

The effect of individual differences on searching the web¹

Ihadjadene Madjid

Université Paris X-Nanterre
Centre de recherche en information spécialisée
200 Avenue de la République
92001 Nanterre
ihadjade@u-paris10.fr

Chaudiron Stéphane

Université Paris X-Nanterre
Centre de recherche en information spécialisée
200 Avenue de la République
92001 Nanterre
stephane.chaudiron@u-paris10.fr

Martins Daniel

Université Paris X-Nanterre
Laboratoire de psychologie cognitive
200 Avenue de la République
92001 Nanterre
daniel.martins@u-paris.10.fr

This paper reports results from a project which sought to investigate the influence of two types of expertise – the knowledge of the search domain and the experience of the Web search engines - on the use of a Web search engine, called *Exalead*, by a panel of students. Forty six students (twenty four undergraduated students in psychology and twenty two undergraduates in other disciplines) were asked to give correct answers to eight questions about definitions of psychology concepts, without any time constraint. Results show that participants with good knowledge in the domain on the one hand and participants with high experience of the Web on the other had the best performances. Participants with low experience of the Web showed less effectiveness than the other

participants. Future research is proposed to know the best aids to users of information retrieval systems.

Introduction

This paper describes initial findings from a large research project exploring the effects of search experience and subject knowledge on using portals and search engines. The paper focuses on a single search engine, called *Exalead*, which proposed a categorization tool. The purpose of the study was twofold: investigating the effect of individual differences in searching the Web with *Exalead*, and analyzing the consequence of using or not the *Navigation Window* allowing users to refine their searches by using keywords or concepts offered by the software. We both

¹ IHADJADENE M., CHAUDIRON S. et MARTINS D., « The Effect of Individual Differences on Searching the Web », in *Proceedings of the 66th Annual Meeting of the American Society for Information Science and Technology*, Long Beach, October 19-22, 2003, p. 240-246.

adopted a quantitative and a qualitative approach to achieve this goal.

We first present the context in which the study has been conducted and, in particular some results which will be relevant for further discussion. We describe then the experiment (brief description of *Exalead*, task and the procedure adopted for the investigation before pointing out the main results. The last part of the paper is dedicated to the discussion, comparing some results with previous studies.

Context of the study

The use of Web search engines and bibliographical search systems may improve information retrieval but may also result in an overwhelming amount of matching documents. Most people using information retrieval systems (IRS) try to find information using only one or two word queries. As a consequence, the results of the querying process often generate very large volume of hits, frequently too large to be of any practical use (Silverstein, 1998). We know that the task of browsing and reading the results displayed on the monitor is not a trivial one. Various studies have shown that a user usually scans the only first 10 to 20 documents in a list (Spink, 2000). One limitation of the current search engines interfaces is the lack of a concise representation of the content of all retrieved documents. Exploring can be time consuming when the number of potentially relevant records becomes large.

The IR community has explored document clustering and automatic categorization as an alternative method of organizing the retrieved documents. In order to address the information overload problem, many alternative output displays have been proposed by the HCI searchers. Visualization interfaces such *VIBE*, *Scatter/Gather* or *Cat_at One* allow the user to browse simultaneously through a large hierarchy of categories and a set of documents (Korfhage, 1997). Rao and al (1995) describe a range of interfaces and tools that help users to manipulate the search results.

The need to categorize the documents has been also addressed by various search engines (*Northern Light* and *Vivísimo* for example). *Northern Light* uses data-mining techniques by grouping search results into “folders” that have been created by librarians. These search results are dynamically sorted according to one of the four following types: subject, type, source, and language. Subject folders use a hierarchy of over 200,000 keywords and phrases

created by librarians at *Northern Light*. The *Vivísimo2* metasearch engine automatically clusters search results into hierarchical categories. Unlike *Northern Light* folders, *Vivísimo* creates its clusters “on the fly” using words and phrases contained in the search results themselves. The *Vivísimo* interface displays the hierarchical clusters as folders which are opened and closed as the user navigates through the search results.

