

NATURAL LANGUAGE PROCESSING (NLP)

Kateryna Riabova

Lecturer,

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Natural language processing (NLP) is a general area of computer science, artificial intelligence and mathematical linguistics. It studies the problem of computer analysis and synthesis of natural language. Artificial intelligence analysis means understanding language and synthesis and generating intelligent text. Solving this problem means creating a more convenient form of computer-human interaction.

The development of computer technologies due to an increasing the issues of natural processing. Voice commands, instructions, chatbot, translation programs all of these applications work with NLP. Leading experts and researchers in mathematics, computer science and linguistics work to solve the problem of computer language analysis.

Natural language processing is a computerized approach to text analysis based on a number of theories and a set of technologies. This branch does not have a common definition, because it is in a state of constant research and development. However, there are certain aspects that would combine all existing definitions.

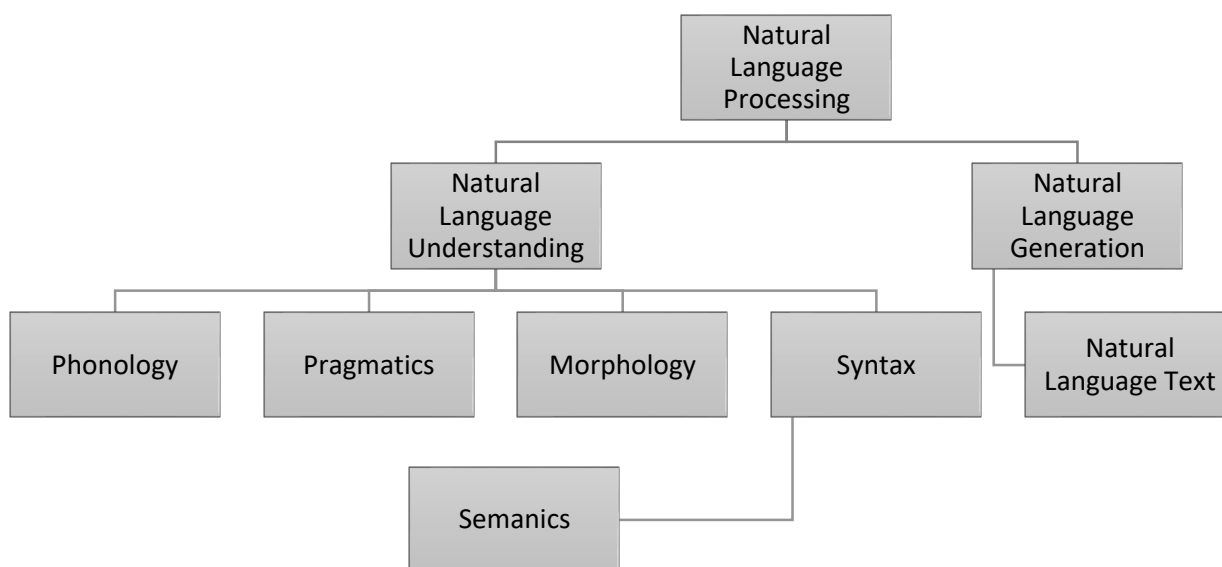


Figure1

Foreign scientists present a model of Natural Language Processing (Figure 1), they claim that NLP can be divided into 2 main parts – NL Understanding and NL Generation. However, other researchers added to these two categories Speech or Voice Recognition, Machine Translation and Spelling Correction.

The complexity of Natural Language, its description and processing leads to the division of this process into separate stages. These stages correspond to the languages' levels. Most language process belong to the modular type in which each level of linguistic analysis or synthesis corresponds to a separate processor module. Phonology refers to the systematic arrangement of sound, the sound is analyzed and encoded into a digital signal for interpretation. Morphology is to determine the morphological characteristics of a word. It helps to identify the parts of speech in the sentence and words that interact with each other. Syntax corresponds for word order and construction of grammatically correct sentences. Semantics is related to the understanding of the text which includes developments in the field of artificial intelligence. It is mainly useful for pragmatic analysis in order to determine which meaning was intended. Pragmatics is an analysis of the real meaning of sentence in natural language.

Natural Language Generation is a process that create from the phrases to the abstracts which are meaningful from an internal representation. Natural Language Generation happens in 4 stages:

- Identifying the goals
- Planning on how these goals can be achieved by evaluating the situation or available communicative sources
- Realizing the plans as a text

Natural Language Processing is a general area that spans many subsections. All of them usually use machine learning models, mostly neural networks, and data from many conversations between people. Since human languages are constantly and spontaneously evolving and computer needs clear and structured data, certain problems arise during processing and accuracy suffers. In addition, text analysis methods are highly dependent on language, genre, topic – additional configuration is always

required. However, today many tasks of NLP are still solved using deep learning of neural networks.

Reference

Joseph, S. R., Hlomani, H., Letsholo, K., Kaniwa, F., & Sedimo, K. (2016). Natural language processing: A review. *Natural Language Processing: A Review*, 6, 207-210.

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2017). Natural language processing: State of the art, current trends and challenges. *arXiv preprint arXiv:1708.05148*.

Onyshchenko, K. H., Daniil, Ya., & Kameniev, R. (2020). *Analiz metodiv obrobky pryrodnoi movy [Analysis of natural language processing methods]*.

Reshamwala, A., Mishra, D., & Pawar, P. (2013). Review on natural language processing. *IRACST Engineering Science and Technology: An International Journal (ESTIJ)*, 3(1), 113-116.