SearchInstruct: Enhancing Domain Adaptation via Retrieval-Based Instruction Dataset Creation

Anonymous ACL submission

Abstract

Supervised Fine-Tuning (SFT) is essential for training large language models (LLMs), significantly enhancing critical capabilities such as instruction-following and in-context learning. Nevertheless, creating suitable training datasets tailored for specific domains remains challenging due to unique domain constraints and data scarcity.

In this paper, we propose *SearchInstruct*, an innovative method explicitly designed to construct high-quality instruction datasets for SFT. Our approach begins with a limited set of domain-specific, human-generated questions, which are systematically expanded using a large language model. Subsequently, domain-relevant resources are dynamically retrieved to generate accurate and contextually appropriate answers for each augmented question.

Experimental evaluation demonstrates that *SearchInstruct* notably enhances both the diversity and quality of SFT datasets, leading to measurable improvements in LLM performance within specialized domains. Additionally, we demonstrate that beyond dataset generation, the proposed method can also effectively facilitate tasks such as model editing, enabling efficient updates to existing models.

To facilitate reproducibility and community adoption, we will release the full implementation details, the complete set of generated instruction–response pairs, and the source code in a public repository upon acceptance.[1]

1 Introduction

The rapid progress of LLMs and their success in solving general-purpose natural language processing (NLP) tasks (Wei et al., 2021; Sanh et al., 2022) have led to a growing focus on adapting these models to specialized domains. Achieving this goal requires not only access to cleaned raw corpora, but also the construction of structured datasets suitable for the SFT stage (Ouyang et al., 2022).

Generating high-quality SFT data—especially for expert domains—presents several major challenges. The data must be both diverse and capable of reflecting the inherent complexity of real-world user queries. Existing approaches either rely on extensive manual annotation, as in InstructGPT (Ouyang et al., 2022), suffer from limited coverage and task diversity (Ziegler et al., 2024), or fail to simulate real user needs accurately. Furthermore, many methods depend solely on the internal knowledge of LLMs, which may be outdated or insufficient for domain-specific adaptation (Lewis et al., 2020; Lazaridou et al., 2022).

At the same time, LLMs have shown that—when provided with appropriate context—they can generate accurate and informative responses to even complex domain-specific questions (Izacard et al., 2022). This observation suggests that, by systematically generating diverse and realistic queries and retrieving relevant, up-to-date contextual information, it is possible to construct powerful training datasets without relying on human annotation. Such an approach enables the integration of external domain-specific knowledge into the training process, offering a path to updating and extending model capabilities (Guu et al., 2020; Wang et al., 2024).

In this paper, we introduce *SearchInstruct*, a novel framework for automatically generating supervised training data. The method operates by expanding a small set of seed questions using an LLM, then dynamically retrieving relevant documents to construct accurate, context-grounded answers. Our empirical results demonstrate that this approach not only increases the quality and diversity of SFT data, but also effectively enhances and updates LLM behavior in specialized domains by leveraging up-to-date external knowledge.

Our contributions are as follows:

[1] Repository link will be provided after review.

1. We introduce *SearchInstruct*, an automated framework for constructing SFT datasets by combining targeted document retrieval with grounded answer generation;
2. We design and implement a four-stage pipeline for query expansion, diversification, document retrieval, and answer synthesis;
3. We empirically show that, in certain specialized domains, this pipeline leads to improvements in model performance and instruction-following accuracy;
4. We demonstrate the effectiveness of *SearchInstruct* in model editing scenarios, enabling efficient updates to model outputs using newly acquired knowledge.

2 Related Work

Instruction tuning has emerged as a key strategy for adapting LLMs to follow user instructions. The FLAN approach by Wei et al. (Wei et al., 2021) demonstrated that fine-tuning on a broad range of instruction-formatted tasks leads to strong zero-shot generalization. Similarly, T0 by Sanh et al. (Sanh et al., 2022) employed multitask prompted training to enable cross-task generalization.

Wang et al. (Wang et al., 2022b) extended this paradigm with the Super-NaturalInstructions dataset, collecting over 1,600 crowdsourced tasks paired with declarative instructions to promote task diversity. InstructGPT (Ouyang et al., 2022) introduced a method for aligning LLMs with human preferences using instruction–response pairs ranked via human feedback. Despite their utility, these datasets are static, expensive to construct, and often insufficient for domain-specific applications or ongoing model updates.

To address limitations in scalability and recency, Self-Instruct (Wang et al., 2022a) bootstraps a base LLM to generate new instructions and responses from a small seed set, filtering outputs using heuristics. While it enhances general instruction-following capabilities, it still relies entirely on the model's internal knowledge, which may be outdated or incomplete for specialized domains.

Recent work has focused on automating instruction dataset creation to reduce dependence on manual annotation. The Alpaca project (Taori et al., 2023) used GPT-3.5 to generate 52,000 instruction–response pairs, enabling effective fine-tuning of a LLaMA model. Evol-Instruct by Xu et al. (Xu et al., 2023) introduced an evolutionary strategy to rewrite and complicate instructions, leading to the creation of WizardLM. InstructZero (Chen et al., 2023) proposed a label-free tuning method using reinforcement learning with black-box models. While these approaches improve scale and diversity, they generally lack grounding in real-world evidence and do not facilitate domain-specific adaptation.

Retrieval-augmented methods aim to improve factuality and relevance by incorporating external documents. RAG (Lewis et al., 2020) combines dense retrieval with generation to condition responses on retrieved passages. REALM (Guu et al., 2020) pre-trains language models with latent document retrieval to support open-domain QA. More recently, InstructRetro (Wang et al., 2024) integrates retrieval with instruction tuning to enhance grounding, and RA-DIT (Lin et al., 2024) jointly trains a retriever and generator using dual instruction signals. However, these methods often require costly retriever training and may not be ideal for rapid, low-resource domain updates.

SearchInstruct addresses the limitations of static and hallucinated instruction datasets by incorporating real-time document retrieval during data generation. Unlike Self-Instruct, which depends solely on internal knowledge, our method grounds both the instruction and the response in up-to-date, domain-specific documents. Unlike methods such as RAG and REALM, which apply retrieval at inference time, SearchInstruct applies it during dataset construction—allowing the fine-tuned model to internalize retrieved knowledge. This enables scalable, lightweight domain adaptation and continual knowledge refresh, bridging the gap between static SFT and real-world evolving needs.

3 Methodology

The core idea behind our method is that an LLM, when provided with relevant contextual information—even for domain-specific queries—can generate accurate and informative answers. These instruction–response pairs can then be used as high-quality data for SFT, improving the model's ability to follow instructions and generalize in specialized domains.

While prior work such as Self-Instruct (Wang et al., 2022a) demonstrated the value of generating large quantities of instruction–response

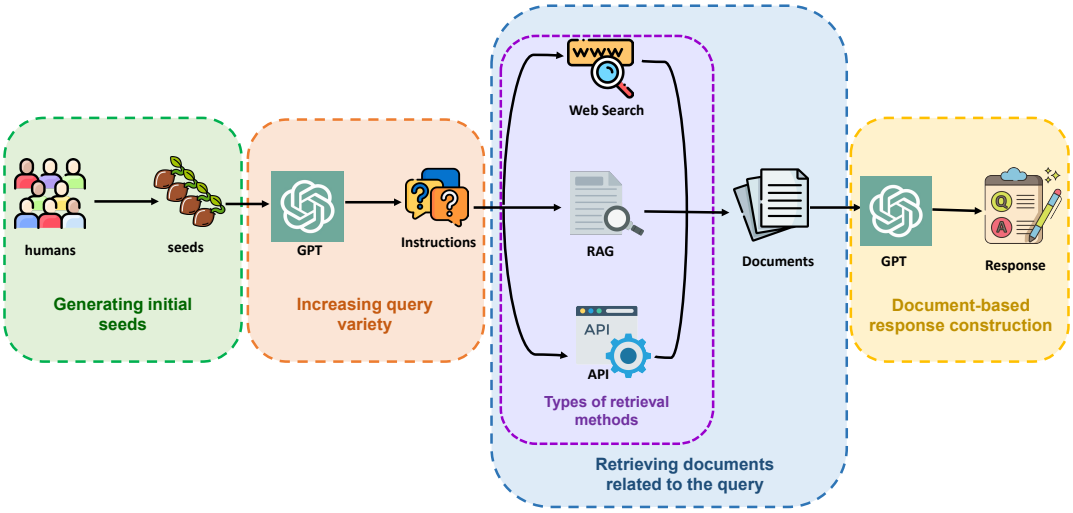


Figure 1: The four-stage *SearchInstruct* pipeline: seed generation, query expansion, document retrieval, and response construction.

examples, we argue that in the current landscape—where LLMs often outperform crowd-annotated answers—the primary focus should shift from generating more answers to diversifying the questions themselves. A broader and more realistic query distribution enables better coverage of domain-specific tasks, especially when answers are grounded in retrieved, up-to-date evidence.

