Evaluation of Web Based Information Systems: Users' Informing Criteria

Stefka Toleva–Stoimenova
State University of Library Science and Information Technologies, Sofia, Bulgaria

s_toleva@yahoo.com

Abstract

This paper examines a popular concept for information systems success through a review of criteria and measures in the relative literature. These measures are combined in a framework for usability study of information systems in four aspects: technological measures, effectiveness measures, user-oriented measures, and feedback provided from the system. The paper focuses on information system's quality to inform their users successfully. An empirical experiment for evaluation of websites, as information systems, is presented.

Keywords: information systems, effectiveness, quality of informing, evaluation

Introduction

Advancement of information technologies in the last decades switched the focus of interest from technical performance to service effectiveness and user satisfaction of information systems. Additional factors such as globalization of the market and rapid growth of the volume of accessible data also influenced this transition. The key term marking this development is Informing Science as the discipline dedicated to study how information systems serve their clients. This development caused changes in the criteria of performance evaluation of information systems as well. In the current society Information Systems play a crucial role in mediation between information resources and the clients of information. Their evolution is marked by development of new approaches to access information influenced by evolution of languages, culture, behavior and growing information needs, caused by development of information and communication technology. In the modern society every facet of human life is heavily influenced by information obtained via information technologies. Availability of relevant information is of key importance in decision-making to allow obtaining rational decisions.

This new social environment raises the problem of evaluation of information systems as well. In addition to technical measures, developed in the past, nowadays the assessment of information system performance requires addressing the issue of information quality according to users' need, taking into account the psychological and behavioral aspects of human mental processes.

The paper shares the experience obtained in assessment of university web sites according to system effectiveness and users' satisfaction. An empirical survey is used to highlight students' reaction on provided information.

The paper is organized in three sections. The first section represents the problem of overall usability and success of the information systems in respect to its effectiveness, efficiency, and user satisfaction.

The second section of the paper examines the traditional methods and measures for evaluation of information systems quality. Combined, these measures form a framework for usability and success regarding the ability of information systems to inform the users. This evaluation framework contains measures in four aspects: technological measures, effectiveness measures, user-oriented measures, and feedback provided by the system.

In the third section this evaluation framework is applied to a set of university websites, considering them as information systems. Data collected in this empirical experiment are presented, analysed, and commented.

The conclusion summarizes the experience obtained in evaluation of information systems.

Evaluation of Information Systems Success

How to measure usability is an important question in information systems research and evaluation. Different approaches to measure usability are explored and even defined in international standards.

The ISO's definition of usability (Brooke, Bevan, Brigham, Harker, & Youmans, 1990) is usage, user, and contextually oriented: "the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in a particular environment."

The term effectiveness relates to the accuracy and completeness with which users achieve certain tasks and sub-tasks. Indicators of effectiveness include quality of solution and error rates.

Measures of efficiency relate the level of effectiveness achieved to the resources consumed in performing tasks. Indicators of efficiency include task completion time and learning time.

Satisfaction is composed of user's comfort (whether the user feels good and pleased, or tense and uncomfortable) and attitudes towards the use of the system. If satisfaction is low when efficiency is high, it is likely that the user's goals do not match the goals selected for measurement of efficiency.

The overall system consists of users, tasks, equipment (hardware, software and materials), and physical and organisational environments that influence the interaction.

These four aspects are functions of the context in which the system is used. The characteristics of the context may be as important in determining usability as the characteristics of the IS itself. The relationship between the factors is illustrated in Figure 1 (Bevan & Macleod, 1994).

