網路資訊再尋獲之檢索行為初探

An Exploratory Study on the Re-finding Behavior on the Web

江信昱*　卜小蝶**

Xin-Yu Jiang*, Hsiao-Tieh Pu**

摘　要

上網重複查詢先前瀏覽資源之資訊再尋獲行為十分常見，但由於網路環境之變動性及資訊尋獲情境之多元性，資訊再尋獲與一般尋獲或查詢行為可能存在差異。本研究以實驗、觀察、訪談、問卷等方法，嘗試瞭解網路使用者之資訊再尋獲檢索行為特性；同時，也分析比較尋獲與再尋獲兩階段之檢索歷程特性。研究結果顯示，相較資訊尋獲行為，使用者於資訊再尋獲過程中所使用之檢索工具更為多元、也會交互使用多種檢索策略尋得先前情境線索、其所採用之檢索技巧較為複雜、與系統的互動也較為頻繁。根據實驗結果，資訊再尋獲雖較尋獲階段所花費時間較短，但其所需認知心力及檢索困難度也增加不少。一般而言，受試者對再尋獲的結果多感滿意，但也表示若系統能提供更多回想支援功能，再尋獲的檢索成效其實應該更高。此外，本研究也提供不同檢索任務類型對資訊再尋獲影響之初步分析結果。最後，本研究針對網路資訊檢索系統及個人資訊管理層面提出一些改善建議。

關鍵字：網路資訊檢索行為、資訊尋獲、資訊再尋獲、個人資訊管理

Abstract

It is common for users to relocate information previously found on the web. However, their search behaviors in initial finding and the subsequent re-finding may differ due to the dynamic nature and contextual diversity of the web. This study used experiment, observation, interview, and questionnaires to investigate the characteristics of re-finding behavior and compare users' performance in finding and re-finding. Though not significantly different, the study participants used more search tools, combined various strategies to obtain contextual clues of finding process, utilized more complex search tactics, and had more interactions with search engines used. Findings also show that participants spent less time in re-finding than in finding, yet the cognitive loading and difficulties increased in re-finding. Participants were satisfied with the results obtained in re-finding, but they also claimed that the search performance would be better if the system offered more functions to support recall of previous search results. Participants' satisfaction with search performance also varied by task type. Based on the findings, this study recommends that re-finding efficiency may be improved by enhancing recall functionalities in browsers and by using personal information management tools.

Keywords: Web Information Search Behavior; Information Finding; Information Re-finding; Personal Information Management

*　國立臺灣師範大學圖書資訊學研究所研究生
Graduate Student, Graduate Institute of Library and Information Studies, National Taiwan Normal University.
**　國立臺灣師範大學圖書資訊學研究所教授(通訊作者)
Professor, Graduate Institute of Library and Information Studies, National Taiwan Normal University. (To whom all correspondence should be addressed.)
E-mail: htpu@ntnu.edu.tw

臺灣大學學術期刊資料庫

壹、前言

透過搜尋引擎查詢資訊已是網路使用者日常生活的一部分。使用者除了查詢相關資訊，也經常重複查詢曾點選或造訪過的網頁或網站。根據多項研究，個人造訪的網頁中有超過半數以上是再訪的網頁（如Cockburn, Greenberg, Jones, Mckenzie, & Moyle, 2003; Obendorf, Weinreich, Herder, & Mayer, 2007; Tauscher & Greenberg, 1997）。而就搜尋引擎的檢索語句中（註1），一項針對Yahoo!的檢索語句分析研究，39%是來自使用者曾查詢及造訪網頁所使用的檢索語句（Teevan, Adar, Jones, & Potts, 2007），顯見搜尋引擎的流量中，有頗高比例是來自重複查詢。這種同一使用者在一段時間後進行重複查詢的行為即稱之為資訊再尋獲行為（Information Re-finding Behavior）（註2）。相較於網路使用者的一般性檢索行為研究已頗具成果（Spink & Zimmer, 2008），資訊再尋獲行為的研究則是近幾年才開始受到重視（Capra & Pérez-Quiñones, 2005; Teevan, Alvarado, Ackerman, & Karger, 2004）。

資訊再尋獲行為與首次查詢或尋獲（註3）資訊的模式並不相同。就後者情境，使用者在查詢時，並無法明確預期檢索結果，其檢索需求充滿不確定性。但前者因對之前的檢索結果已有期待，因此其檢索需求較為確定；同時，使用者也經常需藉由回想（Recall）與再認（Recognition）等回憶機制，來獲取前次造訪網頁的所在位置；此外，使用者於再尋獲的過程中，也較有目標性及專注性（Capra, Pinney, & Pérez-Quiñones, 2005）。換言之，使用者對於再尋獲資訊多有一定的期望，包括如前次尋獲資訊時所使用的檢索語句、尋獲過程中所經過的路徑（Path）、或路徑中曾遇到有助獲取目標資訊的導航點（Waypoints）等。但面對變動的網路環境，資訊更新頻率高及檢索結果排序的變化常不利於再尋獲資訊（Teevan et al., 2007）。雖然使用者可利用如瀏覽器中的個人書籤（Bookmark）或歷史記錄列表（History list）來保存所尋獲的資訊，但根據調查，其使用率並不高（Bruce, Jones, & Dumais, 2004）。面對上述困難，目前搜尋引擎對再尋獲功能的支援並不多（Aula, Jhaveri, & Käki, 2005）。

由於網路環境之變動性及資訊尋獲情境的多元性，資訊再尋獲與一般資訊尋獲的檢索行為可能存在差異，值得進一步探究。本研究即嘗試瞭解網路使用者資訊再尋獲之檢索行為特性，並探討資訊尋獲與再尋獲行為之關聯。主要包括兩方面：一方面分析使用者進行再尋獲資訊的過程中，所呈現之檢索詞彙、檢索語句與檢索歷程（註4）特性，及所使用的檢索功能與技巧特性；另一方面則分析比較資訊尋獲與再尋獲之檢索成效。研究結果除提供資訊再尋獲行為相關研究之參考，同時也有助發展協助使用者再尋獲資訊之功能設計參考。

臺灣大學學術期刊資料庫

貳、相關研究

一、資訊再尋獲之意涵

資訊尋獲與資訊再尋獲屬於不同的搜尋活動類型。依據Capra, Pinney, & Perez-Quinones（2005）之研究指出（如表一所示），資訊尋獲偏重探索性活動，在檢索過程中，使用者並無法預期檢索結果內容，因此較具不確定感；而資訊再尋獲的檢索過程，則因使用者曾看過檢索結果內容，相較之下，其對檢索結果多有預期，目標資訊也較為明確。此外，在這兩類活動中，使用者的認知活動及使用策略也略有差異。就資訊尋獲行為，使用者多依靠個人既有之知識與直覺，針對檢索結果進行再認，以決定是否符合所需；而資訊再尋獲則常借助回想與再認，強調回憶線索的建構與目標導向式的瀏覽。如Lansdale（1988）提及，回想與再認會相互影響，即使用者會回想目標資訊的部分訊息（如前次使用的關鍵詞、檢索結果的標題、儲存的位置等），藉此縮小檢視範圍，再加以確認。

Jones（2007）以個人資訊管理角度，依資訊之尋獲階段及使用者控制資訊之程度，將個人保存之資訊類型區分如表二所示之四種類型。

表一　資訊尋獲與再尋獲之差異性比較

資訊尋獲	資訊再尋獲
不確定性：是否存在所需資訊？在搜尋引擎中所使用的關鍵字是否妥適？	確定性：曾看過所需資訊，但不知放在何處？在搜尋引擎中所使用的關鍵字是否是之前使用的關鍵字？
再認：此資訊是否正是所搜尋的資訊？	再認與回想：在何處看到此資訊？此處是否就是看到該資訊的位置？情境是重要考量。
策略：直覺、關鍵字搜尋、隨意瀏覽	策略：目標導向式之瀏覽

資料來源：Capra, R.G., Pinney, M., & Pérez-Quiñones, M.A.（2005）. *Refinding is Not Finding Again*. Technical Report TR-05-10, Computer Science, Virginia Tech. Retrieved February 12, 2010, from http://eprints.cs.vt.edu/archive/00000715/01/RefindingIsNotFinding-final.pdf

表二　個人資訊管理之保存資訊類型

資訊是…	使用者能控制	使用者不能控制
使用者之前看過（資訊再尋獲）	A	B
使用者之前沒看過（資訊尋獲）	D	C

資料來源：Jones, W.（2007）. *Keeping Found Things Found: The Study and Practice of Personal Information Management*. San Francisco, CA: Morgan Kaufmann, p.3.

