

Web information seeking by pages: an observational study of moving and stopping

[Jarkko Kari](#)

Department of Information Studies
University of Tampere, Tampere, Finland

Abstract

The intention of this paper is to look at how the World Wide Web is used in looking for information in the domain of personal development. The theoretical aim of the paper is to elaborate conceptual tools for understanding better the content of Web pages, as well as navigation through the Web. To obtain detailed and valid data, totally free-form Web searches by fifteen individuals were observed and videotaped. The 1,812 pages visited by the informants, along with their actions therein, were examined and coded. The study explores the subject, language and content type of the viewed pages, as well as the tactics, strategies, interfaces and revisitation in moving from one page to another. Correlations between the variables are also analysed. One of the most interesting discoveries was the wide variety of different tactics for moving around the Web, albeit that only clicking on links and pushing the Back button stood out from the rest. The paper ends by presenting sundry theoretical, methodological and practical contributions of the research to the field of Web searching.

Introduction

The impetus for this paper is given by the statement that "the Internet is increasingly recognized for the vast array of information, services, meeting places, and communities-of-interest that it offers" (Scull *et al.*, [1999](#): 17; see also Hölscher & Strube [2000](#)). Although studies dealing with Web searching have proliferated in recent years, research of this kind has been limited to only some aspects of the phenomenon.

For one thing, there has been a neglect of *content* - the information which is provided by Web pages (see Wang *et al.*, [2000](#)). This is baffling, since content "is the basis of success in finding needed information" ([2000](#): 234; cf. Odlyzko [2001](#)). For instance, a "GOMS [[1](#)] model of Web users", which only paid attention to structural Web elements, did not correlate with people's actions (Pirolli & Fu [2003](#): 1). Khan and Locatis ([1998](#)) took a very narrow slice of Web content by investigating the impact of links on information retrieval. The results indicated that the correspondence between the terminology in search tasks and links increased the efficiency, but not the accuracy of searching. On the other hand, O'Neill ([1998](#)) has proposed a general, semantic typology of Web pages (see [Results and discussion: content type](#)). Light ([2001](#)) even conducted an empirical examination concerning the role of Web content. It seems that the theme of Web sites affected the participants' disposition to interact with those sites.

For another thing, research has exhibited an infatuation with stereotyped information search *strategies* like browsing (e.g., Brown & Sellen [2001](#); Catledge & Pitkow [1995](#)) or using search engines (e.g., Ford *et al.*, [2002](#); Ford *et al.*, [2003](#); Moukdad & Large [2001](#); Ozmutlu *et al.*, [2003](#); Spink & Ozmutlu [2002](#); Spink *et al.*, [2001](#); Su [2003](#); White *et al.*, [2003](#); Wolfram *et al.*, [2001](#); Xie [2003](#)). Scrutinizing Web searching in all its variety and freedom - as in the

work of Choo and his colleagues (1999), Cothey (2002), Fidel *et al.*, (1999), Hill and Hannafin (1997), Hölscher and Strube (2000), as well as Wang *et al.*, (2000) - has not been a popular approach, especially after the turn of the millennium. Even these inquiries have not succeeded in capturing the whole range of Web movement activities. Yet, such holism would be crucial for the validity of empirical data.

These two states of affairs are symptoms of biases from which information research in general suffers: a tendency to favour structure or process at the cost of content, and an obsession with information systems that eclipses real-world phenomena. Taking the above remarks into due consideration, this paper focuses on free-form, everyday life information seeking through the Internet, especially the World Wide Web. The purpose of the report is twofold. The *empirical* intention is to look at how the Web is used in looking for information in the domain of personal development. The *theoretical* aim of the paper is to elaborate conceptual tools for understanding better the nature and types of Web pages, as well as information seeking tactics and strategies.

It is appropriate to define the *Web* as an interactive and collaborative "information environment" (e.g., Catledge & Pitkow 1995: 1066; Tewksbury & Althaus 2000: 128) that is mainly composed of hypermedia and hypertext documents linked to one another (see e.g., Catledge & Pitkow 1995; Lazonder *et al.*, 2000), and distributed over the Internet (see Choo *et al.*, 1999; cf. O'Neill 1998). A *Web page* is conventionally seen as a discrete, electronic document "that is identified by a unique universal resource locator (URL)" (O'Neill, 1998: 115; see also Catledge & Pitkow 1995).

An information search *tactic* can be demarcated as a concrete method of moving from one page to another in the Web. After each (tactical) move, the seeker has to stop at some page. As Choo *et al.*, (1999) found out in their multimethod study of Web information acquisition, a pattern of Web moves is indicative of a certain *strategy* of information searching. In the current investigation, this concept means the broad way in which the individual looks for information in the Web. Thus, the construct does not denote here the overall strategy of information **seeking**, which may concern sources other than the Internet. It is helpful to see the strategy as still relating to a particular move, but being a generalization of the specific tactic (cf. e.g., Marchionini 1995).

Parallel to the evolving information horizon - particularly the Internet - is the human being, who him/herself is also in the process of becoming. If we accept the view that development is a fundamental characteristic of living organisms (Deci & Ryan 1985; Piaget 1971), this sounds only natural. By *personal development* or growth we mean that the individual improves his/her own abilities, skills, knowledge or other qualities by working on them. It is therefore a matter of the actor augmenting and realizing his/her own potential (Deci & Ryan 1985; cf. Magnusson 1995). This is not a solitary phenomenon, but comes about through reciprocal interaction with one's environment (Deci & Ryan 1985; Magnusson 1995). Self-development arises from our needs, and affects our behaviour (Deci & Ryan 1985). Psychological research suggests that personal growth is a variety of coping (King 2002) and well-being (Compton 2001). Especially in these times of rapid change, the continual developing of oneself has almost become a necessity, even in free time. On the other hand, information society with its networked services provides quite innovative solutions to the need for development. In the current research project, personal development represents a pervasive domain-context which extends to the spheres of both work and leisure.

This article seeks answers to three research questions which are further specified below:

1. What kinds of Web pages do personal developers (PDs) visit?
 - What is the main subject of the pages?
 - Which language are the pages primarily in?
 - What is the dominant content type of the pages?
2. How do the PDs move from one page to another?
 - What is their principal tactic?
 - What is their chief strategy?
 - Which interface do they basically use for moving?
 - Have the pages been viewed before?
3. Do the type of page and manner of moving correlate?

Methods

Because relatively little was still known even about the basics of Web information seeking when this project

commenced (in 2001), an exploratory mode of research was in order. As a mostly descriptive work, for example, like that of Hill and Hannafin (1997), the inquiry chiefly probes the study object in an inductive manner, allowing ideas to surface from the data. This enables the valid specification of the concepts of Web page and movement.

