

# 網頁資料的搜尋與判讀

周學政

國立台灣師範大學地理系

以資訊為核心的技術革命，正以前所未有的速度，不斷的型塑我們的社會。地理學與地理教育自然也不能置身於這股潮流之外。一方面要發展新的理論視野來瞭解資訊化所型塑出的網絡化、全球化的新地景；另一方面也要善用地理資訊系統、網際網路等資訊科技，來掌握資訊化社會的動態。在最近的教育改革中。也反映出這個趨勢，不僅要求靜態的資訊的獲取，更強調能力的發展。如何利用網際網路上大量的資訊來學習，也成為重要的學習目標。本文將介紹網頁搜尋工具的運作原理，以及如何利用搜尋工具找到有用的資料。

## 一、網頁搜尋工具的種類

網頁的搜尋工具可以分為兩大類，主題目錄以及搜尋引擎。

### 主題目錄

如果將網路上的資料當程式一個大型的虛擬圖書館，主題目錄就像是這個圖書館的目錄。提供主題目錄的網站，蒐集了大量網站的網址，然後依照各網站的性質，分別將網站歸類在特定的主題之下。通常主題目錄分成不同的階層、以方便使用者搜尋。要尋找網頁資料的使用者，可以循著主題目錄的階層，逐步找到需要的網站（圖1）。例如YAHOO ([www.yahoo.com](http://www.yahoo.com))、雅虎奇摩 ([www.kimo.com.tw;tw.yahoo.com](http://www.kimo.com.tw;tw.yahoo.com))、蕃薯藤 ([www.yam.com.tw](http://www.yam.com.tw))、新浪網 ([www.sina.com.tw](http://www.sina.com.tw))、pchome ([www.pchome.com.tw](http://www.pchome.com.tw))等原本都是以提供主題目錄為主的入口網站，後來才發展出，電子信箱、網路硬碟、線上拍賣等種種的服務。

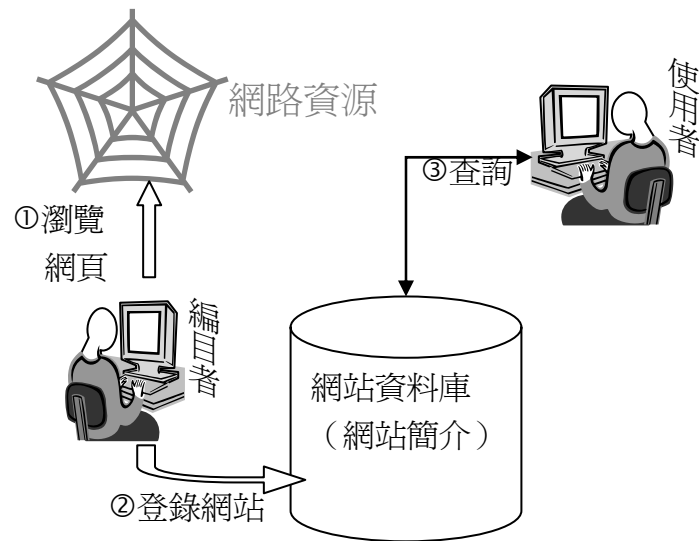


圖 1 主題目錄

這些主題目錄所登錄的網站有兩個來源，一是開放網頁創作者自由的登錄。二是由網站負責編目的人員自行瀏覽與編目。在主題目錄資料庫中所記載的資訊，通常包括網站的名稱、網址、網站內容簡介、網站所屬的主題類別、網站負責人及聯絡方式。

過去在比較主題目錄網站時，常以它們所登錄網站的數量及齊全性為標準；而隨著網頁資料數目的不斷增加，目前的另一個趨勢是強調所選錄網站的品質，例如 YAHOO 及蕃薯藤都會以特別符號推薦的網站。有些網站目錄則特別強調僅提供高品質的連結，例如 About.com ([www.about.com](http://www.about.com)) 的每一個主題都是由該領域的專門人士所整理，有很高的參考價值。在中外各大學的圖書館及研究機構也常編制具有高學術價值網頁的主題目錄。例如加州大學 Riverside 分校所編制的 infomine (<http://infomine.ucr.edu>) 是全面性的學術主題目錄；Utrecht 大學的 KartLab 所編制的, .Oddens' Bookmark (<http://oddens.geog.uu.nl/index.html>) 則整理了 2 萬多個和地圖學有關的網址。

