisement could be a strategic choice by managers and HR departments.

The data in this study support the arguments of Marjorie Rush Hovde and Corinne C. Renguette (2017) that technological literacy should aim to teach students how to learn new tools quickly and evaluate their suitability for specific tasks. Similarly, Nadya Shalamova et al. (2019) found that technical communicators benefit from a general ability to work with technology, to adapt and stay current. The findings also align with the conclusions of Ann Hill Duin and Jason Tham (2018) who have asserted that, for technical communicators, code literacy should be treated "as a holistic competency [that] includes the ability to distinguish between code-as-language, code-as-tool, and code-as-structure" (p. 55). Many technical writers may find themselves in positions where it is valuable or necessary to work with code, but the ability to write code on its own is unlikely to be sufficient.

◼ Proposal and Grant Writing

The separation of grant writing (Topic 2) from proposal writing (Topic 14) points to a clear divide between non-profit and for-profit positions. Topic 2 includes positions for organizations and programs that depend on grants and donors. These positions include charities and advocacy groups that rely heavily on proposals and reports associated with obtaining and sustaining funding for their work. Topic 14 is about proposals in a traditional sense, with ads focused on direct responses to RFPs. The emergence of distinct topics around the seemingly similar roles of grant and proposal writing shows that these roles are considered, at least in the marketplace, to be distinct in significant ways. Our initial analysis suggests that the positions that fit grant writing and proposal writing show clearly distinct workplace cultures and knowledge domains.

▪ Pursuit Writing and Pursuit Strategist

A related, but unexpected, category emerged in the topic model. Topic 27 points to the emergence of a role not previously known to our team called 'pursuit writer' that is a strategic writing position focused on developing business opportunities or pursuing funding. A search in our academic databases did not reveal scholarly discussions of pursuit writing or pursuit strategy; however, there are several companies using these labels for key positions. Initially, this role sounds similar in nature to the work of proposal and grant writers, but there are some important distinctions. Language in job ads under Topic 27 often focused on audience analysis, or gaining insight into the client's needs in order to write a compelling proposal. One job ad in this group specifically focused on dissecting discussions, capturing information accurately, and reducing jargon for increased clarity. Better understanding the function of a pursuit writer or pursuit strategist would likely reveal valuable insights about the ways businesses develop a competitive edge when seeking relationships and funding.

▪ Marketing to Potential Applicants

While it is undoubtedly the case that job advertisements are often written to entice applicants, the unique facets of this strategy in technical communication have several potential implications worth further investigation. Topic 4, workplace culture, revealed several strategies in the advertisements to attract applicants and introduce the type of work environment. Ads associated with Topic 4 often include language related to passion and inspiration with phrases like "leads by example and is passionate" and "write about your passions while educating other developers." Teamwork and collaboration are viewed as highly valued and positive in some of these ads with statements like "Our Customer Success team is one of the most tight-knit groups at the company" and "You will work with some of the smartest and most interesting people in the industry." Others talk about a culture of growth: "join a large team and get the chance to grow and learn from the experienced and passionate in-house team." Many of the advertisements emphasize how exciting and appealing the position is with claims like "Get paid well to write about your passions while educating other developers with your own voice." The emphasis on the relationship between personal investment, passion, collaboration, and an exciting work environment creates a positive tone and a sense of excitement. However, further investigation is needed to understand how 'culture marketing' and 'definable job tasks' are parsed out within the meta-genre of job ads.

The use of these appeals could be a strategy to motivate applicants, or it could be a strategy to find applicants with valuable character traits for certain positions. Workplace culture correlates with UX writer and with user support (see Figure 10.4), which indicates that the language about creating a positive workplace

relates to creating a positive experience for customers. The kind of enthusiasm that is presented in the ads aligns well with the kind of investment and excitement that companies want to see from customers. The positivity and impact-oriented language could be seen as underscoring the role of technical writers as user advocates (Jones, 2016; Cleary & Flammia, 2012). Additionally, some employers may expect that technical writers that are motivated by dynamic, fast-paced, collaborative workspaces are more likely to be self-motivated and more prepared to gather needed information and feedback from their teams. In this way, the topic may point to the technical writer's role as a 'relationship-builder,' as found previously by Tammy Rice-Bailey (2016). Further research should be done to understand the use of workplace culture in job advertisements.