Few studies have been undertaken on the evaluation of the effectiveness of the automatic organization of information. Zamir and Etzioni (1999) report an empirical comparison of users Web search behavior on a ranked list presentation versus a clustered presentation (called outline view). The results showed that participants (users) were able to understand a version of the outline view and found it easy to use. When participants were timed on eight question-answering tasks, the average time for outline view was 72.4 seconds/query while for the list view it was 99.7 seconds/query. Pratt (2000) investigates whether dynamic categorization is more useful than the two existing organizational techniques, relevance ranking and clustering. *DynaCat* uses a knowledge-based approach to organize the search results. In the usefulness evaluation, Pratt demonstrates that users could find more answers in a fixed amount of time, and were more satisfied with their search experience when they used *DynaCat* than when they used either the cluster tool or the ranking tool.

A study made by Drori (2001) indicates that the display of the document title, the lines which contain the search terms, and the documents’ categories is more useful than displaying the information without including the documents’ categories. Dumais & Chen (2000, 2001) compared a category interface with a traditional ranked list interface for presenting web search results. They pointed out that users prefer the category interface and they were 50% faster in finding information that was organized into categories.

One way to develop “intelligent” interfaces is to study how users differ in their use of IRS and to design the interfaces according to these differences. Many researchers have shown the importance of cognitive styles in designing interfaces for information retrieval systems. There are some studies on information retrieval behaviour that try to indicate the reasons why there are so large individual differences in terms of time spent and number of errors made. Ford and al (1994) found significant correlation between cognitive style and CD-ROM online searching.

² <http://www.vivisimo.com>

Leader and Klein (1996) also revealed a significant interaction between search tool and cognitive style in hypermedia database search. Marchionini examined the effects of search and subjects expertise on full-text hypertext-based searching. They found that search specialists exhibited a more varied approach to searching than subject specialists. Both subject and search experts outperformed novices. Chen and Ford (1998) show that users select and use the different access facilities and perform different navigation patterns according to their cognitive style and personal characteristics.

Borgman (1989) examined correlations between more than a dozen of characteristics which contribute to individual differences in information retrieval performance. Borgman, (1989) , Chen and al (2000) provide an overview of such characteristics which have shown to be related to information retrieval behaviour. The few studies that have investigated cognitive style as a factor when using search engines found performance differences between field independents and field dependents. Palmquist and Kim (2000) examine the effects of cognitive style and online database search experience on Web search. An interesting finding of their study is that online search experience can greatly reduce the effect of field-dependence on Web search performance. Wang & Tenpir (1999) have studied the factor of cognitive, affective and physical on user interaction with World Wide Web. Moss and Hale (1999) studied in details the cognitive styles associated to linguistic factors in Internet searching.

The experiment

Brief presentation of the Exalead Web search engine

Exalead3 is a French Web search engine provided by the *Exalead* company, which provides scalable search and navigation platforms based on statistical analysis. At the moment of the experimentation, *Exalead* used information-mining techniques in addition to the classical ranking solution. The information-mining techniques allowed users to interactively refine their search results. For each search request, a dynamic table of contents was generated that summarized the content and concepts contained in the search results. This table of contents, called *Navigation Window*, was made of two parts: the keywords, which are frequent groups of words automatically selected and extracted from the search results; and the categories which classified the documents according to a human-based classification.

By using the *Navigation Window*, users were able to refocus their search according to the search result by content (clicking a keyword) or by concept (clicking a category).

The illustration below shows the *Exalead* search interface with the *Navigation Window* on the left side. On the top, the hyperlinked categories represent frequent concepts used in the retrieved documents and, below, the hyperlinked keywords represent the frequent groups of words found in the search result.

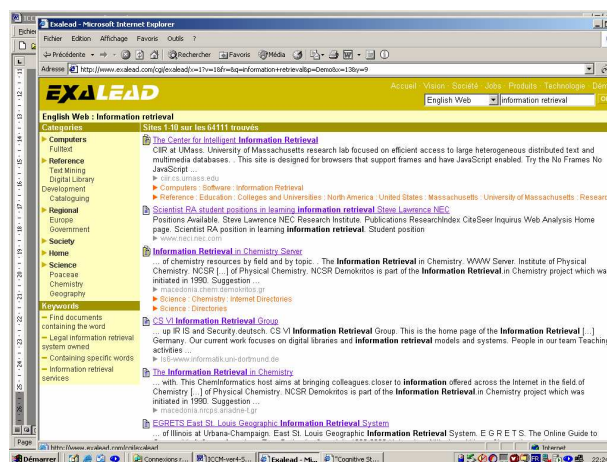


Figure 1: The *Exalead* interface

The *Navigation Window* may be expanded and collapsed, and columns can be resized and sorted. When a user drills down to another node on the taxonomy in the left part of the main window, the interface displays the set of results found within the selected category in the right part of the window. The drop-down box located at the top left allows the user to switch between various categorization schemes. The efficiency of this automatic categorization can be measured in terms of how long it takes for a user to find sufficient relevant information.