## 搜尋引擎

主題目錄的建立，必須依賴人力進行編目。當網站的出版蔚為風氣，網頁數量成幾何級數的成長時，以人工的方式整理網站資源，成本越來越高，並且時間拖的，越來越長；因此利用電腦程式進行網站資料的搜尋與整理也為必然的趨勢。搜尋引擎利用搜尋程式，根據網站間的超連來搜尋網頁，抓取網頁內容；然後利用程式分析網頁，建立索引。因為目前程式並不能像人一樣的判讀網頁的內容，只能機械性的分析網頁中的文字，所以沒有辦法將主題目錄一樣的將網頁歸屬於某一類別，所建立的索引，僅是標示網頁內文中，出現了哪些文字。使用者則以輸

入關鍵字的方式檢索哪些的網頁內文出現了這些關鍵字（圖 2）。

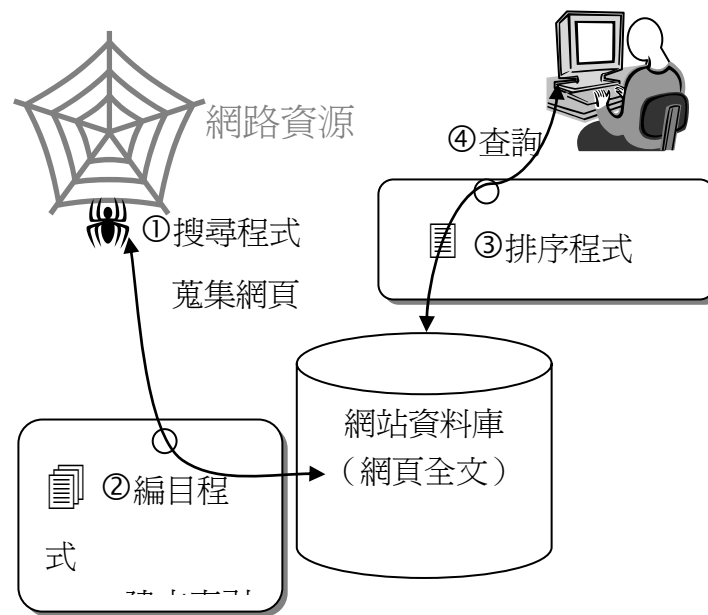


圖 2 搜尋引擎

相較於主題目錄，通常搜尋引擎經常會找到許多符合關鍵字的網頁。但是這些網頁的品質以及和查詢主題的相關性，僅有實際連到這些網站，進行閱讀後才能真正評估。這種依賴使用者來評估網頁價值的方式，在回應的網頁不多時，還可以運作；如果回應網頁的數量多到上百、上千、甚至上萬時，真正對使用者有價值的資料，反而被淹沒在其中。因此要如何使搜尋到的網站依照相關性質及資料品質排序，成了搜尋引擎重要的問題。過去用來安排搜尋結果的方式包括，

- 網頁、網站的名稱的字典順序，例如 A 在 B 之前
- 關鍵字出現的位置，如出現網頁標題較網頁內文重要
- 關鍵字出現的次數或密度。
- 網頁的新舊
- 網站以付費方式提升排行

但是這些排序方法都無法提供令人滿意的結果，並且有關很容易受到操弄。例如在SARS流行時，可能使用者要找尋有關SARS防疫的消息；而口罩、消毒產品的相關業者甚至完全不相關的行業，都可以針對搜尋引擎的特性，提昇自己的排行。例如建立一個網站以“00000123SARS防疫站”為名稱是、在網頁中加入50個SRAS字眼，(可以將文字顏色設成和背景一樣，使讀者看不出來)，並且經常做一些小變動，以提升網站的排序。而搜尋引擎網站，則會針對這些惡意操弄，進行反制；因此展開了一連串的攻防戰。因為沒法解決這個問題，雖然搜尋引擎如Altavista ([www.altavistacom](http://www.altavistacom)) 早在1995年，就能夠搜尋到許多的網頁，但是使用者必須有很豐富的經驗，才能找到有價值的資料。因此主題目錄一直是大眾主

要的網頁搜尋工具。

這個難題的必須要等到 **GOOGLE** 發展出所謂 **PageRank** 的技術，才使情況改觀。**PageRank** 分析網頁間的連結關係，計算網站的重要性。**PageRank** 的基本概念是被連結越多的網頁，越重要；例如圖 3 中網頁 A 被連結的次數多，因此重要性較高；其次是被重要的網站連結所得的權重較高，例如網頁 B，雖然只有一個連結，但是這 A 所做連結的，因此比一般的連結有較高的價值。事實證明利用 **PageRank** 來安排網頁出現的順序，非常的有效；使得 **Google** 成爲目前使用率最高的搜尋網站。

