

PA-RAG: RAG Alignment via Multi-Perspective Preference Optimization

Jiayi Wu¹, Hengyi Cai², Lingyong Yan³, Hao Sun⁴,
Xiang Li^{1*}, Shuaiqiang Wang³, Dawei Yin³, Ming Gao¹

¹School of Data Science and Engineering, East China Normal University

²Chinese Academy of Sciences ³Baidu Inc ⁴Peking University

jiayiwu@stu.ecnu.edu.cn

Abstract

The emergence of Retrieval-augmented generation (RAG) has alleviated the issues of outdated and hallucinatory content in the generation of large language models (LLMs), yet it still reveals numerous limitations. When a general-purpose LLM serves as the RAG generator, it often suffers from inadequate response informativeness, response robustness, and citation quality. Past approaches to tackle these limitations, either by incorporating additional steps beyond generating responses or optimizing the generator through supervised fine-tuning (SFT), still failed to align with the RAG requirement thoroughly. Consequently, optimizing the RAG generator from multiple preference perspectives while maintaining its end-to-end LLM form remains a challenge. To bridge this gap, we propose Multiple Perspective Preference Alignment for Retrieval-Augmented Generation (PA-RAG), a method for optimizing the RAG generator to align with RAG requirements comprehensively. Specifically, we construct high-quality instruction fine-tuning data and multi-perspective preference data by sampling varied quality responses from the generator across different prompt documents quality scenarios. Subsequently, we optimize the generator using SFT and Direct Preference Optimization (DPO). Extensive experiments conducted on four question-answer datasets across three LLMs demonstrate that PA-RAG can significantly enhance the performance of RAG generators. Our code and datasets are available at <https://github.com/wujwyi/PA-RAG>.

1 Introduction

Pre-trained large language models (LLMs) demonstrate impressive question-answering capabilities but also reveal certain drawbacks, such as generating outdated or hallucinatory information (Tonmoy et al., 2024; Turpin et al., 2023). The emergence of

*Corresponding Author

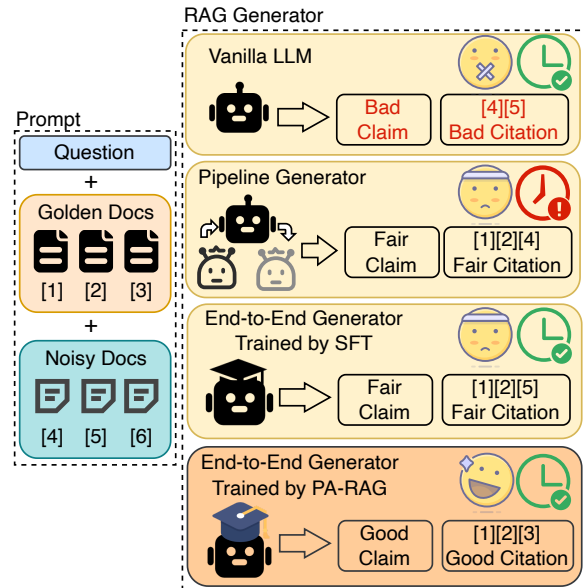


Figure 1: PA-RAG retains the end-to-end form of the generator while enabling the generator to learn multi-perspective preference information.

retrieval-augmented generation (RAG) (Gao et al., 2023b) has alleviated this issue by incorporating retrieved documents relevant to the question into the prompt, thereby providing the LLMs with information they may not know.

However, while RAG systems help improve factual grounding, the role of the generator in previous work is usually far from well-aligned. Despite their impressive pre-training, off-the-shelf LLMs are not optimally suited for the specific requirements of RAG tasks without proper task alignment. 1) **Helpfulness for RAG application (response informativeness)** (Dong et al., 2024b): The generator should possess a refined ability to identify and utilize valuable information from the provided context. When presented with a mix of high-quality and less relevant information, it should actively and effectively leverage the valuable content while disregarding the rest. 2) **Harmlessness for RAG application (response robustness)** (Yoran et al.,

2024a; Zhu et al., 2024): A well-aligned generator should demonstrate robust noise resistance, capable of filtering out misleading information without being adversely affected. This ability ensures the generated responses remain accurate and coherent even when the input context is imperfect. 3) **Honesty for RAG application (citation quality)** (Gao et al., 2023a; Sun et al., 2023): The generator should be capable of producing responses that are firmly rooted in the provided documents while maintaining appropriate attribution. This involves explicitly citing the source documents when necessary, and ensuring that the generated content is not only factually accurate but also traceable to the retrieved documents.

To satisfy the aforementioned RAG requirements, existing RAG systems mainly fall into two categories: end-to-end architecture and pipeline architecture. These approaches attempt to bridge the gap between off-the-shelf LLMs and the specific demands of RAG, but both face significant limitations. End-to-end architectures primarily rely on supervised fine-tuning (SFT) of general LLMs (Yoran et al., 2024a; Fang et al., 2024). These approaches focus on constructing high-quality responses in RAG scenarios. However, through further analysis, we found that the requirements for the generators in RAG tasks are highly context-dependent and often interwoven. These varying and intertwined demands make it challenging for a model to meet all RAG objectives through standard SFT alone, as it does not incorporate the necessary preference information required for adapting to different retrieval scenarios. On the other hand, pipeline architectures introduce additional steps beyond generation, such as re-ranking retrieved documents, filtering irrelevant information, or employing post-hoc verification to ensure that citations support the claims (Dong et al., 2024a; Yu et al., 2024; Wang et al., 2024; Sun et al., 2023). However, these additional steps can only satisfy certain RAG requirements and still have a considerable gap in aligning with the global RAG requirements. Moreover, these additional steps introduce extra computational costs and time consumption, making it less practical for large-scale deployment. This further inspires us to think: *Is it possible to fully align the generator with the diverse RAG requirements while retaining the simplicity and efficiency of an end-to-end architecture?*

To this end, in this paper, we propose Multi-Perspective Preference Alignment for Retrieval-

Augmented Generation (PA-RAG), aiming to optimize the generator of RAG systems to align comprehensively with specific RAG requirements. As illustrated in Figure 1, PA-RAG maintains the end-to-end architecture of the generator while enabling it to learn multi-perspective preference information. The training of PA-RAG is divided into two phases. The first phase is foundational capability training, where the generator acquires the basic ability to utilize and cite documents through instruction fine-tuning. To construct high-quality instruction fine-tuning data, we utilize ChatGPT to generate complete and correct answers and employ a *citation rewrite* mechanism to ensure citation quality. The second phase is the multi-perspective preference optimization phase, in which the generator is trained using Direct Preference Optimization (DPO) (Rafailov et al., 2023) to learn preference information from different perspectives. This phase encompasses three sub-stages, sequentially enhancing the generator’s response informativeness, response robustness, and citation quality. To construct high-quality preference data, we sample responses of varying quality from the generator across scenarios with different document qualities in prompt to build preference data for informativeness and robustness. Additionally, we employ the *citation rewrite* mechanism to construct preference data for citation quality.

We conducted extensive experiments on four QA datasets and three LLMs, demonstrating that PA-RAG significantly enhances the generator’s performance. The improvement achieved by PA-RAG far surpasses the performance gains using only SFT or additional steps. LLMs trained with PA-RAG show an average absolute improvement of 13.97% in correctness, 49.77% in citation recall, and 39.58% in citation precision. Our contributions can be summarized as follows:

- We propose PA-RAG, which achieves comprehensive alignment of LLMs with specific RAG requirements through a multi-stage, multi-perspective training while preserving its end-to-end architecture.
- We publicly release our training data, which includes 58.9k instruction fine-tuning data and 48.7k preference optimization data.
- Through extensive experiments, we demonstrate the effectiveness of PA-RAG, showing that the preference optimization from each perspective is effective.

2 Preliminaries

In this section, we describe the optimization objectives of the RAG generator, the motivation for multi-perspective optimization, and the rationale for choosing preference optimization over mere instruction fine-tuning.

2.1 Optimization Objectives of the RAG Generator

As illustrated in Figure 1, when answering a question, the RAG generator outputs the corresponding response after receiving the question and multiple relevant documents retrieved by the retriever. Our optimization goal is to enable the generator to fully utilize valuable documents and accurately cite the references corresponding to claims.

The ability to fully utilize valuable documents corresponds to the correctness of the response, meaning that all answers contained within the documents should be included in the generator’s output. Formally, let G represent the generator, x denotes the input, $D = \{d_1, d_2, \dots, d_n\}$ represent the documents retrieved by the retriever, $A = \{a_1, a_2, \dots, a_n\}$ represent the short answers contained in the documents, and y denote the response generated by the generator. The correct response by the generator can be expressed as:

$$\begin{aligned} y &= G(x, D) \\ \text{s.t. } \forall a_i \in A, C(y, a_i) &= \text{True} \end{aligned} \quad (1)$$

where $C(y, a_i) = \text{True}$ implies that the answer a_i is included in y .

Based on the generator’s ability to correctly answer questions, it is necessary to enhance its capability to accurately cite documents corresponding to claims, referred to as citation quality. Each response y contains multiple statements comprising a claim and multiple citations. We aim for each claim to be fully supported by the cited documents, while avoiding irrelevant citations. Formally, we denote the statement as s , the claim as c , and the citation as t . The correct citation of references corresponding to claims can be expressed as follows:

$$y = \{s_1, s_2, \dots, s_n\} \quad (2)$$

$$s_i = \{\text{“claim”}:c_i, \text{“citation”}:t_i = [t_{i1}, t_{i2}, \dots, t_{in}]\} \quad (3)$$

$$\forall s_i \in y, \phi(\text{concat}(t_i), c_i) = 1 \quad (4)$$

$$\begin{aligned} \forall s_i \in y, \forall t_{ij} \in t_i, \\ \phi(t_{ij}, c_i) = 1 \vee \phi(\text{concat}(t_i \setminus t_{ij}), c_i) = 0 \end{aligned} \quad (5)$$

Here, $\text{concat}(t_i)$ denotes the concatenation of all cited documents, and $\phi(t, c) = 1$ indicates that the claim c is fully supported by the citation t . Formula 4 indicates that all claims can be fully supported by their citations, while Formula 5 ensures that no irrelevant documents are cited.

