The task-oriented Smart Tourism Chatbot Service

Hoon-chul Kang[1], Myeong-Cheol Jwa[2], Jeong-Woo Jwa[3].

[1]*Jeju Campus of Korea Polytechnic, Jeju, Korea*

[2]*Korea University of Technology and Education, Cheonan-si, Korea*

[3]*Jeju National University, Jeju, Korea; E-mail: lcr02@jejunu.ac.kr*

Abstracts: The smart tourism service provides tourists with a travel planner service before the trip and a tour guide service during the trip. Smart tourism chatbot service can easily and conveniently provide smart tourism services to tourists along with smart tourism apps. In this paper, we develop the smart tourism platform and propose the task-oriented smart tourism chatbot service that efficiently provides tourism information provided by smart tourism apps to users. The smart tourism platform consists of the smart tourism chatbot and smart tourism information systems. The smart tourism chatbot system identifies the intention of the user's question and searches for the tourism information ID in the tourism information knowledgebase. The smart tourism information system provides tourism information of the smart tourism app to the user using the searched tourism information ID. The proposed smart tourism chatbot system applies the tourism information NER (Named Entity Recognition) model to accurately understand the intention of the user's question to the existing DST (Dialogue State Tracking) model-based chatbot system. We create tourist information named entities of the NER model based on the values of the domain and slot in the DST model defined in the 4W1H method. The NER model determines the domain, slot, and value to identify the intention of the user's question, and the DST model manages the dialogue state and retrieves the tourism information ID from the tourism information knowledgebase.

Keywords: Task-Oriented Chatbot System, Smart Tourism Chatbot System, Tourism Information DST and NER Datasets, Pre-Trained Language Model, Smart Tourism App.

1. INTRODUCTION

Smart tourism services provide travel planner and tour guide services to tourists using Internet of Things (IoT), communication infrastructure such as WiFi7 and 6G, Big Data, Artificial Intelligence (AI), Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR) technologies [1]-[4]. We are developing smart tourism apps [5], TTS server-based tour guide services [6], smart tour chatbot services [7], and Instagram and YouTube services to provide smart tourism services to individual tourists. Users can use the smart tourism app to select recommended travel products and tour information and create personalized travel itineraries. The AI-based chatbot service can provide travel planner service and tour guide service to tourists easily and conveniently using smart tourism chatbot app.

We are developing tourism information DST (Dialogue State Tracking), NER (Named Entity Recognition), QA (Question Answering) datasets [8]-[10] for transfer learning of pre-learning language models (PLMs) [11]-[13] to develop a task-oriented smart tourism chatbot service. The tourism information DST dataset is created in JSON format of the WOS (Wizard-of-Seoul) dataset [14] with the domain, slot, and value defined by the 4W1H method. The tourism information NER dataset is created using a tourist information NER dictionary for BIO (begin, inside, outside) tagging. We create a tourist information NER dictionary to correct various expressions and typos in tourist destination names. The tourism information QA dataset is created in the JSON format of KLUE (Korean language understanding evaluation) MRC (machine reading comprehension) dataset. We developed the DST model-based tourism information chatbot system [15] by applying the tourism information DST algorithm to the existing rule-based chatbot system. The DST model-based chatbot system defined the domain, slot, and value with the 4W1H method and modified the attribute data of the Neo4J graph DB's tourism information knowledgebase to understand the user’s question more accurately. In this paper, we develop the task-oriented smart tourism chatbot system using the NER model. The NER model-based tourism information chatbot system uses the values of the domain and slot in the DST model as tourist information NER dictionaries to understand user questions accurately.

2. RELATED WORKS

Smart tourism service provides tourists with tourist information, creation of my tourist itinerary, and context aware tour guide service. We are developing the smart tourism platform and the smart tourism app to provide smart tourism services to individual tourists. The smart tourism platform consists of the smart tourism chatbot system that provides chatbot services and the smart tourism information system that provides smart tourism app services. In this chapter, we describe the smart tourism mobile app, the DST model-based tourism information chatbot system, and the tourism information NER dataset.