Backgrounds and Physical Requirements

The LDA analysis identified several topics based on the 'boilerplate' content that we would have initially overlooked, for example, Topic 3 (EEO), Topic 7 (Experience Preferred), Topic 11 (Background), Topic 13 (Physical Requirements), and Topic 16 (Benefits package). Our initial assumptions about some content in the advertisements was that the information is corporate/human resources language that has little to do with the positions being described. However, when reviewing positions that had contained some of these generic topics, some interesting questions emerged about the value of certain backgrounds for technical writers and the physical requirements that employers list for applicants, which we discuss below.

Encouraging Veterans and People with Criminal Backgrounds

Topic 3 and Topic 11 are closely related, sharing several words with some important differences. Topic 3 appears to be a general topic representing equal employment opportunity statements. By contrast, several of the ads matching Topic 11 include statements about supporting people with disabilities and disabled veterans, but they also include language about employee background checks. In reviewing these ads, an advertisement included a statement that 'we value diverse experiences, including prior contact with the criminal legal system, and applicants with criminal histories are encouraged to apply." This statement follows the EEO section, and could be easy to overlook, but it points to an interesting question about criminal backgrounds and technical writing. A search for the word 'criminal' found 291 tokens over 235 job ads, and a search for the phrase 'background check' found 204 occurrences in 166 ads. The short review found that most instances of criminal activity appeared to encourage applicants to apply regardless of their criminal background, and many of these positions were in technology companies. Further investigation about the expectations for background checks could provide insights into the boundaries between the expectations of human resources and required qualifications for technical writing roles.

Unexpected Physical Requirements

Additionally, Topic 13, which is about physical requirements, may seem out of place, but several of the advertisements with this section are for companies that involve manufacturing and may include requirements to engage with subject matter experts and to work with products in ways that matter to the position. For example, one advertisement says, "the employee is regularly required to use hands to finger, handle, or feel and reach with hands and arms…the employee must regularly lift and / or move up to 10 pounds and occasionally lift and/or move up to 25 pounds… while performing the duties of this Job, the employee is occasionally exposed to fumes or airborne particles." Another ad explains, "the employee is regularly required to sit; use hands/fingers to handle, or feel and talk or hear…specific vision abilities required by this job include close vision and ability to adjust focus. The employee must have sufficient mobility to enable travel to industrial sites, offices, and facilities." A third describes the physical requirements as, "the employee is required to talk, hear, and see. The employee frequently is required to stand, walk, sit, use hands to finger, handle or feel objects, tools, or controls; and reach with hands and arms to read dials, graphs, and procedures." Some of the physical abilities may be associated with operating a computer in an office setting, and the physical requirements described may have the effect of discouraging applications from well qualified candidates. This language may point to a poor understanding of how effective communication can be accomplished with a range of assistive technologies. However, there are also indications that the physical requirements extend further and some of the environmental factors of a manufacturing environment and the types of cross-functional work expected necessitate these expectations.

Conclusion

Texts are a product of technical writing, but texts also define what technical writing is. The genres that describe our work provide the features that can be used to map the most salient features and expectations. As a meta-genre for the field of technical writing, job ads present an opportunity to consider how the language, although often written by non-technical communicators (e.g., human resources departments and hiring managers), shapes roles, defines responsibilities, and establishes expectations. The process of mapping the texts that describe technical writing creates a data-driven opportunity for investigation that has several implications.

Definitional: Mapping the gaps and boundaries can help create moments of reflection and negotiation. If there's a disconnect between what the field is saying technical communication is versus what hiring managers think a technical communicator is, we need to identify the differences and learn to cross disciplinary and cultural boundaries.

Theoretical: Job advertisements often include fragmented descriptions that are a copy of a copy of a copy, yet they play an essential function. The recurring

language that is used to solicit work tells us how others understand technical communication, and these descriptions give a glimpse of what the marketplace expects technical writers to know, to be, and to do.

Methodological: Methods developed as machine learning and natural language processing have numerous potential uses in TPC scholarship. Conceptualizing text as data, specifically topic models, provides a path for how we might begin to shift the ethos of the field toward a more quantifiable, mixed method understanding of ourselves. When we treat text as data, we open new analytical tools. We gain the ability to understand the field, including new descriptions of the field that have the potential for better representation than other means provide.