Subjects

The subjects for this study were forty-six students. Twenty-four of them were students in psychology (thereafter called experts in the domain) and the other twenty-two were students in other matters (thereafter called novices in the domain). In the Expert group in the domain, twelve were Experts in experiencing the Web search techniques and ten were Novices in the Web search

³ <http://www.exalead.fr>

techniques. We had the same distinction in the other group representing the Novices in the domain.

Table 1: The subjects of the experimentation

	Experts in Psychology	Novices in Psychology
Experts in Web search techniques	Group A: 12	Group B: 12
Novices in Web search technique	Group C: 12	Group D: 10

The Experts in the domain were students in Psychology and the experts in searching the Web were students in Library and Information Science.

Task and Procedure

The task to be performed by the panel was to give correct definitions to eight questions concerning the field of experimental psychology by searching the information on the Web. The task was to define the following notions:

- human conditioning,
- intelligence measuring tools,
- types of human memory,
- stages of cognitive development,
- various stages of intelligence,
- cognitive habit,
- influence of ageing on cognitive activities,
- description of ocular movements,
- measuring the working memory.

For the four first questions and only for them, users had the possibility to use the *Navigation Window* with the “keywords” and “categories” lists proposed by the system.

Time duration to perform the task was one hour but it was strongly recommended to perform it as soon as possible.

The task was achieved as soon as the participant considered that the information he/she found was correct. For some questions, participants gave up before finding any answer. Before the evaluation starts, subjects were asked to attend a one-hour tutorial which included how to use Web search engines, and a brief description of the main features of search engines.

A capture program was plugged in each computer that captured the action from screen and saved it in a dribble file as an AVI (Audio-Video Interleaved) movie. After having performed the searching and navigational tasks, several open-ended questions were asked as part of the user-satisfaction questionnaire. Data compiled for each user included demographic data such as gender, age, education, computer efficiency, web search engines efficiency, etc. Additional objective data were also compiled for each user. Objective data included duration of each search session compiled by means of a computerized system.

For the experimental analysis, the full combination of situations, experts in the domain (novices in psychology versus experts in psychology) and experts in Web searching techniques (novices versus experts), produced a 2 x 2 matrix. There were ten subjects in each of the four experimental groups. Five dependant variables have been considered in the experimentation:

- 1) Type of response: the judgment of relevance could be of four types: failure (the user did not perform the search), not relevant response partially-relevant response and relevant response.
- 2) Number of times the categorical information located in the left of the screen has been used.
- 3) Number of accessed Web pages.
- 4) Number of reformulations of the query.
- 5) Total time allowed to the task.

The assessors who formulated the relevance judgments to check the responses were two professors and a post-graduated student in Cognitive science.

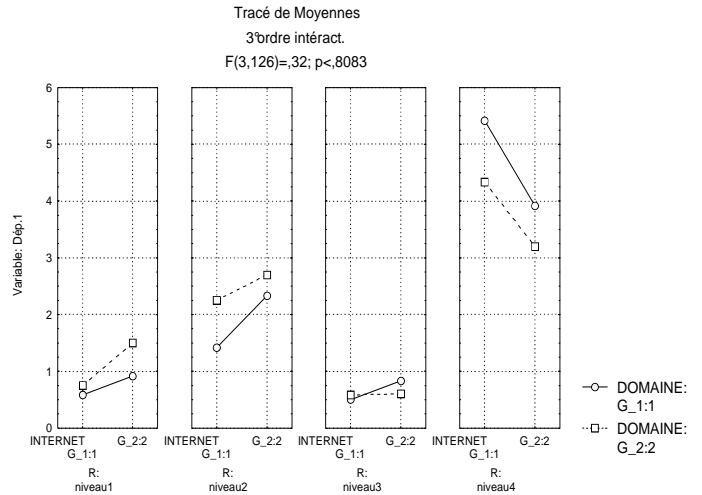
Results

Distribution of the types of responses according the level of expertise

Statistical analysis of results are made by ANalyses Of VAriance (Anova), F of Snedecor. The first result shows, not surprisingly, that experts in the domain realized a significant greater performance than the novices in the domain (5,330 versus 4,358), ($F_{1/42} = 6,365$ $p < .015$). Participants who are Web experts also obtained better results than Web novice ones (5,416 versus 4,275), ($F_{1/42} = 8,727$ $p < .005$).