2.1 Smart Tourism Apps

The smart tourism apps provide users with chatbot services provided by the smart tourism chatbot system as well as smart tourism app services provided by the smart tourism information system. The tourism information chatbot service is performed in three steps: (1) searching for tourism information ID in the smart tourism chatbot system, (2) searching for tourism information provided to users by using the tourism information ID in the smart tourism information system, (3) the user can check tourist information and save it to places of interest to create his travel itinerary. The smart tourism app consists of six menus as shown in Fig. 1: Best Tour Packages, My Travel Itinerary, Hallasan Beacon, Tourist Attractions, Accommodation, and Chatbot. The Best Tour Products menu provides users with recommended travel products created by travel experts, local tour guides, and users. When a user selects a recommended travel product, it is stored in My Travel Itinerary and used to create his travel itinerary. The My Travel Itinerary menu creates my travel itinerary using recommended travel products selected from the Best Tour Products menu and tourist information selected from the Tourist Attractions and Accommodation menus. The Tourist Attractions menu provides tourist information by classifying tourist attractions into three levels: large, medium, and small. Figure 1(d) shows the Hallasan tourism information in the medium category menu.

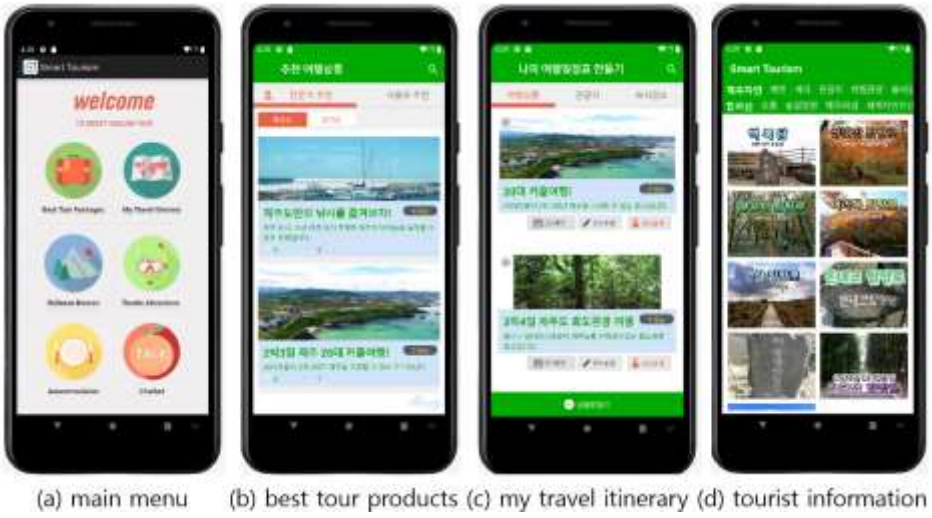

Fig. 1. Smart tourism application.

The Chatbot menu provides the tourism information chatbot service proposed in this paper to users. As shown in Fig. 2(b), the Chatbot menu provides Hallasan tourism information as shown in Fig. 1(d) when the user inputs "Please tell me information about Hallasan in Jeju Island?" in the chatbot input window. When the user selects the Baengnokdam tourist information provided by the Chatbot menu, the user receives the Baengnokdam tourist information from a Tourist Attractions submenu as shown in Fig. 2(c). After checking the tourist information as shown in Fig. 2(c), the user can save the tourist information to the place of interest in the My Travel Itinerary menu with the cart icon. The user can check the tourist information of the place of interest in the My Travel Itinerary submenu in Figure 2(d). The user can create his travel itinerary using the tourist information of the place of interest in the My Travel Itinerary menu.

2.2 DST Model-Based Tourism Information Chatbot System

The multi-domain tourism information DST model-based chatbot system accurately identifies the intent of the user's question and efficiently manages the dialogue state for the rule-based chatbot system. DST model-based tourism information chatbot system consists of DST and Neo4J DB servers. The tourism information knowledgebase of the Neo4J DB server has nodes of Area, Contents, and Services, and has attribute data that can provide tourism information suitable for user's questions. The tourism information DST dataset defines domain, slot, and value in the 4W1H method that matches the attribute data of the tourism information knowledgebase.

