大學圖書館
12 卷 1 期（民國 97 年 3 月） 頁 111-131

佛學數位圖書館詞彙建議介面之評估研究

Evaluation of a Term Suggestion Interface for the Digital Library of Buddhist Studies

唐牧群 **Muh-Chyun Tang***
國立臺灣大學圖書資訊學系助理教授
Assistant Professor, Department of Library and Information Science, National Taiwan University
E-mail: mctang@ntu.edu.tw

鄧雅文 **Ya-Wen Deng**
國立臺灣大學圖書資訊學系學生
Student, Department of Library and Information Science, National Taiwan University
E-mail: jazz74109@hotmail.com

鄭瑋 **Wei Jeng**
國立臺灣大學圖書資訊學系學生
Student, Department of Library and Information Science, National Taiwan University
E-mail: cerises@gmail.com

謝宜瑾 **Yi-Jin Shie**
國立臺灣大學圖書資訊學系學生
Student, Department of Library and Information Science, National Taiwan University
E-mail: sallyke41@yahoo.com.tw

【摘　要】

近年來，有愈來愈多的檢索系統增加互動性的功能，但是對於評估互動性檢索系統仍然未有一套標準的典範。本文根據一項針對佛學數位圖書館所設計的辭彙建議介面所進行的使用者研究提出報告。實驗介面將使用者檢索詞共現的關鍵詞呈現給使用者，以作為使用者修改檢索詞彙之參考。實驗結果顯示，不同的介面與不同的搜尋任務之間有互動的關係。本文也針對本研究在互動性檢索系統評估方法上的意涵進行討論。

【**Abstract**】

In recent years, more and more interactive features have been incorporated

臺灣大學學術期刊資料庫

into information retrieval systems, but an evaluation paradigm for highly interactive systems has yet to emerge. This study reports the results of an experiment conducted to test the effectiveness and usefulness of a Term Suggestion Interface for a Buddhism study collection. Terms co-occurring with users' original queries were extracted and presented to users in order to refine the search. The results suggest that the performance of the experimental interface is influenced by the number of potentially relevant records in the collection. The methodological implications for interactive information retrieval evaluation are also discussed.

關 鍵 詞：互動式資訊檢索；資訊檢索評估；數位圖書館
Keywords：Interactive information retrieval; Information retrieval evaluation; Digital library

1. Introduction

The traditional information retrieval (IR) model is largely based on the match between document representations and queries. In this model, users' information needs are assumed to be static and are fairly well represented by their queries. The assumption has been challenged by the interactive or cognitive perspective of IR (see, for example, Spink & Cole, 2005; Ingwersen, 1992) , which puts in the foreground the difficult issue of representing users' needs (Belkin, 1982, 2005). Techniques such as relevance feedback, where alternative terms are suggested to the users based on their interactions with the information space, has been proved to generate better performance (Harman, 1988, 1992). Lately there have also been attempts to utilize user behaviors such as reading, scrolling or saving as evidence for implicit feedback in order to improve performance (White et al., 2005). Another approach to term suggestion that has gained recent popularity is to simply extract from texts or metadata terms that co-occur with query terms (e.g. Hearst et al., 2002; Joho& Jose, 2006; Anick, 2003).

In this paper we present a study into the use and effectiveness of a term suggestion interface for a Buddhism study collection. Unlike traditional IR evaluation, where user involvement is minimized and task effect strictly controlled, the evaluation of interactive IR system inevitably entails complicated interactions among systems, users, and tasks. To get a full picture of the complexity, the study records and analyzes the participants' on-

臺灣大學圖書館 期刊資料庫

line search behaviors with the systems, as well as their perceptions of the interfaces. Of special interest is how the nature of the tasks might influence the performance of the experimental interface.

2. The experimental interface and the collection

To better understand users' interaction with the term suggestion device, an interface with a term suggestion feature was created for the Digital Library of Buddhist Studies (DLBS) at the National Taiwan University Library. The collection includes bibliographies and full-texts of books, journals, periodicals, research reports and theses in the area of Buddhist Studies. It currently (as of July, 2007) contains 145,299 bibliographic records, 7054 of them with links to full-text. Aside from Chinese, it includes a small portion of English and Japanese texts.

Data from three fields, namely, year of publication, publication types and keywords assigned by the authors and the indexers, are extracted from the search results to provide further suggestions for users to refine their information needs. The extracted subject keywords are simply ranked by their frequency of concurrence with the query (see Appendix I for a screenshot of the experimental interface).

3. Experiment

3.1. Participants

A total of 18 participants (11 males, 7 females) were recruited though NTU's bulletin board system. Among them two were graduate students; the rest were either NTU undergraduates or recently graduated students.

