

- **Specific and Answerable**: The question to be created SHOULD BE FREE-FORM and have a unique, specific answer (a single word, or short phrase) without access to the article. The answer to the question should be definite, well-defined and NOT NUMERIC. IT SHOULD ALSO NOT BE UNCERTAIN like "above XYZ" OR A RANGE LIKE "between XYZ and ABC". Avoid creating binary questions (yes/no, either/or) or questions with a list of specific options (multiple choice).
- **Answerable based on article**: Each question must have a CLEAR AND DEFINITE answer based on information stated in the article. Given the question, the content of the article should be able to resolve the answer to the question INDISPUTABLY WITHOUT ANY AMBIGUITY OR UNCERTAINTY. THE ARTICLE SHOULD NOT STATE THAT THE ANSWER IS TENTATIVE OR AN ESTIMATE OR LIKELY. The answer SHOULD HAVE HAPPENED BY NOW.
- **Temporal Information**: The question should not be about recall of (past) facts or events known before the article publish date. Include any temporal information necessary to answer the question (like by which month, year, etc.) in the question. The question should always be posed in a forward-looking manner.
- **Direct and Precise**: Titles must be straightforward and unambiguous, avoiding vague terms. Use future tense when appropriate.
- **Resolution Criteria**: ALWAYS INCLUDE A BRIEF RESOLUTION CRITERIA in the question title. This is often the date by which the question will be resolved. For example, resolution dates such as "by {{month_name}}, {{year}}?" or "in {{month_name}}, {{year}}?". THE RESOLUTION DATE SHOULD BE BASED ON (AND FAITHFUL TO) THE CONTENT OR PUBLICATION DATE OF THE ARTICLE.
- **No references to article or future information**: DO NOT refer to the specific article, such as by saying "in the article". The forecaster does not have access to the article, its metadata or any information beyond the article publish date.
- **Question Types**: Focus on "Who", "What", "When", "Where" questions that have concrete answers.
- **Understandability**: The question title should have ALL the information to be understandable by a 10 year old. It should be independently understandable without the article.
- **Tense**: ALWAYS POSE THE QUESTION IN A FORWARD-LOOKING MANNER. THE QUESTION SHOULD BE IN FUTURE TENSE. Try to use phrases like "What will", "Who will", "When will", "Where will", "How much/many will" etc. It should appear as a forecasting question and not past prediction.

Answer Guidelines

- **Faithfulness to Article**: The answer should be based on information explicitly stated in the article, and not implications or your own knowledge. IT SHOULD BE STATED VERBATIM IN THE ARTICLE.
- **Non-Numeric**: The answer should not be a number or a percentage. It can be a word, phrase, date, location, etc BUT NOT MORE THAN 3 WORDS.
- **Definite** - Given the question and the article, the answer should be CLEAR, CONCRETE, CERTAIN AND DERIVABLE from the article. It should be short, WELL-DEFINED TERM and not uncertain or vague. IT SHOULD NOT BE A RANGE like "between XYZ and ABC" or "above XYZ" or "below PQR".
- **Resolved** - The answer MUST be something that has already happened or is happening now. It should be resolved given today's date and not be something that will happen in the future.
- **Specificity**: The answer should be specific enough to be unambiguous. Avoid overly general answers.
- **Conciseness**: Keep answers short - typically 1-3 words, occasionally a short phrase if necessary.
- **Exactness**: For names, use the exact names mentioned (full name, if possible).
- **Uniqueness**: The answer should be unique and THE ONLY CORRECT ANSWER to the question.
- **No Ambiguity**: The answer should be indisputable and not be open to multiple interpretations. IT SHOULD BE PRECISE AND NOT A RANGE OR UNCERTAIN ESTIMATE.

****Background Guidelines****

- ****Mention Question Opening Date**:** ALWAYS INCLUDE THE START DATE OF THE QUESTION IN THE BACKGROUND. IT SHOULD BE AT LEAST A FEW DAYS (OR WEEKS IF THE QUESTION IS ABOUT A LONG-TERM EVENT) BEFORE THE ARTICLE'S PUBLISH DATE AND ALSO BEFORE THE RESOLUTION DATE OF THE QUESTION. CONSEQUENTLY, THE BACKGROUND SHOULD NOT CONTAIN ANY INFORMATION WHICH HAS HAPPENED AFTER THE START DATE OF THE QUESTION.
- ****Necessary Context**:** The answerer does not have access to the article, so include MINIMAL CONTEXT required to understand the question keeping in mind the question opening date. Do not give (extra) details of the event from the article as background. If required, EITHER pose the event as a hypothetical scenario as if it were to happen in the future OR describe it as happening (unfolding) in real time. Describe any unfamiliar terms or concepts in the question title.
- ****SHOULD NOT HELP ANSWER**:** WHILE PROVIDING THE CONTEXT, DO NOT REFER OR MENTION OR LEAK THE ACTUAL ANSWER. The background must not help answer the forecasting question. DO NOT INCLUDE ANY INFORMATION from the article or elsewhere that either directly or indirectly (even partially) reveals the answer.
- ****No Additional Knowledge**:** Do not add any knowledge beyond what is required to understand the question. Only include information necessary to understand the question and its context.
- ****Tense**.** ALWAYS POSE THE BACKGROUND INFORMATION IN CURRENT TENSE. Only provide minimal information which is known until the question opening date.

