

Review

Natural Language Processing in Competitive Intelligence: A Systematic Review

Dan Lu¹ , Burin Srisomthawin² , Xiawei Chen^{3,4}  and Zheyun Zheng^{5,*} 

¹ Graduate School, University of the Thai Chamber of Commerce, Bangkok 10400, Thailand

² School of Humanities, University of the Thai Chamber of Commerce, Bangkok 10400, Thailand

³ College & Hospital of Stomatology, Guangxi Medical University, Nanning 530021, China

⁴ Faculty of Liberal Arts, Thammasat University, Khlong Luang 12120, Thailand

⁵ Thai-Chinese International School of Management, University of the Thai Chamber of Commerce, Bangkok 10400, Thailand

* Correspondence: zheng_zhe@utcc.ac.th

Received: 26 March 2026; **Revised:** 15 April 2026; **Accepted:** 23 April 2026; **Published:** 15 October 2026

Abstract: This study examines how natural language processing (NLP) supports the competitive intelligence (CI) process and how its integration affects organizational decision-making. The review also identifies major research trends, practical applications, and key gaps in NLP-enabled CI research. A systematic literature review was conducted using the PRISMA approach. Relevant studies were retrieved from Scopus and Web of Science through structured Boolean search strings combining NLP-, CI-, and decision-making-related terms. After duplicate removal, title and abstract screening, full-text eligibility assessment, and quality appraisal, 94 peer-reviewed studies published between 2022 and 2026 were included in the final synthesis. Although NLP has been widely discussed in business analytics and artificial intelligence research, its specific role in CI remains fragmented across disciplines. This review addresses that gap by consolidating evidence from information systems, management, and computational research and by proposing an integrated framework linking NLP techniques, CI stages, decision effects, and risk pathways. The review shows that NLP supports CI through automated information gathering, content summarization, sentiment analysis, trend detection, weak signal identification, competitive landscape mapping, predictive analytics, and insight communication. However, the organizational value of NLP is not determined by technological adoption alone. It depends on source quality, model interpretability, human oversight, and alignment between analytical outputs and managerial needs. The review also identifies emerging practical applications in competitor monitoring, reputation analysis, market sensing, and risk early-warning. This study contributes a framework that explains both value creation and risk propagation in NLP-enabled CI systems. It also outlines concrete directions for future research, including longitudinal outcome studies, CI-specific benchmarks, human-AI collaboration design, hallucination mitigation, and ethical governance mechanisms.

Keywords: Natural Language Processing; Competitive Intelligence; Text Mining; Sentiment Analysis; Systematic Literature Review; Decision-Making

1. Introduction

In today's hypercompetitive business environment, organizations operate under conditions of rapid technological change, volatile markets, and increasing informational complexity. To remain strategically responsive, they

must continuously monitor their competitive environment and convert external signals into actionable insight [1]. Competitive Intelligence (CI) has therefore emerged as a systematic practice through which organizations collect, analyze, and transform external information into practical knowledge that supports strategic decision-making and helps sustain competitive advantage [2]. Traditionally, the CI process involves identifying intelligence needs, gathering information from multiple sources, analyzing that information, and communicating the resulting insights to organizational decision-makers [3].

As commercial ecosystems have become increasingly digitalized, organizations are now confronted with an unprecedented volume of unstructured textual data generated through online news portals, social media platforms, corporate reports, customer feedback, and professional forums. These sources contain valuable signals about market developments, competitor activities, and technological change, but they also create major challenges for intelligence analysts who must process and interpret large-scale textual corpora efficiently [4]. Conventional manual approaches to data collection and analysis are becoming progressively less effective in dealing with the scale, speed, and complexity of information produced within contemporary digital ecosystems [5]. Without more advanced analytical capabilities, organizations risk overlooking important indicators of market change, competitor intentions, and emerging opportunities [6].

Natural Language Processing (NLP), a specialized branch of artificial intelligence concerned with enabling machines to understand and process human language, has become an important technological response to these challenges [7]. NLP enables the automated processing of large text collections, the extraction of relevant information, the identification of patterns, and the transformation of unstructured content into structured knowledge that can support organizational decision-making [8]. Through techniques such as sentiment analysis, text classification, named entity recognition, and automated summarization, NLP allows organizations to uncover insights that would otherwise remain hidden within massive volumes of text data [9].

Although the use of artificial intelligence in business analytics and intelligence systems has expanded considerably, research specifically examining the role of NLP in CI remains fragmented across disciplines such as information science, computer science, and strategic management. Prior studies have explored the broader application of artificial intelligence and machine learning in decision-support systems, but the specific contribution of NLP techniques to CI processes has not yet been examined in a sufficiently systematic manner [10].

To address this research gap, this paper presents a systematic literature review (SLR) based on PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). To capture recent developments in NLP and its use in competitive intelligence systems, the review focuses on peer-reviewed studies published between 2022 and 2026. The following research questions guide the study:

- 1) RQ1: What role does Natural Language Processing play in the competitive intelligence process?
- 2) RQ2: How does the application of NLP in the CI process affect organizational decision-making?
- 3) RQ3: What are the major NLP techniques and tools used in CI applications?

Through a systematic review of the available literature, this study aims to provide a broad perspective on the application of NLP technologies in the competitive intelligence process, identify research trends and gaps, and propose avenues for future research. The findings contribute to both scholarship and practice by demonstrating how NLP can strengthen organizational intelligence capabilities in data-rich environments.

The remainder of this manuscript is structured as follows. Section 2 outlines the research methodology, including the search strategy, inclusion criteria, and analytical methods. Section 3 presents the review findings organized around the research questions and discusses their implications, the study limitations, and future research directions. Section 4 concludes the study by summarizing its main contributions.

2. Methodology

2.1. Systematic Literature Review Approach

This study used a systematic literature review (SLR) to provide a rigorous and transparent analysis of the available literature on the use of Natural Language Processing (NLP) in Competitive Intelligence (CI). The SLR method enhances methodological transparency and minimizes selection bias by enabling the systematic identification, analysis, and synthesis of relevant academic studies. Systematic reviews are widely used in evidence-based research

to consolidate fragmented knowledge, trace research trends, and identify gaps in the literature [11]. In competitive intelligence research, such structured approaches are particularly important for advancing knowledge and strengthening the practical relevance of intelligence frameworks in organizational contexts [12]. Therefore, the SLR methodology provides a robust basis for systematically examining scholarship on NLP-enabled competitive intelligence and generating insights relevant to both theory and practice.

2.2. Formulation of Research Questions

The research questions were developed systematically to ensure clarity and alignment with the objectives of the review. Following established guidance for formulating questions in evidence-based studies, the process considered the scope of the investigation, the technological intervention, and the anticipated outcomes in organizational settings [13]. This study examines how Natural Language Processing (NLP) is used in competitive intelligence activities, with particular attention to how NLP tools support information extraction, analysis, and strategic decision-making. Accordingly, three research questions were formulated to investigate the role of NLP in the competitive intelligence process, its impact on organizational decision-making, and the principal NLP methods and tools used in CI applications, as shown in **Table 1**.

Table 1. Research Questions, Focus Areas, and Expected Contributions.

Research Question	Focus Area	Expected Contribution
RQ1: How does NLP contribute to the CI process?	Functional contributions	Identify specific ways NLP enhances CI activities
RQ2: What is the impact of NLP integration on decision-making?	Organizational outcomes	Understand how NLP-enabled CI influences strategic choices
RQ3: What are the key NLP techniques employed?	Technical dimensions	Catalog and analyze NLP methods used in CI applications

2.2.1. Database Selection

To achieve comprehensive, high-quality coverage of the literature, two major academic databases—Scopus and Web of Science—were selected for this review. These databases are widely recognized for indexing peer-reviewed publications in fields such as computer science, information systems, and management. Their broad coverage and rigorous indexing standards make them suitable for identifying studies relevant to Natural Language Processing (NLP) and Competitive Intelligence (CI).

2.2.2. Search String Development

The search strings were developed through a structured process that involved identifying key concepts from the research questions, compiling relevant synonyms, and refining the queries using Boolean operators (AND, OR) to ensure accurate retrieval of relevant studies. The literature search in Scopus and Web of Science was conducted on 26 March 2026, and all retrieved records were exported on the same day for duplicate removal and screening. The final database queries used in this study are presented in **Table 2**.

Table 2. Search Strings for Database Queries.

Database	Search String
Scopus	TITLE-ABS-KEY(("natural language processing" OR "NLP" OR "text mining" OR "sentiment analysis" OR "topic modeling" OR "text analytics") AND ("competitive intelligence" OR "strategic intelligence" OR "business intelligence" OR "environmental scanning" OR "competitor analysis") AND ("decision making" OR "strategic decision" OR "decision support"))
Web of Science	TS = (("natural language processing" OR "NLP" OR "text mining" OR "sentiment analysis") AND ("competitive intelligence" OR "strategic intelligence" OR "business intelligence") AND ("decision making" OR "strategic decision"))

2.2.3. Time Frame and Language

The literature search was limited to articles published between 2022 and 2026 to capture recent developments in Natural Language Processing (NLP) technologies and their applications in Competitive Intelligence (CI). Only English-language publications were reviewed to ensure consistency and accessibility. The search was also restricted to peer-reviewed journal articles and conference proceedings, as these sources are generally regarded as validated, high-quality scholarly contributions.

2.3. Inclusion and Exclusion Criteria

2.3.1. Inclusion Criteria

Studies were included in the review if they satisfied the criteria presented in **Table 3**.

Table 3. Inclusion Criteria.

Criterion	Description	Rationale
Relevance	Studies addressing NLP applications in competitive intelligence or closely related domains	Ensures alignment with research objectives
Empirical foundation	Studies providing empirical evidence, case studies, or systematic analysis	Ensures methodological rigor and practical relevance
Publication type	Peer-reviewed journal articles or conference proceedings	Ensures scholarly quality
Language	English	Ensures accessibility and consistency
Time frame	2022–2026	Captures recent technological developments in NLP

2.3.2. Exclusion Criteria

Studies that did not meet the inclusion criteria were excluded based on the conditions summarized in **Table 4**.

