Available online at www.sciencedirect.com
ScienceDirect
Procedia - Social and Behavioral Sciences 148 (2014) 110 – 118
Procedia
Social and Behavioral Sciences

ICSIM

Tourist Destination Marketing Supported by Electronic Capitalization of Knowledge

George Stalidis*, Dimitrios Karapistolis

[a]*Alexander Technological Educational Institute of Thessaloniki, 17 Km Sindos-Thessaloniki, 57400 Thessaloniki, Greece*

Abstract

Marketing for tourist destinations, either as strategic planning or as short-term marketing-mix specification, has a lot to gain from computerized information tools. In this paper, we present knowledge engineering methods for the extraction and modeling of findings from market surveys and data analytics in the form of reusable and sharable knowledge. In contrast to information-based systems which are able to store and provide high quality information but their user relies on his own interpretation and decision abilities, we focus on a knowledge-based approach where data analysis and reasoning are consolidated and the system is thus able to provide solutions to common marketing problems. Data analysis methods suitable for discovering factors, associations, clusters and in general hidden patterns that explain a market phenomenon or customer behavior, are applied on multiple surveys related to tourist destination marketing. In addition, a computerized knowledge management process, utilizing ontologies and a rule-based inference engine, is developed to capitalize the extracted knowledge and offer it to support marketing planning problems. Preliminary results are presented on capturing the image of Thessaloniki as tourist destination and suggesting important factors for improving its promotion to individual visitor groups.

Selection and peer-review under responsibility of the 2nd International Conference on Strategic Innovative Marketing.
Keywords: Marketing decision support; knowledge engineering; multidimensional data analysis; tourist destination marketing;

1. Introduction

The access to high quality market information and the ability to make use of this information to take successful marketing decisions may be of decisive importance for the competitiveness of a destination. To this end, the analysis of survey data is also a challenging field, where specialized expertise and information technology offer advanced capabilities, such as to discern qualitative patterns and hidden dependencies, as well as to study phenomena evolving over time. Since more than two decades ago, it has been supported that

* Corresponding author. Tel.: +30-2310791246; fax: +30-2310791563.
E-mail address: stalidgi@mkt.teithe.gr

Selection and peer-review under responsibility of the 2nd International Conference on Strategic Innovative Marketing.
doi:10.1016/j.sbspro.2014.07.024

successful marketing in tourism depends on the extent to which more specialized consumer demands or lifestyles can be identified, as opposed to massive generic approaches (Weinstein, A., 1994). The positioning strategy of a tourism product can thus be devised following the measurement of the customer's image of the tourism product (Etchner, C. and Ritchie, B., 1993) and his satisfaction from product attributes, in correlation with his specialized needs and desires (Cho, B., 1998), (Kavoura, A., 2013). Such measurements of a destination's or product's image as well as the identification of the customer's needs are typically performed through primary surveys and statistical analyses. An interesting challenge would be to progress from a data or information-based approach to a knowledge-based approach, where the marketer would not be concerned with datasets, statistical tables or reports but with the knowledge derived from those. Moreover, it would be desirable to consolidate knowledge from different sources in order to solve a broader problem and to accumulate this knowledge in a reusable and sharable form, building a knowledge capital available to the marketer.

Strategic planning tools for improving the competitiveness of tourism in selected areas have been reported in the literature (Bousset, J.P., et al, 2007), based mainly on information management and less on sophisticated analysis and knowledge extraction. Knowledge-based decision support systems applied to tourism marketing have also been reported (Moutinho, L., Rita, P. and Curry, B., 1996). In the current paper we present knowledge engineering technologies (Schreiber, G., 2008) that are used to extract knowledge from survey data and to store it in an electronic Knowledge Base, so that it can be used within intelligent computerized systems as an advanced marketing decision support tool. The main focus is on the design of a knowledge model suitable for tourist destination marketing and the illustration of a proposed methodology for managing knowledge derived from questionnaire-based primary surveys. Multidimensional factor and clustering analysis methods (Benzecri, J.-P,, 1992) are used as a powerful knowledge extraction method and original results are presented from a recent marketing survey on the image of Thessaloniki as a tourist destination.

