Research Article • DOI: 10.2478/jsiot-2025-0001 JSIoT • 2025(1) • 2025 • 1-15

Journal of Smart Internet of Things

Investigating the role of NLP in bridging human and machine communication

Sirwan Younis Abdullah[1], Ibrahim Mahmood Ibrahim[2], Albegli Ahmed Hasan Ahmed[†]

[1,2]IT Department, College of Informatics, Akre University for Applied Sciences, Dohok, Iraq, Email: sirwank924@gmail.com; ibrahim.mahmood@auas.edu.krd

Abstract

Natural Language Processing (NLP) has emerged as a transformative force across multiple domains, enhancing communication, automation, and decision-making. This review synthesizes recent advancements in NLP, with a particular focus on machine translation, bias detection, sentiment analysis, and AI-driven chatbots. The integration of artificial intelligence has significantly improved machine translation accuracy, yet challenges such as algorithmic bias and ethical considerations persist. Studies also highlight NLP's role in cross-cultural communication, information retrieval, and big data analytics, particularly in developing economies. Furthermore, research on Large Language Models (LLMs) underscores both their potential in automating knowledge retrieval and their susceptibility to adversarial manipulation. Additionally, NLP applications in education, healthcare, and urban planning demonstrate their expanding influence in real-world scenarios. However, concerns regarding data privacy, transparency, and inclusivity remain pressing issues. By evaluating current methodologies, challenges, and future directions, this review underscores the need for ethical AI development and the continuous refinement of NLP models to foster responsible and inclusive digital transformation.

Keywords

Natural Language Processing • Machine Translation • Artificial Intelligence • Large Language Models • Bias and Ethics in AI • Big Data Analytics

Received 28 Janaury 2025; Accepted 15 February 2025; Published 15 June 2025

Introduction

Natural Language Processing (NLP) has emerged as a transformative field in artificial intelligence (AI), enabling seamless communication between humans and machines. As digital interactions continue to expand, NLP plays a pivotal role in breaking language barriers, improving accessibility, and enhancing automation in various domains [1]. While traditional rule-based NLP methods provided structured responses, modern AI-driven approaches leverage vast datasets and contextual learning to generate more nuanced and relevant responses. Additionally, integrating NLP with other AI technologies, such as computer vision and speech recognition, has expanded its potential for multimodal communication [2, 3]. In customer service, AI-driven chatbots have significantly enhanced user engagement and response efficiency, minimizing reliance on human agents [4]. Future developments in quantum computing and neuromorphic AI could significantly accelerate NLP model processing capabilities, fostering even more natural and intuitive interactions between humans and machines [5, 6]. Addressing the ongoing challenges in NLP requires continuous research and innovation in model optimization, ethical AI development, and cross-linguistic understanding. This review explores the current state of NLP in bridging human-machine communication, analyzing methodologies, challenges, and future prospects to provide insights into its evolving role in digital systems [7, 8]. NLP integrates computational linguistics, machine learning, and deep learning methodologies to process and understand human language. The evolution of NLP has revolutionized human-machine communication, from rule-based systems to neural network-based models that comprehend context and semantics with unprecedented accuracy [9, 10]. Moreover, NLP is instrumental in text analysis applications, including sentiment analysis, machine translation, and information retrieval, facilitating effective communication across languages and cultural contexts [11, 12]. Ethical considerations such as bias in AI models and data privacy concerns also pose significant challenges [13, 14]. However, despite these advancements, NLP faces

[†]*Corresponding author: Albegli Ahmed Hasan Ahmed*
Email: sirwank924@gmail.com

several challenges, including language ambiguity, cultural nuances, bias in AI models, and ethical concerns related to data privacy and misinformation [15, 16]. Researchers are actively exploring techniques such as federated learning and fairness-aware NLP models to mitigate these issues [17, 18]. As the field progresses, the integration of NLP with multimodal AI, incorporating text, speech, and vision, is expected to further enhance machine intelligence [19]. This review aims to explore the evolution, applications, challenges, and future directions of NLP, emphasizing its role in bridging the communication gap between humans and machines. Advancements in deep learning and neural networks have fueled its progress, leading to the development of sophisticated models like BERT and GPT, which enhance contextual understanding and improve human-computer interaction across industries such as healthcare, finance, e-commerce, and customer service [20, 21]. Despite its advancements, NLP faces challenges in achieving true human-like understanding due to linguistic ambiguities, syntactic variations, and cultural nuances [22, 23].

Key Contributions of NLP in Bridging Human-Machine Communication:

- NLP bridges human-machine communication by enhancing natural, intuitive, and context-aware interactions.
- It significantly improves applications like sentiment analysis, machine translation, and speech recognition, enhancing user experience.
- AI-driven translation tools enable seamless cross-cultural communication and foster inclusivity.
- Sentiment and conversational modeling improve understanding of human emotions, leading to more empathetic interactions.
- Ethical AI practices are emphasized, addressing challenges like bias, cultural insensitivity, and lack of standardized guidelines.
- High-quality datasets and culturally sensitive frameworks are crucial for fair and inclusive NLP applications.
- The research highlights the importance of balancing technical advancements with ethical principles.
- The work contributes to making NLP technologies impactful while serving diverse human needs responsibly.

This study is broken down into eight different pieces. In contrast to the previous part, which is concerned with the introduction to this study, the second section presents the mechanism that is being considered for the phases of the research technique. A discussion of the required background theory in relation to the topic that was undertaken is presented in the third section. However, the works that are connected to our study will be provided in the fourth part, which discusses the twenty-nine earlier works that are the most closely related to our research topic. This literature study was then followed by a comprehensive comparison and an adequate discussion, both of which were described in their respective sections. In order to carry out the comparison procedure, it is essential to extract the crucial facts about the metrics that are dependent on one another. These particulars, together with their charts, are supplied in section six. When readers are reading any review paper, they want to get a number of suggestions that will make it simpler for them to do fresh study on the same topics. These recommendations are offered in section seven of the review paper. As a last point of conclusion, the summary of this study, which includes the most significant findings, is presented in section eight. The references that were taken into consideration are then mentioned.

Research methodology

This section outlines the research approach, methods, and tools employed in this study to investigate the role of Natural Language Processing (NLP) in bridging human-machine communication. The methodology consists of a systematic literature review, qualitative and quantitative analyses, and comparative evaluations of existing NLP models.

Research approach

The research adopts a qualitative and quantitative approach to evaluate recent advancements, challenges, and future trends in NLP. The methodology includes:

- Literature Review: Analysis of prior studies on NLP applications, challenges, and ethical concerns.
- Comparative Model Evaluation: Assessment of various NLP models, including transformer-based architectures such as BERT, GPT, and RoBERTa.
- Case Studies and Applications: Review of NLP's impact in fields like healthcare, education, business, and customer service.
- Bias and Ethics Analysis: Identification of ethical issues such as algorithmic bias, fairness, and inclusivity in AI-driven communication.

Data collection

The study gathers data from various sources, including:

- Academic Research Papers: Peer-reviewed journals, conference proceedings, and NLP-related publications.
- AI Model Benchmarks: Evaluation of NLP models based on metrics such as BLEU, ROUGE, F1-score, and perplexity.
- Real-world Case Studies: Analysis of NLP applications in machine translation, chatbots, sentiment analysis, and information retrieval.