Table 4. Exclusion Criteria.

Criterion	Description	Rationale
Document type	Books, book chapters, dissertations, editorials, and opinion papers	Focus on peer-reviewed research
Length	Articles shorter than four pages	Ensures sufficient analytical depth
Relevance	Studies mentioning NLP or CI only superficially (e.g., in keywords or abstracts)	Ensures substantive contribution
Availability	Full text not available electronically	Required for detailed analysis
Duplication	Duplicate records retrieved from multiple databases	Prevents redundant evaluation

2.4. Study Selection Process

The study selection process followed the PRISMA framework to ensure transparency and methodological rigor. It comprised four stages: identification, screening, eligibility assessment, and final inclusion.

Phase 1: Identification. The database search retrieved 687 records, comprising 412 records from Scopus and 275 from Web of Science. All records were exported to reference-management software for organization and duplicate detection.

Phase 2: Screening. After 34 duplicate records were removed, 653 unique records remained. Titles and abstracts were screened against the predefined inclusion criteria. A total of 542 records were excluded because they did not substantively address Natural Language Processing, competitive intelligence, or related analytical applications.

Phase 3: Eligibility. The remaining 111 articles underwent full-text review to assess their methodological rigor, relevance to NLP-enabled intelligence processes, and empirical value. Articles were excluded if they lacked sufficient methodological detail, did not substantively examine NLP applications, or failed to meet the inclusion criteria.

Phase 4: Inclusion. Following the eligibility assessment, 94 studies that met all inclusion criteria were included in the final systematic review and qualitative synthesis.

Figure 1 presents the PRISMA-based study selection process used in this review. The initial database search identified 687 records, including 412 from Scopus and 275 from Web of Science. After removing 34 duplicate records, 653 unique studies remained for title and abstract screening. At this stage, 542 records were excluded because they did not address NLP, CI, or related decision-support applications in a substantive way. The remaining 111 studies underwent full-text eligibility assessment. After detailed evaluation, 17 articles were excluded because they lacked sufficient methodological detail, did not focus substantively on NLP or CI, or did not meet the inclusion criteria. The final sample consisted of 94 studies included in the qualitative synthesis.

2.5. Quality Assessment

To ensure the reliability and methodological rigor of the included studies, a quality-assessment process was conducted using criteria adapted from established systematic-review guidance [14]. Each article was evaluated on several dimensions, including the clarity of its research objectives, methodological rigor, data quality, validity of

findings, and relevance to the research questions, as shown in **Table 5**. These criteria helped ensure that the studies included in the analysis were credible and suitable for examining NLP applications in Competitive Intelligence.

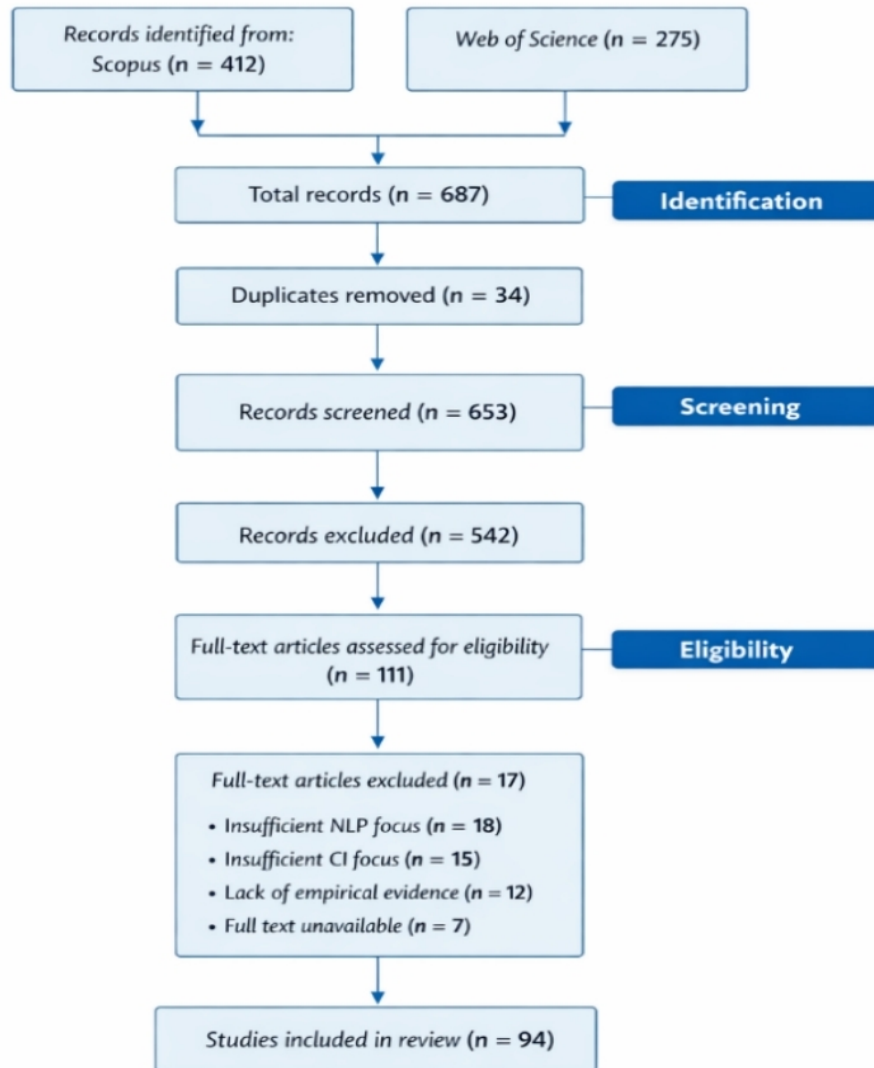


Figure 1. PRISMA Flow Diagram.

Table 5. Quality Assessment Criteria.

Quality Criterion	Assessment Questions	Scoring (0–2)
Research clarity	Are the research aims clearly stated?	0 = No, 1 = Partially, 2 = Yes
Methodological rigor	Is the research methodology appropriate and well described?	0 = No, 1 = Partially, 2 = Yes
Data quality	Are the data collection and analysis methods sound?	0 = No, 1 = Partially, 2 = Yes
Validity of findings	Are the findings clearly presented and supported by evidence?	0 = No, 1 = Partially, 2 = Yes
Relevance to research questions	Does the study address NLP in the competitive intelligence context?	0 = No, 1 = Partially, 2 = Yes

The quality score for each study was calculated by averaging the independent assessments of two reviewers. Studies scoring below 5 out of 10 were classified as low quality and excluded. All 94 included studies met the quality threshold, with scores indicating sufficient methodological rigor and relevance for the final analysis. To enhance review reliability, two reviewers independently conducted the screening and quality-appraisal procedures. Titles and abstracts were first assessed against the inclusion and exclusion criteria, followed by full-text evaluation of potentially relevant studies. Disagreements concerning study inclusion, exclusion, or quality scores were resolved

through discussion until consensus was reached. This dual-review procedure reduced individual selection bias and improved the consistency of study appraisal.

2.6. Data Extraction and Synthesis

To extract relevant information systematically, a standardized data-extraction form was developed for all included studies. This approach ensured consistency and facilitated comparisons across the reviewed literature. The extracted data included bibliographic information, research characteristics, NLP methods, competitive intelligence applications, key findings, effects on decision-making, and reported limitations. **Table 6** summarizes the data-extraction categories.

Table 6. Data Extraction Categories.

Category	Data Extracted
Bibliographic information	Authors, publication year, article title, journal or conference source
Research characteristics	Research objectives, methodology, study context, and sample characteristics
NLP techniques	Specific NLP methods used (e.g., sentiment analysis, topic modeling, text mining)
CI applications	Ways in which NLP techniques were applied within competitive intelligence tasks
Key findings	Major results related to NLP contributions
Decision-making impact	Effects on organizational or strategic decision-making
Limitations	Study limitations identified by the authors

Data Synthesis Approach

A thematic synthesis method was used to analyze the extracted data and identify the principal patterns and insights across the selected studies. The process involved four steps: open coding of relevant findings; grouping similar codes into broader analytical themes; identifying overarching themes related to NLP applications and decision-making effects; and refining those themes through discussion and comparison across studies. This systematic synthesis enabled the identification of major trends, relationships, and research gaps concerning the use of NLP in competitive intelligence.

3. Results and Discussion

3.1. Descriptive Analysis of Included Studies

3.1.1. Publication Trends

The 94 papers included in this systematic review were published between 2022 and 2026, reflecting the rapid growth of research on NLP and Competitive Intelligence. **Table 7** presents the number of publications by year. Scholarly output increased notably after 2023, alongside advances in transformer-based language models, large language models, and automated text-analytics methods that have expanded the potential use of NLP in business intelligence.

Table 7. Distribution of Publications by Year.

Year	Number of Publications	Percentage
2022	12	12.8%
2023	21	22.3%
2024	25	26.6%
2025	23	24.5%
2026	13	13.8%
Total	94	100%

Figure 2 shows the yearly distribution of the 94 publications included in the study. The highest number of studies were published in 2024 (26.6%), followed by 2025 (24.5%) and 2023 (22.3%), indicating increased research activity during these years. Publications in 2022 (12.8%) and 2026 (13.8%) represent comparatively smaller proportions.

The increase in research after 2023 indicates growing academic and practical interest in using NLP techniques to support competitive intelligence through market monitoring, competitor analysis, and strategic decision support.

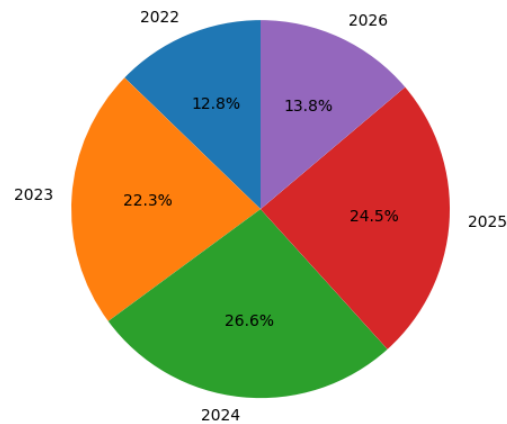


Figure 2. Distribution of Publications by Year (2022–2026).

