Soft Comput (2010) 14:887–897
DOI 10.1007/s00500-009-0472-7

FOCUS

A quality evaluation methodology for health-related websites based on a 2-tuple fuzzy linguistic approach

J. M. Moreno · J. M. Morales del Castillo · C. Porcel · E. Herrera-Viedma

Published online: 29 July 2009

Abstract Nowadays, the patients and physicians use the health-related websites as an important information source and, therefore, it is critical the quality evaluation of health-related websites. The quality assessment of health-related websites becomes especially relevant because their use imply the existence of a wide range of threats which can affect people's health. Additionally, website quality evaluation can also contribute to maximize the exploitation of invested resources by organizations in the development of user-perceived quality websites. But there is not yet a clear and unambiguous definition of the concept of website quality and the debate about quality evaluation on the Web remains open. In this paper, we present a qualitative and user-oriented methodology for assessing quality of health-related websites based on a 2-tuple fuzzy linguistic approach. To identify the quality criteria set, a qualitative research has been carried out using the focus groups technique. The measurement method generates linguistic quality assessments considering the visitors' judgements with respect to those quality criteria. The combination of the linguistic judgements is implemented without a loss of information by applying a 2-tuple linguistic weighted average operator. This methodology means an improvement on quality evaluation of health websites through the commitment to put users first.

Keywords Quality evaluation · Health websites · Computing with words · Linguistic modelling

1 Introduction

In recent years, we have witnessed an unprecedented growth and popularity of websites that have taken a central role in diverse fields such as in finance, education, medicine, industry and business. In the globalized world, citizens have greatly increased their demand for information, and websites are being used as an important information platform. Some typical characteristics of the World Wide Web, as for example its fast and uncontrollable growth, its heterogeneity, lack of publishing control, or freshness requirements has arisen the preoccupation on the quality of the websites. To cope with this situation, some initiatives aimed at assessing quality on the Web have been developed. On the other hand, since organisations throughout the world invest time and money in order to develop and maintain user-perceived websites, evaluating their quality is necessary to understand whether websites comply with user needs and expectations (Grigoroudis et al. 2008).

However, there is neither an agreed theoretical framework to be taken as reference nor a methodology for the evaluation of quality of the websites, at least not one which had obtained a wide consensus. The different approaches on the concept of quality show that it is a complex entity

J. M. Moreno
Department of Information and Communication Engineering, University of Murcia, 30071 Murcia, Spain
e-mail: jmmoreno@um.es

J. M. Morales del Castillo · E. Herrera-Viedma (✉)
Department of Computer Science and A.I, University of Granada, 18071 Granada, Spain
e-mail: viedma@decsai.ugr.es

J. M. Morales del Castillo
e-mail: jmmc@ugr.es

C. Porcel
Department of Computer Science, University of Jaén, Jaén, Spain
e-mail: cporcel@ujaen.es

and non-easily identifiable, as well as the multidimensional nature in what concerns measure and evaluation (Aladwani and Palvia 2002; Provost et al. 2006; Yang et al. 2005). There exist many different definitions of the term quality, and each one with a different translation in practical terms. It is furthermore a concept in ongoing evaluation which has been shaping and enriching its form along time. The debate about how to evaluate quality on the websites remains open and we can find different approaches and models to evaluate it (Aladwani and Palvia 2002; Olsina and Rossi 2002; Olsina et al. 2006; Kirakowski and Cierlik 1998; Negash et al. 2003; Powell et al. 1998; Sellito and Burgess 2005; Yang et al. 2005; Dhyani et al. 2002).

The majority of suggested Web evaluation models and methodologies tend to be more objective than subjective, quantitative rather than qualitative and do not take into account the user perception (Dhyani et al. 2002). We consider that any website quality evaluation methodology should include the direct participation of the users to evaluate the quality of websites, because the concept of quality is typically consumer-dependent, and the consumer must be the ultimate judge of their quality. To evaluate websites focusing on the user-perceived quality is a difficult task that has seldom been studied, and there does not exist a Web quality framework as a reference point (Aladwani and Palvia 2002; Rieh 2002; Yang et al. 2005). Furthermore, the aspects of quality to be considered and their relative importance vary according to the application domain and the goals to be achieved by the users. For example, criteria, as paying easiness or security, are not very important in a website of a government organisation that provides information to citizens while they are specially important in an online shop website. Therefore, we find final users who can use a given website with different purposes: communication, information, entertainment or commerce (Korgaonkar and Wolin 1999). In this sense, much attention must be paid when evaluating the quality of a website establishing a specific context of use and the final users' characteristics.

In the field of health, patients and their families demand to be involved in both medical decision making and actual care, so the Web has become an important tool for online consumers of information and services (Eysenbach et al. 2002). Health-related websites are listed among the most visited on the Internet and they are mainly aimed at providing information (Eysenbach et al. 1999). However, despite that people possess a new way to access information on health, it does not mean that they are covering their needs. The evaluation of the quality of health-related websites becomes especially relevant because, although it offers a potential to improve many health-related activities, it also implies the existence of a wide range of threats which can affect people's health (Risk and Dzenowagis 2001; Childs 2005; Sellito and Burgess 2005; Provost et al. 2006).

The main objective of this paper is to present a user-oriented quality evaluation methodology of health-related websites. This methodology generates linguistic assessments about the quality of websites that provide information or services on a given health-related matter or some specific clinical condition. It is qualitative and user-oriented for two main reasons:

1. We have carried out a qualitative research with real users to identify those criteria which characterize the quality of health-related websites from the users' point of view. These quality criteria are easily understandable by visitors and reflect their quality perspective.
2. To generate the linguistic quality assessments of health-related websites, our proposed method uses the linguistic judgements provided by the users after visiting the websites.

