Specifying Quality Characteristics and Attributes for Websites

L. **Olsina**[*], D. **Godoy**, G.J. **Lafuente**
GIDIS, Department of Computer Science,
Faculty of Engineering, at UNLPam,
[*]*also at UNLP - Argentina*
E-mail *[olsinal, godoyd, lafuente]*
@ing.unlpam.edu.ar

G. **Rossi**
LIFIA, Ciencias Exactas
at UNLP, also at UNM,
and CONICET - Argentina
E-mail *grossi@info.unlp.edu.ar*

ABSTRACT

In this paper, we outline more than a hundred characteristics and attributes for the domain of academic sites in order to show the quality requirement tree and a descriptive framework to specify them. These elements are used in a quantitative evaluation, comparison, and ranking process. The proposed Web-site Quality Evaluation Method (QEM) is a useful approach to assess the artifact quality in the operational phase of a Web Information System (WIS) lifecycle. Hence, we have analyzed three different audiences regarding academic visitor profiles: current and prospective students, academic personnel, and research sponsors. Particularly, the aim of this work is to show a hierarchical and descriptive specification framework for characteristics, sub-characteristics and attributes regarding the student's viewpoint. Finally, partial results are presented and concluding remarks are discussed.

KEYWORDS: Web-site QEM, Quantitative Evaluation, Quality, Characteristics, Attributes.

1. INTRODUCTION

The age of Web-site artifacts for domains as academic sites, museums, and electronic commerce range on an average from one year for the latter, to four years for the former. In addition, existing sites in these domains are not just-document oriented but are becoming application oriented and, as a well-known consequence, they are increasingly complex systems. Hence, to understand, assess, and improve the quality of Web-based systems we should increasingly use software engineering methods, models, and techniques. In this direction, we propose to utilize the Web-site QEM as a powerful quantitative approach to assess the artifact quality in the different phases of a WIS lifecycle. The core models and procedures for logic aggregation and evaluation of characteristics and attributes are supported by the Logic Scoring of Preference (LSP) approach [2]. Particularly, we focus on the evaluation and comparison of quality in the operational phase for academic sites.

Evaluation methods and techniques can be categorized in qualitative and quantitative. Even though software assessment has more than three decades as a discipline [5, 10, 11], the systematic and quantitative quality evaluation of Hypermedia applications and in particular the evaluation of Web sites is rather a recent and frequently neglected issue. In the last three years, quantitative surveys and domain-specific evaluations have emerged [9, 12]. Particularly, in a recent evaluation work [9], the authors identified and measured 32 attributes that influence store traffic and sales. However, in this direction we need flexible, well-defined, engineering-based evaluation methods, models, and tools to assist in the assessment process of complex Web quality requirements.

Specifically, when using Web-site QEM we take into account a set of activities [14, 15]. The main process steps can be summarized as follows: (a) selection of an evaluation and comparison domain; (b) determination of assessment goals and user standpoint; (c) definition and specification of quality requirements; (d) definition and implementation of elementary evaluation; (f) aggregation of elementary attributes to produce the global quality preference; and (g) analyses and assessment of partial and global quality preferences.

In order to illustrate aspects of steps (c) and (d), we include some results of a recently finished case study about academic sites [16]. We have selected six typical, internationally or regionally well-known academic sites to carry out the case study embracing regions of four different continents. In addition, they were published more than three years ago.

With regard to the selected quality characteristics and attributes for assessment purposes, up to eighty direct metrics were found in the process. We group and categorize Web-site sub-characteristics and attributes starting from six standard characteristics [6, 7], which describe with minimal overlap, software quality requirements. As stated in these standards, software quality may be evaluated in general by the following

characteristics: usability, functionality, reliability, efficiency, portability, and maintainability. These high-level characteristics provide a conceptual foundation for further refinement and description of quality. However, the relative importance of each characteristic in the quality requirement tree, varies depending on the user standpoint and application domain considered. The ISO standard, defines three views of quality: users´ view, developers´ view, and managers´ view. Specifically, in the academic domain there are three general audiences regarding the user (visitor) view, namely: current and prospective students (and visitors like parents), academic personnel such as researchers and professors, and research sponsors. Therefore, visitor are mainly concerned in using the site, i.e., its performance, its searching and browsing functions, its specific user-oriented content and functionality, its reliability, its feedback and aesthetic features, and ultimately, are interested in its quality of use. However, maintainability and portability are not visitor concerns. Some student-oriented questionnaires were conducted to help determining the relative importance of characteristics, sub-characteristics, and attributes. Discussions among involved parties took place (i.e. students, academic personnel, and evaluators).

The final aim of the Web-site academic study, is to evaluate the level of accomplishment of required characteristic such as usability, functionality, reliability, and efficiency comparing partial and global preferences. This allows us to analyze and draw conclusions about the state-of-the-art of academic sites quality, from the current and prospective student's point of view.