2.2 Preference Optimization Under Different Document Qualities

As described in § 2.1, enhancing the correctness of the RAG generator is the primary objective. However, optimizing correctness is highly complex. Specifically, the probability of a model correctly answering a question is $P(y|x)$. In the RAG framework, this probability depends on the distribution of the retrieved documents and can be formulated as $P(y|x) = \sum_D [P(D|x) \cdot P(y|D, x)]$. However, the retriever in a RAG system is not perfect, leading to variability in document quality under different circumstances, which results in a non-uniform distribution of $P(D|x)$. Therefore, it is essential to optimize the generator under different document quality conditions.

We categorize document quality into two types. The first type involves scenarios where all documents contain relevant answers. In this scenario, we expect the generator to fully utilize the valuable documents and produce complete answers. Therefore, when the documents are of high-quality, we need to focus on optimizing response informativeness. The second type includes scenarios where the document set contains noisy documents. In this case, we expect the generator to avoid the interference of noisy documents and still maintain the completeness of its answers. Thus, when dealing with low-quality documents, the emphasis shifts to optimizing response robustness. However, optimizing for informativeness and robustness pulls the generator in different directions. The former encourages referencing more documents, while the latter demands ignoring more. Standard instruction fine-tuning is insufficient to teach the model how to balance these competing demands (documents encouraged to be referenced are valuable, while those that should be ignored are considered noisy, see § 4.5 for details), making it difficult to optimize the generator effectively. To address this challenge, we introduce preference information and perform multi-perspective preference optimization to guide the generator in focusing on valuable documents while ignoring noisy ones.

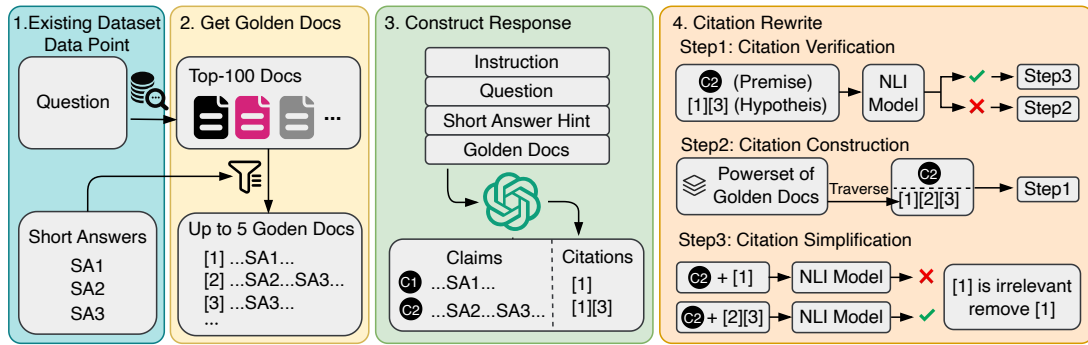


Figure 2: Overview of instruction fine-tuning data construction.

3 Methodology

After defining the optimization objectives, we trained the generator in two phases. The first phase enabled a general LLM to acquire basic RAG capabilities. The method for constructing the training data for this phase is detailed in § 3.1. The second phase involved multi-perspective preference optimization, further enhancing the generator’s response informativeness, response robustness, and citation quality. The method for constructing the training data for this phase is described in § 3.2.

3.1 Instruction Fine-tuning for Basic RAG Capabilities

During the instruction fine-tuning phase, we aimed to equip the generator with fundamental abilities to utilize and cite documents following the optimization objectives outlined in § 2.1. When constructing the training data, we employ ChatGPT-3.5 (GPT-3.5-Turbo-1106) and introduce a *citation rewrite* mechanism to create near-perfect responses. An overview of the instruction fine-tuning data construction is illustrated in Figure 2.

Firstly, we need to acquire high-quality documents after obtaining the questions and short answers provided by existing datasets. We use the RAG retriever to retrieve the top 100 most relevant documents from the retrieval corpus, and then filter to retain the golden documents that contain the short answer. Subsequently, we select up to five documents from all the golden documents as prompt documents, ensuring that all short answers are included in the prompt documents.

Secondly, we need to construct responses that include all short answers and exhibit high citation quality. Specifically, we use ChatGPT-3.5 to generate responses. To enhance the quality of the responses, we include instructions and short answer hints in the ChatGPT prompt. Detailed prompt can

be found in Appendix D.

Although ChatGPT-3.5 is quite powerful and the prompt includes short answer hints, the quality of the citations in the answers is unsatisfactory. Therefore, we introduce a *citation rewrite* mechanism to improve the quality of substandard citations, which is divided into the following steps: **First**, verify the citation. We use a Natural Language Inference (NLI) model to validate whether the document corresponding to each citation (premise) in the statement can satisfy the claim (hypothesis). **Second**, construct the citation. If the citation in the first step cannot support the claim, we traverse the power set of all prompt documents as citations to explore a feasible citation scheme. **Third**, simplify the citation. If the citation can support the claim, we check one by one whether the citation is irrelevant to the claim and remove irrelevant citations.

After generating answers with ChatGPT-3.5 and performing *citation rewrite*, we filter to retain responses that include all short answers and contain only accurate citations to serve as the training data for the instruction fine-tuning of the generator, which is then subsequently fine-tuned.

3.2 Preference Optimization of the Generator through DPO

Once the generator has acquired the fundamental ability to utilize and reference documents, we sequentially optimize the generator from the perspectives of response informativeness, response robustness, and citation quality during the preference optimization phase. Preference optimization requires constructing data with preference information, including an input, a superior answer (chosen output), and an inferior answer (rejected output). We will progressively explain the methods for constructing preference data from each perspective. An overview of the preference data construction is illustrated in Figure 3.

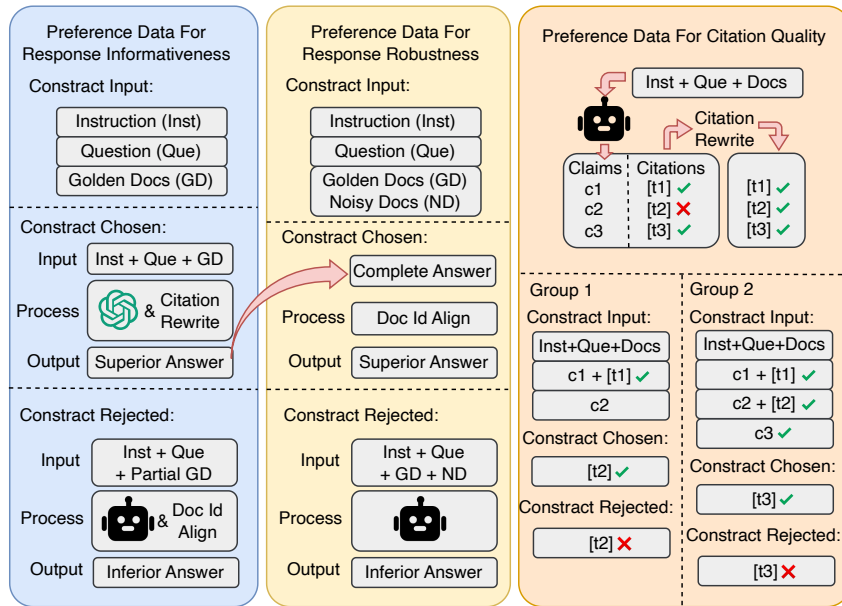


Figure 3: Overview of preference data construction.

3.2.1 Response Informativeness

Response informativeness refers to the completeness of the answer in the response. The optimization objective from this perspective is to make the generator fully utilize the document that contains the short answer.

Input Construction: Similar to constructing the input component of the instruction fine-tuning data, the input includes the instruction, the question, and the high-quality prompt documents containing all short answers from up to 5 golden documents.

Chosen Output Construction: The construction method is consistent with the output component of the instruction fine-tuning data, using ChatGPT-3.5 and the *citation rewrite* mechanism to construct a response that includes all short answers and the accurate citation.

Rejected Output Construction: To simulate the scenario where the generator ignores parts of the golden documents, we delete some golden documents from the prompt, generate responses with the generator, and filter to retain inferior answers that are incomplete. Since some documents are removed from the prompt, we need to adjust the citation numbers after obtaining the inferior answers. Specific examples of citation number adjustments are shown in Appendix E.

3.2.2 Response Robustness

Response robustness refers to the ability of the generator to resist interference. The optimization objective from this perspective is to enable the gen-

erator to avoid interference from noisy documents.

Noisy document construction: We categorize noisy documents into two types. The first type is documents related to the question but does not contain the answer, from which we randomly select two documents that do not contain short answers from the top 100 relevant retrieved documents. The second type is irrelevant documents to the question, from which we randomly select two documents that do not contain short answers from the retrieval documents of other questions.

Input construction: The input includes the instruction, the question, and the low-quality prompt documents that mix up to five golden documents and four noisy documents.

Chosen output construction: To simulate a scenario where the generator ignores all noisy documents, we use the chosen output generated in the absence of noisy documents in § 3.2.1. Since there are no noisy documents, we need to adjust the citation numbers. Specific examples of citation number adjustments are shown in Appendix E.

Rejected output construction: Use inputs containing low-quality documents, generate responses with the generator, and filter to retain inferior answers that are incomplete.

3.2.3 Citation Quality

Citation quality refers to the generator’s ability to cite documents correctly. The optimization objective from this perspective is to enable the generator to cite documents related to the claim correctly and

to avoid citing irrelevant documents.

Data construction is divided into two steps: First, we use the generator to generate responses and filter to retain those containing all short answers. Second, we use the *citation rewrite* mechanism to identify incorrect citations that fail the NLI model verification or cite irrelevant documents as rejected output and then correct them as chosen output.

Construction of input: The input includes the instruction, the questions, the prompt documents, and the response up to the incorrect citation.

Construction of chosen output: Correct citation after citation rewrite.

Construction of rejected output: Incorrect citation that fails the NLI model verification or cites irrelevant documents.