3.1.2. Publication Sources

The selected studies were published across a wide range of peer-reviewed journals and conference proceedings covering information systems, artificial intelligence, management science, and business analytics, as shown in **Table 8**. This diversity of publication outlets demonstrates the interdisciplinary nature of NLP-enabled competitive intelligence. Journals focusing on artificial intelligence, decision-support systems, information management, and business analytics were major publication sources, while several studies also appeared in international conference proceedings on emerging technologies and data-driven decision-making.

Table 8. Distribution of Publication Sources.

Source Category	Number of Articles
Artificial Intelligence and NLP journals	28
Information systems and decision support journals	21
Business and management journals	18
Interdisciplinary technology journals	14
International conference proceedings	13
Total	94

Such a distribution implies that the study of the applications of NLP to competitive intelligence is not limited to one field but incorporates the viewpoints of computer science, strategic management, and data analytics.

3.1.3. Research Methodologies

The reviewed studies used a wide range of methodologies, reflecting both the technical complexity of Natural Language Processing (NLP) development and its practical application in organizations. Empirical quantitative methods were the most common, particularly in studies that developed or tested NLP models for text analytics and intelligence extraction, as shown in **Table 9**. Other studies used case-study and mixed-method designs to examine how NLP-based systems can be applied in competitive intelligence settings.

Table 9. Research Methodologies Employed.

Methodology	Number of Studies	Percentage
Empirical quantitative (algorithms, experiments)	39	41.5%
Case studies	18	19.1%
Conceptual frameworks	14	14.9%
Survey-based studies	9	9.6%
Mixed methods	8	8.5%
Design science research	6	6.4%
Total	94	100%

The high proportion of empirical, algorithm-focused studies indicates a strong technological orientation in

NLP scholarship. However, the inclusion of case studies and conceptual research also reflects growing efforts to translate NLP capabilities into strategic intelligence and organizational decision-support systems.

3.2. RQ1: Contributions of NLP to the Competitive Intelligence Process

Natural Language Processing contributes to the competitive intelligence (CI) process by enabling the analysis of large volumes of textual data across multiple stages of the intelligence cycle. NLP technologies transform unstructured data into structured knowledge, thereby supporting strategic awareness, environmental scanning, and decision-making and making NLP an increasingly important component of modern CI systems [15].

3.2.1. Enhanced Information Gathering and Source Coverage

NLP enhances information gathering and source coverage by automating the monitoring and retrieval of intelligence from diverse textual sources, including news outlets, social media, corporate reports, and online forums. NLP-based information-extraction systems identify key entities, relationships, and events in large document collections, enabling raw textual information to be transformed efficiently into actionable intelligence [16]. Advances in language-processing technologies also improve automated systems' ability to interpret complex textual data across different domains and contexts. Moreover, NLP-assisted workflows can integrate data from multiple sources and automate intelligence-collection tasks, thereby improving the efficiency and scalability of CI activities [17].

3.2.2. Automated Content Processing and Summarization

Wang et al. [18] showed that NLP can support CI through automated content processing and summarization, allowing organizations to analyze large volumes of textual data efficiently. NLP-based models can automatically identify useful web-based information and transform unstructured data into structured intelligence for further analysis [19]. Visual-analytics solutions simplify the interpretation of complex textual data and facilitate the identification of meaningful insights within large document collections [20]. Entity-centric knowledge-extraction systems can reveal technological advances and emerging trends in the scientific literature, while document-parsing methods can extract structured information from complex documents more effectively [21]. In addition, AI-powered NLP systems can generate visual representations and contextual insights from textual information [22]. Automatic text-summarization algorithms reduce information overload by producing concise summaries that retain essential information, while context-aware summarization systems improve the relevance and coherence of automatically generated intelligence reports.

3.2.3. Sentiment Analysis and Public Perception Monitoring

Ovalle [23] identified sentiment analysis as one of the most widely used NLP techniques in CI because it enables organizations to monitor consumer perceptions, opinions, and stakeholder attitudes toward products, brands, and competitors. NLP-based sentiment-monitoring systems can process large volumes of text from social media and online discussions to detect emerging shifts in public sentiment [24]. Deep-learning-based sentiment models further improve the accuracy of opinion analysis across complex social-media datasets [25]. Sentiment analysis has also been used to evaluate public attitudes toward technological projects and social change; for example, text-mining and sentiment-analysis techniques can identify changing perceptions of smart-city initiatives and other digital innovations.

Sentiment analysis also supports brand monitoring and reputation management by tracking changes in stakeholder attitudes and identifying potential reputational risks across digital communication channels [26]. Recent advances in machine learning and deep learning have substantially improved the scalability and accuracy of sentiment-classification systems used in CI applications [27,28]. Nayak [29] showed that integrating sentiment analysis with business-intelligence systems enables organizations to combine qualitative and quantitative performance measures, thereby strengthening strategic decision-making [30,31]. Overall, sentiment analysis enhances organizational situational awareness by enabling the continuous monitoring of public perceptions. However, accurately interpreting context-dependent language, sarcasm, and domain-specific expressions remains challenging.

Trend Detection and Pattern Recognition: Natural Language Processing is highly valuable for competitive intelligence because it helps identify trends and hidden patterns in large volumes of textual data, including news articles, social-media posts, and corporate communications. NLP-based analytics can automatically detect semantic

patterns and changes over time, enabling organizations to monitor market forces, technological developments, and competitor activities more efficiently. Shukla et al. [32] showed that AI-driven deep-learning methods can enhance market analysis by detecting patterns associated with industry change, while Lobanova et al. [33] demonstrated that NLP-based trend-detection models can support the identification of technological and market shifts.

Several studies also emphasize the importance of social-media analytics for identifying behavioral and competitive trends. Carvalho et al. [34] found that NLP-based analysis of social-media data can reveal patterns in consumer behavior and competitor strategies, while Novikova [35] emphasized that integrating NLP into organizational systems supports knowledge processing and organizational flexibility. Similarly, Hassani and Mosconi [36] reported that social-media analytics strengthens competitive intelligence, particularly in small and medium-sized enterprises. NLP can also be used to monitor competitor activity: Yuan et al. [37] proposed a text-mining method for identifying competitor events in social-media content, and Kim et al. [38] demonstrated that firm-generated content can be analyzed to identify communication strategies and engagement patterns. In addition, Ju [39] proposed a social-media intelligence system that uses NLP-based topic detection to monitor brand conversations and predict engagement patterns. Collectively, these studies show that NLP-based trend detection can strengthen CI by helping organizations identify emerging opportunities and anticipate competitive threats.

Weak Signal Detection: Weak-signal detection is a critical capability in the competitive intelligence (CI) process because it helps organizations identify early indications of technological, market, or competitive change. NLP strengthens this capability by enabling the automated analysis of large volumes of unstructured text from sources such as scholarly publications, industry reports, and social media. Zhao et al. [40] noted that weak-signal identification supports corporate foresight by detecting emerging patterns within information streams. Similarly, Ebadi et al. [41] introduced an NLP-based text-mining system specifically designed to identify weak signals in large textual datasets, while Albarrak et al. [42] showed that deep-learning-based text analysis can identify emerging technological trajectories through the analysis of scientific publications. Further evidence is provided by Griol Barres and Lopez-Pujalte et al. [43,44], who reported that NLP models can detect early indicators of cyberthreats in Industry 4.0 settings and that AI-based monitoring tools can identify anomalies that may signal future disruptions.

Combining NLP with big-data analytics and machine learning further improves weak-signal detection. Sharma et al. [45] emphasized that NLP and big-data platforms can process large, heterogeneous datasets, while deep NLP models can extract early indicators of economic behavior from unstructured financial data. Related research on big-data ecosystems indicates that NLP can process complex information streams from IoT systems, social media, and digital environments to identify emerging signals [46,47]. Machine-learning-based predictive analytics further strengthens forecasting by detecting latent patterns in large datasets [48]. Practical examples include the NLP-based risk-monitoring system proposed by Chen et al. [49] and social-media sentiment-analysis systems used to identify emerging economic trends. Together, these studies demonstrate that NLP can substantially improve weak-signal detection in CI; however, data noise, contextual interpretation, and model transparency remain important research challenges [50].

Natural Language Processing also contributes to competitive landscape mapping by enabling the systematic, automated analysis of large volumes of textual data. Silva and Bacao [51] demonstrated that NLP-based analysis of professionally generated content can identify major competitors, determine their strategic positions, and reveal industry structures. Their findings show that expert discussions and professional reports contain valuable competitive signals that can be uncovered through advanced text-mining techniques. Similarly, Qian et al. [52] introduced a competitive-intelligence acquisition model that extracts user-generated information to examine how customers perceive other organizations, thereby helping firms understand brand positioning and competitors' reputations from the consumer perspective.

Recent studies highlight the growing importance of AI-based analytics and generative AI systems for improving competitive landscape mapping. Mapintel, a system developed by Silva and Bacao [51] that integrates embedding models with visual analytics, supports competitive-intelligence acquisition and strategic mapping. Research on generative AI further indicates that advanced AI systems can automatically derive large-scale knowledge from textual sources, enabling deeper market and competitor analysis [53,54]. Collectively, these approaches demonstrate the potential of NLP and AI-driven analytics to transform competitive landscape mapping by providing scalable, data-driven insights into competitor behavior, market dynamics, and strategic positioning.

Lin et al. [55] reviewed the role of NLP in predictive analytics and argued that it has become integral to

forecasting-oriented competitive-intelligence applications. Recent research indicates that combining NLP with big-data and machine-learning methods can improve forecasting accuracy and strategic decision-making. Adewale et al. [56] reported that integrating big-data processing with machine learning strengthens predictive analytics in management information systems by improving data preprocessing and pattern identification. Similarly, Faheem et al. [57] found that NLP-based predictive models can enhance incident-management systems by processing textual reports and predicting potential operational problems. Additional studies also emphasize the role of advanced analytics in improving the forecasting capabilities of business-intelligence systems.

