IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 August 2025, accepted 30 August 2025, date of publication 2 September 2025, date of current version 9 September 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3605354

SURVEY

A Comprehensive Review of NLP Techniques for Military Terminologies and Information Operations on Social Media

TAMARA ZHUKABAYEVA[1,2], (Member, IEEE), ZULFIQAR AHMAD[3], AIGERIM YERIMBETOVA[1,4], (Member, IEEE), MADINA SAMBETBAYEVA[1,2], DUMAN TELMAN[1,4], ABDYGALYM BAYANGALI[1,2,5], AND ELMIRA DAIYRBAYEVA[1,5]

[1]Laboratory of Computer Engineering of Intelligent Systems, Institute of Information and Computational Technologies, Almaty 050010, Kazakhstan
[2]Department of Information Systems, L. N. Gumilyov Eurasian National University, Astana 010000, Kazakhstan
[3]Department of Computer Science and Information Technology, Hazara University, Mansehra 21300, Pakistan
[4]Global Education and Training, University of Illinois Urbana–Champaign, Champaign, IL 61820, USA
[5]Department of Software Engineering, Satbayev University, Almaty 050046, Kazakhstan

Corresponding authors: Zulfiqar Ahmad (zulfiqarahmad@hu.edu.pk) and Aigerim Yerimbetova (aigerian8888@gmail.com)

This work was supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan under Grant AP23484329.

ABSTRACT This paper presents a comprehensive study aimed at systematically analyzing and evaluating natural language processing (NLP) techniques for military information operations, with a special focus on social media intelligence. Among an ever-growing complicated information environment, NLP methods like sentiment analysis, named entity recognition, and topic modeling have been essential in tracking online propaganda efforts, discovering emerging issues and threats globally with dialogues on military operations. These techniques make an impact on available decision making via situational awareness and getting the added extraction from volumes of unstructured data outputs thus increasing the overall strategic benefits to military organizations. There are technical and operational challenges concerning the use of NLP in a military context such as requirements for real-time data processing; language diversity; and maintaining data privacy while preserving ethical standards. To address these challenges, the study conducts an exhaustive survey of NLP methods, reviewing their range of applications, and highlights the relevance of several approaches for military information operations, with special emphasis on social media intelligence. The work further provides discussion on the comprehensive adoption of artificial intelligence (AI), edge computing, and multilingual NLP models for enhancing adaptability, efficiency, and transparency of the systems. It also extols the need for explainable AI (XAI) to improve accountability and trust by making term or even whole early warning systems derived from NLP analyses, transparent and interpretable for these military research applications with significant financial consequences. The paper also emphasizes the strategic importance of multilingual and multimodal analysis and the integration of specialized military lexicons to improve the contextual understanding of military discourse in social media environments. We also elucidate the important capabilities of NLP in enabling military operations to be responsive, rapid and data-driven while also adapting to the evolving nature of warfare. Key conclusions suggest that applying advanced NLP tools enhances situational awareness, enables timely threat detection, and supports more agile, data-informed decision-making within modern military operations. The paper shows a perspective to optimize NLP and AI technologies, leveraging various perspectives to benefit the operational needs of military and defense sectors in more data-rich environments.

The associate editor coordinating the review of this manuscript and approving it for publication was Giacomo Fiumara.

© 2025 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

INDEX TERMS Natural language processing, military information operations, social media analysis, sentiment analysis, edge computing, data privacy, multilingual analysis, explainable AI.

I. INTRODUCTION

Natural Language Processing (NLP) has become a key technological innovation that allows machines to understand, interpret and generate human language in a manner that is both valuable and contextually relevant [1], [2], [3]. There is no doubt that NLP techniques are indispensable in military scenarios. As the volume of unstructured data has exploded from sources such as social media, military organizations are under greater pressure to extract meaningful intelligence from vast seas of information [4], [5], [6], [7]. Real-time data analytics is a core aspect of military effectiveness as it drives decision-making, situational awareness and ultimately operational success for the military. NLP enables a lot of processing at different levels, for example it can be used to identify trends from huge volumes of social media data, analyze public perception on various topics and also gauge narratives elements that might influence military operations [8], [9], [10], [11]. NLP techniques used in advanced applications (Sentiment analysis and Emotion detection) provide sizable knowledge of how the world perceives/receiving a military activity and can be some of the building blocks for strategic communication programs and counter misinformation activities [5], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22]. Besides that, it can also be used to improve internal communications for military units, enhance intelligence reporting accuracy and level coordination between disparate operational teams. NLP integration into information processing systems is necessary for improving operational effectiveness and efficiency as military operations are becoming increasingly complex, integrated Military organizations can use NLP to leverage a more fine-grained, actionable understanding of the information environment thereby enabling higher information decision-making ability and better operational logic [3], [4], [5], [13], [23], [24], [25], [26], [27], [28], [29], [30].

At present, NLP tools are part and parcel of modern military operations, for they aid in intelligence gathering, threat analysis, and decision-making. One such example is Babel X by Babel Street, an AI-driven linguistic analytics platform to analyze global communication, social media, and the dark web as a potential threat being used by defense agencies [31], [32]. The Deep Exploration and Filtering of Text (DEFT) program funded by DARPA also advances the capabilities of NLP through the improvement of text analytics, providing automatic extraction of key insights from huge amounts of unstructured data for military intelligence use [7]. The U.S. Army's AI-enabled system, MADCAT (Machine Foreign Language Translation System), translates and analyzes foreign texts, including handwritten and printed documents, to enable battlefield operations [33]. Raytheon's Translingual Automatic Language Exploitation (TALE) is another advanced system that provides real-time, multilingual speech and text translation to help military personnel understand intercepted communications and foreign documents [34]. Figure 1 shows the important NLP techniques that will support the processing of information by military capabilities.

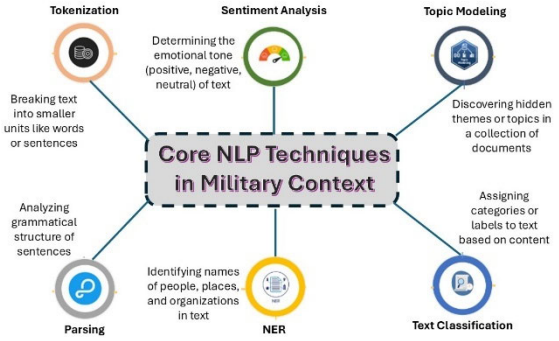


FIGURE 1. Core NLP techniques applicable to military and defense contexts.

The objective of this study is to provide a systematic review of existing NLP techniques and information processing systems for military terminology and information operations in social media. The focus is on how different NLP methods can be used to strengthen military capabilities in building, running and analyzing information carried out by social media. The paper embraces various NLP techniques including sentiment analysis, (Named Entity Recognition) NER, topic modeling and text classification to evaluate their applicability in the military [4], [14], [26], [35], [36], [37], [38]. These techniques have unique strengths as reflected from Table 1 that can be tailored to the specific issues military analysts face such as complex military language and terminology, not necessarily represented well with commercial off-the-shelf (COTS) NLP tools. In addition, this review will introduce real-life military applications of these techniques. Sentiment analysis can be used to determine how the public feels about military actions whereas NER provides insights on prominent individuals, places, and groups that are part of popular military efforts.

The aim of this study is to fill the gaps in research so that potential insights can be drawn for future advancements in military NLP. Despite rapid progress in NLP for commercial and general-purpose applications, there is a lack of systematic evaluations focused specifically on military contexts that involve domain-specific language, multilingual environments, and social media-based operations. Existing studies often overlook the challenges of real-time deployment, integration of edge computing, and the need for explainable AI in sensitive military scenarios [9], [10], [11], [13], [23], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51]. Part of the examination is deemed necessary because of the increasing demand for military organizations to embrace

TABLE 1. Core NLP techniques, their functions, and military applications.

NLP Technique	Primary Function	Military Application
Tokenization	Breaks text into words or phrases	Prepares unstructured data for analysis
Parsing	Analyzes grammatical structure	Enhances understanding of complex military language
NER	Identifies key entities (names, places)	Recognizes relevant entities in military contexts
Sentiment Analysis	Detects emotional tone	Assesses public sentiment on military actions
Topic Modeling	Identify themes in document collections	Uncovers thematic content in intelligence and operations reports
Text Classification	Categorizes text by content	Organizes and retrieves relevant military communications

advanced competitively advantage data analytics solutions. Figure 2 shows a conceptual framework of the integration of NLP in military information processing platforms.

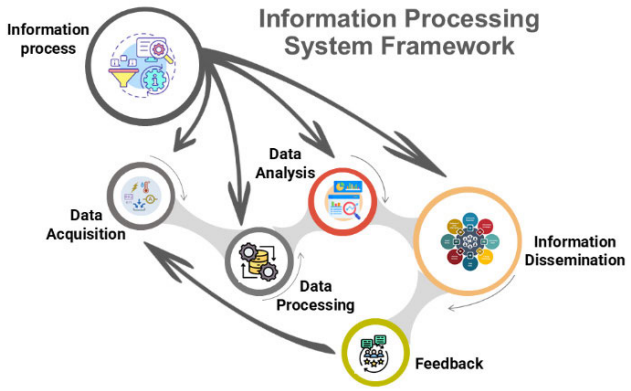


FIGURE 2. A conceptual framework of the integration of NLP in military information processing platforms.

This research focuses only on NLP and information operations with a military purview understood through social media contexts. As the implications of information manipulation extend to operational security, public opinion, and deterrence outside the military sphere, social media has become a vital space for militaries to track narratives which may contribute to shaping such narrative environments [52], [53]. This rapid spread features more and it brings unique challenges to military operations and strategic communication through social media platforms [42], [45]. The challenge is successfully processing army-relayed language and terms. For example, a natural language processing (NLP) based approach relies on linguistic analysis between words and will not be able to give reliable insights if the data contains military jargons or other forms of application specific language where NLP systems need to be capable of analyzing such these languages [23]. As military discourse evolves, so too must its language and communication tactics, requiring innovative NLP models that consider the specificity inherent in these unique linguistic forms.

The main contributions of the study are highlighted below:

1. The study provides a systematic review of NLP methods such as sentiment analysis, NER, topic modeling, and text classification, emphasizing their role in military information operations and social media analysis.
2. By leveraging NLP for real-time data extraction and analysis, the research highlights how military organizations can improve situational awareness, threat detection, and strategic decision-making in dynamic information environments.
3. The paper discusses the role of artificial intelligence and edge computing in enhancing the adaptability and efficiency of NLP-based military intelligence systems, reducing latency in real-time data processing.

1)]4.] The research emphasizes the need for specialized military lexicons and ontologies to improve NLP accuracy when dealing with unique military terminology, jargon, and abbreviations.

5. The study identifies key challenges such as language diversity, ethical concerns, data privacy, and the necessity for explainable AI (XAI) to ensure transparency and trust in AI-driven military applications.
6. The paper outlines future research areas, including multilingual and multimodal NLP approaches, improved interpretability of AI models, and advancements in cybersecurity for NLP-driven military intelligence systems.

This study has relevance beyond theoretical exploration; its execution presents practical applications that can be integrated into military strategy, preparedness, and operational capability. To further reinforce the focus and academic rigor of this work, we present a set of clearly defined research questions. These questions are intended to frame the scope and objectives of the survey while aligning with the core thematic areas examined throughout the paper. The research questions are as follows:

RQ1: What are the most relevant and applicable NLP techniques for processing military-specific language and terminologies, especially in the context of social media?

RQ2: How can domain-specific ontologies and thesauri improve the performance of NLP models in military information operations?

RQ3: What are the major technical, operational, and ethical challenges in deploying NLP systems for real-time military use, and how can they be addressed through current advances in AI?

RQ4: What future directions and open research areas exist for enhancing NLP adoption in defense and military intelligence systems?