The methodology is made of two elements: a set of user-oriented quality criteria and dimensions that will be considered to assess the quality of the health websites; and a computational method for determining the global linguistic quality assessment of a website from the linguistic individual judgements made by users on the quality criteria applied to that website. To generate the global linguistic quality assessment, we consider different linguistic important degrees associated with the different visitor categories, quality criteria, and quality dimensions. Therefore, we need to carry out aggregations of weighted linguistic information. We use a 2-tuple fuzzy linguistic approach (Herrera and Martínez 2000) to represent and manage the linguistic information (importance degrees, judgements and quality assessments). This study was conducted with the participation of a representative sample of two specific classes of users that access regularly health-related websites: general visitors (chronic patients that do regular searches of websites) and expert visitors (doctors with a high level of knowledge in their medical fields). With this methodology, we increase the user participation in the quality evaluation process of websites with respect to other user-based approaches (Aladwani and Palvia 2002; Herrera-Viedma et al. 2006; Herrera-Viedma and Peis 2003; Rieh 2002; Yang et al. 2005), because it is based on the linguistic evaluation judgements provided by the users, but furthermore, the set of considered quality criteria is obtained from the real users' needs and expectations.

The paper is organized as follows. In the next section, we introduce the preliminaries. Section 3 defines the health-related website quality evaluation methodology. And finally, Sect. 4 shows the concluding remarks.

2 Preliminaries

In this section, we analyze the evaluation problem of health-related websites and the basis of the 2-tuple fuzzy linguistic approach.

2.1 Background on quality evaluation of health-related websites

Changes in society have created a growing, seemingly limitless demand for more access to medical information on which to base health decisions (Sieving 1999):

1. The increase in number and proportion of the older population, which is the heaviest user of the medical system.
2. The continuing increase in general educational levels attained and literacy rates. This creates an increase in the general ability of the population to read and act on information, an increased confidence in doing so, and increased sophistication in evaluating the information and using it to make decisions about medical care.
3. Increased comfort in dealing with new technologies, particularly those that are computer-based. Changes in technology enabling consumers of medical information to demand and use the information.

The Web, the largest repository of information, is more widely available at a time when greater initiative and participation are expected by consumers (Stavri et al. 2003). In this sense, the Web has become an important source for patients to gain access to health information and on-line services. A survey from the Pew Internet and American Life Project found that 86% of adult internet users had searched for health information (Fox and Rainie 2003), and between 36 and 55% of Internet users access online health information (Breckons et al. 2008).

But we can find very varied sources of information on the Web, from personal web pages to patients' association, private organizations, pharmaceutical industry, public institutions or scientific communities (Childs 2004). This situation has arisen a general concern on the quality of health information on the World Wide Web (Childs 2005; Risk and Dzenowagis 2001; Sellito and Burgess 2005; Provost et al. 2006). However, the scale of the problem and the epidemiology (distribution and determinants) of poor health information on the Web are still unclear, as is their impact on public health and the question of whether poor health information on the Web is a problem at all (Eysenbach et al. 2002; Coiera 1998). Some researches have been published describing, evaluating and analysing the quality on health-related websites, mainly focused on content quality. These studies tend to take as objective the evaluation or assessment of the quality of those websites which provide information about a given health-related matter or about some specific clinical condition, for instance, cancer (Eysenbach 2003; Meric et al. 2002; Berland et al. 2001), depressions (Lissman and Boehnlein 2001; Christensen and Griffiths 2000; Griffiths and Christensen 2000), or gynaecology (Diering and Palmer 2001; Galimberti and Jain 2000).

On the other hand, others researches published regarding health matters tend mainly to describe the situation of information quality on the Internet. For this purpose they make use of a wide variety of evaluation strategies. These works can be classified into two big different groups:

1. Studies which base website quality evaluation on a particular group of criteria (Eysenbach et al. 2002; Meric et al. 2002; Kim et al. 1999; Berland et al. 2001; Provost et al. 2006). This group includes the most part of the studies published so far. They are based on the selection of a group of quality criteria, mainly referred to the way the information is presented, that is, to the reliability of the websites as potential vehicles of health messages rather than to the content of the messages themselves. The majority of these criteria have been elaborated aimed at being applicable to the evaluation of any health-related website, regardless of the type of information contained. In general, research concludes that few websites fulfil the quality criteria, an also that the most popular sites offer extremely variable information (Eysenbach et al. 2002; Bedell et al. 2004).
2. Studies which deal with quality evaluation through a detailed analysis of the information content provided by websites (Cooke and Gray 2002; Li et al. 2001; Pandolfini et al. 2000; Impicciatore et al. 1997). This strategy to evaluate information quality is related to value the accuracy, exhaustivity and completion of the contents which provide the website using the best evidence available. These studies normally carry out the content evaluation by means of comparing medical protocols and based on experts' opinion. Impicciatore et al. (1997) were among the first to assess the reliability of Web page information by comparing it against a gold standard. Others have followed this approach, but in each case, they have been able to focus on specific pieces of information or advice that have an available gold standard. For example, Pandolfini et. al. (2000) compared information on the management of cough in children against a gold standard. These studies present certain limitations: time-cost, expert and source availability, and it is not a feasible option in websites that present information on a broader range of topics (Craigie et al. 2002).

Furthermore, most of them do not provide details about how or which sources have been used to contrast the contents (Eysenbach et al. 2002).

We should point out that in the increasingly competitive world of managed care, an effective Web presence is not a trivial matter (Sieving 1999). Different organizations are investing time and money in creating and maintaining sophisticated user perceived quality health websites to attract and retain consumers. These providers need to check the quality of what they are presenting, and users of their websites would give them the most important information. Quality assessment can significantly contribute to develop websites that serve users' need and meet the user expectations to the maximum possible extent (Grigoroudis et al. 2008). In order to achieve high-quality websites, designers have to first understand the different quality dimensions that affect users' expectations, and then relate these quality characteristics to specific design features (Zhang and von Dran 2001). For this reason, quality evaluation of health website has evolved as an important activity and several research efforts have been carried out.

