

# Quantitative Text Analysis in International Relations: A Review and Perspective

Takuto Sakamoto

Graduate School of Arts and Sciences, University of Tokyo

sakamoto@hsp.c.u-tokyo.ac.jp

## Abstract

This article reviews the development and current landscape of quantitative text analysis in the field of international relations (IR). Over the past two decades, the rapid digitization of diplomatic documents, parliamentary records, and online communications has generated unprecedented volumes of textual data for social science research. In parallel, advances in natural language processing (NLP) have enabled researchers to analyze large-scale text corpora in ways that go far beyond traditional qualitative reading and manual coding.

The article surveys foundational approaches to quantitative text analysis, including dictionary-based content analysis, topic modeling, and machine-learning-based classification, and examines how these methods have been applied to key questions in IR, such as international norms, policy preferences, and diplomatic discourse. Particular attention is given to research using United Nations documents, which constitute one of the largest and most systematic collections of publicly available texts in global politics.

The article then discusses recent methodological developments associated with deep learning, especially Transformer-based models and large language models (LLMs), and illustrates their analytical potential through state-of-the-art applications to United Nations Security Council debates. These examples demonstrate how contextualized language models enable more fine-grained analyses of concepts such as sovereignty and humanitarian norms, as well as the identification of agenda dynamics and actor-specific patterns at an unprecedented scale.

Finally, the article reflects on the broader implications of integrating LLMs and AI-driven tools into social science research. While these technologies substantially expand the scope and efficiency of quantitative text analysis, they also raise important challenges related to data quality, model bias, interpretability, and the assumptions linking language to political preferences and behavior. The article concludes by outlining future directions for quantitative text analysis in IR in an era of rapidly evolving AI technologies.

## 1 Introduction

Over the past quarter century, textual information generated across virtually all domains of social life has been digitized at an unprecedented pace and scale. Parliamentary debates, government

policy documents, diplomatic records, and official publications of international organizations—once primarily available in printed form—are now routinely stored, shared, and archived as digital text. At the same time, everyday communication through email, social networking services (SNS), and online forums has produced vast quantities of text data that are preserved and accessible for analysis. Alongside this expansion of digital text, techniques in natural language processing (NLP) have advanced dramatically, enabling researchers to analyze large collections of text in systematic and computationally efficient ways. These developments have made it possible to move beyond the limitations of traditional content analysis based on close reading or manual coding, and to conduct large-scale, reproducible analyses of textual data.

Against this background, a research approach commonly referred to as *quantitative text analysis* or *text as data* has rapidly diffused across the social sciences. This approach treats text not only as a source of qualitative interpretation, but also as data that can be systematically quantified, modeled, and analyzed using statistical and computational methods. By transforming words, sentences, and documents into numerical representations—such as word frequencies, co-occurrence patterns, or high-dimensional vector embeddings—researchers can identify latent patterns, topic structures, and stylistic or semantic differences within large text corpora. Quantitative text analysis has been widely adopted in sociology and political science, and international relations (IR) is no exception.

In IR research, quantitative text analysis has been applied to a broad range of substantive questions, including foreign policy preferences, international norm diffusion, diplomatic rhetoric, and the evolution of security discourse. The growing availability of large-scale textual data has played a crucial role in this development. In particular, official documents and speech records produced by the United Nations—such as resolutions, meeting transcripts, and formal statements by member states and international officials—have become central resources for empirical research. These materials are publicly available, systematically archived, and span several decades, making them uniquely well suited for longitudinal and comparative analysis. Recent years have also seen increasing use of textual data from media reports and social media posts by diplomats and policymakers, further expanding the empirical scope of text-based research in IR.

The purpose of this article is to provide a structured overview of these developments by reviewing major methodological approaches and empirical applications of quantitative text analysis in international relations. Rather than offering a comprehensive technical manual, the article aims to situate key methods within the broader trajectory of IR research and to highlight what kinds of substantive insights they have enabled. Particular attention is paid to studies that analyze United Nations texts, as this literature exemplifies both the opportunities and the challenges of large-scale text analysis in global politics.

The article also addresses recent advances in NLP associated with deep learning, especially the emergence of Transformer-based architectures and large language models (LLMs).<sup>1</sup> Since the introduction of the Transformer architecture in 2017, NLP models have achieved remarkable im-

---

<sup>1</sup>For clarity, “deep learning” refers to a subfield of machine learning that relies on multi-layer neural networks for feature extraction and classification. “Machine learning” more broadly denotes computational methods for learning, classification, and prediction, of which deep learning constitutes one important category.

provements in contextual understanding, multilingual performance, and task generalization. These developments have significantly expanded the analytical possibilities of quantitative text analysis, enabling more precise classification, stance detection, and semantic analysis than was previously feasible with dictionary-based or traditional machine-learning approaches. At the same time, the increasing accessibility of LLMs—through open-source implementations and application programming interfaces—has lowered the technical barriers to applying advanced language models in social science research.

This article situates these recent innovations within the longer tradition of quantitative text analysis in IR. It reviews the evolution of core methods, surveys major applications in the field, and illustrates cutting-edge approaches through empirical examples based on United Nations Security Council debates. Beyond methodological considerations, the article also reflects on the broader implications of AI-driven text analysis for social science research, including questions of scalability, reproducibility, interpretability, and the relationship between language and political behavior.

The structure of the article is as follows. Section 1 introduces the basic concepts and methodological foundations of quantitative text analysis, outlining its key approaches and analytical logic. Section 2 reviews how these methods have been applied in international relations, with particular emphasis on large-scale analyses of United Nations documents and related datasets. Section 3 discusses recent developments in deep learning and large language models, and presents illustrative applications using United Nations Security Council speech data. Section 4 considers the broader implications of these technological advances for IR research and outlines future directions. The final section concludes by summarizing the main arguments and reflecting on the prospects and limitations of quantitative text analysis in the study of international relations.

## 2 What Is Quantitative Text Analysis?

In contemporary societies characterized by rapid digitalization, an enormous variety of textual data has accumulated as potential research resources. Emails, social media posts, online forums, and digitized books are now joined by large collections of parliamentary debates, administrative records, and official documents that were once available only in printed form. Quantitative text analysis—often referred to as *text as data*—is an approach that seeks to extract systematic and reproducible patterns from such large-scale textual information using computational methods (Grimmer and Stewart, 2013; Lucas et al., 2015; Evans and Aceves, 2016; Wilkerson and Casas, 2017).

At its core, quantitative text analysis involves transforming linguistic units—such as words, sentences, and documents—into numerical representations that can be analyzed statistically. Common representations include word frequency counts, co-occurrence matrices, and high-dimensional vector embeddings. Once text is represented in numerical form, researchers can apply statistical models and machine-learning algorithms to identify patterns, classify documents, estimate latent structures, or trace changes over time. Compared to traditional qualitative content analysis, this approach offers key advantages in terms of scalability, reproducibility, and transparency, enabling

researchers to analyze text corpora that are far too large to be processed manually.

The analytical techniques used in quantitative text analysis can be broadly divided into two categories. The first consists of dictionary-based methods and frequency-based content analysis, while the second relies on supervised or unsupervised machine-learning approaches for document classification and topic discovery.

A classic example of the former is sentiment analysis, which uses predefined dictionaries to identify emotional tone or evaluative content in text. By counting the occurrence of words associated with positive or negative affect, moral values, or specific frames, researchers can quantify aspects of language that were previously examined qualitatively. For instance, Golder and Macy (2011) analyzed large-scale Twitter data to demonstrate that collective emotional expressions vary systematically across daily and seasonal cycles. Dictionary-based sentiment analysis has since been widely applied to social media data, legislative debates, and political speeches to measure evaluative tone and moral framing (Takikawa and Sakamoto, 2019). While these methods are relatively transparent and easy to implement, their effectiveness depends heavily on the quality and contextual appropriateness of the dictionaries used.

