

# **BOLT IR Assessment Guidelines**

**Version 2.9, October 14, 2013**

**Linguistic Data Consortium**

<b>1. Introduction.....</b>	<b>3</b>
<b>2. Context and Motivation.....</b>	<b>3</b>
<b>3. BOLT IR Decision Tree .....</b>	<b>3</b>
<b>4. Rules of Thumb.....</b>	<b>5</b>
<b>5. BOLT IR Queries and Citations .....</b>	<b>7</b>
5.1 The Query.....	7
5.2 The Citation .....	8
<b>6. DECISION: Do you need to see the source text to proceed?.....</b>	<b>9</b>
6.1 When to choose <i>YES</i> .....	9
6.2 When to choose <i>NO because the translation is incomprehensible</i> .....	9
6.3 When to choose <i>NO because I need to see the source text</i> .....	10
6.4 Examples .....	10
<b>7. Decisions if you are not viewing the source text .....</b>	<b>15</b>
7.1 DECISION: Does the English citation satisfy the rules of interpretation? .....	15
7.2 DECISION: Is the English citation useful?.....	18
<b>8. Decisions if you are viewing the source text.....</b>	<b>20</b>
8.1 DECISION: Does the red/nearby source text satisfy the rules of interpretation? .....	21
8.2 DECISION: Is the red/nearby source text useful? .....	27
8.3 DECISION: Does the English Citation contain any relevant information? .....	32
<b>9. DECISION: Determine if the citation contains extraneous material .....</b>	<b>37</b>
9.1 Examples that do not contain large amounts of extraneous material .....	38
9.2 Examples that do contain large amounts of extraneous material .....	40
<b>10. Conclusion.....</b>	<b>41</b>

## **1. Introduction**

The BOLT (Broad Operational Language Translation) Program will create new techniques for automated translation and linguistic analysis for informal, conversational genres in Egyptian Arabic, Chinese and English.

In BOLT Information Retrieval (IR) task, systems are required to process natural language queries in English and return relevant answers from a large corpus of Egyptian Arabic, Chinese and English discussion forum threads harvested from the web. Systems must return a list of responses (known as citations) for each query, translating responses into English where required. Human assessors then review citations for correctness.

This document describes the process for assessing the relevance of citations in the BOLT Phase 2 IR Evaluation.

## **2. Context and Motivation**

To help you understand the motivation for this task, imagine a busy monolingual English researcher who has a complex question. This researcher has access to a large, multilingual collection of documents where the answers can be found. Because the researcher is busy, there isn't time to manually look at all the documents in the collection. And because the researcher is a monolingual English speaker, it wouldn't be possible to find all the answers even if there was plenty of time.

So instead, the researcher types the question into the BOLT IR machine, and gets back a list of short, relevant answers drawn from the document collection. Answers are automatically translated into English, and contain pointers to the original document so the researcher can follow up if desired.

Your job as an assessor is to assess the quality of the answers returned by the BOLT IR machine.

## **3. BOLT IR Decision Tree**

Assessment is done by following a decision tree. A decision tree depicts each question to be answered along with all possible outcomes and implications.

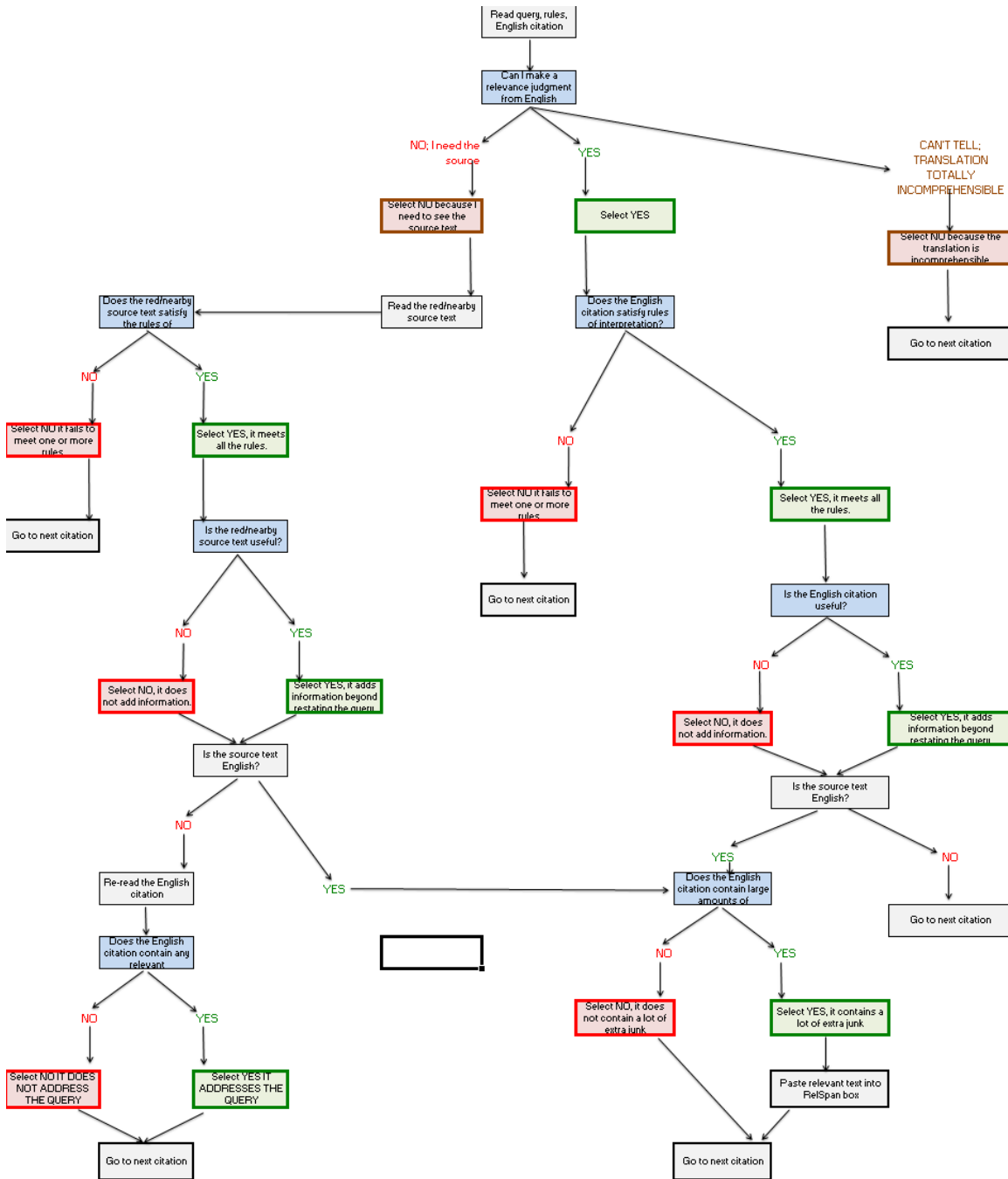
The BOLT IR relevance assessment annotation tool has been designed to support this decision tree model. You will be assigned a "kit" containing multiple citations for a given query. The tool will present one question at a time for a given citation. You must answer each question fully before proceeding to the next question, and you must answer all required questions for a citation before proceeding to the next citation.

