

A Multi-Agent Framework for Synthetic Review Generation in Tourism

Margarita Reyes-Sierra^{1,*}

¹Centro de Investigación en Matemáticas A.C., Sede Monterrey

Abstract

Recent advances in Large Language Models (LLMs) have enabled the generation of realistic synthetic text for numerous Natural Language Processing applications. However, producing high-quality synthetic tourism reviews involves multiple complementary tasks, including content generation, consistency verification, and sentiment validation. Performing all these activities through a single prompting stage may lead to inconsistent reviews, duplicated information, or incorrect sentiment assignments. This paper presents a multi-agent framework for synthetic tourism review generation designed for the REST-MEX 2026 shared task. Instead of relying on a single language model interaction, the proposed framework distributes the generation process among three specialized agents. A Generator Agent produces the initial review according to the desired metadata, a Reviewer Agent evaluates linguistic coherence and structural quality, and a Validator Agent verifies that the generated review satisfies the required sentiment polarity and formatting constraints. Reviews that do not satisfy the validation criteria are automatically regenerated. The proposed framework provides a simple and modular workflow for synthetic corpus construction. By separating generation, review, and validation into independent stages, the methodology facilitates reproducibility while maintaining the flexibility required for different Large Language Models and tourism datasets.

Keywords

Multi-Agent Systems, Large Language Models, Synthetic Reviews, Tourism, REST-MEX, Natural Language Processing

1. Introduction

The rapid evolution of Large Language Models has considerably expanded the possibilities of automatic text generation. Modern language models are capable of producing coherent documents, answering complex questions, summarizing information, and generating domain-specific content from relatively simple textual instructions. These capabilities have motivated their adoption in numerous scientific and industrial applications.

Within the tourism industry, user-generated reviews constitute one of the most valuable sources of information for understanding visitor experiences, customer satisfaction, destination image, and service quality. Consequently, Natural Language Processing has become an essential component of tourism analytics, supporting applications such as sentiment analysis, recommendation systems, conversational assistants, destination image analysis, hotel reputation monitoring, and opinion mining [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11].

The REST-MEX shared task has contributed significantly to this research area by providing standardized evaluation benchmarks for Spanish-language tourism texts. Previous editions addressed recommendation systems, sentiment analysis, and tourism text classification [12, 13, 14, 15, 16]. More recently, REST-MEX 2026 introduced a synthetic review generation task in which participants are required to construct balanced collections of tourism reviews suitable for training sentiment classification models [17, 18].

Generating synthetic reviews is not simply a matter of producing grammatically correct sentences. Every review should satisfy multiple requirements simultaneously. Besides expressing the requested sentiment polarity, the generated text should describe an appropriate tourism experience, maintain

IberLEF 2026, September 2026, León, Spain

*Corresponding author.

✉ maria.sierra@cimat.mx (M. Reyes-Sierra)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

linguistic coherence, avoid duplicated content, and preserve a natural writing style consistent with authentic online reviews. Managing all these aspects within a single generation step can be challenging, particularly when large synthetic corpora are required.

A practical alternative consists of decomposing the generation process into several specialized tasks. Rather than expecting a single language model interaction to solve every aspect of review generation, different components can independently focus on content creation, quality revision, and validation. This modular perspective simplifies the overall workflow and facilitates the incorporation of additional verification stages without modifying the generation process itself.

Inspired by this idea, this paper proposes a simple multi-agent framework for synthetic tourism review generation. The framework consists of three collaborative agents. The Generator Agent produces the initial review according to a set of metadata describing the desired sentiment polarity, attraction category, municipality, and approximate review length. Subsequently, the Reviewer Agent evaluates the generated review from a linguistic perspective, identifying structural inconsistencies or formatting problems. Finally, the Validator Agent verifies that the generated review satisfies the requested constraints before incorporating it into the final synthetic corpus.

Unlike more sophisticated autonomous multi-agent architectures, the proposed framework emphasizes simplicity, modularity, and reproducibility. Each agent performs a single well-defined task, allowing the complete workflow to be easily implemented using different language models without requiring complex coordination mechanisms. The main contribution of this work is a modular multi-agent framework for synthetic tourism review generation that separates the generation process into three complementary stages: content generation, linguistic review, and final validation. By assigning each task to an independent agent, the proposed workflow simplifies corpus construction while improving reproducibility and facilitating its integration with different Large Language Models without requiring architectural modifications. The resulting synthetic corpus is subsequently evaluated according to the official REST-MEX 2026 evaluation protocol.

This paper is organized as follows: Section 2 describes the motivation behind collaborative multi-agent generation. Section 3 presents the proposed framework and the interaction among agents. Section 4 reports the experimental evaluation. Finally, Section 5 concludes the paper and outlines future research directions.

