

E-commerce Site Evaluation: a Case Study

Luis Olsina¹², Guillermo Lafuente¹, Gustavo Rossi²³

- ¹ GIDIS, Department of Computer Science, Faculty of Engineering, at UNLPam, Calle 9 y 110, (6360) General Pico, LP, Argentina
E-mail [olsinal, lafuentel]@ing.unlpam.edu.ar
- ² LIFIA, Informatics School at UNLP, Calle 50 y 115, (1900) La Plata, Argentina
- ³ CONICET - Argentina
E-mail grossi@info.unlp.edu.ar

Abstract. Recent surveys on e-commerce sites confirm the increased use of the Web for shopping. The reasons of these trends have been attributed to different factors such as convenience, saving time, absence of sales pressure, among others. However, an essential site characteristic that should also be taken into account is the quality in use. In this paper, we present a case study on five e-bookstores in order to assess characteristics and attributes that influence quality in use utilizing for such an end the Web-site Quality Evaluation Methodology (QEM). The main goal of this work is to show the level of accomplishment of required quality characteristic like usability, functionality, reliability, efficiency and derived sub-characteristics and attributes, regarding the user standpoint. In addition, we focus on sub-characteristics and attributes concerning e-commerce site functionality.

1 Introduction

E-commerce (EC) is nowadays a growing reality in the Web. Interestingly, recent field studies on EC sites confirm the increased use of the Web for shopping. For instance, the tenth edition of GVU's WWW User Survey [4] reports that both personal and professional shopping are up by significant percentages (10.3% and 23.7%, respectively) since the ninth survey. The major reasons for using the Web for personal shopping was convenience factor (21% of reason given), followed by saving time (18.8%), the presence of vendor information (18.7%), and the absence of sales pressure (16%).

Besides, surveys and evaluations on EC sites' attributes and factors have also drawn initial and interesting outcomes. For example, to cite only some, Lohse and Spiller [6] measured 32 attributes on 28 Web stores that influence store traffic and sales. The authors identified six principal categories for a Web store, namely: Merchandise, Service, Convenience, Store Navigation, Promotion, and Checkout, regarding a previous classification [1]. In addition, Tilson et al [11] surveyed factors and principles affecting the decision to purchase by analyzing four sites; the 16 participants rated the importance of fifty factors on a scale from 1 to 7.

However, considering the evaluation viewpoint, an essential site (product) issue that should be emphasized is the characteristics and attributes that influence the

quality in use. Process quality contributes to improving product quality, and product quality contributes to improving quality in use. Thus, one of the main goals to website evaluation and comparison is to understand the extent which a given set of product characteristics and attributes fulfills a set of explicit and implicit needs in consideration of specific audiences. In this direction, the proposed Web-site QEM methodology [10] can be an useful approach in providing this understanding, in a rather objective and quantitative way. In addition, by analyzing the intermediate and final product quality indicators recommendations can be suggested. That is, the evaluation process generates elemental, partial, and global quality indicators (or preferences) that can be easily backward and forward traced, justified, and efficiently employed in recommendations tasks.

Therefore, the aim of this paper is to show the level of accomplishment of selected quality characteristic like *Usability, Functionality, Reliability, Efficiency* and derived sub-characteristics and attributes in five e-bookstores, from a general user standpoint. With regard to quality characteristics and attributes for assessment purposes, over a hundred and forty ones were taken into account, considering attributes directly or indirectly quantifiable. Furthermore, we build the quality requirement tree starting from the same high-level quality characteristics as those prescribed in ISO/IEC 9126 standard [5], -a new version of this standard is in draft state.

In this work, sub-characteristics and attributes concerning EC functionality and a recursive decomposition mechanism are also discussed. For instance, the *Functionality* characteristic is split up in *Searching and Retrieving, Navigation and Browsing*, and *EC Functionality and Content* sub-characteristics. Regarding the latter, attributes for *Product Information, Purchase, Customer* and *Store Features*, and *Promotion Policies* sub-characteristics were in turn specified.

The structure of this paper is as follows: in the next section, we put the e-bookstores case study in context, and in Section 3, we present the evaluation process regarding Web-site QEM's main steps. In sub-section 3.1, the decomposition mechanism and the EC functionality sub-characteristic are analyzed. In the two following sub-sections the elementary and global criteria and methods are outlined. In sub-section 3.4, we discuss and compare partial and global outcomes. Finally, in Section 4, concluding remarks are considered. An Appendix depicts the whole requirement tree employed in this case study.

2 The E-Commerce Case Study on Bookstores

In order to prepare the study on e-bookstores domain five established sites were selected (the reader can also refers to the case studies carried out on museums and academic sites domains [8, 9]). Likewise, the chosen sites should be typical and well known regionally and/or internationally as well as they should specifically allow domestic and international shopping. One of the primary goals of this study is the understanding and comparison of the current level of fulfillment of essential quality characteristics and attributes given a set of requirements with regard to a general audience. Particularly, we assessed the level of accomplishment of standardized characteristics as previously mentioned (and derived sub-characteristics and attributes), and compare partial and global preferences in order to analyze and draw

conclusions about the state-of-the-art of e-bookstores quality. Important conclusions (mainly, domain specific as well as general ones) can emerge as it will be seen, later.

