Machine Learning Analysis Advancing Security Planning

Yulian Hristov[1[0009-0003-8840-2134]], and Zlatogor Minchev[1,2[0000-0003-2479-5496]]

[1]Institute of ICT, Bulgarian Academy of Sciences,
Acad. Georgi Bonchev Str., Bl. 25A, Sofia, 1113, Bulgaria
[2]Institute of Mathematics & Informatics, Bulgarian Academy of Sciences,
Acad. Georgi Bonchev Str., Bl. 8, Sofia, 1113, Bulgaria
yulian.hristov@iict.bas.bg, zlatogor@bas.bg

Abstract. The rapid advancement of Artificial Intelligence and Large Language Models (LLMs) presents new opportunities for analyzing large-scale unstructured textual data, such as national security strategies. However, effective, machine-assisted methodologies for the qualitative analysis of such documents still remain underdeveloped. This study proposes opportunities for application of a hybrid methodology that integrates computational text analysis with the sociological-linguistic framework of "graphing the graphs" by John W. Mohr and the dramatistic pentad of "grammar of motives" by Kenneth Burke. Applying natural language processing (NLP) techniques, including: named entity recognition (NER), sentiment analysis via specialized dictionaries, and contextual relationship extraction using tools like Voyant Tools – the framework enables the machine-assisted identification of key Actors, Actions, and Scenes within strategic texts. Experimentally applied to three Bulgarian national security documents (1998–2018), the approach demonstrates the extraction of dominant terminology, key actors (Bulgaria, NATO, EU), and their contextual relationships. The results validate the potential of AI to support strategic analysis while highlighting limitations such as model hallucinations and data sensitivity risks. The proposed setup explicitly calls for further integration of machine learning, including advanced models for relation extraction and hybrid validation systems, aiming to increase automation, accuracy, and scalability in the analysis of strategic narratives.

Keywords: Security Planning, NLP, Machine Learning, Holistic System Modelling, Smart Security Analysis, Digital Humanities.

1 Introduction

The rapid evolution of artificial intelligence (AI) and computational tools presents new paradigms for analyzing large-scale, unstructured textual data. In fields such as international relations and security studies, this is particularly relevant for processing national strategic documents-complex texts that result from formal planning processes and contain critical narratives about state priorities, threats, and actions. Despite existing operational solutions in natural language processing (NLP), there remains a significant lack of effective, machine-assisted methodologies capable of performing a

Research Paper
DOI: https://doi.org/10.46793/BISEC25.098H
Part of ISBN: 978-86-89755-40-4

nuanced, qualitative analysis of such documents. Traditional approaches often fail to systematically extract the deep narrative structures and contextual relationships that define strategic intent.

The core objective of this study is to propose and demonstrate a novel hybrid methodology that integrates AI and machine learning (ML) into the analysis of strategic texts. Specifically, we aim to bridge qualitative theoretical frameworks from the social sciences with quantitative computational techniques to enable a more precise, efficient, and scalable analysis.

To achieve this, our research adapts and operationalizes the sociological-linguistic framework of "graphing the graphs" proposed by John W. Mohr et al. [9], which itself builds upon the dramatistic pentad of "grammar of motives" by Kenneth Burke [1]. Burke's model focuses on five key elements: Act, Scene, Agent, Agency, and Purpose, providing a lens to deconstruct narrative action. For computational tractability, we adopt Mohr et al's simplification, focusing on core relational tuples such as Actor-Action and Actor-Action-Actor. This approach complements and computationally implements the concept of "distant reading" in the digital humanities [2], which advocates for the analysis of macroscopic patterns and meta-textual levels across large corpora through visualization and quantitative mapping.

The primary subject of our analysis is a set of national strategic security documents. These texts are crucial for identifying the policy vectors and priorities of a state, both nationally and internationally. Automating parts of their analysis - traditionally done through expert, close reading in political science and sociology - promises faster and more consistent insights without replacing deep expert understanding.

