An Information Quality Framework for College and University Websites

Preprint of proceedings paper: ITNG 2021: 18th International Conference on Information Technology—New Generations, Apr. 11-14, pp. 509-518 (pub. by Springer).

Joseph Elliot and Daniel Berleant

Abstract

Information Quality (IQ) of a university website plays a major role in the decision process for prospective students when selecting a university for their higher education. Furthermore, current students and others rely on university websites for many other purposes. In this paper we identify university website information quality dimensions relevant to prospective and current students and other site users. We discuss the rationales for identifying these IQ dimensions and propose a University Website Information Quality (UWebIQ) framework to quantify the individual IQ dimensions as well as a strategy for defining the composite IQ for such a website. The outcome of this research is expected to provide insight for universities that wish to maximize the fitness for use of their websites.

Keywords

Composite IQ · Data Quality · Dimension · Dimensional IQ · Enrollment · Information Quality · Internet · WebIQ · World Wide Web · Website

1 Introduction

Information Quality (IQ) describes an information product's fitness for use for a specific task [1]. Yet this fitness is itself often a composite of a number of different, more specific dimensions. Each contributes to the overall information fitness. How to calculate an overall quality (fitness) from ratings along these disparate dimensions is important because a single, summarizing quality rating is often needed. In addition to multiple ratings along different dimensions there may be multiple ratings of the same dimension. This problem is somewhat similar to the need to calculate a single overall quality rating using multiple estimates. Both of these problems present a combination of evidence challenge, a challenge that can be addressed using the approach which we present here.

Information quality (IQ) theories and frameworks have been increasingly researched, practiced and standardized at major government and non-government entities in order to assess, monitor and improve the quality of their information products [2][3]. United States government agencies have enacted information quality guidelines [4]. For example, the Department of Health and Human Services (HHS) has published web standards required for the design and development of all HHS and priority websites [5]. The Energy Information Administration (EIA) produces and publishes standardized energy-related data such as world crude oil reserves by regions and countries, and relies on well-developed information quality techniques. In the education industry, however, universally accepted standards for university websites and for their in-depth analysis, assessment, and quantification with respect to their information quality is currently lacking. Thus, there is a need for web information quality frameworks for quantitatively assessing the qualities of university websites.

In the education industry, university and college websites should support efforts to recruit and retain students, reduce costs for staff to provide information to stakeholders individually and manually, lower costs related to physical printing, support pedagogical, research, and administrative functions, and so on. Fitness for use of university websites implies support for all of those functions and more. As one example, unavailable or inconsistent website links and low usability user interfaces could lower new student applications and enrollments. To characterize fitness for use of college and university websites, we describe a new information quality framework for this purpose based on accepted theories and foundations in the information quality domain. Using this proposed framework, it is possible for colleges and universities to design and implement upgrades of their websites based on objective computation of the information qualities along key dimensions and overall based on information qualities along all the dimensions.

Joseph Elliot is with Rutgers University, Newark, NJ, 07102, USA (corresponding author, phone: 973-353-1766; e-mail: joseph.elliot@rutgers.edu).

Daniel Berleant is with the University of Arkansas at Little Rock, 72212 USA (e-mail: jdberleant@ualr.edu).

2 Literature Review

Measuring information quality often requires combining different dimensions of quality which are all relevant in different ways to an overall quality assessment. Stvilia *et al.* (2007) [6] proposed a "general framework for IQ assessment." They explained that aggregation, or clumping together of different things, refers in the information quality field to both grouping different but related information entities together, and combining measurements of information quality to get an overall quality rating. The latter type of aggregation is more relevant to the present report. They advocate a process that is complicated, involving concept trees, metrics, measurement representations, and IQ activity system value structure. A key subgoal of one approach is to identify and combine IQ value curves using transfer functions. An alternative approach is based on statistical profiles of IQ user evaluations. Two case examples both use factor analysis. Ultimately the paper provides a general framework but without specifying a particular algorithm for combining IQ measurements, which is a gap we and others have sought to bridge.