臺灣大學學術期刊資料庫

A類資訊是指使用者從個人資訊空間中，尋找曾看過的資訊；B類是指使用者重返曾經看過的網路資訊；C類是指使用者尋找新的網路資訊；D類指使用者保存但未曾再檢索的資訊。本研究範疇以B及C類資訊為主，特別是B類以使用者不能控制的網路資訊之再尋獲行為為主。而根據Capra（2006）對資訊再尋獲行為之描述，其認為需包括三項基本要素：曾經尋找過的（something was found）、再次需要（is needed again）及一定要找到（must be located）。由此可知，資訊再尋獲與尋找新資訊的尋獲行為仍有相當差異。其實資訊尋獲與資訊再尋獲也可視為同一資訊尋求的兩階段搜尋活動，即使用者進行資訊再尋獲前，會先有資訊尋獲階段。簡言之，本研究之資訊再尋獲是指同一使用者在一段時間後，重複查詢、點選或造訪之前曾點選或造訪之檢索結果或網頁的行為。而資訊再尋獲與尋獲的最大差別在於使用者已知道目標資訊的部分線索。

二、網路資訊再尋獲之實証研究

有關網路資訊再尋獲之相關研究，因網路檢索環境之變動性大，研究情境較難控制，如尋獲與再尋獲之間的時間長短、使用者需求的多樣性及搜尋引擎的變動性等，使得研究設計十分困難。目前相關研究多以探討某特定議題或某因素變項之實証性研究為主，研究方法也以實驗法或訪談法居多。以下簡要敘述一些重要研究主題，並提供實証研究舉例說明。

首先有關資訊再尋獲的行為特性及因應策略是較常見的研究議題。在Teevan（2008）觀察使用者如何重返網頁及所遭遇困難的研究中，其發現使用者常會以「where'd it go?」（網頁跑去哪？）來表示未能造訪先前網頁的困惑。換言之，使用者在重返網頁卻未能及時獲取該網頁時，其最初的反應多希望獲得路徑（path）而非標的（target）資訊，顯示路徑在資訊再尋獲中是一重要線索。此研究結果與Maglio & Barrett（1997）先前研究發現頗為一致，即使用者在回想其檢索歷程時，多會記住路徑中經過的關鍵節點（key nodes）及其檢索路線。這些關鍵節點也就是所謂的導航點（Waypoints），即使用者在尋獲資訊的路徑中所出現的網頁或網站。使用者雖無法記得明確的尋獲路徑，但常能再認或回想路徑中的導航點。Capra & Pérez-Quiñones（2003）依資訊明確度也將導航點分為三種：網頁／網站、網址URL、網頁內容細節等。

再尋獲資訊的過程中，常遭遇許多困難，例如一些非個人所能控制的環境變動，包括網頁內容變動、檢索結果排序改變等。根據Obendorf et al.（2007）研究指出，使用者造訪過的網頁，一天後的更新比例高達69%。檢索結果的改變雖有助尋找新資訊，但其實並不利於再尋獲。根據Teevan et al.（2007）的研究指出，檢索結果排序的改變會造成再尋獲的速度變慢，而點選檢索結果的時間也會變久。面對上述困難，一些使

臺灣大學圖書館 期刊資料庫

用者會嘗試使用瀏覽器所提供的功能來克服，包括使用瀏覽器的上一頁、個人書籤（我的最愛）、瀏覽器中的歷史列等。其中搜尋引擎也是使用者最常利用的再尋獲工具之一（Capra & Pérez-Quiñones, 2005）。

有關資訊再尋獲的影響因素，也是重要的研究範疇。根據Capra（2006）之研究，任務熟悉度與任務類型皆會影響資訊再尋獲之完成時間及搜尋引擎的使用方式。例如使用者對任務的熟悉度越高，就越少使用搜尋引擎進行再尋獲。其進一步將尋獲與再尋獲之相似度分為四類任務：路徑型（Path）、精確型（Exact）、附屬型（Subset）、更動型（Moved）。其發現附屬型較路徑型及精確型任務花費較長的檢索時間、且執行過程也較為困難；就更動型任務，使用者則可能因網頁變動太大，而放棄或根本不想嘗試再尋獲。此外，記憶力也是重要的影響因素之一。Bruce, Jones, & Dumais（2004）曾進行一延遲線索回憶測試（Delayed Cued Recall Test），觀察22位受試者在3至6個月後，依研究者所擬定的線索，回想之前造訪的相關資訊，並使用任何方式重返網頁。其研究結果顯示，多數受試者自認可以成功重返網頁，但仍有部分受試者不知從何下手。依據Teevan et al.（2007）要求119人一小時內回覆之前輸入的檢索語句測試中，有30%受試者回想的檢索語句是錯誤的。根據Elsweiler, Baillie, and Ruthven（2009）針對email之再尋獲研究，再尋獲的確與個人記憶力有關，但記憶愈多並不表示再尋獲的表現就會更好，有時反而表現更差。研究者進一步分析，若記憶的內容是有關email的時間資訊，則表現較佳；但若記得的是email的內容文字，則反而表現較差。

簡言之，上述實証研究提供了一些再尋獲行為的分析結果。而本研究屬於小型實証，其中主要參考Capra（2006）研究的檢索任務設計，希望進一步瞭解網路資訊再尋獲之檢索行為特性；並比較不同檢索任務之尋獲與再尋獲的檢索表現。

參、研究設計

本研究對象採立意抽樣，共有9位國內圖書資訊領域之碩士研究生參與本研究。以圖資碩士生為對象的主要考量，包括圖資領域背景研究生多具備相當資訊素養，其資訊檢索知識與經驗較一般使用者為多；同時，碩士生進行論文研究時，多需大量搜尋及收集資訊，甚且進行個人資訊管理，因此常有再尋獲網路資訊之需求與經驗。

本研究主要採用實驗法，並輔以觀察、問卷及訪談等方法。實驗設計主要參考Bruce, Jones, and Dumais（2004）之架構，分兩階段之實驗方式進行。考量受試者若時間過久未查詢先前尋獲之資料，或甚至完全遺忘，可能等於重新查詢，而較難觀察一般尋獲與再尋獲之關聯。因此本研究以Capra（2006）為參考基礎，採用一週為實驗間隔。此外，本研究也參考Capra（2006）之檢索任務設計，依尋獲與再尋獲任務之相似度擬定路徑型、精確型、附屬型、更動型等

臺灣大學圖書館 期刊資料庫

四種再尋獲之指定檢索任務。精確型任務是指尋獲與再尋獲的目標資訊完全相同，如尋獲某電話號碼，再尋獲時仍以查詢取得該電話號碼為主；路徑型任務則是指尋獲過程中所經過的路徑，如尋獲某字字義是經過某網路字典才取得字義，再尋獲時會以取得該電子字典之網址為主；而附屬型任務，再尋獲時是以目標資訊所包含的資訊為主，如尋獲某旅遊行程資訊，而再尋獲是以查詢其中的飛機航班資訊為目標；更動型任務是指尋獲與再尋獲的資訊有所更動，如尋獲某新聞資訊，而再尋獲是瀏覽此新聞的更新資訊。

實驗開始時會先引導受試者尋獲並取得目標資訊；間隔一週後，再要求受試者再尋獲並取得目標資訊。實驗過程中使用LogSquare（註5）螢幕操作錄製軟體，記錄受試者之尋獲與再尋獲之檢索歷程。再尋獲過程中，本研究並不限制受試者所選擇之再尋獲策略，以客觀觀察受試者可能採用之再尋獲策略，如重新尋找、檢視瀏覽器所保存之暫存記錄、檢視個人保存之檔案記錄等。最後本研究也與受試者進行訪談，以進一步瞭解其再尋獲。

有關指定檢索任務之規劃，為貼近受試者日常生活情境，本研究先設計一網路調查問卷，蒐集受試者相關網路活動，主要區分為生活及工作取向之網路活動。生活取向係指受試者利用網路從事非學術性之活動類型，如購物消費、娛樂休閒、金融交易等；工作取向係指受試者利用網路進行個人學術研究相關活動，如檢索學術文章、查詢圖書館館藏等。問卷內容主要詢問填答者經常及很少進行之生活及工作取向網路活動。網路問卷調查於2009年3月1日起至3月15日，共取得226項網路活動描述。之後再將所收集之活動，依活動類型（生活取向／工作取向）及接觸頻率（經常接觸／不常接觸）進行歸納如表三。

