

以網站流量資料 探討使用者的網路使用行爲

Using Traffic Data of Websites to Explore Users' Behavior in Network Use

林頌堅

Sung-Chien Lin

世新大學資訊傳播學系 助理教授

Assistant Professor

Department of Information and Communications
Shih-Hsin University

【摘要 Abstract】

人類的生活與網路使用愈來愈不可分，分析使用者的網路行為可以了解他們的生活面貌。本研究建議研究者可以利用網站流量資料做為網路使用研究的初步分析資料。在本論文，我們以Alexa.com所提供的網站流量資料為例，說明利用這些資料時的應用方法與限制。

We have not been able to take out use of network from our everyday life anymore. Analyzing users' online behavior will benefit the understanding of various ways and aspects in their lives. This study suggests that researchers can use the traffic data of websites as the analytic data for primary studies when they need to perform some researches of network use. Taking the data retrieved from Alexa.com as an example, we proposed several possible applications and restrictions in the use of this kind of data in this paper.

[關鍵字 Keywords]

網路使用行為；網站流量資料

Network Use Behaviors; Traffic Data of Web Sites



壹、前言

全世界的網路使用人口及普及率正不斷地提升，2007年九月的全球上網人數已超過十二億四千萬人，普及率（penetration）也達到18.9%，相較於2000年的統計資料，使用人數增加了244.7%，而且各地區的上網人口也同樣地都不斷增加，比方說，亞洲的網路使用者比2000年增加了302.0%，總數達到四億五千萬，歐洲和北美洲則分別有三億三千多萬和二億三千多萬的網路使用人口，跟2000年相比各增加了221.5%和117.2%（Internet World Statistics, n. d.）。不僅是網路使用人口所佔的比率持續增加，民眾在生活中上網時間的比重也相當高，以臺灣及中國為例，FIND的「2006我國家庭寬頻、行動與無線應用現況與需求調查分析報告」（孫鴻業，2006）指出我國上網民眾每週的平均上網時數為19小時，另外，中國互聯網絡信息中心（2007）則指出中國網路使用者的每週上網時數已經達到18.6小時。而且隨著資訊與通訊技術的進步，嶄新的應用服務也如雨後春筍般紛紛出現，FIND對臺灣民眾最近一個月內在家曾進行的網路活動所做的調查統計（孫鴻業，2006）指出，民眾的上網活動項目除了瀏覽資訊（86.2%）、收發電子郵件（77.5%）、上傳或下載檔案（55.5%）、傳送即時簡訊（52.1%）、玩線上遊戲（36.4%）等早已普及的網路應用服務之外，近來許多新興的服務也愈來愈普遍，比方說：網路日誌（blog）、網路電話等的使用人數比例正快速地增加中。另外，從Pew Internet & American Life Project針對美國網路使用者所做的多項網路使用行為調查，包括觀看線上視訊（Madden, 2007）、使用社會網絡網站（social networking websites）（Lenhart & Madden, 2007）、使用金融交易（Fox & Beier, 2006）以及尋找居住資訊（Fallows, 2006）、尋找科學資訊（Horrigan, 2006）、尋找政治與選舉資訊（Horrigan, 2006）、尋找疾病與保健資訊（Madden & Fox, 2006; Fox, 2006）等各項線上活動（Horrigan & Rainie, 2006），也顯示愈來愈多的美國網路使用者已經能夠在網際網路上進行相當多的活動。

從上面各種的網路調查結果可以發現網路的普及率以及應用範疇愈來愈擴大，現代人在生活上對於網路依賴愈來愈深。在實際生活上從事的許多社會活動都可以在網路上找到類似或者甚至更為方便的類型，比方說，交友聯誼、查詢資訊、訓練學習、買賣交易與遊戲娛樂等等都是目前網路上盛行的活動。事實上，對許多網路使用者而言，線上與線下的活動同樣來得重要，許多人們在他們的日常生活裡早已離不開網路，並且在使用網路的過程中，學習到或發展出增進生活的便利、樂趣與品

質的各種技能。為了瞭解網路使用情形對於日常生活中的社會現象與個人行為的影響，研究者必須實地蒐集與分析使用者在使用網路進行相關活動的現象。本研究的目的便是建議利用現有的網站流量資料，作為網路使用行為研究時的參考資訊，更具體地來說，本研究以Alexa.com (<http://www.alexa.com/>)所提供的網站流量資料為例，說明利用這些資料時的方法與限制。

第貳節將簡要地說明目前蒐集網路行為資料的方法；第參節將說明Alexa所提供的流量統計資料；第肆節將應用Alexa的流量資料，介紹幾種網站流量的使用方法；第伍節將探討這個方法的限制；最後，第陸節則是本研究的結論。

貳、網路使用行為的資料蒐集方法

社會科學中常用問卷等調查研究 (survey) 方法或是記錄觀察、質性訪談、民族誌 (ethnography) 等田野研究 (field research) 方法，蒐集人們在社會環境中所產生的各種現象，取得相關的資料來進行分析。如前所述，人們在網路上發生的行為與關係已成為社會科學研究的主要標的之一，因此許多學者也將各種社會科學的研究方法應用在各種網路現象的分析，比方說以時間日誌 (time diary) 的資料分析使用者在使用電腦時發生挫折感 (frustration) 的時機、頻率與影響 (Ceaparu, Lazar, Bessiere, Robinson, & Shneiderman, 2004)；以記錄、觀察或統計使用者在進行網路資源尋求時的各種認知、情感和身體行為等各種表現 (Wang, Hawk, & Tenopir, 2000)；或者利用參與線上社群 (online community) 所記錄的田野筆記 (fieldnote) 以及線上蒐集的互動資料，分析與詮釋該線上社群的成員間對特定商品發表的看法與交換的意見 (Kozinets, 2002)。許多研究妥善地運用了這些資料蒐集方法解決了它們的問題，然而這些源自社會科學實證研究、以訪談或觀察為基礎的資料蒐集方法也都承襲了它們以往被批評的問題，特別是在資料所代表的使用者來源廣度以及在資料蒐集過程中研究者的介入程度等兩方面。比方說，調查研究類的方法較容易大規模地對網路使用者蒐集各種統計資料，準確地描述與預測網路使用的狀況，但事實上由於題項設計無法反應需要測量概念的程度，可能導致調查的效度 (validity) 較弱 (Babbie, 2002/邱泯科等譯, 2004)，並且會有記憶偏差所導致的誇大 (inflated) 以及不正確估計等問題 (Ceaparu, et. al, 2004)。田野研究類的資料蒐集方法能揭露網路使用的情境脈絡，但由於研究者的差異而導致這類方法的信度較

