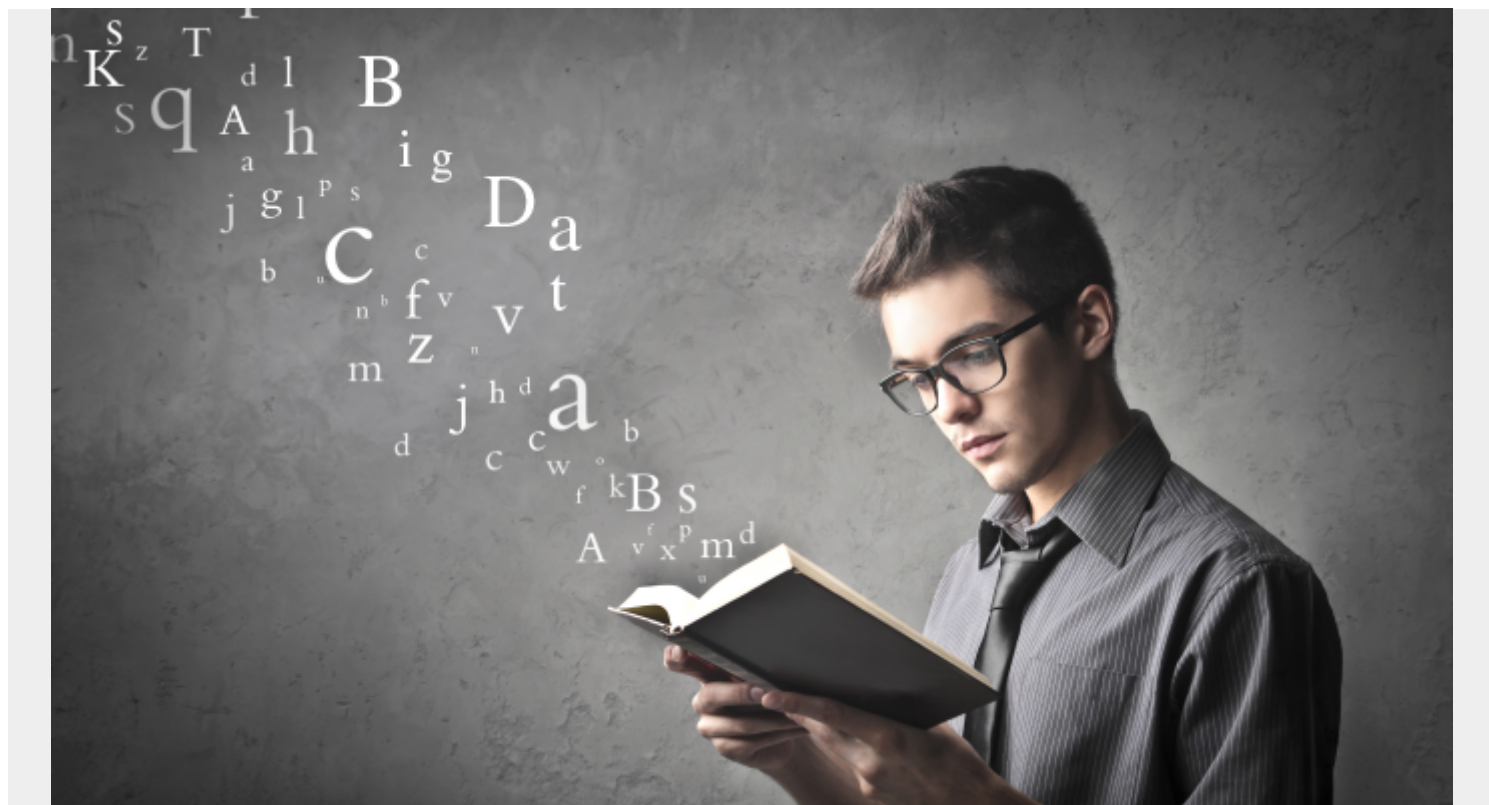# NLU VS NLP: WHAT'S THE DIFFERENCE?

In the 21st century, computers can analyze all sorts of data, providing insights and performing tasks based on the learned outcome. When that data is language, however, it is a whole different world.

Asking a computer to process real-world language is more complicated and difficult to mine in an efficient manner that offers productive results. Regardless which language a computer is learning, it must understand the syntax, semantics, discourse, and purpose of that language to make sense of it all.

To that end, we are exploring the differences between natural language understanding and natural language processing. Sometimes used interchangeably, they are actually two different concepts that do have some overlap.

Defining Natural Language Processing and Natural Language Understanding

NLP is short for natural language processing while NLU is the shorthand for natural language understanding. Similarly named, the concepts both deal with the relationship between natural language (as in, what we as humans speak, not what computers understand) and artificial intelligence.