扣除非搜尋類型任務（如收發E-mail、上傳或下載檔案等），本研究參考上述收集到之網路活動描述，共設計8項指定檢索任務（如表四）。

正式實驗進行前，受試者須填寫檢索任務調查表，包括對任務之熟悉度及接觸次數等。在兩階段的實驗中，每位受試者須完成16項指定檢索任務，共取得144筆任務結果，扣除無效樣本16筆（如花費時間太久、或無法尋獲目標資訊而自行放棄等），有效

表三　受試者之網路活動問卷調查整理結果

接觸頻率 活動類型	經常接觸	不常接觸
生活取向	97	78
工作取向	33	18

臺灣大學學術期刊資料庫

表四　指定檢索任務說明

生活取向任務		
資訊尋獲階段	任務相似性	資訊再尋獲階段
任務1：請至youtube尋找公共電視拍攝的「我們的島 – 我的綠建築」	路徑型任務	任務1：請先試著回想上週您在哪裡找影片，接著找到上週的公視影片
任務2：請尋找蘇打綠樂團，吉他手家凱的無名相簿	精確型任務	任務2：請試著找到上週所找到的家凱相簿
任務3：請至2個不同網站尋找3種不同款式的上衣	附屬型任務	任務3：請試著回想上週您在哪裡找衣服，接著請找到上週所點選的這件衣服，並找到直接購買價及剩餘時間
任務4：請尋找並閱讀Library View發表於2009年2月的文章	更動型任務	任務4：請試著尋找上週您閱讀過的圖書館blog文章
工作取向任務		
任務5：請至Wikipedia尋找資訊素養的名詞解釋	路徑型任務	任務5：請試著回想上週您在哪個網頁查資訊素養，接著以該網頁查找知識管理的定義
任務6：請尋找Thomas H. Davenport最新發表於哈佛商業評論的文章	精確型任務	任務6：請試著回想上週您在哪裡尋找Thomas H. Davenport的期刊文章，接著找到上週點選的文章
任務7：請至師大圖書館館藏目錄查詢2008年出版的2本知識管理相關國內書籍	附屬型任務	任務7：請試著回想上週您在哪裡找什麼書，接著找到上週所點選的2本書其中一本的索書號
任務8：請至conferencealerts中的Internet and World Wide Web類別尋找2009年4月進行的DMA Social media研討會相關資訊	更動型任務	任務8：請試著尋找上週所找到的2009年4月22日舉行的DMA Social media研討會官方網站

臺灣大學學術期刊資料庫

樣本共計128筆。實驗結束後，會請受試者填寫檢索結果評估表，包括其滿意度、認知難易度、是否成功找到目標資訊等。最後再進行一對一訪談，主要瞭解其對再尋獲過程的主觀感受及對支援工具之改善建議。

就資料分析，本研究以量化方式分析實驗中所收集之檢索歷程記錄，分析項目如表五所示，主要包括檢索詞彙、檢索語句及檢索歷程之統計分析，並從中歸納資訊尋獲與再尋獲之檢索行為特性。此外，訪談資料則以質化方式分析，藉以瞭解受試者之主觀感受。

因此本研究主要採用實驗法，設計兩階段實驗，蒐集分析受試者之網路檢索記錄並輔以訪談，嘗試歸納網路使用者資訊再尋獲之檢索行為特性，並比較資訊尋獲階段與資訊再尋獲階段之檢索行為差異。

本研究屬小型實証，在研究範圍、實驗樣本及研究設計上仍有諸多限制，簡述如下。首先，再尋獲行為與人類認知心理有相當關聯，例如受試者之認知風格、記憶能力等，本研究僅就受試者之檢索表現進行分析，未能針對受試者之動機與心理層面進行探討。再者，在實驗樣本的選擇，除數量上十分有限，同時，也未能涵蓋其它背景使用者，研究結果無法推論至整體網路使用者。最後，在研究方法的設計上，兩階段的間隔時間以一週為限，也可能影響到資訊尋獲與再尋獲階段所得資料之完整性；同時，指定檢索任務之實驗設計，非屬自然情境，所取得資料之真實性也會受限；此外，為便於收集受試者之尋獲與再尋獲使用記錄，所有受試者皆使用研究者所提供之同一部筆記型電腦，實驗過程中，研究者首先提供一些練習題，讓受試者習慣實驗所使用之筆電操作環境後，才開始進行正式實驗，但使用者未能使用個人習用之電腦，多少可能影響受試者之使用行為。

肆、研究結果

本研究透過兩階段之網路資訊尋獲與再尋獲實驗，初步分析網路資訊再尋獲之檢索行為特性，主要關注其檢索工具的使用特性及檢索技巧；並以檢索詞彙、檢索語句及檢索歷程等三方面的分析，嘗試比較資訊尋獲與再尋獲之檢索歷程與成效差異；此外，也進一步分析不同檢索任務類型之檢索歷程特性與檢索成效。分述如下。

一、網路資訊再尋獲之檢索行為特性

(一) 檢索工具與途徑

根據實驗結果，受試者再尋獲資訊時，仍與資訊尋獲階段一樣使用相同的瀏覽器。受試者所使用的瀏覽器種類不盡相同，而針對再尋獲的支援功能，受試者表示包括可記憶最常瀏覽網站、可設定常用網頁、網址列具助記檢索詞彙或URL功能、可查詢檢索歷史等。就再尋獲之檢索工具與途徑，多數受試者最常利用搜尋引擎進行再尋獲，且不會更換資訊尋獲階段所使用之搜尋引擎；其次為直接輸入URL擷取網頁，即受試者會利用平日熟悉之途徑再尋獲資訊；最後則是使用

臺灣大學圖書館 期刊資料庫

表五　檢索歷程資料之分析項目說明

類　別	分析項目	說　　明
網路瀏覽器使用	網路瀏覽器種類	受試者於每一檢索任務中所使用之網路瀏覽器種類，如Internet Explorer、Mozilla Firefox等。
	功能使用	檢索過程中，受試者所點選之瀏覽器相關功能，如回上頁、回下頁、重新整理、回首頁、停止等按鍵。
網站/網頁取用分析	原有取用途徑	受試者檢索時所使用之網站/網頁取用途徑，如搜尋引擎、bookmark等。
	更換原有取用途徑	檢索過程是否更換原有取用途徑、新取用途徑為何。
檢索詞彙分析	檢索詞彙平均長度	統計受試者輸入中文與英文檢索詞彙之字元數量。
	中文詞彙字數統計	統計中文檢索詞彙之使用字數，以每個單字數計次。
	檢索詞彙概念組合	歸納每一檢索任務之相關檢索詞彙，通常以一完整名詞所表達之概念為主。
檢索語句分析	檢索語句平均長度	統計受試者輸入檢索詞彙之使用數量。
	布林邏輯使用	受試者輸入之檢索語句是否包含布林邏輯運算元，如AND/and、OR/or、NOT/not等。
	修飾符號使用	受試者輸入之檢索語句是否包含修飾符號，如""、site:（限定）、+（即and）、-（即not）、( )（即限定）等。
	檢索語句更動	過程中，受試者是否更動檢索語句，如增加、刪減、更改檢索詞彙等。
檢索歷程分析	平均包含檢索語句數	統計受試者平均每段檢索歷程所輸入之檢索語句數量。
	檢索結果頁瀏覽數量	輸入檢索語句後，受試者瀏覽檢索結果之頁數。
	任務完成時間	統計受試者從執行任務開始至結束之執行時間。
	點選次數	受試者於每一檢索任務期間所點選之總次數。
	相關網頁瀏覽數量	受試者於每一檢索任務期間所開啟並瀏覽之網頁總次數。
	網頁平均停留時間	受試者於每一檢索任務期間停留於每一網頁之時間。
	檢索路徑	受試者於每一檢索任務期間所找到相關網頁之執行動作。

臺灣大學學術期刊資料庫

瀏覽器中的檢索歷史，這類途徑可作為輔助回想之相關線索、或作為檢索來源，如同Komlodi, Marchionini, and Soergel（2007）指出，檢索歷史具有記憶支援、檢索系統使用、或資訊使用等輔助功能。值得注意的是，超過八成之檢索任務，受試者是以單一檢索途徑即成功再尋獲資訊，但仍有二成檢索任務使用超過二種以上之檢索途徑。簡言之，相較尋獲階段集中於搜尋引擎的使用，再尋獲的檢索工具顯得較為多元。