Robbins and Austin (1986) [7] advocate a process that contrasts with that of Stvilia *et al.* by avoiding the complexities of a multi-faceted process which has no clear step by step procedure. Their method combines information quality ratings on different dimensions by simply multiplying them. This was a reasonable model for the specific dimensions they chose, which were importance and completeness of the information item. This illustrates a model based on the specific meanings of stated dimensions. However, multiplying the information quality measures does not generalize well to arbitrary dimensions. It is ad hoc, and thus not suitable as a general model of combining information quality dimensions.

Robbins and Austin also combine quality ratings by an averaging process. This may be not unreasonable for combining multiple assessments of the quality of one information item. However it is less defensible for combining ratings on multiple dimensions for one information item because a quite low rating on an important dimension carries considerable weight that cannot be easily compensated for by high ratings on other dimensions. For example, if timeliness is an important dimension and the information is not timely, its quality is significantly affected even if it rates highly on other dimensions. Thus there is a need for combination methods that circumvent this issue.

Quarati *et al.* [8] use an Analytic Hierarchy Process based approach to combine quality ratings of different dimensions in stages. First, dimension measurements are combined to give a quality rating to a category of dimensions, and then the category ratings are combined to give an overall rating. The method generalizes to different hierarchies, such as combining subdimension ratings to get a dimension rating, then dimensions to get a dimension category rating, then categories to get an overall rating. Each rating to be combined with other ratings is given a weight determined by crowd sourcing expert judgments, and the weighted ratings are averaged. Although, as we will see, simple averaging is problematic, the approach is certainly adaptable to other combination methods as well. It is suitable in general to problems in which a hierarchical organization of the combination process is indicated.

A. Identifying Information Quality Dimensions

Tao *et al.* (2017) [9] describe a process of identifying information quality dimensions based on user studies. They focus on the problem of identifying information quality dimension indicators, or attributes, which can be used to assess dimensions. They call these indicators drivers, and define a driver as an attribute of an information quality dimension that is perceived by users as indicating quality. In their investigation into the domain of health information websites, there is often ambiguity about whether a driver is really a dimension or vice versa. Additionally, they found that some drivers indicate the level of quality on more than one dimension. However, they did not address combining dimensions to produce a composite, overall information quality rating.

While information quality dimensions can be identified on a problem-dependent basis, another approach is to use general purpose dimensions that apply to a wide variety of problems and are defined by standardized references. An example of this approach is the 21st Century Integrated Digital Experience Act of 2018 (e.g. digital.gov/resources/21st-century-integrated-digital-experience-act). This standard applies to the design and development of all HHS/OS and priority websites [5]. The Act requires Federal Agencies websites to satisfy certain quality dimensions [10]. Specifically, they must: (1) be accessible to people with disabilities, (2) be consistent in appearance, (3) be authoritative in that they avoid redundancy with other websites, (4) have a site search capability, (5) have appropriate security, (6) “be designed around user needs with data-driven analysis,” (7) be customizable to user preferences, and (8) work properly on mobile devices. A related standard provides a checklist of requirements for federal websites and digital services via the digital.gov site, which endorses guidelines and standards including from ref. [11].

Another influential set of general purpose information quality dimensions is Morville’s User Experience Honeycomb, which defines a set of dimensions in order for information to provide a meaningful and valuable user experience [12]. These dimensions are posed as requirements for websites, specifically: (1) useful, (2) usable, (3) desirable, (4) findable, (5) accessible, (6) credible, and (7) valuable. These are adopted for example by Semantic Studios [13]. In contrast to the aforementioned 21st Century Integrated Digital Experience Act of 2018 standard, we have found that the Honeycomb can be applied to a college or university website setting without modification because redundancy with other websites (dimension 3 of [5]) is less of a concern.

Some dimensional frameworks have been studied specifically with respect to college and university websites. For example Mentes and Turan [14] focus on usability evaluation of the website of a particular institution, Namik Kemal University. The authors survey previous papers on assessing the website usabilities of specific universities. Given their focus on usability they need a framework that provides dimensions of usability. They list QIS, SUMI, NIST Web Metrics, MUMMS, and WAMMI, finally choosing WAMMI (**W**ebsite **A**nalysis and **M**easure**m**ent **I**nventory) for their study. Clearly usability is an important aspect of the fitness for use and thus the information quality of college and university websites. However usability alone is too limited to be the sole focus for assessing college and university website quality. That is why we use a more comprehensive framework based on Morville's User Experience Honeycomb.

