This is a preprint version of an article published in *Computers in Human Behavior*. When referring to this publication, please cite the published version:

Pajić, D. (2014). Browse to search, visualize to explore: Who needs an alternative information retrieving model? *Computers in Human Behavior*, *39*, 145-153. doi:10.1016/j.chb.2014.07.010

Browse to search, visualize to explore: Who needs an alternative information retrieving model?

Dejan Pajić

Department of Psychology, Faculty of Philosophy, University of Novi Sad

Dr Zorana Đinđića 2, 21000 Novi Sad, Serbia

dpajic@ff.uns.ac.rs

Abstract

This article presents the results of the evaluation of SCIViS - a visualization-based scientific information retrieval (IR) system. SCIViS is based on the logic of concept maps and enables the visualization of relationships among descriptors and authors of scientific papers. It creates an interactive interface between the user's cognitive space and document information space. The system was evaluated by the group of 138 psychology students which have performed a variety of search tasks, using both the classic text-based and the visualization-based IR systems. The SCIViS model has proved to be effective and intuitive. Participants were more efficient using the visual IR system, particularly when performing tasks requiring modifications of the initial query and finding alternative keywords. User responses indicated that system's speed and ease of use are the most important attributes of the overall assessment. They also revealed the impact of users' previous experience with IR systems on users' satisfaction and perception of usefulness. Previous experience may be regarded both as the rate of success in performing search tasks, as well as the familiarity with popular search engines. The later one seems to be an important factor in modelling users' information seeking behavior and their attitudes towards alternative IR models.

Keywords: information retrieval; information visualization; visual search; bibliographic databases; usability; user satisfaction

1. Introduction

If we are to believe the leading dictionaries, it is possible *to xerox* without Xerox, but not *to google* without Google. Both trademarks have become generic terms, however the second one implies not only the action, but also the instrument. Popular web search engines have obviously influenced our perception of information retrieval (IR) as an important aspect of human behavior in the era of ubiquitous computing. In search for knowledge, we expect the IR systems to be simple and fast in recognizing our information needs. And what is even more important, we have confidence in them. Consequently, searchers are now rarely using advanced search options, they use short and simple queries, and in more than 75% of search sessions, they view only the first page of search results (Jansen & Spink, 2006; Treharne & Powers, 2009). Although this shift in users' searching behavior could be attributed to the radical improvement of information retrieval algorithms, we should ask ourselves a few important questions. First, can we really solve the problem of global information overload simply by relying on IR systems' relevance ranking algorithms? And second, in the time of touch-screen gadgets and graphical user interfaces, why are we still searching mainly by browsing textual result lists?

1.1. Information retrieval and information visualization

Information retrieval can be considered as a set of problem-solving situations that occur during the interactive communication process between the user and the information carrier system (Ingwersen, 1984). The first step in this process is the articulation of information need, usually in the form of a query. A query produces search results that should help user to overcome the inconsistency or deficiency in his or her knowledge. Hence, the IR process is more efficient if it is based on a shared knowledge, primarily the common vocabulary of the user and the system (Furnas, Landauer, Gomez, & Dumais, 1987). In this sense, information visualization (IV) techniques have been recognized as powerful tools for building representations of the IR system's information space that are more consistent with the user's cognitive space and human perception of relations among data (Newby, 2001). This cognitive space is believed to be organized in the form of semantic networks of related concepts (Collins & Loftus, 1975) which are usually visualized as graphs consisted of nodes and edges. Nodes represent different concepts, while edges represent the relationships among them (e.g. "similar to", "includes", "broader than"). The IR system's information space can be graphically represented in the same manner. Resulting maps of knowledge, known as *concept maps,* are generally acknowledged by researchers and educational practitioners as an effective tool for studying users' cognitive structures, transferring knowledge, and improving the efficiency of IR process (Gog, Kester, Nievelstein, Giesbers, & Paas, 2009; Nesbit & Adesope, 2006; Newman et al., 2010).

Information retrieval and information visualization are frequently referred to as relatively young but rapidly developing multidisciplinary research fields. They are closely related to different topics in the fields of information science, human-computer interaction, and cognitive psychology. However, the overlap among those fields is still not substantial (Boyack, Wylie, & Davidson, 2002). This has, among other things, reflected in ambiguous terminology. For example, the concept of *visual search* is usually associated with researches in experimental psychology and processes in human perception, but it is also used to refer to the specific type of IR in which the user gets a "visual support" in the form of graphical metaphors or representations of retrieved documents, web pages, keywords, and other types of information (Fagan, 2006;

Spoerri, 2004; Xu, Jin, & Lau, 2009). Different names for this type of IR model are also in use, such as *visual information retrieval* (Zhang, 2010), *visualization of information retrieval results* (Veerasamy & Belkin, 1996), *visualization-based information retrieval* (Koshman, 2006), and *visualization-based search* (Treharne & Powers, 2009). Further confusion is caused by the fact that most of those terms are often used to denote the process of retrieving multimedia information, such as images or videos. In this paper, *visual search* is used to designate the IR process which involves transformation of search results (i.e. semantic features of retrieved documents) into graphic representations. These representations should act as a form of cognitive externalization tool, expected to *amplify the cognition* of the user by increasing his or her cognitive resources, reducing search time, enhancing recognition of patterns, facilitating data manipulation and, hence, enabling the *holistic exploration* of information (Card, Mackinlay, & Shneiderman, 1999). Such designation is certainly not inconsistent with the notion of visual search as a task in which the subject needs to locate a target (e.g. relevant term or document) within the large set of distractors (e.g. polysemic terms or irrelevant documents).

1.2. Visual (database) search engines

There is a considerable agreement among researchers on the possible advantages of IV and IR integration. However, a large discrepancy between research and practice in these fields is also evident. Extensive literature is available on (allegedly) successful implementations of visualization techniques in IR process[1]. Researches have shown that visualization-based search engines improve the efficacy and accuracy of IR, particularly when common, non-specific queries are used (Kules, Wilson, & Shneiderman, 2008; Turetken & Sharda, 2005; Xu et al., 2009). Learners, novice users, and users who need to explore less familiar terminology are those who benefit the most from the visualization-based IR systems (Wu, Chuang, & Joung, 2008). Graphic representations of retrieved documents and/or keywords in the form of (concept) maps, have been shown to improve the search efficiency, especially when search tasks require comparing documents, grouping similar documents and locating related documents (Huang et al., 2006; Shen, Ogawa, Teoh, & Ma, 2006).