To capture more complex patterns in text, machine-learning-based approaches have become indispensable. A landmark example in the social sciences is the study by King et al. (2013), which examined online censorship in China using millions of social media posts collected from more than 1,400 platforms. By automatically classifying posts that were deleted by government censors, the authors showed that critical expressions directed at the state or political leaders were not systematically suppressed. Instead, censorship targeted content with the potential to facilitate collective action. This finding—made possible only through large-scale automated text classification—illustrates the distinctive analytical power of quantitative text analysis when applied to massive text corpora.

Topic modeling represents another foundational technique in quantitative text analysis. Topic models infer latent themes within a corpus by analyzing patterns of word co-occurrence, and estimate the degree to which each document is associated with multiple topics. Grimmer (2010), for example, developed a Bayesian hierarchical topic model to analyze more than 24,000 press releases issued by U.S. senators, demonstrating how individual legislators differ in their policy agendas and communication styles. Among the most widely used topic modeling approaches are Latent Dirichlet Allocation (LDA) and the Structural Topic Model (STM). Both rely on probabilistic frameworks to estimate topic distributions, while STM additionally allows researchers to incorporate document-level metadata and covariates in estimating topic prevalence and content (Blei et al., 2003; Roberts et al., 2014; Lucas et al., 2015). These models have been applied to a wide range of political texts, including policy documents, news articles, legislative debates, and social media posts.

An illustrative application is provided by Catalinac (2016), who analyzed approximately 7,500 Japanese electoral manifestos using LDA. Her study showed how candidates' issue emphasis shifted from pork-barrel appeals toward policy-oriented competition following electoral reform, demonstrating how topic models can reveal systematic changes in political communication over time.

Network-based approaches have also gained traction in recent years. Rule et al. (2015), for instance, constructed word co-occurrence networks from U.S. State of the Union addresses spanning more than two centuries, from George Washington to Barack Obama. By examining the evolving structure of these lexical networks, they traced both continuity and transformation in American political discourse. Network analysis thus provides an alternative lens through which semantic relationships and discursive change can be examined.

Overall, these methods constitute what may be described as the “classical” foundations of quantitative text analysis, developed prior to the widespread adoption of Transformer-based deep learning models and large language models (LLMs). Despite important methodological differences, they share a common analytical logic: textual data are first converted into numerical representations, which then allow researchers to identify patterns such as frequently used terms, similarities between documents, or shifts in thematic emphasis across time and actors. These foundational techniques have underpinned a wide range of empirical applications in the social sciences, including analyses of agenda-setting, issue competition, and the temporal dynamics of political discourse.

### **3 Applications of Quantitative Text Analysis in International Relations: What Can Be Done?**

#### **3.1 The Expansion of Quantitative Text Analysis in International Relations**

The methods of quantitative text analysis outlined above have been increasingly applied across a wide range of research areas in international relations. Traditionally, the study of diplomacy, foreign policy, and international institutions relied heavily on qualitative analyses of selected cases, often based on close reading of historical documents or elite interviews. While such approaches remain indispensable, advances in natural language processing and statistical text analysis have made it possible to examine international phenomena at a much larger scale and with greater systematicity.

One prominent line of research focuses on the comparative analysis of diplomatic documents, policy statements, and international agreements. For example, Allee et al. (2017) conducted a large-scale textual comparison of preferential trade agreements and World Trade Organization (WTO) agreements, revealing systematic differences in their legal structure and issue coverage. Similarly, Bunea and Ibenskas (2015) applied quantitative text analysis to lobbying documents submitted to European Union institutions, shedding light on the dynamics of policy debates and the characteristics of interest group participation.

Media coverage and public discourse constitute another major domain of application. Trubowitz and Watanabe (2021) analyzed newspaper articles to extract changing geopolitical threat narratives, while Terman (2017) examined representations of Muslim women in U.S. media to trace the formation and diffusion of international norms and stereotypes. In recent years, sentiment analysis and topic modeling have also been applied to social media posts and online news in order to capture the diffusion of international narratives and the dynamics of external image formation (Deutschmann et al., 2020).

Quantitative text analysis has also been used to study crisis management, elections, and decision-making processes in international politics. Katagiri and Min (2019), for instance, applied machine-learning techniques to Cold War-era U.S. White House documents to assess the credibility of public and private signals in foreign policy crises. Yukawa and Sakamoto (2024) analyzed a large corpus of international election observation reports to trace changes in monitoring standards and issue emphasis over time using topic models and related techniques. Other studies have estimated policy positions from textual data, such as Barnum and Lo (2020)’s analysis of statements at Nuclear Non-Proliferation Treaty (NPT) review conferences, or have examined racialized language in diplomatic documents (Carson et al., 2024).

In sum, these studies illustrate how quantitative text analysis has become an established methodological tool for examining a broad range of international phenomena, including diplomacy, security, international political economy, human rights, media discourse, and civil society. From fine-grained analyses of specific policy debates to global overviews of normative and discursive trends, text-based approaches now play a central role in contemporary IR research.

### 3.2 The Development and Release of United Nations Text Datasets

Among the many applications of quantitative text analysis in international relations, the study of United Nations documents has become one of the most vibrant and rapidly expanding research areas. Since its founding, the United Nations has produced a vast number of resolutions, meeting records, and official statements, most of which are in the public domain. Through platforms such as the Official Document System of the United Nations (ODS) and the United Nations Digital Library, researchers can access documents systematically across time periods and issue areas.<sup>23</sup>

Building on this growing accessibility, a number of large-scale, machine-readable corpora of UN documents have been constructed in recent years, enabling diverse forms of quantitative analysis. One prominent example is the United Nations General Debate corpus, which covers speeches delivered during the General Assembly’s annual general debate sessions (Baturu et al., 2017; Jankin et al., 2025). These datasets have made it possible to analyze longitudinal changes in diplomatic rhetoric, issue salience, and policy orientations across member states. Extending this approach beyond high-level speeches, Turco (2024) compiled the UNGA Speech Corpus, which includes statements from formal General Assembly meetings between 1993 and 2018, thereby capturing more routine forms of diplomatic discourse by government representatives.

Security Council documents constitute another major focus of data collection efforts. Several projects have compiled corpora of Security Council meeting records and speeches, providing the empirical foundation for large-scale analyses of international security deliberations (Sakamoto et al., 2025; Schönfeld et al., 2025). Resolution texts have also been systematically collected and analyzed. Hanania (2021) assembled a comprehensive corpus of Security Council resolutions and applied topic modeling to trace the emergence of humanitarian and human rights norms after the Cold

---

<sup>2</sup>Official Document System of the United Nations (ODS), <https://documents.un.org> (accessed May 28, 2025).

<sup>3</sup>United Nations Digital Library, <https://digitallibrary.un.org> (accessed May 28, 2025).



War. Other datasets cover both Security Council and General Assembly resolutions, enabling comparative analyses across organs and issue areas (Bethke et al., 2024; Knapp, 2024; Mesquita and Pires, 2025; Fobbe et al., 2025). In addition, specialized datasets have been developed for particular institutional contexts, such as speeches in the First Committee of the General Assembly on disarmament and international security (Finke and Risse, 2024).