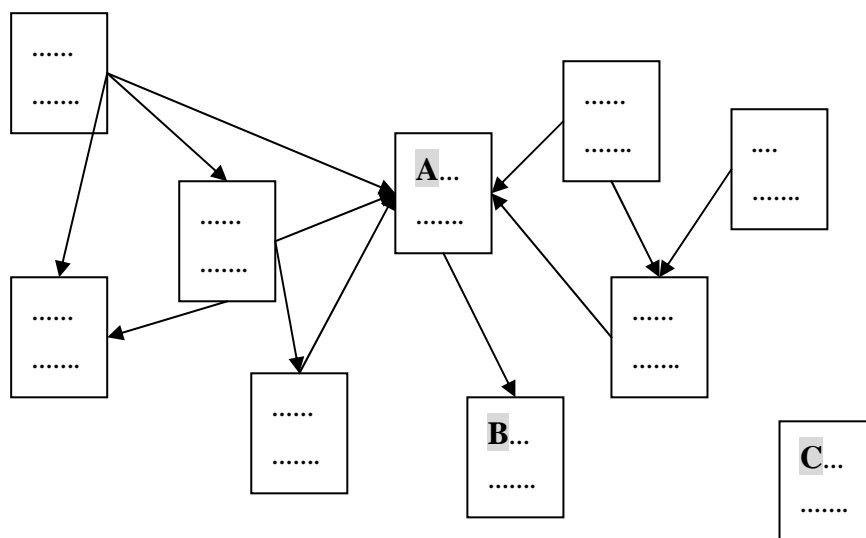


圖 3 相互連結的網頁

針對主題目錄及搜尋引擎的特性，可以做出以下的幾點分析：

- 1 不論主題目錄及搜尋引擎，都不是及時搜尋網路中的各網站，而是搜尋事前建立好的資料庫。主題目錄的資料庫通常是以網站作爲分類的對象，僅記錄網站的摘要資訊、所包含的網站數較少；搜尋引擎的資料庫，儲存網頁全文，並且以全文作爲索引的對象，網頁的數目龐大。
- 2 主題目錄的使用者，可利用層級性的目錄，循序尋找相關網站；搜尋引擎的使用者則輸入關鍵字，進行網頁全文的搜尋。雖然主題目錄也提供關鍵字搜尋，不過關鍵字所尋找的只是網站的名稱和對網站的描述，而不是網站的真正內容。
- 3 過去因爲搜尋引擎的排序的問題無法解決，主題目錄是主要的搜尋工具。隨著 **PageRank** 的成功，類似的技術廣泛被採用；目前搜尋引擎已經逐漸成爲搜

尋工具的主流。主題目錄將由強調一般化、大型化轉變成專門化並強調連結網站的品質。

### 隱身網頁(invisible web)

雖然搜尋引擎資料庫有個龐大網頁的資料庫，如 Google 宣稱在它的資料庫中有高達 4,285,199,774 個網頁(2003 年 4 月)，但是在 INTERNET 上還有許多資源沒有被含括在搜尋引擎的資料庫中。可能的原因有以下數項：

1. 不屬於網頁類型的 INTERNET 資源，例如 BBS、新聞群組、電子郵件等資源。
2. 禁止搜尋的網頁，有些網站不允許搜尋引擎去分析它的網頁。
3. 有些網頁，並沒有其他的網頁連結指向它，也此無法被搜尋程式找到，如圖 3 中的 C 網頁。
4. 沒有和 INTERNET 連結的網路資源，例如機構內部的網頁。
5. 因為網頁太新或其他理由，而還沒有被蒐集到的網頁。
6. 需要收費的網頁。有相當多網站，如報紙、期刊等，需要會員帳號登入後才能使用。搜尋引擎因為沒有帳號，所以無法進入查詢。
7. 通常編目程式所處理的是 html, htm, xml 及 txt 等文字型態的資料格式；其他類型的文字檔案，例如 Microsoft word PowerPoint、Excel、Acrobat 等類型的文件，就不是每一個搜尋引擎都能夠分析其中內容並建立索引。因此就算搜尋引擎能夠透過連結找到這些格式的資料，也無法建立索引以提供給使用者查詢。至於像圖片、聲音 (MP3、WAVE)、影片、FLASH 或者是互動式地圖網站等非文字的資料類型。搜尋引擎通常僅能處理有關檔名、檔案大小、網路連結等資訊，完全無法分析這些資料的實質內容。
8. 許多網站其中的資訊，是放在資料庫系統中 (如 access、mysql、oracle 等)，等到使用者輸入查詢條件後，由程式代為查詢資料庫，得到結果後動態的產生網頁，因為這些網頁，原本並不存在，因此也沒有辦法事前建立索引。圖書館的館藏查詢最常見的資料庫的範例 (圖 4)。