弱；若是要蒐集大規模的訪談或觀察資料也非常困難；此外，許多研究為了能夠比較使用者各方面的差異對網路使用情形造成的影響，通常會以模擬的任務與情境，進行實驗，作為蒐集使用行為表現的方法，但畢竟這些情境與表現和在真實情況下不會完全相同。因此，一種能夠實際反應一般性網路活動的狀況並且容易實行的資料蒐集方法對於網路使用的研究將有很大的幫助。

網路活動與其他社會活動的不同處在於過程中系統會自動留下豐富的記錄資料。舉例而言，在網頁瀏覽過程中，使用者每一次的資源需求或連結會在提出資源需求的用戶端電腦（client terminals）、網站伺服器（web servers）或是ISP（information service providers）的代理伺服器（proxy）中留下要求資源的URL位址、用戶端電腦主機的IP位址、需求時間、訊息長度等等的記錄資料（林頌堅，2000）。另外，使用者在透過網路進行各種資料查詢與交易，也會在存取的線上資料庫上留下交易記錄（transaction log）。為了求得更為接近真實的網路使用情形，便可以利用使用者在進行網路活動時電腦自動儲存的各種記錄資料。由於這些資料的產生是使用者實際使用網路資源所留下的記錄，容易蒐集到足以達到具有統計意義的資料量，而且在資料記錄的時候，對使用者實際的使用行為幾乎沒有影響，如果能夠善用這些資料進行分析與解讀，對使用者在網路上的活動可以有客觀的認識。目前已經有許多經營網路服務的組織與企業，從它們擁有的網站伺服器中取得使用者進行各項網路活動時留下的使用紀錄，藉由WebTrends（<http://www.webtrends.com/>）或Analog（<http://www.analog.cx/>）等網站流量分析軟體進行資料分析與整合，分析出上站的人數、瀏覽的網頁、上站時間以及瀏覽路徑等網站流量資訊，並從中推測使用者的網路使用行為，作為網路行銷與客戶關係管理的參考資訊。例如，Nicholas等（Nicholas, Huntington, Lievesley, & Wasti, 2000）針對英國的The Times/The Sunday Times新聞網站，配合訂閱者（subscribers）的資料，以三個月的網站工作紀錄進行分析，結果發現使用者較多在星期二到四的日子使用The Times，而週末（weekend）的時間則較少使用，而且從在全天任何一個時段的流量都很平均的現象，可以了解這個網站的讀者分布於多個時區。流量分析的結果還顯示，極少數的重度使用者使用了24.5%的使用量，並且比其他的使用者有較高的使用頻率。林頌堅（2000）也曾利用世新大學圖書館網站於1999年5月4日至2000年5月5日間約一年的工作紀錄進行分析，結果發現連上這個網站的用戶端電腦大多位於圖書館，透過取用網頁以及取用路徑的分析則發現使用者連上大學圖書館網站的目的主要為webpac查

詢。另外這個研究還觀察到大學圖書館的流量和當學期的課程時間表有密切的關係。

由前述的研究可以發現，這個研究方法雖然能夠分析單一網站與它所提供的功能與資源的使用情形，但無法比較同一時段內多個網站的流量統計資料，對於全面了解相近類型的網站或整體全球資訊網的使用情形有所不足。另外，這些資料的蒐集、分析與解讀，都需要長期地耗費極大的人力與設備資源，才能獲得可靠而有效的使用者行為資訊。並且這些研究也都指出了這種研究方法的限制（林頌堅，2000；Nicholas, Huntington, Williams, Lievesley, Dobrowolski, & Withey, 1999; Nicholas, et al., 2000），包括這些資訊可能因為許多人利用公用電腦上站、個人可能從多處電腦上站或是由於浮動IP、暫存檔（caching）等網路技術，使得網站的使用人數在統計上不精確；而最大的問題是分析者無法了解使用者上站或點選某一網頁的動機為何，因此無法賦予統計資訊實在的意義。

參、網路流量資料的取得

從上一節對於現有的網路使用行為資料蒐集方法的探討，我們可以了解利用網路使用工作紀錄分析某一特定網站的使用情形，具有符合網站使用情境、不會干擾使用者以及容易取得更有統計意義的資料量等優點。如果要運用這種方法分析同一類型網站的使用行為，必須要從分析的各個網站伺服器中取得所儲存的網路使用工作紀錄。然而，要從數目龐大而且可以彼此有利害衝突的網站取得工作紀錄並且進行分析，對一般研究者實際上是一件相當困難的工作。

現在有許多網路調查公司透過各種方法從用戶端或ISP的電腦中取得使用者進行網路活動的記錄，分析使用者進入的網站、使用的資源與時間等資訊，提供網路為主要通路的企業了解自己與競爭者的使用趨勢，比方說Nielsen/NetRatings（<http://www.nielsen-netratings.com/>）、comScore（<http://www.comscore.com/>）及HitWise（<http://www.hitwise.com/>）或是臺灣的創市際市場研究顧問公司（<http://www.insightxplorer.com/index.html>）。雖然不像直接利用網路使用工作紀錄具有彈性而有廣泛的應用範圍，研究者仍然可以從這些網路調查公司所發布的調查資料中獲取對研究有幫助的資源。但是，如果研究者有特別的研究需求，需要網路調查公司提供特定網站的流量統計，則需要付費申請這些調查公司進行資料查詢。所幸，網路上也有許多機構組織會公開它們蒐集與分析的重要網站流量資料，雖然這些資料可能不像前述的資料一般全面而

且容易解讀，但是研究者如果能夠善用這些資料，也能夠做為發現問題或者驗證研究發現時的參考。其中最負盛名的網站流量資料來源是Alexa網站以及Compete.com (<http://www.compete.com/>)。Compete的優點是它利用了相當多種的管道取得網路的流量資訊並且給予適當的加權，求出較接近實際狀況的網路流量 (Compete: FAQ, n. d.)¹，但由於Alexa網站所提供的網站流量資料較為精細，而且這些資料已經能夠表現出大多數的網站使用行為，因此本研究採用Alexa所提供的資料為例。

Alexa網站上的流量資料估算方法是利用安裝在用戶端電腦瀏覽器上的Alexa工具列 (toolbar)，當使用者提出一個網頁需求時，會由Alexa工具列在Alexa的資料庫留下一筆記錄²。藉由對這些安裝Alexa工具列使用者連線的情形，Alexa每日針對各個網站計算前一天、前七天平均及前三個月平均連線到該網站的使用者佔全體安裝工具列使用者的比率 (Alexa稱為reach) 以及這些使用者在該網站上的平均每天瀏覽網頁數目 (page views)，再藉由這兩種資料估算全球各個網站的網頁流量排名 (traffic rank)。