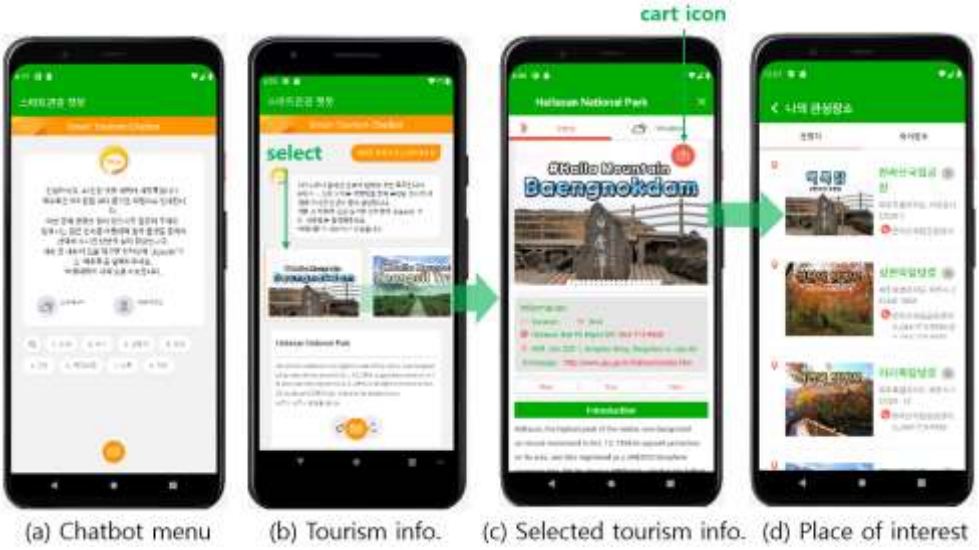

Fig. 2. Chatbot, Tourist Attractions, My Travel Itinerary submenus in the smart tourism app.

2.3 Tourism Information Named Entity Recognition (NER) Datasets

We developed Korean and English tourist information NER datasets to analyze tourist destination names from user's questions in a smart tourism chatbot system and performed transfer learning of the KoBERT-CRF NER model. The tourism information NER dataset was created by BIO tagging process for the tourism information source data using the tourism information NER dictionary as shown in Fig. 3. In this paper, we create the tourism information NER dataset to identify the intent of user's question using the newly defined BIO tagging dictionaries for the source data to which the tourism information DST dataset is added to the tourism information source data.

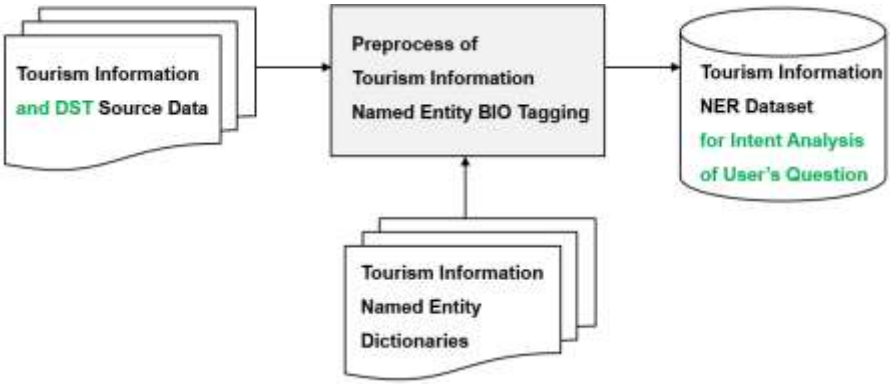

Fig. 3. Preprocessing of the NER dataset using the named entity BIO tagging dictionaries [9].

3. The Task-Oriented Smart Tourism Chatbot System

We have been developed tourism information DST, NER, and QA datasets and performing transfer learning to develop the task-oriented smart tourism chatbot system. We previously proposed the rule-based tourism information chatbot system and the tourism information DST model-based smart tourism chatbot system. The proposed tourism information chatbot system identifies the intention of the user's question, retrieves tourism information from the tourism information knowledge base built with Neo4J graph DB, and provides an answer to the user. In this paper, we propose the task-oriented tourism information chatbot system that provides tourism information provided by the smart tourism app. The proposed tourism information chatbot system is performed in three stages as follows:

(1) The smart tourism chatbot system identifies the intention of the user's question in the NER and DST servers and searches the tourism information ID in the tourism information knowledgebase.