Raeesi Vanani et al. [58] argued that predictive analytics can help identify digital capabilities associated with organizational performance, while Bharadiya [59] described the broader transformation of business intelligence through machine-based and machine-enhanced predictive models. Yu et al. [60] also proposed a hybrid NLP and deep-learning model for predicting corporate performance, demonstrating that textual data extracted from corporate reports and online sources can substantially improve prediction quality. Together, these studies suggest that NLP-driven predictive analytics can provide organizations with data-based insights that support proactive strategic planning and competitive advantage.

Visualization and Insight Communication: The final stage of the competitive intelligence process also benefits from Natural Language Processing because insights must be presented in forms that decision-makers can readily understand. Modern NLP systems can transform large volumes of textual data into visual representations that are easier to interpret and analyze strategically. Recent studies have examined natural-language interfaces for data visualization, which allow users to query complex datasets and generate visual-analytics outputs automatically. Shen et al. [61] provided a comprehensive overview of interfaces that connect NLP and visualization systems, enabling intuitive exploration of analytical results. Uddin [62] highlighted the value of NLP-driven visualization in real-time analytics and showed how textual-data processing can support interactive dashboards and decision-support systems.

Additional research indicates that integrating NLP with visual analytics substantially improves access to intelligence insights. Islam et al. [63] showed that natural-language interaction with visualization tools can help analysts explore large datasets and identify risk patterns more effectively. Similarly, Kavaz et al. [64] emphasized the growing importance of chatbot-based natural-language interfaces that allow users to create visualizations through conversational queries. With the emergence of large language models (LLMs), visual-analytics systems are gaining greater automation capabilities and can derive interpretable insights from complex textual sources [65]. Perren [66] further showed that AI-based visualization methods can reveal latent relationships in textual data and support strategic intelligence analysis.

3.2.4. Illustrative Case Applications of NLP in CI

The reviewed literature shows that NLP is already being applied in practical CI settings rather than remaining only a conceptual or experimental tool. For example, Yuan et al. [37] demonstrated a text-mining approach for monitoring competitor events through large volumes of firm-generated social-media content, showing how NLP can support event detection and communication tracking. Ju [39] proposed a social-media competitive intelligence framework for identifying brand topics and predicting customer engagement, illustrating how NLP can strengthen market sensing and timely strategic response. Qian et al. [52] further showed that large-scale professional-generated content can be mined to acquire competitive intelligence about competitor positioning and customer perception. In a risk-oriented application, Chen et al. [49] developed an NLP-based early-warning system for cross-border mergers and acquisitions, highlighting the usefulness of NLP for identifying emerging threats in strategic decision environments. Together, these studies demonstrate that NLP-enabled CI is already being used in competitor monitoring, reputation analysis, opportunity detection, and strategic risk surveillance. **Table 10** summarizes the major contributions of NLP across the competitive intelligence process.

This overview illustrates how NLP systems support multiple stages of the competitive intelligence lifecycle, including data collection, sentiment analysis, predictive analytics, and the visualization of findings. In doing so, they help organizations transform unstructured textual data into actionable strategic intelligence.

Table 10. Summary of NLP Contributions to the CI Process.

Contribution Theme	Description	Representative Studies
Enhanced information gathering	Automated collection from diverse textual sources	Adepoju et al. [15]; Somaya et al. [16]; Zhang et al. [17]
Automated content processing	Summarization and categorization of text	Loiko [5]; Taherdoost and Madanchian [8]; Wang et al. [18]
Sentiment analysis	Monitoring public perception and attitudes	Zhang et al. [67]; Zhang et al. [19]; Liu [20]; Khan et al. [21]
Trend detection	Identifying patterns and emerging topics	Tsiu et al. [6]; Sharma et al. [22]; Ovalle [23]; Sheng et al. [24]
Weak signal detection	Early identification of potential changes	Huang et al. [25]; Olujimi and Ade-Ibijola [9]
Competitive landscape mapping	Understanding competitor positioning	Adepoju et al. [15]; Zhang et al. [17]
Predictive analytics	Forecasting market developments	Yue et al. [26]; Mohamed and Bayraktar [27]; Bordoloi and Biswas [28]
Visualization	Communicating insights effectively	Sheng et al. [24]; Nayak [29]

3.3. RQ2: Impact of NLP Integration on Decision-Making

The reviewed literature indicates that integrating NLP into the competitive intelligence (CI) process substantially affects how organizations make strategic decisions. By processing large volumes of textual information from diverse sources, NLP makes decision-relevant intelligence more accessible, timely, and reliable. The reviewed studies identify four major effects of NLP integration on organizational decision-making.

Improved Decision Quality and Timeliness: NLP integration can improve the quality and timeliness of decision-making by accelerating the intelligence process and enabling real-time information processing [68,69]. Rawas et al. [70] showed that NLP-based analytical systems enable organizations to monitor digital information streams and generate insights that support rapid responses to emerging market trends. Abdalrazzaq et al. and Baliga et al. [71,72] highlighted the capacity of contemporary NLP technologies to strengthen decision-support systems by automating the extraction and interpretation of large volumes of textual data. Similarly, Chatterjee et al. [73] concluded that organizations adopting NLP and big-data analytics can achieve better strategic decision outcomes through improved information-processing capacity. NLP also strengthens the knowledge-management practices that support timely managerial decisions. Pan and Tiwari [74] reported that NLP-based knowledge-control systems improve access to and organization of enterprise knowledge resources, thereby accelerating decision-making. Goswami and Rainy [75] further demonstrated that AI-enabled customer-relationship-management systems incorporating NLP analytics provide real-time customer-behavior insights that help organizations respond more quickly to market trends.

Reduction of Information Overload: Another major contribution of NLP integration is the reduction of information overload, a common challenge in competitive intelligence environments characterized by large volumes of online information. Automated filtering, summarization, and semantic-classification methods enable organizations to transform extensive textual datasets into concise, relevant intelligence outputs. These capabilities allow decision-makers to focus on the most important insights rather than spending substantial time processing excessive or irrelevant information. Recent studies of NLP analytics indicate that automated language processing can improve the efficiency of information-management systems and support more structured decision environments [76].

Reducing Uncertainty and Risk: NLP technologies also help organizations manage uncertainty and mitigate strategic risk by identifying emerging signals in textual sources at an early stage. By examining large volumes of news articles, social-media posts, and technical documents, NLP systems can detect potential threats, shifts in stakeholder sentiment, and changing market trends. Rameesa and Veeramanju [77] and Myllynen et al. [78] showed that NLP can strengthen the predictive-analytics and risk-identification capabilities of AI-based decision-support systems in complex environments. Applications in healthcare and intelligent systems further illustrate how language-based analytics can identify patterns associated with operational risks or strategic challenges.

Alignment with Decision-Maker Needs: NLP technologies also enable intelligence outputs to be aligned more closely with decision-makers' requirements [74,75]. Advanced NLP systems can produce personalized summaries, analytical dashboards, and situational insights that are more accessible to managers and more relevant to specific intelligence needs. Studies of NLP-based organizational knowledge systems indicate that tailored intelligence products can increase decision-makers' engagement with analytical information and improve its practical use in strategic planning.

3.4. RQ3: Key NLP Techniques and Tools in CI Applications

The analysis identified several prominent NLP techniques that underpin competitive intelligence (CI) applications. These methods enable organizations to transform large volumes of unstructured text into structured knowledge that supports strategic monitoring, competitive analysis, and decision-making. The literature indicates that text mining, sentiment analysis, topic modeling, named entity recognition, and automatic text summarization are particularly important for extracting intelligence from digital information ecosystems.

3.4.1. Text Mining and Information Extraction

CI systems extensively use text-mining methods to extract structured knowledge from unstructured textual information. These techniques enable organizations to identify entities, relationships, and events relevant to market competition and strategic decision-making [79,80]. Research on NLP-driven data-science methods shows that text-mining algorithms can process large textual datasets and derive meaningful information for analytical tasks and organizational interventions [81]. Studies of academic and domain-specific text mining similarly emphasize the importance of automated information-extraction methods for discovering patterns and knowledge in large document collections.

Recent research also shows that NLP-based text-mining tools play an important role in analyzing industry reports, scientific publications, and business-intelligence data. Thakur and Kumar [82] emphasized that text-mining approaches can efficiently extract domain-specific information from biomedical, technical, and scholarly data, demonstrating their applicability across diverse fields. In addition, information-retrieval and linguistic-preprocessing techniques, such as stemming and tokenization, can substantially improve the precision and effectiveness of text-mining systems used for knowledge discovery [83].

3.4.2. Sentiment Analysis and Opinion Mining

Sentiment analysis is one of the most widely used NLP methods in CI because it enables organizations to examine stakeholders' attitudes, opinions, and emotional responses toward products, services, brands, and competitors. Studies show that sentiment-analysis techniques have been applied to educational, social-media, and commercial datasets to generate useful insights into public perceptions and behavioral trends [84].

Opinion-mining research further shows that analyzing user-generated content, including online reviews and feedback, can reveal valuable information about customer satisfaction and brand reputation. Araujo et al. [85] demonstrated that sentiment-analysis models applied to app reviews can identify customer-feedback patterns relevant to product development and competitive positioning. Bharadwaj and Sarkar [86,87] likewise highlighted the usefulness of sentiment-classification methods for deriving meaningful insights from product reviews and online discussions. Moreover, combining sentiment analysis with visualization and topic-discovery models enables deeper analysis of textual data by identifying both emotional tone and thematic structure [88].

3.4.3. Topic Modeling and Theme Discovery

Topic-modeling methods are widely used to identify latent themes and emerging topics in large text collections, particularly for trend detection and strategic intelligence. Probabilistic algorithms such as Latent Dirichlet Allocation (LDA) can reveal underlying thematic patterns in scientific publications, news articles, and social-media datasets. Yu and Xiang [89], for example, demonstrated that topic modeling can identify technological trends and research directions in artificial-intelligence scholarship by analyzing large academic datasets. Similarly, Mustaqim and Suryono [90] emphasized the value of topic modeling for uncovering hidden themes in user-generated content and other large textual datasets. These methods enable organizations to monitor emerging issues, shifts in research interest, and industry trends that may influence strategic decisions.