3.2.4 Staged Preference Optimization

Given the substantial discrepancies in preference information from different perspectives, we divide the preference optimization process into multiple sub-stages to independently optimize the RAG preferences from each perspective. Specifically, the first and second sub-stages optimize response informativeness and response robustness, respectively. The rejected outputs in these stages are generated by the generator that has undergone instruction fine-tuning. The third sub-stage optimizes citation quality, with preference data generated by the generator that has undergone optimization in the first and second sub-stages.

4 Experiment

4.1 Experimental Setup

4.1.1 Dataset and Evaluation Methodology

Datasets The questions used for constructing our training data are sourced from the ASQA (Stelmakh et al., 2022), WebQuestions (Berant et al., 2013), and Natural Questions (Kwiatkowski et al., 2019) training splits. Detailed statistics for the training set can be found in Appendix C. The evaluation data is sourced from the test splits of the three datasets mentioned above, along with TriviaQA (Joshi et al., 2017), which serves as an unseen dataset to assess the **out-of-domain** generalizability of the generator.

Evaluation Metrics To align with the optimization objectives discussed in § 2.1, we evaluate the generator’s performance regarding correctness and citation quality. Following ALCE (Gao et al., 2023a) and VTG (Sun et al., 2023), for correctness,

we assess whether the short answers (provided by the dataset) are exact substrings of the generation to calculate the **exact match** (EM) score. For citation quality, we use **citation recall** to evaluate whether the output is fully supported by the cited documents and **citation precision** to assess whether irrelevant documents are cited.

4.1.2 Implementation Details

We selected three general LLMs as the base RAG generator: LLAMA2-7B-CHAT, LLAMA2-13B-CHAT, and LLAMA3-8B-INSTRUCT. For the NLI model of citation rewrite and citation quality evaluation, we follow the works of Gao et al. (2023a), Sun et al. (2023), and Huang et al. (2024), utilizing the NLI model TRUE (Honovich et al., 2022), a T5-11B model that is fine-tuned on a collection of NLI datasets. Meanwhile, ALCE has verified through extensive human evaluations that this NLI model correlates with human judgment. For retrieval, we employed the Wikipedia dump from December 20, 2018, as our retrieval corpus and used GTR (Ni et al., 2022) as our dense retriever.

We performed full fine-tuning on the generators, with the training hyperparameters detailed in Appendix A. All generators were trained on the same dataset, in which the RAG generator used to generate preference data was the LLAMA2-7B-CHAT at different stages of fine-tuning.

4.1.3 Baselines

We adopt the following baseline systems. 1) **Base generator**, the general LLM without further optimization. 2) **RetRobust-13B** (Yoran et al., 2024b), a method that enhances the robustness of the RAG generator through SFT. 3) **Self-RAG-13B** (Asai et al., 2024), a method that improves the performance of the RAG generator through a rank-then-generate pipeline. 4) **SFT on chosen**, to compare preference optimization and SFT, we use the chosen output of preference data for SFT on the LLAMA2-7B-CHAT. All the methods above were evaluated using the same prompt documents used for PA-RAG to facilitate a fair comparison.

4.2 Main Results

The evaluation results of PA-RAG and the baseline across four QA datasets are presented in Table 1. We have made the following observations: First, PA-RAG consistently and significantly improves the performance of various LLMs in RAG scenarios, with an average increase of 13.97% in EM

Datasets	ASQA			WebQuestions			Natural Questions			TriviaQA		
	Correct	Citation		Correct	Citation		Correct	Citation		Correct	Citation	
Generators	EM	Rec	Prec	EM	Rec	Prec	EM	Rec	Prec	EM	Rec	Prec
Base Generators												
LLAMA2-7B-CHAT	35.61	20.51	34.51	40.91	24.36	40.27	48.82	20.06	34.71	67.69	22.81	37.27
LLAMA2-13B-CHAT	27.24	16.02	24.87	25.45	13.75	21.03	34.47	14.58	22.57	54.55	15.31	23.31
LLAMA3-8B-INSTRUCT	34.77	50.23	51.67	39.69	49.86	51.55	49.71	47.82	49.80	66.98	45.78	45.72
Other Training-Needed Methods												
RetRobust-13B	26.41	-	-	33.71	-	-	43.08	-	-	73.13	-	-
Self-RAG-13B	31.61	58.77	68.88	38.64	60.46	71.75	48.09	63.65	72.75	66.95	63.62	65.24
SFT on Chosen	34.49	72.61	69.29	39.07	80.81	77.66	50.30	73.77	70.76	64.59	66.15	62.75
PA-RAG (Ours)												
LLAMA2-7B+PA-RAG	46.16	77.66	76.82	48.37	79.36	78.16	63.27	76.60	76.49	72.52	70.86	69.89
LLAMA2-13B+PA-RAG	46.53	80.68	77.73	48.72	80.90	78.41	62.71	78.72	75.89	73.56	71.96	68.95
LLAMA3-8B+PA-RAG	47.03	81.97	79.93	48.59	84.31	79.84	62.29	81.26	79.08	73.79	74.07	71.05

Table 1: Main results of PA-RAG and baselines on four QA datasets. The best performances within the same backbone generator are marked in **bold**.

Datasets	ASQA				WebQuestions				Natural Questions				TriviaQA			
	Correct	Citation			Correct	Citation			Correct	Citation			Correct	Citation		
Generators	EM	Rec	Prec	F1	EM	Rec	Prec	F1	EM	Rec	Prec	F1	EM	Rec	Prec	F1
LLAMA2-7B-CHAT																
Base Generator	35.61	20.51	34.51	25.73	40.91	24.36	40.27	30.36	48.82	20.06	34.71	25.43	67.69	22.81	37.27	28.30
+IFT	37.98	75.69	70.17	72.83	42.76	75.85	70.50	73.08	54.59	74.34	68.95	71.54	68.51	65.41	60.34	62.77
+IFT+RI	43.95	77.68	68.32	72.70	47.75	78.93	69.32	73.81	60.11	76.19	66.22	70.85	71.61	69.00	60.23	64.32
+IFT+RI+RR	46.37	75.25	71.58	73.37	48.03	79.00	74.31	76.58	62.99	75.55	71.12	73.27	72.36	69.23	64.74	66.91
+IFT+RI+RR+CQ	46.16	77.66	76.82	77.24	48.37	79.36	78.16	78.76	63.27	76.60	76.49	76.54	72.52	70.86	69.89	70.37
LLAMA2-13B-CHAT																
Base Generator	27.24	16.02	24.87	19.49	25.45	13.75	21.03	16.63	34.47	14.58	22.57	17.72	54.55	15.31	23.31	18.49
+IFT	38.23	76.28	70.82	73.45	42.04	78.87	73.05	75.85	54.85	75.75	69.91	72.71	69.54	66.89	61.81	64.25
+IFT+RI	42.46	80.24	70.05	74.80	46.66	81.74	71.59	76.33	58.68	78.33	68.09	72.85	72.15	70.26	60.88	65.24
+IFT+RI+RR	45.50	77.72	75.57	76.63	48.16	80.23	77.40	78.79	61.58	76.82	74.17	75.47	72.78	68.79	66.57	67.66
+IFT+RI+RR+CQ	46.53	80.68	77.73	79.18	48.72	80.90	78.41	79.64	62.71	78.72	75.89	77.28	73.56	71.96	68.95	70.42
LLAMA3-8B-INSTRUCT																
Base Generator	34.77	50.23	51.67	50.94	39.69	49.86	51.55	50.69	49.71	47.82	49.80	48.79	66.98	45.78	45.72	45.75
+IFT	38.52	76.78	71.74	74.18	42.80	79.75	74.03	76.78	54.95	81.80	71.40	76.25	69.35	69.00	63.58	66.18
+IFT+RI	44.16	82.48	72.82	77.35	47.52	83.54	73.62	78.27	61.13	82.94	76.01	79.33	72.78	73.35	63.39	68.01
+IFT+RI+RR	47.35	81.17	74.65	77.77	48.48	84.11	77.12	80.46	62.71	82.91	75.35	78.95	73.62	75.14	68.18	71.49
+IFT+RI+RR+CQ	47.03	81.97	79.93	80.94	48.59	84.31	79.84	82.01	62.29	81.26	79.08	80.16	73.79	74.07	71.05	72.53

Table 2: Evaluation results for each training phase. IFT stands for instruction fine-tuning, RI denotes the optimization of response informativeness, RR indicates the optimization of response robustness, and CQ represents the optimization of citation quality. The best scores are highlighted in **bold**.

score, 49.77% in citation recall, and 39.58% in citation precision. This demonstrates that our training data is applicable to different LLMs, showcasing strong generalizability. Meanwhile, PA-RAG does not compromise the fluency of the generator’s outputs.¹ Second, PA-RAG significantly outperforms the baselines, indicating that compared to end-to-end generators trained via SFT and pipeline generators, the end-to-end generator trained through preference optimization aligns better with RAG preferences, resulting in superior performance.

4.3 Ablation Study

The evaluation results for each training phase are presented in Table 2. We observed that each train-

¹A detailed analysis of response fluency is presented in Appendix H.

ing phase generally contributes to improvements in both correctness and citation quality, indicating that optimization of RAG requirements in each perspective is necessary. Furthermore, correctness and citation quality are complementary to each other. Detailed Analysis of why optimizing citation quality might have a negative impact on citation recall is presented in Appendix G.

To more intuitively demonstrate the impact of preference optimization on the generator, we present detailed LLM evaluation in Appendix J and case studies in Appendix K.

4.4 Impact of Preference Optimization Order

In the main experiment, we first optimized response informativeness, followed by response robustness. We further investigated the impact of altering the

Generators	ASQA	WebQ	NQ	TQA
LLAMA-7B-CHAT+IFT	37.98	42.76	54.59	68.51
+RI	43.95	47.75	60.11	71.61
+RR	25.91	32.23	52.52	37.52
+RR+RI	44.74	47.09	61.78	71.28
+Mix RI&RR	33.72	37.84	61.43	45.66
+RI+RR	46.37	48.03	62.99	72.36

Table 3: EM score results of different preference optimization orders. Best scores are highlighted in **bold**.