The structure of this paper is as follows: In section 2, general indications about questions and assumptions for the academic study are made. In section 3, we represent quality characteristics and attributes. Next, in section 4, a hierarchical and descriptive specification framework are discussed. In addition, characteristics and attributes are modeled. Finally, some partial outcomes are analyzed, and concluding remarks are considered.

2. SOME CONSIDERATIONS ON THE ACADEMIC STUDY

We have selected six academic operational sites aging four years on an average. In order to carry out the study, the selected sites were typical and well-known academic organizations including universities like Stanford University (USA) [21], Chile University [18], the National University of Singapore [20], the University Technological of Sydney (Australia) [23], the Catalunya Polytechnic University (Spain) [19], and the University of Quebec at Montreal, Canada [22]. Figure 1, shows a snapshot of two home pages.

Figure 1: *From left to right, Stanford University and University Technological of Sydney home pages. These pictures were dumped within data collection period (from Jan. 22 to Feb. 22, 1999).*

One of the primary goals for this academic-site quality assessment is to understand the current level of fulfillment of essential characteristics given a set of quality requirements. The assessment process focus on the prospective and current student viewpoint.

Speaking in a wide sense, software artifacts are generally produced to satisfy specific user's needs, and Web-site artifacts are not the exception. In designing Web-site artifacts, there are many challenges which are not always taken into account. 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 and/or

searching. Then, to get a time-effective mental model of the overall site (i.e., its structure and content), there are attributes like a site map, an index, or a table of contents, that help in getting a quick global site understandability, facilitating browsing. On the other hand, a global searching function provided in the main page could effectively help retrieving the desired piece of information and avoid browsing. Moreover, both functions could be complemented. There are a lot of such attributes and complex characteristics that contribute to site quality as usability, functionality, reliability among others that a designer should take into account when designing for intended audiences.

On the other hand, we should take into account that Web sites are artifacts that can evolve dynamically and users always access the last on-line version. By the time of data collection (which began on January 22, and finished on February 22, 1999), we did not perceive changes in these Web sites that could have affected the evaluation process.

Lastly, we should make an important consideration with regard to data collection. In fact, the data collection activity could be done manually, semi-automatically, and automatically. Most of the attributes values were collected manually because there is no way to do it otherwise. However, automatic data collection is in many cases the more reliable and almost unique mechanism to collect data for a given attribute. This should be the case if we want to measure the *Dangling Links*, *Image Title,* and *Page Size* attributes, among others. These attributes were automated with an integrated tool called SiteSweeper (in the fourth section, we discusss these attributes).

3. OUTLINING THE QUALITY REQUIREMENT TREE FOR THE ACADEMIC DOMAIN
In this section, we outline over a hundred and twenty quality characteristics and attributes for the academic site domain. Among them, up to eighty were directly measurable. The primary goal is to classify and group the elements that might be part of a quantitative evaluation, comparison, and ranking process in a requirement tree. As previously said, to follows well-known standards we use the same high-level quality characteristics like usability, functionality, reliability, and efficiency. These characteristics give evaluators a conceptual framework of quality requirements and provide a baseline for further decomposition. A quality characteristic can be decomposed in multiple levels of sub-characteristics, and finally, a sub-characteristic could be refined in a set of measurable attributes.

In order to effectively select quality characteristics we should consider different kind of users. Specifically, in the academic domain, there are three different audiences regarding the visitor standpoint as studied elsewhere [10, 21]. The visitors were categorized in current and prospective students, academic personnel, mainly researchers and professors, and research sponsors. (This audience-oriented division is clearly established in the structure of UTS site). Figure 2, outline the major characteristics and measurable attributes regarding current and prospective students. Likewise, as in museum sites evaluation and for the student view, high-level artifact characteristics such as maintainability and portability were not included in the requirements.

Following we comment some characteristics and attributes and the decomposition mechanism. The *Usability* high-level characteristic is decomposed in sub-factors such as *Global Site Understandability*, *On-line Feedback and Help Features*, *Interface and Aesthetic Features*, and *Miscellaneous Features*. The *Functionality* characteristic is split up in *Searching and Retrieving Issues*, *Navigation and Browsing Issues*, and *Student-oriented Domain-related Features*. The same decomposition mechanism is applied to *Reliability* and *Efficiency* factors. For instance, *Efficiency* high-level characteristic is decomposed in *Performance* and *Accessibility* sub-characteristics. (A hierarchical and descriptive specification framework for each characteristic or attribute will be presented in the next section).

For *Global Site Understandability* sub-characteristic (within *Usability*), in turn we have split up in *Global Organization Scheme* sub-characteristic, and in quantifiable attributes as *Quality of Labeling*, *Student-oriented Guided Tours,* and *Campus Image Map*. However, *Global Organization Scheme* sub-characteristic is still too general to be directly measurable, so we derive attributes such as *Site Map*, *Table of Content*, and *Alphabetical Index*.

