

# Evaluate information-seeking conversations

In this task, you will read a **passage** and a **conversation** about this passage. In the conversation, one person asks questions to get information about the passage while the other one answers the questions using evidence from the passage. Your job is to evaluate **the validity of the questions** and the **correctness of the answers**.

You will be given the passage and the set of question-answer pairs from the conversation. You will see the passage the entire time, and you will see the question-answer pairs one-by-one. For each question-answer pair, you will need to do the following three steps:

**Step 1:** Verify that whether the question (marked as "Q") is invalid or unanswerable or answerable given the passage.

A question is invalid when it is not in a question format, namely, it is not a question. For example, a statement sentence (like "you are right") or some chit-chat (like "hi") is invalid.

A question is unanswerable when the question **cannot be answered or supported by a single consecutive span** from the passage. To check if the question is answerable or not, you need to **read the passage** and **look for evidence** that can answer this question. If you find the answer, keep it in mind as you may need it for Step 3. If you are not able to find an answer for this question, then the question is unanswerable. Here is an example for illustration. Please study it carefully:

Passage:

In May 1983, she married Nikos Karvelas, a composer, with whom she collaborated in 1975 and in November she gave birth to her daughter Sofia. After their marriage, she started a close collaboration with Karvelas. Since 1975, all her releases have become gold or platinum and have included songs by Karvelas. In 1986, she participated at the Cypriot National Final for Eurovision Song Contest with the song Thelo Na Gino Star ("I Want To Be A Star"), taking second place. This song is still unreleased up to date. In 1987, she and Karvelas had their second child Haris. In 1984, Vissi left her record company EMI Greece and signed with CBS Records Greece, which later became Sony Music Greece, a collaboration that lasted until 2013.

Q1: What happened in 1983?

(**Answerable**/Unanswerable/Invalid)

[This question is answerable because the answer is the first sentence of this passage.]

Q2: What else is interesting about this article?

(**Answerable**/Unanswerable/Invalid)

[An open question like this is answerable. The answer could be anything interesting in this passage that wasn't mentioned before in the conversation. Most open questions in this task should be answerable.]

Q3: Did she have any children?

(**Answerable**/Unanswerable/Invalid)

[It is tricky to annotate questions that are usually answered with "Yes" or "No", because the agent's answer is always a span. This question is answerable, because the span "in November she gave birth to her daughter Sofia" implies that she has children.]

Q4: How many children did she have?

(Answerable/**Unanswerable**/Invalid)

[Although by reading the whole passage you might figure out that the answer is "two" (Sofia and Haris), because this question cannot be answered with a single span from the passage, it is unanswerable.]

Q5: Did Vissi retire?

(Answerable/**Unanswerable**/Invalid)

[This question is unanswerable because there is no single span in this passage that mentions whether Vissi retired or not.

**Note:** Even if it might seem like the answer is "No" to you using commonsense reasoning, we ask that you only use the evidence from the passage. If there is no **DIRECT** evidence in the passage, then the question is unanswerable.]

Q6: I know she retired in 2014.

(Answerable/Unanswerable/**Invalid**)

[This question is invalid because it is not in question format.]

**Note:** When judging the answerability of a question, please base your evaluation on the agent's prediction in the conversation history, even if it is an incorrect answer. For example:

Q: When did she marry?

A: She had a son, Mike. [Incorrect answer]

Q: When did he come to NYC?

[Even though the last answer is incorrect, you should still take it as the context and interpret "he" in the question as "Mike".]

**Step 2:** If a question is answerable, you will then see the answer. Your job is to select whether the answer is correct or incorrect. **Important note:** the answer is a span selected from the passage, not a natural answer that resembles daily human speech. The standard for a correct answer is that the agent's response *contains* the content that can answer the question. Take the above examples again:

Q1: What happened in 1983?

**Correct answers:**

- In May 1983, she married Nikos Karvelas,  
[It's OK to have extra punctuation and redundant information.]
- In May 1983, she married Nikos Karvelas
- she married Nikos Karvelas

**Incorrect answers:**

- In May 1983  
[It does not answer the question.]
- In May 1983, she married Nikos Karvelas, a composer, with whom she collaborated in 1975 and in November she gave birth to her daughter Sofia. After their marriage, she started a close collaboration with Karvelas. Since 1975, all her releases have become gold or platinum and have included songs by Karvelas. In 1986, she participated at the Cypriot National Final for Eurovision Song Contest with the song Thelo Na Gino Star ("I Want To Be A Star"), taking second place. This song is still unreleased up to date.  
[The answer is too long and contains too much information that is irrelevant to the question.]
- Sorry I cannot answer the question.  
[Because in Step 1 you selected that the question should be answerable, if the AI agent cannot answer this question, then its answer is incorrect.]

**Note:** For subjective questions like "What happens next?" or "Why did this happen?", you might have different answers than the agent. Please base the answer of these questions on the **agent's** previous answers.

**Step 3:** If you selected that the answer was incorrect in Step 2, you will now be asked to provide the answer span that you think is correct to the question. To simplify the process and

avoid typos, we ask you to **copy-paste** the answer span directly from the passage to the text box. Any text that is not directly from the passage will not be accepted by the system.

### **Important note:**

Thank you for all your hard work! Your annotation will be used for building better conversational AI agents, and the development of our AI system will heavily rely on your annotation quality. We will evaluate your annotations, and those with an unacceptable error rate will be rejected.