Ventana a la Verdad (Window to the Truth): A Chatbot Application for Navigating Colombian Civil Conflict Archives

Anonymous Author(s)

Abstract

We present Ventana a la Verdad, a chatbot designed to make the Clarification Archive and the reports of the Colombian Truth Commission [3] more accessible to a wider audience. These archives contain a wealth of documents, interviews, and testimonies from Colombia's civil conflict, but navigating them can be challenging due to their volume and complexity. Using existing large language models (LLMs) and natural language processing techniques, our chatbot allows users to interact with the archives through natural language queries, receiving relevant and contextually appropriate responses. In the sensitive context of peace and reconciliation, where misinformation or hallucinations can have significant adverse effects, ensuring the accuracy and reliability of information is paramount. This tool aims to facilitate better understanding and engagement with historical content, supporting educational and research efforts. We discuss the development of the chatbot, the challenges encountered, and its potential impact on making the Colombian Truth Commission's archives more accessible. The chatbot is available by anonymous link here: https://divine-adversely-lioness.ngrok-free.app/

CCS Concepts

• **Information systems** → **Question answering**.

Keywords

Conversational AI, Generative AI, Retrieval-Augmented Generation, Large Language Models

ACM Reference Format:

Anonymous Author(s). 2024. Ventana a la Verdad (Window to the Truth): A Chatbot Application for Navigating Colombian Civil Conflict Archives. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 4 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

The Colombian Truth Commission compiled a comprehensive archive documenting the country's civil conflict (1964-2018). This archive includes a large collection of documents, interviews, and testimony that are invaluable for understanding the historical, social, and political aspects of the conflict. The Commission is part of Colombia's comprehensive system of truth, justice, reparation and nonrepetition [5] [2], which began in 2018, culminating in the publication of its final report on 28 June 2022, following extensive research and interviews with 28520 people, including victims, armed participants, and government officials [2]. The presentation of this report was a significant event, marked by an emotional ceremony in Bogotá and subsequent discussions with international organizations such as the Office of the High Commissioner of the United Nations in Geneva, Switzerland [6]. The reports and subsequent dialogues emphasized the importance of continuing social dialogues to achieve reconciliation and prevent a recurrence of conflict in Colombia [6].

However, the massive volume of information and the complexity of the content present significant challenges for those wishing to access and comprehend the material, especially non-Spanish speakers or individuals unfamiliar with the context. To address these challenges, we developed Ventana a la Verdad (Window to the Truth), a chatbot application that enables users to interact with the Colombian Truth Commission's archives in a conversational manner. Users can ask questions to the chatbot and receive relevant information extracted from the documents with reference to the chapter or the interview with the source, simplifying the process of finding specific details and enhancing engagement with the archives. Our work serves as a blueprint for similar initiatives that aim to enhance access to historical archives.

Our interdisciplinary team designed, built and tested Ventana a la Verdad to represent a wide range of perspectives. Powered by OpenAI's large language models [4], the chatbot integrates data from reputable sources to help users make informed decisions and gain a deeper understanding of historical content. The development of the system involved the preparation of data, the generation of embedded data and the implementation of retrieval-augmented generation (RAG) techniques to ensure accurate and reliable responses with references.

We aim to address the following research questions.

RQ1: How can a chatbot application be used effectively to facilitate access to extensive historical archives like those of the Colombian Truth Commission?

RQ2: What challenges arise for information retrieval in complex and sensitive historical documents?

Focusing on these questions, we explore the practical implementation of a chatbot and its effectiveness in increasing accessibility to archives. We also consider the technical and ethical issues associated with handling sensitive historical information.