Focusing on *Student-oriented Domain-related Features* characteristic (where *Functionality* is the super-characteristic), we have observed two main sub-characteristics, namely: *Content Relevancy* and *On-line Services*. As the reader can appreciate we evaluate aspects ranging from academic units, degree/courses, enrollment and from services information, to ftp, news groups, and web publication provided for undergraduate and graduate students.

1. Usability
1.1 Global Site Understandability
   1.1.1 Global Organization Scheme
      1.1.1.1 *Site Map*
      1.1.1.2 *Table of Content*
      1.1.1.3 *Alphabetical Index*
   1.1.2 *Quality of Labeling System*
   1.1.3 *Student-oriented Guided Tour*
   1.1.4 *Image Map (Campus/Buildings)*
1.2 On-line Feedback and Help Features
   1.2.1 Quality of Help Features
      1.2.1.1 *Student-oriented Explanatory Help*
      1.2.1.2 *Search Help*
   1.2.2 Web-site Last Update Indicator
      1.2.2.1 *Global*
      1.2.2.2 *Scoped (per sub-site or page)*
   1.2.3 Addresses Directory
      1.2.3.1 *E-mail Directory*
      1.2.3.2 *Phone-Fax Directory*
      1.2.3.3 *Post mail Directory*
   1.2.4 *FAQ Feature*
   1.2.5 On-line Feedback
      1.2.5.1 *Questionnaire Feature*
      1.2.5.2 *Guest Book*
      1.2.5.3 *Comments*
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*
      1.3.2.2 *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*
      1.3.3.3 *Global Style Guide*
   1.3.4 *Aesthetic Preference*
1.4 Miscellaneous Features
   1.4.1 *Foreign Language Support*
   1.4.2 *What's New Feature*
   1.4.3 *Screen Resolution Indicator*

2. Functionality
2.1 Searching and Retrieving Issues
   2.1.1 Web-site Search Mechanisms
      2.1.1.1 *Scoped Search*
         2.1.1.1.1 *People Search*
         2.1.1.1.2 *Course Search*
         2.1.1.1.3 *Academic Unit Search*
      2.1.1.2 *Global 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 *Quality of Link Phrase*
2.3 Student-oriented Domain-related Features
   2.3.1 Content Relevancy
      2.3.1.1 Academic Unit Information
         2.3.1.1.1 *Academic Unit Index*
         2.3.1.1.2 *Academic Unit Sub-sites*
      2.3.1.2 Enrollment Information
         2.3.1.2.1 *Entry Requirement Information*
         2.3.1.2.2 *Form Fill/Download*
      2.3.1.3 Degree Information
         2.3.1.3.1 *Degree Index*
         2.3.1.3.2 *Degree Description*
         2.3.1.3.3 *Degree Plan/Course Offering*
         2.3.1.3.4 Course Description
            2.3.1.3.4.1 *Comments*
            2.3.1.3.4.2 *Syllabus*
            2.3.1.3.4.3 *Scheduling*
      2.3.1.4 Student Services Information
         2.3.1.4.1 *Services Index*
         2.3.1.4.2 *Healthcare Information*
         2.3.1.4.3 *Scholarship Information*
         2.3.1.4.4 *Housing Information*
         2.3.1.4.5 *Cultural/Sport Information*
      2.3.1.5 Academic Infrastructure Information
         2.3.1.5.1 *Library Information*
         2.3.1.5.2 *Laboratory Information*
         2.3.1.5.3 *Research Results Information*
   2.3.2 On-line Services
      2.3.2.1 *Grade/Fees on-line Information*
      2.3.2.2 *Web Service*
      2.3.2.3 *FTP Service*
      2.3.2.4 *News Group Service*

3. Site Reliability
3.1 Nondeficiency
   3.1.1 Link Errors
      3.1.1.1 *Dangling 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 (unexpectedly) under Construction*

4. Efficiency
4.1 Performance
   4.1.1 *Static Page Size*
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 *Non-frame Version*

Figure 2: *Quality Requirement Tree for Academic Websites*

Finally, and regarding quantifiable attributes we might easily see that no necessarily all attributes of a given characteristic should exist simultaneously in a Web site. This is the case for *On-line Feedback* characteristic, where a *Questionnaire Feature,* a *Guest Book,* or *Comments* attribute could alternatively exist. However, in many cases, modeling the simultaneity relationship among attributes and characteristics might be an essential requirement of the evaluation system. For instance, for the *Content Relevancy* characteristic might be mandatory the existence of *Academic Unit Information*, and *Enrollment Information*, and *Degree Information* sub-characteristics. Ultimately, it is important to stress that in the evaluation process we use the LSP model, which allows to deal with logic relationships taking into account weights and levels of and/or polarization [2, 3]. (In [16] this is widely illustrated)

4. A HIERARCHICAL AND DESCRIPTIVE SPECIFICATION FRAMEWORK

We present here some attributes for the academic study following a regular structure, i.e., title, code, element type, high-level characteristic, super and sub-characteristics, definition/comments, elementary criteria, preference scale, data collection type, and example components. Figure 3 shows the three templates that model characteristics, sub-characteristics and attributes respectively.