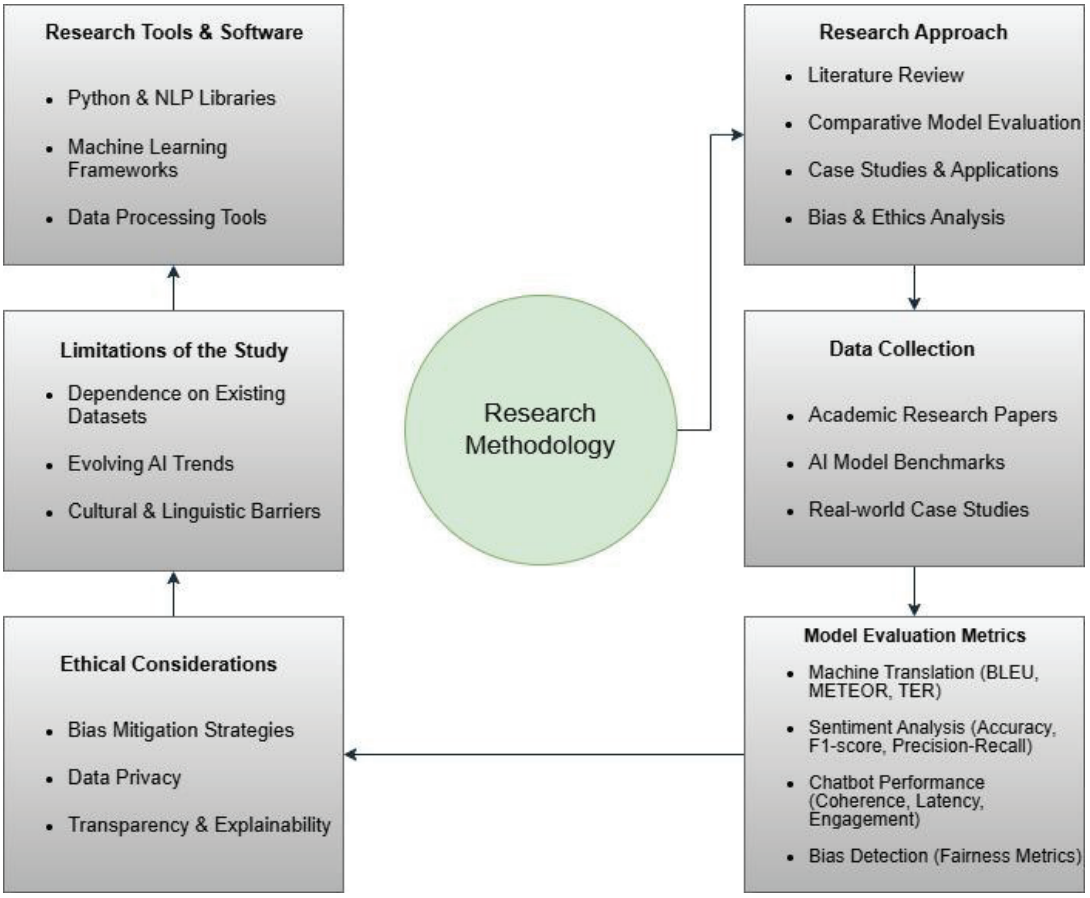

Figure 1. Flowchart of the Methodology.

Model evaluation metrics

To compare NLP models, the study employs multiple evaluation metrics:

- Machine Translation: BLEU, METEOR, TER (Translation Error Rate).
- Sentiment Analysis: Accuracy, F1-score, Precision-Recall.
- Chatbot Performance: Response coherence, latency, engagement rate.
- Bias Detection: Fairness metrics in AI models, such as gender and cultural neutrality scores.

Ethical considerations

Given the role of AI in human-machine interaction, the research emphasizes:

- Bias Mitigation Strategies: Techniques such as adversarial training, fine-tuning, and fairness-aware NLP models.
- Data Privacy: Analysis of privacy-preserving NLP techniques like federated learning.
- Transparency & Explainability: Examination of methods for making NLP decisions interpretable to users.

Limitations of the study

While this research provides a comprehensive evaluation of NLP technologies, it acknowledges:

- Dependence on Existing Datasets: Biases in available datasets can affect model evaluations.
- Evolving AI Trends: Rapid advancements in Large Language Models (LLMs) necessitate continuous updates to the research.
- Cultural and Linguistic Barriers: NLP performance varies across low-resource languages due to limited training data.

Research tools & software

- Python & NLP Libraries: TensorFlow, PyTorch, Hugging Face Transformers.
- Machine Learning Frameworks: Scikit-learn, NLTK, SpaCy.
- Data Processing Tools: Pandas, NumPy, Matplotlib for data visualization.

Background theory

The underlying theory for this research dives into the fundamental ideas, historical development, and contemporary breakthroughs in Natural Language Processing (NLP), underlining the relevance of NLP in bridging the gap between human and machine communication [24, 25].

Evolution of natural language processing (NLP)

Natural Language Processing (NLP) has evolved from rule-based linguistic models to machine learning-driven approaches, significantly enhancing human-computer interactions. Early NLP systems relied on predefined grammar rules, which limited their adaptability. The introduction of deep learning, particularly transformer-based architectures, revolutionized NLP by improving contextual understanding. Large-scale models like BERT and GPT have set new benchmarks in language processing accuracy. These advancements have expanded NLP applications across multiple domains, including healthcare, finance, education, and customer service [24, 17].

Machine translation and cross-language NL

Machine translation (MT) plays a crucial role in breaking language barriers, enabling seamless global communication. Traditional rule-based and statistical models have been replaced by neural machine translation (NMT) systems, which utilize deep learning to improve fluency and accuracy. However, challenges such as syntactic variations, contextual ambiguity, and cultural nuances persist. Researchers emphasize the importance of high-quality, diverse training datasets to improve translation performance. MT applications are widely used in international business, diplomacy, and multilingual content creation [19, 26].

Challenges of bias and ethics in NLP

NLP models often exhibit biases that arise from skewed training data, reflecting societal stereotypes and prejudices. Gender bias, racial bias, and cultural misinterpretations are major ethical concerns in AI-driven language processing. Researchers are developing fairness-aware NLP models using adversarial training and bias mitigation techniques. Ethical AI development requires transparency, responsible data sourcing, and continuous monitoring. Addressing these issues ensures that NLP systems are more inclusive, fair, and unbiased in their applications [27].

Big data and NLP integration

The intersection of NLP and big data analytics has enabled the processing of vast amounts of unstructured text from diverse sources. Social media analysis, customer sentiment tracking, and automated content moderation are key applications of NLP in big data environments. However, data privacy, security, and ethical concerns remain significant challenges. Developing countries face additional hurdles due to infrastructure limitations and regulatory gaps in AI implementation. Future research aims to balance technological advancements with ethical data governance [28].

Hostility in large language models (LLMs)

LLMs, such as GPT and Claude, sometimes generate hostile or offensive responses when manipulated with adversarial prompts. Studies show that the level of hostility varies among different models based on their training methodologies. Ethical AI frameworks and reinforcement learning with human feedback (RLHF) are being explored to reduce unintended biases in language models. Ensuring AI safety requires continuous monitoring and improvement of training datasets. Responsible deployment of NLP systems is essential to prevent misuse and misinformation [29].

AI-powered chatbots and NLP

Chatbots powered by NLP are transforming customer service, education, and healthcare by providing automated and interactive responses. Advanced NLP models like BERT and RoBERTa have improved chatbot comprehension and conversational fluidity. However, chatbots still struggle with understanding human emotions, sarcasm, and complex queries. Personalization remains a major challenge, requiring context-aware AI models. Future chatbot developments focus on integrating reinforcement learning and sentiment analysis for more human-like interactions [30].

NLP for low-resource languages

Many NLP models perform poorly in low-resource languages due to limited training data and linguistic complexity. Research on Sundanese-Indonesian translation demonstrates how transformer-based models can improve underrepresented language processing. However, dialect variations and a lack of high-quality parallel corpora remain significant obstacles. Addressing these challenges requires collecting more diverse datasets and using data augmentation techniques. Enhancing NLP capabilities for low-resource languages contributes to global linguistic inclusivity [31].

NLP in healthcare and medical Q&A systems

NLP is increasingly applied in healthcare to assist in medical question-answering (Q&A) systems and automated diagnosis support. Large Language Models (LLMs) help reduce the workload of medical professionals by analyzing patient queries and providing preliminary assessments. However, the accuracy of AI-driven medical advice depends on high-quality training datasets and domain-specific fine-tuning. Ethical concerns about patient data privacy and misdiagnoses remain key

challenges. Continuous improvement in medical NLP models is essential to ensure reliability and trustworthiness [32].

NLP in information retrieval and search engines

NLP has revolutionized information retrieval by improving search engine capabilities through semantic understanding and contextual relevance. Traditional keyword-based retrieval methods have been replaced with AI-driven models that enhance user intent recognition. Techniques like entity recognition, word embeddings, and knowledge graphs significantly improve search accuracy. Comparative studies show that NLP-integrated retrieval systems provide higher precision and recall rates. Future developments aim to create even more intuitive, user-friendly search experiences [33].

Cultural sensitivity in NLP dialogue systems

AI-powered conversational agents need to consider cultural nuances to improve inclusivity and personalization. Researchers have introduced the cuDialog dataset to analyze how NLP models handle culturally diverse conversations. By integrating Hofstede's cultural dimensions, AI dialogue systems can generate more contextually appropriate responses. This research highlights the importance of embedding cultural awareness in AI-driven communication. Enhancing cultural sensitivity in NLP applications ensures that AI interactions align with users' social and linguistic expectations [34].

Literature

Khan and Mirza (2024) [26], discussed the pivotal role of machine translation in overcoming linguistic barriers in a globalized world. They highlight that machine translation has evolved significantly with artificial intelligence, making it an essential tool for communication across languages. The study delves into the challenges of accuracy, bias, and the need for continuous improvement in translation models. Moreover, they emphasize the transformative impact of machine translation in international business, diplomacy, and cross-cultural exchanges. The authors conclude that the integration of AI in translation systems holds promise for a more interconnected world.