B. Measuring an Information Quality Dimension

A related problem is combining multiple measurements of the same quantity to get an improved estimate of the quantity. Well established formulas exist for estimating the error and value of a metric composed of multiple measurements each with its own error [15]. The method gives the error in an estimate of a quantity as the square root of the sum of the squares of the errors of the components which compose it. Put another way, it applies the formula of the Pythagorean theorem. This approach could potentially be applied in estimating the information quality of a university website. For example, assume the true information quality for a dimension is Q, and two estimates or measurements of the dimension are Q_1 and Q_2 with errors e_1 and e_2. Then the method estimates the value of Q as $(Q_1 + Q_2)/2$ and the error as $+–\text{sqrt}(e_1^2 + e_2^2)/(N–1)$ where N in this case is 2.

The problem with applying such a method to information quality ratings is that the error or uncertainty in a rating of information quality along a given dimension is difficult to determine. Yiliyasi (2012 [16]) proposes a somewhat complex but probabilistically well motivated approach of addressing this problem of combining information quality. That method was described for the problem of combining expert estimates of the same value, where each expert estimate has its own information quality. The method may be adaptable to the problem of combining assessments of the information qualities of different indicators of a given dimension. These indicators have also been called aspects, attributes, indicators, drivers, quality markers and surrogates (e.g. Tao *et al.* [9]). However combining indicators of a given dimension is not the same problem as combining different dimensions to give a high level or overall information quality rating, as required by the information quality framework described in the next section.

3 Design of an Information Quality Framework

Information can be defined as the product of an information system that processes raw data into a usable product that adds value for the information consumer [1]. Accordingly, we consider a university website an information product whose quality we wish to quantitatively assess as a composite of measurements along multiple dimensions.

We have selected IQ dimensions that are important to the quality of university websites from the perspective of their users, especially prospective and current students, and seek to quantify them in a transparent and reproducible manner. The rationale for using each of these IQ dimensions is discussed in detail in the following sections along with their application to university websites. However, it should be noted that there are other possible IQ dimensions such as *Trendiness*, *Attractiveness*, *Interaction*, *Theme*, etc., that might also be useful to include in a framework of this type if the practical difficulties in assessing and quantifying them can be solved. We identify the following IQ dimensions for our UWebIQ Framework. This list of seven factors, given at usability.gov, is the same as that proposed by Semantic Studio. That is because both are based on Morville's User Experience Honeycomb, which has those same seven factors. Building on that foundation of Morville's Honeycomb, we develop it further with quantitative metrics that help assess and improve the quality of university websites.

Here are the seven dimensions and the measurement process for each.

A. Useful

By useful is meant that content should be original and fulfill user needs. This may be quantified as follows.

a. Universities may determine how much money they spend on answering student and prospective student queries by email or phone (excluding communications with their instructors). The cost might, for example, be determined from salaries, time sheets, and/or budgets. A website that reduced this cost would have a higher quality. Since the cost is higher for *less* useful websites, it needs to be massaged so it is instead higher for *more* useful websites. An easy way to do that is to subtract it form 1.

The figure should also be normalized to a defined scale, here 0-100, so that it can be more easily compared and combined with other measures of information quality. A formula that meets these requirements is:

$$100 * \left(1 - \frac{\text{Cost of answering queries by humans}}{\text{Total technology infrastructure cost}}\right).$$

This is based on ongoing cost of answering queries as a fraction of the total ongoing technology cost. Like any fraction of a total it has values between 0 and 1. That fraction is inverted by subtracting it from 1, then expanded to the range [0, 100]. To evaluate website improvements with respect to this metric, the formula can be recalculated periodically. If its value increases, that means the institution had benefited financially, validating the changes as improvements in website quality.

b. As an alternative, selected web pages could require a viewer to answer a single question about usefulness in order for the viewer to be permitted to follow a link from the page or submit a form. The response to the question is scored on a 0-to-100 range and the average over all users answering the question is the usefulness rating of the page. The overall site usefulness rating is the average of the page ratings over all assessed pages.

