Evaluating the Quality of Social Web Applications Using the LSP Method

Tihomir Orehovački
Juraj Dobrila University of Pula, Pula, Croatia
`tihomir.orehovacki@unipu.hr`

Abstract—Quality is an essential determinant of the success of every type of software and social Web applications are not an exception. It is therefore of great importance that the examination of the degree to which social Web applications meet predefined requirements related to particular facets of quality is performed effectively and frequently. With an objective to facilitate evaluation procedure and enable comparison of social Web applications at all levels of the quality model, we initiated a research into development of a methodology that will aggregate quality requirements into a single score. The work presented in this paper draws on the employment of the logic scoring of preference (LSP) method and outlines only some parts of the aforementioned methodology. After identifying quality attributes that constitute the requirement tree, elementary criteria for both objective and subjective performance variables were introduced. As a follow up, field experts were included in the study in order to determine weights of performance variables within particular performance subsystem. Finally, the appropriate logic aggregation operators were selected based on the relevance of performance variables.

Keywords—Social Web Applications, Quality Evaluation, Objective and Subjective Metrics, Preference Subsystems, Requirements Aggregation, Logic Scoring of Preference

1 Introduction

Social Web application is an umbrella term for recent generation of software breed that facilitates interaction among users and enables them to actively participate in design and promotion of various kinds of content and artefacts. Social networks, mashups, wikis, blogs, e-portfolios, virtual worlds, microblogs, podcasting applications, social bookmarking sites and online office suites are just some of the many examples of social Web applications [1]. Taking into account their particularities, social Web applications are widely employed in various aspects of human endeavor.

Quality represents the extent to which particular piece of software meets a set of predefined requirements [2]. It is an important feature of every social Web application that significantly affects its acceptance by end-users. Considering the dynamic life cycle of social Web applications, quality evaluation should be carried out fast and

often [3]. Currents advances in the field include variety of standards, models, methods, and techniques meant for examining facets of quality by means of subjective or objective measuring instruments of only a few types of social Web applications but lacks studies that propose an approach that would enable evaluation of both pragmatic and hedonic dimensions of quality in the context of all types of social Web applications and facilitate their comparison. Our aim is to develop a methodology that will fill this void in the software evaluation practice. Work presented in this paper is one of the essential steps towards this goal. The remainder of the paper is structured as follows. Theoretical background to our work is offered in the second section. Employed evaluation method is explained in the third section. Quality attributes that constitute the requirement tree are described in the fourth section. Performance variables are introduced in the fifth section. Elementary criteria for objective and subjective metrics are proposed in the sixth section. Aggregation of requirements in preference subsystems of a requirement tree is explained in the seventh section. Conclusions are drawn in the last section.

2 Related Work

Considering the subject of the evaluation, current approaches can be, in the context of Web engineering, divided into four groups. The first one are those meant for the assessment of Web sites. In that context, Palmer [4] proposed a set of Web site usability, design, and performance metrics including those related to download delay (initial access speed, speed of display between pages, Alexa speed), navigation/organization (arrangement, sequence, links, layout, Alexa organization), interactivity (customization, interactivity), responsiveness (feedback, FAQ), information/content (amount of information, variety of information, word count, content quality), and Web site success (satisfaction, likelihood of return, frequency of use). Mich et al. [5] introduced 2QCV3Q model which aims to evaluate Web site quality from both owner and user viewpoint through following dimensions: identity (identification, characterization), content (coverage, accuracy), services (functionalities, control), location (reachability, interactivity), management (currentness, maintenance), usability (accessibility, navigability, understandability), feasibility (resources, information and communication technology). Drawing on exhaustive literature review, Hasan and Abuelrub [6] proposed a framework for assessing quality of Web sites that consist of following dimensions: content quality (timely, relevant, multilanguage/culture, variety of presentation, accuracy, objective, authority), design quality (attractive, appropriateness, color, image/sound/video, text), organization quality (index, mapping, consistency, links, logo), and user-friendly quality (usability, reliability, interactive features, security/privacy, customization).

