In-Depth Analysis of Selected Topics Related to the Quality Assessment of E-Commerce Systems

Antonia Stefani[1], Dimitris Stavrinoudis[1], and Michalis Xenos[1,2]

[1] School of Sciences & Technology,Hellenic Open University,23 Sachtouri Str, Patras, Greece
stefani@eap.gr,stavrino@eap.gr,xenos@eap.gr
[2] Research Academic Computer Technology Institute, 61 Riga Feraiou Str, Patras, Greece

Abstract. This paper provides an in-depth analysis of selected important topics related to the quality assessment of e-commerce systems. It briefly introduces to the reader a quality assessment model based on Bayesian Networks and presents in detail the practical application of this model, highlighting practical issues related to the involvement of human subjects, conflict resolution, and calibration of the measurement instruments. Furthermore, the paper presents the application process of the model for the quality assessment of various e-commerce systems; it also discusses in detail how particular features (data) of the assessed e-commerce systems can be identified and, using the described automated assessment process, lead to higher abstraction information (desiderata) regarding the quality of the assessed e-commerce systems.

Keywords: E-commerce Systems, Quality Assessment, Bayesian Networks.

1 Introduction

In the past few years, a large number of e-commerce systems have been developed. To ensure the production of high quality e-commerce systems, it is important for developers to be able to assess the quality of such systems. The latter is inevitably linked with the receivers' perception of quality. It must be noted that e-commerce systems differ from other web applications in that a basic condition of their success is the total involvement of the end-user at almost every stage of the purchasing process (Henfridsson & Holmstrom, 2003), which is not the case with other web applications. The growth that Business to Consumer (B2C) e-commerce systems have experienced in the past few years has given rise to the problem of identification of those factors that determine end-user acceptance of such systems (Chen et al., 2004).

The work presented in this paper is based on a Bayesian Network model (Stefani et al., 2003). The attributes of this model are quality characteristics. Quality assessments using this model can take the form of either a probability value for the abstract 'quality', or a vector of values for each quality characteristic. To be able to interpret this vector of values in a way that provides conclusions about e-commerce systems' quality, one should have collected and analyzed a significant volume of data that will aid in calibrating the measurement scales. This is what this paper focuses on: the presentation of the process used to conduct the quality assessment. Moreover, to help

J. Filipe et al. (Eds.): ICETE 2005, CCIS 3, pp. 39–48, 2007.

make the discussion clearer, this paper presents and explains selected practical cases, serving as examples, related to the process of quality assessment.

This paper is structured in five sections. After this brief introduction, section 2 presents related work. Section 3 presents the model and describes the process used for quality assessment of e-commerce systems. Then, section 4, presents practical cases of quality assessment, used to explain the application of the model and how its results can lead to higher abstraction conclusions. Finally, section 5 summarizes the main conclusions of the paper.

2 Related Work

A number of approaches towards assessing the quality of e-commerce systems focus on the technological aspects of such systems, thus providing a technology-oriented view of quality (Zwass, 1996; Elfriede & Rashka, 2001). Other approaches assess the quality of e-commerce systems as perceived by the end-user, but focusing mainly on the usability of such systems. Such approaches use software evaluation methods such as inspection (Nielsen, 1994) and inquiry methods (Shaw & DeLone, 2002) in order to record end-users' perception of usability. Studies on e-commerce systems quality also focus on more specific quality characteristics such as issues that warrant successful transactions (Bidgoli, 2002), maximize the perceived trustworthiness (Egger, 2001; Slyke et al., 2004), or ensure e-commerce systems reliability (Elfriede & Rashka, 2001).

Although, all the above factors are affecting the quality of e-commerce systems and are prerequisites for their success, they are not the only ones that relate to e-commerce systems quality. In order to model e-commerce systems quality, a global approach, such as the one discussed in this paper, is required combining all factors affecting quality.

Some related works are using questionnaires to detect users' opinions, the data from which are statistically analyzed in order to lead in values measuring quality characteristics such as usability (Sauro & Kindlund, 2005). This is a common practice, since users' opinion is very important for the assessment of e-commerce systems (Julian & Standing, 2003), as well as the active involvement of users into the evaluation process (Henfridsson & Holmstrom, 2003; Chen et al., 2004).

The work presented in this paper, differs from questionnaire-based surveys in that it uses a process aiming to limit subjectivity and frequent errors in similar surveys. Furthermore, thanks to the nature of the used model, the assessment process can be used forwards and backwards, i.e. during the quality design phase for setting the quality goals of an e-commerce system.