In 1996, the earliest initiatives aimed at controlling the health-related information contained in the Internet were oriented to deal with those ethical aspects through behaviour codes. Nowadays, there is a wide variety of initiatives and tools that have tackled the issue of assessing and controlling the quality of health information on the Internet (Risk and Dzenowagis 2001; Kim et al. 1999), most of them designed through an ad hoc model approach and guided by quality criteria. These initiatives use different types of quality assessment schemes, and could be categorised as follows, though there are overlaps between these different categories (Childs 2005): Principles/codes of conduct (AMA Guidelines, HONCode; Checklist/ scoring instrument (DISCERN Netscoring); Quality label/ seal/ trustmark (HONCode TRUSTe, URAC); Third part accreditation (EU Quality Criteria, MedCERTAIN/ MedCIRCLE, TNO QMIC, TRUSTe); Metadata, electronically readable (MedCERTAIN/ MedCIRCLE); Gateway/ quality filter (Canadian Health Network, MedlinePlus, NHS Direct On-Line, OMNI, TNO QMIC).

Although some researches have proposed the development of automatic evaluation tools, so as to avoid the user's participation in the evaluation process (Ivory 2003), according to other authors (Childs 2004; Marshall and Willians 2006; Bernhardt and Felter 2004; Quintana et al. 2001), the users' participation is necessary as they are the end recipients of the websites' information and services. For this reason, one of the most important aspects to be considered by the quality evaluation tools is the users' point of view, both regarding the identification of the characteristics to be evaluated and also the freedom to let them judge directly the website quality. Considering that the user's experience summarises the global perception during a website visit, the development of website evaluation models based on perceived quality and user satisfaction is highly desirable.

2.2 The 2-tuple fuzzy linguistic approach

Many problems in the real world cannot be assessed precisely in a quantitative form, but it may be done in a qualitative one. In that case a better approach may be to use linguistic assessments instead of numerical values. The fuzzy linguistic approach is an approximate technique, which represents qualitative aspects as linguistic values by means of linguistic variables, that is, variables whose values are not numbers but words or sentences in a natural or artificial language (Zadeh 1975). This approach has been applied successfully to different areas as politics (Arfi 2005), decision-making (Herrera and Herrera-Viedma 2000), information retrieval (Herrera-Viedma 2001), information quality on the Web (Herrera-Viedma et al. 2006; Herrera-Viedma and Peis 2003), etc. The use of this fuzzy linguistic tool facilitates the user participation and it is a way to endow the evaluation process with flexibility and precision.

The 2-tuple fuzzy linguistic approach (Herrera and Martínez 2000) is a continuous model of representation of linguistic information that allows to reduce the loss of information typical of other fuzzy linguistic approaches (classical and ordinal (Herrera et al. 1996; Zadeh 1975)).

Let $S = \{s_0, \ldots, s_g\}$ be a linguistic term set with odd cardinality, where the mid term represents a indifference value and the rest of the terms are symmetrically related to it. We assume that the semantics of labels is given by means of triangular membership functions, represented by a 3-tuple (a, b, c), and consider all terms distributed on a scale on which a total order is defined, $s_i \leq s_j \Longleftrightarrow i \leq j$. An example of linguistic term set may be $\{P=(0.83,1,1), VH=(0.67,0.83,1), H=(0.5,0.67,0.83), M=(0.33,0.5,0.67), L=(0.17,0.33,0.5), VL=(0,0.17,0.33), N=(0,0,0.17)\}$, which is graphically shown in Fig. 1:

In this fuzzy linguistic context, if a symbolic method (Herrera et al. 1996; Herrera and Herrera-Viedma 1997) aggregating linguistic information obtains a value $\beta \in [0, g]$, and $\beta \notin \{0, 1, 2, \ldots, g\}$, then an approximation function is used to express the result in S, called Δ :

$$\Delta : [0, g] \longrightarrow S \times [-0.5, 0.5)$$

$$\Delta(\beta) = \begin{cases} s_i & i = \text{round}(\beta) \\ \alpha_i = \beta - i & \alpha_i \in [-0.5, 0.5) \end{cases}$$

In such a way, β is represented by means of a 2-tuple $(s_i, \alpha_i), s_i \in S$ and $\alpha_i \in [-0.5, 0.5)$, where s_i represents the

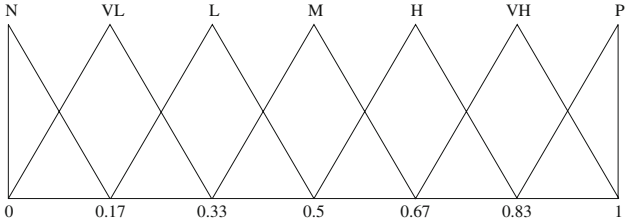

Fig. 1 A set of seven terms with its semantics

linguistic label of the information, and α_i is a numerical value expressing the value of the symbolic translation from the original result β to the closest index label, i, in the linguistic term set $(s_i \in S)$. On the other hand, as it is shown in Herrera and Martínez (2000), there is always a function Δ^{-1}, such that, from a linguistic 2-tuple (s_i, α_i) it returns its equivalent numerical value $\beta \in [0, g] \subset \mathcal{R}$, i.e., $\Delta^{-1}(s_i, \alpha) = i + \alpha_i$. It is obvious that the conversion of a linguistic term s_i into a linguistic 2-tuple consists in adding a value 0 as symbolic translation: $(s_i 0)$.