Among these efforts, the *UNSC Meetings and Speeches* dataset constructed and released by the author and collaborators (Sakamoto and Matsuoka, 2023; Sakamoto et al., 2025) represents one of the largest text datasets currently available in international relations research. Covering the entire period from 1946 to 2024, the dataset includes basic information on all official Security Council meetings—such as dates, agenda items, and presiding countries—as well as nearly all recorded speeches by member states and major international organizations. In total, the dataset comprises approximately 170,000 individual speeches containing roughly 88 million words, making it the largest publicly available corpus of deliberative texts in global politics to date.<sup>4</sup>

The availability of such large-scale datasets has fundamentally expanded the analytical possibilities of research on the United Nations. Whereas earlier studies were often limited to selected cases or short time periods, researchers can now examine long-term patterns of norm formation, policy preferences, and discursive dynamics in a comprehensive and reproducible manner. At the same time, it is important to note certain characteristics of UN texts that shape their suitability for quantitative analysis. Most official UN documents are produced in a shared working language—primarily English—and are typically drafted by professional diplomats or international civil servants. As a result, they often contain standardized formulations, diplomatic conventions, and references to established precedents. While this degree of standardization may limit the expression of individual idiosyncrasies, it also facilitates the identification of shared frames, institutional norms, and recurring patterns in international discourse. The implications of these textual characteristics for interpretation and inference are revisited in the concluding section.

### 3.3 Substantive Contributions of Quantitative Text Analysis Using United Nations Data

As emphasized in influential methodological syntheses such as *Text as Data* (Grimmer et al., 2022), quantitative text analysis can be understood in terms of its functions within research design, including discovery, measurement, and inference. In this article, however, a different organizing principle is adopted. Rather than focusing on methodological functions per se, this subsection classifies existing studies according to the *substantive contributions* they have made to international relations research through the analysis of United Nations texts.

Specifically, the literature can be grouped into five major areas of contribution, each of which addresses core questions in IR concerning norms, policy preferences, actors, and outcomes.

---

<sup>4</sup>The dataset is publicly available via Harvard Dataverse (<https://doi.org/10.7910/DVN/CKPTRB>).

### **3.3.1 Contribution 1: Quantifying the Interpretation and Evolution of Concepts and Norms**

First, quantitative text analysis has enabled researchers to systematically examine how abstract international norms and policy concepts are interpreted and how their meanings evolve over time. By analyzing patterns of language use across large corpora of UN documents, scholars can trace shifts in conceptual associations, framing, and semantic emphasis that would be difficult to detect through qualitative reading alone.

For example, Sakamoto (2023) analyzed official Security Council speeches in the post–Cold War period to examine how the concept of “threat” has been articulated in relation to other issues such as terrorism and climate change. Using word embedding techniques, the study demonstrated that the meaning of “threat” varies significantly across time periods and actors, reflecting changing understandings of security in global discourse. Similarly, Bethke et al. (2024) traced the language of “responsibility” in Security Council resolutions, documenting how both the frequency and the targets of responsibility-related language shifted in the post–Cold War era, thereby illuminating processes of normative transformation.

A growing body of work has applied similar approaches to other core concepts debated within the United Nations. Studies have examined the discursive construction of gender (Brun-Mercer, 2021), climate change (Arias, 2022), health (Voss et al., 2022), humanitarianism (Knapp, 2024), and security threats (Sundström et al., 2025), among others. Taken together, these analyses demonstrate how quantitative text analysis can make the evolution of international norms empirically visible and comparable across issue areas and institutional contexts.

### **3.3.2 Contribution 2: Visualizing the Acceptance and Rejection of Norms and Policies**

Second, quantitative text analysis allows researchers to systematically identify patterns of acceptance, contestation, and rejection of international norms and policies by different actors. By extracting evaluative language, framing devices, or explicit expressions of support and opposition, scholars can map how states and other actors respond to emerging norms within multilateral forums.

Kentikelenis and Voeten (2020), for instance, used UN General Assembly general debate speeches to trace historical shifts in member states’ attitudes toward the liberal international order, revealing periods of convergence as well as growing contestation. Focusing on the Responsibility to Protect (R2P), Scherzinger (2023) analyzed Security Council debates to examine patterns of support and resistance, showing how acceptance of R2P expanded after its introduction while remaining subject to persistent contestation and reinterpretation.

Related work has also examined more critical or confrontational forms of discourse. Verbeek (2024), for example, investigated practices of blaming and criticism in Security Council debates, with particular attention to China and other emerging powers. These studies highlight how quantitative text analysis can capture not only consensus-building but also conflictual dynamics in institutionalized international discourse.



### **3.3.3 Contribution 3: Estimating Policy Preferences and Actor Differences from Text**

Third, recent advances in quantitative text analysis have made it possible to estimate policy preferences and positions directly from textual data and to visualize differences among actors as distances in a policy space. This represents a major methodological advance in the study of international politics, where preferences are often difficult to observe directly.

Medzihorsky et al. (2017) analyzed Security Council debates on the Syrian conflict using scaling techniques to identify policy dimensions and patterns of alignment among major powers. Similarly, Baturu et al. (2017) used the UN General Debate corpus to estimate similarities and differences in states' policy positions based on speech content. Turco (2024) extended this approach by analyzing speeches from formal General Assembly meetings, revealing changes in countries' stances on climate change and other issues over time.

More recently, Finke and Risse (2024) applied multidimensional scaling techniques to speeches in the First Committee of the General Assembly, demonstrating that conflicts over disarmament and international security cannot be captured by a single policy dimension. These studies illustrate how text-based measures of policy preferences can complement—and in some cases substitute for—traditional indicators such as voting behavior or alliance patterns.

### **3.3.4 Contribution 4: Tracing the Diffusion and Dynamics of Norms and Policies**

Fourth, large-scale text analysis of UN documents has enabled systematic examination of how specific norms and policy themes diffuse and evolve over time. By applying topic models and related techniques to resolutions and speeches, researchers can identify long-term trends, turning points, and cycles of attention within international institutions.

Hanania (2021) provided a seminal contribution by applying topic modeling to Security Council resolutions, showing how humanitarian and human rights themes gained prominence after the Cold War. Knapp (2024) analyzed UN policy documents related to humanitarian emergencies to trace the evolution of framing around civilian protection. Bethke et al. (2024) similarly documented the expansion and diversification of responsibility-related language across UN resolutions, capturing both quantitative growth and qualitative transformation in normative discourse.

Overall, these studies demonstrate how quantitative text analysis can move beyond static snapshots to capture the dynamic life cycle of norms and policy ideas within international organizations.

### **3.3.5 Contribution 5: Linking Discourse to Policy Outcomes and Behavior**

Finally, a growing body of research seeks to connect discursive patterns in UN texts to concrete policy outcomes and behavioral measures, such as resolutions, authorizations of force, or aid allocation. This line of work addresses a central concern in IR: whether and how language matters for political action.

Scherzinger (2022), for example, examined the relationship between rhetorical strategies in Security Council debates and the authorization of force under Chapter VII, showing how discourse

can function as a precondition for, or a constraint on, coercive action. Mitra (2024) analyzed General Assembly speeches to demonstrate that the content of states' rhetoric is systematically associated with patterns of development aid allocation, offering new evidence on how language influences international resource distribution.

Complementing these efforts, Lundgren et al. (2024) introduced the Intergovernmental Policy Output Dataset (IPOD), which systematically codes policy outputs across international organizations. When combined with large-scale text data, such datasets open new avenues for analyzing correlations and causal relationships between discourse and institutional outcomes.

Together, these five areas of contribution illustrate how quantitative analysis of United Nations texts has generated substantive insights into some of the most fundamental questions in international relations, including norm evolution, policy preferences, institutional dynamics, and political outcomes. In recent years, rapid advances in natural language processing have further expanded the analytical toolkit available to researchers. The next section turns to these developments, focusing on deep learning and large language models and their implications for the analysis of UN texts.

