AI & Soc (2006) 20: 221–231
DOI 10.1007/s00146-005-0017-2

ORIGINAL ARTICLE

Asako Miura · Nobuhiko Fujihara · Koji Yamashita

Retrieving information on the World Wide Web: effects of domain specific knowledge

Received: 1 December 2004 / Accepted: 2 June 2005 / Published online: 16 September 2005

Abstract In this study, we intend to examine information retrieval behaviors from a psychological point of view using a search engine on the World Wide Web (WWW). We investigated information retrieving behaviors in detail based on both the recorded data of retrievers' web browsing actions and their thinking processes by the "think aloud" method. We focused on selected keywords for retrieving and compared them between retrievers who had enough knowledge about their task and those who did not. Our goal was to learn about the literacy needed for finding required information efficiently on the WWW.

Keywords Domain specific knowledge · Information retrieval · Search engine · Think aloud

1 Information retrieval on the WWW

Information and network technologies are being applied to every aspect of our everyday life. An enormous amount of information has been accumulated in electronic media by some huge projects such as e-library and the information range available to us has been extended more than ever before. Especially after the appearance of the World Wide Web (WWW) in 1992, patterns of information distribution via the Internet changed completely. The WWW enables us to transmit privately any information we have to everyone logging on the Internet. As a result, the Internet has turned out to be a medium for not only

A. Miura (✉)
Kobe Gakuin University, 518, Arise, Ikawadani-cho,
Nishi, Kobe, 651-2180, Japan
E-mail: asarin@team1mile.com · Tel.: +81-78-9746154 · Fax: +81-78-9746154

N. Fujihara
Naruto University of Education, 748 Nakajima, Takashima,
Naruto-cho, Naruto, Tokushima, 772-8502, Japan

K. Yamashita
National Institute of Information and Communications Technology,
3-5 Hikaridai, Seika, Soraku, Kyoto, 619-0289, Japan

collecting but also transmitting information for us. In other words, the quality of information available on the Internet has drastically changed its characteristics from "classified and organized" to "unclassified and amorphous." The WWW has at once become a garbage dump of poor or incorrect information as well as a gold mine of valuable information. It can be a useful tool for knowledge acquisition and creation depending on how we use it. It is necessary for Internet users to discriminate truth from falsehood and find their required information efficiently (i.e., rapidly and exactly).

When we search for information on the WWW, we usually try to retrieve it by using search engines such as Google[1] . For rapid and exact information retrieval, we have to attain certain literacy about how to treat information that we have acquired as well as a skill or how to use these search engines. To describe such literacy concretely, information searchers on the WWW always need to pay attention to collecting sufficient information and extracting only the information necessary to them from enormous retrieval results that include results irrelevant to them. In traditional studies of information retrieval systems, there are two standard measures of performance. The first is precision that is defined as the number of relevant documents retrieved divided by the total number of documents retrieved. The second is recall that is defined as the number of relevant documents retrieved divided by the total number of relevant documents in the collection. A system attempts to maximize both recall and precision simultaneously. If we apply these standards to human behavior, precision would be to collect sufficient information and recall would be to extracting necessary information. Recent research on searching behavior on the WWW using search engines (Kim 2001; Kim and Allen 2002; Palmquist and Kim 2000; Wang et al. 2000; Wolfram and Dimitroff 1998) has shown that people's WWW search behaviors are different from traditional information retrieval patterns. For considering such literacy from a psychological perspective or developing a supporting system for information retrieval on the WWW, we have to examine retrieving behaviors thoroughly and verify how they associate with efficient/inefficient search results.

2 Our previous studies

Considering information searching behavior from a psychological point of view, it can be regarded as a perfect example of a decision-making or problem-solving process in a highly networked information society. It is particularly interesting that there appears to be some dynamic interactions between preexisted knowledge of searchers and external information of retrieved results by search engines. In such an interaction, information retrievers may find some unexpected novel knowledge as well as acquire information they needed.

