ICIC Express Letters
Volume **16**, Number **4**, April **2022**

ICIC International ©2022 ISSN 1881-803X
pp. **409–417**

A HISTORICAL TOURISM RECOMMENDATION SYSTEM FOR THE ELDERLY TOURIST USING NATURAL LANGUAGE PROCESSING AND THE ONTOLOGY TECHNIQUE

Anongporn Salaiwarakul

Department of Computer Science and Information Technology
Faculty of Science
Naresuan University
99 Moo 9, Muang, Phitsanuloke 65000, Thailand
anongporns@nu.ac.th

Received August 2021; accepted October 2021

Abstract. *This paper presents a proposed methodology and framework of a recommendation system for elderly tourists which enables semantic inferencing by the inclusion of an ontology, together with natural language processing for analyzing user queries to increase the efficiency and accuracy of the recommendation system. Given that the target user demographic are Thai tourists, the natural language utilized is Thai. The physical limitations experienced by elderly tourists are specified as significant factors when suggesting the choices of tourist destinations presented and making recommendations appropriate to that demographic. To enhance the efficiency of the system, a new "facility" class is defined in the Elderly Historical Tourism Ontology which provides information regarding the availability of elder-related facilities which can satisfy the physical limitations of the elderly tourist. A 95% precision of retrieved information was achieved in the proposed system.*

Keywords: Elderly tourism recommendation system, Elderly tourist limitations, Historical tourism ontology, Natural language processing

1. **Introduction.** A recommendation system is an approach to suggesting and proposing items of interest to a user, based on that user's profile or preferences [1-3]. Such systems have been implemented in various domains, such as e-commerce [4,5], education [6,7], and tourism [8-10]. Some recommendation systems apply techniques such as content-based filtering which recommend items based on users' profiles or behaviors or techniques of collaborative filtering which predict desired items for the user based on similarity with other users, or a combination of techniques to prefer suggestions appropriate to the user.

An enormous volume of data is available on the Internet, making manual processing of that data a challenge for users, particularly elderly tourists who are the subject of this discussion. Some published research are related to recommendation systems that provide tourism information for elderly tourists and which address the requirements and limitations of elderly tourists that are often different to those of the general tourist [11-13]. The limited mobility is often demonstrated by elderly tourists and other special requirements for facilities, such as elder-specific restrooms, rest areas located at reasonable distances apart or good accessibility by transportation are key factors that influence the choice of preferred tourism attractions for the elderly tourist in addition to the usual tourism activities, tourism categories, and geographical locations which are implemented in general tourism recommendation systems. Some attractions even if they are matched to requirements such as type of attractions, activities, or location, do not provide suitable facilities that support the elderly tourist's limitations. As well, importantly, elderly tourists are

DOI: 10.24507/icicel.16.04.409

unlikely to be able to specify appropriate keywords that would elicit the information that would be useful to them, if the query interface was keyword driven.

Not addressing these factors would result in misleading or inappropriate suggestions and recommendations for the elderly tourist. To overcome these problems, and provide a useful and useable tourism recommendation system, this research included the specification and development of a system that applies an ontology-based information semantic search approach, together with a natural language-based query facility that overcomes the problem of keyword-based querying, thereby simplifying the manner in which the elderly tourist can retrieve information from the system. The developed system implemented a historical tourism ontology, termed the Elderly Historical Tourism Ontology, as the domain tourism information for the system.

2. **Background.** Recommender systems apply machine learning algorithms in order to recommend products or services associated with the user's profile and/or relative environment variables. In order to recommend and predict items based on the user's needs, the sequential logs of the profiles of users with similar interests are applied to recommending the next best item particular to that user, which is referred to as being a memory-based approach in [14]. However, having sequential logs of a sufficient number of similar users would imply a very large database, which would be needed to ensure sufficient prediction accuracy. Current recommendation systems that applied a memory-based approach often have limited information about the current user, and a limited number of similar users and item collections, meaning that they are unable to provide the good level of accuracy that is necessary in current e-commerce systems. From this point of view, the study discussed in [15] suggested a hybrid approach that combines this memory-based approach with a model-based approach to overcome those problems. The result of their experiments with a hybrid approach showed that accuracy is improved when compared to the model-based approach, with significantly fewer database requirements than the memory-based approach.

