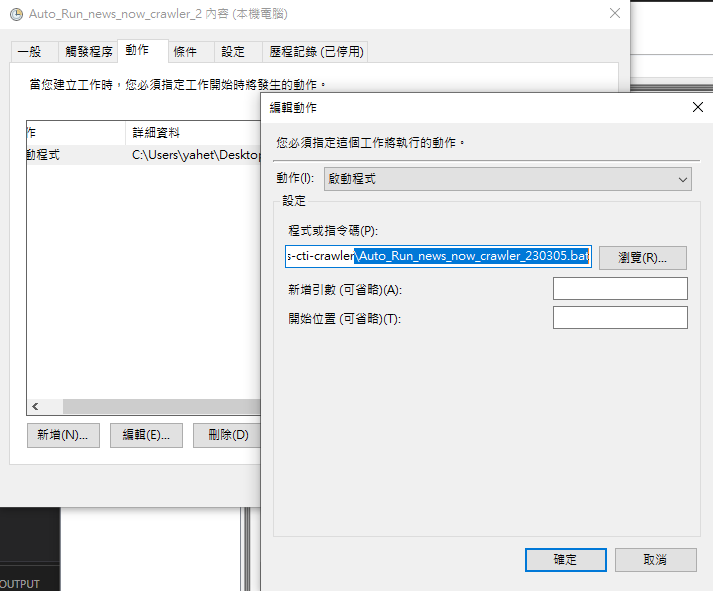
注意 bat 檔路徑！！　如有移動檔案，一定要修改 bat 檔內容

記得去工作排程器設定，並執行一次進行檢查！



一張含有 文字 的圖片

自動產生的描述

**cmd bat檔內容**

cd /d C:\Users\yahet\Desktop\ids\_cti\_web\ids-cti-crawler

call conda activate py37-ioc-extractor

call python ./news\_now\_crawler\_230305.py -d 1

REM pause

----------------------------------------------------------------------------------

python ./news\_now\_crawler\_230305.py -d [爬取的次數(一次八小時，因為工作排程器)]

----------------------------------------------------------------------------------



依照時間進行的爬蟲程式：news\_now\_crawler\_230305\_Specific\_Time.py

“news\_now\_crawler\_230305\_Specific\_Time.py” 指令

-d 一次一天

conda activate py37-ioc-extractor

python ./news\_now\_crawler\_230305\_Specific\_Time.py -d [天數] -t [Unixtimestamp]

Unixtimestamp 如167XXXXXXX (從資料庫獲取)

**SQL查詢**

* 查詢日期區間

SELECT \* FROM newsnow\_security WHERE post\_date BETWEEN '2023-02-02' AND '2023-02-03'

* 查詢重複值

SELECT title, COUNT(\*) FROM newsnow\_security GROUP BY title HAVING COUNT(\*) > 1;

* 刪除重複值

CREATE TEMPORARY TABLE temp\_table AS

SELECT MIN(id) AS keep\_id

FROM newsnow\_security

GROUP BY trans\_url; 根据需要查看重複值的row

DELETE FROM newsnow\_security

WHERE postID NOT IN (SELECT keep\_id FROM temp\_table)

DROP TEMPORARY TABLE temp\_table