2. Methods

2.1. Multidimensional data analysis

In order to uncover the relations among different aspects of the visitors' image for a destination, we applied multi-dimensional factor analysis and clustering methods. In specific, a combination of Multiple Correspondence Analysis (MCA) (Benzecri, J.-P, 1992), (Greenacre, M., 2007) and Hierarchical Cluster Analysis (CHA) based on Benzecri's chi-square distance and Ward's linkage criterion (Benzecri, J.-P, 1992) were used in a multistep analysis procedure, in order to identify distinct perceptions in terms of the respondents' views to specific aspects of their visit, to cluster visitors in terms of their preferences and viewpoints and to discover the relations among different variables in order to promote our understanding as regards key satisfaction factors. The specific analysis methods are produce results in qualitative form, allowing graphical exploration and the formation of patterns that involve classes, properties and associations (Van de Geer, John, P., 1993). The data analysis process was performed using the MAD analysis software (Méthodes d' Analyses des Données – MAD, 2012) and comprised the following steps:

1. For each block of items in the questionnaire (corresponding to an individual aspect of the survey), a combination of MCA and CHA were applied on the variables included in the specific block. The clustering method was used to divide the respondents into homogeneous groups in terms of their responses and the factor analysis was performed to project the profiles of the respondent clusters on the factorial axes, together with the properties (i.e. categories or modalities) of the main variables, allowing to match the groups of respondents with groups of properties.

2. After completing the above clustering process separately for each block, MCA was applied on all the cluster membership variables of the previous step. On the resulting factorial planes it was possible to observe an overall picture of the dependencies among all structural variables.
3. The analysis was focused on specific groups of interest by applying a similar methodology on selected subsamples, in order to identify in more detail the profiles of target segments.

2.2. Knowledge engineering

Knowledge Engineering (KE), was defined in 1983 by Edward Feigenbaum, and Pamela McCorduck as an engineering discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise (Feigenbaum, E. A. & McCorduck, P., 1983). The ability to infer and suggest solutions by utilizing knowledge is passed from the human analyst to an inference engine. Such an approach has several strong points. Firstly, KE facilitates the storage and consolidation of large amounts of knowledge, which can be applied simultaneously by an intelligent system to solve complex problems, while on the contrary it would be impossible for an expert to consider as a whole. Continuous updates of electronic knowledge are possible by introducing new findings and by correcting or discarding obsolete ones, taking advantage of algorithmic methods for automatic checking against inconsistencies. The knowledge capital becomes coherent, maintainable and reusable. Thirdly, and maybe most importantly, the knowledge accumulated in a Knowledge Base (KB) can be used through intelligent decision support tools – incorporated into a Knowledge-Based System (KBS) - to provide solutions to complex problems, without requiring expertise or deep comprehension by the user. By adopting the appropriate formalism, knowledge can be exchanged/accessed between systems or over the Internet.

The stages involved in acquiring and using knowledge in computerized form are several (Ligêza, A., 2006). The most important aspects are Knowledge Extraction, Knowledge Modeling and Inference, corresponding respectively to producing knowledge from data (e.g. through statistical analysis or data mining), expressing knowledge in a formal, standardized and machine-understandable form, and using the accumulated knowledge for problem solving (Schreiber, G., Akkermans, et. al., 1999). Available modeling approaches include ontologies, statistical models, neural networks, rule-based models, case-based reasoning models, each one offering different level of expressiveness and suitability to different kinds of problems. The modeling frameworks that became prominent within KE are Common KADS (Schreiber, G., Akkermans, et. al., 1999), Model-based and Incremental Knowledge Engineering (MIKE) and the ontology-based framework Protégé (2013), which is adopted in this work. An Ontology provides the basis for building a model of a domain, defining the terms inside the domain and the relationships between them (Gruber,T., 1993) and is thus a powerful and widely adopted tool, not only for developing models but also for communicating structured knowledge e.g. through semantic web or otherwise. In a recent review of existing ontologies in the tourism sector (Prantner, K., Ding, Y., Luger, M., Yan, Z., 2007) it was found that there are considerable number of efforts, such as the QUALL-ME (Ou, S., Pekar, V., Orasan, C., Spurk, C., Negri M., 2008) and the DERI e-tourism (DERI, OnTour Ontology, 2011) ontologies.

