

「開放目錄專案」 Open Directory Project 介紹

「開放目錄專案」 Open Directory Project：ODP (<http://dmoz.org/>)，原稱為 Dmoz，是網景公司 (Netscape) 所主持的一項大型公共網頁目錄。由全世界各地的義務編輯人員來審核挑選網頁，並依照網頁的性質及內容來分門別類。

ODP 是一個和 Yahoo 類似的由人工建立的網上網站目錄，不同的是 ODP 是一個非營利的組織，所有的內容都是由志願的編輯員建立的，而且 ODP 將它的資料庫公開提供給其它的網站目錄使用。在經過超過三萬六千名來自世界各地志願的編輯員數年的同心協力之下，ODP 目前已經成為了世界上最大的網上人工網站目錄，它包含了三十七萬多種類別，兩百六十四萬多個網站的資料。目前已經有七百多個網站採用 ODP 的網站目錄資料庫，其中較著名的包括：

網站名稱	網址
蕃薯藤	http://odp.yam.com/
AskJeeves	http://www.ask.com/channels/
AT&T WorldNet Service	http://www.att.net/
Google	http://directory.google.com/
HotBot	http://hotbot.lycos.com/
Netscape Search	http://search.netscape.com/

ODP 的一個重要的特色是它的涵蓋性。它的資料庫包含了七十種語文的網站，其中包括了中文和台文。三萬六千名志願的編輯員依照他們的文化、專業、經驗及興趣來分工合作進行目錄的擴展。這樣的背景促成 ODP 向每一個領域不斷的延伸，也因為這樣的背景，ODP 經常的徵求更多的志願的編輯員來更進一步擴大目錄的涵蓋性。

其中，在搜尋有關於 ODP 也發現了另外一個相似的網站「中文在線開放目錄」 Online Open Directory Project：OODP (<http://www.oodp.org/>)，由於根據網址的判斷 OODP 不屬於 ODP 的分類下，因此猜測這應該是中國大陸仿效 ODP 自行建立的類似 ODP。

檢索工具分析

資料收集

●收錄資料種類：

web pages，所有的內容都是由志願的編輯員建立而成的。

●收錄資料的質與數量 (coverage and quality)

與一般的搜尋引擎有很大的不一樣的地方是 ODP 資料庫內所有的資料都是經過人為的篩選堆積起來的，不像一般的搜尋引擎是藉由 spider 的自動搜索機器人程式到各網站上抓資料傳回資料庫。也因為這樣的特性，在數量方面遠遠不及 google、yahoo...etc 這些強大的搜尋引擎，但由於每筆資料幾乎都是經過人為的篩選，因此在品質方面較不會流於大量的廣告網頁或者是根本毫不相關的網頁，在搜尋品質方面是可以肯定的。例如在 ODP 上輸入 SEIKO 這個品牌，所出現的幾乎都是全球各地的 SEIKO 網頁或者是相關的訊息，絕對不會出現一堆購物網站的相關連結。

●更新頻率：

在更新頻率上，相對於 spider 的自動搜索機器人程式，ODP 由於需透過人為的編輯，在整體更新頻率上，無疑地一定會緩慢很多，不像 spider 日以繼夜地搜尋。

資料索引及比對技術

●搜尋結果的數量？與問題相關的比例有多少？

如同上述所說，ODP 由於人為的編輯，在數量上遠不及一般的搜尋網站，而關於搜尋結果與問題相關的比例為何？其實非常難以斷定，因為此比例會受到所輸入或抽樣的關鍵詞而不一樣，然而在關鍵詞的涵蓋範圍亦會影響搜尋的結果，例如搜尋 APPLE 電腦，但是所搜尋出的結果可能會涵蓋蔬菜水果類、作業系統、電腦廠商...等等，因此無法實際估算出一比例還斷定 ODP 的回收率與精確率。不過可以肯定的是，由於資料庫的品質是經過控管過的，因此所搜尋出的結果大部分都能符合搜尋者的需要，相對地，會比一般搜尋引擎一堆雜而不精的資料優異許多。

