

中研院斷詞後去標點符號程式說明

注意！

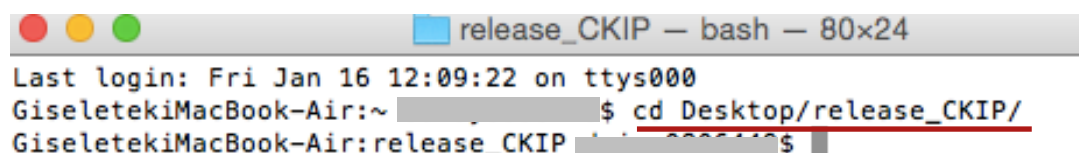
1. 此程式適用於Macintosh系統。
2. 所有txt檔的編碼均須為UTF-8。
3. 每次執行，輸出的資料夾內容會被覆蓋。

步驟一、整理文本，放入正確資料夾中

1. 可將中研院斷詞結果去表點符號的程式，在「release_CKIP」的資料夾中。
2. 將「release_CKIP」的資料夾放在桌面(Desktop)。
3. 執行去標點的動作前，將要斷詞的所有文本放在同一個資料夾(ex: input_dir)，這個資料夾可以自己設定，只要在之後的指令做修改即可。

步驟二、打開終端機(Terminal)，執行程式

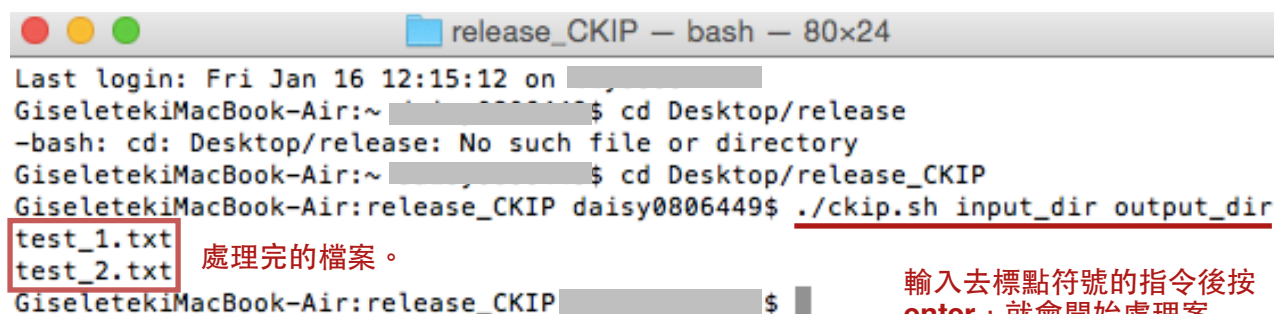
1. 帶開電腦的終端機，先輸入指令進入「release_CKIP」的資料夾中：
`cd Desktop/release_CKIP`



```
release_CKIP — bash — 80x24
Last login: Fri Jan 16 12:09:22 on ttys000
GiseletekiMacBook-Air:~$ cd Desktop/release_CKIP/
GiseletekiMacBook-Air:release_CKIP$
```

輸入進入資料夾的指令
後，按**enter**。

2. 進入「release_CKIP」的資料夾中後，再輸入執行去標點符號的指令：
`./ckip.sh input_dir output_dir`



```
release_CKIP — bash — 80x24
Last login: Fri Jan 16 12:15:12 on 
GiseletekiMacBook-Air:~$ cd Desktop/release
-bash: cd: Desktop/release: No such file or directory
GiseletekiMacBook-Air:~$ cd Desktop/release_CKIP
GiseletekiMacBook-Air:release_CKIP daisy0806449$ ./ckip.sh input_dir output_dir
test_1.txt
test_2.txt
GiseletekiMacBook-Air:release_CKIP$
```

處理完的檔案。

輸入去標點符號的指令後按
enter，就會開始處理案

3. 處理完的檔案會放在output_dir 的資料夾裡，檔名會對應。
範例中「test_1.txt」、「test_2.txt」，處理完的檔案會放在output_dir的資料夾中，每次執行此資料中的檔案都會被覆蓋過去。

FAQ：如有出現：permission denied: ./<xxx>.sh
請下指令：chmod 755 ckip.sh