2 Literature Review

Chatbots are becoming increasingly common in diverse fields like education [13, 18], healthcare [8, 12, 15], and research [9] etc. Their ability to offer personalized and interactive experiences has made them a valuable tool for accessing information, although there are current limitations [7, 17]. The RAG approach represents a significant step forward in addressing these limitations. By combining information retrieval and generative models, RAG allows chatbots

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

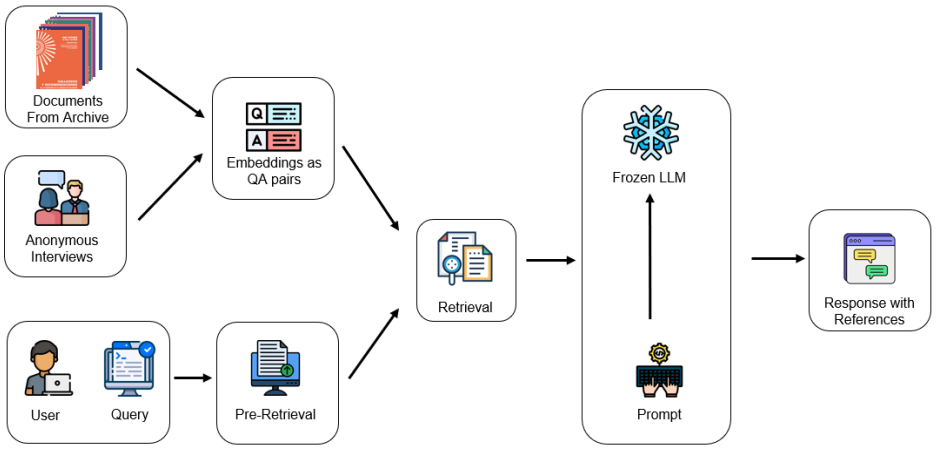

Figure 1: Figure 1: Workflow of the chatbot

to provide more contextually relevant and accurate responses [14]. We will use advanced approaches [11] of RAG to enhance both response accuracy and the breadth of queries chatbots can effectively address, from general overviews to highly specific details. This truly detailed information retrieval capability is particularly valuable in contexts such as the Colombian civil conflict, where detailed and precise information is essential to provide meaningful support and understanding [10, 16].

3 Methods

To illustrate our processes for development and testing, we have divided this section into four parts. In the first part, we explain the architecture and operational processes of Ventana a la Verdad (addressing RQ1). In the second part, we discuss the integration of cultural context into the chatbot's response framework. In the third part, we outline the challenges we faced in information retrieval from complex and sensitive historical documents (addressing RQ2). In the fourth part, we describe our user research and testing methods to ensure the chatbot's effectiveness and usability.

3.1 Architecture of the Chatbot

We collected documents, interviews, and testimonies from the Colombian Truth Commission's digital archives during the data ingestion phase. Each document was pre-processed by segmenting it into individual pages. For each page, LLM GPT-4o mini [1] generates approximately ten potential questions for books and twenty questions per interview that could be answered by the content of that page. This resulted in a collection of question-answer pairs that encapsulate key information from the archives. We wanted to mitigate potential limitations such as the oversimplification of topics, hallucination, the occurrence of nonsensical or irrelevant questions, and the introduction of biases. By creating QA pairs based on the actual content, we ensured that the attachments captured the full complexity and depth of the documents.

The generated questions were converted into embeddings using an embedding model (text-embedding-3-large) [4], producing numerical representations that capture semantic meaning. These embeddings, along with the corresponding answers, were stored in a vector database, allowing for efficient searches for similarity based on cosine similarity.

When users input queries into the chatbot, these queries may be automatically rewritten to better match the format of questions generated during indexing, enhancing retrieval accuracy. The system transforms the user's query into an embedding and then searches for relevant question-answer pairs. By calculating the cosine similarity between the user's query embedding and the stored embeddings, the system retrieves the most relevant matches, ensuring contextually appropriate responses.

A "frozen" LLM, which remains unchanged during this process, generates the final response by combining the retrieved data with the original query. This approach ensures that the response is both comprehensive and grounded in the content of the archives. We begin with a user query to determine if the existing knowledge base is sufficient or if more information is required. If needed, the system retrieves additional data using cosine similarity and then generates and evaluates the completeness response. If the response is insufficient, it refines and processes it further until it meets the user's needs.