Next are approaches proposed for examining Web applications. In that respect, Oztekin et al. [7] proposed UWIS, a methodology for measuring usability of web-based information systems composed of following dimensions: efficiency, effectiveness, satisfaction, reliability, integration of communication, navigation, controllability, assurance, responsiveness, and quality of information. Zhao and Zhu [8] developed

WebQM, the web quality model focused of following quality aspects: Web source quality (availability, accessibility, durability, and timeliness of information), Web information quality (reliability, correctness, objectivity, understandability, and validity of information), and Web application-specific quality (relevance, presentation, information acquisition).

The third group are approaches for exploring quality of social Web applications in general. Majority of current studies have focused on assessing data quality of social Web applications. For instance, Han [9] measured data quality with respect to Web 2.0 applications with following characteristics: accuracy, completeness, consistency, credibility, currentness, accessibility, compliance, confidentiality, efficiency, precision, traceability, understandability, availability, portability, and recoverability. On the other hand, Pang et al. [10] presented a novel model that examines following facets of quality in the context of Web 2.0 applications: emotional quality (assurance, empathy, interaction, playfulness, and emotion), information quality (completeness, timeliness, comprehensibility, trustworthy, presentation variability, architecture, and search capability), interface quality (proximity, compatibility, navigation, appearance, and layout), service quality (customization, support, channel diversity, responsiveness, incentive, and compensation), and system quality (availability, efficiency, reliability, and security).

Last group refers to approaches introduced for evaluating specific types of social Web applications such as mashups. In that respect, Cappiello et al. [11] proposed a model which includes: data quality (accuracy, timeliness, completeness, availability, and consistency), presentation quality (usability and accessibility), and composition quality (added value, component suitability, component usage, consistency, and availability). As a follow up, Orehovački et al. [12] empirically examined model in which following facets of quality with respect to mashups were measured: system quality (efficiency, effectiveness, response time, and compatibility), service quality (availability, reliability, and feedback), content quality (accuracy, completeness, credibility, timeliness, and added value), composition quality (component suitability, composition added value, effectiveness of integrated visualization), effort (minimal memory load, accessibility, ease of use, learnability, and understandability), and user experience (usefulness, playfulness, satisfaction, and loyalty).

3 Logic Scoring of Preference

Logic scoring of preference [13] (LSP) is a quantitative evaluation method based on continuous preference logic [14] that represents generalization of decision-making techniques. The application of the LSP method consists of three steps [15]. The first one is focused on development of a requirement tree where quality attributes are decomposed to quality indicators and items that can be easily measured directly. In the context of LSP, quality indicators and items are called performance variables and are denoted as

$$X_1, X_2, \cdots, X_n \tag{1}$$

The values of performance variables are commonly real numbers [15]. The second step deals with defining an elementary criterion, a function

$$G_i : \left[X_1^{min}, X_i^{max} \right] \rightarrow [0,1] \qquad (2)$$

that for each value of performance variable X_i generates corresponding elementary preference

$$E_i = G_i(X_i), X_i^{min} \leq X_i \leq X_i^{max}, 0 \leq E_i \leq 1, i = 1, \dots, n \qquad (3)$$

The elementary criterion is the range of values that a particular performance variable can take and is expressed through a preference scale. The value of elementary preference is normalized between 0 and 1 (or 100%), and represents the degree of meeting the specific requirement defined in the tree. In that respect, $E_i = 0$ indicates that the performance variable X_i does not meet the predefined requirement, $E_i = 1$ represents the complete satisfaction of the requirement while the partial satisfaction of the requirement is in the range $0 < E_i < 1$. The last step in applying LSP method is aggregation of preferences, a process of calculating the global preference

$$E_0 = L(E_1, \dots, E_n) = L\big(G_1(X_1), \dots, G_n(X_n)\big) = G(X_1, \dots, X_n), 0 \leq E_0 \leq 100\% \quad (4)$$

that represents the extent to which all requirements in a tree are met. The aggregation function

$$L : [0,1]^n \rightarrow [0,1] \qquad (5)$$

is an iterative stepwise process that follows the structure of the requirement tree, going from the leaves towards the root [15]. Theoretically related preferences are aggregated with the appropriate logical operator which results in preference subsystems. The process is continued by aggregating preference subsystems until a single global preference E_0 is calculated. In each step of the aggregation process, it is necessary to select the appropriate logic aggregation operator and determine the relative significance of preferences. Logic aggregation operators are special cases of the generalized conjunction/disjunction function that is implemented by means of the weighted power mean [15].

