Website Error Analysis of Colleges and Universities on Long Island in New York

A. Ebrahimi, C. Schweikert, S. Sayeed,
S. Parham, H. Akibu, A. Saeed, W. Parris
MIS – School of Business
SUNY College at Old Westbury
Old Westbury, New York 11568 USA
{ebrahimia, schweikertc, ssayeed, sparham,
hakibu, asaeed, wparris}@oldwestbury.edu

Abstract

Websites are an important means of communication for colleges and universities. Web-based presentations are replacing traditional ways of disseminating information. This study investigates the websites of 21 colleges and universities located on Long Island (Nassau, Suffolk) as well as five additional schools, four of which are Ivy League institutions. This research developed a technique of web site analysis known as 7X2 C's criteria, which are applied to seven separate layers of website application and development. The study concludes that the web sites of all the included institutions have errors across all the layers of the analysis to some degree that need to be corrected. Some of these errors are crucial, especially with regard to website search engines.

Keywords: Error, Web Quality Assurance, Search Engine, College Websites, Web Programming, HTML, Web Design, Information Technology

1. Introduction

College and university web pages are an important forum for reaching out to prospective students and the school community; these web pages also provide valuable information to the public [1,5]. This economical and convenient means of disseminating information, which is replacing traditional methods of communication by college catalogs, mail, and phone, has tremendous implications. As websites increasingly become a primary source of information, extra care should be considered to avoid any crucial errors [4]. However, how accessible, accurate, helpful, and up to date are college websites?

This study investigates 21 Long Island universities and colleges and five institutions from other locations, four of which are Ivy League. The web sites of these schools are analyzed based on criteria known as 7X2 C's analysis in 7 separate layers of web application and development. The 7 pairs of C's are: (content, context), (correctness, credibility), (currency, continuity), (completeness, coverage), (consistency, conciseness), (community, customization), and (compelling, creativity). The 7 layers of website application and web development will look at school identity (who, what, where, when, why), school website appearance (clarity, animation, features), web search engine (efficiency), html (tags, errors), client side script programming (JavaScript: features, coding), server and database programming (appropriateness, speed), privacy, security (validation), and ethics (integrity). The examples provided in this paper mostly focus on the search engine layer in conjunction with 7X2 C's analysis; since a correctly and efficiently operating search engine is crucial to providing accurate and meaningful information. While this study is not exhaustive, its findings are suggestive that schools must fully take into consideration the implications and responsibilities of commitment to website usage as the primary means of information transfer.

2. Study Background, Long Island, and Motivation

The subject of errors on the web pages of Long Island colleges and universities was the class project for students of Management Information Systems in a course titled Topics in MIS during the Fall 2006 semester at SUNY, Old Westbury in Long Island, New York.

Long Island is the home of many unique colleges and universities and is known for its advancement in education. It produces some of the most qualified graduates in the nation [6]. Each student was assigned

two colleges, one which was worked on as a team. Students discussed their findings during the classes with live and interactive demonstrations. Their work was presented at the Second Annual Computer Human Error Conference (CHE 2006) at the CUNY Graduate Center. In order to expand the geographical scope of this study, an additional five institutions are included, four of which are Ivy League schools. By studying the profiles of web pages, one can improve the design of existing web pages and the results of the analysis can be used to make suggestions for a high-quality site [3].

3. Search Engine Layer Study Categories

The search engine layer is one of the seven layers of web analysis and plays a vital role in accessing information through the web. Instead of manually browsing each page, going from link to link, one can simply type the required information in the search box and receive the appropriate results. The 21 Long Island colleges and universities included in this study are divided into the following categories according to their search engines:

(a) Three schools did not have a search engine box at all and to find a particular item on these web pages, a user must navigate through links, menu bars, or indices
(b) Eight schools did not have a search engine box on their home pages (main page) in order to enter a search key directly
(c) Three schools had a search engine that was powered by Google and did not have an in-house search engine
(d) Ten schools had only in-house search engines as part of their web design
(e) One school had both an in-house search engine and Google
(f) Two schools had both an in-house and a commercial search engine besides Google (Ultraseek and FreeFind)

The five non-LI schools included in this study all had a mixture of both Google and in-house search engines on their site.