The question "what to evaluate?" (van Rijsbergen, 1979) boils down to what can we measure that will reflect the ability of the system to satisfy the user. In the context of IS, as early as 1966, Cleverdon (Cleverdon & Keen, 1966) gave an answer to this with list of the following six measurable quantities that reflect users' ability to use the system:
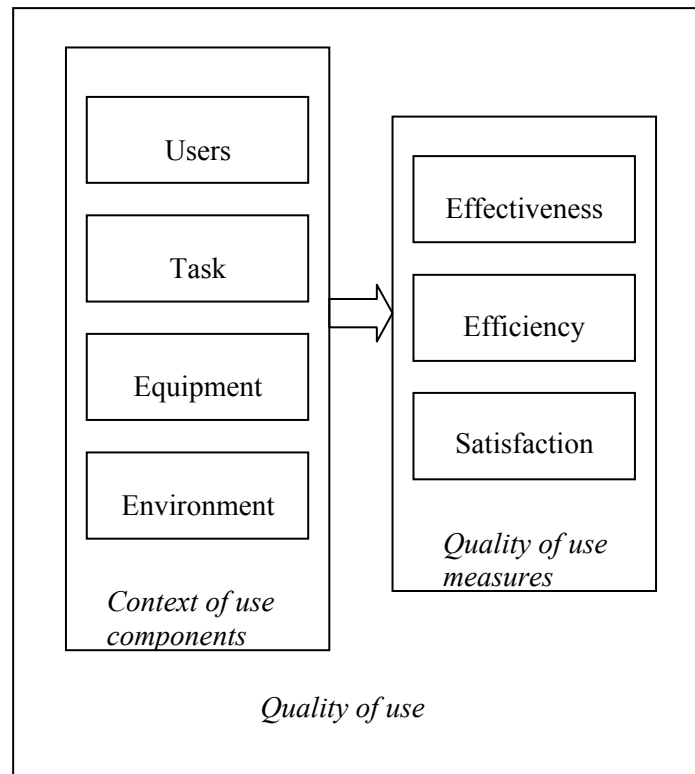
Figure1. Usability factors

(1) The *coverage* of the collection, that is, the extent to which the system includes relevant matter;

(2) the *time lag*, that is, the average interval between the time the search request is made and the time an answer is given;

(3) the form of *presentation* of the output;

(4) the *effort* involved on the part of the user in obtaining answers to his search requests;

(5) the *recall* of the system, that is, the proportion of relevant material actually retrieved in answer to a search request;

(6) the *precision* of the system, that is, the proportion of retrieved material that is actually relevant.

Currently these six values are equally important indicators in assessing the work of information systems. But the amount of information that is operating today, especially on the Internet is enormous and incomparable with the amount of information before 40, even 10 years ago. This determines the need to change the criteria used to measure the success of information retrieval systems on the Internet, which is expressed by placing a different emphasis on these 6 variables or their derivatives.

Recall and precision are the most widely used indicators for the measurement of effectiveness of the retrieval systems. Thus, the notion of relevance and users' relevance judgments are critical to the theory and research of IR. However the user should be interested in the ability of the system to retrieve relevant documents while at the same time holding back non-relevant one. This determines the fact that today more important feature of information found as a result of a specific query is its precision, not recall.

Furthermore early computer systems were built mainly for scientific purposes and often built by those who would ultimately use them. Modern methods for developing systems consider the involvement of other, non-technical professionals, in designing systems. It is assumed that the more effective system will more satisfy the user. However, user perceptions do not always coincide with system designers' notions of quality. For example, user satisfaction commonly depends very strongly on user interface design issues, including the layout, clarity, and responsiveness of the user interface, which are independent of the quality of the results returned. In order to reduce the efforts made by the user to obtain response to its query, nowadays the user friendly systems are designed.

This requires the evaluation of IS in terms of their ability to inform users, including the following aspects:

- Measures for the effectiveness of the system;

- Technological measures for evaluating the system;

- User-oriented measures;

- Feedback.

The next section illustrates, through a brief review, all of these criteria.

Information Systems Evaluation Criteria

Measures for the Effectiveness of the System.

The two most frequent and basic measures for information retrieval effectiveness are precision and recall (Baeza-Yates & Ribiero-Neto, 1999; Manning, Raghavan, & Schütze, 2008; van Rijsbergen,1979).

Precision (P) is the ratio of relevant documents retrieved to the number of documents retrieved.