Considering that the complexity of the knowledge to be engineered is usually higher than class hierarchies and relations between objects, additional expressiveness is offered by rule-based models. A rule-based knowledge framework consists of production rules (Studer R., Benjamins R. & Fensel D., 1998), which are generally expressed in the form:

$$C_1 \text{ AND } C_2 \text{ AND } \ldots \text{ AND } C_n \text{ -> } E \qquad (1)$$

where C_1, C_2, . . . , C_n constitute the conditions of the rule and E is the consequent, which can be a prediction or suggestion. In our case, rules of this form mainly result from the data analysis and correspond to conditional associations between classes or individuals (e.g. if visitor is young and his purpose of visit is vacation, a first priority decision factor is nightlife) or to classifications (e.g. to characterize a trip as a low-budget trip).

2.3. The proposed Knowledge model for destination marketing decision support

The current work includes the construction of a rule-based knowledge framework, to represent the relations among destination image and visitor needs. The proposed model, as shown diagrammatically in Figure 1, includes the following components:

(a) An ontology for the tourist domain, providing the necessary terminology regarding the concepts found in the addressed problem (visitor, destination, trip, hotel etc.) and their properties (e.g. a visitor has as properties his age, country, education, etc.). This terminology provided a standardized basis for expressing knowledge content in a hierarchical structure that reflects the conception of the corresponding real-world domain.

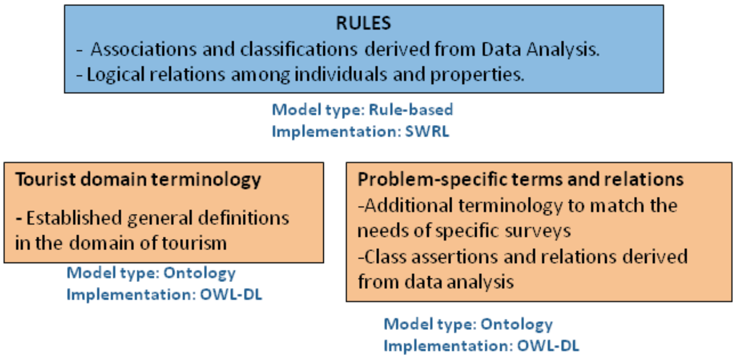

Fig. 1. Architecture of the Knowledge Model(b) A problem-specific ontology component that expands the above basic domain ontology to support the special terminology needs of individual knowledge source and some of the analysis findings, such as visitor profiles (e.g. an identified category of visitor who has negative image regarding prices and quality of services but positive regarding sights and culture).

(c) A rule-based component to express more complex associations among concepts, individuals and their properties, incorporating logic and operations.

The full model and the corresponding Knowledge Base were implemented in Protégé OWL 4.2 (Studer R., Benjamins R. & Fensel D., 1998). The language used for defining ontologies was OWL - DL (Ontology Web Language - Description Logic) in XML syntax (OWL, W3C Recommendation, 2013) and the rules were compiled in SWRL (Semantic Web Rule Language, 2004). An illustration of the constructed model with its preliminary contents is given in section 4.3.

3. Application in marketing decision support for destinations

3.1. Survey and Data Analysis

The purpose to be served by the presented knowledge capitalization approach was the capturing of trends, facts, customer needs and perceptions in the field of tourism, so that these can be employed by a Destination Management Organization (DMO) to enhance marketing planning. In order to precise the problem, the authors collaborated with the DMO of Thessaloniki (Organization for the Touristic Development and Marketing of Prefecture of Thessaloniki), who was involved as an end user. Input data were collected through a primary survey in the city of Thessaloniki, specifically for the purposes of the current research. The instrument was a 3-page structured questionnaire containing 43 questions, organized in 8 blocks, including questions regarding their satisfaction, the reasons for choosing this destination, the perceived image of the city and the country, as well as personal/demographic information. The sample, which was available for this study was 391 respondents from the period from June to mid-July 2013. The survey is in progress, aiming at 2000 questionnaires around a full year.