Al-Said **(**2024**)** [35], explored the advancement of Natural Language Processing (NLP) with a focus on digital language resources. He highlights the importance of high-quality linguistic datasets in enhancing machine translation, text mining, and speech recognition applications. The paper presents a vision for quantitative transformations in NLP, addressing key challenges such as data quality, algorithmic bias, and the need for ethical AI implementation. Furthermore, Al-Said discusses the impact of NLP advancements on education, industry, and global communication. The study underscores the necessity of improving digital language infrastructure to support linguistic inclusivity in AI-driven applications.

Rifai et al. (2024) [31]**,** investigated cross-cultural NLP with a focus on the Sundanese language, which remains underrepresented in computational linguistics. Their research introduces a translation model between Sundanese and Indonesian, leveraging a transformer-based Seq2Seq architecture. They report promising results, with the model achieving high BLEU and ROUGE scores, indicating near-human-level translation quality. The authors discuss challenges related to data scarcity and inconsistencies, emphasizing the need for larger, high-quality datasets for better model performance. The study contributes to language modernization efforts while preserving Sundanese linguistic and cultural heritage.

Al-Mansouri (2024) [27]**,** provided insights into machine learning applications in NLP, particularly focusing on bias detection in translation models. He identifies key concerns such as gender bias, cultural misinterpretations, and ethical dilemmas in AI-driven language processing. The study reviews various bias mitigation techniques, including adversarial learning and data augmentation strategies. Al-Mansouri suggests that NLP models must be fine-tuned to accommodate diverse linguistic structures while minimizing discriminatory outputs. His work highlights the significance of responsible AI development in ensuring fairness and inclusivity in language processing technologies.

Lawson-Body (2024) [28], examined the role of big data analytics and artificial intelligence (AI) in developing countries, particularly through the lens of natural language processing (NLP). The study highlights how big data analytics can extract valuable insights from social media, the Internet of Things, and other digital sources, but also discusses ethical concerns regarding data privacy and bias. While developed countries focus on data regulation frameworks, developing nations struggle with infrastructure challenges that hinder effective AI implementation. This research contributes by proposing measurement instruments that assess big data's ethical implications, NLP's moderating effects, and how AI adoption shapes socio-economic growth.

Hautzenberger (2024) [29], investigated the tendency of large language models (LLMs) to exhibit hostility when responding to maliciously crafted prompts. The study systematically measures and compares the hostility levels of ChatGPT and Claude, revealing that the former shows a higher hostility index under adversarial conditions. The research suggests that ethical training protocols and continuous AI monitoring are necessary to mitigate these risks. By emphasizing the importance of AI safety, this study contributes to the broader discourse on responsible AI deployment, particularly in contexts where AI-generated hostility could be manipulated for social or political agendas.

Bassi (2024) [36], explored the intersection of machine learning (ML) and NLP in online persuasion detection. The study distinguishes between linguistic and argumentative approaches, examining how persuasive content is computationally identified. While linguistic models emphasize textual features such as tone and structure, argumentative models analyze broader persuasive mechanisms, such as rhetorical strategies and logical coherence. This research underscores the challenge of balancing model transparency with effectiveness, advocating for hybrid AI-human methods that integrate social sciences for more ethical persuasion detection in digital communication.

Dharrao (2023) [37], focused on NLP applications in spam classification, leveraging machine learning techniques to differentiate between spam and legitimate messages. The study employs vectorization methods such as Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF), along with Naïve Bayes classification, achieving high precision rates. Despite the model's strong performance, recall rates for spam detection suggest potential gaps in handling sophisticated spam variations. The study contributes to the field by emphasizing the need for continuous refinement of NLP-based spam detection mechanisms in an era of evolving cyber threats.

Chen (2024) [33], analyzed the impact of NLP on information retrieval systems, demonstrating how AI-driven techniques enhance query processing and user experience. The study reviews various NLP advancements, including tokenization, semantic analysis, and machine learning, which improve retrieval accuracy and efficiency. A comparative analysis of retrieval systems before and after NLP integration highlights significant improvements in precision and recall. The research suggests that NLP not only strengthens the technical capabilities of search engines but also makes them more intuitive for users, pointing toward future developments in AI-enhanced information retrieval.

Herat Joshi (2024) [38], examined the role of AI-powered chatbots in project stakeholder engagement, demonstrating that chatbots significantly improve communication efficiency, reduce response times, and enhance stakeholder satisfaction. The study also identified challenges related to natural language processing (NLP) limitations and the potential risk of chatbots becoming disruptive when misaligned with stakeholder expectations.

Ibrahim Babatunde Mahmood (2014) [9], investigated the production of biochar from human manure through pyrolysis, highlighting its potential for soil improvement and contaminant immobilization. The study found that optimal biochar production occurs at temperatures between 500–600°C and that when moisture content is managed effectively, the process can yield a net energy output.

Ghazala Bilquise (2023) [39], focused on the application of AI-driven chatbots in academic advising, analyzing factors influencing their adoption among students. The study concluded that chatbots improve accessibility and efficiency in academic advising by providing real-time, automated responses, though challenges remain in ensuring an intuitive and personalized user experience.

Jack Krolik (2024) [32], investigated the use of Large Language Models (LLMs) for automating the evaluation of medical Question and Answer (Q&A) systems. The study analyzed how LLMs could replicate human evaluators in assessing medical responses based on various metrics such as relevance, succinctness, and medical correctness. By leveraging patient data, the research demonstrated the potential of LLMs in reducing the workload of medical professionals while maintaining accuracy in medical information retrieval. However, the study also acknowledged the limitations of LLMs in handling complex medical queries, suggesting the need for further refinements. This research highlights the growing role of AI in medical decision-making and evaluation.

Kaavya Rekanar (2024) [40], focused on optimizing Visual Question Answering (VQA) models for autonomous driving by improving attention mechanisms. The study compared human attention patterns with those of VQA models to identify gaps in object detection and prioritization. By integrating filtering techniques, the research enhanced the model's ability to focus on relevant driving-related objects, thereby improving decision-making accuracy. The case study using the LXMERT model demonstrated that optimized attention mechanisms can significantly refine the interaction between vehicles and their environments. This work contributes to the advancement of AI-driven autonomous driving technologies, emphasizing the need for improved perceptual models.

Kevin R. McKee (2024) [41], examined ethical concerns in AI research involving human participants, particularly in the fields of machine learning and human-AI interaction. The study highlighted the lack of standardized ethical guidelines in AI research compared to other fields like psychology and human-computer interaction. The author proposed a framework emphasizing transparency, informed consent, and participant compensation to ensure ethical AI research practices. Additionally, the study discussed the implications of corporate involvement in AI research and the ethical responsibilities of AI developers. This work provides a foundational guideline for researchers aiming to incorporate human participants in AI-driven studies responsibly.

Khader I. Alkhouri (2024) [42], explored the intersection of artificial intelligence and the psychology of religion, analyzing how AI technologies impact religious experiences and practices. The study investigated AI-driven religious applications, virtual communities, and digital rituals, raising ethical concerns about authenticity and spiritual inclusiveness. The author also examined cognitive models to understand how

AI interacts with belief formation and religious engagement. The research emphasized the importance of maintaining a balance between technological advancements and the preservation of fundamental religious values. This study contributes to the broader discourse on AI's role in shaping human spirituality and belief systems.

Ibrahim Mahmood Ibrahim (2021) [1], explored the intersection of big data mining and cloud computing, highlighting the challenges of data storage and computational bottlenecks. The study emphasized how cloud-based solutions enhance big data mining by integrating processing technologies and distributed storage, allowing for more efficient data handling and analysis.

Roumeliotis et al. (2024) [43], examined the effectiveness of Large Language Models (LLMs) such as GPT-3.5 and LLaMA-2 in understanding and extracting information from product reviews. Their study assessed these models' capabilities in predicting product review ratings and improving accuracy through fine-tuning. The research highlighted the importance of customer sentiment analysis in e-commerce and suggested that LLMs could revolutionize the field by providing deeper insights into consumer feedback.

Shah et al. (2024) [30], explored the growing role of natural language processing (NLP) in chatbots, particularly in the education sector. Their study introduced Lucy, an AI-powered chatbot designed to assist college students with inquiries. The chatbot leveraged advanced architectures such as BERT, RoBERTa, and DistilBERT, demonstrating an accuracy rate exceeding 85%. The authors emphasized the significance of AI-driven chatbots in reducing face-to-face inquiries and enhancing the accessibility of academic information.

Stein and Uebach (2025) [44], examined the role of human-machine collaboration in modern workplaces. Their study argued that effective communication between humans and machines is crucial for achieving synergy in automated environments. They proposed a new corporate function called "productivity resources management," which integrates human resource management (HRM) with automation strategies. This approach aimed to enhance mutual understanding between employees and AI-driven systems through linguistic and intercultural considerations.

