# Labadain Chat: A Conversational Agent for the Tetun Language

Gabriel de Jesus
INESC TEC
Porto, Portugal
gabriel.jesus@inesctec.pt

Sérgio Nunes
INESC TEC / University of Porto
Porto, Portugal
sergio.nunes@fe.up.pt

## Abstract

Large language model (LLM)-based conversational assistants are designed for general-purpose conversation tasks and are primarily optimized for high-resource languages. Although these systems support some low-resource languages (LRLs), their responses often fall short of user expectations. Consequently, speakers of LRLs remain marginalized and unable to fully benefit from advances in LLMs. These challenges underscore the need for targeted, language-specific solutions that can effectively serve underrepresented language communities. This study presents Labadain Chat, a conversational agent for Tetun, a low-resource language spoken by over 932,000 people in Timor-Leste. We adapt existing LLMs to Tetun using language-specific prompting strategies and report on the system's architecture, features, applications, and utility for the Tetun-speaking community. Results from the user study show a high task success rate for Labadain Chat (91%, with substantial inter-annotator agreement, Cohen's $\kappa = 0.67$) and high user satisfaction (4.30 out of 5, with Cohen's *weighted* $\kappa = 0.75$), demonstrating the effectiveness of language-specific LLM customization for Tetun. Overall, this study provides a practical pathway toward promoting equitable access to AI-powered information services for the Tetun-speaking community and suggests an adaptable methodology that can be applied to other under-resourced languages in similar contexts. The system is publicly available at https://www.labadain.com, with mobile applications for both iOS and Android.

## CCS Concepts

• **Information systems** → **Information retrieval**; • **Artificial intelligence** → *Natural language processing*.

## Keywords

Conversational Search, Large Language Model, Information Retrieval, Low-Resource Language, Tetun Language.

**ACM Reference Format:**
Gabriel de Jesus and Sérgio Nunes. 2026. Labadain Chat: A Conversational Agent for the Tetun Language. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26), July 20–24, 2026, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3805712.3808372




*SIGIR '26, Melbourne, VIC, Australia*

ACM ISBN 979-8-4007-2599-9/2026/07
https://doi.org/10.1145/3805712.3808372

## 1 Introduction

Large language model (LLM)-powered conversational assistants, such as ChatGPT, Gemini, Claude, and Perplexity, have fundamentally reshaped information access, enabling tasks such as question answering, summarization, multi-step reasoning, and complex problem solving [1, 28]. Although these systems improve access to information for many users, they continue to present significant challenges for speakers of low-resource languages (LRLs) due to inconsistent performance, data scarcity, and linguistic bias [3, 18]. Consequently, these communities remain marginalized, emphasizing the need for targeted, language-specific solutions to enable equitable access to information.

Tetun-speaking users face similar challenges, as LLM-based assistants provide limited support for the language, restricting users' ability to access information and benefit from AI-powered services. To address these gaps, we present **Labadain Chat**, a conversational agent developed for Tetun speakers. The system implements a practical, language-specific prompting strategy that adapts existing LLMs to Tetun within an agentic conversational framework. The agent's workflow is illustrated in Figure 1.

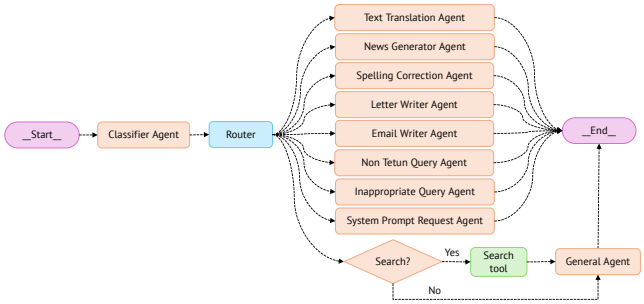


**Figure 1: Workflow of the Labadain Chat Agent.**

To design effective system prompts, we account for key characteristics of the Tetun language, including its extensive use of loanwords. Prior studies indicate that up to 40% of Tetun verbs, nouns, and adjectives are Portuguese loanwords [12, 14, 24], and their presence in web documents has steadily increased [7]. To guide the LLM in producing accurate and coherent responses, we incorporate: (1) Portuguese to Tetun word conversion rules established by the *Instituto Nacional de Linguística* (INL), the official body responsible for Tetun standardization [20, 21]; (2) Examples of official Tetun usage texts; and (3) A pair of curated lexical lists to enrich vocabulary coverage.