Users can continuously interact with the chatbot if there are not enough information in the first query. Query rewriting ensures better alignment with the indexed data for more relevant retrieval, enhancing the overall performance of the chatbot. Using the LLM primarily to generate potential questions during the ingestion phase and to generate responses during user interaction. The complete pipeline is shown in Figure 1.

3.2 Cultural Context Integration Methods

We integrated cultural context into Ventana a la Verdad to accurately reflect the nuances and sensitivities of the Colombian conflict. The documents and interviews in our repository already cover a wide range of perspectives from different groups affected by the conflict, including victims, former combatants, and diverse communities. By generating question-answer pairs directly from these original texts, we capture the specific language, expressions, and

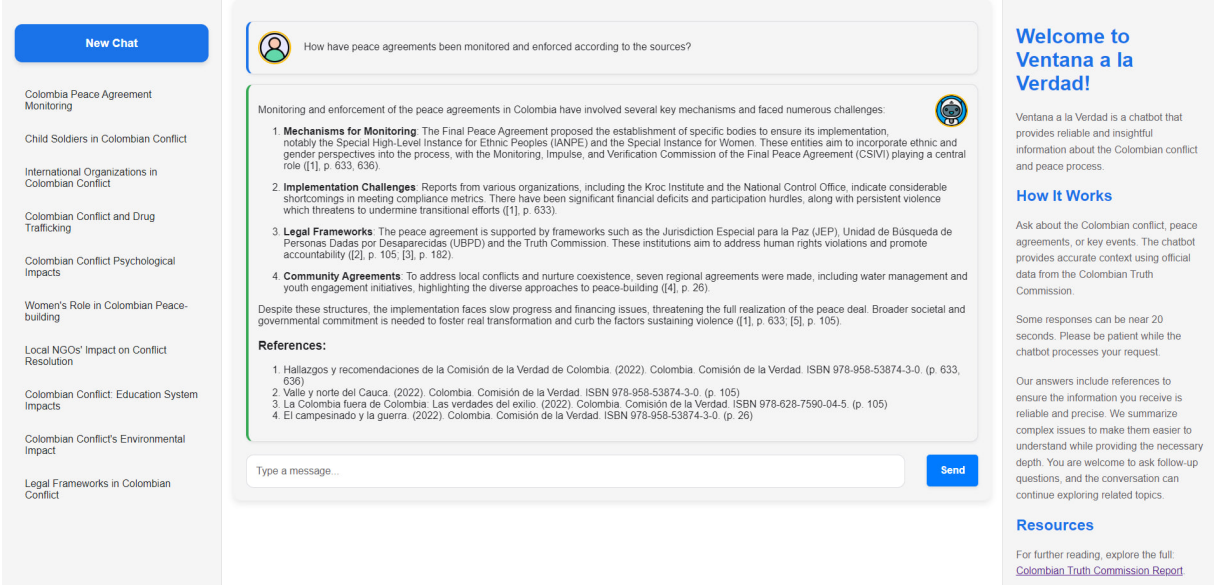

Figure 2: Interface of Ventana a la Verdad Chatbot

cultural references present in the archives. References are an important part here because they allow users to look at the specific chapters or interviews and find the source information directly, which also helps them navigate the archives more effectively. Using embeddings based on these question-answer pairs helped us maintain the semantic integrity of the content and preserve its inherent cultural nuances.

We handled sensitive information with care, ensuring that personal data was anonymized where necessary by the commission itself. For example, we ensured that there is no attribution of individual fault, focusing instead on collective responsibility. If users inquire about the responsibility of specific individuals, such as the president of Colombia, the chatbot is programmed to clarify that the blame cannot be placed on any single individual. This approach maintains the integrity of the historical record and promotes a balanced understanding of the conflict, ensuring that responses remain objective and respectful of all parties involved. Ethical considerations were paramount in our development process to ensure respectful and responsible handling of the historical material.

While we plan to collaborate with bilingual experts in Colombian history and culture in the future to further refine the chatbot's handling of cultural context, our current focus is on leveraging the richness of the original texts to provide culturally sensitive responses. This approach helps maintain the authenticity and integrity of the information presented to users.