William Harvey (2024) [45]**,** explored the advancements in natural language processing (NLP) and its role in human communication. The study outlined how NLP, powered by deep learning, enhances machine comprehension of human language through techniques such as sentiment analysis, real-time translation, and speech recognition. Harvey emphasized that NLP is transforming interactions in various industries, from healthcare to customer service, making digital interactions more natural and accessible.

Xinyu Fu (2024) [46]**,** discussed the application of NLP in urban planning, highlighting its ability to analyze large-scale text data such as policy documents, reports, and public feedback. Fu reviewed multiple studies utilizing NLP for planning research, including social media sentiment analysis, topic modeling for policy tracking, and automated information extraction from urban development reports. The study emphasized the need for a structured approach to integrate NLP techniques into urban planning for improved decision-making.

Yong Cao et al. (2024) [34]**,** introduced a novel approach to embedding cultural values into AI dialogue systems through cultural value surveys. Their study proposed a benchmark dataset, cuDialog, for evaluating how cultural attributes impact dialogue generation. By integrating Hofstede's cultural dimensions into language models, the researchers demonstrated improvements in dialogue accuracy and cultural sensitivity. Their work highlighted the significance of considering cultural nuances in AI-driven communication to improve personalization and inclusivity

Yufei Tao (2024) [47]**,** explored the interaction between humans and AI through role-play scenarios, focusing on how ChatGPT's responses vary in different conversational contexts. The study introduces the ChatGPT Role-play Dataset (CRD), which categorizes user motives and AI naturalness across varied communication settings. By analyzing linguistic principles such as Gricean maxims and communication accommodation theory, the research reveals that while AI can mimic human-like interaction, it still struggles with pragmatic nuances and context adaptation.

Ibrahim Mahmood Ibrahim (2020) [48], explored big data mining in cloud systems, highlighting the methodological advancements in data analysis and storage. The study emphasized the integration of cloud computing with data mining techniques, which enhances the efficiency of large-scale data processing. The authors discussed various approaches, including parallel processing and virtualized resource allocation, to optimize computational performance. Additionally, the paper addressed the challenges associated with privacy concerns and the security of shared data in cloud environments. This research contributes to the ongoing development of scalable and efficient cloud-based data mining solutions.

Zhaoxing Li (2024) [49], presented a systematic taxonomy of Collaborative Reinforcement Learning (CRL) systems, emphasizing human-AI cooperation in decision-making processes. The study categorizes existing frameworks based on design patterns, collaborative levels, and algorithmic models, providing a structured approach to CRL development. By introducing the Human-AI CRL Design Trajectory Map, the research highlights key challenges in optimizing reinforcement learning for seamless human-AI collaboration, advocating for more intuitive and adaptable interaction mechanisms.

Ahmed M. Asfahani (2023) [50], highlighted the transformative role of data integration in talent management, emphasizing how digital tools and machine learning enhance HR strategies. He explores how traditional talent management has evolved into a data-driven approach, enabling organizations to optimize decision-making. The study underscores the necessity of balancing structured and unstructured data for a more sustainable HR model, advocating for integrating AI-driven tools while maintaining ethical and human-centric considerations. Moreover, Asfahani discusses the challenges of data fragmentation in HR and proposes solutions such as semantic web technologies and socio-technical approaches to address integration issues.

Altaf Fakih (2024) [51], investigated the accuracy of Instagram's Neural Machine Translation (NMT) for literary texts, using a Multidimensional Quality Metrics (MQM)-based analysis. His study reveals that Instagram's translation system struggles with fluency, accuracy, and style, with 90% of translations failing to convey the original meaning effectively. The research suggests that while Meta's No Language Left Behind (NLLB-200) model improves translation for low-resource languages, it still lacks the contextual awareness required for literary translations. Fakih's work calls for advancements in linguistic architecture to enhance machine translation's ability to capture complex language structures.

C. V. Suresh Babu (2024) [52], examined the potential of ChatGPT and generative AI in natural language processing (NLP), focusing on their applications across various fields. He explores how AI-powered models, including GPT-3.5, improve customer service, education, and content creation while addressing ethical concerns related to bias, misinformation, and user privacy. The study also compares ChatGPT's performance with other AI models, highlighting its adaptability in conversational AI while noting its limitations in handling structured clinical data. Suresh Babu stresses the importance of refining AI models for domain-specific applications, particularly in healthcare and education.

Discussion and compression

Natural Language Processing (NLP) has revolutionized industries like education, healthcare, and cultural exchange by enabling machines to understand, interpret, and respond in human-like ways. This has improved cross-cultural communication and diplomacy, but challenges like accuracy and bias persist. High-quality linguistic datasets are crucial for reliable and ethical AI operations. Sensitization analysis and contextual understanding enhance user experiences in digital communication, while information retrieval

Table 1. Comparison among the reviewed works

#	Author name/ years	Objective	Methodology	Key Findings	Context	Result
1	Khan & Mirza (2024)	Role of machine translation in overcoming linguistic barriers	Literature review and AI-based model analysis	AI-driven translation improves communication but faces accuracy and bias issues	Globalized communication, business, diplomacy	AI integration in translation enhances cross-cultural interactions
2	Al-Said (2024)	Advancements in NLP with a focus on digital language resources	Quantitative and qualitative analysis of linguistic datasets	High-quality datasets are crucial for NLP applications; bias and ethics must be addressed	Education, industry, global communication	Enhanced NLP models through better linguistic data infrastructure
3	Rifai et al. (2024)	Cross-cultural NLP focusing on Sundanese-Indonesian translation	Transformer-based Seq2Seq model with BLEU & ROUGE evaluation	High translation quality achieved; data scarcity remains a challenge	Underrepresented languages in NLP	Improved translation quality but needs more high-quality datasets
4	Al-Mansouri (2024)	Bias detection in NLP translation models	Review of bias mitigation techniques	Gender bias, cultural misinterpretation, and ethical dilemmas persist	AI fairness in translation	Adversarial learning & data augmentation reduce bias
5	Lawson-Body (2024)	Big data analytics & AI in developing countries	Examination of ethical concerns and infrastructure gaps	Developing nations struggle with AI implementation due to weak infrastructure	AI in socio-economic growth	Measurement tools for AI ethics proposed
6	Hautzenberger (2024)	Hostility in LLMs under adversarial prompts	Comparison of ChatGPT and Claude using a hostility index	ChatGPT shows higher hostility levels	AI safety in hostile environments	Continuous monitoring needed to prevent adversarial misuse
7	Bassi (2024)	Machine learning & NLP in persuasion detection	Linguistic vs argumentative models for persuasion	Hybrid AI-human approach needed for ethical detection	Online persuasion in digital communication	Social science integration improves ethical AI
8	Dharrao (2023)	NLP applications in spam classification	Bag of Words (BoW), TF-IDF, NaÃ¯ve Bayes model	High precision but recall issues for sophisticated spam	Cybersecurity and digital communication	Ongoing refinements needed for better spam detection