Title: Code: Type: Characteristic	**Title:** Code: Type: Attribute
Sub-characteristic/s:	Higher level Characteristic:
Definition / Comments:	Super-characteristic:
Model to determine the Global/Partial Computation:	Definition / Comments:
Employed Tool/s:	Elementary Criteria Type:
Preference Scale:	Preference Scale:
Example/s:	Data Collection Type:
	(Employed Tool/s:)
	Example/s:

Title: Code: Type: Sub-characteristic		
Super-characteristic:	Sub-characteristic/s:	Attribute/s:
Definition / Comments:		
Model to determine the Global/Partial computation:	Employed Tool/s:	
Preference Scale:		
Example/s:		

Figure 3 **:** *Templates to specify a higher level characteristic; an attribute; and a sub-characteristic*

We next use the above specification cards to exemplify one characteristic and five attributes for the academic study.

Title: *Usability*; Code: 1 *;* Type: Characteristic
Sub-characteristic/s: Global Site Understandability, On-line Feedback and Help Features, Interface and Aesthetic Features, Miscellaneous Features.
Definition / Comments: It is a high-level quality characteristic -that can be indirectly measured-; it represents the level of effort that requires a given set of users to operate, understand, and communicate with the software artifact. It includes features like global understandability, operability, and communicativeness, among others, as well as aesthetic and style issues.

It is important to cite as comments the standard ISO definition [7], that states (in pp. 3): *"A set of attributes that bear on the effort needed for use, and on the individual assessment of such use, by a stated or implied set of users"*. In addition, to the one given by IEEE [6], in the A Annex (-informative-), say: *"An attribute that bears on the effort needed for use (including preparation for used and evaluation of results), and on the individual assessment of such use by users"*.
Model to determine the Global/Partial Computation: LSP model;
Employed Tool/s: Automatic, developed to compute LSP logic operators.

Preference Scale:
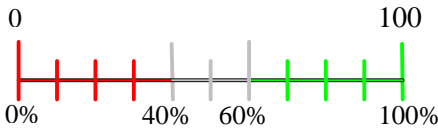
Example/s: It has been used, as a constituent part of the evaluation requirements in two cases studies and a survey. Also, in two WIS development projects.

**Title:** *Table of Content; Code: 1.1.1.2; Type:* Attribute
Higher-level characteristic: Usability
Super-characteristic: Global Organization Scheme
Definition / Comments: It is an attribute that permit structuring the content of the whole site allowing the navigation mainly by means of linked text. It is usually available in the home page and emphasizes the information hierarchy so that users can become increasingly familiar with how the content is organized in subsites. Also, it facilitates fast and direct access to the contents of the Web site [17].
Elementary Criteria: is an absolute and discrete binary criterion: we only ask if it is available (1) or it is not available (0).

Preference Scale:
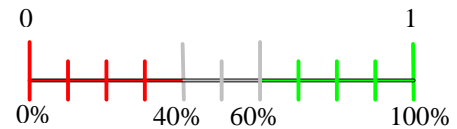
Data Collection Type: Manual, Observational
Example/s: Examples of table of content availability are NUS, UTS, Stanford, and UPC sites. The computed elementary preference is 100%. Besides, in the subsite organization of UTS´s table of content an audience-oriented division is clearly established (e.g. for students, for staff, and for researchers and sponsors).
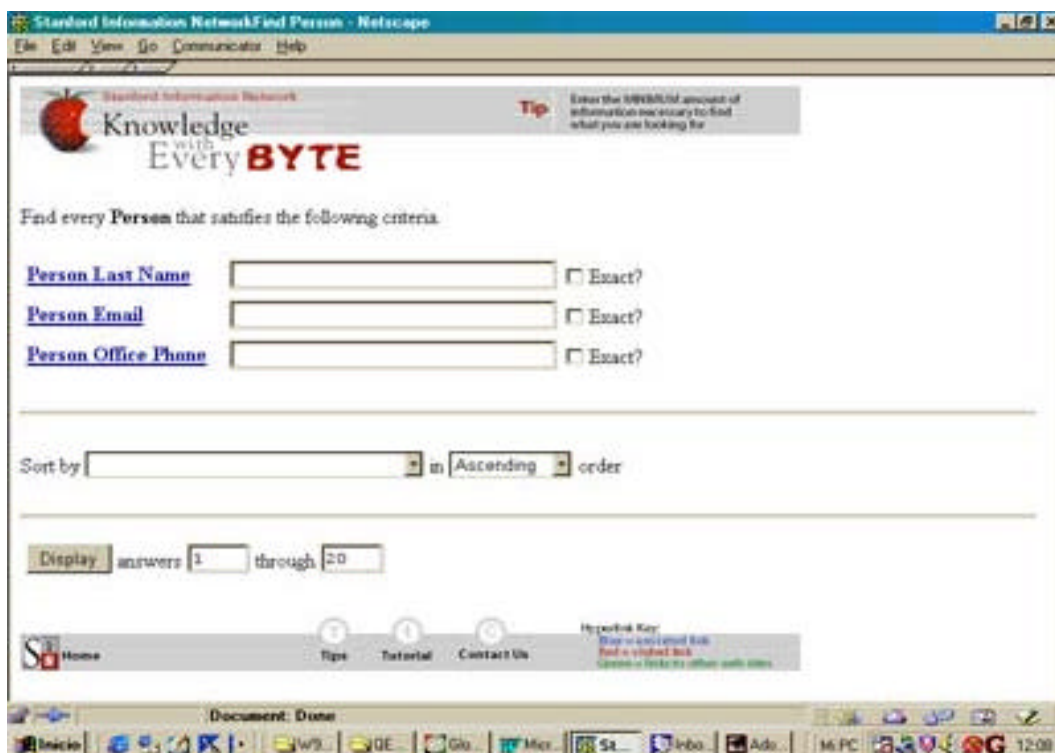

