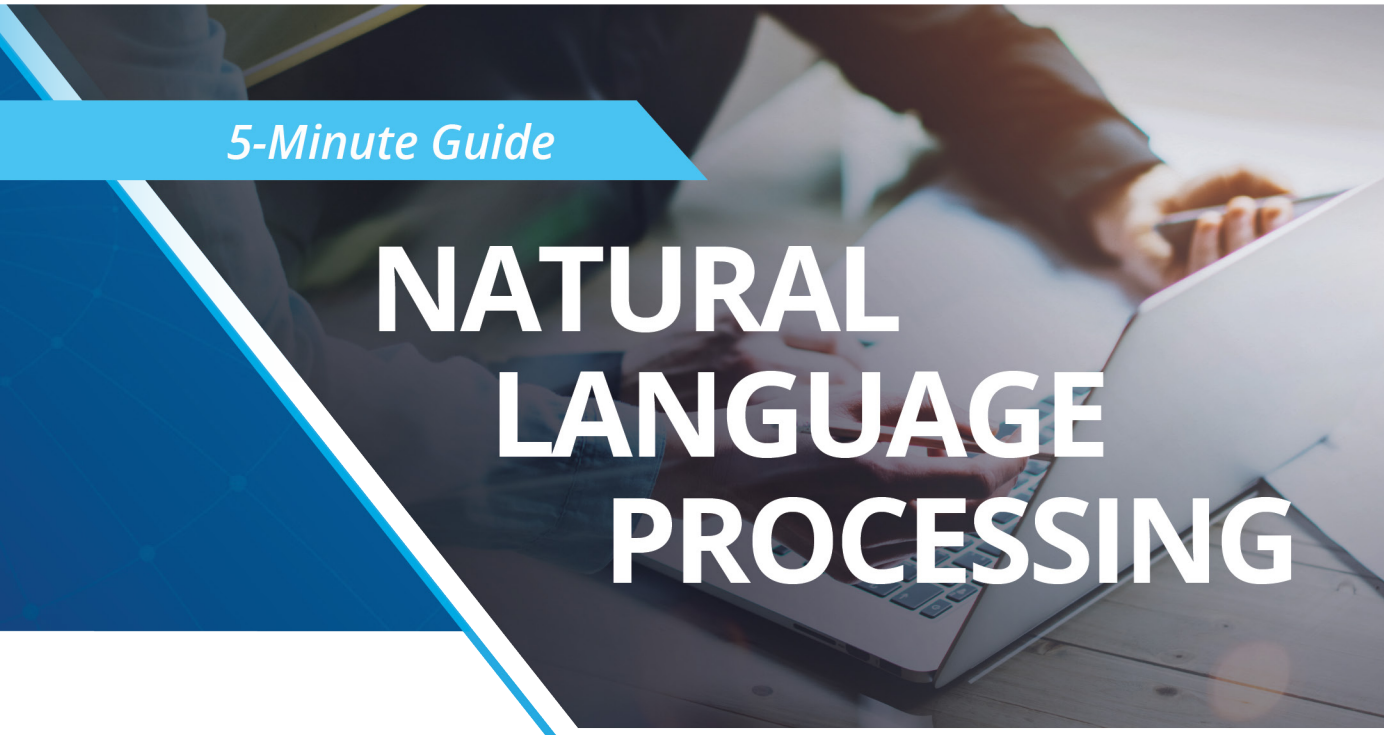
5-Minute Guide

NATURAL LANGUAGE PROCESSING



ATTIVIO®

ATTIVIO®

5-Minute Guide to

Natural Language Processing

What I Meant Was

A 2016 paper by researchers at Google explored the value of context and conversational history in determining the correct response to a query.[1] Early in the paper, the authors offered a simple example. The query, "Do you plan to watch the bulls?" had two possible contexts. In the first, respondents were residents of a village. In the second, they lived in Chicago. Possible responses were, "I am planning to visit the farm soon," and "I am going to watch them on TV."

When a search query is ambiguous, Google uses context to determine what you mean. NLP and text analytics help search engines put queries in a relevant context.