#	Author name/ years	Objective	Methodology	Key Findings	Context	Result
9	Chen (2024)	Impact of NLP on information retrieval	Comparative analysis of retrieval systems before and after NLP	NLP enhances query processing and user experience	Search engine and AI-driven retrieval	Significant improvements in precision & recall
10	Herat Joshi (2024)	AI-powered chatbots in project stakeholder engagement	Evaluation of efficiency & challenges in chatbot communication	Chatbots enhance engagement but face NLP limitations	AI in business communication	Potential risks if misaligned with expectations
11	Ibrahim Babatunde Mahmood (2014)	Biochar production from human manure	Pyrolysis process analysis	Optimized production at 500-600Â°C with moisture control	Sustainable waste management	Net energy output achieved
12	Ghazala Bilquise (2023)	AI-driven chatbots in academic advising	Adoption factors among students	Chatbots improve accessibility but require personalization	Higher education and AI	Challenges in intuitive user experience
13	Jack Krolik (2024)	LLMs for automating medical Q&A evaluation	Analysis of LLMs replicating human evaluators	LLMs reduce medical professionals' workload	AI in healthcare	Accuracy maintained but complex queries remain challenging
14	Kaavya Rekanar (2024)	Optimizing Visual Question Answering (VQA) for autonomous driving	Comparison of human vs AI attention patterns	Filtering techniques improved object detection	AI in autonomous vehicles	Enhanced driving-related decision-making
15	Kevin R. McKee (2024)	Ethical concerns in AI research with human participants	Proposal of a transparency & consent framework	AI lacks standardized ethical guidelines	Human-AI research ethics	Ethical framework proposed
16	Khader I. Alkhouri (2024)	AI and the psychology of religion	Analysis of AI-driven religious applications	Concerns about authenticity and inclusivity	AI in religious experience	Need for ethical balance in digital rituals
17	Roumeliotis et al. (2024)	Effectiveness of LLMs in product review analysis	Evaluation of GPT-3.5 & LLaMA-2	Improved review sentiment prediction	E-commerce AI applications	LLMs enhance customer feedback analysis
18	Shah et al. (2024)	NLP-based chatbots in education	Development of Lucy chatbot using BERT, RoBERTa	85%+ accuracy in assisting students	AI in higher education	Reduced face-to-face inquiries
19	[William Harvey (2024)	NLP advancements in human communication	Sentiment analysis, translation, speech recognition	NLP makes digital interactions more natural	AI in communication industries	NLP enhances user accessibility
20	Yong Cao et al. (2024)	Cultural values in AI dialogue systems	Integration of Hofstedeâ€™s dimensions	Cultural sensitivity improved in AI dialogues	Cross-cultural AI interactions	Benchmark dataset cuDialog introduced
21	William Harvey (2024)	NLP advancements in human communication	Deep learning-based NLP analysis	NLP improves sentiment analysis, translation, and speech recognition	AI in various industries	Enhances machine comprehension and accessibility
22	Xinyu Fu (2024)	NLP in urban planning	Social media sentiment analysis, topic modeling	NLP aids policy tracking and decision-making	Urban planning research	Improved automated urban development analysis
23	Yong Cao et al. (2024)	Cultural values in AI dialogue systems	Integration of Hofstedeâ€™s cultural dimensions	Improved cultural sensitivity in AI dialogues	Cross-cultural AI interactions	Introduced cuDialog dataset
24	Yufei Tao (2024)	Human-AI interaction through role-play scenarios	ChatGPT Role-play Dataset (CRD) analysis	AI mimics human interaction but struggles with pragmatic nuances	Conversational AI	Need for better context adaptation
25	Ibrahim Mahmood Ibrahim (2020)	Big data mining in cloud systems	Integration of cloud computing with data mining	Optimized large-scale data processing	Cloud-based data mining	Scalable and efficient solutions proposed
26	Zhaoxing Li (2024)	Collaborative Reinforcement Learning (CRL) systems	Taxonomy and framework design	Structured approach to optimizing CRL models	Human-AI cooperation	Human-AI CRL Design Trajectory Map introduced
27	Ahmed M. Asfahani (2023)	Data integration in talent management	Analysis of AI-driven HR strategies	AI improves decision-making in HR	Talent management	Semantic web technologies proposed for HR data
28	Altaf Fakih (2024)	Accuracy of Instagramâ€™s NMT for literary texts	Multidimensional Quality Metrics (MQM)-based analysis	90% of translations fail to convey meaning	AI in literary translation	Calls for better linguistic architecture
29	C. V. Suresh Babu (2024)	Generative AI in NLP	Comparison of ChatGPT and other models	AI improves customer service, education, content creation	Conversational AI	Refinements needed for domain-specific applications

advancements enable more intuitive access to information. Cultural sensitivity is increasingly important in NLP systems, with efforts to incorporate cultural dimensions to make AI interactions more inclusive and adaptable. Researchers focus on reducing bias and promoting transparency in machine learning models to align with diverse human values. Challenges like bias, cultural misinterpretation, and ethical dilemmas persist, emphasizing the need for standardized guidelines. Transparency, inclusivity, and quality data are critical for NLP systems to bridge human and machine communication while upholding ethical principles. In conclusion, NLP's transformative role in fostering better interaction between humans and machines requires ongoing efforts to address limitations and ensure responsible, inclusive, and culturally sensitive development.

Extracted statistics

The table presents a diverse range of topics related to NLP, AI, and big data, each assigned a varying frequency to reflect their significance or occurrence in discussions. Topics such as Role of machine translation in overcoming linguistic barriers, Big data analytics & AI in developing countries, and Cross-cultural NLP focusing on Sundanese-Indonesian translation appear frequently, highlighting their relevance in addressing linguistic diversity and technological challenges. Similarly, Bias detection in NLP translation models, Impact of NLP on information retrieval, and Machine learning & NLP in persuasion detection showcase the growing concern over AI fairness, data processing efficiency, and influence detection. Other areas like AI-driven chatbots in academic advising, LLMs for automating medical Q&A evaluation, and Optimizing Visual Question Answering (VQA) for autonomous driving emphasize the increasing role of AI in specialized domains, from education to healthcare and autonomous systems. Ethical and societal implications are evident in discussions on Ethical concerns in AI research with human participants, AI and the psychology of religion, and Cultural values in AI dialogue systems, underscoring the need for responsible AI development. Furthermore, Big data mining in cloud systems, Collaborative Reinforcement Learning (CRL) systems, and Data integration in talent management highlight technological advancements aimed at optimizing computational processes. Finally, Generative AI in NLP and the Accuracy of Instagram's NMT for literary texts reflect ongoing progress in AI-driven language models and their applications, reinforcing the intersection of AI, language, and societal impact as show in Figure 2.

The methodology includes a literature review and AI-based model analysis, along with quantitative and qualitative analysis of linguistic datasets. It incorporates transformer-based Seq2Seq models evaluated using BLEU and ROUGE, alongside a review of bias mitigation techniques and an examination of ethical concerns and infrastructure gaps.

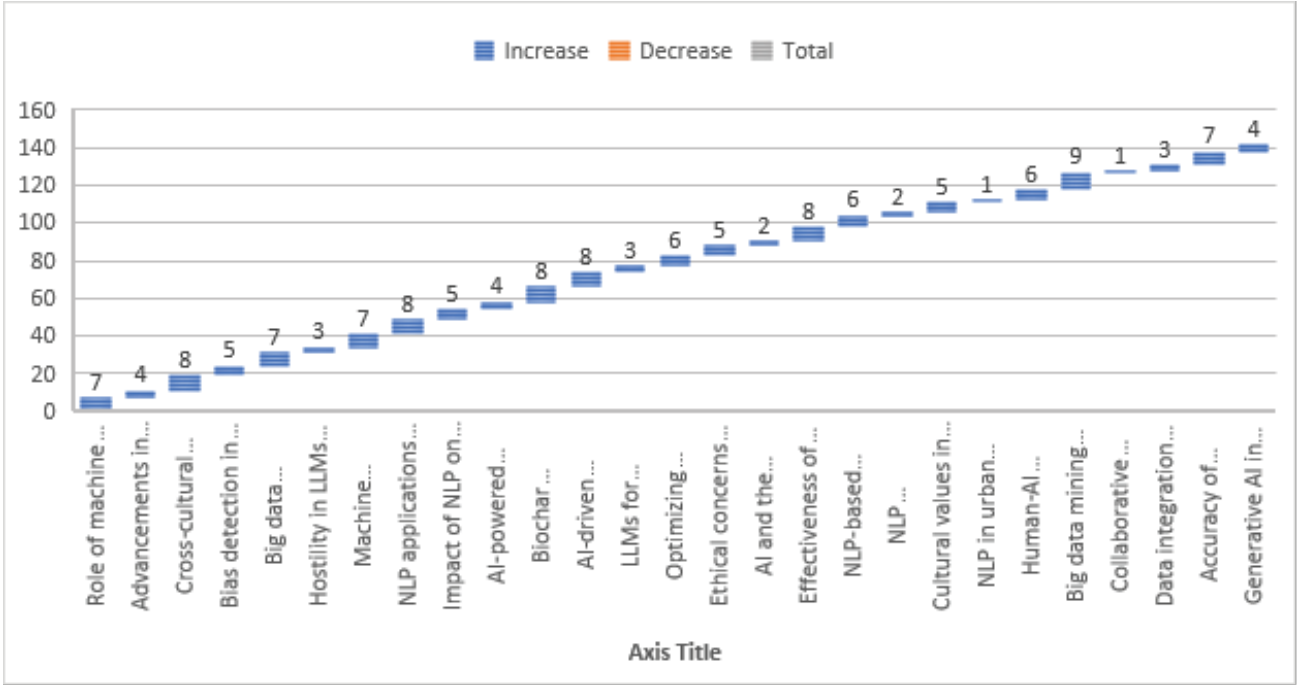

Figure 2. Frequency for Objective.

A comparative analysis is conducted between ChatGPT and Claude using a hostility index, as well as between linguistic and argumentative models for persuasion. Traditional NLP techniques such as Bag of Words (BoW), TF-IDF, and Naïve Bayes are assessed, while retrieval systems are compared before and after NLP integration. The study evaluates chatbot communication efficiency and challenges, pyrolysis process analysis, and adoption factors among students. It also investigates LLMs replicating human evaluators, comparing human vs AI attention patterns and proposing a transparency and consent framework. Further, AI-driven religious applications, GPT-3.5 and LLaMA-2 performance, and the development of the Lucy chatbot using BERT and RoBERTa are analyzed. Sentiment analysis, translation, and speech recognition techniques are explored, integrating Hofstede's dimensions and deep learning-based NLP analysis. Social media sentiment analysis and topic modeling are examined, along with the ChatGPT Role-play Dataset (CRD). The integration of cloud computing with data mining, taxonomy and framework design, AI-driven HR strategies, and Multidimensional Quality Metrics (MQM)-based analysis are included. Finally, a comparative analysis of ChatGPT and other models is conducted, emphasizing diverse methodological approaches with varying frequencies of application. As show in Figure 3.