Figure 4: *The Stanford University Scoped-people Search and Retrieval Customization facilities*

Title: *People Search; Code: 2.1.1.1.1; Type:* Attribute
Higher-level characteristic: Functionality
Super-characteristic: Scoped Search
Definition / Comments: Sometimes, specific areas of a site are highly coherent and distinct from the rest of the site that makes sense to give a scoped or restricted search to users [12].

For instance, for a museum visitor can often be better counting with both scoped and global search; i.e., it could be necessary a customized Scoped Search to search a (museum) collection by author and school as long as a Global Search could also be necessary to search general issues.
Elementary Criteria: is a multi-level discrete absolute criterion defined as a subset, where: 0=no search mechanism is available; 1=search mechanism by name/surname; 2 = 1 + expanded search: search mechanism by academic unit and/or subject area or discipline, and/or phone etc.

Preference Scale:
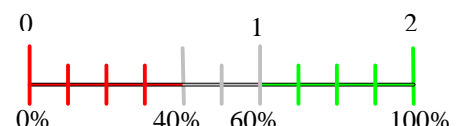
Data Collection Type: Manual, Observational

Example/s:
1) An outstanding example is the Stanford people search (http://sin.stanford.edu:2000/frame?person) as illustrated in the figure 4. The computed elementary preference is 100%.
2) Other examples are at University of Chile (http://www.sisib.uchile.cl/docentes/) and at UQAM (http://www.repertoire.uqam.ca/).

Title*: **Dangling 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 broken links)

The following comment shows some survey results about broken links: Jakob Nielsen's Alertbox [12] (June 14, 1998: http://www.useit.com/alertbox/980614.html), said "*6% of the links on the Web are broken according to a recent survey by Terry Sullivan's All Things Web. Even worse, linkrot in May 1998 was double that found by a similar survey in August 1997. Linkrot definitely reduces the usability of the Web, being cited as one of the biggest problems in using the Web by 60% of the users in the October 1997 GVU survey. This percentage was up from "only" 50% in the April 1997 survey. 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*".
Elementary Criteria: is an absolute and continuous single-variable criteria., where: BL=number of broken links found. TL=number of total site links. The formula to compute the preference is: $X = 100 - (BL * 100/TL) * 10$ where, if $X < 0$ then $X = 0$.

Preference Scale:

X_{min} 100
0%   40%   60%   100%
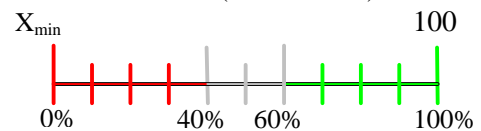
Data Collection Type: Automated.
Example/s: For instance, the National University of Singapore produces a preference of 68.06 %. The real value was computed from the above formula: $100 - ((970*100)/30883) * 10 = 68.06$

Title*: **Static Page Size***;* Code: 4.1.1*;* Type: Attribute
Higher-level characteristic: Efficiency
Super-characteristic: Performance
Definition / Comments: It measures the total size of each static page regarding textual and imaged components. We specify a total download size limit (or threshold) of 35.2 Kbytes per page. A page of this size requires about 20 seconds to download at 14,400 bps. (as a limit of acceptable period of time that a user might wait).

IEEE Web Publishing guide [5], in *Performance* section, comments*: "Users tend to become annoyed when a page takes longer than 20 seconds to load. This means it is best to limit the total of the file sizes associated with a page, including graphics, to a maximum of 30 – 45 kilobytes to assure reasonable performance for most users.*"
Elementary Criteria: is an absolute and continuous multi-variable criterion. The formula to compute the preference is: $X = ((X_1 - 0.4 X_2 - 0.8 X_3) / (X_1 + X_2 + X_3)) * 100$; where X_1 represents the number of pages within a download time ranging from $0 < X_1 <= 20$ seconds, and X_2 represents the number of pages within a download time ranging from $20 < X_2 <= 40$, and X_3 represents the number of pages within a download time where: $X_3 > 40$ sec.

