

F74056247 曾大瑋

Programming Language HW2 報告

執行環境

python 3.7.3

jupyter-notebook 5.7.8

建議使用 jupyter-notebook 開啟程式碼

執行步驟

開啟 jupyter-notebook 後選擇此篇.ipynb 開啟

程式碼說明

雖然可能要用到 Regular expression 結果都沒用到

建議使用 jupyter-notebook 開啟程式碼

getPaperNumber(author): int

取得此 author paper 的搜尋結果數量

取下關鍵字時,showing ...of 多少 result 的數字

用來算頁面,爬每一頁

verifyAuthor(element,author): bool

驗證此 paper 是否為 author(剔除中間名等)

getPaperResult(author): int

取得此 author paper 的真實數量

這裡是經過 verifyAuthor 的結果,所以數量會比 getPaperNumber 較少

getPaperYear(author): list,list

第一個 list 是年份,[2012,2013,2014]等等的 list

第二個是數量[3,4,10]這樣的 list

取得 author 每年 paper 數

getPaperAuthor(author):pandas.core.series.Series

取得 author 所有 paper 中的 co-author

這裡是指 `getPaperResult` 中 `co-author` 不包含自己的其他人
這裡用 `pandas` 的 `dataframe` 純粹是因為有排序好了,方便使用