The 2-tuple linguistic computational model is defined by presenting a negation operator, comparison of 2-tuples and aggregation operators (Herrera and Martínez 2000):

1. A 2-tuple negation operator:

   $Neg((s_i, \alpha)) = \Delta(g - (\Delta^{-1}(s_i, \alpha)))$

2. A 2-tuple comparison operator
   Let (s_k, α_1) and (s_l, α_2) be two 2-tuples. Then
   - if $k < l$ then (s_k, α_1) is smaller than (s_l, α_2)
   - if $k = l$ then
     1. if $\alpha_1 = \alpha_2$ then $(s_k, \alpha_1), (s_l, \alpha_2)$ represents the same information
     2. if $\alpha_1 < \alpha_2$ then (s_k, α_1) is smaller than (s_l, α_2)
     3. if $\alpha_1 > \alpha_2$ then (s_k, α_1) is bigger than (s_l, α_2)

3. Aggregation of 2-tuples
   The aggregation of information consists of obtaining a value that summarizes a set of values; therefore, the result of the aggregation of a set of 2-tuples must be a 2-tuple. Using the functions Δ and Δ^{-1} that transform numerical values into 2-tuples and viceversa without loss of information, we can extend many of the classical aggregation operators defined in other contexts, as for example the following ones:

   - *Arithmetic mean*
     The arithmetic mean is a classical aggregation operator. Its equivalent operator, for linguistic 2-tuples, is defined as
     Definition 1 Let $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ be a set of 2-tuples, the 2-tuple arithmetic mean $\overline{x}^e$ is computed as,

$$\overline{x}^e = \Delta\left(\sum_{i=1}^{n} \frac{1}{n} \Delta^{-1}(r_i, \alpha_i)\right) = \Delta\left(\frac{1}{n} \sum_{i=1}^{n} \beta_i\right)$$

   - *Weighted average operator*
     The weighted average allows different values x_i have a different importance in the nature of the variable x. To do so, each value x_i has a weight associated, w_i, indicating its importance in the nature of the variable. The equivalent operator for linguistic 2-tuples is defined as
     Definition 2 Let $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ be a set of 2-tuples and $W = \{w_1, \ldots, w_n\}$ be their associated weights. The 2-tuple weighted average $\overline{x}^w$ is

$$\overline{x}^w = \Delta\left(\frac{\sum_{i=1}^{n} \Delta^{-1}(r_i, \alpha_i) \cdot w_i}{\sum_{i=1}^{n} w_i}\right) = \Delta\left(\sum_{i=1}^{n} \beta_i \cdot w_i \sum_{i=1}^{n} w_i\right)$$

   - *Linguistic weighted average operator*
     This operator is an extension of the $\overline{x}^w$ introduced in Definition 2; in this case the weights are expressed by means of linguistic values Herrera and Herrera-Viedma (1997)
     Definition 3 Let $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ be a set of 2-tuples and $W = \{(w_1, \alpha_1^w), \ldots, (w_n, \alpha_n^w)\}$ be their associated weights represented in a 2-tuple linguistic representation model. The 2-tuple linguistic weighted average $\overline{x}_l^w$ is

$$\overline{x}_l^w([(r_1, \alpha_1), (w_1, \alpha_1^w)] \ldots [(r_n, \alpha_n), (w_n, \alpha_n^w)]) = \Delta\left(\frac{\sum_{t=1}^{k} \beta_i \cdot \beta_{W_i}}{\sum_{i=1}^{n} \beta_{W_i}}\right),$$

with $\beta_i = \Delta^{-1}((r_i, \alpha_i))$ and $\beta_{W_i} = \Delta^{-1}(w_i, \alpha_i^w))$.

3 A quality evaluation methodology for health-related websites

In this section, we present a methodology to assess quality of websites that provide information or services about a given health-related matter or about some specific clinical condition. The goal of this methodology is to allow a relatively easy assessment of website quality from the users' perspective, that will be helpful to visitors in their search process and website owners in the development of their websites.

There are two different parts on this evaluation methodology:

- On the one hand, we establish the set of quality criteria and dimensions that will be considered from the users' perspective to assess health websites. The participation

of the users is necessary to identify those criteria which better characterize the website quality according to the users.
- On the other hand, we propose a computational method for determining the linguistic quality assessment of a website which is based on 2-tuple linguistic weighted aggregation operators.

3.1 Quality criteria and dimensions for evaluating health-related websites

Given the great number of different definitions and shades existing in what concerns the concept of quality we must establish a referential framework for its evaluation. The variables, factors, criteria or dimensions which are measurable to characterize the different levels of quality are a direct consequence of the specific operational conceptualization that we adopt.

Websites must fulfil the end-user needs and expectations and these needs and expectations may differ from the developers' and providers' perception. To assess the quality of websites under the end users' perspective, we must know the most important aspects of the websites assessed by them. In order to investigate on the most relevant quality aspects of the websites from the users' point of view, a qualitative research has been carried out using the focus groups technique (Morgan 1997; Stewart and Shamdasani 1990). As a result of this research, we have obtained a user-driven quality criteria framework composed of 30 criteria and grouped in five dimensions.

3.1.1 Qualitative research design through technique of focus groups

Focus groups is a qualitative research technique based on group discussions to obtain in-depth information on a specific topic of interest. Qualitative research through the focus groups technique will be helpful to find out which range of needs and expectations the users have. People involved in these groups are encouraged to talk to one another, exchanging anecdotes, experiences and points of view (Kitzinger 1994). The participants become an active part of the process of analysis and they may actually develop particular perspectives as a consequence of talking with other people who have similar experiences (Kitzinger 1995).

In order to carry out a focus group study, certain decisions must be made on certain questions like sampling and group composition, participant recruitment and data analysis (Krueger 1997; Morgan 1997; Stewart and Shamdasani 1990). We choose adequate participants for our research. We work with a representative sample of two specific classes of users that periodically access health-related websites: general visitors and expert visitors.