Recall (R) is the ratio of retrieved documents that are relevant to the user's information need. to the number of relevant documents in the collection.

A popular alternative pair includes recall and fall-out (the proportion of non-relevant documents retrieved). A single measure that trades off precision versus recall is the F measure, derived by van Rijsbergen (1979). Many other modifications of precision and recall for evaluating the effectiveness of information retrieval systems have been proposed as well.

These measures require a collection of documents and a query. All these measures assume a ground truth notion of relevancy: every document is known to be either relevant or non-relevant to a particular query.

The effectiveness of the indexing techniques, in practice, is measured by their average retrieval recalls and precisions for the top documents in the document rank list. The small number of documents to evaluate is used because we are particularly interested in measuring the effectiveness the techniques were for quick retrieval, where the user did not want to spend time checking the relevance of too many retrieved documents.

The result of the information retrieval process is a compromise between recall and precision.

Overall, the purpose of the system is to fulfill users' information needs by retrieving relevant documents and not to return results based on simple keyword matches without any regard for the user context.

Practically in the process of searching specific information various difficulties arise, such as:

- unappropriate method - the method does not correspond to the data structure of the system;

- inadequate model (language) for presentation of information;

- inability to pursue a dialogue with information system;

- lack of correspondence with the terminology, using "fuzzy" concepts;

- ambiguity of human language;

- lack of complete information as a result of extracting only fragments.

Advanced semantic technologies enable computers to reach a higher level of understanding of the meaning of information being processed. Ontologies try to resolve the problem of ambiguity in natural language and the problems that arise due to the use of transmission meanings, analogy, comparison or metaphor.

Technological Measures for Evaluating the System

The performance evaluation of an information retrieval system is the process that obtains the time and space measures. Usually it consists in estimating the factors, which are in direct relationships with the system technology:

- database coverage and volume;

- the response time for the search process – it is a function of indexing scheme;

- expressive power of query language;

The first two of these main criteria are measurable. Nevertheless, the response time is not an easily estimable parameter because it depends on many other parameters: disk space used by the index data structures, CPU speed, disk speed, system workload, accumulation query statistics, etc.

As an instrument to improve the information retrieval process are used different indexing techniques. The majority of them rely on selecting good document descriptors, such as keywords, or *terms*, to represent the content of documents and helps differentiate the document from other documents in the collection.

Kobayashi and Takeda (2000) consider that there is a relationship between the speed (or the response time), recall and precision.

In general, the requirement is as faster as possible response to the query. According to some studies, however, the user needs an optimal response time, i.e., the system must provide the user ability to respond and to obtain feedback.

User-oriented Measures

The communication process between the systems and users is an important indicator in the evaluation of information systems, which is a priority in the early stage of their design. Usability of information systems are evaluated based on various indicators that present the user's perception of its efficiency and effectiveness.

Users' expertise (knowledge about the specific subject area and knowledge about the used system and its capabilities), user profile, graphical user interface, and also some psychological, behavioral and cultural aspects related to the pair user-system are taken into account to determine the sensitivity of the system toward the user.

Quality of information depends on the time and efforts required to reach a certain level of user experience, ease of use, system flexibility, positive attitude, which the system inspires to the users.

In the area of information system discipline a number of studies have attempted to capture the overall user satisfaction and its factors. Doll, Xia, and Torkzadeh (1994) developed a 12-item model providing five determinants framework of user satisfaction: content, accuracy, format, ease of use and timeliness. Bates (1989) and Belkin, Marchetti, and Cool (1993) developed models of browsing and searching strategies on online systems. Saracevic, Kantor, Chamis, and Trivison (1988) studied cognitive factors and user characteristics. Borgman (1989), Hsieh-Yee (1993) and many recent researchers investigate the user expertise as a very important factor. They considered the different information needs and search approaches of experienced versus novice users of age factors in information seeking and of users with various professional experience.