(2) The smart tourism chatbot app searches for tourism information of the medium category in the smart tourism information system with the tourism information ID. The smart tourism chatbot app provides users with tourism information and tourism information of the medium category as the answer.

(3) The user can check tourist information with the smart tourism app and select and save it as a tourist destination for creating my travel itinerary as shown in Fig. 2(d).

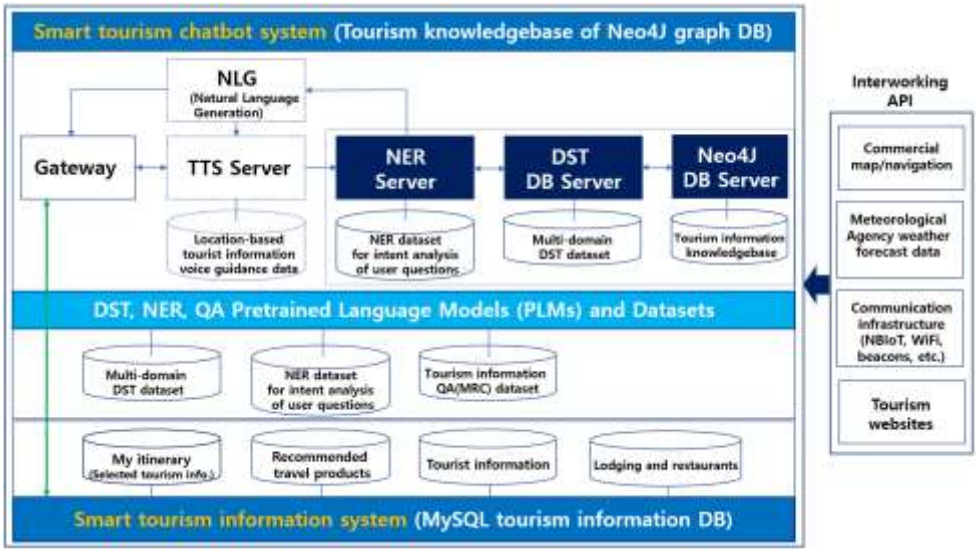

Fig. 4. Smart tourism chatbot and information systems.

3.1 Smart tourism chatbot and information systems

We are developing a smart tourism service platform to provide smart tourism services with the chatbot system and the mobile app. The smart tourism service platform consists of the smart tourism chatbot and information systems as shown in Fig. 4. The smart tourism information system provides best tour packages, my travel itinerary, tourist attractions, accommodation, and chatbot services using the smart tourism mobile app as shown in Fig. 1. The smart tourism app provides tourism information services interworking with commercial map and navigators, weather forecasts from the Korea Meteorological Administration, tourist destination websites.

The smart tourism chatbot system consists of the gateway, TTS, NER, DST, and Neo4J DB servers, and provides the tourism information chatbot service proposed in this paper. The TTS server with location-based tour information voice guidance data can be used to provide tour guide services according to my itinerary to tourists during travel. When a user enters a question into the chatbot input window, the intent of the user's question is

identified by the NER and DST servers through the gateway of the smart tourism chatbot system. The tourism information NER dataset for identifying the intention of the user's question in the NER server is created based on values of the domain-slot defined in the DST server using the existing 4W1H method. The tourism information knowledgebase implemented with Neo4J graph DB has values of the domain-slot as attribute data to search and provide appropriate tourism information for the user's question. The tourism information chatbot service proposed in this paper is performed in two steps: (1) The smart tourism chatbot system analyzes user's questions and searches for tourism information ID in the tourism information knowledgebase. (2) The smart tourism information system uses the tourism information ID to retrieve tourism information from the MySQL DB and provides it to the user using the Chatbot submenu.

3.2 Smart Tourism Chatbot System Using the Tourism Information NER And DST Models

Multi-domain DST chatbot systems use the predefined domains, slots, and values for dialogue management and understanding the intent of user's questions. The smart tourism chatbot system that provides tourism information, recommended travel products, my travel products, and tour guide services has many values in one domain-slot. In a task-oriented chatbot system where many values exist in one domain-slot, such as the smart tourism chatbot system, it is more efficient to find the value of the domain-slot with the NER algorithm. In this paper, we apply the tourism information NER model to the DST model-based tourism information chatbot system to accurately identify the intention of the user's question. Figure 5 shows the service scenario of the proposed smart tourism chatbot system using the tourism information DST and NER models. When a user enters a question with the Chatbot menu, the smart tourism chatbot system searches for values (tourism information named entities) of the predefined domain-slots in the NER server. We use the values of the domain-slot defined by the 4W1H method in the multi-domain DST chatbot system as tourist information named entities in the NER model. The DST server performs dialogue management using the user's question and creates a Neo4J query statement to retrieve the tourism information id from the tourism information knowledgebase.