We selected the following e-bookstore sites, namely: Amazon (US, <http://www.amazon.com>), Cúspide (Argentina, <http://www.cuspide.com.ar>), Barnes and Noble (US, <http://www.bn.com>), Díaz de Santos (Spain, <http://www.diazdesantos.es>), and Borders (US, <http://www.borders.com>). The data collection activity was performed from the 15th Sep. to the 20th Oct., 1999. In this period, Díaz de Santos site changed its look & feel, so we evaluate the new interface.

Speaking in a wider sense, software artifacts are generally produced to satisfy specific user's needs, and Web artifacts are not the exception. In designing e-bookstore sites, there are many challenges that should not be neglected. For instance, when users enter the first time at a given home page they may want to find a piece of information quickly. There are two ways to help them in doing that: browsing or searching. Moreover, to get a time-effective mental model of the overall site, i.e., its structure and content, there are attributes like a table of contents or an index (a catalog) that help in getting a quick global site understandability, facilitating browsing. On the other hand, a searching function (quick and advanced ones), provided in the main page and remaining permanent in all pages can effectively help retrieving the desired piece of information and avoid browsing. So, the two functions can be complemented. In addition, first time visitor help (or a guided tour) and general information attributes can allow quick understanding about content information, common procedures, and policies. Fig. 1, shows a screenshot of a home page and highlights some attributes. There are a lot of such attributes both general and domain specific that contribute to quality, so designers should take into account when building for intended audiences (for a wider specification, see the list in Appendix A).



Fig. 1. Partial view of Cúspide's home page, where some attributes are highlighted. These are generally available in e-bookstore sites.

3 The Evaluation Process

The proposed stepwise, expert-driven methodology, Web-site QEM, is essentially quantitative, flexible and robust, and it covers most of the activities in the evaluation, comparison, and ranking process of websites. These steps are grouped in the following major technical phases:

1. *Quality Requirement Definition and Specification,*
2. *Elementary Evaluation: Definition and Implementation,*
3. *Global Evaluation: Definition and Implementation,*
4. *Analysis, Conclusions and Documentation*

As said above, these phases are composed by activities. Even if we do not describe and discuss thoroughly the steps here (for a broader description, see [9]), the next fourth sub-sections deal with some activities, strategies, models, methods, and tools employed in the EC study regarding the previous listed stages.

3.1 The E-bookstore Requirement Tree

In the first phase, the evaluators should clarify the evaluation goals and user standpoint. They should select the sites to assess and compare, and, finally, agree and specify the quality requirement tree. In the latter step, the ISO-prescribed characteristics give evaluators a conceptual model for quality requirements and provide a baseline for ulterior decomposition. Thus, a quality characteristic can be decomposed in a list of sub-characteristics recursively, and so, a sub-characteristic can be refined in a set of direct or indirect measurable attributes. Therefore, regarding the domain, the specific goals, and the user standpoint (i.e., the user implicit and explicit needs), quality characteristics and attributes should be specified in a requirement tree. The whole-intervened characteristics and measurable attributes for this study are outlined in Appendix A. Next, we discuss some characteristics and attributes and the decomposition mechanism.

The *Usability* high-level characteristic is decomposed in sub-characteristics such as *Global Site Understandability*, *Feedback and Help Features*, and *Interface and Aesthetic Features*. The *Functionality* characteristic is split up in *Searching and Retrieving Issues*, *Navigation and Browsing Issues*, and *EC Functionality and Content* sub-characteristics. The same decomposition mechanism is applied to *Reliability and Efficiency* factors.

Focusing on *EC Functionality and Content* characteristic, we have grouped five main components: *Product Information* (2.3.1 coded), *Purchase Features* (2.3.2), *Customer Features* (2.3.3), *Store Features* (2.3.4), and *Promotion Policies* (2.3.5). Regarding *Product Information* sub-characteristic, we see at the same level *Product Description*, *Price Evaluation*, and 2.3.1.3, 2.3.1.4, 2.3.1.5 attributes. For example, the *Product Description* (2.3.1.1 coded), is compounded by a *Basic Description* attribute, *Book Content & Structure*, and *Product Image* sub-characteristics. The *Basic Description* attribute takes into account the availability of information such as book title, author, edition, format, pages, size/weight, ISBN, price and availability (e.g., we can use a discrete multi-variable criterion in order to evaluate this attribute). On the other hand, the *Price Comparison Availability* (2.3.1.2.1) mechanism allows

users to compare prices in different stores. According to the recent GVV's WWW User Survey [4] it informs that for professional shopping *Detailed Information* was cited as important when selecting a product or service on the Web for 30.6% of responses out of a choice of 6 possible ones; *Information about Availability* was cited as the next most important (26.7%), followed by ability to make *Price Comparisons* (26.5%) by the user. Unfortunately, no evaluated site had this desirable attribute.

Focusing on *Purchase Features* (2.3.2 coded), we can observe two main sub-characteristics: *Purchase Mode* (2.3.2.1), and *Purchase Policies* (2.3.2.2). Regarding the *Purchase Mode* sub-factor, online and offline modalities are feasible, however, the former is becoming more preferred as long as confidence in security is increasing. For purchase online, the *Shopping Basket*, *Quick Purchase*, and *Checkout* features are modeled. The shopping basket mechanism (also known as a design pattern in [12]) is generally used to decouple the selection process from the checkout process of products or services. For example, we evaluate the availability, the continue buying feedback, and the edit/recalculate feature of the *Shopping Basket* sub-characteristic. On the other hand, the *Purchase Policy* (2.3.2.2) sub-factor, should contain a sound and concise information for potential and current customers, as cancellation and return policies information, shipping and handling costs (separating domestic from international information), payment policy information, and, optionally, a resent purchase or gift service facility.

