

Evaluating Websites of Specialized Cultural Content Using Fuzzy Multi-Criteria Decision Making Theories

Katerina Kabassi, Athanasios Botonis and Christos Karydis

Department of Environment, Ionian University, Minotou Giannopoulou 26, 29100 Zakynthos, Greece

E-mail: kkabassi@ionio.gr, nasbotonis@gmail.com and c.karydis@ionio.gr

Keywords: website evaluation, cultural informatics, multi-criteria decision making

Received: February 19, 2019

The museums' conservation labs and the treatments on the artifacts many times are overlooked and are not obvious for the public. Nevertheless, their content, which is more specialized than the content of the main museum, may be of interest to students, researchers, archaeologists, tourists, artists for further education and preservation guidelines purposes. In this paper, we evaluate the electronic presence of museums' conservation labs using both empirical and inspection methods of evaluation. For this purpose, a combination of Analytic Hierarchy Process (AHP) and Fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is used to implement an evaluation experiment that combines inspection and empirical methods of evaluation. The proposed scheme of evaluation that implements a combination of methods and decision making theories for the evaluation of websites with specialized cultural content has been used for evaluating the 29 websites of museum's conservation labs and ranks them taking into account their content, usability, and functionality.

Povzetek: Prispevek z metodami umetne intelligence ocenjuje spletne strain muzejev.

1 Introduction

Museums' main role is connected with the exhibition of their artifacts. For this reason, museum websites are mainly concerned with this museum function. Another major work that is being done in a museum environment and is often overlooked by the public is the work carried out within a museum's conservation lab. Consequently, the electronic presence of museum conservation labs is also neglected as a result of ignoring or diminishing the public's consciousness of the important work of preserving collections. Despite this fact, there are museums that have invested in the electronic presentation of their conservation labs. The website of a museum's conservation lab differs from the main website of the museum as it contains more specialized information about the artifacts, the equipment used and the research conducted in the labs.

The existence of a website does not guarantee success. Sometimes the websites are poorly developed. As a result, the interaction is made difficult and the museums may lose attention instead of gaining. Indeed, Dyson and Moran (2000) discussed the importance of creating accessible and usable information resources for online museum projects. Therefore, many researchers have highlighted the need for evaluating websites with cultural content (Cunliffe et al., 2001, van Welie & Klaasse 2004). As a result, most of the evaluations of websites of cultural content are about e-museum websites' evaluations.

There is a plethora of methods and theories that could be used in order to evaluate a museum website (Kabassi 2017), however, not many solutions have been proposed for evaluating websites of specialized cultural

content. A rather common categorization of the proposed evaluation methods is made taking into account the participants of the experiment (Kabassi 2017). Indeed, Lewis & Rieman (1994) as well as Davoli et al. (2005), distinguish methods to empirical methods and inspection methods. Inspection methods are used in experiments that the participants are experts. Empirical methods, on the other hand, are implemented with the participation of different categories of potential users of a museum's website (Kabassi 2017). Each method has different advantages and disadvantages. For example, expert-based evaluations are easier and cheaper compared to empirical ones (Reeves 1993; Karoulis et al. 2006). Empirical methods, on the other hand, may be more successful in capturing end user's perceptions as real users participate in the experiment (Kabassi 2017). However, in this case, the experiment needs a large group of evaluators. This is more complicated and expensive compared to inspection methods but their results are undeniable.

In view of these advantages and disadvantages, some evaluation experiments use both users and experts (Garzotto et al. 1998, Harms & Schweibenz 2001, Vavoula et al. 2009, Sylaiou et al. 2014). In this paper, we have used a combination of inspection and empirical method to evaluate the websites of specialized cultural content, such as the websites of the museums' conservation labs. More specifically, we have used experts to evaluate the importance of the criteria used in the evaluation experiment and estimate their weights and real users for evaluating the different alternative websites. The inspection and the empirical methods are

combined with multi-criteria decision making theories for processing the input and making the essential estimation. The Multi-Criteria Decision Making (MCDM) theories that are used are AHP (Analytic Hierarchy Process) (Saaty 1980) and Fuzzy TOPSIS (Fuzzy Technique for Order of Preference by Similarity to Ideal Solution) (Chen 2000).

AHP aims to analyze a qualitative problem through a quantitative method (Saaty 1980). TOPSIS, on the other hand, aims at ordering evaluation items, which in our case are museum websites, through detecting distance between evaluated objects and optimal solutions (Hwang & Yoon 1981). In the particular evaluation experiment, Fuzzy TOPSIS is used instead of TOPSIS because the theory is used in combination with an empirical method, where real users, and not just experts, participated in the experiment. The empirical method involved users answering a questionnaire with linguistic terms, which is easier for users to comprehend and use. Therefore, Fuzzy TOPSIS (Chen 2000) was used to convert linguistic terms to fuzzy numbers, process the data, making estimations and rank the alternatives.

Taking into account the above, AHP is used to implement the inspection method and fuzzy TOPSIS is used for the implementation of the empirical method. These two theories have different reasoning but seem rather complementary. This combination is mainly reported in the evaluation of websites in e-commerce and, more specifically, in the evaluation of websites of travel agencies (Soleymaninejad et al. 2016) or group-buying (Zhang 2015). Furthermore, Fuzzy AHP has been combined with Fuzzy TOPSIS for evaluating university websites (Nagpal et al. 2015) and e-government sites (Büyüközkan & Ruan 2007). This combination has never been used before in the cultural domain.