Pedagogy and professionalization: While classroom practices must balance several demands, employability is one dimension of what this type of analysis strives for. To meet the needs of our students, our courses and programs must provide opportunities to develop knowledge and skills that prepare them to meet the challenges ahead of them. Providing job seekers themselves with topic models can create opportunities to strategize skill development and examine career options.

Ultimately, an analysis of job ads results in one perspective on the field and there are some things this data set can't show. However, technical communication scholars and practitioners can both benefit by examining the patterns that are revealed in an ongoing effort to understand the topics, skills, and boundaries that mark our field. Keeping an eye on the defining features of our field seems especially important as we continue to adjust from a wave of remote work, as new language technologies are disrupting the work of content production, and as higher education continues to face funding and enrollment challenges. No matter what position we play in the field, we need strategies to situate our work, to describe how we make an impact, and to communicate our value in both industry and academia.

▌ References

Adler-Kassner, Linda, & Wardle, Elizabeth A. (Eds.). (2015). *Naming what we know: Threshold concepts of writing studies*. Utah State University Press.

Aggarwal, Charu C. (2018). *Machine learning for text* (1st edition). Springer Science.

Allen, Jo. (1990). The case against defining technical writing. *Journal of Business and Technical Communication*, *4*(2), 68–77. https://doi.org/10.1177/1050651990000400204

Andersen, Rebekka. (2014). Rhetorical work in the age of content management: Implications for the field of technical communication. *Journal of Business and Technical Communication*, *28*(2), 115–157. https://doi.org/10.1177/1050651913513904

Batova, Tatiana, Andersen, Rebekka, Evia, Carlos, Sharp, Matthew R., & Stewart, John. (2016, September 23). Incorporating component content management and content strategy into technical communication curricula. *Proceedings of the 34th ACM International Conference on the Design of Communication*. https://doi.org/10.1145/2987592.2987631

Boettger, Ryan K., & Friess, Erin. (2016). Academics are from Mars, practitioners are from Venus: Analyzing content alignment within technical communication forums. *Technical Communication*, *63*(4), 314–327.

Bowdon, Melody A. (2014). Tweeting an ethos: Emergency messaging, social media, and teaching technical communication. *Technical Communication Quarterly*, *23*(1), 35–54. https://doi.org/10.1080/10572252.2014.850853

Bridgeford, Tracy. (2020). *Teaching content management in technical and professional communication*. Routledge.

Brumberger, Eva, & Lauer, Claire. (2015). The evolution of technical communication: An analysis of industry job postings. *Technical Communication*, *62*(4), 224–243.

Brumberger, Eva, & Lauer, Claire. (2020). A day in the life: Personas of professional communicators at work. *Journal of Technical Writing & Communication*, *50*(3), 308–335. https://doi.org/10.1177/0047281619868723

Clark, Dave. (2008). Content management and the separation of presentation and content. *Technical Communication Quarterly*, *17*(1), 35–60. https://doi.org/10.1080/10572250701588624

Cleary, Yvonne, & Flammia, Madelyn. (2012). Preparing technical communication students to function as user advocates in a self-service society. *Journal of Technical Writing and Communication*, *42*(3), 305–322. https://doi.org/10.2190/TW.42.3.g

Clegg, Geoffrey, Lauer, Jessica, Phelps, Johanna, & Melonçon, Lisa. (2021). Programmatic outcomes in undergraduate technical and professional communication programs. *Technical Communication Quarterly*, *30*(1), 19–33. https://doi.org/10.1080/10572252.2020.1774662

Coffelt, Tina, Cosgrove, Samantha, & Vance, Bremen. (2022). Measuring business and professional communication skills. *Business Communication Quarterly*, *85*(2), 169–191. https://doi.org/10.1177/23294906221082235

Duin, Ann Hill, & Tham, Jason Chew Kit. (2018). Cultivating code literacy: Course redesign through advisory board engagement. *Communication Design Quarterly*, *6*(3), 44–58. https://doi.org/10.1145/3309578.3309583

Fan, Weiguo, Wallace, Linda, Rich, Stephanie, & Zhang, Zhongju. (2006). Tapping the power of text mining. *Communications of the ACM*, *49*(9), 76–82. https://doi.org/10.1145/1151030.1151032

Giammona, Barbara. (2004). The future of technical communication: How innovation, technology, information management, and other forces are shaping the future of the profession. *Technical Communication*, *51*(3), 349–366.