In a decision tree, your answers to one question will determine future questions for that citation. For instance, if you judge that a citation fails to meet the rules of interpretation for a query you will not be asked whether the citation is useful, whereas you will be asked about usefulness if you judged the citation to meet the rules.

Note that you cannot climb “up” the decision tree, only down. That means you cannot go back and change your answers for a given citation. Therefore it is very important that you consider each question carefully before recording your answer.

The complete decision tree is shown on the following page. The decision tree is intended to give you an overview of the decision making process for this task. It is not necessary to memorize the decision tree. The annotation tool has been designed to present questions to you one at a time, with all of the logical dependencies applied automatically.

We will examine each question in turn, in the sections that follow.



**BOLT Phase 2 - IR Evaluation Query Assessment Decision Tree**

**4. Rules of Thumb**

This task is all about relevance assessment. Relevance is somewhat subjective: what matters to me may not matter to you. But it's important that we apply a fair and

reasonable standard when assessing the relevance of BOLT IR system citations. Follow these rules of thumb when judging relevance.

### **Rule 1: Be reasonable**

Do not be excessively literal or overly strict when judging citations. For instance, if the query asks “What are people saying about XYZ?”, citations may include some statements of fact about XYZ. For instance:

#### **Example**

**Query:** What do people think about Philadelphia's ban the box legislation?

**Rules:**

- 1: The answer must refer to Philadelphia.
- 2: The answer must refer to the ban the box legislation.

**Citation:**

"Ban the Box" - also known by its less-catchy official name, the Philadelphia Fair Criminal Screening Standards Ordinance - was signed into law by Mayor Nutter in April.

Although the citation above is a statement of fact about Philadelphia’s ban the box legislation, it still addresses the query and satisfies the rules of interpretation.

### **Rule 2: Be generous**

If you are on the fence about how to answer a question, give the machine the benefit of the doubt. This means that you should answer questions as follows when you can’t decide:

**Question:** Can you answer relevance questions based on this English citation alone, without looking at the source text?

**Possible answers:**

- YES
- NO because the translation is incomprehensible.
- NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

→ Choose “**NO because I need to see the source**” if you can’t decide among these options.

**Question:** Does the citation (source) meet the rules of interpretation?

**Possible answers:**

- YES, it meets all the rules.
- NO, it fails to meet one or more rules.

→ **Choose “YES”** if you can’t decide between these options.

**Question:** Is the English citation (source) useful?

**Possible answers:**

- YES, it adds information beyond restating the query.
- NO, it does not add information.

→ **Choose “YES”** if you can’t decide between these options.

**Question:** Does the English citation contain any relevant information?

**Possible answers:**

- YES, it addresses the query.
- NO, it does not address the query.

→ **Choose “YES”** if you can’t decide between these options.

**Question:** Does the English citation contain large amounts of extraneous material?

**Possible answers:**

- YES, it contains a lot of extra junk.
- NO, it does not contain a lot of extra junk.

→ **Choose “NO”** if you can’t decide between these options.

## 5. BOLT IR Queries and Citations

### 5.1 The Query

Before you begin answering any questions you must read the natural language query in the right-hand panel of the annotation tool, along with the rules of interpretation for the query. Make sure you understand the query and its rules before you proceed.

You will be assessing responses to many different kinds of queries. Queries generally ask for information about persons, organizations, locations, facilities, events, movements, practices or customs, products, publications, laws, awards, diseases, or abstract entities. All queries are written in natural language, as a single English sentence.

Each query is accompanied by one or more rules of interpretation. Rules of interpretation are intended to provide some common-sense guidance about the kind of responses that should be considered relevant. Rules of interpretation should

be predictable from the query itself and should not provide any new information or constraints on the query.

Some examples of queries and their rules of interpretation:

**Query:** Describe Russia's relations with the nation of Georgia.

**Rules of interpretation:**

- 1: The answer must refer to the nations of Russia and Georgia.
- 2: The answer must describe relations between Russia and Georgia.

**Query:** Why do people to leave the Church of Latter Day Saints?

**Rules of interpretation:**

- 1: The answer must refer to the Church of Latter Day Saints or Mormonism.
- 2: The answer must refer to causes of people leaving the Church of Latter Day Saints.

**Query:** What are the effects of outsourcing jobs?

**Rules of interpretation:**

- 1: The answer must be about outsourcing jobs.
- 2: The answer must be about the effects of outsourcing.

**NOTE:** If you don't understand the query after reading it and reviewing the rules of interpretation, consult your team leader right away.

## 5.2 The Citation

Citations are answers to the query, automatically generated by the BOLT IR systems. Before you begin answering any questions, you must read the citation carefully.

You will judge multiple citations for each query. The citations are always in English, which means that in some cases they have been automatically translated from Arabic or Chinese by the BOLT IR system. Because the translation is done automatically, it may contain errors. Citations are always short – no more than 250 characters. This means that citations are sometimes missing context that you will need to accurately assess relevance. In this case, you will need to check the original source document to gain a better understanding<sup>1</sup>.

Some examples of citations:

**Citation:** I've already said that I think outdoor smoking bans are silly and unnecessary, I also think any kind of smoking ban on private vehicles or homes is intrusive and unenforceable.

**Citation:** Most tourists enjoy the country from the comfort of their tour buses but on the streets it's a different story.

---

<sup>1</sup> If you are an English annotator, all of the source documents for your citations will be in English. If you are an Arabic annotator, all of the source documents for your citations will be in Arabic. If you are a Chinese annotator, all of the source documents for your citations will be in Chinese.



**Citation:** According to a report by the South Korea's "cultural daily" NBA: recently, a South Korean ancestry player Jeremy Lin owns 1/4.

**Citation:** Cirsiium Governor Tan spandex initially demanded " Wu to the imperial government , Vietnam War ( " Ming the Shih - wen , vol. II tan spandex Amorphophallus thing

As you can see, some citations are very clear even without additional context, while others require additional context to understand their meaning. Some citations are obviously machine translations but are still comprehensible, while other translated citations are completely incomprehensible.

## 6. DECISION: Do you need to see the source text to proceed?

After reading the query, rules and citation carefully, you are ready to make your first decision about the citation.

There are three possible answers to the question "Can you answer relevance questions based on this English citation alone, without looking at the source text?"

- YES.
- NO because the translation is incomprehensible.
- NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

### 6.1 When to choose *YES*

You should choose "YES" if it is possible for you to make a relevance judgment from the English citation alone, without looking at the source text at all. Note that if you choose YES, the annotation tool will not display the source text and all subsequent decisions will be made based on the **citation** alone.

### 6.2 When to choose *NO because the translation is incomprehensible*

Recall that for Arabic and Chinese kits, the citation has been automatically translated into English. In some cases this automatic translation will be so garbled that it is impossible to tell whether or not it is related to the query. You should choose "NO because the translation is incomprehensible" in cases where the English citation is obviously a machine translation, and the quality is so poor that you cannot understand the meaning, and therefore cannot make a relevance judgment one way or the other<sup>2</sup>.

Note that if you choose "NO because the translation is incomprehensible", you will immediately proceed to the next citation.