(二) 檢索策略與技巧

根據研究者觀察及訪談結果，如同資訊尋獲階段，受試者在再尋獲階段也以關鍵字檢索與瀏覽為主要檢索策略。但其同時也會回想資訊尋獲階段之相關情境資訊。部分受試者會記得導航網頁、但忘記目標資訊；部分受試者則是記得目標資訊、但卻忘記導航網頁。若為前者情境，受試者多會採取依循資訊尋獲階段之檢索路徑再尋獲資訊；若為後者情境，受試者則多選擇重新尋找。除上述檢索策略外，受試者亦會交互使用關鍵字搜尋及瀏覽策略，從中尋找相關情境線索（如時間、地點等），逐步接近目標資訊，即所謂的定向式（Orienteering）檢索策略（Teevan et al., 2004）。就檢索技巧之運用，由於再尋獲較尋獲階段來得複雜，受試者使用進階檢索功能之比例較高。如表六所示，受試者尋獲與再尋獲資訊過程中經常使用「空格」以表達AND的交集功能。扣除空格之使用情況後，兩階段實驗實際運用布林邏輯運算元與修飾符號之使用比例，皆低於5%，與一般使用者搜尋引擎檢索行為特性頗為相似（Spink, Wolfram, Jansen, & Saracevic, 2001）。整體而言，資訊再尋獲階段使用之進階檢索功能比例較高。

表六　受試者進階檢索功能之使用次數

階段 / 項目	資訊尋獲階段（n＝122）		資訊再尋獲階段（n＝144）	
	使用次數	百分比	使用次數	百分比
空格	44	36.07%	67	46.53%
AND/and/And	－	－	5	3.47%
OR/or/Or	－	－	－	－
NOT/not/Not	－	－	－	－
＋ (加號)	－	－	4	2.78%
－ (減號)	－	－	－	－
“ ” (雙括號)	2	1.64%	6	4.17%
小計	46	37.71%	82	56.95%

臺灣大學學術期刊資料庫

二、網路資訊再尋獲之檢索歷程特性

(一) 檢索詞彙分析

本研究於資訊尋獲與再尋獲二階段各取得199及239筆檢索詞彙，再尋獲所使用之檢索詞彙數量較尋獲階段多了20%。而就檢索詞彙之平均長度，尋獲階段所使用之平均中文詞彙長度為3.21字（Characters），而再尋獲階段則降至2.78字，基本上都屬於單一組合詞（Compound word）；而英文詞彙平均長度則從1.94詞（Words）提高至2.16詞，是由單詞（Uni-gram）轉變為雙詞（Bi-gram）形式。進一步分析檢索詞彙所包含之概念，「網頁來源」與「內容主題」是經常使用的概念（註6）。同時，再尋獲階段所使用之概念數量也多於尋獲階段。究其原因，可能與再尋獲時，因任務描述中的主題概念少且模糊，加上個人記憶限制等影響，受試者較難判斷有哪些可用的檢索詞彙，而須不斷嘗試各種檢索詞彙的組合。換言之，受試者在進行資訊再尋獲時，仍需花費相當精力來形成檢索詞彙。

(二) 檢索語句分析

就資訊尋獲與再尋獲階段所取得之122及144筆檢索語句，再尋獲所使用之檢索語句數量較尋獲階段多了18%。而就檢索語句之平均長度，再尋獲之檢索語句平均包含1.86筆檢索詞彙，略高於尋獲階段的1.63筆，但皆相當簡短。就檢索語句之更動情形，超過七成之檢索語句中的檢索詞彙仍與尋獲階段相同；但再尋獲階段之更動比例仍高於尋獲階段，其中增加與更換檢索詞彙之比例較為明顯，而刪減檢索詞彙之比例則較低（如表七所示）。

上述結果顯示再尋獲之檢索語句相當簡短且不複雜，此與一般使用者之搜尋引擎檢索行為相關研究結果頗為相似（Spink et al., 2001）。但受試者在找不到先前檢索結果時，仍會嘗試增加或更換檢索詞彙來提升尋獲率。

(三) 檢索歷程分析

就資訊尋獲與再尋獲階段所取得之59及

表七　受試者檢索語句之更動次數

階段 / 項目	資訊尋獲階段（n＝122）		資訊再尋獲階段（n＝144）	
	使用次數	百分比	使用次數	百分比
維持原始輸入檢索詞彙	104	85.25%	107	74.31%
增加（add）檢索詞彙	9	7.38%	20	13.89%
刪減（delete）檢索詞彙	7	5.74%	4	2.78%
更換（change）檢索詞彙	2	1.64%	13	9.03%

註：更動次數是指受試者在指定任務檢索中更動檢索語句的次數。

臺灣大學學術期刊資料庫

57筆檢索歷程，二階段之檢索歷程數量差異不大。但就檢索歷程之包含檢索語句數量，再尋獲之檢索歷程平均包括2.67筆檢索語句，高於尋獲階段的2.15筆。從檢索語句使用數量分佈來看，二階段平均每筆檢索歷程中，使用1至2筆檢索語句者居多；但再尋獲階段使用3筆檢索語句以上之比例則高於尋獲階段。由上述分析，二階段的檢索語句數量雖不算多，但顯然再尋獲時，受試者有較多的嘗試，與系統的互動較為頻繁。

(四) 檢索成效分析

本研究之檢索成效主要是指受試者尋獲與再尋獲過程中之完成時間、點擊次數、瀏覽頁數及停留時間。如表八所示，就9位受試者共完成有效的尋獲與再尋獲各64次檢索歷程，在平均完成時間部分，尋獲階段平均花費130秒即尋獲目標資訊；而再尋獲時則更為快速，降低至111秒。此外，再尋獲階段之點擊次數、相關網頁瀏覽數量及網頁平均停留時間皆低於尋獲階段。此外，分析受試者平均瀏覽檢索結果頁數，尋獲與再尋獲階段分別為1.42及2.65頁，顯示再尋獲時瀏覽較多檢索結果，究其原因，可能與受試者為成功再尋獲資訊，會多瀏覽檢索結果以尋找更多回想線索。綜合上述資料，受試者再尋獲時，能以較少點擊次數及較短瀏覽時間即尋獲目標資訊，顯示再尋獲較尋獲階段來得有效率。特別是其能以較短的停留時間瀏覽更多檢索結果，顯示受試者可以很容易且快速地再認及判斷目標資訊，因此其停留時間得以降低許多。

但由前述檢索歷程相關分析來看，受試者花費較多心力於形成有效的檢索詞彙與語句，也與系統有較多的互動，顯示其所花費心力並不亞於資訊尋獲階段，只是花費心力的內涵有所差異。根據問卷評估及訪談，一般而言，受試者對再尋獲的結果多感到滿意。但受試者表示有時受個人記憶限制或目標資訊改變之影響，需不斷回憶及辨別相關線索，造成一些認知負擔，也增加檢索的難度，若能有適當回想支援功能，再尋獲的檢索成效其實應該更高。

表八　資訊尋獲與再尋獲之檢索成效比較

階段 / 項目	資訊尋獲階段（n＝64）	資訊再尋獲階段（n＝64）
平均完成時間	130秒（sd＝88.24）	111秒（sd＝98.79）
平均點擊次數	9.78（sd＝8.53）	7.67（sd＝6.54）
平均瀏覽相關網頁數	7.77（sd＝7.37）	6.48（sd＝5.84）
每一網頁平均停留時間	10秒（sd＝5.88）	6秒（sd＝3.29）

臺灣大學學術期刊資料庫

三、不同檢索任務類型之檢索成效

根據相關研究指出（Xie, 2009），檢索任務類型對檢索行為具有相當影響力。因此本研究進一步分析不同任務類型與再尋獲檢索行為之關聯。研究中主要有四種類型分析，包括：(1)任務之目的屬性：即尋找生活或工作取向資訊之檢索任務；(2)對檢索任務之接觸程度；(3)對檢索任務之熟悉程度；及(4)尋獲與再尋獲階段之相似性，即依尋求目標資訊之相似性分為路徑、精確、附屬及更改等四種類型。由於本研究樣本有限，因此本研究嘗試綜合實驗、問卷及訪談結果，來呈現尋獲與再尋獲之檢索成效。但因結果不易以量化數據呈現，同時僅以文字說明亦不易分辨各項成效之關聯，因此本文嘗試設計表九之呈現方式，初步提供研究者之主觀評估比較結果。此表僅作為呈現本實証研究中，受試者於實驗中之檢索成效表現，並無法推論至其它或全體網路使用者。同時，主觀評估之完整性與正確性仍需更多系統化及科學化之實証研究驗証。

以任務的目的屬性來看，生活取向任務不論在尋獲及再尋獲階段，整體檢索表現皆佳；但在完成時間及停留時間等項目，再尋獲較尋獲階段為短，顯示受試者再尋獲時，能更快速獲取目標資訊。而就工作取向任務，整體檢索表現不若生活取向任務來得有效率，再尋獲時，其點擊次數及瀏覽網頁數變少，顯示受試者花費更多心力於瀏覽及回想尋獲時的線索。根據訪談，受試者表示在生活取向任務所花費之認知心力遠低於工作取向任務。根據問卷中對檢索結果之滿意度調查，生活取向任務之滿意度也高於工作取向任務，且受試者認為，再尋獲時的工作取向任務之檢索困難度也高於生活取向任務。簡言之，對受試者而言，尋獲與再尋獲工作取向資訊較生活取向資訊來得不易，且再尋獲更顯困難。