To operationalize this idea, we propose a four-stage pipeline, illustrated in Figure 1, consisting of the following steps:

1. **Seed Generation**: Start with a small set of domain-relevant, human-written or curated instruction prompts;
2. **Query Expansion**: Use an LLM to generate diverse variations and paraphrases of the seed prompts;
3. **Document Retrieval**: Retrieve domain-specific and up-to-date documents relevant to each expanded query (using web search, RAG, or external APIs);
4. **Response Construction**: Generate accurate, context-aware answers grounded in the retrieved evidence.

In the following sections, we describe each stage of the pipeline in detail.

3.1 *SearchInstruct* Dataset Formulation

We formally define our instruction dataset $\mathcal{D}$ as a set of instruction–response pairs constructed through a four-stage pipeline.

Let $Q = \{q_1, q_2, \ldots, q_{|Q|}\}$ denote the initial set of domain-specific seed queries. From Q, we generate a set of additional instructions via LLM-based expansion. Let I denote the set of newly generated instructions, and define the final instruction pool as:

$$I^{\text{total}} = Q \cup I$$

For each instruction $i \in I^{\text{total}}$, we retrieve a set of relevant textual contexts (e.g., documents, passages), denoted $C_i = \{c_{i,1}, \ldots, c_{i,k}\}$. We then use an LLM to generate an answer A_i grounded in the retrieved context:

$$A_i = \text{LLM}_{\text{answer}}(i, C_i^{\text{filtered}})$$

The final instruction dataset is constructed as:

$$\mathcal{D} = \{(i, A_i) \mid i \in I^{\text{total}}\}$$

During training, the model M learns the conditional instruction-following mapping:

$$M(i) \rightarrow A_i$$

Note that while C_i is essential for answer generation during dataset construction, it is not required at inference or training time. This enables M to internalize external knowledge through SFT and improve response quality in dynamic or specialized domains.

3.2 Stage 1: Seed Generation

The *SearchInstruct* pipeline begins with the creation of a diverse set of seed queries $Q =$

$\{q_1, q_2, \ldots, q_{|Q|}\}$ that serve as the foundation for subsequent query expansion. Because all downstream instructions are generated based on these seeds, ensuring high quality and topical diversity in Q is critical.

We consider two approaches for seed construction:

Human-Crafted Seeds. In the fully manual setup, domain experts are given a detailed instruction-writing guideline that includes:

- Requirements for seed diversity and realism,
- Types of questions that are difficult to synthesize from retrieved documents (e.g., abstract reasoning, subjective judgment),
- Subcategories within the target domain.

Each annotator is asked to write several domain-specific instructions based on this guideline. This ensures that Q contains high-coverage, realistic questions that may not emerge from automatic generation methods.

Human–LLM Collaborative Seeds. In the hybrid setup, annotators are guided to use powerful LLMs (e.g., GPT-4o, Claude 3.5 Sonnet, Gemini 2 Pro) to generate seeds. A separate guideline is provided, which instructs annotators to:

1. Ask the LLM to suggest a comprehensive list of subtopics within the target domain;
2. For each subtopic, prompt the LLM to generate multiple instruction-style queries based on predefined types;
3. Select, edit, or rewrite these outputs into final seeds.

This method enables rapid scaling while retaining human supervision and correction. In practice, we collected approximately twice as many seeds per annotator compared to the fully manual setting.

After initial training, we identify weaknesses in specific subdomains or instruction types. Additional seeds are then generated—using either method above—to target these weaknesses and enrich low-performing areas.

Appendix C provides a detailed example of how seed queries were constructed in the domain of Iranian culinary and tourism.

3.3 Stage 2: Query Expansion

To increase the diversity and coverage of the instruction dataset, we expand the initial seed set Q using a LLM. This process is inspired by the Self-Instruct framework (Wang et al., 2022a), but adapted to focus exclusively on instruction generation rather than full instruction–response pairs.

At each expansion step, a subset of n seed queries $S = \{q_1, q_2, \ldots, q_n\} \subset Q$ is selected, and the LLM is prompted to synthesize k new instruction-style queries:

$$I_S = \{i_1, i_2, \ldots, i_k\}$$

Each instruction $i_j \in I_S$ is either a paraphrase of a single q_i or a novel combination of multiple seed queries in S. This process is repeated over m iterations to generate a sufficiently diverse instruction pool.

We denote the union of all generated instructions across iterations as:

$$I = \bigcup_{t=1}^{m} I_{S_t}$$

The final instruction set used for downstream processing is defined as:

$$I^{\text{total}} = Q \cup I$$

Appendix A.1 shows the prompt template used to diversify queries during expansion.

3.4 Stage 3: Document Retrieval

In this stage, we retrieve domain-relevant content to ground each instruction $i \in I^{\text{total}}$ in external evidence. The goal is to identify high-quality sources that can support accurate and informative answer generation in the next stage.

Depending on the application setting and available resources, we consider multiple retrieval strategies:

- **RAG-style retrieval:** If access to a structured document collection is available, we apply dense or hybrid retrieval to identify the top-k text chunks $C_i = \{c_{i,1}, \ldots, c_{i,k}\}$ that are semantically similar to the instruction i. This is done using retriever models such as DPR or OpenAI embeddings.

- **Web search:** For open-domain instructions or in the absence of curated corpora, we use web search engines to collect a set of relevant URLs and associated snippets as external context.

- **External APIs:** In some cases, domain-specific APIs (e.g., tourism databases, legal knowledge graphs) are queried to extract structured or semi-structured evidence.

To improve the effectiveness of retrieval, we do not use the instruction i directly. Instead, we construct a search-oriented query q_i^{search} from i using an LLM-based rewriting module:

$$q_i^{\text{search}} = \text{LLM}_{\text{rewriter}}(i)$$

This transformation helps reduce ambiguity and match the query to the structure and vocabulary of the underlying retrieval source.

Prompt template used for query rewriting is provided in Appendix A.2.

3.5 Stage 4: Response Construction

In the final stage of the *SearchInstruct* pipeline, we synthesize high-quality answers for each instruction $i \in I^{\text{total}}$ using retrieved evidence. Each instruction is paired with its corresponding set of retrieved contexts $C_i = \{c_{i,1}, \ldots, c_{i,k}\}$ obtained from Stage 3.

To ensure efficient and accurate generation, we pre-process the retrieved documents to filter out irrelevant or noisy content. This step is especially important when dealing with long documents, as excessive input length can degrade LLM performance and increase computational cost. We apply one of the following filtering strategies:

- **LLM-based chunk filtering:** A lightweight language model is used to rank or extract segments of each document that are most relevant to the instruction;
- **Rule-based filtering:** Heuristics are applied to remove common noise sources such as HTML tags, ads, metadata, or user comments.

After filtering, we concatenate the instruction i with the cleaned context C_i^{filtered} and prompt a powerful LLM to generate a contextually grounded answer:

$$A_i = \text{LLM}_{\text{answer}}(i, C_i^{\text{filtered}})$$

This produces the final instruction–response pair (i, A_i) to be included in the training dataset $\mathcal{D}$.

Prompt templates for this stage are provided in Appendix A.3.

4 Applications of *SearchInstruct*

The advantages of our proposed approach can be summarized in three key aspects:

1. It enables the generation of accurate and contextually grounded responses for questions that require domain-specific knowledge—particularly in cases where no existing LLM is capable of providing a correct answer.
2. The instruction–response pairs generated by this method provide a close approximation of real end-user queries.
3. The method remains effective even when using smaller or open-source LLMs. As long as relevant documents are retrieved and appropriately incorporated, it is possible to construct high-quality responses to complex queries without relying on large proprietary models.

In this work, we apply *SearchInstruct* in two primary settings:

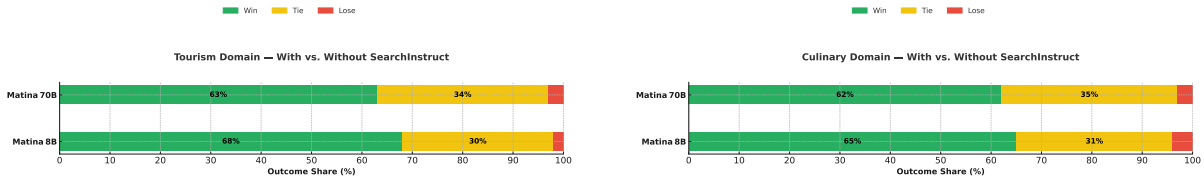
- Constructing SFT datasets in the domain of Iranian culture, with a focus on two subdomains: traditional cuisine and domestic tourism;
- Updating model knowledge with up-to-date information, to enhance its ability to respond to recent or evolving queries in specialized domains.

In the following sections, we provide a detailed analysis of both application scenarios.

4.1 SFT Dataset Construction for Iranian Culture

Iran is a vast and culturally diverse country, home to a wide range of historical sites, regional customs, and intangible heritage. Many of these cultural and geographical details are either underrepresented or entirely absent in current LLMs. To support this claim, Appendix B presents examples of domain-specific queries for which existing models failed to provide correct answers, while the *SearchInstruct* framework successfully generated accurate responses.

