



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 7, July 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# Natural Language Processing and its Applications

Keshava P<sup>1</sup>, Sridhara N<sup>2</sup>, Hemanth Kumar G<sup>3</sup>, Jyothi T<sup>4</sup>

P.G. Student, Department of Master of Computer Applications, RNS Institute of Technology, Channasandra,  
Bangalore, India<sup>1 2</sup>

Assistant Professor, Department of Master of Computer Applications, RNS Institute of Technology, Channasandra,  
Bangalore, India<sup>3 4</sup>

**ABSTRACT:** Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that investigates robot reading and interpret human language. With the digital age's data explosion, natural language processing (NLP) has become increasingly vital in facilitating communication between humans and machines. This research paper discusses NLP, including its history, applications, methodologies, and problems. The study also addresses the latest advances in NLP employing AI, Like deep learning, and their potential societal implications

**KEYWORDS:** Natural Language Processing, Language Processing, Speech Recognition, Natural Language, Oral History, Language Processing, Automatic Speech Recognition, NLP, AI Artificial Intelligence

## I. INTRODUCTION

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) concerned with the interaction of computers and humans through the use of natural language. Since the 1950s, academics have been creating rule-based systems to parse and analyse human language. NLP has improved greatly in recent years, thanks to the emergence of machine learning and deep learning, allowing intelligent computers to understand, interpret, and synthesise human language.

## II. NLP APPLICATIONS

NLP has a wide range of applications in various fields including:

1. Sentiment Analysis: Natural language processing (NLP) can be used to analyse and understand the sentiment of textual data, such as customer feedback or social media posts.
2. Machine Translation: Natural Language Processing (NLP) can translate one language into another, allowing people who speak various languages to converse with one another.
3. Chatbots: Natural language processing (NLP) can be used to create chatbots that can interact with humans in natural language.
4. Speech Recognition: Natural language processing (NLP) can be used to recognise and transcribe human speech.
5. Text Summarization: Natural language processing (NLP) can be used to summarise extensive pieces of text, such as articles or reports

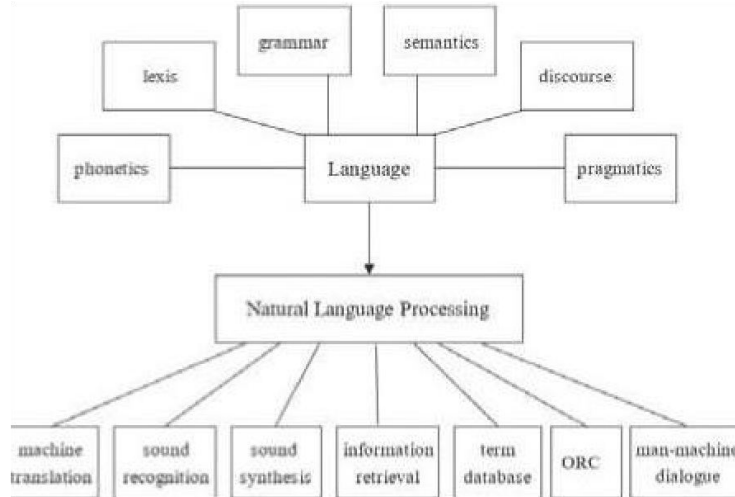


Fig. 1 The components of language and applications of NLP

### III. THE KEY CONCEPTS AND MAIN CONTENT OF NATURAL LANGUAGE PROCESSING

#### A. NLP's Key Concepts

Natural Language Processing (NLP) is the study of language communication problems between humans and machines. "Natural language processing is defined as a discipline that studies language issues in human-to-human and human-to-machine communication," wrote Bill Manaris in *Advanced in Computers* in 1998. Its duties are classified as natural language understanding (NLU) and natural language generating (NLG).

#### B. NLP's Main Content

Linguists divide language into the following levels: phonetics, lexis, grammar, semantics, discourse, and pragmatics. Natural language processing applications on the aforementioned levels can be further classified into the following sections: Machine translation, sound recognition, sound synthesis, automatic information retrieval, term database, optical character recognition, and man-machine communication are all examples of machine translation.

Natural language processing now has a wealth of theoretical foundations, language resources, and important technologies. The applications listed above have also advanced significantly. Sound recognition is utilised in voice translation and interrogation systems at unmanned service centres such as railway stations and airports, for example. Optical character recognition technology, which can be used in scanning software, is used to recognise printed typefaces and even handwritten scripts, and then generate electronic documents. Natural language processing is also widely used in many parts of translation technologies