As previously said, there are a lot of such attributes that contribute to site quality in use that designers should consider when building for some domain and intended audiences. Finally, focusing on *Feedback and Help Features* characteristic (where Usability is the super-characteristic), we have grouped five main components: two of them are *Link-based Feedback* (1.2.3) and *Form-based Feedback* (1.2.4). *FAQ* and *What's New Feature* are attributes of 1.2.3 sub-characteristic; *Questionnaire Feature*, *Comments/Suggestions* and *Subject-Oriented Feedback* (like in Barnes & Noble and Borders sites) are attributes of 1.2.4 sub-characteristic. They partially contribute to the communicativeness and learning process.

3.2 The Elementary Evaluation

In the second phase, the evaluators should define the basis for elementary evaluation criterion (for each attribute), and perform the measurement and mapping process. For each direct or indirect attribute A_i , we can associate a variable X_i , which can take a real measured or calculated value. Besides, for each variable it is necessary to establish a scale and unit [3], and set a criterion function, called the elementary criterion function. This function models a mapping among the measured or calculated value to the value of the new numerical representation, resulting afterwards in an elementary quality preference. The elementary preference EQ_i is frequently interpreted as the percentage of satisfied requirements for a given attribute, and it is defined in the range between 0, and 100% [2]. In this way, the scale and unit become normalized.

There are two kinds of general criteria: absolute and relative. Regarding absolute criteria, a basic taxonomy decomposes preference variables in continuous and discrete. Continuous variable criteria are classified in direct preference assessment,

single, normalized and multi-variable criteria. Discrete variable criteria are classified in binary, multi-level, subset-structured, and multi-variable criteria. In the following specification two types of criteria are shown. On the other hand, to ease the preferences interpretation, we group them in three categories or acceptability levels, namely: unsatisfactory (from 0 to 40%), marginal (from 40 to 60%), and satisfactory (from 60 to 100%) -or red, gray, and green ranges respectively. Next, for the e-bookstore study two attributes are documented following a hierarchical and descriptive specification framework as in previous case studies.

Title: Customized Recommendations; **Code:** 2.3.3.2; **Type:** Attribute
Higher-level characteristic: Functionality ; **Super-characteristic:** Customer Features
Definition / Comments: It is a mechanism that provides customers with personalized recommendations of products according previous buying preferences. These recommendations are automatically generated each time he/she enters to the e-store.
Elementary Criteria: It is an absolute and discrete binary criterion. We only ask if it is available (1), or not (0).
Preference Scale: see Fig. 2; **Data Collection Type:** Manual, Observational
Example/s: Only Amazon site provided this facility.

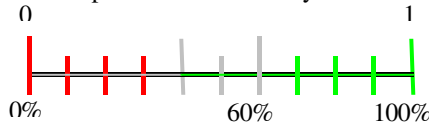


Fig. 2. Customized Recommendations attribute represented by a preference scale.

Title: Broken Links; **Code:** 3.1.1.1; **Type:** Attribute
Higher-level characteristic: Reliability; **Super-characteristic:** Link Errors
Definition/Comments: It represents found links that lead to missing destination nodes (also called dangling links). "*Users get irritated when they attempt to go somewhere, only to get their reward snatched away at the last moment by a 404 or other incomprehensible error message*" [7].
Elementary Criteria: It is an absolute and continuous normalized-variable criterion, where if BL = number of broken links found, and TL = number of total site links, then, the formula to compute the variable is: $X = 100 - (BL * 100/TL) * 10$; where, if $X < 0$ then $X = 0$;
Preference Scale: see Fig. 3; **Data Collection Type:** Automated.
Example/s: Cúspide yielded an elementary preference of 99.83 %, Amazon, 98.40 %, and Barnes and Noble site 97.45 %.

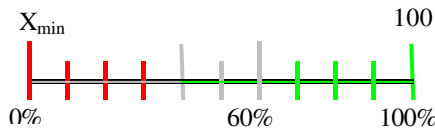


Fig. 3. Broken Links attribute represented by a preference scale.

Ultimately, once all elementary criteria were defined and data collected, the elementary preferences can be yielded and documented. The final hyperdocument links the requirement tree with the specification cards and elementary outcomes. Table 1, shows the preferences of some attributes for each e-bookstore site.

Table 1. Partial outcomes of elementary preferences for the five e-bookstores.