****Resolution Criteria****

- ****Necessary Criteria**:** State the EXACT conditions by which the outcome will be judged. Include the criteria which determines how the question will be resolved. state the conditions by which the outcome will be judged.
- ****Date and Source of Resolution**:** Always state the date and the source by which the question will be resolved. For example, resolution dates such as "by {{month_name}}, {{year}}?" or "in {{month_name}}, {{year}}?", and potential source(s) of resolution such as "based on {{news source}}", "reports from {{official name}}", etc. THE RESOLUTION DATE SHOULD BE CHOSEN THOUGHTFULLY AS THE ANSWER'S VALIDITY AND SOUNDNESS DEPENDS ON IT. THE RESOLUTION DATE SHOULD BE SUCH THAT THE ANSWER CAN BE RESOLVED DEFINITELY AND INDISPUTABLY FROM THE CONTENT OR PUBLICATION DATE OF THE ARTICLE. IT SHOULD MENTION BY WHEN IS THE OUTCOME OF THE QUESTION EXPECTED TO HAPPEN. HOWEVER, IT SHOULD NOT LEAK OR MENTION ANYTHING ABOUT THE ARTICLE.
- ****Details**:** Be as detailed as possible in creating the resolution criteria for resolving the question as cleanly as possible. There should be no ambiguity in the resolution criteria.
- ****Expectation and Format of Answer**:** Based on the actual answer, the resolution criteria should state how precise the expected answer should be and in what format it should be. For example, if the actual answer is a date, the resolution criteria should specify how detailed the expected date should be -- only year, or both month and year, or day, month, and year all together. DO NOT GIVE THE ACTUAL DATE (ANSWER). If the actual answer is a percentage, then the criteria should state the expected answer should be a percentage. DO NOT GIVE THE ACTUAL PERCENTAGE. If the actual answer is in certain unit, then the criteria should specify that. THE RESOLUTION CRITERIA SHOULD MAKE IT EXACTLY CLEAR AND PRECISE WHAT IS EXPECTED FROM THE ANSWERER AND IN WHAT FORMAT AND HOW IT WILL BE CHECKED LATER. IF GIVING AN EXAMPLE, IT SHOULD BE VERY GENERIC AND AS FAR AWAY FROM THE ACTUAL ANSWER AS POSSIBLE.
- ****SHOULD NOT HELP ANSWER**:** The resolution criteria must not directly help answer the forecasting question. DO NOT INCLUDE ANY INFORMATION from the article or elsewhere that either directly or indirectly (even partially) reveals the answer. DO NOT REFER OR MENTION OR LEAK THE ACTUAL ANSWER HERE.

****Answer Type Guidelines****

- ****Expected Format**:** The answer type should be either "numeric (XYZ)" if the answer is a number (of any kind) or "string (XYZ)" in all other cases. In

numeric cases, XYZ should be the exact type of number expected. For example, "numeric (integer)", "numeric (decimal)", "numeric (percentage)", "numeric (whole number)", etc. In string cases, XYZ should broadly be the category of string expected. For example, "string (name)", "string (date)", "string (location)", etc. If the category is not clear, use "string (any)". HOWEVER, ALWAYS TRY TO CREATE QUESTIONS WHERE THE ANSWER CATEGORY IS CLEAR AND PRECISE.

****Question Quality Criteria****

- ****Forecastable**:** The question should be something that could reasonably be predicted or forecasted before the article's publication.
- ****Towards the future**:** THE QUESTION SHOULD BE POSED IN A FORWARD-LOOKING MANNER.
- ****Interesting**:** The question should be about a meaningful event or outcome, not trivial details.
- ****Impactful**:** The question should be such that if its answer is forecasted ahead of time, it should have significant (downstream) impact (relevant to high number of people).
- ****Difficulty**:** While the question should be hard to answer without access to the article, it should also not be unreasonably difficult.
- ****Verifiable**:** The answer should be something that can be EXACTLY verified from the article itself.
- ****Time-bound**:** Include clear timeframes or deadlines when relevant.
- ****Free-form**:** If possible, avoid creating binary questions (yes/no, either/or) or questions with a list of specific options (multiple choice).

Generate `{self.num_questions_per_article}` high-quality, DIVERSE short answer forecasting questions based on the provided article. Use the XML format with question_id value "0", "1", "2", etc. DO NOT INCLUDE ANY ANALYSIS, RANKING, OR ADDITIONAL COMMENTARY.

Article:

`{source_article}`

****Required Output Format****

```
<q1>
<question_id>0</question_id>
<question_title>[Question 1]</question_title>
<background>[Background 1]</background>
<resolution_criteria>[Resolution Criteria 1]</resolution_criteria>
<answer>[Answer 1]</answer>
<answer_type>[Answer Type 1]</answer_type>
</q1>
..
<q{self.num_questions_per_article}>
<question_id>{self.num_questions_per_article - 1}</question_id>
<question_title>[Question {self.num_questions_per_article}]</question_title>
<background>[Background {self.num_questions_per_article}]</background>
<resolution_criteria>[Resolution Criteria
    {self.num_questions_per_article}]</resolution_criteria>
<answer>[Answer {self.num_questions_per_article}]</answer>
<answer_type>[Answer Type {self.num_questions_per_article}]</answer_type>
</q{self.num_questions_per_article}>
```

Stage 2 — Individual Validation

****Task:**** You will be provided with a news article and a question WHOSE ANSWER IS SUPPOSED TO BE BASED ON THE ARTICLE. Your job is to validate whether the answer to the question is valid by being faithful to the article (content, title, or description).