3.2. Information seeking tasks

A researcher in Buddhism studies was asked to provide six search tasks that reflect his information needs. After initial testing, two of the tasks were screened out because they were too difficult to answer using the DLBS collection, which left four tasks that were actually used in the experiment (see appendix II for the task narratives).

3.3. Design and procedures

3.3.1. Design

A within-subject design was adopted where all 18 participants were asked to carry out all four search tasks alternately on two interfaces, one with the term suggestion feature (the experimental interface) , the other without (the conventional interface). Thus a total of 72 search se-

臺灣大學學術期刊資料庫

Table 1 Graphic presentation of the experimental design, A = conventional interface, B = experimental interface

Topic / subject	1	2	3	4
1	A	B	A	B
2	A	A	B	B
3	B	A	B	A
4	B	B	A	A
5	A	B	B	A
6	B	A	A	B
⋮				
18				

ssions were carried out by the 18 participants. To avoid an order effect, the interfaces were rotated in a way that each task had an equal chance to be searched with either interface (see Table 1 for a graphic presentation of the design).

3.3.2. Research procedures

Before the experiment began, the participant was given an introduction that explained the experiment protocol and a short demonstration of the interfaces. For each task, the participants were asked to find and save five records that they believed best answered the question. They were given a maximum of 15 minutes for each task but were told that they could stop whenever they thought they had finished their search. Their search behaviors, such as inputting and changing queries, as well as browsing and saving the records, were logged by the Morae screen capture software. Instances of the participants' querying behaviors (i.e. inputting, refining, and selecting terms) were marked by a remote observer. After each task, they were asked to fill out a post-search questionnaire concerning their perceived satisfaction with the search results and search experiences (see appendix III for the texts of the questionnaires).

The bibliographic records saved by the 18 participants were then mixed and presented to the Buddhism researcher who had created the search tasks for judgment. Each record was given two scores: topical relevance and novelty, on a 0-7 scale. For topical relevance, the researcher assessed the degree to which the record

臺灣大學圖書館 期刊資料庫

was relevant to the topics specified in the search task. As for novelty, the record was judged specifically on how much new information a record contains that the expert had not been aware of.

3.3.3. Performance criteria

The two interfaces were compared on both objective and subjective performance criteria. The objective criteria were the average topical relevance and novelty scores of the records selected by the participants. As for subjective criteria, each participant was asked to indicate after each search session, how satisfied s/he was with the search results and search experience, how much s/he has learned through the search process, as well as the extent to which the search was supported by the assigned interface.

4. Results

Comparison of two interfaces

One-way ANOVA tests were performed to assess the differences between the conventional and the experimental interfaces in terms of participants' "satisfaction with the results", "satisfaction with the search experience", "learning gained in the search process", and "degree of support received".

As table 2 shows, the conventional interface generated higher scores on participants' satisfaction with the results and the experience, whereas the experimental interface did better on learning and support provided, though none of the differences were significant.

The numbers of unique records retrieved across the four tasks by the two in-

Table 2 System comparisons on subjective measures. A = conventional, B = experimental interfaces

	System	Mean	SD
Satisfaction with the search	A	3.86	1.53
	B	3.75	1.38
Experience with the search	A	3.80	1.71
	B	3.75	1.38
Learning gained in the search	A	3.66	1.68
	B	3.69	1.54
Support provided by the system	A	3.91	1.55
	B	4.11	1.61

臺灣大學學術期刊資料庫

terfaces were then calculated (see table 3). On this basis we can compare how capable each interface is of retrieving potentially relevant records, as an equivalent of the recall measure. The result shows records retrieved by the experimental interface contain less duplicates. Among the 180 records selected by participants who used the conventional interface, 108, or 60% were unique; as for the 177 records retrieved by the experimental interface, 119, or 67% were unique. Notice also that the participants selected slightly more unique records in task 3 and task 4, which seems to suggest a lesser degree of agreement in these tasks. Also notice that the experimental interface retrieved more unique relevant records in task 1 and task 2, but not so in tasks 3 and 4.

When using the experimental interface, the participants also tended to select records that were ranked lower in the returned set, which is consistent with results found in Joho & Jose (2006). As table 4 shows, the average rank position of records retrieved by the conventional and experimental interfaces were 24.59 and 30.59, respectively.