Participants

The research effort was centred on individuals who were sufficiently motivated to go through the various phases of the investigation (see Savolainen 1998b). Accordingly, persons were sought who were interested in developing themselves, and who also used the Internet in connection with their development. For practical and financial reasons, it was sensible to limit the dispersion of the informants to the province of Pirkanmaa (in Finland), for the study is performed in its capital, Tampere. Since depth rather than generality was our aspiration, we deemed a sample of twenty individuals adequate. Because there was no exhaustive or even representative register of self-developers, the only option was to seek out volunteers from a number of quarters in the hope of reaching at least some degree of coverage. Considering the theme of the project - "Self-development and Internet use" - the Internet was probably the best vehicle for contacting participants. Therefore, a notification about informants wanted was sent by e-mail to the local public library, adult education centres, and a computer club for senior citizens (altogether five organizations), who forwarded our message to their own people, as well as putting our hyperlink on their Websites.

As a result, eighteen individuals were persuaded to take part in the research. Because authenticity is a central element in the project at hand, each participant was studied on his/her own terms, that is, where and when s/he wished. To the surprise of the researchers, a preponderance (fifteen) of the recruits opted for the departmental facilities, whereas two informants selected their home, and one preferred his workplace. The research space at the university was a standard meeting room furnished with a computer and video recording apparatus for the purposes of this inquiry. If an individual wanted to perform a Web search somewhere else, the immobility of the video equipment prevented screen capturing on site. Because video data is crucial to the current report, the three "dissidents" had to be excluded here.

Curiously, only five of the fifteen informants were male. This bias towards women may have been inadvertently caused by an appeal on the Web form saying "Help men out now!". The participants represented all age groups between ten and seventy years, their average age being thirty-seven years ($n=14$) [2]. The educational distribution of the subjects did not come as a surprise: one of them had no degree at all (since he was still in the primary school), none had just a high-school diploma, eight had an intermediate (college, upper secondary school, or vocational school) diploma, and four had a university degree ($n=13$). Their occupational status was such that seven individuals were studying, four were working, and three did neither of those ($n=14$): one of these was a pensioner, another was on vacation, and the third person was unemployed. The high proportion of students could be explained by the recruitment channels used. The participants' Internet experience varied between two and ten years, with a mean of five years. It did seem that both novice and expert searchers were rare in this group.

Data collection

The main part of the data was collected during November 2001 - January 2002. Following the example set by the bulk of earlier Internet studies (e.g., Choo *et al.* 1999; Rieh 2002; Wang *et al.* 2000), this investigation used multiple methods of data collection, for different research questions demanded different procedures. Interviewing was the core technique, covering the context of situational Web information seeking. The real-time scrutiny of Web interaction, in turn, addressed the research questions in this paper, and necessitated observation (by the researcher) and thinking aloud (by the participants). Wherever a Web session took place, the participant always had the liberty to select any available browser program s/he wished, and to carry out the search as s/he saw fit, on a subject of his/her own choice. There was no real time limit, either. The duration of the sessions fell somewhere between half an hour and two hours. The sole restriction was that the search topic had to concern personal development.

Sometimes, a participant looked for information on more than just one subject within a single Web session; the maximum was as many as seven topics. In a few other cases, the person searched the Web for the same thing on two separate occasions, under surveillance. A great majority of the informants felt that it would take more than three (some even spoke of dozens) rounds of Web searching to satisfy their then need for information (cf. Fidel *et al.*, 1999). For understandable reasons, it would have become impossible to document all of these instances, and therefore only one search session per person was usually deemed sufficient. Two sittings took place with three

participants (separately), as this appeared more appropriate in their situations. Pharo (1999) also observed one search session per individual, but this is far less than, say, the three sessions employed by Fidel *et al.*, (1999). The minimal line in this regard was justified by aspirations to profound, in-depth data gathering and analysis.

When a participant came to the Web laboratory, the search session was captured on videotape. This was enabled by a computer-to-TV converter ([A VerKey500](#), which had been inserted between the PC and its monitor. A microphone was also attached to the video cassette recorder (VCR), which arrangement synchronized the events on the computer screen with the searcher's speech. The video tapes came to hold a total of some seventeen hours of data.

Data analysis

All empirical material was then transformed into computer-readable text. The data processing involved transcribing the audio recordings, as well as interpreting the taped video films. The current article examines the observational (video) data, supplemented by thinking-aloud. The video cassettes were played back, and the participants' navigational paths were manually logged (see Hill & Hannafin 1997), page by page. A Web page may be composed of various (visual and audio) elements (see Pharo 2002), but here an entire page is taken as the smallest meaningful unit of analysis (see Rieh 2002; cf. Cothey 2002) for Web information-seeking researchers.

Operationalizing a Web *page* became a critical issue. In this study, a Web page included everything that could be displayed by scrolling in a single browser window at a particular moment (cf. Pharo 2002). Whenever the content of the selected window changed altogether, by whatever means, this was regarded as moving on to the next page. It was ordinarily a question of activating (one way or another) a different URL. Revisiting a previously-viewed address was counted as a separate page, but going to another part within the same page was not. The number of visited pages ranged between fifty-nine and 236, depending on the individual. The average was 121 pages, whereas the total number of Web pages in the corpus was 1,812. In each case, the manner of moving from one page to another could be discerned by following the mouse pointer, scroll bars, and text boxes on the screen. So for every Web page that an informant visited, certain information—particularly that pertaining to the research questions above—was noted.

The first phase of actual analysis was coding, which laid the foundation for and merged with later stages. In this basically qualitative task, the major methods were content analysis (as in Rieh 2002, for example), and classification. Typologies were constructed according to single dimensions, so that the various types became mutually exclusive. Page topic was the sole concept that was categorized deductively. The top choice for a classification scheme would have been the *Universal Decimal Classification* (UDC), but it is not freely and fully available (see <http://www.udcc.org/outline/outline.htm>). Therefore, the second best alternative was chosen, that is, the *Finnish Public Libraries Classification System* (PLC), which in fact resembles UDC.

The second phase involved extracting quantitative, descriptive measures, such as frequency and percentage (cf. Jansen & Pooch 2001), from the coded material. In the third phase, simple statistical operations were performed between pairs of variables. Since these are all of nominal scale, contingency table analyses were carried out. As the size of the table was always bigger than 2 X 2 cells, the proper correlation coefficient was Cramér's V (see Elifson *et al.*, 1990). The accompanying significance level (p) was also calculated. The quantitative work was done by the help of statistical software - *StatView SE+Graphics 1* (the latest version is [JMP 5](#)).

Quality

The validity of the research results was affected by several factors, both negative and positive. It was most of all weakened by the fact that the Web searches studied in this article were conducted somewhere other than in the participants' natural environment. Moreover, the resolution of the video recordings left something to be desired, as the body text on the screen often remained unreadable. These shortcomings are above all compensated for by having collected the data in real time, thus avoiding pitfalls of retrospection. The laboratory approach is also balanced by grounding the analysis in the data. If something about a Web page seemed unclear or ambiguous, the original document (on the Internet) could usually be consulted. Hence, the project should yield moderately valid findings.

It was not to be expected that the reliability of the results would pose serious problems, as the empirical material was gathered and coded quite systematically. However, it was still given a boost by double-checking some of the codes

(see Klobas & Clyde [2000](#)). Due to the small and self-selected sample, as well as the time lag of two years (between data collection and reporting), the quantitative generalizability of the results may be low, which is almost a hallmark of Internet usage research (see Savolainen [1998b](#)). This is not a dilemma, however, because the aim is mainly theoretical.