2. Collaborative Generation in Large Language Models

The remarkable capabilities of Large Language Models (LLMs) have considerably simplified automatic text generation across numerous Natural Language Processing tasks. Modern LLMs are capable of producing coherent documents from relatively simple textual instructions while adapting their responses to different writing styles, application domains, and user requirements. These characteristics have motivated their adoption in a wide variety of applications including conversational assistants, automatic summarization, educational content generation, software development, and scientific writing.

Within tourism, these advances have opened new opportunities for generating synthetic reviews that can be used for data augmentation, privacy preservation, and corpus balancing. Tourism reviews constitute a particularly interesting application because they naturally combine subjective opinions, descriptions of places, customer experiences, recommendations, and sentiment expressions. Consequently, Natural Language Processing techniques have been successfully applied to sentiment analysis, recommendation systems, destination image assessment, topic discovery, hotel reputation analysis, conversational assistants, gastronomy, and tourism intelligence [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11].

The REST-MEX initiative has significantly contributed to this research area by providing standardized datasets and evaluation campaigns focused on Spanish-language tourism texts. Previous editions investigated recommendation systems, sentiment analysis, and text classification for tourism reviews [12, 13, 14, 15]. More recently, REST-MEX 2026 introduced a synthetic review generation task in which participants are required to generate balanced tourism corpora suitable for training sentiment classification models [17, 18].

2.1. Motivation for Collaborative Generation

Generating a realistic tourism review involves satisfying several requirements simultaneously. Besides producing grammatically correct text, the generated review should express the desired sentiment polarity, describe an appropriate tourism experience, preserve coherence throughout the document, and satisfy the formatting requirements established by the shared task [17, 18].

Although these objectives can be addressed within a single interaction with a language model, separating them into smaller independent tasks simplifies the overall generation process. Instead of expecting one prompt to control every aspect of the review, individual components may focus on specific responsibilities such as content generation, quality inspection, or constraint verification. Similar modular workflows have been successfully adopted in conversational systems and intelligent assistants, where independent components perform complementary tasks during the interaction process [7, 1, 2].

This modular perspective offers several practical advantages. Each stage can be modified independently without affecting the remaining components of the workflow. Likewise, additional validation stages may be incorporated as required by future applications without redesigning the complete generation process.

2.2. Specialized Agents

The framework proposed in this work consists of three lightweight collaborative agents that sequentially cooperate to generate and validate synthetic tourism reviews.

The first agent, referred to as the *Generator Agent*, receives the metadata describing the desired sentiment polarity, attraction type, municipality, and approximate review length. Using this information, the agent produces an initial tourism review whose objective is to satisfy the requested characteristics while maintaining a natural writing style.

The second component, called the *Reviewer Agent*, performs a linguistic inspection of the generated review. This stage evaluates structural properties such as sentence completeness, writing fluency, readability, and formatting consistency. Reviews containing incomplete ideas, grammatical inconsistencies, or structural problems are marked for regeneration before continuing through the workflow.

Finally, the *Validator Agent* verifies that the generated review satisfies the metadata originally requested by the user. This verification includes checking the expected sentiment polarity, attraction category, language consistency, and minimum length requirements before incorporating the review into the final corpus, following the evaluation criteria established by the REST-MEX 2026 shared task [17].

Unlike autonomous multi-agent systems, where agents cooperate through complex planning, negotiation, or distributed decision-making mechanisms [19, 20, 21, 22], the proposed framework adopts a simple sequential collaboration strategy. Every agent performs a single well-defined task and forwards the resulting review to the next stage of the workflow. This design considerably simplifies implementation while preserving the flexibility required to integrate different Large Language Models or validation modules.

2.3. Design Principles

The proposed framework was designed according to four general principles. First, each agent is responsible for only one task, reducing implementation complexity and facilitating independent improvements. Second, communication among agents is performed exclusively through textual information, making the framework independent of the underlying language model. Third, validation is carried out incrementally, allowing potential inconsistencies to be detected before the review becomes part of the final synthetic corpus. Finally, the complete architecture remains modular, enabling individual agents to be replaced or extended without modifying the remaining stages of the generation process.

These principles facilitate the construction of reproducible synthetic datasets while maintaining a lightweight architecture suitable for practical deployment within the REST-MEX 2026 shared task [17, 18].

3. Proposed Multi-Agent Framework

The proposed framework decomposes the synthetic review generation process into three sequential and independent agents. Each agent performs a single specialized task and transfers its output to the next stage. This collaborative strategy simplifies the generation workflow while allowing each component to focus on a well-defined objective.

Let $\mathcal{A} = \{A_G, A_R, A_V\}$, where

- A_G denotes the *Generator Agent*,
- A_R denotes the *Reviewer Agent*,
- A_V denotes the *Validator Agent*.