Another challenge is that Persian speakers, particularly in informal contexts, often use colloquial and context-rich language that differs significantly from formal documents. Except for a few major companies that serve Persian-speaking users, there



(a) Tourism domain

(b) Culinary domain

Figure 2: Human evaluation results comparing baseline models with those fine-tuned using the *SearchInstruct* framework across two domains.

is limited access to large-scale, real-world Persian queries. Attempting to synthesize such instructions purely from documents using LLMs often fails to capture the diversity of real user intent and task formulations. For example, a typical user might describe their own travel constraints and ask for a multi-day itinerary for a specific region of Iran—a type of instruction that is rarely (if ever) found explicitly in documents or websites.

To evaluate the practical effectiveness of our method, we applied *SearchInstruct* in two cultural subdomains: traditional Iranian cuisine and domestic tourism. As discussed in the following sections, our results demonstrate the value of this approach in generating realistic, high-quality SFT data in resource-limited domains.

4.1.1 Improvements in Cuisine and Tourism

In the MATINA project (Hosseinbeigi et al., 2025), multiple LoRA-tuned LLaMA-based models were trained across various domains, including traditional cuisine and tourism. The training data was generated using a mix of document-based question–answering and Evol-Instruct methods.

However, during human evaluation, it became evident that the models failed to perform adequately in several cultural subdomains. Upon analysis, these shortcomings were largely attributed to gaps in training data—particularly the absence of certain types of queries. These included questions such as:

- "List multiple examples of..."
- "Imagine the following scenario..."
- "Recommend something based on personal constraints..."

Such instructions are rarely present in documents and are not well-covered by typical synthetic data pipelines.

To address these gaps, we curated a specialized seed set as described in Section 3.2 and Appendix C. The SearchInstruct pipeline was then applied to produce high-quality SFT data specifically targeting underrepresented question types and subdomains. The number of generated instances for each category using this method is reported in Table 1. Each stage in the table corresponds to one full iteration of the SearchInstruct pipeline—including query expansion, document retrieval, and answer generation—which was repeated in response to human feedback or model performance analysis. In each stage, the pipeline was refined to specifically address the issues and gaps identified during the previous round, leading to incremental improvements in coverage, quality, and diversity.

Domain	Stage 1	Stage 2	Stage 3	Overall
Culinary	4273	2886	1773	**8932**
Tourism	3932	2378	1250	**7560**

Table 1: Number of SFT samples generated in culinary and tourism domains across three iterative stages. Each stage reflects one cycle of the SearchInstruct pipeline applied to address feedback from the previous round.

To evaluate the effectiveness of the proposed method, we conducted a blind human evaluation. Five independent annotators—none of whom were involved in seed construction—were asked to each design 20 diverse questions, resulting in a 100-question benchmark. This benchmark was used to evaluate two versions of the MATINA model: one trained prior to SearchInstruct augmentation, and one trained with our additional SFT data. Annotators were blind to the model identities[2].

As shown in Figure 2, models trained with SearchInstruct data performed significantly better on previously weak areas. These results validate the method's utility in real-world settings and highlight its ability to fill content gaps that standard data generation techniques miss.

This setup also enables iterative refinement: once specific weaknesses in the model's behavior

[2]Model identifiers were hidden from annotators during evaluation.

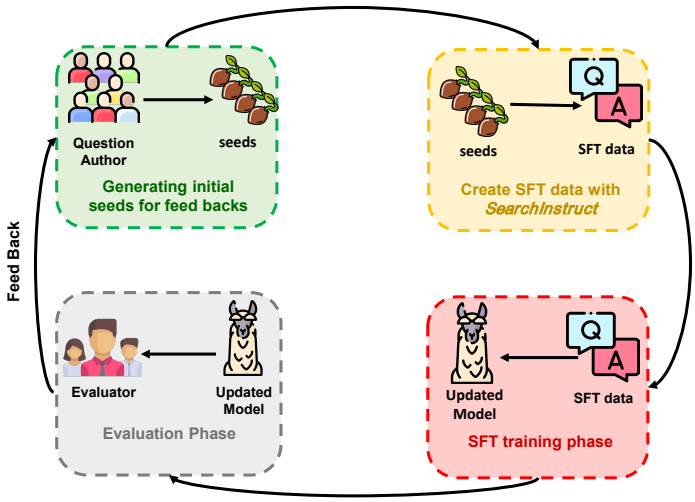


Figure 3: Iterative refinement loop enabled by the *SearchInstruct* framework. After initial fine-tuning, specific model weaknesses are identified through targeted evaluation. New instruction–response data is then generated to directly address these shortcomings, creating a feedback loop that leads to focused and incremental improvements in model performance.

are identified through human evaluation or domain-specific testing, additional instruction–response pairs can be generated using the SearchInstruct framework to target those areas. As shown in Figure 3, this creates a continuous improvement loop in which domain-specific shortcomings are systematically addressed. The process can be repeated until performance stabilizes or marginal gains diminish, allowing for efficient and focused model enhancement over time.

4.2 Model Updating via SearchInstruct

Another practical use case for the SearchInstruct framework is model updating—specifically, targeted modification of model outputs based on recent information. The underlying intuition is that by retrieving up-to-date documents relevant to a query, it is possible to revise the model's response in a lightweight and localized manner.

To constrain the task for evaluation purposes, we focused on a narrow update scope: tracking recent political changes and selected current events. Our goal was to update only specific knowledge while preserving the rest of the model's internal representations, effectively performing controlled *model editing*.

We applied this strategy to the `Gemma 27B` model. First, using a set of manually curated seed instructions (see Appendix C.3), we generated query–response pairs via SearchInstruct. For each instruction, we obtained the model's original output and retrieved relevant, up-to-date documents. These documents included new facts—such as political appointments, resignations, or verified current affairs.

To incorporate the updated knowledge, we used a secondary model such as `DeepSeek` to minimally revise the original answer from Gemma. The editing model was guided to only modify incorrect or outdated information based on the retrieved documents, without altering unrelated content. The prompt system used for editing is described in Appendix A.4.

As a result, only selected spans within the model's responses were modified, ensuring that the updates were localized, accurate, and minimally invasive. A schematic overview of the update-specific data generation process is shown in Figure 4, and examples of before-and-after outputs are provided in Appendix D.

4.2.1 Evaluation of the Edited Model

To evaluate the effectiveness of model updating via SearchInstruct, we constructed a preference dataset for fine-tuning using the `ORPO` (Hong et al., 2024). For each instruction $i \in I^{\text{edit}}$ (instructions related to recent knowledge), we defined a pair:

$$(i, A_i^{\text{reject}}, A_i^{\text{chosen}})$$

where A_i^{reject} is the original output from the base model (e.g., Gemma), and A_i^{chosen} is the edited response generated by a secondary editing model (e.g., DeepSeek) based on up-to-date documents.

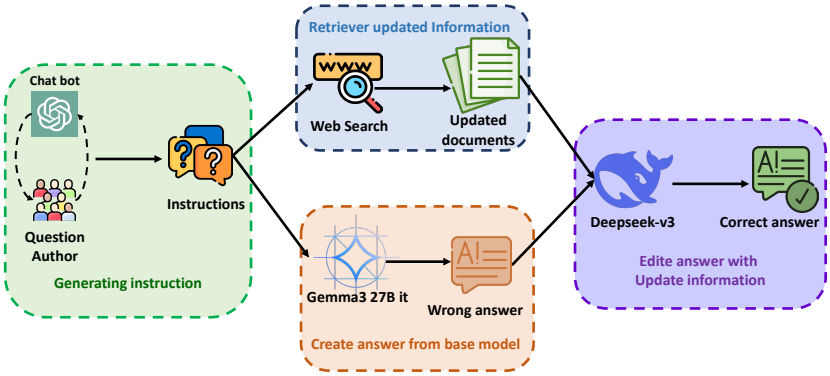


Figure 4: Pipeline for constructing update-specific instruction data used in model editing. Starting from user-provided queries, relevant documents are retrieved and used to construct grounded instruction–answer pairs.

Category	Gemma3 27B (original)	Gemma3 27B (updated)
STEM	66.04	62.86 (-3.18)
Social Sciences	79.30	77.87 (-1.43)
Humanities	60.60	59.02 (-1.58)
Other	73.66	71.65 (-2.01)
Average	68.88	66.89 (-1.99)

Table 2: Performance of Gemma3 27B before and after the update. Score decreases are highlighted in red.

Using this preference data, we fine-tuned the model with ORPO to align its outputs toward the corrected information. We then evaluated the updated model in two ways:

- **Targeted evaluation:** A new set of instructions related to the same domain (e.g., political changes) was submitted to both the original and edited models. Human evaluation confirmed that updated responses were correct and aligned with recent information.
- **General evaluation:** We also measured model performance on the MMLU benchmark (Hendrycks et al., 2021) to assess whether general knowledge was degraded. Results in Table 2 show no significant drop in accuracy, suggesting that the editing process preserved the broader knowledge of the model.