---

<sup>2</sup> English annotators work on kits containing English source documents only, and so the citations they judge are never machine translations. Therefore, English annotators should never choose "Translation incomprehensible".

### 6.3 When to choose **NO** because *I need to see the source text...*

In some cases, the English citation lacks sufficient context for you to make a decision and you must check the source to proceed. This can happen for a few reasons. For instance, the citation may contain a word (*she, that, it*) or phrase (*the famous musician, what the company said*) whose meaning cannot be understood from the citation alone.

For Arabic and Chinese kits, where the citation has been automatically translated into English, the translation may be somewhat unclear – not completely incomprehensible, but a bit confusing – and you need to check the source to get a better sense of its meaning.

And in some cases, the citation could possibly be relevant but you simply need more context to understand the citation. For instance, the citation might mention *regulations* but you need more context to understand whether this refers to financial regulations, safety regulations, etc.

If any of the following conditions apply:

- You need additional context from the source document to understand the citation
- The citation contains a word or phrase and you can't tell what it refers to
- The translated citation is somewhat unclear or has some errors: it satisfies some rules possibly not all of them, and you need to check the source to be sure

Then you should choose “NO because I need to see the source text...”

Note that if you choose “NO because I need to see the source text...”, the annotation tool will display the source document and subsequent relevance decisions should be based on the **source** rather than the **citation**.

#### **RULE OF THUMB**

**If you can't make up your mind, choose NO because I need to see the source text....**

### 6.4 Examples

Let us consider several examples to illustrate how to make a decision for this first question.

#### **Example**

**Query:** Should smoking be allowed in public areas?

**Rules of interpretation:**

1: The answer must be about smoking.

2: The answer must focus on public areas.

**Citation:**

Yet every country in the world is allowed to damage whatever ecosystem they want.

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

**YES.**

NO because the translation is incomprehensible.

NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The citation is perfectly understandable in isolation, and it is obviously unrelated to the query. There is no need to look at the source.

**Example**

**Query:** Should smoking be allowed in public areas?

**Rules of interpretation:**

1: The answer must be about smoking.

2: The answer must focus on public areas.

**Citation:**

I've already said that I think outdoor smoking bans are silly and unnecessary, I also think any kind of smoking ban on private vehicles or homes is intrusive and unenforceable.

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

**YES.**

NO because the translation is incomprehensible.

NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The citation is perfectly understandable in isolation, and it is obviously related to the query. There is no need to look at the source.

**Example**

**Query:** What do people say about Jeremy Lin's nationality?

**Rules of interpretation:**

1: The answer must be about Jeremy Lin.

2: The answer must be about Jeremy Lin's nationality.

**Citation:**

According to a report by the South Korea's "cultural daily" NBA: recently, a South Korean ancestry player Jeremy Lin owns 1/4.

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

**YES.**

NO because the translation is incomprehensible.

NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The translation isn't perfect, but I can understand it and the citation contains sufficient information for me to judge relevance – it is about Jeremy Lin and discusses his ancestry, which could have bearing on his nationality. There is no need to look at the source.

### Example

**Query:** What do people think about Pope Shenouda III?

**Rules of interpretation:**

1: Answers must be about Pope Shenouda III, not the current pope.

**Citation:**

Delivering a farewell ceremony Pope Shenouda III from inside the Saint Mark Cathedral in Abbasiya on 2012 17 March Al-lahid , a state of extreme sadness controlled Assiut churches after the news of

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

**YES.**

NO because the translation is incomprehensible.

NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The translation isn't perfect, but I can understand it, and the citation contains sufficient information for me to judge relevance – it is about Pope Shenouda III and it mentions a reaction of "extreme sadness", which has some bearing on what people think about him. There is no need to look at the source.

### Example

**Query:** What do people think of the Egyptian economic crisis?

**Rules of interpretation:**

1: Answers must be about Egyptian economy only.

**Citation:**

But the Egyptian society is now in a real crisis of the spread of neglect the upbringing of children (the new generation who will judge

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

**YES.**

NO because the translation is incomprehensible.

NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The translation isn't perfect, but I can understand it, and the citation contains sufficient information for me to judge relevance – it says nothing about the Egyptian economy. There is no need to look at the source.

### Example

**Query:** Is it safe to travel to Ethiopia?

**Rules of interpretation:**

1: The answer must be about Ethiopia.

2: The answer must focus on the issue of safety.

3: The answer must refer to travelling, not living, in Ethiopia.

**Citation:**

Most tourists enjoy the country from the comfort of their tour buses but on the streets it's a different story.

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

YES.

NO because the translation is incomprehensible.

**NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.**

Discussion: The citation says "the country" but doesn't name the country. It could be Ethiopia, so I need to check the source to find out.

### Example

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

1. The answer must be about job hunting.

2. The answer must be about college students.

3. The answer must address the potential difficulty of job hunting among college students.

**Citation:** Cultivated so many college students to find a job place?

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

YES.

- NO because the translation is incomprehensible.  
 NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The citation mentions jobs and college students, but the translation is somewhat unclear. I need to check the source to get a better understanding before I can make a relevance decision.

### Example

**Query:** What are some benefits of taking Omega 3,6,9?

**Rules of interpretation:**

- 1: The answer must be about Omega 3,6,9
- 2: Answers should be about the benefits of taking Omega 3,6,9
- 3: Omega 3,6,9 can also be called Omega 3

**Citation:** Found in fish oil and all the fruits of the sea , and in contrast to cholesterol , which harms human health , fish oil containing omega material prevents objects injured in bloody clots that affect the heart and the brain , fat found in animals from

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

- YES.  
 NO because the translation is incomprehensible.  
 NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The citation mentions “omega material” and “prevents objects injured in blood clots”, but the translation is somewhat unclear. I need to check the source to get a better understanding before I can make a relevance decision.

### Example

**Query:** How does war affect soldiers?

**Citation:** Cirsium Governor Tan spandex initially demanded " Wu to the imperial government , Vietnam War ( " Ming the Shih - wen , vol. II tan spandex Amorphophallus thing

**Can you answer relevance questions based on this English citation alone, without looking at the source text?**

- YES.  
 NO because the translation is incomprehensible.  
 NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation.

Discussion: The translation is completely garbled. I can't understand its meaning at all, and so it's impossible for me judge relevance one way or the other.

## 7. Decisions if you are not viewing the source text

This section covers the questions you will answer if you have chosen not to see the source text. All answers are based on the English citation alone. There are two possible questions to answer in this case. First, you must decide whether the English citation satisfies the rules of interpretation. If you answer NO to this question, then you will immediately proceed to the next citation. If you answer YES to that question, then you must decide whether the English citation is useful. These questions are covered in detail in the following sub-sections.

### 7.1 DECISION: Does the English citation satisfy the rules of interpretation?

If you have chosen not to see the source text, the next question for you to answer is whether the English citation reasonably satisfies the rules of interpretation. There are two possible answers to this question.

Does the English citation satisfy the rules of interpretation?

- YES, it meets all the rules.
- NO, it fails to meet one or more rules.

