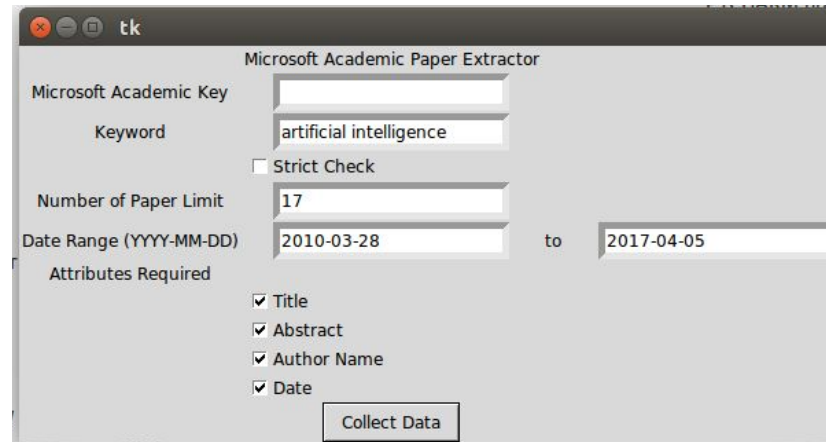# Microsoft Academic Extractor Tool Documentation

- Creating Microsoft Academic API Key:
    - Microsoft Academic Extractor requires an API key to access and collect data.
    - Process to create a key is available at Microsoft official website on: https://labs.cognitive.microsoft.com/en-us/subscriptions

- GUI of the tool:

- Options available in the tool:
    - Microsoft Academic Key - The key created above needs to be entered in the text box here. This key validates the user and gives access to the API to collect the data.
    - Keyword - The keyword for which we need to collect the papers (data).
    - Strict Check - When marked "checked", it searches for papers matching the exact keyword. By default, it is left "unchecked". When "unchecked", it searches for all papers matching with any word in the keyword.
    - Number of Paper Limit - This keeps an upper cap on the number of papers to be collected.
    - Date Range (YYYY-MM-DD) - Date range for which the papers need to be searched. Date should be entered in the form of YYYY-MM-DD.
    - Attributes Required - It provides option to choose from the four main attributes of a paper i.e. Title, Abstract, Author Name, Date. All 4 attributes are marked checked by default. Abstract is collected in the form of "Bag of Words" i.e. it contains list of words along with their index/position in the abstract.

    - Collecting Data:
        - All text boxes available in the GUI are mandatory to be filled.
        - After entering all options, "Collect Data" button needs to be clicked.

- - ○ Depending on the number of papers required, it may take from few seconds to minutes or more to collect the data.

  - Messages/Warnings:
    - ○ "Success","Data has been collected in the file name paper_data.csv":
      - If papers/data is collected successfully.
    - ○ "Warning","Please enter correct input":
      - If all text boxes are not filled before clicking on "Collect Data" button.
      - If no relevant papers are found for the entered search options.

An example file with the name "paper_data.csv" is attached for the example screenshot above.