以任務之接觸次數來看，高接觸任務的二階段表現皆佳，但再尋獲階段之花費時間明顯降低，且檢索歷程所包含之檢索語句數量及檢索結果頁瀏覽數量也都變少，顯示受試者愈常接觸的任務，其檢索成效愈佳。就低接觸任務而言，二階段之檢索語句長度較短、也較少使用進階檢索功能；此外，其更動檢索語句之比例也低於高接觸任務，顯示受試者對於較少接觸的任務，其檢索歷程中，受試者花費較多時間在瀏覽檢索結果，而較少花費在檢索功能的利用上。根據問卷中對檢索結果之滿意度調查，高接觸任務於尋獲階段之滿意度較高，而再尋獲階段則在低接觸任務的滿意度較高，且再尋獲之低接觸任務之困難度也高於高接觸任務。簡言之，任務接觸程度對尋獲及再尋獲皆有一些影響，但影響項目略有差異。此與Capra（2006）研究結果頗為一致，即經常接觸的任務，使用者多已發展出可用的方法來取得資訊，因此再尋獲並不困難；但極少接觸的任務，再尋獲其實等於重新尋獲，難度是相同的。

以任務熟悉度來看，高熟悉任務之二階段表現皆佳，各項目表現相較高接觸任務表

臺灣大學圖書館 期刊資料庫

表九　不同檢索任務類型之尋獲與再尋獲檢索成效比較

任務類型 / 評估項目	目的屬性				任務接觸次數				任務熟悉度				尋獲與再尋獲之相似性							
	生活取向		工作取向		常接觸		少接觸		高熟悉		低熟悉		路徑型		精確型		附屬型		更改型	
	C	B	C	B	C	B	C	B	C	B	C	B	C	B	C	B	C	B	C	B
完成時間	◎	●				●	◎		◎	●			◎	●				●	◎	
點擊次數		●	◎		◎			●	◎	●			◎	●				●	◎	
瀏覽網頁數		●	◎			●	◎		◎	●			◎	●				●	◎	
停留時間	◎	●			◎	●			◎	●			◎	●				●	◎	
檢索語句長度	◎	●					◎	●		●	◎		◎	●			◎	●		
檢索歷程平均檢索語句數量	◎	●			◎	●			◎	●			◎	●		●			◎	
結果頁瀏覽數	◎	●			◎	●			◎	●				●			◎	●	◎	
進階功能	◎	●					◎	●		●	◎		◎	●			◎			●
檢索語句更動		●	◎		◎			●	◎	●			◎	●			◎	●	◎	
檢索滿意度	◎	●			◎			●	◎	●			◎	●					◎	●
檢索難易度	◎	●			◎	●			◎	●			◎	●				●	◎	

註：1. C表資訊尋獲階段；B表資訊再尋獲階段。◎表資訊尋獲階段有較好的檢索表現；●表資訊再尋獲階段有較好的檢索表現；若◎●二者並存表二者檢索表現皆佳；若二者皆空白，則表檢索表現皆差。

2. 本研究之接觸程度及熟悉程度皆由使用者自行判斷所得。

現更佳。但整體而言，受試者尋獲與再尋獲高熟悉度任務所花費之認知心力遠低於低熟悉度任務。值得注意的是，根據問卷中對檢索結果之評估，尋獲及再尋獲高熟悉度任務之滿意度遠高於低熟悉度任務，且再尋獲低熟悉度任務之困難度遠高於高熟悉度任務。顯示受試者對不熟悉任務所感到的困難度更甚於少接觸的任務。

以四類不同目標資訊之再尋獲任務來看，路徑型任務在二階段的表現皆佳。其次為附屬型任務，且再尋獲階段之表現較尋獲時來得突出。而更改型任務在尋獲階段有不錯的檢索表現，但在再尋獲階段則表現不佳，主要來自目標資訊之更動，再尋獲須花費較多的時間及心力。精確型任務的表現最差，根據研究者觀察，此與受試者對任務內

臺灣大學圖書館 期刊資料庫

容的熟悉度有所關聯，當受試者對該任務內容主題不熟悉時，其尋獲已有一定難度，當再尋獲時其困難度並未降低。反映在問卷之檢索結果評估，也相當一致。即受試者對路徑型任務之滿意度最高，附屬型任務次之，且對更改型及精確型任務感到困難度較高。

伍、結論

本研究主要藉由指定檢索任務之實驗情境，分析網路使用者之資訊再尋獲檢索行為特性。同時，也分析比較尋獲與再尋獲兩階段之檢索歷程特性。此外，也進一步分析比較不同任務類型與兩階段之關聯。研究結果顯示，就檢索行為特性，搜尋引擎雖同時為資訊尋獲與再尋獲之最主要檢索工具，但再尋獲的檢索工具顯得較為多元；再尋獲階段也以關鍵字檢索與瀏覽為主要檢索策略，但其同時也會交互使用這些策略，從中尋得相關檢索情境線索；再尋獲所使用之檢索技巧也較為複雜。就檢索歷程，雖然再尋獲較尋獲所花費時間較短、較有效率，但其認知心力及檢索困難度也增加不少。一般而言，受試者對再尋獲的結果多感到滿意；但若系統能提供更多回想支援功能，再尋獲的檢索成效其實應該更高。就檢索任務類型的影響而言，尋獲與再尋獲工作取向資訊較生活取向資訊來得不易，且再尋獲更顯困難；經常接觸的任務，使用者多已發展出可用的方法來取得資訊，因此再尋獲並不困難，但極少接觸的任務，再尋獲其實等於重新尋獲，難度是相同的；受試者對不熟悉任務所感到的困難度更甚於少接觸的任務；路徑型任務之再尋獲滿意度最高，附屬型任務次之，對更改型及精確型任務則感到困難度較高。

Capra（2006）比較多個檢索任務之相關變項發現，資訊尋獲與再尋獲行為受使用者對檢索任務之熟悉度（familiarity）、接觸次數（frequency）及相似度（similarity）之影響。故本研究以此為基礎，以指定檢索任務方式進行實驗，藉以觀察、分析與歸納資訊再尋獲之檢索行為特性，並嘗試比較本研究結果與Capra（2006）之異同。由於其任務設計以生活取向任務為主，未涉及工作取向任務。基於網路使用者之資訊需求多元，本研究另增加工作取向之檢索任務，以期獲得較完整之觀察結果。就生活取向任務，本研究結果與Capra（2006）頗為一致，重要結果包括資訊尋獲與再尋獲確實存在差異，特別是再尋獲階段之任務困難度有增加趨勢。同時，受試者在兩階段皆未更換所使用之搜尋引擎，顯示受試者之檢索習慣相當穩定。此外，任務熟悉度越高者，較少使用搜尋引擎。整體任務執行情況，資訊再尋獲階段之平均花費時間、點擊次數、瀏覽相關網頁數、及網頁平均停留時間相較於資訊尋獲階段呈下降之趨勢，也符合Capra（2006）實驗所得結果。

就與Capra（2006）相異之處，在四種不同相似性類型任務中，本研究使用者在精確型任務之檢索表現較差，而Capra（2006）則是以附屬型任務表現較差。造成此差異，可能與本研究之實驗設計有關，如

臺灣大學圖書館期刊資料庫

本研究之附屬型任務中出現一些圖片，受試者得以作為再尋獲之相關線索；而精確型任務所提供之資訊相當少，增加受試者再尋獲的難度。究竟此四種類型任務之檢索難度如何，仍需進一步探究。此外，本研究增加工作取向之任務設計，其整體檢索表現不若生活取向任務來得有效率。再尋獲時，其點擊次數及瀏覽網頁數變少，顯示受試者花費更多心力於瀏覽及回想尋獲時的線索。根據訪談，受試者表示在生活取向任務所花費之認知心力遠低於工作取向任務。根據問卷中對檢索結果之滿意度調查，生活取向任務之滿意度也高於工作取向任務，且受試者認為，再尋獲時的工作取向任務之檢索困難度也高於生活取向任務。簡言之，對受試者而言，尋獲與再尋獲工作取向資訊較生活取向資訊更為不易，且再尋獲更顯困難，值得更進一步研究。