Results and discussion

Pages

Subject

The topic of the Web pages was determined by looking at the title and other cues on the page. The subject classification and the shares of the various topics are listed below in Table 1. The categories are self-explanatory, except for "none", perhaps. This means that some pages did not have any content whatsoever, as they were empty. Many of these were initial pages, chiefly since the (Web) browsers were configured to show a blank window on start-up, or when pushing the Home button. Other pages looked void, because they did not get a chance to load their content.

Subject*	f	%
0 General works, books and publishing, libraries, general cultural policy, mass communication (e.g., Makupalat [Titbits])	298	16.5
1 Philosophy, psychology, paranormal phenomena (e.g., Aforismer. ordspråk och citat [Aphorisms, proverbs and quotations])	83	4.6
2 Religion (e.g., Mystinen yhtyminen ja meditaatio [Mystical union and meditation])	26	1.4
3 Society (e.g., Työelämä [Working life])	225	12.4
4 Geography, travel, ethnology (e.g., Travel)	86	4.8
5 Natural sciences, mathematics, medicine (e.g., Expedition Three Space Walks)	94	5.2
6 Technology, industry, handicraft, agriculture and forestry, domestic science, business economy, traffic (e.g., Uranusfin.comin asiakkaita [Uranusfin.com's customers])	392	21.6
7 Arts, sports (e.g., Alariesto-galleria [Alariesto Gallery])	143	7.9
8 Fiction and poetry, literary studies, linguistics (e.g., Langue française [French language])	133	7.3
9 History (e.g., Sonkajärven historiaa [History of Sonkajärvi])	79	4.4
More than one** (e.g., SpiritLinks: Miscellaneous [sic!] Sites)	55	3.0
None	56	3.1
Unknown***	142	7.8
Total	1812	100.0
* Source of classification: Ministry... 2003 . ** These pages exhibited two or more equally strong topics instead of one major theme. *** These pages remained unclassifiable.		

Table 1: Subject of pages

On the basis of Table 1, it appears that technology and economy (Class 6) was the most typical thematic area with its portion of roughly one fifth, followed by general works (Class 0; one sixth) and society (Class 3; one eighth). It also seems that religion (Class 2; one 70th) was clearly the least popular page topic. Only two of the fifteen informants viewed some religious pages.

The prevalence of Class 6 cannot be explained by generic navigational (later referred to as "meta-informational") pages, for these belong to Class 0. Instead, the category of technology and economy is so broad in scope that this peculiarity alone must have magnified its occurrence. Moreover, the medium of information seeking in this study is suspect: the Internet as a technology is, after all, an ideal instrument for obtaining technical information. Religion, on the other hand, is a sort of antithesis of technology, thereby perhaps discouraging religious Internet searching. Given the domain of interest here, the scarcity of Class 1 pages, which would include the matter of human development, can appear baffling. This observation, however, is probably explained by the participants' tendency to focus on a specific, substantial aspect of self-development, rather than personal growth in general.

Overall, the distribution of the page topics across the classification scheme may have been influenced by the scheme itself. One must keep in mind that such a scheme does not equally divide different subject areas into the major classes. This hypothesis is supported by the correlation between the breadth and frequency of Class 6 in Table 1.

Language

The Web page languages are listed in Table 2. The language typology presented here is, of course, culture-specific in that it is tailored for Finns. A universal taxonomy could comprise the classes of native language, second language, and other.

Language	f	%
Finnish	985	54.4
English	547	30.2
Other*	109	6.0
More than one**	34	1.9
None***	56	3.1
Unknown	81	4.5
Total	1812	100.1
* e.g., French or Swedish.		
** These pages exhibited two or more equally dominant tongues.		
*** The pages were empty.		

Table 2: Language of pages

Table 2 tells us that over half of the viewed pages were in the Finnish language, whereas English pages were placed second with nearly one third of the pages. Other languages appeared to be marginally relevant (1/17) only. Finnish was presumably the mother tongue of all participants, therefore its predominance. English, in turn, has long been the first foreign language in Finland. So the distribution in Table 2 suggests that searchers, in part, choose Web pages on grounds of their language proficiency, which sounds rational. Considering the widespread utilization of the English language in the world at large, it is logical to anticipate that in countries where English is the first tongue, the manifestation of all other languages in Web searching would be much less conspicuous than in Table 2, even among self-developers.

Content type

The data held four types of page content: information, meta-information, forms and hybrids (Table 3). An *information* page predominantly represents some part of the perceived reality (see Kari 1996). *Meta-information*, in turn, betokens hyperlinks or references to other pages or sources, as in a "manually generated" index, search engine, or search result list (Pharo 1998). In Pharo's terms (e.g., 2002), meta-informational pages are "information surrogates". These two kinds of content have also been called primary and secondary information, respectively. A third page type turned out to be Web *form*, an interactive document by which the actor can ask questions or provide answers to another party through the Internet. Content of this sort greatly differs from the previous two in that a form casts the user in the role of information source instead of seeker. A *hybrid* page is some balanced combination of the three major varieties of content (cf. Pharo's (1999: 211) "conglomerates").

Type	f	%
Information (e.g., Math in HTML (and CSS))	499	27.5
meta-information (e.g., Tampere tietokaupungiksi - Infoengas [Tampere for a knowledge city - Info Ring])	956	52.8
Form (e.g., Kysy ja anna palautetta [Ask and give feedback])	18	1.0
Hybrid (e.g., NASA Human Space Flight)	277	15.3
Empty	56	3.1
Unknown	6	0.3
Total	1812	100.0

Table 3: Content type of pages

According to Table 3, meta-informational pages were viewed by far most frequently (528%). The share of informational pages was only 27.5%, whereas forms made up a mere one per cent of the corpus. The high proportion of meta-information is a curious finding. The reason for this is not immediately obvious, but it might have something to do with the study setting. One could hypothesize that because these informants conducted their search in an unfamiliar physical environment, their ability to take advantage of familiar, local resources (such as bookmarks) was inhibited. As a result, they would have been "forced" to resort more frequently to Web-based finding aids, and this would have inflated the occurrence of meta-information. But does the rarity of forms denote that the participants seldom turned to search engines? The answer is "no": search engine interfaces were usually one element on meta-informational or hybrid pages, and this is why they do not stand out in the typology above.

A rare and quite different alternative to the Web page content typology above was presented by O'Neill (1998), who devised the classes of non-fiction, fiction/entertainment, reference/index, institutional and personal. While "reference/index" is obviously analogous to meta-information, the other types can be considered as varieties of information proper. O'Neill's taxonomy does not seem to be very usable, however. As a combination of various dimensions, the categories are mutually inclusive: for instance, Sonera's (a Finnish company) service form in Table 3 is both "non-fiction" and "institutional". They are also severely skewed against meta-information.

Moving

Tactic

The tactics of moving in the World Wide Web manifested considerable versatility, as revealed by Table 4a. Linking (with its share of almost 45%) was definitely the most popular, followed by backtracking (25%). Other navigation techniques occupied more or less marginal positions. The frequent use of hyperlinks and the Back button is nothing new, for this tendency has also been found in several earlier inquiries (e.g., Catledge & Pitkow 1995; Fidel *et al.*, 1999; Wang *et al.*, 2000; cf. Savolainen 1998a).