Table 3 Comparison of unique records retrieved by two systems. A = conventional, B =experimental

System / Task	A	B
Task 1	24	29
Task 2	20	23
Task 3	28	30
Task 4	36	35
Total	108/180	119/177

Task Difficulty

To assess how comparable the four tasks were with each other, a one-way analysis of variance was performed to

Table 4 Average rank position of the records retrieved by two interfaces

	Mean	SD	N
A	24.59	36.66	180
B	30.59	47.04	177
Total	27.57	42.17	357

Table 5 Average topical relevance scores across four tasks

Search task	Mean	SD
Task 1	3.98	2.35
Task 2	5.10	1.98
Task 3	2.06	1.48
Task 4	3.35	1.56

臺灣大學學術期刊資料庫

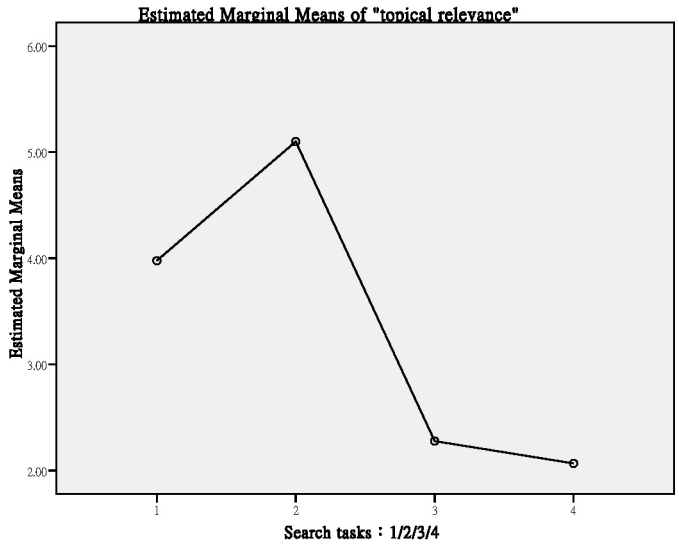


Figure 1　Graphic presentation of topical relevance among four tasks

evaluate whether the topic relevance scores differed significantly. The ANOVA was significant F (3, 353) = 52.16, p<.05, indicating that the four tasks were not a homogenous group.

According to the expert judge, the reason that topical relevance turned out to be significantly lower in task 3 and task 4 was that these were more obscure topics with less published literature in the DLBS. In other words, these tasks were more difficult in the sense that there were few records about the topics in the collection, therefore records with high relevant scores were less likely to be found. This conclusion is further supported by a comparison of the average rank position of the retrieved records among the four tasks. The ANOVA teat of the average rank position among the four tasks was significant F (3, 353) = 3.31, p <.05. The results indicate that when searching for tasks 3 and 4, the participants had to dig deeper into the returned sets in order to find the five relevant records required by the task (see table 6). However, as shown earlier, even with greater expenditure of effort, records selected by the participants in tasks 3 and 4 actually scored lower in topical relevance.

Such findings suggest that these tasks were distinguishable in terms of their a priori likelihood of finding relevant records. Previous studies have shown the

臺灣大學圖書館期刊資料庫

Table 6 Average rank position of the records retrieved across four tasks

Task \ Rank	Mean	SD	N
Task 1	19.47	34.98	88
Task 2	23.19	24.07	90
Task 3	29.68	53.93	90
Task 4	37.87	47.55	89
Total	27.57	42.17	357

nature of tasks is an important factor with respect to system performance when term suggestion devices were tested (Joho & Jose, 2006; White et al., 2005). In Joho and Jose (2006) , the interface with a term suggestion function was found to perform better in high complexity tasks in several subjective measures. In White et al. (2005) , two relevance feedback approaches, implicit and explicit, were compared. Task complexity was again found to be an intervening factor, with the implicit feedback method judged to be more "effective" and "useful" by the participants when searching for tasks with higher complexity. These findings provide reasonable grounds to investigate whether task characteristics might influence the performance of the interfaces. Therefore in the following analyses, the four tasks were classified into two groups: "easy" (task 1 and 2) and "difficult" (task 3 and 4). It should be noted that they are labeled "easy" and "difficult" not in terms of how they were perceived by the participants, but mainly by the amount of possible answers in the DLBS.

Comparison of two interfaces in different task types

Two-way ANOVAs were conducted to evaluate the effects of the interface and task difficulty on participants' perception of the interfaces. The means and standard deviations for these subjective measures as a function of the two factors are presented in Table 7. Even though neither of the main effects, interface and task difficulty, were significant, an interesting pattern emerges that seems to suggest interaction between the two factors. All four subjective measures consistently demonstrated that the conventional interface performed better where the tasks were "difficult", whereas the experimental interface did better when the tasks were "easy" (See figure 2-5 for graphic presentation of the

臺灣大學圖書館 期刊資料庫

Table 7 Participants' perception of the interfaces with different task types (N=72)