Generators	ASQA	WebQ	NQ	TQA
LLAMA-7B-CHAT+IFT	37.98	42.76	54.59	68.51
+SFT RI(Chosen)	38.93	43.69	54.84	69.10
+SFT RI+RR(Chosen)	32.17	37.02	47.01	62.56
+DPO RI	43.95	47.75	60.11	71.61
+DPO RI+RR	46.37	48.03	62.99	72.36

Table 4: Further comparison of DPO and SFT under response informativeness (RI) and response robustness (RR). Results of the EM score.

optimization order on performance, with results shown in Table 3. Our findings are as follows: First, optimizing response robustness before response informativeness resulted in lower performance than the main experimental setup. Second, skipping informativeness and directly optimizing robustness or merging the data to optimize simultaneously can lead to a decline in performance. We believe this is because optimizing informativeness enables the generator to learn how to effectively utilize retrieved documents, which is the fundamental capability in the RAG system. In contrast, optimizing robustness requires the generator to learn to reject irrelevant documents, which is a more advanced skill.² Learning skills at different levels requires a reasonable arrangement in learning order, which is similar to curriculum learning (Bengio et al., 2009). Neglecting the foundational skills or mixing different skill levels during learning may hurt the model’s original capabilities.

4.5 Further Comparison of DPO and SFT

As mentioned in § 2.2, there is a significant disparity in the optimization directions of response informativeness and robustness. We further explore the differences between DPO and SFT when optimizing these two directions sequentially. As shown in Table 4, we observe that, when using SFT for training, optimizing response informativeness first can enhance the generator’s performance. However, subsequent optimization for response ro-

²Detailed analysis of why rejecting irrelevant documents is a more advanced skill is showcased in Appendix F.

business lead to a significant performance decline, exposing SFT’s vulnerability to catastrophic forgetting (French and Chater, 2002) when undergoing two relatively different optimizations consecutively. In contrast, DPO handles this situation well, consistently improving the generator’s performance.

5 Related Work

5.1 Retrieval-Augmented Generator

To satisfy the requirements of RAG system, existing retrieval-augmented generators are primarily categorized into end-to-end and pipeline frameworks. In end-to-end frameworks, models such as RetRobust (Yoran et al., 2024a) and RAAT (Fang et al., 2024) enhance the robustness of the generator by performing SFT on high-quality data. In pipeline frameworks, models like DPA-RAG (Dong et al., 2024a), REAR (Wang et al., 2024), Self-RAG (Asai et al., 2024), and RankRAG (Yu et al., 2024) enhance the robustness of the generator by explicitly re-ranking the retrieved documents. Furthermore, VTG (Sun et al., 2023) improves citation quality by introducing explicit citation verification and modification. However, these methods still struggle to align the generator with the RAG requirements fully.

5.2 Fine-tuning Approaches for LMs

The mainstream approaches for fine-tuning language models in generative tasks include SFT (Radford, 2018) and reinforcement learning from human feedback (RLHF) (Christiano et al., 2017). SFT involves constructing input-output pairs to teach the model how to complete tasks. RLHF methods like DPO (Rafailov et al., 2023) incorporate training data that includes both superior and inferior outputs, allowing the model to learn preference information. RLHF is more effective than SFT in aligning the model with task preferences.

6 Conclusion

In this work, we propose PA-RAG, a method for optimizing the generator of RAG systems to align with specific RAG requirements comprehensively. The training process includes instruction fine-tuning and multi-perspective preference optimization. We conducted extensive experiments on four QA benchmarks and three LLMs, demonstrating that PA-RAG can significantly enhance the generator’s response informativeness, response robustness, and citation quality.

Limitations

Our method requires fine-tuning in four stages, including one instruction fine-tuning stage and three preference optimization stages. This results in a cumbersome search for the optimal hyperparameter settings during training. We have presented the hyperparameter details in Appendix A to prevent other researchers from duplicating the search for the best hyperparameter settings.

Ethics Statement

This work was conducted in strict compliance with the ACL Ethics Policy. All datasets and models used for experiment are publicly available. Furthermore, our work aims to explore a RAG generator training method. We do not foresee any negative ethical impacts arising from our work.

Acknowledgement

This work was supported by the Shanghai "Science and Technology Innovation Action Plan" Project (No.23511100700).

References

- Akari Asai, Zeqiu Wu, Yizhong Wang, Avirup Sil, and Hannaneh Hajishirzi. 2024. [Self-rag: Learning to retrieve, generate, and critique through self-reflection](#). In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net.
- Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. 2009. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*, pages 41–48.
- Jonathan Berant, Andrew Chou, Roy Frostig, and Percy Liang. 2013. [Semantic parsing on freebase from question-answer pairs](#). In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, EMNLP 2013, 18-21 October 2013, Grand Hyatt Seattle, Seattle, Washington, USA, A meeting of SIGDAT, a Special Interest Group of the ACL*, pages 1533–1544. ACL.
- Paul F Christiano, Jan Leike, Tom Brown, Miljan Martić, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30.
- Guanting Dong, Yutao Zhu, Chenghao Zhang, Zechen Wang, Zhicheng Dou, and Ji-Rong Wen. 2024a. [Understand what LLM needs: Dual preference alignment for retrieval-augmented generation](#). *CoRR*, abs/2406.18676.
- Qian Dong, Yiding Liu, Qingyao Ai, Zhijing Wu, Haitao Li, Yiqun Liu, Shuaiqiang Wang, Dawei Yin, and Shaoping Ma. 2024b. [Unsupervised large language model alignment for information retrieval via contrastive feedback](#). In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2024, Washington DC, USA, July 14-18, 2024*, pages 48–58. ACM.
- Feiteng Fang, Yuelin Bai, Shiwen Ni, Min Yang, Xiaojun Chen, and Ruifeng Xu. 2024. [Enhancing noise robustness of retrieval-augmented language models with adaptive adversarial training](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024, Bangkok, Thailand, August 11-16, 2024*, pages 10028–10039. Association for Computational Linguistics.
- Robert M. French and Nick Chater. 2002. [Using noise to compute error surfaces in connectionist networks: A novel means of reducing catastrophic forgetting](#). *Neural Comput.*, 14(7):1755–1769.
- Tianyu Gao, Howard Yen, Jiatong Yu, and Danqi Chen. 2023a. [Enabling large language models to generate text with citations](#). In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, Singapore, December 6-10, 2023*, pages 6465–6488. Association for Computational Linguistics.
- Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Qianyu Guo, Meng Wang, and Haofen Wang. 2023b. [Retrieval-augmented generation for large language models: A survey](#). *CoRR*, abs/2312.10997.
- Or Honovich, Roei Aharoni, Jonathan Herzig, Hagai Taitelbaum, Doron Kukliansky, Vered Cohen, Thomas Scialom, Idan Szpektor, Avinatan Hassidim, and Yossi Matias. 2022. [TRUE: re-evaluating factual consistency evaluation](#). In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, pages 3905–3920. Association for Computational Linguistics.
- Chengyu Huang, Zeqiu Wu, Yushi Hu, and Wenya Wang. 2024. [Training language models to generate text with citations via fine-grained rewards](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024, Bangkok, Thailand, August 11-16, 2024*, pages 2926–2949. Association for Computational Linguistics.
- Mandar Joshi, Eunsol Choi, Daniel S. Weld, and Luke Zettlemoyer. 2017. [Triviaqa: A large scale distantly supervised challenge dataset for reading comprehension](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume*

- I: *Long Papers*, pages 1601–1611. Association for Computational Linguistics.
- Tom Kwiatkowski, Jennimaria Palomaki, Olivia Redfield, Michael Collins, Ankur P. Parikh, Chris Alberti, Danielle Epstein, Illia Polosukhin, Jacob Devlin, Kenton Lee, Kristina Toutanova, Llion Jones, Matthew Kelcey, Ming-Wei Chang, Andrew M. Dai, Jakob Uszkoreit, Quoc Le, and Slav Petrov. 2019. [Natural questions: a benchmark for question answering research](#). *Trans. Assoc. Comput. Linguistics*, 7:452–466.
- Xuechen Li, Tianyi Zhang, Yann Dubois, Rohan Taori, Ishaan Gulrajani, Carlos Guestrin, Percy Liang, and Tatsunori B. Hashimoto. 2023. AlpacaEval: An automatic evaluator of instruction-following models. https://github.com/tatsu-lab/alpaca_eval.
- Jianmo Ni, Chen Qu, Jing Lu, Zhuyun Dai, Gustavo Hernández Ábrego, Ji Ma, Vincent Y. Zhao, Yi Luan, Keith B. Hall, Ming-Wei Chang, and Yinfei Yang. 2022. [Large dual encoders are generalizable retrievers](#). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 9844–9855. Association for Computational Linguistics.
- Krishna Pillutla, Swabha Swayamdipta, Rowan Zellers, John Thickstun, Sean Welleck, Yejin Choi, and Zaïd Harchaoui. 2021. [MAUVE: measuring the gap between neural text and human text using divergence frontiers](#). In *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pages 4816–4828.
- Alec Radford. 2018. Improving language understanding by generative pre-training.
- Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D. Manning, Stefano Ermon, and Chelsea Finn. 2023. [Direct preference optimization: Your language model is secretly a reward model](#). In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.
- Ivan Stelmakh, Yi Luan, Bhuwan Dhingra, and Ming-Wei Chang. 2022. [ASQA: factoid questions meet long-form answers](#). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 8273–8288. Association for Computational Linguistics.
- Hao Sun, Hengyi Cai, Bo Wang, Yingyan Hou, Xiaochi Wei, Shuaiqiang Wang, Yan Zhang, and Dawei Yin. 2023. Towards verifiable text generation with evolving memory and self-reflection. *arXiv preprint arXiv:2312.09075*.
- S. M. Towhidul Islam Tonmoy, S. M. Mehedi Zaman, Vinija Jain, Anku Rani, Vipula Rawte, Aman Chadha, and Amitava Das. 2024. [A comprehensive survey of hallucination mitigation techniques in large language models](#). *CoRR*, abs/2401.01313.
- Miles Turpin, Julian Michael, Ethan Perez, and Samuel R. Bowman. 2023. [Language models don’t always say what they think: Unfaithful explanations in chain-of-thought prompting](#). In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.
- Yuhao Wang, Ruiyang Ren, Junyi Li, Wayne Xin Zhao, Jing Liu, and Ji-Rong Wen. 2024. Rear: A relevance-aware retrieval-augmented framework for open-domain question answering. *arXiv preprint arXiv:2402.17497*.
- Ori Yoran, Tomer Wolfson, Ori Ram, and Jonathan Berant. 2024a. [Making retrieval-augmented language models robust to irrelevant context](#). In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net.
- Ori Yoran, Tomer Wolfson, Ori Ram, and Jonathan Berant. 2024b. [Making retrieval-augmented language models robust to irrelevant context](#). In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net.
- Yue Yu, Wei Ping, Zihan Liu, Boxin Wang, Jiaxuan You, Chao Zhang, Mohammad Shoeybi, and Bryan Catanzaro. 2024. Rankrag: Unifying context ranking with retrieval-augmented generation in llms. *arXiv preprint arXiv:2407.02485*.
- Junda Zhu, Lingyong Yan, Haibo Shi, Dawei Yin, and Lei Sha. 2024. Atm: Adversarial tuning multi-agent system makes a robust retrieval-augmented generator. *arXiv preprint arXiv:2405.18111*.