2 Research aim

Taking into consideration the advantages of the different evaluation methods we have implemented a framework describing an experiment for the evaluation of websites of specialized cultural content combining inspection and empirical methods. For the implementation of the different evaluation methods different multi-criteria decision-making theories have been used. More specifically, we use a combination of different MCDM theories to implement an evaluation experiment that combines inspection and empirical methods of evaluation in order to check the electronic presence of museums' conservation labs.

AHP is combined with an inspection method and Fuzzy TOPSIS with an empirical method of evaluation. This combination is proposed due to the advantages that each method provides. AHP provides the tools to analyse a qualitative problem. The method's ability in making decisions by making pairwise comparison of uncertain, qualitative and quantitative factors and also its ability to model expert opinion (Mulubrhan et al. 2014) are the main reasons for its combination with an inspection method of evaluation. In the particular evaluation

experiment, AHP is used for forming the set of criteria for the evaluation as well as their weights of importance.

Fuzzy TOPSIS, on the other hand, provides adequate tools to analyze the linguistic responses of users in a questionnaire to order the evaluated objects. Indeed, the empirical method that is combined with Fuzzy TOPSIS, involved users answering a questionnaire with linguistic terms, which is easier for users to comprehend and use. For this reason, Fuzzy TOPSIS was considered very suitable for converting linguistic terms to fuzzy numbers, processing the data, making estimations and ranking the alternatives. According to the theory, if the evaluated object is near the optimal solution and far away from the poor solution, it is the best.

Most evaluation experiments of websites in the cultural domain refer to the evaluation of museums' websites not websites of specialized cultural content. The proposed framework that is described in detail in this paper, could be easily applied for the evaluation of other websites of specialized cultural content.

3 Multi-criteria decision making methods

MCDM has evolved rapidly over the last decades (Zopounidis 2009). MCDM theories are devoted to the development and implementation of decision support tools and methodologies to confront complex decision problems involving multiple criteria, goals or objectives of conflicting nature (Zopounidis 2000). Various MCDA methods are available, such as AHP, Fuzzy AHP, TOPSIS, Fuzzy TOPSIS, Data Envelopment Analysis (DEA), Multi-attribute utility theory and many more. All these decision methodology approaches differentiate in the way the objectives and alternative weights are determined (Mohamadali & Garibaldi 2011).

Analytic Hierarchy Process (Saaty 1980) is one of the most popular MCDM theories. The choice of AHP amongst other MCDM theories is because it presents a formal way of quantifying the qualitative criteria of the alternatives and in this way removing the subjectivity of the result (Tiwari 2006). Furthermore, the method's ability in making decisions by making a pairwise comparison of uncertain, qualitative and quantitative factors and also its ability to model expert opinion (Mulubrhan et al. 2014) is another important reason for its selection against other alternatives. This method uses the nine-point scale developed by Saaty for evaluation of the goal with the criterion as well as the criterion with the alternative (Mulubrhan et al. 2014).

AHP can be used to implement all the stages of a decision-making process until having the alternatives shorted. However, the main problem of AHP is that the complexity rises with the increase of alternatives; therefore, it is better used when the number of alternatives is limited. A method to resolve this problem is by combining AHP with another theory that manages to process and sort several alternatives without increasing the complexity disproportionately, such as is TOPSIS. This theory calculates the relative Euclidean distance of the alternative from a fictitious ideal alternative. The

alternative closest to that ideal alternative and furthest from the negative-ideal alternative is chosen as the best.

However, the main problem with the use of TOPSIS is that since the evaluation of the alternatives was part of an empirical method, where real users, and not just experts, participated in the experiment, it is difficult for them to evaluate the websites using numbers. Indeed, in many cases, crisp data are inadequate to model real-life situations. The evaluation experiment is using a questionnaire with the linguistic term then Fuzzy TOPSIS (Chen 2000) should be used to process the data, making estimations and rank the alternatives. In this case, fuzzy numbers are used to access the ratings of each alternative with respect to each criterion and Fuzzy TOPSIS is implemented.

4 Inspection method for the implementation of AHP

In the first part of the evaluation experiment, an inspection method is implemented using AHP. The steps of the implementation of AHP in an inspection evaluation are the following:

1. Developing goal hierarchy

- a. **Forming the overall goal:** The overall goal is to evaluate the museum's conservation labs websites
- b. **Forming the set of criteria:** The criteria for evaluating the websites of the museum's conservation labs have been selected after a review on inspection evaluation experiments of museum websites proposed by Kabassi (2017) and selecting those that seem more appropriate for the particular evaluation.

i. **Category 1: Content.** *In this category, all criteria are related to the content of a website.*

1. **c11: Currency/Clarity/Text comprehension.** This criterion checks the currency and the clarity of the text. Currency refers to how successful is the system in providing up-to-date information, and how successfully it can reflect the current state of the world that it represents. Clarity refers to how comprehensible the texts provided to the users are. For this purpose, the quality and the style are checked as well as the way the content is organized and designed in order to make the website credible and trustworthy.
2. **c12: Completeness/Richness.** This criterion checks whether a website has adequate information on the subject.
3. **c13: Quality Content.** This criterion involves the accuracy and understandability of content.
4. **c14: Support of Research.** Checks whether the website provides information for the support of research.

