Environment. Technology. Resources. Rezekne, Latvia
Proceedings of the 15th International Scientific and Practical Conference. Volume II, 179-182

Rapid Development of Chatbot for Tourism Promotion in Latgale

Sergejs Kodors
Institute of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
sergejs.kodors@rta.lv

Guna Kanepe
Rezekne Academy of Technologies
Rezekne, Latvia
gk11056@edu.rta.lv

Daniēls Zeps
Rezekne Academy of Technologies
Rezekne, Latvia
dz22036@edu.rta.lv

Imants Zarembo
Institute of Engineering
Rezekne Academy of Technologies
Rezekne, Latvia
imants.zarembo@rta.lv

Lienīte Litavniece
Research Institute for Business and Social Processes
Rezekne Academy of Technologies
Rezekne, Latvia
lienite.litavniece@rta.lv

Abstract. **The release of ChatGPT technology identified the large language models as a new disruptive technology, which changes the behaviours of society and its attitude towards the presence of artificial intelligence in everyday life. The tourism industry is one of the economic sectors, which will be impacted by the large language models through personalized marketing and advertisements. A common approach to capture the attention of AI-centric tourists, who want to get answers to their questions without manually researching the topic or using services of the travel advisors, is to integrate a chatbot or virtual assistant in the tourism information system. We applied this approach to the promotion of tourism in East Latvia (Latgale) by rapidly developing a chatbot by using a prompt method with context-oriented material. Two models were prepared for tourism promotion in Latgale. The models were evaluated through a pilot survey to understand the satisfaction of target users. The data analysis was applied. The study identified the importance of trustworthy information and answer saturation. The trade-off between dialog freedom and trustworthiness of answers can be achieved through the development of microservices, which are grouped as one system to direct conversation with chatbot. The appropriate conceptual models are presented in the article.**

Keywords: chatbot, large language models, system modelling, tourism.

I. Introduction

Tourism is an important sector of the national economy, as it constitutes a significant part of the GDP of many countries and employs many people. This encourages fierce competition between different tourism destinations to motivate and interest potential tourists. As a result, significant importance is attached not only to the customer service process when the tourist has already arrived at the final destination but also to interest the tourist in choosing this final destination.

To attract tourists, different tools and ways of presenting information are used to create interest, taking into an account also the behaviour patterns and trends of consumers, when it is necessary to be able to find information quickly, investing as little time as possible.

One of the latest technologies that is used in various fields is artificial intelligence (AI). Speaking about tourism, AI is enabling technology, which can underpin current service innovations that impact firm-customer interactions with implications for service management and marketing. Large language model (LLM) is a type of AI, which can process and generate text. For example, automated teller machines (ATMs) have been in use since the 1960s, whilst chatbots, such as Siri and Alexa are now widely adopted as customer-facing service robots [1].

The aim of study is to evaluate the user-satisfaction with a chatbot rapidly developed by using LLM and prompt method.

Our experiment showed that users are mainly satisfied with the developed chatbots for tourism promotion in Latgale. However, the chatbots sometimes generate false or too abstract information about the region, which strongly decreases the user satisfaction with the chatbots. In result we propose to apply LLMs as microservices to transform tourist personal preferences into data search parameters for

Print ISSN 1691-5402
Online ISSN 2256-070X
https://doi.org/10.17770/etr2024vol2.8060

standard algorithms, which can satisfy the truthfulness of data.

II. MATERIALS AND METHODS

The experiment design is based on a similar study, which was conducted to evaluate question and answer (Q&A) LLM trained on medical data [2].

The experiment is organized in two stages:

1) prepare a chatbot; 2) conduct a pilot survey to understand user satisfaction through the questionnaire.

The chatbot was developed using the ChatPDF tool, which takes contextual prompts as a PDF document. As result, this tool provides a rapid development of a chatbot simply collecting tourism material about the target region.

