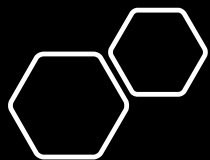


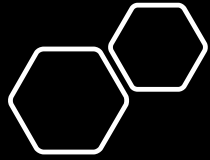
Similaridade

- Similaridade VS Matcher
- Similaridade
 - Retornar um valor (métrica de similaridade)
- Matcher
 - Busca um padrão, retorna se encontra (semelhante a expressão regular)



Similaridade

- Baseada em contexto
- Precisa do modelo Large do Idioma
- Métrica é o cosseno, valor entre zero e 1
- Utiliza representação vetorial das palavras (word embeddings) (word2vec)
- Pode ser aplicado entre Docs, spans e tokens



Melhora de Performance

- Sentenças menores (span)
- **Utilizar outra função de similaridade (ex. `sense2vec`)**
- Remover Stopwords, pontuação, pronomes