The communication among agents follows the sequential workflow

$$A_G \rightarrow A_R \rightarrow A_V. \quad (1)$$

Only reviews successfully validated by all agents are incorporated into the final synthetic corpus.

3.1. Generator Agent

The Generator Agent is responsible for producing the initial tourism review.

For each review, the agent receives the following metadata:

- Sentiment polarity.
- Attraction type.
- Municipality.
- Desired review length.

Using this information, the Large Language Model generates a tourism review describing a realistic visitor experience while satisfying the requested characteristics.

The generated review is then forwarded to the Reviewer Agent.

3.2. Reviewer Agent

The Reviewer Agent evaluates the linguistic quality of the generated review.

Rather than modifying the content, this agent verifies that the generated text satisfies a set of basic writing requirements.

The inspection includes:

- Sentence completeness.
- Writing fluency.
- Formatting consistency.
- Readability.
- Presence of coherent tourism descriptions.

If obvious inconsistencies are detected, the review is rejected and a new generation request is issued to the Generator Agent.

Otherwise, the review is forwarded to the Validator Agent.

4. Experimental Evaluation

4.1. Experimental Setup

The proposed multi-agent framework was evaluated using the official REST-MEX 2026 training corpus. Review metadata describing sentiment polarity, attraction type, municipality, and target review length were used as inputs for the Generator Agent.

Each synthetic review was sequentially processed by the Reviewer Agent and the Validator Agent before being incorporated into the final corpus. Reviews rejected during any validation stage were regenerated until all quality requirements were satisfied.

Following the REST-MEX 2026 guidelines, the final corpus contained 3,000 synthetic reviews uniformly distributed across the five sentiment levels and the three attraction categories.

5. Results

Table 1 summarizes the evaluation results obtained using the synthetic corpus generated by the proposed framework.

Table 1

Evaluation results obtained using the generated corpus.

Metric	Value
Track Score	0.518
Macro Precision	0.544
Macro Recall	0.541
Macro F1-score	0.542

The generated corpus satisfied the evaluation requirements established by the REST-MEX 2026 shared task. The sequential collaboration among agents allowed every generated review to satisfy the requested metadata while maintaining coherent tourism-related descriptions.

The generated reviews present balanced sentiment distributions and realistic descriptions of hotels, restaurants, and tourist attractions. Manual inspection of randomly selected samples indicated that the proposed workflow preserved coherent narratives and natural tourism-related vocabulary.

6. Conclusions

This paper presented a simple multi-agent framework for synthetic tourism review generation designed for the REST-MEX 2026 shared task. The proposed methodology decomposes the generation process into three specialized agents responsible for content generation, linguistic review, and final validation. This sequential collaboration simplifies the construction of balanced synthetic corpora while maintaining a modular and reproducible workflow.

The resulting framework is independent of the underlying Large Language Model and can therefore be adapted to different instruction-tuned architectures with minimal modifications. Furthermore, its modular design facilitates the incorporation of additional agents or validation stages according to the requirements of future applications.

Declaration on Generative AI

Generative AI tools were used solely to improve the language, grammar, and readability of the manuscript. The authors reviewed and edited all AI-assisted content and take full responsibility for the final version of the manuscript.