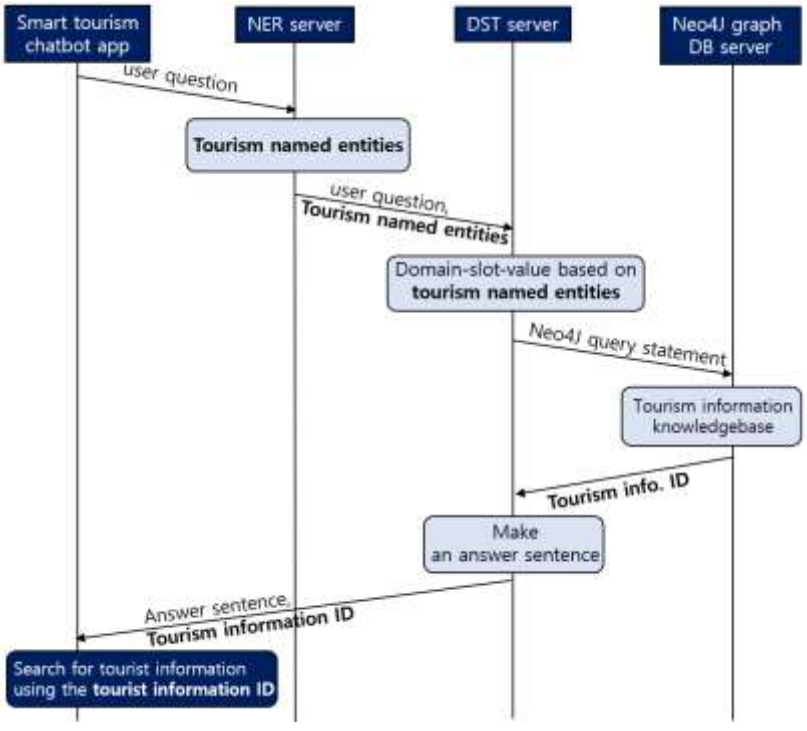

Fig. 5. The service scenario of the proposed smart tourism chatbot system using the NER model.

In the second step, the Chatbot menu searches the tourism information requested by the user in the smart tourism information system using the searched tourism information ID. The tourism information ID includes the large and medium classification codes to which the tourism information belongs according to the menu classification

system of the smart tourism app. The Chatbot menu finds medium-category tourist information that includes tourist information so that users can easily and conveniently select the desired tourist destination and provides them as answers to the user as shown in Fig. 2(b). In the third step, the Chatbot menu provides tourism information to the user, and the user checks the tourism information with the smart tourism app, selects and stores the tourist information of interest as shown in Fig. 2(d). The user can check the selected tourist information from the points of interest submenu in the My Travel Itinerary menu. The user can create my travel itinerary using the selected tourism information and recommended travel products. During the trip, the user can use the tour guide service according to my itinerary, and can change tourist information according to context awareness using the smart tour chatbot app and my itinerary app. After the trip, the user can register the travel itinerary used for the trip as a recommended travel product to provide it to others, including acquaintances.

3.3 Tourism information NER model and dataset in the smart tourism chatbot system

We use the tourism information NER model to efficiently identify the intention of user's questions in the DST model-based smart tourism chatbot system. The existing tourism information NER dataset is created with BIO tagging for smart tourism contents using Korean and English tourist name dictionaries. The Korean and English tourist name dictionaries used for BIO tagging are designed to correct misspellings as well as abbreviations and aliases of tourist destination names. In this paper, we define tourism information named entities based on the domain, slot, and value defined in the DST model with the 4W1H method. Table 1 shows an example of the tourism information named entity defined by the 4W1H method. The tourism information named entities in Table 1 are created based on values of the domain-slot defined in the tourism information DST model. In Table 1, the tourist information named entity belonging to How, which represents the tourist's travel method, is similar to the named entity of What, which represents the purpose of travel, such as experiential tourism.