3 The Model Application Process

In order to assess the quality of e-commerce systems as perceived by the end-users, one must focus on the user oriented quality characteristics of ISO 9126 (ISO/IEC 9126, 2001), which are functionality, usability, reliability and efficiency, and their sub-characteristics. The model used in this process is based on Bayesian Networks,

which are a special category of graphic models where nodes represent variables and the directed arrows the relations between them. In this case, the model's nodes are the above mentioned quality characteristics as well as e-commerce characteristics that are connected to the appropriate quality characteristics, forming a number of relations between them. For each node of this model the dependent probabilities that describe the relations between the variables is determined.

The model can be used both forwards and backwards. In the forward use, the user inserts evidence to the nodes of the e-commerce characteristics, which have only two possible states: 'yes' and 'no'. In this way, the model estimates the system's quality providing the probabilities for the possible states of the nodes that represent the quality characteristics and the overall quality. The backward use of the model provides assessments regarding the child nodes (e.g. nodes of e-commerce characteristics) when the value of a parent node (e.g. node of 'quality' characteristic) is defined. Since the purpose of this paper is to present the process followed for assessing the quality of already existing e-commerce systems, we focus mainly on the forward use of the model.

This process, which is also represented in figure 1, consists of 4 different steps: a) the assignment of an e-commerce system to two evaluators and the filling of an appropriate evaluation sheet by them, b) the examination of the identity between the two evaluation sheets, c) the forward use of the model and d) the classification of the e-commerce system. These steps are described more analytically hereinafter.

The most important benefit of applying this model is the fact that it provides an easy and non-subjective way to rank an e-commerce system according not only to the overall quality, but to each quality characteristic as well. The limitation of the subjectivity while evaluating such a system is achieved because of the values of the possible states of the nodes that represent the e-commerce characteristics in the model. In other words, the evaluators are asked to determine the existence, or not, of these specific characteristics answered in the evaluation sheet by a simple yes or no. Although the contribution of the evaluators to the assessment of the quality of e-commerce systems is trivial and non-subjective, the first step of the process is to assign this task to two evaluators. In this way, possible errors while filling the evaluation sheet, mainly because of careless mistakes or because of the possibility of misunderstanding a question of the sheet, are avoided. It must also be stressed that the evaluators chosen for this process must be experts. This does not necessarily indicate that they should be experienced in judging or estimating the quality of a software product. But they should be expert users of such e-commerce systems and they should also be aware of the used terminology. Besides, they examine these systems only from their customer's viewpoint.

The evaluation sheet is in the form of a simple questionnaire, where the possible answers of each question are only two: 'yes' or 'no'. Although the aspects that it is concerned with are trivial, the questions must be clearly stated, and in some cases more specifically commented on, in order to avoid any misinterpretation by the evaluators. The evaluation sheet is then delivered to the evaluators either electronically or by hard copy. The questions on this sheet are stated as follows, so that they can be answered by a simple 'yes' or 'no':