ii. **Category 2: Usability.** *All the criteria that are related to Usability.*

1. **c21: Consistency.** Consistency means that similar pieces of information are dealt with in similar fashions (Di Blas et al. 2002).
2. **c22: Accessibility.** Accessibility measures how easily and intuitively accessible is the website's information for any user.
3. **c23: Structure/Navigation.** The structure of the information provided plays an important role in the success of a website. Therefore, the organization of the content pieces should be in such a way that the navigation of the user to the content of the website is easy.
4. **c24: Easy to use/simplicity.** The user interface should be simple and easy to use.
5. **c25: User interface-Overall presentation-Design.** This criterion checks whether the overall presentation is attractive and engaging.
6. **c26: Efficiency.** This criterion shows whether actions within the website can be performed successfully and quickly (Di Blas et al. 2002).

iii. **Category 3: Functionality.** *Criteria that are related to the functionality of the website.*

1. **c31: Multilingualism.** the information should be given in more than one language (Di Blas et al. 2002)
2. **c32: Multimedia.** Different media should be used to convey the information (Di Blas et al. 2002)
3. **c33: Interactivity.** This criterion checks whether the content of the website is comprehensive and useful, nicely presented, easy to explore and use.
4. **c34: Adaptivity.** Adaptivity is the ability of the system to adapt to users' characteristics such as needs and interests while adaptability refers to the ability of users to adapt the user interface to their own preferences.

c. **Finding the websites to be evaluated:** In this step, the websites of the museums' conservation labs that are going to be evaluated are 29 and are presented in table 1.

d. **Forming the hierarchical structure:** In this step, the hierarchical structure is formed so that criteria could be combined in pairs.

2. **Form the set of evaluators:** As an inspection method is used the set of evaluators composites of human experts. Indeed, the correct choice of the expert would give reliable and valid results. Therefore, a double expert (software engineers and domain experts) system is proposed may increase the reliability of the results. As a result the group of evaluators contained 4 professional conservators and 4 software engineers, 3 of which had experience in a University Department of Conservation of Antiquities & Works of Art.

1	Archaeological Museum of Thessaloniki
2	Australian Museum
3	Barberini – Corsini Gallery – Roma
4	Benaki Museum
5	Boston Museum of Fine Arts
6	British Museum
7	Brooklyn museum
8	Byzantine & Christian Museum in Athens
9	De Young museum of Fine Arts
10	Galleria Nazionale d'Arte Moderna
11	Getty Institution
12	Guggenheim Museum
13	Hermitage Museum
14	Metropolitan Museum
15	MoMa
16	Museo Del Prado
17	Museum of Byzantine Culture in Thessaloniki
18	Museum of Islamic Art - Doha
19	National Gallery of Greece
20	National Museum New Delhi
21	NTNU University museum
22	Oriental Institute Museum
23	Rijksmuseum
24	Smithsonian museum
25	Tate Modern
26	Tokyo National Museum
27	University of Michigan Museum of Art
28	Vatican Museum
29	Victoria & Albert Museum

Table 1: The websites of museums' conservation labs that are evaluated.

- Setting up a pairwise comparison matrix of criteria:** In this step, a comparison matrix is formed so that the criteria of the same level are pair-wise compared. More specifically, three matrices are formed. The first compares content, usability and functionality, which are dimensions at the same level and then another one is formed for the sub-criteria of each one of the three dimensions. For example, the matrix of combining the three dimensions is presented in table 2. In the comparison process, a V from the scale that is presented in Table 2 is assigned to the comparison result of two criteria

	Content	Usability	Functionality
Content	1	V	X
Usability	1/V	1	Y
Functionality	1/X	1/Y	1

Table 2: Matrix for the pairwise comparison of the three dimensions.

	Content	Usability	Functionality
Content	1.00	0.46	1.99
Usability	2.16	1.00	2.59
Functionality	0.50	0.39	1.00

Table 3: Matrix for the pairwise comparison of the three criteria of the first level.

‘Content’ and ‘Usability’, then the value of comparison of ‘Usability’ and ‘Content’ is a reciprocal value of V, i.e. 1/V. The value of the comparison of ‘Content’ and ‘Content’ is 1.

Each professional expert combines all four (4) matrices and the final values of each matrix are calculated taking into account the geometric mean of the 8 corresponding values of each matrix’s cell. As a result, the final matrices are built. From the pairwise comparison matrix of the dimensions (Table 3) one can easily derive the fact that usability and content are considered more important than functionality. Tables 4, 5 and 6 present the pairwise comparison matrices of the sub-criteria of content, usability and functionality, respectively. The information collected for the creation of the pairwise comparison matrix of the sub-criteria of usability (Table 5) revealed that museum curators thought that the criteria ‘Content Quality’ and ‘Currency/Clarity/Text comprehension’ were very important whereas experts in usability thought that ‘Overall presentation/Design’ and ‘Structure/Navigation/Orientation’ were more crucial. Finally, in functionality, the opinions of the software engineer, the web designer and the museum curators were in agreement and the pairwise comparison matrix of the sub-criteria of functionality is presented in Table 6.