	Amazon	Barnes & Noble	Cúspide	Díaz de Santos	Borders
1. Usability					
1.2.3.1	100	100	100	0	100
1.2.3.2	100	100	80	80	100
1.2.4.1	0	0	0	100	0
1.2.4.2	0	100	0	70	70
1.2.4.3	0	100	0	0	70
2. Functionality					
2.1.1.1	100	100	100	0	100
2.1.1.2	100	100	100	100	70
2.3.1.1.1	100	100	40	24	70
2.3.1.1.2.1	100	100	50	50	0
2.3.1.1.2.2	42	42	12	12	21
2.3.1.2.1	0	0	0	0	0
2.3.2.1.1.1.1	100	100	100	100	100
2.3.2.1.1.1.2	100	70	0	0	40
2.3.2.1.1.1.3	100	100	100	100	100
2.3.2.1.1.2	100	0	0	0	100
2.3.2.1.1.3.1	100	100	100	0	100
2.3.2.1.1.3.2	100	100	100	0	0
3. Reliability					
3.1.1.1	98.4	97.45	99.83	60.07	76.34
4. Efficiency					
4.1.1	100	77.75	100	99.75	99.70

Even if they are only elementary values where no aggregation mechanism and computation were yet applied, some observations can be done all the same. For instance, we see the five sites having *What's New* (1.2.3.2) attribute resolved (the greater preference arises if, to the newest additions -generally in the home page, is also included a summary and links to the information objects). Nevertheless, in the book *Basic Description* (2.3.1.1.1) attribute, two sites fall in the red range. The elementary result is worse in the book *Content Description* (2.3.1.1.2.2), as the reader can observe. We evaluate the availability and richness of content description. The greater richness is for Amazon, B&N, and Borders, but the availability is varying. Notice that a review of a product is considered in the *Customer Revision of a Book* (2.3.3.4) attribute.

Lastly, focusing on the *Shopping Basket* sub-characteristic, we observe the *Shopping Basket Availability* (2.3.2.1.1.1.1) in all selected sites. However, the *Continue Buying Feedback* attribute is absent in Cúspide and Díaz de Santos, and unsatisfactory (40%) in Borders site. The *Quick Purchase* (2.3.2.1.1.2) attribute is implemented efficiently only in Amazon and Borders (it is absent in the other sites). Unfortunately, Díaz de Santos site has no *Checkout Security* attribute, and the *Checkout Canceling Feedback* (2.3.2.1.1.3.2) one is only explicitly considered in Amazon, B&N, and Cúspide sites, in all pages at checking out time.

3.3 The Global Evaluation

In order to obtain a global quality indicator for each website, the evaluators should define, in the third phase, the aggregation process and implement it. In the process, the type of relationships among attributes, sub-characteristics, and characteristics and the relative weights might be considered. For this purpose, it was agreed as in previous case studies the use of a robust and sensible model such as the Logic Scoring of Preference (LSP) model (regarding the amount of intervening characteristics and attributes, i.e., over a hundred and forty ones). LSP is based in a weighted power means mathematical model [2]. However, in simpler cases a merely additive scoring model can be used where indicators can be computed using the following structure: *Global/Partial Indicator* = (*Component Weight* * *Elementary Indicator*). The strength of LSP model over merely additives ones resides in the power to deal with different logical relationships and operators to reflect the evaluation needs. The basic relationships modeled are: 1) Replaceability, when it is perceived that two or more input preferences can be alternated; 2) Simultaneity, when it is perceived that two or more input must be present simultaneously; 3) Neutrality, when it is perceived that two or more input preferences can be grouped independently (neither conjunctive nor disjunctive relationships).

On the other hand, the major LSP operators are the arithmetic means (A) that models neutrality relationship, and weak (C-), medium (CA), and strong (C+) quasi-conjunction functions, that model simultaneity relationships. In addition, we can tune these operators to intermediate values, e.g., C-- is positioned between A and C- operators, and C+ is between CA and C- operators, and so on. The above operators (except A) mean that, given a low quality of an input preference can never be well compensated by a high quality of some other inputs to output a high quality preference. Similarly to conjunctive operators, we can also use the quasi-disjunction operators that model replaceability relationships. That is, a low quality of an input preference can always be compensated by a high quality of some other input.

Regarding the aggregation process, it follows the hierarchical structure of the requirement tree, from bottom to top. Applying a stepwise aggregation mechanism, the elementary preferences can be partially structured; in turn, repeating the aggregation process at the end can be obtained a global schema. This aggregation model allows computing partial and global preferences. For example, Fig. 4a) depicts a partial aggregation structure for the *EC Functionality and Content* (2.3) sub-characteristic. The 2.3 output (0.5 weighted), produced by means of C-- operator, together with 2.1 and 2.2 outputs, are inputs to the C- operator in order to compute the *Functionality* (2) preference (do not shown in the figure). The rectangles on the left side of the figure contain the coded elementary preferences. For instance, the values of 2.3.1.1.2.1 and 2.3.1.1.2.2 preferences yield the 2.3.1.1.2 output. The C-- operator does not model mandatory requirements, i.e., a zero in one input does not produces a zero at the output even though it punishes the outcome. The interpretation is we need a *Book's Content Description* and *Table of Contents* simultaneously (however, in this case, an "and" weak relationship is used). Finally, Fig. 4b) shows the final aggregation of the characteristics coded as 1, 2, 3, and 4 respectively, to produce the global preference. The global quality preference represents the total degree of satisfaction of quality requirements.