We compare three end-to-end systems: the agentic Labadain Chat, the earlier version of Labadain Chat (referred to as Labadain Old) [5], and ChatGPT.[1] Both Labadain systems provide Tetun

[1] https://chatgpt.com

information access, but they differ in interaction paradigm and orchestration strategy, whereas ChatGPT is a general-purpose conversational assistant that accounts for the majority of AI chatbot web traffic in Timor-Leste (approximately 78% market share).[2] We report Labadain Chat utility and perceived usefulness specifically for Tetun-speaking users.

## 2 Background and Related Work

Labadain Chat builds upon Labadain Old, which was launched in early 2024 [5]. The two systems differ substantially in architecture and underlying methodology. While Labadain Old is an LLM-based conversational assistant, the current system is designed as an agentic AI. Both systems support conversation in Tetun, one of Timor-Leste's official languages and a LRL spoken by over 932,000 people in Timor-Leste [4, 25]. Labadain Old was primarily developed to study the search behavior of Tetun-speaking users [16].

Conversational systems have been extensively studied in natural language processing research, including task-oriented dialogue systems, open-domain chatbots, and LLM-powered conversational agents [11, 13, 15]. More recently, research has shifted toward agentic and modular dialogue architectures to improve robustness and multi-step reasoning in complex interactions [22, 27].

However, prior studies have primarily focused on high-resource languages, leaving LRLs underrepresented due to data scarcity, limited evaluation benchmarks, and linguistic diversity [17, 18]. This is particularly challenging, as LRLs often have complex scripts and limited lexical resources for developing effective language technologies. For Tetun, these challenges are compounded by orthographic variation, widespread loanword usage, and limited digital resources, emphasizing the need for tailored system designs.

Several AI-powered tools exist for LRLs in open-domain conversational settings, such as AfricanGPT,[3] which supports more than 20 African languages, Reobot[4] for the Māori language, and Macsen[5] for the Welsh language. However, there is limited publicly available documentation on the design and development of these systems. To the best of our knowledge, Labadain Chat is among the first conversational systems developed for LRLs. Our work documents the end-to-end system design and implementation, demonstrating how LLMs can be effectively adapted for an under-resourced language.

Regarding evaluation, user-centered studies typically emphasize *task success* (whether a system helps users achieve their goals) and *user satisfaction* (subjective perception of usefulness and interaction quality) [11, 23, 26]. Following these principles, we compare task success and user satisfaction across the three systems—Labadain Chat, Labadain Old, and ChatGPT—to assess both the system's objective utility and perceived usefulness for Tetun-speaking users.

## 3 System Architecture and Agent Workflow

The high-level system architecture of Labadain Chat and its agent components and workflow are detailed in the following subsections.

### 3.1 High-Level System Architecture

The high-level system architecture of Labadain Chat is illustrated in Figure 2. The system is organized into four logical layers and an external service layer that provides access to LLMs and a web-based search tool. These layers are described as follows:

- **Client Layer** includes a web interface developed using Next.js[6] and mobile applications built with Flutter.[7] Both clients communicate with the same backend API to ensure consistent functionality and behavior across platforms.
- **API Layer** is implemented using FastAPI[8] and handles HTTP requests, input validation, and RESTful operations to manage data. It also manages server-sent events (SSEs) to stream conversational responses from the agent, and user authentication and registration via OAuth or email.
- **Agent Layer** comprises LangChain, LangGraph, and LangSmith.[9] LangChain is used for prompt orchestration and tool integration, LangGraph for constructing the agent as a stateful graph, and LangSmith for monitoring the agent's performance. Conversational memory is managed explicitly through the LangGraph state to keep multi-node workflows debuggable and easier to tune.
- **Data Layer** uses a PostgreSQL database to store user accounts, conversation histories, and interaction logs.
- **External LLMs / Tools** provide access to LLMs (Gemini 2.5 Pro and Flash, and Claude 3 Haiku) and the Google Search API, which support the agents' reasoning and response generation. Each API is accessed via standard RESTful endpoints with authenticated API keys.