On the other hand, most attempts to bring visualization techniques in popular search engines have proven to be unsuccessful, unprofitable or simply ignored by users. In 2009, Google introduced *Wonder Wheel*, a topic exploration tool used to visualize relations among keywords and queries. Users were able to explore the graph of related terms along with the text-based list of results. By clicking the terms (circles) on the map, it was possible to easily expand or change initial query and automatically update the list of search results. Wonder Wheel was taken offline in 2011. Microsoft's *Bing Visual Search* was launched in 2009 and discontinued three years later. This feature allowed users to search categorized and structured results by using images instead of text links. EBSCO, one of the largest scientific information providers, has also recently ceased the visual search option for all databases. In its first version, *EBSCO Visual Search* was based on the *Grokker* engine which was used to create and visualize clusters of documents. Single documents were displayed as rectangles, while clusters of related documents were represented as circles. The color of the object indicated document's age (blue - old, red - hot). In a later version, the service was substantially changed and the visualization was largely simplified. It was based on hierarchical grouping of documents presented as a tree of descriptors and subheadings related

---

[1] For a more detailed review of visual search engines see, for example, Kules, Wilson, & Shneiderman, 2008.

to the terms in initial query. Documents were depicted as rectangles containing short previews which were arranged in blocks or columns. There are many other examples of discontinued graphical search engines, such as *KartOO*, *Ujiko*, *Vivismo*, and *Viewzi*.

1.3. Evaluation of (visual) search engines

We should consider two possible reasons for the aforementioned inconsistency between theory and practice of visualization-based IR. One is poor information exchange between psychology and information science, and the other are common evaluation procedures. Developers of information visualization tools were usually driven by the mechanical eagerness for what can be done, rather than psychological understanding of what should be done in order to meet users' needs (Chen, 2006). The usability of search engines has often been reduced to simple and harsh indicators of *precision* and *recall*, sometimes computed without any involvement of human users. Such an approach was rationalized by the assumption that user studies in the field of IR were too expensive and unreliable because of the large number of intervening variables that needed to be controlled for (Manning, Raghavan, & Schutze, 2008). This "laboratory" model of evaluation has suffered much criticism. Its major drawback is certainly disregarding both the active role of the user, and dynamic, iterative nature of user's information need (Ingwersen & Järvelin, 2005). Research shows that measures of precision and recall are appropriate for relatively small information systems, but their use is unsuitable if the collections of documents are large and heterogeneous like, for example, multidisciplinary bibliographic databases are (Belkin, Cool, Croft, & Callan, 1993; Chen, 2006).

Another important issue is that the search tasks used to evaluate visualization-based IR systems are often designed to favor certain form of visualization and measure efficiency of specific operations such as locating clusters of related documents (Sebrechts, Cugini, Laskowski, Vasilakis, & Miller, 1999), finding a period in which a document was current (Cui, Zhou, Qu, Zhang, & Skiena, 2010), identifying citation links among authors (Bergström & Whitehead Jr, 2006), detecting co-author groups and collaborations (Bergström & Atkinson, 2009) or locating related terms and concepts (Chalmers, 1993). Generally speaking, visual search engines are primarily designed for solving *spatial tasks* which implies searching by linking, extracting, allocating, ranking and clustering (Tanabe, Oyobe, Sunaoka, Yokoyama, & Miyadera, 2002; Huang et al., 2006). What all these tasks have in common is the perception of IR as an analytic process initiated by the user's desire to learn, reveal, and comprehend, rather than to simply locate a web page or a document. Hence, the difference between text-based and visualization-based information retrieval is often referred to as the difference between *search* and *exploration*, between *finding* and *discovering* (Kules et al., 2008). This is probably the reason why the few remaining visual search engines are actually sophisticated analytic tools, designed to explore large datasets and relationships among different entities, such as words (e.g. *Thinkmap Visual Thesaurus*), researchers (e.g. *Arnetminer*), patent data (e.g. *Thomson Reuters' ThemeScape*) or scientific projects (e.g. *NIH Visual Browser*).

1.4. SCIndeks Visual Search

Serbian citation index (SCIndeks) is a bibliographic database which indexes more than 400 Serbian scholarly journals in all scientific fields. Abstracts and metadata are available for over 150.000 articles of which more than 55.000 are freely available in full-text format. SCIndeks data are globally accessible through the Open Archives Initiative Protocol for Metadata Harvesting, including references cited in all indexed articles. All the above allows SCIndeks to be used in three ways: 1) as a substantial communication channel for local scientists, 2) as a mean of international promotion of Serbian science, and 3) as an essential information source for evaluation of national science (Šipka, 2005). Additionally, SCIndeks represents a valuable research platform in the fields of scientometrics, information science, natural language processing, and human-computer interaction. Several tools and services were designed and successfully implemented during the development of SCIndeks, such as the annual *Journal Bibliometric Report* which contains national impact factor for all indexed journals (Šipka, 2013), *Scriptor* – bibliographic reference parsing software (Pajić, Šipka, & Kosanović, 2002), and *MISH* - multistage indexing method for automatic extraction of subject headings (Jevremov, Pajić, Sotirović, & Šipka, 2012). One of the latest additions to the group of research projects related to SCIndeks is the visualization-based search engine called *SCIndeks Visual Search (SCIViS)*[2] which was introduced in 2012.