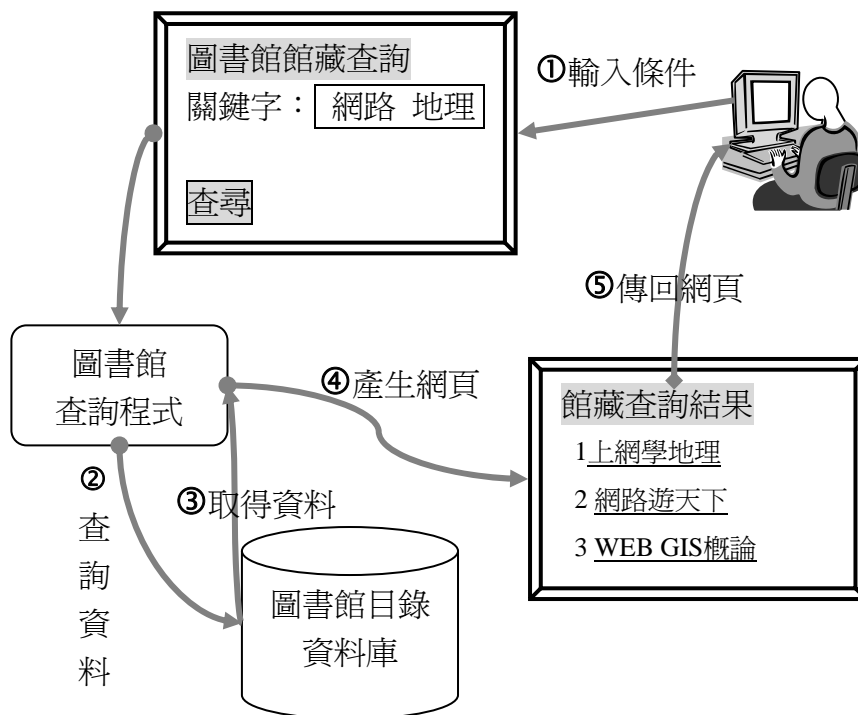


圖 4 由資料庫動態產生網頁

使用搜尋引擎時，必須清楚的理解這些可能的限制，對於資料庫類型的網頁，更要特別留意，因為其中有著大量高品質的資料。以台灣而言，全國博碩士論文資訊(<http://datas.ncl.edu.tw/>) 及各大報社都有豐富的資訊；而國外各種學術期刊及電子書的資料庫，更是重要的資訊來源。有些搜尋網站，如ProFusion (<http://www.profusion.com/>) 就專門蒐集及整理網路上這些資料庫。

## 二、搜尋引擎使用基本技巧

### 瞭解搜尋工具的功能

瞭解所使用的工具是成功搜尋的第一步。本來搜尋主題目錄以及搜尋引擎屬於兩種不同的搜尋方式。但隨著搜尋引擎的功能的進展。有許多原本是主題目錄的網站也加入了搜尋引擎的功能，甚至對於一些原本是隱身的網路資源(invisible resource)，也提供了搜尋的功能；例如幾乎所有入口網站都支援新聞群組 (BBS)、報章新聞及圖片的搜尋 (表 1)。其實在經過一連串的合作與合併後，有許多入口網站使用同樣的技術及資料庫。例如蕃薯藤、新浪網都使用 Google 的搜尋引擎；Yahoo、Altavista、All the web 都使用 Yahoo search 搜尋引擎。

表 1 入口網站提供的各種搜尋工具

	主題目錄	搜尋引擎	BBS 新聞群組	新聞	圖 片	其他
<b>openfind</b>	網頁	分類	BBS	新聞	圖片	音樂 文件
<b>Google</b>	網頁目錄	所有網頁	網上論壇	News(英文 版)	圖片	商品 股市
<b>Yahoo 奇摩</b>	分類網站	所有網站		新聞搜尋	圖片 搜尋	
<b>蕃薯藤</b>	網站	網頁 (google 提 供)	BBS	新聞	圖片	商品 股市
<b>新浪網</b>	所有網頁	分類網站 (google 提 供)		新聞媒體 雜誌報導		期刊論文(中 央圖書館)