LSP method has been used so far in evaluation of different types of Web sites. While majority of researchers (e.g. [16][17]) based their assessment approaches on objective metrics, only few of them (e.g. [18]) carried out subjective evaluation by employing a questionnaire as a measuring instrument. Yip and Mendes [19] have found that in the context of measuring usability of Web sites results of objective evaluation based on the employment of LSP significantly differ from subjective assessment. To overcome the aforementioned drawback, the objective of the comprehensive research this work is part of is to develop a method that would, drawing on both subjective and objective metrics, results in composite quantitative indicator with respect to the quality of social Web applications.

4 Requirement Tree

Quality model is "defined set of characteristics, and of relationships between them, which provides a framework for specifying quality requirements and evaluating quality" [20]. When the quality of social Web applications is considered, the quality model in the form of requirement tree we are proposing is composed of 35 quality attributes. They denote the extent to which particular social Web application: provides various navigation mechanisms (navigability), has uniform interface structure, design, and terminology (consistency), is similar to the previously used applications (familiarity), can be personalized to meet users' needs (customizability), has implemented mechanisms that protect created and stored artefacts from unauthorized use (security), operates properly on different devices and among various environments (compatibility), can exchange files with other applications and use files that were exchanged (interoperability), provides various forms of help to users (helpfulness), is continuously reachable (availability), facilitates handling of created artefacts (artefacts management), contains mechanisms that prevent errors to emerge (error prevention), is unfailing (reliability), can recover from errors and operational interruptions (recoverability), notifies users with appropriate and useful messages (feedback), supports teamwork and enables different types of communication among users (interactivity), enables users to execute tasks accurately and completely (effectiveness), encourages users to quickly perform tasks (efficiency), responds promptly to users' actions (response time), is capable to operate under an increased or expanding workload (scalability), is usable within and beyond initially intended contexts of use (context coverage), consumes small amount of physical and mental effort when employed (minimal workload), can be used by people with the widest range of characteristics and capabilities (accessibility), provides users a full freedom in executing tasks (controllability), is simple for operation (ease of use), enables users to easily become skilled in interaction with its functionalities (learnability), has easy-to-remember interface features (memorability), is clear and unambiguous (understandability), has visually appealing user interface (aesthetics), is beneficial in the context of executing tasks (usefulness), stimulates users' curiosity and creativity (playfulness), is perceived positively by users (attitude towards use), meets users' expectations (satisfaction), arouses users' emotional responses (pleasure), is distinctive among applications with the same purpose (uniqueness), and encourages users to employ it on regular basis and recommend it to others (loyalty).

5 Performance Variables

In the context of evaluating quality of social Web application, we distinguish two types of performance variables: indicators and items. Indicators are performance variables designed for collecting objective data related to estimated facets of quality. They are commonly used when methods such as logging actual use [21] are applied for evaluation purposes. There are six indicators which are relevant for our study. Task completion denotes the proportion of the number of scenario-based tasks partic-

ular user completed and the total number of tasks in the scenario. This indicator is designed for evaluating estimated effectiveness. The distance refers to the number of millimeters user travelled while moving the mouse during scenario-based tasks execution. The mouse clicks denote the sum of left, right, and middle mouse clicks for the purpose of performing scenario-based tasks. The mouse wheel scrolls represent the amount of scrolls the user made while reaching the solution of scenario-based tasks. The keystrokes indicate the total number of keys on the keyboard that the user pressed while completing scenario-based tasks. The set forth indicators are intended for measuring estimated workload by means of specialized software such as Mousotron [22]. The time is the amount of seconds required to complete the scenario-based tasks. This indicator is proposed for the assessment of estimated efficiency.

Items are performance variables in the form of questionnaire statements meant for measuring perceived dimensions of quality. Although literature offers a number of questionnaires meant for evaluating software in general (e.g. [23]) and web sites in particular (e.g. [24]), none of them capture all relevant particularities of social Web applications. In addition, they are commonly designed for examining pragmatic aspects of usability or hedonic dimensions of user experience while measuring instruments which combine both facets of quality are rather scarce. In that respect, we developed a post-use questionnaire that enables assessment of all relevant facets of perceived quality that constitute the requirement tree with between three and five items.