A Training Hyperparameters

We utilized full fine-tuning for all training stages and employed the same hyperparameter settings for all models.

During the instruction fine-tuning phase, we set the batch size to 128, the learning rate to $2e-5$, and trained for one epoch.

In the preference optimization phase, we set the batch size to 64 and trained for one epoch for all stages. For the optimization stages of response informativeness and response robustness, the learning rate is $2e-6$. In the citation quality optimization stage, the learning rate is $2e-7$.

B Compute Resources

Our experiments were conducted on a server equipped with 1TB of memory and 4 Nvidia A800 80G GPUs.

C Training Data Statistics

	IFT	RI	RR	CQ
ASQA	1,714	1,046	962	631
WebQ	1,681	326	357	653
NQ	55,463	10,416	12,080	21,241
Sum	58,858	11,788	13,399	22,525

Table 5: Training data statistics

Detailed information on the data size and composition for each training stage can be found in Table 5.

C.1 Distribution of Citation Complexity in Instruction Fine-tuning

In the instruction fine-tuning training data, the distribution of the number of citations per claim is shown in Table 6.

	1	2	3	4	5
ASQA	70.98%	24.17%	4.24%	0.46%	0.05%
WebQ	71.43%	21.54%	6.26%	0.65%	0.05%
NQ	71.87%	21.44%	6.00%	0.52%	0.16%

Table 6: Distribution of Citation Complexity

C.2 The Disparity Between “Chosen” and “Rejected” in Preference Data.

As shown in Table 7, the disparity in correctness and citation quality between the “Chosen” and “Rejected” in the preference data are presented. We can observe that the preference differences within the preference data are substantial enough to provide relatively rich preference information.

Dataset	Preference	Correct	Citation	
		EM	Rec	Prec
Response Informativeness				
ASQA	Chosen	100	100	100
	Rejected	38.33	77.59	75.51
WebQ	Chosen	100	100	100
	Rejected	21.14	72.90	71.00
NQ	Chosen	100	100	100
	Rejected	9.07	70.53	69.62
Response Robustness				
ASQA	Chosen	100	100	100
	Rejected	37.30	76.62	69.19
WebQ	Chosen	100	100	100
	Rejected	21.14	74.71	67.51
NQ	Chosen	100	100	100
	Rejected	6.08	73.19	67.40
Citation Quality				
ASQA	Chosen	-	100	100
	Rejected	-	27.42	12.13
WebQ	Chosen	-	100	100
	Rejected	-	25.88	13.17
NQ	Chosen	-	100	100
	Rejected	-	27.15	13.13

Table 7: Preference disparity within the preference data

D Prompt

Prompt for ChatGPT-3.5

Instruction: Write an accurate, engaging, and concise answer for the given question using only the provided search results (some of which might be irrelevant) and cite them properly. Use an unbiased and journalistic tone. Always cite for any factual claim. When citing several search results, use [1][2][3]. Cite at least one document and at most three documents in each sentence. If multiple documents support the sentence, only cite a minimum sufficient subset of the documents.

Question: {Question}

The final answer should contain the following short answers: {Short answers}

Documents: {Documents}

Answer:

E Examples of Adjusting the Citation Number

E.1 Rejected Output for Response Informativeness

When constructing the DPO training data for response informativeness, a question and five golden documents were given, numbered [1][2][3][4][5]. When constructing the rejected output, we would remove some golden doc-

uments, such as documents [1] and [3], leaving documents [2], [4], and [5]. To maintain a unified input format, the documents will be renumbered starting from 1. The mapping of the new and old document numbers is shown as Table 8.

Old ID	New ID
[2]	[1]
[4]	[2]
[5]	[3]

Table 8

The generator outputs an answer based on the question and the remaining documents: Claim1[2]. Claim2[1][3].

Since the citation numbers in the answer generated by the generator are the new document numbers, and the documents in the input part of the DPO training data are the old numbers, we need to revert the new numbers to the old ones using to the table above. The restored answer is as follows: Claim1[4]. Claim2[2][5].

E.2 Chosen Output for Response Robustness

When constructing the training data for response robustness, given a question and documents numbered [1][2][3][4][5][6][7][8], where [1][3][4][7] are golden documents and [2][5][6][8] are noisy documents. When constructing the chosen output, we will remove all the noisy documents. The mapping of the new and old document numbers is shown as Table 9.

Old ID	New ID
[1]	[1]
[3]	[2]
[4]	[3]
[7]	[4]

Table 9

ChatGPT outputs an answer based on the question and the remaining documents: Claim1[1][2]. Claim2[3][4].

Restore the new numbers to the old ones according to the table above. The restored answer is as follows: Claim1[1][3]. Claim2[4][7].

F Detailed Analysis of Why Rejecting Irrelevant Documents is a More Advanced Skill

As shown in Table 3, in our experiments, skipping the training for informativeness (learning to utilize relevant documents) and directly conducting the training for robustness (learning to ignore irrelevant documents) leads to a decline in model performance. This is because the model may not realize that the documents it needs to ignore are the irrelevant ones, resulting in indiscriminate neglect of most documents. In contrast, performing training for informativeness first and then training for robustness improves performance. This is because the model has already developed a basic ability to utilize documents and can recognize that it should ignore only the irrelevant documents. Therefore, ignoring irrelevant documents is a more advanced skill than fully utilizing every relevant document. If the results were the opposite, it would be counterintuitive and inexplicable.

G Detailed Analysis of Why Optimizing Citation Quality Might Have a Negative Impact on citation Recall.

It is challenging to simultaneously improve both the precision and recall of citation quality. While enhancing recall, the model may become stricter in assessing the relevance of documents, which can lead to a decrease in precision. Therefore, we introduced the F1 score to evaluate citation quality from a more balanced perspective. Our experimental results indicate that optimizing response informativeness and response robustness can enhance answer correctness, while optimizing citation quality can improve the citation F1 score.

H Further analysis of whether PA-RAG affects the fluency of responses

We follow ALCE (Gao et al., 2023a), using MAUVE (Pillutla et al., 2021) to evaluate the fluency of the model’s responses. The fluency of the text output by models at various training stages on the ASQA dataset is shown in the Table 10:

We found that PA-RAG does not affect the fluency of the model’s output.

MAUVE score	LLAMA2-7B	LLAMA2-13B	LLAMA3-8B
Base Generator	80.10	72.40	82.30
Base+IFT	84.99	84.00	80.86
Base+IFT+RI	83.88	81.41	85.87
Base+IFT+RI+RR	84.77	80.72	80.73
Base+IFT+RI+RR+CQ	85.42	82.20	83.21

Table 10: The results of the MAUVE score for each training stages. IFT stands for instruction fine-tuning, RI denotes the optimization of response informativeness, RR indicates the optimization of response robustness, and CQ represents the optimization of citation quality.

I Specific Criteria and Examples of Inferior Answers that are Incomplete

Inferior answers refer to responses that do not meet the criteria for correct answers. The criteria for the correct answers are shown in § 2.1. Transform these into criteria for inferior answers is as follows: Let G represent the generator, x denotes the input, $D = \{d_1, d_2, \dots, d_n\}$ represent the documents retrieved by the retriever, $A = \{a_1, a_2, \dots, a_n\}$ represent the short answers contained in the documents, and y denote the response generated by the generator. The correct response by the generator can be expressed as:

$$y = G(x, D), \exists a_i \in A, C(y, a_i) = \text{False} \quad (6)$$

where $C(y, a_i) = \text{False}$ implies that the answer a_i is not included in y .

Specific examples of inferior answers are as follows:

Prompt for ChatGPT-3.5
<p>Question: What is the name of the pirate in SpongeBob?</p> <p>Golden short answers: ['Patchy', 'Burger Beard']</p> <p>Inferior long answer: The pirate in SpongeBob is named Patchy the Pirate[1]. Patchy the Pirate is the president of the SpongeBob SquarePants fan club and is portrayed as a tyrannical pirate in the plot of the SpongeBob SquarePants film series[1][2].</p>

In this inferior answer, the answer ‘‘Burger Beard’’ is missing. We determine whether an answer is inferior by checking if any golden short answer is not included in the long answer.

J LLM Evaluation

As an accuracy evaluation metric, Exact Match (EM) cannot directly reflect the response informativeness and response robustness. We follow AlpacaEval (Li et al., 2023) and use the LLM (GPT-3.5-Turbo) evaluation to qualitatively judge

PA-RAG v.s. IFT Only	ASQA	WebQ	NQ	TQA
LLAMA2-7B-CHAT	65.30	65.11	67.09	69.39
LLAMA2-13B-CHAT	73.10	70.72	72.39	74.78
LLAMA3-8B-INSTRUCT	71.41	66.98	69.60	71.09

Table 11: The results of the win rate after preference alignment (PA-RAG) v.s. before preference alignment (instruction fine-tuning only). A higher success rate denotes a larger improvement in response informativeness and response robustness attributed to preference alignment.

whether the response informativeness and response robustness increase after RAG preference alignment. The prompt for the LLM judgement is as follows:

Prompt for ChatGPT-3.5
<p>I want you to create a leaderboard of different of large-language models. To do so, I will give you the instructions (prompts) given to the models, and the responses of two models. Please rank the models according to the correctness of their answers and whether they make use of valuable documents and avoid being interfered with by irrelevant documents. All inputs and outputs should be python dictionaries. Here is the prompt: {instruction}</p> <p>Here are the outputs of the models:</p> <pre>[{"model": "model_1", "answer": "{output_1}"}, {"model": "model_2", "answer": "{output_2}"}]</pre> <p>Now please rank the models by the quality of their answers, so that the model with rank 1 has the best output. Then return a list of the model names and ranks, i.e., produce the following output:</p> <pre>[{'model': <model-name>, 'rank': <model-rank>}, {'model': <model-name>, 'rank': <model-rank>}]</pre> <p>Your response must be a valid Python dictionary and should contain nothing else because we will directly execute it in Python. Please provide the ranking that the majority of humans would give.</p>

The results of the win rate after v.s. before preference alignment are shown in Table 11. We found that on all models and datasets, generators that have undergone RAG preference alignment have better response informativeness and response robustness.