## 4 The Frontier of Quantitative Text Analysis: Illustrations and Applications

### 4.1 Advances in Natural Language Processing: The Impact of Deep Learning and Transformers

Since the 2010s, rapid advances in deep learning have fundamentally transformed the field of natural language processing (NLP).<sup>5</sup> Among these developments, the introduction of the Transformer architecture in 2017 marked a particularly important turning point (Vaswani et al., 2017). Unlike earlier recurrent neural networks (RNNs) or convolutional neural networks (CNNs), Transformers made it possible to model long-range dependencies and complex contextual relationships within text more effectively. As a result, they enabled models to capture both global structure and fine-grained semantic nuances in language data with unprecedented accuracy.

Following the emergence of the Transformer architecture, a series of influential models were developed, including BERT (Bidirectional Encoder Representations from Transformers) and the GPT (Generative Pre-trained Transformer) family. These models established a new paradigm in NLP based on large-scale pre-training on massive text corpora, followed by fine-tuning for specific downstream tasks. This two-stage framework substantially reduced the need for large task-specific labeled datasets, which had previously been a major bottleneck in applying advanced NLP techniques to domain-specific research problems.

The diffusion of pre-trained language models has enabled a wide range of NLP tasks—such as text classification, summarization, machine translation, question answering, and dialogue generation—

---

<sup>5</sup>A large number of textbooks and manuals have been published on deep learning and advanced NLP techniques, including implementation-oriented guides. For accessible introductions, see Raschka (2025).

to be performed with a single model architecture. Moreover, recent large language models (LLMs) exhibit strong generalization capabilities, allowing them to adapt to new tasks through zero-shot or few-shot learning based solely on task descriptions or a small number of examples. These properties have dramatically lowered the technical barriers to applying state-of-the-art NLP methods, not only within computer science but also across the social sciences.

As a consequence, analytical practices that previously required specialized expertise in NLP have become increasingly accessible through open-source libraries and cloud-based application programming interfaces (API). This shift has accelerated methodological innovation and experimentation in fields such as political science and international relations, where text data play a central empirical role.

## 4.2 The Adoption of Deep Learning–Based Text Analysis in Political Science and International Relations

Deep learning–based language models, particularly those built on Transformer architectures, have gradually been incorporated into quantitative text analysis in political science and international relations. Compared to dictionary-based methods or traditional machine-learning approaches, these models offer substantial improvements in contextual understanding, enabling more accurate classification, scaling, and semantic analysis of political texts.

Several studies have systematically evaluated the performance of Transformer-based models in political science applications. Timoneda and Vallejo Vera (2025) compared major Transformer variants across a range of political text corpora and demonstrated that models such as RoBERTa and XLM-RoBERTa consistently outperform earlier approaches in classification tasks. Similarly, Laurer et al. (2024) and Gupta et al. (2020) showed that BERT-based models can achieve high classification accuracy even with relatively small labeled datasets, addressing a long-standing data scarcity problem in supervised text analysis.

Beyond performance benchmarks, recent research has explored the substantive implications of applying deep learning models to political texts. Heseltine and von Hohenberg (2024) examined the use of large language models as substitutes for human experts in annotating political text and found high levels of agreement between LLM-based coding and expert judgments in many tasks.<sup>6</sup> Le Mens and Gallego (2025) proposed an approach to estimating policy positions by prompting LLMs with structured questions and aggregating their responses, illustrating how generative models can be integrated into established scaling frameworks.

Applications to international relations have also begun to emerge. Kertcher and Zwilling (2025), for example, used BERT-based models to analyze emotional polarity and policy stances in United Nations speeches related to the Russia–Ukraine war, revealing systematic differences across regions and actor types. These studies suggest that deep learning–based models can substantially enhance the precision and scalability of quantitative text analysis in IR, while also raising new questions

---

<sup>6</sup>An innovative attempt to outsource political text annotation to large numbers of crowd workers, rather than relying exclusively on experts, is provided by Benoit et al. (2016).

about validation, transparency, and interpretability.

At the same time, the adoption of large language models introduces a set of methodological challenges that require careful consideration. Well-known issues include hallucination—where models generate plausible but unsupported outputs—as well as the reproduction of social and political biases embedded in training data. In addition, the complexity of deep neural networks poses persistent problems for explainability, making it difficult to trace how specific inputs lead to particular outputs. Despite these limitations, the rapid diffusion of LLMs has already reshaped the practical frontier of quantitative text analysis, redefining what levels of scale, granularity, and automation are feasible in social science research.

### 4.3 Illustrative Applications Using United Nations Security Council Data

As discussed above, advances in deep learning and Transformer-based NLP models have substantially expanded the analytical possibilities of quantitative text analysis in international relations. However, applications of these state-of-the-art methods remain relatively recent, and their substantive value must be demonstrated through concrete empirical examples. This subsection presents two illustrative analyses based on the *UNSC Meetings and Speeches* dataset (Sakamoto and Matsuka, 2023; Sakamoto et al., 2025), which covers all official Security Council meetings and speeches from 1946 to 2024. Together, these examples demonstrate how large language models and contextualized embedding techniques can be used to analyze normative change and agenda dynamics in international security deliberations at an unprecedented scale.

#### 4.3.1 Illustration 1: The Transformation of Sovereignty Discourse—From Rights to Responsibility?

In international politics, sovereignty has traditionally been understood as a set of rights associated with state independence, territorial integrity, and non-intervention. This rights-based conception of sovereignty has long served as a foundational principle of the international system. By contrast, the emergence of the Responsibility to Protect (R2P) norm in the early 2000s introduced an alternative framing, emphasizing that sovereignty entails responsibilities, particularly the responsibility of states to protect civilian populations from mass atrocities (International Commission on Intervention and State Sovereignty, 2001; United Nations General Assembly, 2005).

While a growing literature has examined the acceptance and contestation of R2P in Security Council deliberations (Medzihorsky et al., 2017; Scherzinger, 2023; Bethke et al., 2024), few studies have systematically investigated whether and to what extent the introduction of R2P was accompanied by a broader shift in the use of sovereignty-related language itself. In particular, it remains unclear whether sovereignty discourse in the Security Council moved away from a predominantly rights-based conception toward a responsibility-based understanding, and whether such changes varied across time periods and actors.

To address this gap, the analysis focuses on all sentences in Security Council meeting records between 1946 and 2024 that contain sovereignty-related terms such as “sovereign” or “sovereignty.”