根據上述結果，本研究針對系統改善及個人資訊管理等兩層面提出一些建議。有關系統改善部分，研究發現除利用搜尋引擎外，檢索歷史為受試者再尋獲時最常使用之取用途徑，主要包括瀏覽器中的書籤、使用歷史記錄、網頁暫存等功能；同時，受試者再尋獲資訊時，多會嘗試回想及再認相關情境線索；此外，根據記憶之相關研究，回想的情境若能與事件發生時的情境一致，將有助使用者回想及再認。以目前瀏覽器所記錄之使用記錄都十分簡略，若能保存使用者先前實際且完整的檢索歷史，包括當時之檢索過程及結果，則對使用者再尋獲資訊時將有不少助益。有關個人資訊管理層面，許多使用者仍僅依賴關鍵字檢索工具來再尋獲資訊，但研究結果發現，導航點及路徑其實是輔助再尋獲的重要工具，與其指導使用者如何形成良好的檢索詞彙，不如指引其對網站或網頁之資訊架構多做瞭解，加深其對起始點、路徑、目的地之印象，等再尋獲時自然就較能正確地回想。此外，使用者亦可善加利用一些網路註記（Annotation）或標記（Tagging）工具如Clipmarks（註7）、Diigo（註8）、Delicious（註9）等，協助個人註記所保存的資訊，這對之後再尋獲也會有不少助益。

針對未來研究建議，以目前網路搜尋工具多以資訊尋獲觀點設計，鮮少以資訊再尋獲觀點進行探討，如何設計一具有保存個人使用情境之檢索系統，是值得進一步發展的方向。同時，本研究僅以臺灣地區圖書資訊領域研究生之小樣本進行實証，未來若能增加其它領域或不同人口背景之研究對象，將有助瞭解不同類型使用者的資訊尋獲與再尋獲行為差異與檢索成效，並提升研究結果之概括化（Generalization）。此外，本研究之實驗場域為網路之搜尋引擎平台，且著重於較一般性之資訊需求，未來若能以特定之機構組織做為研究對象，將有助瞭解不同機構組織之再尋獲行為特性，並可比較組織與個人情境之再尋獲行為差異。

註釋

註 1：本研究之檢索語句（Query）是指

臺灣大學學術期刊資料庫

使用者在搜尋引擎輸入的單一檢索字串，其中可能包括多個檢索詞彙（Search term）及指令。

註 2：本研究所指之資訊再尋獲為「尋找先前看過的資訊」，與資訊尋獲的「尋找新資訊」有所不同。

註 3：本研究中所指之「查詢」（Search）與「尋獲」（Find）皆指使用者利用搜尋引擎檢索資訊之活動類型。前者如使用者以關鍵字搜尋，獲取大量檢索結果；而後者則是使用者由其中發現感興趣、可能進行保存及利用資訊之活動。二者在檢索過程中雖略有不同，但基本上不論尋獲或再尋獲之前置步驟皆需有查詢活動，而本研究著重於尋獲及再尋獲行為之比較，因此查詢與尋獲在本研究中並不特別區分。

註 4：本研究之檢索歷程（Session）是指使用者為滿足其資訊需求，在一段時間內所進行的檢索過程，主要由一連串檢索語句所構成。在搜尋引擎之檢索環境中，一段檢索歷程一般多以30分鐘為限（Catledge & Pitkow, 1995）。

註 5：http://www.mangold-international.com/products/software/logsquare.html

註 6：「網頁來源」是指受試者尋獲資源之管道，如該網站/網頁之URL或名稱；而「內容主題」則指尋獲資源之內容主題，如該網站/網頁內容出現之關鍵字。

註 7：可分享、分類及註記的網路剪貼簿。網址：http://clipmarks.com/

註 8：可增加註記及標示網頁重點、分類標籤並支援不同網路瀏覽器。網址：http://www.diigo.com/

註 9：網路書籤系統。網址：http://delicious.com/

參考書目

Aula, A., Jhaveri, N., & Käki, M. (2005). Information search and re-access strategies of experienced Web users. *WWW '05*, 583-592. Retrieved February 12, 2010, from http://www2005.org/cdrom/docs/p583.pdf

Bruce, H., Jones, W., & Dumais, S. (2004). Information behavior that keeps found things found. *Information Research, 10*(1). Retrieved February 12, 2010, from http://InformationR.net/ir/10-1/paper207.html

Capra, R. G. (2006). *An investigation of finding and refinding information on the web*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Charlottesville.

Capra, R.G., & Pérez-Quiñones, M.A. (2003). Re-finding found things: An exploratory study of how users re-find information. Retrieved February 12, 2010, from http://arxiv.org/ftp/cs/papers/0310/0310011.pdf

臺灣大學學術期刊資料庫

Capra, R. G., & Pérez-Quiñones, M. A. (2005). Using web search engines to find and refind information. *Computer, 38*(10), 36-42.

Capra, R.G., Pinney, M., & Pérez-Quiñones, M.A. (2005). *Refinding is not finding again*. Technical Report TR-05-10, Computer Science, Virginia Tech. Retrieved February 12, 2010, from http://eprints.cs.vt.edu/archive/00000715/01/RefindingIsNotFinding-final.pdf

Catledge, L., & Pitkow, J. (1995). Characterizing browsing strategies in the World-Wide Web. *WWW '95*, 1065-1073.

Cockburn, A., Greenberg, S., Jones, S., Mckenzie, B., & Moyle, M. (2003). Improving Web page revisitation: Analysis, design and evaluation. *IT & Society, 1* (3), 159-183.

Elsweiler, D., Baillie, M., & Ruthven, I. (2009). On understanding the relationship between recollection and refinding. *Journal of Digital Information, 10*(5). Retrieved February 12, 2010, from http://journals.tdl.org/jodi/article/download/436/542

Jones, W. (2007). *Keeping found things found: The study and practice of personal information management*. San Francisco, CA: Morgan Kaufmann.

Komlodi, A., Marchionini, G., & Soergel, D. (2007). Search history support for finding and using information: User interface design recommendations from a user study. *Information Processing and Management, 43*, 10-29.

Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics, 19*, 55-66.

Maglio, P.P., & Barrett, R. (1997). How to build modeling agents to support web searchers. *The 6th International Conference on User Modeling*, Sardinia, Italy. Retrieved February 12, 2010, from http://wwwcssrv.almaden.ibm.com/wbi/papers/um97.pdf

Obendorf, H., Weinreich, H., Herder, E., & Mayer, M. (2007). Web page revisitation revisited: Implications of a long-term click-stream study of browser usage. *CHI '07*, 597-606.

Spink, A., & Zimmer, M. (eds.) (2008). *Web search: Multidisciplinary perspectives*. Berlin: Springer.

Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology, 52*(3), 226-234.

Tauscher, L., & Greenberg, S. (1997). How people revisit web pages: Empirical findings and implications for the design of history systems. *International Journal of Human-Computer Studies, 47*(1), 97-137.

臺灣大學圖書館 期刊資料庫

Teevan, J. (2004). *How people re-find information when the web changes*. MIT Computer Science and Artificial Intelligence Lab Technical Report. Retrieved February 12, 2010, from http://dspace.mit.edu/handle/1721.1/30480

Teevan, J. (2008). Where'd it go?: How people ask after lost Web information. *ASIST 2008 Annual Meeting*. Retrieved February 12, 2010, from http://people.csail.mit.edu/teevan/work/publications/papers/asist07.pdf

Teevan, J., Adar, E., Jones, R., & Potts, M.A.S. (2007). Information re-retrieval: Repeat queries in Yahoo's logs. *ACM SIGIR' 07*, Amsterdam, The Netherlands.

Teevan, J., Alvarado, C., Ackerman, M.S., & Karger, D.R. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. *ACM CHI'04*. Retrieved February 12, 2010, from http://groups.csail.mit.edu/haystack/documents/papers/2004/chi2004-perfectse.pdf

Xie, I. (2009). Dimensions of tasks: Influences on information-seeking and retrieving process. *Journal of Documentation, 65*(3), 339-366.

（投稿日期：2010年2月22日　接受日期：2010年3月8日）

臺灣大學學術期刊資料庫

An Exploratory Study on the Re-finding Behavior on the Web

Xin-Yu Jiang*, Hsiao-Tieh Pu**

Extended Abstract

1. Introduction

Information re-finding behavior is the repeated query and re-visitation of specific information content that was previously located by the same user. In information re-finding, users have more clues about where to locate the target information than in exploratory information search activities (Capra, Pinney, & Pérez-Quiñones, 2005). In the web environment, the occurrence of information re-finding is no less than finding (i.e., initial searching of information) (Cockburn et al., 2003; Obendorf et al., 2007; Tauscher & Greenberg, 1997), yet the re-finding of information is often impeded by fast content changes as well as the frequent updated result rankings in search engines (Teevan et al., 2007). Also, the current search engines provide insufficient support for information re-finding activities (Aula, Jhaveri, and Käki, 2005). This study thus examined the user behaviors of information re-finding and the characteristics of re-finding search sessions; it then compared the search performances of information finding and re-finding. Results of this study may be used to improve the design of web search engines.