1. *General visitors* These are non-expert users who have any specific interest on a health-related issue. We will focus on the chronic patient visitor because this sort of general user shows the characteristics appropriate to our research objectives. Due to the chronic nature of the illness they suffer, they do regular searches about new treatments, nutritional advises and alternative therapies. They become, therefore, regular users of websites which provide information about their specific illness. As well, they accumulate enough experience about those websites.
2. *Expert visitors* Within the expert visitors we will focus on medical personnel as consumers of online health information. This group possesses a high level of information and uses health-related websites to widen its subject knowledge area.

We choose 23 participants to develop our experience in a hospital of Murcia country. They were 19 men and 4 women in of age 31–68 years. Among them we had 13 real patients and 10 doctors. We developed four sessions of focus groups with four groups of people, respectively. After development of meetings, a full transcription of the group discussions was written down verbatim and we analysed the data obtained in the sessions. The most complex level of analysis was the interpretation of the findings. It meant an explanation of the data and an analysis of the content of the opinions gathered to comprehend all those aspects that the participants stated.

Once the process of analysis was finished, we identified a list of 30 criteria grouped into five dimensions. These criteria and dimensions for assessing the quality of health-related websites are presented in the following subsection.

3.1.2 Quality criteria and dimensions

The exploration carried out through the focus groups technique allows us to choose a sample of those possible factors, particularly those most relevant for the population under study. We should point out that the results of this study indicate that there is a high degree of coincidence on those valuation elements expressed by both types of participants, doctors and chronic patients. In this sense, we could establish a unique group of dimensions and criteria for evaluating the quality of health-related websites by both types of users.

According to the qualitative research results we define the following five quality dimensions together with their quality criteria:

D_1 : *Credibility* This dimension is related with those aspects that offer reasonable grounds for being believed, allowing users to assess the credibility degree of a website. It is the dimension which contains more quality criteria that are associated with the website owners (the identification of the institution or organization), sponsorship (disclosure of sponsorship and the nature of the support provided), objectives (website objectives must be specified), and advertising (advertising and information contents must be differentiated), etc. The quality criteria of this dimension are the following:

C_1: It must be possible to identify the website owner.
C_2: Website origin place must be displayed.
C_3: Owner's contact address must be shown.
C_4: Website objective must be specified.
C_5: Specified aims must be satisfied.
C_6: The website audience must be specified.
C_7: Website sponsors and investors must be disclosed.
C_8: Interest conflict declaration must be shown.
C_9: Advertising and contents must be differentiated.
C_{10}: Advertising should not be contradictory with respect of the website contents.
C_{11}: Updating date must appear.
C_{12}: A declaration of personal data protection must be shown.

D_2: *Content* This dimension is composed of those criteria related to health information contents provided by websites. Accuracy of information is the most obvious criterion for quality of content, and users have the right to expect that websites will provide accurate information. We assess accuracy of website content by considering what visitors think about the information that the website provides. Other criteria belonging to this dimension are the following: authorship (requires to disclose the information authors), bibliography (literature used to gather information content), date (the last update of content posting), etc. Particularly, in this dimension, we have obtained the following quality criteria:

C_{13}: Personal evaluation of website content.
C_{14}: It should be possible to identify authors' names of the website documents.
C_{15}: Publication date must appear in the documents.
C_{16}: Bibliography must appear in the documents.
C_{17}: The language must be understandable.

D_3: *Usability* This dimension refers to the functionality for improving ease-of-use during the users' visit. The user analyzes this dimension by taking into account criteria mainly related to website design (logical organization of elements in such a way that visitors easily understand how to use the website) and website navigation mechanism (navigation tools provided by the website necessary to reach the specific information). We evaluate this dimension by means of the following quality criteria:

C_{18}: The design and organization of a website must be user-friendly.
C_{19}: Website should have options for moving freely along its structure.
C_{20}: Website should not have browsing problems.
C_{21}: The website surfing speed must be high.
C_{22}: Website should not include pop-up advertising.
C_{23}: Website should present a versatile and user-friendly search engine.
C_{24}: Downloading should be user-friendly.

D_4: *External links* This dimension is related to connections from a website to other external sites, forming a web-like structure of information between websites. This category should be assessed by whether the website provides information about the linked source and alerts when visitors move to an external website.

C_{25}: External links should show a full description of the linked website.
C_{26}: External links must be distinguished from those that are not

D_5: *Interactivity services* This dimension refers to the interactive services provided by the website, as for example, web forums or online question services. Then, we identify the following quality criteria in this dimension:

C_{27}: If website has a question service.
C_{28}: If the website has a forum it should be possible to identify the administrator.
C_{29}: If the website has a forum the access should be identified.
C_{30}: Availability of tools for self-diagnosis.

These quality dimensions presented do not play the same role to measure the quality of a website, i.e., some dimensions should be more important than others for assessing the global quality of a website. For example, user opinions on the credibility dimension would be more important than on the dimension of external links.

3.2 Computational method to obtain quality assessments

In this section, we present a computational method to generate global linguistic assessment on the quality of health websites from the linguistic individual judgements provided by different visitors $E = \{e_1, e_2, \ldots, e_m\}$ according to the quality criteria and dimensions showed

previously. It is based on the 2-tuple fuzzy linguistic approach (Herrera and Martínez 2000). We should point out that the use of the 2-tuple fuzzy linguistic approach in our evaluation methodology provides a well-founded mathematical framework to represent and deal directly with linguistic information. In such a way, we can generate linguistic quality assessments from linguistic judgements provided by visitors without loss of information. This is an important limitation in other Web quality evaluation methodologies (Aladwani and Palvia 2002; Huizingh 2000; Katerattanakul and Siau 1999) that also use in some cases labels to represent assessments, because they lack aggregation operators of linguistic information to generate the global quality values.

