

# Theory: Natural language processing

🕒 13 minutes   0 / 5 problems solved

Skip this topic

Start practicing

767 users solved this topic. Latest completion was about 2 hours ago.

Just like everyone around the world, you probably interact with computers and mobile devices daily. Have you ever wondered how these machines manage to understand you? Maybe you are a computer engineer, but common users certainly aren't, so generally, they communicate their needs to machines in a different way. If you are interested in how it is done, get ready to peek under the hood and explore the world of language technology.

## §1. Natural language

Apart from user-friendly interfaces, there's another thing that matters in this communication: computers actually need to understand our human language. It's commonly called **natural**, as opposed to *artificial* like Esperanto or Interlingua, and *formal* languages, which include the programming ones.



Natural languages haven't been constructed on purpose but rather evolved in human communities. The rules we rely on in conversation or writing are not stated explicitly and comprehensively, it's more of a tacit convention. No wonder, it's difficult for computers to analyze "natural" texts while even humans occasionally misunderstand each other. Thus, language understanding may require a complex approach.

Current topic:

[Natural language processing](#) ...

Topic depends on:

✓ [Components of computational thinking](#) 2★ ...

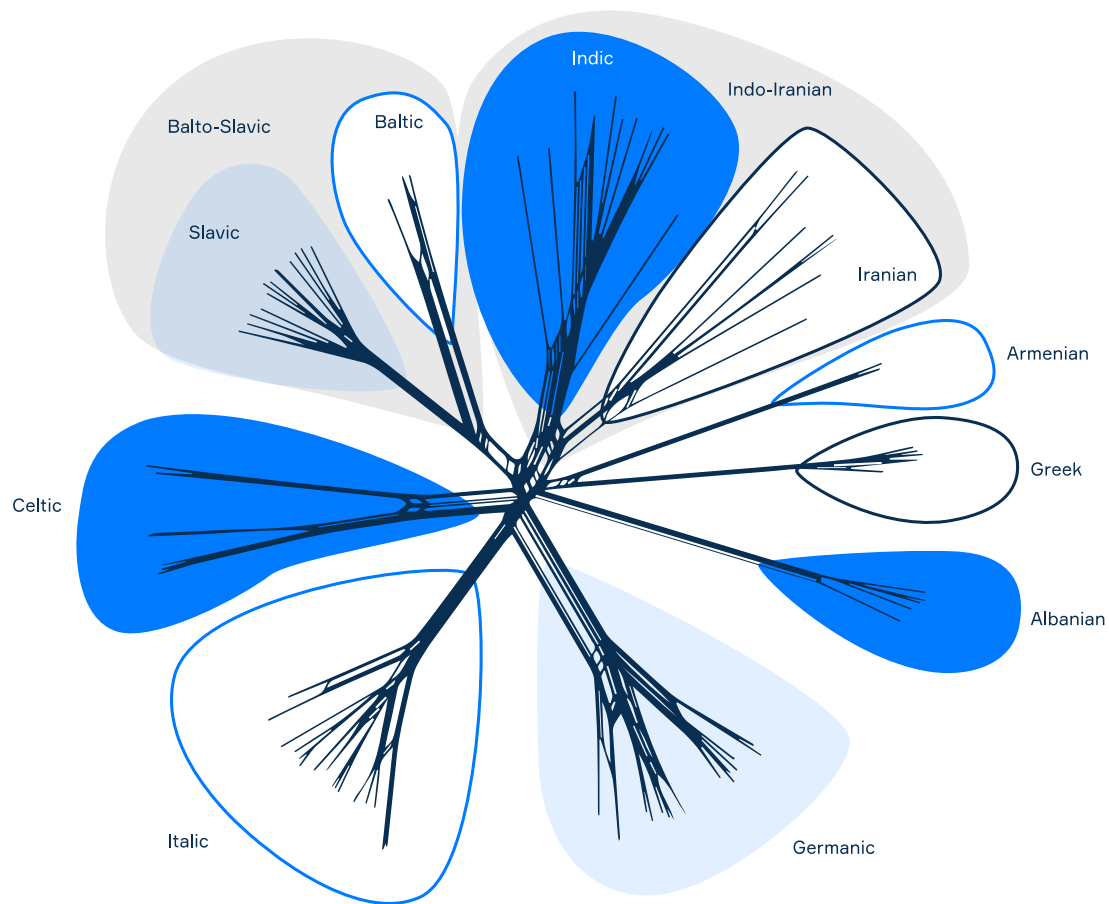
Topic is required for:

[Overview of NLTK](#) ...

