2020-1119 IJOI
https://www.ijoi-online.org/

CONSTRUCTING THE MECHANISM OF KEYWORD EDITOR TOWARDS ONLINE TOURISM INFORMATION DATABASES

Wei-Hsiung Chang
Associate Professor, Department of Tourism Management,
Nanhua University
hector629@gmail.com

Abstract

Recently, e-tourism has been notable for the travel industry around the world. This paper aims to specify and structuralize the appropriate mechanism of keyword editors in online tourism information databases. Also, the appropriate mechanism is constructed for being utilized by tourists and it can be significant for the tourism industry businesses. Our results reveal that the digital tourism industry keyword editing designer (referred to as "mechanism designer") uses the traces of the tourism industry's online keyword library to show the relevance of various chapters and subsections to automatically adjust the standpoint based on the requirements of the tourists. Furthermore, the order of arrangement is adjusted after revision and reprint; the chapter layout is automatically corrected before the first edition, according to the keywords of the basic content provided by the original product supplier (tourism businesses), and the original chapter layout is tracked, according to the trajectories left by the tourists after reading the previous edition. Accordingly, the statements and perspectives interpreted above may be applied and utilized as the main editing mechanisms for automatic editing of online tourism information databases in the travel industry.

Key words: Online Tourism Databases, E-Tourism, E-Commerce, Travel Industry

Research Background and Motivation

When a tourist is traveling in a foreign country, they may wish to ask for directions to reduce the likelihood of losing their way. However, this information could be confusing for the tourists. According to Li and Chang (2016), the integration of digital technology with cultural tourism has resulted in tourists no longer being satisfied with passive, pre-arranged tours. Therefore, there is a need to consider the use of smartphone-based route navigation systems to enhance tourist satisfaction. However, as each person has a different way to express spatial information, they will tend to use different types of information or expressions to help people reach their destination; for example, different people use landmarks, distance, direction, turns, indications, and walking distance (Chang, 2016; Denis, Pazzaglia, Cornoldi, & Bertolo, 1999; Golding, Graesser, & Hauselt, 1996; Mark & Gould, 1995; Vanetti & Allen, 1988; Wright, Lickorish, Hull, & Ummelen, 1995) differently to communicate.

Previous studies have shown that the route directions provided by an individual depends on their cultural background, and that different individuals use different ways to describe their spatial environment. In the future, it is likely that tourists will rely only on their smartphone and locating system (e.g., Google Maps) to perform tourist activities, without using a guide or joining a tour. These tourists will then use a centralized spatial information system that has full access to their location at all times to find their way. The maturation of the current technological environment and emergence of new technologies has resulted in the gradual digitalization of tourism models. Consequently, the general public now expects information to be transferred via human-machine interfaces that facilitate two-way communication and return a variety of information, instead of textual media (Li & Chang, 2016).

As each person has different preferences in the way they use language to describe information, complexity or ambiguity in the directions provided by a

passerby could cause the tourist to move in the wrong direction (Chang, 2016; Dale, Geldof, & Prost, 2002; Hund, Schmettow, & Noordzij, 2012). Several studies have tried to assess the factors that influence tourist preferences in tour route navigation and how this information is used. Such studies include an estimation of tourists' preferred route planning strategies, gender, cultural background, and self-perceived sense of direction (Chang, 2016; Hund & Minarik, 2006; Saucier et al., 2002). On this basis, we created a system that filters and reorganizes the information provided by tourism businesses (keywords) and tourists (via questionnaires) to provide tourists with the information they need. This includes the tour's starting and ending points, number of days, number of people, required arrangements, costs, seasons, and transportation. In this way, the current approach is a holistic mapping of the structural information of tourist systems. According to the above statements, this paper attempts to construct the mechanism of keyword editors toward online tourism information databases from the perspective of automated editing mechanisms for providing directions to tourists.