On the other hand personalization is a very common concept in the development of web-based information systems. The goal of personalization is to facilitate the user's needs and to enable him to obtain relevant information. For this purpose the user behavior and habits are studied and recorded systematically, the history of user's searches is maintained to allow better selection in his next search. For example, Amazon (www.amazon.com) recommended books to its visitors based on their previous purchases and according to judgments of people who liked the same books.

Besides the mentioned above factors other psychological aspects of informing systems are included in this aspect oriented to the user, such as credibility of information sources. In the contemporary information age credibility is supposed to be a very significant attribute of information and quality of informing (Gackowski, 2006a, 2006b).

Sometimes measures of satisfaction can provide a useful indication of the user's perception of system usability, even if it is not possible to obtain measures of effectiveness and efficiency.

System Feedback

In cybernetics, the feedback is the key concept in controlled systems. Users search relevant information regarding their needs in decision making. Information retrieval is the process of matching a representation of an information need, usually via a user-supplied query, to an indexed document representation. Providing relevance feedback (RF) is one of the most powerful methods for improving performance of information retrieval process. It involves the user in the retrieval process to improve the final set of the retrieved documents. This is one of the main measures for evaluating the quality of informing offered by information systems through which actually improvement according almost all other measures can be achieved.

There are numerous studies on this topic which presented different approaches for system feedback. One problem of RF evaluation is what to measure: the quality of the feedback or how well does the system improve the user's query; or the quality of the adaptation to the information need or how well does the algorithm track how the query is changing (Ruthven & Lalmas, 2003)

Empirical Evaluation of informing systems

As the amount of available digital information grows, especially on the World Wide Web, the ability to search it becomes increasingly critical. Google currently claims their index contains over eight billion pages, others also claim index sizes in the billions.

On the Web, usability of particular information system is a significant factor. The common users want to accomplish specific task as timely, accurately and completely as possible, and the interaction experience of doing so is as pleasurable as possible with minimal user frustration.

The main goal of this study is evaluation of the quality of informing of an information system, in particular web site. Web site is not a typical information retrieval system. It may be a gateway to a set of structured databases, textual databases, or other services outside of the Web itself. Regarding evaluation of web pages, Alexander and Tate (1999) defined five main criteria: accuracy, authority, objectivity, currency and coverage of web sites.

A frequently used method for collecting and recording of information from the user's perspective is questionnaire. Such approaches for information system evaluation typically assess the user satisfaction (http://hcibib.org/perlman.). They help to obtain the user's perceptions and expectations for a specific service's characteristics

In this paper the goal is to assess the quality of information systems to inform their users – informativeness of the system.

A questionnaire was designed to survey user opinions regarding university web sites in respect to their quality to inform users by considering the major listed above measures. There are some limitations in this study - we do not have real indicators for system efficiency. Efficiency measure provides a more general work rate by trading off the quantity and quality of output against time. Task time itself is sometimes used as a usability measure. However it is hard to measure and compare similar systems, when used by the same users for the same tasks in the same environments; or users when using the same system for the same tasks in the same environment; or tasks when carried out by the same users on the same system in the same environment. It is also impossible to evaluate resources spent on task such as labour hours, process expense and materials consumed. Also it was impossible to detect role of the feedback in those systems.

Some of the important details about c*ontext of use* include:

Users: As a representative part of the general population of Internet users have been used students from SULSIT (State University of Library Studies and Information Technologies, Sofia). They are in their first year of bachelor degree in Information Technology and the typical age range is 19 - 21.

Tasks: This questionnaire was designed as a part of the course Processes and Models for Informing in SULSIT, Sofia (Christozov, Denchev, & Toleva–Stoimenova, 2009). The respondents are offered a five-level scale to answer the questions related to the particular task.

It is assumed that they are looking for a university to study. Actually they did this a year ago. Choosing a university is based on collection of typical information about it. One way to do it today is university's website.

In the case being considered, every student received two universities with two concrete specialties for evaluation and comparison. The specialties are similar to their present specialty in SULSIT, namely Information Technology.

