Web Engineering: A Quantitative Methodology for Quality Evaluation and Comparison of Web Applications

Dr. Luis Olsina[1] , *Dr.* Gustavo Rossi[2]
[1] Doctoral Thesis presented at Facultad de Ciencias Exactas,
Universidad Nacional de La Plata, Argentina. Date: 2000, April 28th
E-mail olsinal@ing.unlpam.edu.ar
[2] Thesis Director
E-mail grossi@info.unlp.edu.ar

Abstract. This research is aimed at defining and discussing a quantitative methodology, the one integral, robust and flexible, for the quality evaluation and comparison of Web sites and applications (WebApps). The approach, from now on called *Web-site Quality Evaluation Methodology* (Web-site QEM) is intended to perform an engineering contribution proposing for such an end a systematic, disciplined and quantitative strategy customized to the evaluation, comparison, and analysis of the quality of somewhat complex WebApps. We analyze their phases and activities, we describe the produced deliverables, and we present models, methods, procedures, principles and tools to apply in these activities. We outline the approach and its solutions, using along the thesis two case studies for WebApps, and other carried out field studies. In addition, we present a conceptual framework for metric validation and we discuss the theoretical validation of some utilized metrics in the evaluation process. Finally, we describe thoroughly an evaluation process model which Web-site QEM adheres. In this report, just the main methodology's steps are shown. Besides, some specific models, criteria and procedures to apply in these activities are discussed by quoting case studies already performed (e.g., in the academic domain [4] as well as in museums [3] and e-commerce domain, i.e., international e-bookstores [5]).

Introduction. The quick pace of the Internet infrastructure has marked a sudden growth, particularly, in Web-based products both document oriented and software application oriented. However, defined process models and methodologies that leverage the development and evaluation activities, mainly in medium and large-scale projects, have not been accompanied by that growth. One of the main goal to quantitative evaluations is understanding to what extent a given set of quality characteristics and attributes fulfills a set of requirements regarding specific audiences. For this aim, the proposed Web-site QEM methodology can be an useful tool in providing this understanding, in an objective, systematic and quantitative way as well as recommendations for improvements can be drawn.

The aim of performing the aforementioned case studies was the assessment of the level of fulfillment of required characteristic such as *Usability*, *Functionality*, *Reliability*, and *Efficiency* given a user viewpoint. In addition, elemental, partial and global quality indicators were compared. This allowed us to understand and draw conclusions about the state-of-the-art on the quality of domain-specific sites. On the one hand, the same quality model as that prescribed in ISO 9126 standard [2] was used mainly for the highest level characteristics. On the other hand, the model and procedures for attributes and characteristics aggregation and computation were based on a non-linear multi-attribute decision model [1]. However, in simpler cases where the amount of intervening characteristics and attributes are less than forty, a merely additive scoring model could be used.

Lastly, even though software evaluation is rather an old task in the software engineering field, the quantitative and systematic quality evaluation of WebApps has often been neglected.

An Overview of Web-Site QEM. In order to evaluate and compare the product quality (e.g., in the operational phase of a WebApps lifecycle), the major process steps that evaluators should perform are described:

- *The selection of a WebApps domain*
- *The specification of the evaluation goal and the user standpoint*
- *The definition of quality characteristics, sub-characteristics and attributes*
- *The definition of elementary quality metrics, criteria, and the determination of preferences*
- *The aggregation of elementary preferences to yield the global quality preference*
- *The analysis and conclusion of the evaluation process*

Step one. *The selection of a WebApps domain:* First, the evaluators should know the WebApps domain to evaluate or compare. For instance, it should be emphasized more a characteristic or sub-characteristic than others regarding the domain. On the other hand, to perform a case study, typical sites should be selected in order to be successful throughout the evaluation process (e.g., in the academic case study six typical sites were selected, and five to the e-bookstore study).

Step two. *The specification of goals and the user standpoint:* In these activities, the decision-makers should define the goals and scope of the evaluation process. The results can be useful to understand, control, or improve the quality of WebApps. The evaluators could assess a Web development project or an operational one; the quality of a component, a whole artifact, or compare quality preferences of selected artifacts. Also, can be

compared two version of sites, e.g., an old version and a newer one. On the other hand, the relative importance of characteristics varies depending on the different users and application domains. According to this, three views of quality are defined, namely: visitor, developer, and manager viewpoints. The visitor category can be decomposed, in turn, in two sub-categories: *general visitors* and *expert visitors*. The former represents casual or intentional audience maybe having a general interest and/or minimum domain knowledge. The latter represents a specialist or expert in the domain. For instance, general visitors were selected to the case studies.