Applying advanced NLP techniques can strengthen military institutions' decision-making capability in responding to emerging threats, understanding oscillation of public

sentiment, and countering misinformation more efficiently. As modern warfare and information operations grow increasingly complex and data-driven, the insights derived from NLP can serve as a foundation for developing strategic hypotheses and guiding future research into its evolving role within military paradigms [54], [55].

To ensure the objectivity and comprehensiveness of this study, a systematic literature review (SLR) methodology was employed [56], [57], [58], [59]. The main aim of the review was to identify scientific research and reliable sources that discuss NLP applications in military, defense, and social media fields. Searches were made in IEEE Xplore, Scopus, SpringerLink, and Web of Science using keywords like ''NLP,'' ''military information operations,'' ''social media analysis,'' ''named entity recognition,'' ''multilingual NLP,'' and ''explainable AI.'' At the beginning, 310 papers were analyzed, and only those relevant and written in English were accepted. The inclusion criteria stated that studies had to be released from 2015 onwards, have a focus on security, defense, or military uses of NLP, and provide enough information on their approaches. Certain papers were excluded because they dealt with healthcare or topics not related to the study. Applying the filters, the chosen articles were sorted by the NLP technique, what they are about, and how relevant they are to present-day military issues.

II. TECHNICAL FOUNDATIONS OF NLP IN MILITARY AND SOCIAL MEDIA CONTEXTS

A. INTRODUCTION TO NLP IN MILITARY CONTEXTS

NLP is one of the combat support capabilities in military intelligence, and one of the most important types of technology for structuring unstructured data analysis in real-time [23]. Military organizations are finding themselves increasingly faced with a large amount of information from different sources such as social media, intelligence reports and internal communications that need to be interpreted. NLP is critical for military applications not only because it automates extraction of insights but also enhances decision-making and operational effectiveness.

The defense sector has a language of its own, which is one of the biggest challenges military analysts have to deal with. Such as abbreviations, jargons, and phrases that are specific to certain context which are often missed by standard NLP [10]. The speed at which content is created and shared on social media imposes the need for adaptive NLP solutions [60]. A modern military would need to employ state-of-the-art NLP techniques to sift through information and glean actionable intelligence out of it, and that is a necessity in order for a modern military to even be responsive enough to address its operational requirements.

B. CORE NLP TECHNIQUES

The core NLP techniques are briefly discussed below:

- **Tokenization and Parsing:** Tokenization is the first step in NLP and useful when it comes to breaking down a sentence [14], [61]. Such a segmentation is important within the military context, where communications might be hierarchical and rich in meaning. Parsing is a step that complements tokenization, it analyzes the grammatical structure of these segments, recognizes phrases in them and identifies relations between words. Tokenization and parsing help glean critical elements from military communications to help analysts make the inferences needed quickly. As an example, when processing military orders or reports accurate tokenization helps to eliminate ambiguity in orders and intent [62], [63].
- **Named Entity Recognition (NER):** NER is a sub-type which identifies and associating the named entities, it defines the names like military personnel, locations, weapon systems and organizations [24], [64], [65]. This is particularly useful in military cases so the military can visualize how important these entities are. NER is, to some extent differentiating military-specific terminology as soon as it arrives by extracting proper nouns and relations between them which expedite the speed and accuracy of intelligence analysis. NER could help pinpoint the specific people or organizations being discussed on social media and then military analysts could track sentiments regarding these organizations in relation to ongoing operations.
- **Sentiment Analysis:** This refers to the analysis of emotions as they are portrayed in a given sight and sentiment expressed about military level, which can be specifically analyzed [66], [67], [68], [69]. Using this analysis in a wartime setting could prove invaluable in determining the effect of military actions on the views of civilians. Sentiment analysis can identify upward or downward trends in public sentiment positive, negative or neutral toward military actions by analyzing what people are saying on social media and news articles. For example, understanding a society's reaction to a military operation can help the communication effort by determining how military organizations can best counteract concerns or false information received in various forms.
- **Topic Modeling and Classification:** Topic modeling is also a technique that used to find abstract topics in the corpus of text whereby military analysts are able to place major themes and trends by opening mid-discussion on particular issues associated with military operations [55], [70], [71]. Topic modeling helps provide context as to how topics belong in the same category by clustering similar documents together, all of which are beneficial for military. In contrast, classification techniques can assign texts to predefined topics so the relevant information can be retrieved and analyzed rapidly. Both techniques are critical for dealing with the massive amounts of data faced by military organizations daily. For instance, topic modeling allows analysts to discover new threats

being discussed on social media, and classification can rank intelligence reports by urgency or importance.

C. MACHINE LEARNING VS. DEEP LEARNING IN NLP

The progress of NLP methods has been greatly driven by the advances in machine learning and deep learning techniques. Instead, traditional machine learning methods relied on the previously prevalent but ragged paradigm of manual feature extraction which may not fully capture language nuances and is very labor-intensive. On the other hand, deep learning models, especially with neural networks, can learn hierarchical representations from raw data and perform much better on NLP tasks. Advances like BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer) have revolutionized NLP applications, allowing systems to better understand context and semantics [54]. Such advanced models have a lot of potential applications in the military including sentiment analysis and entity recognition enabling the analysts to perform even better. If we use BERT for sentiment analysis, the accuracy of interpretation in overall public discussion regarding military operations would be greatly improved, allowing leaders to gain insight on opinion trends and plan strategically based on them. Table 2 details various machine learning techniques and their respective applications in military natural language processing.

TABLE 2. Machine learning techniques and their military NLP applications.

Technique	Strengths	Weaknesses	Military NLP Use Cases
Naive Bayes	Simple, effective for small datasets	Less accurate with complex data	Simple text categorizatio n
Support Vector Machines (SVM)	Good for text classification	Limited interpretability	Entity identification in reports
Recurrent Neural Networks (RNN)	Good for sequence-based data like text	Requires a large dataset and computational power	Sentiment analysis over time
Convolutional Neural Networks (CNN)	Effective at extracting spatial patterns	Not designed for sequential text data	Image-based text processing

NLP techniques applied in military operations represent a paradigm shift for processing and analyzing information. Through the effective application of core techniques including but not limited to tokenization, NER, sentiment analysis and topic modeling military institutions may be able find new utility for those systems in more effectively processing the complexities of modern warfare and information operations [61]. Machine Learning vs Deep Learning Difference: The change in methodology once again represents the evolution of methodologies, but deep learning having advanced capabilities has been tailored towards complex military communications. With military sectors adopting the natural language processing, we have the opportunity to increase decision-making, boost situational awareness and mitigate resource-wasting. This layout prepares the bottom for a deeper examination with susceptibility and usability of NLP in armed forces domains, at some point of basic and serious attention in quasi-imminent seamless society [72]. Not only does this serve to bolster military readiness, but it also cultivates an understanding of the information environment in which such forces are active. Since global threats are becoming more complex in nature, smart applications of NLP can help maintain operational advantage and facilitate timely decisions by the customer organizations.

III. THESAURI AND ONTOLOGIES FOR MILITARY TERMINOLOGY

A. IMPORTANCE OF DOMAIN-SPECIFIC THESAURI

Military-related NLP systems heavily depend on domain-specific thesauri for high-quality and fast output of NLP functions. Given the peculiar lexicon of its environment, these thesauri offer systematic vocabularies for specifying, defining, and interpreting military terms. Domain-specific thesauri help reduce ambiguity which enhances the capabilities of NLP applications to process and analyze military communications. This can aid information retrieval systems with the incorporation of a military thesaurus which allows an individual to find relevant data within a few seconds thereby facilitating rapid decision-making in situations that call for immediacy [73]. Use of domain-specific terminology is critical in military engagements as misinterpretations may lead to catastrophic outcomes. Accurate terminology ensures that personnel, whatever their functions (medical or non-medical), are able to communicate with each other which is a prerequisite for coordinating any operational efforts in the context of mission success [39]. Finally, thesauri contribute directly to the training and onboarding of new military personnel by providing a standardized reference point for understanding military jargon and acronyms.

B. EXISTING MILITARY THESAURI AND ONTOLOGIES

Several established military thesauri and ontologies, tailored to the specific language needs of military organizations, are depicted in Figure 3.

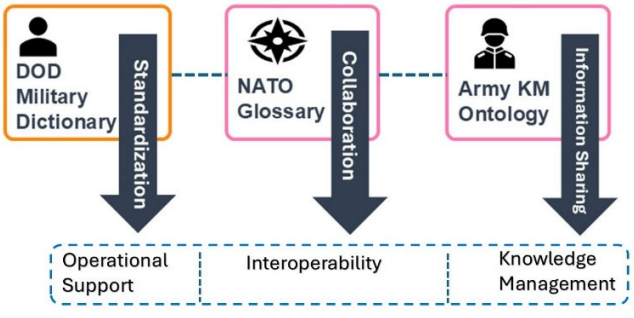


FIGURE 3. Key Military Thesauri and Ontologies Supporting Standardized Communication and Knowledge Management.

TABLE 3. Military thesauri and ontologies, including key features and limitations.

Thesaurus/Ontology	Provider	Comprehensive Military Terminology	Multilingual Support	Standardized for Joint Operations	Frequent Updates	Limitations
DOD Military Dictionary	Department of Defense	✔	✘	✘	✔	U.S.-centric, limited to specific terms
NATO Glossary	NATO	✔	✔	✔	✔	Limited to NATO-affiliated countries
UNTERM	United Nations	✘	✔	✔	✘	Limited military specificity
Jargon File	Various	✘	✘	✘	✘	Informal, lacks standardization

The details of each component are provided below:

- **DOD Military Dictionary**: It was created by the U.S. Department of Defense; this dictionary provides thorough definitions and categorizations of military terms to help standardize the use of military language throughout all branches of the armed services. The large number of terms in its database is integral to enabling a common language between personnel and systems [74].
- **NATO Glossaries**: NATO has created multiple glossaries of nomenclature used among allied forces. By providing these glossaries that communicate in a standard fashion across nations, it improves interoperability when forces are brought together for coalition operations. The NATO glossary consists of definitions, acronyms, and terms that are pivotal for cooperation in military operations [49], [74].
- **The Army's Knowledge Management (KM) Ontology**: This ontology assists with Army KM efforts and provides a high-level construct for organizing military knowledge, storage mechanisms, and information sharing across operations/units [74].

They are required not only for standardization but also to have the definition and precision in military communications such as Joint Operations where multiple stakeholders may be part. Table 3 provides an overview of existing military thesauri and ontologies.

C. CHALLENGES IN DEVELOPING MILITARY THESAURI

There are some challenges of creating military thesauri that make them an important subject matter. One major challenge is the complexity of military language: jargon, acronyms, and rapidly evolving terms. Military terminology often contains words with contextual meanings (during war/saga) that a thesaurus could never capture. An example of where a thesaurus might not be helpful is that acronyms can change meaning based on the operational context. However, if we are not meticulous in addressing this it may lead to misinterpretations.

The fast-paced landscape of military operations requires regular revisions in the thesauri to accommodate evolving terminologies, technological advancements, and innovative operational concepts. As tools evolve continuously, it can put pressure on resources and make the maintenance of such vital tools a more challenging task. Further, gaining input and agreement from differing military branches that use the terms and those that would develop them can be lengthy and arduous [75].

The above challenges can be addressed relatively effectively through military thesauri development and integration into NLP models with certain requirements that are featured in Table 4.

TABLE 4. Checklist of attributes for military thesauri and ontologies, comparing key features.

Criteria	Present in Thesaurus?	Assessment
Handles Military Jargon	✔	Essential for understanding terms
Includes Acronyms	✔	Critical for accurate interpretation
Dynamic Updates	✘	Needs regular updates
User-Friendly Interface	✔	Facilitates ease of use
Compatibility with NLP Models	✔	Ensures effective integration
Comprehensive Coverage	✘	Gaps in coverage identified

D. INTEGRATION OF THESAURI WITH NLP MODELS

Domain-specific thesauri are a fundamental component which helps to determine the effectiveness of military information processing systems and achieve high accuracy in interpreting specialized texts, thus integrating domain-specific thesauri with NLP models might be critical. It thereby enables NLP applications to make use of structured vocabularies, thereby enhancing their performance in understanding military communications and extracting pertinent information [76], [77], [78]. Figure 4 shows a military thesaurus incorporated into the NLP model that can be used to identify and classify entities such as locations, organizations, and personnel involved in military operations.