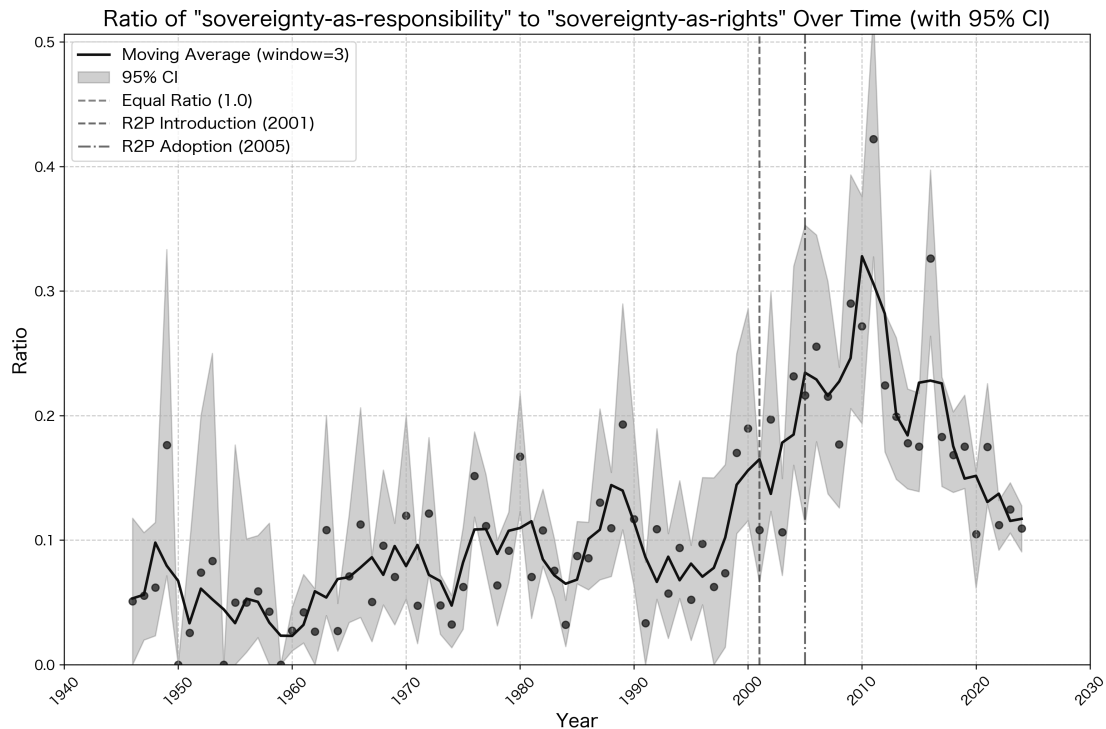


Figure 1: Sovereignty Discourse in UNSC Debates: Responsibility vs. Rights over Time

Out of approximately 3.34 million sentences in the full corpus, 32,944 sentences met this criterion. Each sentence was classified into one of three categories: **sovereignty as rights**, **sovereignty as responsibility**, or **neutral**.

The classification was conducted using a state-of-the-art open-weight large language model operating in a local environment (**phi-4**).<sup>7</sup> The model was prompted using a few-shot approach, in which five manually selected prototypical examples were provided for each category. Based on these examples, the model automatically classified the remaining sentences. This approach enables scalable and consistent coding of large text corpora that would be prohibitively costly to analyze through expert manual coding alone.

Using the sentence-level classification results, the analysis examines temporal trends and actor-specific patterns in sovereignty discourse. Figure 1 shows the relative frequency of sovereignty-as-responsibility compared to sovereignty-as-rights over time. While rights-based sovereignty language dominates throughout the entire period, responsibility-based references increase noticeably around the late 1990s and early 2000s, coinciding with the emergence of the R2P norm, before declining again in the 2010s.

Figure 2 disaggregates these patterns by actor type, distinguishing between Western permanent members, non-Western permanent members, other states, and UN Secretariat officials. The results

<sup>7</sup>**phi-4** is a lightweight open-weight large language model released by Microsoft. Other local LLMs, including **mistral-small-3.1** and **gemma-3:27B**, were also tested, but **phi-4** offered the best balance of classification accuracy and computational efficiency for this task.

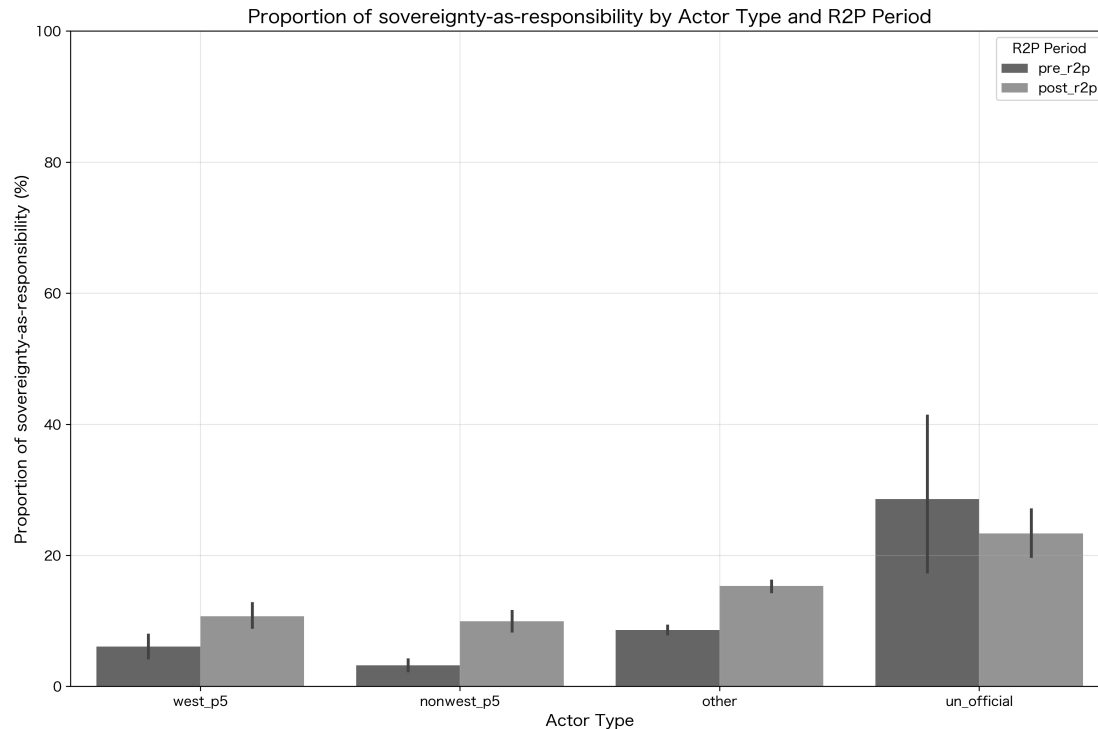


Figure 2: Sovereignty Discourse by Actor Type

indicate that references to sovereignty as responsibility are most frequently articulated by officials from the UN Secretariat, including the Secretary-General and senior representatives. By contrast, permanent members of the Security Council—both Western and non-Western—exhibit comparatively lower and more stable levels of responsibility-based sovereignty language. These patterns suggest that international bureaucratic actors played a particularly important role in promoting responsibility-oriented interpretations of sovereignty, while major powers remained more cautious in their rhetorical adoption.

Overall, this analysis demonstrates that while the R2P norm temporarily reshaped sovereignty discourse in the Security Council, its impact was limited in scope and durability. More broadly, the illustration shows how large language models can be used to classify normative language consistently across decades of deliberative text, enabling systematic analysis of conceptual change in international politics.

#### 4.3.2 Illustration 2: The “Humanitarian Turn” and Its Proponents

A second illustration focuses on what has been described as the “humanitarian turn” in the Security Council’s agenda—that is, the growing prominence of humanitarian, human rights, and civilian protection issues in post–Cold War deliberations (Hanania, 2021; Bethke et al., 2024; Knapp, 2024). Much of the existing literature on this transformation has relied on analyses of policy outputs, particularly resolutions. While such studies have yielded important insights, they provide only a