一些入口網站除了透過搜尋工具的選擇，讓使用者直接選擇搜尋的對象及方式外，也會自動的轉換搜尋工具。例如蕃薯藤及MSN會先使用主題目錄查詢，如果沒找到相關網站，則使用搜尋引擎查詢。另一個例子是如果輸入美國住址，如「107 5th Ave, New York」，作為關鍵字，會自動的轉換到地圖網站行進查詢，並繪出現地圖。查詢Yahoo或Google時，在關鍵字前加上 define，僅會查詢網上英文字典網站，例如「define geography」會搜尋網路的字典，列出geography的定義。各大入口網站，為了提供更好服務，提供的功能越來越多。要有效的使用這些網站來搜尋資料，必須先仔細閱讀使用說明，以瞭解搜尋工具的特性。

### 關鍵字的選擇

一般而言，使用關鍵字查詢時，必須要越精確愈好，例如想查詢人口資料，用「人口」作為關鍵字，如果找到太多不相關的資料。可以用「人口成長」做關鍵字來縮小資料範圍；要進一步縮小尋找範圍，則使用「出生率」、「死亡率」、「性比例」等關鍵字，那麼找到的都會是和人口統計有關的網頁。其次，使用同義詞，可以找到更多資料，例如「甘薯」又稱「蕃薯」、「紅薯」、「地瓜」。因為網路上目前還是以英文資料最為豐富，如果找不到適當的中文資料，可以用英文關鍵字查詢英文網站。如果不清楚英文名詞的拼法，除了使用字典，電子辭典、翻譯軟體等工具外，也可以使用網路上的漢英字典，例如林語堂當代漢英詞典 (<http://humanum.arts.cuhk.edu.hk/Lexis/Lindict/>) 英漢/漢英 線上電子字典

(<http://www.ee.tku.edu.tw/~rexchen/cdict/cdict.html>)或者是altavista的網頁翻譯工具(<http://babelfish.altavista.com/>)。例如使用「蕃薯 OR 甘薯 OR 紅薯 -蕃薯藤」為關鍵字查詢google，可以查到 94,400 筆資料；以英文「"sweet photo"」查詢可以找到 552,000 筆資料。如果先使用「蕃薯 學名」查詢到蕃薯的學名Ipomoea batatas (L.) Lam，然後使用「Ipomoea batatas」作為關鍵字查詢，可以找到 14,000 筆資料。總括來說，選擇關鍵字時，要以作者的角度來思考一個主題中最可能出現的字眼，並且要多試幾種關鍵字的組合。

在查詢時，經常需要使用多個關鍵字，這時要善用搜尋引擎所提供的邏輯運算使查詢更為有效。大部分的搜尋引擎雖然在語法上可能有差異，但都支援下列的邏輯（布林）運算：

- 且 (AND)： 「蕃薯 AND 原產地」表示「蕃薯」及「原產地」兩個字都必須出現。
- 或 (OR)： 「蕃薯 OR 甘薯」表示「蕃薯」或「甘薯」之一出現即可。
- 排除 (NOT)： 「蕃薯 NOT 蕃薯藤」查詢「蕃薯」但是排除「蕃薯藤」。

其中 AND 的用途很明顯，可以用多個關鍵字進行更精確搜尋，例如想要瞭解蕃薯的原產地，使用「蕃薯 AND 原產地」比僅用「蕃薯」作為關鍵字，更為有效。使用 OR 的情況，是多個可能的名詞時，同時搜尋。雖然查詢時可以用兩次搜尋來模擬 OR 的效果；但是如果使用 OR，僅會進行一次的排序，容易比較結果。例如同時使用蕃薯的各種別名「蕃薯 OR 甘薯 OR 紅薯」進行查詢，找到最好的網頁。使用 NOT 可以消除不必要的連結。例如以「蕃薯」為關鍵字查詢時，發現太多的有關蕃薯藤網站的內容，可以使用「蕃薯 NOT 蕃薯藤」來剔除一些不相關的網站。