#### **RULE OF THUMB**

**If you can't make up your mind between YES and NO, choose YES.**

Do not interpret the rules of interpretation too literally or strictly. If a citation can be reasonably inferred to satisfy all the rules of interpretation, then you should answer "YES". If the citation cannot be seen as fitting the rules even when applying a reasonable, generous interpretation, then you should answer "NO".

Note that if you answer "NO" to this question, you will proceed immediately to the next citation.

Let us consider some examples that illustrate how to answer this question.

### Example

**Query:** Should smoking be allowed in public areas?

**Rules of interpretation:**

- 1: The answer must be about smoking.
- 2: The answer must focus on public areas.

**Citation:**

Yet every country in the world is allowed to damage whatever ecosystem they want.

**Does the English citation satisfy the rules of interpretation?**

- YES, it meets all the rules.  
 NO, it fails to meet one or more rules.

Discussion: While an ecosystem could be considered a public place, this citation isn't in the right topic area; there is no mention of smoking whatsoever.

### Example

**Query:** Describe Russia's relations with the nation of Georgia.

**Rules of interpretation:**

- 1: The answer must refer to the nations of Russia and Georgia.
- 2: The answer must describe the relations between Russia and Georgia.

**Citation:**

Tensions over Georgia and uncertainty over the future of Ukraine, whose pro-Western president wants the country to join NATO, are hurdles in efforts by Russia and the U.S. to mend strained ties.

**Does the English citation satisfy the rules of interpretation?**

- YES, it meets all the rules.  
 NO, it fails to meet one or more rules.

Discussion: Even though Georgia and Russia are both mentioned here and the statement is on the topic of international relations, the statement only refers to the relations between Russia and the US, not Russia and Georgia.

### Example

**Query:** Should smoking be allowed in public areas?

**Rules of interpretation:**

- 1: The answer must be about smoking.
- 2: The answer must focus on public areas.

**Citation:**

I've already said that I think outdoor smoking bans are silly and unnecessary, I also think any kind of smoking ban on private vehicles or homes is intrusive and unenforceable.

**Does the English citation satisfy the rules of interpretation?**

- YES, it meets all the rules.  
 NO, it fails to meet one or more rules.

Discussion: The citation mentions both smoking and smoking in public areas, and thus fits both of the rules.

### Example



**Query:** What did people think about Sarah Palin possibly running for president?

**Rules of interpretation:**

- 1: The answer must be about Sarah Palin.
- 2: The answer must be about the possibility of Sarah Palin running for president.

**Citation:**

Sarah Palin in West Palm Beach: "I'll run for president if the American people want me to!"

**Does the English citation satisfy the rules of interpretation?**

- YES, it meets all the rules.**
- NO, it fails to meet one or more rules.

Discussion: The citation is a statement from Sarah Palin about her possible presidential bid, and thus satisfies all the rules.

**TIP:** Do not take the phrases "think about" or "saying about" too literally when they appear in a query. Any citation that satisfies the rules of interpretation should receive a "YES" answer, even if the citation is a statement or fact rather than an opinion.

**Remember: Be reasonable! Be generous!**

**Example**

**Query:** What are the influences of Euro financial crisis on China?

**Rules of interpretation:**

- 1: Answers must be about Euro financial crisis rather than any other country's economic crisis.
- 2: Answers should be about the effects of Euro financial crisis on China rather than other countries.
- 3: Answers must be the effects instead of any other things about Euro economic crisis.

**Citation:**

Due to the spread of the European debt crisis has intensified, the us economic recovery is sluggish, further deterioration of the external environment in the development of the Asian economies.

**Does the English citation satisfy the rules of interpretation?**

- YES, it meets all the rules.**
- NO, it fails to meet one or more rules.

Discussion: The citation discusses the Euro crisis and its effects. While the citation doesn't mention China explicitly, it is reasonable to infer that China is one of the "Asian economies", so we can conclude that the citation fits the rules of interpretation.

## 7.2 DECISION: Is the English citation useful?

If your citation satisfies the rules of interpretation, you will also be asked to judge the citation's utility. There are two possible answers to this question.

Is the English citation useful?

- YES, it adds information beyond restating the query.
- NO, it does not add information.

By “useful”, we mean that the citation contributes to your understanding of the query. The citation can add a very small amount of information and still be useful, but the citation should not simply re-state the query without adding anything new or informative.

Note that if you are working on English kit, you will be presented with one more question about the citation after this (regardless of whether you answered YES or NO). This question is described in section 9 below. If you are working on an Arabic or Chinese kit, you will immediately proceed to the next citation after answering this question (regardless of whether you answered YES or NO).

### **RULE OF THUMB**

**If you can't make up your mind between YES and NO, choose YES.**

Let us consider some examples that illustrate how to answer this question.

### **Example**

**Query:** What did people think about Sarah Palin possibly running for president?

**Rules of interpretation:**

- 1: The answer must be about Sarah Palin.
- 2: The answer must be about the possibility of Sarah Palin running for president.

**Citation:**

"I think there's a really good chance Sarah Palin could become president, and I think that's a really scary thing," Fox quoted Damon, as saying.

**Is the English citation useful?**

- YES, it adds information beyond restating the query.  
 NO, it does not add information.

**Discussion:** The citation is informative. We learn someone's opinion about the prospect of Sarah Palin's potential presidency.

## Example

**Query:** What did people think about Sarah Palin possibly running for president?

**Rules of interpretation:**

- 1: The answer must be about Sarah Palin.
- 2: The answer must be about the possibility of Sarah Palin running for president.

**Citation:**

Palin has been rumored to be considering headquartering her 2012 White House campaign, if there is one, in Scottsdale.

**Is the English citation useful?**

- YES, it adds information beyond restating the query.**
- NO, it does not add information.**

Discussion: The citation is informative. We learn where Palin's presidential campaign may be headquartered.

## Example

**Query:** What did people think about Sarah Palin possibly running for president?

**Rules of interpretation:**

- 1: The answer must be about Sarah Palin.
- 2: The answer must be about the possibility of Sarah Palin running for president.

**Citation:**

WSJ/NBC Poll: Would You Like to See Palin as President Someday?

**Is the English citation useful?**

- YES, it adds information beyond restating the query.**
- NO, it does not add information.**

Discussion: The citation is not informative. It simply restates the query without adding anything new.

## Example

**Query:** Is it safe to travel to Ethiopia?

**Rules of Interpretation:**

- 1: The answer must be about Ethiopia.
- 2: The answer must focus on the issue of safety.
- 3: The answer must refer to traveling, not living, in Ethiopia.

**Citation:**

And, if it is possible and safe, am I likely to be able to get a Sudanese visa here in Addis?

**Is the English citation useful?**

- YES, it adds information beyond restating the query.  
 NO, it does not add information.

Discussion: The citation satisfies the rules of interpretation. It mentions getting a visa, from which we can infer that the person is traveling. It mentions being in Addis (Addis Abbaba), which is the capital of Ethiopia. It mentions that safety is a potential concern. But is the citation useful? It is not. The citation does not add to our understanding of the query; it simply mentions safety as a possible issue without giving any actual information about the safety of traveling in Ethiopia.

### Example

**Query:** Describe Russia's relations with the nation of Georgia.

**Rules of interpretation:**

- 1: The answer must refer to the nations of Russia and Georgia.
- 2: The answer must describe relations between Russia and Georgia.