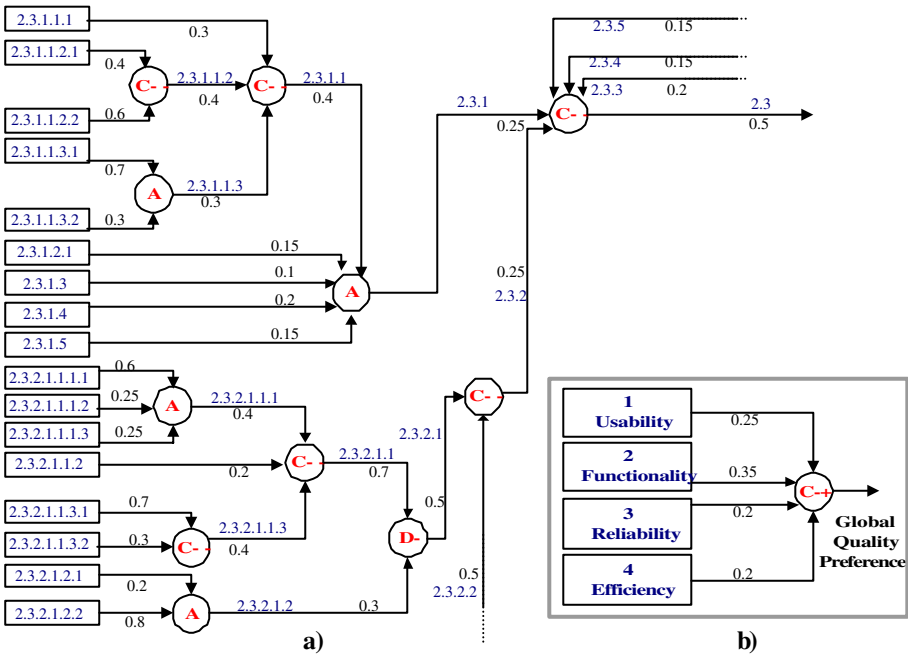


Fig. 4. Aggregation of preferences by using the LSP Model. a) Depicts a partial aggregation structure for the 2.3 sub-characteristic; b) Shows the final aggregation structure.

3.4 The Analysis of E-bookstore Sites

In the fourth phase, the evaluators analyze and compare the elementary, partial and global outcomes regarding the goals and user standpoint. Partial and final values dumped, for instance, in tables 1, and 2, and schemas as in Fig. 4, are useful sources of information to analyze, justify, and draw conclusions about the quality of e-bookstore sites. Besides, Fig. 5 b) represents the final ranking.

The colored quality bars at the left side of the Fig. 5b), indicate the levels of acceptability, as previously commented. For instance, a scoring within a gray bar can be interpreted as though improvement actions should be considered (this is the case for the global preference of Díaz de Santos site), as long as an unsatisfactory rating level can be interpreted as though necessary and urgent change actions must be taken. A scoring within a green bar can be interpreted as a satisfactory quality of the artifact as a whole. However, partial results for each high-level characteristic or sub-characteristic could indicate some kind of improvement.

The global quality preference was satisfactory for Amazon (86.81%), B&N (82.95%), Cúspide (75.52%), and Borders (74.86%) sites (see Fig. 5b). Díaz de Santos site should plan improvement actions due to the marginal score (50.37%). In the following paragraphs, we focus the discussion mainly on the *Functionality* factor.

Table 2. Some results of partial and global preferences for each e-bookstore

Characteristic and Sub-characteristic	Amazon	B&N	Cúspide	Díaz Stos	Borders
1. Usability	76.16	82.62	75.93	56.09	72.87
1.1 Global Site Understandability	74.99	70.58	70.58	53.46	66.22
1.2 Feedback and Help Features	74.99	78.91	67	35.72	74.23
1.3 Interface and Aesthetic Features	98.49	91.14	91.14	88.24	77.55
2. Functionality	83.15	80.12	61.69	28.64	61.45
2.1 Searching and Retrieving Issues	100	100	91	42.67	72.06
2.2 Navigation and Browsing Issues	99.99	69.85	73.25	64.12	51.95
2.3 E-commerce Functionality Content	81.89	76.53	45.81	14.42	61.55
2.3.1 Product Information	63.72	42.20	40.64	10.20	15.98
2.3.2 Purchase Features	91.76	84.84	67.72	17.11	81.92
2.3.2.1 Purchase Mode	83.80	70.64	75.80	14.40	65.25
2.3.2.1.1 On-line	100	66.21	62.57	17.18	77.87
2.3.2.2 Purchase Policies	100	100	60	20	100
2.3.3 Customer Features	100	85	20	28.08	65
2.3.4 Store Features	100	96.80	71.20	33.60	93.57
2.3.5 Promotion Policies	60	100	40	0	100
3. Reliability	99.44	99.11	90.97	78.51	91.66
4. Efficiency	96.88	74.54	90.17	86.01	90.90
4.1 Performance	100	77.75	100	99.75	99.70
4.2 Accessibility	89.74	67.26	68.79	57.08	71.60
Global Quality Preference	86.81	82.95	75.52	50.37	74.86