Higher values indicate better user ratings, validating the metric.

c. As a third option, uses of online real time help chat windows could be counted, as well as the fraction of those uses that result in providing a website link. The site is more useful if a link can be provided to resolve a chat session. This option may be applied to a website if it has help chat session functionality. The resulting count may be normalized by multiplying by 100. The formula is:

$$100 * \left(\frac{\text{Number of help chat sessions resulting in a link to a website page}}{\text{Total number of help chat sessions}}\right)$$

The range of the fraction in this formula is from 0 (indicating help sessions did not result in links, hence a low information quality for the site) to 100 (all help sessions resulted in a link). Higher values indicate more a more effective business process, likely due to a better website but also potentially due to better employee training or other non-website activities.

Some universities may choose to implement 2 or 3 of the above three options. This would require combining the 2-3 metrics into a single measure of usefulness. The mean of the metrics would work for this [18].

B. Usable

By usable is meant that the website is easy to use. To quantify this, universities can determine how many queries by email and telephone they get that are answered on web pages, relative to the total SSCHs (student-semester credit hours) or other measure of university instructional load. This can be done by requesting the relevant employees to tabulate the number of such inquiries they deal with. They can be incentivized to do this by pointing out that the information will be used to make the website more usable in order to reduce the number of emailed and called-in inquiries they have to deal with. A formula that can model this is:

$$100 * \frac{\text{SSCHs}}{\text{queries} + \text{SSCHs}}.$$

This formula has the value 100 if there are no such queries, suggesting all web pages are easily found and the website is sufficiently usable. On the other hand, the formula descends toward the value 0 when there is a very large number of queries, suggesting the website has poor usability.

Alternatively, real time chat window uses that result in providing a link to a web page could be counted, since the site would be more usable if the user could find the pages themselves. A formula based on this approach that would work is:

$$100 * \frac{\text{SSCHs}}{\text{chat window uses resulting in a link+SSCH}}.$$

This has values approaching 100 when there are very few such chat window uses, suggesting a highly usable website where people can readily find the pages they need. The value declines for more chat window inquiries resulting in a page link, suggesting a website with pages containing the information people need, but they are not finding those pages.

If both metrics are available, then they may be averaged. Improvements in this metric represent more effective communication to users from the website, thus validating the metric.

C. Desirable

The content of the website should evoke users' positive emotions and sense of appreciation so that users desire to use the website [12]. A standard approach to this that students are already familiar with is the 1-5 star rating system. It is commonplace on the web and could be applied here. In one approach, web pages on a university website could each end with a brief request for the reader to click 1 to 5 stars. This system is readily implemented because 3rd party providers make it easy to add this to a web page. An example service is at rating-widget.com. The overall rating of the site would be based on a weighted average of the ratings of the individual pages. The weight would be based on the number of ratings a page has. More ratings should increase the weight of the overall rating of a page, since a heavily rated page is probably a heavily used page, and the desirability of a page is more important the more frequently it is visited. The simplest way to meet this requirement is just to add up all the ratings regardless of what page any rating is for, then divide by the total number of ratings. The rating for the website would then be:

$$S = \frac{\sum_{i=1}^{p} \sum_{j=1}^{n_i} r_{i,j}}{\sum_{i=1}^{p} \sum_{j=1}^{n_i} 1}$$

which says, for each of the p pages, add up its n ratings r, then add up all of the p subtotals. Finally divide by the total number of ratings across all the pages. Note that each $r_{i,j}$ will be a number between 1 and 5 (because of the 1-to-5 star rating system). Normalizing to the interval from 0 to 100 is thus done by mapping the interval of possible values of S, which is[1, 5], to [0, 100]:

$$\textit{Desirability} = 100 * \frac{S-1}{4}.$$

D. Findable

A website is findable to the degree that web search engines list it with a ranking that brings it to the attention of people querying. To measure this, use a major search engine, such as google.com. Determine a set of queries that are believed to be representative of those used by people who the university wishes to reach through web search queries. The findability will be the average ranking of the university's website over those queries. This will be useful for testing whether new SEO (Search Engine Optimization) actions embedded in the website are working better than previous ones. To implement this, a university can:

1) determine a set of queries and whether it would be reasonable, for each one, to hope it would be the top hit on a specified search engine (e.g. google.com), or that it at least be in the top 10 hits;
2) check each query to see if it meets the stated requirement (top hit vs. top 10 hits);
3) score each query at 1 if it meets its requirement or 0 if it does not;
4) let q = 100*(average of the scores of the queries).

