

## KWIC (Key Words In Context)

The KWIC Python routine (KWIC\_GUI.py) offers users the option of computing word co-occurrences.

### The co-occurrences computation routine

The routine creates the co-occurrences data from the CoNLL table and saves the results as a csv file.

When creating the csv file the Python script will export co-occurrences of words appearing to the left, right or on both sides of a selected word for different windows sizes: 1 word, 2 words, 3, 4, 5, ... 9, 99 for sentence, 999 for document.

### Searching the KWIC csv table

Once the KWIC table is computed, the table can be searched for specific values. For example, consider a document with the following sentence: “A few weeks ago a house and a warehouse were destroyed by fire in Hinesville, and all the circumstances pointed to its being the work of an incendiary.”

If the user typed the search term: “fire” and the position index 2 s/he would see that the words “destroyed” and “Hinesville” occur respectively at position -3 and +3 with respect to the word “fire” in the document.

The user can specify any position index from 1 to 10 or the special context “Sentence”. All the words co-occurring with the search term in the same sentence will be displayed in the output table if the user selects the “Sentence” context.

### The co-occurrences visualization routine

The routine gives the user various options:

1. Specify the **co-occurrences KWIC csv file to be used** (or create a new **KWIC** file);
2. Specify the **search window size** for the word co-occurrences to be visualized (1, 2, 3.. words)
3. Specify the **search window position** (left, right, or both sides of the selected word)

Co-occurrences of the selected word will be visualized as network graphs using Gephi.