However, further analysis revealed a critical limitation. In some cases, the model was only capable of answering specific updated questions correctly (e.g., "Who is the current president of Iran?"), but lacked deeper knowledge about the subject. For instance, although the edited model correctly answered “Masoud Pezeshkian” as the current president, it failed to provide accurate information about his past roles or political background—facts not present in the updated dataset and not previously encoded in the model’s internal knowledge.

This indicates that the editing process effectively replaces surface-level facts for specific queries but does not deeply integrate new knowledge into the model’s reasoning capabilities. While model editing is efficient and minimally invasive, its ability to introduce deeply connected knowledge remains limited.

5 Conclusion

In this work, we introduced *SearchInstruct*, a framework for constructing high-quality instruction datasets for SFT. The method combines instruction generation with targeted document retrieval to address challenges of diversity, realism, and recency in specialized domains.

We demonstrated the effectiveness of this approach in two key scenarios: (1) improving model performance in cultural domains such as Iranian cuisine and tourism, and (2) updating factual knowledge in large language models through minimally invasive editing. Experimental results show that SearchInstruct generates contextually accurate, domain-aligned data, leading to measurable improvements without degrading existing knowledge.

Overall, SearchInstruct provides a flexible, scalable solution for data construction and model editing, making it a promising tool for developing adaptive, domain-aware language models.

6 Limitations

Despite its effectiveness, the *SearchInstruct* framework has several limitations:

- **Dependency on retrieved documents:** The quality and accuracy of generated answers are directly tied to the reliability of retrieved content. In domains with sparse or noisy resources, this can limit performance.

- **Shallow model editing:** Our findings suggest that editing responses through surface-level substitution does not lead to deep conceptual integration. The model may memorize specific updates without understanding broader context.

- **Limited scalability:** Although semi-automated, applying the method across large or fast-changing domains still requires substantial resources for document retrieval, filtering, and validation.

- **Reliance on strong LLMs:** Several stages, including query expansion and response construction, depend on high-quality LLMs. In low-resource settings, performance may degrade.

- **Risk of bias propagation:** Web-based retrieval introduces potential biases from source documents, which can affect the neutrality and fairness of generated data.

References

Lichang Chen, Jiuhai Chen, Tom Goldstein, Heng Huang, and Tianyi Zhou. 2023. Instructzero: Efficient instruction optimization for black-box large language models. *arXiv preprint arXiv:2306.03082*.

Kelvin Guu, Kenton Lee, Zora Tung, Panupong Pasupat, and Ming-Wei Chang. 2020. Realm: Retrieval-augmented language model pre-training. In *Proceedings of the 37th International Conference on Machine Learning (ICML 2020)*, pages 3929–3938. PMLR.

Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. 2021. Measuring massive multitask language understanding. In *Proceedings of the International Conference on Learning Representations (ICLR 2021)*. Also available as arXiv:2009.03300.

Jiwoo Hong, Noah Lee, and James Thorne. 2024. Orpo: Monolithic preference optimization without reference model. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 11170–11189, Miami, Florida, USA. Association for Computational Linguistics.

Sara Bourbour Hosseinbeigi, MohammadAli SeifKashani, Javad Seraj, Fatemeh Taherinezhad, Ali Nafisi, Fatemeh Nadi, Iman Barati, Hosein Hasani, Mostafa Amiri, and Mostafa Masoudi. 2025. Matina: A culturally-aligned Persian language model using multiple LoRA experts. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 20874–20889, Vienna, Austria. Association for Computational Linguistics.

Gautier Izacard, Patrick S. H. Lewis, Maria Lomeli, Lucas Hosseini, Fabio Petroni, Timo Schick, Jane Dwivedi-Yu, Armand Joulin, Sebastian Riedel, and Edouard Grave. 2022. Atlas: Few-shot learning with retrieval augmented language models. *arXiv preprint arXiv:2208.03299*.

Angeliki Lazaridou, Elena Gribovskaya, Wojciech Stokowiec, and Nikolai Grigorev. 2022. Internet-augmented language models through search. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8460–8478, Dublin, Ireland. Association for Computational Linguistics.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. In *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, pages 9459–9474. Curran Associates, Inc.

Xi Victoria Lin, Xilun Chen, Mingda Chen, Weijia Shi, Maria Lomeli, Rich James, Pedro Rodriguez, Jacob Kahn, Gergely Szilvasy, Mike Lewis, Luke Zettlemoyer, and Scott Wen-tau Yih. 2024. Ra-dit: Retrieval-augmented dual instruction tuning. In *Proceedings of the Eighth International Conference on Learning Representations (ICLR 2024)*. OpenReview.net / ICLR. Originally published as arXiv:2310.01352 (Oct 2, 2023).

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. *arXiv preprint arXiv:2203.02155*.

Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, Manan Dey, M. Saiful Bari, Canwen Xu, Urmish

Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal V. Nayak, and 22 others. 2022. Multitask prompted training enables zero-shot task generalization. In *The Tenth International Conference on Learning Representations (ICLR)*, Virtual Event.

Rohan Taori, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Xinyang Geng, Deepak Narayanan, Percy Liang, and Tatsunori B. Zhang. 2023. Stanford alpaca: An instruction-following llama model. https://github.com/tatsu-lab/stanford_alpaca.

Boxin Wang, Wei Ping, Lawrence McAfee, Peng Xu, Bo Li, Mohammad Shoeybi, and Bryan Catanzaro. 2024. Instructretro: Instruction tuning post retrieval-augmented pretraining. In *Proceedings of the 41st International Conference on Machine Learning (ICML 2024)*, pages 51255–51272. PMLR. Originally available as arXiv:2310.07713 (Oct 11, 2023).

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2022a. Self-instruct: Aligning language models with self-generated instructions. *arXiv preprint arXiv:2212.10560*.

Yizhong Wang, Swaroop Mishra, Pegah Alipoormolabashi, Yeganeh Kordi, Amirreza Mirzaei, Atharva Naik, Arjun Ashok, Arut Selvan Dhanasekaran, Anjana Arunkumar, David Stap, Eshaan Pathak, Giannis Karamanolakis, Haizhi Lai, Ishan Purohit, Ishani Mondal, Jacob Anderson, Kirby Kuznia, Krima Doshi, Kuntal Kumar Pal, and 19 others. 2022b. Super-naturalinstructions: Generalization via declarative instructions on 1600+ nlp tasks. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5085–5109, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Jason Wei, Maarten Bosma, Vincent Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V. Le. 2021. Finetuned language models are zero-shot learners. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Canwen Xu, Qingfeng Sun, Kai Zheng, Xiubo Geng, Pu Zhao, Jiazhan Feng, Chongyang Tao, Qingwei Lin, and Daxin Jiang. 2023. Wizardlm: Empowering large language models to follow complex instructions. *arXiv preprint arXiv:2304.12244*.

Zachary Ziegler, Ethan Perez, Tiago Pimentel, Emily Reif, Tianyi Zhang, Christopher Manning, and Colin Raffel. 2024. Craft: Conceptual retrieval-augmented fine-tuning. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP 2024)*. Association for Computational Linguistics.

Appendix A
System Prompt Designs

Each stage of our pipeline uses a dedicated system prompt tailored to the specific objective of that component. This appendix presents the design of each prompt through structured illustrations.

A.1 System Prompt for Query Expansion

The system prompt used in this stage instructs the model to generate diverse and semantically rich variations of an initial instruction seed. This system prompt is illustrated in Figure 5

A.2 System Prompt for Search-Oriented Rewriting

This prompt transforms general natural language instructions into search-optimized queries that enhance document retrieval performance.

Search-Oriented Rewriting System Prompt

Given an instruction, respond with a search term query
which can be efficiently used to find the answer in search engines.
Do not include any explanations.

Figure 6: System prompt design for search-oriented query rewriting.

A.3 System Prompt for Response Construction

This prompt guides the model to synthesize a grounded and context-aware response using retrieved documents as input.

Response Construction System Prompt

Provide accurate and well-explained answers to the user's questions,
based on the content provided below.
Ensure your response is as complete and comprehensive as possible.

Do not include information or assumptions outside of this content.

Figure 7: System prompt design for evidence-grounded response construction.

A.4 System Prompt for Answer Updating

The final system prompt is designed to apply local edits to existing answers when new or corrected information is introduced.

Appendix B
Failure Cases of Current LLM Models

As discussed in the main body of the paper, current language models often fail to provide accurate answers to questions in specialized domains due to insufficient domain knowledge. However, when provided with relevant supporting documents, their performance improves significantly. This observation reflects one of the core ideas behind our proposed method, *SearchInstruct*.

To demonstrate the effectiveness of our approach in situations where the base model fails to produce a reliable response, we present a set of representative examples in Table 3. For each case, we include the model's original response as well as the revised answer generated using the *SearchInstruct* pipeline. These examples highlight how retrieval-augmented prompting can compensate for gaps in model knowledge and substantially improve response quality in specialized domains.