Step three. *The definition of quality characteristics, sub-characteristics and attributes:* In this step, the evaluators should define, categorize, and specify the quality characteristics and attributes, grouping them into a requirement tree. In order to follow well-known standards the same conceptual characteristics as in the ISO 9126 standard were selected. From these characteristics, sub-characteristics are derived, and, in turn, measurable attributes are specified. For each attribute $A_{i,}$ a variable X_i is associated taking a real value, i.e., the measured value. For example, in the academic case study more than eighty measurable attributes were selected.

Step four. *The definition of elementary quality metrics, criteria and the determination of preferences:* In this task, the evaluators should define the basis for elementary evaluation criteria and perform the measurement process. Elementary evaluation criteria say how to evaluate quantifiable attributes. The result is a rating, which can be interpreted as the degree of satisfied requirement. For each variable X_i , it is necessary to establish an acceptable range of values and define a function, called the elementary criterion. This function is a mapping of the variable value (obtained from the empirical domain) into the new numerical domain, and called the elementary quality preference. The elementary quality preference EQ_i can be interpreted as the percentage of requirement satisfied by the value of X_i. In this sense, EQ_i = 0% denotes a totally unsatisfactory situation whereas EQ_i=100% represents a fully satisfactory situation. For each attribute, the measurement activity and the determination of the elementary preference should be carried out.

Step five. *The aggregation of elementary preferences to yield the global quality preference:* In this step, the evaluators obtain a quality indicator representing the global preference for each evaluated site. Applying a stepwise aggregation mechanism, the *n* elementary quality preferences can be grouped accordingly in order to allow computing the global quality preference. The global quality preference represents the global degree of satisfaction of all involved requirements. In the performed case studies, the Logic Scoring of Preference (LSP) model was used [1]. The strength of LSP resides in the power to model simultaneity, neutrality, replaceability, and other attribute and sub-characteristics relationships using logic aggregation operators and the weighted power mean function.

Step six. *The analysis and conclusion of the evaluation process:* In this final step, the evaluators assess the partial and total quantitative quality preferences regarding the stated goals and user standpoint. Thus, specific recommendations can be given to the requester.

Concluding Remarks. The evaluation process generates elemental, partial, and global indicators or quality preferences that can be easily analyzed, backward and forward traced, justified, and efficiently employed in decision-making activities. The rational utilization of Web-site QEM should help reducing the subjectivity in the process by providing a quantitative basis for quality assessment. It also provides a powerful conceptual framework and tool in order to understand and improve the quality of WebApps.

Currently, we are finishing an integrated collaborative environment to support the whole evaluation process called WebQEM_Tool. On the other hand, the methodology includes a step for the validation of metrics both theoretically and empirically. So, this activity is strengthening the research [6].

References

1. Dujmovic, J.J., 1996, *A Method for Evaluation and Selection of Complex Hardware and Software Systems*, The 22nd Int´l Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems. CMG 96 Proceedings, Vol. 1, pp. 368-378.
2. ISO/IEC 9126-1991 Int´l Standard, *Information technology - Software product evaluation - Quality characteristics and guidelines for their use.*
3. Olsina, L., 1999, *Web-site Quantitative Evaluation and Comparison: a Case Study on Museums*, Workshop on Software Engineering over the Internet, at Int'l Conference on Software Engineering (ICSE), Los Angeles, US. Electronic Proceeding at http://sern.cpsc.ucalgary.ca/~maurer/ICSE99WS/ICSE99WS.html.
4. Olsina, L., Godoy, D; Lafuente, G.J; Rossi, G.; 1999, *Assessing the Quality of Academic Websites: a Case Study*, In New Review of Hypermedia and Multimedia (NRHM) Journal, Taylor Graham Publishers, UK/USA, ISSN 1361-4568, Vol. 5, pp. 81-103.
5. Olsina, L.; Lafuente, G.J; Rossi, G.; 2000, *E-commerce Site Evaluation: a Case Study*. In Springer LNCS Nº 1875; 1st Int´l Conference on Electronic Commerce and Web Technology, UK, pp. 239-252.
6. Olsina, L.; Pons, C; Rossi, G.; 2000, *Towards Metrics and Model Validation in Web-site QEM* Proceedings of VI Congreso Argentino de Ciencias de la Computación. (CACIC), Ushuaia, Arg.