Literature Review

Hund and Padgitt (2010) examined the usefulness of each type of wayfinding information for different individuals, and found that the individuals rated landmark and left-right information more highly than cardinal directions (east, west, south and north). Their experiments indicated that the worst-rated directions resulted in the shortest actual navigation time, whereas the best-rated directions produced the least benefit for the wayfinder (Chang, 2016). Moreover, Hund et al. (2012) proposed that the provider and recipient of route information need to work together to understand the type of directions needed by the recipient, as this would facilitate the delivery of directional information that is actually useful for the way-finder. The findings above begs the following question: when one provides route navigation information to tourists, should the information be based on the tour's starting and ending points, number of days, number of persons, required arrangements, costs, season, modes of transportation, landmarks, tourists' cultural background (Hund et al., 2012) or other types of spatial information? Which type

of spatial information would be most useful for foreign tourists? Currently, no conclusion has been reached in this regard.

Further, all information providers have constructed comprehensive keyword databases for the tourism industry by filtering and reorganizing the relevant spatial information. This is done to enable independent tourists easily search for this information and satisfy their complex needs via the Internet (Lin, Chang, and Chen, 2020).

Since 2006, Taiwan has used "Taiwan, touch your heart" as its tourism slogan to announce the world-famous friendliness and warmth of the Taiwanese people, who have always been willing to assist tourists in finding their way. However, has this friendliness been helpful to the tourists, or has it exacerbated their difficulties? We assess this problem in our study regarding the informatization of tourism systems. Pan and Fesenmaeier (2006) have noted that internet-based vacation planning is a complicated, dynamic, and contingent process, which may be deconstructed into a hierarchical structure of chapters and episodes. Each user will employ their own semantic models to parse complex trip planning problems and handle these problems in independent "chapters." These chapters are then divided into individual "episodes," which relate to alternative solutions for each chapter. When a tourist searches for information on the Internet, they often use authoritative websites (information hubs) that contain large amounts of information about their destinations, such as official tourism websites or business portals, to speed up their browsing processes (Yen, 2014). Lin et al. (2020) proposed three critical points to reflect the crucial implications on the mechanism of keyword editors toward online tourism information databases: (1) online tourism databases with a fixed publication period can be obtained easily by tourists, and they also allow businesses to plan and establish an integrated publication production line, which includes article writing, editing, proofreading, and delivery. (2) To improve the search efficiency of the keyword database and expand its userbase, it is necessary to consider the keywords that are required to present the contents of the keyword database, and the tourism knowledge

structures formed by the keyword groups. (3) By optimizing the correspondence between the keyword groups and the level of tourism information analysis within the data structure, the tourists will then be able to obtain any facet of tourism information, at any level of analysis they desire, by using keywords to search the tourism database. In other words, it implies that the level of tourism information analysis is the level of detail by which the contents of the keyword database describe some topic in the online tourism database.

Research Design and Procedures

For online tourism information databases already in operation, it is extremely critical to construct an effective mechanism for the existing structure of a keyword editor through a feasible and accurate information source. Hence, three issues must be considered before an online tourism database can be established: the optimal level of information analysis in the keyword database, the IT equipment and knowledge structures of the keyword database, and the target audience of the keyword database. To improve the search efficiency of the keyword database and expand its userbase, it is necessary to consider the keywords that are required to present the contents of the keyword database, and the tourism knowledge structures formed by the keyword groups. By optimizing the correspondence between the keyword groups and the level of tourism information analysis within the data structure, the tourists will then be able to obtain any facet of tourism information, at any level of analysis they desire, by using keywords to search the tourism database. Please note that the level of tourism information analysis is the level of detail by which the contents of the keyword database describe some topics in the online tourism database.