6   Elementary Criteria

6.1   Objective metrics

Determining elementary criteria for six indicators meant for objective assessment of three pragmatic dimensions (effectiveness, workload, and efficiency) of quality in the context of social Web applications was based on findings of our prior studies [25][26].

An elementary criterion for evaluating estimated users' effectiveness in performing predefined scenario steps of interaction with social Web applications designed for collaborative writing and mind mapping are shown in Figure 1. Values on the preference scales were obtained on the basis of the minimum and maximum number of scenario steps that the pilot study participants managed to complete during interaction with the aforementioned types of social Web applications.

The elementary criteria for the remaining five indicators were calculated in several steps. First, values obtained by means of specific indicator were divided by the number of scenario steps that individual participants in the pilot study managed to complete by employing particular social Web application. In that manner, indicator values per scenario step for every participant were calculated. Then, the yielded values were multiplied by the total number of scenario steps in order to calculate the values that the indicators would take if a specific participant succeeded in completing all scenario steps by means of specific social Web application. Finally, by applying the lower and upper interquartile to a sample of indicator values for social Web applications of the

same type which the pilot study participants used, minimum and maximum values on the indicators' preference scales were calculated, respectively.

Estimated effectiveness. Top preference scale presented in Figure 1 indicates that particular social Web application for collaborative writing completely meets requirement of estimated effectiveness ($E_i = 100\%$) if users during an interaction with it can complete all 45 scenario steps (X^{max}) and that the same requirement is not met ($E_i = 0\%$) if users can complete 17 or less scenario steps (X^{min}) in the same respect. On the other hand, bottom preference scale shown in Figure 1 denotes that if users can complete 43 scenario steps (X^{max}) when using social Web application for mind mapping, it completely ($E_i = 100\%$) satisfies requirement of estimated effectiveness but fails to meet this requirement ($E_i = 0\%$) if users can complete 13 or less scenario steps (X^{min}).
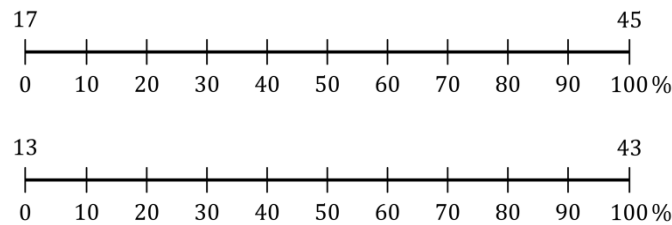

Fig. 1. Elementary criteria defined as preference scales for evaluating estimated users' effectiveness in completing tasks by means of social Web applications for collaborative writing (top) and mind mapping (bottom)

For both types of social Web applications, the elementary preference score related to estimated users' effectiveness can be calculated by means of following increasing function composed of three linear parts [14][15]:

$$E_i = G_i(X_i) =$$
$$\begin{cases} 0, & X \leq X^{min} \\ max\{0, min[1, (X - X^{min})/(X^{max} - X^{min})]\}, & X^{min} < X < X^{max} \\ 1, & X \geq X^{max} \end{cases} \quad (6)$$

Estimated workload. In the context of social Web application for collaborative writing, if users have pressed 591 keyboard keys or less when completing scenario steps, the number of keystrokes pressed as a requirement of estimated workload is considered to be met. On contrary, if the users pressed 1184 keyboard keys or more while performing scenario steps with social Web applications for collaborative writing, this requirement related to estimated workload is not satisfied. When social Web applications for mind mapping are taken into account, number of keyboard keys pressed equal to or less than 642 perfectly meets requirement of estimated workload related to the number of keystrokes pressed whereas number of keyboard keys pressed equal to or greater than 1011 violates the set forth requirement in the context of completing predefined scenario-based tasks.

When the distance as the second requirement of estimated workload is tackled, if users have traveled distance of 75961 millimeters or less while moving the mouse for the purpose of performing predefined scenario steps with social Web application designed for collaborative writing, it has perfectly met this elementary criterion. On the other hand, if users have traveled distance of 169756 millimeters or more in the same respect, social Web application failed to satisfy this requirement of estimated workload. When the aforementioned preference scale is considered in the context of social Web applications for mind mapping, distance traveled of 80834 millimeters or less represents complete fulfillment of the eponymous requirement of estimated workload while distance traveled of 160140 millimeters or more indicates non-compliance with this criterion.