To get more objective results, we have asked two experts to prepare promotional material about the Latgale region. We asked experts to include the material of 8 themes: Latvia, Latgale, Daugavpils (the city in Latgale), restaurants, recreational activities, local producers, cycling routes, and observation towers. The experts collected promotion materials independently one from another. As a result, we developed two chatbots, which can be integrated into the tourism platform through ChatPDF API.

The questionnaire was taken from the study of Singhal et al. (2023) [2]. The following six questions were included in the questionnaire:

Q1: Alignment with question consensus;

Q2: Reading comprehension;

Q3: Knowledge recall;

Q4: Reasoning;

Q5: Inclusion of irrelevant content;

Q6: Omission of important information;

Additionally, we included questions about “Incorrect data” (Q7) and “Satisfaction with chatbot” (Q8).

The pilot group included 12 students (volunteers) and 3 tourism industry experts. Each respondent needed to ask questions to the chatbot about tourism possibilities considering the mandatory themes and evaluate the chatbot using Likert scale 5, where 1 – very weak, 5 – very strong. Each theme was evaluated by the respondent independently to get the mean values after conversation. Meanwhile, the respondents must ask similar questions to both chatbots and fill answers in the comparative style.

III. RESULTS

Overall, both chatbots obtained sufficiently high scores. The mean satisfaction of Chatbot A is 4.29 and Chatbot B – 4.18 (good). However, the most of the respondents preferred Chatbot B (see Fig. 1). The respondents adduced in the open conclusions that they were more satisfied with Chatbot B. However, we can see the smaller mean value, which can be explained by the high deviation of the satisfaction with Chatbot B (see Fig. 1). Let’s investigate the impact factors on the user decisions.

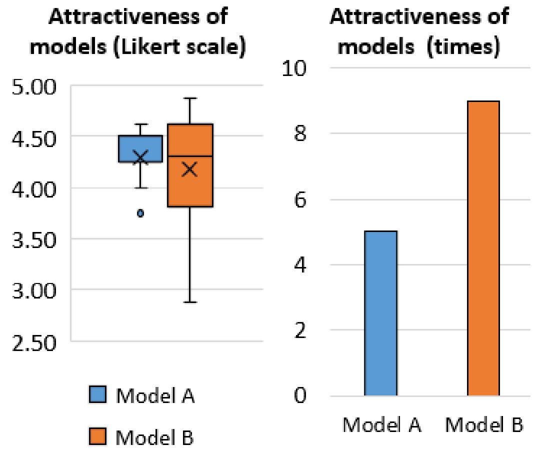

Fig. 1. Satisfaction with chatbots (Q8): left – Likert 5, right – how many times one model was better than the other.

The data analysis showed that this deviation is called by incorrect and abstract answers provided by Chatbot B (see Fig. 2). Fig. 2 is organized in the comparative style: the answers on similar questions are placed near for both chatbots.

The data analysis was completed through anomaly investigation. We can see that the standard deviation of all questions can be separated into three groups: near 1.0, near 0.5, and near 0.3 (see Fig. 2); where the highest group “near 1.0” belongs to Q5, Q6, and Q7. These questions are related to the precision of information provided by the chatbots.

These three questions (Q5-7) show anomalies. The review of the anomaly in questionnaire showed comments about incorrect locations of tourism products, which compromises the user's trust in the chatbot and satisfaction with it. The mean value of Q5-7 was compared with Q8 using Spearman correlation, which showed a strong negative impact of -0.75 and -0.80 for Chatbot A and Chatbot B respectively.

The group “near 0.5” belongs to Q3 (knowledge recall), which describes the abstraction level of answers. In this case, Spearman correlation was very strong -0.92 for Chatbot B and very weak for Chatbot A (-0.07). The review of the comments supports the importance of respondent satisfaction with conciseness and accuracy of answers. However, the abstraction impact is mentioned more rarely than precision.