運用前述的連線使用者比率、平均網頁瀏覽數和流量排名等資料，Alexa網站提供全世界近三個月的前五百名高流量網站，並且根據用戶端電腦的IP估計每個國家或地區三個月流量前一百名網站。表一表示2007年12月30日在Alexa所檢索到近三個月流量最高的前二十名網站排名，另外表二是臺灣使用者的前二十名網站排名。Alexa也提供了個別網站的流量資訊，如表三是Yahoo!在2007年12月30日的網頁流量記錄，包括前一天、前七天平均及前三個月平均的連線使用者比率、平均網頁瀏覽數和流量排名 (見表三)。此外，雖然Alexa沒有對各個網站提供流量歷史資料的檢索功能，但它對各網站提供了以各種時間間距的使用者比率、平均網頁瀏覽數和流量排名等為基礎產生的流量歷史圖 (traffic history graph)，研究者可以根據近七日、近一月、近三月、近六月、近一年、近三年、近五年和最大期限 (從Alexa蒐集資料開始) 等不同時間間距檢視網站近、中、長期的流量變化情形³。

1. 在本文完稿後，Alexa已經在2008年4月15日在其部落格上宣布，使用更多的管道取得網路流量資料以代表更廣泛的使用群眾，並且利用更準確的新方法計算網站的流量排名 (Mack, 2008, April 15)，但由於回溯資料的關係，目前所能提供的資料範圍僅包括2007年8月以後。
2. 如註1所提到的，目前Alexa已經改變蒐集網路流量的方法，這段敘述所描寫是2007年研究時的情況。
3. 由於改變流量資料蒐集方式，目前Alexa的流量歷史圖僅提供近七日、近一月、近三月、近六月和最大期限 (從2007年8月開始) 等時間間距。

表一 Alexa網站所公佈的網站流量排名

排名	網站名稱	網站網域
1	Yahoo!	yahoo.com
2	Google	google.com
3	Windows Live	live.com
4	YouTube	youtube.com
5	Microsoft Network (MSN)	msn.com
6	Myspace	myspace.com
7	Facebook	facebook.com
8	Wikipedia	wikipedia.org
9	Hi5	hi5.com
10	Orkut	orkut.com
11	Rapidshare.com	rapidshare.com
12	Blogger.com	blogger.com
13	Megaupload	megaupload.com
14	Friendster	friendster.com
15	Yahoo!カテゴリ	yahoo.co.jp
16	Fotolog.com	fotolog.net
17	Baidu.com	baidu.com
18	Google.fr	google.fr
19	Microsoft Corporation	microsoft.com
20	EBay	ebay.com

資料來源：檢自Alexa網站，2007年12月30日。

表二 臺灣連線網站流量排名

排名	網站	網站網域
1	Yahoo!	yahoo.com
2	無名小站	wretch.cc
3	PC home電腦報	pchome.com.tw
4	巴哈姆特電玩資訊站	gamer.com.tw
5	Yam天空	yam.com
6	Google	google.com.tw
7	Badongo	badongo.com
8	Windows Live	live.com
9	Hinet	hinnet.net
10	YouTube	youtube.com
11	伊莉心情車站	eyny.com
12	Wefong.com	wefong.com
13	Im.tv	im.tv
14	Gogobox.com.tw	gogobox.com.tw
15	Google	google.com
16	中華電信Xuite	xuite.net
17	臺灣論壇	twbbs.net.tw
18	Oursogo.com	oursogo.com
19	Plus28.com	plus28.com
20	Bbs-tw.com	bbs-tw.com

資料來源：檢自Alexa網站，2007年12月30日。



表三 Yahoo!網站的流量資訊

	昨天	前七天平均	前三個月平均
連線使用者數比率	26.950001%	26.705%	27.629999%
流量排名	1	1	1
平均網頁瀏覽數	12	12.2	13.5

資料來源：檢自Alexa網站，2007年12月30日。

肆、網站流量資料的應用

從Alexa網站所蒐集的網站流量資料能夠用來作為網站選擇與使用研究的參考資訊。在本節中，本研究將提出幾個網站選擇與使用的問題，嘗試利用Alexa提供的網站流量資料進行分析。

一、全球高流量網站反映的網路使用行為

首先，本研究嘗試依據Alexa所提供的全球網站流量排名（表1）作為參考資訊，選取較高流量的網站，分析當時較受到使用者重視的網站以及這些網站的類型，了解當前網路產業的焦點與趨勢，並且對重要的網站利用Alexa所提供的主要連線國家與地區，探討這些網站的國際化情形；對於使用行為較為特殊的網站，本研究將進一步透過使用者較常造訪的子網域（subdomain）來討論這些網站中較為重要的服務，並且利用流量歷史圖進一步了解這些網站的使用週期。

在表1的全世界高流量網站上，首先會注意到流量排名最高的網站包括了Yahoo!、Google、Windows Live、MSN（Microsoft Network）、Baidu（百度）、Yahoo!カテゴリ以及Google.fr等入口網站及搜尋引擎。若以這些入口網站與搜尋引擎的主要連線國家與地區來分析，除了Baidu、Yahoo!カテゴリ與Google.fr的連線分別以來自中國、日本及法國佔絕大多數以外，其他網站的連線使用者大多分散於全球各地。Baidu有85.9%的連線使用者來自於中國，Yahoo!カテゴリ有93.2%來自於日本，Google.fr則有74%來自於法國；但反觀Yahoo!的前五個連線使用者較多的國家與地區依序是美國（19.4%）、臺灣（5.4%）、香港（4.4%）、越南（4.0%）和印度（3.8%），Google的主要連線國家與地區為美國（26.1%）、巴西（3.7%）、印度（3.6%）、智利（2.8%）和英國（2.7%），連線Windows Live和MSN的連線地區也相當分散。

至於在這些入口網站及搜尋引擎中，使用者較常利用的資源與服務為何？從Alexa所提供的連線子網域名稱來看，電子郵件在這些網站的應

用中都是數一數二的，比方說，在Yahoo!的連線裡，mail.yahoo.com佔了47%，hotmail.msn.com佔了msn.com 73%的連線，mail.live.com也佔了Windows Live所有連線的72%，即便是以搜尋引擎聞名的Google，mail.google.com也佔有全部連線的13%，從這個現象可了解電子郵件是入口網站服務中相當重要的一個應用。搜尋引擎重要的應用還有資訊查詢，目前Google及Baidu兩個搜尋引擎除了提供網頁檢索以外，還有各式各樣形式的資訊檢索，以Alexa的資料來看，image.google.com佔所有google.com連線的10%，image.baidu.com和mp3.baidu.com則各佔baidu.com 12%和9%的連線，因此圖像及mp3的檢索是有較多使用量的搜尋服務。