Calculating weights of criteria: After making pairwise comparisons, estimations are made that result in the final set of weights of the criteria. In this step, the principal eigenvalue and the corresponding normalized right eigenvector of the comparison matrix give the relative importance of the various criteria being compared. The elements of the normalized eigenvector are the weights of criteria or sub-criteria. There are several methods for calculating the eigenvector. Multiplying together the entries in each row of the matrix and then taking the nth root of that product approximates the correct answer. The nth roots are summed and that sum is used to normalize the eigenvector elements to add to 1.00. In terms of simplicity, we have used the 'Priority Estimation Tool' (PriEst) (Sirah et al. 2015), an open-source decision-making software that implements the Analytic Hierarchy

Process (AHP) method, for making the calculations of AHP (fig. 2). The weights of the criteria are:

$w_{c1} = 0.292$, $w_{c2} = 0.534$, $w_{c3} = 0.174$, $w_{c23} = 0.214$, $w_{c24} = 0.213$, $w_{c25} = 0.164$,
 $w_{c11} = 0.34$, $w_{c12} = 0.186$, $w_{c13} = 0.325$, $w_{c26} = 0.088$, $w_{c31} = 0.242$, $w_{c32} = 0.315$,
 $w_{c14} = 0.149$, $w_{c21} = 0.172$, $w_{c22} = 0.15$, $w_{c33} = 0.196$, $w_{c3} = 0.247$.

	c11: Currency/Clarity/Text comprehension	c12: Completeness/Richness	c13: Quality Content	c14: Support of Research
c11: Currency/Clarity/Text comprehension	1.00	2.24	0.89	2.20
c12: Completeness/Richness	0.45	1.00	0.60	1.47
c13: Quality Content	1.13	1.67	1.00	1.92
c14: Support of Research	0.46	0.68	0.52	1.00

Table 4: Matrix for the pairwise comparison of the sub criteria of Content.

	c21: Consistency	c22: Accessibility	c23: Structure/Navigation	c24: Easy to use/simplicity	c25: User interface-Overall presentation-Design	c26: Efficiency
c21: Consistency	1.00	1.10	0.76	0.83	0.93	2.37
c22: Accessibility	0.91	1.00	0.73	0.68	0.98	1.51
c23: Structure/Navigation	1.31	1.37	1.00	1.13	1.36	2.06
c24: Easy to use/simplicity	1.21	1.47	0.88	1.00	1.62	2.22
c25: User interface-Overall presentation-Design	1.08	1.02	0.74	0.62	1.00	2.26
c26: Efficiency	0.42	0.66	0.48	0.45	0.44	1.00

Table 5: Matrix for the pairwise comparison of the sub criteria of Usability.

	c31: Multilingualism	c32: Multimedia	c33: Interactivity	c34: Adaptivity
c31: Multilingualism	1.00	0.85	1.10	1.00
c32: Multimedia	1.18	1.00	1.70	1.33
c33: Interactivity	0.91	0.59	1.00	0.75
c34: Adaptivity	1.00	0.75	1.34	1.00

Table 6: Matrix for the pairwise comparison of the sub criteria of Functionality.

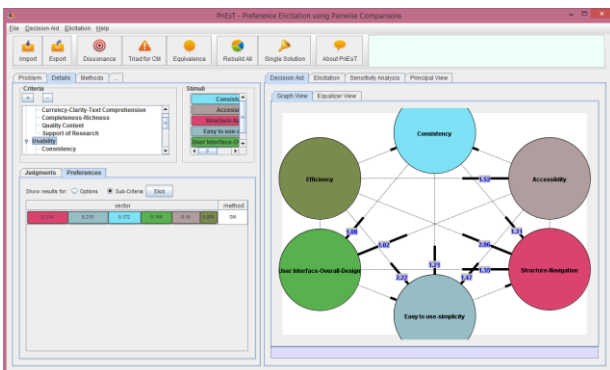


Figure 1: The interface of PriEst.

5 Empirical method with the implementation of fuzzy TOPSIS

In the second phase of the evaluation experiment, an empirical method is implemented. For this purpose, a new set of evaluators is formed to contain not only expert users but other categories of users, as well.

1. **Forming a new set of evaluators:** In this phase of the evaluation experiment, the set of evaluators was formed, following the taxonomy of types of users of cultural websites proposed by Sweetnam et al. (2012). More specifically, the final group of

evaluators, which involved professional researchers in conservation, students at advanced undergraduate and postgraduate level, informed users (researchers who are not professional academics but have knowledge of the subject) and the general public.

2. **Assigning values to the criteria:** In order to make this process easier for the user, especially for those that do not have experience in multi-criteria analysis, a questionnaire has been formed. The questionnaire involves a section of demographic questions and then another 29 section, one for each website that was evaluated. Each section contained 14 questions, one for each of the sub-criteria presented in the previous section. The questions provided only multiple-choice answers using the linguistic terms of table 7. The questionnaire was provided electronically using GoogleDocs (Figure 2).
3. **Linguistic terms are transformed to fuzzy numbers.** Each linguistic term is assigned to a fuzzy number, which is a vector like $\tilde{a} = (a_1, a_2, a_3)$. The matches are presented in table 7 (Chen 2000).