Thesauri assist in expanding the potential of NLP systems and thereby reduce the chances of misinterpretation while also facilitating quality insights from the processed data. Examples of such applications include sentiment analysis where the language can be subtle and therefore important to evaluate how statements can influence public sentiment on military type activities. The other benefit is that thesaurus relationships are used in creating NLP architectures and these can be custom built for specific applications that address the underlying problems which military analysts face with military language genre. Incorporating such domain-specific thesauri into a military NLP system is important in increasing the precision and confidence of military communication and decision-making. These thesauri help to improve the overall efficiency of information processing in military domains by resolving some of the issues related to the military-specific specific terminology [11].

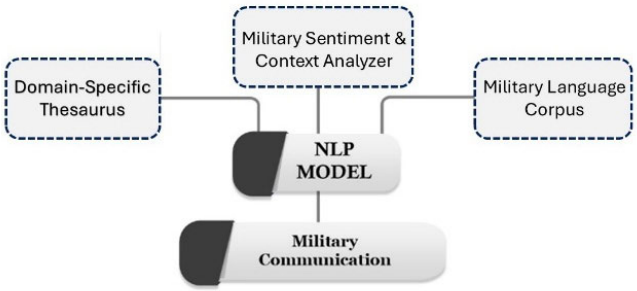


FIGURE 4. Key components with an NLP model for military communication processing.

IV. INFORMATION PROCESSING SYSTEMS FOR SOCIAL MEDIA ANALYSIS

Social media are thus making a completely new environment for military intelligence activities by the speed with which they allow information to be processed. Social media data presents an abundance of volume, variety, and velocity which leads to unique challenges in the analysis and interpretation process. A considerable amount of advanced information processing systems has arisen to aid military situational awareness and strategic decision making in order to mitigate these issues. Here we will be going into the foundational systems and how they undergo productization for social media analysis in military [79].

A. INFORMATION RETRIEVAL (IR) SYSTEMS

The IR (Information Retrieval) systems, as shown in Figure 5, work as a backbone to extract military-related information that is relevant among the huge datasets available in social media. Intelligence analysts often need a way to quickly find needles in haystacks of unwieldy, irrelevant information these systems provide exactly that functionality. With recent advancements IR systems employ advanced natural language processing, semantic search engines and ML models that possess high ager to understanding user intent and extract contextually relevant information [48], [62], [65].

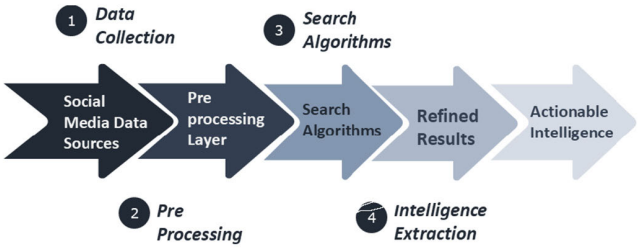


FIGURE 5. Information retrieval systems.

The key techniques and algorithms are highlighted below:

- **Boolean Search and Inverted Indexing:** Boolean Search and Inverted Indexing the basics of how we match user queries against stored data.
- **Latent Semantic Analysis (LSA) and Vector Space Models**: Used in comprehending the semantic relation amidst words, detecting these latent patterns.
- **Neural IR Models**: Traditional IR systems often utilize feature engineering methods to achieve query understanding and relevance score calculation of content, while more recent work has leveraged state-of-the-art learning models such as Transformers to provide a better match between the contextual representation of a given query and relevant content.

Application in Military Contexts: Used to follow conversations, track the spread of disinformation and keep a finger on social sentiment during conflicts or military operations. It makes it easier for military analysts to track which public view, potential threats and geo-political development touches these waters [43]. For instance, IR systems can discover and retrieve information regarding troop movements, public opposition or threats of violence against military operations.

IR systems have some limitations, such as ambiguity of the language, rapid change of colloquial expression and noisy data source over social media. And periodically (every year or so) search criteria must be updated and improved via developments such as reinforcement learning.

B. TEXT PREPROCESSING FOR SOCIAL MEDIA DATA

Unlike other types of data, social media content is informal in nature comprising slang words, emojis, abbreviations and non-standardized grammar. Before feeding the data into any machine learning or NLP models, the process of cleaning the raw data is done in order to prepare it and this cleansing of text data is termed as Text Preprocessing and the same is illustrated by Figure 6. Text preprocessing aims at making the input data more uniform and removing unnecessary details so that downstream processing accuracy may improve [80].

Following techniques can be used for text preprocessing for social media data:

- Tokenization and sentence splitting for dividing every sentence into words or smaller segments for analysis.

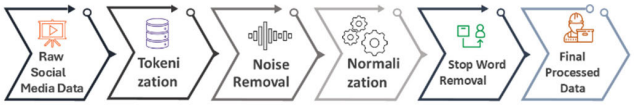


FIGURE 6. Text preprocessing for social media data.

- Lowercasing and normalization for sticking to a common format of the text.
- Removal of noise so that special characters, hyperlinks and other unnecessary data can be removed.
- Expanding abbreviations and handling slang such as changing lol into laughing out loud and other similar entities for better interpret by models.

C. ADVANCED PREPROCESSING WITH CONTEXTUAL EMBEDDINGS

With the advancement in the modern NLP frameworks like BERT and GPT based models, Contextual embeddings come into the picture which takes cares of processing around context to capture better meaning between words. This is important for military intelligence, because a similar phrase may take on meaning depending on the context. Addressing Non-Standard Language and Idiomatic Expressions: Incorporating the ability to recognize or adjust non-standard text san imprecise utilization of phrases helps in preventing oiliness throughout an assortment while taking into account exactness [81]. Efficient data processing increases the precision of sentiment analysis tools, entity recognition models, and topic extraction algorithms which are effectively utilized for public opinion assessment, risk monitoring, and trend prediction [82], [83].

D. AUTOMATED MONITORING AND ALERT SYSTEMS

Automated tracking and alert notice systems are important for the near instantaneous monitoring of social media to detect high-priority threats and military-relevant content. By consistently monitoring social media feeds for keywords, themes, or other patterns that have been defined beforehand (along with looking for anomalies or trends to know when a tipping point has been reached), these systems can be characterized as actuality marketing early warning systems [81], [84], [85]. Figure 7 illustrates the flow of automated monitoring and alert system.

The key features of monitoring systems are given below:

- **Rule-Based Detection**: Tracks military-specific terms and topics from a pre-set list of keywords.
- **Machine Learning for Adaptive Monitoring**: Uses unsupervised learning models to identify unnatural changes or new topics.
- **Anomaly Detection Algorithms**: Deployed to detect rapid increases in activity or language suggesting coordinated disinformation campaigns or threats.
- **Real-Time Applications**: These systems can bring attention to military intelligence information on emerging protests and other threats, or the dissemination of false information related to operations. They can also assist in the monitoring of possible leaks, or breaches into social media of sensitive military information.

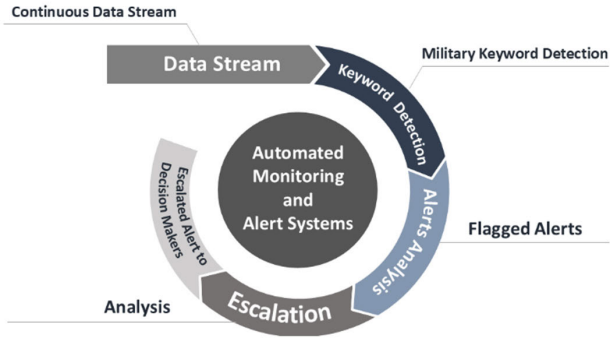


FIGURE 7. Automated monitoring and alert systems.

E. DATA FUSION AND AGGREGATION SYSTEMS

Data fusion and aggregation systems are valuable tools in situations where various data sources must be synthesized into a more comprehensive way as shown in Figure 8. Strategic access to social media data is seldom sufficient without the cooperation of news data, intelligence reporting and other open sources to accurately assess the domestic context. Inferences can be built through personal browsing habits from a variety of platforms that on face are unlikely associates. The key fusion techniques are highlighted below:

- **Rule-Based and Statistical Fusion Model:** Rule-based approaches or statistical aggregations with simple models.
- **Probabilistic Methods**: Combining data based on their source reliability and relevance, assigning Newton scores to the bits of data stacked from a single fluid analysis.
- **Neural Network Models for Deep Fusion**: These models integrate multi-modal data (i.e., text + image analysis) to obtain a deeper understanding of complex situations.
- **Application in Military Intelligence**: Data fusion systems help correlate spikes in social media activity with geopolitical events or corroborate other intelligence data. An AI driven fusion system, on the other hand, can compare different layers of evidence such as real-time social media conversations and satellite imagery against news feeds to correlate troop movements or flag misinformation dissemination.
- **Emerging Trends**: Future innovations in military-grade data fusion will likely include predictive analytics capable of forecasting emerging conflicts or sensing shifts in public opinion before they turn violent.

V. ANALYSIS OF SOCIAL MEDIA DATA RELATED TO MILITARY TOPICS

A. DATA COLLECTION AND PREPROCESSING

Data Collection and Preprocessing from Social Media Platforms is a primary step in analyzing military-related themes.

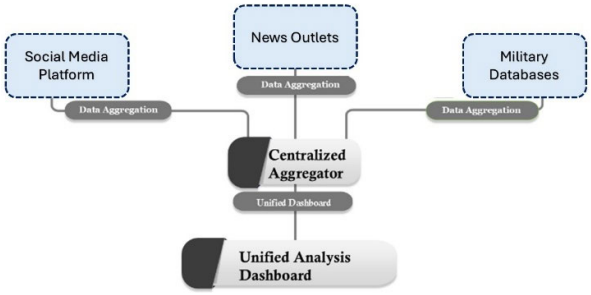


FIGURE 8. Data fusion and aggregation systems.

In the world of military analysis, there are several ways of collecting data, however one of the more common categories would be open-source intelligence (OSINT) based methods which greatly rely on publicly available sources of information. For large-scale data collection, tools like Twitter API, web scraping programs, and data mining frameworks such as Scrapy and Beautiful Soup are commonly used. Furthermore, data extraction tools are available through APIs or third-party tools developed by platforms such as YouTube, Instagram and Facebook for text and video analysis. A necessary step, preprocessing is particularly important for social media data because it can be noisy, unstructured and may contain slang, emoticons and hashtags. Preprocessing Steps like tokenization, stemming, lemmatization and stop word removal for instance Stop word removal eliminates unnecessary words (''and'', ''the''), stemming & lemmatization transform to the root form of any word, so that we can concentrate on some main subjects. Tools such as NLTK, spacy, and the Text Blob toolkit help accomplish these tasks (optimizing the text for effective analysis) [54].

B. KEYWORD AND PHRASE EXTRACTION

Keywords and phrases that can be added in order to help sift through a large dataset. According to the visualization of words using TF-IDF and representation of words as vectors (Word2Vec, GloVe), these techniques help capture names that are most frequently used and are highly relevant. We can improve keyword extraction by applying NER that will recognize whether the keywords is identifying a name, a locations or organizations and apply sentiment analysis to check if post has an emotional tone [86], [87], [88]. When talking about military, it needs special words. words such as ops, RPG, and ETA have domain-specific meanings that will be missed by general-purpose extraction approaches. That is the reason why models are always tuned with military language to increase accuracy. Transformer architecture-based contextual embeddings like BERT and RoBERTa are effective for better detections as they can encapsulate the meaning of a word based on context, differentiating between ambiguous military & non-military [89].