While evaluating estimated workload in interaction with social Web applications for collaborative writing with number of mouse clicks required for completing scenario-based assignments, 484 mouse clicks or less imply perfect compliance with this criterion whereas 1038 mouse clicks or more represent violence of this requirement in the context of estimated workload assessment. When evaluating social Web applications for mind mapping in the same respect, 534 mouse clicks denote that this requirement of estimated workload is completely satisfied while 991 mouse clicks or more signify that this requirement of estimated workload has not been met.

Estimated efficiency. When examining estimated efficiency with respect to social Web applications for collaborative writing, time of 2506 seconds or shorter means perfect fulfillment of the eponymous requirement while time of 5900 seconds or longer uncovers non-compliance with this criterion. In the context of social Web applications for mind mapping, time of 2466 seconds or shorter implies complete satisfaction of this requirement of estimated efficiency while time of 5192 seconds or longer represents violence of this requirement.

Values of elementary preferences for the aforementioned five indicators (number of keystrokes, number of mouse clicks, distance traveled while moving the mouse, number of mouse wheel scrolls, and time required for completing the assignments), can be calculated with following decreasing function composed of three linear segments [14][15]:

$$E_i = G_i(X_i) =$$
$$\begin{cases} 1, & X \leq X^{min} \\ max\{0, min[1, (X^{max} - X)/(X^{max} - X^{min})]\}, & X^{min} < X < X^{max} \\ 0, & X \geq X^{max} \end{cases} \quad (7)$$

6.2 Subjective metrics

The elementary criterion for subjective quality evaluation by means of five-point Likert scale of five degrees (1 - completely agree, 5 - completely disagree) is shown in Figure 2.
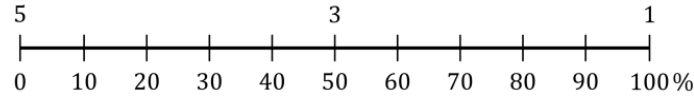
Fig. 2. Elementary criterion defined as preference scale for evaluating perceived quality of social Web applications

Unlike the objective elementary criteria that differ depending on the employed indicator, the subjective elementary criteria are the same for all quality evaluation items. Value 1 on the Likert scale indicates that particular subjective requirement is perfectly met, value 3 implies that 50% of the subjective requirement is satisfied, while value 5 denotes that subjective requirement has not been met. The remaining two values are calculated by linear interpolation. The elementary preference for subjective quality evaluation is calculated using the following function:

$$E_i = G_i(X_i) = \begin{cases} 1, & X \leq X^{min} \\ 0{,}25(5-X), & X^{min} < X < X^{max} \\ 0, & X \geq X^{max} \end{cases} \tag{8}$$

7 Preference Subsystems

The preference subsystems that constitute requirement tree can appear in two forms: quality attributes whose facets are directly measured by means of performance variables (items and/or indicators) and composite latent variables whose preferences are affected by performance variables assigned to them together with latent variables that, according to the conceptual model [3], represent their predictors. Composite quality index of particular preference subsystem indicates the extent to which evaluated social Web application meets quality requirements defined by this preference subsystem. In order to calculate a composite quality index with respect to social Web applications at the level of every preference subsystem, it is necessary to determine the appropriate logic aggregation operator as well as the relative importance of each performance variable within a specific subsystem of preferences it is assigned to. The relative importance of the performance variable is expressed in terms of weights determined by the judgment of a group of domain experts [27].

The process of identifying the appropriate logic aggregation operator consist of several steps. Firstly, it is necessary to determine whether the preference subsystem should consist of predominantly conjunctive, mostly disjunctive, or a combination of conjunctive and disjunctive requirements. Then, the type of relationship between the input preferences needs to be identified. If the relationship appears to be symmetric, type and intensity of the logic aggregation operator should be selected. When evaluation of software products is taken into account, sufficient preferences (disjunctive partial absorption) are much less common than mandatory preferences (conjunctive partial absorption) [14]. Since all performance variables were categorized by field experts as mandatory or desirable [3], the symmetric and asymmetric logic aggregation operators with conjunctive polarization are used when calculating the composite

