

Database Systems Term Project Proposal

0716214 江岳勳

0716222 黃偉傑

1. Data

➤ Introduction

我們的資料將會以各式的職缺資訊組成，包含職位名稱、工時、薪資條件及經驗需求等。因為不同求職網的資訊可能會有不一樣的名稱，我們會統合各個網站的欄位資訊，使不同網站的資料能夠以相同格式呈現。

➤ The meaning of columns and tables

目前預想的table會有：

job：實際職缺資訊

company：公司資訊

position：職位ID和職位名稱對照

fieldtype：領域ID和領域名稱對照

educate：學歷等級和其名稱對照

other：其他資訊

table	columns	primary keys
job	job_id、com_id、pos_id、type_id、edu_id、工作待遇、要求年資、工作性質（FT、PT）、管理責任、出差外派、上班時段、需求人數、休假福利制度	job_id
company	com_id、公司名稱、產業類別、資本額、員工人數、公司地址、公司電話	com_id
position	pos_id、職位名稱	pos_id
fieldtype	type_id、職務類別名稱	type_id
educate	edu_id、學歷等級名稱	edu_id
other	job_id、網頁連結、補充說明、聯絡方式	job_id

➤ Where was the data from

我們將會為每個網站設計爬蟲程式將資料爬下來後整理成為資料庫，因為各網站上的資料說明可能會不太一樣(會計/會計師)，所以可能需要對每個網站做交叉比對，確保同樣的資料不會有不一樣的描述。

➤ **Other info about the data**

因為是使用爬蟲來建立資料庫，所以資料的更新可以半自動化。只要在後台或是使用者介面上加上一個更新選項，更新資料庫將不是難事。

➤ **The source of the data**

各求職網站：1111人力銀行、yes123求職網、104人力銀行、518人力銀行

2. Appication Design

➤ **Main Idea**

在COVID-19肆虐的當下，全球陷入了失業潮。雖然這樣的情景還沒發生在台灣，可是我們難保有一天會自己碰到類似的情況。本來求職就是一個很令人頭痛的問題，職缺百百種，哪一個才是最適合自己的？

求職網站所包含的職缺佔了大部分的實際職缺，但因為求職網站百家爭鳴，求職者在尋找適合自己的職缺時，需要同時上不同網站搜尋，得同時適應不同網站的介面。

我們想到有些好用的工具，例如購物比價網站、旅遊比價網站，他們整合了不同網站的資料，讓消費者不用自己貨比三家也能得到最優惠的價格。於是我們認為或許在職缺這方面也能以統合介面方式，提供給使用者方便的體驗。除此之外，我們的查詢方式還會提供一些有趣的分析方式，例如，輸入求職者的需求(條件與希望工作地點)，得出平均薪資或工時落在哪個區間的多向查詢系統。

➤ **Functionality**

- 職缺搜尋：使用者可以輸入關鍵字，選擇職務類別、工作地區、預期薪資等條件。在資料庫查詢結果產出後，能夠以欄位資訊再分別查看。
- 多向查詢：概念是選擇一個目標變數，輸入其他條件資訊後，分析得到的目標數值。例如輸入地區、職務，展示符合條件的職缺的薪資平均、或是多少比例職缺要求員工派遣。

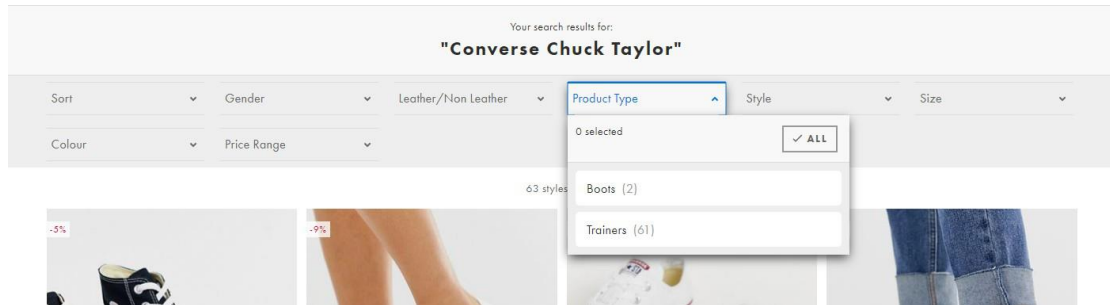
➤ **Interface**

我們會用網頁顯示介面，而搜尋介面會先以簡單的幾個選項讓使用者填入，如下頁第一張示意圖；呈現結果之後，會有更進階的類別能讓使用者調整，如下頁第二張示意圖。(以某知名旅遊比價網示意)



比較各大網站的房價找到您理想的飯店
您可以搜尋城市、特定飯店或甚至著名景點

▲簡易搜尋示意 / ▼進階篩選示意



3. Work Plan

➤ Time Schedule

- 找到能利用的徵才網站
- 對每個網站寫專屬的爬蟲把資料先全部爬下來
- 將資料進行編號化儲存並分類
 1. 將可能有不同名稱但相同意義的工作領域、職位等資料統一並分類
 2. 將工作領域、職位名稱等資料用各自的unique id來表示
- 使用已分類的資料建立資料庫
- 從資料庫發想可以用以與使用者輸入互動的查詢
- 呈現查詢結果

➤ Discussion

我們將會使用Trello來討論細項與問題

Link: <https://trello.com/b/1isqFHkH/db-proj-discussion>

➤ Repository

Github Link: <https://github.com/hm-ysjiang/NCTU-DB-proj>