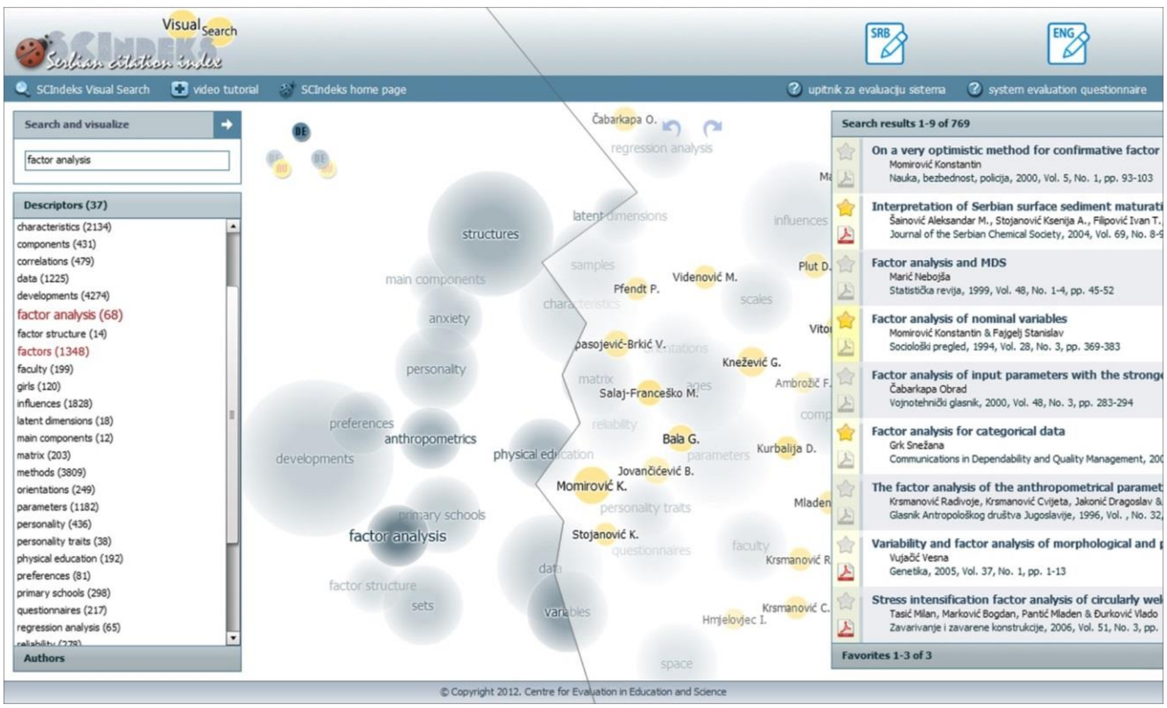

Figure 1. Combined image of SCIViS interface displaying search results for query *factor analysis* in descriptors view (left) and experts view (right)

SCIViS is theoretically based on the cognitive view of information retrieval process. User initiates the process by submitting a query, whereupon the communication with the system is shifted to the visual channel. The co-occurrence matrix of descriptors related to the query are extracted from the set of preprocessed matrices, after which it is subjected to multidimensional scaling algorithms and visualized using a force-directed layout. Although SCIViS displays search results in the form of a text-based list, the main feedback for users is a concept map, that is, a graph of keywords contained in the query along with other closely related descriptors from the database (Figure 1). Descriptors shown on the map are not author-provided keywords, but subject

[2] SCIndeks and SCIndeks Visual Search are available at http://scindeks.ceon.rs.

headings previously extracted from SCIndeks articles in accordance with several disciplinary dictionaries and thesauri. Size, saturation, and position of the graph node indicate the relative frequency of a term, its relevance for the query, and proximity to other terms. The system supports several so-called *task support facilities* (Sutcliffe & Ennis, 1998) designed to help user to modify the initial query and more accurately articulate his/her information need. The user makes modifications by interacting with the graph, that is by panning and zooming the map, or adding, removing, combining, and merging terms. Text-based list of retrieved documents is displayed in a separate window which swipes from the right edge of the screen when the user hovers mouse pointer over it. Both the list and the map are being updated each time the user interacts with the map. Besides the keywords visualization, SCIViS enables creating experts maps, i.e. maps of authors related to the presented keywords. Similarly to descriptors, author names are displayed in circles whose size, saturation, and position represent authors productivity, relevance for the subject terms, and relationships with other descriptors or authors, both in terms of co-authorships and topic similarity. The user can easily switch from the *descriptors view* (left part of Figure 1) to the *experts view* (right part of Figure 1) and extract the author's bibliography or his/her subject terms profile.

1.5. Purpose of the study

Generally speaking, this article is a contribution to the body of knowledge in the intersection of psychology, information science, and web design. More specifically, the purpose of this study is to evaluate and explore the possible purposes of an alternative visual search IR system. The intention was to evaluate the (dis)advantages of using such a system in the context of two different approaches to information retrieval as a cognitive task. One is *browsing*, or searching using text-based user interface, and the other is *exploring* the information space using graphical representations. All this has led to three specific research questions:

1. Is there a difference in search efficiency between the text-based and the visualization-based IR system?
2. What is the structure of perceived usability and which components contribute the most to users' overall assessment of IR systems?
3. How does users' previous experience with IR systems affect their overall satisfaction with IR systems?

2. Methods

2.1. Participants

The sample in this study has consisted of 138 second and third-year undergraduate students and first-year graduate students at the Department of Psychology, Faculty of Philosophy, University of Novi Sad, Serbia. It was decided to use this somewhat biased sample for three reasons. The first was that all students were familiar with the interface and functions of the national citation index since they had several research tasks during the first year of study which required searching for journal articles. The second reason was related to the problems with task conception and interpretation, since it was much easier to adjust the difficulty level and topics of search tasks knowing the contents of students' curriculum. And finally, students' availability and motivation was also a facilitating factor because it was possible to award them points for participating in the study.

2.2. Search tasks

For the evaluation purposes, eight search tasks were prepared and classified into four categories based on two criteria: topic and complexity. Four tasks were selected from the field of psychology and four from other disciplines, regarding the frequency of each particular term in the SCIndeks database so that the number of possible related articles would be approximately the same. Four tasks were designed to be simple, and four to be complex. It was expected that in the case of complex tasks users would have to focus more on exploring the results of an initial query because of the unfamiliar and ambiguous terms, and to perform additional operations in order to modify the query. An example of a simple psychological task would be:

Find article abstracts dealing with the treatment of anxiety.

This task contains a common phrase in psychology and terms that students are familiar with. On the other hand, one of the complex, non-psychological (general) tasks was:

Find article abstracts that mention the name of a discipline dealing with the evaluation of science using citation data.

Apart from the unfamiliar subject, some terms used in the task formulation are rather common and could refer to different areas of knowledge, for example the concept of *citation* in literature, philosophy, communicology or scientometrics.