[Intro to text representations](#) ...

Table of contents:

- 1 [Natural language processing](#)
- [§1. Natural language](#)
- [§2. Natural language processing](#)
- [§3. Main applications](#)
- [§4. Summary](#)
- [Feedback & Comments](#)



Indeed, the language technology currently in use has an interdisciplinary basis. Its foundations lie in linguistics, psychology, computer science, machine learning, and even ethics when it comes to biased data. Let's move on to a bit more formal introduction.

## §2. Natural language processing

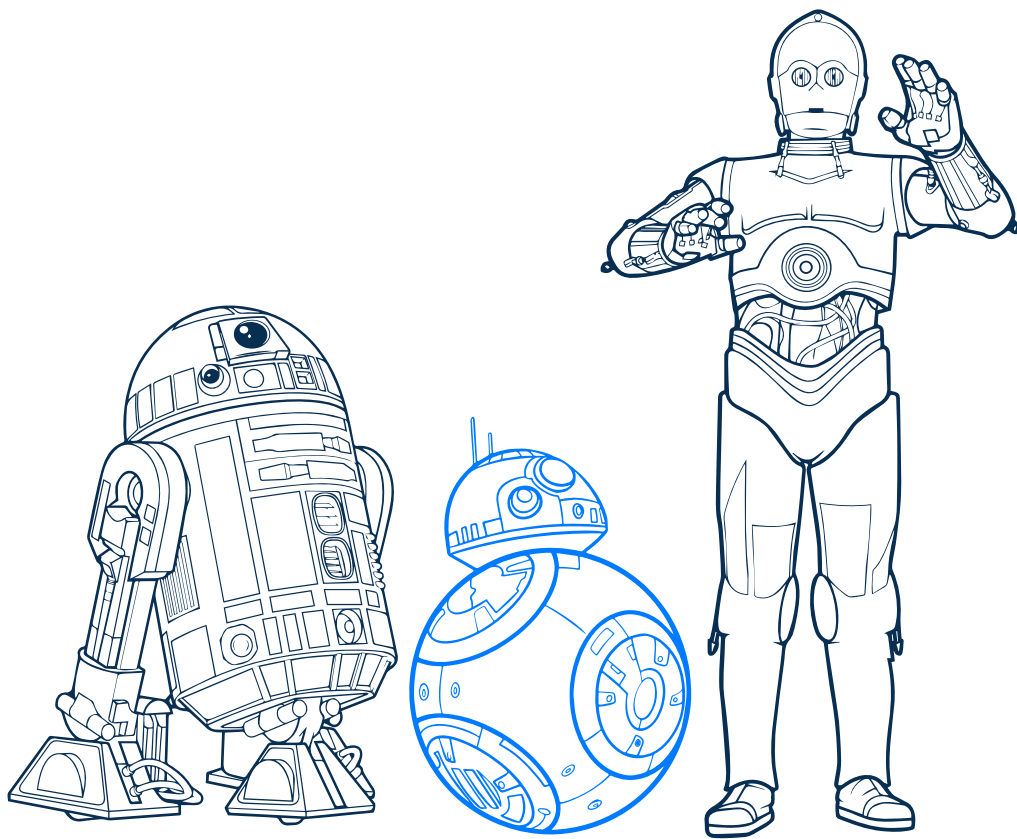
Natural language processing, or NLP for short, is a branch of artificial intelligence concerned with the interaction between computers and human languages. It involves the processing of large amounts of text data. More often than not, texts are grouped into a special collection called *corpus*, or *corpora* in the plural. Text corpora are convenient to work with, but if you want to process language efficiently, it's crucial to know how it is organized.

We share a common linguistic background since we use language every day. For computers, though, the way we choose words, apply grammatical rules and extract meaning from a message is not that obvious. Natural languages have many levels at which they can be analyzed, e.g. morphology, syntax or semantics. The linguistic analysis at multiple levels sheds light on the internal structure of a text, and this can be used to solve various real-life problems.

### §3. Main applications

In NLP, we use computers to solve language-related tasks, and it applies quite impressively to numerous areas.