Appendix C
Seeds Generation

To generate high-quality SFT data, we begin with a curated set of instruction seeds Q designed to ensure realism, topic diversity, and domain coverage. We employ two complementary strategies for constructing these seeds:

(1) Human-Crafted Seeds. In the fully manual setup, domain experts were given a detailed instruction-writing guideline that included:

- Requirements for seed diversity and contextual realism;
- Emphasis on question types that are difficult to synthesize from retrieved documents (e.g., abstract reasoning, subjective judgment);
- A comprehensive taxonomy of subcategories within each domain.

Each annotator was tasked with writing several domain-specific instructions that reflect real-world user needs and domain-specific reasoning challenges. This ensures that Q contains high-coverage, high-quality questions that may not emerge from automatic generation methods.

(2) Human–LLM Collaborative Seeds. In the hybrid setup, annotators used powerful LLMs (e.g., GPT-4o, Claude 3.5 Sonnet, Gemini 2 Pro) to assist with ideation and rewriting. Annotators were instructed to:

1. Ask the LLM to suggest a list of subtopics within the domain;

Query Expansion System Prompt

You are an expert in question generation.
You will be provided with sample questions.
Your task is to create exactly **{args.number_created_questions}** new questions inspired by the sample questions.
Each question should retain the structure and intent of the original sample questions but include variations by modifying specific details.

The generated questions must:
1. Be realistic, logical, and coherent.
2. Showcase creativity while maintaining a challenging nature.
3. Avoid being overly simple, nonsensical, or repetitive.

Provide the output in JSON format with the following structure:

```
{{
  "questions": [
    "Generated question 1",
    "Generated question 2",
    ...
  ]
}}
```

Do not include any additional explanations, comments, or extra content. Only return the JSON output as specified.

Figure 5: System prompt design for query expansion.

2. For each subtopic, prompt the LLM to generate instruction-style questions based on predefined types;

3. Select, edit, or rewrite the outputs into final seeds that match our quality criteria.

This collaborative setup enabled faster idea generation, controlled diversity, and linguistic variation, while maintaining human oversight to ensure instructional clarity and domain relevance.

In the following sections, we describe the seed construction process for each target domain in detail.

C.1 Culinary Domain

To construct seeds for the culinary domain, we followed a dual strategy: fully human-written instructions and human–LLM collaborative generation. Our objective was to capture real-world cooking challenges that require reasoning beyond standard recipes, while ensuring topical diversity and realistic phrasing.

C.1.1 Human-Crafted Seeds

In the manual setup, domain experts wrote a wide range of instruction-style questions grounded in common cooking situations. Rather than relying on simple recipe queries, the focus was placed on mistakes, decision-making, ingredient constraints, and reasoning-oriented prompts. Example questions include:

- "My Fesenjan stew turned out too sour. How can I fix the taste?"

- "How can I grill kebab without using a charcoal grill?"

To encourage topic diversity, annotators were provided with a guide listing more than 30 thematic categories. These included examples such as food preservation, regional dishes, dietary limitations, scientific explanations of cooking phenomena, ingredient reuse, children-oriented meals, and presentation aesthetics. These categories served purely as inspiration, not constraints, and annotators were encouraged to propose novel or cross-cutting question types that go beyond the listed themes.

C.1.2 Human–LLM Collaborative Seeds

In the collaborative setup, annotators were encouraged to interact with large language models (e.g., GPT-4o, Claude 3.5) to co-develop seeds. This included using the model to generate candidate questions, rewrite drafts in different tones, or suggest prompts within a topic area.

Common interaction patterns included:

- "Suggest a question about common mistakes in cooking rice."

Answer Updating System Prompt

You will receive two string variables:

• CURRENT_ANSWER: {current_answer} – the existing answer text
• NEW_CONTEXT: {new_context} – newly retrieved context that may confirm, update, or extend the original answer

Instructions (Minimal-Change Policy)

Your job is to update CURRENT_ANSWER using NEW_CONTEXT with the **least possible modification**. Follow these rules:

1. **Preserve the wording, tone, and sentence structure** of CURRENT_ANSWER unless NEW_CONTEXT explicitly requires a change.
2. If only names, dates, numbers, or similar facts change, **replace those elements only**—leave the rest untouched.
3. If NEW_CONTEXT introduces new, non-conflicting information, **append** it naturally using the same writing style.
4. If a statement in CURRENT_ANSWER is contradicted, **overwrite only the smallest necessary part** to reflect the updated fact. Always favor information from NEW_CONTEXT.
5. Do **not** rewrite, paraphrase, or restructure unless absolutely necessary for grammatical correctness after an edit.

Temporal Reference

When interpreting date expressions (e.g., "this year"), assume the current year is **2025 CE / 1404 SH**.

Output

Return a **single plain-text string** with the fully updated answer.
Do **not** include any metadata, brackets, or explanations—only the final revised answer.

Figure 8: System prompt design for answer refinement and updates.

- "Rewrite this to be more natural: Why is the center of my cake undercooked?"

Human reviewers selected and refined the outputs to ensure clarity, relevance, and natural phrasing. This collaborative process enhanced topical coverage and linguistic diversity, while accelerating seed creation.

C.2 Tourism Domain

To construct seeds in the tourism domain, we aimed to capture real-world questions that reflect the complex needs, constraints, and reasoning challenges of travelers within Iran. The objective was to go beyond basic location lookups or generic suggestions and instead focus on experiential, cultural, seasonal, and logistical aspects of domestic tourism.

We adopted a dual strategy: (1) fully human-authored seeds and (2) human–LLM collaborative construction.

C.2.1 Human-Crafted Seeds

In the manual setup, expert annotators were instructed to write instruction-style questions that simulate authentic user concerns in diverse travel scenarios. These included comparisons between destinations, cultural experiences, regional infrastructure challenges, weather-related planning, and cost-aware decisions.

To ensure coverage, annotators were given a guide listing more than 30 example categories. These included: historical landmarks, local cuisines, nature tourism, rural and nomadic travel, eco-tourism, handicrafts and souvenirs, accommodation types, seasonal destinations, transportation, adventure activities, underexplored sites, wellness tourism, and even environmental concerns such as climate impact on desert and coastal areas.

These categories were intended for inspiration only, not as limitations. Annotators were encouraged to mix themes (e.g., cultural + seasonal + budget-focused) and explore niche or unconventional travel scenarios.

Representative examples of well-formed questions include:

- "If I want to travel in winter to a warm region with eco-lodges and low tourist density, where should I go?"
- "What makes staying in a traditional Yazdi house different from a modern hotel, and which one would be better for a three-day cultural trip?"
- "During my visit to Kandovan, the cave-like architecture stood out. How does it compare to stepped villages like Masuleh in terms of tourism experience?"
- "Why is Hormuz Island, although less famous than Qeshm or Kish, often described as more unique and immersive?"