3.4.4. Named Entity Recognition (NER)

Named Entity Recognition (NER) is a core NLP technique designed to recognize and classify entities such as organizations, people, products, locations, and dates in written text [91]. This capability is particularly relevant to competitive intelligence applications that require the identification of competitors and the monitoring of market participants. Sharma et al. [92] showed that NER technologies are central to extracting structured information from

textual sources and automatically identifying key entities in large datasets. Baigang and Yi [93] further emphasized that NER systems increasingly use machine-learning and deep-learning methods to improve entity-detection accuracy and contextual understanding. These advances have expanded the application of NER in specialized areas such as aeronautical intelligence and technical information systems. Lagutina et al. [94] also reported that recent developments have improved NER performance in multilingual and complex textual environments.

3.4.5. Automatic Text Summarization

Automatic text summarization condenses lengthy documents while retaining their most important information. In competitive intelligence, summarization tools are particularly useful when decision-makers need to understand large volumes of textual information without reading every document in full. Research on NLP-based information processing emphasizes that summarization systems improve information accessibility by converting extensive textual material into concise intelligence reports [79,80]. These systems are also increasingly integrated into advanced NLP pipelines that combine information extraction, topic modeling, and sentiment analysis. Such integrated systems enable organizations to analyze large volumes of textual data efficiently and produce concise intelligence outputs for strategic decision-making [81,82]. **Table 11** summarizes the principal NLP techniques used in CI applications.

Table 11. Summary of NLP Techniques in CI Applications.

Technique	Description	CI Applications	Representative Studies
Text mining	Extracting structured information from unstructured text	Competitor monitoring, event detection	Malik et al. [79]; Aleqabie et al. [80]; Lan [81]
Sentiment analysis	Classifying opinions and emotions	Brand perception, stakeholder attitudes	Shaik et al. [84]; Araujo et al. [85]; Bharadwaj [86]
Topic modeling	Identifying latent themes	Trend detection, emerging topics	Yu and Xiang [89]; Mustaqim and Suryono [90]
Named entity recognition	Identifying and classifying entities	Competitor identification	Sharma et al. [45]; Jehangir et al. [91]
Text summarization	Generating concise summaries	Information reduction, intelligence briefing	Malik et al. [79]; Thakur and Kumar [82]

3.5. Synthesis of Findings

3.5.1. Convergence of NLP Contributions across the Intelligence Cycle

The synthesis of the 94 reviewed studies shows that Natural Language Processing is not a single-stage intervention in competitive intelligence but a pervasive enabling technology that can support the entire intelligence cycle, from automated information collection and content processing to sentiment monitoring, trend detection, weak-signal identification, competitive landscape mapping, predictive analytics, and the visualization of strategic insights [13,16,17,20,31,39,50]. This cross-stage applicability distinguishes NLP from more limited analytical tools and positions it as an architectural component of contemporary CI systems rather than an auxiliary method. However, the synthesis also reveals a marked asymmetry in the strength of evidence across these stages. Information gathering [15,16] and sentiment analysis are supported by substantial empirical literature and multiple studies across different domains, whereas visualization and predictive analytics remain comparatively underdeveloped in terms of CI-specific validation. Predictive models, in particular, are often evaluated using general business data rather than competitive-intelligence scenarios; consequently, it remains premature to conclude that NLP-based forecasting reliably improves strategic CI decisions at the organizational level [59,60]. This uneven evidence base represents an important structural weakness that is insufficiently acknowledged in the reviewed literature.

3.5.2. Tension between Technical Performance and Organizational Decision Value

A critical synthesis of the findings related to RQ1, RQ2, and RQ3 reveals an unresolved tension between the technical performance of NLP systems and their demonstrated value for organizational decision-making. The analyzed literature reports numerous NLP contributions using computational metrics such as classification accuracy, summarization quality, entity-extraction performance, and topic-coherence scores. However, empirical evidence linking these technical outputs to measurable improvements in strategic decision quality remains scarce and methodologically weak. Studies examining decision-related effects—such as improved timeliness, data-driven decision-making, and risk reduction—typically rely on self-reported organizational perceptions or theoretical arguments rather than

controlled assessments of decision outcomes. This creates a substantial evidentiary gap: causal links between NLP-system performance and organizational competitive advantage are generally inferred rather than demonstrated. Moreover, reported risks such as large-language-model hallucination, limited interpretability in deep-learning sentiment classifiers, and the suppression of weak signals through automated relevance filtering [40,43,44] suggest that the organizational risks of NLP integration have been systematically underestimated relative to its promoted benefits. An integrated account of NLP in competitive intelligence must therefore model not only the functional contributions of NLP techniques but also the pathways through which technical failures can become strategic-intelligence failures.

3.5.3. An Integrated Framework of NLP in Competitive Intelligence

Synthesizing the findings from all three research questions, this study proposes an integrated conceptual framework describing the structural relationships among NLP techniques, CI process stages, organizational decision-making effects, and associated risk dimensions. As shown in **Figure 3**, the framework maps NLP contributions across the intelligence cycle—from data acquisition and analytical processing to insight generation and strategic communication—and links the techniques identified in RQ3 to their corresponding CI functions identified in RQ1. The organizational decision effects identified in RQ2 are modeled as outcomes of the intelligence-output stage, including improved decision quality and timeliness, reduced information overload, lower uncertainty, and closer alignment with decision-makers' needs. Importantly, the framework also incorporates a risk-feedback layer connecting technical constraints—such as hallucination, limited interpretability, poor source quality, and weak-signal suppression—to analytical processing and organizational decision outcomes. This representation reflects the evidence that NLP risks are not confined to technical systems but can propagate into strategic intelligence products and, ultimately, organizational decisions. The framework therefore moves beyond descriptive lists of NLP methods by emphasizing that the strategic value of NLP-enabled CI depends on input-data quality, model interpretability, the design of human-AI collaboration, and the organization's capacity to critically evaluate algorithmically generated intelligence.

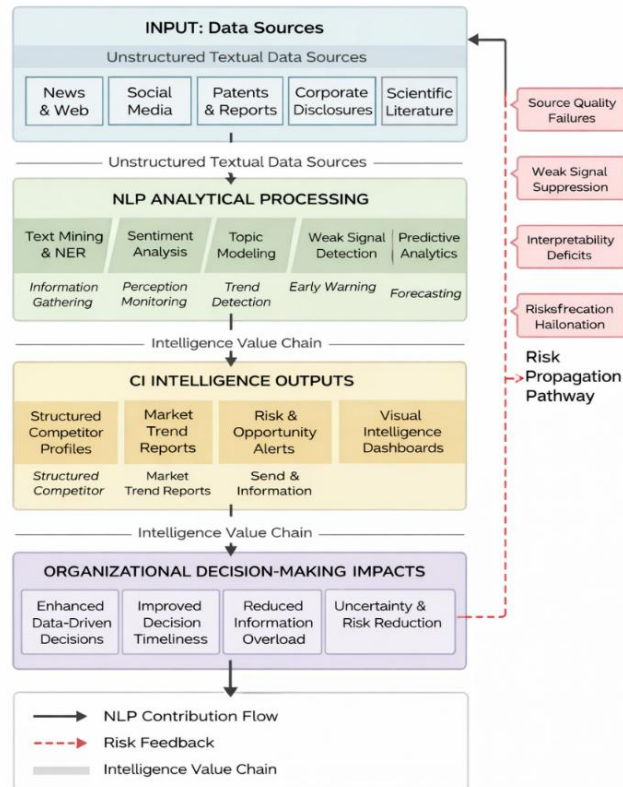


Figure 3. Integrated Framework of NLP in Competitive Intelligence.

Figure 3 presents the integrated framework through which NLP methods transform unstructured textual data into competitive-intelligence products that support organizational decision-making. It also illustrates a risk-propagation pathway showing how data quality, model interpretability, and weak-signal suppression can affect intelligence analysis and subsequent strategic decisions.

3.6. Discussion

3.6.1. NLP as a Structural Enabler of CI Transformation

NLP has transitioned from a supplementary analytical tool to a structural enabler of competitive intelligence, expanding the scale, speed, and analytical depth of CI operations beyond what manual approaches can achieve [13,15,16]. However, a critical examination of the findings reveals an uncomfortable asymmetry: studies demonstrating technical performance gains in text mining [79,80], NER [91,92], and topic modeling substantially outnumber studies demonstrating corresponding improvements in actual organizational intelligence outcomes. This gap indicates that the CI field remains in a transitional phase where NLP capabilities are technically available but organizationally underutilized. Successful CI transformation therefore demands attention not merely to NLP deployment but to the analyst training, process redesign, and output validation conditions that convert technical capability into strategic value [74,75].

3.6.2. The Conditionality of NLP's Organizational Value

The evidence across all three research questions makes clear that NLP's organizational value in CI is conditional, not inherent. The intelligence value chain contains multiple points of potential degradation: source quality failures corrupt downstream outputs regardless of model sophistication; interpretability deficits in deep learning classifiers and LLM summarization systems [53,54] prevent analysts from critically evaluating generated intelligence; and automated filtering risks suppressing the novel weak signals that represent the highest-value CI outputs [43]. The predominantly self-reported decision quality improvements in the RQ2 literature [70,72–74] cannot substitute for controlled evidence of changed strategic behavior. The dominant narrative framing NLP integration as straightforwardly beneficial therefore requires critical qualification—NLP enhances CI only when input data quality, model interpretability, and organizational AI literacy are simultaneously adequate.

3.6.3. Complementarity of Techniques and Integration Costs

The RQ3 findings consistently show that NLP techniques achieve maximum CI value through integration rather than isolated deployment: NER establishes entity extraction foundations [79], sentiment analysis adds attitudinal dimensionality [84,87], topic modeling supplies thematic context [89,90], and summarization produces decision-ready outputs. This complementarity is analytically compelling but organizationally expensive. The reviewed literature systematically underweights the infrastructure, expertise, and maintenance costs of integrated multi-technique pipelines, presenting technically sophisticated systems as straightforwardly feasible without adequate treatment of implementation complexity. Organizations should treat pipeline integration cost as a primary variable in NLP-CI investment decisions, not an afterthought.