Table 1. An example of a tourism information named entity defined by the 4W1H method.

4W1H	Tourism information named entity
Who	Travel commentator, Alone, Friend, Family, Child, Couple, Colleague, Pet(puppy, cat)
What	Dark tourism, Cultural tourism, Experiential tourism, Marine tourism, Wellness, etc.
When	Season(Spring, Summer, Autumn, Winter), Time(Weekdays, Weekends, Now, Morning, Noon, Evening, Day, Night), Duration(Festival, Flower, Open time), etc.
Where	Tourist attractions, Tourism Information Classification Category(Large, Medium, Small), Photo(Flowers, Sunrise/Sunset), Area(address), etc.
How	Climbing, Walking, Driving, Biking, Personal transportation, etc.

We create tourism information NER dictionaries for BIO tagging of tourist information contents based on the tourist information named entities in Table 1. Figure 6 shows tourism information NER dictionaries for tourism information BIO tagging. In Fig. 6, the tourist destination NER dictionary is used for BIO tagging for the tourist information classification codes to tourist destination names. The tourist information named entity belonging to Who is the same as that in Table 1 and represents a travel companion. The tourist information named entity belonging to What is classified into general experiential tourism and sea experiential tourism. The tourism information NER dictionary of How is not created separately but is included in that of What. The tourist information named entity belonging to When is used for selecting tourist destinations for tour guide service during travel as well as tourist information for making my itinerary.

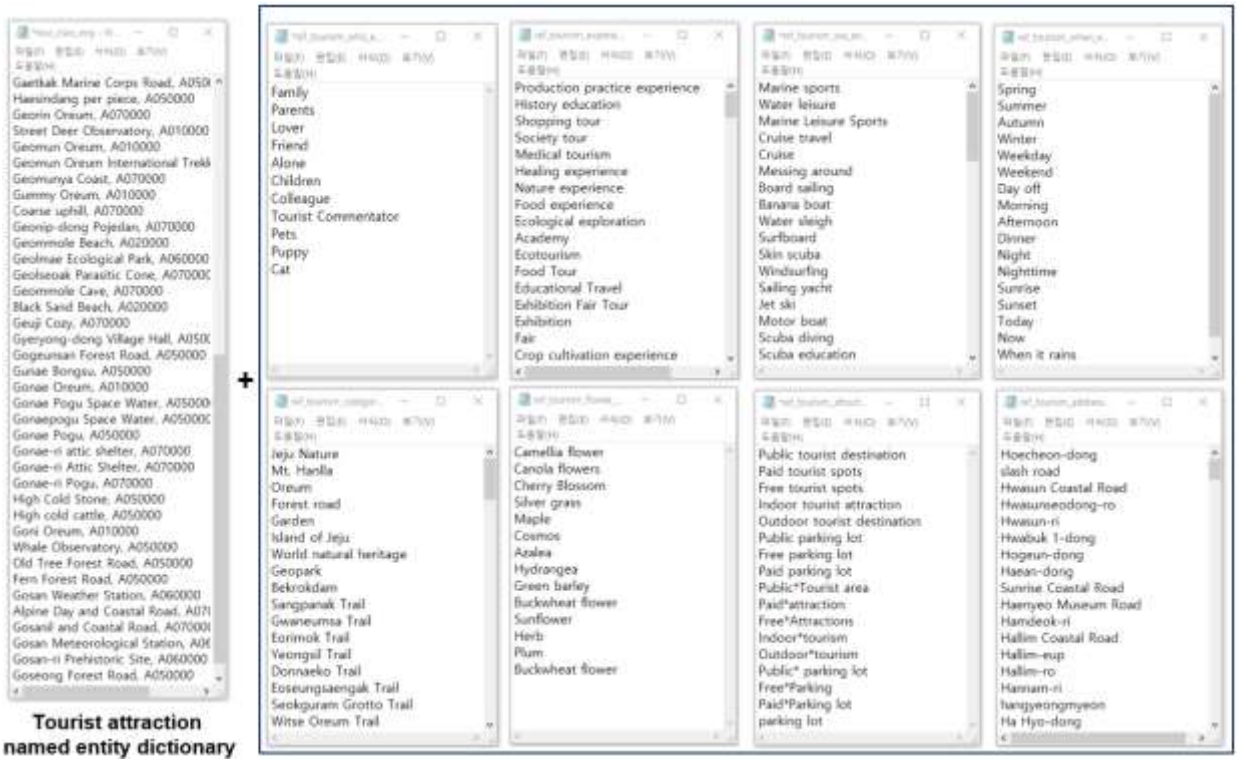

Fig. 6. The tourism information NER dictionaries based on the values in the DST model.