Not surprisingly, residents of Chicago more often said they were going to watch the bulls on TV rather than plan a visit to a farm. The opposite was true for residents of a village.

It's easy to use this example as a proxy for the value of natural language processing (NLP) and text analytics in AI-powered search. Type "I want to watch the bulls," into the Google search box. The results will depend on where you live and your prior search history—in other words, context.

When a search query is ambiguous, Google uses context to determine what you mean. NLP and text analytics help search engines put queries in a relevant context.

NLP and Text Analytics: What's the Difference?

The techniques of NLP and text analytics overlap a great deal. The differences mainly lie in the problem that each tries to solve. In the search world, natural language processing analyzes user inputs (queries) to understand their intent.

1 https://arxiv.org/pdf/1606.00372.pdf

Pattern recognition is where NLP and text analytics overlap quite a bit. NLP uses pattern recognition to parse a query. Text analytics uses pattern recognition to find answers to a query.

It allows a user to communicate with a machine in a way that is natural for the user, which, of course, is not natural for the machine.To accomplish this, NLP operates on data so that a computer can understand a document—and the relationships it may infer—in the same way a user understands it. It's wise here to remember that infer means to "make an educated guess." This is where NLP and text analytics use many of the same methods.

Text analytics uses a variety of techniques to extract information from text data and then classify, group, cluster, and mine that information for concepts and patterns as they're expressed via key words and phrases.

Pattern recognition is where NLP and text analytics overlap quite a bit. NLP uses pattern recognition to parse a query. Text analytics uses pattern recognition to find answers to a query.

When a computer executes text analytics to parse a sentence, it pulls out or "tokenizes" words and parts of speech so that it can map them to other sentences.

AI-powered search solutions provide increasingly relevant answers to implicit and explicit queries by combining self-learning technologies such as indexing, natural language processing, and machine-learning.

How Does AI-Powered Search Use NLP and Text Analytics?

A simple way to think about NLP and text analytics is that text analytics happens as information is indexed, tagged, and entities extracted and annotated. NLP is an interaction that occurs when analyzing a user query and trying to match that to the work done by the text analytics engine. NLP is the human face of text analytics.

NLP ensures that when you ask for directions to a restaurant, the first result is a map, not a list of restaurants. NLP understands that "how do I get to" is the semantic equivalent of "where is" and "I can't find." And if it doesn't understand a query, you'll often see, "Did you mean?" followed by a list of alternate queries. That's much more satisfying for the user than a "no match" response.

Faceted navigation provides multiple filters, one for each aspect of the information. It also provides a structure to help users understand the information space and give them ideas about what is available and how to search for it. [2]

In the enterprise, A-I powered search helps individuals find answers to questions they encounter doing their jobs. It can also suggest relevant information sources without a query based on a user's prior search history—a new report, for example. Using artificial intelligence, such as machine learning and pattern matching, the search application gets to "know" its users. NLP provides a semantic interface between the user and the application.

The more specific the query, the easier it is for the NLP interface. To narrow broad queries, NLP employs faceted search, which is a way for NLP to break down a complex query into manageable components.

The Quest for an Intelligent Search Interface

The authors of the Google paper cited earlier in this guide point out, "Designing conversational systems is a challenging task and one of the original goals of Artificial Intelligence." [3] Like machine learning, another foundational component of AI, NLP and text analytics get better as they encounter more data from which to learn.

As algorithms and compute power continue to improve and increase, the accuracy and responsiveness of AI-powered search interfaces will follow suit. More and more often, they will understand what you meant.

ATTIVIO®

Attivio is the leading AI-powered search & insight platform company. Our mission is to put search at the core of every enterprise. By easily integrating every data source into one always-learning platform, Attivio enables Fortune 500 enterprises to answer the most complex questions, by providing immediate insight across all of their information. Our solutions power innovation, drive operational efficiencies, and transform business outcomes at scale, and feature industry-leading capabilities including natural language processing, machine learning, analytics, and knowledge graphing.

For more information, **please visit www.attivio.com.**

2 https://www.nngroup.com/articles/filters-vs-facets/
3 https://arxiv.org/pdf/1606.00372.pdf