The students have to collect the following information for Bulgarian universities from their websites:

- curriculum of the course, including mandatory and optional subjects that are thought during 4-yaer study (bachelor program);

- additional conditions, including course assignments and projects, internships;

- conditions to complete the degree – examinations, thesis.

In the paper are presented the results about two universities and their websites – State University of Library Science and Information Technologies (www.svubit.org) and Sofia University (http://www.uni-sofia.bg).

Organisational and technical environment: Students are familiar with information technologies and have a good computer and Internet literacy. They are provided the requisite equipment for this task, such as computer and Internet.

The questionnaire is composed by 10 questions (see the Appendix).

The first two questions represent a self assessment about user's expertise. As it is shown in Tables 1 and 2 below students have enough knowledge and ability to use computers and Internet technology efficiently.

Figures 1a, 1b, 2a, and 2b show the results graphically.

Table 1: Computer Literacy

	SULSIT			Sofia University		
	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent
satisfactory	2	7,7	7,7	4	12,9	12,9
good	8	30,8	38,5	12	38,7	51,6
very good	16	61,5	100,0	15	48,4	100,0
Total	26	100,0		31	100,0	

Table 2: Internet Literacy

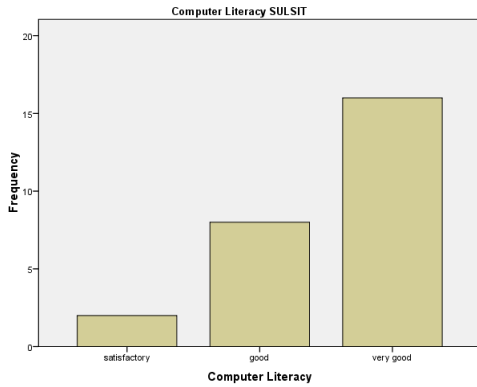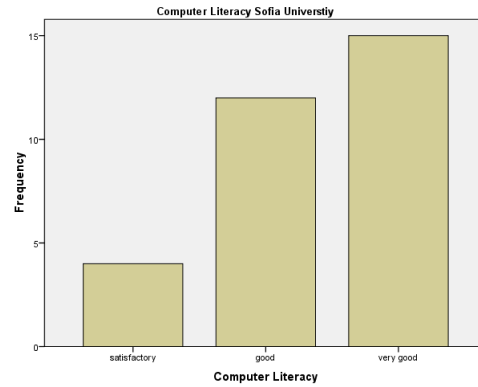|	SULSIT			Sofia University		
	Frequency	Percent	Cumulative Percent	Frequency	Percent	Cumulative Percent
bad				1	3,2	3,2
satisfactory	1	3,8	3,8	2	6,5	9,7
good	6	23,1	26,9	10	32,3	41,9
very good	19	73,1	100,0	18	58,1	100,0
Total	26	100,0		31	100,0	

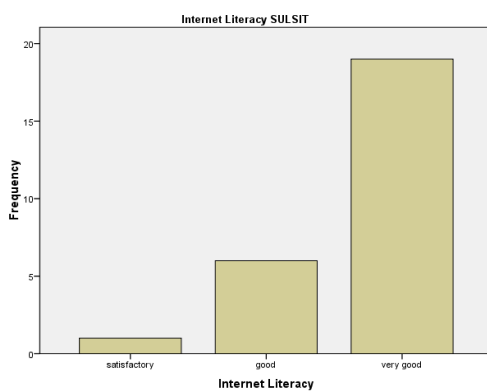**Figure 1a SULSIT** **Figure 1b Sofia University**
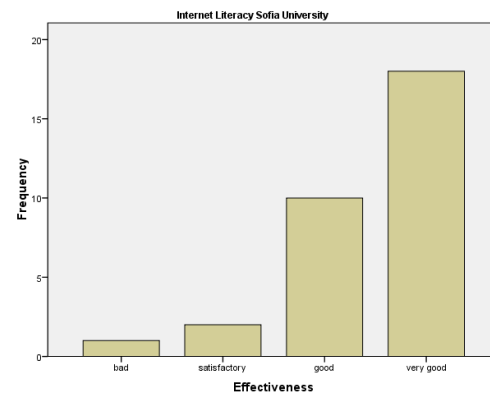

Figure 2a SULSIT **Figure 2b Sofia University**

The next eight questions were designed to measure system effectiveness and user satisfaction.