C. CONTEXTUAL CHALLENGES

Contextual difficulties like sarcasm and differences due to ambiguity make social media analysis of military topics complex. An important yet challenging task is sarcasm detection, the sarcastic sentences typically have a strong contrast to literal meanings. Deep learning-based approaches that utilize various deep architectures, such as Recurrent Neural Networks (RNNs) and attention matrix mechanisms for sarcasm detection and context finding, are still viable to improve prediction results albeit demand more computing power with longer training times and larger datasets [90], [91].

Tools such as multilingual BERT (mBERT) and cross-lingual embeddings can be utilized to work with multilingual data, providing analysts with the ability to read in any number of languages or dialects of text. Dealing with ambiguity is similarly difficult, as words can have more than one meaning depending on context [92]. For example, in military analysis and practice where the same term, such as deterrence or influence may carry different meanings by virtue of local geopolitical or cultural factors. Laws regarding privacy and data protection like the GDPR impose stringent regulations on how data can be handled and the privacy of users as shown in Table 5. Military data analysts are required to adhere to these regulations, ensuring that every user identity is protected, and nobody has access to private information unauthorizedly. During data collection and processing, techniques like differential privacy and data anonymization are often used to preserve identities [93], [94].

TABLE 5. Compliance considerations.

Compliance Guideline	Met	Details
Data Anonymization	✔	Applies techniques to remove personal identifiers
Privacy Compliance (e.g., GDPR)	✔	Ensures all methods align with GDPR requirements
Avoiding Unauthorized Access	✔	Access restricted to authorized analysts only
Transparency in Data Collection	✔	Clear information on data use provided
Public Communication	✔	Ensure users understand analysis intentions

D. MILITARY-LANGUAGE PECULIARITIES IN NLP APPLICATIONS

Military communications differ linguistically in a way that can be very challenging and interesting to the Natural Language Processing (NLP) applications. Among the most notable is the use of brevity codes, often used to represent complex operational instructions in a small amount of space, often highly compressed phrases or alphanumeric strings (e.g., LZ represents Landing Zone, and CAS represents Close Air Support). Although such codes facilitate quick communication when other conditions may be time-sensitive, their polysemy (e.g., CAS could also be Collision Avoidance System) can lead to ambiguity in automated analysis unless backed by context modeling that can withstand that type of ambiguity. The other distinguishing feature is the

intentionality of the deception patterns in the opposing communications, such as the switching of code between languages, insertion of deceptive words, and use of false identifiers to conceal intent. Such strategies are meant to confound both human and computer analysis, so NLP models must be able to identify anomalies, detect inconsistencies, and reshape themselves to fit changing word sets. Also, use of military language tends to mix technical terminologies, action terms, and slang terms, and in some cases in multilingual or dialect-rich forms, which further complicate tokenization and semantic analysis. Military NLP systems should be able to do this by integrating specialized ontologies, contextual disambiguation algorithms, and ongoing lexicon updates based on domain expert knowledge to interpret this kind of communication in a real-time context.

VI. APPLICATIONS FOR MILITARY INFORMATION OPERATIONS

Over the years, the integration of Natural Language Processing (NLP) in information warfare operations has gradually risen in importance to encompass the tasks of monitoring/analysis as well as the ability to affect and shape narratives across cyber space. The following is a detailed account of how sentiment analysis, a sub-field of NLP, is conducted in the various levels of military information operations.

A. MONITORING INFORMATION OPERATIONS

NLP contributes significantly to the monitoring and evaluation of campaigns of propaganda, disinformation, and other forms of influence. Scarce or scarce amounts of truthful information is often disseminated using social media in order to win people over, sway the general public, or construct fictitious tales. NLP tools assist in locating and keeping track of the dispersion of such content on sites such as Twitter, Facebook, Instagram and many others. Emotion detection, text categorization and NER are among the tactics covered, which are useful for locating content consistent with misinformation, propaganda techniques or campaigns known to the user. NLP techniques such as mining of sentiment have proven to be very useful in analysis of emotions in regard to information put up or later circulated. Emotion analytics fills the gaps that forever are present in the analysis of formal communications strategy document or social media strategies. It contributes towards understanding the level of attention a specific post can elicit. By identifying specific terms, keywords, and entities associated with propaganda (such as groups, leaders, or movements), NLP can aid in tracking the development and spread of false narratives. This also helps in determining the geographical and social impact of such campaigns [95], [96]. However, such deployments are not without risk. NLP systems may suffer from misclassification errors, such as incorrectly labeling benign posts as threats or failing to detect subtle propaganda, which can result in flawed situational awareness or unjustified responses. Sentiment analysis tools may be biased because of their training data, which can result in incorrect interpretations of what people or enemies are feeling. In the military, such errors can lead to higher risks in operations and unforeseen troubles in fighting information battles. It is important to add confidence scoring, human involvement, and explanation methods into the way AI is used in companies. With XAI frameworks, people can see the reasoning behind how NLP systems come to their conclusions, mainly in important contexts where decisions are vital for ethics and strategy.

B. DETECTION OF POTENTIAL THREATS AND RISKS

Another key application of NLP in military information operations is the detection of early indicators of threats through social media analysis. Monitoring vast amounts of unstructured data from online platforms allows military analysts to identify patterns that may signal emerging threats or destabilizing actions.

- **Social Media Monitoring**: NLP can help detect unusual or risky patterns of discourse, such as aggressive rhetoric, calls for violence, or signs of mobilization related to terrorist activities.
- **Emerging Threat Detection**: NLP techniques such as anomaly detection can identify deviations from normal online behavior, flagging potential threats or risks. For example, a sudden increase in posts about a specific military target or the use of certain keywords could signal the initiation of an operation or impending attack [97].

C. SENTIMENT AND OPINION MINING IN MILITARY CONTEXTS

Public perception and public sentiment are the meat and potatoes of military Information Operations, but especially in regions of conflict. Mood of both domestic and external audiences about military action, policy or political developments is ascertained through sentiment analysis.

- **Public Opinion Analysis**: Sentiment mining can be used to know the perception of general public for military operations or political events, by analyzing social media posts, news articles and public statements. This can help to craft public relations, sway international perspectives and contain internal morale.
- **Impact of Military Actions**: Mil Ops Sentiment mining also allows very important feedback of the effects of military operations. Sentiment analysis, for instance, can analyze public sentiment after a successful mission and indicate whether the general public opinion is positive or negative to provide insight into how communications should proceed [9].

D. CASE STUDIES FOR SUCCESSFUL APPLICATIONS OF NLP IN MILITARY INFORMATION OPERATIONS AND SOCIAL MEDIA MONITORING

There are several case studies which show that NLP could effectively work in ways useful to military information

operations. These case studies shed light on how NLP systems could be leveraged for online military information campaign monitoring, detection, and response.

1) CASE STUDY 1

During the 2016 U.S. Presidential Election, Natural Language Processing (NLP) methods employed by governmental agencies to detect foreign interference in online discussions. Agencies identify potential threats arising from misinformation campaigns by tracking key terms and hashtags as well as analyzing sentiment to take preemptive action [9]. By using sentiment analysis, named entity recognition (NER), and topic modeling, analysts kept an eye on newly developing stories, tracked emotional posts, and found clusters of similar fake actions. NLP tools could identify hashtags that caused people to divide harsh language, and false stories distributed by computer networks and people pretending to be real users. Agencies used data from online discussions to see where disinformation campaigns came from and where they affected people, which led to the discovery of how some groups were being targeted. As a result of using NLP, intelligence agencies were able to identify and isolate suspicious accounts, put out more informed messages to the public, and encourage social media companies to disrupt misleading content. The case study reveals that NLP has an important role in keeping democratic processes and national security safe by improving awareness of events, identifying threats from information, and providing quick advice during online battles.

2) CASE STUDY 2

For the cases of Syria and Iraq, comparable NLP models were utilized by the U.S. DoD as a social media monitoring mechanism to identify terrorist activity and online recruitments [32]. Using certain triggers, military analysts were able to identify trends in online activity that could provide an early warning and react accordingly [9]. With the help of keywords and phrases, NER, sentiment analysis, and topic modeling, analysts could find extremist messages, spot the actions of known dangerous actors, and find unusual rises in online talk about violence or mobilization. Specific words related to weapons, key locations, or recruitment codes were constantly examined with both rule-based and machine learning methods adapted to fit the region. Military intelligence could find out about terrorist plans, propaganda, and recruitment of people with easy targets before they became serious threats. Often, insights from NLP made it possible to react faster by matching what people posted online with ground information and satellite images, which raised the effectiveness of counterterrorism measures. The use of military-grade NLP systems is well demonstrated in the Syria and Iraq case, given the presence of many different languages. It demonstrates that using automated tools on social media can help people understand what's happening, react in the moment, and make decisions in battles against enemies of differing strengths. This case not only demonstrates the effectiveness of NLP in detecting and countering large-scale misinformation campaigns but also offers clear opportunities to adapt these techniques for military intelligence and operational use.

3) MILITARY-SPECIFIC ADAPTATIONS OF NLP TECHNIQUES

The adaptations demonstrated in both case studies enhance situational awareness and enable proactive threat detection and timely response in contested information environments. For instance, integrating NER with geospatial intelligence systems can create automated alerts when certain location codes or base names appear in hostile communications. Sentiment analysis can monitor shifts in public opinion following key military events. Topic modeling, when fused with other intelligence sources such as satellite imagery or SIGINT, can corroborate emerging threats and assist in prioritizing resource allocation. Together, these capabilities form a multi-layered analytical approach that allows defense organizations to counter adversarial information campaigns, protect critical assets, and maintain operational advantage. The case studies thus underscore the value of applying advanced NLP methods not only for monitoring but also for shaping the information environment in alignment with military objectives.

E. CHALLENGES AND LIMITATIONS IN MILITARY SOCIAL MEDIA ANALYSIS USING NLP

Despite the importance of NLP in military social media analysis, there are various challenges and limitations that hinder its effective use. Many of these challenges include technical and data related issues, as well as matters associated with security, privacy, equity, and bias. In this section we outline the basics relevant to NLP applications uncertainties for military. There are a few technical challenges with military NLP for social media analysis, which impacts its performance as well:

- The amount of data on social media is immense, and to gain any operational benefit from it for military purposes there is a need for real-time data analysis. However, real-time processing is a major technical challenge especially when the streams of data we are analyzing become high-volume. This requires a high-performance computing infrastructure and algorithms to deliver fast insights.
- The military-related social media data comes from various sources with a diversity of languages and dialects, including English, Arabic, Chinese, and Russian. The NLP models may not be able to handle data from other languages if trained in one. To rectify this problem, models such as mBERT and XLM-R have been built to run more than one language yet barely end up being cross-dialects or able to manage military terms that are regionally written in unbiased data.
- Another big technical challenge comes from the amount of tonnage social media produces millionnually worldwide. Most NLP systems require text

categorization post processing as they are expected to scale up and handle millions of posts in very little time at higher accuracy. Scaling up models while remaining efficient presents an important technical challenge [98].

- Military terminologies are rich in abbreviations, acronyms, and context-specific jargon, that can significantly impact NER) performance by causing misclassification or omission of critical entities. Standard NER models trained on general-domain data may struggle to accurately recognize and label these specialized terms.
- Cross-cultural sentiment analysis in military coalitions has also some shortcomings because of the linguistic peculiarities, cultural conventions, and divergent displays of emotion. A phrase that is considered motivational by one culture will be interpreted as aggressive or critical in another, which causes incorrect evaluation of morale levels. The use of sarcasm, idioms, and suchlike indirect forms of communication makes it even more difficult to program automated sentiment detection, particularly where there are multiple languages involved. Culturally sensitive lexicons, context-sensitive translation, and sentiment models fine-tuned with military-relevant data in each force are therefore necessary to achieve accurate assessment.