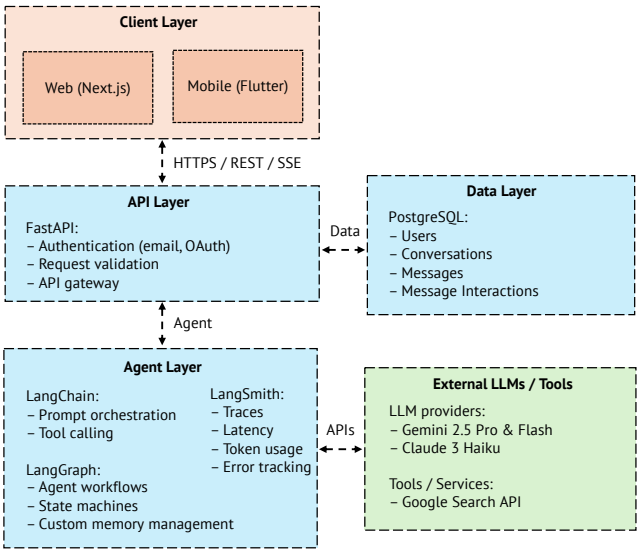


**Figure 2: General Architecture of Labadain Chat.**

When a user submits a query from the client layer, it is forwarded to the API layer, where it is handled by FastAPI and passed to the

[2] https://gs.statcounter.com/ai-chatbot-market-share/all/timor-leste
[3] https://www.africangpt.org
[4] https://reobot.co.nz/
[5] https://techiaith.cymru/index.html-p-1151-and-lang=en.html
[6] https://nextjs.org
[7] https://flutter.dev
[8] https://fastapi.tiangolo.com
[9] https://www.langchain.com

agent layer. The agent processes the request and, when necessary, performs an external web search to retrieve additional context before generating a response using an LLM. The response is then returned to the API layer and streamed back to the user. Both the user query and the generated response are stored in the database.

### 3.2 Agent Components and Workflow

Labadain Chat consists of eleven agents and a web-search tool, as illustrated in Figure 1. The system comprises eight specialized agents for different conversational tasks and three exception-handling agents designed to manage system prompt requests, inputs in non-Tetun languages, and inappropriate content.

The agent design is informed by insights from the interaction logs of Labadain Old [5] and prior analyses of Tetun-speaking users' search behavior [16]. Based on these insights, eight specialized agents were designed: *(1) Classifier Agent* identifies the intent of the user query; *(2) Translation Agent* handles text translation requests; *(3) News Generator Agent* produces structured news articles from a user-provided draft; *(4) Spelling Correction Agent* performs spelling correction and text normalization; *(5) Letter Writer Agent* generates formal letters according to user specifications; *(6) Email Writer Agent* composes professional email correspondence; *(7) Search Decision Agent* determines whether a web search is required for the user query; *(8) General Agent* processes requests that fall outside the scope of all other agents.

When a user request is received, the *Classifier Agent* first determines the user's intent and automatically routes the input to the appropriate specialized agent. For general conversations, the system evaluates whether a web search is necessary. This process uses a hybrid approach, combining rule-based logic with the *Search Decision Agent*. If a search is performed, the retrieved results are provided as context for the *General Agent* to generate a context-aware response, with corresponding sources included only when the system uses them as part of its response.

## 4 Implementation

The implementation of Labadain Chat includes prompt engineering, backend API development, database design, and the development of web and mobile client applications. Each component is described in detail in the following subsections.

### 4.1 Prompt Engineering

In prompt engineering, explicit instructions were designed for each agent. In most cases, agents are instructed to act as Tetun linguists, and all responses are strictly required to be in Tetun, except for translation tasks. For the three exception-handling agents, we provide pre-defined responses in Tetun. For the eight specialized agents (except *Search Decision Agent*), each LLM call includes at least one guiding example, along with Portuguese–Tetun transformation rules and a curated lexical list to ensure linguistically accurate and contextually coherent outputs. The system prompts used to instruct the *General Agent* are presented in Prompt 4.1.