Subjective measures	Task / Sys-	Easy		Difficult		Interaction effect
		Mean	SD	Mean	SD	
Satisfaction with result	A	3.62	1.61	4.11	1.57	NS (p =.06)
	B	4.17	.99	3.33	1.65	
Satisfaction with the search experience	A	3.50	1.86	4.11	1.60	*
	B	4.22	1.06	3.28	1.56	
Learning from the search	A	3.39	1.65	3.94	1.77	NS
	B	3.89	1.45	3.50	1.69	
Support provided by the system	A	3.89	.37	3.94	.37	NS (p =.06)
	B	4.78	.37	3.44	.37	

*: p<.05, NS : Non-significant

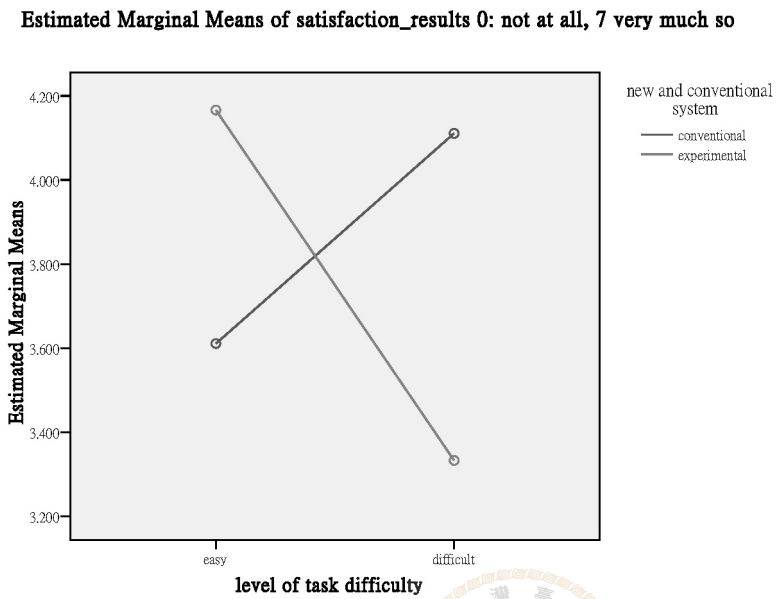


Figure 2 Graphic presentation of interaction effect on satisfaction with the results

臺灣大學圖書館 期刊資料庫

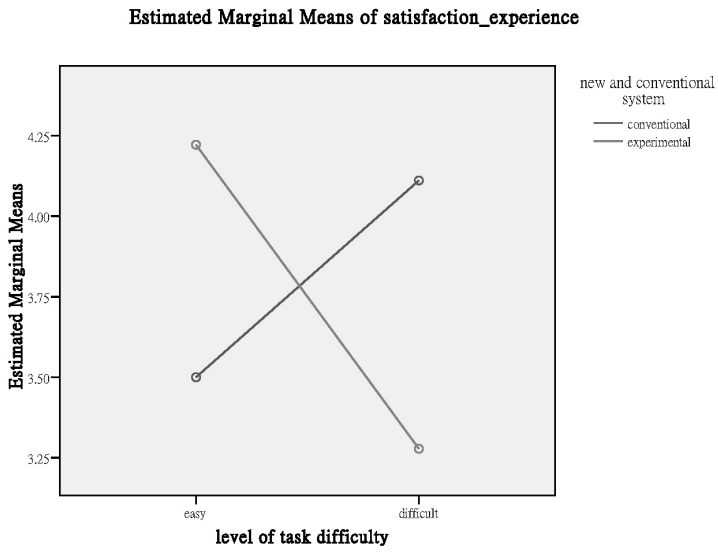


Figure 3　Graphic presentation of interaction effect on satisfaction with the experience