index	Language	Instruction	gpt4o	serachInstruct (gpt4o-mini)
One	Persian	طرز تهیه ی کله کنجی	**کله‌گنجی** یکی از غذاهای سنتی **استان کرمان** است که بیشتر با کله‌پاچه فرق دارد و از گوشت و استخوان سر گوسفند یا گاو و آرد تهیه می‌شود. در ادامه طرز تهیه سنتی کله کنجی را برایت توضیح می‌دهم: **مواد لازم:** • کله گوسفند یا گاو: ۱ عدد (شامل مغز، زبان، گوشت صورت، گونه و...) • آرد گندم سبوس‌دار: حدود ۲ لیوان • پیاز: ۲ عدد بزرگ	**کله کنجی (یا گمگه)** یکی از شیرینی‌های سنتی و مقوی **استان ایلام** است که از ترکیب کنجد و خرما تهیه می‌شود. این شیرینی به دلیل طبع گرم و انرژی بالایی که دارد، در فصول سرد سال به‌ویژه شب یلدا و عید نوروز محبوب است. در ادامه، طرز تهیه این شیرینی خوشمزه را برایتان آورده‌ام: **مواد لازم:** • کنجد: ۵۰۰ گرم • خرمای خشک (زاهدی یا زرّی): ۱ کیلوگرم
	Translated into English	Recipe for Kaleh Konji	**Kaleh Konji** is a traditional dish from **Kerman Province** in Iran. It differs from regular *kaleh pacheh* and is made from the meat and bones of a sheep's or cow's head, along with whole wheat flour. Below is the traditional method of preparing Kaleh Konji: **Ingredients:** • Sheep or cow head: 1 (including brain, tongue, cheek meat, etc.) • Whole wheat flour: about 2 cups • Onion: 2 large	**Kaleh Konji (or Gomgeh)** is a traditional and nutritious sweet from **Ilam Province** in Iran. Made from sesame seeds and dates, it's especially popular during cold seasons—particularly on Yalda Night and Nowruz—due to its warming nature and high energy content. Below is the method for making this delicious sweet: **Ingredients:** • Sesame seeds: 500 grams • Dried dates (Zahedi or similar variety): 1 kilogram
Tow	Persian	من گوشت قرمز، تخم مرغ، خلال بادام و برگ بو دارم چه غذای ایرانی باهاش درست کنم؟	با توجه به موادی که در اختیار دارید، می‌توانید یک غذای ایرانی خوشمزه به نام **خورش خلال بادام کرمانشاهی** درست کنید. این غذا مخصوص مناطق غرب ایران (به‌ویژه کرمانشاه) هست و بسیار مجلسی و معطره. **مواد لازم:** • گوشت قرمز (خورشتی) – ۳۰۰ گرم • خلال بادام – ۱/۲ پیمانه • پیاز – ۱ عدد بزرگ • زعفران دم‌کرده – ۲ قاشق غذاخوری • رب گوجه‌فرنگی – ۱ قاشق غذاخوری (اختیاری) • گلاب – ۱ قاشق غذاخوری • برگ بو – ۱ تا ۲ عدد • تخم‌مرغ – فقط زرده (برای تزئین، اختیاری)	با توجه به موادی که در اختیار دارید، می‌توانید یک غذای ایرانی خوشمزه به نام **«پیچاق قیمه»** (خورش بادام اردبیلی) تهیه کنید. این خورش سنتی و مجلسی از استان اردبیل است و با طعم منحصر به فرد در مهمانی‌ها و مجالس سرو می‌شود. **مواد لازم برای ۴ نفر** • گوشت گوسفندی خورشتی: ۴۰۰ گرم • پیاز متوسط: ۴ عدد • خلال بادام: ۱/۲ پیمانه • تخم‌مرغ: ۴ عدد • برگ بو: ۲ عدد • کره محلی: ۱۰۰ گرم • زعفران دم‌کرده غلیظ: ۳ قاشق غذاخوری • آبغوره یا آبلیمو: ۱ قاشق غذاخوری
	Translated into English	I have red meat, eggs, slivered almonds, and bay leaves. What Persian dish can I make with them?	With the ingredients you have (red meat, eggs, slivered almonds, bay leaves), you can make a delicious traditional Persian dish called **Khoresht-e Khalal Badam Kermanshahi**. This dish is native to western Iran (especially Kermanshah) and is considered very festive and aromatic. **Ingredients:** • Red meat (stewing cuts) – 300 grams • Slivered almonds – 1/2 cup • Onion – 1 large • Brewed saffron – 2 tablespoons • Tomato paste – 1 tablespoon (optional) • Rose water – 1 tablespoon • Bay leaves – 1 to 2 leaves • Eggs – yolks only (for garnish, optional)	Based on the ingredients you have, you can make a delicious Persian dish called **"Pichagh Gheimeh"** (Ardabili Almond Stew). This is a traditional and festive stew from Ardabil Province, known for its unique flavor and typically served at gatherings and special occasions. **Ingredients for 4 people:** • Lamb stew meat: 400 grams • Medium onions: 4 • Slivered almonds: 1/2 cup • Eggs: 4 • Bay leaves: 2 • Local butter: 100 grams • Strong brewed saffron: 3 tablespoons • Verjuice or lemon juice: 1 tablespoon
Three	Persian	مسلمونی نقل نخ کشی سوغات کجاس؟	بله، "نقل نخ‌کش" یکی از سوغاتی‌های معروف شهر تبریز است. این نقل‌ها با نخ‌هایی به هم متصل‌اند و ظاهرشان شبیه به تسبیح یا زنجیره‌ای از نقل‌های کوچک است. طعم آن معمولاً شیرین و گاهی با هل یا گلاب معطر می‌شود. در گذشته بیشتر در مراسم سنتی و عروسی‌ها در تبریز استفاده می‌شد.	بله، نقل نخ کش یکی از سوغاتی‌های معروف **استان همدان** است که به‌ویژه در شهرهای **ملایر** و **همدان** تولید می‌شود. این شیرینی سنتی با طعم شیرین و بافت نرم خود، محبوبیت فراوانی دارد و به‌عنوان یکی از صنایع دستی معروف این منطقه شناخته می‌شود.
	Translated into English	Do you know where *Nogh-e Nakhkesh* (thread-pulled nougat) is a local souvenir from?	Yes, *Nogh-e Nakhkesh* (thread-pulled nougat) is one of the famous souvenirs of **Tabriz**. These sweets are connected with threads and resemble a rosary or a chain of small sugar-coated pieces. Their taste is usually sweet and sometimes flavored with cardamom or rose water. In the past, they were commonly used in traditional ceremonies and weddings in Tabriz.	Yes, *Nogh-e Nakhkesh* (thread-pulled nougat) is one of the well-known souvenirs of **Hamadan Province**, especially produced in the cities of **Malayer** and **Hamadan**. This traditional sweet is highly popular for its sweet flavor and soft texture, and it is recognized as one of the region's famous handmade delicacies.
Four	Persian	یزد، بهم معرفی کن پیست اسکی نوی	یزد به‌دلیل آب‌وهوای کویری و اقلیم گرم و خشک، **پیست اسکی طبیعی یا دائمی ندارد**. برخلاف استان‌های شمالی یا غربی ایران مانند **البرز، همدان یا آذربایجان** که دارای پیست‌های اسکی واقعی هستند، یزد **شرایط اقلیمی لازم برای ایجاد پیست اسکی برفی را ندارد**.	در **استان یزد**، پیست **اسکی سخوید** در **شهرستان تفت** واقع شده است. این منطقه در دامنه‌های جنوبی **کوه شیرکوه** قرار دارد و به دلیل بارش‌های برف در فصل زمستان، به محلی محبوب برای تفریحات زمستانی تبدیل می‌شود. با این حال، باید توجه داشت که این مکان به طور رسمی به‌عنوان **یک پیست اسکی استاندارد شناخته نمی‌شود** و بیشتر برای بازی‌های برفی مانند **تیوب‌سواری** مناسب است.
	Translated into English	Tell me a ski resort in Yazd	Due to its desert climate and hot, arid conditions, **Yazd does not have a natural or permanent ski resort**. Unlike northern or western provinces of Iran such as **Alborz, Hamedan, or Azerbaijan**, which have actual ski resorts, **Yazd lacks the necessary climatic conditions** for establishing a snow-based ski resort.	In **Yazd Province**, the **Sakhvid ski slope** is located in **Taft County**. This area lies on the southern slopes of **Shirkooh Mountain** and, due to snowfall in the winter season, has become a popular spot for winter recreation. However, it's important to note that this location is **not officially recognized as a standard ski resort**, and is more suitable for snow-related activities such as **snow tubing**.

Table 3: Examples of partial outputs from GPT-4o and SearchInstruct (GPT4o mini):This demonstrates that in certain instructions, the SearchInstruct method is capable of producing better responses than larger models, as it benefits from access to highly relevant contextual information, even when using a smaller model

These questions require the model to reason across multiple dimensions, including cultural insight, regional comparison, travel logistics, and user preference.

C.2.2 Human–LLM Collaborative Seeds

In the collaborative setup, annotators interacted with large language models (e.g., GPT-4o, Claude 3.5) to generate, refine, or extend questions. They could prompt the model with a general topic or theme and request multiple formulations, rewrites, or angle shifts.

Typical prompts used in the process included:

- "Generate questions about seasonal travel in northern Iran."
- "Give me a few ideas on how to ask about nomadic tourism experiences."

Annotators reviewed and adjusted the model's output for realism, clarity, and tone. This approach supported exploration of lesser-covered subdomains, stylistic variation, and faster seed creation, especially in cases where human inspiration plateaued.

Together, these two methods enabled the creation of a rich, diverse, and high-quality set of instruction seeds for the tourism domain, covering both typical and edge-case user needs.

C.3 Model Updating Domain

Language models are often unable to respond accurately to questions involving real-world facts that have changed after the model's training cut-off. These include updates to public figures, international events, institutional changes, or recent sports outcomes. In this section, we focus on designing instruction seeds that explicitly aim to identify and correct outdated model knowledge in such contexts.

The goal was to create seed questions that prompt the model to revise its internal responses, particularly in scenarios where factual updates are essential. These seeds support fine-tuning or model editing workflows and serve as a targeted evaluation tool for temporal generalization.

Seeds were manually constructed by identifying areas where the model is likely to return outdated or incorrect answers. Each question was designed to be direct, fact-seeking, and anchored in well-known public events. Special care was taken to select topics that are newsworthy and important yet unlikely to be politically sensitive in the context of scientific publication.

Representative examples include:

- "Who won the 2024 U.S. presidential election?"
- "Who is currently the president of Iran in 1404 (2025)?"
- "Has the war between Russia and Ukraine ended or is it still ongoing?"
- "What was Iran's medal ranking in the 2024 Summer Olympics?"

Such questions are simple in structure but powerful in purpose: they help surface blind spots in the model's temporal awareness and provide clear entry points for knowledge injection. The resulting dataset serves both as an editing benchmark and as fine-tuning input for time-sensitive model correction.

Appendix D
Examples from the Updated Dataset

This section presents examples from the final dataset constructed for model updating. These instances are designed to help the model revise outdated or incorrect knowledge in a localized and controlled manner, without disrupting its broader behavior.

Each entry consists of a time-sensitive question accompanied by two responses:

- **Rejected:** An outdated, incorrect, or uncertain response—typically resembling what an unedited, pre-trained model might produce.
- **Chosen:** A minimally edited and factually updated version of the rejected answer.