0   100
0%   40%   60%   100%

Preference Scale:
Data Collection Type: Automated.
Example/s: As an example we may consider UTS site, where the tool reported "*You specified a total download size limit of 35.2K bytes per page. A page this size requires about 20 seconds to download at 14.4K bps. Of the 18.872 pages on your site, 2.210 pages (12%) have a total download size that exceeds this threshold*". Regarding the above formula and the values reported by the tool (see also figure 5), the following computation: (16662-0.4*1850-0.8*440)/18872, yield a preference of 82 %. Amazingly, Stanford Web-site drew an elementary preference of 100% (no page overflow the threshold of 35.2K bytes)
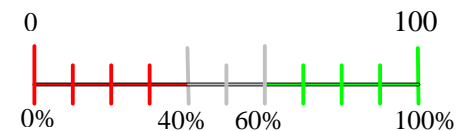
Title*: **Image Title***;* Code: 4.2.1.2.1*;* Type: Attribute
Higher-level characteristic: Efficiency
Super-characteristic: Readability by deactivating Browser Image Feature
Definition / Comments: Alternative texts for each image or graphic component should be provided since they convey visual information. It measures the percentage of <ALT> tag presence that includes replacement text for

the image. This attribute favors the readability feature when the user can not use the browser's image feature. However, the measure of this attribute does not guarantee the quality of alternative text. Some text could be generated automatically when editing with tools like FrontPage, etc.

See the guides provided by the W3C in the WAI Accessibility Guidelines [25], specifically "*A.1 Provide alternative text for all images, applets, and image maps*". Among others things says: "*Text is considered accessible to almost all users since it may be handled by screen readers, non-visual browsers, Braille readers, etc. It is good practice, as you design a document containing non-textual information (images, graphics, applets, sounds, etc.) to think about supplementing that information with textual equivalents wherever possible*".

Elementary Criteria: is a continuous absolute single-variable criterion, where AAR= absent ALT reference. TAR=number of ALT inline references. The formula to compute the preference is: $X = 100 - (AAR * 100/TAR)$

Preference Scale:
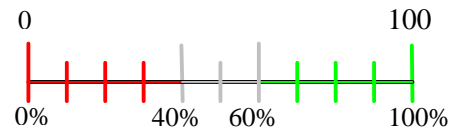
Data Collection Type: Automated.

Example/s: An example is shown in the figure 5 for the UTS site. The tool gives us directly the percentages, and in this case reported "*Of the 63,882 inline references on your site that should specify an ALT attribute, 11,721 references (18%) are missing the attribute. The missing ALT attributes appear on 3,338 different pages*". The elemental preference drew 81.65%
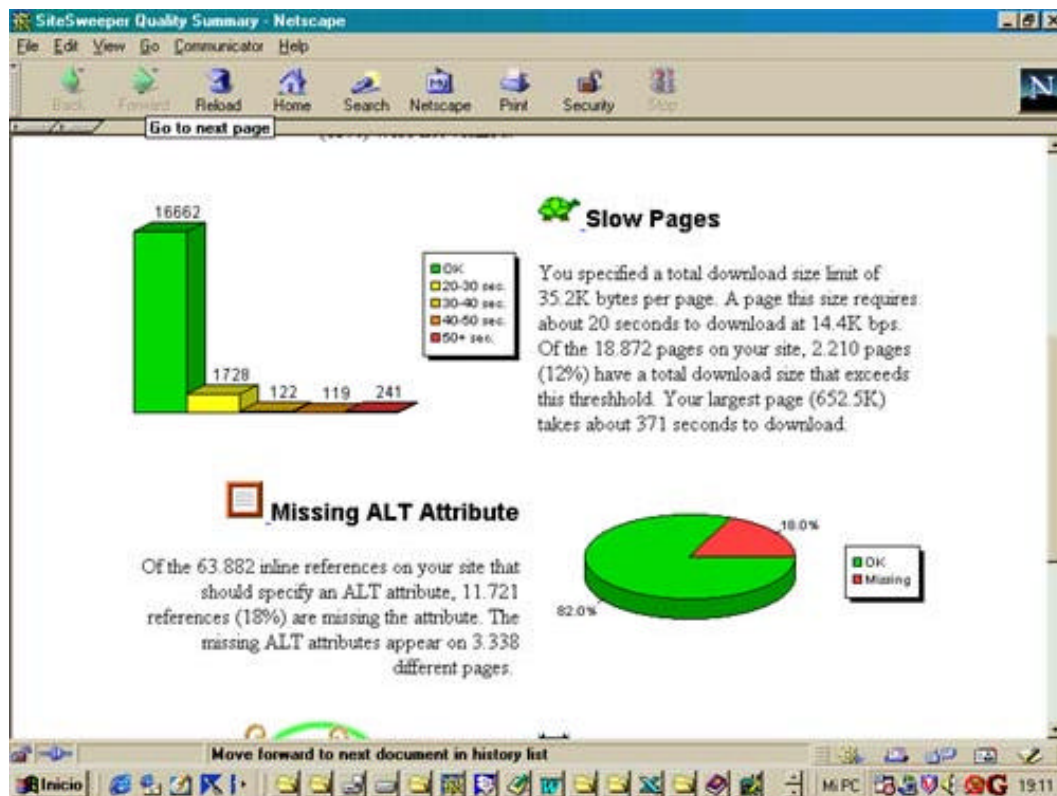