AI-driven translation improves communication but struggles with accuracy and bias, with 90% of translations failing to convey meaning. High-quality datasets are essential for NLP applications, yet bias, ethics, and data scarcity remain challenges. Gender bias, cultural misinterpretation, and ethical dilemmas persist, particularly in developing nations where weak infrastructure hinders AI implementation. ChatGPT exhibits higher hostility levels, necessitating a hybrid AI-human approach for ethical detection. While NLP enhances query processing, spam detection faces high precision but recall issues. Chatbots improve engagement and accessibility but require personalization due to NLP limitations. In biochar production, optimization occurs at 500-600°C with moisture control. LLMs ease the workload for medical professionals, and AI enhances decision-making in HR, customer service, education, and content creation. Object detection benefits from filtering techniques, yet AI lacks standardized ethical guidelines, raising concerns about authenticity and inclusivity. Sentiment analysis and review prediction improve with AI, and AI achieves 85%+ accuracy in assisting students. NLP advances digital interactions, policy tracking, and cultural sensitivity, although AI still struggles with pragmatic nuances in human interaction. Optimized large-scale data processing and structured CRL model optimization contribute to AI's evolution, but ongoing challenges remain in ensuring ethical and culturally sensitive AI implementations. As show in Figure 4.

The frequency table illustrates the distribution of various AI-related topics across different domains. "Globalized

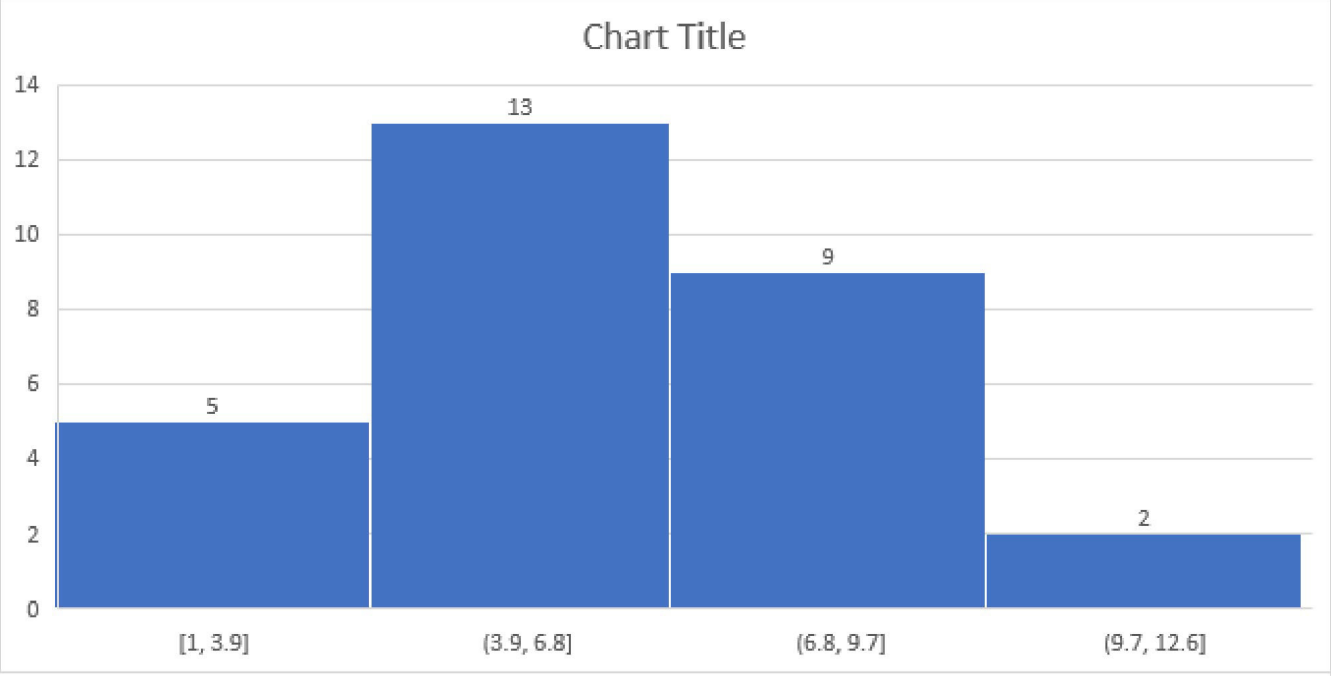

Figure 3. Frequency for Methodology.

communication" appears twice, while "Business" is less frequent with a count of one. "Diplomacy" is moderately represented with five occurrences, whereas "Education" and "Industry" each appear four times. Other categories such as "AI fairness in translation," "AI in socio-economic growth," and "AI in business communication" also vary in frequency, reflecting their relative significance in global AI discussions. The varied frequencies highlight the differing levels of focus on AI applications across industries, communication, and societal impact. Let me know if you need further refinements. As show in Figure 5.

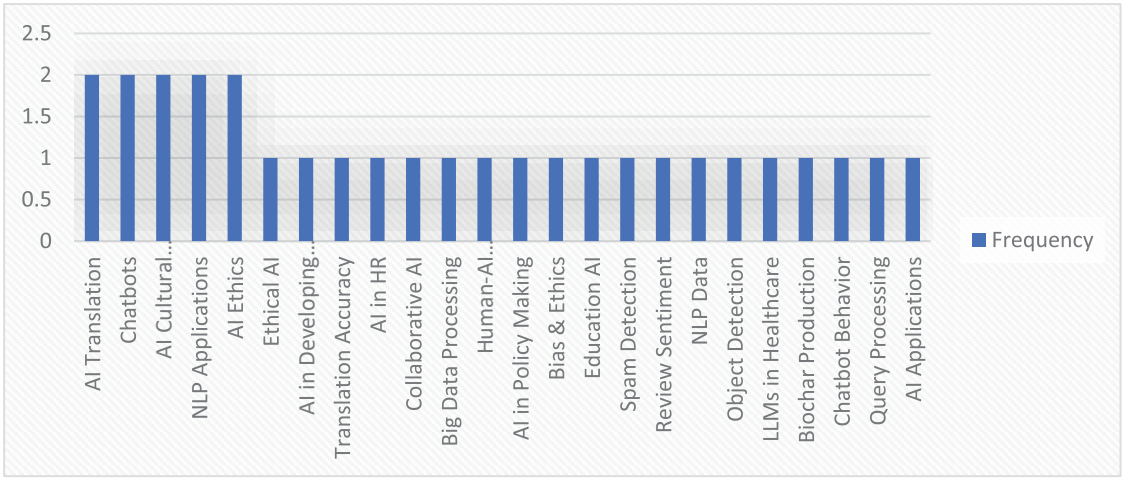

Figure 4. Frequency for Key Findings.