社會網絡網站是近來最引人注目的網站。表1包括了MySpace、Orkut、Facebook、Hi5、Friendster和Fotolog等等社會網絡網站，這類型的網站加上Google、YouTube影音分享網站、Wikipedia維基百科網站以及Blogger部落格網站等都被歸類為Web 2.0概念的網站。最近一兩年來，Web 2.0的概念在許多領域都是相當沸沸揚揚的話題，但此一名稱所代表的究竟是一個可能改變網路使用行為的革命理念？或僅止於一個商業炒作的噱頭（O'Reilly, 2005; Madden & Fox, 2006）？Pew Internet & American Life Project最近幾篇報告，即是特別針對美國網路使用者使用線上視訊（Madden, 2007）、社會網絡網站（Lenhart & Madden, 2007）以及部落客（Lenhart & Fox, 2006）等情形進行調查，其他產業分析以及學術研究（Ahn, Han, Kwak, Moon, & Jeong, 2007）也同樣有針對這個主題進行探討，可見這個問題的重要性。若是以Alexa的流量排名資料做為網路使用影響的指標來看，可以發現前二十名流量最高網站中已經有上述的十一個網站屬於這個概念，似乎可以推論Web 2.0概念網站已經對網路使用產生了一些影響，並且具有廣大使用者的Web 2.0概念網站，以影音分享和社會網絡網站為主。

就使用者的分布情形而言，影音分享網站的使用者集中於YouTube網站上，安裝Alexa工具列的使用者近三個月來每天平均有15.64%會進入這個網站，並且它的使用者分布於各個國家與地區，除了有13.7%的使用者來自於美國之外，其餘每個國家都僅低於5%。社會網絡網站的使用者則分散於不同的網站上，MySpace、Orkut、Facebook、Hi5、Friendster和Fotolog等網站過去三個月的連線使用者比率如表4所示。表4也呈現了這些網站的使用者國家與地區，可以發現各網站使用者的國家與地區都不相同：以安裝Alexa工具列的使用者來計算，MySpace的美國使用者接近全體使用者的一半；Orkut的使用者則是有60%以上來自於巴西，接近20%來自印度；Facebook則有三分之一以上的使用者來自北美的美國與

加拿大；Hi5的前五個最多使用者的國家中，有四個以西班牙語為主要語言，並且都位居於中美洲；大多數的Friendster使用者來自於東南亞國家；Fotolog的使用者則是大部分位於南美洲，智利、阿根廷和烏拉圭等國家，共佔大約四分之三。另外，使用者在利用這類網站上的行為特徵還包括使用者每天閱讀的網頁數較多，以Alexa所提供的平均網頁瀏覽數而言，上述這些網站三個月平均的每日瀏覽網頁數都在二十頁以上。

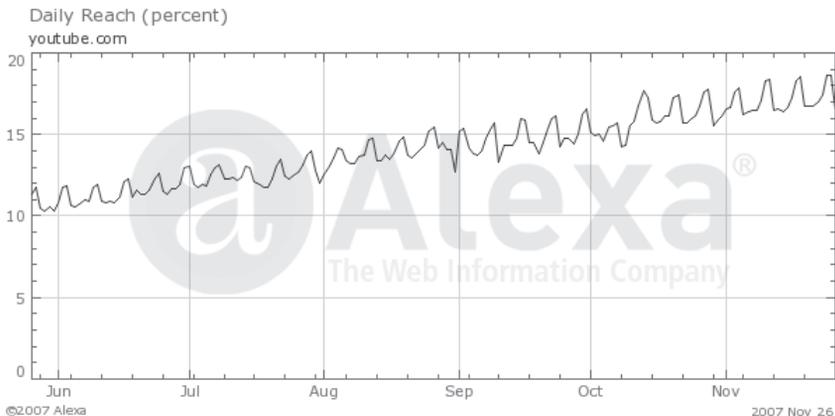
表四 主要社會網絡網站的流量資訊

社會網絡網站	前三個月 連線使用者 比率	使用者主要所屬國家或地區	前三個月 平均網頁 瀏覽數
MySpace	5.46%	美國 (46.9%)、波多黎各 (6.6%)、英國 (4.1%)、法國 (3.2%)、墨西哥 (3.2%)	31.3
Orkut	3.17%	巴西 (62.0%)、印度 (19.8%)、美國 (3.4%)、巴拉圭 (2.3%)、巴基斯坦 (1.9%)	26.4
Facebook	4.41%	美國 (27.1%)、加拿大 (9.1%)、土耳其 (8.5%)、英國 (4.2%)、波多黎各 (3.1%)	30.6
Hi5	3.24%	祕魯 (13.1%)、泰國 (7.8%)、多明尼加 (7.0%)、哥斯大黎加 (6.7%)、墨西哥 (6.1%)	30.8
Friendster	1.89%	菲律賓 (38.9%)、馬來西亞 (18.8%)、新加坡 (14.7%)、印尼 (13.6%)、美國 (3.3%)	37.8
Fotolog	1.84%	智利 (56.6%)、阿根廷 (15.9%)、巴西 (6.5%)、西班牙 (5.1%)、烏拉圭 (5.1%)	23.4

資料來源：檢自Alexa網站，2007年11月22日。

再進一步，針對YouTube等網站在不同日期的使用情形來進行分析。圖一是從Alexa網站上所檢索的YouTube網站近六月流量歷史圖。從此圖以及近三年的流量歷史圖上看來，YouTube網站的使用者正持續地增加，2007年6月初的連線使用者比率大約在10-12%之間，到了11月底則有17-19%。另外，這個網站的上站人數有週期性的增減情形，以近三月或近一月等更少天數的流量歷史圖來觀察，可以發現這個週期大約是七天，在週一到週五之間使用者較少，週六與週日則有較多的使用