4. **Construction of the MCDM matrix.** A fuzzy multi-criteria group decision-making problem can be expressed in matrix format. Each element of the matrix is a fuzzy number. However, in order to aggregate all the values of the decision-makers in one single value the geometric mean is used. The geometric mean of two fuzzy numbers $\tilde{a} = (a_1, a_2, a_3)$ και $\tilde{b} = (b_1, b_2, b_3)$ is calculated as follows:

$$\tilde{c} = (\sqrt{a_1 b_1}, \sqrt{a_2 b_2}, \sqrt{a_3 b_3}).$$

$$\tilde{D} = \begin{matrix} & C_1 & C & C_n \\ A_1 & \begin{pmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & \tilde{x}_{2n} \\ \tilde{x}_{m1} & \tilde{x}_{m2} & \tilde{x}_{mn} \end{pmatrix} \\ A_2 & & & \\ & & \tilde{x}_{ij} & \\ A_m & & & \end{matrix}$$

$$i = 1, 2, \dots, m; j = 1, 2, 3, n,$$

$$\tilde{x}_{i,j} = (a_{i,j}, b_{i,j}, c_{i,j})$$

where i shows the alternative and j shows the criterion. Each $\tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$ is a triangular fuzzy number.

5. **Normalisation of fuzzy numbers.** To avoid the complicated normalization formula used in classical TOPSIS, Chen (2000) proposes a linear scale transformation in order to transform the various criteria scales into a comparable scale. The particular normalization method aims at preserving the property that the ranges of normalized triangular fuzzy numbers belong to

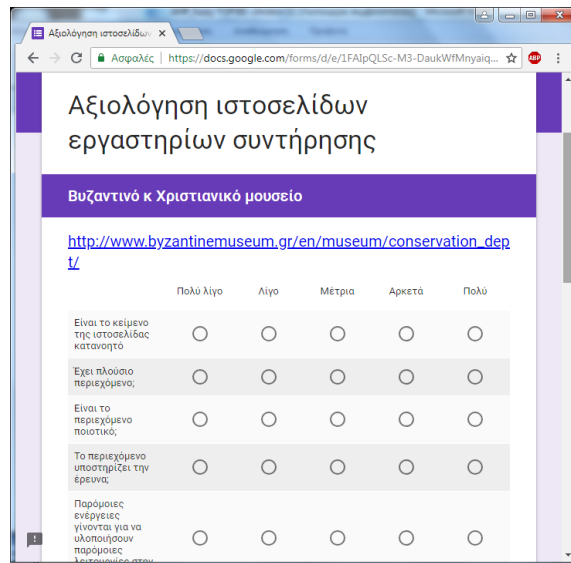


Figure 2: The questionnaire (in greek) of the empirical method.

[0,1]. The normalization of a fuzzy number $\tilde{x}_{ij} = (a_{ij}, b_{ij}, c_{ij})$ is given by the formula:

$$\tilde{r}_{ij} = \left(\frac{a_{ij}}{c_j^*}, \frac{b_{ij}}{c_j^*}, \frac{c_{ij}}{c_j^*} \right), \text{ where } c_j^* = \max_i c_{i,j}$$

6. **Calculating the weighted normalized fuzzy numbers of the MCDM matrix.** Considering the different importance of each criterion, which is imprinted in the weights of the criteria, the weighted normalized fuzzy numbers are calculated: $\tilde{u}_{i,j} = \tilde{r}_{i,j}(\bullet)\tilde{w}_j$ and these values are used to construct the weighted normalized fuzzy MCDM matrix $\tilde{V} = [\tilde{u}_{ij}]_{M \times N}, i = 1, 2, \dots, m; j = 1, 2, 3, \dots, n$

7. **Determination of the Fuzzy Positive-Ideal Solution (FPIS) and the Fuzzy Negative-Ideal Solution (FNIS).** The Fuzzy Positive-Ideal Solution (FPIS) and the Fuzzy Negative-Ideal Solution (FNIS) are calculated as follows:

a. FPIS: $A^* = \{\tilde{u}_1^*, \tilde{u}_2^*, \dots, \tilde{u}_i^*, \dots, \tilde{u}_n^*\},$

$$\tilde{u}_j^* = (1,1,1)$$

b. FNIS: $A^- = \{\tilde{u}_1^-, \tilde{u}_2^-, \dots, \tilde{u}_i^-, \dots, \tilde{u}_n^-\},$

$$\tilde{u}_j^- = (0,0,0)$$

Linguistic term	Fuzzy number
Very Poor	(1,1,3)
Poor	(0,1,3)
Fair	(3,5,7)
Good	(7,9,10)
Very Good	(9,10,10)

Table 7: Linguistic terms assigned to fuzzy numbers.