In our previous studies, Fujihara and Miura (2003), and Miura and Fujihara (2001), we carried out a psychological experiment on information retrieval and investigated the retrieval process of average Internet users. We analyzed exhaustively all kinds of retrievers' actions and classified them into some categories. In these studies, graduates and undergraduates were asked to retrieve information for two tasks using a search engine. They retrieved information on

[1] URL: http://www.google.com

the WWW regularly but were not expert users. One task was well defined and had only one answer and the other was ill-defined and had several possible answers. There was no difference in task-related knowledge among the participants. Actions for information retrieval by search engine were divided into four categories: selection of keywords, operation on web pages, operation on browser, and miscellaneous (detailed definition and subcategories are shown in Table 1). Results revealed that the average of the number of actions for an ill-defined task was larger than that for a well-defined task ($P < 0.05$). On the other hand, there were no significant differences for the number of web pages browsed, the search times between tasks, the frequencies of appearance of each action category, and the number of keywords used for retrieval. These results suggest that average Internet users have only a simple strategy and apply it to any task for information retrieval on the WWW. In this way, some basic actions for information retrieval on the WWW were categorized and we were able to estimate various retrieval behaviors based on these categories. There was no difference among tasks when participants have nearly equal task-related knowledge and skill for information retrieval.

According to these studies, based on these categorized retrieval actions, we compared information retrieval behavior between those who have enough task-related knowledge and those who do not. As it can be predicted that task-related knowledge would have some effects on the process of selecting keywords for information retrieval, we focus on how an effective information retriever behaves.

3 Experiment

3.1 Participants

Twelve graduates and undergraduates of Osaka University and Naruto University of Education (Eight female and four males, mean age = 21.33 years)

Table 1 Categories and subcategories of retrieval actions

1. Selection of keywords
1-1. Input new keywords
1-2. Add keywords
1-3. Delete keywords
1-4. Substitute keywords
2. Operation of web pages
2-1. Select hyperlinks
2-2. Select item of pull-down menu
3. Operation to browser
3-1. Press forward button
3-2. Press back button
3-3. Press home button
3-4. Select URL from jump menu
3-5. Input URL directly
3-6. Select other windows
4. Miscellaneous

participated in this experiment. Six of them were psychology majors and the remaining six were education majors. As experimental tasks described later, were associated with psychology, the former were assumed to be "knowledgeable" participants and the latter were "unknowledgeable." Some attributes of each group are summarized in Table 2. Table 2 indicates that the participants in the knowledgeable group had used a PC and the Internet longer and made informational retrievals on the WWW more frequently than those in the unknowledgeable group. Though we have to consider this difference when we investigate our results, we also need to take into consideration the fact that high-frequency informational retrievals do not necessarily correspond to their level of proficiency directly. Some rewards were given for their participation.

3.2 Experimental task

The participants were asked to search for web pages to perform the following task and to provide the URLs of the pages that provided the answer. The task was associated with perceptual psychology. The psychological phenomena introduced in the task were easy to understand but the psychological technical term that represented the phenomena was not generally known.

Task: Task was related to one of the color after-effects in human visual perception. Before beginning information retrieval, the participants were asked to gaze at stimulus figures (one test stimulus and two adaptation stimuli shown in Fig. 1) for 2 min and they experienced this color after-effect. After that, they were asked to retrieve the term by which this effect was named, after the name of the researcher, by using a search engine on the WWW. The answer was "McCullough Effect."

3.3 Procedure

The participants conducted information retrievals using a notebook PC (SONY VAIO; Windows XP OS). The PC was connected to the Internet through a LAN. They were directed to vocalize all of their thoughts and passing ideas (i.e., the think aloud method) through their information retrieval. Google, regarded as the most familiar search engine in Japan, was selected as the retrieval tool in this experiment. To give them time to adjust to thinking aloud, the participants

Table 2 Attributes of Participants

	Knowledgeable	Unknowledgeable
Number of participants	6	6
Age	21.67	21.00
Years of computer use	4.67	3.33
Years of Internet use	4.67	3.33
Hours of Internet use a week	8.50	2.17
Frequency of web retrieval a week (persons)		
Almost every day	3	0
4–5 days a week	1	0
2–3 days a week	1	4
1 day a week	1	2