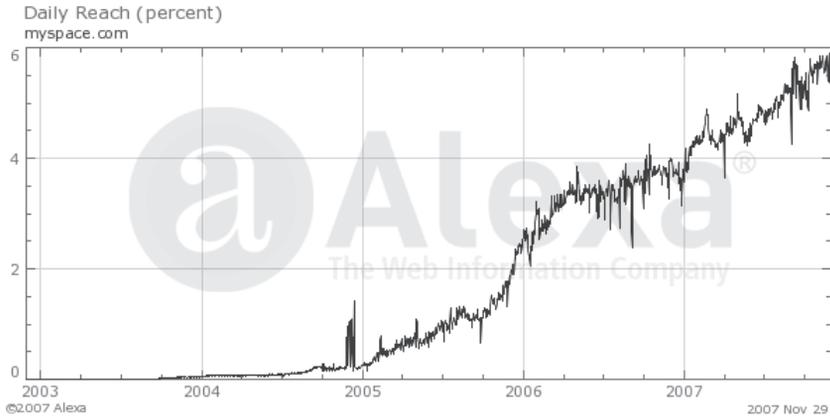
者，此外週六日也有較平日多的網頁瀏覽量。這些現象可以和Madden (2007) 的研究報告相比較，在這份報告裡指出2007年2到3月間已經有57%的美國網路使用者曾收看過網路視訊，其中有29%的人最常收看的網路視訊是YouTube，而使用者收看網路視訊的地方，在他們自己的家中(59%)遠較於在工作場所(24%)來得多。接下來，觀察社會網絡網站的歷史流量圖。前述之社會網絡類型的各種網站大都成立於2004年後，因此本研究觀察它們近五年的reach歷史流量圖，例如圖2是Myspace的近五年歷史流量圖，從這些歷史流量圖上可以發現，除了Orkut之外，其餘的網站在2007年的連線使用者比率都有大幅的成長情形，其中的Facebook便是最為明顯的例子，另外諸如Myspace、Hi5和Friendster於2006年的成長情形原本已經比較緩慢，但在2007年連線使用者比率也有相當規模的增加，因此最近一年社會網絡網站的使用者數目大幅增加。而且從近三月的流量歷史圖可以發現，這些網站的連線使用者數比較沒有規律的週期變化，在週六日時僅略為增加一些，並不像YouTube網站明顯地在週六日有較多的使用者，而且社會網絡網站的使用者在週末時瀏覽的網頁數量也僅比平日時稍微多一些，差別不大，這些都表示以發展和維繫友誼為主要功能的社會網絡網站與以消遣娛樂為主的影音分享網站在使用行為上有所不同。



圖一 YouTube網站六個月間的連線使用者比率變化

資料來源：檢自Alexa網站，2007年11月28日。





圖二 Myspace網站五年間的連線使用者比率變化

資料來源：檢自Alexa網站，2007年12月1日。

最後，在表一上的網站類型還有網路空間網站，這種類型的網站提供使用者儲存、傳播與分享資料與檔案的網路服務。使用者可以透過網路空間網站所提供的超大型記憶空間，將大量的資料與檔案儲存在網路上，便於日後在其他地點使用，同時使用者也可以利用這樣的方式將資料與朋友分享，甚至傳播給其他不相識的使用者。由於這樣的應用服務對使用者而言，相當便利，因此不僅有許多網路公司提供這樣的服務，而且不少使用者也已經非常嫻熟地將它們應用在生活與工作上。在Alexa所提供的全球前二十名最高流量網站中，RapidShare和MegaUpload都是屬於提供網路空間服務的網站，它們的三個月平均連線使用者比率分別為5.81%與5.80%。觀察Alexa的流量歷史圖可以發現這兩個網站的連線使用者比率在週末時會比平日時有少量的增加。

二、臺灣流量較高網站反映的網路使用行為

接下來，本研究將分析臺灣使用者連線量最高的二十個網站（表二），並且與全球高流量網站相比較，尋找臺灣網路使用的特殊現象。在表二所列出的網站中，只有Yahoo!、YouTube、Windows Live和Google等四個網站，屬於全球最高流量的網站者，正好是全球最高流量網站的前四名。從Alexa所提供的資料上看來，Google在臺灣的名次是第六名，但主要是由於Alexa將臺灣Google網站（google.com.tw）和Google網站（google.com）兩個網站分開計算，臺灣大部分的使用者在利用Google的時候大部分習慣進入具有繁體中文介面的臺灣Google網站。

臺灣連線量最高的二十個網站中，具備Web 2.0概念的網站有無名小站、臺灣Google、YouTube、Im.tv、Xuite與Google。依據它們的分類來說，臺灣Google及Google是搜尋引擎，YouTube和Im.tv是屬於影音分享服務，無名小站和Xuite則是提供部落格服務。在上述的觀察中，本研究也發現臺灣連線量最高的二十個網站並沒有許多國家或地區相當盛行的社會網絡網站。此一現象相當值得深入探究。

無名小站、Im.tv和Xuite三個網站分別屬於部落格提供者以及影音分享網站，它們的平均連線使用者比率都有週期性增減的情形，這三個網站通常在週六日時會有比平日多的使用者，這個現象和YouTube網站相似，似乎表示網路使用者於週末時會有多的空閒時間，因此得以連上這些網站編輯與上傳自己製作的部落格與視訊短片，或者瀏覽他人所提供的部落格與影音內容。無名小站在近三個月的平均連線使用者比率为0.83%，並且每天平均網頁瀏覽數為24.2頁，因為有較多的連線使用者以及較多的網頁閱讀數量，使得這個網站成為全球網頁流量的第三十八名。Xuite的前三個月平均連線使用者比率大約為0.16%，是全球網頁流量的第四百九十二名。從Xuite網站的次網域，我們可以了解使用者進入這個網站所利用的主要功能，最多為網路相簿服務（44%），其次為部落格服務（40%）。

將其餘的網站類型進一步分析，我們可以發現屬於入口網站的有PChome電腦報、Yam天空和Hinet，屬於網路空間服務的網站有Badongo和Gogobox，最為特殊的應該是有許多網站是屬於論壇（forum）性質的網路服務，比方說：巴哈姆特電玩資訊站、伊莉心情車站、Wefong、臺灣論壇、Oursogo和Plus28.com等。甚至先前被歸類為入口網站的PChome電腦報，其實從它被使用較多的應用服務來說，如個人新聞台等等，被歸類為網路論壇可能較為合適。論壇性質的網站具有和電子佈告欄（bulletin board systems, BBS）和新聞群組（newsgroups）等網路服務相似的功能，提供網路使用者發表與分享個人思想與創作、抒發情感、維繫友誼、建立社群等。這些網站所提供的服務是否在某種程度上取代了社會網絡網站？並且成為前述之社會網絡網站在臺灣不盛行的因素？同樣也值得進行深入的研究。而愈來愈重視資訊取得與分享（information acquiring-and- sharing）議題的學術研究（Rioux, 2004）也應該可以從這些網站的內容與使用中獲得豐富的研究題材與資源。

論壇類型網站在網路流量上的特色，正如社會網絡網站與部落格網站一般，使用者每天在這些網站上所瀏覽的網頁內容相當多，例如巴哈姆特電玩資訊站、伊莉心情車站、Wefong和Plus28.com在近三個月的使