A key principle in constructing the chosen responses is that the edits are minimal and highly targeted. Instead of rewriting the entire answer, only the incorrect portion is replaced with updated information, while preserving the original structure, phrasing, and tone as much as possible. This design allows the model to learn localized corrections through fine-tuning or model editing, rather than general retraining.

Table 4 provides several representative samples. Each row shows a real-world question, the outdated (rejected) model response, and the minimally revised, accurate version (chosen). These examples span topics such as elections, leadership updates, international conflicts, and sports results.

Appendix E – Training Hyperparameters

This appendix provides a consolidated overview of the most relevant training hyperparameters used across different stages of our experiments. We categorize the configurations into three groups:

- **SFT (Supervised Fine-Tuning)** on domain-specific tasks using the Matina-8B and Matina-70B models, applied separately to culinary and tourism instruction datasets.
- **ORPO (Optimal Response Preference Optimization)** training on the `Gemma-3-27B-IT` model using LoRA adapters and a custom ORPO objective.

Table 5 presents a unified summary of these hyperparameters, including key factors such as LoRA configuration, learning rate, dropout rate, batch size, training duration, and preference optimization parameters (where applicable)

Appendix F: Future Work

While *SearchInstruct* demonstrates promising results, several directions remain open for future exploration:

index	Language	Instruction	Gemma27b-it Response	Response Updated
One	Persian	وزیر ورزش ایران در سال ۲۰۲۴ کیست؟	در حال حاضر (خرداد ۱۴۰۳)، **کیومرث هاشمی** وزیر ورزش و جوانان ایران است. ایشان از مرداد ۱۴۰۱ این سمت را بر عهده دارند و در دولت سیزدهم به عنوان وزیر ورزش مشغول به کار هستند.	در حال حاضر (شهریور ۱۴۰۳)، **احمد دنیامالی** وزیر ورزش و جوانان ایران است. ایشان از ۳۱ مرداد ۱۴۰۳ این سمت را بر عهده دارند و در دولت سیزدهم به عنوان وزیر ورزش مشغول به کار هستند. پیش از ایشان، کیومرث هاشمی از ۲۷ شهریور ۱۴۰۲ تا ۳۱ مرداد ۱۴۰۳ عهده‌دار این سمت بود.
	Translated into English	Who is the Minister of Sports of Iran in the year ۲۰۲۴?	At present (June 2024), **Kioumars Hashemi** is the Minister of Sports and Youth Affairs of Iran. He has held this position since August 2022 and is serving as the Minister of Sports in the thirteenth administration.	At present (August–September 2024), **Ahmad Donyamali** is the Minister of Sports and Youth Affairs of Iran. He has held this position since August 21, 2024, and is serving as the Minister of Sports in the thirteenth administration. Before him, Kioumars Hashemi held this position from September 18, 2023, to August 21, 2024.
Tow	Persian	رئیس جمهور ایالات متحده کیه	رئیس جمهور فعلی ایالات متحده آمریکا جو بایدن است. او از ۲۰ ژانویه ۲۰۲۱ این سمت را بر عهده دارد. برای اطلاعات بیشتر می‌توانید به منابع زیر مراجعه کنید: • **وب‌سایت رسمی کاخ سفید**: https://www.whitehouse.gov/ : **وب‌سایت رسمی ریاست جمهوری آمریکا**: https://www.usa.gov/president :	**ژانویه ۲۰۲۵** به‌عنوان چهل‌وهفتمین رئیس‌جمهور آمریکا فعالیت خود را آغاز کرد. رئیس‌جمهور کنونی ایالات متحده آمریکا دونالد ترامپ می‌باشد. وی از تاریخ ۲۰
	Translated into English	Who is the President of the United States?	The current President of the United States is Joe Biden. He has held this position since January 20, 2021. For more information, you can visit the following sources: • **Official White House website**: https://www.whitehouse.gov/ • **Official U.S. government presidential page**: https://www.usa.gov/president	The current President of the United States is Donald Trump. He began his term as the 47th President of the United States on **January 20, 2025.**
Three	Persian	برگزاری المپیک تابستانی ۲۰۲۴ در پاریس چه رکوردهای جدیدی را به ثبت رساند؟	المپیک تابستانی ۲۰۲۴ پاریس با ثبت رکوردهای متعددی در جنبه‌های مختلف، یک دوره به یاد ماندنی و تاریخی را رقم زد. در اینجا به تفکیک این رکوردها می‌پردازیم: **رکوردهای ورزشی:(Sporting Records)** • **تعداد مدال‌ها:** • **ایالات متحده آمریکا** ثبا کسب ۱۲۱ مدال (۳۷ طلا، ۴۱ نقره، ۴۳ برنز) رکورد بیشترین تعداد مدال کسب شده در یک دوره المپیک را به نام خود ثبت کرد. • **چین** ثبا ۸۸ مدال (۳۸ طلا، ۳۲ نقره، ۱۸ برنز) در رتبه دوم قرار گرفت و عملکرد بسیار خوبی را به نمایش گذاشت. • **بریتانیا** ثبا ۶۷ مدال (۲۰ طلا، ۲۳ نقره، ۲۴ برنز) در رتبه سوم ایستاد. • **رکوردشکنی‌های انفرادی:** • **دنیل جونز (شنا)** :رکورد جهانی ۱۰۰ متر آزاد را با زمان ۴۶.۹۷ ثانیه شکست. • **آریانا آریانی (دوومیدانی)** :رکورد جهانی پرتاب وزنه را با ۲۱.۴۲ متر به نام خود ثبت کرد. • **کایلی مک‌کینون (دوچرخه‌سواری)** :در رشته BMX رکورد جدیدی را به ثبت رساند. • **بسیاری از ورزشکاران دیگر** :در رشته‌های مختلف مانند تیراندازی، ورزه‌برداری، قایقرانی و ... رکوردهای المپیک و یا رکوردهای ملی خود را بهبود بخشیدند (به دلیل تعداد زیاد، ذکر همه آن‌ها در این پاسخ امکان‌پذیر نیست.) • **رکوردشکنی تیمی:** • **تیم استافتا ۴ ۱۰۰متر آزاد مردان آمریکا** :رکورد جهانی را با زمان ۳:۰۷.۴۵ دقیقه شکست. • **تیم استافتا ۴ ۲۰۰متر آزاد زنان استرالیا** :رکورد جهانی را با زمان ۷:۳۳.۱۵ دقیقه شکست. • تیم‌های مختلف در رشته‌های تیمی مانند والیبال، بسکتبال، هندبال و ... نیز عملکردی فراتر از انتظارات داشتند و به رکوردهای جدیدی دست یافتند. **رکوردهای مربوط به سازماندهی و فناوری:** • **پایداری** :المپیک پاریس به عنوان "گالمپیک سبز" شناخته شد و با کاهش کربن و استفاده از انرژی‌های تجدیدپذیر، رکوردی در زمینه پایداری زیست‌محیطی به ثبت رساند. • **فناوری** :استفاده از فناوری‌های نوین مانند هوش مصنوعی، واقعیت مجازی و اینترنت اشیا برای بهبود تجربه تماشاگران و ورزشکاران، یک رکورد در زمینه استفاده از فناوری در المپیک محسوب می‌شود. • **امنیت** :با به‌کارگیری سیستم‌های امنیتی پیشرفته و همکاری گسترده نیروهای امنیتی، المپیک پاریس با کمترین میزان حادثه امنیتی برگزار شد که این نیز یک رکورد محسوب می‌شود. • **تعداد ورزشگاه‌های جدید و بازسازی شده** :پاریس برای برگزاری المپیک، تعدادی ورزشگاه جدید ساخت و ورزشگاه‌های قدیمی را بازسازی کرد که این امر، رکوردی در زمینه زیرساخت‌های ورزشی به شمار می‌رود.	لمپیک تابستانی پاریس ۲۰۲۴ از ۲۶ ژوئیه تا ۱۱ اوت ۲۰۲۴ در پاریس برگزار شد و میزبان ورزشکارانی از ۲۰۶ کمیته ملی المپیک در ۳۲ رشته و ۳۲۹ رویداد بود. **رکوردهای ورزشی:(Sporting Records)** • **تعداد مدال‌ها:** ○ **ایالات متحده آمریکا** ثبا کسب ۱۲۶ مدال (۴۰ طلا، ۴۴ نقره، ۴۲ برنز) صدر جدول را به خود اختصاص داد. این رکورد، یکی از بالاترین مجموع مدال‌های تاریخ المپیک به شمار می‌رود. ○ **چین** ثبا ۹۱ مدال (۴۰ طلا، ۲۷ نقره، ۲۴ برنز) در رده دوم قرار گرفت. ○ **بریتانیا** ثبا ۶۵ مدال (۱۴ طلا، ۲۲ نقره، ۲۹ برنز) در رتبه سوم ایستاد. • **رکوردهای جهانی و المپیک:** ○ بیش از ۱۲۵ رکورد المپیک و حدود ۳۲ رکورد جهانی در این دوره ثبت شد. ○ از جمله رکوردهای برجسته: • ارمند دوپلانتیس (سوئد) در پرش با نیزه رکورد جهانی خود را ارتقاء داد. • سیدنی مک‌لاگلین-لیورون (آمریکا) در ۴۰۰ متر با مانع، با ثبت زمان ۵۰.۳۷ ثانیه رکورد جهانی را شکست. • تیم رله ۴ر کیلو ۴۰۰×۴ متر مختلط آمریکا در مرحله مقدماتی با زمان ۳:۰۷.۴۱ رکورد المپیک را به ثبت رساند. • تامیرات تولا (اتیوپی) در دوی ماراتن مردان، با زمان ۲:۰۶:۲۶ رکورد جدیدی در این ماده ثبت کرد. • فیت کیپگن (کنیا) در ۱۵۰۰ متر زنان با زمان ۳:۵۱.۲۹ رکورد المپیک را شکست. • بابی فینک (آمریکا) در شنای ۱۵۰۰ متر آزاد مردان، با زمان ۱۴:۰۶.۶۷ رکورد جهانی جدیدی را ثبت کرد. ○ همچنین در رشته‌هایی مانند دوچرخه‌سواری، تیراندازی، ورزه‌برداری، قایقرانی و دیگر رشته‌ها، رکوردهای جهانی، المپیک یا ملی متعددی به ثبت رسید. **رکوردهای مربوط به سازماندهی، فناوری و پایداری:** • مراسم افتتاحیه برای نخستین‌بار در تاریخ المپیک خارج از ورزشگاه و در کنار رودخانه سن برگزار شد و حدود ۳۲۶ هزار بلیط برای آن فروخته شد. • المپیک پاریس با هدف کاهش انتشار کربن، استفاده از انرژی‌های تجدیدپذیر، و ارائه غذاهای گیاهی در دهکده المپیک، به عنوان "المپیک سبز" شناخته شد. • استفاده از فناوری‌های نوین مانند هوش مصنوعی، واقعیت مجازی و اینترنت اشیا برای ارتقای تجربه تماشاگران، ورزشکاران و رسانه‌ها، یکی از ویژگی‌های برجسته این دوره بود. • با بهره‌گیری از سیستم‌های امنیتی پیشرفته و همکاری نیروهای بین‌المللی، این المپیک با کمترین میزان رخدادهای امنیتی برگزار شد.
	Translated into English	What new records were set during the Paris 2024 Summer Olympics?	The Paris 2024 Summer Olympics set numerous new records across various aspects, making it a memorable and historic event. Here's a breakdown of those records: **Sporting Records:** • **Medal Counts:** ○ **United States:** With 121 medals (37 gold, 41 silver, 43 bronze), the U.S. set the record for the most medals won in a single Olympic Games. ○ **China:** Secured second place with 88 medals (38 gold, 32 silver, 18 bronze), showcasing an outstanding performance. ○ **Great Britain:** Ranked third with 67 medals (20 gold, 23 silver, 24 bronze). • **Individual World Records:** ○ **Daniel Jones (Swimming):** Set a new world record in the 100m freestyle with a time of 46.97 seconds. ○ **Ariana Ariani (Athletics):** Broke the world record in shot put with a distance of 21.42 meters. ○ **Kylie McKinnon (Cycling):** Set a new world record in BMX. ○ **Many other athletes** also broke Olympic or national records in sports such as shooting, weightlifting, rowing, and more (due to the high number, not all can be listed here). • **Team World Records:** ○ **USA Men's 4x100m Freestyle Relay Team:** Broke the world record with a time of 3:07.45 minutes. ○ **Australia Women's 4x200m Freestyle Relay Team:** Set a world record with a time of 7:33.15 minutes. ○ Teams in sports such as volleyball, basketball, and handball also exceeded expectations and achieved new records. **Organizational and Technological Records:** • **Sustainability:** Known as the "Green Olympics," Paris 2024 set a benchmark in environmental sustainability by reducing carbon emissions and using renewable energy. • **Technology:** Advanced technologies like artificial intelligence, virtual reality, and the Internet of Things enhanced the experience for athletes and spectators — a first in Olympic history. • **Security:** With cutting-edge security systems and coordinated efforts among security forces, Paris hosted one of the safest Games in history. • **New and Renovated Venues:** Numerous new sports venues were built and existing ones renovated, marking a record in sports infrastructure development.	The Paris 2024 Summer Olympics were held from July 26 to August 11, 2024, hosting athletes from 206 National Olympic Committees across 32 sports and 329 events. **Sporting Records:** • **Medal Counts:** ○ **United States:** Topped the medal table with 126 medals (40 gold, 44 silver, 42 bronze), marking one of the highest total medal counts in Olympic history. ○ **China:** Ranked second with 91 medals (40 gold, 27 silver, 24 bronze). ○ **Great Britain:** Finished third with 65 medals (14 gold, 22 silver, 29 bronze). • **World and Olympic Records:** ○ Over **125 Olympic records** and approximately **32 world records** were set during the Games. ○ Notable highlights include: • **Armand Duplantis** (Sweden) improved his own world record in pole vault. • **Sydney McLaughlin-Levrone** (USA) broke the world record in the women's 400m hurdles with a time of 50.37 seconds. • **USA's mixed 4×400m relay team** set a new Olympic record in the heats with a time of 3:07.41. • **Tamirat Tola** (Ethiopia) set a new record in the men's marathon with a time of 2:06:26. • **Faith Kipyegon** (Kenya) broke the Olympic record in the women's 1500m with a time of 3:51.29. • **Bobby Finke** (USA) set a new world record in the men's 1500m freestyle swimming with a time of ۱۴:۰۶.۶۷. ○ Many other world, Olympic, and national records were also broken in disciplines such as cycling, shooting, weightlifting, rowing, and more. **Records in Organization, Technology, and Sustainability:** • The **opening ceremony** was held **outside a stadium for the first time in Olympic history**, along the Seine River, with about **326,000 tickets sold**. • Paris 2024 was recognized as the "**Green Olympics**" due to efforts to reduce carbon emissions, use renewable energy, and provide plant-based food options in the Olympic Village. • The **use of advanced technologies** such as artificial intelligence, virtual reality, and the Internet of Things significantly enhanced the experience for spectators, athletes, and media. • With **cutting-edge security systems** and international cooperation, this edition of the Olympics had the **lowest number of security incidents** recorded

