WEB DESIGN QUALITY VERSUS WEB INFORMATION QUALITY

(Research-in-progress)

Anne Matheus
Marist College
Anne.Matheus@marist.edu

Abstract: In the literature on web site quality, there is frequently a lack of distinction between the design qualities of the web site versus the information quality of the web site. This study is an exploratory effort to develop independent instruments to measure design quality as separate and distinct from information quality of the same web site. The intermixing of the two aspects of web site quality has resulted in confusing outcomes. This research will explore and define these two different aspects of web site quality. Assessing the quality of information found on the World Wide Web is of critical importance in the educational arena, therefore this study is being conducted in an academic setting.

INTRODUCTION

In practice and in the literature on the development of web sites, the content of a web site is one aspect of development [1, 10]. There is a difference between the information contained or content of the web site and the design of the web site. It is possible to have a well designed web site that has poor or useless information. When looking for information quality of a web site, the user is looking for accurate, reliable, believable, complete, and timely data. When looking at the design quality of the web site, the user is looking for the structure, text, graphics, style, navigation, and promotion (advertising) dimensions [10]. Do the navigation links work, does the contrast make it easy to read, and are there too many advertisements? These are very different components of web site development. Design quality is usually under the purview of the web developer or programmer while the information quality is determined by experts in that field of study [12].

The problem manifests itself as confusion and errors on the part of the user, who is fooled into thinking that if it looks good, it is good. Is there information in all that data [2]? "Modern computers 'scale' well…the amount of information they can receive, display and store goes up almost without limit. Human beings don't scale." [2, p.BU5] Fuld [6] described it as "Data Slam": the user is overwhelmed by the amount of data that is presented. If there is information in the web site, the user is unable to interpret it or understand it due to the excessive amount displayed. The World Wide Web has changed the flow of information in organizations [15]. Previously,
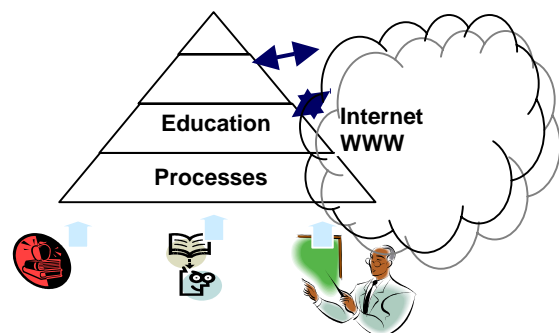
Figure 1: WWW Information Flow

information was filtered by the management information systems group; they decided what information was used and who received it. In academia, this problem is exacerbated by the online learning venue. Figure 1 illustrates the new pedagogy of the online classroom. In the current Internet environment there is no longer a filtering process for information. The professor, textbook, and classroom are no longer between the student and the content. As in industry, students have access to the Internet and are able to download whatever information they feel is important, without the benefit of those traditional filters. How are they judging the quality of that information? Are they judging it at all?

