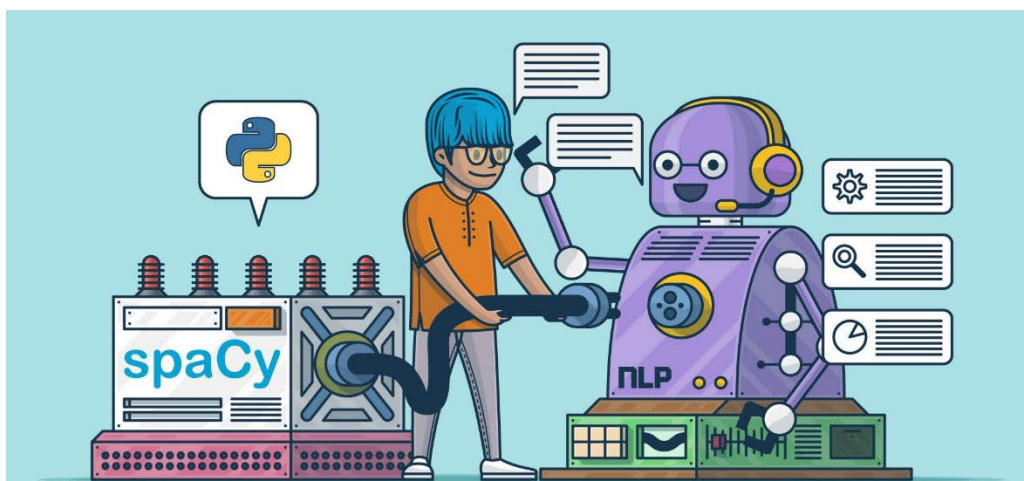


## Software Lab:

# Natural Language Processing (NLP): legal documents analysis

### Setting

The information included in the different legal documents, norms, and client requirements is described in unstructured text documents. Automatic digitalization of these documents requires a manual parsing and extraction of nouns, verbs, adjectives..etc. Additionally, to get more insights on the meaning formed by these extracted parts, different patterns are usually detected and analyzed.



Picture from: <https://realpython.com/natural-language-processing-spacy-python/>

Natural language processing (NLP) and text mining are machine-based solutions for text analysis. Through NLP techniques, it is possible to split the sentences to different tokens and then infer different meanings from them. There are many advanced NLP techniques implemented and available today. Those techniques make it easy to measure the text similarity, dependency as well as produce a text summary.

### Task

- Explore the different techniques of text extraction and summarization.
- Evaluate their application on legal documents.

The final prototype should be capable of extracting nouns, verbs, and the relationships between them.

### Supervisor

Jimmy Abualdenien, jimmy.abualdenien@tum.de, Chair of Computational Modeling and Simulation