使用中文查詢時，還必須注意字與詞的分別，例如，在英文中，使用「population growth」做為關鍵字查詢，只要在網頁中出現 population 及 growth 兩個字就可以；而使用「"population growth"」查詢時，population 及 growth 兩個字必須相鄰。因為中文不靠空白分隔，因此有字詞分辨的問題。例如輸入「人口成長」為關鍵字查詢時，不同的搜尋引擎可能有不同的查詢方式，例如 altavista 會將「人口成長」當成一個詞來查詢。而在中文 GOOGLE 及 Yahoo Search 中，會將「人口成長」拆解成「人口」及「成長」兩個詞來查詢。因此造成查詢結果的差異。

簡體字和繁體字的差別是查詢中文資料的另一個問題。對讀者而言，簡體字和繁體字的差異主要在外形上的不同；對電腦來說，繁體字和簡體字的主要差別並不是文字圖像（字型）的不同，而是繁簡字的內碼不同。在電腦中是以數字代碼來



記錄文字，數字代碼和文字的對應規則稱為內碼。內碼隨著文字種類、年代及地區有所不同，台灣、香港所通用的內碼是 Big5 碼。在 BIG5 碼中，「人」這個字是用 16 進位數字 A448 來代表。中國大陸所使用的 GB 碼則是以 C8CB 這個 16 進位數字代表「人」字。如果使用錯誤的內碼來看文字，就會呈現亂碼（圖 5）。因此輸入繁體字的「人口成長」，當然就只能夠查到繁體(Big5 編碼)的資料，要查詢簡體字的資料必須輸入 GB 碼的「人口成長」。要在繁體的 MS Window 中輸入 GB 碼，必須利用一些特殊的輸入法，如自然輸入法、譯點通的多語輸入；或者採用繁簡字轉換的方式。MS Window 的附屬應用程式中提供了轉碼器，可以將純文字繁簡互換，在新版的 office 軟體中，也提供繁簡字的互換的功能。Google 則替中文使用者做了貼心的繁體字及簡體字自動互換的功能；只要輸入繁體字查詢，Google 會自動將關鍵字轉換成簡體字，同時搜尋繁體及簡體網頁。

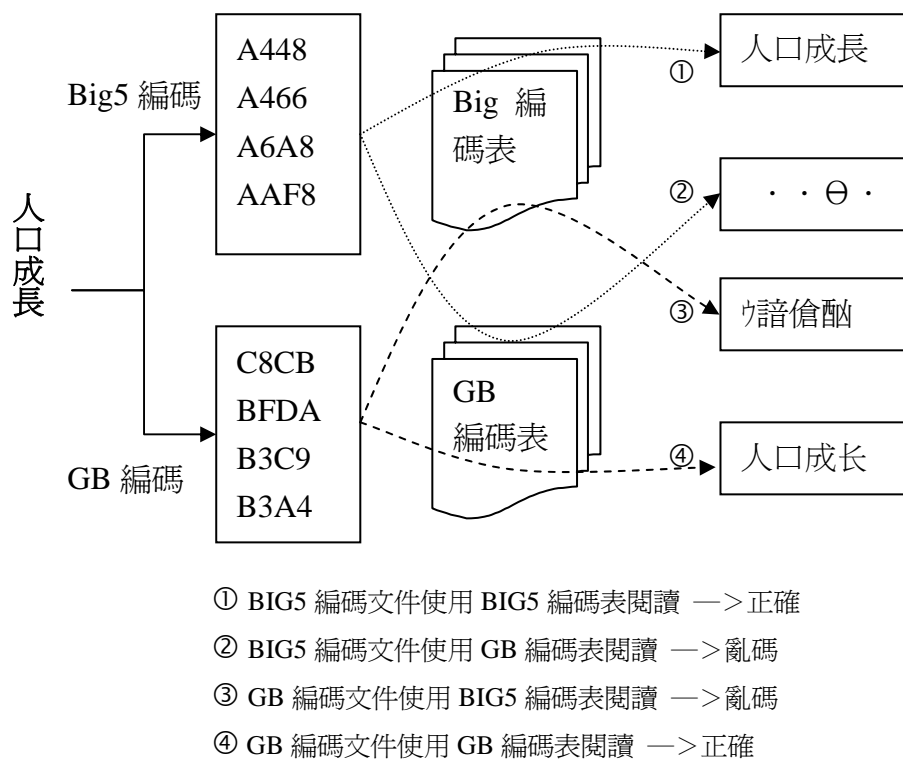


圖 5 文字的編碼

由上面的說明中可以發現，搜尋引擎不僅僅是單純的接受關鍵字進行查詢。經常會對關鍵字進行一些處理，讓查詢更容易及更有效（表 2）。

表 2 搜尋引擎的排序方式及關鍵字處理

排序方式	關鍵字處理
<b>Openfind</b> 關鍵字出現的密度越高，網頁優先出現的機率越大。 關鍵字位於標題的網頁會比關鍵字	自動加入同義字。例如查詢「榮總」，Openfind 會自動將它最佳化成「榮總   榮民總醫院」，