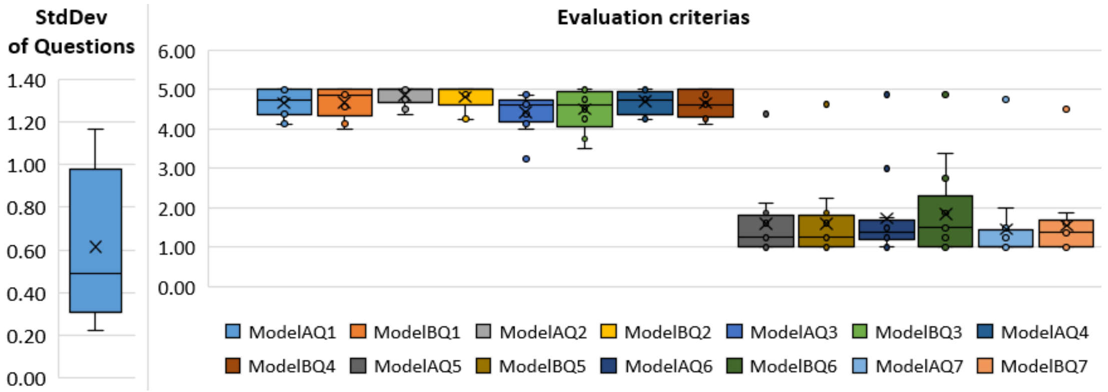

Fig. 2. Answers of respondents: Model A belongs to Chatbot A, Model B – to Chatbot B, Q1-7 identify questions respectively.

It must be noted that we completed data cleaning through the investigation of anomalies and reviewing the standard deviation of questions to scope the summary opinions. We excluded one questionnaire, which contained very good results without comments. The other two were partially cleaned based on the comments, which identified that the respondents wanted to get specific answers from chatbot related to their subjective preferences.

Analysing the comments, the reason why respondents prefer Chatbot B, they mention the more structured answers of the chatbot. Reviewing the PDF document of Chatbot B, the second expert used material with a similar structure: lists of tourism products instead of descriptive text. Document A contained 477 pages, but Document B – 712 pages. It is possible that too long prompts can call for unstable answers of Chatbot B. Therefore, it is recommended to make compact and structured documents (catalogues) for chatbot prompts.

IV. DISCUSSION

So, the prompt-based chatbot is sufficiently usable for the conversation with clients. It can be applied to attract attention and interest to tourists. However, it is not a sufficiently stable and trustworthy solution to generate personalized travel plans for tourists, additionally, it can lead to the opposite effect creating a negative experience to its users. In result it is required to discuss enhancement solutions to overcome this problem.

The success of LLMs lies in their ability to capture the statistical patterns and linguistic nuances present in the training data [3]. By processing and analysing vast amounts of text, LLMs gain a comprehensive understanding of language to be able to generate coherent and contextually relevant responses [4]. Considering to Chinchilla scaling law for LLM, which was developed by Hoffman et al. (2022) [5], the LLM size and the amount of training data should be increased in approximately equal proportions. It means that the training from scratch of LLM is a costly process, especially for multi-language tasks, that includes the cost of computing and data collection.

Another approach is fine-tuning of LLM and retraining it on the target domain data. E.g. G. Trichopoulos et al. (2023) fine-tuned GPT-4 to assist museum visitors in providing textual data about locations and exhibition narratives [6]. However, it is hard to guarantee the precision of location due to the black-box principle of neural networks as well as the textual description of orientations of the whole geographical region will be too challenging.

The non-real-time information, incomplete knowledge, and insufficient spatial awareness are mentioned by other researchers too [7]. As a solution, LLM is presented as a microservice, which transforms the tourist request into JSON parameters applied by the itinerary planning system to provide the personalized content (see Fig. 3).

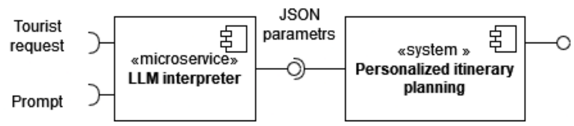

Fig. 3. LLM as microservice for tourist request interpretation with chatbots.

The systems with user profile data can apply LLMs in opposite directions. E.g. LLM can add emotions to the personalized messages generated by the recommendation system based on the tourist profile data (see Fig. 4) [8].