4. Indirect Anonymous Error Reporting

The aim of this study is to identify and report website errors in order to improve the quality of college web pages. Extra care has been taken to mask the identity of any specific college or university. However, by examining the error messages included in this study, a college web master and/or an interested person will be able to identify the school in question. From the original snapshots, the images, white spaces, and extra lines have been removed and replaced with punctuation marks to indicate formatting. Punctuation such as a semicolon (;) is used as a line separator and dots have been used to show continuity. The following are the semi-snapshot errors on the web pages analyzed based on the 7X2 C's in different layers.

5. Content and Context

A web page should have content, referring to the volume of information that is helpful and informative, within a context that gives the content meaning and appropriateness. An example error in this category would be a web page that is fully empty or the context is not relevant to the user accessing it.

(a) Search for Smith

Searching for "Smith" in one college's site resulted in the following links. No information was provided about this search key and the information provided was outdated.

"Web Pages
77.4% ……FUNDED GRANTS AND PROGRAMS;... 77.4% ..... || CAMPUS OFFICES | HUMAN RESOURCES; Newsletters, Calendar Events and Press Releases; 77.4% Calendar ;"

(b) Professor Search

When searching for a known professor "Kei..." the following message was displayed:

"*There are no records matching your criteria.;"*

(c) What is in a Blank Space?

When a blank space is entered into the search box every possible file was returned.

"*100% ...WORKSHOP; 100% ...FUNDED GRANTS ...PROGRAMS; .....;.;.;.;.;.;.;.;*"

(d) Meaning of Special Characters

Searching for special characters such as "?", gave random results for other sections of the university's website. However, this may nonetheless be one of the best in-house search engines that were analyzed.

".....University:Admissions;http://admissions.....edu/;Ignored: ?;Your search for did not return any matches (0 documents were searched):No documents were found.;"

(e) META Tag Relevancy

A problem was found in the HTML layer of content and context. The HTML code of one school listed several irrelevant terms in their META TAG Keywords, i.e. Desktop Publishing, Southern Folk Art and sculpture, while relevant terms such as "Public College" or "business", which should have been amongst the first terms in the list were missing.

(f) Significance of a Page

One college's website on Long Island had very little content on its index page. The content that was on the main page had very little to do with the campus itself and gave prominence to the "Network Use Policy".

6. Correctness and Credibility

A web page should provide correct information and build trust among its users. The user should rely on the information provided by the site and should believe the information to be correct and persuasive [2].

(a) Looking for Smith, not O...., J…. Carol

Using a university's website a search for "Smith" in the people directory was conducted. The results included records that had Smith as part of their last name and one result that did not have Smith in the name at all - "O...., J…. Carol". The intended person to be found was the tenth option in the list.

(b) Server Not Responding; Page Cannot Be Displayed;

When clicking on the Health or Career Services tags on the navigation bar on one college's webpage, the site ceased to respond. This problem belongs to the server and database layer.

7. Currency and Continuity

A web page must be current, up-to-date, and provide a vision that will progress with time. A webpage that was last updated several years ago would be an example of error in this category.

(a) "Classes begin soon"

In one of the colleges analyzed we found that there was a lack of currency. The "News" and Events" section could only be accessed from one of the satellite campus' website. On the main page, under the New and Events section the message **"***Classes begin soon***"** is permanently displayed and this is the only event in the calendar. It was also found that the copyright display of this college was outdated as follows:

"*© 2001-2006 ..... College.; All rights reserved.***"**

(b) ....Last update: April 17, 2005;

A website was analyzed and under prospective students, an update had not been performed in quite some time. The website itself had the following line:

"*....Last update: April 17, 2005;*"

8. Completeness and Coverage

A web page should be complete and cover all areas necessary to access. An incomplete search or a broken link to a webpage would be an error that fits in this category.

A search for "Smith" using a university's directory search resulted in the following

(a) Size Limit Exceeded

The keyword "Smith" was typed in a university's search engine box and resulted in the following display:

"*PHonebook… Size limit exceeded: your search smith returned over 300 entries! Please try a different search;"*

(b) Too Many Entries to Print

Another university's search engine rejected the search criteria "Smith" as being too broad and gave the following message:

"*Too many entries to print; Your query was rejected by the server at directory…edu as being too broad…You need to narrow your query.*"

(c) Only First Name Last Name

For a university, the search succeeded when the criteria were given in the format First name Last Name (J….. Smith). However, when we searched in the format Last name, First name, the following message was generated even though a person with the name "J…… Smith" exists.