Additionally, using prompt engineering techniques, we have ensured that the chatbot generates precise references to the source material when answering questions. This allows the user to the manually verify the truthfulness of the chatbot's responses.

3.3 User Testing Methods

To test the usability and effectiveness of Ventana a la Verdad, we have conducted initial internal testing within the development team. This allowed us to identify technical issues and ensure that the system retrieves and generates appropriate responses. We focused on verifying the accuracy of the retrieval process, the coherence of the generated answers, and the overall user experience. For anonymity, we provided the link to chat. Figure 2 shows Ventana a la Verdad's interface, designed for easy interaction with the Colombian Truth Commission archives. Three sections improve the user experience. On the left is the conversation history. In the center, users can type questions and receive chatbot responses. The right side provides helpful information about chatbot.

In the future, we plan to conduct pilot testing with bilingual experts in Colombian history and culture. Their feedback will be crucial in identifying areas where cultural context may not be adequately preserved and in refining the system accordingly. This collaboration will help enhance the chatbot's ability to handle nuanced queries and provide more insightful responses.

We also intend to expand testing to include educators, students, researchers, and members of the public interested in the Colombian conflict. Participants will interact with the chatbot, perform specific tasks, and provide feedback on their experience. Users will have the option to rate responses and provide comments directly through the chatbot interface. This feedback will be used to iteratively improve the system, addressing any shortcomings and enhancing its usability.

Throughout the user testing process, we will ensure compliance with ethical guidelines, including informed consent and data privacy protections. Participants will be informed about the purpose of the testing, how their data will be used, and their rights to withdraw at any time. By adhering to these principles, we aim to conduct responsible and respectful research that contributes to the effectiveness of Ventana a la Verdad.

4 Findings

Our approach can potentially provide researchers, educators, and the public with trustworthy information on the Colombian conflict, ultimately facilitating better understanding and informed discussions. Our approach underscores the value of integrating advanced natural language processing techniques to enhance the performance of LLMs in specialized domains like historical archives. By leveraging the extensive data from the Colombian Truth Commission, we aim to create a tool that provides more accurate and reliable answers to questions related to the conflict (RQ1).

We encountered several challenges in information retrieval from the archives. The volume and complexity of the data, sensitivity of the content, language nuances, and technical limitations made the process difficult. By identifying and addressing these challenges, we improved the chatbot's ability to provide accurate and contextually appropriate responses (RQ2).

While our tool is effective in making complex historical documents more accessible to a broader audience, it is crucial to clarify that it does not aim to replace scholarly analysis or official interpretations. The tool serves solely as a supplementary resource that helps distill and summarize key information, thereby supporting—but not substitute for—the complex and multifaceted process of historical research and reconciliation efforts. Making these archives more accessible can contribute to educational initiatives and promote a deeper understanding of the conflict's impact. For example, easier access to personal testimonies and reports can enable more meaningful engagement with the material. In summary, Ventana a la Verdad aims to make the information more accessible and the exploration process more efficient, without overstepping into the domain of official historical interpretation.

5 Discussion

Using a chatbot to present information from the Colombia Truth Commission archives brought unique challenges. One of the main concerns was the risk of "hallucinations," where the language model could generate incorrect or misleading information. This is especially problematic with sensitive historical content, as inaccuracies can distort events and impact users' understanding.

To address this, we took steps to minimize hallucinations. We added specific instructions to guide the chatbot's responses, helping it stay accurate and focused on the content. Also, we added the function to show references with pages or interviews that users could double-check the information.

Ethical considerations were also critical. The archives include personal stories and accounts of traumatic events. We needed to handle this information carefully, respecting privacy and cultural sensitivities. We programmed the chatbot to avoid blaming specific individuals and to provide balanced responses focused on the collective experience.

Looking forward, we plan to work with human experts in Colombian history and culture to further improve the chatbot. This collaboration will help ensure that the chatbot continues to provide users with accurate, respectful, and useful information.