quality indices of preference subsystems which form the requirement tree related to the quality of social Web applications. The manner of aggregating desirable requirements is illustrated in Figure 3. Since all three items proposed for evaluating familiarity of social Web applications were perceived as desirable by field experts, the arithmetic mean (A) was used to create a neutral logic polarization among them.
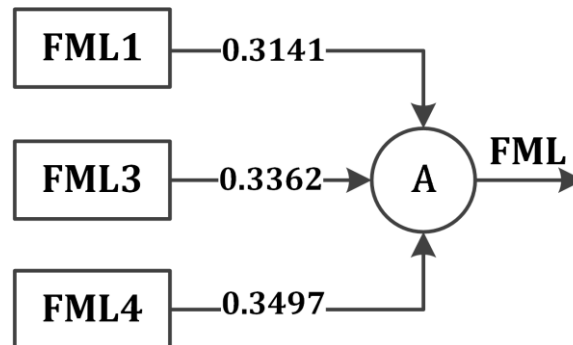

Fig. 3. Aggregating desirable requirements with the neutral logic polarization for quality attribute meant for measuring familiarity of social Web applications

Items designed for examining the reliability of social Web applications were all recognized by field experts as mandatory performance variables. As shown in Figure 4, to achieve the simultaneity of requirements, strong quasi-conjunction was employed for aggregation purposes.
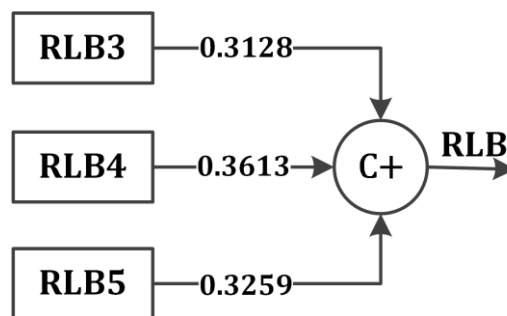

Fig. 4. Aggregating mandatory requirements with the conjunctive polarization for quality attribute meant for measuring reliability of social Web applications

In the context of the quality attribute introduced for evaluating the controllability of social Web applications, two items (CTR1 and CTR3) were identified as mandatory and one item (CTR2) was determined as desirable. In that respect, the conjunctive partial absorption (shown in Figure 5) was applied for aggregation purposes.
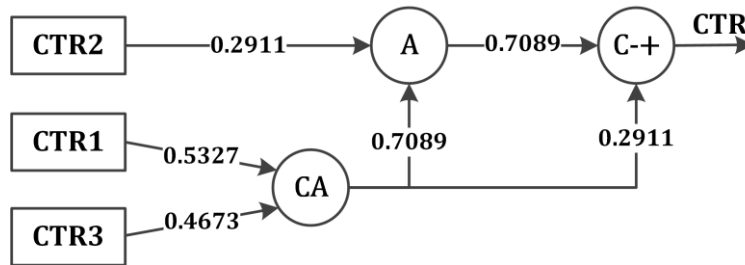
Fig. 5. Aggregating mandatory and desirable requirements with the conjunctive partial absorption for quality attribute meant for measuring controllability of social Web applications

8 Conclusion

The outcomes of the study presented in this paper provide several contributions to the extant body of knowledge in the software evaluation field. Based on the comprehensive literature review, the requirements tree with both objective and subjective performance variables related to the quality of social Web applications was proposed. As a follow up, several empirical studies were conducted in order to determine elementary criteria for performance variables. Field experts were then involved in the study in order to uncover weights of performance variables within particular performance subsystem. Finally, based on the relevance of performance variables, the appropriate logic aggregation operators were selected. All the aforementioned can be used by researchers as a backbone for future advances in the field while practitioners can use proposed performance variables for the purpose of evaluating social Web applications from both subjective and objective perspective. Although findings discussed in this paper provide important implications for evaluating quality of social Web applications, the limitation dealing with the generalizability of the elementary criteria needs to be acknowledged. Considering that reported elementary criteria refer only to social Web applications meant for collaborative writing and mind mapping, their reference values could be different for other types of social Web applications. Keeping that in mind, reported elementary criteria should be interpreted and used carefully. On the other hand, the manner of identifying the reference values of both objective and subjective metrics can be employed for all other types of social Web applications. In our future work, we are going to determine weights of preference subsystems and select the appropriate logical operators for their aggregation which would allow us to calculate a global preference that reflect an overall quality of social Web applications in the form of a single score and enable us to compare various social Web applications at different levels of granularity in the requirement tree.