System effectiveness is measured by the following questions: q7,q8, q9.

User satisfaction – by the following: q3, q4, q5, q6, q10.

Table 3 and Table 4 summarize the experience obtained in evaluation of websites to the two major criteria effectiveness and users satisfaction, according to SULSIT and Sofia University.

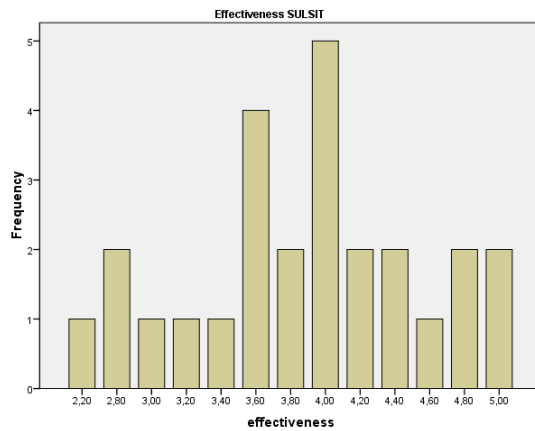The distribution of answers for the two web sites are presented in Figures 3 and 4.
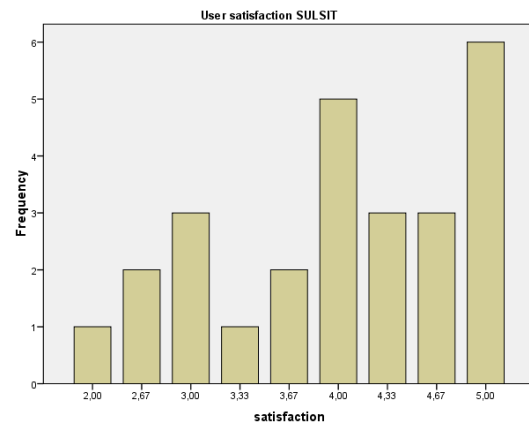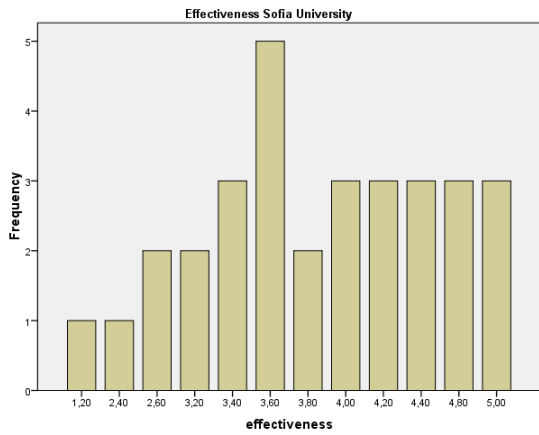

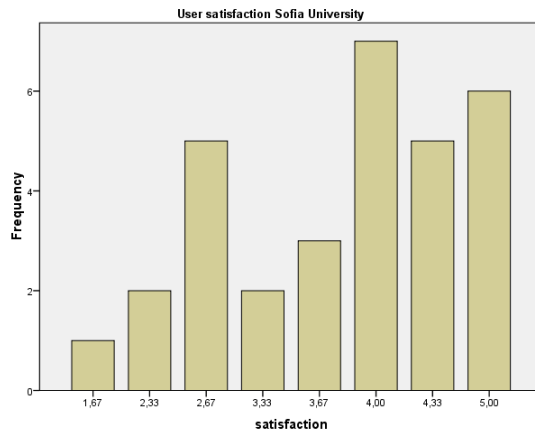**Figure 3a SULSIT** **Figure 3b Sofia University**