F. DATA AVAILABILITY AND QUALITY

Data quality and availability are major constraints in military social media analysis:

- **Limited Access to Data**: A primary challenge in focusing on data relevant to military applications is that access, both due to privacy considerations and governmental restraints, or even platform limitations make it difficult at best. Although large amounts of data may be available on news, examination papers, or speeches, the datasets themselves can be a challenge to obtain in their entirety
- **Noisy and Unstructured Data**: Social media data is often unstructured, containing informal language, slang, abbreviations, and non-standard expressions, making it challenging for NLP models to accurately interpret and extract meaningful information. Cleaning and preprocessing this data for analysis is a time-consuming task [98].
- **Data Bias and Representation**: Social media data is often limited to specific demographics, ideologies or geographies causing biased insights. This may lead to corrupted data and affect the validity and generalization of NLP models, especially in sensitive military environments.

G. PRIVACY-PRESERVING TECHNOLOGIES IN MILITARY NLP SYSTEMS

It is significant in military NLP that enough data is available while at the same time ensuring privacy to meet both operational and ethical and legal requirements. While it is important to use large-scale textual data to obtain useful details, this data often holds private information and sensitive material. Privacy-preserving technologies provide helpful answers to this problem. Applying differential privacy to datasets or model outputs introduces a little random noise, which allows us to study groups of people without sharing their individual information. This method is especially useful in analyzing sentiment and detecting threats, since it does not violate the privacy of either users or sources. Federated learning is useful since it allows collaborative model training on separate, decentralized data from military or allied sources without exposing any raw data. During military analytics, joint analysis enjoys greater privacy because of secure multi-party computation and homomorphic encryption. Using these technologies allows military NLP systems to continue using data wisely, even while keeping sensitive information secure. NLP systems used in military social media analysis must navigate significant security and privacy challenges:

- **Data Privacy and Compliance:** The personal information presented within the social media data could be classified as personally identifiable information (PII), which is governed by strict privacy regulations, such as the General Data Protection Regulation (GDPR) or other national privacy laws. Special attention must be paid by the analysts to avoid a disclosure of the sensitive data, meaning that it should be ensured it its anonymized enough to prevent others accessing this data and abusing it.
- **Sensitive Information Handling**: Military social media analysis tends to focus on discussions relating to politics, society and security matters. However, the collection and analysis of such sensitive information creates ethical challenges requiring necessary security measures to protect against abuse, unauthorized access or leaks.
- **Cybersecurity**: Cyber-attacks on NLP systems themselves. However, if the social media analysis is powered by a model and it can manipulate or change the classification result via adversarial attacks, this may potentially lead to leaks of classified information. To be impactfully used militarily, these systems need to be secure against such attacks [99].

H. BIAS AND ACCURACY IN MILITARY CONTEXTS

NLP models for military social media analysis must contend with issues of bias and accuracy:

- **Bias in Training Data**: They are usually trained on massive datasets, which by the nature of their constructions could have in-built biases. These might be sort of political, cultural or social biases. In military applications, biased data may result in inaccurate analysis and misreading of the public mood, or even erroneous threat detection [41]

- **Maintaining Accuracy**: It is especially important for NLP models to be accurate in military contexts. Defense applications require high precision in terms of threat detection, sentiment analysis, and entity recognition. A wrong decision or an error in analysis can lead to serious consequences especially in military decisions.
- **Impact of Bias**: Social media analysis of public perception may be affected by biased models, which in turn can affect military operations. If the model is inclined to a specific political or social view, it can lead to misinterpretation of enemy feeling or population issues [100].

VII. RESEARCH GAPS

It is important to explore emerging technologies and strategic advancements regarding NLP in military applications. Accordingly, the following subsections outline research gaps and recommendations to enhance operational capabilities and decision-making in complex environments.

A. ADVANCEMENTS IN NLP FOR MILITARY CONTEXTS

Natural Language Processing (NLP) techniques have seen significant advancement recently pertaining to military use. This progress leans on better and more robust transformers like GP-4, multi-modal NLP where they have included not only text but also images and video data. Additionally, If we continue building in the areas of unsupervised learning and transfer learning, this will provide the potential to adapt even low-resourced NLP models in a way that is appropriate for military applications across many use cases including intelligence collection, tracking operations and more broader battlefield communication systems. Deep learning, reinforcement learning and other emerging technologies could also benefit decision-making automation and accuracy in operational processes [101], [102].

B. ROLE OF AI AND EDGE COMPUTING

AI technologies with NLP are likely to be a considerable addition in military applications like providing military over system processing capabilities in real time and supporting quick and timely decisions as shown in Figure 9. Edge computing, which entails carrying out computation close to data generated (at field devices or local servers), can minimize delay resulting from sending data to central servers. This enables military forces to gain access to real time intelligence, a critical element for tactical decisions on the abstracted over drone battlefield. Merging AI and edge computing also helps ensure military systems stay operational in austere environments, whether individual or collapsed satellite locations that cannot support extensive cloud infrastructure [103], [104].

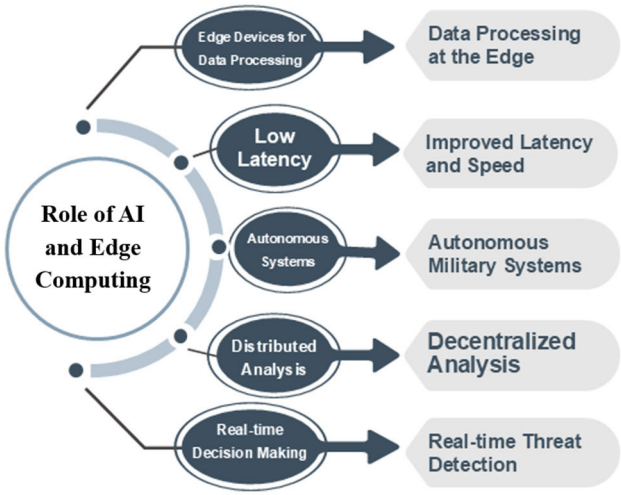


FIGURE 9. Role of AI and Edge Computing.

C. MULTILINGUAL AND MULTIMODAL ANALYSIS

Military operations are executed in a variety of linguistic work environments and typically include many dialects/methods (text, pictures, recordings in that general area). Multilingual analysis is key to effective intelligence as shown in Figure 10.

In recent times, models including XLM-R and mBERT enable NLP systems to work with several languages such as low-resource languages often present in conflict areas. This will help in a more contextual understanding of what is happening when we analyze not just text but also audio, video and images called Multimodal analysis helping the detection of military operations and threats in an ever-wider perspective. These capabilities can provide improved drivers in understanding public opinion, media coverage and military reaction in real time [105], [106], [107]. To enhance the performance of such systems, multimodal fusion techniques are increasingly being employed, including cross-modal attention mechanisms. The cross-modal attention mechanisms allow models to dynamically align and interpret correlated information across different modalities such as text and imagery. For example, attention-based fusion can help to associate textual sentiment from social media posts with visual cues from accompanying images or videos. It enables richer situational awareness in information operations. Approaches such as early fusion (feature-level integration) and late fusion (decision-level integration) are also being explored in military contexts to combine heterogeneous data sources for robust threat detection and narrative analysis [108].

To support real-time multilingual processing in operational environments, recent advancements include the use of lightweight transformer models (e.g., DistilBERT, TinyBERT) and quantized versions of large language models for deployment on edge devices [109]. These models minimize memory footprint and computational overhead, making them suitable for field environments. Edge computing frameworks enable low-latency responses (under 500 ms) and high-throughput processing (exceeding 1,000 tokens

per second), which are important for timely threat detection and situational assessment in military missions [110], [111], [112].

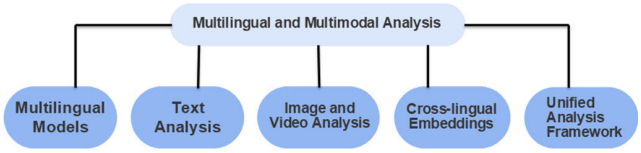


FIGURE 10. Multilingual and Multimodal Analysis.

D. IMPROVING INTERPRETABILITY AND EXPLAINABILITY

In the military domain, AI and NLP decisions can have far-reaching consequences, so interpretability and explainability of these systems is key consideration as shown in Figure 11.

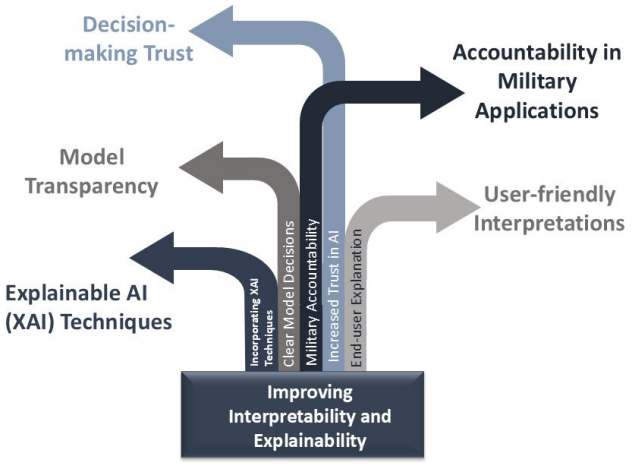


FIGURE 11. Improving Interpretability and Explainability.

Transformers and neural networks, commonly used models for deep learning now, are often treated as "black boxes," meaning that it is unclear how they arrived at a specific decision. Explainable AI (XAI) is already working to remedy this problem; approaches are being developed that can provide insight into model behavior. It is an important step towards securing confidence from the military end, where transparency and accountability matter. Attention mechanism, model agnostic interpretability methods and post-hoc explanation frameworks are some methods which will aid the developing of transparent and trustworthy systems [44].

Table 6 summarizes key future directions in applying advanced technologies to military NLP systems. It highlights emerging technologies such as advanced transformers, AI integration with edge computing, multilingual and multimodal analysis, and explainable AI. Each direction is paired with its prospective benefits.

TABLE 6. Emerging Technologies and Their Benefits in Future Military NLP Applications.

Future Direction	Technology	Prospective Benefits
Advancements in NLP for Military Contexts	Transformers (e.g., GPT-4), unsupervised and transfer learning, deep/reinforcement learning	Enhanced intelligence gathering, adaptable low-resource models, improved operational decision-making
Integration with AI and Edge Computing	AI with Edge Computing	Real-time processing, reduced latency, continued functionality in disconnected environments
Multilingual and Multimodal Analysis	XLM-R, mBERT, multimodal NLP (text, audio, image, video)	Context-aware intelligence, threat detection across diverse languages and media formats
Improving Interpretability and Explainability	Explainable AI (XAI), attention mechanisms, model-agnostic methods	Increased transparency, accountability, and trust in AI-driven decisions

VIII. CONCLUSION

This survey has demonstrated how information processing methods for NLP can aid military operations, particularly in social media intelligence gathering. Our research focused on NLP tools for tracking and combating disinformation, detecting propaganda, and identifying emerging threats. Through techniques such as sentiment analysis, NER, and keyword extraction, military analysts can efficiently process vast volumes of social media data to identify resonant narratives and assess public sentiment toward military actions. Coupled with advancements in AI and edge computing, this capability enables near real-time insights that support faster and more informed decisions in dynamic environments. A key takeaway is the strategic value of multilingual and multimodal analysis, especially in operations across linguistically and culturally diverse regions. By integrating models like mBERT and Transformer-based architectures, which accommodate contextual language variations, the military can significantly enhance intelligence precision and broaden situational awareness. These applications have immediate

operational implications, from threat forecasting to decision support in mission-critical scenarios. However, to unlock the full potential of NLP in defense contexts, several priorities must be addressed:

1. Scalability and real-time deployment of NLP models, especially for edge environments, where latency and bandwidth are constrained.
2. Data quality and ethical compliance, military NLP systems must rely on accurate, unbiased, and secure data, raising the need for privacy-preserving mechanisms like differential privacy and anonymization.
3. Model interpretability and trustworthiness, XAI techniques should be further developed and integrated to support transparent decision-making in high-stakes applications.

We have outlined a three-part strategy for future advancement:

- Develop field-ready NLP systems with multilingual and multimodal capabilities tailored to conflict zones, especially those involving low-resource languages.
- Prioritize funding for edge AI infrastructure, ethical NLP frameworks, and robust data governance models.
- Focus on interpretable, domain-adaptive models that address context variability and support human-in-the-loop systems for oversight.