9 Acknowledgement

The author would like to extend his sincerest gratitude and appreciation to The British University in Egypt (BUE) for the excellent mutual cooperation and for facilitating the publication of this article.

10 References

[1] Orehovački, T., Kovačić, A., Bubaš, G. (2012). Taxonomy of Web 2.0 Applications with Educational Potential. In Transformation in Teaching: Social Media Strategies in Higher Education, C. Cheal, J. Coughlin, S. Moore, Ed. Informing Science Press, Santa Rosa, pp. 43-72.

[2] ISO/IEC 25010:2011. Systems and software engineering - Systems and software quality requirements and evaluation (SQuaRE) - System and software quality models. https://doi.org/10.3403/30215101u

[3] Orehovački, T. (2013). Methodology for Evaluating the Quality in Use of Web 2.0 Applications, Ph.D. thesis. University of Zagreb, Faculty of Organization and Informatics

[4] Palmer, J.W. (2002). Web Site Usability, Design, and Performance Metrics. Information Systems Research 13 (2): 151-167. https://doi.org/10.1287/isre.13.2.151.88

[5] Mich, L., Franch, M., Gaio, L. (2003). Evaluating and Designing Web Site Quality. IEEE MultiMedia 10 (1): 34-43. https://doi.org/10.1109/mmul.2003.1167920

[6] Hasan L., Abuelrub, E. (2011). Assessing the quality of web sites. Applied Computing and Informatics 9 (1): 11-29. https://doi.org/10.1016/j.aci.2009.03.001

[7] Oztekin, A., Nikov, A., Zaim, S. (2009). UWIS: An assessment methodology for usability of web-based information systems. Journal of Systems and Software 82 (12): 2038-2050. https://doi.org/10.1016/j.jss.2009.06.047

[8] Zhao, B., Zhu, Y. (2014). Formalizing and validating the web quality model for web source quality evaluation. Expert Systems with Applications 41 (7): 3306-3312. https://doi.org/10.1016/j.eswa.2013.11.027

[9] Han, W-M. (2018). Evaluating perceived and estimated data quality for Web 2.0 applications: a gap analysis. Software Quality Journal 26 (2): 367-383. https://doi.org/10.1007/s11219-017-9365-7

[10] Pang, M., Suh, W., Hong, J., Kim, J., Lee, H. (2010). A New Web Site Quality Assessment Model for the Web 2.0 Era. In Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications, S. Murugesan, Ed. IGI Global, Hershey, pp. 387-410. https://doi.org/10.4018/978-1-60566-384-5.ch022

[11] Cappiello, C., Daniel, F., Koschmider, A., Matera, M., Picozzi, M. (2011). A Quality Model for Mashups. In Web Engineering, S. Auer, O. Díaz, G.A. Papadopoulos, Ed. ICWE 2011. Lecture Notes in Computer Science vol. 6757. Springer, Heidelberg, pp. 137-151. https://doi.org/10.1007/978-3-642-22233-7_10

[12] Orehovački T., Cappiello C., Matera M. (2016). Identifying Relevant Dimensions for the Quality of Web Mashups: An Empirical Study. In Human-Computer Interaction. Theory, Design, Development and Practice, M. Kurosu, Ed. HCI 2016. Lecture Notes in Computer Science vol. 9731. Springer, Cham, pp. 396-407. https://doi.org/10.1007/978-3-319-39510-4_37

[13] Dujmović, J.J. (1996). A Method for Evaluation and Selection of Complex Hardware and Software Systems. In Proceedings of the 22nd International Computer Measurement

Group Conference (San Diego, CA, USA, December 10-13, 1996) CMG 1996. Computer Measurement Group, pp. 368-378.