8. Calculation of the distance of each alternative from FPIS and FNIS

The distances (d_i^* και d_i^-) of each weighted alternative $i = 1, 2, \dots, m$ from FPIS and FNIS is calculated as follows:

$$d_i^* = \sum_{j=1}^n d_u(\tilde{u}_{ij}, \tilde{u}_j^*) \quad , \quad i = 1, 2, \dots, m$$

$$d_i^- = \sum_{j=1}^n d_u(\tilde{u}_{ij}, \tilde{u}_j^-) \quad , \quad i = 1, 2, \dots, m$$

where $d_u(\tilde{a}, \tilde{b})$ is the distance between two fuzzy numbers \tilde{a}, \tilde{b} . The distance of two fuzzy numbers $\tilde{a} = (a_1, a_2, a_3)$ and $\tilde{b} = (b_1, b_2, b_3)$ is calculated

$$d(\tilde{a}, \tilde{b}) = \sqrt{\frac{1}{3}[(\alpha_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]}$$

9. Calculation of the closeness coefficient of each alternative. The closeness coefficient of each alternative j , is given by the formula

$$CC_i = \frac{d_i^-}{d_i^* + d_i^-}, \quad 0 \leq CC_i \leq 1.$$

According to

the values of the closeness coefficient, the ranking order of all the alternatives is determined. The alternative that is closer to FPIS and further from FNIS as CC_i approaches 1. The values of the closeness coefficient of each alternative and the final ranking of the evaluated websites are presented in table 8.

6 Discussion

Museum websites and especially museum conservation labs play an important role in promoting culture. However, a website has to be evaluated so that its effectiveness is verified. Despite its importance, this phase is often omitted by the website life-cycle especially when several criteria are to be checked (Nilashi & Janahmadi 2012). In order to make the evaluation experiment easier for professionals, researchers and students to implement we present in detail the steps that one has to take in order to combine different evaluation methods and different multi-criteria decision-making theories.

The proposed method uses a combination of inspection and empirical method. More specifically, the evaluation experiment is implemented into two phases. In the first part, the inspection method is implemented and in the second part, an empirical method is used. In the first phase, in which the criteria and the weights of the criteria are estimated, expert users can more effectively provide such information. The conclusions are even stronger because both domain and computer experts are used. The

#	Museum Conservation Labs	CC_i
1	National Gallery of Greece	0.165445
2	Benaki Museum	0.163536
3	Metropolitan Museum	0.162603
4	Hermitage Museum	0.160753
5	Byzantine & Christian Museum in Athens	0.158755
6	Museo Del Prado	0.150677
7	Vatican Museum	0.150438
8	Archaeological Museum of Thessaloniki	0.148635
9	Victoria & Albert Museum	0.146742
10	Boston Museum of Fine Arts	0.145825
11	Guggenheim Museum	0.144999
12	MoMa	0.144952
13	De Young museum of Fine Arts	0.144765
14	Tokyo National Museum	0.142368
15	Smithsonian museum	0.139851
16	British Museum	0.138384
17	Tate Modern	0.138328
18	Australian Museum	0.134923
19	Rijksmuseum	0.132292
20	Brooklyn museum	0.131841
21	Oriental Institute Museum	0.130455
22	NTNU University museum	0.128826
23	Getty Institution	0.127965
24	University of Michigan Museum of Art	0.126843
25	Museum of Byzantine Culture in Thessaloniki	0.122426
26	Museum of Islamic Art - Doha	0.112785
27	Barberini – Corsini Gallery – Roma	0.105871
28	National Museum New Delhi	0.102325
29	Galleria Nazionale d'Arte Moderna	0.094584

Table 8. The final ranking of the websites based on the values of closeness coefficient of all alternatives.

implementation of the experiment using an inspection method is easier and cheaper than empirical. Despite the advantages of inspection methods, these methods are not appropriate for all kinds of evaluation experiments. For example, in the second part of the experiment, the perception of real users is needed. Therefore, for the second part of the experiment larger group of potential users of the websites was used. This method was more complicated and expensive compared to the previous method but it was considered essential due to the conclusions that had to be extracted.

The inspection method was implemented using AHP. AHP has the ability to model expert opinion

and, therefore, was considered ideal for being combined with an inspection method of evaluation. As a result, AHP was used for the calculation of the weights of the criteria. But AHP is a time-consuming technique because of the mathematical calculations and number of pairwise comparisons which increases as the number of alternatives and criteria increases or changes (Jadhav & Sonar 2011). Since complexity rises with the increase in websites, the number of alternatives that can be compared is limited. This is one of the main reasons for selecting to combine AHP with another theory.

The theory that was selected to implement the empirical method in the second phase of the evaluation experiment was Fuzzy TOPSIS. The complexity of TOPSIS's application does not increase with the same rate of AHP when the number of alternative websites increases. Therefore, the suitability of TOPSIS for the second phase of the website evaluation is inevitable. A main drawback of TOPSIS is that it does not provide a specified way for calculating the weights of criteria as AHP does. Taking into account the advantages and disadvantages of AHP and TOPSIS, these two theories have different reasoning but seem rather complementary. Furthermore, in the case of an empirical method, in which several evaluators are involved that do not have experience in implementing, Fuzzy TOPSIS seems more appropriate. The linguistic terms used in Fuzzy TOPSIS are easier for users to comprehend and use.

The results of the first part of the evaluation revealed that the most important criterion of the first level is Usability, followed by Content. Within the sub-criteria of Content, the Quality of Content was considered the most important criterion. Regarding Usability, the sub-criteria Structure/Navigation and Easy to use/Simplicity are considered almost equally important. As concerned Functionality, the existence of Multimedia is considered the most important criterion.

The results of the second phase of the evaluation revealed that the best website was considered to be the website of the conservation lab of the National Gallery of Greece. The particular website provided rich content related to the activities of the department, the different departments, the equipment, and the staff. Its content is enriched with multimedia. The user interface is well designed and generally, the website is well structured and usable. The website of the Benaki Museum in Athens was also rated high. However, one may be concerned with the fact that two Greek websites were rated first. Although the language is a factor that may have influenced the evaluators, one can also observe that other Greek sites have been ranked in the last five.