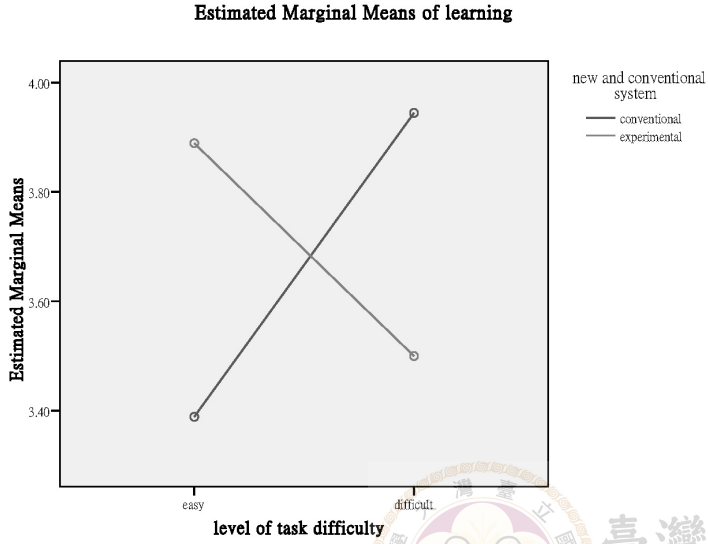


Figure 4　Graphic presentation of interaction effect on learning

臺灣大學學術期刊資料庫

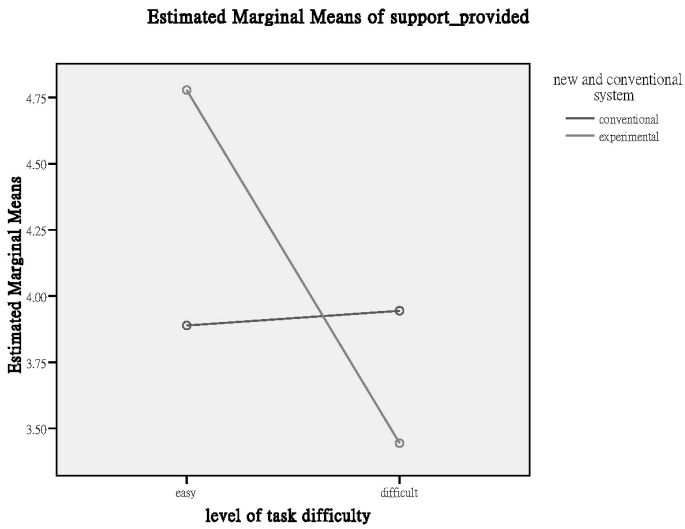


Figure 5 Graphic presentation of interaction effect on amount of support provided

Table 8 Average topical relevance and novelty scores of the records retrieved (N=360)

Objective measure	Task / Sys-	Easy		Difficult		Interaction effect
		Mean	SD.	Mean	SD.	
Topical relevance	A	4.44	2.31	2.32	1.47	NS
	B	4.65	2.17	2.01	1.57	
Novelty	A	2.33	1.53	1.89	1.44	NS
	B	2.38	1.60	1.57	1.36	

NS: none significant

interaction effect).

The same pattern can be observed, though less evident, in the two objective measures. The conventional interface performed better for the "difficult" topics, while the experimental interface did slightly better for the "easy" topics, in both topical relevance and novelty scores (see figure 5 and 6 for graphic presentations of the interaction effect).