To solve the problem of a large volume of data in the recommender system, some researchers used social networks as the data collection source. In [2], an ontology was constructed from a data collection of user interactions in Twitter accounts related to an academic conference. Their recommendation system recommended appropriate academic conferences based on relationship interactions in the social network and environment. Other entities, such as the location of each conference, and the location of users, were added to promote greater accuracy of recommendations.

The use of an ontology to enhance the accuracy of semantic searching and to reduce the searching of similarity items in a database, was proposed in [17]. The ontology was constructed from user profiles in the recommender system together with a hybrid approach that used content-based and collaborative-based filtering. The method in their study used collaborative filtering to collect interesting research papers from an ontology-based research paper database and then applied content-based filtering to suggesting the topics of interest that matched the user's profile.

Several natural language processing techniques have been applied to promoting efficiency in a recommendation system, overcoming the difficulties often experienced in a keyword-based query interface [7,18]. In [18], Wordnet was implemented, together with keyword-based features, to improve the performance of the system. Several researchers have shown the importance of understanding the concept of subjective ratings rather than presumably objective keywords. The system should be able to comprehend the meaning of the subjective ratings provided by the user in order to understand the implied negative or positive sentiments. Wordnet was applied as the natural language processing approach in a recommendation system by [19,20], together with the implementation of an ontology for semantic information retrieval.

There are several research directions related to tourism recommendation systems, to the best of our knowledge, and it is generally accepted that the elderly tourist limitations are important considerations when these tourists choose their preferred attractions. However, there are very few published papers that consider these factors as variables, thereby limiting their usefulness for this demographic. This paucity of research also extends to the application of natural language processing in recommendation systems. Both of these matters should be explored in regard to tourism attractions to enhance the efficiency and accuracy of tourism recommendation systems for the elderly tourist. The next section illustrates the research framework that solves this problem.

3. **Research Framework.** The proposed research framework contains two major parts: the first being the natural language processing approach, and the second being the development of a historical tourism ontology appropriate for the elderly tourist: historical tourism being a particular category of tourism. The natural language processing approach enhances the power of the recommendation system in understanding the queries entered by the interested tourist, while providing a simpler and more convenient method for enabling the elderly tourist in acquiring useful information from the recommendation system. The ease of use of the system is an essential element in any system. As the intended user base for this system is Thai, this research considered only the Thai language for queries. This did introduce a level of difficulty in the natural language processing, given the lack of significant research into, or application of, natural language processing of the Thai language.

Another component in the proposed framework is **the Elderly Historical Tourism Ontology**, containing historical tourism information appropriate for the elderly tourist, which is composed of classes related to historical tourism attractions likely to be of interest to the elderly tourist, which provide the necessary information for them. This knowledge base filtering is an efficient method for retrieving information semantically in a recommendation system. The ontology is a methodology to support the structure of information and relevant information using inferencing properties which in turn enable the recommendation system to understand the context of the information provided by users instead of relying on keyword matching. The proposed research framework is shown in Figure 1.

Figure 1 illustrates how the user provides a natural language query, specifically in Thai language, to the system. The *natural language processing approach* is applied in the first stage in order to understand the meaning and specific name entity that is supplied to the system by users. The extracted name entity and information from the user's query is further processed to acquire information from the database. The Elderly Historical Tourism Ontology is used inferentially to acquire the information semantically to be able to recommend tourism attractions that are suitable to elderly tourists. Detailed descriptions of approaches that are applied in the recommendation system are illustrated in the following section.

Natural Language Processing Approach. For the purpose of understanding queries from users in recommendation system, natural language processing is used in the system. The proposed recommendation system allows the elderly tourist to acquire tourism information by entering a query in natural Thai language. The natural language processing algorithm then starts by analyzing the sentence from the user to identify words from the query's sentences and then the antecedent extraction process is then applied to defining the specific attraction and user's requirements.