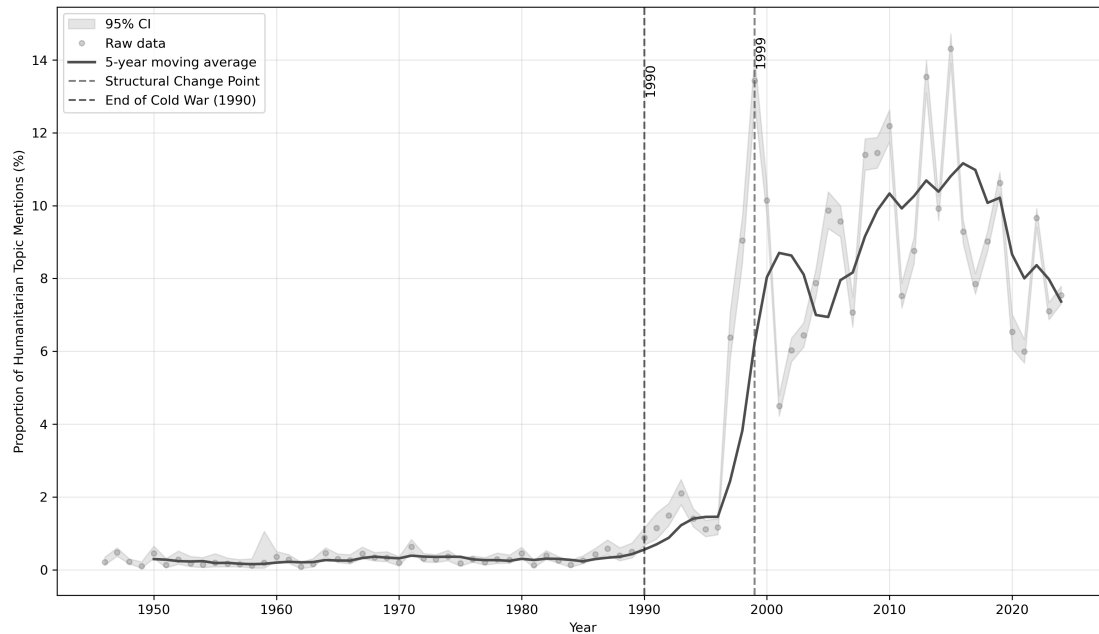


Figure 3: Humanitarian Topics in UNSC Debates over Time (BERTopic)

partial view of how humanitarian norms were debated and promoted in the day-to-day deliberative practices of the Council.

To examine these dynamics at the level of deliberation, this analysis applies a deep learning-based topic modeling approach to the full corpus of Security Council speeches. Specifically, the analysis uses BERTopic, a document embedding-based topic model that leverages Transformer-based sentence representations to cluster texts based on semantic similarity (Grootendorst, 2022). Compared to traditional topic models such as LDA or STM, BERTopic offers substantially improved contextual sensitivity, allowing it to capture complex thematic structures in long and heterogeneous political texts.

The model was applied to approximately 1.06 million paragraphs extracted from Security Council meeting records. From the resulting topic clusters, a set of humanitarian-related topics was identified, including civilian protection, women and peace, children in armed conflict, and the protection of journalists and media workers.

Figure 3 presents the temporal dynamics of humanitarian topic references across the entire period from 1946 to 2024. During the Cold War, humanitarian topics appear only sporadically. Following the end of the Cold War, however, their frequency increases steadily, with a particularly sharp rise around 1999, coinciding with the adoption of key resolutions on civilian protection (S/RES/1265) and children in armed conflict (S/RES/1261). Statistical change-point analysis confirms that this period represents a structural break in the Council's agenda.

More recent years show a relative decline in the salience of humanitarian topics, mirroring the patterns observed in the sovereignty analysis above. This suggests that the post-Cold War expansion of humanitarian norms was not linear, but subject to periods of consolidation and re-

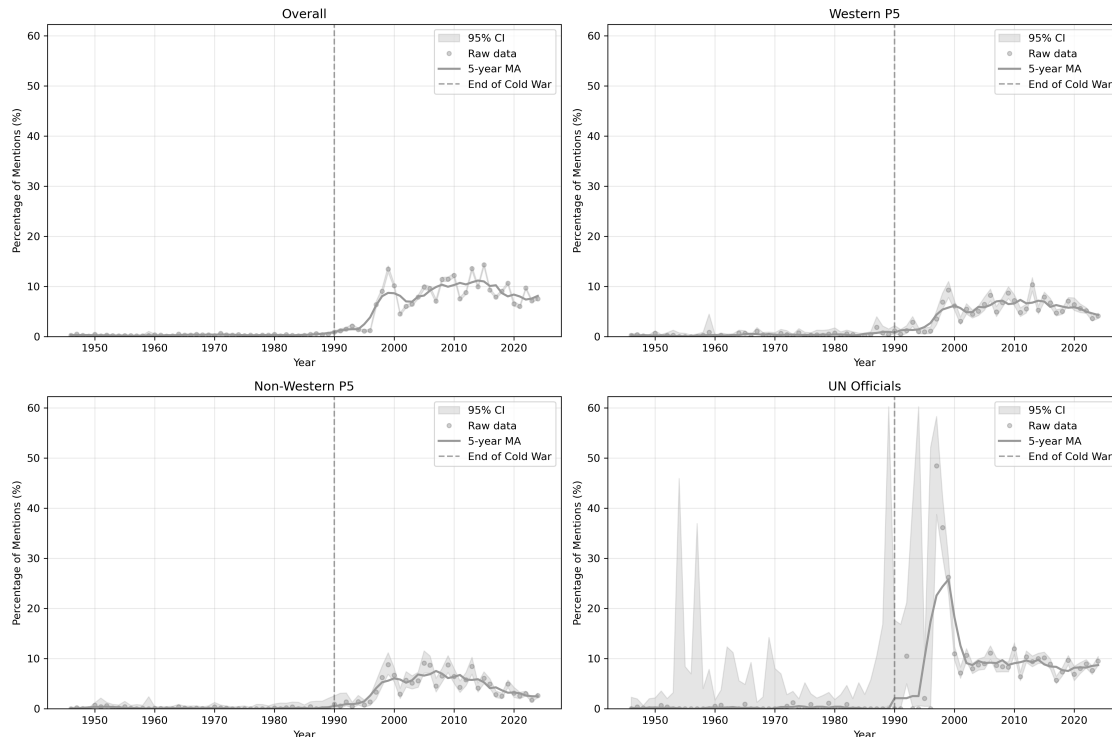


Figure 4: Humanitarian Topics by Actor Type

trenchment.

Figure 4 disaggregates humanitarian topic references by actor type. As with sovereignty discourse, officials from the UN Secretariat exhibit markedly higher levels of engagement with humanitarian topics than permanent members of the Council. Differences among permanent members are comparatively modest, indicating that the humanitarian turn was not driven primarily by any single major power. Instead, the findings point to the central role of international bureaucratic actors in sustaining humanitarian agendas within Security Council deliberations.

By shifting the analytical focus from resolutions to deliberative speech, this illustration complements existing studies and provides a more fine-grained view of how humanitarian norms were articulated, promoted, and contested in practice. More broadly, it demonstrates how deep learning–based topic models can uncover long-term agenda dynamics that remain difficult to capture through manual coding or traditional text analysis methods.

## 5 The Future of Quantitative Text Analysis: Beyond Analytical Tools

Recent advances in quantitative text analysis are inseparable from the rapid evolution of artificial intelligence (AI), particularly deep learning–based natural language processing. NLP has become one of the central application domains of contemporary AI, and innovations in this field have

increasingly spilled over into other areas of AI research. As a result, text analysis is no longer confined to written language alone, but is increasingly integrated into multimodal systems capable of processing images, audio, and video. The growing sophistication of NLP thus reflects, and contributes to, broader transformations in data-driven research.

Much of the discussion so far has focused on the use of NLP as an analytical tool—for classification, topic modeling, or semantic analysis. However, the implications of AI for social science research extend far beyond methodological efficiency. The emergence of generative AI, multimodal large language models, and agent-based AI systems has begun to reshape the organization of research itself, raising fundamental questions about how social science knowledge is produced.

One increasingly prominent development is the use of AI systems as collaborative research partners rather than as mere analytical instruments. In the natural sciences and computational fields, experimental platforms have already been developed in which AI systems assist with literature review, hypothesis generation, experimental design, data analysis, and report writing, often in an integrated workflow. Gottweis et al. (2025), for example, propose the concept of an “AI co-scientist,” illustrating how AI systems can be embedded into the research process to reduce costs and time requirements while maintaining analytical rigor.