An interaction is observed between expertise in the domain and response type ($F_{3/126} = 2,721 = p < .047$) in the one hand, and, in the other, another interaction is also observed between Web expertise and response type ($F_{3/126} = 5,216 p < .001$). These two interactions show that the two types of experts give correct responses more

Figure 2: Mean scores of relevant responses according to the two types of expertise



frequently than their corresponding novices. So, the two types of experts and the two types of novices are different in the number of the correct responses answered but also in the type of given responses (partially-correct and correct versus false responses and failures).

The following figure displays the means of correct responses according to the two types of expertise. In all the following figures, the four groups of people are represented in the same way.

- in the low line of the figure, G_1:1 represents the Web expert subjects and G_2:2 represents the Web novices subjects.
- dotted line in figure refers to Novices in the domain and continued line refers to Experts in the domain.

The following figure presents the distribution of the participants according to the two types of expertise (Web vs Domain) and the four types of responses from left to right:

- failure (mentioned as niveau 1),
- not relevant response (mentioned as niveau 2),
- partially-relevant response (mentioned as niveau 3),
- relevant response (mentioned as niveau 4).

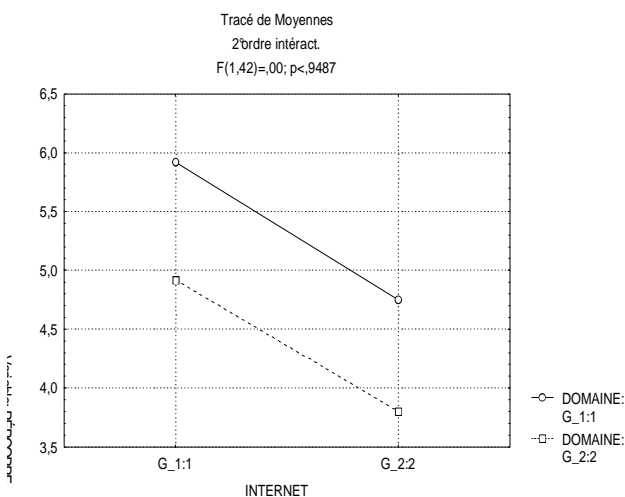


Figure 3: Performance of the panel according to the level of expertise and the types of responses

Use of the categories list

The level of expertise (both in searching the Web and in the domain) does not really influence the use of the categories, even if the experts in the domain tend to use the categories more frequently than the other participants ($F_{1/42} = 3,784 p < .058$). In addition, the results show that when the categories have been used, the relevance of the responses was not really improved.

Other results have been pointed out from the user-satisfaction questionnaire. It seems that users appreciated the semantic organization of retrieved results through the categories. Being asked to comment on this option, most of them answered that:

- it was fast, time-saving,
- displaying “keywords” and « categories » were helpful,
- they appreciated the possibility to refine query without typing another request.

Contrary to our expectations, the number of times this categorical information is used is very low, i.e. 8% (the Navigation Window has been accessed only 14 times on a maximum of 184). Most of the users indicated in the questionnaire that they had found difficult to understand the process of categorization. They judged that the labels of categories were not meaningful. Some of the users were also confused by the heading “categories” and they did not understand the difference between the “keywords” and the categories.

Average number of accessed Web pages

Anova shows that the Web experts opened more Web pages than their corresponding novices ($F_{1/42} = 18,178$ $p < .000$) and the Experts in the domain access less pages than Novices in the domain ($F_{1/42} = 4,10$ $p < .049$). Moreover, figure 2 shows that the interaction between the two types of expertise is also significant ($F_{1/42} = 28,902$ $p < .000$). This interaction suggests that participants who are experts in the domain but novices in the Web have great difficulty with the activity of clicking on web pages (mean clicks = 4.9) comparing to the other participants (mean clicks = 25).