The interaction effect between interfaces and task types is further supported

臺灣大學學術期刊資料庫

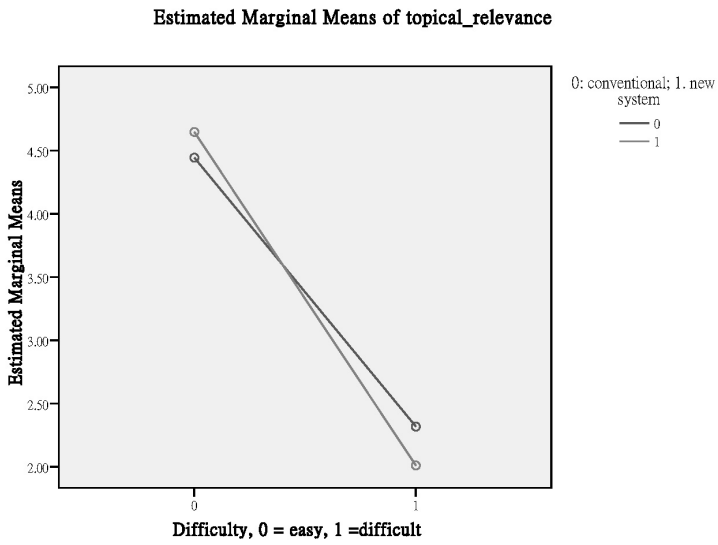


Figure 6　Graphic presentation of interaction effect on topical relevance

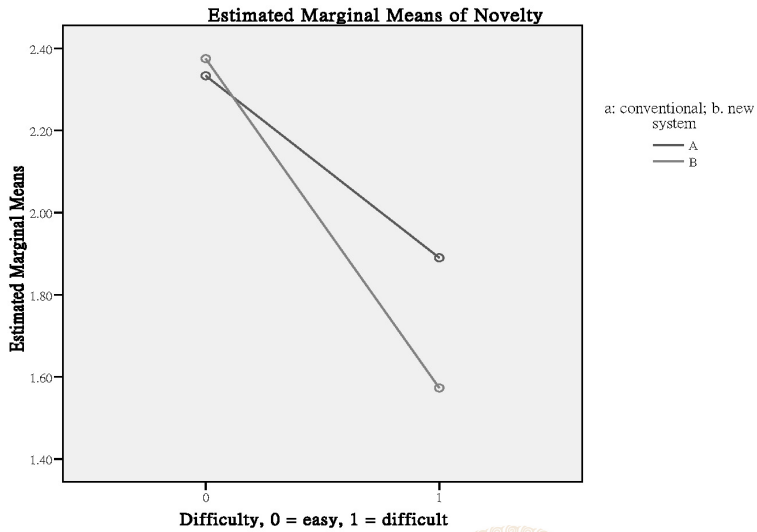


Figure 7　Graphic presentation of interaction effect on topical novelty

臺灣大學學術期刊資料庫

by analyzing the participants' experience with making use of the suggested terms in different task types. Among the 180 search sessions where the experimental interface was available, the participants opted to use the terms suggested by the interface in 125 sessions. After the sessions where the suggested terms were used, the participants were asked how useful the experimental interface was. A t-test was performed to evaluate whether the usefulness of the experimental interface differed in different task types. The result was not significant, t (123) = 1.92, p value was at the borderline of .057. Table 9 shows that the participants held a more positive opinion toward the experimental interface when searching for "easy" tasks. Also noteworthy is how the experimental interface was used slightly more frequently in "easy" tasks.

5. Discussion

From the interactive or cognitive IR perspective, query formation, that is, the representation of the users' needs, is arguably the most crucial link for successful retrievals. This line of IR research has been focusing not mainly on the ranking algorithms that have been the primary interest of classic IR research, but on how to facilitate user interaction with the systems (Qu & Furnas, 2008). Various tools and techniques, such as relevance feedback, document clustering (Pirolli et al., 1996) , faceted metadata-directed search (Hearst et al., 2002), and more recently, text mining based approaches (Swanson & Smalheiser, 1999; Swanson, 1986, 1988) , have been proposed with a view to eliciting active user participation in the IR process. Underlying these approaches is the belief that the "perfect" algorithm is not enough to answer all the users' needs, implying that for certain kinds of search tasks it is beneficial to increase user's level of interaction with the information space. Along with the development of the interaction techniques comes recognition of the variety of search activities with which users are engaged. Of particular interest is

Table 9 Comparison of the usefulness of the new interface in different task types

	Mean	SD	N
Easy (Task 1&2)	3.77	1.59	65
Difficult (Task 3&4)	3.25	1.43	60
Total	3.52	1.53	125

臺灣大學學術期刊資料庫

"exploratory search systems (ESSs) ", a recently emerging research agenda that mainly concerns itself with how to support searches that are ill-defined and open-ended, in other hand, "exploratory" by nature (White et al., 2007; Marchionini 2006). The growing interests in experimenting ESSs and other highly interactive IR systems have posed a challenge to evaluation methodology. It has been pointed out that the traditional Cranfiled-TREC evaluation paradigm is ill-equipped to assess the performance of the interaction-based systems, given the necessity of user involvements (see, for example, Kekalainen & Jarvelin; White et al., 2007). The inclusion of user and task elements into the equation induces variance that makes it difficult to separate the system effects. Even though there have been proposals (Borlund, 2000; Hersh & Over, 2000) regarding how to balance experimental control and realism in IR evaluation, a new evaluation paradigm and proper metrics have yet to emerge. In this paper we presented the design and results of an evaluation of a simple term suggestion device. The performance of the experimental interface was measured both subjectively, by the participants' opinions, and objectively, by the relevance of the records retrieved. The results of our study revealed interaction effects between types of search tasks and interfaces. The experimental interface was found to perform better, in terms of both subjective and objective metrics, where the tasks were categorized as "easy," and worse where the tasks were categorized "difficult." The level of difficulty was mainly determined not by how complicated they were or how they were perceived by the participants, but by the amount of potentially relevant records in the collection. In other words, the level of difficulty was a ctually a reflection of the abundance of relevant records associated with each search topic. We suspect that "sparsity" of relevant records was the reason why the experimental interface was used less and generated worse results in the "difficult" tasks (tasks 3 and 4). One of the advantages of using the suggested terms is to help users manage a large amount of initial results. With the suggested terms, users are able to "slice and dice" (to borrow the phrase used by Joho and Jose, 2006) a large returned set and bring up potentially relevant records that might have otherwise been buried too deeply in the returned set to be seen, which explains why the experimental interface retrieved more unique records then the conventional interface in "easy" tasks (see table 3). Yet, bringing up records that were initially assigned lower relevant scores by the system was actually detrimental in searches where the relevant records were more sparse because those

臺灣大學圖書館 期刊資料庫

low ranking records were likely to contain little relev ant information (see table 4).