They share a common goal of making sense of concepts represented in unstructured data, like language, as opposed to structured data like statistics, actions, etc. To that end, NLP and NLU are

opposites of a lot of other data mining techniques. But that's where the comparisons stop: NLU and NLP aren't the same thing.

Natural language processing

NLP is a decades-old field that sits at the cross-section of computer science, artificial intelligence, and, more and more, data mining. It focuses on how we can program computers to process large amounts of natural language data, such as a poem or novel or a conversation, in a way that is productive and efficient, taking certain tasks off the hands of humans and allowing for a machine to handle certain processes – the ultimate "artificial intelligence".

NLP can refer to a range of tools, such as speech recognition, natural language recognition, and natural language generation. Common NLP algorithms are often manifest in real-world examples like online chatbots, text summarizers, auto-generated keyword tabs, and even tools that attempt to identify the sentiment of a text, such as whether it is positive, neutral, or negative.

Natural language understanding

Considered a subtopic of NLP, natural language understanding is a vital part of achieving successful NLP. NLU is narrower in purpose, focusing primarily on machine reading comprehension: getting the computer to comprehend what a body of text really means. After all, if a machine cannot comprehend the content, how can it process it accordingly? (But, drawing distinct, clear insights from data that is anything but clear or distinct – often governed by only half-rules and exceptions, as is common for language– is tricky; more on that later.)

Natural language understanding can be applied to a lot of processes, such as categorizing text, gathering news, archiving individual pieces of text, and, on a larger scale, analyzing content.

Real-world examples of NLU range from small tasks like issuing short commands based on comprehending text to some small degree, like rerouting an email to the right person based on a basic syntax and decently-sized lexicon. Much more complex endeavors might be fully comprehending news articles or shades of meaning within poetry or novels.

It's best to view NLU as a first step towards achieving NLP: before a machine can process a language, it must first be understood.

The Turing Test: Language, computers, and artificial intelligence

Computers and language have gone together for decades. Natural language processing can be traced back to the 1950s, as many computer programmers began experimenting with simple language input to train computers to complete tasks. Then, in the 1960s, natural language understanding began developing out of a desire to get computers to understand more complex language input.

The most known example of artificial intelligence and language is likely the Turing Test, developed by Alan Turing in the 1950s as a way to determine whether a computer could be considered intelligent. Moral and philosophical debates aside, the Turing Test has forever created the idea – and the chase – that a human being may be able to communicate with a human being with the human knowing it's a machine, not a human.

Today, big data can sort, organize, and provide insights into data in a way that the human eye

wouldn't immediately see. But in the pursuit of artificial intelligence, we now know that it takes a lot more for computers to actually understand the data of language.

Machines can find patterns in numbers and statistics, but understanding language takes a lot more: an understanding of a language's syntax, the difference in how a language is spoken or written, pure context and definitions, shifting language patterns and ever-developing new definitions, picking up on subtleties like sarcasm which aren't inherently readable from text, or understanding the true purpose or goal of a body of content.

This is where NLP and NLU come in. Though the concepts might feel straightforward, getting into the nitty-gritty details and their applications, you may need a crash course in computer programming and linguistics to really understand how complicated these concepts can be.

Therein lies some complexity, experts say. After all, language understanding is often broken down into three linguistic levels:

- Syntax – understanding the grammar of the text
- Semantics – understanding the meaning of the text
- Pragmatics – understanding what the text is trying to achieve

Language is hard enough for a person to learn – we don't have a single way to ensure language acquisition. More complicated are the ways that languages are always shifting, adding and subtracting from a vast lexicon, incorporating ways that emails, texts, social media are used to affect language.

NLP and NLU Today

A good way to understand the difference between NLP and NLU? It is natural language understanding if you're only referring to the ability of a machine to understand our language and what we say. If it's more than that, such as making decisions based on what the text is, responding to the content as in a chatbot talking with a human, it likely encompasses the wider concept of natural language processing.

Unfortunately, understanding and processing natural language isn't as simple as providing a huge set of vocabulary and training your machine on it. Successful NLP must blend techniques from a range of fields: language, linguistics, data science, computer science, and more.

This is why NLP has been so elusive – an academic advance in a small part of NLP may take years for a company to develop into a successful tool that relies on NLP. Such a product likely aims to be effortless, unsupervised, and able to interact directly with customers in an appropriate and successful manner. Plus, companies who are working on these projects might struggle to find the right combination of talent, as a person may specialize in one or two of these fields.

Just as we don't have a single method or process for teaching or acquiring human language, we haven't yet cracked the code on how machines can understand and respond to such a human mystery – but companies and academics are investing more in getting there soon.