Figure 5: *A dumped screen of the Quality Page report (using a trial version of SiteSweeper 2.0) showing both the different page size categories and missing ALT attribute.*

Finally, once all elementary criteria were prepared and agreed, and necessary data collected, we can compute the elementary quality preference for each competitive system. Table 1, shows partial results of preferences after computing the corresponding criteria function for each academic site attribute.

We include some elementary results for *Usability* characteristic as well as *Functionality*, *Reliability*, and *Efficiency* characteristics; mainly, values for the aforementioned specified attributes. Even if they are only elementary values where no aggregation mechanisms were still applied (the f-step of our methodology, as commented in the Introduction section), and no global outcomes produced, however, some important conclusions can be obtained.

Table 1: *Partial results of elementary quality preferences for the six academic sites*

	UPC Spain	Info	Uchile Chile	Info	UTS Australia	Info	NUS Singapore	Info	Stanford USA	Info	UQAM Canada	info
Usability												
1.1.1.1	100	1	0	0	0	0	0	0	0	0	0	0
1.1.1.2	100	1	0	0	100	1	100	1	100	1	0	0
1.1.1.3	0	0	0	0	100	1	0	0	100	1	0	0
1.1.2	90		90		90		80		90		80	
1.1.3	0	0	0	0	100	1	0	0	100	1	0	0
1.1.4	100	1	100	1	100	1	100	1	50	0.5	100	1
Functionality												
2.1.1.1.1	60	1	100	2	60	1	100	2	100	2	100	2
2.1.1.1.2	0	0	0	0	100	2	0	0	100	2	0	0
2.1.1.1.3	0	0	0	0	0	0	0	0	100	2	100	2
2.1.1.2	60	1	60	1	60	1	0	0	100	2	100	2
Reliability												
3.1.1.1	0	-29	75.02	75.02	74.1	74.1	68.06	68.06	58.32	58.32	0	-10
Efficiency												
4.1.1	75.3		50.46		82		51.46		100		83.44	

For instance, we can see that two out of six sites have no resolved *Global Organization Scheme* (i.e. neither *Site Map*, nor *Table of Content*, and nor *Alphabetical Index* attributes available). As previously said, when users enter at a given home page for the first time, the direct or indirect availability of these attributes may help them in getting a quick *Global Site Understandability* both for the structure and the content. Likewise, attributes like *Quality of Labeling*, *Student-oriented Guided Tours,* and *Campus Image Map* contribute to global understandability. Nonetheless, and regarding attributes of the *Global Organization Scheme* feature, we see that no necessarily all of them might exist at the same time (the replaceability relationship); a *Table of Content*, an *Index* attribute, or a *Site Map* could be required; however, others arrangements could be possible. This is the case with UPC, UTS, NUS, and Stanford sites, where only some attributes are present (and should not be punished for the absence of one another).

On the other hand, only Stanford and UTS universities have *Student-oriented Guided Tours*; both are excellent tours (accomplishing the 100% of the quality preference), but the one in UTS is simply outstanding. Not only it has student-oriented tour but it also contains a personalized guide for each academic unit. (The visitor can access it in the table of content's "*For Students*" label, in the "*Virtual Open One day*" link).

Besides, all universities have the necessary *Campus Image Map* feature; only the Stanford campus imagemap is not easy to access it (goes out of context), and is not well structured (getting 50% of the preference). Let us recall that a scoring within gray lines of the preference scale (among 40 and 60%) can be interpreted as an improvement actions should be considered as long as an unsatisfactory rating level, within the red lines (among 0 and 40%), can be interpreted as a necessary change actions must be taken [14]).

Regarding *Functionality,* there are two main functions to move into a site in order to find information, i.e. browsing and/or searching. In addition, from the point of view of current and prospective students, scoped searching functions as outlined in the requirement tree are necessary attributes. For instance, we found all sites having at least the basic feature of *People Search* attribute; however, not all sites have *Course Search* facilities. In addition, the reader can appreciate the elementary results of *Dangling Links* (3.1.1.1), and *Static Page Size* (4.1.1) attributes, which were automated with a sweeper tool, as commented in previous sections.