●搜尋引擎的特色：特殊檢索功能？同音(科林頓or柯林頓)？容錯？多國語言？

雖然 ODP 的資料庫包含了七十種語文的網站，但是截至目前為止，主要的資料庫內容還是以英文為主，筆者也嘗試了一下搜尋引擎的檢索功能，不過似乎

並沒有那樣的多功能。

● 是否容易使用？是否有足夠的說明來協助你建立查詢策略？

使用起來與一般的搜尋引擎沒有太大的差異，不過另外要讚賞的是 ODP 的分類目錄，分類極其細微，難怪就連一般的搜尋公司（Ex：yahoo）也要借助於 ODP 的資料庫分類目錄。

結果顯示

● 是否包括摘要、檔案大小或日期、相關性評分顯示、查詢資料結果是否重覆？

在搜尋結果方面，所列出的每筆資料皆有相關的摘要解說，但並沒有相關的日期或者是檔案大小，在搜尋資料的重覆性問題也因為經過篩選而降低。在相關性評分這部分以ODP的主頁面並沒有這項服務，不過筆者倒是發現了另外更有趣的事情，「ChefMoz」（<http://chefmoz.org/>）附屬在ODP（part of the open directory project）內，透過Welcome to the ChefMoz dining guide beta!的敘述可以知道此搜尋功能還在測試階段，這個搜尋引擎可以搜尋有登入在其內的所有相關的餐廳，不過目前台灣餐廳的資料少得可憐，只有一筆由瑞士人登錄的台北縣餐廳，以大陸廣州的餐廳目錄（<http://chefmoz.org/China/Guangzhou/>）為例，在每筆餐廳旁有\$\$的符號，是由編輯員評論消費昂貴與否，其次每個網友可以加上自己的評論，針對食物、服務、氣氛(ambience)、整體評價四種項目進行評分，除此之外，也可以加上自己的心得和推薦的食物，此項功能實在非常的不錯，可透過編輯者的推薦將全球的美食餐廳一覽無遺。

● 是否足夠讓你判斷網頁與問題的相關性？

可以的！這又要說起ODP的另一項優點，在每筆搜尋資料的結果連結最後，ODP都會貼心的附上此連結的相關分類，很容易可以判斷此資料是否為自己所需要的。例如若要搜尋APPLE電腦，那使用這就不會傻傻的去點選「Home: Cooking: Fruits and Vegetables: Apples」的連結了。