"*501:No matches to your query,*"

This error belongs to the server and database layer.

(d) Different Last Name, Different Division

A University's general website search engine was not complete. A search was conducted for a member of the staff in the career services center "Lisa K.....". The results included staff from the law school with a different last name, but did not give us any information about "Lisa K.....".

(e) College Name Not Found in Search Engine

Searching for the full name of the College or acronym in the site's search engine resulted in no matches found.

(f) No Data Checking in Script Layer

On a college's website the faculty web portal lacked a script to check for input accuracy. We typed in the domain incorrectly on the login box, but the authentication process did not give a message regarding the incorrect input. This error belongs to the script layer.

(g) Marquee Stops Scrolling

On the top center of the main page of one college's website there is a marquee box that is similar to a search box. When the text box is clicked, the marquee stops scrolling and disappears. Any text can be typed in the marquee but it would not start scrolling again without refreshing the page.

(h) Only Last Name

On a college web directory, only search by last name is permitted. Any addition to the last name results in the following message:

"*No match found - Click the back button and refine search*"

9. Consistency and Conciseness

A web page should be consistent in the structure, style, and words used, and to the point in delivering its message. An example of a problem in this category would be a search engine that returns results for some entries but does not work for others; however, with some manipulation the search will work.

(a) Inconsistent Display of Contact Information

In one of the college sites studied, the faculty directory results were listed in the format

"*NAME; TITLE; QUALIFICATION; EMAIL,*";

However, for institution executives a small window pops up with the following information

"*NAME; TITLE; EMAIL.*"

(b) Extra Tag with No Title

The navigation tag bar across the top of the College's website had extra space to the right side of the page. There is an empty tag bar with no information in it.

10. Community and Customization

A web site should be designed and tailored to its clients needs.

(a) Contact for Claim Infringement

Community Colleges should consider the mindset of prospective students in their website design, especially when the audience is mainly recent high school graduates. On the left center of the main page of the college's website under the "Quick Links" section there is a link for

"*Contact for Claim Infringement.*"

Clicking on the link gives a full page PDF starting with the following:

> "*Interim Designation of Agent to Receive Notification of Claimed Infringement; Full Legal Name of Service Provider:.....;... __Executive Director of Legal Affairs___;"*

(b) Main Home, Home, Home

The university web page had many different links to the main page entitled "Home." In the main page itself there are three links to the main page.

11. Compelling and Creativity

A website should be impressive and appealing to users as well as reflect the image and cause of its institution. The following are some of the examples of the compelling and creativity errors due to the search box itself or result of search activation: a) There is no search box or content at all, b) The search box or its information are hidden or misplaced, c) The search box color or its content are the same as the web background color, d) Unrelated information is attached to the search box or to its search result.

(a) Smith; Warning openDir

The keyword "Smith" was typed in a College search engine box and resulted in the following display. A page of warnings were generated and at the end there were a listing of links. On activation of the first link the keyword "Smith" was found after scrolling to the third page.

> "*1- smith; warning: OpenDir: No such file or directory (errno 2).....; 548 ;*
> *Warning: readdir(): supplied argument is not a valid Directory resource in /usr/local/www/chris/searchprocess.php on line 555;*
> ;;;
> *http://www.....edu/Alumni/ey10.php Occurrences: 2; ..... – Web.EY – 10 Years Consecutive Report;*"

(b) Invisible Search Box- Unsafe Color

In one college, the HTML background color and the color of the search box were not color safe. The college designed these two elements as the same color which made the search box difficult to distinguish.

12. Conclusion and Future Remarks

The study of all college and university websites on Long Island and a sampling of some non-Long Island institutions indicate that there has been a lack of quality assurance testing to certify the accuracy of information and quality of presentation. The finding of this study has been based on the examination of 7 layers and 7X2 C's web analysis criteria, especially in the area of the search engine layer. These web layers range from school identity, appearance, search engine, HTML, and script programming to server, database manipulation, security, privacy and ethics. The 7X2 C's web analysis criteria range from content, correctness, and completeness to currency, consistency, and compelling. A college web site should consider a quality test plan in these 49 areas.