[14] Dujmović, J.J. (2007). Continuous Preference Logic for System Evaluation. IEEE Transactions on Fuzzy Systems 15 (6): 1082-1099. https://doi.org/10.1109/tfuzz.2007.902041

[15] Dujmović, J.J., Nagashima, H. (2006). LSP method and its use for evaluation of Java IDEs. International Journal of Approximate Reasoning 41 (1): 3-22. https://doi.org/10.1016/j.ijar.2005.06.006

[16] Buckley, G., Dujmović, J.J. (2008). Interfacing the System Evaluation Method LSP with E-commerce Web Sites. Computer Science and Information Systems 5 (1): 25-39. https://doi.org/10.2298/csis0801025b

[17] Olsina, L., Rossi, G. (2002). Measuring Web Application Quality with WebQEM. IEEE MultiMedia 9 (4): 20-29. https://doi.org/10.1109/mmul.2002.1041945

[18] Alva, M.E., Martínez, A.B., del Carmen Suárez, M., Labra, J.E., Cueva, J.M., Sagástegui, H. (2010). Towards the evaluation of usability in educative websites. International Journal of Technology Enhanced Learning 2 (1/2): 145-161. https://doi.org/10.1504/ijtel.2010.031265

[19] Yip, M.C.L., Mendes, E. (2005). Web Usability Measurement: Comparing Logic Scoring Preference to Subjective Assessment. In Web Engineering, D. Lowe, M. Gaedke, Ed. ICWE 2005. Lecture Notes in Computer Science, vol. 3579. Springer, Heidelberg, pp. 53-62. https://doi.org/10.1007/11531371_10

[20] ISO/IEC 25040:2011. Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Evaluation process. https://doi.org/10.3403/30205731u

[21] Nielsen, J. (1993). Usability Engineering. Morgan Kaufmann, San Francisco.

[22] Blacksun Software. Mousotron: Mouse and Keyboard Usage Measurement. URL= http://www.blacksunsoftware.com/mousotron.html

[23] Kirakowski, J., Corbett, M., (1993). SUMI: the software measurement inventory. British Journal of Educational Technology 24 (3): 210-212. https://doi.org/10.1111/j.1467-8535.1993.tb00076.x

[24] Aladwani, A.M., Palvia, P.C. (2002). Developing and validating an instrument for measuring user-perceived web quality. Information & Management 39 (6): 467–476. https://doi.org/10.1016/s0378-7206(01)00113-6

[25] Orehovački, T., Granić, A., Kermek, D. (2013). Evaluating the perceived and estimated quality in use of Web 2.0 applications. Journal of Systems and Software 86 (12): 3039-3059. https://doi.org/10.1016/j.jss.2013.05.071

[26] Orehovački, T., Babić, S. (2018). Identifying the Relevance of Quality Dimensions Contributing to Universal Access of Social Web Applications for Collaborative Writing on Mobile Devices: An Empirical Study. Universal Access in the Information Society 17 (3): 453-473. https://doi.org/10.1007/s10209-017-0555-7

[27] Orehovački, T. (2018). Employing Fuzzy AHP to Determine Relevance of Criteria Constituting the Quality of Social Web Applications. In Fuzzy Systems and Data Mining IV, A.J. Tallón-Ballesteros, L. Kaicheng, Ed. Frontiers in Artificial Intelligence and Applications, vol. 309. IOS Press, pp. 351-357. https://doi.org/10.3233/978-1-61499-927-0-351

11   Author

Tihomir Orehovački is an associate professor at the Faculty of Informatics, Juraj Dobrila University of Pula. He took part in 10 scientific projects that were financed

by European and Croatian funds. Tihomir was co-organizer of 9 workshops and webinars related to the international projects management and implementation of social Web applications in educational settings. He has been a reviewer for many international scientific conferences (e.g. IUI, CHI, CSCW, INTERACT, CECIIS, ITI, ISD, etc.) and journals (e.g. ACM Transactions on the Web, Behaviour & Information Technology, Computers & Education, Computer Speech & Language, IETE Technical Review, International Journal of Human-Computer Interaction, Information Sciences, JMIR mHealth and uHealth, Journal of Telemedicine and Telecare, Telematics and Informatics, The Journal of Systems and Software, Universal Access in the Information Society, etc.). Tihomir served as a Program Committee member of International Conference on Software and Information Engineering in 2019, ACM International Conference on Intelligent User Interfaces (2012-2015), International Conference of Information Systems Development (2014-2017), and International Conference on the Quality of Information and Communications Technology (track on Quality in Web Engineering) in 2014. He is an author of more than 100 papers published in conference proceedings and peer reviewed journals. Tihomir is a member of the following professional associations: Association for Computing Machinery, European University Information Systems E-Learning Task Force, Institute of Electrical and Electronics Engineers, and International Society for Web Engineering.