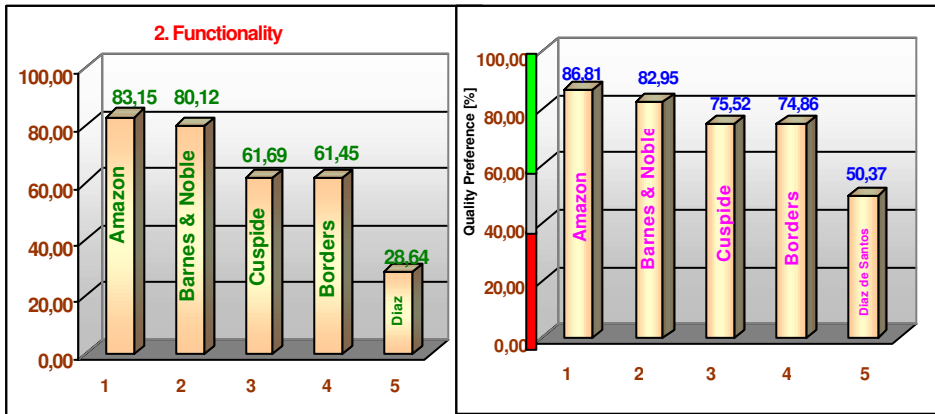


Fig. 5. Scores and rankings for the five e-bookstores: a) to the Functionality characteristic; b) to the final Quality

Regarding the *Functionality* characteristic, the highest score was to Amazon site (83.15%), and the lowest to Díaz de Santos (28.64%); i.e., the quality is unsatisfactory for this factor (see Fig 5a). For instance, Amazon and B&N sites have reached the outstanding score of 100% in *Searching and Retrieving* sub-characteristic (e.g., in-site quick and advanced searching, and level of retrieving customization and feedback attributes). Nonetheless, Díaz de Santos site has no quick search, and the retrieving customization and feedback mechanisms are weak (drawing a partial preference of 42.67% -see Table 2). On the other hand, the bigger differences among sites are

observed in the *EC Functionality and Content* (2.3) sub-characteristic. Only Amazon, B&N, and Borders sites show a satisfactory score of 81.99, 76.53, and 61.55% respectively. The quality for Cúspide is below the satisfactory level (45.81%), and Díaz de Santos site is very poor (14.42%).

The strength of Web-site QEM methodology is that by generating elemental, partial, and global quality indicators the evaluation process can be easily analyzed, traced, and justified, as well recommendations can be made. For example, comparing the strongest and the weakest sites, we observe 63.72% of the preference to Amazon, and 10.20% to Díaz de Santos, considering *Product Information* sub-characteristic. Looking for the causes in Table 1, the book *Basic Description* (2.3.1.1.1) score was 100% to the former and 24% to the latter. Besides, Díaz de Santos has lower scores in 2.3.1.1.2.1, 2.3.1.1.2.2, and 2.3.1.1.3.1 attributes. In addition, it has neither *Related Product/Author Recommendations*, nor *Product Rating Availability*. On the other hand, looking at the schema in Fig. 4a), the 2.3.1.1 output (Product Description) is produced by means of a C-- operator, which punishes low input preferences, as previously commented.

Likewise, in the *Online Purchase Mode* (2.3.2.1.1), we observe tremendous differences among both sites. Unfortunately, Díaz de Santos site has available neither *Quick Purchase*, nor *Checkout Security*, nor *Checkout Canceling Feedback* among other needed or desirable attributes. Conversely, Amazon site has an excellent *Online Purchase* preference (100%). Besides, in Table 2, the reader can observe and compare the partial preferences for *Customer Features* (2.3.3), *Store Features* (2.3.4), and *Promotion Policies* (2.3.5) sub-characteristics, among intervening e-bookstores. Analyzing each attribute that composes these sub-characteristics, recommendations to improvement can be made, when necessary.

Finally, considering the *Efficiency* characteristic, the highest score was to Amazon (96.88%), and the lowest was to B&N (74.54%), however, all scores have fallen in the satisfactory range. The reader can observe that four out of five sites have surpassed the 99% in the *Performance* sub-characteristic. Here, we measure the *Quick Access Pages* attribute [9]. With regard to the *Accessibility* sub-characteristic, the highest preference was again to Amazon (89.74%). Unfortunately, one out of five sites have implemented the *Text-only Version* (4.2.1.1) attribute. This capability would be necessary because users would have total information accessibility on pages, mainly for people with disabilities, or when speed is a drawback [13].

4 Final Remarks

As commented in the Introduction Section, there are some factors that influence quality in use, particularly, internal and external product attributes. Nowadays, it is widely recognized that process quality contributes to improving product quality, and product quality contributes to improving quality in use. However, the evaluation of the Web as a product has often been neglected in the web engineering community. Hence, to understand, control, and improve the quality of Web-based products we should increasingly use software engineering methods, models, techniques, and tools. In this direction, the proposed expert-driven Web-site QEM methodology, can be a

useful approach to assess the quality in the different phases of a Web product lifecycle. As frequently said, if we don't know where we are standing, a map won't help.

Specifically, in this work, we have illustrated the Web-site QEM steps, by analyzing the state-of-the-art of the quality of five e-bookstores sites. The case study has resulted that the state-of-the-art of the quality of typical e-bookstores sites, from a general user standpoint, is rather high (particularly, for four out of five evaluated sites, which surpass the lower satisfactory limit of 60%). The sites satisfied globally a range from 50% and 87% of the specified quality requirements. However, the partial and final results show that the wish list because of poor-designed or absent attributes, is not empty. Each assessed site has at least a sub-characteristic in the red and/or gray quality bar, where improvement actions should be considered. Hence, engineers and designers have a powerful and flexible tool to redirect their efforts on the weaker characteristics and attributes of the site. Moreover, the strengths and weaknesses can be well understood, so, in this case, a map would help.