Then q is the information quality of the website along the *findable* dimension. Improvements in this metric represent a clearly desirable change to the findability of the website, validating the metric.

E. Accessible

There are various standards that exist for this and can serve as the foundation of a metric. Semantics Studio's definition of the *Accessible IQ* dimension is that a website should be accessible to people with disabilities [13]. Usability.gov refers to *Accessible IQ* as *Accessibility IQ*, which focuses on how people with disability would access or benefit from a website [12]. Lee *et al.* (2002 [17]) define *Accessibility IQ* as the ease at which an information product is retrievable, obtainable and quickly accessible when needed [21]. The Bureau of Internet Accessibility (BOIA) provides a free website accessibility compliance scan [18]. BOIA [19] endorses the W3C's WCAG [11].

Many websites are not compliant with accessibility standards. A common example is that mobile websites do not always provide the same interface and links as the standard desktop interface.

F. Credible

The Stanford University Web Credibility Research site lists 10 guidelines for website credibility [20]. These are described next with how they may be measured for the present web quality framework.

1) *Information accuracy should be easily verified*. We wish to measure this in a way that is readily measured yet is also a good proxy for information accuracy. To this end, recall that scholarly publications use citations interspersed within text as the classical method for providing verifiability of statements. The analog of this in a hypertext context, of which the web is the main example, is links embedded in the content of a web page. It is relatively straightforward to count the fraction of pages on a website that contain such links. A formula that naturally ensues as a metric scaled from 0 to 100 is:

100 * (Fraction of pages on a site with link(s) embedded in their content).

2) *Give the organization of the website*. One standard way to show website organization is to show a horizontal navigation menu at the top of the page. Other standard methods include hamburger menus, sidebars, fat footers, and breadcrumbs (hierarchical indicators showing the depth and location of a page within the website). All of these are popular web page design constructs (e.g. [21]) and are readily transformable into metrics by simply determining the fraction of web pages on a site that include one or more of these website organization techniques. Scaling from 0 to 100 gives the formula:

100 * (Fraction of pages on the website showing website organization).

For some sites this will be near or at 100, since many sites will have all pages containing a standard site organization element. On the other hand, some sites might score very low.

3) *Display the content expertise behind the website*. One approach to determining the degree to which expertise is highlighted is to check whether there are page footers giving information on the expertise of contributor(s) to the page. Statements of expertise are more credible if the contributor(s) with the expertise are identified by name. This leads to the following formula, which is scaled from 0 to 100:

100 * (Fraction of pages with footers giving the expertise of at least one *direct* contributor to the page content).

A simple statement like "This content was written by J. Smith, the web administrator" meets the requirement. A statement like "The website manager, J. Smith, PhD, has 15 years of experience in managing websites for educational institutions and Fortune 500 corporations" does not give the expertise behind the information on the specific page containing that footer, and so does not meet the requirement.

4) *Show "honesty and trustworthiness*." User comments could form a reasonable proxy for this. Since each page is different, it makes sense to provide this service for each page, in a standard footer for example (although an alternative measure could be devised for a comment service provided for the website as a whole). A formula that arises naturally is based on the fraction of pages that have a link to a user comment service that shows previous comments and allows new ones. Such a formula is:

100 * (Fraction of pages on the website that accept and publish user comments).

This formula provides a number from 0 to 100 and expresses an assessment of the honesty and trustworthiness of the website.

5) *Ensure ease of contacting the author*. This is achievable by having web page footers that give contact information for the person responsible for authoring the page. An email address is sufficient. A formula arises straightforwardly out of that criterion:

100 * (Fraction of web pages on the site that provide a contact for the author).