Based on the above statements, this paper aims to specify and structuralize the appropriate mechanism of keyword editors in online tourism information databases: (1) The themes and chapter names of the online tourism database are determined by the specifications of the configuration document. The contents of each chapter will then be provided by experts in the tourism industry (who will be compensated), and they will be made members of the author group. The de-

gree of overlap between the keywords provided by the various authors will be used to form a mechanism to automatically arrange the chapters. This mechanism will be used to edit the first edition of the online tourism database; the order of the chapters will be revised according to the usage patterns of the tourists in subsequent editions. (2)

According to the perspectives proposed by Lin et al. (2020), we utilize a keyword database for the tourism industry, which is filled with various tourism businesses with their keywords. A classification mechanism will be designed to organize keywords into an information architecture. Further, a questionnaire will be designed for tourists. Based on the demanders' responses, the information architecture will be filtered and reorganized to provide the starting/ending points, number of days, number of people, required arrangements, costs, season, transportation, landmarks, and cultural background of the tour. In light of these perspectives, the content provided by two professional travel operators who are familiar with the online keyword databases of the tourism industry have many highly relevant keywords overlapping, which means that the content provided by the two tourism operators is highly relevant and should be arranged together or closely as much as possible and will be decided automatically based on the real-time feedback from the tourist. However, the ordering of the contents of the same chapter of the online keyword database of the tourism industry involves the issue of the structure of the preliminary knowledge attribute. Before determining the structure of the preliminary knowledge attribute, the issue of who should be arranged first and who is arranged later should be further discussed. The overlapping concepts based on the above keywords should be arranged together; however, they may encounter difficulties in actual operation. For instance, if Content A and Content B and Content C all have a high degree of keyword overlap, but Content B and Content C do not have a high degree of keyword overlap, then Content A should only be close to Content B and Content C separately. If the three should lean together simultaneously, then we think that, in the previous example, Content A should appear before Content B, and Content C simultaneously as Content A to reflect the level of information analy-

sis under the content theme, as a search option for readers to choose whether to read or not.

Results

According to the procedures of analysis and construction, the above editing concepts are interpreted and presented mathematically as follows: Assume that W = {□1, □2, □3, ..., □n} is all keyword sets S = {□1, □2, □3, ... □m} is the basic content provided by all prospective authors; where the □ (□□, □□) value (between 0 and 1) is the degree of correlation between the i-th basic content and the j-th keyword (the relative number of times that the keyword □□ appeared in the i-th basic content □□).

To make the chapter ordering of the editor of the tourism online keyword database more flexible, we can use the dichotomy to dictate the relationship between the two based on what is "relevant" and "irrelevant." At this time, we can choose a value □ between 0 and 1. When □ (□□, □□) ≥ □, we recognize that □□ is related to □□ and record it as □□ (□□, □□) = 1; otherwise it is considered to have nothing to do with □□. Let □□ (□□, □□) = 0. It is easy to know that the larger the threshold, the fewer keywords related to the basic content.

It may be supposed that there are 9 keywords written as w = {□1, □2, □3, ... □9} and the basic content provided by 4 prospective authors is written as S = {□1, □2, □3, s4}. If a certain value of □ is determined and the relationship of □□ is used, and it is considered that there are more than two common basic contents (denoted as q = 2), they should be combined (organized) together, as shown in Figure 1.

As shown in Figure 1, since □3 overlaps with two or more keywords □2 and s4 simultaneously, □3, 2, and s4 should be combined together (although □2 and □4 do not overlap with keywords). As □1 and □2 have □3 keywords □3, □4, and □5 overlapping, □1 and □2 must be combined together; thus, □1, □2, □3, and □4 should all be arranged together into sub-contents under heading A (See Figure 1). At this time, we call {□1, □2, □3, □4} a combination of q = 2. If you want to add several chapters between the title A and its basic content in the combined diagram in Figure 1, you can change the

relationship between q = 2 and q = 3 as follows: as □1 and □2, and □3 and □4—all 3 keywords—overlap, the basic content □1 and □2 are combined under the relationship of q = 3 and together become the sub-content under heading B; □3 and □4 are combined and become the sub-content under heading C. However, the groups □1 and □2 and the groups □3 and □4 are not combined under the relationship of q = 3 (see Figure 2).