In the social sciences, and particularly in international relations, full automation of the research process remains limited. Nevertheless, important elements of this transformation are already observable. One notable trend is the increasing use of large language models and Transformer-based classifiers to replace or complement expert manual coding. Studies such as Heseltine and von Hohenberg (2024) demonstrate that LLM-based annotations often achieve high levels of agreement with expert judgments, especially in well-defined classification tasks. While such developments may appear incremental, they substantially expand the scale and granularity of data that can be analyzed, enabling research designs that were previously infeasible due to resource constraints.

Beyond coding and classification, AI systems are increasingly used for brainstorming, hypothesis generation, drafting research manuscripts, and even assisting with complex quantitative analyses. Many recent models claim to possess reasoning capabilities comparable to those of graduate-level researchers in mathematics and statistics, and in practice, researchers have begun to delegate portions of model specification, estimation, and diagnostic analysis to AI-assisted workflows. These developments lower the barriers to entry for advanced quantitative research, contributing to a partial democratization of social science methodology. At the same time, they raise new questions about authorship, accountability, and the validation of AI-generated outputs.

More radical possibilities are suggested by recent work on agent-based AI systems. Lu et al. (2024) introduce an “AI Scientist” capable of autonomously conducting large portions of the scientific research process, from hypothesis generation to experimental execution and manuscript preparation. Similarly, Swanson et al. (2024) present a “Virtual Lab” in which AI agents design and validate biomedical experiments, while Schmidgall et al. (2025) propose an “Agent Laboratory” framework in which multiple AI agents collaborate as research assistants. Although these systems have thus far been developed primarily in machine learning and natural science contexts,

they point toward a future in which research workflows themselves may be partially automated or reconfigured.

For international relations and other social sciences, these developments present both opportunities and challenges. On the one hand, AI-driven tools promise dramatic gains in efficiency, scale, and analytical scope, particularly in text-rich domains such as diplomacy, international organizations, and global governance. On the other hand, the increasing opacity and autonomy of AI systems complicate core epistemological and methodological assumptions. Quantitative text analysis typically rests on the assumption that language reflects underlying preferences, beliefs, or policy positions. Yet political speech and official documents are often shaped by strategic considerations, institutional constraints, and standardized diplomatic conventions. The use of increasingly complex AI systems may further obscure the relationship between observed language and underlying political processes.

Moreover, the widespread use of UN documents in quantitative text analysis raises additional interpretive challenges. Most official UN texts are written in a shared working language—primarily English—by speakers operating within professional diplomatic and bureaucratic settings. This results in a high degree of formulaic expression and precedent-based phrasing. While such standardization may limit the expression of individual nuance, it also facilitates the identification of shared frames, institutional norms, and recurring discursive patterns. Understanding how these textual characteristics shape analytical results remains an important task for future research.

Despite these challenges, the technological foundations of quantitative text analysis continue to advance rapidly. Techniques that were considered impractical only a few years ago—such as large-scale automated coding, AI-assisted hypothesis generation, or model design based on natural language instructions—are increasingly becoming feasible. These developments expand the space of possible research designs and encourage more experimental and creative approaches to the study of international politics.

Looking ahead, quantitative text analysis is likely to become even more firmly embedded as a core research tool in international relations, while simultaneously prompting reflection on the nature of research practice itself. As Grimmer et al. (2022) emphasize, text analysis should remain embedded within a carefully designed research framework and subject to human oversight. Yet as NLP and AI technologies continue to evolve at an accelerating pace, maintaining such control will become an increasingly complex challenge. Navigating this tension—between analytical empowerment and methodological responsibility—will be central to the future development of quantitative text analysis in international relations.

## 6 Conclusion

This article has reviewed the development, applications, and future prospects of quantitative text analysis in international relations. Against the backdrop of the large-scale digitization of text and rapid advances in natural language processing, quantitative approaches to textual data have

become an indispensable component of contemporary social science research. In international relations in particular, the systematic analysis of text has enabled scholars to examine norms, policy preferences, institutional dynamics, and political behavior in ways that were previously difficult to achieve through qualitative methods alone.

A central focus of the review has been the growing body of research that analyzes United Nations documents, including resolutions, meeting records, and official speeches. The availability of large, publicly accessible, and longitudinal UN text corpora has substantially expanded the empirical foundations of IR research. By leveraging these data, scholars have been able to trace the evolution of international norms, estimate policy positions of states and organizations, identify agenda dynamics, and explore the relationship between discourse and concrete policy outcomes. The examples discussed in this article illustrate how quantitative text analysis has generated substantive insights into core questions of international politics.

The article has also highlighted how recent advances in deep learning and large language models have transformed the methodological frontier of text analysis. Transformer-based models and LLMs enable more nuanced contextual understanding, scalable classification, and semantically rich topic discovery than earlier approaches. Through illustrative analyses of Security Council debates, the article has shown how these techniques can be applied to examine normative change—such as shifting conceptions of sovereignty—and agenda transformations, including the rise and partial retrenchment of humanitarian concerns.

At the same time, the expanding use of AI-driven text analysis raises important challenges. Methodological concerns related to bias, hallucination, and interpretability remain salient, as do more fundamental questions about how language relates to political preferences and behavior. Official texts produced in diplomatic and bureaucratic settings are shaped by institutional conventions, strategic considerations, and standardized formulations. While these characteristics impose limits on inference, they also create opportunities for identifying shared frames and institutional norms. Careful attention to these trade-offs is essential when interpreting results from large-scale text analysis.

Looking forward, quantitative text analysis is likely to play an increasingly prominent role in international relations research. Advances in NLP and AI are rapidly expanding the range of feasible research designs, enabling analyses at scales and levels of detail that were previously unattainable. At the same time, these developments challenge researchers to reflect on the role of automation, human judgment, and responsibility in the production of social scientific knowledge. As quantitative text analysis continues to evolve, its greatest contribution may lie not only in new empirical findings, but also in prompting renewed reflection on how international politics can be studied in an era of data abundance and artificial intelligence.

## Acknowledgements

The author is grateful to Professor Atsushi Ishida for his careful reading of an earlier draft of this paper and for his valuable comments. The ideas underlying the analysis of sovereignty discourse benefited greatly from discussions with Tomoyuki Matsuoka, who also prepared the few-shot examples and validation data used in the large language model-based classification. This research was supported by JSPS KAKENHI Grant No. 23K22087 and by the University of Tokyo's Institute for Digital Observatory.

## References

- Allee, T., Elsig, M., and Lugg, A. (2017). The ties between the world trade organization and preferential trade agreements: A textual analysis. *Journal of International Economic Law*, 20(2):333–363.
- Arias, S. B. (2022). Who securitizes? climate change discourse in the united nations. *International Studies Quarterly*, 66(2).
- Barnum, M. and Lo, J. (2020). Is the npt unraveling? evidence from text analysis of review conference statements. *Journal of Peace Research*, 57(6):740–751.
- Baturo, A., Dasandi, N., and Mikhaylov, S. J. (2017). Understanding state preferences with text as data: Introducing the UN general debate corpus. *Research & Politics*, 4(2):1–9.
- Benoit, K., Conway, D., Lauderdale, B. E., Laver, M., and Mikhaylov, S. (2016). Crowdsourced text analysis: Reproducible and agile production of political data. *American Political Science Review*, 110(2):278–295.
- Bethke, F. S., Haass, F., and Niemann, H. (2024). The language of responsibility in the united nations security council, 1946–2020. *International Studies Quarterly*, 68(2).
- Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022.
- Brun-Mercer, N. (2021). Women and men in the united nations: A corpus analysis of general debate addresses. *Discourse & Society*, 32(4):443–462.
- Bunea, A. and Ibenskas, R. (2015). Quantitative text analysis and the study of eu lobbying and interest groups. *European Union Politics*, 16(3):429–455.
- Carson, A., Min, E., and Van Nuys, M. (2024). Racial tropes in the foreign policy bureaucracy: A computational text analysis. *International Organization*, 78(2):189–223.
- Catalinac, A. (2016). *Electoral Reform and National Security in Japan: From Pork to Foreign Policy*. Cambridge University Press, Cambridge.