3.7. Theoretical Implications

3.7.1. Revising CI Frameworks for the NLP Era

Traditional CI frameworks conceptualize the intelligence cycle as a sequential, periodically executed, human-driven process [1,3]—a model structurally incompatible with NLP systems operating continuously and generating outputs without direct human involvement at extraction and classification stages. Three revisions are required. The intelligence cycle must shift from periodic to continuous [32]; the analyst role must be reconceptualized from primary processor to critical evaluator of algorithmic outputs [65,76]; and CI theory must incorporate algorithmic risk—hallucination, distributional shift, and systematic bias—as first-order theoretical concerns. The reviewed literature's near-total silence on these theoretical revisions represents a significant scholarly gap that limits the field's capacity to guide NLP-CI implementation in organizationally responsible ways.

3.7.2. Contribution to Information Processing Theory

NLP represents a qualitatively new organizational information processing capability, not merely an efficiency improvement, by making previously inaccessible information categories—unstructured text at scale, real-time sentiment streams, cross-lingual content—available for systematic competitive analysis. This expansion shifts the organizational bottleneck from information scarcity to interpretation capacity. The reviewed studies do not adequately theorize this shift: they treat expanded intelligence volume as straightforwardly beneficial without examining the absorptive capacity constraints that determine whether decision-makers can actually incorporate larger intelligence outputs into improved strategic choices [68,72].

3.8. Practical Implications

For CI Practitioners: Technique selection should be driven by intelligence requirements: NER and text mining for entity-centric competitor tracking, sentiment analysis for brand and stakeholder monitoring [27,28], topic modeling for environmental scanning, and summarization for high-volume information management. Critically, practitioners must resist treating NLP outputs as authoritative. Given documented risks of hallucination, domain-specific sentiment errors, and weak signal suppression, human review protocols for high-stakes intelligence outputs are operationally essential. Practitioners should also develop explicit frameworks for assessing whether NLP-generated intelligence actually alters decision behavior—a measurement practice almost entirely absent from the reviewed literature.

For Technology Developers: Domain specificity is the most critical and most consistently neglected design dimension in CI-oriented NLP development. General-purpose models underperform in specialized competitive intelligence contexts, where cross-source information fusion, temporal trend tracking, and strategic entity disambiguation impose requirements that generic benchmarks do not capture. Investment in CI-specific training corpora, explainability mechanisms, real-time processing architectures, and modular pipeline designs represents substantially higher practical value than marginal benchmark improvements on standard NLP tasks.

3.9. Research Gaps and Future Directions

Future research should move beyond reporting technical performance and address a more focused set of implementation-oriented questions. First, under what conditions does NLP-enabled CI improve measurable strategic outcomes such as forecast accuracy, market responsiveness, or competitive positioning? Second, which CI tasks benefit most from domain-specific models compared with general-purpose large language models? Third, how should human analysts and NLP systems divide work in high-stakes CI settings to maximize both speed and judgment quality? Fourth, which validation mechanisms—such as retrieval augmentation, fact-checking pipelines, explainability tools, or human-review checkpoints—most effectively reduce hallucination and weak-signal suppression? Fifth, what ethical and governance frameworks are required for privacy-sensitive CI applications that rely on large-scale textual monitoring? Addressing these questions would help move the field from descriptive enthusiasm toward empirically grounded, implementation-oriented knowledge. **Table 12** summarizes these priority research gaps and recommended directions.

Table 12. Priority Research Gaps and Future Directions.

Research Gap	Critical Issue	Recommended Direction
Organizational decision outcome evidence	Strategic outcomes unmeasured; technical metrics dominate	Longitudinal field studies linking NLP-CI to competitive performance
Domain-specific NLP for CI	General models systematically underperform	CI-specific corpora, benchmarks, and evaluation protocols
Human-AI collaboration in CI	Analyst validation roles undertheorized	Ethnographic and experimental workflow studies
LLM hallucination mitigation	High reliability risk in high-stakes CI contexts	RAG architectures and CI-specific fact-verification pipelines
Ethical and privacy frameworks	Near-absent from reviewed literature	GDPR compliance studies and ethical guidelines for NLP-CI practice

The most urgent priority is longitudinal outcome evidence. The current evidence base cannot establish whether NLP-CI integration generates sustained competitive advantage or merely replicates existing capabilities at lower cost [68,73]. Until this gap is addressed, the field's strategic value claims remain inferential rather than empirically grounded.

3.10. Limitations of the Review

Four limitations warrant explicit acknowledgment. Limiting the search to Scopus and Web of Science, while helping ensure publication quality, may have excluded relevant practitioner research and gray literature. The English-language restriction introduces geographic bias because substantial NLP-CI research is published in Chinese, German, and other languages. The 2022–2026 time frame captures recent advances in NLP but excludes foundational pre-transformer CI studies that could provide important historical context. Finally, although thematic synthesis is appropriate for an interdisciplinary and heterogeneous body of literature, it introduces interpretive subjectivity into the categorization of findings. Future reviews of more methodologically homogeneous studies could partly address this limitation through structured meta-analytic procedures.

4. Conclusion

This systematic review examined 94 peer-reviewed studies published between 2022 and 2026 to evaluate the role of NLP in competitive intelligence across three research questions. The review identified eight major contribution themes spanning the entire CI cycle, critically assessed the conditions under which NLP affects organizational decision-making, and summarized five core techniques together with their domain-specific constraints. It also proposed an integrated framework linking NLP techniques, CI process stages, decision outcomes, and risk-propagation pathways.

4.1. Key Findings

The review yields five principal findings. First, NLP functions as a pervasive architectural component across the CI cycle rather than as a discrete analytical tool. Second, the evidence base is structurally skewed toward technical performance metrics, while evidence concerning actual organizational decision outcomes remains limited and largely inferential. Third, the value of NLP depends on source quality, model transparency, and the design of human–AI collaboration—factors that the literature often underestimates. Fourth, integrated, multi-technique pipelines generally provide greater analytical value than isolated tools, but their implementation costs are insufficiently recognized. Fifth, risks such as LLM hallucination, weak-signal suppression, and limited interpretability are not peripheral technical constraints; they are first-order competitive-intelligence concerns.

4.2. Concluding Remarks

NLP has emerged as an essential competitive-intelligence capability; however, the central conclusion of this review is that technological capability cannot be equated with organizational value. Organizations that treat NLP adoption alone as sufficient to improve CI—without robust validation frameworks, analyst capability development, and ethical governance—risk introducing algorithmic failure modes while weakening the human judgment that previously served as a safeguard. The field therefore needs not only further documentation of technical capabilities but, more urgently, robust longitudinal evidence demonstrating the strategic value of NLP-enabled CI and the organizational conditions under which that value can be realized responsibly.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable. This study is a systematic literature review based exclusively on published sources and did not involve human participants, animals, or identifiable personal data.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data supporting the findings of this study consist of published articles indexed in Scopus and Web of Science. Bibliographic records and study selection details are available from the corresponding author upon reasonable request, subject to database access restrictions.

Acknowledgments

The authors would like to thank the editor and anonymous reviewers for their valuable comments and suggestions, which helped improve the quality of this manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

The authors used an AI-assisted language tool during drafting and revision for language refinement and structural editing. All literature screening, interpretation, verification, and final manuscript decisions were performed by the authors, who take full responsibility for the content of the manuscript.

References

1. Elkaabi, K.; Mamouny, A.; Elmaallam, M. The impact of artificial intelligence tools on the competitive intelligence process: A systematic literature review. In Proceedings of the 2025 International Conference on Circuit, Systems and Communication (ICCSC), Fez, Morocco, 19–20 June 2025; pp. 1–7. [CrossRef]
2. Punukollu, M. NLP applications in mining pharmaceutical text data for competitive intelligence and market insights. *Essex J. AI Ethics Responsible Innov.* **2023**, *3*, 594–635. Available online: <https://www.ejaei.org/index.php/publication/article/view/86>
3. Alafi, K.K.; Ismaeel, B.; Almarshad, M.N.; et al. Analysis of competitive intelligence in retail management in the Jordanian market from the consumer's perspective. *J. Intell. Stud. Bus.* **2024**, *13*, 24–38.
4. Adil, B.; Abdelhadi, F. A big data analytics framework for competitive intelligence systems. *J. Theor. Appl. Inf. Technol.* **2022**, *100*, 149–169.
5. Loiko, A.I. Technology of digital ecosystems. *Vestn. Samar. Gos. Tekh. Univ. Ser. Filos.* **2022**, *4*, 49–56. [CrossRef]
6. Tsiu, S.V.; Ngoben, M.; Mathabela, L.; et al. Applications and competitive advantages of data mining and business intelligence in SMEs performance: A systematic review. *Businesses* **2025**, *5*, 22. [CrossRef]
7. Kalyampudi, P.L. Natural language processing for business intelligence and market analysis. In *Artificial Intelligence and Machine Learning in Management Science: Emerging Research and Applications*; NC Publishers: Karachi, Pakistan, 2025; p. 97.
8. Taherdoost, H.; Madanchian, M. Artificial intelligence and sentiment analysis: A review in competitive research. *Computers* **2023**, *12*, 37. [CrossRef]
9. Olujimi, P.A.; Ade-Ibijola, A. NLP techniques for automating responses to customer queries: A systematic review. *Discov. Artif. Intell.* **2023**, *3*, 20. [CrossRef]
10. Wiggins, W.F.; Tejani, A.S. On the opportunities and risks of foundation models for natural language processing in radiology. *Radiol. Artif. Intell.* **2022**, *4*, e220119.
11. Hosseini, M.-S.; Jahanshahloo, F.; Akbarzadeh, M.A.; et al. Formulating research questions for evidence-based studies. *J. Med. Surg. Public Health* **2024**, *2*, 100046. [CrossRef]
12. Madureira, L.; Popovič, A.; Castelli, M. Competitive intelligence empirical validation and application: Foundations for knowledge advancement and relevance to practice. *J. Inf. Sci.* **2023**. [CrossRef]
13. Tyagi, N.; Bhushan, B. Demystifying the role of natural language processing (NLP) in smart city applications: Background, motivation, recent advances, and future research directions. *Wireless Pers. Commun.* **2023**, *130*, 857–908. [CrossRef]
14. Vysotska, V. Modern state and prospects of information technologies development for natural language content processing. In Proceedings of the 8th International Conference on Computational Linguistics and Intelligent Systems, Lviv, Ukraine, 12–13 April 2024; pp. 198–234.
15. Adepoju, A.H.; Austin-Gabriel, B.; Eweje, A.; et al. Framework for automating multi-team workflows to maximize operational efficiency and minimize redundant data handling. *IRE J.* **2022**, *5*, 663–664.