The study recommends that, given the importance ascribed to search engines, a college search engine should prioritize its search on the official web page of the institution. An official college web page consists of a limited list of important entries and personnel on which the search engine can concentrate and accurately extract. The search engine should look for the entire entry on the official institution site only. When there is no match found in the official web page, the search engine should use some intelligence and do a partial search to locate the desired entry. Use of commercial search engines on college web pages provides an overwhelming amount of information, much of which is irrelevant. Future studies are needed to analyze more websites and involve more colleges and universities nationwide or worldwide. A profile for each of these schools should be created based on each of the seven layers and criteria.

Although the purpose of this study was not to compare Long Island institutions to other institutions, a finding reveals that all non-Long Island colleges and universities including Ivy League schools in this study use Google in combination with the in-house search. A question arises as to whether or not choosing a specific search provider such as Google will affect the neutrality of the net as well as academia?

13. Acknowledgements

Special thanks to students of SUNY Old Westbury, who took the Topics in MIS course in the Fall semester 2006, for participating in this project in and outside of class. Our gratitude is extended to Dr. John Impagliazzo, Chief Editor of SIGCSE who encouraged the publication of this work after seeing the students' presentations at CHE 2006.

References

[1] Cavanaugh, C. & Cavanaugh, T. (2002) College Website Review and Revision. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2002* (pp. 33-34). Chesapeake, VA: AACE.

[2] Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., and Treinen, M. (2001) What makes Web sites credible?: a report on a large quantitative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, United States). CHI '01. ACM Press, New York, NY, 61-68.

[3] Ivory, M. Y. and Megraw, R. (2005) Evolution of web site design patterns. *ACM Trans. Inf. Syst.* 23, 4 (Oct. 2005), 463-497.

[4] Miller, E. (2002) The Web site quality challenge. *The Software Practitioner* (Mar. 2002).

[5] Whitelaw, K. (2003) Why make websites accessible?: and how?. In *Proceedings of the 31st Annual ACM SIGUCCS Conference on User Services* (San Antonio, TX, USA, September 21 - 24, 2003). SIGUCCS '03. ACM Press, New York, NY, 259-261.

[6] Long Island Colleges. http://www.longislandcolleges.com, 2006

Some College and University Websites Used in Developing this Paper

Adelphi University, http://www.adelphi.edu, 2006
Briarcliffe College, http://www.bcl.edu, 2006
C.W. Post, http://www.cwpost.liu.edu, http://www.liu.edu, 2006
Dowling College, http://www.dowling.edu, 2006
Farmingdale State University, http://www.farmingdale.edu, 2006
Five Towns College, http://www.fivetowns.edu, 2006
Hofstra University, http://www.hofstra.edu, 2006
Molloy College, http://www.molloy.edu, 2006
Nassau Community College, http://www.ncc.edu, 2006
New York College of Health Professions, http://www.nycollege.edu, 2006
NYIT - New York Institute of Technology, http://www.nyit.edu, 2006
Polytechnic University, http://www.poly.edu, 2006
St. John's University, www.stjohns.edu, 2006
St. Joseph's College, http://www.sjcny.edu, 2006
Stony Brook University, http://www.sunysb.edu, 2006
Suffolk County Community College, http://www.sunysuffolk.edu, 2006
SUNY Empire State College, http://www.esc.edu, 2006
SUNY Old Westbury, http://www.oldwestbury.edu, 2006
Touro Law School, http://www.tourolaw.edu, 2006
US Merchant Marine Academy, http://www.usmma.edu, 2006
Webb Institute, http://www.webb-institute.edu, 2006
Brooklyn College, http://www.brooklyn.cuny.edu, 2006
Princeton University, http://www.princeton.edu, 2006
Yale, http://www.yale.edu, 2006
Cornell, http://www.cornell.edu. 2006
Stanford, http://www.stanford.edu, 2006

Editor's Note

Among the authors, S. Sayeed, S. Parham, H. Akibu, A. Saeed, and W. Parris are students.