

# 政府電子採購網爬蟲使用說明

---

## Introduction

本程式為政府電子採購網爬蟲工具，透過關鍵字篩選標案名稱與機關名稱，自動擷取並彙整公開標案資訊，輸出為 Excel 格式，便於後續資料分析與追蹤。

## 一、準備步驟

### 1. 複製專案資料夾

- 將整個專案資料夾複製到桌面。

### 2. 安裝 Anaconda

- 資料夾裡的 **Anaconda3-2023.07-Windows-x86\_64.exe** 為 Anaconda 安裝檔，也可雙擊此檔按安裝，若已安裝可跳過此步驟。
- 安裝完成後，即可使用 **conda** 指令進行虛擬環境與套件管理。

### 3. 關鍵字設定

- 程式會根據使用者在 **spidar.py** 中設定的關鍵字進行篩選與擷取標案資料，若需擴充或變更關鍵字，可直接編輯 **spidar.py** 中對應的 **keywords1** 與 **keywords2** 變數內容；及 **fulltext\_spider.py** 中對應的 **keywords**。

## 二、第一次執行

### 1. 建立並啟用虛擬環境

- **雙擊 run1.bat**
  - 此步驟將自動執行下列操作：
    - 使用 Python 3.11 建立名為 **spidar** 的 Conda 虛擬環境。
    - 啟用虛擬環境。
    - 安裝 **requirements.txt** 中所列之所有相依套件。

### 2. 執行主程式

- **雙擊 script.bat**
  - 程式將自動執行 **spidar.py** 及 **fulltext\_spider.py** 並完成資料處理。

## 三、後續使用

- 日後使用時，僅需**雙擊 script.bat**，即可再次執行主程式，無需重新建立虛擬環境。

## 四、執行結果

- 程式執行後，會自動在相同資料夾中產生 **Excel** 檔案作為輸出結果。
  - 範例如：
    - **全文檢索結果.xlsx**

- 政府採購標案彙整.xlsx
- 

## 作者

- Chen-Yi Wu
- GitHub: [wstone302/GovTenderSpidar](#)