

Supplementary Materials

Supplement A. Source code of get prompt template string for RAG

```
def get_template()->str:
    return """#### Instruction

Your task involves accurately selecting pharmaceutical terms from a list based on their
relevance to a specific query. This selection process requires a detailed understanding of
the query's fundamental components, including the active ingredient, dosage, and dosage form.
Your goal is to identify and prioritize terms from the options that best match the query's
specified criteria.

#### Key Steps for Selection:

1. **Dissect the Query** Break down the query to identify its essential elements:

    Ingredient
    Dosage (in mg or other units)
    Dosage Form (e.g., tablet, capsule, suspension)

2. **Examine the Options** Carefully review each option in the list and evaluate how they
align with the query's core components. Focus on finding matches in terms of ingredients,
doses, and dosage forms.

3. **Apply Selection Criteria**

    First relevance: Prioritize options that match all three components: ingredient, dose, and
dosage form.

    Second relevance: If no complete match is available, consider options that match two
components.

    Last relevance: Only ingredient.

4. **Maintain Accuracy and Relevance**

    Avoid assumptions or inferences beyond the given information. Ensure the selections are
directly relevant to the query in terms of unit, quantity, and dosage form.

5. **Response Format**

List your selections in order of relevance, starting with the closest match.

Always maintain the original format of the terms.

In cases where multiple terms are equally relevant or where none of the options constitutes a
complete match (e.g., several options align with two of the three criteria), all such
relevant options should be included in the response.

Provide a response as a comma-separated list: example1, example2, example3.

If there are no selections, do not provide any output.

6. **Clarification on Dosage Forms**

When selecting a medication based on a query, it is essential that the term selected matches
exactly with the dosage form mentioned in the query. For instance, if the query specifies
'Oral Tablet,' then a similar form like 'Oral Capsule' would not be acceptable. A slight
variation, such as 'Chewable Tablet,' may only be considered if there are no exact matches
available for the dosage form. Crucially, the dosage of the medication must match precisely
with that specified in the query. When a query specifies a particular dosage, it is
imperative that the available options match the exact specified dosage. In cases where the
query demands 5 mg, and the available options offer dosage strengths of 4 mg and 6 mg, these
options should not be considered or included in the selection process. Variations in dosage
strengths, regardless of their proximity to the specified dosage or common therapeutic
ranges, should be excluded from consideration. Only options that precisely match the
specified dosage should be selected.

7. **Handling of Different Dosage Units**

When confronted with unit conversions (e.g., from mg to mL), these conversions should only be
entertained when there are no available options that precisely match the original unit
specified in the query. In cases where the query specifies a dosage in mg, considering
options in mL should be reserved solely for instances where there are no exact matches
available in mg. The primary objective is always to locate an option that aligns precisely
with the original unit, and unit conversions should only be employed as a secondary measure
when an exact match in the original unit is unavailable. Furthermore, when performing these
unit conversions, it is imperative that they are executed with utmost precision and accuracy;
approximations should be strictly avoided unless no other alternative exists.
```

8. ****Handling Partial Information in Query Responses****

Options providing only partial information, such as listing just the active ingredient without specifying the dosage or form, are secondary choices. They should be considered only when there are no other options offering complete or nearly complete information. We prioritize evaluating options that fully match the query's requirements, including the active ingredient, dosage, and form, before considering those with partial information. If an option with partial information doesn't align with any aspect of the query, it should not be selected.

9. ****Handling Additional Ingredients in Options****

Options that contain the correct active ingredient, dose, and form but include additional, unmentioned ingredients should be excluded to align with the instruction to avoid unlisted information. We should strictly adhere to the query's specified requirements, including the active ingredient, dosage, and form, without introducing any unmentioned components in the options. Only options that precisely meet the query's criteria, as specified, should be considered for selection.

10. ****Treatment of Variants of the Same Ingredient****

When the query specifies a particular variant of an active ingredient, such as a specific salt or ester, only that exact variant will be considered relevant. Variants, even if chemically related or derived from the same basic molecule, should not be treated as equivalent to the specified active ingredient in the query. This policy applies regardless of matching dosage and form. The selection process mandates an exact match with the named active ingredient variant in the query, without substitution or consideration for related variants, salts, or esters.

11. ****Brand Names and Generic Names****

When a medication query specifies a drug by its generic name, options presented must strictly adhere to this specification. If both generic and brand name options are available, and the brand name options contain the exact same active ingredient, dosage, and dosage form as the generic name specified in the query, these brand name options should still be excluded from selection. The preference is always given to the generic name specified in the query, with no exceptions. This means that even if a brand name is an exact match in terms of the active ingredient, the dosage amount, and the form of administration (e.g., tablet, capsule), it should be ignored in favor of the generic name. The rationale behind this strict adherence is to ensure that the selection process remains consistent and in line with the exact terms specified in the query. Any deviation from this approach, even if the brand name matches perfectly with the generic name in all specifications, renders the brand name option irrelevant and it should not be considered in the selection process.

12. ****Exclusion of Unlisted Information****

If an option in the list contains additional information that is not specified in the query, such as extra ingredients or indications, this option must be excluded. This means that the selection process requires strict adherence to the information provided in the query, and any option that goes beyond this information by including additional elements is not to be considered relevant. Only those options that strictly match the query's specifications, without introducing any extra, unlisted details, should be included in the selection.

13. ****Exact Match Importance****

The primary goal is to find an exact match for all three query components (ingredient, dose, and dosage form) where possible, and only consider partial matches if no exact match exists.

Example:

Query: Tofisopam 50 mg Oral Tablet

Options:

```
-----  
Paxipam 40 MG Oral Tablet  
tofisopam  
tofisopam Oral Product  
tofisopam Oral Tablet  
tofisopam 50 MG  
tofisopam 50 MG Oral Tablet  
-----
```

Expected Answer:

tofisopam 50 MG Oral Tablet
tofisopam 50 MG
tofisopam Oral Tablet
tofisopam

Example2:

Query: Nitrazepam 2 mg Oral Tablet

Options:

nitrazepam
nitrazepam 10 MG Oral Tablet
nitrazepam 5 MG Oral Tablet
nitrazepam 5 MG Oral Capsule
nitrazepam 10 MG
nitrazepam 5 MG
nitrazepam 2 MG
nitrazepam 0.5 MG/ML Oral Suspension
nitrazepam 1 MG/ML Oral Suspension
nitrazepam 1 MG/ML
nitrazepam Oral Tablet
nitrazepam 0.5 MG/ML

Expected Answer:

nitrazepam Oral Tablet
nitrazepam 2 MG
nitrazepam

{format_instructions}

QUESTION:

Based on the provided list of pharmaceutical options, select the term(s) that most closely align with the query. Your response should strictly adhere to the terms as they appear in the list. Consider each option's relevance to the query's active ingredient, dose, and dosage form. Avoid creating or referencing terms not listed in the options.

Query: {source}

Options:

{targets}

Answer:

""