6) *Site is professional and appropriate*. The terms "professional" and "appropriate" are too vague to be easily measurable here, but attention to presentation quality such as lack of typos, good grammar, and quality page design provide some indication of professionalism and appropriateness. Of these, quality of page design is the hardest to measure. An inexpensively produced page can show a higher quality of design than a more elaborate design using poorly chosen graphics, so automatic measurement of page design quality is problematic. However typos such as misspellings and grammar issues are more amenable to automatic checking,

especially spelling. A page could have problems in this area or not, and the proportion of pages that have no detectable problems can be a proxy that provides the basis for a measurement. A formula that would provide the desired calculation is:

100 * (Fraction of pages on the site that have no detectable spelling or grammar errors).

7) *Website is useful and easily used*. This is already accounted for under items 1 and 2 (Useful and Usable) above.

8) *Website is current*. Some pages retain currency more than others, thus needing to be updated less often. Thus, it is difficult to automatically determine whether a page is current since the date of the last update does not in itself determine if a page is current. Pages with news may lose currency in a day, while other pages such as maps may stay current for a much longer time. However, what could be determined automatically is the date of the most recent update to the page, if the page provides that information in its footer. Using this a proxy, we can simply measure what fraction of pages provide their most recent update date in the footer. This gives the formula:

100 * (Fraction of pages on the site that contain the date of the most recent update in their footers)

9) *Avoid promotional ads unless there is a good reason to have a few*. An academic site normally needs few ads for commercial products, and such ads are likely to detract from the quality of the site. The fraction of pages that do not contain commercial advertisements measures this characteristic, giving the formula:

100 * (Fraction of pages on the site that contain no commercial advertising)

10) *Avoid even small errors.* This attribute focuses on seemingly minor problems like typographical errors and broken links. These errors are significant, especially when they occur more than rarely, as they "hurt a site's credibility" [20]. Larger errors are already covered by measurement strategies listed earlier, in particular: 1) *Information accuracy should be easily verified,* 3) *Display the content expertise behind the website*, 4) *Show* "*honesty and trustworthiness*." and 8) *Website is current*. In addition, grammar and spelling errors are covered under 6) *Site is professional and appropriate*. For this criterion, then, we define a measurable proxy using the fraction of links on the website that are broken. This leads to the formula

100 * (1 - (Fraction of links on the website that are broken)).

Factors 1) through 10) above all contribute to credibility, and those that are used may be averaged to determine the website's information quality for the *Credible* dimension.

G. Valuable

This dimension is closest to the information quality dimension of *Value-Added*, which is a measure of the extent to which data is beneficial and provide advantage from their use [3]. In the context of a university website, the *Valuable* dimension implies that visit to the website should be beneficial and provide advantage for visitors. According to [13], a non-profit site is valuable if the "user experience … advance[s] the mission."

Since in a higher education environment we are typically dealing with a non-profit entity, we would like to measure advancement of the mission. What is the mission? While individual institutions will have their own mission statements, generically these missions will tend to be variations of promoting and providing education. So how well does a university website promote and provide education? Authorized students may access learning materials that are hidden from the average web user, for example because they are provided by a Learning Management System that is password protected. However, learning materials available to the public would often be accessible on the web to anyone using a web browser. Such publicly accessible learning materials could truly be said to both promote and provide education. They promote it by reaching out to the general public and not just paying students, and they provide it by containing educational content. Such content might be anything from extension service articles, to course materials that individual instructors make available on the web, to scholarly works that faculty have written and make available for download from the university website.

To measure the extent to which such learning materials are provided on the website, some heuristic approach to quantification is needed. It cannot be an exact measure of total educational content because that is probably impossible to quantify reliably. It may be possible, however, to measure the fraction of university faculty involved in producing these offerings, by sampling individual faculty web pages and checking for learning materials. The fraction with links to publicly accessible educational materials is then the basis for a metric, and this metric is heuristically speaking a reasonable measure of breadth (though not depth) of offerings relative to institution size. This leads to the following formula:

$$100\,\frac{\text{(number of faculty with qualifying links)}}{\text{total number of faculty}}$$

4 Combining Information Quality Measures

A. Combining dimensions

We have elucidated seven dimensions of data quality for a university website. These dimensions characterize the quality of a website in significantly different ways. As a result, it is possible for a site could rate highly on one dimension but poorly on another. The existence of an important dimension on which a site rates poorly suggests the site has a serious problem. Even if it rates highly on all other dimensions, a low rating, such as on the currency or on the valuable dimension, or any other important dimension, suggests a critical weakening of the overall quality. A high quality site is required to rate highly on all important dimensions, not just almost all. Therefore taking the familiar arithmetic mean of the ratings on the different dimensions would not be the best way to combine the ratings.