BACKGROUND

This research is an exploratory effort to look at the relationship between the quality of the information content and the quality of the web design. It is an effort to determine if students are able to make the distinction between good web design and good information content. There are three fields of study that have looked at these issues; in psychology, the study of human factors; in information science, the study of information literacy, and in information systems, the study of data quality. From the information systems field, [11] Klein's research looked at the difference in how students evaluated the information quality of web sites based on whether they had received training or not. Significant differences were found between those who had received training versus those who did not. One of the limitations indicated by the author was that students were not assigned the same web sites to evaluate.
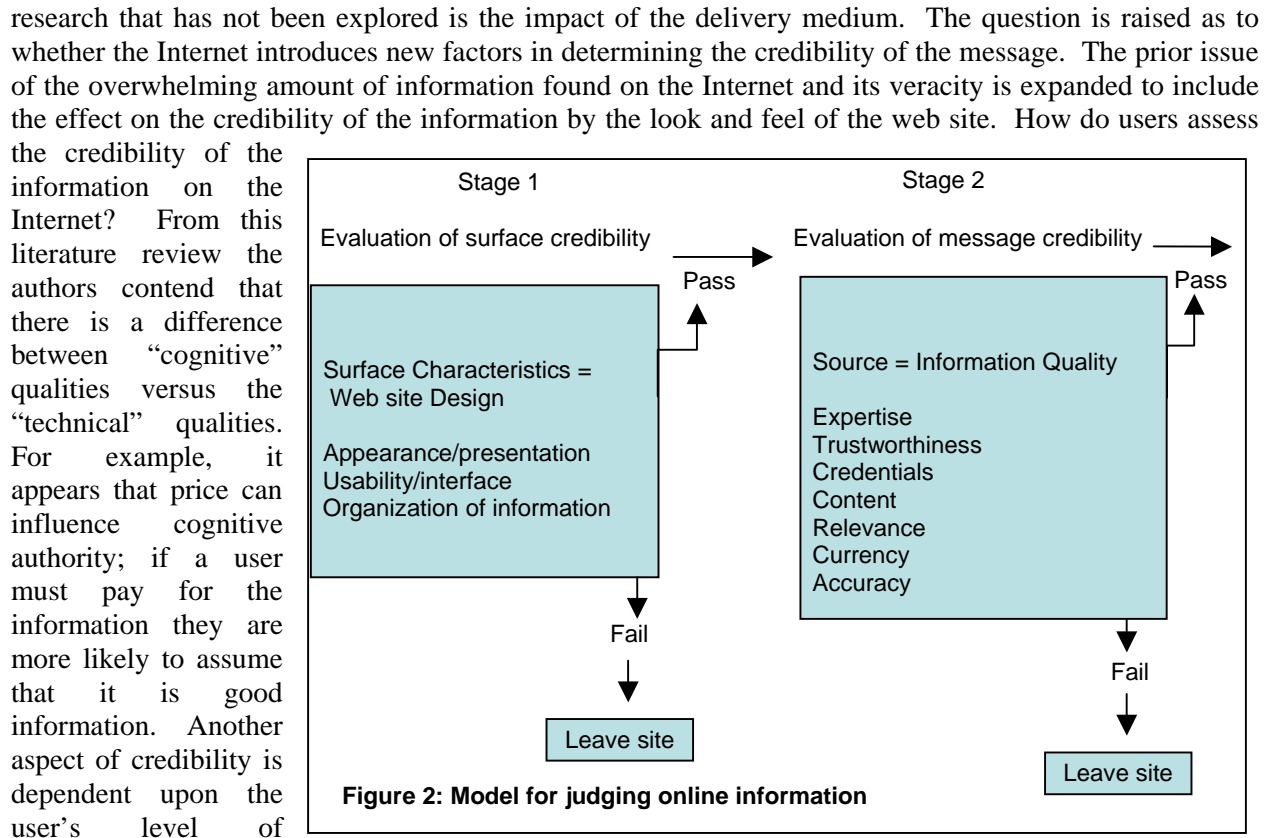
Another study from the information systems arena [9] started out as data quality research, but ended up as a reinterpretation of the data quality dimensions. Katerattanakul and Siau began their study using the data quality framework developed by Wang, et al [16, 17, 18, 19]. However, instead of using the dimensions as they are, the authors modified them to reflect design issues. For example, within the intrinsic category Katerattanakul and Siau [9] redefined it to mean grammatical and spelling errors, workable and relevant hyperlinks, and broken links; thereby reinterpreting the data quality dimensions as web design quality dimensions. Instead of looking at the data quality dimensions they were measuring design issues. The results were inconclusive and confusing. In a study by Zhang, et al [21], the authors again stated they were looking at quality and instead measured the presentation and the navigation functionality of the web sites.

A small body of work that most closely approximates the study of information quality of web sites is in the human factors literature. Initially, this research focused on two factors: trustworthiness and expertise [5]. According to the authors, both factors were components of the dimension called credibility. Subjects evaluated the credibility of a web site by measuring its trustworthiness and the expertise of the site. Fogg, et al [4] later expanded their credibility criteria to include:
- does the site have a real world feel
- is it easy to use
- is it tailored to the audience
- does it have commercial implications
- is it amateurish.

The more work they did [3], the closer they approximated the approach to the data quality conventions developed by Strong, [16], Wand [17], Wang [18, 19] et al.