K Case Study

To more intuitively demonstrate the improvements of preference optimization on response informativeness, response robustness, and citation quality, we randomly sampled several outputs from the LLAMA2-13B-CHAT before and after preference alignment (PA) for comparison. The examples are

as follows.

Take Case 1 as an example, we found that the generator without PA-RAG failed to utilize the document containing the answer and was misled by the noisy document, mistakenly identifying Erika Flores as the actress who played Ingrid, resulting in the incorrect citation. In contrast, the generator with PA-RAG training correctly used the document containing the answer and provided a correct response, demonstrating better response informativeness, response robustness, and citation quality.

Case 1: Sampled from the ASQA

Question: Who played Ingrid on Dr. Quinn, Medicine Woman?

Documents:

[1] she was featured throughout the sixth season, when her character fell in love with Jake Slicker. The role of Anthony (Grace and Robert E.'s adopted son) was played by Brenden Jefferson for four episodes in season four. He was replaced by Brandon Hammond, who continued in the role throughout seasons five and six. Jennifer Youngs did not begin playing Ingrid until the character's second appearance; the first time the character appeared, she was played by Ashley Jones. "Dr. Quinn" was best known for its large, supporting cast, and its high-concept storytelling. The series often used its semi-historical setting as a

[2] Erika Flores Erika Flores (born November 22, 1979 in Grass Valley, Nevada County, California) is an American former child actress. She is known for her role as the first Colleen Cooper in the "Dr. Quinn, Medicine Woman" TV series; Flores left the role in 1995, in the middle of the third season, and the role was then recast with Jessica Bowman taking over as Colleen. Flores guest-starred on the "" episode "". She played Marissa Flores, one of three young science-fair winners aboard the USS "Enterprise". In January 2009, Flores guest starred in "Big Baby", a fifth-season episode of the

[3] Heidi Kozak Heidi Karin Kozak (born June 22, 1963) is a Danish-born American actress. She is perhaps best known for her roles in several late-1980s horror films, including "Slumber Party Massacre 2", "" (both 1987), and "Society" (1989). She also starred in the first season of "Dr. Quinn, Medicine Woman" in the supporting role of Emily Donovan. Although she received an on-screen departure, it is unknown why she was written out of the series. Additionally, she appears in the behind-the-scenes documentary "Sleepless Nights: Revisiting the Slumber Party Massacres". Starting in 2005 she began charity work that involved baking different kinds

[4] Jessica Bowman Jessica Robyn Bowman (born November 26, 1980) is an American actress known for her role as Colleen Cooper on "Dr. Quinn, Medicine Woman". Bowman was born in Walnut Creek, California. She began acting in community theater and commercials, and appeared as Darcy on "The Road Home". After guest appearances on "Boy Meets World", "NYPD Blue", and "Baywatch", Bowman assumed the role of Colleen Cooper on the popular drama "Dr. Quinn, Medicine Woman" in season three, episode 15. She took over this role after the original portrayal, Erika Flores, did not renew her contract. Bowman won the Best Performance

Short answers: ['Ashley Jones', 'Jennifer Youngs']

Output from the generator without PA: Erika Flores played Ingrid on "Dr. Quinn, Medicine Woman"[2][3].

Output from the generator with PA: Ashley Jones and Jennifer Youngs both played the character Ingrid on "Dr. Quinn, Medicine Woman." [1]. Ashley Jones portrayed Ingrid in her first appearance, while Jennifer Youngs took over the role for subsequent appearances [1]. Additionally, Erika Flores played the first Colleen Cooper in the series before Jessica Bowman took over the role [2].

Case 2: Sampled from the WebQuestions

Question: where did bristol palin go to school?

Documents:

[1] aunt and uncle where she attended West Anchorage High School. After returning to Wasilla, she attended and then graduated from Wasilla High School in May 2009. Palin first became the subject of media attention when her pregnancy was announced during her mother's unsuccessful run for Vice President. In February 2009, she told Fox News that abstinence is "not realistic at all," but that she would like it to become more accepted among people her age. At age 18 in May 2009, Palin appeared on the "Today" show and "Good Morning America" in recognition of the National Day to Prevent Teen

[2] season, where she was eliminated in the fourth week of competition. Palin was born on October 18, 1990, in Wasilla, Alaska to parents Todd and Sarah (née Heath) Palin. She was named "Bristol" for the Bristol Inn where her mother Sarah was employed; Bristol, Connecticut, the headquarters city of ESPN, where her mother Sarah hoped to work as a sportscaster; and the Bristol Bay area of Alaska, where her father Todd grew up. Palin was raised largely in Wasilla and attended Juneau-Douglas High School when her mother was governor of Alaska. During 2008, she briefly lived in Anchorage with her

[3] Wasilla High School Wasilla High School (WHS) is a public secondary school in Wasilla, Alaska, United States, serving students in grades 9–12. The school is part of the Matanuska-Susitna Borough School District, with admission based primarily on the locations of students' homes. The school earned widespread media attention in 2008 following former pupil Governor Sarah Palin's nomination as the Republican vice-presidential running mate to John McCain in the 2008 United States presidential election. As of 22 February 2018, there are fifteen AP classes offered. WHS also participates in the University of Alaska Anchorage's Tech Prep program, which allows students to

[4] school (Palmer Jr. Middle School) and one high school (Palmer High School). Palmer High is of a modest size with about 900 students and approximately 60 staff members. In 1999, Palmer High became the first school in Alaska to offer the International Baccalaureate program. About ten minutes from Palmer are "cross-town" rivals, Colony Middle and High Schools. Palmer is also home to Alaska Bible College, Alaska's only accredited school of theology and ministry, and Matanuska-Susitna College, an extended college of the University of Alaska Anchorage. There is also the Palmer Alaska Job Corps, 800 E. Lynn Martin Drive, across from

[5] Pacific University in Honolulu for a semester in the fall of 1982 and then to North Idaho College, a community college in Coeur d'Alene, for the spring and fall semesters of 1983. She enrolled at the University of Idaho in Moscow for an academic year starting in August 1984 and then attended Matanuska-Susitna College in Alaska in the fall of 1985. Palin returned to the University of Idaho in January 1986 and received her bachelor's degree in communications with an emphasis in journalism in May 1987. In June 2008, the Alumni Association of North Idaho College gave Palin its Distinguished

Short answers: ['Wasilla High School', 'West Anchorage High School', 'Juneau-Douglas High School']

Output from the generator without PA: Bristol Palin attended West Anchorage High School and graduated from Wasilla High School in May 2009[1][3].

Output from the generator with PA: Bristol Palin attended West Anchorage High School and later graduated from Wasilla High School in May 2009[1]. Additionally, she briefly lived in Anchorage and attended Juneau-Douglas High School during her mother's governorship of Alaska[2].

Case 3: Sampled from the Natural Questions

Question: name the process of fusion of an egg with a sperm

Documents:

[1] the oocyte. Fusion between the oocyte plasma membranes and sperm follows and allows the sperm nucleus, the typical centriole, and atypical centriole that is attached to the flagellum, but not the mitochondria, to enter the oocyte. The protein CD9 likely mediates this fusion in mice (the binding homolog). The egg "activates" itself upon fusing with a single sperm cell and thereby changes its cell membrane to prevent fusion with other sperm. Zinc atoms are released during this activation. This process ultimately leads to the formation of a diploid cell called a zygote. The zygote divides to form a blastocyst and,

[2] fertilization process. Human fertilization is the union of a human egg and sperm, usually occurring in the ampulla of the fallopian tube. The result of this union is the production of a zygote cell, or fertilized egg, initiating prenatal development. Scientists discovered the dynamics of human fertilization in the nineteenth century. The process of fertilization involves a sperm fusing with an ovum. The most common sequence begins with ejaculation during copulation, follows with ovulation, and finishes with fertilization. Various exceptions to this sequence are possible, including artificial insemination, "in vitro" fertilization, external ejaculation without copulation, or copulation shortly

[3] and once bound together, the sperm and the egg fuse together to form two pronuclei. These pronuclei supply the zygote with the genetic material necessary for the formation of an embryo. Additionally, once the fusion of the sperm and oocyte is complete, phospholipase C zeta is released from the sperm. As the sperm arrive at the egg, it is the egg that chooses the sperm and pulls it towards her. The selected sperm actually tries to swim away from the egg but is tethered to the egg by female hormones. The membrane around the egg literally opens up and swallows

[4] (microgametophyte) containing two haploid generative (sperm) cells and a tube nucleus. When a pollen grain makes contact with the female stigma, the pollen grain forms a pollen tube that grows down the style into the ovary. In the act of fertilization, a male sperm nucleus fuses with the female egg nucleus to form a diploid zygote that can then develop into an embryo within the newly forming seed. Upon germination of the seed, a new plant can grow and mature. The adaptive function of meiosis is currently a matter of debate. A key event during meiosis in a diploid cell

[5] Human fertilization Human fertilization is the union of a human egg and sperm, usually occurring in the ampulla of the fallopian tube. The result of this union is the production of a zygote cell, or fertilized egg, initiating prenatal development. Scientists discovered the dynamics of human fertilization in the nineteenth century. The process of fertilization involves a sperm fusing with an ovum. The most common sequence begins with ejaculation during copulation, follows with ovulation, and finishes with fertilization. Various exceptions to this sequence are possible, including artificial insemination, "in vitro" fertilization, external ejaculation without copulation, or copulation shortly after ovulation.