However, it may appear somewhat surprising that carrying out a search - albeit placed third - only happened with about every 13th page. This figure hides the fact that most of the time, leaving a page in a search engine Web site took place by some means other than querying. It is interesting to note, though, that the browser's built-in Search function was used only once. This is a good example of the detected overall disregard for the browser's own navigational tools (cf. Choo *et al.*, 2000). Among these, entering a URL was the most common tactic, but even that ranked as low as number five in Table 4a.

Tactic	Definition	f	%
Answering	submitting a non-search Web form	5	0.3
Automatic	another page loaded by itself in the same window (e.g., redirecting)	26	1.4
"Back"	pushing the browser's Back button	459	25.3
Button	pushing some other than a standard browser button	5	0.3

Closing	shutting the current window	70	3.9
"Forward"	pushing the browser's Forward button	6	0.3
Hiding	concealing the window	1	0.1
History	using the browser's History list	4	0.2
"Home"	pushing the browser's Home button	14	0.8
Launching	starting a computer program	16	0.9
Linking	clicking on a hyperlink	810	44.7
Menu	choosing a menu item	32	1.8
New	a new (e.g., pop-up) window opened automatically	24	1.3
Quitting	exiting the program	6	0.3
Resetting	rebooting the computer	3	0.2
"Search"	pushing the browser's Search button	1	0.1
Searching	executing a query with a search engine	140	7.7
Slipping away	the window automatically moved behind another window	4	0.2
Switching	activating another window	93	5.1
URL	entering a URL in the address box	83	4.6
None	the person did not go any further	9	0.5
Unknown	the manner of movement could not be observed	1	0.1
Total		1812	100.1

Table 4a: Tactic of moving

According to Table 4b, most of the twenty techniques of Web navigation have not been investigated in prior research. Even so, there were some possible tactics missing: at least Bookmarks or Favorites, cloning (the window), crashing (of the program), disappearing (the window closes by itself), Go (using the browser's Go list) and opening (a new window). In this study, the reason for avoiding bookmarks is obvious: the searchers did not use their own computer with their own bookmark list. The results of some previous inquiries (Savolainen [1998a](#); Wang *et al.* [2000](#)) indicate that bookmarks are not utilized too much anyway. As for the Go menu, its use was apparently replaced by employing the History list.

This study (2004)	Catledge & Pitkow (1995)*	Kim (2001)
Answering		
Automatic		
"Back"	Back	Back button
Button		
Closing	Close window	
"Forward"	Forward	
Hiding		
History		
"Home"	Home document	Home button
Launching		
Linking	Anchor	Embedded link
Menu		
New		
Quitting	Exit program	

Resetting		
"Search"		[Keyword search]
Searching		Keyword search
Slipping away		
Switching		
URL	Open URL	
	Clone window	
	Hotlist [=Bookmarks]	
	New window	
	Open local** (file)	
	Source document*** (view programming code)	
		Jump options (e.g., Go & History)****
<p>* Here, only those tactics are enumerated which conform to their definition in this study. That is, actions that are strategic or not about moving from page to page were left out.</p> <p>** This could come about in connection with Button, Launching, Menu or URL (on the left).</p> <p>*** One result of Menu.</p> <p>**** Also includes Catledge's and Pitkow's (1995) "Hotlist"</p>		

Table 4b: Movement tactics across versatile Web searching studies

Strategy

It became feasible and even necessary to see the multitude of specific tactics as instances of more general strategies of navigation. Three such methods emerged: pointing, typing and following. When a person *points*, s/he moves on to the next page by essentially pushing a symbol of one kind or another. In the investigation at hand, this was ordinarily done by clicking a mouse button, but sometimes a keyboard shortcut or the computer's Reset switch was a more viable solution [3]. Advancement as a consequence of inputting a character string into one or more text fields is called *typing*. At a conceptual level, the most fundamental difference between those two methods is that pointing is a close-ended strategy, because the searcher chooses one of the predestined routes, whereas typing is an open-ended strategy, since the person essentially creates a path of his/her own through the Web (cf. Kari's (2001) closed vs. open-ended information needs).

The third strategy - *following* - signifies that the individual momentarily loses control over manoeuvring, while the computer takes him/her to another page. In this project, it was not a matter of artificial intelligence acting or the machine getting muddled; in all likelihood, it happened because the programming behind some of the Web pages instructed the browser to jump to a designated address. Automation brings a new, previously unnoticed member to the family of Web search strategies. Thus, we have pointing and typing as active, user-driven strategies, and following as a passive, computer-governed strategy. These represent various degrees of freedom in searching for information: typing gives the user the "freest hands", whereas following means that his/her "hands are tied".

Upon examining the quantitative aspects of Table 5a, one can immediately perceive the multiplicity of pointing tactics in contrast to the other strategies. This is presumably brought about by the fact that modern browsers (and operation systems) have a graphical interface which favours the mouse as the primary input device (cf. Smith's and others' (1997) parallel statement on browsing). In moving between pages, pointing was also the preferred strategy, since it accounted for an overwhelming five sixths of the traffic. This observation resembles earlier results on browsing in the Web (e.g., Iivonen & White 2001; Ylikoski 2003). With its share of one eighth, typing was left far behind. 'Following' was even more exceptional, as this only took place every 34th time on an average. Considering the absence of personal bookmarks and settings, the absolute supremacy of pointing was remarkable. From this, one

may speculate that pointing is even more common when a person searches the Web on his/her own computer.

Strategy	Includes these tactics	f	%
Pointing	"Back" Button Closing "Forward" Hiding History "Home" Launching Linking Menu Quitting Resetting "Search" Switching	1520	83.9
Typing	Answering Searching URL	228	12.6
Following	Automatic New Slipping away	54	3.0
None	None	9	0.5
Unknown	Unknown	1	0.1
Total		1812	100.1

Table 5a: Strategy of moving

Table 5b compares various typologies of Web movement strategies. The rows exhibit strategies that are similar across different pieces of research. By reading the chart this way, one gets the impression that *pointing* is roughly the same as browsing by following links, whereas *typing* refers to searching. The results above suggest that this is often true, but not always. For instance, a person may proceed by querying, when s/he is really just browsing to see what can be found in the Web with particular words. So there is a noticeable difference here: the strategies of pointing and typing are physical motion methods that can be directly observed, and concern a single page. In turn, browsing and searching are more like cognitive modes of operation which cannot be directly observed, and concern multiple pages (i.e. a single search episode; see Marchionini 1995). It is precisely because of this partial incompatibility, or disparity in viewpoints, that many of the older strategic categories do not seem to match the new ones (in Table 5b).

This study (2004)	Choo et al. (1999 & 2000)	Hawk & Wang (1999)*	Hölscher & Strube (2000)	Pharo (1999)	Puskala (2002)
Pointing	Browsing Chaining	Back & forward going Exploring Link following	Browsing	Linking Scanning	Link-oriented
Typing	Extracting	Engine using Loyal engine using Metasearching (URL)	Searching	Searching	Search-oriented
Following					
	Differentiating	Shortcut seeking	Access (site directly)		
	Monitoring				

	Starting				
		Double-checking			
					Undifferentiated
* Here, only those strategies are enumerated that conform to their definition in this study. That is, methods that are tactical or not about moving from page to page were left out.					