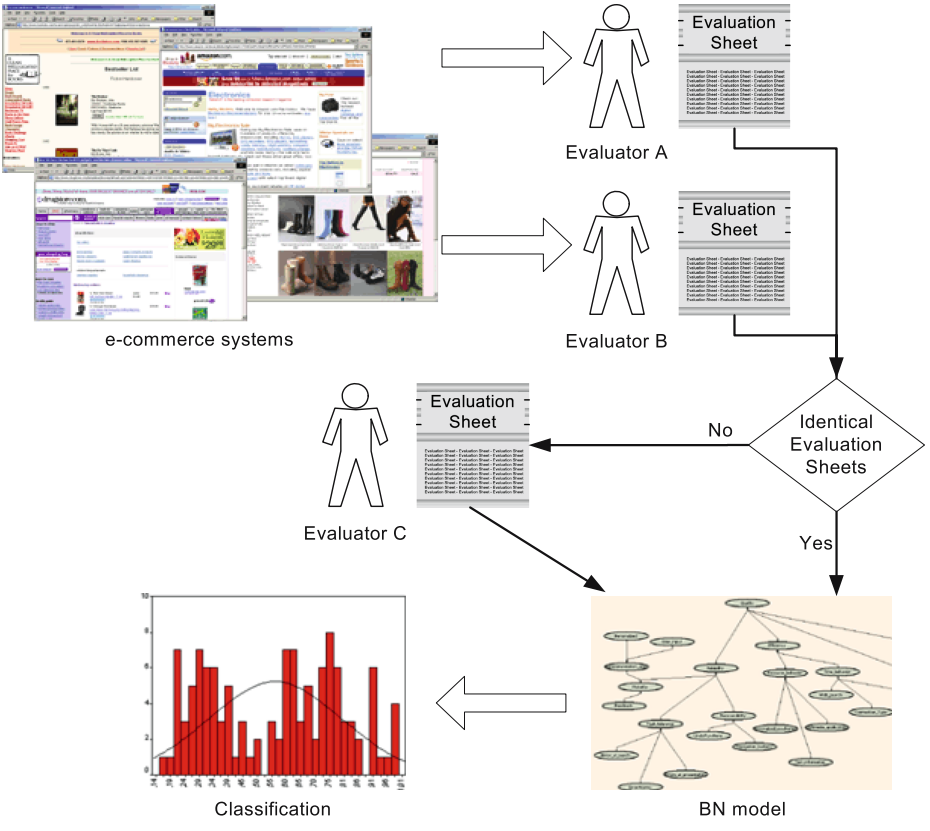

Fig. 1. The model application process

"Does the e-commerce system provide frequently asked questions (FAQs) to the user?",

"Is there a shopping cart available to the users in the e-commerce system?",

"Can the user sort automatically the results of a search by various parameters (e.g. order by price, order by manufacturer, alphabetical order, etc.)?".

The questionnaire is structured in such a way, in order to be clear to the evaluators which questions concern each quality factor. Furthermore, the sequence of the questions is in tune with the most possible sequence of actions that a user of the e-commerce system will follow when he visits the site and has a transaction with it.

The second step of the process is the examination of the answers given by the evaluators. The two evaluation sheets must be identical, since their questions are quite trivial and with obvious answers. However, because of possible mistakes while filling them, as mentioned above, a third evaluator is needed in this process. His role is to provide conflict resolution in such cases. Of course, these possible differences are easy to be solved by the first two evaluators if this is possible for them in terms of time and place. But, if this is not possible, then a third person must determine the appropriate answer when a difference between the evaluation sheets appears. The

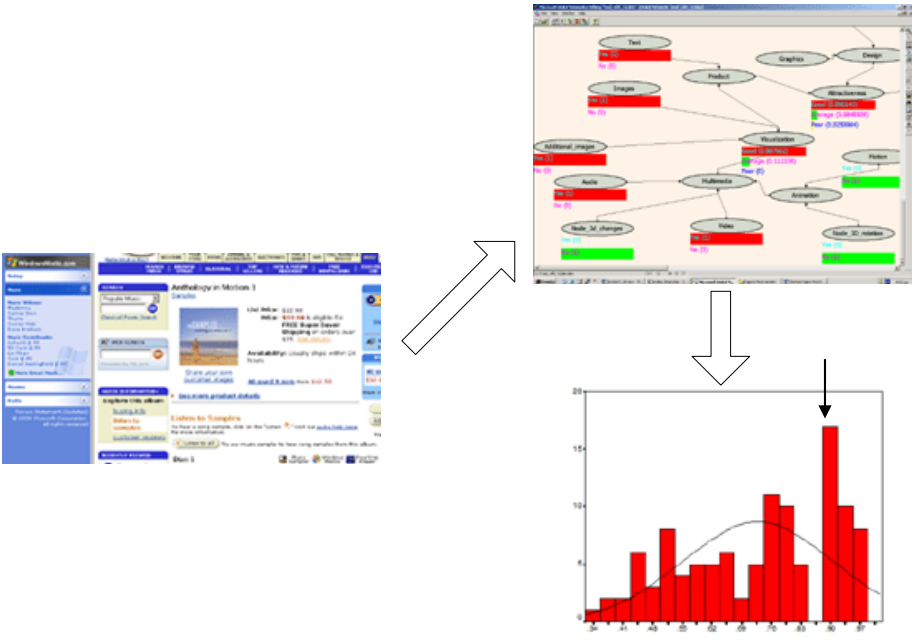

Fig. 2. Case of attractiveness

purpose of the third evaluator is also to eliminate the subjectivity if this appears in cases of significant differences between the opinions of the first two evaluators.

Having a fixed and correct evaluation sheet, the model can be used to assess the quality of the e-commerce system. In the next step of the process the answers of the sheet are inserted into the model as evidence to the nodes of the e-commerce characteristics, which are the leaf nodes of the Bayesian Network. Thus, by the forward use of the model, the corresponding probability values of all the parent nodes can be estimated automatically, since all the Node Probability Tables have already been specified. In this way, one can easily gather the provided results not only for the overall quality of the system, but also for all the quality characteristics and sub-characteristics used in the model.