Figure 1 and Figure 2 are the combination diagrams of the given value. If the value is smaller, the related keywords involved in each basic content will increase. Therefore, we can say that according to the chapters, subsections, and arrangement of subsections of the digital tourism online keyword database, a corresponding (□, q) pair of values is recorded as (□□, □□) a group. The pairing value also corresponds to a chapter arrangement method.

For a certain tourism theme, the higher the level of information analysis, the more detailed the content of the knowledge describing the tourism theme. The detailed agreement can be divided into two types: breadth and depth. The smaller the value of □, the more basic the content (see the definition of □□ (□□, □□)), implying a more detailed breadth is. The depth of detail belongs to the linear ordering problem of each sub-content under the same combined group (same chapter). An example is as follows: In Figure 1, although □2 and □4 do not have common keywords, □1, □2, □3, and □4 are combined together. This is because simultaneously □2 and 4 have more than two keywords overlapping with □3 and they are indirectly combined; therefore, we can say that □2 and □3 are a (direct) combination of distance 1, and □2 and □4 are an indirect combination of distance □2.The content relationship of such a combined group (i.e., the same travel chapter) can be measured by the total weighted distance of the relationship between one basic travel content and other basic content. According to the strength of the weighted distance indicator, the sub-contents are arranged in a straight line from small to large. The purpose of the straight-line distance arrangement is as follows: if the tourist does not want to read all the content of a certain travel chapter but only analyze it in terms of certain information, they may read some of its tourism content at a

standard level; then, the first half of the above-mentioned linear arrangement content can be chosen to be read by tourists. That is, the s to which any of the first half of this straight line belongs corresponds to the knowledge of a certain information analysis level (this analysis level can be marked next to the letters A, B, and C in Figure 2). The higher the level of information analysis, the greater the length of the corresponding straight-line value in the first half.

The choice of the value of □ (□□, □□) ≥ □ in the foregoing inequality, and the choice of a certain point on the linear arrangement of the same combination content can become the options for tourists to reflect the level of information analysis required by them. Theoretically, according to the above concepts, a retrieval software for the level analysis of tourist demand information can be designed. The above is the arrangement of the chapters in the first edition of a digital tourism online keyword database based on the overlap of various content keyword groups provided by various tourism companies.

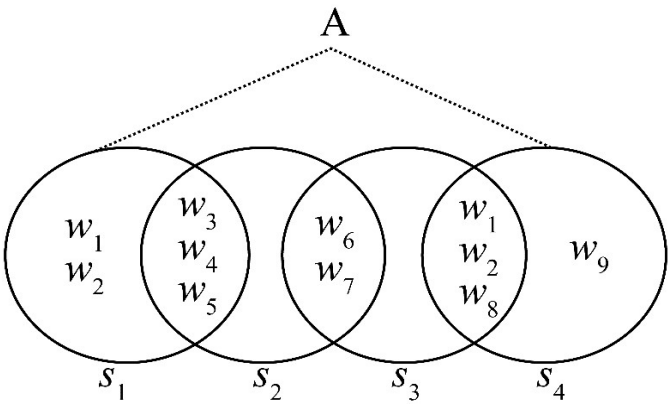

Figure 1. Combination of the basic content of the keyword group through the tourism industry (in the case of q = 2)
Source: Obtained and integrated by researchers in this study.