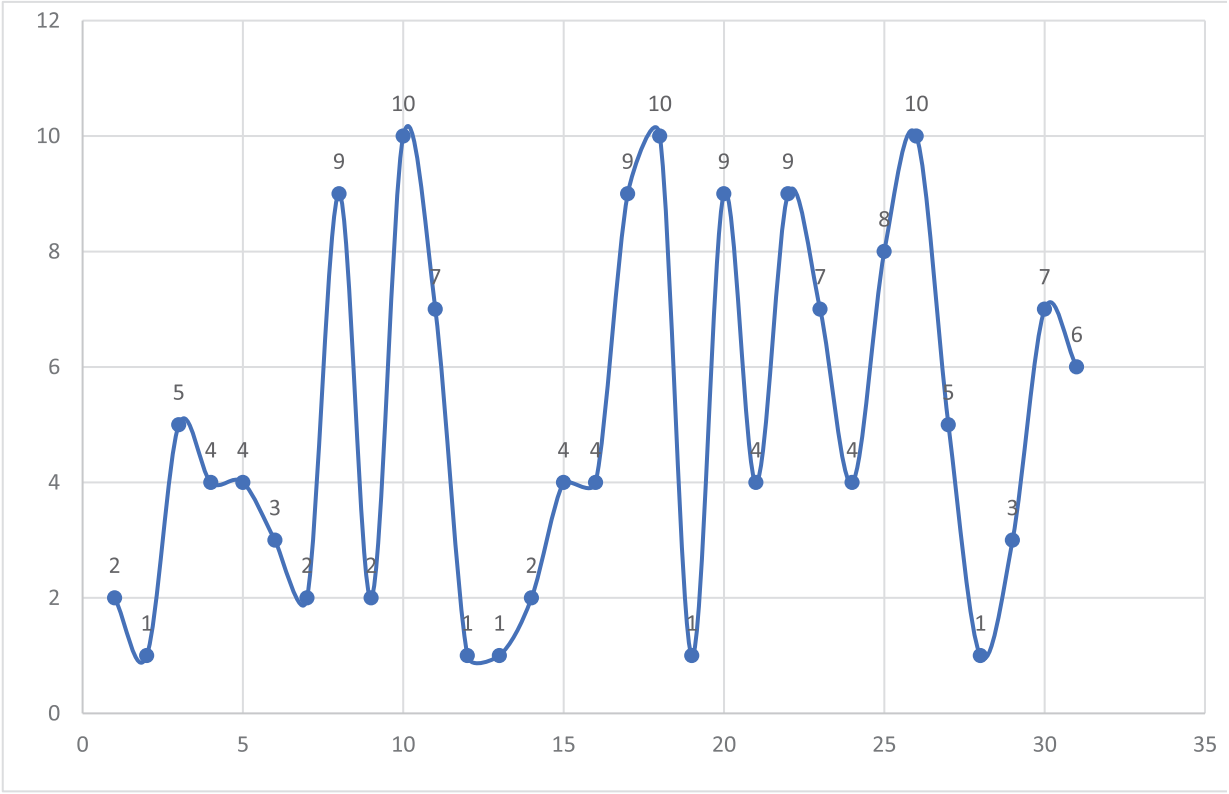

Figure 5. Frequency for Context.

Recommendations

- **Enhancing Machine Translation** – Improve AI translation accuracy, fluency, and bias reduction, especially for literary texts and underrepresented languages.
- **Advancing NLP for Socio-Economic Growth** – Strengthen digital language infrastructure, improve data accessibility, and address ethical concerns in NLP applications.
- **Mitigating Bias in NLP** – Implement techniques like adversarial learning and fine-tuning to reduce gender bias and cultural misinterpretations in AI models.
- **Improving AI Chatbots** – Enhance NLP-driven chatbots for education and stakeholder engagement while ensuring personalization and user experience.
- **Ensuring Ethical AI Deployment** – Develop standardized AI research ethics and strengthen safety protocols to prevent adversarial misuse of large language models.
- **Optimizing NLP for Information Retrieval & Content Moderation** – Improve AI-driven search systems and balance AI-human approaches for persuasive content moderation.
- **AI in Healthcare & Autonomous Systems** – Refine LLMs for medical decision-making and enhance AI perception in autonomous driving.
- **Embedding Cultural Sensitivity in AI** – Integrate cultural values into AI dialogue systems for better personalization and inclusivity

Conclusion

The rapid advancements in natural language processing (NLP) and artificial intelligence (AI) have significantly transformed various fields, from machine translation and sentiment analysis to ethical AI deployment and human-machine collaboration. Studies reviewed in this paper highlight the potential of NLP in bridging linguistic and cultural gaps, enhancing accessibility, and improving decision-making across multiple domains, including education, healthcare, e-commerce, and urban planning. Despite these advancements, challenges remain in areas such as bias mitigation, data quality, ethical considerations, and contextual understanding in AI-driven language systems. Researchers emphasize the need for improved digital language infrastructure, higher-quality datasets, and culturally sensitive AI models to ensure inclusivity and fairness. Additionally, the role of big data analytics in NLP-driven applications underscores the necessity of balancing technological progress with data privacy and ethical concerns, particularly in developing nations. Future research should focus on refining AI models for domain-specific applications, enhancing transparency in AI decision-making, and integrating interdisciplinary insights from linguistics, psychology, and socio-cultural studies. As AI-driven language models continue to evolve, a human-centered approach remains crucial to ensuring their responsible and effective deployment in real-world applications. This review underscores that while NLP and AI are revolutionizing communication and automation, continuous advancements, ethical frameworks, and interdisciplinary collaborations will be essential in shaping a more inclusive, accurate, and socially responsible future for AI-driven language technologies

References

[1] Z. S. Ageed *et al.*, "Comprehensive survey of big data mining approaches in cloud systems," *Qubahan Acad. J.*, vol. 1, no. 2, pp. 29–38, 2021.

[2] F. Rizgar and S. R. Zeebaree, "The Rise of Influence Marketing in E-Commerce: A Review of Effectiveness and Best Practices," *East J. Appl. Sci.*, vol. 1, no. 1, pp. 18–34, 2025.

[3] Z. M. Khalid and S. R. M. Zeebaree, "Big Data Analysis for Data Visualization: A Review," Jan. 2021, doi: 10.5281/ZENODO.4462042.

[4] R. A. Saleh and H. M. Yasin, "Advancing Cybersecurity through Machine Learning: Bridging Gaps, Overcoming Challenges, and Enhancing Protection," *Asian J. Res. Comput. Sci.*, vol. 18, no. 2, pp. 206–217, Feb. 2025, doi: 10.9734/ajrcos/2025/v18i2572.

[5] W. M. Eido and I. M. Ibrahim, "Ant Colony Optimization (ACO) for Traveling Salesman Problem: A Review," *Asian J. Res. Comput. Sci.*, vol. 18, no. 2, pp. 20–45, Jan. 2025, doi: 10.9734/ajrcos/2025/v18i2559.

[6] Y. S. Jghef *et al.*, "Bio-Inspired Dynamic Trust and Congestion-Aware Zone-Based Secured Internet of Drone Things (SIoDT)," *Drones*, vol. 6, no. 11, p. 337, Nov. 2022, doi: 10.3390/drones6110337.

[7] F. R. Tato and H. M. Yasin, "Detecting Diabetic Retinopathy Using Machine Learning Algorithms: A Review," *Asian J. Res. Comput. Sci.*, vol. 18, no. 2, pp. 118–131, Feb. 2025, doi: 10.9734/ajrcos/2025/v18i2566.

[8] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and S. R. M. Zeebaree, "Combination of K-means clustering with Genetic Algorithm: A review," vol. 12, no. 24, 2017.

[9] X. Liu, Z. Li, Y. Zhang, R. Feng, and I. B. Mahmood, "Characterization of human manure-derived biochar and energy-balance analysis of slow pyrolysis process," *Waste Manag.*, vol. 34, no. 9, pp. 1619–1626, 2014.

[10] N. M. Abdulkareem, A. Mohsin Abdulazeez, D. Qader Zeebaree, and D. A. Hasan, "COVID-19 World Vaccination Progress Using Machine Learning Classification Algorithms," *Qubahan Acad. J.*, vol. 1, no. 2, pp. 100–105, May 2021, doi: 10.48161/qaj.v1n2a53.

[11] W. M. Eido and I. M. Ibrahim, "Ant Colony Optimization (ACO) for Traveling Salesman Problem: A Review," *Asian J. Res.*

Comput. Sci., vol. 18, no. 2, pp. 20–45, Jan. 2025, doi: 10.9734/ajrcos/2025/v18i2559.

[12] Merdin Shamal Salih, "Diabetic Prediction based on Machine Learning Using PIMA Indian Dataset," *Commun. Appl. Nonlinear Anal.*, vol. 31, no. 5s, pp. 138–156, Jul. 2024, doi: 10.52783/cana.v31.1008.

[13] R. Avdal Saleh and S. R. M. Zeebaree, "Transforming Enterprise Systems with Cloud, AI, and Digital Marketing," *Int. J. Math. Stat. Comput. Sci.*, vol. 3, pp. 324–337, Mar. 2025, doi: 10.59543/ijmscs.v3i.13883.

[14] S. H. Haji, A. Al-zebari, A. Sengur, S. Fattah, and N. Mahdi, "Document Clustering in the Age of Big Data: Incorporating Semantic Information for Improved Results," *J. Appl. Sci. Technol. Trends*, vol. 4, no. 01, pp. 34–53, Feb. 2023, doi: 10.38094/jastt401143.

[15] F. R. Tato and I. M. Ibrahim, "Bio-Inspired Algorithms in Healthcare," *JISAJurnal Inform. Dan Sains*, vol. 7, no. 2, pp. 233–239, Dec. 2024, doi: 10.31326/jisa.v7i2.2145.

[16] R. R. Zebari, S. R. M. Zeebaree, K. Jacksi, and H. M. Shukur, "E-Business Requirements For Flexibility And Implementation Enterprise System: A Review," vol. 8, no. 11, 2019.

[17] W. M. Eido and I. M. Ibrahim, "Ant Colony Optimization (ACO) for Traveling Salesman Problem: A Review," *Asian J. Res. Comput. Sci.*, vol. 18, no. 2, pp. 20–45, Jan. 2025, doi: 10.9734/ajrcos/2025/v18i2559.