Sentence Analysis. In order to analyze the user's query, this research applies the POS-based trigram model which combines the advantages of both the trigram model and the POS-based model. The methodology combines the rule-based model and the n-gram model in order to predict which part of speech is being used in the sentence and to remove the words that are unnecessary for understanding the meaning of the sentence. As its

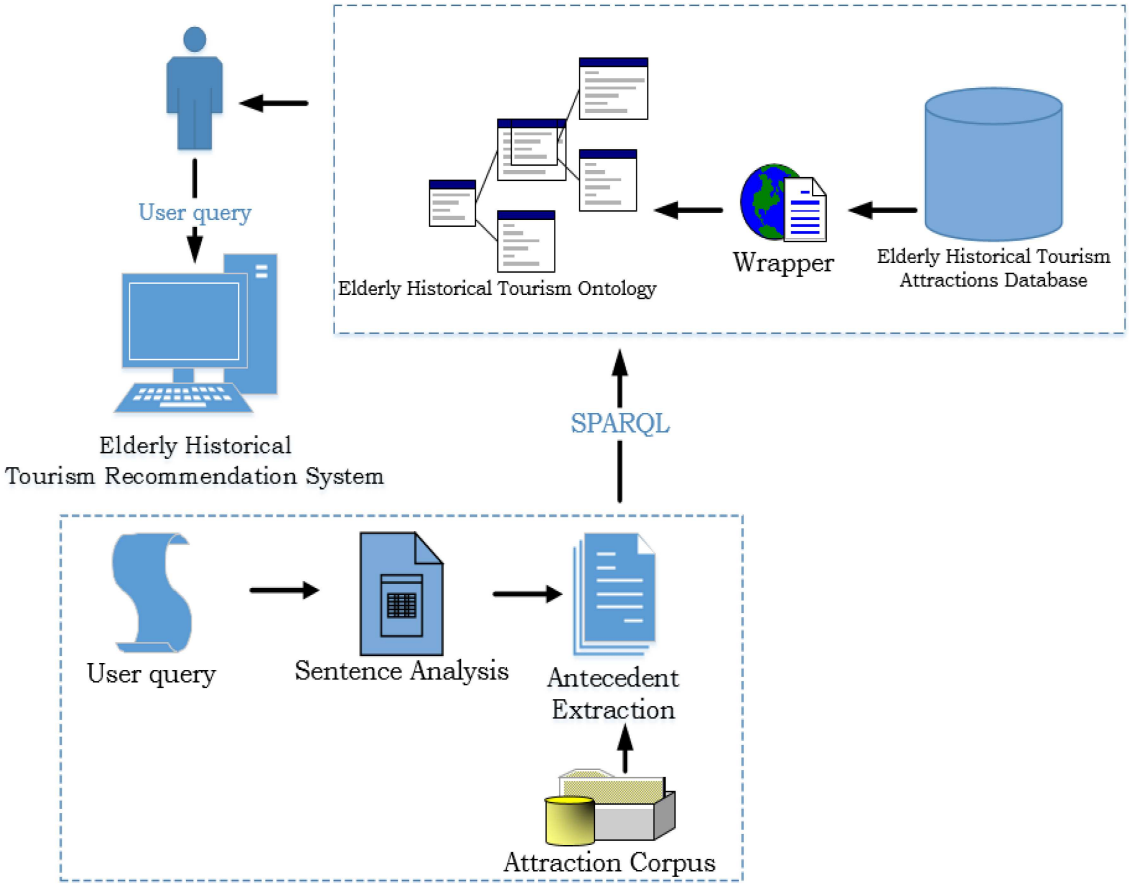

FIGURE 1. The elderly historical tourism recommendation system

name implies, the rule-based model uses a set of rules that identify the various parts of speech such as noun, verb, adjective or preposition. The n-gram model is a probability-based language model that predicts the next items in a sequence. Trigrams allow the prediction of, and probability of, a word based on the preceding 2 words in combination. Among bigrams, trigrams and quadrigrams, the trigram model was shown to produce the highest accuracy. Applying this algorithm does not require a large corpus to be stored in a database. Figure 2 shows the POS-based trigram model algorithm which is used in the proposed recommendation system.