Giltrow, Janet. (2002). Meta-genre. In Richard M. Coe, Lorelei Lingard, & Tatiana Teslenko (Eds.), *The rhetoric and ideology of genre: Strategies for stability and change* (pp. 187–205). Hampton Press.

Grimmer, Justin, Roberts, Margaret E., & Stewart, Brandon M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.

Heilker, Paul, & Vandenberg, Peter (Eds.). (2015). *Keywords in writing studies*. Utah State University Press.

Henning, Teresa, & Bemer, Amanda. (2016). Reconsidering power and legitimacy in technical communication: A case for enlarging the definition of technical communicator. *Journal of Technical Writing and Communication*, *46*(3), 311–341. https://doi.org/10.1177/0047281616639484

Hovde, Marjorie Rush, & Renguette, Corinne C. (2017). Technological literacy: A framework for teaching technical communication software tools. *Technical Communication Quarterly*, *26*(4), 395–411. https://doi.org/10.1080/10572252.2017.1385998

Ignatow, Gabe, & Mihalcea, Rada. (2018). *An introduction to text mining: Research design, data collection, and analysis*. SAGE.

Jones, Natasha N. (2016). The technical communicator as advocate. *Journal of Technical Writing and Communication*, *46*(3), 342–361. https://doi.org/10.1177/0047281616639472

Kimball, Miles A. (2013). Visual design principles: An empirical study of design lore. *Journal of Technical Writing and Communication*, *43*(1), 3–41. https://doi.org/10.2190/TW.43.1.b

Kimball, Miles A. (2017). The golden age of technical communication. *Journal of Technical Writing and Communication*, *47*(3), 330–358. https://doi.org/10.1177/0047281616641927

Lam, Chris, & Boettger, Ryan. (2017). An overview of research methods in technical communication journals (2012–2016). *2017 IEEE International Professional Communication Conference (ProComm)*. https://doi.org/10.1109/ipcc.2017.8013953

Lanier, Clinton R. (2009). Analysis of the skills called for by technical communication employers in recruitment postings. *Technical Communication*, *56*(1), 51–61.

Lauer, Claire, & Brumberger, Eva. (2016). Technical communication as user experience in a broadening industry landscape. *Technical Communication*, *63*(3), 248–264.

Lindstedt, Nathan C. (2019). Structural topic modeling for social scientists: A brief case study with social movement studies literature, 2005–2017. *Social Currents*, *6*(4), 307–318. https://doi.org/10.1177/2329496519846505

Melançon, Lisa, & Schreiber, Joanna. (2022). Introduction: Promoting a sustainable collective identity for technical and professional communication. In *Assembling Critical Components: A Framework for Sustaining Technical and Professional Communication* (pp. 3–16). The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/tpc-b.2022.1381.1.3

Melançon, Lisa, & St.Amant, Kirk. (2019). Empirical research in technical and professional communication: A 5-year examination of research methods and a call for research sustainability. *Journal of Technical Writing and Communication*, *49*(2), 128–155. https://doi.org/10.1177/0047281618764611

Moretti, Franco. (2013). *Distant reading*. Verso.

Mueller, Derek N. (2017). *Network sense: Methods for visualizing a discipline*. The WAC Clearinghouse ; University Press of Colorado. https://wacclearinghouse.org/books/writing/network/

Murakami, Akira, Thompson, Paul, Hunston, Susan, & Vajn, Dominik. (2017). "What is this corpus about?": Using topic modelling to explore a specialized corpus. *Corpora*, *12*(2), 243–277. https://doi.org/10.3366/cor.2017.0118

Rainey, Kenneth T., Turner, Roy K., & Dayton, David. (2005). Do curricula correspond to managerial expectations? Core competencies for technical communicators. *Technical Communication*, *52*(3), 323–352.

Rice-Bailey, Tammy. (2016). The role and value of technical communicators: Technical communicators and subject matter experts weigh in. *Technical Communication Quarterly*, *25*(4), 230–243. https://doi.org/10.1080/10572252.2016.1221140

Robinson, Joy, Dusenberry, Lisa, Hutter, Liz, Lawrence, Halcyon, Frazee, Andy, & Burnett, Rebecca E. (2019). State of the field: Teaching with digital tools in the writing and communication classroom. *Computers and Composition*, *54*, 102511. https://doi.org/10.1016/j.compcom.2019.102511

Rosselot-Merritt, Jeremy. (2020). Fertile grounds: What interviews of working professionals can tell us about perceptions of technical communication and the viability of technical communication as a field. *Technical Communication*, *67*(1), 38–62.