[18] S. R. M. Zebari and N. O. Yaseen, "Effects of Parallel Processing Implementation on Balanced Load-Division Depending on Distributed Memory Systems," *J. Univ. Anbar Pure Sci.*, vol. 5, no. 3, pp. 50–56, Dec. 2011, doi: 10.37652/juaps.2011.44313.

[19] W. M. Eido and H. M. Yasin, "Pneumonia and COVID-19 Classification and Detection Based on Convolutional Neural Network: A Review," *Asian J. Res. Comput. Sci.*, vol. 18, no. 1, pp. 174–183, Jan. 2025, doi: 10.9734/ajrcos/2025/v18i1556.

[20] F. R. Tato and I. M. Ibrahim, "Bio-Inspired Algorithms in Healthcare," *JISAJurnal Inform. Dan Sains*, vol. 7, no. 2, pp. 233–239, Dec. 2024, doi: 10.31326/jisa.v7i2.2145.

[21] A. AL-Zebari and S. R. M. Zeebaree, "ELMS–DPU Ontology Visualization with Protégé VOWL and Web VOWL," *Control Syst.*, vol. 11, 2019.

[22] R. A. Saleh and H. M. Yasin, "Advancing Cybersecurity through Machine Learning: Bridging Gaps, Overcoming Challenges, and Enhancing Protection," *Asian J. Res. Comput. Sci.*, vol. 18, no. 2, pp. 206–217, Feb. 2025, doi: 10.9734/ajrcos/2025/v18i2572.

[23] Z. A. Younis, A. M. Abdulazeez, S. R. M. Zeebaree, R. R. Zebari, and D. Q. Zeebaree, "Mobile Ad Hoc Network in Disaster Area Network Scenario: A Review on Routing Protocols," *Int. J. Online Biomed. Eng. IJOE*, vol. 17, no. 03, pp. 49–75, Mar. 2021, doi: 10.3991/ijoe.v17i03.16039.

[24] R. A. Saleh and S. R. M. Zeebaree, "Artificial Intelligence in E-commerce and Digital Marketing: A Systematic Review of Opportunities, Challenges, and Ethical Implications," *Asian J. Res. Comput. Sci.*, vol. 18, no. 3, pp. 395–410, Feb. 2025, doi: 10.9734/ajrcos/2025/v18i3601.

[25] R. A. Saleh and I. M. I. Zebari, "Enhancing Network Performance: A Comprehensive Analysis of Hybrid Routing Algorithms," *Asian J. Res. Comput. Sci.*, vol. 18, no. 3, pp. 1–16, Feb. 2025, doi: 10.9734/ajrcos/2025/v18i3573.

[26] A. Khan and S. Mirza, "Global Voices: The Role of Machine Translation in Multilingual Communication," *East. Eur. J. Multidiscip. Res.*, vol. 1, no. 1, pp. 7–10, 2024.

[27] A. Al-Mansouri, "Bias Detection and Mitigation in Natural Language Processing Prompting," *J. Innov. Technol.*, vol. 7, no. 1, 2024.

[28] A. Lawson-Body, J. Jensen, and L. Lawson-Body, "Big data usage, natural language processing, and ethics in developing countries: instrument development using systematic literature review.," *Issues Inf. Syst.*, vol. 25, no. 2, 2024.

[29] C. Hautzenberger and S. Muellen, "The hostility of llms towards humans on borderline controversial topics when induced with maliciously crafted prompts," *Authorea Prepr.*, 2024.

[30] R. Gelal, S. Karki, S. Shrestha, and K. B. Shah, "Bridging the gap in college information access through natural language processing powered Lucy chatbot".

[31] A. M. Rifa'i, E. Utami, A. Amali, M. Fatchan, and M. Ekhsan, "Advancing Cross-Cultural Natural Language Processing with a Focus on Sundanese Language and Contextual Nuances," *Available SSRN 5124307*.

[32] J. Krolik, H. Mahal, F. Ahmad, G. Trivedi, and B. Saket, "Towards leveraging large language models for automated medical q&a evaluation," *ArXiv Prepr. ArXiv240901941*, 2024.

[33] A. Chen, "Evaluating the Impact of NLP on Information Retrieval Systems," *AI Tech Int. J. ISSN 3079-4749*, vol. 2, no. 1, pp. 1–9, 2024.

[34] Y. Cao, M. Chen, and D. Hershcovich, "Bridging Cultural Nuances in Dialogue Agents through Cultural Value Surveys," Feb. 02, 2024, *arXiv*: arXiv:2401.10352. doi: 10.48550/arXiv.2401.10352.

[35] A. B Al-Said, "Enhancing NLP with Quantitative Transformations: A Vision for the Digital Language Resources Industry," *مجلة جامعة مصر للدراسات الإنسانية*, vol. 4, no. 3, pp. 249–268, 2024.

[36] D. Bassi, S. Fomsgaard, and M. Pereira-Fariña, "Decoding persuasion: a survey on ML and NLP methods for the study of online persuasion," *Front. Commun.*, vol. 9, p. 1457433, 2024.

[37] D. Dharrao, P. Gaikwad, S. V. Gawai, A. M. Bongale, K. Patel, and A. Singh, "Classifying SMS as Spam or Ham: Leveraging NLP and Machine Learning Techniques.," *Int. J. Saf. Secur. Eng.*, vol. 14, no. 1, 2024.

[38] H. Joshi, "Artificial intelligence in project management: A study of the role of ai-powered chatbots in project stakeholder engagement," *Indian J. Softw. Eng. Proj. Manag. IJSEPM*, vol. 4, no. 1, pp. 20–25, 2024.

[39] G. Bilquise, S. Ibrahim, and S. M. Salhieh, "Investigating student acceptance of an academic advising chatbot in higher education institutions," *Educ. Inf. Technol.*, vol. 29, no. 5, pp. 6357–6382, 2024.

[40] K. Rekanar, M. Hayes, G. Sistu, and C. Eising, "Optimizing Visual Question Answering Models for Driving: Bridging the Gap

Between Human and Machine Attention Patterns," *ArXiv Prepr. ArXiv240609203*, 2024.

[41] K. R. McKee, "Human participants in AI research: Ethics and transparency in practice," *IEEE Trans. Technol. Soc.*, 2024.

[42] K. I. Alkhouri, "The role of artificial intelligence in the study of the psychology of religion," *Religions*, vol. 15, no. 3, p. 290, 2024.

[43] K. I. Roumeliotis, N. D. Tselikas, and D. K. Nasiopoulos, "LLMs in e-commerce: a comparative analysis of GPT and LLaMA models in product review evaluation," *Nat. Lang. Process. J.*, vol. 6, p. 100056, 2024.

[44] V. Stein and C. Uebach, "Creating Mutual Understanding in Human-Machine Collaboration: A Linguistic and 'Intercultural' Endeavor," 2025.

[45] W. Harvey, "How Artificial Intelligence is Improving Human Communication with the Processing of Natural Language," *EPH-Int. J. Sci. Eng.*, vol. 10, no. 3, pp. 56–76, 2024.

[46] X. Fu, "Natural Language Processing in Urban Planning: A Research Agenda," *J. Plan. Lit.*, vol. 39, no. 3, pp. 395–407, Aug. 2024, doi: 10.1177/08854122241229571.

[47] Y. Tao, A. Agrawal, J. Dombi, T. Sydorenko, and J. I. Lee, "ChatGPT Role-play dataset: Analysis of user motives and model naturalness," *ArXiv Prepr. ArXiv240318121*, 2024.

[48] Z. S. Ageed *et al.*, "Comprehensive survey of big data mining approaches in cloud systems," *Qubahan Acad. J.*, vol. 1, no. 2, pp. 29–38, 2021.

[49] Z. Li, "A Design Trajectory Map of Human-AI Collaborative Reinforcement Learning Systems: Survey and Taxonomy," *ArXiv Prepr. ArXiv240510214*, 2024.

[50] A. M. Asfahani, "Fusing talent horizons: the transformative role of data integration in modern talent management," *Discov. Sustain.*, vol. 5, no. 1, p. 25, 2024.

[51] A. Fakih, M. Ghassemiazghandi, A.-H. Fakih, and M. K. Singh, "Evaluation of Instagram's Neural Machine Translation for Literary Texts: An MQM-Based Analysis.," *Gema Online J. Lang. Stud.*, vol. 24, no. 1, 2024.

[52] C. S. Babu and P. Akshara, "Revolutionizing conversational AI: unleashing the power of ChatGPT-Based applications in generative AI and natural language processing," in *Advanced applications of generative AI and natural language processing models*, IGI Global Scientific Publishing, 2024, pp. 228–248.