2.3. Procedure

Students performed search tasks in the computer classroom of the Faculty of Philosophy in Novi Sad. They were divided into seven groups and each group have received 10 minute long training on using SCIViS, during which students have searched the database and learned basic features of the new interface. For the purposes of the study, a custom application was designed. When started, the application displays browser window with the randomly selected task and randomly selected database interface (textual or visual). Students were instructed to find and save up to five abstracts they considered relevant for the task. They were encouraged to use alternative search terms and methods if they have encountered a problem, but they were allowed to quit the task if they felt it was too difficult. The application was designed to enable students to easily save the entire abstract with a single mouse click. After completing the first task (or withdrawal from it) the procedure was repeated for a different type of task and another type of interface. Search time was not limited. After each group of students finished both tasks, prior probability of tasks to appear were modified in order to ensure the uniform distribution of task occurrences.

Search efficiency was expressed as the ratio between search effectiveness and search speed. Search effectiveness was measured both as the number of relevant articles found and as a dichotomous outcome, i.e. whether the user were able to find at least one relevant article or not. Abstract relevance was assessed after all data have been collected by two independent raters and expressed binary (relevant / irrelevant). Search speed was measured both as the time needed to complete each task and the time taken to reach the first relevant article (Downing, Moore, & Brown, 2005; Kummamuru, Lotlikar, Roy, Singal, & Krishnapuram, 2004). Other relevant human-computer interaction (HCI) variables were also recorded, such as the idle time, number of keystrokes, number of mouse clicks, and total length of the mouse cursor motion path on the screen. The application was designed to take a screenshot each time the user was idling more than 10 seconds, which made it possible to determine the probable reasons for users' inactivity.

After completion of the second task, students were automatically redirected to the short online questionnaire on preferences and attitudes towards the two IR systems. Hence, it was possible to concatenate search efficiency data and questionnaire data. The questionnaire was designed for the purposes of this study and it was based on two reference sources. The first is the recommendations given in international standards regarding the usability of visual computer interfaces (ISO, 1998). The second source was studies based on the most popular questionnaires for assessing usability of information technology products[3]. The first page of the questionnaire contained seven items regarding users' previous experience with different IR systems, primarily the frequency of use and usual search tasks they perform. The second page contained eight questions intended to measure three key components of IR systems' perceived usability (effectiveness, efficiency, and user satisfaction) and a number of specific attributes: ease of use, feedback quality, relevance of search results, and query manipulation options. Questionnaire have also included a free-text field for user comments.

2.4. Data analysis

Most usability questionnaires require users to answer the same set of questions for each system being evaluated. This method significantly increases the assessment time and has a negative impact on users' motivation and response rates (Deutskens, De Ruyter, Wetzels, & Oosterveld, 2004). Furthermore, if users are not obliged to choose between two systems, they tend to evaluate positively both of them (Nielsen & Levy, 1994). In this sense, it was decided to make the questionnaire as simple as possible. Questions were formulated to "force" students to choose between systems on each item regarding the usability, offering them tree possible options to answer which system is better: *textual, visual* or *both equally*. This approach has influenced the selection of statistical methods which had to be suitable for nominal and ordinal levels of measurement and skewed distributions. Most of the analyses were performed using nonparametric tests: Yates corrected Chi-square, Wilcoxon Matched Pairs Test, Mann-Whitney U, generalized nonlinear models, and multiple correspondence analysis. Data were analyzed using *Statistica 10.0* and *FactoMineR* package for R.

3. Results

3.1. The efficiency of IR systems

[3] For a more detailed review of the most popular usability questionnaires see, for example, Dumas & Salzman, 2006.

The first goal of this research was to examine the differences in efficiency between the text-based and the visualization-based IR system. Both the time needed to perform search tasks and the number of saved relevant abstracts were analyzed as components of search efficiency. Regarding the search speed, there was no statistically significant difference in total search duration between two IR systems ($T(129) = 3735$; $Z = 1.07$; $p > .05$)[4]. It took about five minutes on average for students to finish the task. However, individual differences were very large, with the standard deviation of almost 3 minutes. Some students gave up after only a few minutes without even completing the task, while others have succeeded after almost twenty minutes of searching. Furthermore, students who were unable to solve the search task have spent less time searching when compared to those who were successful ($U = 4565$; $Z = -4.61$; $p < .01$) which raises the question of their motivation as a potentially important determinant of search efficiency. Students were slightly faster in locating the first relevant article using the classic search interface, but this difference was not statistically significant ($T(95) = 1933$; $Z = 1.29$; $p > .05$). The only significant time difference was in idle duration ($T(129) = 3152$; $Z = 2.45$; $p = .01$). Screenshots have indicated that students had more periods of inactivity when using the visual search interface, particularly during the observation of keyword maps. However, they have managed to locate more relevant articles using SCIViS ($T(129) = 1201$; $Z = 3.77$; $p < .01$). The same effect is evident, although not statistically significant, when the effectiveness is defined binary (successful / unsuccessful), considering that more students have successfully completed tasks using the visual interface ($\chi^2(1, 129) = 2.88$; $p = .09$).

Table 1. Parameter estimates for logit model of search efficiency

	Level of Effect	Estimate	Std. Error	Wald Stat.	p
search length		**-0.003**	**0.001**	**18.465**	**0.000**
interface type	**textual**	**0.420**	**0.117**	**12.874**	**0.000**
task topic	psychology	-0.053	0.116	0.205	0.651
task complexity	**simple**	**-0.277**	**0.116**	**5.724**	**0.017**
IR type * topic	1	-0.025	0.117	0.046	0.831
IR type * complexity	1	-0.138	0.115	1.428	0.232
topic * complexity	**1**	**-0.464**	**0.118**	**15.400**	**0.000**
IR type * topic * complexity	1	0.022	0.116	0.037	0.847

Since the effect of interface type on search efficiency may vary depending on the task characteristics, the data were also analyzed using ordinal multinomial logistic regression. The number of saved relevant abstracts was used as response variable, interface type, task topic, and task complexity as categorical predictors, and total search duration as a continuous covariate. Parameter estimates are presented in *Table 1*. The model estimates the probability of a student being in the lower category, that is, being less effective in performing search tasks. In this sense, negative values of parameters indicate that students are more likely to be successful for a certain level of effect. As might be expected, students were more efficient in performing simple search tasks. On the other hand, they were equally successful when performing task from the field of psychology and non-psychological tasks. Interactive effect of task type and interface type on search efficiency was also non-significant. This is mainly because students were generally more

[4] Nine out of 138 students failed to save at least one abstract on both tasks, hence their data were excluded from all further analyses.

effective using the visual search interface, regardless of the task type. However, the advantage of using visual interface is more evident in complex tasks, which is indirectly manifested through the significant higher-order effect of task topic and task complexity. For example, students were more efficient in solving complex non-psychological tasks using SCIViS, than in solving simple psychological tasks using the classic interface. In response to the first research question, we can conclude that the visualization-based IR system is more efficient, especially for solving complex tasks, i.e. in situations which require query modifications and locating alternative search terms.