**Citation:**

Russia is currently cleaning its military bases in Georgia.

**Is the English citation useful?**

- YES, it adds information beyond restating the query.  
 NO, it does not add information.

Discussion: The citation satisfies the rules of interpretation. It mentions Russia and Georgia, and the fact that Russia has military bases in Georgia provides allows us to infer a very small amount of new information about the relationship between the two countries (namely that they have some sort of military relationship). This does not add greatly to our understanding of the relations between Russia and Georgia, but since we are trying to be reasonable and generous in our assessments, we can conclude that this citation does adds information beyond what is stated in the query.

## 8. Decisions if you are viewing the source text

This section covers the questions you will answer if you have chosen to see the source text. There are three possible questions to answer in this case. Answers to the first two questions are based on the source text rather than the English citation. First, you must decide whether the source text satisfies the rules of interpretation. If you answer YES to that question, then you must decide whether the source text is useful. If you have a non-English kit and answer YES to the first question, you will be asked one additional question based on the English citation (not the source text): Does the English citation above contain any relevant information?

**NOTE:** Because you have chosen to use the source text to judge relevance, the annotation tool will display the source text in the middle panel. The piece of source text corresponding to the English citation will be **highlighted in red**. You can use any

part of the document to resolve pronouns or other referents, to clarify the meaning of a translated English citation, or to understand the context of the citation.

**NOTE:** When making judgments about rules of interpretation and the usefulness of the source text, you should focus on the red text. However, it may happen that a correct answer to the query appears outside of the red text. **If the correct answer to the query appears close to the red text – within a sentence or two in either direction – you should also consider that content when judging the source text for rules and usefulness.** In many cases, the source text actually contains more or different information than the English citation. If you are viewing the source text, be sure to make your relevance decisions based on the red and nearby text – not just the portion of the source text that corresponds to the English citation.

**NOTE:** You should read however much source document context is necessary to understand the citation's meaning. This can include the headline of a post or thread. If checking the headline of a post or thread contributes to your understanding of a citation or its context, you should take this information into account when making your assessments.

The questions you will answer if you have chosen to see the source text are covered in detail in the following sub-sections.

### **8.1 DECISION: Does the red/nearby source text satisfy the rules of interpretation?**

If you have chosen to see the source text, the next question for you to answer is whether the source text reasonably satisfies the rules of interpretation. There are two possible answers to this question.

Does the red/nearby source text satisfy the rules of interpretation?

- YES, it meets all the rules.
- NO, it fails to meet one or more rules.

#### **RULE OF THUMB**

**If you can't make up your mind between YES and NO, choose YES.**

Do not interpret the rules of interpretation too literally or strictly. If the red/nearby source text can be reasonably inferred to satisfy all the rules of interpretation, then you should answer “YES”. If the red/nearby source text cannot be seen as fitting the rules even when applying a reasonable, generous interpretation, then you should answer “NO”.

Note that if you answer “NO” to this question, you will proceed immediately to the next citation.

Let us consider some examples that illustrate how to answer this question.

NOTE: In these examples, we provide a partial human translation where necessary, to help illustrate the point being made by the example. However, human translations are not available during the actual relevance assessment task.

### Example

**Query:** Is it safe to travel to Ethiopia?

**Rules of interpretation:**

- 1: The answer must be about Ethiopia.
- 2: The answer must focus on the issue of safety.
- 3: The answer must refer to travelling, not living, in Ethiopia.

**Citation:**

One myth i don't buy is that travelling in a group, with or without armed guards, in an organized tour booked well in advance is safer.

**Source Text:**

The trap tourists in Ethiopia always fall into is the whole "safety in numbers" thing. *One myth i don't buy is that travelling in a group, with or without armed guards, in an organized tour booked well in advance is safer.*

**Does the red/nearby source text satisfy the rules of interpretation?**

- YES, it meets all the rules.**  
 NO, it fails to meet one or more rules.

Discussion: The citation itself mentions safety and travel, but we don't know whether it is in Ethiopia. The nearby text makes it clear that the citation is about traveling in Ethiopia. Based on this contextual information, we can conclude that the citation meets the rules and is informative.

### Example

**Query:** How is Obama reacting to gun violence in the U.S.?

**Rules of interpretation:**

- 1: The answer must be about Obama's reactions.
- 2: The answer must be about reactions to gun violence in the U.S.

**Citation:**

Obama has said he supports more restrictions on sales.

**Source Text:**

*Obama has said he supports more restrictions on sales.* Given the tobacco industry's influence in this country though, there's not much he can do to stop kids from buying cigarettes.

**Does the red/nearby source text satisfy the rules of interpretation?**

- YES, it meets all the rules.  
 NO, it fails to meet one or more rules.

Discussion: The red/nearby source text makes clear that the Obama’s reaction of supporting “more restrictions on sales” is referring to restrictions on the sale of tobacco, not firearms. Thus, we can conclude that the red/nearby source text doesn’t satisfy the rules of interpretation.

### Example

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

1. The answer must be about job hunting.
2. The answer must be about college students.
3. The answer must address the potential difficulty of job hunting among college students.

**Citation:** Cultivated so many college students to find a job place?

**Source Text:** 培养出那么多找不到工作的大学生有意义吗？

**Partial Human Translation:** *Is it meaningful to cultivate so many college students who can’t find jobs?*

**Does the red/nearby source text satisfy the rules of interpretation?**

- YES, it meets all the rules.  
 NO, it fails to meet one or more rules.

Discussion: The source text clearly states that many college students can’t find jobs, so we can conclude this citation fits the rules of interpretation.

### Example

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

- 1: The answer must be about job hunting.
- 2: The answer must be about college students.
- 3: The answer must address the potential difficulty of job hunting among college students.

**Citation:** In the choice of employment in cities , college students are concentrated in Shanghai , Beijing and other places

**Source Text:** 在就业城市的选择上, 大学生的目光集中在上海、北京等地。求职的工资底线的平均值为 2390. 38 元, 与 2001 年的大学生求职的工资底线平均值 2244. 6

元相比，依然偏高。专家分析说，现在大学生在求职时选择企业的标准相当一致，就业目标相对集中，这也是近年来大学生就业难的重要原因之一。

**Partial Human Translation:** *Currently when college students are seeking employment, their consideration of options is quite similar and their targets are relatively too focused. This is one of the important reasons for college students' employment difficulty in recent years.*

**Does the red/nearby source text satisfy the rules of interpretation?**

**YES, it meets all the rules.**

NO, it fails to meet one or more rules.

Discussion: The red/nearby source text satisfies the rules of interpretation. Although the red text doesn't mention job hunting difficulties for college students, the very next sentence states "This is one of the important reason for college students' employment difficulty." When we consider the red source text plus the nearby source text, we conclude that the rules of interpretation have been satisfied.

## Example

**Query:** What are people's opinions about the effects of gay marriage's legalization?

**Rules of interpretation:**

- 1: Answers must be about people's opinions about the effects of the legalization of gay marriage rather than other things.
- 2: Answers should be about the effects instead of the causes of gay marriage's legalization
- 3: Answers should be about gay marriage rather than other types of marriages.