Since everyday Tetun writing practices, particularly in web documents, commonly fail to adhere to standard conventions [7, 8], incorporating examples of official Tetun usage helps ensure the correct application of diacritics and hyphenation for monosemantic compound words. Including these examples enhances the model's knowledge of accurate Tetun grammar per INL standardization.

To further support linguistic accuracy, we incorporated examples of Portuguese loanwords used in Tetun with their corresponding conversion rules documented in the INL standard orthography for Tetun [21]. The rules are implemented as a set of explicit character-level transformation rules (e.g., *nh*→*ñ*, *lh*→*ll*) for handling Portuguese-derived words in Tetun.

Additionally, we construct a curated lexical list containing 33 native Tetun words that differ substantially from Portuguese and English, for which LLMs often fail to infer correct meanings through generalization. This list is periodically updated based on observed model outputs. Example entries (English, Portuguese, **Tetun**) such as (rice, arroz, **etu**) and (milk, leite, **susubeen**).

**Prompt 4.1: Details of the General Agent Prompt.**

You are Labadain, a Tetun linguist expert and helpful assistant. Your task is to assist the user with their request in Tetun, providing clear and accurate information. Assume the question is in Tetun and respond using accurate Tetun grammar.

The output must follow these examples of accurate official Tetun writing:
`{tetun_official_example}`

To improve Tetun accuracy, follow the rules for converting Portuguese-derived characters into Tetun below:
`{pt_loanwords_tetun_rules}`

Incorporate the provided vocabulary pairs to improve your Tetun orthography:
`{vocabulary_enrich_list}`

### 4.2 LLM Selection and Usage

Labadain Chat uses the Claude 3 Haiku model from Anthropic[10] and the Gemini 2.5 Pro and Flash models from Google.[11] Claude 3 Haiku is selected for its strong performance on Tetun, cost-effectiveness, prior use in Labadain Old [5], and in several Tetun-focused studies [6, 10]. Gemini 2.5 Pro and Flash are selected based on a preliminary quality assessment conducted across multiple LLMs.

Given that Gemini 2.5 Pro is designed for complex reasoning tasks and demonstrates stronger performance on Tetun than the Flash variant and Claude 3 Haiku, it is assigned to agents handling higher-complexity tasks, including the *Classifier*, *Translation*, *News Generator*, *Spelling Correction*, *Email Writer*, and *Letter Writer* agents. The Gemini 2.5 Flash variant is used for the *Search Decision* and *General Conversation* agents, while Claude 3 Haiku is employed for the exception-handling agents.

### 4.3 System Features

Labadain Chat offers core features such as user registration, chat interaction, file upload, web search, and access to chat history. The

[10] https://www.anthropic.com
[11] https://deepmind.google/models/gemini/

system interface is shown in Figure 3, illustrating the home page and an example conversation view.[12] Authentication is required to access Labadain Chat. Users may register or log in via email or OAuth (Google or Apple). Email registration requires verification through a confirmation email containing a validation token.