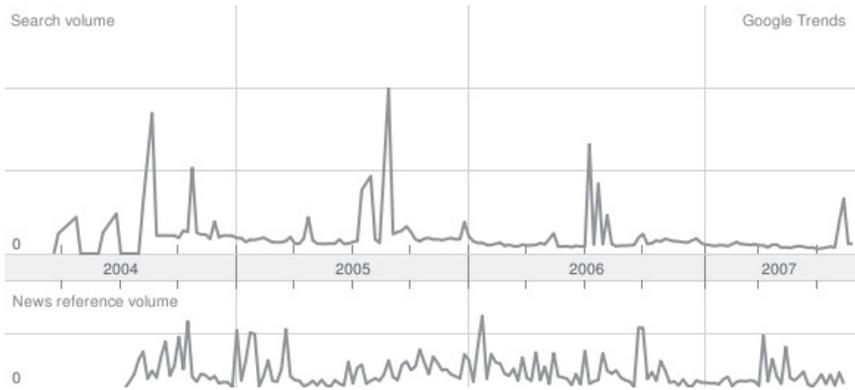
用者平均每天的瀏覽頁數都在16頁到18頁之間，Oursogo則更多達25.0頁，並且在觀察論壇類型網站的歷史流量圖時，同樣也可以發現到各網站在周六日的平均連線使用者比率較多，產生週期性的變化。

從觀察臺灣較高連線量網站的連線情形可以發現，除了全球連線量前五名的網站外，連線到這些網站的主機位址大多來自臺灣，舉例來說，諸如無名小站有79.3%的連線來自臺灣，PChome電腦報則有83.2%。在連線上這些網站可以發現它們所提供的內容大部分是以繁體中文撰寫的。上面的這個現象很明顯地可以看出網站內容所使用的語言是臺灣網路使用者在選擇網路資訊與服務的一個很重要的因素，所以Badongo及Gogobox取代了RapidShare及MegaUpload成為臺灣連線較多的網路空間服務提供者，Im.tv也成為臺灣網路使用者在YouTube之外的另一種選擇。此一推測值得研究者與網站公司作為進一步研究與企劃行銷上的參考。

三、個別網站流量所反映的網路使用行為

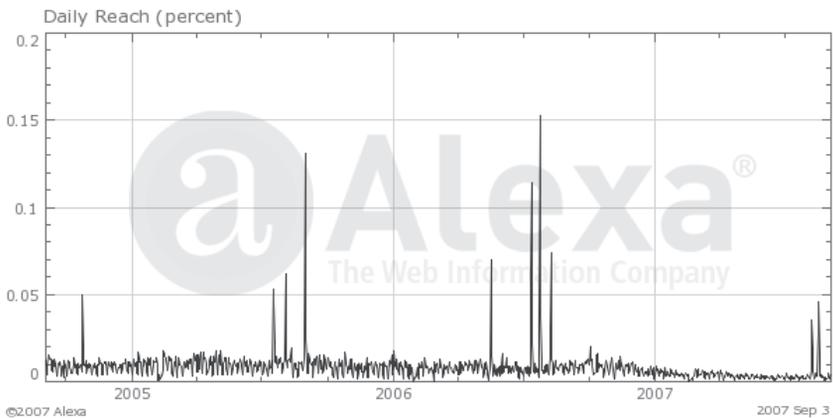
雖然從前面對於高流量網站的分析，研究者可以知道較多數使用者偏好的網路資源，以推估使用者上站的需求動機，但有時研究者想分析探討某些因素對特定網站使用的影響，例如重要事件的發生是否影響使用者從某些網站來獲取他們所需的資訊。如果研究者採用網路流量統計，此時便需要設法取得該網站的工作紀錄來進行分析。然而，誠如之前所提到的，網站工作紀錄的取得並不容易。本研究針對這個困難，建議研究者可以利用Alexa所提供的網站流量資訊，比較在這些事件發生前後日期中這些網站流量的變化情形，從上網的使用者資料與下載的網頁量來推斷使用者從這些網站中取得資訊的可能性。以下先描述一個筆者曾遭遇的問題，再說明解決問題的思考過程。

在某次瀏覽Google Zeitgeist (<http://www.google.com/press/intl-zeitgeist.html>) 所提供的查詢詞語排行中，筆者發現2006年7月的資料中有一個詞語是「人事行政局」，在利用Google Trends (<http://www.google.com/trends>) 了解這個詞語被查詢的情況時，發現這個詞語被查詢次數最多的時候大多是在每年第三季（7到9月間）的時候（如圖三所示）。因此，接著利用Alexa查詢人事行政局的網站 (<http://www.cpa.gov.tw/>)，查看近三年的連線使用者比率變化情形，結果如圖四所示，結果發現這個網站較高的連線使用者比率也是大多發生在這三年的第三季之間。



圖三 人事行政局在Google搜尋引擎內的查詢變化

資料來源：檢自 Google Trends，2007年9月6日。

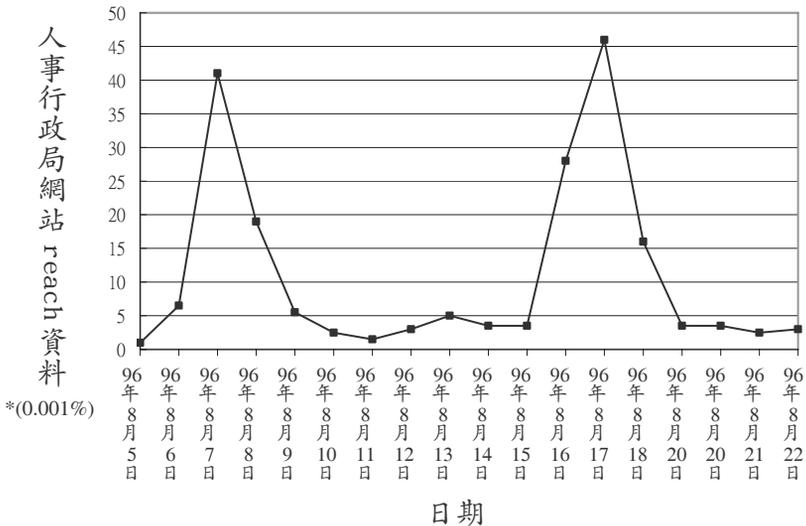


圖四 人事行政局網站在三年間連線使用者比率的變化

資料來源：檢自 Alexa網站，2007年9月6日。

究竟是什麼因素使得網路使用者會在每年的第三季大量地利用 Google 搜尋引擎檢索人事行政局，並且進入該網站？從時間上考慮，每年的7到9月是臺灣的颱風季節，人事行政局網站在颱風前會發布天然災害停止上班上課訊息，雖然民眾可以藉由廣播及電視等傳播媒體收聽或收看這項訊息，但是否也有網路使用者直接利用人事行政局取得第一手