Finally, the proposed quantitative methodology can be used to assess diverse application domains regarding different user viewpoints and evaluation project goals. It should be noticed though, that the definition and the specification of quality requirements are key activities in the evaluation process. For instance, from an EC site developer viewpoint, some other internal and external attributes should be considered; besides, the ISO-prescribed characteristics like *Portability* and *Maintainability* should be taken into account. Lastly, according to our recent studies, many attributes and sub-characteristics can be reused among different website domains, if we consider a particular kind of user. Some others are unavoidably domains specific. On the other hand, the precision of the assessment process can be adjusted selecting the appropriate metrics and, ultimately, the appropriate evaluation criteria and procedures. Moreover, the subjectivity can substantially be minimized in the evaluation process but can not be eliminated at all.

References

1. Arnold, S.; Ma, S; Tigert, D.: A comparative analysis of determinant attributes in retail store selection, *Advances in Computer Research*. Assoc. for Consumer Research (5), (1977) 663-667.
2. Dujmovic, J.J.; Bayucan, A.: A Quantitative Method for Software Evaluation and its Application in Evaluating Windowed Environments, *IASTED Software Engineering Conference*, San Francisco, US (1997).
3. Fenton, N.E.; Pfleeger, S.L.: *Software Metrics: a Rigorous and Practical Approach*, 2nd Ed., PWS Publishing Company (1997).
4. GVU's WWW User Survey, The tenth edition, http://www.gvu.gatech.edu/user_surveys/survey-1998-10/tenthreport.html
5. ISO/IEC 9126-1991 International Standard, "Information technology - Software product evaluation - Quality characteristics and guidelines for their use" (1991).
6. Lohse, G.; Spiller, P.: Electronic Shopping, *Communications of the ACM* 41,7 (1998) 81-86.
7. Nielsen, Jakob; *The Alertbox*, <http://www.useit.com/alertbox/>
8. Olsina, L.: 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, <http://sern.cpsc.ucalgary.ca/~maurer/ICSE99WS/ICSE99WS.html>, Los Angeles, US (1999).
9. Olsina, L., Godoy, D; Lafuente, G.J; Rossi, G.: Assessing the Quality of Academic Websites: a Case Study, In: New Review of Hypermedia and Multimedia (NRHM) Journal, Taylor Graham Publishers, UK/USA Vol 5, ISSN 1361-4568 (1999).
 10. Olsina, L.; Rossi, G.: Towards Web-site Quantitative Evaluation: defining Quality Characteristics and Attributes, Proceedings of IV Int'l WebNet Conference, World Conference on the WWW and Internet, Hawaii, USA, Vol.1, (1999) 834-839.
 11. Tilson, R., Dong, J., Martin, S., Kieke, E.: Factors and Principles Affecting the Usability of Four E-commerce Sites", 4th Conference on Human Factors & the Web, Baking Ridge, NJ, US (1998).
 12. Rossi, G., Schwabe, D.; Lyardet, F.: Improving Web Information Systems with navigational Patterns, Proceedings of WWW8 Congress, Toronto, Canada, Elsevier Ed. (1999) 589-600.
 13. W3C, W3C Working Draft, WAI Accessibility Guidelines: Page Authoring, <http://www.w3c.org/TR/WD-WAI-PAGEAUTH/> (1999).