Antecedent Extraction. The output from the sentence analysis algorithm shown in Figure 2 still would not identify the specific words, such as attraction name or location. The antecedent extraction algorithm is further applied in order to specify attraction names and locations. Figure 3 shows the antecedent extraction algorithm that is implemented in the recommendation system.

The output from these two algorithms gives an attraction name or other specific words such as physical limitations likely for an elderly tourist that are significant factors when suggesting tourism attractions that offer different facilities. The extracted keywords from the proposed natural language processing approach allow the recommendation system to understand the user's query and then the ontology is applied for gathering information from the knowledge base semantically.

Elderly Historical Tourism Ontology. The ontology is constructed with information related to **historical tourism attractions** which enables sufficient information for elderly tourists to be retrieved. Figure 4 illustrates the Elderly Historical Tourism Ontology that is implemented as data inferences for retrieving the tourism information semantically.

```
1.  Data = word
2.  Result = Valid object concepts(Split word into Trigram)
3.  For each word
4.          While(Initialize Concepts to Null)
5.          POS base Trigram;
6.                  if adj noun Then
7.                          add to Concepts: adj+noun
8.                  else if noun noun Then
9.                          add to Concepts: noun+noun
10.                 else if Stopword noun Then
11.                         add to Concepts: noun
12.                 else if adj Stopword Then
13.                         Continue
14.                 else if Stopword adj Then
15.                         Continue
16.                 else
17.                         add to concepts: entire Trigram
18.                 end
19.         end
20. end
```

FIGURE 2. POS-based trigram model algorithm

```
1.  Data = array(word,position)
2.  For each Data
3.          If Data_N = Antecedent Then
4.                  While(true)
5.                      Data_N = Data_N+1
6.                      Data_N(matching search) attraction name
7.                      If matching result = true
8.                          Return Data_N
9.                          break
10.                 end
11.         end
12. end
```

FIGURE 3. Antecedent extraction algorithm

Additionally, facilities of attractions are highlighted in this ontology since it is a significant factor that will be mapped with the extracted information from user's queries. This research treats the *facility* class differently from other tourism ontologies that also provide this type of class. The *facility* class is emphasized to match with the elderly tourist's physical limitations. This increases the efficiency in delivering recommendations. This class highlights the facilities that the elderly tourist requires and which are therefore important key factors that the elderly tourist considers when selecting their preferred attractions.

The Elderly Historical Tourism Inference. The historical tourism attraction ontology created for this purpose applies three relationships in order to infer appropriate data from the database. The inferences that are implemented in this proposed ontology