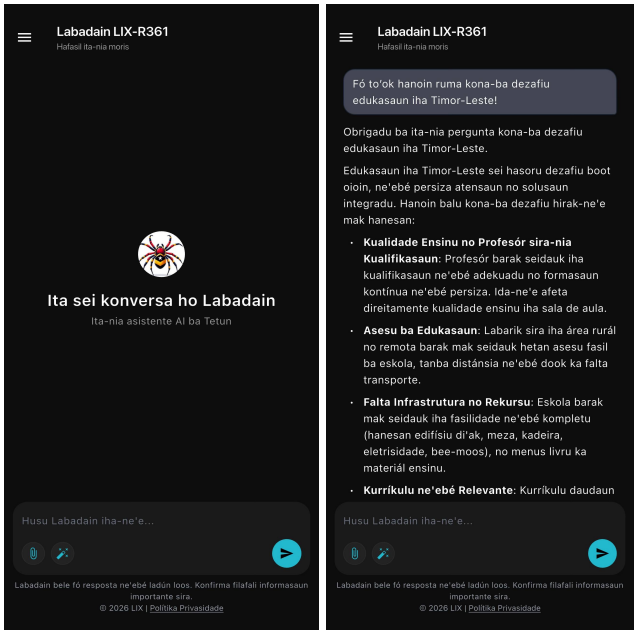


**Figure 3: Interface of Labadain Chat.**

The system allows users to attach supplementary information via file upload to support tasks such as translation, orthography correction, and structured news article generation. Supported file types are limited to Word, Excel, and PDF, with a maximum size of 256 KB, and only one file can be uploaded at a time. The uploaded file is processed and incorporated directly into the user prompt.

Users can also enable the web search feature to force the system to perform a web search using the Google Search API, retrieving the top-10 ranked documents and incorporating them into the user prompt as contextual information for general conversational queries. Each retrieved document includes a title, snippet, and URL, and the original language of the content is preserved. When this context is used to generate a response, the corresponding external source link is also provided.

For each new conversation, the system creates a new topic that stores the full conversation history. For the *General Agent*, only the most recent turn is used as context (i.e., a $k$=1 memory window). This memory mechanism is exclusive to the *General Agent* and is not applied to other agents due to operational cost constraints. Users can also delete their conversation history.

## 5 User Study

To assess Labadain Chat's objective utility and perceived usefulness for Tetun-speaking users, we compared it against Labadain Old [5] and ChatGPT. We collected over 8,000 query log entries from Labadain Chat (2,101 entries from January 1–27, 2026) and Labadain Old (6,210 entries drawn from daily log files within three randomly selected months in 2025). From this pool, three native Tetun-speaking annotators independently selected candidate queries and applied spelling corrections following the INL standard, then jointly curated a final set of 60 queries from 180 candidates based on local information needs, topical diversity, and query complexity.

The queries cover three categories: **fact-seeking queries** (direct information requests), **multi-turn queries** limited to two turns (tasks requiring interaction across two conversation turns), and **ambiguous queries** (under-specified requests requiring clarification), with 20 queries per category. The same annotators assessed the systems' responses, with each query reviewed by two annotators to ensure reliability. Responses were precompiled, and system identities were anonymized to avoid bias. Annotators rated task success (whether the system helped achieve the user's goal) and user satisfaction (perceived usefulness and interaction quality).

Queries were distributed among annotators to ensure coverage while keeping individual workload manageable. Annotators evaluated: **(1) Task Success**: a binary judgment (1 if the system provided a correct or sufficiently informative answer, or asked an appropriate clarification for ambiguous queries; 0 otherwise); **(2) User Satisfaction:** a Likert scale from 1 to 5 (1 = very dissatisfied, 5 = very satisfied), reflecting perceived usefulness and interaction quality.

Inter-annotator agreement was measured using Cohen's $\kappa$ [2], with agreement strength interpreted using the scale of Landis and Koch [19]. The assessment results are summarized in Table 1.

**Table 1: Task Success Rate and User Satisfaction for Labadain Chat (LBC), Labadain Old (LBO), and ChatGPT (CGPT).**

| Query Type | Task Success Rate | | | User Satisfaction | | |
|---|---|---|---|---|---|---|
| | LBC | LBO | CGPT | LBC | LBO | CGPT |
| Fact-seeking | 95% | 73% | 48% | 4.60 | 3.35 | 2.60 |
| Multi-turn | 88% | 65% | 58% | 4.05 | 2.75 | 2.80 |
| Ambiguous | 90% | 58% | 53% | 4.25 | 3.10 | 3.10 |
| *Overall* | **91%** | 65% | 53% | **4.30** | 3.07 | 2.83 |

Table 1 shows that Labadain Chat (LBC) outperforms Labadain Old (LBO) and ChatGPT (CGPT) across all query types. LBC achieves high task success and user satisfaction for fact-seeking, multi-turn, and ambiguous queries, demonstrating its advantage over general-purpose open-domain assistants like ChatGPT. ChatGPT underperforms compared to both LBC and LBO due to linguistic biases (e.g., generating non-Tetun responses), inconsistent performance, and lower quality of Tetun text, which limit its usefulness for Tetun-speaking users.