Table 4: Representative examples of minimally revised responses with factual updates. Each row includes a user instruction, a rejected response containing outdated or incorrect information, and a chosen response with concise, targeted edits. The examples cover domains such as politics, international affairs, and sports, demonstrating factual correction with minimal changes to phrasing and tone.

- **Deeper exploration of RAG methods and external APIs:** Our current pipeline only briefly references the use of retrieval-augmented generation (RAG) techniques and external APIs, without a thorough analysis of their practical impact. Future work could investigate how different RAG configurations, API sources, and retrieval strategies affect data diversity and factual accuracy across domains.

- **Leveraging agent-based multi-step interactions for RL training:** Large language model agents are increasingly used to perform multi-step tasks such as guided retrieval, source evaluation, answer synthesis, and refinement. Capturing these rich, interactive sequences offers an opportunity to create environments or trajectories for reinforcement learning (RL). This could facilitate training of more capable agents in real-world, information-seeking scenarios.

- **Integration with structured knowledge sources:** Incorporating structured resources—such as knowledge graphs,

Model / Task	LoRA Rank	LoRA Alpha	DeepSpeed Stage	Learning Rate	Dropout	Batch Size	Epochs	Pref. β
Matina-8B (SFT)								
Culinary	128	256	Z0	1×10^{-4}	0.05	64	3	–
Tourism	256	512	Z0	1×10^{-4}	0.05	64	4	–
Matina-70B (SFT)								
Culinary	128	256	Z3	1×10^{-4}	0.05	8	2	–
Tourism	256	512	Z3	1×10^{-4}	0.05	8	3	–
Gemma-3-27B-IT (ORPO)								
Update	16	32	Z3	2×10^{-5}	0.005	4	6	0.05

Table 5: Unified view of key hyperparameters for all training stages. Rows are grouped into three sections: SFT with Matina-8B, SFT with Matina-70B, and ORPO on Gemma-3-27B-IT.

ontologies, or semantic lexicons—could improve answer precision and coherence, especially in knowledge-intensive domains like law, medicine, or education. Future pipelines may combine retrieved text with structured representations to enhance factual grounding.

- **Automated optimization of the feedback loop:** Our current framework includes an iterative improvement loop driven by human feedback and manual adjustments. A promising future direction is to formalize this loop using reinforcement learning or other optimization techniques, allowing the system to continuously refine its query generation, retrieval, and synthesis strategies based on performance signals.

- **Automated quality assessment tools:** Evaluating the quality of generated training data remains a bottleneck. Future work could explore automatic metrics or learned models that assess dimensions such as evidential coverage, factual accuracy, and stylistic consistency, reducing reliance on manual review.