- Deutschmann, E., Lorenz, J., Nardin, L., Natalini, D., and Wilhelm, A., editors (2020). *Computational Conflict Research*. Springer, Cham.
- Evans, J. A. and Aceves, P. (2016). Machine translation: Mining text for social theory. *Annual Review of Sociology*, 42(1):21–50.
- Finke, D. and Risse, T. (2024). Multidimensional conflicts over disarmament and international security: Analyzing speeches in the first committee of the UN general assembly. *Political Science Research and Methods*.
- Fobbe, S., Gasbarri, L., and Ridi, N. (2025). Words of power: Introducing a comprehensive corpus of UN security council resolutions. Preprint.
- Golder, S. A. and Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051):1878.
- Gottweis, J., Weng, W.-H., Daryin, A., et al. (2025). Towards an ai co-scientist. arXiv preprint.
- Grimmer, J. (2010). A bayesian hierarchical topic model for political texts. *Political Analysis*, 18(1):1–35.
- Grimmer, J., Roberts, M. E., and Stewart, B. M. (2022). *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press, Princeton.
- Grimmer, J. and Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21:267–297.
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based tf-idf procedure. arXiv preprint.
- Gupta, S., Bolden, S. E., Kachhadia, J., Korsunskaya, A., and Stromer-Galley, J. (2020). PoliBERT: Classifying political social media messages with BERT. In *Social, Cultural and Behavioral Modeling (SBP-BRiMS 2020) Conference*, Washington, DC. Conference paper.
- Hanania, R. (2021). The humanitarian turn at the United Nations Security Council: Explaining the development of international norms through machine learning algorithms. *Journal of Peace Research*, 58(4):655–670.
- Heseltine, M. and von Hohenberg, B. C. (2024). Large language models as a substitute for human experts in annotating political text. *Research & Politics*, 11(1).
- International Commission on Intervention and State Sovereignty (2001). *The Responsibility to Protect*. IDRC, Ottawa.
- Jankin, S., Baturo, A., and Dasandi, N. (2025). Words to unite nations: The complete United Nations general debate corpus, 1946–present. *Journal of Peace Research*, 62(4):1339–1351.



- Katagiri, A. and Min, E. (2019). The credibility of public and private signals: A document-based approach. *American Political Science Review*, 113(1):156–172.
- Kentikelenis, A. and Voeten, E. (2020). Legitimacy challenges to the liberal world order. *The Review of International Organizations*, 16(4):721–754.
- Kertcher, C. and Zwillig, M. (2025). The meaning of sentiment analysis of UN speeches on the Russia–Ukraine war: a comparative study using VADER and BERT NLP techniques. *Frontiers in Political Science*, 7:1546822.
- King, G., Pan, J., and Roberts, M. E. (2013). How censorship in china allows government criticism but silences collective expression. *American Political Science Review*, 107(2):326–343.
- Knapp, A. (2024). Speaking of civilians: Protection, harm, and the politics of humanitarian language at the United Nations. *International Interactions*, 50(2):300–331.
- Laurer, M., van Atteveldt, W., Casas, A., and Welbers, K. (2024). Less annotating, more classifying: A systematic comparison of annotation strategies for supervised text classification. *Political Analysis*, 32(1):84–100.
- Le Mens, G. and Gallego, A. (2025). Positioning political texts with large language models. *Political Analysis*.
- Lu, C., Lu, C., Lange, R. T., et al. (2024). The ai scientist. arXiv preprint.
- Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, D., and Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, 23(2):254–277.
- Lundgren, M., Squatrito, T., Sommerer, T., and Tallberg, J. (2024). Introducing the intergovernmental policy output dataset (ipod). *The Review of International Organizations*, 19(1):117–146.
- Medzihorsky, J., Popovic, M., and Jenne, E. K. (2017). Rhetoric of civil conflict management. *Research & Politics*, 4(2).
- Mesquita, R. and Pires, A. (2025). The references of the nations: Introducing a corpus of United Nations general assembly resolutions since 1946 and their citation network. *Journal of Peace Research*, 62(4):1279–1291.
- Mitra, A. (2024). Words that matter. *Research & Politics*, 11(3).
- Raschka, S. (2025). *Build a Large Language Model (From Scratch)*. Manning, Shelter Island, NY.
- Roberts, M. E., Stewart, B. M., Tingley, D., et al. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4):1064–1082.

- Rule, A., Cointet, J.-P., and Bearman, P. S. (2015). Lexical shifts, substantive changes, and continuity in state of the union discourse. *Proceedings of the National Academy of Sciences*, 112(35):10837–10844.
- Sakamoto, T. (2023). Threat conceptions in global security discourse: Analyzing the speech records of the United Nations Security Council, 1990–2019. *International Studies Quarterly*, 67(3).
- Sakamoto, T. and Matsuoka, T. (2023). The UNSC meetings and speeches. Harvard Dataverse.
- Sakamoto, T., Matsuoka, T., and Ito, H. (2025). The security council in its entirety: Unveiling 80 years of deliberation through the UNSC meetings and speeches dataset. APSA Preprints.
- Scherzinger, J. (2022). Acting under chapter 7. *International Relations*, 17(1).
- Scherzinger, J. (2023). Unbowed, unbent, unbroken? *Cooperation and Conflict*, 58(1).
- Schmidgall, S. et al. (2025). Agent laboratory. arXiv preprint.
- Schönfeld, M. et al. (2025). Security council debates corpus. arXiv preprint.
- Sundström, K., Lundgren, M., and Klamberg, M. (2025). Debating threats in the UNSC: A new dataset on explanations of votes 1989–2019. *International Studies Perspectives*, page ekaf010.
- Swanson, K. et al. (2024). The virtual lab. bioRxiv.
- Takikawa, H. and Sakamoto, T. (2019). The moral–emotional foundations of political discourse. *Quality & Quantity*.
- Terman, R. (2017). Islamophobia and media portrayals of muslim women. *International Studies Quarterly*, 61(3).
- Timoneda, J. C. and Vallejo Vera, S. (2025). BERT, RoBERTa, or DeBERTa? comparing performance across transformers models in political science text. *The Journal of Politics*, 87(1):347–364.
- Trubowitz, P. and Watanabe, K. (2021). The geopolitical threat index: Measuring geopolitical competition through news text. *International Studies Quarterly*, 65(3).
- Turco, L. R. (2024). Speaking volumes: Routine diplomacy and the United Nations General Assembly. *International Studies Quarterly*, 68(1).
- United Nations General Assembly (2005). 2005 world summit outcome. A/RES/60/1.
- Vaswani, A. et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- Verbeek, N. (2024). ‘Wolf Warriors’ in the United Nations Security Council?: Blame, criticism, and status signaling in multilateral diplomacy. *Global Policy*, 15(S2).

- Voss, M., Kump, I., and Bochtler, P. (2022). Unpacking the framing of health in the UNSC. *Australian Journal of International Affairs*, 76(1).
- Wilkerson, J. and Casas, A. (2017). Large-scale computerized text analysis in political science: Opportunities and challenges. *Annual Review of Political Science*, 20(1):529–544.
- Yukawa, T. and Sakamoto, T. (2024). The evolution of monitoring. *Foreign Policy Analysis*, 20(1).