**Citation:** A grandmother level same - sex partners in Niagara Falls first time buried the sunroof, becoming New York 's first legal gay couples

**Source Text:** 一对祖母级同性伴侣第一时间在尼亚加拉瀑布拉埋天窗，成为纽约第一对合法的同性恋新人。54 岁的兰伯特和 53 岁的拉德昨日凌晨身处宏伟的尼亚加拉瀑布前，在象征同性恋者尊严的彩虹色彩衬托下，由市长主持婚礼。一对新人在交换誓词后深情一吻，在场见证的数百名群众不禁热烈欢呼。

**Partial human translation:** *A pair of same-sex partners who are old enough to be grandparents became the first legal same-sex couple in New York State. The couple exchanged vows and kissed. Hundreds of people who witnessed this moment cheered for them.*

**Does the red/nearby source text satisfy the rules of interpretation?**

**YES, it meets all the rules.**

NO, it fails to meet one or more rules.



Discussion: The citation is in the right topic area but the translation is a bit funny (“buried the sunroof”), so checking the source text is necessary. Reading both the red source text and the nearby text, we understand the rules of interpretation are satisfied. The source describes the crowd’s reaction to the first legal gay marriage in New York State (namely, cheering). This reaction can be reasonably understood to be an opinion about legalization of gay marriage. Therefore, the source text satisfies the rules.

## Example

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

1. The answer must be about job hunting.
2. The answer must be about college students.
3. The answer must address the potential difficulty of job hunting among college students.

**Citation:** Her job is to provide accommodation for job - seeking college students came to Shanghai , information services , there are also psychological comfort

**Source Text:**

铺设了 58 张床位，她的工作就是为来沪求职的大学生房客提供住宿、信息服务，还有心理安慰。非大学生不能入住 不是“圈内人”很难找到“求职驿站”的地址。它的真实位置在龙吴路一家部队招待所的 3 楼和 5 楼。所有寻上门的住户，多是通过同学介绍，口耳相传，一下火车慕名而来。还有些人是在上海的人才招聘会上看

**Partial human translation:** Her job is to provide boarding, information and consoling to college students seeking jobs in Shanghai.

**Does the red/nearby source text satisfy the rules of interpretation?**

- YES, it meets all the rules.  
 NO, it fails to meet one or more rules.

Discussion: The source text is about college students hunting for jobs, which satisfies Rules 1 and 2. However, the source text doesn’t mention anything (directly or indirectly) about the potential difficulty of job hunting, so it fails on Rule 3.

## Example

**Query:** What are some benefits of taking Omega 3,6,9?

**Rules of interpretation:**

- 1: The answer must be about Omega 3,6,9
- 2: Answers should be about the benefits of taking Omega 3,6,9
- 3: Omega 3,6,9 can also be called Omega 3

**Citation:** Found in fish oil and all the fruits of the sea , and in contrast to cholesterol , which harms human health , fish oil

containing omega material prevents objects injured in bloody clots that affect the heart and the brain , fat found in animals from

**Source Text:**

وجود مادة الأوميغا -3. الموجودة في دهن السمك و جميع ثمار البحر، وعلى عكس الكوليسترول الذي يضر بصحة الانسان ، ، فإن احتواء دهن السمك على مادة اوميغا يمنع إصابة الاجسام بالجلطات الدموية التي تصيب القلب و الدماغ فإن كانت الدهون الموجودة في الحيوانات من غنم و بقر تضر بالصحة ، فإن د

**Partial Human Translation:** *The presence of Omega 3 found in fish oil and all the fruits of the sea, and in contrast to cholesterol, which harms human health, fish oil containing omega material prevents objects injured in blood clots that affect the heart and the brain, fat found in animals from*

**Does the red/nearby source text satisfy the rules of interpretation?**

**YES, it meets all the rules.**

NO, it fails to meet one or more rules.

Discussion: The source text states that fish oil containing “omega material” prevents blood clot, which is clearly about the benefits of Omega 3. Thus we can conclude that the source fits the rules of interpretation.

## Example

**Query:** What do people think about Pope Shenouda III?

**Rules of interpretation:**

1: Answers must be about Pope Shenouda III, not the current pope.

**Citation:** A church source of crime and insulting Islam and leave the law is Pope Shenouda III

**Source Text:**

مصدر الإجرام الكنسى وسب الإسلام والخروج على القانون هو شنودة الثالث.. مصدر الخراب والفتنة

**Partial human translation:** *Pope Shenouda III is a person who insults Islam, violates the law and is the source of church crimes.*

**Does the red/nearby source text satisfy the rules of interpretation?**

**YES, it meets all the rules.**

NO, it fails to meet one or more rules.

Discussion: The English citation is a bit confusing, but the red/nearby source text contains a statement of someone’s opinion about Pope Shenouda III, namely that he “violates the law and is the source of church crimes”. Thus we can conclude that the source satisfies the rules.

## Example

**Query:** What happens if you get addicted to the internet?

**Rules of interpretation:**

1: The answer must be about the effects of internet addiction, not its causes

**Citation:**

This warning is directed to all the mothers of the internet addicts who like my situation and especially the guys Al-raghayat charming here.

**Source Text:**

هذا التحذير اوجهه لكل الأمهات مدمنى النت اللى زى حالاتى و خصوصا جماعة الرغايات الفاتنات اللى عندنا

**Partial human translation:** *This warning is directed to all mothers who are addicted to internet, like in my case, especially those who are talkative and pretty.*

**Does the red/nearby source text satisfy the rules of interpretation?**

- YES, it meets all the rules.  
 NO, it fails to meet one or more rules.

Discussion: The English citation seems to be in the right topic area, but the translation is a bit confusing. When we check the source text we see that it discusses people who are addicted to the internet, but it doesn't say anything about the effects of internet addiction. Therefore the citation fails to meet the rules.

### 8.2 DECISION: Is the red/nearby source text useful?

If the red/nearby source text satisfies the rules of interpretation, you will also be asked to judge the its utility. There are two possible answers to this question.

Is the red/nearby source text useful?

- YES, it adds information beyond restating the query.
- NO, it does not add information.

By “useful”, we mean that the red/nearby source text contributes to your understanding of the query. The text can add a very small amount of information and still be useful, but it should not simply re-state the query without adding anything new or informative.

Note that if you are working on English kit, you will be presented with one more question about the English citation after this (regardless of whether you answered YES or NO), described in section 9. If you are working on an Arabic or Chinese kit,

you will be presented with a different question about the English citation, (regardless of whether you answered YES or NO), described in section 8.3.

### **RULE OF THUMB**

**If you can't make up your mind between YES and NO, choose YES.**

Let us consider some examples that illustrate how to answer this question.

#### **Example**

**Query:** Is it safe to travel to Ethiopia?

**Rules of interpretation:**

- 1: The answer must be about Ethiopia.
- 2: The answer must focus on the issue of safety.
- 3: The answer must refer to travelling, not living, in Ethiopia.

**Citation:**

One myth i don't buy is that travelling in a group, with or without armed guards, in an organized tour booked well in advance is safer.

