Natural Language Processing is a branch of Artificial Intelligence that focuses on understanding and generating natural languages, where a natural language is one spoken by humans in conversation with each other as opposed to more formal languages such as context free grammars. Natural language understanding and natural language generation generally advance together and share categories of approaches (rule-based, probabilistic, and machine learning), but they are separate problems with unique solutions and approaches. Natural language processing has become widely used in modern technology, from personal assistants like Alexa and Siri to less obvious uses such as automated sentiment analysis and accessibility software.

Rules-based natural language processing generally attempts to model natural languages using more formal languages, such as regular expressions or context free grammars. While simplistic, rules-based approaches are relatively easy to implement, cheap to execute, and do not require vast amounts of data. Use-cases for rules-based approaches include grammar checking, generation of grammatically correct sentences, modification of tense or plurality of words, etc.

Statistical and probabilistic approaches base their understanding and generation of language on the probability of words, phrases, or sentences to exist near each other. While these models require data to function, they typically only require enough to calculate accurate probabilities, for example, the likelihood that the phrases “drive a car” vs. “drive a plane” are to appear, and the algorithms involved in generating those probabilities are well-understood and optimized. Statistical and probabilistic approaches outperformed rules-based approaches in a number of areas, for example, translation, which leverages both natural language understanding and generation. Statistical and probabilistic models are more capable of creating correct translations of phrases where the language, strictly speaking, is ambiguous. The phrases “little dog” and “little brother” would both be recognized as an adjective attached to a noun by a rules-based translation program, but the meaning of the word “little” describes different properties of the respective nouns, a distinction that cannot be properly translated without context on the wider use of both the English language and the output language.

Deep learning approaches leverage modern neural networks and truly massive amounts of data to train models capable of understanding the context of the text being processed, rather than just the correctness of the language or the most likely configuration of words. As deep learning approaches to natural language processing are a developing science, we do not fully understand how they may be used. We do know that models capable of processing natural language in context to a conversation or to a larger body of text has the potential to improve results of several common applications of natural language processing, including search, chatbots, and digital assistants.

My personal interest in natural language processing is its potential to revolutionize how we think about search. Currently, services like Google and Bing offer search functions that, given a search string, find the page we most likely want to see. As the ability of these models to

understand context including and beyond immediate conversation grows, we may be able to ask questions in natural language and receive factual generated answers, the language of which is tailored to our individual level or need, and which can be fact-checked through a provided references list.