As aforementioned, we consider two categories of visitors of a health-related website: general visitors and expert visitors, and both with different importance degrees, which are assessed on a linguistic term set S, i.e., $I_1 \in S$ and $I_2 \in S$, respectively. When a visitor $e_i \in E$ finishes his/her visit to a website Web_t which provided him/her some information about a specific health-related topic, he/she is invited to complete a quality evaluation questionnaire as per the quality criteria, and therefore, this questionnaire contains 30 questions $\{q_1, q_2, \ldots, q_{30}\}$. The concept embedded in each question is also rated on a linguistic term set S, i.e., visitor provides linguistic judgements assessed on S.

On the other hand, we consider that the quality criteria do not play the same role and some criteria should be more influential than others to measure the quality dimension to which they belong. Similarly, as aforementioned, some quality dimensions are more important than others to measure the global quality of a website. Therefore, a set of relative linguistic importance degrees is assigned to the set of quality criteria and group of dimensions, i.e.,

$$\{I(C_1), \ldots, I(C_{30})\}, I(C_v) \in S,$$
$$\{I(D_1), \ldots, I(D_5)\}, I(D_v) \in S.$$

The particular linguistic importance degrees associated with each visitor category, quality criterion or quality dimension could be determined and established by the system administrator or an external expert. Another possible way of obtaining these linguistic importance degrees would be the design and application of a previous user questionnaire to assess this aspect.

The linguistic global quality assessments are obtained by aggregating the linguistic evaluation judgements by means of aggregation operators of weighted linguistic information, considering the different importance degrees with respect to visitors category, quality criteria and dimensions. In particular, the computational method is based on the application of the 2-tuple linguistic weighted average operator presented in Definition 3. The computation method generates the global linguistic quality assessment for a website Web_t in three steps:

Step 1: *Aggregation per quality criteria*. In this step we calculate the linguistic quality assessment for each criterion C_v by aggregating all the judgements provided by website visitors on that criterion. Let $\{q_1^{i,t}, \ldots, q_{30}^{i,t}\}$ be a set of linguistic evaluation judgements $(q_j^{i,t} \in S)$ provided the visitor $e_i \in E$ after to visit the website Web_t. Then, the linguistic quality assessment of the website Web_t according to the criterion C_v, called $Q_v^t \in S \times [-.5, .5)$, is obtained using the 2-tuple linguistic aggregation operator $\bar{x}_l^w$ applied on the set of 2-tuple linguistic evaluation judgements provided by all visitors $\{(q_v^{1,t}, 0), \ldots, (q_v^{m,t}, 0)\}$:

$$Q_v^t = \bar{x}_l^w([(q_v^{1,t}, 0), (I(e_1), 0)], \ldots, [(q_v^{m,t}, 0), (I(e_m), 0)]),$$

where $(I(e_i), 0) \in \{(I_1, 0), (I_2, 0)\}$ are the linguistic importance degrees associated with each visitor.

Step 2: *Aggregation per quality dimension*. In this step we calculate the linguistic quality assessment for each dimension D_h by aggregating the respective linguistic quality assessments obtained in step 1 for each one of its criteria. Once the linguistic quality assessments for the 30 quality criteria have been calculated $\{Q_1^t, \ldots, Q_{30}^t\}, Q_v^t \in S \times [-.5, .5)$ on website Web_t, we calculate the linguistic quality assessment of the website Web_t according to the dimension D_h, called $Q_h^t \in S \times [-.5, .5)$, by aggregating the linguistic quality assessments obtained for each one of its criteria (see Table 1) using again the operator $\bar{x}_l^w$:

$$Q_h^t = \bar{x}_l^w([Q_r^t, (I(C_r), 0)], [Q_{r+1}^t, (I(C_{r+1}), 0)], \ldots, [(Q_p^t, (I(C_p), 0)])$$

where $\{C_r, C_{r+1}, \ldots, C_p\}$ are the set of quality criteria grouped in dimension D_h and $\{(I(C_r), 0), (I(C_{r+1}), 0), \ldots, (I(C_p), 0)\}$ their linguistic importance degrees, respectively.

Step 3: *Aggregation per website*. In this step, we calculate the global linguistic quality assessment of a website by aggregating the linguistic quality assessments for each dimension obtained in the step 2. Let $\{Q_1^t, \ldots, Q_5^t\}, Q_h^t \in S \times [-.5, .5)$ the linguistic quality assessments obtained in step 2 on the website Web_t. Then, using the 2-tuple linguistic weighted average operator, $\bar{x}_l^w$, we obtain the global linguistic quality assessment $Q^t \in S \times [-.5, .5)$ as follows:

$$Q^t = \bar{x}_l^w([Q_1^t, (I(D_1), 0)], \ldots, [Q_5^t, (I(D_5), 0)]).$$

If we had a set of websites $\{Web_1, \ldots, Web_T\}$ which offer information with respect to an specific health-related topic, with our methodology we could establish a classification of these websites which could be very useful for users in their information search processes on the World Wide Web.

Table 1 Quality criteria grouped by dimensions

Quality dimensions	Quality criteria
Credibility	$C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}, C_{12}$
Content	$C_{13}, C_{14}, C_{15}, C_{16}, C_{17}$
Usability	$C_{18}, C_{19}, C_{20}, C_{21}, C_{22}, C_{23}, C_{24}$
External links	C_{25}, C_{26}
Interactive services	$C_{27}, C_{28}, C_{29}, C_{30}$

4 Concluding remarks

There exists a need for assessing the variable quality among websites in the health field. Furthermore, it is necessary to understand whether websites comply with users' needs and expectations and by this reason the user-perceived quality measurement of health-related websites becomes especially relevant given their potential implications on public health. Although many initiatives for evaluating the quality of health websites have been developed, there exist a few initiatives that deal appropriately with the user participation in the quality evaluation methodology.