Appendix A: Quality Requirement Tree for E-Bookstores

1. Usability

1.1 Global Site Understandability

1.1.1 Book-domain Organization Scheme

1.1.1.1 Table of Content

1.1.1.2 Alphabetical Subject Index

1.1.2 Quality of Labeling System

1.1.3 Guided Tour for First Time Visitors

1.2 Feedback and Help Features

1.2.1 Quality of Help Features

1.2.1.1 Global Help for First Time Visitors

1.2.1.2 Specific Help

1.2.1.2.1 Search Help

1.2.1.2.2 Purchase Help

1.2.1.2.3 Check-out Help

1.2.2 Addresses Directory

1.2.2.1 E-mail Directory

1.2.2.2 TE/Fax Directory

1.2.3 Link-based Feedback

1.2.3.1 FAQ Feature

1.2.3.2 What's New Feature

1.2.4 Form-based Feedback

1.2.4.1 Questionnaire Feature

1.2.4.2 Comments/Suggestions

1.2.4.3 Subject-Oriented Feedback

1.2.5 Miscellaneous Features

1.2.5.1 Foreign Language Support

1.3 Interface and Aesthetic Features

1.3.1 Cohesiveness by Grouping Main Control Objects

1.3.2 Presentation Permanence and Stability of Main Controls

1.3.2.1 Direct Controls Permanence

1.3.2.1.1 Main Controls Permanence

1.3.2.1.2 Search Control Permanence

1.3.2.1.3 Browse Permanence

1.3.2.1.4 Account Control Permanence

1.3.2.1.5 Shopping Basket Control Permanence

1.3.2.2 Indirect Controls Permanence

1.3.2.3 Stability

1.3.3 Style Issues

1.3.3.1 Link Color Style Uniformity

1.3.3.2 Global Style Uniformity

1.3.4 Aesthetic Preference

2. Functionality

2.1 Searching and Retrieving Issues

2.1.1 Web-site Search Mechanisms

2.1.1.1 Quick Search

2.1.1.2 Advanced Search

2.1.2 Retrieve Mechanisms

2.1.2.1 Level of Retrieving Customization

2.1.2.2 Level of Retrieving Feedback

2.2 Navigation and Browsing Issues

2.2.1 Navigability

2.2.1.1 Orientation

2.2.1.1.1 Indicator of Path

2.2.1.1.2 Label of Current Position

2.2.1.2 Average of Links per Page

2.2.2 Navigational Control Objects

2.2.2.1 Presentation Permanence and Stability of Contextual (sub-site) Controls

2.2.2.1.1 Contextual Controls Permanence

- 2.2.2.1.2 *Contextual Controls Stability*
- 2.2.2.2 *Level of Scrolling*
- 2.2.2.2.1 *Vertical Scrolling*
- 2.2.2.2.2 *Horizontal Scrolling*
- 2.2.3 *Navigational Prediction*
- 2.2.3.1 *Link Title (link with explanatory help)*
- 2.2.3.2 *Link Comment or Description*
- 2.2.4 *Book Browse Mechanisms*
- 2.2.4.1 *Quick Browse*
- 2.3 **E-commerce Functionality and Content**
- 2.3.1 *Product Information*
- 2.3.1.1 *Product (Book) Description*
- 2.3.1.1.1 *Basic Description*
- 2.3.1.1.2 *Book Content & Structure*
- 2.3.1.1.2.1 *Book's Table of Contents*
- 2.3.1.1.2.2 *Content Description*
- 2.3.1.1.3 *Product Image*
- 2.3.1.1.3.1 *Image Availability*
- 2.3.1.1.3.2 *Zooming*
- 2.3.1.2 *Price Evaluation*
- 2.3.1.2.1 *Price Comparison Availability*
- 2.3.1.3 *Product Rating Availability*
- 2.3.1.4 *Related Titles /Authors Recommendation*
- 2.3.1.5 *Catalog Download Facility*
- 2.3.2 *Purchase Features*
- 2.3.2.1 *Purchase Mode*
- 2.3.2.1.1 *On -line*
- 2.3.2.1.1.1 *Shopping Basket*
- 2.3.2.1.1.1.1 *Shopping Basket Availability*
- 2.3.2.1.1.1.2 *Continue Buying Feedback*
- 2.3.2.1.1.1.3 *Edit/Recalculate Feature*
- 2.3.2.1.1.2 *Quick Purchase (1-click or similar)*
- 2.3.2.1.1.3 *Check-out Features*
- 2.3.2.1.1.3.1 *Check-out Security*
- 2.3.2.1.1.3.2 *Canceling Feedback*
- 2.3.2.1.2 *Off-line*
- 2.3.2.1.2.1 *Printable Check-out Form*
- 2.3.2.1.2.2 *Fax/TE/E-mail Purchase*
- 2.3.2.2 *Purchase Policies*
- 2.3.2.2.1 *Purchase Cancellation Policy*
- 2.3.2.2.2 *Return Policy Information*
- 2.3.2.2.3 *Shipping & Handling Information*
- 2.3.2.2.4 *Payment Policy Information*
- 2.3.2.2.5 *Resent Purchase (Gift service)*
- 2.3.3 *Customer Features*
- 2.3.3.1 *E-subscriptions*
- 2.3.3.2 *Customized Recommendations*
- 2.3.3.3 *Account Facility*
- 2.3.3.3.1 *Account Availability*
- 2.3.3.3.2 *Account Security*

- 2.3.3.3.3 *Account Configuration*
- 2.3.3.3.3.1 *Order History/Status*
- 2.3.3.3.3.2 *Account Settings*
- 2.3.3.3.3.3 *Address Book*
- 2.3.3.4 *Customer Revision of a Book*
- 2.3.4 *Store Features*
- 2.3.4.1 *Title Availability Rate*
- 2.3.4.2 *Store Ranking*
- 2.3.4.2.1 *The Top Books*
- 2.3.4.2.2 *The Best Seller Books*
- 2.3.5 *Promotion Policies*
- 2.3.5.1 *With-sale Promotion Availability*
- 2.3.5.2 *Appetizer Promotion Availability (Contests, Miles, etc.)*
- 3. **Site Reliability**
- 3.1 **Non-deficiency**
- 3.1.1 *Link Errors*
- 3.1.1.1 *Broken Links*
- 3.1.1.2 *Invalid Links*
- 3.1.1.3 *Unimplemented Links*
- 3.1.2 *Miscellaneous Errors or Drawbacks*
- 3.1.2.1 *Deficiencies or absent features due to different browsers*
- 3.1.2.2 *Deficiencies or unexpected results (e.g. non-trapped search errors, frame problems, etc.) independent of browsers*
- 3.1.2.3 *Dead-end Web Nodes*
- 3.1.2.4 *Destination Nodes under Construction*
- 4. **Efficiency**
- 4.1 **Performance**
- 4.1.1 *Quick Pages*
- 4.2 **Accessibility**
- 4.2.1 *Information Accessibility*
- 4.2.1.1 *Support for text-only version*
- 4.2.1.2 *Readability by deactivating Browser Image Feature*
- 4.2.1.2.1 *Image Title*
- 4.2.1.2.2 *Global Readability*
- 4.2.2 *Window Accessibility*
- 4.2.2.1 *Number of panes regarding frames*
- 4.2.2.2 *Version without frames*