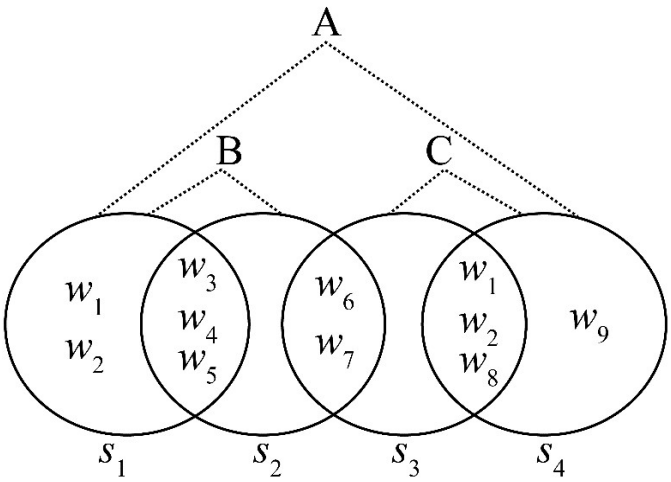

Figure 2. Combination of the basic content si of the keyword group of the tourism industry network (in the case of q = 2 and q = 3)
Source: Obtained and integrated by researchers in this study.

As a keyword may be related to several basic contents, a basic content may also be related to several keywords. Therefore, the above-mentioned concept of combining basic content through related keywords can be similarly applied to the basic content through keywords (as shown in Figure 3).This means that under a certain threshold value, through the calculation methods shown in Figure 1, 2, and 3, the combination of basic content can be displayed for different q values, and the combination of keywords in the online tourism industry. That is, the "combination" here is two-fold: one is the combination of basic content, and the other is the combination of keywords.

From this two-way combination, a concept called "threshold dual combination" is extended as follows. Assume that given the threshold value, the basic tourism content group {s1, s3, s5} is a combination of q1, and the tourism industry keyword group {w2, w3, w7, w8} is a combination of q2; The keyword set of the content group {s1, s3, s5} is "relevant" (if the key value □ (□□, □□) ≥ □, then wi is related to sj, otherwise it is irrelevant), which is just {□2, □3 , □7, □8}. Simultaneously, the basic content set related to one of wi of {□2, □3, □7, □8} is just {s1, s3, s5}, then, {s1, s3, s5} and the keyword group {□2, □3, □7, □8} (q1, q2) includes dual combination groups of □ values.

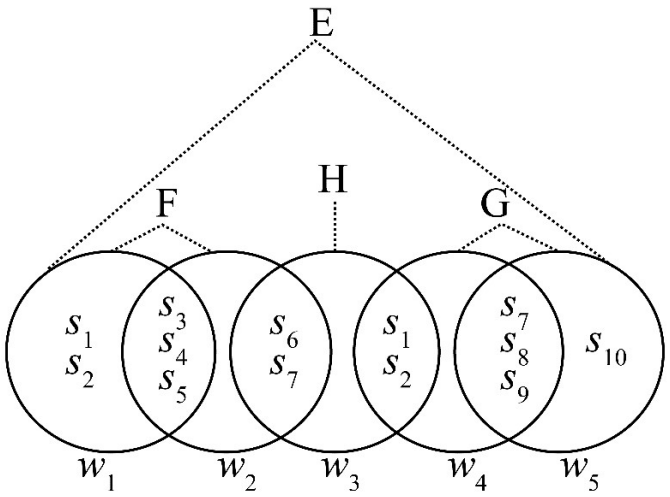

Figure 3. Tourism network keyword combination diagram
(given a certain threshold, q = 2 and q =3).
Source: Obtained and integrated by researchers in this study.

As the number of basic content m is different from the number of keywords n, we can consider the relative size of m and n to calculate the q1 combination between the basic content groups, which is equivalent to the q2 (m, n) combination of the keyword groups, where q2 (m, n) is obtained by adjusting the original q2 by the relative sizes of m and n. After adjusting the relative sizes of q1 and q2 in this normalization, then q2 in the aforementioned "value of (q1, q2) dual combination group" is: if q2 = q2 (m, n), then {s1 , S3, s5} are self-reflecting q1 binding groups below the threshold, also known as {□2, □3, □7, □8} are self-reflecting q2 binding groups below the threshold.