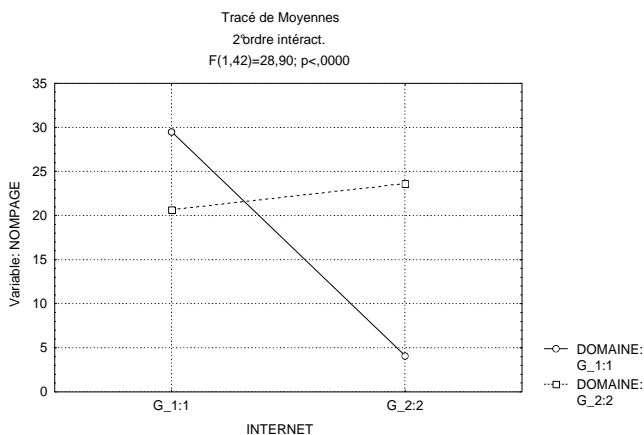


Figure 4: Means of accessed Web-pages according to the two types of expertise

An empirical indication of efficiency has been constructed in order to evaluate the effectiveness of the activity “web pages opening”. So, for each participant, the number of the web pages he/she opened has been divided by the number of his/her correct responses. Results show that Web novices (both experts and novices in the domain) are less efficient than Web experts (both experts and novices in the domain) ($F_{1/42} = 5,459$ $p < .024$).

Number of reformulations

Participants who are experts in the domain do more reformulations (modifying, dropping or canceling the query without using the *Navigation Window*) than novices in the domain (15,350 versus 12,35 ; $F_{1/42} = 4,293$ $p < .044$). The effectiveness of the reformulations has been calculated using the following formula: number of reformulations divided by the number of correct responses). It shows that Web novices (both experts and novices in the domain) are less efficient than Web experts (both experts and novices in the domain), ($F_{1/42} = 3,878$ $p < .055$).

Discussion

The two types of experts (in the domain and in the Web) have better scores than their corresponding novices, both in the number on the correct responses and in the type of responses. These results may suggest that the experts (both in the domain and in using the Web) used a kind of cognitive compensation to perform the task. It means that experts in the domain - but novices in Web - use their domain knowledge to find the correct response despite the technical difficulty of the task. In the other hand, Experts in Web – but novices in the domain - use their procedural and technical know-how to compensate their lack of knowledge in Psychology. These results agree with those of Marchionini and al. (1990) and with their interpretation. Experts have general knowledge in their long-term memory, under the form of situation model or mental model which they apply to the new situations. It is interesting to observe that these mental models (in the form of procedural knowledge or in the form of particular domain knowledge) may compensate each other.

Experts in the domain - but novices in the Web experience- open less Web pages than the other participants. However, experts in the domain do more reformulations and use more frequently the categories and keywords lists displayed on the screen. These results,

considered together, suggest that experts in the domain use a conceptual and verbal strategy (reformulating the query, using semantic categories) that facilitates the solution to the problem of information research. By contrast, if they are novices in Web, they have difficulties with technical and procedural knowledge like clicking with the mouse to open new Web pages.

The experimentation shows that Web novices are less efficient to perform the task in the allowed time, due to the number of reformulations and the number of accessed Web pages. These three results show that the lack of technical and procedural knowledge impair both the experts and novices in the domain when they have to retrieve information. Moreover, we can note that this negative impact appears not only in the novices in the domain but also in the experts in the domain. This suggests that the experts in the domain are less efficient than experts in the Web.

Unlike several studies tend to point out (Dumais,2001) (Pratt,2000),(Zamir,1999) our results seem to show that categories and keywords lists displayed on the screen are rarely used apart from the experts in the domain. One explanation of this result is that the participants, particularly when they are novices in the domain, do not consider the categories as relevant semantic cues for information access. On the other hand, we may assume that Web experts perform other strategies. A last hypothesis would be that the information categorization is not as suitable for the Web as it could be for vertical portals.

Conclusion

End users of IRS are often forced to shift through a long ordered list of documents. The IR community has explored document clustering as an alternative method of organizing retrieval result. Grouping of retrieved documents may be especially important when the user has issued a very short or vague query. In this study, we investigated, through a cognitive psychological experiment and its protocol analysis, human cognitive processes of seeking information on the WWW and the effects of subject's knowledge and experience on the information seeking processes and performance. Results show that participants with high knowledge in the domain on one hand and participants with high experience in Web on the other hand had the best performances. However participants with low experience in the Web are less efficient than the other participants. In addition, it seems that the displayed categories proposed by the *Navigation Window* in *Exalead* are not really used and appear to be not efficient.

Future research plans to investigate the role of conceptual aid for IR and the effects of cognitive compensation between the two types of expertise.