2. Research Design

This study employed a combination of experiments, observations, interviews, and questionnaires to obtain data needed for analyses. Purposive sampling was used

* Graduate Student, Graduate Institute of Library and Information Studies, National Taiwan Normal University

** Professor, Graduate Institute of Library and Information Studies, National Taiwan Normal University (To whom all correspondence should be addressed.)
E-mail: htpu@ntnu.edu.tw

Note. This extended English abstract is supplied by the JLIS editors and approved by the authors.
To cite this article in APA format: Jiang, X. Y., & Pu, H. T. (2010). An exploratory study on the re-finding behavior on the web. *Journal of Library and Information Studies, 8*(1), 29-57. [Text in Chinese].
To cite this article in Chicago format: Jiang, Xin-Yu, and Hsiao-Tieh Pu. "An exploratory study on the re-finding behavior on the web." *Journal of Library and Information Studies* 8, no. 1 (2010): 29-57. [Text in Chinese].

臺灣大學學術期刊資料庫

to recruit nine masters' students in library and information science to participate in a two-staged experiment which required the participants to locate information on the web in the first stage and relocate the same information after one week's interval.

The information search and re-finding tasks assigned to the participants were designed based on Capra's (2006) categorization of four relationship types between initial search and re-finding: "exact" – tasks are exactly the same in both sessions; "path" the re-finding tasks involves looking for information on the same path as the information that was first found; "subset" – the re-finding task involves looking for a subset of information; "moved" the re-finding tasks involves looking for information that has moved since it was found.

In this study, each participant carried out eight tasks, four of which represented everyday information search, while the other four represented serious work-related search (see Table 1).

LogSquare, a screen logging software, was used to record participants' search processes. Prior to the experiment, the participants were asked to answer a questionnaire about their familiarity with the search tasks and whether they frequently conducted such search tasks in real life situations. After the experiment they were surveyed again by questionnaire as well as interviewed individually to answer questions regarding the perceived level of difficulty of the search tasks, satisfaction with the search results, and how they felt the tools should be improved for better results. The aforementioned procedures produced 144 search sessions, among which 128 constituted a valid sample for analysis. The following aspects of each search sessions were examined (see Table 2).

3. Findings

(1) Characteristics of Information Re-Finding Behaviors

A. Search tools & path of access: in both stages, participants employed mainly search engines to look for target information. But the tools and paths of access appeared to be more diverse in information re-finding, for example, participants may type the URL into the browser or use the browser's search history to relocate information. Participants found it useful for information re-finding activities when the browsers were equipped with functionalities such as remembering most frequently browsed pages, storing frequently used pages, automatic completion of URLs, and searchability of browse history.

臺灣大學圖書館 期刊資料庫

Table 1: the Eight Tasks Assigned to the Participants

Everyday Information Search Tasks		
Initial Information Search Tasks	Task Types	Information Re-finding Tasks
Task 1: Please find the film X produced by the Taiwan PTS in YouTube.	Path	Task 1: Please recall where you had found the film X produced by the Taiwan PTS and relocate it.
Task 2: Please go to the Taiwanese social networking site, *Wretch* (www.wretch.cc), and find the online photo album of the guitarist of the band, Sodagreen.	Exact	Task 2: Please recall the album of the Sodagreen guitarist that you found last week and relocate it.
Task 3: Please go to two different online audition sites and find three different outfits in those sites.	Subset	Task 3: Please relocate the three outfits that you found in the audition sites and find out the prices and the time left for placing bids.
Task 4: Please go to the *Library View* blog and read the postings of the February, 2009.	Moved	Task 4: Please recall the *Library View* postings that you've read in the last week and relocate them.
Work-related Information Search Tasks		
Task 5: Please find the definition of "information literacy" in Wikipedia.	Path	Task 5: Please recall where you had found out the definition of "information literacy" in the last week, and use the same resource to find the definition of "knowledge management."
Task 6: Please find the most recent article by Thomas H. Davenport in *Harvard Business Review*.	Exact	Task 6: Please recall the article by Thomas H. Davenport that you found last week and relocate it.
Task 7: Please find two books on knowledge management National Taiwan Normal University Library's OPAC.	Subset	Task 7: Please recall which books you had found last week and where you found them. Relocate one of them and tell me what the call numbers is.
Task 8: Please go to the *Conference Alerts* web site and find information about *DMA Social Media*, to be held in April, 2009, under the category of Internet and World Wide Web.	Moved	Task 8: Please relocate the information that you saw last week about the DMA Social Media conference to be held on April, 22, 2009.

臺灣大學學術期刊資料庫

Table 2: Aspects Examined in Each Search Session

Aspect	Item	Description
Web browser usage	Web browser used	The various web browsers used by the participants such as Internet Explorer, Mozilla Firefox, etc.
	Web browser functions used	The web browser functions used during each search session such as moving back, moving forward, reloading the current page, moving to the browser default page, stop, etc.
Path of access	Original path of access	The sequential use of search engines, bookmarks, or other waypoints by the participants to locate a certain web content
	Alternated path of access	Whether participants changed the sequential use of waypoints in accessing the target web content and the new path of access
Query terms	Average length of query terms	The average number of the Chinese and English characters in the query terms input by the participants
	Average number of Chinese characters in query terms	The average number of Chinese characters in the query terms
	Combination of terms	The average number of concepts in each search task, usually in the form of nouns combined by query terms
Query syntax	Average length of each query	The average number of query terms input by the participants
	Boolean logic usage	Whether a search query contains Boolean operators such as And/and, Or/or, or Not/not
	Refined delimiters usage	The delimiters used to refine a search query, e.g., “” for phrase search, “site:” for search within a specific website, etc.
	Alternation in search query	Whether a participant add, delete, or change terms during a search query
Search session	Average number of queries	The average number of queries input by the participants in each search session
	Number of web pages browed	The number of web pages browsed after the participant enters each search query
	Duration of the search	The time used by the participant to complete the search task
	Number of clicks	The total number of clicking by the participant during each session
	Number of browsed web pages	The total number of web pages accessed and browsed by the participants during each session
	Average duration of web page browse	The average time spent on each accessed web page during each session
	Search path	The search path participants used during each re-finding session

臺灣大學學術期刊資料庫

Search Performance in Different Task Types: Initial Search vs. Re-Finding

Task features / Observed items	Information Search type				Experiences with similar tasks				Task familiarity				Similarity between initial search & re-finding by task types							
	Everyday Info		Work related		Many		Few		High		Low		Path		Exact		Subset		Moved	
	C	B	C	B	C	B	C	B	C	B	C	B	C	B	C	B	C	B	C	B
Duration of the search task	◎	●				●	◎		◎	●			◎	●				●	◎	
Number of clicks		●	◎		◎			●	◎	●			◎	●				●	◎	
Number of Web pages browsed		●	◎			●	◎		◎	●			◎	●				●	◎	
Time spent on each web page	◎	●			◎	●			◎	●			◎	●				●	◎	
Length of search query	◎	●					◎	●		●	◎		◎	●			◎	●		
Average number of search queries in each session	◎	●			◎	●			◎	●			◎	●		●			◎	
Number of page browse in search result	◎	●			◎	●			◎	●				●			◎	●	◎	
Advanced function usage	◎	●					◎	●		●	◎		◎	●			◎			●
Alteration of search query		●	◎		◎			●	◎	●			◎	●			◎	●	◎	
Satisfaction with the search result	◎	●			◎			●	◎	●			◎	●					◎	●
Perceived difficulty with the search task	◎	●			◎	●			◎	●			◎	●				●	◎	

Note:

i. C: the initial information search stage; B: the information re-finding stage; ◎: participants performed well in the initial information search stage; ●: participants performed well in the re-finding stage; ◎ ● (co-exist): participants performed similarly well in both stages; blank: participants performed similarly unwell in both stages.

ii. The past experiences with the similar search tasks and task familiarity were self-reported by the participants.

臺灣大學學術期刊資料庫

B. Search strategies & skills: in both stages participants employed mainly keyword search and browsing to look for information. But in re-finding activities the two strategies were used in combination to obtain contextual clues to the tasks. Advanced search skills, mainly the use of spaces between keywords (meaning a Boolean logic "AND" search), are used more often in information re-finding.

