

# IS344 Information Retrieval

## Homework #2

**Due April 19<sup>th</sup> 2024 [6 Marks]**

Extend the project of the HW#1 To have the following two features:

### **1- BI word index [3 Marks]**

You should enter each word alone and with an underscore with the next word

For example

the doc: automated retrieval of specific information

will be entered

automated, retrieval, automated\_retrieval, of, retrieval\_of, specific,

of\_specific, information, specific\_information

the query: will be

“specific information” and will return the document

i.e. if in quotes insert underscore between the 2 words

Bonus: mix single words with phrase

i.e the query: *automated “specific information”*

will return the last document

however if a document contains the three word but separated will not be returned

### **2- include a positional index such that**

given doc1: Cairo University Zayed City is a new CU branch

and doc2: Zayed attending in AinShams University in Cairo

The query “Cairo University Zayed City” would return doc 1 but not doc 2