The provided results from the model cannot be directly exploitable to determine the quality of the system. In fact, they are probability values of the possible states of the nodes. For example, a result of 0.88 for the usability node does not imply directly the level of the usability of the system. Thus, in the final step of the process the classification of the system for each quality characteristic and sub-characteristic is specified. This classification can be found by the means of the scale calibration table and the accompanying histograms that has been presented in (Stefani et al., 2004). In this way, using the boundaries and the scales of the probability values of the model, one is able to determine the specific category (good, average or poor) in which the e-commerce system belongs according to each quality characteristic. In other words, he can identify the cluster to which each quality characteristic belongs and detect the possible drawbacks of the system.

4 Practical Cases of Quality Assessment

For the evaluation process, two evaluators for each e-commerce system are selected who have previously used (e.g. for purchasing an item online) at least two distinct e-commerce systems, regardless of the time spent on them. This shopping experience is a pre-requisite for the evaluation process, because it ensures that the evaluator is familiar with e-commerce systems' use.

Initially an evaluation pair was established for each e-commerce system. Evaluators have worked individually and answered with 'yes' or 'no' on the evaluation sheet. The sheets, for every pair of evaluators were checked in order to find any conflicts in the answers between the two evaluators. Conflicts were the different answers that we have found on the evaluation sheets. Less than 5% of the evaluation sheets have revealed conflicts; the maximum number of conflicts of each evaluation pair was two, and the questions that have presented conflicts were the same for additional evaluation pairs. Therefore it is logical to assume the results are quite objective. Reasoning for the existence of conflicts was either the time spent at the evaluation process or misunderstanding of the questions.

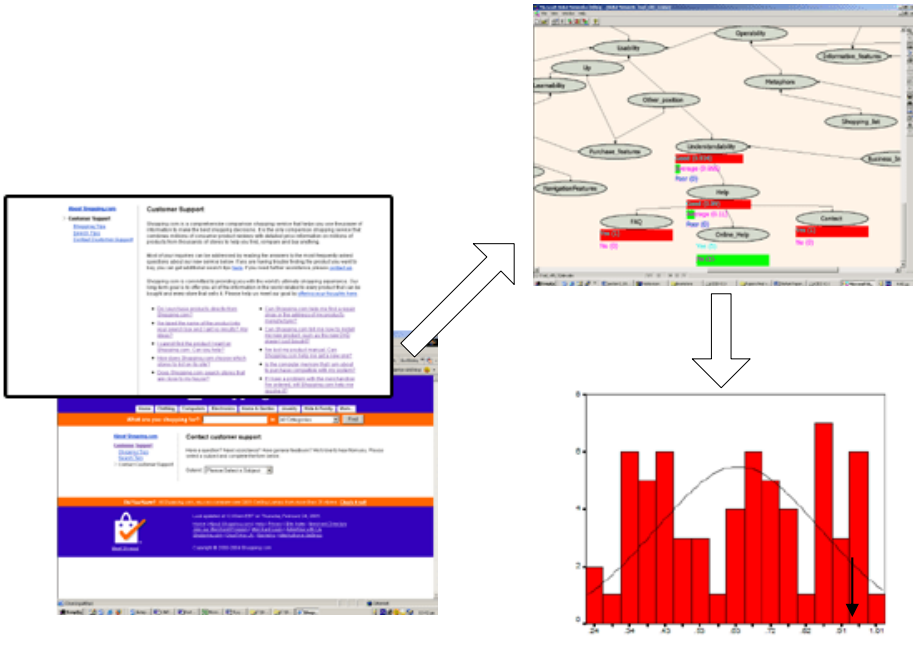

Fig. 3. Case of understandability

For example, the evaluators were addressed to examine if the e-commerce system provides video applications for the presentation of the product. The first evaluator answered 'no', because he had not found video application in the categories of Video and DVD at the e-commerce system, but the second evaluator had found video application at the CD category of the same e-commerce system and answered 'yes'.

In order to avoid these kind of conflicts, we asked the evaluators to use the ten most popular products of the e-commerce system or the products that the system sells at its home page. Additionally we asked the evaluators to proceed in a check out process in order to have a complete shopping experience.