Table 5b: Movement strategies across versatile Web searching studies

In the still-young research tradition of Web searching, there is no equivalent to the strategy of *following*. The closest approximation is Erdelez's (2000) "information encountering", which also implies the lack of a strategy. However, the individual meets with useful information by pure chance, which is not necessarily the case with following. That is to say, encountering is about seizing the opportunity, whereas following is about "going with the stream".

Interface

The analysis exposed four interfaces that were used for moving around the Web: page, browser, operating system, and hardware (see Table 6). These layers constitute a hierarchy in which pages are at the top, and the hardware is at the bottom. In other words, the Web pages are shown in a browser program, the browser works within the operating system, and the operating system runs on the computer hardware. Tampering with the hardware was the very last means of continuing the search, when nothing else could be done. Table 6 also enumerates which tactics of moving were typically applied with each interface. This connection is indicative only, for due to the small number of several tactics, no correlation could be calculated. Contrary to a finding by Pharo (2002: 103), the table nevertheless shows that "following link" and "entering queries" are by no means the only search tactics that are not dependent on the particular browser software.

Interface	Typical tactics	f	%
Page (e.g., AltaVista)	Answering Automatic Button* Linking New Searching Slipping away	1019	56.2
Browser (e.g., Netscape 6)	"Back" Button* Closing "Forward" History "Home" Menu Quitting "Search" URL Unknown	682	37.6
Operating system (e.g., Windows 98)	Hiding Launching Switching	99	5.5
Hardware (e.g., desk-top PC)	Resetting	3	0.2
None**	None**	9	0.5
Total		1812	100.0
* Button-pushing occurred just as rarely on pages as on the browser level.			
** The person did not go any further.			

Table 6: Interface in moving

As attested by Table 6, the participants mostly (in over half of the cases) moved to the next Web page by activating something on the previous page. At the other end of the continuum, hardware-level motion was quite exceptional (every 604th time). Indeed, one can discern a distinct pattern here: the more general the interface, the less often it was employed for the purpose of moving between pages. There is nothing abnormal about this, because Web pages regularly incorporate requisite navigation tools, after all. The other interfaces provide the reader with additional functions and shortcuts.

Wang and her colleagues' exploratory study is apparently the only antecedent inquiry which has examined the user interface in Web searching. They distinguished "access methods, navigation tools, access results/objects, messages/clues, and input/output (I/O) devices" (2000: 233-234). However, that typology divides the interface elements according to their function, whereas this investigation sees the interface as a layered structure. Without taking a stand on the potential superiority of one perspective over another, it is a fact that two different sets of empirical findings reflecting those two classifications cannot be meaningfully compared.

Revisiting

The process of Web information searching was measured with the elementary gauge of visiting and revisiting pages. In this regard, every viewed page was denominated as either a different, same or similar one (see Table 7). Going to a *different* page means that the person has not been to that page before, under the period of data collection. Its opposite is a *same* page which looks exactly like a page seen earlier in the process. A *similar* page is an intermediate form of these two: it is not really a new page, nor is it identical with another page, but an adaptation thereof. Most of the content remains the same, but there are some noticeable, minor changes. If major changes are perceived, the page would be treated as a different one (cf. Cockburn & McKenzie 2001). The similar pages in this study were almost always products of dynamic content, but many other dynamic pages were also different or same ones.

Difference in pages	f	%
Different (e.g., Medicinsk ordbok [Medical dictionary])	895	49.4
Same (returning to e.g., Medicinsk ordbok)	893	49.3
Similar (modified version of e.g., Medicinsk ordbok after querying with "kardiovaskulär" [cardiovascular])	24	1.3
Total	1812	100.0

Table 7: Visiting pages

By reading the figures in Table 7, we learn that Web searching took the informants to different and same pages equally often (about every second time), whereas similar pages were something of a curiosity (1/76). This distribution communicates that on the average, each distinct page was viewed twice. However, the top frequency of returning to a unique (different or similar; cf. Cockburn & McKenzie 2001) page was seventeen times.

The observed high rate of revisitation runs parallel to, but is still lower than in prior findings, according to which between 58% and 81% of all page viewings are in fact re-views (Cockburn & McKenzie,). The prevalence of revisiting could be partially explained by the structure of Websites (Choo *et al.*, 2000). On the other hand, the inconsistencies between the percentages might depend on the duration of the data collection period in different studies. One would expect that the longer the interval, the greater the share of same (or similar) pages. It is probably not a coincidence that the highest rate was indeed perceived in Cockburn and McKenzie's (2001) investigation which incorporated four months of continuous data, as opposed to the lowest rate in the present piece of research using roughly one hour of search data (per informant).

Correlations

In order to figure out meaningful covariations that would meet the technical requirements of the Cramer's V test, uncertain data had to be excluded. Thus the "unknown" categories, and classes with too few cases (e.g., "hardware" as interface; see Table 6), were ruled out here. Table 8 reports the correlations (on the scale of zero to one) between

the six variables for which they could be legitimately determined. Since causality was not analysed, the direction of influence cannot be determined. It appears that although the connections were not very strong, their statistical significance was generally high. When we look at the two variable groups, it seems that - relatively speaking - page (subject, language & content type) correlations were *high*, page vs. moving covariations (e.g., upper right quadrant below) were *moderate*, and moving (strategy, interface & revisiting) correlations were *low*.

Apparently, the covariation was most intense between the subject and language of Web pages, whereas there was no evidence of any dependency between search strategy and page revisitation. One would expect that in comparison with different pages, people return to same pages more often by straightforward pointing, but this was not the case here. Two interpretations present themselves: the participants accidentally arrived at the selfsame pages, or they did not fully exploit navigational aids (such as the Back button) supportive of the pointing strategy.

	Variable X						
Variable Y	Subject	Language	Content type	Tactic	Strategy	Interface	Revisiting
Subject		0.53**	0.30**	—	0.32**	0.25**	0.15**
Language	0.53**		0.15**	—	0.11**	0.21**	0.09*
Content type	0.30**	0.15**		—	0.18**	0.29**	0.13**
Tactic	—	—	—		—	—	—
Strategy	0.32**	0.11**	0.18**	—		0.13**	0.03
Interface	0.25**	0.21**	0.29**	—	0.13**		0.13**
Revisiting	0.15**	0.09*	0.13**	—	0.03	0.13**	
— The test statistic could not be computed, owing to the lack of observations.							
* p<.01							
** p<.001							

Table 8: Correlations (Cramér's V) between variables (n varies)

Lastly the sole prominent correlation - that between the Web page subject and language - is examined in more detail. Upon perusing Table 9, it turns out that *Finnish* pages dealt with paranormal phenomena (#1), religion, society, geography and history more frequently than on the average, but less frequently with arts and linguistics (#8). Pages in *English*, on the other hand, emphasized general works, natural sciences, technology, and arts at the cost of religion, society, geography, linguistics, history, and miscellaneous topics ("more than one"). When *other* languages were concerned, the pages treated of religion, linguistics, and miscellaneous topics more often than average, but less often of general works, paranormal phenomena, society, geography, natural sciences, technology, and arts. Society was the most common theme of Finnish pages, technology in English pages, and linguistics with pages in other tongues.