6. Conclusion

In this paper we presented an investigation of the usefulness and effectiveness of a simple-term suggestion interface. Our overall conclusion was that the performance of the experimental interface was influenced by the amount of potentially relevant records in the collection. When the relevant records are abundant, the suggested terms help unearth more unique relevant records. On the other hand, when searching for topics that have fewer answers in the first place, the experimental interface was used less and produced worse results. One of the limitations of the study was that many of the comparisons were not statistically significant due to the limited number of cases. Yet it was felt that the fact that all the differentials pointed in the same direction entitles us to a certain degree of confidence in affirming the existence of the interaction effects. Our experience with the study also points to the need to collect rich user data when it comes to assessing the performance of interactive systems. Participan ts' perceptions and learning increased markedly through the searching process, (Qu and Furnas, 2007; Pirolli, 2004), and have therefore become important dimensions to consider when investigating the interactive IR process that entails the complicated interactions between users, tasks, and systems. (Received: 2007/12/28; Accepted: 2008/2/25; Revised: 2008/3/10)

致謝：本文作者希望在此向台灣大學佛學數位圖書館暨博物館與陳平坤先生，在研究過程中所提供的協助致上謝意。

References

Anick, P. (2003). Using terminological feedback for web search refinement- A log-based study. In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, (pp. 88-95).

Belkin, N. J. (2005). Anomalous State of Knowledge. In K. E. Fisher, S. Erdelez, & E. F. McKechnie (Eds.) , Theories of information behavior: A researchers' guide. (pp. 44-48) Medford, NJ: Information Today.

Belkin, N. J., Oddy, R., & Brooks, H. (1982). ASK for Information Retrieval. Journal of Documentation, 38, 61-71 (part 1) & 145-164 (part 2).

臺灣大學學術期刊資料庫

Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. Journal of Documentation, 56 (1) , 71-90.

Harman, D. (1988). Towards interactive query expansion. In Proceedings of the 11th International Conference on Research and Development in information retrieval, (pp. 32-331). Trance: Grenoble.

Harman,D. (1992). Relevance feedback revisited. In Proceedings of the 15th annual ACM SIGIR International Conference on Research and Development in information retrieval, (pp. 1-10). Demark: Copenhagen.

Hearst, M., Elliott, A., English, J., Sinha, R., Swearingen, K., & Lee, K.-P. (2002). Finding the flow in web site search. Communications of the ACM, 45 (9) , 42-49.

Hearst, M. (2006).Clustering versus faceted categories for information exploration. Communications of the ACM, 49 (4) , 59-61.

Hersh, W., & Over, P. (2000). SIGIR workshop on Interactive Retrieval at TREC and beyond. ACM SIGIR Forum, 34 (1) , 24-27.

Ingwersen, P. (1992). Information retrieval interaction. London: Taylor.

Joho, H., & Jose, J. M. (2006). Slicing and dicing the information space using local contexts. In 1st International Conference on Information Interaction in Context, 18-20 October 2006 (pp. 66-74). Denmark: Copenhagen.

Pirolli, P., Schank, P., Hearst, M., & Diehl, C. (1996). Scatter/gather browsing communicates the topic structure of a very large text collection. In Proceedings of the SIGCHI conference on Human factors in computing systems: common ground, (pp. 213-220).

Pirolli, P. (2004). The InfoCLASS model: Conceptual richness and inter-person conceptual coherence about information collections. Journal of the Japanese Cognitive Science Society, 11 (3) , 197-213.

Qu, Y., & Furnas, G. W. (2008). Model-driven formative evaluation of exploratory search: a study under a sensemaking framework. Information Processing & Management, 44 (2), 534-555.

Spink, A., & Cole, C. (2005). New Directions in Cognitive Information Retrieval. Dordrecht: Springer.

Swanson, D. R. (1986). Fish Oil, Raynaud's syndrome, and undiscovered public knowledge. Perspectives in Biology and Medicine, 30, 7-18.

Swanson, D. R. (1988). Migraine and magnesium: Eleven neglected connections. Perspectives in Biology and Medicine, 31, 526-557.

Swanson, D. R. & Smalheiser, N. R. (1999). Implicit text linkages between Medline re-

臺灣大學圖書館 期刊資料庫

cords: Using Arrowsmith as an aid to scientific discovery. Library Trends, 48, 48-59.

White, R. W., Ruthven, I., & Jose, J. M. (2005). A study of factors affecting the utility of implicit relevance feedback. In Proceedings of the 28th annual ACM SIGIR conference on Research and development in information retrieval, (pp. 35-42).

