Expert Systems with Applications 41 (2014) 3306–3312

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Formalizing and validating the web quality model for web source quality evaluation

Bo Zhao, Yan Zhu *

School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China

ARTICLE INFO

Keywords:
Formalization of WebQM with Z
Structural Equation Model
Fitness analysis of model
Goodness-of-fit criteria
Web source quality evaluation process

ABSTRACT

The proposed web quality model (WebQM) is formalized with ISO/IEC Z language and empirically studied based on the Structural Equation Modeling (SEM) approach. By building the sample data set and constructing the structure equation model, the goodness-of-fit of WebQM is analyzed based on generalized least square method. A web source quality evaluation process based on validated WebQM is implemented and verified as more objective and credible, because the weights of quality criteria are automatically produced in the validation procedure, which avoids the subjective weight assignment in some classic assessment approaches. The model validation and implemented evaluation show that WebQM fits the real web source quality data and is feasible, reliable, and effective for web source quality evaluation.

1. Introduction

The web source quality is the most critical factor for the performance of E-business and E-government and can contribute to the success of a variety of web-based applications. Quality oriented web resource evaluation and selection is vital for efficient use of web information, knowledge discovery, information analysis and decision making.

Web source quality evaluation and selection should be guided by model and criteria. However, a number of information evaluation approaches used in the past have not integrated some critical features of web resource qualities yet, such as autonomy, dynamics, openness, and heterogeneous data structure. It is necessary to build a scientific and reasonable web source quality evaluation model, which fulfills the quality requirements of web source, web content and web application context. Additionally, the model should be feasible, effective and fit to the actual web quality status. A web source quality evaluation model (WebQM, Web quality model) had been proposed in our previous work (Zhu, 2004). This model covers 3 quality dimensions, such as autonomy and dynamics of web source, openness and heterogeneity of web content and extensive context of various applications. 13 evaluation criterias are used to define quality dimensions. In this paper we discuss the formal specification of WebQM and the validation of feasibility and effectiveness of the model.

Based on literature review, Aladwaniam and Palvia (2002) assessed the web quality based on a 4 dimensions and 25 criteria model: Technical adequacy, specific content, content quality and web appearance. However, some of the criteria contribute less to the web quality, e.g. finding people without delay and finding site maintainer. A few of them conflict with each other, e.g. uniqueness and extensiveness. The web quality model of Lowry, Vance, Moody, Beckman, and Read (2008) has 6 dimensions: Responsiveness, competence, quality of information, empathy, web assistance and callback systems. The Structural Equation Modeling analysis was adopted to establish the factorial validity of their model. However, some of their 6 dimensions cannot be implemented easily for assessing web quality, for example, competence or empathy. Sun and Lin (2009) proposed 3 quality dimensions and 12 criteria to evaluate the competitive advantages of on-line shopping sites. Then, they employed the fuzzy TOPSIS method to determine the weights of different criteria for the online shopping websites, therefore the results were significantly influenced by the experts who evaluate the websites. To make the web quality assessment more objective, we use questionnaires rather than expert evaluation to the web source quality assessment score. Yoo and Donthu developed SITEQUAL (Yoo & Donthu, 2001) to evaluate on-line shopping websites based on 9 dimensions. They gathered fifty-four unique items and then reduce the items to nine for measuring four factors (ease to use, design, processing speed, and security) by using exploratory factor analysis. But SITEQUAL's original item set was too narrowly based, and most of its final factors were measured by only two items. In another model WEBQUAL (Loiacono, Watson, & Goodhue,

* Corresponding author. Tel.: +86 13880957071.
E-mail addresses: zhaobo1987@gmail.com (B. Zhao), yzhu@swjtu.edu.cn (Y. Zhu).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.11.027

2007), the Theory of Reasoned Action and the Technology Acceptance Model were employed, which has 12 dimensions. To check reliability and discriminant validity, a confirmatory factor analysis was conducted by using LISREL. Cao, Zhang, and Seydel (2005) used five dimensions to evaluate e-commerce website and the exploratory factor analysis to check for factorial structures. In Orehovacki, Granic, and Kermek (2013), the authors employed logging actual use method to measure the estimated website quality. They combined retrospective thinking aloud method with an online questionnaire to assess the perceived quality. However, the model orients just a few kinds of websites, such as websites providing services for mind mapping and diagraming. It is hard to generalize their model to all types of web applications, because each kind of web 2.0 applications has its specific features. Baloglu and Pekcan (2006) applied content analysis to study the websites of hotels in terms of site design and marketing characteristics using a binary variables of yes-or-no (one-or-zero). The shortcoming of this approach is that they cannot express the quality performance on each criterion.