	Language***			
Subject**	Finnish	English	Other	Average
0 General works, books and publishing, libraries, general cultural policy, mass communication	16.8	23.0	11.9	18.6
1 Philosophy, psychology, paranormal phenomena	6.1	4.0	2.0	5.1
2 Religion	2.6	0.0	3.0	1.7
3 Society	23.1	2.3	4.0	14.6
4 Geography, travel, ethnology	5.6	1.5	1.0	3.9
5 Natural sciences, mathematics, medicine	5.0	6.5	0.0	5.2
6 Technology, industry, handicraft, agriculture and forestry, domestic science, business economy, traffic	21.1	36.3	0.0	24.9
7 Arts, sports	2.8	21.5	0.0	9.1
8 Fiction and poetry, literary studies, linguistics	5.4	2.1	68.3	8.4
9 History	7.4	1.0	5.0	5.0

More than one	4.2	1.9	5.0	3.5
Total	100.1	100.1	100.2	100.0
* n=1524, Cramer's V=0.53, p=.0001				
** Classes "none" and "unknown" were discarded.				
*** Classes "more than one", "none" and "unknown" were discarded.				

Table 9: Percentages (%) of page subjects by language*

Table 9 does not, however, report the percentages of page language by subject. From this angle, the pages about society - which was the most frequent topic in the informants' native language - were nearly always (92.8%) in Finnish. The technology pages, in turn, were equally often in Finnish (49.7%) and English (50.3%). More surprisingly, some four fifths (81.9%) of the pages on arts were in English. As far as linguistics (#8) is concerned, other languages held dominance with their share of over half (53.9 %).

It is not likely that the correlational pattern between page subject and language would reflect an objective state of affairs in which certain topics were better covered in some tongues than others. Instead, I believe that the observed relationship is in large part explained by contextual factors like the participants' search topic or language proficiency. Some themes could even more naturally tally with particular languages. For example, most of the pages in other languages were visited for the purpose of learning those tongues.

Conclusions

Summary

This paper has scrutinized information searching on the World Wide Web in terms of what sort of pages were visited by self-developers, how they moved from one page to another, as well as whether there was a dependency between these two phenomena. The research questions were successfully answered, albeit the analyses could not dig very deep. The most central findings were:

- the most typical subject area of the viewed Web pages was technology and economy
- about half of the pages were in the searchers' native language (Finnish);
- Web form added a new category to typologies of content, but the majority of the pages were meta-informational;
- a wide variety of different tactics for moving around the Web was discovered, although only clicking on links and pushing the Back button stood out from the rest;
- following (automatic transition to another page) was identified as a novel strategy, but pointing (selecting an option) was used by far most;
- Web navigation occurred by exploiting all layers of the user interface, page being the most popular;
- on average, every page was revisited once, but there were also some similar, modified versions of those pages; and
- statistically significant correlations between the page and movement variables were common, but not particularly strong; the most intense covariation was discovered between the subject and language of Web pages.

As explained in the method section, caution must be exercised when generalizing the quantitative results obtained in this study. Its true contributions are more of the qualitative kind, in developing theoretical constructs for making sense of Web searching.

Implications

The fact that all of the ten *subject* categories in a general library classification system used to classify pages got hits is a demonstration of how heterogeneous the searches were - at least in semantic terms - and how well the classification scheme works with the sphere of personal development. It could be desirable, though, to have a set of fewer themes form the categorization, reflecting the particular domain, but no way was found to do this with self-development. On the whole, it would seem that Web page subject is a central variable, with connections to other facets of content, and also navigation. The absence of library classifications in Web search studies is a major

deficiency, since such results could be more directly applied to improving library services.

Then again, the present inquiry has something to say about potential shortcomings of library classification schemes, which were developed prior to the Internet. Although they may adequately reflect the range of information content in the Web, it could be practical to move the Internet from the category of technology to the general category, which usually includes books, publishing, libraries, mass communication, etc. It would be more logical to have all information tools under one heading.

Overlooking *language* has been another failure in research on Web searching. Given the fundamental nature of language for the human being, there is little doubt about its effect on information seeking. Assuming that English is **the** language of transactions is simply a wrong idea in a world with hundreds of languages. By comparing Web search activities across various languages, we would be in a much better position to pinpoint the role of language in information seeking.

Describing the multifarious Web search *tactics* can be illuminating. Apart from a few exceptions, scientific knowledge about their majority remains rather shallow, in some cases even non-existent. There would be plenty of fertile research terrain to be covered here. However, statistical analyses of covariation would require either focusing on just a few tactics, or collecting much—at least twenty times—more data (some 40,000 movements) than in the investigation at hand.

Search tactics may make better sense when set against the larger *strategic* background (cf. Choo *et al.*, [1999](#)). The ease of juxtaposing tactics and strategies depends on how they are defined and operationalized. The current study sees the relationship between those two concepts as a matter of straightforward generalization. Such a "translation" is quite problematic in the traditional tendency to regard tactics and strategies as referring to fairly different sides of information seeking. There, they probably have a correlational instead of logical interconnection. This distinction just goes to show that there is more than one way to define information search strategy.

In the present study—like numerous others—the observed frequency of many categories was low. The rarity of those events does not automatically mean, however, that they would be unimportant. For example, this inquiry supports prior findings on the prevalence of the browsing or 'pointing' strategy, as opposed to searching or 'typing'. Such results have been highly consistent over the years, yet it seems that most of the research conducted on information seeking in the Web has been about querying in search engines. In spite of what was just said at the beginning of this paragraph, the situation in the field does raise questions: how can such an imbalance possibly be justified? Should there not be a pronounced move towards analysing the strategy of pointing or browsing?

The *revisitation* typology proposed in this article resolves "the issue of repeated refreshing of the 'same' Web page that presents 'different' content each time, as to whether this is the 'same' Web page or 'different' Web pages". The rapidly changing Internet environment has led into a situation in which going to Web pages gives the searcher mere snapshots of information-in-progress. (Brooks [2003](#).) Moreover, some international Websites are nowadays "mirrored", so that the sites are duplicated on different servers on different continents. This means that there are multiple copies of the same Web pages in different locations. The anomalies of mirroring and dynamic content lead to a breakdown in operationalizing a Web page: a page can no longer be reliably identified by its URL, but instead by its content.

Future directions

The work started in this study can be carried on along several avenues. In addition to the suggestions above, there are many fruitful, related *questions* that are worth exploring. To mention but a few:

- Selecting: Why are particular Web pages and ways of movement chosen?
- Sites: To which Web site do the visited pages belong? Are the documents just Web pages, or the home pages of a Web site?
- Succession: What is the relationship between two consecutive pages? Is there a connection between their content and navigation?
- Trails: What kinds of "paths" do people take through Web space, page by page? How do they travel on these paths?
- Context: How do context and information searching in the Web interact? For instance, how is the adopted strategy influenced by the situation of the person? Or, what are the roles of the subject domain in Web

information seeking?

