

系統簡介

組 別：第 112510 組

專題名稱： iCrawler

指導教師： 陳信宏

專題學生： 10856010 郭宗翰、10856030 彭鈺程、10856031 彭鈺達、10856032 陳永祥、
10856040 鍾皓年、

一、前言

在現在這個大數據時代，科技的進步日新月異，不管是 5G、AI、大數據等科技都在不斷的成熟。而網路上的資訊量非常龐大，因此需要一種自動化的方式來收集、分析和整理這些資料。

二、系統功能簡介

1. 爬蟲系統:

工具(有一定的框架):

使用智慧型爬蟲機器人，利用下拉式選單來選擇要查詢的網站，再去選擇要爬取的版面，最後選擇爬取內容的細項、爬取的頁數

訂製(客製化):

使用智慧型爬蟲機器人，根據客戶的要求進行爬蟲，輸出客戶要的資料，並且做資料分析、文本整理

2. 會員系統:登入與註冊

3. 資料集:可以查詢、下載爬取的歷史資料

4. 程式上傳:根據會員系統的創作者來上傳、更新自己的爬蟲程式碼給客戶使用

5. 系統維護

三、系統使用對象

企業、研究機構、廠商等

四、系統特色

使用智慧型機器人，根據不同網頁自動分辨網頁的框架，快速抓取網站內容

五、系統開發工具

Visual Studio Code、GitKraken、MySQL

六、系統使用環境

Python、CSS、HTML、Javascript

七、結論及未來發展

不受限特定網站，結合 AI 技術，利用智慧型機器人自動偵測網站結構，分辨爬取內容，利用關鍵字搜尋，爬取更精準的內容，增加程序效能。

提供一個平台，能讓創作者能丟出自己的程式代碼，藉此來彌補平台所沒有的一些功能