From the information science field, Wathen and Burkell [20] reviewed the literature, including some authors previously cited, to bring together the current thinking on how people interact with the Internet and suggest research goals. Credibility is determined to be the result of the characteristics or expertise of the message as well as the content and is also dependent upon the perception of the user. A feature of this

research that has not been explored is the impact of the delivery medium. The question is raised as to whether the Internet introduces new factors in determining the credibility of the message. The prior issue of the overwhelming amount of information found on the Internet and its veracity is expanded to include the effect on the credibility of the information by the look and feel of the web site. How do users assess the credibility of the information on the Internet? From this literature review the authors contend that there is a difference between "cognitive" qualities versus the "technical" qualities. For example, it appears that price can influence cognitive authority; if a user must pay for the information they are more likely to assume that it is good information. Another aspect of credibility is dependent upon the user's level of

Stage 1

Evaluation of surface credibility

Pass

Surface Characteristics = Web site Design

Appearance/presentation
Usability/interface
Organization of information

Fail

Leave site

Stage 2

Evaluation of message credibility

Pass

Source = Information Quality

Expertise
Trustworthiness
Credentials
Content
Relevance
Currency
Accuracy

Fail

Leave site

Figure 2: Model for judging online information

expertise; the more technologically savvy user will be less gullible in evaluating web sites. The authors have created a staged process of web site evaluation that forms the basis for this current study. They stress that this is an untested process. In stage one; the user determines the surface characteristics of the web site, the appearance, interface design, usability, and organization. These factors describe web site design characteristics. If this phase is passed, the second aspect is the evaluation of message credibility or information quality. The message is evaluated for expertise, competence, trustworthiness, content, relevance, currency, and accuracy. The final phase in the model proposed by these authors is the user's ability to critically evaluate the web site. Figure 2 is an interpretation of the model proposed by Wathen and Burkell.

In information systems, extensive work has been done by Wang, and others [19, 17, 16, 14] in defining data quality. In their initial work in defining data quality dimensions, Wang and Strong [19] used a factor analytical technique to determine the characteristics that users attribute to data. The initial process resulted in 179 dimensions; using an additional sorting process, the data quality dimensions were broken down into four categories. Category one is intrinsic data quality, two is contextual, three is representative, and four is accessibility. The ontological foundations were established in a theoretical framework to validate these dimensions [17]. In subsequent work [16] the authors delineate the dimensions and categories as indicated in Table 1.

These dimensions of data quality categories are defined as; intrinsic, quality in its own right; contextual, quality within the context at hand; representational, describes ease of use; and accessibility is security and accessibility. This research was done from the perspective of the consumer, trying to answer the question, "is the information fit for the use the consumer has for it?" This perspective implies that information is a product and should be evaluated in terms of whether it has a value to the customer or not. The question needs to be asked: Does the data on the web site add to my knowledge? "A web site is an information product and site visitors are information consumers [18 p. 102]."

DQ Category	DQ Dimensions
Intrinsic	Accuracy, Objectivity, Believability, Reputation
Accessibility	Accessibility, Access security
Contextual	Relevancy, Value-added, Timeliness, and Completeness, Amount of data
Representational	Interpretability, Ease of understanding, Concise representation, Consistent representation

Table 1: Data Quality Categories and Dimensions

Web design measurement uses different categories. Measurement is taken of usage patterns and traffic analysis; metrics include word count, graphics, download time, page size, links, font colors, and text positioning. In their book on designing quality web sites, Tate and Alexander [1] described the categories that should be considered in the development of good web sites. There is some overlap with the data quality dimensions described above and the categories described in [1]. The first category of Authority [1] includes: the extent to which the material is created by a person or organization that is recognized as having knowledge in the given area; and include the following design issues:

- contact information
- copyright indication
- outside authorities listed and linked to
- is there a supervisory organization.