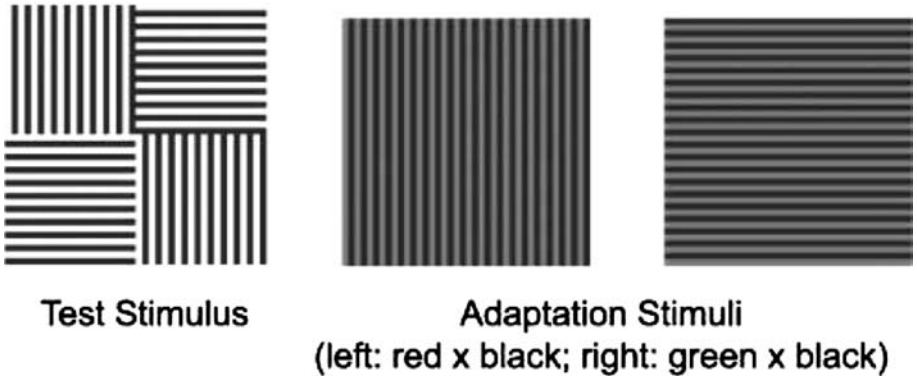

Fig. 1 Stimulus figures for Task

performed a trial information retrieval session. The trial task was ''What is the longest river in Hokkaido, Japan and how long is that?'' and the answer is ''Ishikari-river, 268 km long.''

The participants were asked to answer the question and provide the URL that contained the answer within 20 min. If they had not found the answer by then, the retrieval session was forced to terminate. The participants performed both tasks and the order of solving the tasks was counterbalanced. If some participants could not think aloud sufficiently or clearly, the experimenter called their attention to thinking aloud during the session. After the end of the task, participants were asked to answer some questions about their previous knowledge of the task and the degree of confidence in their answer. After finishing the first task, they were told the other task and asked to answer it in the same way as the first task. To estimate the participants' retrieval performance, we applied the following two recorded data. The first was a sequential motion picture of retrieval actions that was recorded with screen capturing software named ''Koushi-Daiko XP (by Lifetree Corp.).'' The second was audio data of thinking aloud that was recorded by microphone or IC recorder. After finishing the two tasks, participants were asked questions on their demographic attributes and experience regarding computer use, Internet use, and search engine use.

4 Results

According to the post-survey, it was demonstrated that no participant had known the correct answer of the task before their information retrieval in this experiment. There were six of six knowledgeable participants and two of six unknowledgeable participants who could obtain the correct answer.

4.1 Categorization of retrieval actions

Based on a preliminary examination of the screen-captured data, information retrieval behavior of each participant was described as a series of actions and

those actions were divided into four categories determined in our previous study (Miura and Fujihara 2001). The number of web pages browsed, the search times, and the number of actions taken by each participant are shown in Table 3. All the knowledgeable participants found the correct answer. Half of them did it in the shortest time and the number of their actions and keywords were also very small. These three participants seemed to obtain positive results with nearly minimum efforts. On the other hand, the unknowledgeable participants took a fairly long time and performed numerous actions for information retrieval even if they got the correct answer. They seemed to repeat trial and error many times and check around various contents of the WWW during their retrieval session (and not a few participants' activities were in vain).

The average rates at which actions of each category appeared among those with the subject attributes knowledgeable and unknowledgeable are shown in Table 4. Whereas, there was a distinct difference of quantitative retrieval performance between the groups, the average frequency for each action category was not so different between the groups. (In previous studies, we found that there was no significant difference in the frequencies of appearance of each action category). Furthermore, in this study, it was suggested that there was no difference between knowledgeable and unknowledgeable subjects in a general strategy for information retrieval on the WWW.

4.2 Input keywords

The total number of keywords input by each participant (some of them were repeatedly input) is also shown in Table 3.