The advent of Large Language Models (LLMs) has significantly expanded the capabilities of computer-aided text analysis, enabling not just term tracking but also contextual positioning to reach meta-textual levels [10]. However, their application in critical domains like security and strategic decision-making carries substantial risks, including model hallucinations and escalation [8], data leakage with classified information, and vulnerabilities such as prompt injection attacks [11]. Studies caution against over-reliance on models like GPT for decision-support, noting their potential for dangerously erroneous predictions and logical shortcomings that could lower standards in strategic planning [5], [6], [7]. In some cases, predictive scenarios may even escalate towards arms race dynamics [8].

Consequently, understanding the place of LLMs in these analytical processes, along with realistic expectations of their strengths and weaknesses in the context of international relations research, is crucial for their effective and critical application [7], [13], [14].

Therefore, to harness the potential of AI while mitigating these distorting effects, there is a pressing need for a scientifically grounded methodology. Such a framework must ensure the reliability, relevance, and critical validation of automated analysis outputs when applied to large, sensitive document sets.

2 Problem Statement

The identified challenge reveals a clear research gap: while powerful computational tools (NLP, LLMs) and rich qualitative frameworks (dramatistic analysis) exist, they operate in isolation. There is no established, reproducible methodology that systematically integrates narrative theory with transparent and auditable ML techniques specifically for the analysis of strategic security documents. Current applications tend to be either:

- Purely computational, relying on opaque LLMs that risk generating unverifiable or misleading outputs for high-stakes analysis [5], [7], [8], or

- Purely manual, relying on expert close reading that lacks scalability and computational rigor.

This gap necessitates a hybrid approach that bridges theoretical depth with computational scalability while maintaining analytical transparency. Consequently, this study is guided by the following central research question: *How can a hybrid methodology, combining Burke's dramatistic framework with modern NLP techniques, be systematically applied to extract and visualize key actors, actions, and narrative scenarios from national security strategies?*

To address this question, the study has two primary objectives: (1) to design a structured methodological pipeline that operationalizes narrative theory for computational analysis, and (2) to demonstrate and validate this pipeline on a corpus of national security documents, showing its utility for augmenting expert-led strategic analysis.

The rest of this paper is structured as follows: Section 3 reviews the relevant theoretical and computational literature. Section 4 details the proposed hybrid methodology. Section 5 presents the results of applying this methodology to a corpus of Bulgarian national security documents. Section 6 discusses the findings, implications, limitations, and outlines directions for future work, emphasizing the necessary further integration of advanced ML models.

3 Related Work

This section reviews relevant work in three key areas that form the foundation for our proposed methodology: (1) advanced techniques for Named Entity Recognition (NER) and semantic understanding, (2) information retrieval and semantic role labeling for domain-specific texts, and (3) the application of NLP for the analysis of strategic policy documents.

3.1 Advanced Named Entity Recognition & Semantic Enrichment

Traditional NER, a core NLP task for locating and classifying entities into predefined categories (e.g., LOC, PER, ORG), is often treated as a preprocessing step beneficial for downstream applications like machine translation and information extraction [15].

However, conventional sequence modeling approaches that rely on static word embeddings often fail when encountering rare or ambiguous entities, as they tend towards pattern memorization rather than contextual understanding and suffer from out-of-domain generalization issues [15].

To address these limitations, more sophisticated models leverage contextual embeddings and external knowledge. The KARL-Trans-NER model exemplifies this trend by utilizing a transformer encoder architecture (similar to BERT) to process large-scale knowledge bases represented as factual triplets (e.g., "Bulgaria – capital – Sofia") [15]. The model constructs a graph from these triplets to capture relational context and generates a contextualized triplet representation - a specialized, relation-aware embedding for entities. This approach demonstrates how injecting structured world knowledge can significantly enhance NER performance, a principle relevant even when using more accessible tools like spaCy, which can be extended with transformer-based embeddings via libraries like spacy-transformers.

The process of deepening text understanding often involves a pipeline where syntactic analysis precedes semantic interpretation. Dependency Parsing identifies grammatical relationships within a sentence (e.g., subject, object). In contrast, Semantic Role Labeling (SRL) analyzes the predicate-argument structure, assigning semantic roles (e.g., Agent, Patient, Instrument) to the participants in an event [16]. For instance, in "The doctor prescribed antibiotics to the patient", dependency parsing identifies "doctor" as the subject, while SRL labels it as the Agent and "antibiotics" as the Instrument. This layered analysis – from text to dependency parsing, then to SRL, and finally to event/relation extraction - provides a robust framework for extracting structured, interpretable narrative units like Actor-Action-Actor tuples [16].