In general, the smaller the value of □, the looser the correlation between wi and sj, and therefore the more wi there is for the same sj; for the same wi, the sj related to it is higher. This means that the smaller the selected threshold, the weaker the binding force between different members of the q-binding group for the same q-value.

According to the above analysis and discussion, it is learned that through various (□, q) choices, the various self- reflections combined with the basic content

group are obtained. In practice, there may not be a fully self-reflective combination group, but there will always be a near-self-reflective combination group. These are the editors and decision makers of the tourism industry online keyword library. They can consider dividing the content into separate chapters, sections, or subsections. They may consider publishing a continuous tourism industry online keyword library series. For example, the content of each different episode can be divided in the first, second, third sections, and so on of the travel guide series.

Conclusion and Recommendations

The digital tourism industry keyword editing designer (mechanism designer) uses the traces of the tourism industry's online keyword library to show the relevance of various chapters and subsections to automatically adjust the demand of the tourist in real time. The order of arrangement is decided after revision and reprint, based on how to automatically correct the chapter layout before the first edition, according to the keywords of the basic content provided by the original product supplier (tourism businesses), and how to automatically track the original chapter layout according to the trajectories left by the tourism buyers after reading the previous edition. These are the two main editing mechanisms for automatic editing of an article. The first is called the tourism industry-oriented editing mechanism, which has been shown in the previous section, while the second called tourist demand-oriented automatic editing mechanism, introduced below.

According to the above-mentioned idea of building an automatic editing mechanism in this article, it is understood that the tourism industry may come from different social classes and may have different educational backgrounds. Therefore, we are unable to determine the consistency of their expressions of knowledge objects, terms, tone, and other expressions; we may also not be sure of the different chapters of the digital tourism online keyword library degree.

This means that we need to develop a computer program that can automatically record the time trajectory of each tourist who browses each chapter as well as the order of clicks on each chapter. With its recording and statistical measurement function, this program can be

used as a reference for the rearrangement of each chapter when the digital tourism industry keyword library is republished. The more times a chapter is clicked, the longer the chapter is. The order of the chapters should be ranked first. If two different chapters are clicked back and forth by the same reader, it means that the contents of the two chapters are highly related and should be arranged together. It may even melt into the same chapter, but due to the content of different chapters, the issue of preliminary knowledge attributes can be raised (see the above section). Therefore, the adjustment of the arrangement of the chapters cannot be completely adjusted according to the reading trajectory of the tourist. However, the demand for reading various chapters of a completed digital tourism online keyword database by tourists is a piece of information that cannot be ignored. The distance between the chapters of a completed digital tourism online keyword database can be defined as follows: as shown in Figure 4, the distance between Content A and Content B is 1, the distance between Content A and Content C is 2, and the distance between Content C and Content D is 3.

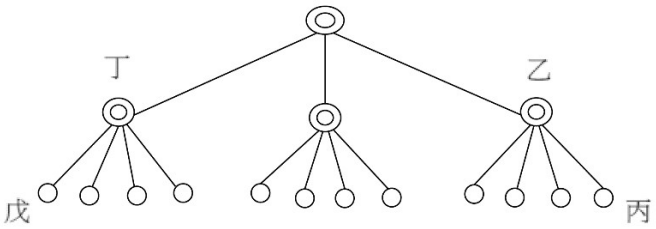

Figure 4. The Conceptual Framework of Chapter and Keyword Editing
Source: Obtained and integrated by researchers in this study.

When tourists use a completed digital book, the travel trajectory of each tourist can be used to calculate the trajectory distance of different chapters. Basically, the distance structure organized according to the chapters that are effective for tourist demand should be consistent with the distance formed by calculating the reading trajectory of tourist demand. In addition, the fewer the number of clicks by the traveler, the shorter the browsing time will be in the back, and even if it is lower than the set number of clicks and browsing time, the

content will be automatically removed.