A few *methodological* issues need to be considered, as well. To begin with, the research reported here treats Web searching as an objective phenomenon observable by anyone, in principle. It is quite possible, however, to approach the activity as a subjective phenomenon (i.e., as experienced by the informants), or a combination of objective and subjective phenomena. One must be aware that this choice has an effect on the results. For the sake of illustration, let us take an example. From the *objective* point of view, a Web page is best seen as a discrete, electronic document that is located at a unique Internet address (cf. O'Neill 1998), but identified by its content. It includes everything that can be displayed by scrolling in a single browser window at a particular moment. From the *subjective* angle, on the other hand, a Web page can be something else. For instance, when a person clicks on an internal link (which takes him/her to another place on the same page), s/he may surmise that s/he opened a new page. In cases like this, URL has little relevance.

The findings of a study are dependent on the *setting* of data collection, too. Every environment has its strengths and weaknesses. Our provisional "laboratory" was certainly ideal for gathering material on Web navigation, but it undermined contextuality and validity, which are important considerations. These factors were taken into account when interpreting the results above. Then again, collecting data somewhere else would have enhanced naturalism, but would have detracted from the accuracy of Web data. In this project, selecting the setting was determined by the unwieldy video equipment, which was judged the only viable alternative when the investigation was embarked on. The ideal tool of data gathering in the context of Web searching would be a small, portable apparatus that could be easily connected to any Internet device, capture the screen as a digital video, and simultaneously record audio. An instrument of this kind would greatly boost validity, and facilitate longitudinal research.

Acknowledgements

The author wishes to acknowledge the perspicacious contributions of Reijo Savolainen, and the suggestions of the anonymous referees. He is also grateful to the Information Society Institute at the University of Tampere for funding this research.

Notes

1. GOMS stands for "Goals-Operators-Methods-Selection Rules" (Marchionini 1995: 74).
2. The number of observations fluctuates here, because some participants did not answer all of the questions.
3. Describing the use of the Reset button as a navigational strategy may seem unusual, but, if, for example, the computer 'hangs' in the middle of a search and no other action can allow the searcher to move on, pushing Reset is the only possible action to take.

References

- Brooks, T.A. (2003). [Web search: how the Web has changed information retrieval](http://InformationR.net/ir/8-3/paper154.html). *Information Research*, 8(3), paper 154. Retrieved 2 February, 2004 from <http://InformationR.net/ir/8-3/paper154.html>
- Brown, B. & Sellen, A. (2001). [Exploring users' experiences of the Web](http://firstmonday.org/issues/issue6_9/brown/). *First Monday*, 6(9). Retrieved 2 February, 2004 from http://firstmonday.org/issues/issue6_9/brown/
- Catledge, L.D. & Pitkow, J.E. (1995). Characterizing browsing strategies in the World-Wide Web. *Computer Networks and ISDN Systems*, 27(6), 1065-1073.
- Choo, C.W.; Detlor, B. & Turnbull, D. (1999). [Information seeking on the web - an integrated model of browsing and searching](http://Web.firstmonday.dk/issues/issue5_2/choo). *First Monday*, 5(2). Retrieved 2 February, 2004 from http://Web.firstmonday.dk/issues/issue5_2/choo. Also *ASIS '99: Proceedings of the 62nd ASIS annual meeting*, 36, 3-16.
- Choo, C.W.; Detlor, B. & Turnbull, D. (2000). *Web work: information seeking and knowledge work on the World Wide Web*. Dordrecht: Kluwer Academic Publishers. (Information science and knowledge management 1).
- Cockburn, A. & McKenzie, B. (2001). What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies*, 54(6), 903-922.

- Compton, W.C. (2001). Toward a tripartite factor structure of mental health: subjective well-being, personal growth, and religiosity. *The Journal of Psychology*, **135**(5), 486-500.
- Cothey, V. (2002). A longitudinal study of World Wide Web users' information-searching behavior. *Journal of the American Society for Information Science and Technology*, **53**(2), 67-78.
- Deci, E.L. & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- Elifson, K.W.; Runyon, R.P. & Haber, A. (1990). *Fundamentals of social statistics*. (2nd ed.). New York: McGraw-Hill.
- Erdelez, S. (2000). Towards understanding information encountering on the web. *ASIS 2000: Proceedings of the 63rd ASIS annual meeting*, **37**, 363-371.
- Fidel, R.; Davies, R.K.; Douglass, M.H.; Holder, J.K.; Hopkins, C.J.; Kushner, E.J.; Miyagishima, B.K. & Toney, C.D. (1999). A visit to the information mall: Web searching behavior of high school students. *Journal of the American Society for Information Science*, **50**(1), 24-37.
- Finland. Ministry of Education. (2003). [WebPLC](http://ykl.kirjastot.fi/). Helsinki: Ministry of Education. Retrieved 2 February, 2004 from <http://ykl.kirjastot.fi/>
- Ford, N.; Miller, D. & Moss, N. (2002). Web search strategies and retrieval effectiveness: an empirical study. *Journal of Documentation*, **58**(1), 30-48.
- Ford, N.; Miller, D. & Moss, N. (2003). Web search strategies and approaches to studying. *Journal of the American Society for Information Science and Technology*, **54**(6), 473-489.
- Hawk, W.B. & Wang, P. (1999). Users' interaction with the World Wide Web: problems & problem-solving. *ASIS '99: Proceedings of the 62nd ASIS annual meeting*, **36**, 256-270.
- Hill, J.R. & Hannafin, M.J. (1997). Cognitive strategies and learning from the World Wide Web. *Educational Technology Research and Development*, **45**(4), 37-64.
- Hölscher, C. & Strube, G. (2000). Web search behavior of Internet experts and newbies. *Computer Networks*, **33**(1-6), 337-346.
- Iivonen, M. & White, M.D. (2001). The choice of initial web search strategies: a comparison between Finnish and American searchers. *Journal of Documentation*, **57**(4), 465-491.
- Jansen, B.J. & Pooch, U. (2001). A review of web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology*, **52**(3), 235-246.
- Kari, J. (1996). [Rajatieon hankinnan arkipäivää - osa I: kyselytutkimus rajatieon harrastajien paranormaaleihin ilmiöihin liittyvistä tiedontarpeista ja tiedonhankinnasta arkielämän tiedonhankinnan viitekehyyksessä](http://Web.rajatie.org/paradocs/tutkimuksia.html) [Seeking information on the paranormal in everyday life - part I: a survey on needs and seeking of paranormal information in the framework of everyday life information seeking]. *ParaDocs*. (Studia paranormalia 1). Retrieved 2 February, 2004 from <http://Web.rajatie.org/paradocs/tutkimuksia.html>. Master's thesis, University of Tampere, Tampere, Finland.
- Kari, J. (2001). [Information seeking and interest in the paranormal: towards a process model of information action](http://acta.uta.fi/teos.phtml?5557). Tampere: University of Tampere. (Acta Electronica Universitatis Tamperensis 118; also Acta Universitatis Tamperensis 826). Retrieved 2 February, 2004 from <http://acta.uta.fi/teos.phtml?5557>. Doctoral dissertation, University of Tampere, Tampere, Finland.
- Khan, K. & Locatis, C. (1998). Searching through cyberspace: the effects of link cues and correspondence on information retrieval from hypertext on the World Wide Web. *Journal of the American Society for Information Science*, **49**(14), 1248-1253.
- Kim, K.-S. (2001). Information seeking on the Web: effects of user and task variables. *Library & Information Science Research*, **23**(3), 233-255.
- King, L.A. (2002). Personal growth and personality development: a foreword to the special section. *Journal of Personality*, **70**(1), 1-3.
- Klobas, J.E. & Clyde, L.A. (2000). Adults learning to use the Internet: a longitudinal study of attitudes and other factors associated with intended Internet use. *Library & Information Science Research*, **22**(1), 5-34.
- Lazonder, A.W.; Biemans, H.J.A. & Wopereis, I.G.J.H. (2000). Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science*, **51**(6), 576-581.
- Light, A. (2001). The influence of context on users' responses to websites. *The New Review of Information Behaviour Research*, **2**, 135-149.
- Magnusson, D. (1995). *Individual development: a holistic integrated model*. Stockholm: Stockholm University. (Reports from the Department of Psychology 796).
- Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge: Cambridge University Press. (Cambridge series on human-computer interaction 9).