Roundtree, Aimee Kendall. (2018a). From engineers' tweets: Text mining social media for perspectives on engineering communication. *2018 IEEE International Professional Communication Conference (ProComm)*, 6–15. https://doi.org/10.1109/ProComm.2018.00009

Roundtree, Aimee Kendall. (2018b). Hospital tweets on H1N1 and death panels: Text mining the situational crisis communication response to health crises and controversies. *KOME*, *6*(1), 32–62. https://doi.org/10.17646/KOME.2018.13

Rude, Carolyn D. (2009). Mapping the research questions in technical communication. *Journal of Business and Technical Communication*, *23*(2), 174–215. https://doi.org/10.1177/1050651908329562

Shalamova, Nadya, Rice-Bailey, Tammy, & Wikoff, Katherine. (2019). Evolving skill sets and job pathways of technical communicators. *Communication Design Quarterly*, *6*(3), 14–24. https://doi.org/10.1145/3309578.3309580

Shen, Zixing, Tan, Songxin, & Pritchard, Michael J. (2022). Understanding the effects of visual cueing on social media engagement with YouTube educational videos. *IEEE Transactions on Professional Communication*, *65*(2), 337–350. https://doi.org/10.1109/TPC.2022.3156225

Sievert, Carson, & Shirley, Kenneth. (2014). LDAvis: A method for visualizing and interpreting topics. *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, 63–70. https://doi.org/10.3115/v1/W14-3110

Society for Technical Communication. (n.d.). Defining technical communication. *Society for Technical Communication*. Retrieved September 19, 2022, from https://www.stc.org/about-stc/defining-technical-communication/

Spyridakis, Jan H. (2015). Identifying new topics in TC curricula: Preparing students for success in a changing world. *Communication Design Quarterly*, *3*(2), 27–37. https://doi.org/10.1145/2752853.2752857

Stanton, Rhonda. (2017). Do technical/professional writing (TPW) programs offer what students need for their start in the workplace? A comparison of requirements in program curricula and job ads in industry. *Technical Communication*, *64*(3), 223–236.

Swarts, Jason. (2012). Communication design. *Communication Design Quarterly*, *1*(1), 12–15. https://doi.org/10.1145/2448917.2448920

Tham, Jason Chew Kit (Ed.). (2022). *Keywords in design thinking: A lexical primer for technical communicators & designers*. The WAC Clearinghouse ; University Press of Colorado. https://doi.org/10.37514/TPC-B.2022.1725

Tham, Jason, Howard, Tharon, & Verhulsdonck, Gustav. (2022). Extending design thinking, content strategy, and artificial intelligence into technical communication and user experience design programs: Further pedagogical implications. *Journal of Technical Writing and Communication*, *52*(4), 428–459. https://doi.org/10.1177/00472816211072533

Wang, Xiaobo, & Gu, Baotong. (2016). The communication design of WeChat: Ideological as well as technical aspects of social media. *Communication Design Quarterly*, *4*(1), 23–35. https://doi.org/10.1145/2875501.2875503

Wright, David, Malone, Edward A., Saraf, Gowri G., Long, Tessa B., Egodapitiya, Irangi K., & Roberson, Elizabeth M. (2011). A history of the future: Prognostication in technical communication: An annotated bibliography. *Technical Communication Quarterly*, *20*(4), 443–480. https://doi.org/10.1080/10572252.2011.596716

Yu, Han, & Buehl, Jonathan (Eds.). (2023). *Keywords in Technical and Professional Communication*. The WAC Clearinghouse; University Press of Colorado. https://doi.org/10.37514/TPC-B.2023.1923

Zheng, Pei, Liang, Xuan, Huang, Guanxiong, & Liu, Xun. (2016). Mapping the field of communication technology research in Asia: Content analysis and text mining of SSCI journal articles 1995–2014. *Asian Journal of Communication*, *26*(6), 511–531. https://doi.org/10.1080/01292986.2016.1231210