Figure 4a SULSIT

Figure 4b Sofia University

Table 3 SULSIT

	N	Minimum	Maximum	Mean	Std. Devia-tion	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
q3	26	2	5	4,04	,871	,758	-,865	,456	,570	,887
q4	26	2	5	3,77	1,142	1,305	-,555	,456	-1,069	,887
q5	26	3	5	4,38	,697	,486	-,703	,456	-,575	,887
q6	26	1	5	3,62	1,023	1,046	-,338	,456	,200	,887
q7	26	2	5	3,85	,967	,935	-,532	,456	-,500	,887
q8	26	2	5	3,96	1,038	1,078	-,615	,456	-,759	,887
q9	26	2	6	4,19	1,096	1,202	-,807	,456	-,185	,887
q10	26	1	5	3,58	1,270	1,614	-,884	,456	-,075	,887
Valid N (list-wise)	26									

Table 4 Sofia University

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
q3	31	1	5	4,03	,875	,766	-1,342	,421	3,491	,821
q4	31	1	5	3,55	1,028	1,056	-,336	,421	-,130	,821
q5	31	1	5	3,68	1,107	1,226	-,719	,421	-,252	,821
q6	31	2	5	4,10	,978	,957	-,889	,421	-,105	,821
q7	31	2	5	4,19	,910	,828	-,693	,421	-,728	,821
q8	31	1	5	3,58	1,119	1,252	-,445	,421	-,593	,821
q9	30	2	5	3,63	,964	,930	-,159	,427	-,833	,833
q10	31	1	5	3,71	1,131	1,280	-,561	,421	-,468	,821
Valid N (list-wise)	30									

Discussion

Most of respondents were able to find information they needed successfully using the university website. The respondents are able to find information they needed more quickly (q8) from SULSIT website than SU website.

Students' answers to three of the questions (q3, q4, q10) have very similar values about Sofia university and SULSIT, their means do not differ very much. The rest of the questions about website navigation and interface (q5 and q6) obtained various answers. Users liked the SULSIT website navigation (mean 4,38) more than the website navigation of Sofia university (mean 3,68). But they assessed Sofia university website interface (mean 4,19) like more pleasant than SULSIT website interface (mean 3,68).

Conclusion

The paper shares the experience obtained in assessment of university web sites according to system usability. An empirical experiment for evaluation of websites is presented. A questionnaire surveyed user opinions regarding university web sites in respect to their quality to inform users. The future work will address further improvement and development of questionnaire, such as measuring efficiency and examining some extra issues of system effectiveness and user's satisfaction.

References

Alexander, J., & Tate, M. A. (1999). *Web wisdom: How to evaluate and create information quality on the web*. Mahwah, NJ: Lawrence Erlbaum. Retrieved from http://www2.widener.edu/Wolfgram-Memorial-Library/webevaluation/webeval.htm

Baeza-Yates, R., & Ribiero-Neto, B. (1999). *Modern information retrieval.* Addison-Wesley.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407–424.

Belkin, N. J., Marchetti, P. G., & Cool, C. (1993). BRAQUE: Design of an interface to support user interface in information retrieval. *Information Processing & Management, 29*(3), 237–251.

Bevan, N., & Macleod, M. (1994). Usability measurement in context. *Behaviour and Information Technology, 13*, 132-145.

Borgman, C. L. (1989). All users of information systems are not created equal: An exploration into individual differences. *Information Processing and Management, 25*(3), 237-252.

Brooke, J., Bevan, N., Brigham, F., Harker, S., & Youmans, D. (1990). Usability statements and standardisation -work in progress in ISO. In: *Human Computer Interaction - INTERACT'90*, D Diaper et al. (Ed.), Elsevier.