3.2. The structure of users' assessment of IR systems

The second purpose of this study was to explore the structure of users' overall assessment of IR systems and to identify main components of perceived usability. Users' responses to eight questions about different attributes of usability were analyzed using multiple correspondence analysis (MCA). The resulting map is shown in Figure 2. The position of attributes on the map indicates both their contribution to the characterization of two principal dimensions and their mutual proximity. For example, correlation of *general satisfaction* with *perceived usefulness* is very high and statistically significant ($C = 0.70$; $p < .01$), while the correlation between the ease of *query modification* and general *satisfaction* with IR systems is not significant ($C = 0.25$; $p > .05$). At the level of individual answers, this means that students were generally more satisfied with the system they have perceived as more useful. On the other hand, not a small number of users who have stated that SCIViS enables easier query modifications, were actually more satisfied with the classic search interface.

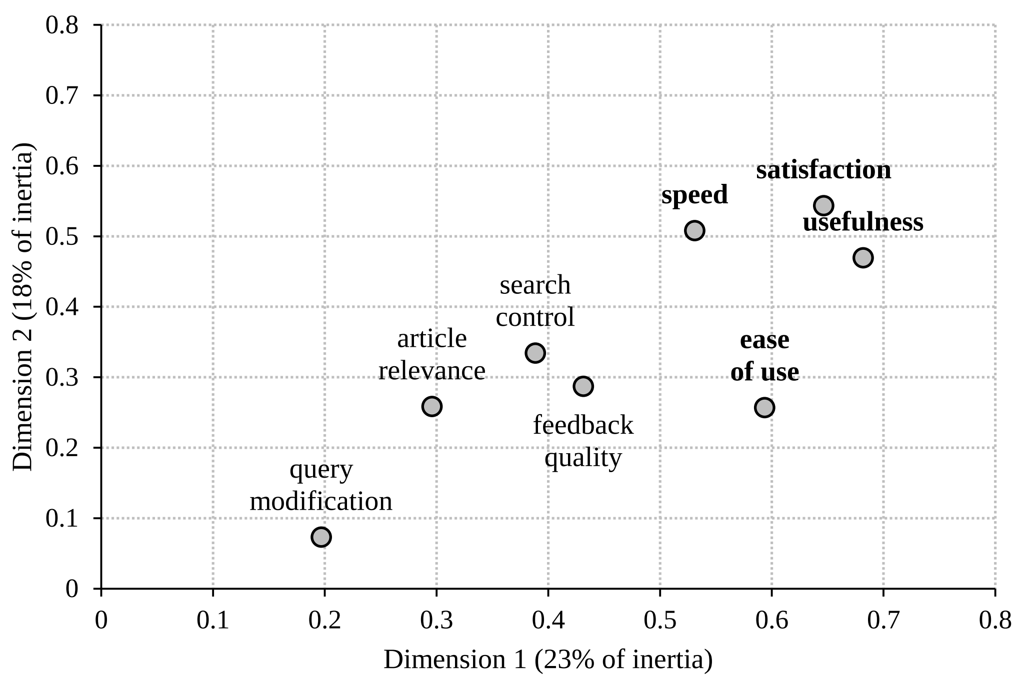

Figure 2. MCA map of eight usability attributes of IR systems (variables

Presented data basically does not support the view of perceived usability as a true multidimensional phenomenon. As we can see from Figure 2, users' *satisfaction* and *perceived* *usefulness* of IR systems seem to be the most important determinants of user preferences. Those are the two most closely related attributes and correlates highly with both latent dimensions. It seems that system's *speed* and *ease of use* are two slightly distinct aspects of usability (or usefulness) that are most closely related to users' general satisfaction. Advanced features, especially those related to query modification, are the least relevant aspects of system's usability. As we said before, they does not necessarily correlate with perceived usefulness nor users' satisfaction. Minor differences between the x and y coordinates of each specific attribute indicate that the "results management options" are more related to the ease of use, rather than system's efficiency. Those options are perceived as useful, but not as crucial for the overall user-friendliness of IR systems. In response to the second research question, we may conclude that users' satisfaction, as an emotional manifestation of perceived usefulness, is basically the principal measure of usability which very well represents users' attitude towards IR systems. Additionally, those attitudes are mostly related to perceived ease of use and speed of IR systems.

3.3. Relationships between users' satisfaction and previous experience with IR systems

The third goal of this study was to explore the relationships between users' satisfaction with the evaluated IR systems and their previous experience with those systems, as well as the other online scientific information services, like Google Scholar, ScienceDirect, and SCIndeks. The measures of users' direct experience during the search tasks were task topic (psychological / general), task complexity (simple / complex), and search performance (number of relevant articles found). All variables were analyzed separately for tasks performed using the textual system and those performed using the visualization-based system. Students' prior experience was operationalized as the summative score of responses to questions regarding the frequency of bibliographic database use. A total of seven independent variables were used to compare the group of students who reported that they were more satisfied with the visualization-based IR system to others. Data were analyzed using the Yates Chi-square and simple binary logistic regression, depending on the type of predictor variable. Results are presented in Table 2.

Table 2. The effects of different independent variables on users' satisfaction with the visualization-based IR system

Independent variable	Statistical test		
	χ^2		p
Task topic (classic interface)	0.01		0.94
Task topic (visual interface)	0.06		0.81
Task complexity (classic interface)	**4.24**		**0.04**
Task complexity (visual interface)	0.31		0.58
	Estimate	Wald St.	p
No. of relevant articles found (classic interface)	-0.22	3.57	0.06
No. of relevant articles found (visual interface)	**0.25**	**5.25**	**0.02**
Frequency of use of other IR systems	**-0.47**	**18.38**	**0.00**

It was shown earlier that task topic was not related to search efficiency. Now we can also see that there is no significant difference in satisfaction with different IR systems between students who were assigned psychological tasks and those who were assigned non-psychological tasks. This is true regardless of the interface type that was used. When it comes to the complexity of search task, there is a significant effect of the type of task performed using the textual search interface. Students who were assigned the combination of textual interface and complex task tended to evaluate visual search system more positively. It seems that students' satisfaction with the visual search system was affected by the possible difficulties they have encountered while using textual search. This was also reflected in the effect of textual search effectiveness which is close to statistical significance ($p = .06$). Namely, students who have managed to find more relevant articles using the classic interface were more satisfied with it. Likewise, students who were more effective using the visual search system have preferred this system over textual, and this difference is statistically significant ($p = .02$).