The tourist information classification named entities that belongs to Where are defined using the menu of the smart tourism app. The smart tourism chatbot system provides answers to users based on tourism information classification codes. Users can receive the same tourism information as smart tourism apps according to the tourism information classification code with the smart tourism chatbot app, so they can easily and conveniently check tourist information as shown in Fig. 2(c). The Area (address) named entity belonging to Where includes not only the name of the administrative district but also the commonly used name of the area. Administrative district named entities are created from tourist destination addresses.

Table 2. The KoBERT NER performance on the tourism information NER dataset

Named entity	Precision	Recall	F1-score	
TOUR	0.98	0.99	0.99	19307
TOUR_I	0.96	0.98	0.97	7535
CAT	1.00	0.98	0.99	3684
EXP	0.99	0.98	0.99	1839
FLOW	0.99	1.00	0.99	616
HOW	1.00	0.98	0.99	357
LOC	0.99	1.00	1.00	7875
WHEN	0.99	1.00	1.00	1432
WHO	1.00	0.99	1.00	861

The tourism information NER dataset for transfer learning of NER model is created by including the tourism information content of the smart tourism information system and the conversation between the user and the system of the DST learning dataset. Table 2 shows the performance of the KoBERT NER model on the tourism information NER dataset as micro average, macro-average, and weighted average for F1, recall, and precision scores. The total number of tourism information entities tagged with TOUR (tourist attractions), CAT (category), EXP (experiential tourism), FLO (flower), HOW, LOC (location) WHEN, WHO are 19307, 3684, 1839, 616, 357, 7875, 1432, 861, respectively. The precision, recall, and F1 scores of the KoBERT NER model are 0.985, 0.982, 0.987,

respectively. The task-oriented smart tourism chatbot system proposed in this paper uses the NER model to predict the value of the DST model's domain and slot because the performance of the NER model is generally better than that of the DST model.

CONCLUSION AND FURTHER WORKS

We are developing the smart tourism platform and mobile apps that provides context-aware and personalized travel itineraries to individual tourists and provides tour guide services. In this paper, we propose the task-oriented smart tourism chatbot service that efficiently provides tourism information using a chatbot service in the smart tourism platform and mobile apps. The developed smart tourism platform consists of the smart tourism chatbot system composed of the NER, DST, and Neo4J graph DB servers, and the smart tourism information system that provides tourist information, recommended travel products, and personalized travel itinerary using mobile apps. The proposed smart tourism chatbot system efficiently and accurately identify the intention of the user's question by applying the NER model to the DST model-based chatbot system. The NER model predicts the values of domain-slot to identify the intention of the user's question by using the tourist information NER dictionaries created with the values of the DST model. We use the values defined by the 4W1H method in the DST model as tourist information named entities in the NER model to determine the domain-slot-value by analyzing a user's question in the NER server. Since the tourism information NER dictionaries include the menu name and classification code of the smart tourism app, the smart tourism chatbot service can search for the tourism information of the smart tourism app.

The task-oriented chatbot service proposed in this paper can be applied to various task-oriented mobile app services such as creating my travel itinerary and making hotel and restaurant reservations. We plan to develop the function of the smart tourism platform so that individual tourists can easily and conveniently create their itinerary using the smart tourism chatbot service. We also plan to develop a smart tourism chatbot service that provides tourism information QA services through smart tourism platforms and mobile apps.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2021R1A2C1093283).

REFERENCES

[1] Yunchuan Sun, Houbing Song, Antonio J. Jara, Rongfang Bie, "Internet of Things and Big Data Analytics for Smart and Connected Communities", IEEE Access, Vol. 4, pp. 766-773, Feb. 2016.