Overall, LBC reaches a 91% task success rate and 4.30/5 user satisfaction, substantially outperforming LBO (65%, 3.07) and CGPT (53%, 2.83). Inter-annotator agreement is substantial, with Cohen's $\kappa = 0.67$ for task success and Cohen's *weighted* $\kappa = 0.75$ for user satisfaction. The observed gains suggest that language-specific LLM customization can enhance both effectiveness and user experience for Tetun-speaking users, particularly when incorporating linguistic knowledge such as orthographic normalization and loanword transformation rules. These gains may be partly attributed to the

[12] Copilot was used to assist in the development of some client-side code, including UI components. All code was reviewed by the authors.

model's knowledge of Portuguese and the extensive use of Portuguese loanwords in Tetun.

The remaining 9% of failures are primarily attributed to several factors, including the retrieval of irrelevant documents, incorrect search decisions (e.g., when search is required but not triggered, leading to ungrounded responses), and hallucinations in ambiguous queries. Failures appear more frequently for queries related to local contexts where limited relevant information is available on the web. For reproducibility, we make the query logs, system responses, and annotation results publicly available [9].

## 6 Conclusion and Future Work

We presented Labadain Chat, an agentic conversational system for Tetun speakers that builds on Labadain Old. The system demonstrates a practical application of agentic AI for open-domain interaction in Tetun, supporting tasks such as text translation, structured news generation, and general conversation.

This study highlights the benefits of language-specific LLM customization for Tetun, leveraging the extensive use of Portuguese loanwords and curated lexical resources to improve the quality of LLM-generated responses. Results from the user study show that this approach achieves high task success and user satisfaction, contributing to more accessible AI-powered information services for the Tetun-speaking community. More broadly, the findings suggest a practical pathway toward developing inclusive conversational systems for under-resourced languages, with potential applicability to other LRLs in similar contexts.

Future work will focus on fine-tuning open-source models for Tetun to improve cultural awareness and support more sustainable system deployment. Such approaches may also enable the generation of more accurate, relevant, and contextually appropriate responses for Tetun-speaking users.

## 7 Limitations

Labadain Chat relies on the Google Search API for retrieval and uses document titles, snippets, and URLs as contextual input for the *General Agent*, which may limit the completeness and depth of the information presented. Additionally, the *General Agent* employs limited conversational memory ($k = 1$) due to cost constraints, which may affect response coherence in extended multi-turn interactions.

The system also depends on commercial LLM APIs, which may influence sustainability as service conditions evolve. While the language-specific prompting strategies may remain applicable when transitioning to alternative models, some prompt adaptation may be required to maintain performance under different model behaviors.

The assessment is based on a small-scale, user-centered study with 60 queries emphasizing end-to-end task success and user satisfaction. While this demonstrates practical utility, it is limited to the *General Agent* and does not cover other agents due to the labor-intensive and costly nature of the evaluation process.

## Acknowledgments

This work is funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the support UID/50014/2025 (https://doi.org/10.54499/UID/50014/2025). We acknowledge Fanézio Pinto and Altedio Araújo for their contributions to dataset annotation. The authors also acknowledge the people of the Woi Wurrung and Boon Wurrung language groups of the eastern Kulin Nation on whose unceded lands ACM SIGIR 2026 was hosted. We pay our respects to their Elders past and present, and extend that respect to all Aboriginal and Torres Strait Islander peoples today and their continuing connection to land, sea, sky, and community.

## References

[1] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S. Yu, Qiang Yang, and Xing Xie. 2024. A Survey on Evaluation of Large Language Models. *ACM Trans. Intell. Syst. Technol.* 15, 3, Article 39 (March 2024), 45 pages. doi:10.1145/3641289

[2] Jacob Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20 (1960), 37–46. doi:10.1177/001316446002000104

[3] Cohere for AI. 2024. *The AI Language Gap: Considerations on the Multilingual Capabilities of AI Language Models.* Policy Primer. Cohere. https://cohere.com/research/papers/the-ai-language-gap.pdf

[4] Instituto Nacional de Estatística Timor-Leste (INETL). 2015. Census 2015 Priority Table Population By Language. https://inetl-ip.gov.tl/2023/03/09/census-2015-priority-table-population-by-language/ Accessed on January 22, 2026.

