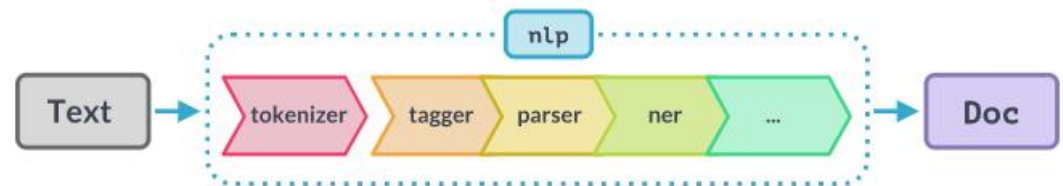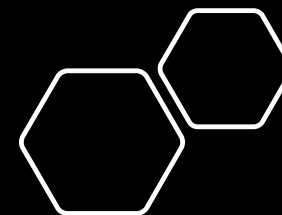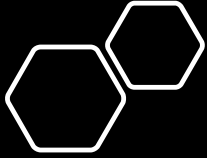# Pipeline

- Você pode personalizar o Pipeline para ter apenas as etapas que você precisa

- Etapas podem ser prontas (buitl in) ou personalizadas

STRING NAME	COMPONENT	DESCRIPTION
tagger	Tagger	Assign part-of-speech-tags.
parser	DependencyParser	Assign dependency labels.
ner	EntityRecognizer	Assign named entities.
entity_linker	EntityLinker	Assign knowledge base IDs to named entities. Should be added after the entity recognizer.
entity_ruler	EntityRuler	Assign named entities based on pattern rules and dictionaries.
textcat	TextCategorizer	Assign text categories: exactly one category is predicted per document.
textcat_multilabel	MultiLabel_TextCategorizer	Assign text categories in a multi-label setting: zero, one or more labels per document.
lemmatizer	Lemmatizer	Assign base forms to words.
morphologizer	Morphologizer	Assign morphological features and coarse-grained POS tags.
attribute_ruler	AttributeRuler	Assign token attribute mappings and rule-based exceptions.
senter	SentenceRecognizer	Assign sentence boundaries.
sentencizer	Sentencizer	Add rule-based sentence segmentation without the dependency parse.
tok2vec	Tok2Vec	Assign token-to-vector embeddings.
transformer	Transformer	Assign the tokens and outputs of a transformer model.

Modo Pipe

- `documento = nlp("texto")`
  - `Cria um objeto Doc com todo o texto processado`

- Modo Pipe
  - Python Generator
- Suporta Multiprocessamento e tamanho de batch personalizados