REFERENCES

[1] F. Aslam, "The impact of artificial intelligence on chatbot technology: A study on the current advancements and leading innovations," *Eur. J. Technol.*, vol. 7, no. 3, pp. 62–72, Aug. 2023, doi: 10.47672/ejt.1561.

[2] E. Kasthuri and S. Balaji, "Natural language processing and deep learning chatbot using long short term memory algorithm," *Mater. Today, Proc.*, vol. 81, pp. 690–693, Jan. 2023, doi: 10.1016/j.matpr.2021.04.154.

[3] I. Awasthi, K. Gupta, P. S. Bhogal, S. S. Anand, and P. K. Soni, "Natural language processing (NLP) based text summarization—A survey," in *Proc. 6th Int. Conf. Inventive Comput. Technol. (ICICT)*, Jan. 2021, pp. 1310–1317, doi: 10.1109/ICICT50816.2021.9358703.

[4] P. Lavanya and E. Sasikala, "Deep learning techniques on text classification using natural language processing (NLP) in social healthcare network: A comprehensive survey," in *Proc. 3rd Int. Conf. Signal Process. Commun. (ICPSC)*, May 2021, pp. 603–609, doi: 10.1109/ICSPC51351.2021.9451752.

[5] A. Albladi, M. Islam, and C. Seals, "Sentiment analysis of Twitter data using NLP models: A comprehensive review," *IEEE Access*, vol. 13, pp. 30444–30468, 2025, doi: 10.1109/ACCESS.2025.3541494.

[6] P. A. Olujimi and A. Ade-Ibijola, "NLP techniques for automating responses to customer queries: A systematic review," *Discover Artif. Intell.*, vol. 3, no. 1, p. 20, May 2023, doi: 10.1007/s44163-023-00065-5.

[7] C. Stokes, A. Kumar, F. Choi, and R. Weischedel, "Scaling NLP algorithms to meet high demand," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Oct. 2015, p. 2839, doi: 10.1109/BIGDATA.2015.7364095.

[8] C. Gunasekara, T. Carryer, and M. Triff, "On natural language processing applications for military dialect classification," in *Proc. 20th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2021, pp. 211–218, doi: 10.1109/ICMLA52953.2021.00040.

[9] A. Quilez-Robres, M. Acero-Ferrero, D. Delgado-Bujedo, R. Lozano-Blasco, and M. Aiger-Valles, "Social networks in military powers: Network and sentiment analysis during the COVID-19 pandemic," *Computation*, vol. 11, no. 6, p. 117, Jun. 2023, doi: 10.3390/computation11060117.

[10] P. H. L. Rettore, P. Zißner, M. Alkhowaiter, C. Zou, and P. Sevenich, "Military data space: Challenges, opportunities, and use cases," *IEEE Commun. Mag.*, vol. 62, no. 1, pp. 70–76, Jan. 2024, doi: 10.1109/MCOM.001.2300396.

[11] D. Aminova, "Military terminology as a special layer of the lexical system of language," *Western Eur. J. Historical Events Social Sci.*, vol. 1, no. 3, pp. 23–25, 2023.

[12] S. Bradshaw, U. Campbell-Smith, A. Henle, A. Perini, S. Shalev, H. Bailey, and P. N. Howard, *Country Case Studies Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*. Oxford, U.K.: Oxford Internet Institute, 2021. Accessed: Oct. 24, 2024.

[13] Y. Liang, C. Zhang, and L. Xu, "Detecting terrorism-related content using NLP techniques," *J. Mil. Intell.*, vol. 7, no. 1, pp. 25–38, Jun. 2020.

[14] M. Horvat, G. Gledec, and F. Leontić, "Hybrid natural language processing model for sentiment analysis during natural crisis," *Electronics*, vol. 13, no. 10, p. 1991, May 2024, doi: 10.3390/electronics13101991.

[15] H. T. Phan, V. C. Tran, N. T. Nguyen, and D. Hwang, "Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model," *IEEE Access*, vol. 8, pp. 14630–14641, 2020, doi: 10.1109/ACCESS.2019.2963702.

[16] M. Lepelaar, A. Wahby, M. Rossouw, L. Nikitin, K. Tibble, P. J. Ryan, and R. B. Watson, "Sentiment analysis of social survey data for local city councils," *J. Sensor Actuator Netw.*, vol. 11, no. 1, p. 7, Jan. 2022, doi: 10.3390/jsan11010007.

[17] A. D. Dubey and S. Tripathi, "Analysing the sentiments towards work-from-home experience during COVID-19 pandemic," *J. Innov. Manage.*, vol. 8, no. 1, pp. 13–19, Apr. 2020, doi: 10.24840/2183-0606_008.001_0003.

[18] M. H. Hoti and J. Ajdari, "Sentiment analysis using the vader model for assessing company services based on posts on social media," *SEEU Rev.*, vol. 18, no. 2, pp. 19–33, Dec. 2023, doi: 10.2478/seeur-2023-0043.

[19] K. B. Hansen and C. Borch, "Alternative data and sentiment analysis: Prospecting non-standard data in machine learning-driven finance," *Big Data Soc.*, vol. 9, no. 1, Jan. 2022, Art. no. 205395172110707, doi: 10.1177/20539517211070701.

[20] S. Kumi, C. C. Snow, R. K. Lomotey, and R. Deters, "Uncovering concerns of citizens through machine learning and social network sentiment analysis," *IEEE Access*, vol. 12, pp. 94885–94913, 2024, doi: 10.1109/ACCESS.2024.3426329.

[21] M. Naldi, "A review of sentiment computation methods with R packages," 2019, *arXiv:1901.08319*.

[22] J. Yin, Z. Qiao, L. Han, and X. Zhang, "EEG-based emotion recognition with autoencoder feature fusion and MSC-TimesNet model," *Comput. Methods Biomechanics Biomed. Eng.*, vol. 2025, pp. 1–18, Mar. 2025, doi: 10.1080/10255842.2025.2477801.

[23] T. Zatonatska, G. Kharlamova, V. Pakholchuk, and A. Syzov, "Natural language processing methods application in defense budget analysis," *Stud. Bus. Econ.*, vol. 19, no. 2, pp. 290–307, Aug. 2024, doi: 10.2478/sbe-2024-0039.

[24] A. Mohamed, N. Allam, and M. H. Haggag, "The question answering systems: A survey," *Int. J. Research Reviews Inf. Sci.*, vol. 2, no. 3, 2012.

[25] N. Patwardhan, S. Marrone, and C. Sansone, "Transformers in the real world: A survey on NLP applications," *Information*, vol. 14, no. 4, p. 242, Apr. 2023, doi: 10.3390/info14040242.

[26] S. Gurrapu, A. Kulkarni, L. Huang, I. Lourentzou, and F. A. Batarseh, "Rationalization for explainable NLP: A survey," *Frontiers Artif. Intell.*, vol. 6, 2023, Art. no. 1225093, doi: 10.3389/frai.2023.1225093.

[27] S. P. Lende and M. M. Raghuwanshi, "Closed domain question answering system using NLP techniques," *Int. J. Eng. Sci. Res. Technol.*, vol. 5, pp. 632–639, May 2016.

[28] R. Uppal, "Natural Language Processing (NLP) can automate military intelligence and provide hands-free communication between a soldier and robots," *Int. Defense Secur. Technol.*, Oct. 2022. Accessed: Jan. 6, 2025. [Online]. Available: https://idstch.com/technology/ict/natural-language-processing-nlp-can-automate-military-intelligence-and-provide-hands-free-communication-between-a-soldier-and-robots/

[29] R. Patil, S. Boit, V. Gudivada, and J. Nandigam, "A survey of text representation and embedding techniques in NLP," *IEEE Access*, vol. 11, pp. 36120–36146, 2023, doi: 10.1109/ACCESS.2023.3266377.

[30] C. Barata, V. Rotemberg, N. C. F. Codella, P. Tschandl, C. Rinner, B. N. Akay, Z. Apalla, G. Argenziano, A. Halpern, A. Lallas, C. Longo, J. Malvehy, S. Puig, C. Rosendahl, H. P. Soyer, I. Zalaudek, and H. Kittler, "A reinforcement learning model for AI-based decision support in skin cancer," *Nature Med.*, vol. 29, no. 8, pp. 1941–1946, Aug. 2023, doi: 10.1038/s41591-023-02475-5.

[31] K. E. Wu, K. E. Yost, H. Y. Chang, and J. Zou, "BABEL enables cross-modality translation between multiomic profiles at single-cell resolution," *Proc. Nat. Acad. Sci. USA*, vol. 118, no. 15, Apr. 2021, Art. no. e2023070118, doi: 10.1073/pnas.2023070118.

[32] M. Müller, S. Choudhary, C. Chung, A. Mouchtaris, and S. Kunzmann, "In pursuit of babel–multilingual end-to-end spoken language understanding," in *Proc. IEEE Autom. Speech Recognit. Understand. Workshop (ASRU)*, Dec. 2021, pp. 1042–1049, doi: 10.1109/ASRU51503.2021.9688263.

[33] N. Schörnig, "Artificial intelligence as an arms control tool: Opportunities and challenges," in *Armament, Arms Control and Artificial Intelligence* (Studies in Peace and Security), T. Reinhold and N. Schörnig, Eds., Cham, Switzerland: Springer, 2022, doi: 10.1007/978-3-031-11043-6_5.

[34] L. Qian, I. Zitouni, and J. F. Pitrelli, "Making the case for an open, unified system architecture in response to rapid developments in the natural language industry: Translingual automatic language exploration the system (TALES)," in *Where Humans Meet Machines*. New York, NY, USA: Springer, 2013, pp. 3–25, doi: 10.1007/978-1-4614-6934-6_1.

[35] Y.-H. Zhang and X. F. Liu, "Satellite broadcasting enabled blockchain protocol: A preliminary study," in *Proc. Inf. Commun. Technol. Conf. (ICTC)*, May 2020, pp. 118–124, doi: 10.1109/ICTC49638.2020.9123248.

[36] I. El Alaoui, Y. Gahi, R. Messoussi, Y. Chaabi, A. Todoskoff, and A. Kobi, "A novel adaptable approach for sentiment analysis on big social data," *J. Big Data*, vol. 5, no. 1, p. 12, Dec. 2018, doi: 10.1186/s40537-018-0120-0.

[37] A. Sadia, F. Khan, and F. Bashir, "An overview of lexicon-based approach for sentiment analysis," in *Proc. Int. Electr. Eng. Conf.*, vol. 1, 2018, pp. 1–6.

[38] H. Allam, L. Makubvure, B. Gyamfi, K. N. Graham, and K. Akinwolere, "Text classification: How machine learning is revolutionizing text categorization," *Information*, vol. 16, no. 2, p. 130, Feb. 2025, doi: 10.3390/info16020130.

[39] S. Lee, E. Kim, K. Kim, J. Lee, and W. Choi, "Standardization trends and military communication applications of core technologies of 5G NR," *J. Korea Inst. Mil. Sci. Technol.*, vol. 27, no. 5, pp. 611–618, Oct. 2024, doi: 10.9766/kimst.2024.27.5.611.

[40] K. H. Abdullah and M. A. Ismail, "Analyzing social media approaches in military campaigns: A scientometric and scoping perspective," *J. Social Sci.*, vol. 2, no. 3, pp. 109–122, 2025, doi: 10.61552/SJSS.2025.03.001.

[41] K. Vyas, D. Murphy, and N. Greenberg, "Cognitive biases in military personnel with and without PTSD: A systematic review," *J. Mental Health*, vol. 32, no. 1, pp. 248–259, Jan. 2023, doi: 10.1080/09638237.2020.1766000.

[42] O. Marukhovskyi, O. Marukhovska-Kartunova, S. Kolybabiuk, L. Pavlova, and O. Karchevska, "The role of online media in shaping public opinion about the prospects for peace and military conflict," *Pakistan J. Criminol.*, vol. 16, no. 2, pp. 431–446, 2024, doi: 10.62271/pjc.16.1.431.446.