Generally speaking, there are similarities between their work and ours, especially in the quality evaluation model construction. However, the above studies mainly focused on the criteria design and concepts, few of them formalized the model. Besides, some of the model criteria cannot be quantized, e.g. past experience and proficiency. The most outstanding feature of our work is that the WebQM takes the differential weighting into consideration and each criterias weight can be automatically generated. Additionally, the web quality models described above mostly were developed on the perspectives of web developers and designers, but not on that of web users. It is undeniable that the users play an increasingly important role in web quality. Our model defines web quality from user's perspective. The goodness-of-fit of the model is analyzed by Structure Equation Model approach based on the questionnaire made by 213 users.

As to the model formalization, there are few related literature. Authors of paper (Cheng & Wang, 2008) formally modeled the combination pattern of the semantic web services based on Colored Petri Net, and proposed an algorithm to validate the syntax of the combined model. Different from their work, we discuss the formal specification of WebQM and validate the feasibility and effectiveness of the model.

The major contributions of this paper are as follows:

1. Specify the WebQM using formal Z language to avoid ambiguity, which has precise semantics and is standardized by ISO/IEC.
2. Verify the feasibility and effectiveness of WebQM using the Structural Equation Modeling (SEM) approach.
3. Generate the quality criteria weights of WebQM in goodness-of-fit analysis, which improve the objectiveness of the quality assessment based on WebQM.
4. Validate the reliability, objectiveness and feasibility of WebQM by implementing a web source quality evaluation process.

2. Formal specification of WebQM

Z is a formal specification language of ISO/IEC JTCI/SC22. It is based on the first-order logic and Zermelo–Fraenkel set theory. It usually contains 3 parts:

1. Grammar, which specifies the representation method of the language.
2. Semantics, which defines the full domain of objects in the description system.
3. A set of relationships, which defines rules, with which the specific objects should be consistent.

In order to precisely define and explicitly represent WebQM (as shown in Fig. 1), we use Z language to formally define the structure, data type and evaluation pattern of the WebQM.

Definition 1. The tree structure of WebQM is expressed as Z child function. The child function represents the partial function (1:N relationship) between the node sets, i.e. child: NODE $\mapsto$ NODE.

Given WebQM∈dom (child), WebSQ∈dom (child), WebIQ∈dom (child), WebAQ∈dom (child), then dom {WebQM $\mapsto$ WebSQ} = {WebSQ}, and so on.

The general formal definition of WebQM is: child = {WebQM $\mapsto$ WebSQ, WebQM $\mapsto$ WebIQ, WebQM $\mapsto$ WebAQ, WebSQ $\mapsto$ Availability, WebSQ $\mapsto$ Accessibility, WebSQ $\mapsto$ Durability, WebSQ $\mapsto$ Timeliness of Information, WebIQ $\mapsto$ Reliability, WebIQ $\mapsto$ Correctness, WebIQ $\mapsto$ Completeness, WebIQ $\mapsto$ Objectivity, WebIQ $\mapsto$ Understandability, WebIQ $\mapsto$ Validity of Information, WebAQ $\mapsto$ Relevance, WebAQ $\mapsto$ Presentation, WebAQ $\mapsto$ Information Acquisition}

$$
\begin{array}{l}
\textit{QualityEvaluation} \\
\hline
quality : \mathbb{P}\ Dimensions \\
WebSQ, WebIQ, WebAQ : \mathbb{P}\ Criteria \\
Availability, Accessibility, \cdots, InformationAcquisition : Z \\
Weight : Z \\
correlation\ \ quality \nrightarrow Dimensions \\
correlation\ \ Dimensions \nrightarrow Criteria \\
correlation\ \ Criteria \rightarrowtail\!\!\!\!\!\!\nrightarrow weights \\
\hline
\mathrm{dom}\ WebSQ \subset quality\ \wedge \\
\mathrm{dom}\ WebIQ \subset quality\ \wedge \\
\mathrm{dom}\ WebAQ \subset quality\ \wedge \\
\mathrm{dom}\ Criteria \subseteq Dimensions\ \wedge \\
weight > 0
\end{array}
$$