16. Somaya, J.; Hasna, E.E.; Abdelwahed, H.E.; et al. Automating information extraction from textual documents using NLP techniques: A review of integration in document management systems. In Proceedings of the 4th International Conference on Advances in Communication Technology and Computer Engineering, Marrakech, Morocco, 29–30 November 2024; pp. 507–518. [CrossRef]
17. Zhang, Z.; Qin, W.; Xu, H. A visual webpage information extraction framework for competitive intelligence system. *Scalable Comput. Pract. Exper.* **2024**, *25*, 4138–4152. [CrossRef]
18. Wang, B.; Ma, K.; Wu, L.; et al. Visual analytics and information extraction of geological content for text-based mineral exploration reports. *Ore Geol. Rev.* **2022**, *144*, 104818. [CrossRef]
19. Zhang, Q.; Wang, B.; Huang, V.S.; et al. Document parsing unveiled: Techniques, challenges, and prospects for structured information extraction. *arXiv preprint* **2024**, *arXiv:2410.21169*. [CrossRef]
20. Liu, D. Design information extraction and visual representation based on artificial intelligence natural language processing techniques. In Proceedings of the 2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 12–14 May 2023; pp. 154–158. [CrossRef]
21. Khan, B.; Shah, Z.A.; Usman, M.; et al. Exploring the landscape of automatic text summarization: A comprehensive survey. *IEEE Access* **2023**, *11*, 109819–109840. [CrossRef]
22. Sharma, N.; Singh, S.; Kapoor, M. Context-based review on auto text summarization techniques using natural language processing for performing evaluation and comparative analysis of ATS techniques. *SN Comput. Sci.* **2026**, *7*, 15. [CrossRef]
23. Ovalle, C. Application for monitoring sentiment analysis. In Proceedings of the 13th Computer Science Online Conference 2024, Zlín, Czech Republic, 25–28 April 2024; p. 340. [CrossRef]
24. Sheng, L.; Wang, Z.; Zhang, L.; et al. Application of sentiment analysis based on deep learning in public opinion monitoring in healthcare networks. *Int. J. High Speed Electron. Syst.* **2025**. [CrossRef]
25. Huang, Z.; Tay, E.; Wee, D.; et al. Public perception of the use of digital contact-tracing tools after the COVID-19 lockdown: Sentiment analysis and opinion mining. *JMIR Formative Res.* **2022**, *6*, e33314. Available online: <https://formative.jmir.org/2022/3/e33314/>
26. Yue, A.; Mao, C.; Chen, L.; et al. Detecting changes in perceptions towards smart city initiatives on Chinese social media: A text mining and sentiment analysis approach. *Buildings* **2022**, *12*, 1182. [CrossRef]
27. Mohamed, K.; Bayraktar, Ü.A. Analyzing the role of sentiment analysis in public relations: Brand monitoring and crisis management. *Int. J. Humanities Soc. Sci.* **2022**, *9*, 116–126.
28. Bordoloi, M.; Biswas, S.K. Sentiment analysis: A survey on design framework, applications, and future scopes. *Artif. Intell. Rev.* **2023**, *56*, 12505–12560. [CrossRef]
29. Nayak, S. Leveraging natural language processing and machine learning for sentiment analysis in fintech: Enhancing customer insights and decision-making. *Int. J. Appl. Eng. Technol.* **2022**, *4*, 242–260.
30. Revathy, G.; Alghamdi, S.A.; Alahmari, S.M.; et al. Sentiment analysis using machine learning: Progress in machine intelligence for data science. *Sustainable Energy Technol. Assess.* **2022**, *53*, 102557. [CrossRef]
31. Oloyede, P. Integrating NLP techniques to analyze transactional data for business insights. *SSRN Electron. J.* **2025**. [CrossRef]
32. Shukla, S.; Singh, J.; Nassa, V.K.; et al. Artificial intelligence-driven deep learning for competitive intelligence to enhance market analysis and strategic positioning. In Proceedings of the 2024 4th Asian Conference on Innovation in Technology (ASIANCON), Pimari Chinchwad, India, 23–25 August 2024; pp. 1–5. [CrossRef]
33. Lobanova, P.A.; Kuzminov, I.F.; Karatetskaia, E.Y.; et al. Trend detection using NLP as a mechanism of decision support. *Sci. Tech. Inf. Process.* **2023**, *50*, 440–448. [CrossRef]
34. Carvalho, D.; Picoto, W.; Busch, P. Organizational experience of social media: Impacts on competitive intelligence. *VINE J. Inf. Knowledge Manage. Syst.* **2022**, *52*, 161–183. [CrossRef]
35. Novikova, E. Influence of NLP on Organizational Design and Development of Dynamic Capabilities. Master's Thesis, Politecnico di Milano, Milan, Italy, 2023. Available online: <https://hdl.handle.net/10589/217941>
36. Hassani, A.; Mosconi, E. Social media analytics, competitive intelligence, and dynamic capabilities in manufacturing SMEs. *Technol. Forecasting Soc. Change* **2022**, *175*, 121416. [CrossRef]
37. Yuan, H.; Deng, W.; Ma, B.; et al. Monitoring events of market competitors: A text mining method for analyzing massive firm-generated social media. *J. Theor. Appl. Electron. Commerce Res.* **2023**, *18*, 908–927. [CrossRef]
38. Kim, W.H.; Park, E.; Kim, S.B. Understanding the role of firm-generated content by hotel segment: The case of Twitter. *Current Issues Tourism* **2023**, *26*, 122–136. [CrossRef]
39. Ju, X. A social media competitive intelligence framework for brand topic identification and customer engagement prediction. *PLoS ONE* **2024**, *19*, e0313191. [CrossRef]
40. Zhao, D.; Tang, Z.; He, D. A systematic literature review of weak signal identification and evolution for corpo-

- rate foresight. *Kybernetes* **2024**, 53, 3160–3188. [CrossRef]
41. Ebadi, A.; Auger, A.; Gauthier, Y. Detecting emerging technologies and their evolution using deep learning and weak signal analysis. *J. Informetr.* **2022**, 16, 101344. [CrossRef]
42. Albarrak, M.; Salonitis, K.; Jagtap, S. Natural language processing (NLP)-based frameworks for cyber threat intelligence and early prediction of cyberattacks in Industry 4.0: A systematic literature review. *Appl. Sci.* **2026**, 16, 619. [CrossRef]
43. Griol Barres, I. Modelling of a System for the Detection of Weak Signals through Text Mining And NLP: Proposal of Improvement by a Quantum Variational Circuit. PhD Thesis, Universitat Politècnica de València, Valencia, Spain, 2022. Available online: <https://riunet.upv.es/server/api/core/bitstreams/68f241b7-e4c7-4697-b100-6f4ce8fe4d29/content>
44. Lopez-Pujalte, C.; Tena-Mateos, M.J.; Muñoz-Cañavate, A. A technology watch/competitive intelligence-based decision-support system optimised with genetic algorithms. *J. Inf. Sci.* **2022**, 51. [CrossRef]
45. Sharma, R.; Agarwal, P.; Arya, A. Natural language processing and big data: A strapping combination. In *New Trends and Applications in Internet of Things (IoT) and Big Data Analytics*; Springer: Cham, Switzerland, 2022; pp. 255–271. Available online: https://link.springer.com/chapter/10.1007/978-3-030-99329-0_16
46. Dash, B. Information Extraction from Unstructured Big Data: A Case Study of Deep Natural Language Processing in Fintech. PhD Thesis, University of the Cumberland, Williamsburg, KY, USA, 2022. Available online: <https://proquest.com/openview/0a7fccfbd798ac514d956893df11810>
47. Taha, K. Big data analytics in IoT, social media, NLP, and information security: Trends, challenges, and applications. *J. Big Data* **2025**, 12, 150. [CrossRef]
48. Rane, N.L.; Paramesha, M.; Choudhary, S.P.; et al. Machine learning and deep learning for big data analytics: A review of methods and applications. *Partners Universal Int. Innov. J.* **2024**, 2, 172–197. [CrossRef]
49. Chen, Y.; Wei, G.; Lee, E.T. Sentiment analysis and risk early-warning system for cross-border M&A based on natural language processing. *Int. J. Comput. Vision Robotics* **2025**, 15, 1–22. [CrossRef]
50. Fishekqiu, N.S. Exploring the impact of social media sentiment on market dynamics and economic growth: A big data and natural language processing perspective. In *Proceedings of the 8th International Scientific Conference on Business and Economics*, Tetovo, North Macedonia, 8–9 May 2025; pp. 343–359. [CrossRef]
51. Silva, D.; Bacao, F. Mapintel: Enhancing competitive intelligence acquisition through embeddings and visual analytics. In *Proceedings of the 21st EPIA Conference on Artificial Intelligence*, Lisbon, Portugal, 31 August–2 September 2022; pp. 599–610. [CrossRef]
52. Qian, Y.; Ling, H.; Meng, X.; et al. Voice of the professional: Acquiring competitive intelligence from large-scale professional generated contents. *J. Bus. Res.* **2024**, 180, 114719. [CrossRef]
53. Wu, J.; Gan, W.; Chen, Z.; et al. AI-generated content (AIGC): A survey. *arXiv preprint* **2023**, *arXiv:2304.06632*. [CrossRef]
54. Zhang, C.; Zhang, C.; Zheng, S.; et al. A complete survey on generative AI (AIGC): Is ChatGPT from GPT-4 to GPT-5 all you need? *arXiv preprint* **2023**, *arXiv:2303.11717*. [CrossRef]
55. Lin, J.; Jiang, X.; Li, Q.; et al. A competitive intelligence acquisition framework for mining user perception from user-generated content. *Appl. Soft Comput.* **2023**, 147, 110764. [CrossRef]
56. Adewale, G.T.; Victor, A.U.; Sylvia, A.E.; et al. Integrating big data and machine learning in management information systems for predictive analytics: A focus on data preprocessing and technological advancements. *World J. Adv. Res. Rev.* **2024**, 24, 774–789. [CrossRef]
57. Faheem, M.; Awais, M.; Iqbal, A.; et al. Enhancing IT incident management with natural language processing and predictive analytics. *Int. J. Sci. Res. Archive* **2025**, 15, 224–237. [CrossRef]
58. Raeesi Vanani, I.; Taghavifard, M.T.; Yalpanian, M.A. Predictive analytics solution for digital capabilities identification towards business performance improvement. *SN Comput. Sci.* **2025**, 6, 85. [CrossRef]
59. Bharadiya, J.P. The role of machine learning in transforming business intelligence. *Int. J. Comput. Artif. Intell.* **2023**, 4, 16–24. [CrossRef]
60. Yu, Q.; Xing, C.; He, Y.; et al. Intelligent information processing for corporate performance prediction: A hybrid natural language processing (NLP) and deep learning approach. *Electronics* **2026**, 15, 443. [CrossRef]
61. Shen, L.; Shen, E.; Luo, Y.; et al. Towards natural language interfaces for data visualization: A survey. *IEEE Trans. Visual. Comput. Graphics* **2023**, 29, 3121–3144. [CrossRef]
62. Uddin, M.K.S. A review of utilizing natural language processing and AI for advanced data visualization in real-time analytics. *Int. J. Manag. Inf. Syst. Data Sci.* **2024**, 1, 34–49. [CrossRef]
63. Islam, M.R.; Razzak, I.; Wang, X.; et al. Natural language interactions enhanced by data visualization to explore insurance claims and manage risk. *Ann. Oper. Res.* **2024**, 339, 1569–1587. [CrossRef]