Short answers: [['fertilization'], ['Human fertilization'], ['fertilization'], ['fertilization']]

Output from the generator without PA: The process of fusion of an egg and a sperm involves several stages[1]. The fusion between the oocyte plasma membranes and sperm occurs, which allows the sperm nucleus, mitochondria, and centriole to enter the oocyte [1]. This process leads to the formation of a diploid cell called a zygote [1]. The zygote divides to form a blastocyst, marking the beginning of prenatal development [1]. The sperm and egg fuse together to form two pronuclei, supplying the zygote with the genetic material necessary for the formation of an embryo [3]. This fusion ultimately results in the formation of a fertilized egg, known as zygote, initiating prenatal development [2][5].

Output from the generator with PA: The process of fusion of an egg with a sperm is called fertilization[2]. This process involves the union of a human egg and sperm, typically occurring in the ampulla of the fallopian tube, resulting in the production of a zygote cell, or fertilized egg, which initiates prenatal development [2][5]. The fusion between the oocyte plasma membranes and sperm allows the sperm nucleus, centriole, and other cellular structures to enter the oocyte, ultimately leading to the formation of a diploid cell called a zygote [1]. This fusion is crucial for the transfer of genetic material from the sperm to the egg, enabling the formation of an embryo [3].

Case 4: Sampled from the Natural Questions

Question: who was kat slater's sisters in eastenders

Documents:

[1] Slater family (EastEnders) The Slater family are a fictional family from the BBC soap opera "EastEnders", that have appeared since episode 1997, first broadcast in September 2000. The family are introduced as a central family unit consisting of father Charlie Slater (Derek Martin), grandmother Mo Harris (Laila Morse), Charlie's daughters, Kat Slater (Jessie Wallace), Lynne Slater (Elaine Lordan), Little Mo Morgan (Kacey Ainsworth) and Zoe Slater (Michelle Ryan), and Lynne's partner Garry Hobbs (Ricky Groves). A fifth sister, Belinda Peacock (Leanne Lakey), was introduced a year later. The family were the eighth to be introduced in the soap's history, replacing

[2] Kat Slater Kathleen "Kat" Moon (also Slater) is a fictional character from the BBC soap opera "EastEnders", played by Jessie Wallace. She was also played by Kate Peck in a flashback in 2001. Kat is the daughter of Viv Slater and Charlie Slater (Derek Martin), and arrived with her sisters Lynne Hobbs (Elaine Lordan) and Little Mo Morgan (Kacey Ainsworth). Her daughter Zoe Slater (Michelle Ryan) is initially believed to be her sister as well until Kat reveals she was raped by her uncle Harry Slater (Michael Elphick) as a child. She later marries Alfie Moon (Shane Richie). It is

[3] ends of the country, and in some cases beyond." Slater family (EastEnders) The Slater family are a fictional family from the BBC soap opera "EastEnders", that have appeared since episode 1997, first broadcast in September 2000. The family are introduced as a central family unit consisting of father Charlie Slater (Derek Martin), grandmother Mo Harris (Laila Morse), Charlie's daughters, Kat Slater (Jessie Wallace), Lynne Slater (Elaine Lordan), Little Mo Morgan (Kacey Ainsworth) and Zoe Slater (Michelle Ryan), and Lynne's partner Garry Hobbs (Ricky Groves). A fifth sister, Belinda Peacock (Leanne Lakey), was introduced a year later. The family were the

[4] Discussing the character's death, Digital Spy's Kris Green praised Crace's acting. He compared Danielle's revelation that Ronnie was her mother to the "EastEnders" storyline several years previously, when Kat Slater (Jessie Wallace) revealed she was actually the mother of her supposed sister Zoe (Michelle Ryan). Green opined: "Although it didn't quite capture the magic of [the Kat-Zoe scene] it definitely comes an "extremely" close second." He concluded: "I'd probably go as far as to say that it's one of the best episodes "EastEnders" has produced in a long time". The episode was selected as recommended viewing by "The Guardian" Sarah

[5] that she wants to get back together with him, but he tells her that he is not ready as he does not want to hurt her again. This causes Kat to reevaluate her life, so she visits the convent where she gave birth to Zoe to seek some answers, though it is now a nun's care home. A nun, Sister Ruth (June Whitfield), remembers her going into labour and they talk about the past and the future. Kat then leaves and goes to the police station to make a statement about Harry. Ruth checks their records and discovers that Kat

Short answers: [['Lynne', 'Little Mo'], ['Lynne', 'Little Mo', 'Zoe'], ['Lynne', 'Little Mo', 'Zoe']]

Output from the generator without PA: Kat Slater had several sisters in EastEnders[1][3]. She was introduced as part of the central family unit consisting of father Charlie Slater and his daughters, including Lynne Slater and Zoe Slater [1][2][3]. Additionally, a fifth sister, Belinda Peacock, was introduced a year after Kat's arrival [1].

Output from the generator with PA: Kat Slater had four sisters in "EastEnders." [1]. Her sisters were Lynne Slater, Little Mo Morgan, Zoe Slater, and later, Belinda Peacock [1]. Kat, portrayed by Jessie Wallace, was the daughter of Viv Slater and Charlie Slater, and she arrived with her sisters Lynne Hobbs and Little Mo Morgan [2]. Zoe Slater, played by Michelle Ryan, was initially believed to be her sister until Kat revealed that she was her daughter [2]. Belinda Peacock was introduced a year later [1].

Case 5: Sampled from the TriviaQA

Question: England won the 1966 World Cup beating Germany 4-2. Hurst scored 3 of England's goals who scored the other?

Documents:

[1] the game went to extra time. In the 98th minute, Hurst found himself on the scoresheet again; his shot hit the crossbar, bounced down onto the goal line, and was awarded as a goal. Debate has long raged over whether the ball crossed the line, with the goal becoming part of World Cup history. England's final goal was scored by Hurst again, as a celebratory pitch invasion began. This made Geoff Hurst the only player ever to have scored three times in a single World Cup final. BBC commentator Kenneth Wolstenholme's description of the match's closing moments has gone down

[2] over France in their final group game. Roger Hunt scores both of England's goals. 23 July 1966: England beat Argentina 1-0 in the World Cup quarter-final thanks to a goal by Geoff Hurst. 26 July 1966: England reach the World Cup final by beating Portugal 2-1 in the semi-final. Bobby Charlton scores both of England's goals. 30 July 1966: England win the World Cup with a 4-2 win over West Germany in extra time. Geoff Hurst scores a hat-trick, with Martin Peters scoring the other goal. !style="width:15em"|Competition!!style="width:15em"|Winner!!style="width:15em"|Runner-up Liverpool, FA Cup winners the previous season and league champions in 1964, won

[3] described at the time as a 4-4-2. The match is remembered for England's only World Cup trophy, Geoff Hurst's hat-trick - the first, and to date, only one ever scored in a FIFA World Cup Final - and the controversial third goal awarded to England by referee Gottfried Dienst and linesman Tofiq Bahramov. Both teams were strong throughout the tournament. Each won two and drew one of their three matches in the group stages. England did not concede a goal until their semi-final against Portugal. One of the enduring images of the celebrations in Wembley immediately after the game was

[4] 1962 FIFA World Cup. Ramsey confidently predicted that at the next tournament, England would win the trophy, and they did just that. The 1966 World Cup saw England win the World Cup in a controversial 4-2 victory over West Germany. The three goals scored by Geoff Hurst within 120 minutes, of which some are controversial, are the only hat trick to be achieved in a World Cup final to date. Bobby Moore was the captain on that day, whilst Munich air crash survivor Bobby Charlton also played. Moore's West Ham colleagues Geoff Hurst and Martin Peters scored that day. The [5] over Germany in the final held on home soil, at the old 1923 built Wembley stadium (since replaced by the new Wembley Stadium opened in 2007). In that final, Moore was the team captain and a central defender, while Wilson served as full back alongside him. Peters in midfield, and Hurst as one of the strikers, were the England goalscorers on the day of the final, with Hurst scoring the first goal in the first half, Peters scoring in the second half, and Hurst scoring two more in extra time to complete a famous World Cup hat-trick and secure victory.

Short answers: ['Martin Peters']

Output from the generator without PA: England won the 1966 World Cup by beating Germany 4-2[2]. Geoff Hurst scored 3 of England's goals [1].

Output from the generator with PA: Geoff Hurst scored 3 of England's goals in the 1966 World Cup final, while Martin Peters scored the other goal in England's 4-2 victory over Germany[2]. Hurst's hat-trick in the final remains the only one ever scored in a FIFA World Cup Final[3].

Case 6: Sampled from the TriviaQA

Question: Who was the only British Prime Minister to hold office during the reign of King Edward VIII?