5. CONCLUDING REMARKS

In this paper, standardized characteristics, and about eighty directly measurable attributes for the sites on the academic domain were considered. The main goal was to establish quality requirements to arrange the list of characteristics and attributes that might be part of a quantitative evaluation, comparison, and ranking process. The proposed Web-site QEM methodology, grounded in a logic multi-attribute decision model and procedures, is intended to be a useful tool to evaluate artifact quality in the operational phase of a WIS lifecycle. In addition, it could be also used in earlier stages as exploratory and development phases.

The evaluation process generates elemental, partial, and global quality preferences that can be easily analyzed, backward and forward traced, justified, and efficiently employed in decision-making activities. The outcomes should be useful to understand, and potentially improve the quality of Web artifacts in medium and large-scale projects.

Finally, we have shown a hierarchical and descriptive specification framework to represent characteristics, sub-characteristics, and attributes. We have shown some attributes for the academic case study following a regular structure, i.e., title, code, element type, high-level characteristic, super and sub-characteristics, definition/comments, elementary criteria, preference scale, data collection type, and example components. Some data were collected manually and some others automatically. It is important to stress the valuable help and high confidence that provided by automatic tools.

At this moment, we have finished the academic case study, and we are working on the evaluation and comparison of well-known electronic commerce sites. As an anecdotal end, in the final ranking we find Stanford University with the 79.76 % of the global quality preference, UTS with 69.61%, UQAM with 66.05%, UPC with 65.06%, UChile with 56.551%, and NUS with 54.46% [16]. Finally, these case studies will allow us to strength the validation process on quality metrics as long as our experience grows.

REFERENCES

1. Botafogo, R. Rivlin, E., Shneiderman, B.*,* 1992, *"Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics,* ACM Transactions on Office Information Systems, 10(2), pp. 142-180.
2. Dujmovic, J.J., 1996, *"A Method for Evaluation and Selection of Complex Hardware and Software Systems",* The 22nd International Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems. CMG 96 Proceedings, Vol. 1, pp.368-378.
3. Dujmovic, J.J.; Bayucan, A., 1997, *"A Quantitative Method for Software Evaluation and its Application in Evaluating Windowed Environments",* IASTED Software Engineering Conference, San Francisco, US.
4. Fenton, N.E.; Pfleeger, S.L., 1997, *"Software Metrics: a Rigorous and Practical Approach",* 2nd Ed., PWS Publishing Company.
5. Gilb, T., 1969, *"Weighted Ranking by Levels"*, IAG Journal, Vol 2 (2), pp. 7-22
6. IEEE Web Publishing Guide, http://www.ieee.org/web/developers/style/
7. IEEE Std 1061-1992, *"IEEE Standard for a Software Quality Metrics Methodology"*
8. ISO/IEC 9126-1991 International Standard, *"Information technology – Software product evaluation – Quality characteristics and guidelines for their use"*.
9. Lohse, G.; Spiller, P., 1998, *"Electronic Shopping",* CACM 41,7 (Jul-98); pp. 81-86.
10. McCall, J.A; Richards, P.K.; Walters, G.F.; 1977, *"Factors in Software Quality"*, RADC TR-77-369.
11. Miller, J.R.; 1970, *"Professional Decision-Making",* Praeger Publisher.
12. Nielsen, Jakob; The Alertbox, http://www.useit.com/alertbox/
13. Olsina, L, 1998, *"Building a Web-based Information System applying the Hypermedia Flexible Process Modeling Strategy";* 1st International Workshop on Hypermedia Development, at ACM Hypertext 98, Pittsburgh, US (The paper is available at http://ise.ee.uts.edu.au/hypdev/).
14. Olsina, L., 1998, *"Web-site Quantitative Evaluation and Comparison: a Case Study on Museums"*, ICSE ´99 Workshop on Software Engineering over the Internet
15. Olsina, L., Rossi, G.; 1998, *"Toward Web-site Quantitative Evaluation: defining Quality Characteristics and measurable Attributes",* Submitted paper to WebNet '99 Conference.
16. Olsina, L., Godoy, D; Lafuente, G.J; Rossi, G.; 1999, *"Assessing the Quality of Academic Websites: a Case Study",* Submitted paper.
17. Rosenfeld, L., Morville, P., 1998, *"Information Architecture for the WWW"*, O`Reilly.
18. Universidad de Chile: http://www.uchile.cl
19. Universidad Politécnica de Cataluña: http://www.upc.es
20. University of Singapore: http://www.nus.sg
21. University of Stanford: http://www.stanford.edu
22. University of Quebec: http://www.uqam.ca
23. University Technological of Sydney: http://www.uts.edu.au
24. Webby, R.; Lowe, D., 1998, *"The Impact Process Modeling Project",* 1st International Workshop on Hypermedia Development, at ACM Hypertext 98, Pittsburgh, US (The paper is available at http://ise.ee.uts.edu.au/hypdev/).
25. W3C, 1999, W3C Working Draft, *"WAI Accessibility Guidelines: Page Authoring"*, http://www.w3c.org/TR/WD-WAI-PAGEAUTH/