The analysis process was performed as outlined in Section 2.1. As regards the image of Thessaloniki, the analysis resulted in 5 groups of visitors. The main factors were visualized using MCA, as shown in Figure 2. The 1st factorial axis (explaining 28,9% of inertia) represented the contrast between the fully negative and fully positive properties. The 2nd factorial axis (22,8% of inertia) represented the degree of indifference. The 3rd factorial axis (5,7% of inertia) expressed the contrast between the image for security and the image for sights and leisure, positioning between these two ends, the image regarding practical issues such as prices, cleanness and hotels. As regards the priorities expressed by the respondents for selecting their destination, four groups were identified: (1) the ones who were mainly interested in visiting museums, (2) those who were attracted by night-life and life-style, (3) attracted by the natural beauty of the area, the climate, natural environment and excursions and (4) the summer tourists who arrived for the beaches/swimming.

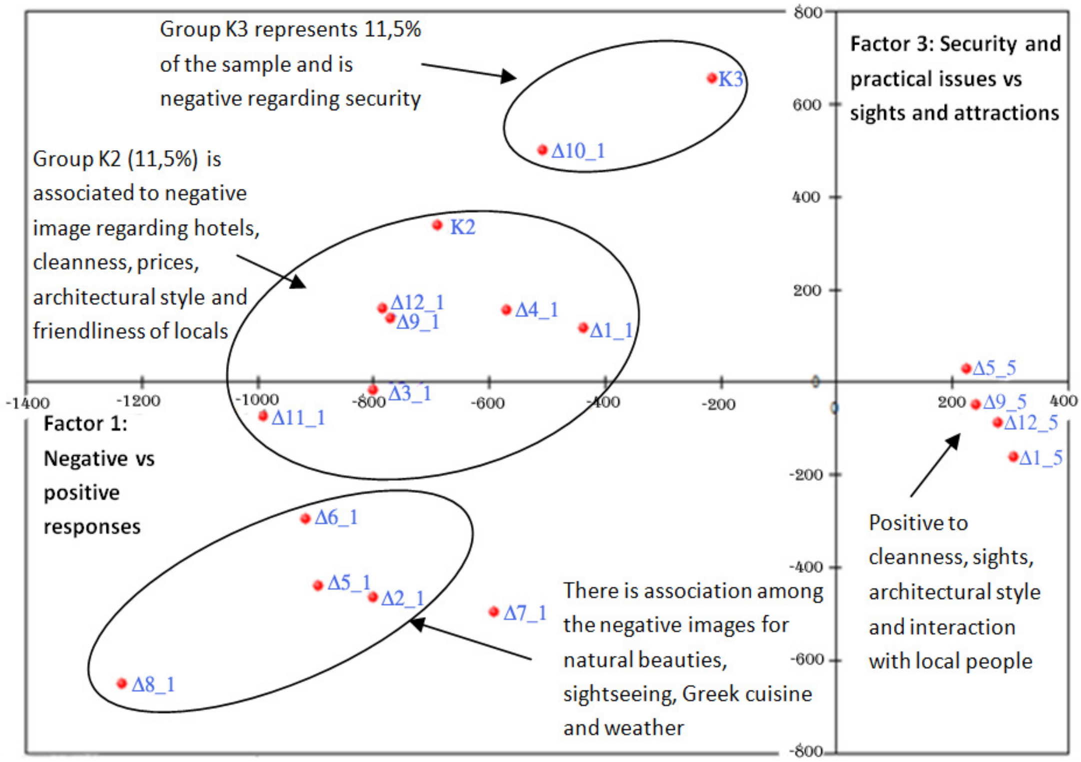

Fig. 2. Factorial plane 1X3 (34,6% of inertia) projecting the responses on the city's image. F1 differentiates negative from positive responses, while F3 differentiates concerns with security (top) from negative image for hotels, cleanness and prices (center) up to distaste for natural beauties, sights, cuisine and weather (bottom). Notation: e.g. point Δ7_1 represents response 1 (lowest/negative) to question 7 of block Δ.