Two of the last ranked websites of museums' conservation labs are the websites of the National Museum of New Delhi and the Galleria Nazionale d'Arte Moderna. Their content was poor and there was no information about the staff, the facilities and the equipment. Furthermore, the websites had only a few

photos and no other multimedia. Finally, both websites did not appear to be updated until nowadays.

7 Conclusions

Websites of cultural content are targeted to a variety of users (Wubs & Huysmans 2006, Purday 2009, Sweetnam et al. 2012). Therefore, these websites have to address the needs and interests of a variety of users. In order to confirm that a website meets its goals, an evaluation experiment should be implemented. The evaluations are usually complicated procedures that focus on the examination of several different criteria. The particular paper focuses on the evaluation of the websites of museums' conservation labs. The conservation labs in the Museums serve a unique and separate scope and goal inside each institution (i.e. different staff, particular equipment, etc.). Therefore, these websites may differ from the main websites of the museum in terms of content and structure. The framework presented in this paper aims at the evaluation of websites of specialized cultural content in general. The websites of museums' conservation labs that contain specialized cultural content have been used as a testbed to test its functionality.

The main contribution of the particular paper is that it presents a framework for the evaluation of websites of specialized cultural content. This framework combines different methods and different multi-criteria decision-making theories in order to evaluate websites of specialized cultural content. More specifically, it is shown in detail the combination of inspection and empirical methods for evaluating the websites of specialized cultural contents such as the websites of conservation labs in museums. The combination of these two methods is made so as to benefit from the advantages of each method and restrict the disadvantages of each method.

Furthermore, the proposed approach shows how a multi-criteria decision-making theory, namely AHP, is combined with a fuzzy multi-criteria decision-making theory, namely Fuzzy TOPSIS, to evaluate websites of cultural content. AHP's main advantage is that it uses pairwise comparisons of criteria for estimating their weights. However, these pairwise comparisons increase complexity dramatically when the number of alternative websites increases. Therefore, AHP does not seem appropriate for the evaluation of 29 websites. A solution to this problem is given through the use of Fuzzy TOPSIS. The complexity of Fuzzy TOPSIS applications does not increase so dramatically with the increase of alternatives. Furthermore, Fuzzy TOPSIS uses linguistic terms and seems ideal for an experiment where real users, without prior experience in the implementation of multi-criteria decision-making theories, are involved.

It is among our future plans to use this framework in the evaluation of other websites of different specialized cultural content. Furthermore, we aim at trying other MCDM theories and comparing them to find the best combination for the purposes of

evaluation of cultural websites with specialized cultural content.