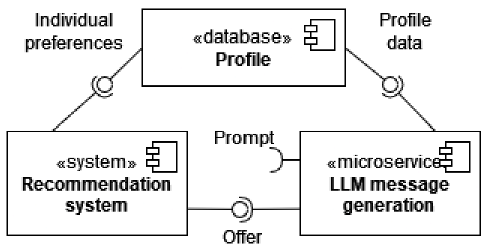

Fig. 4. LLM generates personalized message.

Speaking about microservice architecture, it is typical and preferable for the tourism domain due to the possibility to extend recommendation systems with different data sources and services (see Fig. 5) [9].

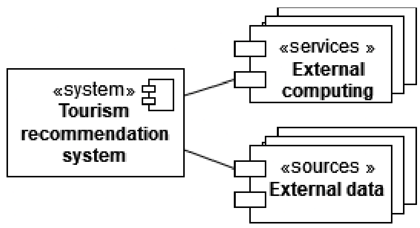

Fig. 5. Microservice architecture of tourism recommendation systems.

Additional factor is real-time data processing by LLMs. LLM simply can not be retrained to use these data due to the short time of their actuality, e.g., weather data, tickets, events, etc. Another example is speech generation using LLM output text. E.g. Yoshimaru et al. (2023) applied one LLM to get search parameters for sightseeing DB, meanwhile, the second LLM for speech generation to sound out recommendations [10].

The limits of a single LLM are mentioned by different authors. Meanwhile, Wu et al. (2023) presented a multi-agent framework, which connects together multiple LLMs where each is prepared for specific tasks. Additionally, LLMs can communicate with one other to complete complex tasks [11].

Considering the rapid development, the optimal solution is expressed by the possibility to combine multiple LLMs into one system completing the minimal LLM training and combining them with data sources and secure algorithms. The sketch of a tourism chatbot, which can solve the problems with spatial awareness, abstraction, and multiple language support for East Latvia's promotion to tourists, is depicted in Fig. 6. That is our next frontier for research and development.

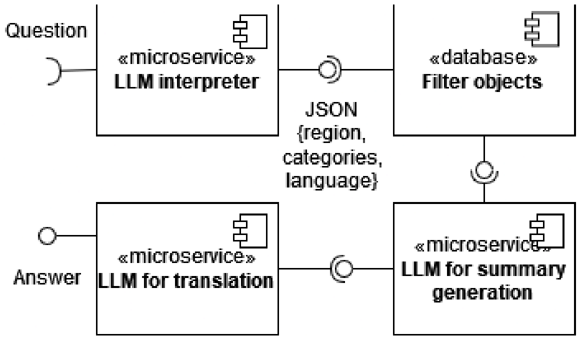

Fig. 6. Promotion chatbot based on microsevice architecture with LLMs.

CONCLUSIONS

The experiment has shown that while LLM based chatbots are a very interesting and prospective direction for future technological advancements in the tourism industry, there are a number of challenges yet to overcome before confidently integrating virtual assistants into publicly available tourism information systems. Lack of factual correctness when it comes to tourism object details and locations requires the addition of more mature technologies to support LLMs to be able to come close to generating travel plans or itineraries. Microservice approach proposed in this paper is designed to take the planning load off of the LLM and leave only the input interpretation, textual content translation, and output decoration functions. The microservice approach can be extended by additional service providers to enable additional features, like booking, weather forecasting, and external image scraping.

ACKNOWLEDGEMENT

This research is funded by the Latvian Council of Science project "Digital twin to promote tourism competitiveness and complementarity development: a Latgale region use case", project No. lzp-2022/1-0350.