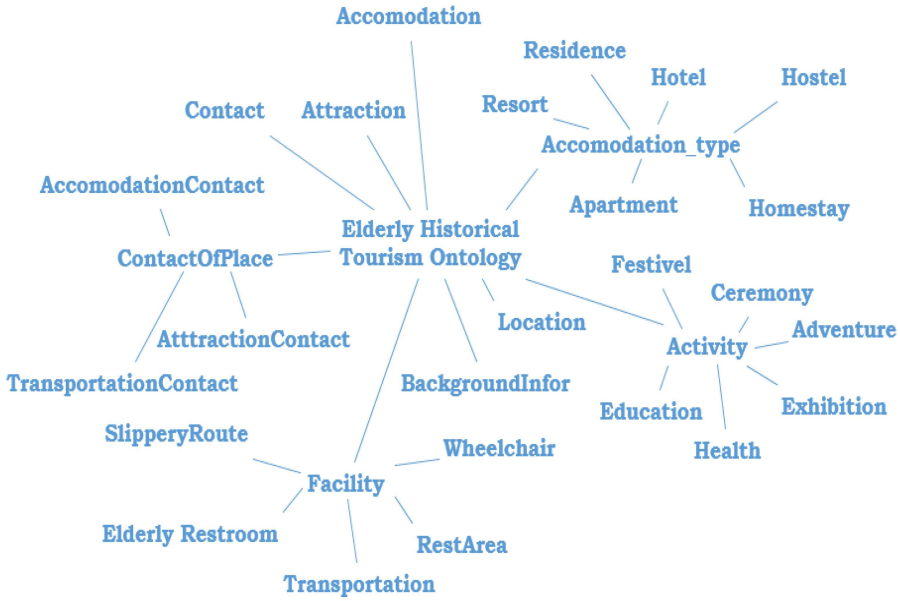

FIGURE 4. The Elderly Historical Tourism Ontology

enable automatic inferencing of new arrivals of data instances in the ontology. The three inference properties are as the following:

InverseOf Property. This property is applied to promoting the relationship for classes in the ontology that can swap between Domain and Range. Figure 5 shows an example of the InverseOf property that is used in the design Elderly Historical Tourism Ontology. It can be illustrated that the *Attraction* class has a *hasActivity* relationship with the *Activity* class whereas the *Activity* class has a *CanbeDoneAt* relationship with the *Attraction* class. This property allows the domain and range to have an inverse relationship. The ontology has the instance *Wanatara* (the name of an attraction) which has an activity of *healthHealing*. Therefore, the InverseOf property allows the automatic inference that *healthHealing* is an activity that can be done at *Wanatara*.

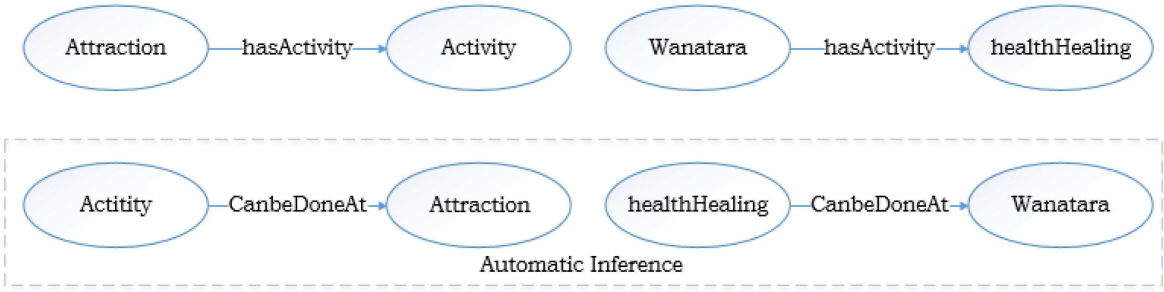

FIGURE 5. InverseOf property inference for the Elderly Historical Tourism Ontology

Symmetric Property. This type of relationship is applied for inferring that the subject has the same meaning or symmetric meaning with the object. Figure 6 illustrates that *Wat Nang Phaya* has a *nearBy* relationship with *Wat Rat Ratchaburana*; therefore, the system can automatically infer that *Wat Rat Ratchaburana* has a *nearBy* relationship with *Wat Nang Phaya*. According to this property, if there is a new arrival of a data instance that has a *nearBy* relationship with any of the stored values, the recently added value will automatically have the *nearby* relationship with the previously stored value. Applying this property promotes better inferencing when gathering data from the tourism ontology.

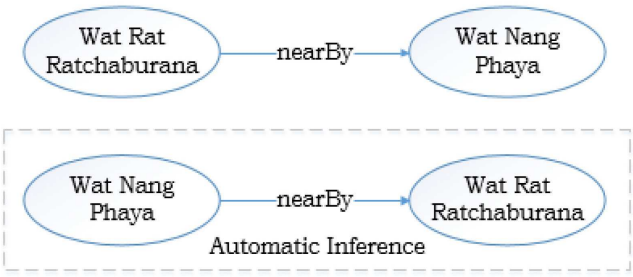

FIGURE 6. Symmetric property inference for the Elderly Historical Tourism Ontology

Transitive Property. This property is defined in the tourism ontology so that if a class or an instance has transitivity, then it can be inferred that where a transitive class A has a relationship with a class B, and class B has a relationship with class C, then class A is transitively considered to have a relationship with class C.