The second category is Accuracy, which is also a data quality dimension, but they [1] describe it somewhat differently. The following questions are raised: Is the information error free, grammatically correct, and are sources provided if it is original. Included in the Accuracy category are design issues; are the graphics well done, are the links to a site working. The third category is Bias, which also overlaps with the DQ dimension of Objectivity; but in this context, bias is interpreted as possible distortion from advertising; and is there a distinction between advertising and information. Currency is the fourth category; can you determine if the information is up to date, is the date listed? In the DQ dimensions, timeliness is an important issue, but from a user's perspective the question is, did I receive the information in the time frame for which I needed it; not is the date on the report current.

In addition to the above categories, there are more categories for coverage and intended audience, interaction and transaction features, navigational features, and the last category is considered non-text features that include Logos and other non-text features like sound or video. There are awards given to the best web designers on a monthly basis. An example is the Judging Criteria at American Design Awards (*http://www.americandesignawards.com/judging.html*), based on the participant's talent, creativity, and potential. The focus is clearly on the artistic efforts.

RATIONALE AND PURPOSE

The fundamental reasons for this exploratory research are to first determine if it is possible to create assessment forms that measure information quality and design quality of web sites. An effort will be made to determine if information quality is separate and distinct from design quality and to measure the relationship between the two. In addition, this research will attempt to determine if there are other factors that affect users' judgments; including, training, field of study, and experience with using the Internet.

The purpose of this research is to find out if users can reasonably assess the information quality of web sites and if that assessment is either positively or negatively affected by the design of the site. By

reasonably is meant and that users will be able to quickly answer some questions and determine to their satisfaction if the site does contain information that has high/usable quality to suite their needs; also, to determine if the credibility of a web site is impacted by the design quality.

METHODS

Web sites were selected with the assistance of a reference librarian using the Google Search engine. The web sites contain content in subject areas that are used by college students from various fields. These fields of study were taken directly from the list of majors offered. For a complete list of the selected web sites see Appendix A. The sites were selected from non-identifiable sites; that is they were not from educational sites or recognized authorities like CNN or a not-for-profit-organization. The goal was to select sites that would not give the user clues as to the quality of the site.

Students (novices in their chosen field of study) were administered a three-part questionnaire. The first part included only demographic data (See appendix B). The second and third parts were interchanged to prevent the outcome from being possibly influenced by the administration of the testing part (Appendix C). The information quality questions were taken form Wang and Strong's work [19]. Web sites were assigned to students based on their recorded major field of study; if a student's major was undecided; a random selection of a site was assigned. Experts in the field who volunteered to participate were randomly assigned web sites to evaluate for information quality, using the same procedures as students. Experts included faculty members and upper-level students including seniors and graduate students.

The same sites that were evaluated for information quality were evaluated for design quality by faculty members in that discipline area and upper level students and graduate students in information technology, who were considered experts in the field. The questionnaire can be found in Appendix D, it is based on the book by Alexander and Tate [1]. The same demographic part and testing part were administered to this sample as well. For purposes of consistency in the design and to attempt to preclude questions of confounding of variables, students who will be studying web programming participated as well as novices in the design quality testing.

RESULTS

There are four groups in this study, novice information users (Group A), expert information users (Group B), novice design users (Group C) and expert design users (Group D). Testing has not yet taken place. Each quality questionnaire will result in a composite score based on the ordinal scale of Strongly Agree = 6 to Strongly Disagree = 1. In order to determine if the question itself is valid and useful in this context, two null possibilities were included in the questionnaires ("Not applicable" and "Could not be determined"); questions that receive the majority of responses as "Not applicable" or "Could not be determined" will be removed from the subsequent research. Expected results will answer the following questions.

- Is there a difference between composite scores of novice information users (Group A) and expert information users (Group B)?
- Is there a difference between composite scores of novice design users (Group C) and expert design users (Group D)?

- What is the correlation between novice information users (Group A) and novice design users (Group C) scores?
- What is the correlation between expert information users (Group B) and expert design users (Group D) scores?
- Additional questions will be asked based on the demographics of age, experience, and field of study as well as the test results of Section 2 in Appendix C.

LIMITATIONS

Since the sample size is small and the representation of the sample may not be generalizable to the general population, findings should be considered within the context of the research venue. However, the importance of the questions in view of the problem presented by information overload on the World Wide Web has far reaching implications.