3.2 Domain-Specific Information Retrieval Using NER and SRL

The combination of NER and SRL has proven effective for content-based information retrieval in specialized domains. Antony and Mahalakshmi [16] applied this methodology to Tamil Siddha medical texts. Their process involves two main steps: first, identifying domain-relevant named entities and natural medical terms; second, labeling sentences based on their verbal clause structure using SRL. The SRL phase matches sentence patterns against predefined role libraries like PropBank and FrameNet, assigning role-based labels (e.g., Agent, Theme) to entities based on their semantic function rather than just grammatical position. This demonstrates a practical application where moving from syntactic to semantic analysis is crucial for extracting meaningful, domain-specific information.

3.3 NLP for Analysis of Strategic and Policy Documents

(i) The analysis of unstructured textual data, which constitutes a significant portion of global information [17], is increasingly applied to policy documents. Papadopoulos and Charalabidis [18] provide a seminal example by analyzing national AI strategies using

NLP. Their methodology involves: (a) extracting key terms and summarizing documents, (b) discovering latent topics within each document using Latent Dirichlet Allocation (LDA), a probabilistic model for uncovering hidden thematic structures in document collections [19], and (c) clustering strategies based on pairwise similarity.

(ii) To compute semantic similarity between documents, they employed a hybrid approach combining: (1) cosine similarity based on LDA topic distributions, (2) similarity based on average word embeddings (e.g., Word2Vec), and (3) similarity derived from the Word Mover's Distance [18]. This moves beyond simple Bag-of-Words representations, which create term-document matrices weighted by schemes like TF-IDF or BM25. While these weighting methods improve upon raw term frequency by assessing a term's informativeness, LDA offers a higher-level abstraction by grouping co-occurring terms into interpretable thematic clusters (e.g., terms like: "war", "peace", "conflict" might form a "Security and Defense" topic).

Our study builds upon these foundations. We adopt the goal of extracting structured narratives from strategic texts, similar to [18], but instead of topic modeling, we operationalize a dramatistic narrative framework [1], [9]. We utilize pragmatic, accessible tools for NER and contextual analysis (spaCy, Voyant Tools) while drawing inspiration from advanced methods for semantic role labeling [15] and knowledge-aware representations [16] to outline a clear path for future methodological enhancement towards automated relation extraction.

4 Methodology

4.1 Theoretical Framework & General Approach

Tools from the field of Natural Language Processing (NLP), as a sub-discipline based on machine learning and artificial intelligence, significantly facilitate working with large document sets for extracting relevant information. The method for labeling textual elements - Named Entity Recognition (NER) - is particularly useful for extracting nominations for actors, actions, and scenes of action. There are freely available developed tools [3], [4], which are leading platforms for building Python programs for working with human speech and text. They include built-in libraries for classification, token and stem separation, and are widely used in computational linguistics [3].

For the purposes of this study, a hybrid methodology has been proposed, consisting of two main analytical steps: 1) determining the terminological flow in the text, and 2) extracting the main actors and their relationships. This approach operationalizes Burke's dramatistic framework [1] and his understanding of scene, actor, and act, adapted for computational analysis by Mohr et al. [9].

4.2 Determining Terminological Flow (Scene Analysis)

Determining the terminological flow is applicable for determining the contextual window (scene) and positioning the main actors relative to it in the Actor - Scene setup. For this process, the bag of words approach was used. It has its drawbacks but can be optimized through active updating of pre-created dictionaries in .json format. For this purpose, three specialized dictionaries were developed, containing terminology with positive, negative, and neutral connotations, relevant to the national security discourse. Through them, the most frequently used terms in the documents are identified and counted.

The final result allows for the determination of the essence of the scene of action, as well as its contextual charge (positive, negative, or neutral) regarding the actors and their actions. An important limitation of this approach is that the compiled dictionaries are relevant only to the specific set of documents subjected to analysis and do not guarantee automatic applicability in other contexts.