The frequency of use of online IR systems has the strongest effect on users' satisfaction. This means that students who reported to use bibliographic databases less frequently prefer the visual search interface, while those who are more familiar with textual search interfaces seem to be more critical of the alternative search model. In addition to questions about the frequency of use, students were asked about the usual search tasks they perform when using popular IR systems. The intention was to explore the possible correlation between typical search tasks and user preferences. As expected, students usually search databases by using one or two keywords. More than 70% of students use popular IR systems primarily to locate a specific article, knowing its title and/or author. On the other hand, less than 25% of them use advanced search options, such as reference search and related articles search. However, most of these variables do not affect users' overall assessment of the two evaluated IR systems. The only statistically significant difference is that regarding the use of bibliographic databases to browse, i.e. to search for a known article. Students who reported to use IR systems primarily to locate a known article prefer classic search over visual ($\chi 2 = 4.38$; $p = .04$).

3.4. HCI indicators and user comments

The analysis of basic HCI indicators suggests that different IR systems also involve different motor activities. All the differences in keyboard and mouse use between textual and visual search interfaces are statistically significant. When using the visual IR system, students have had 33% more mouse clicks ($U = 5907.0$; $Z = -4.03$; $p < .01$) and 38% less keystrokes ($U = 6005.5$; $Z = 3.86$; $p < .01$) compared to the tasks preformed using the classic search interface. Accordingly, total length of the mouse path was 42% longer in the case of visual search interface. If we consider the fact that the students were more efficient using the visual search interface and that there was no significant difference in total search length between the two systems, it seems that the SCIViS visualization model would primarily be appropriate for adapting IR system interfaces for touchscreens. This would potentially help users to complement and modify search queries more easily using swipes, pinches, flicks, and other interaction gestures.

Students have generally evaluated the SCIViS system very positively. They have preferred visual over textual search on all aspects of usability. The ease of query modification is the highest rated attribute of the visualization-based system with more than 70% students considered SCIViS to be more usable than the classic textual search interface. Students comments indicate that they prefer visualization because it makes search results easier to remember. They recognize SCIViS as a "valuable educational tool" which enables more coherent view of a particular topic. On the other hand, the lowest rated attribute was the ease of use with 60% of students preferring visual search. The comments of students who were more satisfied with the classic textual search, indicate that they preferred the simplicity, clarity, and efficacy of the classic model, which, according to them, invalidates the need for any interface changes. It is valuable to note that some students didn't evaluate the new system only from their own perspective, but also bearing in mind the wider context of its use. For example, several students with a negative attitude towards the visualization-based system have expressed concern about its usefulness for the population of "visually impaired persons" and "older and more conservative users".

4. Discussion

Textual rank lists are generally accepted and easy to use form of displaying search results. However, they are not designed to fully interact with users which makes them incompatible with the psychological comprehension of the user's active role in the search process as a problem solving situation (Ingwersen, 1984). Consequently, users are becoming more (cognitively) involved in browsing long lists, rather than truly searching (or exploring) the information space. The presented research was an attempt to point out the value of information visualization techniques in making IR systems more efficient in transferring information to users. Most visualization-based IR systems were based on 2D graphical representation of retrieved documents or clusters of documents (Sutcliffe, Ennis, & Hu, 2000; Koshman, 2006). In many cases, inadequate choice of visual metaphors have led to very complex models that required special training and even some special user skills (e.g. Shen et al., 2006; Costagliola & Fuccella, 2011). A psychologist could ask whether these systems were designed to demonstrate the technological achievements or to truly facilitate the retrieval of relevant information.

The logic of conceptual mapping have proven to be very effective and intuitive model of visual information representation (Landauer, Laham, & Derr, 2004; Nesbit & Adesope, 2006; Newman et al., 2010). The presented research has confirmed such results. The preliminary evaluation of SCIndeks Visual Search supports the view that comprehensible and simple information visualization techniques may serve as a useful interface between the information space of the IR system and cognitive space of the user. Visualization of article descriptors in the form of semantic networks was generally well accepted and positively evaluated by the students. They have quickly adapted to the new search model and had no major problems using it and learning its basic functionalities. Furthermore, they were generally more efficient in solving search tasks using the visualization-based system. This was particularly evident in situations when users were solving complex search tasks which required query modifications and finding alternative search terms. Concept maps have obviously helped users overcome the so-called *vocabulary problem*, often caused by the incompatibility between user's terminology and the vocabulary of the information carrier system.

Although the alternative IR model was positively evaluated, it would be superficial to claim that the visual search interface is better than the classic one. This is mainly because general usefulness of an IR system should not be equated to perceived usefulness. For example, our data have shown that, although visual support for query modifications have helped users solve complex search tasks, it was not perceived as a relevant attribute of the system's overall usability. Perceived usefulness and users satisfaction, as the basic components of system's usability, are most closely related to system speed and ease-of-use. This confirms the results of previous studies which have shown that user preferences towards different IR systems mostly depend on their overall satisfaction and is not necessarily associated with their greater efficiency (Rivadeneira & Bederson, 2003; Sutcliffe et al., 2000). Furthermore, since the user satisfaction is a psychological dimension, it is subject to the influence of many factors. We have shown that users' motivation, typical search strategies they use, and task complexity are just some of those factors.