**Source Text:**

The trap tourists in Ethiopia always fall into is the whole "safety in numbers" thing. **One myth i don't buy is that travelling in a group, with or without armed guards, in an organized tour booked well in advance is safer.**

**Is the red/nearby source text useful?**

**YES, it adds information beyond restating the query.**

NO, it does not add information.

Discussion: The citation itself doesn't mention Ethiopia, but the red/nearby source text indicates that one is not safe traveling in Ethiopia just because one is in a group. Therefore, the red/nearby source adds information beyond what is stated in the query.

#### **Example**

**Query:** Why is climate change happening?

**Rules of interpretation:**

- 1: The answer must reference causes.
- 2: The answer must be about climate change.

**Citation:**

focus more on the causes of the issue

**Source Text:**

Jonathan Foley, the director of the Institute on the Environment at the University of Minnesota, has said that leaders need to

concentrate less on the politics of climate change, and **focus more on the causes of the issue**

**Does the red/nearby source text satisfy the rules of interpretation?**

- YES, it adds information beyond restating the query.  
 NO, it does not add information.

Discussion: Because the citation refers to “the issue”, we check the source text to learn what the issue is. When we check we find that the red/nearby source text does indeed mention climate change and the importance of focusing on its causes. However, it doesn’t state what those causes are. Thus, we must conclude that the red/nearby text doesn’t add information beyond what is stated in the query.

### Example

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

1. The answer must be about job hunting.
2. The answer must be about college students.
3. The answer must address the potential difficulty of job hunting among college students.

**Citation:** Cultivated so many college students to find a job place?

**Source Text:** 培养出那么多找不到工作的大学生有意义吗？

**Partial Human Translation:** *Is it meaningful to cultivate so many college students who can't find jobs?*

**Is the red/nearby source text useful?**

- YES, it adds information beyond restating the query.  
 NO, it does not add information.

Discussion: The source text clearly states that many college students can’t find jobs, so we must conclude that it provides information beyond just what is stated in the query.

### Example

**Query:** What do people say about Jeremy Lin's nationality?

**Rules of interpretation:**

- 1: The answer must be about Jeremy Lin.
2. The answer must be about Jeremy Lin’s nationality.

**Citation:** Wang Peng , " said Jeremy Lin 's success shows that in the correct mode , as long as their own efforts , everything is possible

**Source Text:**

王一鹏说，“林书豪的成功说明，在正确的模式下，只要自己努力，一切都有可能。网上一个很久不打篮球的哥们说，看完林书豪，明天就要去打篮球。” 国际体育流派众多，不同民族运动员的特长项目不尽相同。长期以来，人们认为亚洲人更适合技巧性运动，在身体对抗激烈的足球和篮球等项目相对弱勢。

**Partial human translation:** Wang Yipeng said, "Jeremy Lin's success demonstrated that if you work hard, anything is possible in the right mode. A netizen friend who hasn't played basketball said that he planned to plan basketball tomorrow after watching Jeremy Lin." For a long time, people believe that Asian people are more suitable for sports requiring skills and are relatively weaker in sports like soccer and basketball which requires body contacts.

**Is the red/nearby source text useful?**

**YES, it adds information beyond restating the query.**

NO, it does not add information.

Discussion: The citation mentions Jeremy Lin (satisfying one rule) but doesn't discuss his nationality. Because the translation is awkward we check the source text to see whether it contains additional information, and we see that the source text does meet the rules of interpretation. Because we can infer from the source text that Jeremy Lin is an Asian basketball player, and can conclude that the red/nearby source text provides information beyond what is stated in the query.

## Example

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

1. The answer must be about job hunting.
2. The answer must be about college students.
3. The answer must address the potential difficulty of job hunting among college students.

**Citation:** In the choice of employment in cities , college students are concentrated in Shanghai , Beijing and other places

**Source Text:** 在就业城市的选择上，大学生的目光集中在上海、北京等地。求职的工资底线的平均值为 2390. 38 元，与 2001 年的大学生求职的工资底线平均值 2244. 6 元相比，依然偏高。专家分析说，现在大学生在求职时选择企业的标准相当一致，就业目标相对集中，这也是近年来大学生就业难的重要原因之一。

**Partial Human Translation:** *Currently when college students are seeking employment, their consideration of options is quite similar and their targets are relatively too focused. This is one of the important reasons for college students' employment difficulty in recent years.*

**Is the red/nearby source text useful?**

**YES, it adds information beyond restating the query.**

NO, it does not add information.

Discussion: Although the red text doesn't mention job hunting difficulties for college students, the very next sentence is "This is one of the important reasons for college students' employment difficulty in recent years." When we consider the red source text plus the nearby source text, we conclude that the red/nearby text adds information beyond restating the query.

### Example

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

1. The answer must be about job hunting.
2. The answer must be about college students.
3. The answer must address the potential difficulty of job hunting among college students.

**Citation:** survey students enter in job market if employment difficult

**Source Text:**

最近中国青年报对即将进入就业市场的学生问卷调查，询问他们是否有就业困难。

**Partial human translation:** *Recently Chinese Youth Daily conducted a survey among graduating college students, asking whether they have difficulty in seeking employment.*

**Is the red/nearby source text useful?**

- YES, it adds information beyond restating the query.  
 NO, it does not add information.

Discussion: The source and nearby text is about college students looking for jobs, but it only restates the query without adding new information.

### Example

**Query:** What are some benefits of taking Omega 3,6,9?

**Rules of interpretation:**

- 1: The answer must be about Omega 3,6,9
- 2: Answers should be about the benefits of taking Omega 3,6,9
- 3: Omega 3,6,9 can also be called Omega 3

**Citation:** Found in fish oil and all the fruits of the sea , and in contrast to cholesterol , which harms human health , fish oil containing omega material prevents objects injured in bloody clots that affect the heart and the brain , fat found in animals from

Source Text:

وجود مادة الأوميغا -3. الموجودة في دهن السمك و جميع ثمار البحر، وعلى عكس الكوليسترول الذي يضر بصحة الانسان ، ، فإن احتواء دهن السمك على مادة اوميغا يمنع إصابة الاجسام بالجلطات الدموية التي تصيب القلب و الدماغ فإن كانت الدهون الموجودة في الحيوانات من غنم و بقر تضر بالصحة ، فإن د

*Partial Human Translation: The presence of Omega 3 found in fish oil and all the fruits of the sea, and in contrast to cholesterol, which harms human health, fish oil containing omega material prevents objects injured in bloody clots that affect the heart and the brain, fat found in animals from*

Is the red/nearby source text useful?

YES, it adds information beyond restating the query.

NO, it does not add information.

Discussion: The source text states that fish oil containing Omega 3 prevents blood clots, which is clearly about the benefits of Omega 3. Thus we can conclude that the source adds information beyond just what is stated in the query.

### Example

**Query:** What do people think about Pope Shenouda III?

**Rules of interpretation:**

1: Answers must be about Pope Shenouda III, not the current pope.

**Citation:** A church source of crime and insulting Islam and leave the law is Pope Shenouda III

Source Text:

مصدر الإجرام الكنسى وسب الإسلام والخروج على القانون هو شنودة الثالث.. مصدر الخراب والفتنة

*Partial human translation: Pope Shenouda III is a person who insults Islam, violates the law and is the source of church crimes.*

Is the red/nearby source text useful?