In the next step, an overall picture of the existing dependencies was obtained by applying factor and clustering analysis on all group-membership variables simultaneously. According to the factorial plane 1X2, the most positive image for the city was associated to visits for professional reasons, shopping or health/medical reasons. On the opposite side of the plane we found those who were negative regarding the security in the city, the reasons for selecting this destination were entertainment, night-life and lifestyle and they responded that they were influenced by advertisements. A third grouping of properties on the same plane showed that the response "a future visit is possible" characterized the group of visitors who were mainly attracted by the natural beauties of the area and the climate, they arrived for excursions and contact with nature and that the same group was related to "attracted by the beaches/swimming" and "influenced by the fame of the destination". After introducing as complementary variables the country, age, income and gender, some overall trends were visible, such as: the profile of visitors from Russia was similar to Australia and South America and included the properties of positive image, high proportion of ages 46 to 65, wish to visit museums, negative to night-life, purpose of visit either touristic or religious. In the negative direction we found North-Europeans from Baltic countries and the age group above 65 years old. No pattern was found at this broad scale of observation regarding gender and income, neither was it possible to perform more detailed analysis per country because of the small sample. In the near future it is planned to obtain a much larger sample, so that clear profiling per country will be possible.

In order to investigate the feasibility of consolidating knowledge from different sources, the results of a previous survey (Stalidis, G., 2012) on a related subject were used as additional input to the Knowledge Base. This research was carried out during the summer of 2010 in Northern Greece, to evaluate the perceived quality of the hotels in the area and to study the factors explaining satisfaction, in order to identify representative visitor classes and to associate them with priority expectation attributes. The dataset included a sample of 400 visitors, most of who stated that the purpose of their visit was vacation and just 8% mentioned professional reasons. The data underwent multidimensional Factor and Cluster analysis an extract of the results is illustrated in Table 1.

Table 1. Classification of visitors in terms of their requirements from their hotel

Group	Associated Class and its Properties	Demographic properties
K1 N=174, 43,4%	**Low budget** Willing to spend up to 50€, Low expectations overall.	
K2 N=102, 25,5%	First time visit to this destination, Willing to spend up to 50€	Age 26-35
K3 N=32, 8,1%	**Basic quality vacation** Purpose of visit is vacation or entertainment, Duration 2 weeks, Willing to spend 50 to 100€, High expectations for security, cleanness, materials and consumables (i.e. shampoos, towels, etc.), hairdryer, restaurant in the hotel.	Age 56-65, Profession: retired
K4 N=16, 4%	**Recreational vacation** Has been in this destination once or twice before, The hotel to belong to a group of hotels, swimming pool, entertainment activities, comfortable lobby.	Age 56-65, Profession: retired
K5 N=76, 18,9%	**Activities and wellness** Duration 3 weeks or more, Willing to spend 100-150€ per night, Requirements for spa and wellness services, facilities for persons with special needs, sports facilities, special diet menu.	Age 46-55, Profession: freelancer, Medium income

3.2. Knowledge elicitation

The transformation of data analysis results into contents of a Knowledge Base relies on the knowledge model. The latter keeps building up as the coverage of the knowledge to be introduced becomes wider. Figure 3 illustrates the consolidated ontology class tree, which reflects both the main problem of destination marketing support based on the destination's image and the additional research on hotel satisfaction. The classes were defined to reflect the concepts found in both cases, the most important ones being Visitor, Travel, Destination and Accommodation, with their subclasses defined in hierarchical form, e.g. Destination has the subclasses RuralArea and UrbanArea, subclasses of UrbanArea are City and Town, while individuals belonging to the class

City are Thessaloniki and Athens. Subclasses were also defined to match survey variables e.g. ArchitecturalStyle and PeopleFriendliness as subclasses of ImageAttribute. An important step in the ontology construction was to introduce the classes that resulted from Data Analysis, in order to be able to describe them in the KM and to assign properties to them, for example VisitorForNature was defined as the subclass of visitors who stated as decision factor for selecting this destination the NaturalBeauty, Climate and Tours. Additionally, OWL properties were defined to represent relations, such as hasRequirement that links a Visitor with an Activity or HotelFeature.

The next step was rule formation, where more complex data analysis results with interpretation value were formulated in the form (1). The following rule written in SWRL contains classes, individuals and properties defined in the ontology:

Visitor(?v),VisitorForNightlife(?v),hasVisitorImage(?v,NegSecurityDestinationImage)
-> hasDecisionFactor(?v,InfluencedByAdvertisement)

The rule declares that IF v (v is an arbitrary variable name) belongs to class Visitor AND v belongs to class VisitorForNightlife (already defined in the ontology to contain those who responded that the reason for selecting the destination was nightlife, entertainment or lifestyle) AND v is connected through the property hasVisitorImage to the object NegSecurityDestinationImage (the latter is the name given to the negative response to the questionnaire item "security" regarding destination image) THEN v is connected through the property hasDecisionFactor to the object FirstPriorityInfluencedByAdvertisement (meaning that the visitor is predicted to have been influenced by an advertisement in order to select this destination).