4.3 Extracting Actors (Actor Identification)

For the automated identification of actors (actors) as key entities in the text, the Named Entity Recognition (NER) method was used. Entity extraction was performed using the spaCy module [4], which offers a high-quality, pre-trained solution for labeling objects. Within the analysis, we focused on entity categories particularly significant for strategic documents, such as GPE (countries, cities), ORG (organizations, such as NATO and the EU), and PERSON.

This stage generates a structured list of all identified actors, their types, and frequency of occurrence, which serves as the foundation for the subsequent relationship analysis.

4.4 Extracting Relationships (Relation Extraction)

To discover the specific narrative relationships corresponding to the Actor-Act and Actor-Act-Actor models, the web-based textual analysis environment Voyant Tools [12] was used. The primary tool applied here is KWIC (Key Words In Context), which visualizes each occurrence of a predefined key actor (via NER) within its immediate context.

This method allows for the manual or semi-automatic identification of actions (acts) associated with a given actor, as well as other actors with which it interacts within a sentence or close context. In this way, specific scenarios are extracted and the semantic networks of interactions embedded in the strategic texts are mapped. This process complements and validates the results from the previous stages by translating them into concrete, interpretable narrative units.

5 Results

5.1 Terminological Analysis & Scene Definition

The approach for extracting the most frequently used terminology is essential for determining the mode of the scene and the context within which Actor-Act and Actor-Act-Actor relationships develop. This approach combines the concepts of close reading and distant reading through data mapping and statistical analysis. The diagram (Fig. 1) presents the results of the analysis of three Bulgarian national security strategies, where the absolute count of terms has been transformed into the percentage ratio of their usage relative to the total volume of terms in the respective text.

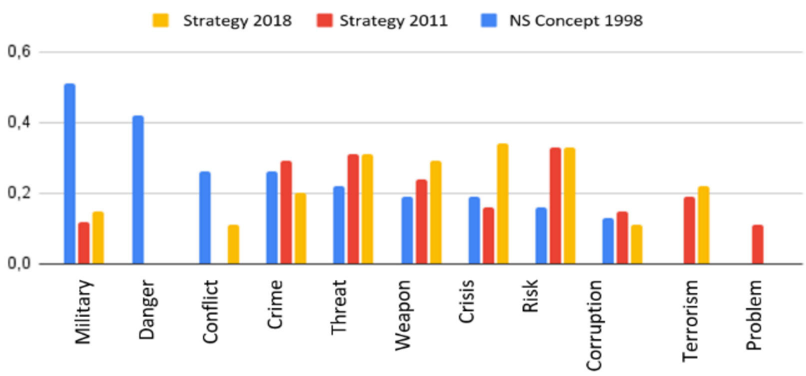

Fig. 1. Frequency of used terms in the three documents as a percentage relative to the text volume.

The data reveal clusters of terms that maintain temporal stability across all three documents, covering the period from 1998 to 2018. For this specific analysis, the dictionary with negative terminology was used. From the results, distinct cores can be observed that define the environment (scene) in its negative context. With certainty, we can assert that terms such as "crime", "weapons", "crisis", and "risk" (and their derivatives, such as re/armament, for example) are the modes that have relative weight in determining the scene within this specific set of studied documents.

5.2 Terminological Analysis & Scene Definition

The data reveal clusters of terms that maintain temporal stability across all three documents, covering the period from 1998 to 2018. For this specific analysis, the dictionary with negative terminology was used. From the results, distinct thematic cores can be observed that define the environment (scene) in its predominantly negative context. With certainty, we can assert that terms such as "crime", "weapons", "crisis", and "risk" (and their derivatives, such as re-armament) are the dominant modes that carry significant weight in determining the scene within this specific set of studied documents.

5.3 Identification of Key Actors

In extracting the actors, we can perform a similar visualization for those entities that are given the most prominent space in the texts. The results, presented in diagram (Fig. 2), demonstrate this distribution.