	位於內容中的網頁優先出現。使用多個字詞查詢時，字詞彼此間隔得越近，越有機會優先出現。網頁被其他網站連結的次數越高，優先出現的機率也較大	同音校正輸入「通訊計術」(技打成計)，則 Openfind 會自動最佳化成「通訊計術   通訊技術」。
<b>Google</b>	PageRank 關鍵詞接近的程度	自動加入相近的字,如查詢「book」也會查詢「books」 中文字詞分析(限中文介面) 繁簡互換:輸入繁體字會同時查詢繁體及簡體網頁 英文拼字檢查 英文同義字
<b>Yahoo search</b>	分析網頁的內文、標題、描述、出處及連結。	中文字詞分析 英文拼字檢查 雅虎奇摩的圖片搜尋有繁簡字互換功能

### 三、資料搜尋策略

#### 選擇適當的搜尋工具

進行搜尋時，首先要根據問題的特性選擇搜尋適當的搜尋引擎。對於單純、有直接的答案的問題，搜尋引擎是較恰當的搜尋工具。例如、「世界最高的三座山」、「鐵路局的網址」等問題。只要多嘗試幾組關鍵字，就可以找到資料。

如果我們不是找一個特定的資料，而是要對一個問題做較全面性的瞭解，則考慮使用主題目錄；特別是經過專家精選的主題目錄。例如「溫室效應」、「印度的經濟發展」等問題，都不是三言兩語可以說的清楚的。如果在欠缺相關的知識背景下使用搜尋引擎，經常會迷失在搜尋引擎所提供的多樣化資料中。對於這類問題，可以先嘗試由 [www.about.com](http://www.about.com) 這類專業的主題目錄中所介紹的網站作為起點；或者直接在網路上的百科全書中尋找初步的答案（表 3）。

表 3 一些網路上的百科全書

- **Wikipedia** (<http://www.wikipedia.org>) 是最新類型的線上百科全書。完全由志願者透過網路所編寫的百科全書，只要有意願、任何人都可以修改、新增條目。
- **大英百科全書** (<http://www.britannica.com>) 是最具知名度的百科全書，

但是要收費；不過它針對於各主題所推薦的網站，可以免費使用。

- 哥倫比亞百科全書是 **Columbia Encyclopedia** (<http://www.encyclopedia.com/>) 是網路上免費的百科全書。有不少網站都使用這個資料，如 **infoplease** (<http://www.infoplease.com>)。
  - 微軟的 **Encarta** (<http://encarta.msn.com/encnet/features/reference.aspx>) 則將部分條文公開免費使用，查詢完整內容則必須付費。
  - 大美百科全書 (<http://go.grolier.com/>) 必須付費使用，不過國內許多學校圖書館都有訂閱。
  - 智慧藏 (<http://www.wordpedia.com/>) 提供大英百科全書、中國大百科的付費查詢。
  - 中華百科全書 (<http://living.pccu.edu.tw/chinese/index.asp>) 是較完整的免費中文百科全書，但編寫的年代是民國 70 年。
- 

有些問題，如”新竹市歷年的人口”、”蕃薯的起源地及傳播路徑”，雖然有單純的答案，卻可能要綜合許多網頁的資料，才能得到解答。搜尋這類問題時，通常要有相當程度的背景知識及熟練的搜尋技巧，才能由眾多網站中挖掘並整理出答案。例如要查詢新竹市歷年的人口，必須瞭解人口資料可能出現在人口統計、政府都市、交通等規劃報告中，而這些資料的可能出版者包括行政院主計處、內政部、省政府及新竹市政府等單位。有了這些基本的知識，才能規劃有效率的查詢方式。