An example of this property in the ontology is shown in Figure 7 which shows that *Phu Hin Rong Kla National Park* has a relationship *Location: relatedTo* with *Lan Hin Pum*, and *Lan Hin Pum* has a relationship *Location: relatedTo* with *Lan Hin Taek*. This transitive property allows *Phu Hin Rong Kla* to have a relationship *Location: relatedTo* with *Lan Hin Taek*.

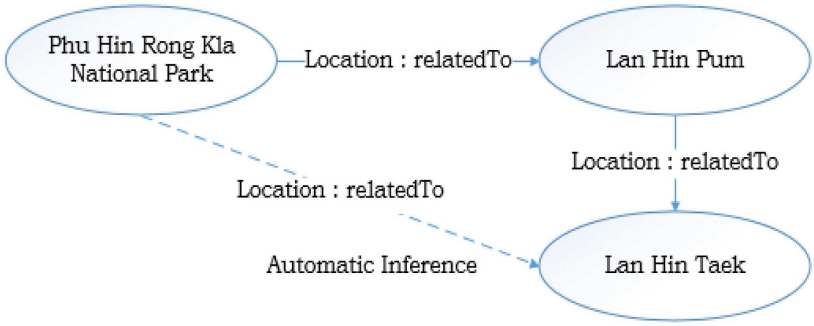

FIGURE 7. Transitive property inference for the Elderly Historical Tourism Ontology

The three properties as described allow the recommendation system to have a better understanding of the semantics inherent in the knowledge base of tourism attractions, allowing the recommendation system to make better recommendations. The recommendation choices are also enhanced by including the *facility* class in the ontology related to the tourism activity.

4. **Experimental Results and Discussion.** The precision of the proposed recommendation system is calculated using (1):

$$precision = tp/(tp + fp) \tag{1}$$

where tp is the total number of retrieved instances that are relevant to the user query and fp is the total number of retrieved instances that are not relevant to the user query. The queries were collected from three popular web sites that are widely used in Thailand for searching for tourism information. A selection of 100 queries that were made in natural language was randomly selected from those websites and used as test sentences for the recommendation system.

The demonstrated precision of the system in recommending attractions that match the user's limitation was 95% which illustrates the effectiveness of both natural language query input and the inclusion of users' limitations as input information. This level of

precision surpasses any of the other recommendation systems that were identified in the literature review.

5. **Conclusions and Future Work.** A historical tourism recommendation system relevant to elderly tourists, that applies a historical tourism ontology, is presented. The system contains classes relevant to historical tourism information as well as information that can be applied to matching with elderly tourists' limitations, especially the *facility* class. The ontology is embedded in the system to promote semantic inferences of tourism information held in the database. As well, the natural language processing approach that is implemented in the system, enables ease of use of the system by elderly tourists, by allowing queries to be entered in natural Thai language. The POS-based Trigram is used as a technique for sentence analysis. The antecedent extraction is applied for specifying specific attraction names and identifying information related to the probable physical limitations of the elderly tourist. This is information not previously seen in such systems. This additional, relevant, information improves the efficiency of providing suggestions for the best suitable tourism attractions for the elderly tourist who has physical limitations. The precision of the retrieved information from the user's query was 95%, achieved in the tests of the proposed recommendation system. However, the attraction corpus which was applied in this research was limited. Further studies should consider a more comprehensive real-world corpus and more query collections to improve the efficiency and accuracy of the system. Future work could include separating physical limitations as another elders-limitations ontology and merging the Elderly Historical Tourism Ontology and elders-limitations ontology to produce more suitable recommendations appropriate to this class of tourist.

REFERENCES

[1] F. O. Isinkaye, Y. O. Folajimi and B. A. Ojokoh, Recommendation systems: Principles, methods and evaluation, *Egyptian Informatics Journal*, vol.16, no.3, pp.261-273, 2015.