Christozov, D., Denchev, S., & Toleva–Stoimenova, S. (2009). Informing processes and models: A core course in IS curriculum. *Issues in Informing Science and Information Technology, 6*, 143-151. Retrieved from http://iisit.org/Vol6/IISITv6p143-151Christozov624.pdf

Cleverdon, C., & Keen, M. (1966). Factors Affecting the Performance of Indexing Systems, Vol 2., pp. 37-59. ASLIB, Cranfield Research Project. Bedford, UK

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument, *MIS Quarterly*, December 1994, 453-461.

Gackowski, Z. J. (2006a). Quality of informing: Credibility. Provisional model of functional dependencies. *Proceedings of the 2006 Informing Science and IT Education Conference*, 2006. Retrieved from http://proceedings.informingscience.org/InSITE2006/ProcGack194.pdf

Gackowski, Z. J. (2006b). Quality of informing: Disinformation and bias – Philosophical background and roots. *Journal of Issues in Informing Science and Information Technology*, *3*, 731-744. Retrieved from http://informingscience.org/proceedings/InSITE2006/IISITZbig250.pdf

Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on online search behavior: Measuring the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science, 44*, 161-174

Kobayashi M., & Takeda, K. (2000). Information retrieval on the web, IBM Research. *ACM Computing Surveys*, *32*(2), 144-173.

Manning C., Raghavan, P., & Schütze, H. (2008). *An introduction to information retrieval.* Cambridge, England: Cambridge University Press. [Preliminary Draft]. Retrieved from http://www-csli.stanford.edu/~hinrich/information-retrieval-book.html

Ruthven, I., & Lalmas, M. (2003). A survey on the use of relevance feedback for information access systems. *The Knowledge Engineering Review*, *18*(2), 95-145. Retrieved from http://inex.is.informatik.uni-duisburg.de:2004/pdf/ker_ruthven_lalmas.pdf

Saracevic, T., Kantor, P., Chamis, A. Y., & Trivison, D. (1988). A study of information seeking and retrieving; I. Background and methodology. *Journal of the American Society for Information Science 39*(3), 161–176.

van Rijsbergen, C. J. (1979). *Information Retrieval* (2nd ed.). London: Butterworths.

Appendix

Questionnaire

1 - "very bad"; 2 - "bad"; 3 - "satisfactory"; 4 - "good"; 5 - "very good".

1) Try to assess your computer literacy   1 2 3 4 5

2) Try to assess your Internet literacy   1 2 3 4 5

1 - "strongly disagree"; 2 - "disagree"; 3 - "not agree, neither disagree";
4 - "agree"; 5 - "strongly disagree".

3) The information provided from the website is easy to understand.   1 2 3 4 5

4) The organization of information on website screen is intuitive.   1 2 3 4 5

5) The website navigation is intuitive.   1 2 3 4 5

6) The interface of website is pleasant.   1 2 3 4 5

7) I successfully found information I needed using the website.   1 2 3 4 5

8) It was simple to find information I needed using the website.   1 2 3 4 5

9) I was able to find information I needed.quickly using the website.   1 2 3 4 5

10) Overall I liked this site/ I felt comfortable using the website.   1 2 3 4 5

Biography
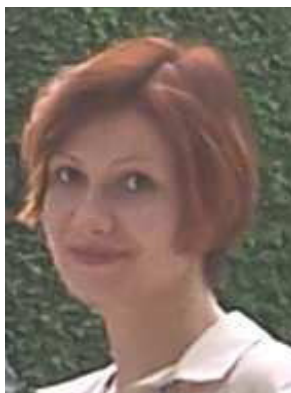
Stefka Toleva-Stoimenova is a Senior Assistant Professor of Mathematics and Informatics at the State University of Library Studies and Information Technologies, Sofia, Tzarigradsko shosse Blvd. 119, Bulgaria 1784, s_toleva@yahoo.com. She has obtained her MSc degree in Industrial Automation from the Faculty of Automation and System Design, Technical University – Sofia. Her main research interests are in the field of Automation and Informatics. She is currently researching to obtain his PhD degree on Informing Science and Automated Information Systems. She is a member of Bulgarian Informing Science Society