In this paper, we have presented a user-oriented methodology for evaluating the quality of health-related websites that improves the user participation in the quality evaluation process mainly because the set of considered quality criteria is obtained directly from real users through of a qualitative research tool as the technique of focus groups (Kitzinger 1994, 1995). This is an important and differential aspect because in other user-based approaches (Aladwani and Palvia 2002; Herrera- Viedma et al. 2006; Herrera-Viedma and Peis 2003; Rieh 2002; Yang et al. 2005) we find that the considered quality criteria are established by the experts using their experience or user-oriented quality evaluation models defined for other frameworks.

We should point out that some limitations of our website quality evaluation methodology are precisely related to the user participation, i.e., it is a user-dependent methodology. Consequently:

1. As it happens in (Herrera-Viedma et al. 2006; Herrera-Viedma and Peis 2003) the quality of websites can only be evaluated if user perceptions can be gathered, which normally is not an easy task.
2. Furthermore, in our case, we need the user participation to identify the quality criteria to be considered in the quality evaluation process of health websites, and this presents to main problems: (a) Any user cannot participate in the sessions of focus groups; we require users familiarized with the use of new technologies and the Web; (b) The organization of the sessions of focus groups requires that users spend more time to interact with the chosen health websites and provide their preferences and opinions about the aspects of visited websites.

On the other hand, our website quality evaluation methodology shows some benefits:

1. It can be easily adapted to different health-related websites, and also, to different domains, as for example education.
2. Users that participate in our methodology see easily rewarded their effort with a better knowledge on the websites that provide information on their health. This increases the user motivation and allows to extend our methodology in many different health frameworks.
3. It uses fuzzy linguistic techniques to model user-evaluation judgements and the quality ratings and, in such a way, the subjectivity typical of user-system interactions can be managed efficiently and in a linguistic way which is closer to the humans. Additionally, the 2-tuple fuzzy linguistic approach provides us additional advantages, as for example, the combination of linguistic information without loss of information.
4. It allows to introduce in the computation process of linguistic quality assessments important differences among user judgements, quality criteria and quality dimensions.

Finally, we should point out that with this methodology, the evaluation possibilities on the Web are increased, and its application means an improvement in the quality assessment of health-related websites that could be helpful to patients, health care professionals, web developers and website owners. This methodology is proposed to generate linguistic recommendations on such websites that can help other users in their future search processes, discriminating among websites by means of their quality ratings. When a user requires health information on the Web, to know about the quality of the health websites could be of interest to the users to make their choice. On the other hand, the website owners could find the quality assessments very helpful to develop and enhance their websites, thus maximising the exploitation of invested resources.

Acknowledgments This work has been supported by the research project EMCA08-05 from "Programa EMCA", Murcia, Spain. It has also been developed with the financing of FEDER funds in FUZZYLING project (TIN2007-61079), PETRI project (PET2007-0460), and project of Ministry of Public Works (90/07).

References

Aladwani A, Palvia P (2002) Developing and validating an instrument for measuring user-perceived web quality. Inform Manag 39:467–476