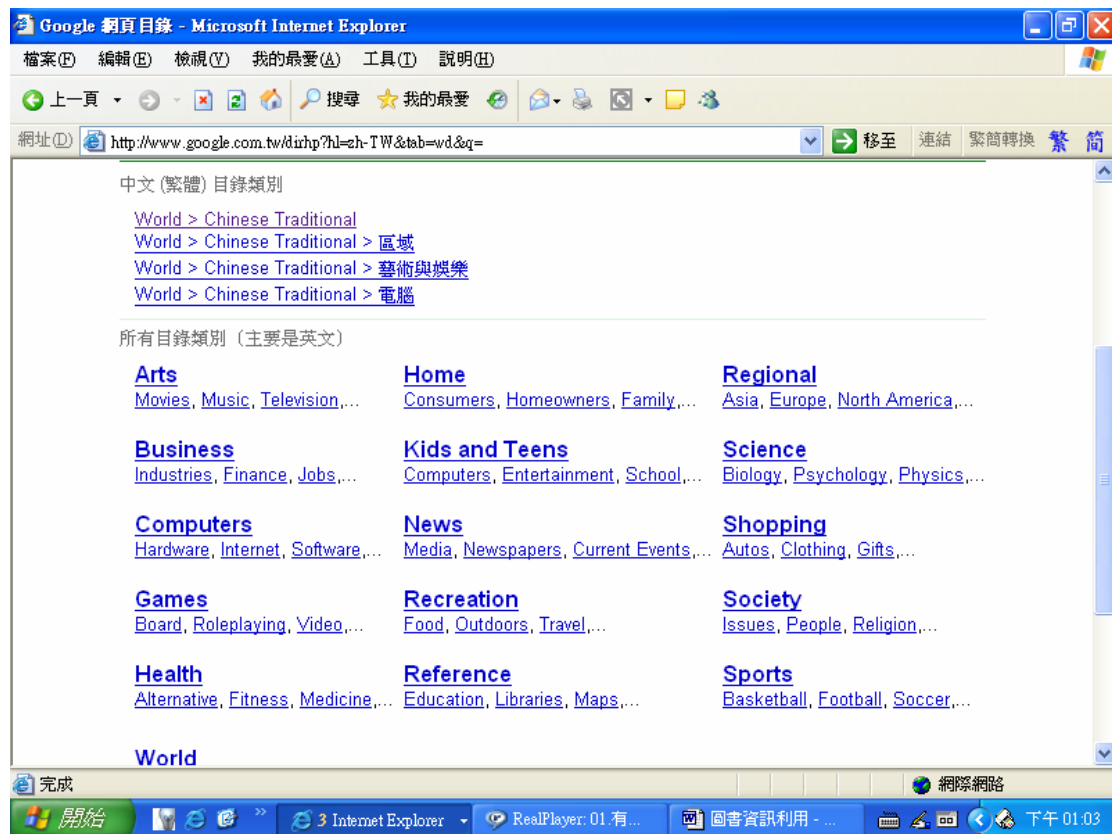
● 排列是否合理？

經過筆者仔細地觀察與瞭解，發覺 ODP 的排列並沒有特殊的邏輯，既不是按照分類的項目也不是依據資料比數的分類目錄多寡更不是依據英文字母的開頭排列。

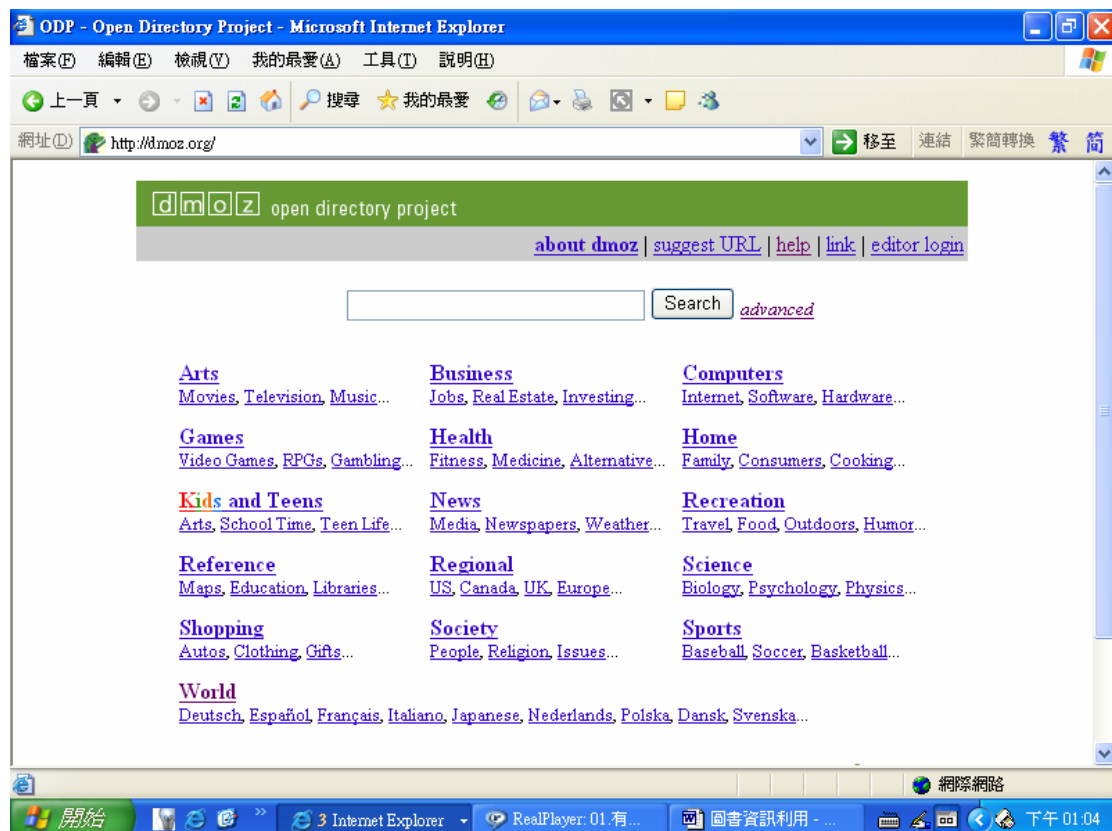
Google 網頁目錄與 ODP 之比較

Google	屬性	ODP
依網頁內容分類	分類	依網頁內容分類
將重要網頁列於目錄的前方	編排	無一定邏輯順序
查詢結果的範圍涵括太廣時，使用網頁目錄可以縮小搜尋於特定的範圍內	使用時機	查詢結果的範圍涵括太廣時，使用網頁目錄可以縮小搜尋於特定的範圍內
較高，常常可見一樣的網頁在 google 的搜尋結果出現了兩個一模一樣的連結，另外也常出現所搜尋出的網頁一則是用中文表示，另一則則是用英文表示	重複性	較低，所有 ODP 內的網頁全透過編輯人員篩選，因此同一個網站重複出現的機率幾乎為零，另外 ODP 有語言上的區隔，不同的語言查詢詞彙只會出現該語言的相關連結
較粗略	資料區分	較精細（有人員審核）
多，全部都放	數量	少，英文只出現英文的資料
可以跳頁瀏覽	缺點	必須一頁一頁按
網頁簡介內有一條長短不等的綠色橫線，從綠線的長短，能輕易明瞭這網頁的評價	提示	是公司首頁會有星星的符號出現
有	頁庫存檔	沒有
有	類似網頁	沒有
不能	編輯自由	可申請成為編輯者
沒有註明	語言版本	有寫出目前有的語言
可連結搜尋到相關圖片、新聞等資訊	連結	列出在哪些搜尋器可找到相關資料
簡單大方	外觀設計	簡單大方
一百五十萬個以上網站	收錄網站數	兩百六十四萬多個網站
由於 google 是透過 spider 自動到各網頁作搜尋的動作傳回 google 資料庫，因此在資料上的更新頻率會較高	更新頻率	ODP 是透過編輯員經過網頁的篩選之後才作更新的動作，因此其更新頻率必須視其編輯人員的數量多寡或者是所推薦的網站數目而定，在頻率的更新上，若在未來編輯人員或者是開放目錄的普遍使用達到一定的可觀數量時，在更新的頻率上可能會比 google 來的快速與準確

