# **Why do computers have difficulty with NLP?**



1. Computers have mostly been dealing with structured data. This is data that's organized, indexed and referenced, often in databases.
2. In NLP, we often deal with unstructured data. Social media posts, news articles, emails, and product reviews are examples of text-based unstructured data.
3. To process such text, NLP has to learn the structure and grammar of the natural language. Importantly, 80% of enterprise data is unstructured.
4. Human languages have plenty of complexities such as ambiguous phrases, colloquialisms, metaphors, puns, or sarcasms.
5. The same word or text can have multiple meanings depending on the context.
6. Language evolves with time. Worse still, we communicate imperfectly (spelling, grammar or punctuation errors) but still manage to be understood. These variations, so natural to human communication, are complex for computers.
7. Ambiguities in natural languages can be classified as lexical, syntactic or referential.
8. When the source of information is speech, more challenges arise: accent, tone, loudness, background noise or context, pronunciation, emotional content, pauses, and so on.

**Some examples of the complexities of English:**

1. Consider the sentence, "One morning I shot an elephant in my pajamas". The man was in his pajamas but grammatically it's also correct to think that the elephant was wearing his pajamas.
2. Likewise, a person may say, "Listening to loud music slowly gives me a headache". Was she listening to music slowly or does the headache develop slowly?
3. A more confusing example is, "The complex houses married and single soldiers and their families". Confusion arises because we may initially interpret "complex houses" as an adjective-noun combination. The sentence makes sense only when we see that "complex" is a noun and "houses" is a verb. NLP addresses this via part-of-speech tagging.
4. Consider this one, "John had a card for Helga, but couldn't deliver it because he was in her way". Was John in Helga's way? In fact, "he" refers to an earlier reference to a third person. NLP calls this coreference resolution.
5. "The Kiwis won the match" is an example that requires context to make sense. New Zealand nationals are referred to as "Kiwis", after their national bird. Natural language is full of metaphors like this.