Arfi B (2005) Fuzzy decision making in politics: A linguistic fuzzy-set approach (lfsa). Political Anal 13(1):23–56
Bedell S, Agrawal A, Petersen L (2004) A systematic critique of diabetes on the world wide web for patients and their physicians. Int J Med Inform 73:687–694
Berland G, Elliott M, Morales L, Algazy J (2001) Health information on the internet: Accessibility, quality, and readability in english and spanish. J Am Med Assoc 285(20):2612–2621
Bernhardt J, Felter E (2004) Online pediatric information seeking among mothers of young children: results from a qualitative study using focus groups. J Med Internet Res 6:e7
Breckons M, Jones R, Morris J, Richardson J (2008) What do evaluation instruments tell us about the quality of complementary medicine information on the internet? J Med Internet Res 10(1):e3. doi:10.2196/jmir.961
Childs S (2004) Developing health web site quality assessment guidelines for the voluntary sector: outcomes from the judge project. Health Inform Libr 21:14–26
Childs S (2005) Judging the quality of internet-based health information. Perform Meas Metr 6(2):80–96
Christensen H, Griffiths K (2000) Sites for depression on the web: a comparison of consumer, professional and commercial sites. Aust N Z J Public Health (24):396–400
Coiera E (1998) Information epidemics, economics, and immunity on the internet. Br Med J 317:1469–1470
Cooke A, Gray L (2002) Evaluating the quality of internet-based information about alternative therapies: development of the biome guidelines. J Public Health Med 24:261–267
Craigie M, Loader B, Burrows R, Muncer S (2002) Realibility of health information on the internet: an examination of expert ratings. J Med Internet Res 4(1):e2
Dhyani D, Keong Ng W, Bhowmick S (2002) A survey of web metrics. ACM Comput Surv 34(4):469–503
Diering C, Palmer M (2001) Professional information about urinary incontinence on the world wide web. J Wound Ostomy Cont Nurs 28:55–62
Eysenbach G (2003) The impact of the internet on cancer outcomes. CA Cancer J Clin (53):356–371
Eysenbach G, Ryoung E, Diepgen T (1999) Shopping around the internet today and tomorrow: towards the millennium of cybermedicine. Br Med J 319:1294
Eysenbach G, Powell J, Kuss O, Sa E (2002) Empirical studies assessing the quality of health information for consumers on the world wide web. A systematic review. J Am Med Assoc 287(20):2691–2700
Fox S, Rainie L (2003) The online health care revolution: how the web helps Americans take care of themselves. Pew Internet and American Life Project, Washington DC
Galimberti A, Jain S (2000) Gynaecology on the net: evaluation of the information on hysterectomy contained in health-related web sites. J Obstet Gynaecol 20:297–299
Griffiths K, Christensen H (2000) Quality of web based information on treatment of depression: cross sectional survey. Br Med J 321:1511–1515
Grigoroudis E, Litos C, Moustakis V, Politis Y, Tsironis L (2008) The assessment of user-perceived web quality: application of a satisfaction benchmarking approach. Eur J Oper Res187:1346–1357
Herrera F, Herrera-Viedma E (1997) Aggregation operators for linguistic weighted information. IEEE Trans Syst Man Cybern Part A Syst Hum 27:646–656
Herrera F, Herrera-Viedma E (2000) Linguistic decision analysis: steps for solving decisions problems under linguistic information. Fuzzy Sets Syst 115:67–82
Herrera F, Martínez L (2000) A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans Fuzzy Syst 8(6):746–752
Herrera F, Herrera-Viedma E, Verdegay J (1996) Direct approach processes in group decision making using linguistic owa operators. Fuzzy Sets Syst 79:175–190
Herrera-Viedma E (2001) An information retrieval system with ordinal linguistic weighted queries based on two weighting elements. Int J Uncertain Fuzziness Knowl Based Syst 9:77–88
Herrera-Viedma E, Peis E (2003) Evaluating the informative quality of documents in sgml-format using fuzzy linguistic techniques based on computing with words. Inform Process Manag 39(2):195–213
Herrera-Viedma E, Pasi G, Lopez-Herrera A, Porcel C (2006) Evaluating the information quality of web sites: a methodology based on fuzzy computing with words. J Am Soc Inform Sci Technol 57(4):538–549
Huizingh E (2000) The content and design of web sites: an empirical study. Inform Manag 37(3):123–134
Impicciatore P, Pandolfini C, Casella N, Bonati M (1997) Reliability of health information for the public on the world wide web. Br Med J 314:1875–1879
Ivory M (2003) Automated web site evaluation. Kluwer, Dordrecht
Katerattanakul P, Siau K (1999) Measuring information quality of web sites: development of an instrument. In: Proceedings of 20th international conference on information systems, pp 279–285
Kim P, Eng T, Deering M, Maxfield A (1999) Published criteria for evaluating health-related web sites: a review. Br Med J 318:647–649
Kirakowski J, Cierlik B (1998) Measuring the usability of web sites. In: Proceeding of human factors and ergonomics society 42nd annual meeting, Santa Monica, CA
Kitzinger J (1994) The methodology of focus groups: the importance of interactions between research participants. Sociol Health Illn 16:103–121
Kitzinger J (1995) Introducing focus groups. Br Med J 311:299–302
Korgaonkar P, Wolin L (1999) A multivariate analysis of web usage. J Advert Res 39:53–68
Krueger R (1997) Developing questions for focus groups (focus group kit). Sage, Thousand Oaks
Li L, Irvin E, Guzman J, Bombardier C (2001) Surfing for back pain patients: the nature and quality of back pain information on the internet. Spine 26:545–557
Lissman T, Boehnlein JK (2001) A critical review of internet information about depression. Psychiatr Serv 52(8):1046–1050
Marshall L, Willians D (2006) Health information: does quality count for the consumer? how consumers evaluate the quality of health information materials across a variety of media. J Librariansh Inform Sci 38:141–156
Meric F, Bernstam E, Mirza N, Hunt K, Ames F, Ross M, Kuerer H, Pollock R, Musen M, Singletary S (2002) Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites. Br Med J 324:577–581
Morgan D (1997) The focus groups guidebook (focus group kit). Sage, Thousand Oaks
Negash S, Ryan T, Igbaria M (2003) Quality and effectiveness in web-based customer support systems. Inform Manag 40:757–768
Olsina L, Rossi G (2002) Measuring web application quality with webqem. IEEE Multimed October–December:20–29
Olsina L, Covella G, Rossi G (2006) Web engineering. In: Mendes E, Mosley N (eds) Web quality. Springer, Berlin
Pandolfini C, Impicciatore P, Bonati M (2000) Parents on the web: risks for quality management of cough in children. Pediatrics 105:e1

Powell T, Jones D, Cutts D (1998) Web site engineering: beyond web page design. Prentice Hall, Englewood Cliffs

Provost M, Koompalum D, Dong D, Martin B (2006) The initial development of the webmedqual scale: domain assessment of the construct of quality of health web sites. Int J Med Inform 75:42–57

Quintana Y, Feightner J, Wathen C, Sangster L, Marshall J (2001) Preventive health information on the internet—qualitative study of consumers' perspectives. Can Fam Phys 47:1759–1765

Rieh S (2002) Judgment of information quality and cognitive authority in the web. J Am Soc Inform Sci Technol 53(2):145–161

Risk A, Dzenowagis J (2001) Review of internet health information quality initiatives. J Med Internet Res 3(4):e28

Sellito C, Burgess S (2005) Towards a weighted average framework for evaluating the quality of web-located health information. J Inform Sci 31(4):260–272

Sieving P (1999) Factors driving the increase in medical information on the web—one american perspective. J Med Internet Res 1(1):e3

Stavri P, Freeman D, Burroughs C (2003) Perception of quality and trustworthiness of internet resources by personal health information seekers. In: AMIA symposium proceedings, pp 629–633

Stewart D, Shamdasani P (1990) Focus groups: theory and practice (applied social research methods). Sage, Newbury Park

Yang Z, Cai S, Zhou Z, Zhou N (2005) Development and validation of an instrument to measure user perceived service quality of information presenting web portals. Inform Manag 42:575–589

Zadeh L (1975) The concept of a linguistic variable and its applications to approximate reasoning. Inform Sci Part I, II, III 8, 8, 9:199–249, 301–357, 43–80

Zhang P, von Dran G (2001) User expectations and rankings of quality factors in different website domains. Int J Electron Commer 6(2):9–33