的停止上班上課訊息？為了驗證這個想法，本研究從Alexa網站記錄了人事行政局網站在2007年8月5日到22日的每日連線使用者比率。每天的資料變化情形，如圖五的實線表示。



圖五 人事行政局網站的連線使用者比率

資料來源：檢自Alexa網站，本研究整理。

在資料蒐集的這段期間有帕布、梧提和聖帕等三個颱風侵襲臺灣，人事行政局發布的日期共有8月8日、9日、17日、18日和19日共五天，本研究假定使用者會在停止上班上課訊息的前一天連線到人事行政局網站閱讀這項訊息。因此在實驗裡，將訊息發布前一天與平日的人事行政局網站的平均連線使用者比率相等建立為虛無假設（null hypothesis），以變異數分析（analysis of variance）來進行檢定。在計算後，發現組內平均平方和為2550.646，而組間平均平方和為45.370，因此F值為56.219，此一數據大於1並且其顯著性小於0.05，所以此實驗拒絕虛無假設。換言之，訊息發布前一天與平日網站連線使用者比率的平均值並不相等，而且訊息發布前一天的數值通常遠大於平日。所以，從Alexa所蒐集的網站流量顯示，網路使用者極有可能會直接利用人事行政局取得第一手的停止上班上課訊息。



伍、網路流量資料的限制

任何資料蒐集方法都有其先天的特質與限制，只有了解這些特質與限制，才不會導致嚴重的錯誤，並且才能夠將這些方法運用在最適當的時機與問題上，取得豐富的現象，進而獲得確切的研究成果。目前網路產業或者一些資訊科技論壇已經相當廣泛地應用Alexa所提供的流量資料做為網站使用的事證，研究者在閱讀其他人的研究時，必須留意分析者所使用的流量資料種類，並且深究這些流量資料的產生方法、代表意義與可能的應用範圍。研究者在利用這些網站流量資料撰寫研究報告或論文時，同樣也需要對使用的資料加以解釋，詳細地區別上站人數、網頁瀏覽數目等不同流量資料的意義，說明資料的應用方法。以下本研究簡要地探討利用Alexa所提供流量資料的蒐集方法的限制。

首先，這類型資料蒐集方法依賴Alexa網站提供經過統計處理的資料，然而Alexa網站先前所提供的全球網站的流量資料是透過網路使用者在用戶端電腦上安裝的瀏覽器工具列進行蒐集。一般而言，只有電腦與網路技能比較高的使用者才會主動安裝Alexa工具列，這些使用者在使用上偏好的網站，與佔使用人口大多數的一般使用者明顯地會有所不同，這表示Alexa估算的網路使用流量與其他的公司或組織所估算的結果不同，並且也與真正的情形可能有所偏差。這是Alexa受到相當多人批評的地方，比方說，前一陣子有人曾舉出Alexa網站所提供的Youtube網站的流量排名已經超越Google的例子，證明Alexa資料的不準確（Pallas, 2007, October 30）。因此，Alexa在2008年4月特別針對這個情形，改變流量資料的蒐集方法以及網站排名的計算公式（Mack, 2008, April 15）。然而，不論是原先或是現在的流量計算方法，Alexa都將資料蒐集的詳細過程列為商業機密，研究者無法從Alexa獲得詳細說明或原始資料，無法了解資料產生的以及偏差的嚴重程度，以致於運用流量資料所得到的分析結果很難與其它調查研究相互比較，也無法運用加權來調整各種使用情形使得預測的結果更加準確，降低了這些資料的利用價值。

其次，由於無法取得原始資料，研究者無法利用精確與有效的方式解決某些研究問題。Alexa只提供全球流量前十萬名的網站資訊，然而具有地區特色並有特定使用者社群的網站大多屬於流量較低的網站，因此研究者若是要研究這些網站的使用者行為，無法從Alexa取得這些網站的流量資訊。Alexa的另一個問題是利用網站的網域名稱（domain name）來進行流量統計，對於Google等分散於多個不同國家網域的網站相當難得到整體的統計資訊，如果不特別留意的話，這類型網站的流量往往會

被低估。而且Alexa所提供的統計資料是以整個網域名稱作為流量統計單位，除非網站管理者以子網域作為網站內各種資源或功能的命名方式，否則便無法取得這些資源或功能的詳細使用資料，因此要精確地區分各個網站的不同使用行為並不容易。以入口網站為例，這類型的網站提供電子郵件、購物、拍賣或相簿等多種不同功能吸引網路使用者利用。雖然使用者在使用這些功能時的使用情境與所需技能極不相同，但在Alexa僅提供以網域名稱作為流量單位的統計資訊的情形下，這種限制造成研究者很難區分使用者在這些網站中不同功能的使用情形，也因此無法了解使用者每次進入入口網站的目的。

利用Alexa網站蒐集網站長期流量資料也不方便。Alexa每天會更新前一天、前七天與前三個月的連線使用者比率、平均網頁瀏覽數和流量排名等流量資訊以及不同時間間距的歷史流量圖，但並沒有提供按照日期查找過去流量歷史資料的功能。研究者如果想要驗證研究假設的話，只能利用各種歷史流量圖進行大略的估計；或是必需要經常上Alexa網站檢索並記錄所需的資料，要蒐集到具有統計意義的資料量需要相當一段長久的時間與耐心。並且由於這些流量資料都是由Alexa放置在其網站上，Alexa若不提供這些資料時，它們的可信度與穩定性都有堪慮之處。

最後，運用這種方法所蒐集來的資料正像利用網站工作紀錄來分析使用者行為的研究一樣，所能夠獲得的訊息相當表面。事實上，Alexa網站提供各個網站的連線使用者比率及平均網頁瀏覽數等流量資料，是一種相當微觀且具體的行為現象。利用這些資料進行分析，研究者可以知道當時流行的網站，但無法真正得知使用者利用這些網站的動機與目的。因為每一位使用者本身的知識領域、動機和應用目的不同，使得連線到訪同一個網站的用途與價值都有所不同。同時，因為並沒有考慮在網站中瀏覽多少網頁或多久的時間以上才計算為一次連線，所以這些資訊可能會包含錯誤點選、自動彈出式廣告、甚至其他搜尋引擎的到訪，研究者無從得知個別使用者在利用這個網站的實際狀況。因此，若是研究者要進一步探索這些議題，最好能夠配合其他的研究方法，從不同面向進行現象的蒐集與解讀。