[2] R. Mu, A survey of recommender systems based on deep learning, *IEEE Access*, vol.6, pp.69009-69022, 2018.

[3] P. M. Alamdari, N. J. Navimipour, M. Hosseinzadeh, A. A. Safaei and A. Darwesh, A systematic study on the recommender systems in the e-commerce, *IEEE Access*, vol.8, pp.115694-115716, 2020.

[4] X. Wang and C. Wang, Recommendation system of e-commerce based on improved collaborative filtering algorithm, *Proc. of the 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pp.332-335, 2017.

[5] Z. Zhang and S. Qian, The research of e-commerce recommendation system based on collaborative filtering technology, *Proc. of Advances in Computer Science and Information Engineering*, Berlin, Heidelberg, pp.507-512, 2012.

[6] J. Joy, N. S Raj and V. G. Renumol, Ontology-based e-learning content recommender system for addressing the pure cold-start problem, *J. Data and Information Quality*, vol.13, no.3, pp.16:1-16:27, 2021.

[7] S. Renuka, G. S. S. R. Kiran and P. Rohit, An unsupervised content-based article recommendation system using natural language processing, in *Data Intelligence and Cognitive Informatics. Algorithms for Intelligent Systems*, I. J. Jacob, S. K. Shanmugam, S. Piramuthu and P. Falkowski-Gilski (eds.), Singapore, Springer, 2021.

[8] Z. Bahramian and R. A. Abbaspour, An ontology-based tourism recommender system based on spreading activation model, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS)*, vol.XL15, pp.83-90, 2015.

[9] M. E. B. H. Kbaier, H. Masri and S. Krichen, A personalized hybrid tourism recommender system, *Proc. of 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, Hammamet, Tunisia, pp.244-250, 2017.

[10] O. Alnogaithan, S. Algazlan, A. Aljuraiban and A. A. Shargabi, Tourism recommendation system based on user reviews, *Proc. of 2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, Sakhier, Bahrain, pp.1-5, 2019.

[11] A. Kazeminia, G. Del Chiappa and J. Jafari, Seniors' travel constraints and their coping strategies, *Journal of Travel Research*, vol.54, no.1, pp.80-93, 2015.

[12] P. A. S. Utami and N. K. G. Prapti, Recreational activities and travel choices of the elderly, *Indonesian Journal of Global Health Research*, vol.2, no.2, pp.165-172, 2020.
[13] D. Huber, S. Milne and K. Hyde, Constraints and facilitators for senior tourism, *Tourism Management Perspectives*, vol.27, pp.55-67, 2018.
[14] M. Quadrana, P. Cremonesi and D. Jannach, Sequence-aware recommender systems, *ACM Computing Surveys*, vol.51, no.4, pp.66:1-66:36, 2018.
[15] K. Bagherifard, M. Rahmani, M. Nilashi and V. Rafe, Performance improvement for recommender systems using ontology, *Telematics and Informatics*, vol.34, no.8, pp.1772-1792, 2017.
[16] M. Arafeh, P. Ceravolo, A. Mourad, E. Damiani and E. Bellini, Ontology based recommender system using social network data, *Future Gener. Comput. Syst.*, vol.115, pp.769-779, 2021.
[17] S. Middleton, D. De Roure and N. Shadbolt, Ontology-based recommender systems, *Handbook on Ontologies*, vol.32, pp.779-796, 2009.
[18] S. Kamath and M. Kanakaraj, Natural language processing-based e-news recommender system using information extraction and domain clustering, *International Journal of Image Mining*, vol.1, pp.112-125, 2015.
[19] N. Yanes, S. B. Sassi and H. Ghezala, Ontology-based recommender system for cots components, *Journal of Systems and Software*, vol.132, pp.283-297, 2017.
[20] Z. Abbasi-Moud, H. Vahdat-Nejad and J. Sadri, Tourism recommendation system based on semantic clustering and sentiment analysis, *Expert Systems with Applications*, vol.167, DOI: 10.1016/j.eswa.2020.114324, 2021.