[2] Ajaya K. Tripathy, Pradyumna K. Tripathy, Niranjan K. Ray, Saraju P. Mohanty, "iTour: The Future of Smart Tourism: An IoT Framework for the Independent Mobility of Tourists in Smart Cities", IEEE Consumer Electronics Magazine, Vol. 7, No. 3, pp. 32-37, April 2018.

[3] Tsaih, Rua-Huan and Hsu, Chih Chun, "Artificial Intelligence in Smart Tourism: A Conceptual Framework", The 18th International Conference on Electronic Business, pp. 124-133, Dec. 2018.

[4] Wei Wang, Neeraj Kumar, Junxin Chen, Zhiguo Gong, Xiangjie Kong, Wei Wei, Honghao Gao, "Realizing the Potential of the Internet of Things for Smart Tourism with 5G and AI", IEEE Network, Vol. 34, No. 6, pp. 295-301, 2020.

[5] Hyun-Ji Cho, Jin-Yi Lee, Tae-Rang Park, Jeong Woo Jwa, "React Native and Android Mobile Apps for Smart Tourism Information Service to FITs", The International Journal of Internet, Broadcasting and Communication, vol.14, no.2, pp.63-69, 2022.

[6] KiBeom Kang, JeongWoo Jwa, SangDon Earl Park, "Smart Audio Tour Guide System using TTS", International Journal of Applied Engineering Research, pp.9846-9852, 2017.

[7] Jam, F. A., Singh, S. K. G., Ng, B., & Aziz, N. (2018). The interactive effect of uncertainty avoidance cultural values and leadership styles on open service innovation: A look at malaysian healthcare sector. International Journal of Business and Administrative Studies, 4(5), 208-223.

[8] Dong-Hyun Kim, Hyeon-Su Im, Jong-Heon Hyeon, Jeong-Woo Jwa, "Development of the Rule-based Smart Tourism Chatbot using Neo4J graph database", International Journal of Internet, Broadcasting and Communication, Vol.13, No.2, pp 179-186, 2021.

[9] Myeong-Cheol Jwa, Tae-Seung Ko, Byeong-Joo Kim, Jeong-Woo Jwa, "Tourism Information Multi-domain Dialogue State Tracking Datasets for Smart Tourism Chatbot", International Journal of Intelligent Systems and Applications in Engineering, vol.10, no.1S(2022), pp. 192-196, 2022.

[10] Myeong-Cheol Jwa, Jeong Woo Jwa, "Development of Tourism Information Named Entity Recognition Datasets for the Fine-tune KoBERT-CRF Model", International Journal of Internet, Broadcasting and Communication, Vol.14, No.2, pp 55-62, 2022.

[11] Ko, Tae-Seung, Jwa, Myeong-Cheol, Jwa, Jeong-Woo, "Tourism Information QA Datasets for Smart Tourism Chatbot", International Journal

of Membrane Science and Technology, pp. 243-248, June 2023.

[12] SKT KoBERT, https://github.com/SKTBrain/KoBER

[13] Ko, T.-S. ., Jwa, M.-C. ., & Jwa, J.-W. . (2023). Tourism Information QA Datasets for Smart Tourism Chatbot. International Journal of Membrane Science and Technology, 10(1), 243-248. https://doi.org/10.15379/ijmst.v10i1.1451

[14] Kim, Sungdong, et al. "Efficient dialogue state tracking by selectively overwriting memory." Proceedings of the 58th Annual Meetings of the Association for Computational Linguistics, 2020.

[15] KoBigBird model, https://github.com/monologg/KoBigBird

[16] Park, Sungjoon, Jihyung Moon, Sungdong Kim, Won Ik Cho, Jiyoon Han, Jangwon Park, Chisung Song et al., “KLUE: Korean Language Understanding Evaluation.”, arXiv, 2021.

[17] Hoon-chul Kang, Ki-Beom Kang, Dong-Hyun Kim, Myeong-Cheol Jwa, Tae-Seung Ko, Jeong-Woo Jwa, "Smart tourism chatbot system using Multi-domain Tourism Information DST", 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN), pp. 608-612, July 2023.

DOI: https://doi.org/10.15379/ijmst.v10i4.1888