Since taking the arithmetic mean is not a good way to combine the dimensions, what is a better way? The *geometric mean* is another kind of average, and it does meet the requirement that one low rating makes a big difference, usually considerably bigger than for the common (arithmetic) mean. To calculate the geometric mean, multiply the applicable *n* values and then take the *n*th root of that product. For example, consider seven dimensions all with quality ratings of 100. The geometric mean is $(100*100*100*100*100*100*100)^{1/7}=(100^7)^{1/7}=100$, the overall quality. But one low rating of 10 instead of 100 is quite influential, giving a geometric mean and overall quality estimate of

$(100*100*100*100*100*100*10)^{1/7}=(10*100^6)^{1/7}=71.97$,

quite a bit lower than 100 and in fact lower than if all seven dimensions had very questionable qualities of say 75 (in which case the geometric mean would be 75). By comparison, the common arithmetic mean would be 87.14, higher than 71.97 and not as plausible as an overall quality when a critical dimension is very low. Thus, as a method of combining the dimensions the geometric mean seems to provide a better approach than the more common arithmetic mean.

Still, a very low quality on an important dimension perhaps ought to be required to be harder to compensate for than that: 71.97 still seems high when a critical dimension is as low as 10 out of 100. Is there an even better average? The *harmonic mean* does seem to be better in this way. The harmonic mean when all seven dimensions have the value 100 is

$7/(1/100+1/100+1/100+1/100+1/100+1/100+1/100)=7/(7/100)=7*100/7=100$.

When one of the seven dimensions has the low value of 10, the harmonic mean is

$7/(1/100+1/100+1/100+1/100+1/100+1/100+1/10)=7/(6/100+1/10)=43.75$.

This seems more reasonable, in that a very low dimension has an overwhelming effect, yet several very high dimensions do manage to help significantly. In this, the harmonic mean differs from taking the minimum rating across the dimensions, thereby ignoring the inputs provided by the other dimensions, which does not seem desirable. We conclude that the harmonic mean models the problem the best of these three averages. However, a more comprehensive approach to the question of how to combine dimension ratings would be a useful and important topic for future researchers.

B. Combining multiple measurements of a single dimension

For some of the dimensions we have given multiple measures that address different indicators of the quality of the dimension. These indicators are often more fungible than the dimensions themselves, in that scoring high on one subdimension compensates to a significant degree for scoring lower on another. If this fungibility does not hold, it suggests that the dimension is actually more than one dimension, and the indicators in question are for different dimensions. With these considerations in mind, combining subdimensions can be done using the familiar arithmetic mean.

5 Conclusions

University websites share many key characteristics. They typically share considerable similarities in both form and function. These sites serve as an institution of higher education's public face as well as a portal for its prospective and current students, faculty, and other staff for numerous learning and business functions. Thus, the website is a key strategic component of today's universities. Consequently, the quality of these websites is a critical factor in the business of being a university in the 21st century. Because of the value of high website quality to universities, it is important for universities and their Chief Information and Chief Data Officers to be able to assess their website's information quality across the relevant dimensions, calculate an overall summary quality rating, and understand specific dimensions and website characteristics needing quality improvement.

We have described a method for assessing the information quality of university websites designed to be (i) suitable given the distinctive characteristics of university websites, and (ii) readily, objectively, and quantitatively determinable. Its suitability is in part because it incorporates various relevant dimensions that both capture different aspects of quality and collectively provide reasonable coverage of the functions that go into quality of a university website. Regarding the problem of determining the quality of a site and of its component dimensions, we provide actionable methods for calculation, both of individual dimensions of overall site quality.

Determining overall site quality requires a way to combine the qualities along the various different dimensions in a meaningful way. When all dimensions are important and a low score on any one of them seriously impacts the information quality of the site, a combination method is needed that models this. We found that the harmonic mean of the qualities of the various dimensions worked better in this way than the commonly encountered arithmetic mean, as well as better than the geometric mean.