REFERENCES

[1] D. Buhalis, T. Harwood, V. Bogicevic, G. Viglia, S. Beldona, C. Hofacker, "Technological disruptions in services: lessons from tourism and hospitality," Journal of Service Management, vol. 30, no. 4, pp. 484-506, May 2019. https://doi.org/10.1108/JOSM-12-2018-0398

[2] K. Singha, T. Tu, J. Gottweis, R. Sayres, E. Wulczyn, L. Hou, K. Clark, S. Pfohl, H. Cole-Lewis, D. Neal, M. Schaekermann, A. Wang, M. Amin, S. Lachgar, P. Mansfield, S. Prakash, B. Green, E. Dominowska, B. Arcas, N. Tomasev, Y. Liu, R. Wong, C. Semturs, S. Mahdavi, J. Barral, D. Webster, G. Corrado, Y. Matias, S. Azizi, A. Karthikesalingam and V. Natarajan, "Towards Expert-Level Medical Question Answering with Large Language Models," May 16, 2023. [Online]. Available https://doi.org/10.48550/arXiv.2305.09617 [Accessed Feb. 27, 2024]

[3] P. Li, M. Zhang, P. Lin, J. Wan, and M. Jiang, "Conditional embedding pre-training language model for image captioning," Neural Processing Letters, vol. 54, no. 6, pp. 4987–5003, 2022. https://doi.org/10.1007/s11063-022-10844-3

[4] M. Hadi, Q. Al-Tashi, R. Qureshi, A. Shah, A. Muneer, M. Irfan, A. Zafar, M. Shaikh, N. Akhtar, J. Wu, and S. Mirjalili, "A Survey on Large Language Models: Applications, Challenges, Limitations, and Practical Usage," July 2023. [Online] Available https://www.researchgate.net/publication/372258530_Large_Language_Models_A_Comprehensive_Survey_of_its_Applications_Challenges_Limitations_and_Future_Prospects [Accessed Feb. 27, 2024]

[5] J. Hoffmann, S. Borgeaud, A. Mensch, E. Buchatskaya, T. Cai, E. Rutherford, D. Casas, L. Hendricks, J. Welbl, A. Clark, T. Hennigan, E. Noland, K. Millican, G. Driessche, B. Damoc, A. Guy, S. Osindero, K. Simonyan, E. Elsen, J. Rae, O. Vinyals and L. Sifre, "Training Compute-Optimal Large Language Models," Mar 29, 2022. [Online] Available https://doi.org/10.48550/arXiv.2203.15556 [Accessed Feb. 27, 2024]

[6] G. Trichopoulos, M. Konstantakis, G. Alexandridis, G. Caridakis, "Large Language Models as Recommendation Systems in Museums," Electronics, vol. 12, no. 18, p. 3829, 2023. https://doi.org/10.3390/electronics12183829

[7] Y. Tang, Z. Wang, A. Qu, Y. Yan, K. Hou, D. Zhuang, X. Guo, J. Zhao, Z. Zhao, W. Ma, "Synergizing Spatial Optimization with Large Language Models for Open-Domain Urban Itinerary Planning," Feb. 11, 2024. [Online] https://doi.org/10.48550/arXiv.2402.07204 [Accessed Feb. 27, 2024]

[8] M. Remountakis, K. Kotis, B. Kourtzis, G. Tsekouras, "Using ChatGPT and Persuasive Technology for Personalized Recommendation Messages in Hotel Upselling," Information, vol. 14, p. 504, 2023. https://doi.org/10.3390/info14090504

[9] L. Garcia, S. Aciar, R. Mendoza and J. Puello, "Smart Tourism Platform Based on Microservice Architecture and Recommender Services," In: International Conference on Mobile Web and Intelligent Information Systems, Jul. 14, 2018, pp 167–180 [Online]. https://doi.org/10.1007/978-3-319-97163-6

[10] N. Yoshimaru, M. Okuma, T. Iio and K. Hatano, "AsyncMLD: Asynchronous Multi-LLM Framework for Dialogue Recommendation System," Dec 21, 2023. [Online] Available https://doi.org/10.48550/arXiv.2312.13925 [Accessed Feb 27, 2024]

[11] N. Yoshimaru, M. Okuma, T. Iio and K. Hatano, "AsyncMLD: Asynchronous Multi-LLM Framework for Dialogue Recommendation System", Dec 21, 2023. [Online] Available https://doi.org/10.48550/arXiv.2308.08155 [Accessed Feb 27, 2024