# Natural Language Processing

Vishal Vishwakarma[1], Rishabh Bokoliya[2], Akash Shinde[3], Shreyash Bhoinkar[4], Prof. Swati Patil[5]

[1]*Student, Computer Technology, Bharati Vidyapeeth Institute of Technology, Maharashtra, India*
[2]*Student, Computer Technology, Bharati Vidyapeeth Institute of Technology, Maharashtra, India*
[3]*Student, Computer Technology, Bharati Vidyapeeth Institute of Technology, Maharashtra, India*
[4]*Student, Computer Technology, Bharati Vidyapeeth Institute of Technology, Maharashtra, India*
[5]*Prof, Computer Technology, Bharati Vidyapeeth Institute of Technology, Maharashtra, India*

## ABSTRACT

*Natural Learning Processing (NLP) is used to understand the speech of the human and talk back in the same way. In this paper we have elaborate the meaning of NLP, components of the natural language processing i.e. Natural Language Understanding (NLU) and Natural Language Generation (NLG). In this way, the AI intelligence will now replace the need of human in between and will be able to do the work on behalf of human. We have also discussed the steps involved in the process of NLP and how basically a sentence is understood and an appropriate output is given to the user that is meaningful.*

**Keywords:** *Natural Language Processing, NLP, Natural Language Understanding, Natural Language Generation, NLU , NLG*

## 1. INTRODUCTION

Natural Language Processing or NLP is a branch of artificial intelligence that is an interaction between computers and human, it basically tells computer how to process large sentences given as an input and analyze it. NLP is the way of understanding the language of the human and what he is saying and replying him the way the human is saying and responding to it. NLP involves machines or robots to understand and process the language that human speak. It identifies the nouns, pronouns, verbs, adjectives, adverbs, prepositions, conjunctions and interjections and process it understands what basically the sentence is go through number of analysis phase and provide an appropriate meaningful sentence to the user. NLP is now replacing the need of a human to book an appointment and the need of a middle person in the communication process. The computer learns from the interactions between the humans. It connects the words in a network and search for thousands of possibilities for a sentence and provides an appropriate output.

## 2. LITERATURE REVIEW

Natural language processing (NLP) is a major area of artificial intelligence research, which in its turn serves as a field of application and interaction of a number of other traditional AI areas[1]. Until recently, the focus in AI applications in NLP was on knowledge representation, logical reasoning, and constraint satisfaction - first applied to semantics and later to the grammar.

In this paper the author explains some new developments in knowledge for computational linguistics, then discuss about what kind of knowledge is required for a dialogue system[2]. Information retrieval on the Web is an important technology. Estimating the confidence degree of information is a very difficult problem. The author discusses some possible ways of estimating the confidence degree of information.

Natural language processing (NLP) is a wide range of AI for the automatic understanding and representation of human language[3]. This paper focuses on recent developments in NLP research to look at the past, present, and future of NLP technology in a new light.

Processing natural language such as English has always been one of the central research issues of artificial intelligence[4], both because of the key role language plays in human intelligence and because of the wealth of potential applications. Many of the knowledge represents and inference techniques that have been applied successfully in knowledge-based systems were originally developed for processing natural language.

## 3. METHODOLOGY

**Components of NLP:**

### 3.1 Natural Language Understanding (NLU)

NLU involves several other sub-components as

- i) Ambiguity
- ii) Lexical
- iii) Syntactical
- iv) Referential.

i) Ambiguity

Ambiguity can be referred as the ability of having more than one meaning or being understood in more than one way. Natural languages are ambiguous, so computers are not able to understand language the way people understand. Ambiguity can occur at various levels of NLP. Ambiguity could be Lexical, Syntactic, Semantic, Pragmatic etc.

ii) Lexical (word level)

It works on words or noun. It identifies the single words in the sentence and process it further.

iii) Syntactical (passing)

It works on sentence level e.g. Call me a cab. But according to ambiguity this sentence have two meaning, your friend may call you cab by your name or they will call you a cab. This can also confuse the machine with the meaning of the sentence.

iv) Referential

John went to Alexa. He said I am hungry! In this computer doesn't understand who is hungry, who is I referencing the person. References are not properly given. So this may result in getting an inappropriate output.

### 3.2 Natural Language Generation (NLG)

Text Planning - It is an database with the knowledge base that helps the computer to plan the text for providing the output to the user which should be appropriate and meaningful.