Table 3 Retrieval performance of each participant

Knowledgeable group						
Participants	K1	K2	K3	K4	K5	K6
N. of browsed web pages	24	6	8	9	13	34
Times for searches (min)	19 m43 s	1 m52 s	3 m26 s	2 m18 s	12 m48 s	15 m48 s
Answer[a]	CO	CO	CO	CO	CO	CO
N. of actions	75	10	17	16	53	64
N. of keywords-related actions	18	1	3	3	15	11
N. of input keywords	17	1	3	3	13	9
N. of input keywords a kr-action	5.11	1.00	2.00	1.67	2.13	2.18
Unknowledgeable group						
Participants	U1	U2	U3	U4	U5	U6
N. of browsed web pages	47	8	17	22	91	53
Times for searches (min)[b]	TO	8 m4 s	15 m0 s	TO	TO	21 m10 s
Answer[c]	NO	CO	CO	NO	NO	FA
N. of actions	122	20	51	63	218	103
N. of keywords-related actions	20	6	19	14	24	5
N. of input keywords	14	7	16	18	15	7
N. of input keywords a kr-action	1.60	3.67	2.42	2.86	2.00	2.80

[a]The participants finished with *CO* correct answer
[b]*TO* Time over
[c]The participants finished the task with *NO* no answer, *CO* correct answer, and *FA* false answer

Table 4 Average frequency for which actions in each category appeared (%)

	Knowledgeable	Unknowledgeable
1. Selection keywords	0.22	0.15
2. Operation of web pages	0.41	0.42
3. Operation to browser	0.38	0.42
4. Miscellaneous	0.00	0.00

In the next phase, these input keywords were classified into two categories depending on their contexts and meanings. The first criterion for classification was based on the information source of each keyword. We divided the keywords into one of four categories, (1) instruction-related: source from the instruction of the experimenter (e.g., green, pinstripe, and phenomenon), (2) search-results-related: source from certain words appearing during retrieval session, (3) inference/preexisting knowledge: source in inference or pre-existed knowledge of each participant, and (4) miscellaneous. As there were some compound words which were combined multiple nouns, they were divided into single nouns and classified separately. The second criterion was based on the meaning of each keyword. We also divided the keywords into one of four categories, (1) description: directly related to the phenomenon participants had observed (e.g., red and pinstripe), (2) domain specific knowledge: academic term related to the phenomenon (e.g., visual illusion, color after-effect), (3) mistaken notion: approximate but unrelated to the phenomenon (e.g., Fechner, entropy), and (4) unrelated to the phenomenon but possible cue of retrieval (e.g., researcher, effect, and experiment). The result of the categorization of the input keywords by the two criteria is shown in Table 5.

Comparing the criterion of information sources between groups, the participants in the knowledgeable group tended to more frequently introduce their inference or preexisted knowledge into their retrieval. It was also characteristic for the unknowledgeable group that compound words were used as keywords (13 for unknowledgeable vs. 2 for knowledgeable).

Comparing the criterion of meaning between groups, the participants in the knowledgeable group tended to use domain-related keywords (description and domain specific knowledge) more frequently than those of the unknowledgeable group. Especially, the participants who found the correct answer in the shortest time mainly used domain specific keywords from the start of retrieval. On the

Table 5 Categorization for the input keywords on the two criteria (%)

	Knowledgeable	Unknowledgeable
Criteria I: Information sources		
Instruction-related	18(0.37)	39(0.41)
Search results-related	13(0.27)	26(0.28)
Inference / pre-existed knowledge	18(0.37)	29(0.31)
Miscellaneous	0(0.00)	0(0.00)
Criteria 2: Meaning		
Description	21(0.46)	29(0.38)
Domain specific knowledge	13(0.28)	17(0.22)
Mistaken notion	3(0.06)	14(0.18)
Possible cue of retrieval	9(0.19)	17(0.22)

other hand, the frequency of mistaken notion in the unknowledgeable group was relatively larger than that of the knowledgeable group.

4.3 Protocol analysis of think-aloud performance

The think aloud performance of each of the 12 participants was transcribed into protocol data. We present a broad outline of comparison of those protocol data between knowledgeable and unknowledgeable groups. In the knowledgeable group, participants tended to find and focus attention on some academic terms that represented the phenomenon they observed (e.g., after-effect and complementary color) and immediately adopted them as keywords in the process of their retrieval. They also tended to read the contents of web pages linked to search results very carefully and tried to judge whether they were associated with the task or not. In case of the unknowledgeable group, the participants tended to have difficulties specifying the certain domain of the given task and blindly focused on some proper names like researcher names and technical terms without considering their validity for the task.