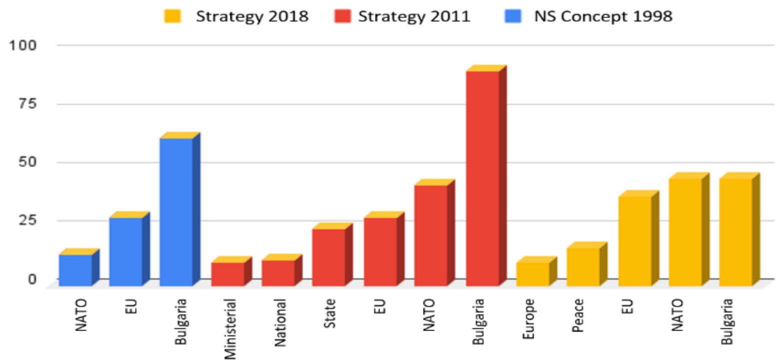

Fig. 2. Most frequently used actors in the studied documents in absolute values, with a count of occurrences above ten.

The data show that across the three strategic documents, three main actors stand out: Bulgaria, NATO, and the EU. However, their relative prominence varies. In the earliest conceptual document, the European Union holds the second position in terms of significance, while in the other two, later documents, NATO emerges as the secondary key actor. This shift suggests an evolution in Bulgaria's primary strategic alignment and referential framework over the two decades analyzed.

5.4 Extraction and Mapping of Narrative Scenarios

The execution of the final part of the analysis focuses on the groups Actor-Action, Action-Actor, and Actor-Action-Actor. For the sake of methodological simplification, these groups are considered as one core type (Actor-Action) and two of its variants (Action-Actor, Actor-Action-Actor).

As indicated, the KWIC functionality of Voyant Tools was applied for this purpose, utilizing its capacity to determine the contextual environment of the actors. For a comprehensive analysis, it is necessary to examine not only the three actors with the highest weight in the three strategic documents (as mentioned, these are Bulgaria, NATO, and the EU) but also all others with lower prominence. However, since mapping all these relationships is extensive, this section will focus solely on the contextual role assigned to the entity Bulgaria within one of the three documents (see Fig. 3).

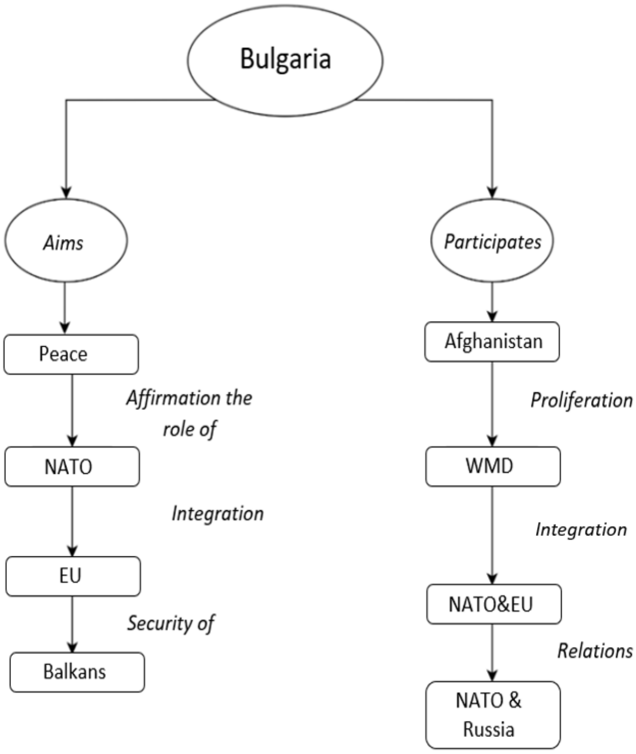

Fig. 3. Actor-Action-Actor relationships in the “Concept for National Security of the Republic of Bulgaria”, 1998.

As it is clearly evident from Fig. 3, Bulgaria aims to support peace on a global scale and also aims to affirm the role of NATO and integration into EU structures. The country has as a strategic objective the security of the Balkans, participates in NATO initiatives in Afghanistan and in efforts for the non-proliferation of Weapons of Mass Destruction (WMD), again within the context of integration into EU and NATO structures, as well as in shaping NATO-Russia relations.

These specific scenarios were extracted manually from the contextual analysis. The ambition, however, is to fully automate this process in future iterations of the methodology to optimize efficiency and scalability while maintaining analytical precision.