White, R. W., Marchionini, G., & Muresan, G. (in press). Evaluating exploratory search systems: Introduction to special topic issue of information processing and management. Information Processing & Management. Retrieved November 5, 2007.

臺灣大學學術期刊資料庫

Appendix I



Appendix II, task narratives

一、以禪宗、三論宗、天台宗為限，請尋找探討「語言與真理之連繫」這一主題的論文資料，並選擇出您覺得最符合主題的五筆書目資料。

二、以「自然」與「因果」為關鍵詞，請找出討論此兩議題關係的佛學資料，並選擇出您覺得最符合主題的五筆書目資料。

三、請找到五筆「比較研究『真如』、『法性』或『實際』之異同」這一主題的佛學研究資料，並選擇出您覺得最符合主題的五筆書目資料。

四、以「『涅槃』或『解脫』，與『世間』的關係」為問題意識，請搜尋參考資料，並選擇出您覺得最符合主題的五筆書目資料。

臺灣博碩士論文知識加值系統 期刊資料庫

Appendix III, data collection instruments

佛學資料庫介面評估專題研究

受訪者	進行檢索的問題

每一個問題檢索前問卷

※請問您對此搜尋主題的熟悉程度為何？

（請依照熟悉的程度勾選，0為完全不熟悉，7為非常熟悉。）

0 1 2 3 4 5 6 7

□ □ □ □ □ □ □ □

每一個問題檢索後的問卷

1. 對於這次的檢索結果您是否滿意？

（請依照滿意的程度勾選，0為非常不滿意，7為非常滿意。）

0 1 2 3 4 5 6 7

□ □ □ □ □ □ □ □

2. 對於這次的檢索經驗您覺得如何？非常有挫敗感或者是個愉快的檢索經驗？

（0為非常不滿意，很有挫敗感，7為非常滿意，檢索愉快。）

0 1 2 3 4 5 6 7

□ □ □ □ □ □ □ □

3. 經過這次的檢索，您對於這個主題有更深一層的的認識嗎？

（請依照學習的程度勾選，0為沒有學習到任何東西，7為學習非常多。）

0 1 2 3 4 5 6 7

□ □ □ □ □ □ □ □

4. 您覺得這個系統在檢索問題過程中，對您的搜尋有多少幫助？

（請依照系統支援的程度勾選，0為完全沒有支援，7為提供非常多支援。）

0 1 2 3 4 5 6 7

□ □ □ □ □ □ □ □

臺灣大學學術期刊資料庫

以下問題只有在檢索完System B（新介面）才會有的問題：

5. 在做檢索過程中，請問您是否有使用系統所提供的關鍵詞？
（有使用到請接第6題，沒用到的話直接作答第8題。）
a. 有　b. 沒有

6. 在這次的搜尋過程中，您覺得系統所提供的關鍵詞提供您多少幫助？
（請依照幫助的程度勾選，0為一點幫助都沒有，7為極有幫助。）
0　1　2　3　4　5　6　7
□　□　□　□　□　□　□　□

7. 若是系統所提供的關鍵詞對您而言有助於檢索，那麼是在哪些方面幫助您能順利進行檢索？請依照您此次的檢索經驗作答。
a. 藉由系統所呈現的關鍵詞，能夠幫助我將心中模糊的概念明確化。
（請依照符合程度勾選，0為一點都不符合，7為非常符合。）
0　1　2　3　4　5　6　7
□　□　□　□　□　□　□　□

b. 系統所提供的關鍵詞，能夠提供給我更多有關於搜尋題目的資訊，使我能夠有更多的想法，嘗試不同的搜尋方式。
（請依照符合程度勾選，0為一點都不符合，7為非常符合。）
0　1　2　3　4　5　6　7
□　□　□　□　□　□　□　□

c. 藉由系統所呈現的關鍵詞，能夠顯現這個資料庫的範圍與架構
（請依照滿意的程度勾選，0為非常不滿意，7為非常滿意。）
0　1　2　3　4　5　6　7
□　□　□　□　□　□　□　□

d. 藉由系統所提供的關鍵詞，能夠幫助我們找到更精確、符合搜尋需求的結果。（請依照符合程度勾選，0為一點都不符合，7為非常符合。）
0　1　2　3　4　5　6　7
□　□　□　□　□　□　□　□

臺灣大學學術期刊資料庫

8. 如果沒有用到系統提供的關鍵詞的話，請問原因是？
   a. 關鍵詞表太難使用
   b. 系統提供的關鍵字無法幫助進一步檢索
   c. 太花時間，所以不想用
   d. 沒有必要
   e. 其他 _____

非常謝謝您的填寫！！！

臺灣大學學術期刊資料庫