

Google 資料搜集 自動化程式

此份程式為我在研究室為教授撰寫的資料搜集程式，以下簡單說明：

動機：

在研究室的工作的其中一項是協助教授搜集個案分析的資料，包含新聞、週刊、TEJ 及觀測站的資料。在搜集資料的過程中，需要輸入搜尋結果後，一筆一筆資料點開，再轉成 PDF 檔的形式儲存下來，極為費時，也讓我有了開發這個專案的想法。

程式說明：

這份程式的優點在於開始爬蟲前，我讓使用者輸入必須要出現在網站標題的關鍵字，來篩選較相關的資料，亦可選擇要下載多少筆資料，避免爬取全部資料導致程式運作過久，造成效率不佳等問題。最後自動下載至指定資料夾，預估可以省下此流程約 90% 的時間。

結論：

此份資料可以讓使用者快速搜集瀏覽器上的各項資料，使用範圍極為廣泛，且可節省大量人力與時間。而完成這份程式後，未來還可結合文字探勘、自然語言等研究，量化搜集下來的資料，並建立預測模型等，有極佳的擴充性。