In an ongoing process we have detected all the questions that presented conflict on the evaluation sheet and have edited them. Finally a third evaluator has been used in order to reduce the impact of the human factor. The third evaluator individually has provided answers for the questions that still presented conflicts and as a result of this process we had evaluation sheets with no subjectivity, because two evaluators agreed individually for a 'yes' or 'no'. So we have the final evaluation sheet for each e-commerce system that was used at the next step of the evaluation process. The answers for each final evaluation sheet were used as evidence in respect of the child node of the model's tool. Hereinafter we present three practical cases extracted from the quality assessment process of several e-commerce systems. In these cases we present the evaluators' answers for the evaluation sheets' criteria and also we present, in an anonymous representation, the functions and services which each e-commerce system offers.

4.1 First Case

For example, the evaluators were asked to examine the presentation of the products at each e-commerce system. The e-commerce system usually presents a product by text description where its properties are described; so each evaluator can have a description for products' characteristics and also information for cost and availability. Complementary e-commerce systems offer photographs, audio, video, graphics and 3-D representation of each product.

Figure 2 shows evaluators' evidence about the product's presentation. For image, and additional images in greater size, audio, and video samples the evaluators have answered 'yes'.

The probability value for the parent node titled 'Visualization', which refers to the visual representation of the product, is 0.88 and the value for product presentation by text and images is 0.94. Finally the probability value for the quality characteristic of Attractiveness is 0.89 and the meaning of these probability values can be explained using the scale calibration from our previous work (Stefani et al., 2004). Figure 2 presents the histogram for Attractiveness. This e-commerce system is characterized as 'Good' but the value of Attractiveness is on the boundaries between 'Good' and 'Average', meaning that the e-commerce system needs improvement at Graphics, 3-D representation, and animation.

4.2 Second Case

At the evaluation process each evaluator used Help functions that each e-commerce system supports. As Help functions we have defined the existence of FAQ (Frequently Asked Questions), contact capability via e-mail, or fax and online help. Figure 3 presents the help evaluation of an e-commerce system. The Understandability of the system as it is perceived by the customer is presented by the probability value 0.93. That means this e-commerce system belongs at Category A of

the model's scale calibration, also the probability value for Usability is 0.87 which means that belongs also at category A.

4.3 Third Case

Another category of questions on the evaluation sheet is related to the search function of the e-commerce system. Usually the search function appears as a form where the evaluator can insert words as keywords for a question. In advance the evaluator can use the search function by defining the products' categories and limits for price in order to have more accurate results. The evaluator could search by keyword but the system did not provide advanced methods. The tool defines the search engine of the e-commerce system as good by the probability value of 0.62. This result means that the search engine of the e-commerce system usually provides correct results according to evaluators' keywords. That is the most common search option at e-commerce systems, but the same system by not providing advanced methods of searching does not offer a completely operable searching function.

According to operability the same e-commerce system offers informative features as compare features for products, and cross selling mechanism for complementary products, but the e-commerce system provides notification services by e-mail to the frequent customers. Finally the e-commerce system offers metaphors like shopping cart but not shopping list where the customer could save his/her shopping preferences. Figure 4 presents these values and the total operability of the e-commerce system, which is 0.55.