References

- [1] Büyüközkan D., D. Ruan. 2007. Evaluating government websites based on a fuzzy multiple criteria decision-making approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15(3): 321–343. <https://doi.org/10.1142/s0218488507004704>
- [2] Chen C.T. 2000. Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets and Systems*, 114: 1-9. [https://doi.org/10.1016/s0165-0114\(97\)00377-1](https://doi.org/10.1016/s0165-0114(97)00377-1)
- [3] Cunliffe D., E. Kritou, D. Tudhope. 2001. Usability evaluation for museum web sites. *Museum Management and Curatorship*, 19(3): 229-252. <https://doi.org/10.1080/09647770100201903>
- [4] Davoli P., F. Mazzoni, E. Corradini. 2005. Quality Assessment of Cultural Web Sites with Fuzzy Operators, *Journal of Computer Information Systems*, 46(1): 44-57.
- [5] Di Blas N., M.P. Guermant, C. Orsini, P. Paolini. 2002. Evaluating the Features of Museum Websites. In: *Museums and the Web 2002: Selected Papers from an International Conference* (6th, Boston, MA), April 17-20.
- [6] Dyson M. & K. Moran. 2000. Informing the design of Web interfaces to museum collections. *Museum Management and Curatorship*, 18: 391–406. <https://doi.org/10.1080/09647770000501804>
- [7] Garzotto F., M. Matera, P. Paolini. 1998. To use or not to use? Evaluating usability of museum web sites. In *Proceedings of Museums and the Web '98*, Toronto, Canada. Retrieved April 2016: http://www.museumsandtheweb.com/mw98/papers/garzotto/garzotto_paper.html <https://doi.org/10.1145/948496.948515>
- [8] Harms I. & W. Schweibenz. 2001. Evaluating the usability of a museum Web site. In D. Bearman, & J. Trant (Eds.), *Museums and the Web* (pp. 43–54). Pittsburgh, PA: Archives and Museum Informatics.
- [9] Hwang C.L., K. Yoon. 1981. *Multiple Attribute Decision Making: Methods and Applications*. Springer-Verlag, New York. <http://dx.doi.org/10.1007/978-3-642-48318-9>
- [10] Jadhav S., R. Sonar, 2011. Framework for evaluation and selection of the software packages: A hybrid knowledge based system approach. *Journal of Systems and Software*, 84: 1394–1407. <https://doi.org/10.1016/j.jss.2011.03.034>
- [11] Kabassi K. 2017. Evaluating Websites of Museums: State of the Art. *Journal of Cultural Heritage* (Elsevier), 24: 184-196. <https://doi.org/10.1016/j.culher.2016.10.016>
- [12] Karoulis S., S. Sylaiou, M. White, 2006. Usability Evaluation of a Virtual Museum Interface, *INFORMATICA*, 17(3): 363–380.
- [13] Lewis C.L., J. Rieman. 1994. *Task-centered User Interface Design: A Practical Introduction*, Boulder: University of Colorado.
- [14] Mohamadali N.A., J. Garibaldi. 2011. Comparing user acceptance factors between research software and medical software using AHP and Fuzzy AHP. In: *The 11th Workshop on Computational Intelligence*, 7 - 9 September 2011, Kilburn Building.
- [15] Mulubrhan F., A. Akmar Mokhtar, M. Muhammad. 2014. Comparative Analysis between Fuzzy and Traditional Analytical Hierarchy Process. *MATEC Web of Conferences* 13. <https://doi.org/10.1051/mateconf/20141301006>
- [16] Nagpal R., D. Mehrotra, P. Kumar Bhatia, A. Sharma, 2015. Rank University Websites Using Fuzzy AHP and Fuzzy TOPSIS Approach on Usability. *International Journal of Information Engineering and Electronic Business*, 1: 29-36. <https://doi.org/10.5815/ijieeb.2015.01.04>
- [17] Nilashi M., N. Janahmadi. 2012. Assessing and Prioritizing Affecting Factors in E-Learning Websites Using AHP Method and Fuzzy Approach. *Information and Knowledge Management* 2(1): 46-61
- [18] Purday J. 2009. Think culture: Europeana.eu from concept to construction. *The Electronic Library*, Vol. 33(2): 170-180. Available: <http://dx.doi.org/10.1108/02640470911004039>
- [19] Reeves T.C. 1993. Evaluating technology-based learning, In G.M. Piskurich (Ed.), *The ASTD Handbook of Instructional Technology*. McGraw-Hill, New York. 15: 1–32.
- [20] Saaty T. 1980. *The analytic hierarchy process*. New York, McGraw-Hill.
- [21] Sirah S., L. Mikhailov, J. A. Keane. 2015. PriEsT: an interactive decision support tool to estimate priorities from pair-wise comparison judgments. *International Transactions in Operational Research*, 22(2): 203–382. <https://doi.org/10.1111/itor.12054>
- [22] Soleymaninejad M., M. Shadifar, A. Karimi. Evaluation of Two Major Online Travel Agencies of US Using TOPSIS Method. *Digital Technologies*, 2(1), (2016), 1-8
- [23] Sweetnam M. S., M. Agosti, N. Orio, C. Ponchia, C.M. Steiner, E.-C. Hillemann, M. Siochrú, S. Lawless. 2012. User needs for enhanced engagement with cultural heritage collections. *Proceedings of Second International Conference TPDL*, Paphos, Cyprus, September, 23 – 27: 64 – 75. Available: <https://dx.doi.org/10.1007/978-3-642-33290-6>. https://doi.org/10.1007/978-3-642-33290-6_8
- [24] Sylaiou S., V. Killintzis, I. Paliokas, K. Mania, P. Patias. 2014. Usability Evaluation of Virtual Museums' Interfaces Visualization Technologies.

- In R. Shumaker and S. Lackey (Eds.): VAMR 2014, Part II, LNCS 8526, (2014), 124–133, 2014. © Springer International Publishing Switzerland.
https://doi.org/10.1007/978-3-319-07464-1_12
- [25] Tiwari N. 2006. Using the Analytic Hierarchy Process (AHP) to identify Performance Scenarios for Enterprise Application, Computer Measurement Group, Measure It, 4(3).
- [26] Vavoula G., M. Sharples, P. Rudman, J. Meek, P. Lonsdale, P. 2009. Myartspace: Design and evaluation of support for learning with multimedia phones between classrooms and museums. *Computers and Education*, 53 (2): 286-299.
<https://doi.org/10.1016/j.compedu.2009.02.007>
- [27] van Welie M., B. Klaasse. 2004. Evaluating Museum Websites using Design Patterns. Technical report number: IR-IMSE-001, December 2004, Vrije Universiteit, Amsterdam
- [28] Wubs H., F. Huysmans. 2006. Click to the past. The Netherland Institute for Social Research. [Online]. Available:
http://www.scp.nl/english/Publications/Summaries_by_year/Summaries_2006/Click_to_the_past
- [29] Zhang W. 2015. Group-Buying Websites Evaluation Model Based on AHP-TOPSIS under the Environment of Multi-Attribute Decision-Making. *International Journal of Multimedia and Ubiquitous Engineering*, 10(7): 31-40.
<https://doi.org/10.14257/ijmue.2015.10.7.04>
- [30] Zopounidis C. 2000. Foreword: Special issue on artificial intelligence and decision support with multiple criteria. *Computers & Operations research*. Vol. 27, 597-599.
[https://doi.org/10.1016/s0305-0548\(99\)00107-0](https://doi.org/10.1016/s0305-0548(99)00107-0)
- [31] Zopounidis C. 2009. Knowledge-based multi-criteria decision support. *European journal of operational research*, 195: 827-828.
<https://doi.org/10.1016/j.ejor.2007.11.026>