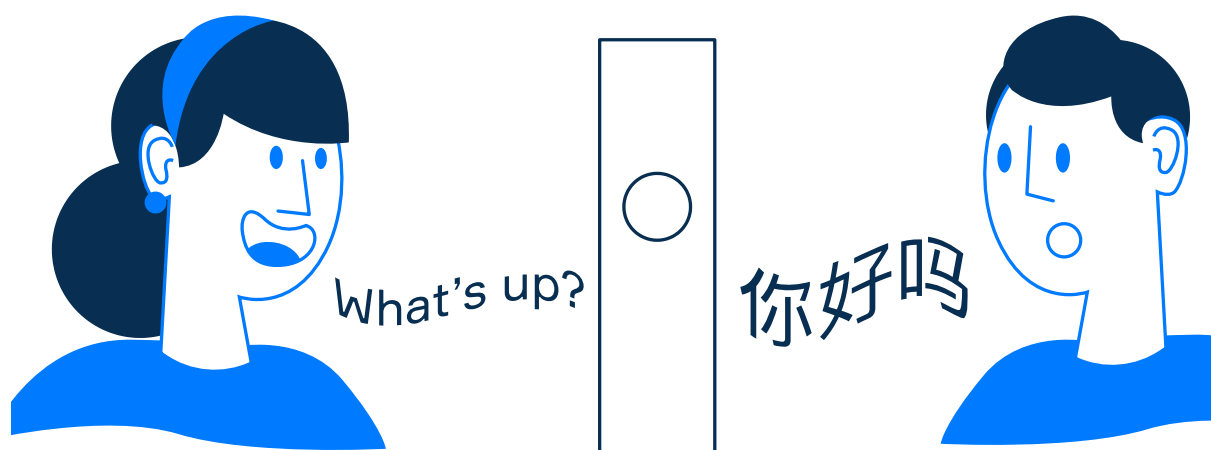
Much beloved in sci-fi, **conversational agents** have made their way into reality. Nowadays they converse with us to guide, train, support or simply entertain. Communication with **dialogue systems** usually consists of several stages. First, an agent recognizes human speech, which may be difficult due to various accents, speech impediments and, lastly, individual pronunciation characteristics. The transcribed text, in turn, should be analyzed and represented semantically, so that the computer understands you and plans its response, which will later find its voice through speech synthesis. That's quite an engaging field indeed, as diverse as human beings themselves, we would say.



The architecture of dialogue systems may include **natural language generation**. At this step, the program selects parts of data to focus on, structures the content and puts ideas back into words. You can generate replies to keep up the flow of conversation or even work with non-textual data, e.g. describe what's shown on a picture (it's called [image captioning](#)).

**Spell checkers** and **writing assistants** may seem a bit down-to-earth, but they certainly make your life easier. They can correct your misspellings, check grammar, catch punctuation mistakes and spot other issues in a text. These tools rely on statistical patterns found in a language to improve your writing.

**Machine translation** seems vitally important in some cases. Indeed, why speak just one language when your ideas can circulate the world? In terms of the task's input and output, a text is translated from the **source** language to the **target** language. Of course, there are some difficulties in between, since it's way fancier than word-for-word replacement. So you might be relieved to know that machine translation, being one of the earliest challenges in computational linguistics, now includes a vast number of techniques and approaches.



**Sentiment analysis** and **opinion mining** are generally applied to texts of a smaller size, for example, tweets or online reviews. The idea is to separate what is said (an opinion) from how it is said (an emotion, or **polarity**). It is useful for identifying trends of public opinion in social media and developing new strategies in marketing.

**Text summarization** helps create a concise summary of a document. First, the task is to identify the most relevant information in the content. Then your action plan depends on the approach: you can extract and compile the original data or transform it into a brand-new text. Either way, this would feel like a superpower in a world overloaded with data.

The above list of natural language processing applications is far from exhaustive. We will recommend resources for further study, but now it's time to sum up.

# §4. Summary

Let's highlight the main points we discussed:

- Natural language processing (NLP) is a branch of artificial intelligence that builds a bridge between computers and human languages.
- This field of research relies on findings from different disciplines, such as linguistics and computer science.
- Natural language processing tackles a wide variety of language-related problems and has many real-world applications.

Hopefully, we have advertised NLP just enough to attract you and spark your interest (that's what we aimed for!). If you are bursting with curiosity now, check out this brilliant book by Dan Jurafsky and James H. Martin, [Speech and Language Processing](#), or try learning Python, a great programming language with lots of NLP tools.

 Report a typo

**102** users liked this theory. **2** didn't like it. What about you?



Start practicing

[Comments \(5\)](#)

[Hints \(0\)](#)

[Useful links \(0\)](#)

[Show discussion](#)