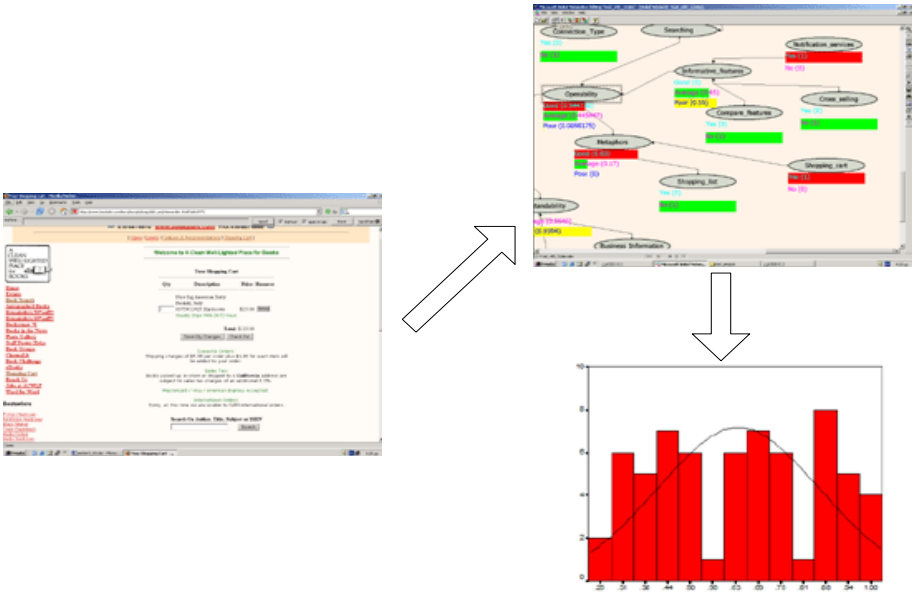

Fig. 4. Case of operability

In another case it is obvious that the absence of search functions that are extremely helpful at the purchasing process reflect the total quality of the system. The evaluator desired to view the search history of his/her searches in order to proceed with the results, and to have alternative ordering options, but the system offers none of these search options. On this evidence the model reveals that the absence of these options indicates search from this source is poor.

5 Conclusions

This paper presents a measurement process for the quality of e-commerce systems. This process uses a model based on a Bayesian Networks and consists of different steps, which are described analytically. The process has been applied in different e-commerce systems and cases of their results are also presented. Expert evaluators were asked to rank e-commerce systems by filling an appropriate evaluation sheet and by determining the existence or not of specific e-commerce characteristics in them. This sheet was formed in such a way in order to minimize subjectivity while evaluating such a system. Moreover, the process itself provides a way to eliminate any possible conflict between the evaluators' opinions.

Having a fixed evaluation sheet of an e-commerce system and by applying the Bayesian Network model the probability values of the overall quality and the quality characteristics can be assessed. Furthermore, by the means of the appropriate boundaries and scale calibration tables, the classification of the system for each quality characteristic and sub-characteristic can be specified.

The model can be used both forwards and backwards. Although the presented process is based on the forward use of the model as a summative evaluation of e-commerce systems, future work includes the application of the model during the design phase of an e-commerce system. In other words, it includes the formative evaluation of such systems. Moreover, the process of using the model should be refined dynamically, due to the continuous evolution and enhancement of the e-commerce systems and the appearance of new characteristics and functions provided by them.

References

Bidgoli, H.: Electronic Commerce Principles and Practice. Academic Press, San Diego (2002)

Chen, L., Gillenson, M., Sherrell, D.: Consumer Acceptance of Virtual Stores: A Theoretical Model and Critical Success Factors for Virtual Stores. The DATA BASE for Advances in Information Systems 35(2) (2004)

Egger, F.: Affective Design of E-Commerce User Interfaces: How to Maximise Perceived Trustworthiness. In: Egger, F. (ed.) International Conference on Affective Human Factors Design, Asean Academic, London (2001)

Elfriede, D., Rashka, J.: Quality Web Systems, Performance, Security, and Usability. Addison – Wesley, New York (2001)

Henfridsson, O., Holmstrom, H.: Developing E-commerce in Internetworked Organizations: A Case of Customer Involvement Throughout the Computer Gaming Value Chain. The DATA BASE for the Advances in Information Systems, 33(4) (2003)

ISO/IEC 9126, Software Product Evaluation – Quality Characteristics and Guidelines for the User, International Organization for Standardization, Geneva (2001)

Julian, T., Standing, C.: The value of User Participation in the E-Commerce Systems Development, Informing and IT Education Conference, Pori, Finland (2003)

Nielsen, J.: Designing Web Usability: The Practice of Simplicity. New Riders Publishing, Indianapolis. Indiana (2000)

Sauro, J., Kindlund, E.: A Method to Standardize Usability Metrics Into a Single Score, CHI2005 Methods and Usability, Portland, Oregon, USA (2005)

Shaw, N., DeLone, W.: Sources of Dissatisfaction in End-User Support : An Empirical Study. The DATA BASE for Advances in Information Systems, 33(2) (2002)

Slyke, C., Belanger, F., Comunale, C.: Factors Influencing the Adoption of Web – Based Shopping: The Impact of Trust. The DATA BASE for Advances in Information Systems, 35(2)

Stefani, A., Xenos, M., Stavrinoudis, D.: Modeling E-Commerce Systems' Quality with Belief Networks. In: VECIMS, IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems, IEEE, Los Alamitos (2003)

Stefani, A., Stavrinoudis, D., Xenos, M.: Experimental Based Tool Calibration used for Assessing the Quality of E-Commerce Systems. In: ICETE2004, 1st IEEE International Conference on E-Business and Telecommunication Networks, Portugal, IEEE Computer Society Press, Los Alamitos (2004)

Zwass, V.: Structure and macro-level impacts of electronic commerce: from technological infrastructure to electronic marketplaces. International Journal of Electronic Commerce, 1 (1996)