Different perceptions of usefulness may partially explain the previously mentioned discrepancy between the supportive results of evaluation studies and poor acceptance of visual IR systems in practice. Besides the computational and perceived usefulness, we need to take into consideration the dimension usually referred to as *contextual* or *situational* usefulness (Ingwersen & Järvelin, 2005). Our results have shown that users' perception of usability is influenced by their direct experience with the two evaluated IR systems. Students who were more successful using visual search interface and those who have had difficulties in solving complex tasks using textual search interface, both tended to evaluate the visual search interface more positively. Other studies have also demonstrated that the evaluation of usability greatly depends on the type and features of user tasks (Mathieson & Keil, 1998; Sebrechts et al., 1999). However, it seems that students' previous long-term experience may be even more important for the assessment of IR systems. It was shown that students who claim to use IR systems more frequently evaluate the textual search engine more favorably. And they generally use web search engines and bibliographic databases for simple search tasks, usually to browse journal archives and search for a specific article from a known title. Hence, this study has basically confirmed the results of some earlier studies (White, Kules, & Bederson, 2005; Kules et al., 2008; Bergström & Atkinson, 2009; Lehmann, Schwanecke, & Dörner, 2010) which suggested that users who are more directed towards exploration search strategies, such as using citation data and combining different search fields, prefer visual search interfaces.

4.1. Limitations of the study

The main limitation of the current research comes from the relatively small and biased sample. This sampling method has enabled controlling for the possible confounding variables, but there is a strong possibility that students responded in a different manner than a "typical" database user. Taking into account the results of some earlier studies, it can be assumed that students are more enthusiastic about technological innovations and graphical representations since they are more often engaged in search tasks that imply learning new topics and solving unfamiliar problems (Conner & Browne, 2013). The future evaluations should probably be more extensive and involve a more heterogeneous sample of users. This also raises the question of whether to include additional variables in the analysis. For example, it was shown that perceived usefulness and the acceptance of new technologies can significantly depend on users' (spatial) abilities (Chen, 2006), personal innovativeness and computer anxiety (van Raaij & Schepers, 2008), as well as gender and general computer self-efficacy (Bao, Xiong, Hu, & Kibelloh, 2013). However, only the future analyses of log files and users statistics will offer a true feedback on system's usefulness and evaluate the actual need for an alternative IR model.

4.2. Conclusion and implications

The main contribution of this article is the preliminary feedback on the fully functional visual IR system. It has demonstrated (once more) how a simple and intuitive psychological concept can significantly improve the efficiency and depth of the information retrieval process. However, the implications of the study should go beyond just answering the proposed research questions. At the time when modern communication technologies and pervasive information overload are largely influencing our online search behavior, we need to think about the consequences of perceiving simple and fast IR systems as more useful. It was already shown that the growing popularity of easy-to-use and "all-inclusive" web search engines have negatively influenced academics and made their information-seeking behavior more biased (Boeker, Vach, & Motschall, 2013; Conner & Browne, 2013; Jamali & Asadi, 2010). In this context, the question to consider is how will users, particularly scholars, know if they have thoroughly explored a research problem? How to know if we are performing poorly at searching when practically any query, in any modern IR system, retrieves at least dozens of results? We seem to be fighting the information pollution problem by simply ignoring large portions of information space. We are not only *googling* for common everyday information, but also *"scholaring"* for relevant scientific results. And this carries the risk of being tendentious and perfunctory. In order to minimize this risk, visualization-based IR systems should complement the functionality of textual ones, and enable users to explore what is hidden and (falsely) categorized as irrelevant. When this happens, we might perceive textual browsing as inefficient and inappropriate as the image zooming on a keypad mobile phone is nowadays.

References

Bao, Y., Xiong, T., Hu, Z., & Kibelloh, M. (2013). Exploring Gender Differences on General and Specific Computer Self-Efficacy in Mobile Learning Adoption. *Journal of Educational Computing Research*, *49*(1), 111–132. doi:10.2190/EC.49.1.e

Belkin, N. J., Cool, C., Croft, W. B., & Callan, J. P. (1993). The effect multiple query representations on information retrieval system performance. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 339–346).

Bergström, P., & Atkinson, D. C. (2009). Augmenting the exploration of digital libraries with web-based visualizations. In *Digital Information Management, 2009. ICDIM 2009. Fourth International Conference on Data Mining* (pp. 1–7). Miami, USA.

Bergström, P., & Whitehead Jr, E. J. (2006). CircleView: scalable visualization and navigation of citation networks. In *Proceedings of the 2006 Symposium on Interactive Visual Information Collections and Activity (IVICA 2006), College Station, Texas*.

Boyack, K., Wylie, B., & Davidson, G. (2002). Information visualization, human-computer interaction, and cognitive psychology: domain visualizations. *Lecture Notes in Computer Science*, *2539*, 145–158.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann.

Chalmers, M. (1993). Using a landscape metaphor to represent a corpus of documents. In *Proceedings of the European Conference on Spatial Information Theory, Elba, September 1993* (pp. 377–390).

Chen, C. (2006). *Information visualization: Beyond the horizon*. London: Springer.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428.

Conner, M., & Browne, M. (2013). Navigating the information-scape: information visualization and student search. *Reference Services Review*, *41*(1), 91–112. doi:10.1108/00907321311300901

Costagliola, G., & Fuccella, V. (2011). CyBiS: A novel interface for searching scientific documents. In *15th International Conference on Information Visualisation* (pp. 276–281).

Cui, W., Zhou, H., Qu, H., Zhang, W., & Skiena, S. (2010). A dynamic visual interface for news stream analysis. In *Proceedings of the First international workshop on Intelligent visual interfaces for text analysis* (pp. 5–8).

Deutskens, E., De Ruyter, K., Wetzels, M., & Oosterveld, P. (2004). Response rate and response quality of internet-based surveys: An experimental study. *Marketing Letters*, *15*(1), 21–36.

Downing, R. E., Moore, J. L., & Brown, S. W. (2005). The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in Human Behavior, 21*(2), 195–209.

Dumas, J. S., & Salzman, M. C. (2006). Usability Assessment Methods. *Reviews of Human Factors and Ergonomics, 2*(1), 109–140. doi:10.1177/1557234X0600200105

Fagan, J. C. (2006). Usability testing of a large, multidisciplinary library database: Basic search and visual search. *Information Technology and Libraries*, *25*(3), 140–150.

Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The vocabulary problem in human-system communication. *Communications of the ACM*, *30*(11), 964–971.

Gog, T. van, Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior*, *25*(2), 325–331. doi:10.1016/j.chb.2008.12.021

Huang, Z., Chen, H., Guo, F., Xu, J. J., Wu, S., & Chen, W. H. (2006). Expertise visualization: An implementation and study based on cognitive fit theory. *Decision Support Systems*, *42*(3), 1539–1557.