Documents:

[1] used government papers in breach of the "Fifty-year rule" then in operation. Cabinet Secretary Sir Norman Brook persuaded the Prime Minister to demand amendments to conceal the degree of Cabinet involvement in the abdication of King Edward VIII (who was still alive in 1961) and the degree to which the civil servants Horace Wilson and Warren Fisher had demanded that the former King "reorder his private life" afterwards. Former Prime Minister Lord Avon, who cherished his (somewhat exaggerated) reputation as an opponent of "appeasement", complained that such a book by a serving Cabinet Minister might be thought to express official

[2] not like: Wilhelm II. Edward's difficult relationship with his nephew exacerbated the tensions between Germany and Britain. In April 1908, during Edward's annual stay at Biarritz, he accepted the resignation of British Prime Minister Sir Henry Campbell-Bannerman. In a break with precedent, Edward asked Campbell-Bannerman's successor, H. H. Asquith, to travel to Biarritz to kiss hands. Asquith complied, but the press criticised the action of the King in appointing a prime minister on foreign soil instead of returning to Britain. In June 1908, Edward became the first reigning British monarch to visit the Russian Empire, despite refusing to visit in

[3] office of Prime Minister has coincided with the reigns of 11 British monarchs (including a Regency during the incapacity of George III from 1811 to his death in 1820), to whom the Prime Minister has been constitutionally head of government to the sovereign's headship of state. Until 1837 the death of a sovereign led to Parliament being dissolved within six months which led to a general election. Results of such elections were: Stanley Baldwin is the only Prime Minister to have served three sovereigns in succession King George V, King Edward VIII and King George VI. Through being in office at

[4] Edward VIII Edward VIII (Edward Albert Christian George Andrew Patrick David; 23 June 1894 – 28 May 1972) was King of the United Kingdom and the Dominions of the British Empire, and Emperor of India, from 20 January 1936 until his abdication on 11 December the same year, after which he became the Duke of Windsor. Edward was the eldest son of King George V and Queen Mary. He was named Prince of Wales on his sixteenth birthday, nine weeks after his father succeeded as king. As a young man, he served in the British Army during the First World

[5] invested by King Edward VII at Buckingham Palace on 8 August 1902. (He was the only prime minister to be knighted during his term of office until Robert Menzies in 1963; various others were knighted after leaving the office; Sir Earle Page was already a knight when he briefly became prime minister in 1939.) He received an honorary LL.D. from the University of Cambridge in 1900, and honorary D.C.L. from the University of Oxford and LL.D. from the University of Edinburgh while visiting the United Kingdom in the summer of 1902. He also received the Freedom of the City of

Short answers: ['Stanley, 1st Earl Baldwin of Bewdley Baldwin', 'Stanley Baldwin, 1st Earl Baldwin of Bewdley, KG, PC', 'Stanley Baldwin', 'Stanley, 1st Earl Baldwin of Bewdley, Viscount Corvedale of Corvedale Baldwin', 'Power without responsibility', 'Stanley Baldwin, 1st Earl Baldwin of Bewdley', 'Baldwin, Stanley, 1st Earl Baldwin of Bewdley']

Output from the generator without PA: Iain Macleod was the only British Prime Minister to hold office during the reign of King Edward VIII[1][3].

Output from the generator with PA: Stanley Baldwin was the only British Prime Minister to hold office during the reign of King Edward VIII[3]. Edward VIII reigned from 20 January 1936 until his abdication on 11 December the same year, after which he became the Duke of Windsor[4]. Stanley Baldwin served as Prime Minister during this period, making him the only Prime Minister to have served three sovereigns in succession: King George V, King Edward VIII, and King George VI[3].

Case 7: Sampled from the TriviaQA

Question: Louis Smith won an individual silver medal in gymnastics at the 2012 London Olympics. For what discipline?

Documents:

[1] 2016 Rio Olympics, this time finishing behind teammate, and 2012 bronze medalist, Max Whitlock. Smith was also part of the Great Britain team that took the bronze in the men's artistic team all-around at the 2012 London Olympics. He is the only British gymnast to win Olympic medals in three separate Games, and only the second gymnast after Marius Urzică to win three successive Olympic pommel horse medals. A 2006 Commonwealth Games champion representing England in the pommel horse, in 2015 he became European champion, his first major international title representing Great Britain. Smith is a member of the Huntingdon

[2] Louis Smith (gymnast) Louis Antoine Smith, MBE (born 22 April 1989) is a retired British artistic gymnast. He received a bronze medal and two silver medals on the pommel horse at the 2008 Beijing Olympics, 2012 London Olympics and the 2016 Rio Olympics respectively, with the former marking the first time a British gymnast had placed in an Olympic event since 1928. He fell just short of gold in 2012, tying with Kristian Berki, but taking silver for a lower E or execution score. He followed this up with a second consecutive silver medal on the pommel horse at the [3] but fell on his dismount, lowering his execution score. Despite this, he still placed 3rd behind Krisztián Berki and Cyril Tommasone to win the Bronze medal — the only medal won by the British men and women at those championships. In 2012, Smith won the silver medal in the pommel horse finals in the European Men's Artistic Gymnastics Championships. Smith was part of the Great Britain team at the London Olympics. The team won a bronze medal at the men's artistic team all-around event on 30 July 2012, marking the first team medal for a British Olympic gymnastics team in

[4] Horse at the World Cup in Moscow. On 9 August 2008, Smith qualified in fifth place for the Olympic final of the men's pommel horse, and on 17 August he won a bronze medal in the finals, becoming first British man to win a medal in the individual gymnastics at the Olympic Games since Walter Tysall won a silver medal in 1908 and the first Briton generally to win a medal in gymnastics at the Olympic Games since the women's team won a bronze medal in 1928. He is also the second black male gymnast to win a medal in

[5] was harder than Gymnastics as everything was a new challenge. I never expected to win – I just went out every week with the aim of doing my best." Smith subsequently participated in the 2014 Christmas special with guest professional, Aliona Vilani. The couple won after dancing a Quickstep that scored a perfect 10 from all judges. Louis Smith (gymnast) Louis Antoine Smith, MBE (born 22 April 1989) is a retired British artistic gymnast. He received a bronze medal and two silver medals on the pommel horse at the 2008 Beijing Olympics, 2012 London Olympics and the 2016 Rio Olympics

Short answers: ['Gymnastics pommel horse', 'Side Horse', 'Pommel Horse', 'Pommel horse', 'Pommel horse (gymnastics)', 'Pommel horse', 'Pommel-Horse']

Output from the generator without PA: Louis Smith won an individual silver medal in gymnastics at the 2012 London Olympics[5].

Output from the generator with PA: Louis Smith won an individual silver medal in gymnastics at the 2012 London Olympics for the pommel horse discipline[2]. He is a retired British artistic gymnast who received a bronze medal and two silver medals on the pommel horse at the 2008 Beijing Olympics, 2012 London Olympics, and the 2016 Rio Olympics, respectively[2].

Case 8: Sampled from the TriviaQA

Question: Which Asian country hosted its first Formula One Grand Prix motor race in October 2010?

Documents:

[1] the inaugural teams in the ASEAN Basketball League which was founded in October 2009. Kranji Racecourse is run by the Singapore Turf Club and hosts several meetings per week, including international races—notably the Singapore Airlines International Cup. Singapore began hosting a round of the Formula One World Championship, the Singapore Grand Prix at the Marina Bay Street Circuit in 2008. It was the inaugural F1 night race, and the first F1 street race in Asia. It is considered a signature event on the F1 calendar. Kranji Racecourse is run by the Singapore Turf Club and hosts several meetings per week,

[2] 2011 Indian Grand Prix The 2011 Indian Grand Prix, formally the 2011 Formula 1 Airtel Grand Prix of India, was a Formula One motor race that was held on 30 October 2011 at the Buddh International Circuit in Greater Noida, Uttar Pradesh, India. It was the seventeenth round of the 2011 Formula One season and the first Formula One Grand Prix to take place in South Asia and first to take place in India. The 60-lap race was won by Red Bull Racing's Sebastian Vettel, after leading every lap of the race from pole position and setting the fastest lap

[3] 2010 Korean Grand Prix The 2010 Korean Grand Prix (formally the 2010 Formula 1 Korean Grand Prix) was a Formula One motor race held on 24 October 2010 at the Korea International Circuit in Yeongam, South Jeolla, South Korea. It was the seventeenth round of the 2010 Formula One season and the first Korean Grand Prix. The 55-lap race was won by Ferrari driver Fernando Alonso, who started from third position. Lewis Hamilton finished second for McLaren and Alonso's teammate Felipe Massa was third. This was the first time since the 2008 Australian Grand Prix that neither Red Bull was

[4] 2011 Korean Grand Prix The 2011 Korean Grand Prix, formally the 2011 Formula 1 Korean Grand Prix was a Formula One motor race that was held on 16 October 2011 at the Korea International Circuit in Yeongam, South Jeolla, South Korea. It was the sixteenth round of the 2011 Formula One season, the second running of the Korean Grand Prix, and the first race after Sebastian Vettel claimed the World Drivers' Championship. The 55-lap race was won by Vettel, after starting from second on the grid. Lewis Hamilton finished in second place for McLaren, and Mark Webber completed the podium

[5] held later than usual, to avoid a clash with the 2006 Commonwealth Games. For the first time, Bahrain hosted the first Grand Prix. Brazil hosted the last race, while Japan and China swapped their original dates. In 2006, the FIA announced the Belgian Grand Prix would not be part of the 2006 Formula One season, since the local authorities had started major repair work in Spa-Francorchamps. The Belgian Grand Prix returned in 2007, when Kimi Räikkönen took pole position and his 3rd Belgian Grand Prix win in a row. The calendar was initially announced as the same as for 2005,

Short answers: ['Nam Chosun', 'Korea, Republic Of', 'So. Korea', 'ISO 3166-1:KR', 'R.O.K.', 'Korea, Republic of', 'Korea, South', 'Daehan Minguk', 'Republic Of Korea', 'S. Korea', 'R. Korea', 'South Kroea', 'South korea', 'Republican Korea', 'Korea (R.O.K.)', 'The republic of korea', 'Korea South', 'ROK', 'Daehanminkook', 'South Korea/Economy', 'Daehanmingook', 'Korea (South)', 'Korea Republic', 'SKorea', 'S Korea', 'South-korea', 'Republican Korea', 'Republic of KOREA', 'Corea Republic', 'Daehanminkuk', 'Korea (Republic of)', 'Republic of korea', 'Korea (ROK)', 'South Koreans', 'Free Korea', 'Daehanminguk', 'South Corea', 'Modern Corea', 'South corea', 'Taehan-min'guk', 'S.Korea', 'South-Korea', 'Republic of Corea', 'S. Koreans', 'S. Korean', 'South Korean', 'Korea (Seoul)', 'Namjoseon', 'Republic of Korea', 'Korea, S', 'Korean Republic', 'Republic of South Korea', 'South Korea', 'Corean Republic']

Output from the generator without PA: The Indian Grand Prix in October 2010 marked India's first Formula One Grand Prix motor race, making it the first Formula One Grand Prix to take place in South Asia and the first to take place in India [2].

Output from the generator with PA: South Korea hosted its first Formula One Grand Prix motor race in October 2010[3]. The 2010 Korean Grand Prix, formally the 2010 Formula 1 Korean Grand Prix, took place on 24 October 2010 at the Korea International Circuit in Yeongam, South Jeolla, South Korea, marking the first Korean Grand Prix in the country's history[3].