[5] Gabriel de Jesus. 2024. Labadain Chat Old: a conversational system for Tetun speakers. URL https://old.labadain.com.

[6] Gabriel de Jesus and Sérgio Nunes. 2024. Exploring Large Language Models for Relevance Judgments in Tetun. In *Proceedings of The First Workshop on Large Language Models for Evaluation in Information Retrieval (LLM4Eval 2024), co-located with the 10th International Conference on Online Publishing (SIGIR 2024)*, C. Siro, M. Aliannejadi, H.A. Rahmani, N. Craswell, C.L.A. Clarke, G. Faggioli, B. Mitra, P. Thomas, and E. Yilmaz (Eds.), Vol. 3752. Washington D.C., USA, 19–30. https://ceur-ws.org/Vol-3752/

[7] Gabriel de Jesus and Sérgio Nunes. 2024. Labadain-30k+: A Monolingual Tetun Document-Level Audited Dataset. In *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages @ LREC-COLING 2024*, Maite Melero, Sakriani Sakti, and Claudia Soria (Eds.). ELRA and ICCL, Torino, Italia, 177–188. https://aclanthology.org/2024.sigul-1.22

[8] Gabriel de Jesus and Sérgio Nunes. 2025. Establishing a Foundation for Tetun Ad-Hoc Text Retrieval: Stemming, Indexing, Retrieval, and Ranking. arXiv:2412.11758 [cs.IR] https://arxiv.org/abs/2412.11758

[9] Gabriel de Jesus and Sérgio Nunes. 2026. LabadainLog-60: A Curated Tetun Query-Response Dataset for Conversational System Evaluation [Dataset]. INESC TEC. doi:10.25747/6FM1-V130

[10] Gabriel de Jesus, Siddharth A.K. Singh, Sérgio Nunes, and Andrew Yates. 2025. Zero-Shot and Hybrid Strategies for Tetun Ad-Hoc Text Retrieval. In *Proceedings of the 2025 International ACM SIGIR Conference on Innovative Concepts and Theories in Information Retrieval (ICTIR) (ICTIR '25).* Association for Computing Machinery, New York, NY, USA, 264–274. doi:10.1145/3731120.3744593

[11] Jan Deriu, Álvaro Rodrigo, Arantxa Otegi, Guillermo Echegoyen, Sophie Rosset, Eneko Agirre, and Mark Cieliebak. 2021. Survey on Evaluation Methods for Dialogue Systems. *Artificial Intelligence Review* 54, 1 (2021), 755–810. doi:10.1007/s10462-020-09866-x

[12] Zuzana Greksáková. 2018. *Tetun in Timor-Leste: The role of language contact in its development.* Ph. D. Dissertation. Universidade de Coimbra, Portugal. http://hdl.handle.net/10316/80665

[13] Shengyue Guan, Jindong Wang, Jiang Bian, Bin Zhu, Jian guang Lou, and Haoyi Xiong. 2026. Evaluating LLM-based Agents for Multi-Turn Conversations: A Survey. arXiv:2503.22458 [cs.CL] https://arxiv.org/abs/2503.22458

[14] John Hajek and Catharina Williams van Klinken. 2019. Language Contact and Gender in Tetun Dili: What Happens When Austronesian Meets Romance? *Oceanic Linguistics* 58 (June 2019), 59–91. doi:10.1353/ol.2019.0003

[15] Islam A. Hassan and Yvette Graham. 2024. Advancing Open-Domain Conversational Agents - Designing an Engaging System for Natural Multi-Turn Dialogue. In *Proceedings of the 1st Workshop on Simulating Conversational Intelligence in Chat (SCI-CHAT 2024)*, Yvette Graham, Qun Liu, Gerasimos Lampouras, Ignacio Iacobacci, Sinead Madden, Haider Khalid, and Rameez Qureshi (Eds.). Association for Computational Linguistics, St. Julians, Malta, 75–79. https://aclanthology.org/2024.scichat-1.8/

[16] Gabriel de Jesus and Sérgio Nunes. 2025. Insights into LLM-Based Conversational Search: A Study of Tetun-Speaking Users' Search Behavior. In *Proceedings of the 2025 International ACM SIGIR Conference on Innovative Concepts and Theories in Information Retrieval (ICTIR)* (Padua, Italy) *(ICTIR '25).* Association for Computing Machinery, New York, NY, USA, 297–306. doi:10.1145/3731120.3744596

[17] Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. 2020. The State and Fate of Linguistic Diversity and Inclusion in the NLP World.