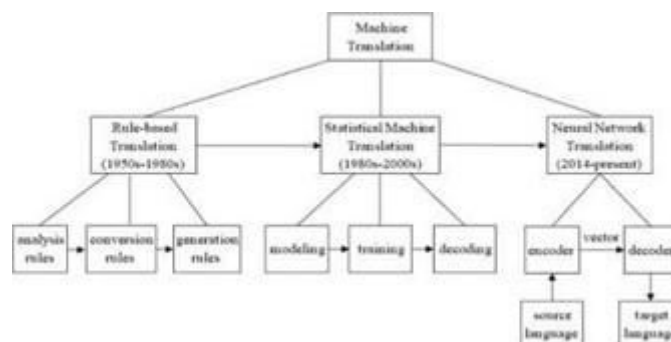


Fig. 2. The history and development of machine translation

IV. THE APPLICATIONS OF NATURAL LANGUAGE PROCESSING IN MACHINE TRANSLATION

In recent years, natural language processing technology has achieved rapid progress, and a series of machine translation systems have been devised and widely used. In 2006, Google began to develop its machine translation system, and finally brought Google Translate to the market. In 2011, Baidu launched its machine translation system that can support 27 languages. These machine translation systems are extensively utilized in our daily work and life, and the quality of translation has been enhanced significantly.

A. Machine Learning

A distinct feature of natural language processing is that it now increasingly adopts machine learning methods to acquire language knowledge. Machine learning is a discipline to study how to exploit experience to enhance the performance of the system through computational means. Specifically, the computer derives the "model" algorithm from a large amount of data, and then provides the empirical data to the computer, from which it generates a new model. Eventually, when new data appear, the computer can make corresponding judgments based on the generated model.

B. Machine Translation

Machine translation is a system that utilizes the computer to translate, that is, a computer program which is designed to translate text from one language (source language) to another language (target language) without the help of human. Machine translation constitutes one of the most important applications of natural language processing—a field which combines the elements of information technology with linguistics. It involves many classic natural language processing issues during the operation, such as data mining and cleansing, word segmentation, part-of-speech tagging, and syntactic analysis. In addition, machine translation is also closely related to the application of machine learning algorithms. Machine translation is broadly divided into rule-based machine translation and corpus based machine translation. According to the different modeling methods, corpus-based machine translation is further divided into example-based machine translation, statistical machine translation and neural machine translation. In machine translation, corpus is referred to as data, which means that machine translation requires a large amount of corpus to train the model. Different types of corpus are utilized to train different models. For instance, target language corpus is used to train language generation models (to improve sentence fluency), and parallel corpus is used to train translation models (to acquire translation knowledge).

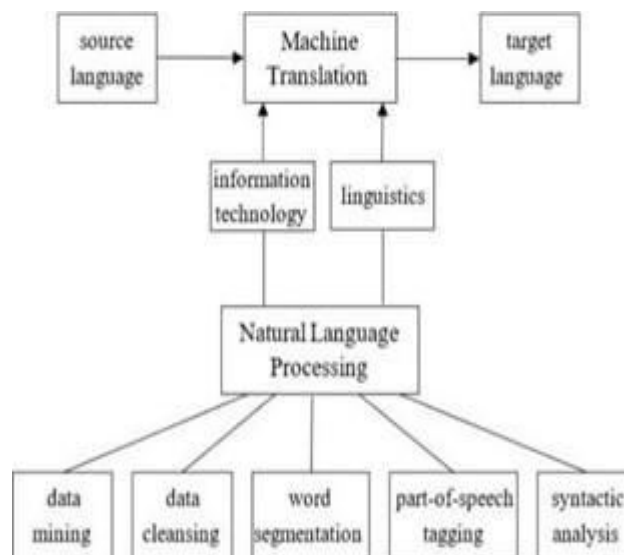


Fig. 3. The application of NLP in machine translation



## V. NLP APPROACHES

Several NLP approaches are utilized, including:

- 1) Tokenization is the process of dividing text into smaller parts known as tokens, such as words or sentences.
- 2) POS tagging: This is the process of labelling each word in a phrase with the appropriate part of speech, such as a noun, verb, or adjective.
- 3) Named Entity Recognition (NER): NER is the process of identifying and extracting named entities from text, such as persons, locations, or organisations.
- 4) Parsing: This is the process of analysing a sentence's grammatical structure.
- 5) Sentiment Analysis: This is the technique of detecting whether a piece of text has a good, negative, or neutral sentiment.

## VI. NLP CHALLENGES

Despite advances in NLP, there are still a number of issues that must be solved, including:

- 1) Ambiguity: Because human language is frequently un-clear, machines find it challenging to grasp and interpret it.
- 2) Contextual Understanding: Because human language is extremely contextual, comprehending the meaning of a statement frequently necessitates knowledge of the surrounding text.
- 3) Domain-Specific Language: Because different domains have different languages and jargons, developing general NLP models that function across all domains is difficult.
- 4) Data Scarcity: Developing good NLP models necessitates significant amounts of annotated data, which is frequently unavailable in many fields. AI breakthroughs in NLP: Recent advances in AI, such as deep learning, have considerably increased the accuracy of NLP models, allowing intelligent systems to understand, interpret, and synthesise human language.

## VII. NLP METHODOLOGY

It has been investigated numerous researches in recent times for the text and speech processing through the NLP. This arena has become one of the most interesting fields during last decade due to multifarious reason. There are various steps involved in the methodology of research which are defined as below. NLP.