References

- [1] M. Á. Álvarez-Carmona, R. Aranda, A. Y. Rodríguez-Gonzalez, D. Fajardo-Delgado, M. G. Sánchez, H. Pérez-Espinosa, J. Martínez-Miranda, R. Guerrero-Rodríguez, L. Bustio-Martínez, Á. Díaz-Pacheco, Natural language processing applied to tourism research: A systematic review and future research directions, *Journal of king Saud university-computer and information sciences* 34 (2022) 10125–10144.
- [2] A. Diaz-Pacheco, M. Á. Álvarez-Carmona, R. Guerrero-Rodríguez, L. A. C. Chávez, A. Y. Rodríguez-González, J. P. Ramírez-Silva, R. Aranda, Artificial intelligence methods to support the research of destination image in tourism. a systematic review, *Journal of Experimental & Theoretical Artificial Intelligence* 36 (2024) 1415–1445.
- [3] R. Guerrero-Rodríguez, M. Á. Álvarez-Carmona, R. Aranda, A. P. López-Monroy, Studying online travel reviews related to tourist attractions using nlp methods: the case of guanajuato, mexico, *Current issues in tourism* 26 (2023) 289–304.
- [4] Á. Díaz-Pacheco, R. Guerrero-Rodríguez, M. Á. Álvarez-Carmona, A. Y. Rodríguez-González, R. Aranda, A comprehensive deep learning approach for topic discovering and sentiment analysis of textual information in tourism, *Journal of King Saud University-Computer and Information Sciences* 35 (2023) 101746.
- [5] Á. Díaz-Pacheco, R. Guerrero-Rodríguez, M. Á. Álvarez-Carmona, A. Y. Rodríguez-González, R. Aranda, Quantifying differences between ugc and dmo's image content on instagram using deep learning, *Information Technology & Tourism* 26 (2024) 293–329.
- [6] A. Diaz-Pacheco, M. A. Álvarez-Carmona, A. Y. Rodríguez-González, H. Carlos, R. Aranda, Measuring the difference between pictures from controlled and uncontrolled sources to promote a destination. a deep learning approach (2023).
- [7] E. P. Ramirez-Villaseñor, H. Pérez-Espinosa, M. A. Álvarez-Carmona, R. Aranda, Design, development, and evaluation of a chatbot for hospitality services assistance in spanish, *Acta universitaria* 33 (2023).
- [8] E. Olmos-Martínez, M. Á. Álvarez-Carmona, R. Aranda, A. Díaz-Pacheco, What does the media tell us about a destination? the cancan case, seen from the usa, canada, and mexico, *International Journal of Tourism Cities* 10 (2024) 639–661.
- [9] I. Castillo-Ortiz, M. Á. Álvarez-Carmona, R. Aranda, Á. Díaz-Pacheco, Evaluating culinary skill transfer: A deep learning approach to comparing student and chef dishes using image analysis, *International Journal of Gastronomy and Food Science* 38 (2024) 101070.
- [10] R. Guerrero-Rodríguez, M. A. Álvarez-Carmona, R. Aranda, Á. Díaz-Pacheco, Ai-driven analysis of hotel digital reputation: insights from a mexican destination, *International Journal of Tourism Cities* 12 (2026) 19–36.
- [11] M. Á. Álvarez-Carmona, R. Aranda, R. Guerrero-Rodríguez, A. Y. Rodríguez-González, A. P. López-Monroy, A combination of sentiment analysis systems for the study of online travel reviews: Many heads are better than one, *Computación y Sistemas* 26 (2022). doi:<https://doi.org/10.13053/CyS-26-2-4055>.
- [12] M. Á. Álvarez-Carmona, R. Aranda, S. Arce-Cárdenas, D. Fajardo-Delgado, R. Guerrero-Rodríguez, A. P. López-Monroy, J. Martínez-Miranda, H. Pérez-Espinosa, A. Rodríguez-González, Overview of rest-mex at iberlef 2021: Recommendation system for text mexican tourism, *Procesamiento del Lenguaje Natural* 67 (2021). doi:<https://doi.org/10.26342/2021-67-14>.
- [13] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, D. Fajardo-Delgado, R. Guerrero-Rodríguez, L. Bustio-Martínez, Overview of rest-mex at iberlef 2022: Recommendation system, sentiment analysis and covid semaphore prediction for mexican tourist texts, *Procesamiento del Lenguaje Natural* 69 (2022).
- [14] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, L. Bustio-Martínez, V. Muñis-Sánchez, A. P. Pastor-López, F. Sánchez-Vega, Overview of rest-mex at iberlef 2023: Research on sentiment analysis task for mexican tourist texts, *Procesamiento del Lenguaje Natural* 71 (2023).

- [15] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, A. Y. Rodríguez-González, L. Bustio-Martínez, V. Herrera-Semenets, Overview of rest-mex at iberlef 2025: Researching sentiment evaluation in text for mexican magical towns, volume 75, 2025.
- [16] A. Díaz-Pacheco, A. B. García-Gutiérrez, A. Y. Rodríguez-González, R. Aranda, M. Á. Álvarez-Carmona, Balancing sentiment analysis datasets through representative-word-guided synthetic review generation: A case study on mexican spanish tourism reviews, *Applied Sciences* 16 (2026) 7398.
- [17] M. Á. Álvarez-Carmona, Á. Díaz-Pacheco, R. Aranda, M. Reyes-Sierra, A. Y. Rodríguez-González, L. Bustio-Martínez, V. Herrera-Semenets, Overview of rest-mex at iberlef 2026: Research on synthetic text modeling for mexican magical towns, *Procesamiento del Lenguaje Natural* 77 (2026).
- [18] A. Bonet-Jover, J. Á. González-Barba, L. Chiruzzo, Overview of IberLEF 2026: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2026)*, co-located with the 42nd Conference of the Spanish Society for Natural Language Processing (SEPLN 2026), CEUR-WS. org, 2026.
- [19] M. Wooldridge, N. R. Jennings, *Intelligent agents: Theory and practice*, *The Knowledge Engineering Review* 10 (1995) 115–152.
- [20] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, 4 ed., Pearson, 2021.
- [21] J. S. Park, J. C. O'Brien, C. J. Cai, M. R. Morris, P. Liang, M. S. Bernstein, Generative agents: Interactive simulacra of human behavior, in: *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST)*, 2023.
- [22] L. Wang, et al., A survey on large language model based multi-agent systems, *arXiv preprint arXiv:2502.01657* (2025).