5 Discussion

In this experiment, based on various indices of behavioral performance data, we compared information retrieval behavior between those who have enough task-related knowledge and those who do not. Results suggest that task-related knowledge had a great impact on the process of information retrieval in many stages. In this study, in accordance with these results, we first consider some effective strategies for information retrieval, in particulars, those based on the relationship between retrieval behavior and preexisting knowledge. Second, we intend to advance some suggestions about future systems and another new perspective for information retrieval. Finally, we refer to some limitations of our experiment.

5.1 Relationship between retrieval behavior and pre-existing knowledge

Summing up our analysis of performance data in the previous and in this experimental study, there are two kinds of knowledge we should use for accurate information retrieval, (1) task-related domain specific knowledge, (2) knowledge (or skill) relevant to search engines or web browsing.

The fact whether participants have some task-related domain specific knowledge would have a much greater impact on various stages of their retrieval behavior. At the first stage of retrieval, if they have no domain specific knowledge relevant to a certain task, they immediately face some difficulties. They might not be able to comprehend which domain the task refers to (e.g., psychology in the case of Task A) nor select sufficient keyword(s) for filtering their retrieval results. In the next stage, browsing search results, domain specific knowledge also influences their retrieval behavior. In the default setting of Google (also in this experiment), search results include some significant cues with hyperlinks for retrieved URLs, such as the title of the web page found, text that is an excerpt from the returned result page showing your query terms in

bold, and the cached link that enables retrievers to see the contents of the web page as of the time Google indexed it. If retrievers have enough task-related domain specific knowledge, they would make full use of these cues to filter the results. Lack of domain specific knowledge would lead to unproductive increment in performing their retrieval trials. Finally, domain specific knowledge has a great impact on the stage of assessing the likelihood of their answer. As the reliability of information on the WWW is quite ambiguous and uncertain, retrievers have to confirm it by themselves. Domain specific knowledge would make it easy for them to judge whether their answer is correct or incorrect.

According to the results of this study, it became clear that participants who had task-related domain specific knowledge (knowledgeable for psychology) performed more efficient (i.e., rapid and exact) information retrieval than those who did not. Furthermore, their retrieval performance tended to become more efficient when knowledgeable participants combine some technical keywords derived from their domain specific knowledge (e.g., visual illusion and accidental color) with keywords that directly related to the task (e.g., stripe and red). This would suggest that, through such combinations, they could effectively filter information from a vast array of search results.

5.2 Strategies and systems for information retrieval

These results suggest some efficient retrieval strategies in consideration of the relationship between the preexisting knowledge of the retrievers and their given task. If retrievers have sufficient domain specific knowledge about a task, they should make positive efforts to use it in their choice of retrieval keywords. That knowledge would make it possible for them to cleverly limit the scope of retrieval and extract more exact information while eliminating unnecessary details. If retrievers do not have sufficient domain specific knowledge about a task, especially when they have some difficulties in retrieving the correct answer with phenomenon-descriptive keywords, they should not expend much effort to get a definite answer, but attempt to comprehend a task-related domain for efficient retrieval. The latter knowledge that we should use for accurate information retrieval is associated with literacy about human information behavior. Compared with possible difficulties in acquiring sufficient domain specific knowledge, it would be easier for us to support information retrievers with some education or training programs and improvement systems or their interfaces.

As Fiske and Taylor (1991) indicated, however, a human being could be a ''cognitive miser'', who usually tends to conserve his energy and to behave in a manner consistent with what was appropriate to gain useful information. The idea of ''cognitive miser'' was supported by our results from both previous studies and this study. For our participants, the way of maximum saving their energy in retrieving information on the WWW was to take and repeat a single strategy, for example, alternation of ''keyword search'' and ''web browsing.'' Most of them tended to spend too little time to think about other strategies, or to avoid devising a new strategy. Even if they could not find their answer successfully by using such a single and simple strategy, they seemed to be reluctant to change the current strategy, which was to repeat ''keyword search'' and ''web browsing'' all in the same key, or to try actions other than usual. Though search engines include such functions that enable them to break a deadlock in infor-

mation retrieval (e.g., domain-specific search, minus-search, and so on), those who adhere to a single and simple strategy skimp on spending their additional time and make little attempt to direct their attention to possibly efficient alternate functions. In sum, users of the WWW in general do not use all the potential functions of search engine though they try to use few of them repeatedly.