- 1) Syntax: Syntax involves applying the rules of the gram- mar of the target language, the job is to evaluate the position of each word in a sentence and arrange this data into a form that is easier
- 2) Grammar: A statement in English consists of a noun phrase, a verb phrase, and a prepositional phrase in some cases. A noun phrase is a subject that a noun can summarize or identify. This sentence can contain articles and adjectives as well as an embedded verb phrase and the noun itself. A verb phrase represents an event and may be preceded by an associated noun phrase. A prepositional phrase is defined in the sentence as a nounor verb. Many natural languages are composed mainly of several parts of speech: verbs, nouns, adjectives, adverbs, conjunctions, pronouns and articles.
- 3) Parsing: It is the process by which a sentence is con-verted into trees representing the syntactic construction of the sentence. If the statement: "the green book is sitting on the desk" is the noun phrase: "The green book" and the action word state: "is perched on the work area". The expression tree will begin at the word stage and separated into the thing and action word state. The articles, the adjectives and the nouns would then be labeled. Parsing determines the validity of a sentence in relation to grammatical rules of the language.
- 4) Semantics: This provides a description of the objects; acts defined in a sentence and involve the particular details provided by adverb, adjective and proposals. In order to determine which purpose the consumer intended, this method gathers information crucial to the realistic analysis .
- 5) Pragmatics: It is the analysis of the original meaning of a pronouncement in a human language, by reparsing and parsing the pronouncement it is done by defining and addressing the systems ambiguities using one or more forms of techniques of disambiguation.
- 6) Ambiguity: It is defined as the issue that a pronounce-ment in a human language has more than one possible meaning. Available Tools for NLP: various methods which deals with the Chinese word selection and to-kenization have been introduced some are connected in powerful equipment kits are given below.

- Fudan NLP instrument in JAVA Language
- Language technology platform (LTP) instrument in C++
- Niu parser tool in C++
- Gensim Python
- Stanford Core NLP

## VIII. CONCLUSIONS

Natural language processing is rooted in multi-disciplines such as linguistics, computer science, and mathematics, and it has now become a major research field of artificial intelligence.

- The rapid advancements in natural language processing provides strong support for machine translation research.
- Natural language processing, a core component Of artificial intelligence, began with machine translation. Historically, the development of natural language eprocessing almost accords with that of machine translation, and the two complement each other.
- Natural language processing has been widely utilized in machine translation, and new breakthroughs have been achieved.
- This not only opens up abroad scope for machine translation research, but also adds vitality to its progress.
- This paper introduces the history and development of natural language processing and machine translation, analyzes the applications of natural language processing in machine translation, and discusses the challenges faced by natural language processing and the future trend of machine translation.
- In the era of knowledge-driven economy, with the continual indepth research on the cognitive mechanism of human brain and the advancements in natural language processing, the quality and efficiency of machine translation would be substantially increased

## REFERENCES

- [1] .D. Maynard and K. Bontcheva, "Natural language processing," in Perspectives on Ontology Learning, 2014.
- [2] J. Hirschberg and C. D. Manning, 'Advances in natural language processing,' Science. 2015, doi:10.1126/science.aaa8685.
- [3] D. Meurers, "Natural Language Processing and Language Learning," in The Encyclopedia of Applied Linguistics, 2012.
- [4] D. Rubinstein and A. Ami Podrebarac, "Introduction to Natural Language Processing," in Fragmentation of the Photographic Image in the Digital Age, 2019.
- [5] J. Li, L. Deng, R. Haeb-Umbach, and Y. Gong, "Fundamentals of speech recognition," in Robust Automatic Speech Recognition, 2016.
- [6] C. Tapsai, P. Meesad, and C. Haruechaiyasak, "TLS-ART: Thai Language Segmentation by Automatic Ranking Trie," 9th Int. Conf. Auton.Syst. (AutoSys 2016), no. October, 2016.
- [7] Y. A. Solangi, Z. A. Solangi, S. Aarain, A. Abro, G. A. Mallah, and A. Shah, 'Review on Natural Language Processing (NLP) and Its Toolkits for Opinion Mining and Sentiment Analysis,' 2019, doi:10.1109/ICETAS.2018.8629198.
- [8] A. Reshamwala, D. Mishra, and P. Pawar, "Review on Natural Language Processing," An International Journal (ESTIJ, 2013.)
- [9] A. Reshamwala, D. Mishra, and P. Pawar, 'Review on Natural Language Processing' Precision Agriculture View Project Image Retrieval View project Alpa Reshamwala Narsee Monjee Institute of Management Studies Review on Natural Language Processing," 2013.
- [10] W. W. Yim, M. Yetisgen, W. P. Harris, and W. K. Sharon, 'Natural Language Processing in Oncology Review,' JAMA Oncology. 2016.
- [11] .S. Sun, C. Luo, and J. Chen, 'A review of natural language processing techniques for opinion mining systems,' Inf. Fusion, 2017.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details