In summary, the creation of Ventana a la Verdad showed us the importance of addressing both technical and ethical issues when using AI to present complex historical information. By focusing on reducing hallucinations and carefully handling sensitive content, we improved the chatbot's ability to provide accurate and relevant responses.

6 Conclusion

We introduced Ventana a la Verdad, a chatbot designed to make the Colombian Truth Commission's extensive archives more accessible. The chatbot allows users to ask questions in natural language and provides answers with references to specific chapters or interviews from the archives. This feature helps users, including non-Spanish speakers and researchers, navigate and understand complex historical documents more easily.

By simplifying access to these important resources, our tool supports education and research related to Colombia's civil conflict. Ventana a la Verdad serves as a helpful aid for those seeking to engage with the archives without replacing scholarly analysis or official interpretations. We plan to continue improving the chatbot to enhance its usefulness for a wider audience.

References

[1] [n. d.]. [2303.08774] GPT-4 Technical Report. https://arxiv.org/abs/2303.08774

[2] [n. d.]. Colombian Armed Conflict. https://justiceforcolombia.org/about-colombia/colombian-armed-conflict/

[3] [n. d.]. Inicio | Informe Final Comisión de la Verdad. https://www.comisiondelaverdad.co/

[4] [n. d.]. OpenAI. https://openai.com/

[5] 2022. Colombian Truth Commission's Final Report. https://www.gicj.org/positions-opinons/gicj-positions-and-opinions/2810-colombian-truth-commission-s-final-report

[6] ABColombia. 2022. Truth Commission of Colombia: Executive Summary. https://www.abcolombia.org.uk/truth-commission-of-colombia-executive-summary/

[7] Joshua Au Yeung, Zeljko Kraljevic, Akish Luintel, Alfred Balston, Esther Idowu, Richard J Dobson, and James T Teo. 2023. AI chatbots not yet ready for clinical use. *Frontiers in digital health* 5 (2023), 1161098.

[8] Chiranjib Chakraborty, Soumen Pal, Manojit Bhattacharya, Snehasish Dash, and Sang-Soo Lee. 2023. Overview of Chatbots with special emphasis on artificial intelligence-enabled ChatGPT in medical science. *Frontiers in Artificial Intelligence* 6 (2023), 1237704.

[9] Yu Chen, Scott Jensen, Leslie J Albert, Sambhav Gupta, and Terri Lee. 2023. Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers* 25, 1 (2023), 161–182.

[10] Yujuan Ding, Wenqi Fan, Liangbo Ning, Shijie Wang, Hengyun Li, Dawei Yin, Tat-Seng Chua, and Qing Li. 2024. A survey on rag meets llms: Towards retrieval-augmented large language models. *arXiv preprint arXiv:2405.06211* (2024).

[11] Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, and Haofen Wang. 2023. Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997* (2023).

[12] Nima Ghorashi, Ahmed Ismail, Pritha Ghosh, Anton Sidawy, and Ramin Javan. 2023. AI-powered chatbots in medical education: potential applications and implications. *Cureus* 15, 8 (2023).

[13] Lasha Labadze, Maya Grigolia, and Lela Machaidze. 2023. Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education* 20, 1 (2023), 56.

[14] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems* 33 (2020), 9459–9474.

[15] Erwin Loh. 2023. ChatGPT and generative AI chatbots: challenges and opportunities for science, medicine and medical leaders. *Bmj Leader* (2023), leader–2023.

[16] Yuchen Shao, Yuheng Huang, Jiawei Shen, Lei Ma, Ting Su, and Chengcheng Wan. 2024. Vortex under Ripplet: An Empirical Study of RAG-enabled Applications. *arXiv preprint arXiv:2407.05138* (2024).

[17] Chris Stokel-Walker and Richard Van Noorden. 2023. What ChatGPT and generative AI mean for science. *Nature* 614, 7947 (2023), 214–216.

[18] Ahmed Tlili, Boulus Shehata, Michael Agyemang Adarkwah, Aras Bozkurt, Daniel T Hickey, Ronghuai Huang, and Brighter Agyemang. 2023. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart learning environments* 10, 1 (2023), 15.