Should we intend to develop such a support system or interface that would enable us to retrieve information efficiently, it would be important to take into account the characteristics of the cognitive and behavioral aspects of human beings. It would be important not only to upgrade and expand the support system's functions but also to devise a user-friendly interface that would make active use of such functions. It might be effective to introduce social information filtering into retrieval results. Social information filtering is a general approach to personalized information filtering. Through social information filtering, items are recommended to a user, based upon the values assigned by other people with similar taste. The system determines which users have similar taste via standard formulas for computing statistical correlations. If we apply such a function in information retrieval systems, retrievers who do not have any preexisting domain specific knowledge would be able to substitute recommended items with preexisting knowledge. Alternatively, in cooperative retrieval, multiple members performing one common retrieval task together might promote the efficiency of information retrieval. It enables retrievers to supplement their lack of knowledge by exchanging their knowledge with each other. This kind of knowledge exchange would make it possible not only to compensate for the lack of knowledge of each retriever but also to have completely novel knowledge emerge through dynamic interaction between multiple members and in human–computer interaction (HCI). We will consider this possibility in the near future.

5.3 Limitation of this experiment

In this experiment, there was a particular limitation in collecting think aloud data. Though we asked all participants to vocalize their "thoughts" and "passing ideas", it seemed they did not necessarily follow our instructions faithfully. Judging from the results of the transcription, a considerable number of participants did not vocalize their thoughts or passing ideas but did vocalize their "actions." Particularly in browsing retrieval results or linked sites they selected, participants seemed to be almost incapable of thinking aloud and there were few transcribed records. Though it could be surmised that participants in such phases thought or judged certain important considerations related to their retrieval, we were obliged to minimize our consideration of that.

There are some possible reasons for this limitation. Those were caused by the artificiality of participants performing "thinking aloud" during their retrieval session. One of the possible reasons was the brief degree of previous practice in thinking aloud and that might cause poor performance. In the next experiment, participants should be trained sufficiently prior to the experiment. Another possible reason could emerge from the shy characteristics of the Japanese. In this experiment, the participants performed their tasks independently and the experimenter monitored their performance from behind. As a result, they might have felt the weight of the eyes of the experimenter and would not have been able to think their ideas or thoughts aloud or adequately because of embar-

rassment. Corresponding to this possibility, it would be effective that we conduct an ''ice-breaking'' session with them prior to the experiment or/and an interview as the occasion demands during the experiment. Anyway, based on the results of this study, further investigation into the details of the thinking process of retrievers is required for us to clarify human behavior and cognition in information retrieval on the WWW.

References

Fiske S, Taylor S (1999) Self-efficacy and personal control. In: Fiske S, Taylor S (eds) Social cognition. McGraw Hill, New York

Fujihara N, Miura A (2003) Patterns of searching for information on the World Wide Web: a pilot study. Psychol Reports 92:1091–1096

Kim KS (2001) Information seeking on the web: Effects of user and task variables. Lib Informa Sci Res 23:233–255

Kim KS, Allen B (2002) Cognitive and task influences on web searching behavior. J Am Soc Informa Sci Technol 53:109–119

Miura A, Fujihara N (2001) Experimental study of information searching behavior on WWW: Categorization of behaviors and examination of difference between tasks. Trans Japanese Soci for Informa Syst Edu, 18:121–128 (In Japanese with English abstracts)

Palmquist RA, Kim KS (2000) Cognitive style and on-line database search experience as predictors of web search performance. J Am Soc Informa Sci 51:558–566

Wang P, Hawk WB, Tenopir C (2000) Users' interaction with World Wide Web resources: an exploratory study using a holistic approach. Informa Process Manage 36:229–251

Wolfram D, Dimitroff A (1998) Hypertext vs. boolean-based searching in a bibliographic database environment: A direct comparison of searcher performance. Informa Process Manage 34:669–679