Ingwersen, P. (1984). Psychological aspects of information retrieval. *Social Science Information Studies*, *4*(2-3), 83–95.

Ingwersen, P., & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context*. Dordrecht: Springer.

ISO. (1998). *ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs)*. Geneve: International Organization for Standardization.

Jamali, H. R., & Asadi, S. (2010). Google and the scholar: the role of Google in scientists' information-seeking behaviour. *Online Information Review, 34*(2), 282–294. doi:10.1108/14684521011036990

Jansen, B. J., & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, *42*(1), 248–263.

Jevremov, T., Pajić, D., Sotirović, M., & Šipka, P. (2012). Mapping science based on keywords of articles antecedences, presences, and consequences: An application of CEON/CEES model of multi-perspective description of articles. Presented at the Fifth Belgrade International Open Access Conference 2012, May 18-19, Belgrade.

Koshman, S. (2006). Visualization-based information retrieval on the web. *Library & Information Science Research*, *28*(2), 192–207.

Kules, W., Wilson, M. L., & Shneiderman, B. (2008). *From keyword search to exploration: How result visualization aids discovery on the web* (Technical Report No. 1516920080208). Retrieved from http://eprints.ecs.soton.ac.uk/15169/

Kummamuru, K., Lotlikar, R., Roy, S., Singal, K., & Krishnapuram, R. (2004). A hierarchical monothetic document clustering algorithm for summarization and browsing search results. In *Proceedings of the 13th international conference on World Wide Web* (pp. 658–665).

Landauer, T. K., Laham, D., & Derr, M. (2004). From paragraph to graph: Latent semantic analysis for information visualization. *PNAS*, *101*(Suppl 1), 5214–5219.

Lehmann, S., Schwanecke, U., & Dörner, R. (2010). Interactive visualization for opportunistic exploration of large document collections. *Information Systems*, *35*(2), 260–269.

Manning, C. D., Raghavan, P., & Schutze, H. (2008). *Introduction to information retrieval* (Vol. 1). Cambridge: Cambridge University Press.

Mathieson, K., & Keil, M. (1998). Beyond the interface: Ease of use and task/technology fit. *Information & Management*, *34*(4), 221–230.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, *76*(3), 413–448.

Newby, G. B. (2001). Cognitive space and information space. *Journal of the American Society for Information Science and Technology*, *52*(12), 1026–1048.

Newman, D., Baldwin, T., Cavedon, L., Huang, E., Karimi, S., Martinez, D., … Zobel, J. (2010). Visualizing search results and document collections using topic maps. *Web Semantics: Science, Services and Agents on the World Wide Web*, *8*(2), 169–175.

Nielsen, J., & Levy, J. (1994). Measuring usability: preference vs. performance. *Communications of the ACM*, *37*(4), 66–75.

Pajić, D., Šipka, P., & Kosanović, B. P. (2002). Skriptor: Bibliographic information parsing program. *Infoteka*, *3*(1-2), 13–21.

Rivadeneira, W., & Bederson, B. B. (2003). *A study of search result clustering interfaces: Comparing textual and zoomable user interfaces* (Technical Report No. HCIL-2003-36). University of Maryland. Retrieved from http://hcil.cs.umd.edu/trs/2003-36/2003-36.pdf

Sebrechts, M. M., Cugini, J. V., Laskowski, S. J., Vasilakis, J., & Miller, M. S. (1999). Visualization of search results: a comparative evaluation of text, 2D, and 3D interfaces. In

Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 3–10). New York, NY, USA: ACM. doi:10.1145/312624.312634

Shen, Z., Ogawa, M., Teoh, S. T., & Ma, K. L. (2006). BiblioViz: A system for visualizing bibliography information. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation-Volume 60* (pp. 93–102).

Šipka, P. (2005). The Serbian citation index: context and content. In *Proceedings of ISSI 2005 - 10th International Conference of the Society for Scientometrics and Informetrics, Stockholm, Sweden, July 24-28, 2005* (pp. 710–711). Stockholm: ISSI and Karolinska University Press.

Šipka, P. (2013). Bibliometric Quality of Serbian Journals 2002-2011: More Than Just a Dress for Success (pp. 161–165). Centre for Evaluation in Education and Science. doi:10.5937/BIOAC-110

Spoerri, A. (2004). Visual search editor for composing meta searches. *Proceedings of the American Society for Information Science and Technology*, *41*(1), 373–382.

Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, *10*(3), 321–351.

Sutcliffe, A. G., Ennis, M., & Hu, J. (2000). Evaluating the effectiveness of visual user interfaces for information retrieval. *International Journal of Human-Computer Studies*, *53*(5), 741–763.

Tanabe, S., Oyobe, K., Sunaoka, N., Yokoyama, S., & Miyadera, Y. (2002). A visualization system of relationships among papers based on the graph drawing problem. In *Proceedings of the Sixth International Conference on Information Visualisation* (pp. 202–210).

Treharne, K., & Powers, D. M. W. (2009). Search engine result visualisation: Challenges and opportunities. In *13th International Conference on Information Visualisation* (pp. 633–638).

Turetken, O., & Sharda, R. (2005). Clustering-based visual interfaces for presentation of web search results: An empirical investigation. *Information Systems Frontiers*, *7*(3), 273–297.

Van Raaij, E. M., & Schepers, J. J. L. (2008). The acceptance and use of a virtual learning environment in China. *Computers & Education*, *50*(3), 838–852. doi:10.1016/j.compedu.2006.09.001

Veerasamy, A., & Belkin, N. J. (1996). Evaluation of a tool for visualization of information retrieval results. In *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 85–92).

White, R. W., Kules, B., & Bederson, B. (2005). Exploratory search interfaces: Categorization, clustering and beyond. *ACM SIGIR Forum*, *39*(2), 52–56.

Wu, L. L., Chuang, Y. L., & Joung, Y. J. (2008). Contextual multi-dimensional browsing. *Computers in Human Behavior*, *24*(6), 2873–2888.

Xu, S., Jin, T., & Lau, F. (2009). A new visual search interface for web browsing. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining* (pp. 152–161).

Zhang, J. (2010). *Visualization for information retrieval*. Berlin: Springer.