It is noted that the compilation of such rules involves the interpretation of the analysis results by a domain expert, an evaluation of which rules would be useful/credible to be introduced in the KB and several more steps such as checks for inconsistencies, optimization, maintenance, etc. The outcome was at the current stage 30 rules and 8 analysis-based class definitions, while 64 more rules were derived from the hotel satisfaction survey. The introduction of findings from two independent surveys into a common Knowledge Base was performed successfully, taking advantage of the ability of the consolidated knowledge model to describe common concepts without inconsistencies, as well as to satisfy the requirements of both sources for terminology definitions and expressiveness.

4. Conclusions

In the presented work, multidimensional data analysis methods were coupled with knowledge modeling, as two important components of an envisaged Knowledge-Based framework for producing and accumulating knowledge for decision support in the destination-marketing domain. It was shown that the proposed analysis methods perform well as a knowledge extraction mechanism, having the ability to reveal qualitative information from survey data in explorative fashion. The application of the methods on primary survey data showed the feasibility of extracting qualitative knowledge from data and consolidating knowledge from different surveys. Limitations of the current work were that the Knowledge Base has not yet been linked to an inference engine and that the knowledge content was not rich enough in order to perform tests on the decision support abilities of the results in realistic marketing tasks. Future steps include the deeper analysis of the destination image survey (e.g. per country and per type of visit) in order to enrich the knowledge content, the expansion of the KM to cater for additional data sources and the incorporation of a query mechanism. The populated KBS will then be tested regarding its problem solving abilities by the DMO of Thessaloniki in pilot marketing actions. An additional challenge is to introduce the time dimension into the model, in order to deal with temporal trends. Based on the feedback to be received, it is planned to eventually build a marketing platform offering solutions to typical marketing application scenaria.

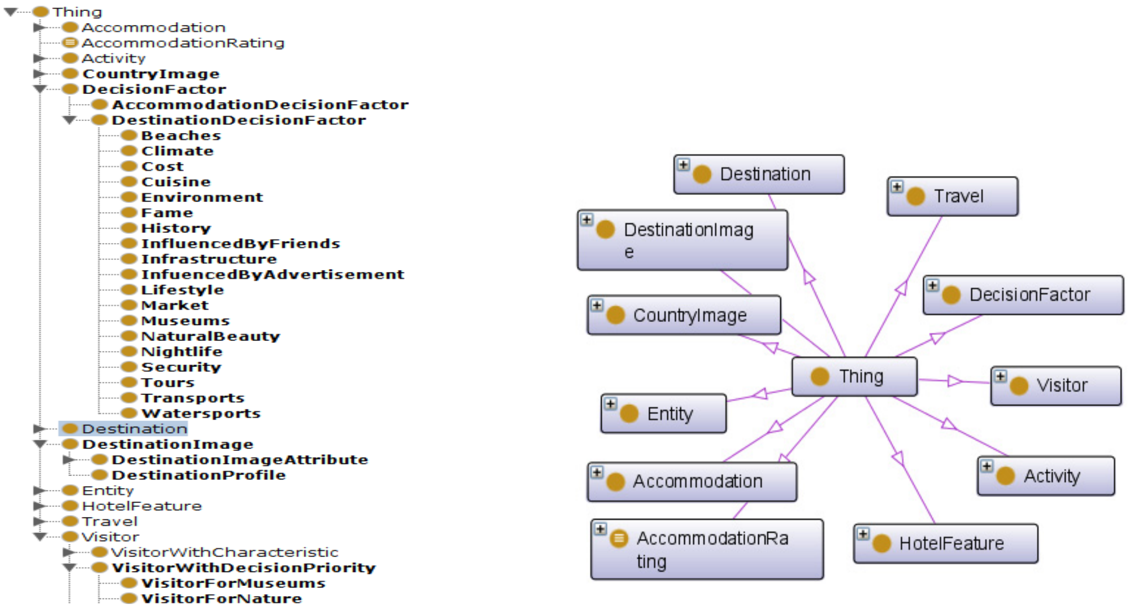

Fig.3. An extract of the class hierarchy (a) and the first level classes (b) of the consolidated ontology for the destination image and the hotel satisfaction surveys

Acknowledgements

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: ARCHIMEDES III.

