

Text Splitting

Assignment: Text Splitter Tool (CLI + Streamlit)

 by mukesh kumar



Split by sentence



Split by sentence

Cras tunc sit erat solum erat.
 porttitor etiam sed ut.
 yestitio, paratititio.

Split by paragraph



Split by paragraph

Castertititio paratitio.
 yestitio, paratitio.
 porttitor solum erat.
 pascititio.

Split by custom delimiters



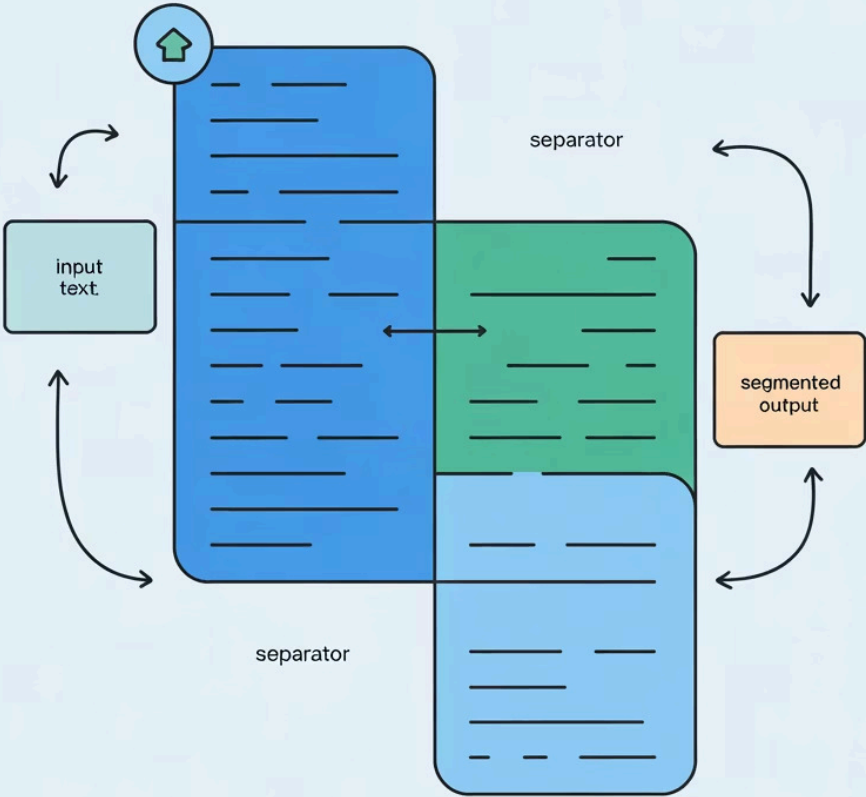
Custom delimiters

Casertitio solum erat.
 porttitor, paratitio.
 yestitio, paratitio.
 porttitor, paratitio.
 porttitor, paratitio.

Text segmentation dladxcing sepmenationg

Lone eorpcos pase deroth ourstune adds nexecctued on
annet des lorahanod attloe liera t inpo terrnesohee dtonun.

Get started



Objective

Practice string splitting and understand how text can be broken into parts using delimiters.

Part A – Command Line Version

Ask the user to enter a paragraph.

Ask how they want to split it:

- By **sentence** (i.e., .)
- By **word** (i.e., space)
- By custom character

Command Line Output Requirements

Show the list of resulting items and their count.

Ask the user to optionally search for a word and show whether it exists in the list.

Part B – Streamlit Version

Use a text area to input a paragraph.

Dropdown to select delimiter: **Sentence (.)**, **Word (space)**, **Custom (user input)**

Display the resulting items in a list below.

Streamlit Display Requirements

Show:

- Total number of items
- Optional search box to look for a word in the list

Implementation Features



Text Input Methods

Support both command line input and web-based text area for maximum flexibility



Multiple Delimiters

Handle sentence splitting by periods, word splitting by spaces, and custom character splitting



Search Functionality

Allow users to search for specific words within the split text results

Expected Outcomes

This assignment will help students understand string manipulation, user input handling, and the differences between command line and web-based interfaces. Students will gain practical experience with text processing and learn how different delimiters can break text into meaningful components.