6 Discussion

6.1 Interpretation of Findings & Strategic Implications

The application of the proposed hybrid methodology to the corpus of Bulgarian national security documents (19982-018) yields interpretable results that align with and extend expert understandings of the country's strategic posture. The terminological analysis reveals a persistently negative scene, dominated by terms such as “crime”, “weapons”, “crisis”, and “risk”. This finding substantiates the primary lens through which national security is conceptualized in these documents - one of threat perception and risk management. It confirms a stable, underlying narrative of vulnerability that transcends specific governments or short-term political cycles.

The evolution in the prominence of key actors – where the European Union is secondary in the earliest (1998) conceptual document, but NATO assumes this role in the later strategies - provides a quantifiable signal of a strategic re-orientation. This shift maps onto Bulgaria’s tangible foreign policy trajectory: from post-Cold War conceptualization towards active membership in NATO (achieved in 2004) and later the EU (2007). The methodology successfully captures this macro-level narrative shift through a simple frequency metric, demonstrating its utility for tracking changes in strategic alignment over time.

Finally, the extracted Actor-Action-Actor scenarios, such as “Bulgaria aims integration into EU and NATO structures” or “Bulgaria participates in NATO initiatives”, transform abstract strategic goals into concrete, relational narrative units. These scenarios do not merely list priorities but reveal the active role Bulgaria envisions for itself (an aspirant, a participant) and the relationships it seeks to cultivate. This moves the analysis beyond “what” is said to “how” strategic intent is narratively constructed, offering a more nuanced layer of insight for policy analysts.

6.2 Methodological Evaluation & Future Directions

The presented model for machine-assisted analysis is demonstrative in nature. To perform a comprehensive “revisionary” reading and analysis of large-scale textual corpora, more detailed and in-depth work on an expanded methodology must be undertaken. This includes the extraction of all possible scenario developments and an evaluation of their relative weight or significance.

Given that the process aims to achieve a greater degree of autonomy, the implementation of state-of-the-art technologies is of paramount importance. In order to take in account the risks of data distortion and analytical bias, a systematic approach for model selection and validation will be essential. This could involve employing advanced machine learning models such as XGBoost or Random Forest for classification and validation tasks. Furthermore, the introduction of a microservices (agent) architecture – comprising independent modules for specific functions – could significantly improve system maintainability, scalability, and robustness. The integration of techniques for handling data imbalances is also crucial for improving model performance.

Equally significant is ensuring the model retains a hybrid character, combining different methodological approaches, particularly in the scenario extraction phase. For instance, extracting complex Actor–Action–Actor scenarios could benefit from an approach that combines the strengths of linear context extraction via KWIC with Dependency Parsing, which reveals the grammatical relationships between entities in the text. This would allow for a more nuanced and syntactically grounded identification of narrative relations, moving beyond co-occurrence towards understanding predicative structures.

6.3 Limitations of the Study

While the methodology shows promising results, several limitations have to be marked:

First, the corpus is limited in size and scope, comprising only three Bulgarian documents. This affects the generalization of the specific findings and necessitates application to a broader, multilingual set of strategies for validation.

Second, the methodological pipeline relies on components with inherent constraints. The bag-of-words approach with static sentiment dictionaries, while transparent, lacks contextual nuance and is domain-dependent. The accuracy of the scene analysis is directly tied to the quality and comprehensiveness of these manually crafted lexicons. Similarly, the extraction of relationships via Voyant's KWIC tool remains a semi-manual process, limiting full automation and scalability.

Finally, the study operates on a specific level of analysis. It excels at identifying explicit actors and their directly stated actions but may not capture implicit assumptions, rhetorical nuances, or intertextual references that a deep qualitative reading would reveal. The model is designed to augment, not replace, expert analysis, and its outputs must be interpreted within these bounded capabilities.

Acknowledgments. The authors of this study are granting special appreciations for the experimental base and partial funding support to the National Scientific Programme "Security & Defense". The analytical results presented in the paper are benefiting the international expert support obtained in the framework of the initiative "Securing Digital Future 21" with more than sixty countries now, spread around the world, https://securedfuture21.org/.