REFERENCES

- Allen, B. (1992) . Cognitive Differences in End User Searching of a CD-ROM Index, 15th Annual Inter-national SIGIR'92, ACM, Denmark.
- Borgman, C L. (1989). All Users of Information Retrieval Systems are not Created Equal: An exploitation into individual differences, *Information Processing & Management*, Vol. 25, No. 3, pp. 237-251.
- Chen, C., Czerwinski, M., Macredie, R. (2000) Individual differences in virtual environments: Introduction and overview. *Journal of the American Society for Information Science*, 51(6), 499-507.
- Chen, S. Y & Ford, N (1998) Modelling User Navigation Behaviours in a Hypermedia-Based Learning System: An Individual Differences Approach. *International Journal of Knowledge Organization* . 25(3), 67-78.
- Drori, O. (2001). Improving Display of Search Results in Information Retrieval Systems - User's Study. In P. T. Isaias (Eds), *Proceedings of the 1st International Workshop on New Developments in Digital Libraries (NDDL)*, Portugal, July 6-7, 2001 , pp. 20-33: ICEIS Press.
- Dumais Susan T., Hao Chen (2000): Hierarchical classification of Web content. *SIGIR 2000*: 256-263
- Dumais, S., Cutrell, E. & Chen, H. (2001). Optimizing Search by Showing Results In Context. In *Proceedings of the SIGCHI conference on Human factors in computing systems in Seattle, Washington, United States, 2001*,pp 277-284, . New York: ACM Press.
- Ford, N., Wood, F. and Walsh, C. (1994) Cognitive Styles and Searching. *Online & Cdrom Review*, 18(2).
- Korfhage, R. (1997). *Information Storage and Retrieval*. New York: Wiley Computer Pub.
- Lazonder, A.W., Biemans, H.J.A. & Wopereis, G.J.H. (2000). Differences between novice and experienced users in searching in the Worl Wide Web. *Journal of the American Society for Information Science*, 51 (6) 576-581.
- Leader, L. F., & Klein, J. D. (1996). The effects of search tool type and cognitive style on performance during hypermedia database searches. *Educational Technology Research and Development*, 44 (2), 5-15.
- Marchionini, G. (1989). Information-seeking strategies of novices using a full- text electronic encyclopedia. *Journal of the American Society for Information Science*, 40 (1) 54-66.
- Marchionini, G., Lin, X., Dwiggins, S. (1990). Effects of the users' knowledge on searching in a hypertext environment. In *Proceedings of the 53rd ASIS Annual Meeting (Vol. 4-8, 1990, Toronto, Canada)* (pp129-142).

- Moss, N.C. and Hale, G.G. (1999) Cognitive style and its effect on Internet search strategies: A quantitative investigation Presented at the European Educational Research Association Annual Conference, Lahti, Finland
- Nahl, D. and Tenopir, C. (1996) Affective and Cognitive Searching Behavior of Novice End-Users of a Full-Text Database. *Journal of the American Society for Information Science* 47
- Palmquist, R. A. and Kim, K. "The Effect of Cognitive Style and Online Search Experience on Web Search Performance. "Journal of the American Society of Information Science": Special Topics Issue on Individual Differences in Virtual Environments. 51(6):558-567, Apr, 2000.
- Pratt W and L. Fagan (2000). The Usefulness of Dynamically Categorizing Search Results. *Journal of the American Medical Informatics Association (JAMIA)*. 2000; 7(6) p. 605-617.
- Ramana Rao and al (1995) . Rich Interaction in the Digital Library. *CACM* 38(4): 29-39 (1995)
- Silverstein C (1998). Analysis of a very large Altavista query log. SRC technical Note 1998-014. Digital, Palo Alto
- Spink, A., Jansen, B. J., & Ozmultu, H. C. (2000). Use of query reformulation and relevance feedback by Web users. *Internet Research: Electronic Networking Applications and Policy*, 10(4): 317-328
- Wang, P. and Tenopir, C. (1999) An Exploratory Study of Users' Interaction with World Wide Web Resoures, Information Skills, Cognitive Styles, Affective States and Searching Behaviors. Proceeding of 19 th Annual National Online Meeting.
- Zamir O, Oren Etzioni (1999): Grouper: A Dynamic Clustering Interface to Web Search Results. *WWW8 / Computer Networks* 31(11-16): 1361-1374