Google 的網頁目錄畫面



ODP 的網頁目錄畫面



結論及心得

現在科技發達，在搜尋資料時，網路成了一個相當便捷的工具。可是在瀏覽的過程中往往因為資料太多導致找不到自己真正想要的東西，而透過網頁目錄可以快速找到自己想要的資訊的所屬類別。

Open Directory project 開放式搜尋目錄專案 (ODP) 是許多搜尋引擎的數據來源，由全世界各地的義務編輯人員來審核挑選不同種語言的網頁，並依照網頁的性質及內容來分門別類，如果自己看到不錯的網頁，卻沒有在其他的搜尋引擎中看到，便可以向開放式搜尋目錄專案提供網頁資訊。

另外，如同上述所言，由於大多數的搜索引擎所提供的搜索結果雜而無章，取而代之的是相關性廣告，而且以前幾個較大的網頁目錄已經被取消或甚至必須要求收費才可以登錄，然而 ODP 則是透過專業人員經過網頁的篩選，使得網站的瀏覽更為便利，ODP 為使用者提供了全球化以及條理化的網頁目錄。隨著網際網路的增長，ODP 編輯人員的數量亦在增長，這些專業的編輯人員可以管理整個網頁目錄的一小部分，並把管理的內容呈現在 ODP 上，剔除雜亂和無用的內容，擷取更具意義的搜尋結果。

在確實使用過 Google 網頁目錄及 ODP 之後，我自己比較喜歡 ODP 的使用，因為雖然 Google 的資料很多，但是分類較不仔細，只有大概的方向區分，每每找資料，如果不是很好找的資料都像是在大海撈針，總是要花不少時間在篩選的工夫上，反倒模糊了 Google 身為強大搜尋引擎的優點；相對地，使用 ODP 可以快速找到所屬的類別，然後一層層去找出自己真正要的資料，不過也有可能因為所包含的資料太少而找不到。

由於有專業審核人員把關，所以在個別的分錄上不會出現低相關的網頁，例如我們想搜尋 apple 電腦，輸入 apple 之後，在其他的搜尋引擎上可能就會出現有關 apple 產品的購物介紹或是推銷，類似的網頁佔據大多的版面，給使用者雜亂無章的感覺，但就 ODP 而言，因為一開始就有明確的分類，所以輸入 apple 後，所出現的就是有關電腦資訊的新聞或是文章，而不會有購物或是拍賣的網頁出現，使整個搜尋結果呈現較一致的水準。