References

Benzecri, J.-P, (1992). Correspondence Analysis Handbook. New-York: Dekker, P.
Bousset, J.P., et al, (2007).A Decision Support System for Integrated Tourism Development: Rethinking Tourism Policies and Management Strategies. Tourism Geographies,9:4,387-404.
Cho, B. (1998), “Assessing tourist satisfaction”, Tourism Recreation Research, Vol. 23 No. 1, pp. 47-54.
DERI, OnTour Ontology (2011). http://etourism.deri.at/ont/index.html, visited: 15/12/2011.
Etchner, C. and Ritchie, B. (1993). The measurement of destination image: an empirical assessment. Journal of Travel Research, pp. 3-13, Spring 1993.
Feigenbaum, E. A. & McCorduck, P. (1983). The fifth generation (1st ed.), Reading, MA: Addison-Wesley, ISBN 978-0-201-11519-2, OCLC 9324691
Greenacre, M., 2007. Correspondence Analysis in Practice, Chapman & Hall.
Gruber,T.,(1993). A Translation Approach to Portable Ontology Specifications, Knowledge Acquisition, 5(2), pp. 199-220.
Kavoura, A. (2013). Politics of Heritage promotion: Branding the identity of the Greek state. Tourism Culture and Communication, 12, 69-83.
Ligêza, A., (2006). Logical foundations for rule-based systems, Studies in Computational Intelligence, vol. 11, 2nd Ed., Springer-Verlag Berlin Heidelberg.
Méthodes d' Analyses des Données - MAD (2012), http://www.mkt.teithe.gr/en/research/research-labs/madlab/mad/ , last visited: 30/6/2013.

Moutinho, L., Rita, P. and Curry, B. (1996). Expert Systems in Tourism Marketing, London and New York: Routledge.

Ou, S., Pekar, V., Orasan, C., Spurk, C., Negri M., (2008). Development and Alignment of a Domain-Specific Ontology for Question Answering, in European Language Resources Association (ELRA) (ed.): Proc 6th Intl Language Resources and Evaluation (LREC'08), Marrakech, Morocco.

OWL, W3C Recommendation (2013), http://www.w3.org/TR/owl-features/, visited 3/5/2013.

Prantner, K., Ding, Y., Luger, M., Yan, Z. (2007). Tourism Ontology and Semantic Management System: State-of-the-Arts Analysis. Proceedings of IADIS International Conference WWW/Internet 2007, pp.111-5.

Protégé, (2013), http://protege.stanford.edu/ , last visited:30/5/2013.

Schreiber, G. (2008). Knowledge Engineering. In: Handbook of Knowledge Representation, F. van Harmelen, V. Lifschitz, B. Porter (Eds.), Elsevier, 2008, pp. 929–946.

Schreiber, G., Akkermans, et. al. (1999), “Knowledge Engineering and Management: The CommonKADS Methodology”, MIT Press, 1999.

Stalidis, G. (2012). Explorative Data Analysis and Knowledge Modeling methods for marketing decision support applied in the tourist sector. ICCMI 2012, Thessaloniki, pp 601-7.

Studer R., Benjamins R. & Fensel D. (1998). Knowledge Engineering: Principles and methods. Data & Knowledge Engineering 25 (1998) 161-197.

SWRL: A Semantic Web Rule Language Combining OWL and RuleML (2004), National Research Council of Canada, Network Inference, and Stanford University, http://www.w3.org/Submission/SWRL/, last visited 30/5/2013.

Van de Geer, John, P. (1993). Multivariate Analysis of Categorical Data: Applications Newbury Park: Sage Publications Inc. Advanced Quantitative Techniques in the Social Sciences Series Vol 3.

Weinstein, A. (1994), Market Segmentation: Using Demographics, Psychographics, and other Niche Marketing Techniques to Predict and Model Customer Behaviour, Probus, Chicago, IL.