64. Kavaz, E.; Puig, A.; Rodríguez, I. Chatbot-based natural language interfaces for data visualisation: A scoping review. *Appl. Sci.* **2023**, *13*, 7025. [CrossRef]
65. Agarwal, N.S.; Sonbhadra, S.K. A review on large language models for visual analytics. *arXiv preprint* **2025**, *arXiv:2503.15176*. [CrossRef]
66. Perren, L. Visualising Textual Connections with AI-Driven Insights. Master's Thesis, Lucerne School of Art and Design, Emmenbrücke, Switzerland, 2023. Available online: https://junior-research.ch/images/databa-se-files/23HS_MADI_Perren_Luiz_thesis_final.pdf
67. Zhang, H.; Zhang, C.; Wang, Y. Revealing the technology development of natural language processing: A scientific entity-centric perspective. *Inf. Process. Manage.* **2024**, *61*, 103574. [CrossRef]
68. Ojadi, J.O.; Onukwulu, E.C.; Somtochukwu, C.; et al. Natural language processing for climate change policy analysis and public sentiment prediction: A data-driven approach to sustainable decision-making. *Iconic Res. Eng. J.* **2023**, *7*, 732–751.
69. Cano-Marin, E.; Sánchez-Alonso, S.; Mora-Cantalops, M. Unleashing competitive intelligence: News mining analysis on technology trends and digital health driving healthcare innovation. *IEEE Trans. Eng. Manage.* **2023**, *71*, 12311–12325. [CrossRef]
70. Rawas, S.; Tafran, C.; AlSaeed, D.; et al. Transforming healthcare: AI – NLP fusion framework for precision decision-making and personalized care optimization in the era of IoMT. *Comput. Mater. Contin.* **2024**, *81*, 4575. [CrossRef]
71. Abdalrazzaq, A.S.; Meftin, N.K.; Al Fawadi, D.M.; et al. Optimizing robotic interaction and decision making using advanced NLP and AI technologies. In Proceedings of the 2024 36th Conference of Open Innovations Association (FRUCT), Lappeenranta, Finland, 30 October 2024–1 November 2024; pp. 692–701. [CrossRef]
72. Baliga, S.; Sadhale, A.; Upadhyaya, D. Future prospects of natural language processing. In *High-Performance Automation Methods for Computational Intelligent Systems*; CRC Press: Boca Raton, FL, USA, 2025; pp. 252–278. [CrossRef]
73. Chatterjee, S.; Chaudhuri, R.; Mikalef, P. Examining the dimensions of adopting natural language processing and big data analytics applications in firms. *IEEE Trans. Eng. Manage.* **2022**, *71*, 3001–3015. [CrossRef]
74. Pan, S.K.; Tiwari, A. Improving organizational knowledge control with the use of NLP. In Proceedings of the 2025 International Conference on Intelligent Control, Computing and Communications (IC3), Mathura, India, 13–14 February 2025; pp. 985–991. [CrossRef]
75. Goswami, D.; Rainy, T. Mechanisms by which AI-enabled CRM systems influence customer retention and overall business performance: A systematic literature review of empirical findings. *ASRC Procedia: Global Perspect. Sci. Scholarship* **2025**, *1*. [CrossRef]
76. Rzig, D.E.; Hassan, F.; Bansal, C.; et al. Characterizing the usage of CI tools in ML projects. In Proceedings of the 16th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), Helsinki, Finland, 19–23 September 2022; pp. 69–79. [CrossRef]
77. Rameesa, K.; Veeramanju, K.T. Analysis of software industry: Natural language processing approach. *Focus* **2023**, *13*, 743–752.
78. Myllynen, T.; Kamau, E.; Mustapha, S.D.; et al. Review of advances in AI-powered monitoring and diagnostics for CI/CD pipelines. *Int. J. Multidisciplinary Res. Growth Evaluation* **2024**, *5*, 1119–1130.
79. Malik, P.; Mittal, V.; Nautiyal, L.; et al. NLP techniques, tools, and algorithms for data science. In *Artificial Intelligence for Signal Processing and Wireless Communication*; De Gruyter Brill: Berlin, Germany, 2022; p. 123.
80. Aleqabie, H.J.; Sfoq, M.S.; Albeer, R.A. A review of text mining techniques: Trends and applications in various domains. *Iraqi J. Comput. Sci. Math.* **2024**, *5*, 9. [CrossRef]
81. Lan, M. Natural language processing and text mining. In *Artificial Intelligence for Drug Design*; Springer: Singapore, 2026; pp. 189–215.
82. Thakur, K.; Kumar, V. Application of text mining techniques on scholarly research articles: Methods and tools. *New Rev. Acad. Librarianship* **2022**, *28*, 279–302. [CrossRef]
83. Jabbar, A.; Iqbal, S.; Tamimy, M.I.; et al. An analytical analysis of text stemming methodologies in information retrieval and natural language processing systems. *IEEE Access* **2023**, *11*, 133681–133702. [CrossRef]
84. Shaik, T.; Tao, X.; Dann, C.; et al. Sentiment analysis and opinion mining on educational data: A survey. *Nat. Lang. Process. J.* **2023**, *2*, 100003. [CrossRef]
85. Araujo, A.F.; Gôlo, M.P.; Marcacini, R.M. Opinion mining for app reviews: An analysis of textual representation and predictive models. *Automated Software Eng.* **2022**, *29*, 5. [CrossRef]
86. Bharadwaj, L. Sentiment analysis in online product reviews: Mining customer opinions for sentiment classi-

- fication. *Int. J. Multidisciplinary Res.* **2023**, 5.
87. Sarkar, S. A comparative study of different natural language processing techniques for sentiment analysis and opinion mining. *Iconic Res. Eng. J.* **2025**, 8.
88. Chu, K.E.; Keikhosrokiani, P.; Asl, M.P. A topic modeling and sentiment analysis model for detection and visualization of themes in literary texts. *Pertanika J. Sci. Technol.* **2022**, 30, 2535–2561. [[CrossRef](#)]
89. Yu, D.; Xiang, B. Discovering topics and trends in the field of artificial intelligence using LDA topic modeling. *Expert Syst. Appl.* **2023**, 225, 120114. [[CrossRef](#)]
90. Mustaqim, I.Z.; Suryono, R.R. A systematic literature review of topic modeling techniques in user reviews. *J. Inf. Syst. Eng. Bus. Intell.* **2025**, 11, 238–253.
91. Jehangir, B.; Radhakrishnan, S.; Agarwal, R. A survey on named entity recognition—Datasets, tools, and methodologies. *Nat. Lang. Process. J.* **2023**, 3, 100017. [[CrossRef](#)]
92. Sharma, A.; Amrita; Chakraborty, S.; et al. Named entity recognition in natural language processing: A systematic review. In Proceedings of the Second Doctoral Symposium on Computational Intelligence, Lucknow, India, 6 March 2021; pp. 817–828.
93. Baigang, M.; Yi, F. A review: Development of named entity recognition (NER) technology for aeronautical information intelligence. *Artif. Intell. Rev.* **2023**, 56, 1515–1542. [[CrossRef](#)]
94. Lagutina, N.S.; Vasilyev, A.M.; Zafievsky, D.D. Tasks in named entity recognition: Technologies and tools. *Autom. Control Comput. Sci.* **2024**, 58, 779–796. [[CrossRef](#)]



Copyright © 2026 by the author(s). Published by UK Scientific Publishing Limited. This is an open access article under the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher's Note: The views, opinions, and information presented in all publications are the sole responsibility of the respective authors and contributors, and do not necessarily reflect the views of UK Scientific Publishing Limited and/or its editors. UK Scientific Publishing Limited and/or its editors hereby disclaim any liability for any harm or damage to individuals or property arising from the implementation of ideas, methods, instructions, or products mentioned in the content.