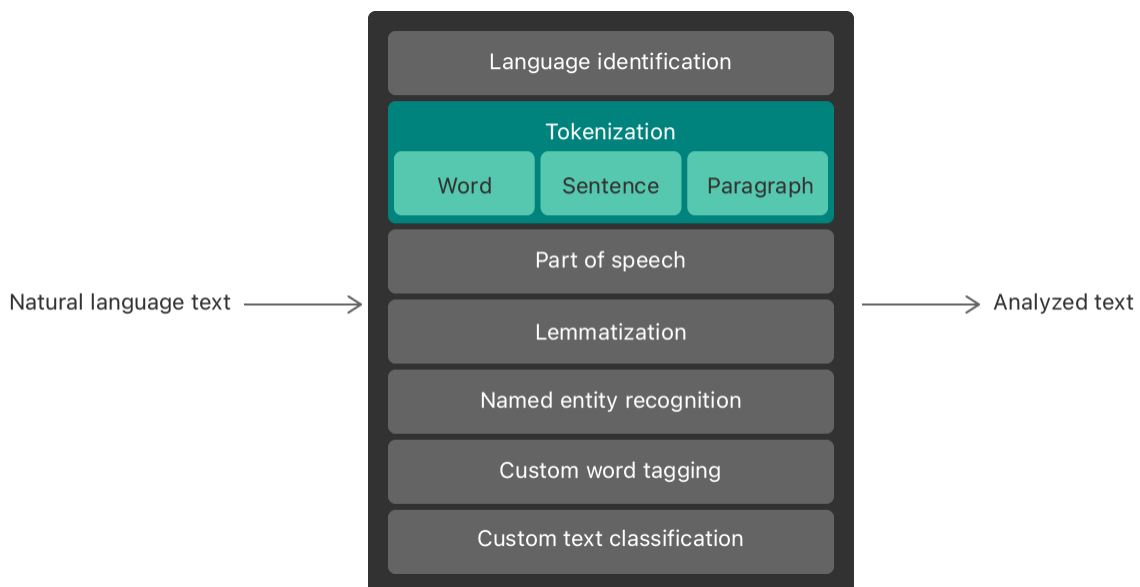Framework

Natural Language

Analyze natural language text and deduce its language-specific metadata.

iOS 12.0+ | iPadOS 12.0+ | Mac Catalyst 13.0+ | macOS 10.14+ | tvOS 12.0+ | visionOS 1.0+ | watchOS 5.0+

Overview

The Natural Language framework provides a variety of natural language processing (NLP) functionality with support for many different languages and scripts. Use this framework to segment natural language text into paragraphs, sentences, or words, and tag information about those segments, such as part of speech, lexical class, lemma, script, and language.

Use this framework to perform tasks like:

- *Language identification*, automatically detecting the language of a piece of text.

- *Tokenization*, breaking up a piece of text into linguistic units or tokens.

- *Parts-of-speech tagging*, marking up individual words with their part of speech.

- *Lemmatization*, deducing a word's stem based on its morphological analysis.

- *Named entity recognition*, identifying tokens as names of people, places, or organizations.

You can also use this framework with Create ML to train and deploy custom natural language models. For more information, see <u>Creating a text classifier model</u> and doc:creating-a-word-tagger-model.

Topics

Tokenization

📄 Tokenizing natural language text

Enumerate the words in a string.

`class` `NLTokenizer`

A tokenizer that segments natural language text into semantic units.

Language identification

📄 Identifying the language in text

Detect the language in a piece of text by using a language recognizer.

`class` `NLLanguageRecognizer`

The language of a body of text.

`struct` `NLLanguage`

The languages that the Natural Language framework supports.

Linguistic tags

📄 Identifying parts of speech

Classify nouns, verbs, adjectives, and other parts of speech in a string.

📄 Identifying people, places, and organizations

Use a linguistic tagger to perform named entity recognition on a string.

`class` `NLTagger`

A tagger that analyzes natural language text.

Text embedding

📄 Finding similarities between pieces of text

Calculate the semantic distance between words or sentences.

class `NLEmbedding`

A map of strings to vectors, which locates neighboring, similar strings.

Contextual embedding

class `NLContextualEmbedding`

A model that computes sequences of embedding vectors for natural language utterances.

struct `NLContextualEmbeddingKey`

Contextual embedding keys.

struct `NLScript`

The writing scripts that the Natural Language framework supports.

Natural language models

📄 Creating a text classifier model

Train a machine learning model to classify natural language text.

📄 Creating a word tagger model

Train a machine learning model to tag individual words in natural language text.

class `NLModel`

A custom model trained to classify or tag natural language text.