[43] N. U. Ahmed, "Integrating machine learning in military intelligence process: Study of futuristic approaches towards human–machine collaboration," *NDC e-J.*, vol. 2, no. 1, pp. 59–89, 2021.

[44] N. G. Wood, "Explainable AI in the military domain," *Ethics Inf. Technol.*, vol. 26, no. 2, pp. 1–13, Jun. 2024, doi: 10.1007/s10676-024-09762-w.

[45] J. Titterton, "Military intelligence: Misdirection, misinformation and espionage," in *Deception in Medieval Warfare: Trickery and Cunning in the Central Middle Ages*. Boydell & Brewer, 2022, pp. 27–52. [Online]. Available: https://doi.org/10.1017/9781800104747.004

[46] A. B. Rashid, A. K. Kausik, A. Al Hassan Sunny, and M. H. Bappy, "Artificial intelligence in the military: An overview of the capabilities, applications, and challenges," *Int. J. Intell. Syst.*, vol. 2023, no. 1, 2023, Art. no. 8676366.

[47] X. Li, W. Zhang, X. Zhao, W. Pu, P. Chen, and F. Liu, "Wartime industrial logistics information integration: Framework and application in optimizing deployment and formation of military logistics platforms," *J. Ind. Inf. Integr.*, vol. 22, Jun. 2021, Art. no. 100201, doi: 10.1016/j.jii.2021.100201.

[48] J. J. Galán, R. A. Carrasco, and A. LaTorre, "Military applications of machine learning: A bibliometric perspective," *Mathematics*, vol. 10, no. 9, p. 1397, Apr. 2022, doi: 10.3390/math10091397.

[49] G. Csanádi, "Information management in NATO (part one)," *Appl. Mil. Sci.*, vol. 1, pp. 136–158, Mar. 2018.

[50] I. Almomani, B. Al-Kasasbeh, and M. Al-Akhras, "WSN-DS: A dataset for intrusion detection systems in wireless sensor networks," *J. Sensors*, vol. 2016, no. 1, 2016, Art. no. 4731953. Accessed: Nov. 1, 2023. [Online]. Available: https://www.kaggle.com/datasets/bassamkasasbeh1/wsnds

[51] M. Haghi Kashani, A. M. Rahmani, and N. Jafari Navimipour, "Quality of service-aware approaches in fog computing," *Int. J. Commun. Syst.*, vol. 33, no. 8, pp. 1–34, May 2020, doi: 10.1002/dac.4340.

[52] S. Miao and Z. Tang, "Utilizing human processing for fuzzy-based military situation awareness based on social media," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2017, pp. 1–6, doi: 10.1109/FUZZ-IEEE.2017.8015709.

[53] Y. Peng, Y. Zhao, J. Dong, and J. Hu, "Adaptive opinion dynamics over community networks when agents cannot express opinions freely," *Neurocomputing*, vol. 618, Feb. 2025, Art. no. 129123, doi: 10.1016/j.neucom.2024.129123.

[54] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *NAACL HLT-Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol. - Proc. Conf.*, 2018, pp. 4171–4186.

[55] J. C. Campbell, A. Hindle, and E. Stroulia, "Latent Dirichlet allocation: Extracting topics from software engineering data," *Art Sci. Anal. Softw. Data*, vol. 3, pp. 139–159, Apr. 2015, doi: 10.1016/B978-0-12-411519-4.00006-9.

[56] N. Cavus, S. E. Mrwebi, I. Ibrahim, T. Modupeola, and A. Y. Reeves, "Internet of Things and its applications to smart campus: A systematic literature review," *Int. J. Interact. Mobile Technol.*, vol. 16, no. 23, pp. 17–35, Dec. 2022, doi: 10.3991/ijim.v16i23.36215.

[57] L. Nicolescu and M. T. Tudorache, "Human–computer interaction in customer service: The experience with AI chatbots—A systematic literature review," *Electronics*, vol. 11, no. 10, p. 1579, May 2022, doi: 10.3390/electronics11101579.

[58] M. Hosseini Shirvani and R. Noorian Talouki, "Bi-objective scheduling algorithm for scientific workflows on cloud computing platform with makespan and monetary cost minimization approach," *Complex Intell. Syst.*, vol. 8, no. 2, pp. 1085–1114, Apr. 2022, doi: 10.1007/s40747-021-00528-1.

[59] Z. Ahmad, A. I. Jehangiri, M. A. Ala'anzy, M. Othman, R. Latip, S. K. U. Zaman, and A. I. Umar, "Scientific workflows management and scheduling in cloud computing: Taxonomy, prospects, and challenges," *IEEE Access*, vol. 9, pp. 53491–53508, 2021, doi: 10.1109/ACCESS.2021.3070785.

[60] Z. Wang, C. Zhang, Z. Chen, W. Hu, K. Lu, L. Ge, and Z. Wang, "ACR-Net: Learning high-accuracy optical flow via adaptive-aware correlation recurrent network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 10, pp. 9064–9077, Oct. 2024, doi: 10.1109/TCSVT.2024.3395636.

[61] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing [review article]," *IEEE Comput. Intell. Mag.*, vol. 13, no. 3, pp. 55–75, Aug. 2018, doi: 10.1109/MCI.2018.2840738.

[62] D. Gupta, "Progress in information retrieval: An extensive analysis," *J. Adv. Res. Libr. Inf. Sci.*, vol. 2023, pp. 37–43, Dec. 2023.

[63] K. Mohiuddin, M. A. Alam, M. M. Alam, P. Welke, M. Martin, J. Lehmann, and S. Vahdati, "Retention is all you need," in *Proc. 32nd ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2023, pp. 4752–4758, doi: 10.1145/3583780.3615497.

[64] A. Sharma, A. Amrita, S. Chakraborty, and S. Kumar, "Named entity recognition in natural language processing: A systematic review," in *Proc. 2nd Doctoral Symp. Comput. Intell.*, 2021, pp. 817–828, doi: 10.1007/978-981-16-3346-1_66.

[65] Y. Peng, X. Chen, D. Miao, X. Qin, X. Gu, and P. Lu, "Unveiling user identity across social media: A novel unsupervised gradient semantic model for accurate and efficient user alignment," *Complex Intell. Syst.*, vol. 11, no. 1, p. 24, Jan. 2025, doi: 10.1007/s40747-024-01626-6.

[66] G. Shi, S. Deng, B. Wang, C. Feng, Y. Zhuang, and X. Wang, "One for all: A unified generative framework for image emotion classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 8, pp. 7057–7068, Aug. 2024, doi: 10.1109/TCSVT.2023.3341840.

[67] O. Araque, G. Zhu, and C. A. Iglesias, "A semantic similarity-based perspective of affect lexicons for sentiment analysis," *Knowl.-Based Syst.*, vol. 165, pp. 346–359, Feb. 2019, doi: 10.1016/j.knosys.2018.12.005.