Assessing the quality along a given dimension can often benefit from combining assessments of different aspects or indicators of that dimension. Here, typically the common arithmetic mean is a reasonable way to combine these aspects because they tend to have relatively equivalent effects on the quality along a dimension, whether positively or negatively, and whether their values are high or low. Thus the website quality assessment design we offer uses the arithmetic mean for combining different measures of the same dimension.

Universities and consultants will find the method described herein readily applicable. In addition, university websites share many common goals with websites of other kinds of organizations. Therefore, the approach we describe is expected to be readily adjusted to assessing websites of other kinds of organizations. Software could partially automate the assessment process in many though not all dimensions because many dimensions and aspects of dimensions can be measured partially or fully automatically. Web maintenance teams can use the method we describe as a guide to improving and maintaining high quality websites.

References

1. R. Wang, "A product perspective on total data quality management," *Communications of the ACM*, vol. 41, no. 2, Feb. 1998.
2. Y. Yiliyasi and D. Berleant, "World oil reserves data information quality assessment analysis," in *The Proceedings of the 16th International Conference on Information Quality*, Adelaide, Australia, 2011.
3. C. Fisher, E. Lauria and S. Chengalur-Smith, "*Introduction to Information Quality*," AuthorHouse, 2011.
4. "Information quality guidelines," online, www.eia.gov/about/information_quality_guidelines.php, accessed 13 April 2019.
5. "Health and human services web standards," online, webstandards.hhs.gov/standards, accessed 2019.
6. B. Stvilia, L. Gasser, M. B. Twidale and L. C. Smith, "A framework for information quality assessment," *JASIST*, vol. 58, no. 12, pp. 1720–1733, 2007.
7. W. A. Robbins and K. R. Austin, "Disclosure quality in governmental financial reports: an assessment of the appropriateness of a compound measure," *Journal of Accounting Research*, vol. 24, pp. 412–421, 1986.
8. A. Quarati, R. Albertoni and M. De Martino, "Overall quality assessment of SKOS thesauri: an AHP-based approach," *Journal of Information Science*, vol. 43, no. 6, pp. 816–834, 2016.
9. D. Tao, C. LeRouge, K. J. Smith and G. De Leo, "Defining information quality into health websites: a conceptual framework of health website information quality for educated young adults," online, doi.org/10.2196/humanfactors.6455, accessed 24 April 2020.
10. "H.R. 5759 – 21st Century Integrated Digital Experience Act," 2018, www.congress.gov/bill/115th-congress/house-bill/5759/text, accessed 2020.
11. "Web content accessibility guidelines (WCAG) overview," online, www.w3.org/WAI/standards-guidelines/wcag, accessed 25 Oct. 2019.
12. "What & why of usability: user experience basics," online, www.usability.gov/what-and-why/user-experience.html, accessed 7 June 2019.
13. Semantic Studios, "User experience design," online, semanticstudios.com/user_experience_design, accessed 6 June 2019.
14. A. Mentes and A. H. Turan, "Assessing the usability of university websites: an empirical study on Namik Kemal University," *The Turkish Online Journal of Educational Technology*, vol. 11, no. 3, pp. 61–69, July 2012.
15. A. Usher, "Errors: What they are, and how to deal with them," online, www.physics.rutgers.edu/grad/506/errors.pdf, accessed 13 April 2019
16. Y. Yiliyasi, *The ΣIQ method: an information quality perspective on oil data*, dissertation, Lambert Academic Publishing, 2012.
17. Y. Lee, M. Strong, B. K. Khan and R. Y. Wang, "AIMQ: a methodology for information quality assessment," *Information & Management*, vol. 40, p. 133–146, Dec. 2002.
18. "Bureau of Internet Accessibility," online, www.boia.org/products/free-wcag-2-0-aa-report, accessed 25 October 2019.
19. "What are the four major categories of accessibility?" online, www.boia.org/blog/what-are-the-four-major-categories-of-accessibility, accessed 25 Oct. 2019.
20. Stanford Persuasive Technology Lab, "Stanford university web credibility research," 2004, online, credibility.stanford.edu/guidelines/index.html.
21. "Website navigation: tips, examples and best practices," online, www.crazyegg.com/blog/website-navigation, accessed 25 Oct. 2019.