

Digital humanities

FREE ACCESS

at your institution

All teachers, students and researchers at your institution can log in with their institutional login details at no charge. Funded by ELEXIS.



Sketch Engine is software for analysing large amounts of texts (text corpora) using linguistic criteria. It contains data of billions of words in 90+ languages and lets users upload their own data for analysis.

compare documents, books, papers, articles or websites

How are two texts or very large collections of texts different? How can I compare them effectively without reading them?

online texts Sketch Engine downloads web pages or the whole websites automatically and compares them. It displays keywords and multiword expressions that are typical of one website compared to the other. Such a list helps to understand how the two sources differ in the coverage of topics.

Learn bit.ly/2V5cMCV offline texts The same functionality can be applied to files the user uploads such documents, stories, descriptions, ► Learn bit.ly/2PlpqBS press releases, literary works or contracts and agreements.

analyse the text and uncover the meaning

What is typical or frequent in my texts or my database? How frequent is it?

Use the wordlist to generate all kinds of frequency lists: all words, nouns, verbs, adjectives, words with a capital letter... Use the word sketch to identify the most typical word combinations. Use the keywords to identify words define their topic, message or subject area.

This is a great way of getting insight into your texts and understanding the main ideas or concepts of short documents as well as large text collections.

► Learn bit.ly/2ISV277



focus on context

It is the context that shapes the meaning of each word. Use the word sketch to see the typical combination or use the collocation tool in the concordance to understand the contexts in which the word or phrase appears.

- ► Learn bit.ly/2V3xUsY
- ► Learn bit.ly/2PGb4So

apply linguistic criteria

Any texts you upload or collect from the web will be automatically annotated (tagged) for parts of speech so statistics can be generated for only adjectives, nouns etc. All texts will also be lemmatized. This means that Sketch Engine is able to treat all the different word forms as one word. Extremely useful in Spanish, French, German, Russian, Polish and many other languages with rich morphology.

learn

Quick Start Guide www.sketchengine.eu/guick-start-guide Face-to-face training bit.ly/2Vr8K3m Training videos bit.ly/2HY0gys

log in

Use your institutional login details on www.sketchengine.eu