CONCLUSION

Based on the results of this pilot study, additional research will be completed to further investigate these issues. The importance of a user's ability to assess the information quality of a web site cannot be understated. The findings of this and other research may have a significant impact on how courses are developed both in the classroom and online.

BIBLIOGRAPHY

[1] Alexander, J., Tate, M. (1999). Web Wisdom: How to evaluate and create information quality on the web. Mahwah, NJ, Lawrence Erlbaum Associates.

[2]Fallows, J. (2004). "Humans vs. Computers, Again. But There's Help for Our Side". New York Times. New York: 5.

[3]Fogg, B. J. (2003). Prominence-Interpretation Theory: Explaining How People Assess Credibility Online. CHI, Ft. Lauderdale, FL, ACM.

[4]Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A. Shon, J., Swani, P., Treinen, M. (2001). What Makes Web Sites Credible? A Report on a Large Quantitative Study. CHI, Seattle, WA, ACM.

[5]Fogg, B. J., Tseng, H. (1999). The Elements of Computer Credibility. CHI, Pittsburgh, PA, ACM.

[6]Fuld, L. M. (1998). The Danger of Data Slam, CIO.com. 2004.

[7]Hawkins, D. T. (1999). What is Credible Information?, Information Today. 2004.

[8]Ivory, M., Sinha, R. R., Hearst, M.A. (2001). Empirically Validated Web Page Design Metrics. CHI, Seattle, WA, ACM.

[9]Katerattanakul, P., Siau, K. (1999). Measuring information quality of web sites: development of an instrument. Proceeding of the 20th International Conference on Information Systems, Charlotte, North Carolina.

[10] Kendall & Kendall (2004). Systems Analysis and Design. Upper Saddle River, NJ, Prentice Hall.

[11]Klein, B. (2002). "Internet Data Quality: Perceptions of Graduate and Undergraduate Business Students." Journal of Business and Management 8(4): 425-432.

[12] Miletsky, Jason (2002). Successful Web Sites. Boston, MA. Course Technology.

[13]Pask, J. M., Saunders, E. S. (2004). Differentiating Information Skills and Computer Skills: A Factor Analytic Approach, Portal: Libraries and the Academy. 2004.

[14]Pipino, L. L., Lee, Y. W., and Wang, R.Y. (2002). "Data Quality Assessment." Communications of the ACM 45(4): 211-218.

[15]Segev, A. (1996). On Information Quality and the WWW Impact: A Position Paper. ICIQ 1996, Cambridge, IQ Conference.

[16]Strong, D. M., Lee, Y. W. and Wang, R. Y. (1997). "Data Quality in Context." Communications of the ACM 40(5): 105-110.

[17]Wand, Y., Wang, R.Y. (1996). "Anchoring Data Quality Dimensions in Ontological Foundations." Communications of the ACM 39(11): 86-95.

[18]Wang, R. Y., Lee, Y. W., Pipino, L.L. and Strong, D. M. (1998). "Manage Your Information as Product." Sloan Management Review: 95-105.

[19]Wang, R. Y., Strong, D.M. (1996). "Beyond Accuracy: What Data Quality Means to Data Consumers." Journal of Management Information Systems 12(4): 5-34.

[20]Wathen, C. N., Burkell, J. (2002). "Believe It or Not: Factors Influencing Credibility on the Web." Journal of the American Society for Information Science and Technology 53(2): 134-144.

[21]Zhang, X., Keeling, K.B., Pavur, R. J. (2000). Information Quality of Commercial Web Site Home Pages: An Explorative Analysis. ICIS, ACM.