In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel R. Tetreault (Eds.). Association for Computational Linguistics, 6282–6293. doi:10.18653/V1/2020.ACL-MAIN.560

[18] Nir Kshetri. 2024. Linguistic Challenges in Generative Artificial Intelligence: Implications for Low-Resource Languages in the Developing World. *Journal of Global Information Technology Management* 27, 2 (2024), 95–99. doi:10.1080/1097198X.2024.2341496

[19] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174. doi:10.2307/2529310

[20] National Institute of Linguistics (INL). 2004. *The Standard Orthography of the Tetum Language: 115 Years in the Making*. Directorate of the National Institute of Linguistics, Dili, Timor-Leste. https://archive.org/details/the-standard-orthography-of-the-tetum-language

[21] Democratic Republic of Timor Leste. 2004. The Standard Orthography of the Tetun Language. Government Decree-Law No. 1/2004 of 14 April 2004. http://mj.gov.tl/jornal/lawsTL/RDTL-Law/RDTL-Gov-Decrees/Gov-Decree-2004-01.pdf Accessed on November 8, 2024.

[22] Timo Schick, Jane Dwivedi-Yu, Roberto Dessí, Roberta Raileanu, Maria Lomeli, Eric Hambro, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2023. Toolformer: language models can teach themselves to use tools. In *Proceedings of the 37th International Conference on Neural Information Processing Systems* (New Orleans, LA, USA) *(NIPS '23)*. Curran Associates Inc., Red Hook, NY, USA, Article 2997, 13 pages.

[23] Clemencia Siro, Mohammad Aliannejadi, and Maarten de Rijke. 2022. Understanding User Satisfaction with Task-oriented Dialogue Systems *(SIGIR '22)*. Association for Computing Machinery, New York, NY, USA, 2018–2023. doi:10.1145/3477495.3531798

[24] Catharina Williams van Klinken and John Hajek. 2018. Language contact and functional expansion in Tetun Dili: The evolution of a new press register. *Multilingua* 37 (2018), 613–647. https://doi.org/10.1515/multi-2017-0109

[25] Pedro Carlos Bacelar de Vasconcelos, Andreia Sofia Pinto Oliveira, Ricardo Sousa da Cunha, Andreia Rute da Silva Baptista, Alexandre Corte-Real de Araújo, Benedita McCrorie Graça Moura, Bernardo Almeida, Cláudio Ximenes, Fernando Conde Monteiro, Henrique Curado, et al. 2011. Constituição Anotada da República Democrática de Timor-Leste. Escola de Direito da Universidade do Minho. http://hdl.handle.net/10400.22/4008

[26] Marilyn A. Walker, Diane J. Litman, Candace A. Kamm, and Alicia Abella. 1997. PARADISE: a framework for evaluating spoken dialogue agents *(ACL '98/EACL '98)*. Association for Computational Linguistics, USA, 271–280. doi:10.3115/976909.979652

[27] Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R. Narasimhan, and Yuan Cao. 2023. ReAct: Synergizing Reasoning and Acting in Language Models. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net. https://openreview.net/forum?id=WE_vluYUL-X

[28] Yutao Zhu, Huaying Yuan, Shuting Wang, Jiongnan Liu, Wenhan Liu, Chenlong Deng, Haonan Chen, Zheng Liu, Zhicheng Dou, and Ji-Rong Wen. 2025. Large Language Models for Information Retrieval: A Survey. *ACM Trans. Inf. Syst.* 44, 1, Article 12 (Nov. 2025), 54 pages. doi:10.1145/3748304