References

1. K. Burke, A Grammar of Motives. Berkeley, CA, USA, Univ. of California Press, 1st edn, (1969)
2. F. Moretti, Graphs, Maps, Trees: Abstract Models for Literary History. London, UK: Verso, 1st edn, (2005)
3. Natural Language Toolkit, version 3.9.1. (2025), https://www.nltk.org/, last accessed 2026/02/11
4. spaCy: Industrial-Strength Natural Language Processing in Python. Explosion, https://spacy.io/, last accessed 2026/02/11

5. D. Levinson, How to Keep Generative AI from Crashing in Combat, War on the Rocks, (2025), https://shorturl.at/HeS5U, last accessed 2026/02/11
6. The White House, Winning the Race: America's AI Action Plan, (2025), https://whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf, last accessed 2026/02/11
7. C. Hunter, B. Bowen, We'll never have a model of an AI major-general: Artificial Intelligence, command decisions, and kitsch visions of war, Journal of Strategic Studies, (2023), https://doi.org/10.1080/01402390.2023.224164, last accessed 2026/02/11
8. J.-P. Rivera, G. Mukobi, A. Reuel, M. Lamparth, Ch. Smith, and J. Schneider, Escalation risks from language models in military and diplomatic decision-making, in Proc. 2024 ACM Conf. Fairness, Accountability, and Transparency (FAccT), 836–898, (2024), https://dl.acm.org/doi/10.1145/3630106.3658942, last accessed 2026/02/11
9. J. W. Mohr, R. Wagner-Pacifici, R. L. Breiger, and P. Bogdanov, Graphing the grammar of motives in National Security Strategies: Cultural interpretation, automated text analysis and the drama of global politics, Poetics, vol. 41, no. 6, 670–700, (2013), https://doi.org/10.1016/j.poetic.2013.08.003, last accessed 2026/02/11
10. G. Shopov and S. Gerdjikov, Consistent Bidirectional Language Modelling: Expressive Power and Representational Conciseness, in Proc. 2024 Conf. Empirical Methods in Natural Language Processing (EMNLP), 5724 –5768, (2024), https://aclanthology.org/2024.emnlp-main.328/, last accessed 2026/02/11
11. X. Liu, Z. Yu, Y. Zhang, N. Zhang, and C. Xiao, Automatic and Universal Prompt Injection Attacks against Large Language Models, (2024), https://arxiv.org/abs/2403.04957, last accessed 2026/02/11
12. S. Sinclair, and G. Rockwell, Voyant Tools, (2016), voyant-tools.org, last accessed 2026/02/11
13. B. Jensen, I. Reynolds, Y. Atalan, Critical Foreign Policy Decisions (CFPD)-Benchmark: Measuring Diplomatic Preferences in Large Language Models, Center for Strategic and International Studies, Washington, DC, USA, https://arxiv.org/html/2503.06263v1, last accessed 2026/02/11
14. W. N. Caballero and P. R. Jenkins, On large language models in national security applications, Stat, vol. 14, no. 2, p. e70057, (2025), https://doi.org/10.1002/sta4.70057, last accessed 2026/02/11
15. A. Chawla, N. Mulay, V. Bishnoi, and G. Dhama, KARL-Trans-NER: Knowledge Aware Representation Learning for Named Entity Recognition using Transformers, (2021), https://arxiv.org/abs/2111.15436, last accessed 2026/02/11
16. B. Antony, and G. Mahalakshmi, Content-based Information Retrieval by Named Entity Recognition and Verb Semantic Role Labelling, J. Universal Comput. Sci., vol. 21, no. 13, 1830–1848, https://www.jucs.org/jucs_21_13/content_based_information_retrieval.html, last accessed 2026/02/11
17. A. Mishra, Information Professionals and Big Data, Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 5, no. 9, 317 – 321, https://www.ijarcsse.com/docs/papers/Volume_5/9_September2015/V5I9-0214.pdf, last accessed 2026/02/11
18. T. Papadopoulos, and Y. Charalabidis, What do governments plan in the field of artificial intelligence?: Analysing national AI strategies using NLP, in Proc. 13th Int. Conf. Theory Pract. Electron. Governance (ICEGOV), 100–111, (2020) https://doi.org/10.1145/3428502.3428514, last accessed 2026/02/11, last accessed 2026/02/11
19. D. Blei, A. Ng, and M. Jordan, Latent Dirichlet Allocation, J. Mach. Learn. Res., vol. 3, 993–1022, (2003), https://dl.acm.org/doi/10.5555/944919.944937, last accessed 2026/02/11