APPENDICES

Appendix A

http://www.theartgallery.com.au/ArtEducation/greatartists/Botticelli/about/index.html
http://www.artchive.com/artchive/B/botticelli.html
http://www.ibiblio.org/expo/1492.exhibit/c-Columbus/columbus.html
http://www.caribbean-connection.com/christopher-columbus/
http://www.legalethics.com/
http://www.hiethics.com/
http://www.philosophypages.com/hy/index.htm
http://www.philosophypages.com/ph/aqui.htm
http://www.philosophypages.com/ph/hobb.htm
http://www.thomas-hobbes.com/
http://www.snpp.com/other/papers/jk.paper.html
http://www.sntp.net/behaviorism/skinner.htm
http://www.criticism.com/md/weber1.html
http://www.marxists.org/reference/subject/philosophy/works/ge/weber.htm
http://www.designerhistory.com/historyoffashion/callot.html
http://www.designerhistory.com/historyoffashion/patou.html
http://www.fragrancex.com/tierjean.html

Appendix B

Section 1: Demographic Characteristics

For the purpose of comparing users' level of expertise, please complete the following:
1. If you are a student, indicate what year (check one):
       ___ Freshman
       ___ Sophomore
       ___ Junior
       ___ Senior
       ___ Graduate
2. How old are you? (check one)
     ___ 18-25

  ___ 26-40
  ___ 41-55
  ___ 56+

3. Your Major Field of study (check one):

___ Accounting	___ Environmental Science
___ American Studies	___ Fashion
___ Art History	___ History
___ Athletic Training	___ Integrative Studies
___ Biology	___ Mathematics
___ Business	___ Modern Languages
___ Chemistry	___ Political Science
___ Communications	___ Psychology
___ Computer Science/IT/IS	___ Social Work
___ Criminal Justice	___ Studio Art

4. Your highest degree earned. (check one):
  ___ Undergraduate
  ___ Bachelors
  ___ Masters
  ___ Doctorate

5. What would you rate your personal computer skills (circle one)?
  Beginner → 1 2 3 4 5 ← Advanced

6. Have you taken CSIS 158 Information Literacy?
  ___ Yes
  ___ No

7. Have you taken CSIS 103 Information and Computer Literacy?
  ___ Yes
  ___ No

Appendix C

Section 2: (Test of sophistication in use of the web)Questions

With the use of the World Wide Web, please find the answers to the following questions.

1. Find the current weather conditions of your town today, using any web source that you are familiar with.
2. Please find a price for the following digital camera, Canon Powershot S400.
3. Please find yesterday's closing price for Dell stock.
4. What year did Chief Justice Brandeis rule on the right to privacy?
5. Please find the phone number of Coyote Grill Restaurant, Poughkeepsie.

Section 3:
Place a mark (X) in the column that best reflects your perception of the quality dimension listed as it applies to the web page.

	Perception	Not Applicable	Could not be determined	Strongly Agree	Agree	Somewhat Agree	Somewhat Disagree	Disagree	Strongly disagree
1	This information is accurate								
2	This information is believable.								
3	This information is objective.								
4	This information has a good reputation.								

5	Representation of this information is compact and concise.
6	This information is relevant to my field of study.
7	The meaning of this information is easy to understand.
8	This information is represented in a consistent format.
9	Using this information increases the value of my work
10	The information can only be accessed by people who should see it.
11	This information is complete.
12	This information is easy to comprehend.
13	This information is quickly accessible when needed.
14	This information is current
15	This information is of sufficient volume for my needs.

Appendix D

Section 3 (for Design questions):
Place a mark (X) in the column that best reflects your perception of the quality dimension listed as it applies to the web page.

	Perception	Not Applicable	Could not be determined	Strongly Agree	Agree	Somewhat Agree	Somewhat Disagree	Disagree	Strongly disagree
1	Are the graphics well done?								
2	Are there links to outside resources?								
3	If under construction, is the expected date of completion indicated?								
4	If the site has different types of pages, are they indicated?								
5	If cookies are used is the user notified?								
6	Is there a feedback mechanism available?								
7	If downloading is restricted, are they listed?								
8	Does the browser title indicate responsible organization?								
9	Are internal directional links consistently placed on each page?								
10	For links not on this site is there and indication that the user is leaving the site?								
11	Is there a site map or index on the home page or on a page directly linked to the home page?								
12	Does the site map or index include at least the main topics?								

13	Is it easy to read?
14	Organized in a logical manner?
15	Is there a text alternative to the logo?
16	If sound or video available, are there directions for downloading the necessary software?
17	If special software is necessary is another file format available?
18	If a specific browser is needed or a specific version is needed is it indicated along with directions how to get it?