### 解讀網頁資料需要批判性的思考

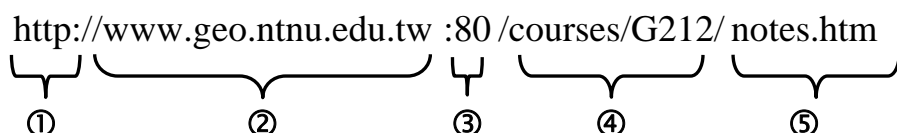
孟子說，「盡信書、不如無書」；對於網頁資料更應該抱持著批判態度來解讀。雖然經由 **PageRank** 的排序或主題目錄網頁的篩選，對網頁的品質已經有了初步的肯定，我們還是必須透過對網站的屬性、作者的權威性、內容的客觀性、時效性及涵蓋性的清楚認知，才能正確的判斷網頁資料的適切性。例如中小學生所製作的網頁，通常不宜直接引用，但可由其中找到更具權威性資料的可能來源；環保團體及民間公司的網站對環境問題常有各自的立場，要瞭解環境問題的全貌，兩者不可偏廢。

利用搜尋引擎所找到資料，更要仔細的分析其品質。因為搜尋引擎是以網頁為搜尋的目標，如果僅看網頁，而不瞭解整個網站的性質，很難判斷該網頁的品質。因此如果該網頁有指向網站首頁的連結時，應該連去該網站的首頁，以便瞭解該網站的性質，最好還能夠閱讀該網站的簡介，如「關於我們」、「網站緣起」、「宗旨」等項目。如果該網頁中沒有指向首頁的連結，那麼可以使用網址列去找到網站的首頁（圖 6）；只要將網址中的檔案路徑及檔案部分消去，通常就是該網站的首頁。舉例來說，如果網址中的內容是

<http://www.geo.ntnu.edu.tw/courses/G212/notes.htm>，那 <http://www.geo.ntnu.edu.tw>

就可能是首頁的網址。

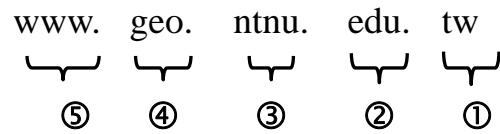
連結到首頁，通常由這些高品質的網站，其中常常還有相關連結，各些連結通常也具有相當的品質。



- ①：傳輸協定的種類。Internet 提供多樣的服務，不同的網路服務使用不同的傳輸協定。如 http 指網頁文件服務、ftp 指檔案傳輸服務、mailto 指 e-mail 服務。
- ②：電腦的網域名稱，也就是電腦在 Internet 網路上的住址。
- ③：通訊協定的 port 編號。因為一台電腦可能提供多種 Internet 服務，因此用 port 來區分服務種類。如果不指定 port，則使用預設則使用該通訊協定預設的 port。例如 http 預設使用 port 80。ftp 預設使用 port 21
- ④：存放檔案的路徑
- ⑤：檔案名稱

圖 6：網址（URL）的意義

其實只要學會閱讀網址中的網域名稱部分（圖 7），就能找到一些有用的資訊。例如 `www.geo.ntnu.edu.tw` 網址，由右端向左看，以「.」分隔成 5 部分，tw 代表這是台灣的網站，edu 代表教育機構，ntnu 是學校的縮寫，因為直接在 edu 之前，並且名稱中有 u（university），大概是大學；geo 代表系所，應該是地理（geography）、地質（geology）或地科系（geoscience）；www 應該代表該系官方網站。在搜尋引擎的搜尋結果頁面中，通常也會列出網址，這樣就可以大致判斷網頁的性質，屬於 edu 網域，通常學術性較高、gov 網站較有權威性、org 網站往往有其特殊宗旨。另外一些搜尋網站也可以用網域來限制搜尋的範圍。例如 Google 及 Yahoo 可以使用 site:這個指令來限制搜尋的範圍，如 `site:tw` 表示僅搜台灣的網址，`site:edu.tw` 表示僅搜尋台灣的教育單位。



- ①：國家、地區代碼；如 tw:台灣、cn:中國、jp:日本，許多在美國註冊的網域名稱沒有國家代碼。
- ②：機構類別；如 edu:教育單位、com:公司、gov:政府機構、org:非營利組織、mil:軍事單位、net: ISP 等網路服務單位、idv:個人、
- ③：機構名稱，例如 ntnu 為台灣師範大學
- ④：機構內單位名稱，例如 geo 為地理系
- ⑤：伺服器名稱

電腦的網域名稱，也就是電腦在 Internet 網路上的住址。通常  
①②③⑤為必要，④視情況可能沒有，或有多層。

圖 7：網域名稱的意義

搜尋過程其實就是學習的過程，透過搜尋及判讀的過程，我們不斷的學習到新的知識。對問題瞭解的越深入，所使用的關鍵字就會越精確，也更能找到相關的資料，也更能認知到某個網頁的真正價值。