(2) Characteristics of Search Sessions

A. Search terms: the total number of search terms used in the re-finding stage (239) is higher than that of the initial search stage (199). The average length of Chinese term is shorter in re-finding (3.21 characters for information re-finding vs. 2.78 for initial search). Also, the use of English terms shifted from using uni-gram to using bi-gram, with term length increasing from 1.94 words for the initial search stage to 2.16 words for the re-finding stage. Search terms were formed based on participants' remembered source pages and the subject topic of the target information. The concepts used in information re-finding were more than those used in initial search stage, which suggests that information re-finding requires rather heavy cognitive efforts from participants.

B. Search queries: the total number of search queries in the re-finding stage (144 queries) was more than that of initial search (122 queries). The average length of queries was also longer (1.83 words for re-finding vs. 1.63 words for initial search). In both stages the search queries were all short and uncomplicated. But the re-finding stage saw more altered search queries. That is, participants added or altered search terms to enhance success rates in information re-finding.

C. Search session: the numbers of search sessions were nearly equal (59 for initial search vs. 57 for re-finding). But the average number of queries in each search session was higher in re-finding (2.67 queries for re-finding vs. 2.15 for initial search). It suggests that participants conducted more trial searches and interacted with the systems more in information re-finding.

D. Search performance: participants were more efficient in information re-finding. It took the participants less time (111 seconds for re-finding

臺灣大學圖書館 期刊資料庫

vs. 130 for initial search), less clicks (7.67 for re-finding vs. 9.78 for initial search), less page browsing (6.48 pages for re-finding vs. 7.77 for initial search), and shorter average duration of each page browse (6 seconds for re-finding vs. 10 seconds for initial search) to complete an information re-finding task.

(3) Search performances in different search tasks (based on quantitative evidence and researchers' subjective evaluation)

A. Information search types: work related information was harder to find than everyday information in initial search stage, and was even harder to relocate in re-finding. Participants were generally more satisfied with the everyday information search results than with the work related searches.

B. Experiences with similar search tasks: participants with more experiences of the similar search tasks have built a repertoire of search strategies, and thus it wasn't difficult for them to relocate the same information in re-finding. Participants with few or none experience with certain tasks found it equally hard to locate information in initial search and in re-finding. But in terms of satisfaction with search results, more experiences yielded higher satisfaction in the initial search stage; in contrast, fewer experiences with certain tasks yielded higher satisfaction in information re-finding.

C. Task familiarity: it required participants with low task familiarity more cognitive efforts to carry out a search task. The perceived difficulty might be higher than that resulted from fewer experiences with similar tasks.

D. Similarity between initial search & re-finding by task types: the four task types demonstrate certain levels of similarity between the initial search stage and the re-finding stage in terms of search performances and satisfaction with search results. The four tasks by descending similarity were path, subset, moved, and exact.

4. Conclusion

This study showed that, although search engines were the major tool used to locate information in both the initial searching and re-finding stages, tools used in the latter appeared to be more diverse (e.g., browser provided history or bookmarking functions). In re-finding, keyword searching and browsing were used in combination to identify contextual clues

臺灣大學學術期刊資料庫

leading to re-locating of the original information content. The search tactics employed to relocate information were more complicated. It took shorter time for participants to re-locate information content than in initial search and the retrieval re-finding appeared to be more efficient. But the search tasks appeared to be more complicated for participants in re-finding, and they spent more cognitive efforts to relocate the information.

Participants were mostly satisfied with the search results. But they also considered it useful if systems provided certain functionalities to support the recall and relocation of information. In terms of search types, re-finding of work-related information was more difficult than finding everyday information. Experiences and task familiarity affected the efficiency of information re-finding. But in the case of few experiences and low task familiarity, searching and re-finding appeared to be similarly difficult for participants. Participants' satisfaction with search results also varies by task type. Satisfaction was highest with the "path" tasks, and then the "subset;" "moved" and "exact" tasks were more difficult for participants and they were less satisfied with their search performance.

Most of the findings were consistent with Capra (2006) except the search performance in the four task types. This might have resulted from the various designs of actual tasks assigned to participants in the two studies.

Based on the findings, the authors recommend that browsers may be improved by providing functionalities that facilitate the recall and the re-finding of information, such as preserving complete search history. Also, users may improve their re-finding efficiency by gaining a deeper understanding of web content architectures and by using annotation and tagging tools to store information located.

References

Aula, A., Jhaveri, N., & Käki, M. (2005). Information search and re-access strategies of experienced Web users. *WWW '05*, 583-592. Retrieved February 12, 2010, from http://www2005.org/cdrom/docs/p583.pdf

Bruce, H., Jones, W., & Dumais, S. (2004). Information behavior that keeps found things found. *Information Research, 10*(1). Retrieved February 12, 2010, from http://InformationR.net/ir/10-1/paper207.html

Capra, R. G. (2006). *An investigation of finding and refinding information on the web.* Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Charlottesville.

Capra, R.G., & Pérez-Quiñones, M.A. (2003). Re-finding found things: An exploratory

臺灣大學學術期刊資料庫

study of how users re-find information. Retrieved February 12, 2010, from http://arxiv.org/ftp/cs/papers/0310/0310011.pdf

Capra, R. G., & Pérez-Quiñones, M. A. (2005). Using web search engines to find and refind information. *Computer, 38*(10), 36-42.

Capra, R.G., Pinney, M., & Pérez-Quiñones, M.A. (2005). *Refinding is not finding again*. Technical Report TR-05-10, Computer Science, Virginia Tech. Retrieved February 12, 2010, from http://eprints.cs.vt.edu/archive/00000715/01/RefindingIsNotFinding-final.pdf

Catledge, L., & Pitkow, J. (1995). Characterizing browsing strategies in the World-Wide Web. *WWW '95*, 1065-1073.

Cockburn, A., Greenberg, S., Jones, S., Mckenzie, B., & Moyle, M. (2003). Improving Web page revisitation: Analysis, design and evaluation. *IT & Society, 1* (3), 159-183.

Elsweiler, D., Baillie, M., & Ruthven, I. (2009). On understanding the relationship between recollection and refinding. *Journal of Digital Information, 10*(5). Retrieved February 12, 2010, from http://journals.tdl.org/jodi/article/download/436/542

Jones, W. (2007). *Keeping found things found: The study and practice of personal information management*. San Francisco, CA: Morgan Kaufmann.

Komlodi, A., Marchionini, G., & Soergel, D. (2007). Search history support for finding and using information: User interface design recommendations from a user study. *Information Processing and Management, 43*, 10-29.

Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics, 19*, 55-66.

Maglio, P.P., & Barrett, R. (1997). How to build modeling agents to support web searchers. *The 6th International Conference on User Modeling*, Sardinia, Italy. Retrieved February 12, 2010, from http://wwwcssrv.almaden.ibm.com/wbi/papers/um97.pdf

Obendorf, H., Weinreich, H., Herder, E., & Mayer, M. (2007). Web page revisitation revisited: Implications of a long-term click-stream study of browser usage. *CHI '07*, 597-606.

Spink, A., & Zimmer, M. (eds.) (2008). *Web search: Multidisciplinary perspectives*. Berlin: Springer.

Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology, 52*(3), 226-234.

Tauscher, L., & Greenberg, S. (1997). How people revisit web pages: Empirical

臺灣大學圖書館 期刊資料庫

findings and implications for the design of history systems. *International Journal of Human-Computer Studies, 47*(1), 97-137.

Teevan, J. (2004). *How people re-find information when the web changes.* MIT Computer Science and Artificial Intelligence Lab Technical Report. Retrieved February 12, 2010, from http://dspace.mit.edu/handle/1721.1/30480

Teevan, J. (2008). Where'd it go?: How people ask after lost Web information. *ASIST 2008 Annual Meeting*. Retrieved February 12, 2010, from http://people.csail.mit.edu/teevan/work/publications/papers/asist07.pdf

Teevan, J., Adar, E., Jones, R., & Potts, M.A.S. (2007). Information re-retrieval: Repeat queries in Yahoo's logs. *ACM SIGIR' 07*, Amsterdam, The Netherlands.

Teevan, J., Alvarado, C., Ackerman, M.S., & Karger, D.R. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. *ACM CHI'04*. Retrieved February 12, 2010, from http://groups.csail.mit.edu/haystack/documents/papers/2004/chi2004-perfectse.pdf

Xie, I. (2009). Dimensions of tasks: Influences on information-seeking and retrieving process. *Journal of Documentation, 65*(3), 339-366.

(Received: 2010/2/22; Accepted: 2010/3/8)

臺灣大學學術期刊資料庫

臺灣大學學術
期刊資料庫