陸、結論

從以上關於網路統計資訊的分析與應用，我們可以了解這種資料蒐集方法可以從大量的實際網路使用記錄中，提供研究所需的資訊。研究者可以經由長期持續地蒐集並且觀察Alexa所提供的網站流量資料，

推測網路上大多的使用者所需要或偏好的網路資源以及其取用方法，進而了解在某一個時空環境中他們的線上行為及生活型態。然而，利用 Alexa 的網站流量資料作為研究資訊也有其本身的限制與適用性，包括資料偏差與校正、可信度、資訊的表現範疇等問題。因此，本研究認為這樣的資料蒐集方法所提供的資料不一定適合做為嚴謹的研究事證，但可以作為其他研究方法的初步研究分析。由於網路使用行為本身具有高度的複雜性，研究者可以針對從網站流量資料分析得到的某些現象，進行調查研究或田野觀察，運用多元的資料蒐集方法從各個面向與層次探討線上的使用行為。研究者可以根據網站流量資訊的分析結果提出研究假設，設計問卷題項，進行統計檢驗；或者研究者可以將分析結果發展為問題意識，進行觀察與質性分析。比方說，前述關於論壇類型的網站提供的功能是否取代社會網絡網站，造成後者在臺灣不甚風行，便相當值得深入研究。

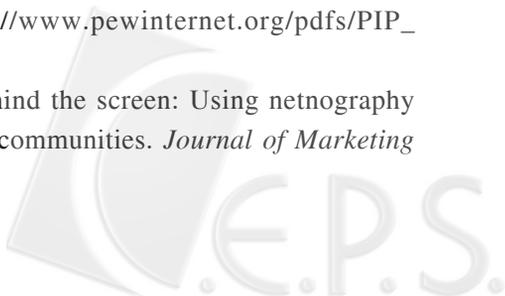
誌謝

本論文由原先發表在「資訊方法論工作坊」的論文修改而成。本論文能夠完成，作者相當誠摯地感謝多位圖書資訊學界的先進給予的指導與鼓勵，包括籌辦「資訊方法論工作坊」的葉乃靜教授、提出相當寶貴建議的吳美美教授與楊美華教授、推薦本論文在修改後投稿本刊的賴鼎銘教授以及本論文的審稿委員。

參考文獻

- 中国互联网络信息中心（2007）。中国互联网络发展状况统计报告。上網日期：2007年8月30日，檢自：<http://www.cnnic.net.cn/uploadfiles/pdf/2007/7/18/113918.pdf>
- 林頌堅（2000）。從網站工作記錄觀察大學圖書館網站使用情形。**大學圖書館**，4(2)，2-22。
- 孫鴻業（2006）。2006年我國家庭寬頻、行動與無線應用現況與需求調查：家戶指標。上網日期：2007年8月30日，檢自資策會FIND網站。（<http://www.find.org.tw/find/home.aspx?page=many&id=149>）
- Babbie, E. (2004)。研究**方法：基礎理論與技巧**（邱泯科等譯）。臺北市：雙葉書廊。（原作於2002年出版）

- Ahn, Y-Y., Han, S., Kwak, H., Moon, S., & Jeong, H. (2007). Analysis of topological characteristics of huge online social networking services. *Proceedings of the Sixteen International World Wide Web Conference* (pp. 835-844). Retrieved May 8, 2008, from <http://www2007.org/papers/paper676.pdf>
- Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J., & Shneiderman, B. (2004). Determining Causes and Severity of End-User Frustration. *International Journal of Human-Computer Interaction*, 17(3), 333-356.
- Compete: FAQ (n. d.). Retrieved December 11, 2007, from <http://www.compete.com/help#snp1>.
- Fallows, D. (2006, December 13). Looking for information about a place to live. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Place_to_Live_2006.pdf).
- Fox, S. (2006, October). Online Health Search 2006. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Online_Health_2006.pdf).
- Fox, S., & Beier, J. (2006, June 14). Online Banking 2006: Surfing to the Bank. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Online_Banking_2006.pdf).
- Horrigan, J. B. (2006, November 20). The Internet as a Resource for News and Information about Science. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Exploratorium_Science.pdf).
- Horrigan, J. B. (2006, September). Politics Online. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Politics%20Aug06_Memo.pdf).
- Horrigan, J., & Rainie, L. (2006, April 19). The Internet's Growing Role in Life's Major Moments. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Major%20Moments_2006.pdf).
- Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research*, 39(1), 61-72.



- Internet World Statistics (n. d.). Internet Usage Statistics. Retrieved November 6, 2007, from <http://www.internetworldstats.com/stats.htm>
- Lenhart, A., & Fox, S. (2006, July 19). Bloggers: A Portrait of the Internet's New Storytellers. Retrieved January 8, 2008, from Pew Internet & American Life Project (<http://www.pewinternet.org/pdfs/PIP%20Bloggers%20Report%20July%2019%202006.pdf>).
- Lenhart, A., & Madden, M. (2007, January 3). Social Networking Websites and Teens: An Overview. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_SNS_Data_Memo_Jan_2007.pdf).
- Mack, G. (2008, April, 15). The Alexa ranking system has been changed. Retrieved May 8, 2008, from <http://awis.blogspot.com/2008/04/alexaranking-system-has-been-changed.html>.
- Madden, M. (2007, July 25). Online Video. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Online_Video_2007.pdf).
- Madden, M., & Fox, S. (2006, May 2). Finding Answers Online in Sickness and in Health. Retrieved August 30, 2007, from Pew Internet & American Life Project (http://www.pewinternet.org/pdfs/PIP_Health_Decisions_2006.pdf).
- Nicholas, D., Huntington, P., Lievesley, N., & Wasti, A. (2000). Evaluating consumer website logs: A case study of The Times/The Sunday Times website. *Journal of Information Science*, 26(6), 399-411.
- Nicholas, D., Huntington, P., Williams, P., Lievesley, N., Dobrowolski, T. & Withey, R. (1999). Developing and testing methods to determine the use of web sites: Case study newspapers. *Aslib Proceedings*, 51(5), 144-154.
- O'Reilly, T. (2005, September 30). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Retrieved November 28, 2007, <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- Pallas, D. (2007, October 30). Google vs. YouTube? Retrieved December 12, 2007, <http://awis.blogspot.com/2007/10/google-vs-youtube.html>
- Rioux, K. S. (2004) Information acquiring-and-sharing in internet-based environments: An exploratory study of individual user behaviors.

Unpublished doctoral dissertation, University of Texas at Austin.

Wang, P., Hawk, W. B., & Tenopir, C. (2000). Users' interaction with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing and Management*, 36, 229-251.