- Moukdad, H. & Large, A. (2001). Users' perceptions of the Web as revealed by transaction log analysis. *Online Information Review*, **25**(6), 349-358.
- Odlyzko, A. (2001). [Content is not king](http://Web.firstmonday.dk/issues/issue6_2/odlyzko/). *First Monday*, **6**(2). Retrieved 2 February, 2004 from http://Web.firstmonday.dk/issues/issue6_2/odlyzko/
 - O'Neill, E.T. (1998). Characteristics of web accessible information. *IFLA Journal*, **24**(2), 114-116.
 - Ozmutlu, S.; Spink, A. & Ozmutlu, H.C. (2003). Multimedia Web searching trends: 1997-2001. *Information Processing & Management*, **39**(4), 611-621.
 - Pharo, N. (1998). [Solving problems on the World Wide Web](http://InformationR.net/ir/4-2/isic/pharo.html). *Information Research*, **4**(2). Retrieved 2 February, 2004 from <http://InformationR.net/ir/4-2/isic/pharo.html>
 - Pharo, N. (1999). Web information search strategies: a model for classifying web interaction? In: T. Aparac, T. Saracevic, P. Ingwersen & P. Vakkari (Eds.), *Digital libraries: interdisciplinary concepts, challenges and opportunities. Proceedings of the third international conference on the conceptions of the library and information science* (pp. 207-218). Zagreb & Lokve: Filozofski fakultet Zagreb amp; Naklada Benja.
 - Pharo, N. (2002). [The SST method schema: a tool for analysing work task-based Web information search processes](http://acta.uta.fi/teos.phtml?6719). Doctoral dissertation, University of Tampere, Tampere, Finland. (Acta Electronica Universitatis Tamperensis 178; also Acta Universitatis Tamperensis 871). Retrieved 2 February, 2004 from <http://acta.uta.fi/teos.phtml?6719>.
 - Piaget, J. (1971). *Biology and knowledge: an essay on the relations between organic regulations and cognitive processes* (B. Walsh. Trans.). Edinburgh: Edinburgh University Press.
 - Pirolli, P. & Fu, W.-T. (2003). [SNIF-ACT: a model of information foraging on the World Wide Web](http://Web2.parc.com/istl/groups/uir/publications/). (UIR-2003-02). Retrieved 2 February, 2004 from <http://Web2.parc.com/istl/groups/uir/publications/>. Also presented at the Ninth international conference on user modeling, Johnstown, PA.
 - Puskala, A. (2002). [Tiedonhakustrategiat World Wide Webissä](http://ethesis.helsinki.fi/julkaisut/hum/psyko/pg/puskala/) [Information search strategies in the World Wide Web]. Retrieved 2 February, 2004 from <http://ethesis.helsinki.fi/julkaisut/hum/psyko/pg/puskala/>. Master's thesis, University of Helsinki, Helsinki.
 - Rieh, S.Y. (2002). Judgment of information quality and cognitive authority in the Web. *Journal of the American Society for Information Science and Technology*, **53**(2), 145-161.
 - Savolainen, R. (1998a). *Tietoverkot kansalaisten käytössä: Internet ja suomalaisen tietoyhteiskunnan arki* [Information networks in citizens' use: the Internet and everyday life in the Finnish information society]. (Finnish information studies 12). Tampere: University of Tampere, Åbo Akademi & University of Oulu.
 - Savolainen, R. (1998b). Use studies of electronic networks: a review of empirical research approaches and challenges for their development. *Journal of Documentation*, **54**(3), 332-351.
 - Scull, C.; Milewski, A. & Millen, D. (1999). Envisioning the web: user expectations about the cyber-experience. *ASIS '99: Proceedings of the 62nd ASIS annual meeting*, **36**, 17-24.
 - Smith, P.A.; Newman, I.A. & Parks, L.M. (1997). Virtual hierarchies and virtual networks: some lessons from hypermedia usability research applied to the World Wide Web. *International Journal of Human-Computer Studies*, **47**(1), 67-95.
 - Spink, A. & Ozmutlu, H.C. (2002). Characteristics of question format web queries: an exploratory study. *Information Processing and Management*, **38**(4), 453-471.
 - 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.
 - Su, L.T. (2003). A comprehensive and systematic model of user evaluation of Web search engines: II. an evaluation by undergraduates. *Journal of the American Society for Information Science and Technology*, **54**(13), 1193-1223.
 - Tewksbury, D. & Althaus, S.L. (2000). An examination of motivations for using the World Wide Web. *Communication Research Reports*, **17**(2), 127-138.
 - 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**(2), 229-251.
 - White, R.W.; Jose, J.M. & Ruthven, I. (2003). A task-oriented study on the influencing effects of query-biased summarisation in web searching. *Information Processing & Management*, **39**(5), 707-733.
 - Wolfram, D.; Spink, A.; Jansen, B.J. & Saracevic, T. (2001). Vox populi: the public searching of the Web. *Journal of the American Society for Information Science and Technology*, **52**(12), 073-1074.
 - Xie, H.I. (2003). Supporting ease-of-use and user control: desired features and structure of Web-based online IR systems. *Information Processing & Management*, **39**(6), 899-922.
 - Ylikoski, T. (2003). *Access denied: patterns of consumer Internet information search and the effects of Internet search expertise*. Helsinki: Helsinki School of Economics. (Acta Universitatis Oeconomicae Helsingiensis A-214).

Find other papers on this subject.

How to cite this paper:

Kari, J. (2004) "Web information seeking by pages: an observational study of moving and stopping" *Information Research*, 9(4) paper 183 [Available at <http://InformationR.net/ir/9-4/paper183.html>]

Check for citations, [using Google Scholar](#)

5 1 4 9
[Web Counter](#)

© the author, 2004.
Last updated: 15 March, 2004

-
- [Contents](#) |
 - [Author index](#) |
 - [Subject index](#) |
 - [Search](#) |
 - [Home](#)
-