Sentence Planning - It is the process of arranging the words into a meaningful order that is understandable by the user and provide it to the user.

Text Realization - It implements the sentence in a structure goes through thousands of combinations to provide an appropriate output to the user.
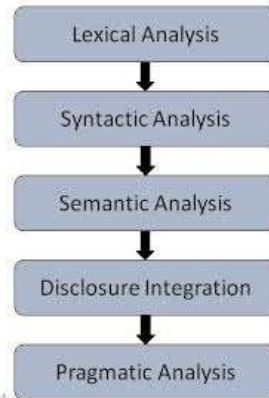
Steps involves in NLP

Fig. Steps in Natural Language Processing

i)   Lexical Analysis

It is the process of checking the words in a sentence and separating the full stop(.), punctuation marks(!). It is the method of breaking sentence into smaller and smaller pieces so that it is easy to understand to the machine.

ii)   Syntactic Analysis

It identifies the words and produces the sentence in a proper manner. It is the process of making the sentence.

iii)   Semantic Analysis

It is the process of knowing the meaning of individual words. E.g. Mango eats John where Mango is a fruit and john is a name. It will check the relation of the words and it checks in the network and connects it and it will identify that the sentence is wrong.

iv)   Disclosure Integration

It is the process of knowing the meaning by knowing the sentence before it and what is the possibility of the next sentence. E.g. John wanted it. He always had. In this sentence the machine will the check what is the header of the first sentence in order to determine the reference to 'He' and it will refer that he refers to john. He will know the meaning whether everything is meaning and appropriate or not. If two sentences are present then only it is used.

v)   Pragmatic Analysis

It is the process of combining all the above thing and re-interpret it. It will give a final result with an appropriate output.

## 4. Future Scope

NLP is going to a big revolution in the world of Artificial Intelligence. Not only it will take over human but it will act as a middle person from getting an appointment to fixing your car. NLP is now been developed by huge tech giant such as Google, IBM, and much more. NLP is going to learn from the human interactions so it will give response in the human way. Natural language processing can deal with text and speech recognition.

## 5. Conclusion

In this we have studied and gathered the information about natural language processing i.e. How NLP works and what are the components involved and the steps involved in it. As we saw each element has its own task and it helps us to

know whether the specific task is being performed properly or not. Each step determines the meaning of the sentence, what it is by breaking it into small parts and help us to know whether the machine has understood it properly or not.

## 7. REFERENCES

[1] A. Gelbukh "Natural language processing" https://ieeexplore.ieee.org/document/1587718/
[2] M. Nagao "Natural language processing and knowledge" https://ieeexplore.ieee.org/document/1598694/
[3] Erik Cambria and Bebo White "Jumping NLP Curves: A Review of Natural Language Processing Research " https://ieeexplore.ieee.org/abstract/document/6786458/
[4] T. Patten and P. Jacobs "Natural-language processing" https://ieeexplore.ieee.org/document/295134/
[5] Gobinda G. Chowdhury "Natural language processing"
https://onlinelibrary.wiley.com/doi/abs/10.1002/aris.1440370103
[6] Elizabeth D. Liddy "Natural Language Processing"
https://surface.syr.edu/cgi/viewcontent.cgi?referer=https://scholar.google.co.in/&httpsredir=1&article=1019&context=cnlp
[7] James F Allen "Natural Language Processing"
https://onlinelibrary.wiley.com/doi/abs/10.1002/0470018860.s00078
[8] Aravind K. Joshi "Natural Language Processing" http://science.sciencemag.org/content/253/5025/1242
[9] Christopher D. Manning and Hinrich Schutze "Foundations of Statistical Natural Language Processing"
https://books.google.co.in/books?hl=en&lr=&id=YiFDxbEX3SUC&oi=fnd&pg=PR16&dq=natural+language+processing+&ots=v_tkqwlJRP&sig=Sp8iQ-sTG6Fmy-
TREjtpYn4YcHg#v=onepage&q=natural%20language%20processing&f=false
[10] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Seven J. Bethard and Davind McClosky "The Stanford CoreNLP Natural Language Processing Toolkit" http://www.aclweb.org/anthology/P14-5010
[11] Adam L. Berger, Vincent J. Della Pietra and Stephen A. Della Pietra "A maximum entropy approach to natural language processing" https://dl.acm.org/citation.cfm?id=234289