[68] S. Deng, L. Wu, G. Shi, L. Xing, M. Jian, Y. Xiang, and R. Dong, "Learning to compose diversified prompts for image emotion classification," *Comput. Vis. Media*, vol. 10, no. 6, pp. 1169–1183, Dec. 2024, doi: 10.1007/s41095-023-0389-6.
[69] W. Song, X. Wang, Y. Jiang, S. Li, A. Hao, X. Hou, and H. Qin, "Expressive 3D facial animation generation based on local-to-global latent diffusion," *IEEE Trans. Vis. Comput. Graphics*, vol. 30, no. 11, pp. 7397–7407, Nov. 2024, doi: 10.1109/TVCG.2024.3456213.
[70] A. M. Shah, R. A. Naqvi, and O.-R. Jeong, "Detecting topic and sentiment trends in physician rating websites: Analysis of online reviews using 3-wave datasets," *Int. J. Environ. Res. Public Health*, vol. 18, no. 9, p. 4743, Apr. 2021, doi: 10.3390/ijerph18094743.
[71] N. Thakur, S. Cui, V. Knieling, K. Khanna, and M. Shao, "Investigation of the misinformation about COVID-19 on YouTube using topic modeling, sentiment analysis, and language analysis," *Computation*, vol. 12, no. 2, p. 28, Feb. 2024, doi: 10.3390/computation12020028.
[72] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, vol. 30, pp. 5998–6008.
[73] N. Loukachevitch and B. Dobrov, "RuThes thesaurus for natural language processing," in *The Palgrave Handbook of Digital Russia Studies*, D. Gritsenko, M. Wijermars, and M. Kopotev, Eds., Cham, Switzerland: Palgrave Macmillan, 2021. [Online]. Available: https://doi.org/10.1007/978-3-030-42855-6_18
[74] P. S. Giannaris, J. Borges, and N. Doukas, "Design of a defense controlled vocabulary, SKOS-based encoding of terms from military press releases," in *Proc. 13th Int. Conf. Dependable Syst., Services Technol. (DESSERT)*, Oct. 2023, pp. 1–6, doi: 10.1109/DESSERT61349.2023.10416475.
[75] R. Smaliukiene, "Sustainability issues in the military: Genesis and prospects," *J. Secur. Sustainability Issues*, vol. 8, no. 1, pp. 19–32, Sep. 2018, doi: 10.9770/jssi.2018.8.1(2).
[76] H. Shi, S. D. Dao, and J. Cai, "LLMFormer: Large language model for open-vocabulary semantic segmentation," *Int. J. Comput. Vis.*, vol. 133, no. 2, pp. 742–759, Feb. 2025, doi: 10.1007/s11263-024-02171-y.
[77] G. Valiyev, M. Piraino, A. Kok, M. Street, I. I. Mestric, and R. Birger, "Initial exploitation of natural language processing techniques on NATO strategy and policies," *Inf. Secur., Int. J.*, vol. 47, no. 2, pp. 187–202, 2020, doi: 10.11610/isij.4713.
[78] H. Shi, M. Hayat, and J. Cai, "Unified open-vocabulary dense visual prediction," *IEEE Trans. Multimedia*, vol. 26, pp. 8704–8716, 2024, doi: 10.1109/TMM.2024.3381835.
[79] E. M. Wegrzyn, "Social media as an intelligence source: Collection and analysis of tweets during the russian military buildup and air campaign in Syria," Master thesis, Amer. Mil.Univ., 2016.
[80] S. Pradha, M. N. Halgamuge, and N. Tran Quoc Vinh, "Effective text data preprocessing technique for sentiment analysis in social media data," in *Proc. 11th Int. Conf. Knowl. Syst. Eng. (KSE)*, Oct. 2019, pp. 1–8, doi: 10.1109/KSE.2019.8919368.
[81] S.-Y. Huang and T. Ban, "Monitoring social media for vulnerability-threat prediction and topic analysis," in *Proc. IEEE 19th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Dec. 2020, pp. 1771–1776, doi: 10.1109/TrustCom50675.2020.00243.
[82] E. Wang, Y. Yang, J. Wu, W. Liu, and X. Wang, "An efficient prediction-based user recruitment for mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 17, no. 1, pp. 16–28, Jan. 2018, doi: 10.1109/TMC.2017.2702613.
[83] H. Li, C. Xia, T. Wang, Z. Wang, P. Cui, and X. Li, "GRASS: Learning spatial–temporal properties from chainlike cascade data for microscopic diffusion prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 11, pp. 16313–16327, Nov. 2024, doi: 10.1109/TNNLS.2023.3293689.
[84] J. Hu, H. Jiang, S. Chen, Q. Zhang, Z. Xiao, D. Liu, J. Liu, and B. Li, "WiShield: Privacy against Wi-Fi human tracking," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 10, pp. 2970–2984, Oct. 2024, doi: 10.1109/JSAC.2024.3414597.
[85] G. Alwakid, M. Humayun, Z. Ahmad, and M. Shaheen, "Development of a sustainable and optimized energy ecosystem with real-time monitoring through innovative IoT-based actuator design for human-centric security in smart cities," in *Human-Centric Computing and Information Sciences*, 2025, p. 15.
[86] T. Pires, E. Schlinger, and D. Garrette, "How multilingual is multilingual BERT?" in *Proc. 57th Annu. Meeting Assoc. Comput. Linguistics*, Jun. 2019, pp. 4996–5001, doi: 10.18653/v1/p19-1493.
[87] C. Zhu, "Research on emotion recognition-based smart assistant system: Emotional intelligence and personalized services," *J. Syst. Manag. Sci.*, vol. 13, no. 5, pp. 227–242, Sep. 2023, doi: 10.33168/jsms.2023.0515.
[88] J. Ding, X. Chen, P. Lu, Z. Yang, X. Li, and Y. Du, "DialogueINAB: An interaction neural network based on attitudes and behaviors of interlocutors for dialogue emotion recognition," *J. Supercomput.*, vol. 79, no. 18, pp. 20481–20514, Dec. 2023, doi: 10.1007/s11227-023-05439-1.
[89] C. Dwork, "Differential privacy: A survey of results," in *Proc. Int. Conf. Theory Appl. Models Comput.*, vol. 4978, 2008, pp. 1–19, doi: 10.1007/978-3-540-79228-4_1.
[90] M. Taddeo, D. McNeish, A. Blanchard, and E. Edgar, "Ethical principles for artificial intelligence in national defence," *Philosophy Technol.*, vol. 34, no. 4, pp. 1707–1729, Dec. 2021, doi: 10.1007/s13347-021-00482-3.
[91] D. Li, K. D. Ortegas, and M. White, "Exploring the computational effects of advanced deep neural networks on logical and activity learning for enhanced thinking skills," *Systems*, vol. 11, no. 7, p. 319, Jun. 2023, doi: 10.3390/systems11070319.
[92] H.-H. Zhao, T.-L. Ji, P. L. Rosin, Y.-K. Lai, W.-L. Meng, and Y.-N. Wang, "Cross-lingual font style transfer with full-domain convolutional attention," *Pattern Recognit.*, vol. 155, Nov. 2024, Art. no. 110709, doi: 10.1016/j.patcog.2024.110709.
[93] Y. Niu, H. Zhang, and J. Li, "A pitman-yor process self-aggregated topic model for short texts of social media," *IEEE Access*, vol. 9, pp. 129011–129021, 2021, doi: 10.1109/ACCESS.2021.3113320.
[94] X. Zou, J. Yuan, P. Shilane, W. Xia, H. Zhang, and X. Wang, "From hyper-dimensional structures to linear structures: Maintaining deduplicated data's locality," *ACM Trans. Storage*, vol. 18, no. 3, pp. 1–28, Aug. 2022, doi: 10.1145/3507921.
[95] R. H. Raven. (2017). *Natural Language Processing of Online Propaganda as a Means of Passively Monitoring an Adversarial Ideology*. [Online]. Available: http://hdl.handle.net/10945/52993%0Ahttps://calhoun.nps.edu/handle/10945/52993
[96] N. Hussain, M. Ali, S. A. Syed, R. M. Ghoniem, N. Ejaz, O. I. Alramli, M. A. Ala'anzy, and Z. Ahmad, "Design and evaluation of Arabic handwritten digit recognition system using biologically plausible methods," *Arabian J. Sci. Eng.*, vol. 49, no. 9, pp. 12509–12523, Sep. 2024, doi: 10.1007/s13369-024-08710-6.
[97] M. Chaudhary and D. Bansal, "Open source intelligence extraction for terrorism-related information: A review," *WIREs Data Mining Knowl. Discovery*, vol. 12, no. 5, pp. 1–35, Sep. 2022, doi: 10.1002/widm.1473.
[98] N. Hussain et al., "Design and evaluation of arabic handwritten digit recognition system using biologically plausible methods," *Arab. J. Sci. Eng.*, vol. 49, pp. 12509–12523, 2024, doi: 10.1007/s13369-024-08710-6.
[99] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," 2014, *arXiv:1412.6572*.
[100] J. Paavola and H. Jalonen, "An approach to detect and analyze the impact of biased information sources in the social media," in *Proc. 14th Eur. Conf. Cyber Warfare Secur. (ECCWS)*, Jul. 2015, p. 213.
[101] S. Ruder, "An overview of multi-task learning in deep neural networks," 2017, *arXiv:1706.05098*.
[102] P.-R. Liu, L. Lu, J.-Y. Zhang, T.-T. Huo, S.-X. Liu, and Z.-W. Ye, "Application of artificial intelligence in medicine: An overview," *Current Med. Sci.*, vol. 41, no. 6, pp. 1105–1115, Dec. 2021, doi: 10.1007/s11596-021-2474-3.
[103] J. Fix, "Integration of AI and edge computing: Exploring the synergy between artificial intelligence and edge computing for enhanced IoT applications," *Distrib. Learn. Broad Appl. Sci. Res.*, vol. 9, pp. 253–260, Jun. 2023.
[104] L. Song, G. Sun, H. Yu, and D. Niyato, "ESPD-LP: Edge service pre-deployment based on location prediction in MEC," *IEEE Trans. Mobile Comput.*, vol. 24, no. 6, pp. 5551–5568, Jun. 2025, doi: 10.1109/TMC.2025.3533005.
[105] J. Wagner, "Multilingual and multimodal interactions," *Appl. Linguistics*, vol. 39, no. 1, pp. 99–107, Feb. 2018, doi: 10.1093/applin/amx058.
[106] A. Wilson, S. Wilkes, Y. Teramoto, and S. Hale, "Multimodal analysis of disinformation and misinformation," *Roy. Soc. Open Sci.*, vol. 10, no. 12, 2023, Art. no. 230964.
[107] Q. Hu, Q. He, H. Zhong, G. Lu, X. Zhang, G. Zhai, and Y. Wang, "VARFVV: View-adaptive real-time interactive free-view video streaming with edge computing," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 7, pp. 2620–2634, Jul. 2025, doi: 10.1109/JSAC.2025.3559140.

[108] L. Wu, Y. Long, C. Gao, Z. Wang, and Y. Zhang, "MFIR: Multimodal fusion and inconsistency reasoning for explainable fake news detection," *Inf. Fusion*, vol. 100, Dec. 2023, Art. no. 101944, doi: 10.1016/j.inffus.2023.101944.

[109] F. Huang, Y. Bei, Z. Yang, J. Jiang, H. Chen, Q. Shen, S. Wang, F. Karray, and P. S. Yu, "Large language model simulator for cold-start recommendation," in *Proc. 18th ACM Int. Conf. Web Search Data Mining*, Mar. 2025, pp. 261–270, doi: 10.1145/3701551.3703546.

[110] I. Afzal, N. U. Amin, Z. Ahmad, and A. Algarni, "A latency-aware and fault-tolerant framework for resource scheduling and data management in fog-enabled smart city transportation systems," *Comput., Mater. Continua*, vol. 82, no. 1, pp. 1377–1399, 2025, doi: 10.32604/cmc.2024.057755.

[111] F. Golpayegani, N. Chen, N. Afraz, E. Gyamfi, A. Malekjafarian, D. Schäfer, and C. Krupitzer, "Adaptation in edge computing: A review on design principles and research challenges," *ACM Trans. Auto. Adapt. Syst.*, vol. 19, no. 3, pp. 1–43, Sep. 2024, doi: 10.1145/3664200.

[112] M. Babar, M. S. Khan, U. Habib, B. Shah, F. Ali, and D. Song, "Scalable edge computing for IoT and multimedia applications using machine learning," *Hum.-Centric Comput. Inf. Sci.*, vol. 11, p. 41, May 2021, doi: 10.22967/HCIS.2021.11.041.



TAMARA ZHUKABAYEVA (Member, IEEE) received the Ph.D. degree in computer science, computer engineering, and control from Kazakh National Technical University. She is currently a Full Professor with the Department of Information Systems, L. N. Gumilyov Eurasian National University. She has published more than 110 scientific articles, books, and monographs, has an innovative patent and copyright certificates for intellectual property rights. She has extensive experience as a Researcher and an Educator, contributing to various national and international projects. Her research interests include artificial intelligence (AI), information security, computer vision, and the Internet of Things (IoT). She has received numerous awards, including the Best University Teacher Award and the "Bolashak" International Scholarship.



ZULFIQAR AHMAD received the M.Sc. degree in CS from Hazara University, Mansehra, Pakistan, in 2012, the M.S. degree in CS from COMSATS University, Abbottabad, Pakistan, in 2016, and the Ph.D. degree in computer science (CS) from Hazara University, in 2022. He is the author of more than 35 research publications in the fields of edge computing, cloud computing, high-performance computing, and scientific workflows execution and management. His research interests include scientific workflows management in cloud computing, the Internet of Things, fog computing, edge computing, cybersecurity, healthcare, and wireless sensor networks (WSNs).



AIGERIM YERIMBETOVA (Member, IEEE) received the Ph.D. degree in technical sciences. She is currently an Associate Professor with the Institute of Information and Computational Technologies, Committee of Science, Ministry of Science and Higher Education of the Republic of Kazakhstan, Almaty. Her current research focuses on the development of linguistic algorithms for semantic search and information retrieval, and the application of mathematical and logical methods in computational linguistics. She is the author of more than 100 articles on computational linguistics and information systems development, including one textbook, two monographs, 41 articles indexed in Scopus, and 30 articles in journals recognized by the Committee for Quality Assurance in the Field of Education and Science (CQAES). She was a recipient of the prestigious 500 Scientists scholarship program and is conducting research in data science at the University of Illinois Urbana–Champaign.



MADINA SAMBETBAYEVA received the Ph.D. degree in information systems from L. N. Gumilyov Eurasian National University, Astana, Kazakhstan. She is a Project Manager of the Grant Financing Project "Development of the Fundamentals of the Technology for Creating Models and Software for Building Distributed Information Systems to Support Scientific and Educational Activities (taking into account the morphology of the Kazakh language)." She is also involved in interstate projects, working together with scientists from Novosibirsk State University and the Institute of Computational Technologies SB RAS. She has more than 70 publications, including one educational-methodical manual, 22 articles in Scopus, 25 articles in the journals of the list of the Higher Attestation Commission of the Republic of Kazakhstan and the Russian Federation, and 15 certificates of official registration of computer programs used in pedagogical and scientific practice. She has an H-index of five. She is working on research and development of the fundamentals of the technology for creating models and software for building distributed information systems to support scientific and educational activities, taking into account the morphology of the Kazakh language, identified in the form of electronic libraries that are consistent with international standards and trends in the development of national and international information infrastructure. Her research interests include information systems, information retrieval systems, information retrieval processes, decision support systems, and linguistic support of information systems.



DUMAN TELMAN has an academic journey and professional experience that have equipped him with a strong foundation in natural language processing, data science, and machine learning. His research interests include misinformation detection on social media, sentiment analysis, and the development of computational methods for analyzing multilingual and low-resource languages. He is particularly focused on advancing NLP techniques for practical applications in information security, data analytics, and social network analysis.



ABDYGALYM BAYANGALI received the master's degree in technical sciences. He is currently pursuing the Ph.D. degree with the L. N. Gumilyov Eurasian National University. He is a Senior Lecturer with the Department of Information Technology and Engineering, Astana International University, and a Junior Researcher with the Institute of Information and Computing Technology. His research interests include machine learning, natural language processing, decision support systems, and scientometrics.



ELMIRA DAIYRBAYEVA received the Ph.D. degree in research program. Previously, she was a Senior Lecturer with Satbayev University, Almaty, Kazakhstan. Her research was carried out at the Institute of Information and Computational Technologies of the Committee of Science, Ministry of Science and Higher Education of the Republic of Kazakhstan (CS MSHE RK), based in Almaty. She is the author of more than 25 publications, including one teaching manual, 12 articles indexed in Scopus, and holds eight certificates for the official registration of computer programs used in educational and scientific activities, and has an H-index of two. Her research interests include image processing, hidden data transmission, and steganography.