Surely, we cannot let users or learners who want to travel, aimlessly find the chapter content that suits them, so there must be a feedback mechanism. For example, the first time one clicks and browses, we can let them see how many chapter contents and time (assuming 5 minutes or 5 pages) are there. If the feedback is satisfactory and recommended, then they can continue to open the full text. Such an automatic editing and publishing mechanism for the online keyword database of the tourism industry can, on the one hand, allow users or learners of tourism to quickly find the latest, most popular, and even the most relevant learning content. On the other hand, through feedback from users or learners of tourism demand, the original supplier (tourism business) of the product that provides the content can continuously update or improve the content.

References

Chang, H. H. (2016). The study of tourist's preferences towards wayfinding directional descriptors. *Journal of Tourism and Leisure Studies*, *22*(2), 123–149. doi: 10.6267/JTLS.2016.22(2)1

Dale, R., Geldof, S., & Prost, J. P. (2002). Generating more natural route descriptions. In *Proceedings of the 2002 Australasian natural language processing workshop* (pp. 41–48).

Denis, M., Pazzaglia, F., Cornoldi, C., & Bertolo, L. (1999). Spatial discourse and navigation: An analysis of route directions in the city of Venice. *Applied Cognitive Psychology, 13*(2): 145–174. doi:10.1002/(SICI)1099-0720(199904)13:2<145::AID-ACP550>3.0.CO;2-4

Golding, J. M., Graesser, A. C., & Hauselt, J. (1996). The process of answering direction-giving questions when someone is lost on a university campus: The role of pragmatics. *Applied Cognitive Psychology, 10*(1), 23–39. doi:10.1002/(SICI)1099-0720(199602)10:1<23::AID-ACP357>3.3.CO;2-8

Hund, A. M., & Minarik, J. L. (2006). Getting from here to there: Spatial anxiety, wayfinding strategies, direction type, and wayfinding efficiency. *Spatial Cognition and Computation, 6*(3), 179–201. doi: 10.1207/s15427633scc0603_1

Hund, A. M., & Padgitt, A. J. (2010). Direction giving and following in the service of wayfinding in a complex indoor environment. *Journal of Environmental Psychology, 30*(4), 553–564. doi: 10.1016/j.jenvp.2010.01.002

Hund, A. M., Schmettow, M., & Noordzij, M. L. (2012). The impact of culture and recipient perspective on direction giving in the service of wayfinding. *Journal of Environmental Psychology, 32*(4), 327–336. doi: 10.1016/j.jenvp.2012.05.007

Lin, T. Y., Chang, W. H., & Chen, M. S. (2020). Construction of The Optimal Sales Model for The Digital Navigation Systems of Online Tourism Databases. The *International Journal of Organizational Innovation, 13*(1), 165–177.

Mark, D. M., & Gould, M. D. (1995). Wayfinding directions as discourse: A comparison of verbal directions in English and Spanish. *Journal of Cross-Cultural and Interlanguage Communication, 11*(3), 267–292. doi: 10.1515/mult.1992.11.3.267

Saucier, D. M., Green, S. M., Leason, J., MacFadden, A., Bell, S., & Elias, L. J. (2002). Are sex differences in navigation caused by sexually dimorphic strategies or by differences in the ability to use the strategies? *Behavioral Neuroscience, 116*(3), 403–410. doi:10.1037//0735-7044.116.3.403

Vanetti, E. J., & Allen, G. L. (1988). Communicating environmental knowledge: The impact of verbal and spatial abilities on the production and comprehension of route directions. *Environment and Behavior, 20*(6), 667–682. doi: 10.1177/0013916588206001

Wright, P., Lickorish, A., Hull, A., & Ummelen, N. (1995). Graphics in written directions: Appreciated by readers but not by writers. *Applied*

Cognitive Psychology, 9(1), 41–59.
doi: 10.1002/acp.2350090104

Yen, R. Y. (2014). A study of the usability of official tourism websites. *Bulletin of Library and Information Science*, *6*(1), 24–50.