YES, it adds information beyond restating the query.

NO, it does not add information.

Discussion: The source text is someone's opinion about Pope Shenouda III, namely that he "insults Islam, violates the law and is the source of church crimes", which adds information beyond just what is stated in the query.

### 8.3 DECISION: Does the English Citation contain any relevant information?

As stated above, this question only applies to Arabic and Chinese annotators who are working on kits containing non-English source documents. The annotation tool will not display this question for English kits.



When you choose to see the source text and answer YES to the question “Does the red/nearby source text satisfy the rules of interpretation?”, you are required to answer one additional question about the **English citation**.

Does the English citation above contain any relevant information?

- YES, it addresses the query.
- NO, it does not address the query.

Before answering this question, re-read the English citation. Then decide whether the English citation (not the source text) addresses the query at all.

**NOTE:** You are not trying to decide if the English citation satisfies all the rules and is useful. After all, if the English citation satisfied all the rules and was useful, there would have been no need to look at the source text in the first place!

Instead, you are trying to decide whether the English citation is at all relevant to the query. A good rule of thumb: if you had been searching on this query, would you feel that the English citation addressed your query at all.

Although the English citation has been automatically translated from Arabic or Chinese, note that you are **not** judging the translation quality directly. Instead, you are judging whether the translated citation addresses the query. It is possible to have a poor translation but still answer YES to the question “Does the English citation contain any relevant information”. Conversely, it is possible to have a good translation but still answer NO to the question “Does the English citation contain any relevant information”.

Note that after you answer this question, you will immediately proceed to the next citation (regardless of whether you answer YES or NO).

#### **RULE OF THUMB**

**If you can't make up your mind between YES and NO, choose YES.**

Let us consider some examples that illustrate how to answer this question.

#### **Example**

**Query:** Is job hunting tough for college students?

**Rules of interpretation:**

1. The answer must be about job hunting.
2. The answer must be about college students.
3. The answer must address the potential difficulty of job hunting among college students.

**Citation:** In the choice of employment in cities , college students are concentrated in Shanghai , Beijing and other places

**Source Text:** 在就业城市的选择上, 大学生的目光集中在上海、北京等地。求职的工资底线的平均值为 2390. 38 元, 与 2001 年的大学生求职的工资底线平均值 2244. 6 元相比, 依然偏高。专家分析说, 现在大学生在求职时选择企业的标准相当一致, 就业目标相对集中, 这也是近年来大学生就业难的重要原因之一。

**Partial Human Translation:** *Currently when college students are seeking employment, their consideration of options are quite similar and their targets are relatively too focused. This is one of the important reasons for college students' employment difficulty in recent years.*

**Does the English citation contain any relevant information?**

YES, it addresses the query.

NO, it does not address the query.

Discussion: While the source text is relevant to the query, the English citation does not contain any information that discusses college students' difficulties in finding employment. Thus we must conclude that the English citation does not address the query.

## Example

**Query:** What do people say about Jeremy Lin's nationality?

**Rules of interpretation:**

1: The answer must be about Jeremy Lin.

2. The answer must be about Jeremy Lin's nationality.

**Citation:** How will " voices " into reality , the Chinese Basketball Association to make practical efforts , needs more Jeremy Lin determined to seize the opportunity to attract talent

**Source Text:**

如何将“呼声”转化为现实, 需要中国篮协为吸引人才作出切实努力, 更需要林书豪下定决心, 抓住机遇。对这位祖籍中国浙江的华裔球员实力十分了解的姚明接受采访时评价说, 林书豪如果真能加盟的话, 对中国男篮肯定会有帮助。 不过, 即便林书豪最终未能加盟中国男篮, 他在 NBA 的这次爆发, 也给包括郭艾伦、杨鸣、于澍龙等众多年轻的中国球员的成长提供了成功的借鉴。

**Partial Human Translation:** *Yaoming, who knows this player whose ancestors were born in Zhejiang, China very well, commented when he was interviewed that if Jeremy Lin could join CNBT, ...*

**Does the English citation contain any relevant information?**

YES, it addresses the query.

NO, it does not address the query.

Discussion: The source text is relevant to the query because it discusses Lin's ancestry, which has bearing on the question of his nationality. The English citation

on the other hand says nothing about Lin’s ancestry or nationality, it only mentions him by name. Therefore, we must conclude that the citation does not address the query.

### Example

**Query:** What do people say about Jeremy Lin's nationality?

**Rules of interpretation:**

1: The answer must be about Jeremy Lin's nationality, not other entity.

**Citation:** In the eyes of one thousand people has one thousand Hamlet , but if we go beyond nationality , ethnic vision , put aside the debate of these external environment , to carefully explore

**Source citation:**一千个人眼中有一千个哈姆雷特，但是如果我们超出国籍、族裔的视野，放下这些外部环境的争论，去细细探求林书豪的成功之路，就会发现，出身也好、国籍也好，都不是最重要的东西，他的成功，首先得益于自身的努力。人前的风光，源自背后的汗水。娴熟的运球，精准的投篮，都来自极端刻苦的训练。即便是在联赛停摆、饭碗无着之时，他的训练也没有中断。而来大陆做活动时，他问得最多的是能否解决训练场所、条件如何、离宾馆多远……天分固然重要，但是如果

**Partial Human Translation:** Jeremy’s nationality is not the important factor in his success. His success comes from his hard work.

**Does the English citation contain any relevant information?**

YES, it addresses the query.

NO, it does not address the query.

Discussion: The English citation mentions “nationality” but it does not mention Jeremy Lin at all. Therefore, we must conclude that it does not address the query.

### Example

**Query:** What are some benefits of taking Omega 3,6,9?

**Rules of interpretation:**

1: The answer must be about Omega 3,6,9

2: Answers should be about the benefits of taking Omega 3,6,9

3: Omega 3,6,9 can also be called Omega 3

**Citation:** Found in fish oil and all the fruits of the sea , and in contrast to cholesterol , which harms human health , fish oil containing omega material prevents objects injured in bloody clots that affect the heart and the brain , fat found in animals from

Source Text:

وجود مادة الأوميغا -3. الموجودة في دهن السمك و جميع ثمار البحر، وعلى عكس الكوليسترول الذي يضر بصحة الانسان ، ، فإن احتواء دهن السمك على مادة اوميغا يمنع إصابة الاجسام بالجلطات الدموية التي تصيب القلب و الدماغ فإن كانت الدهون الموجودة في الحيوانات من غنم و بقر تضر بالصحة ، فإن د

*Partial Human Translation: The presence of Omega 3 found in fish oil and all the fruits of the sea, and in contrast to cholesterol, which harms human health, fish oil containing omega material prevents objects injured in bloody clots that affect the heart and the brain, fat found in animals from*

Does the English citation contain any relevant information?

YES, it addresses the query.

NO, it does not address the query.

Discussion: The English citation has some translation problems, but it is still possible to understand that it is discussing the benefits of “fish oil containing omega material”. The English citation contrasts omega material with cholesterol “which harms human health” and seems to say that omega acids can prevent blood clots. Thus, we must conclude that the English citation does address the query in some way.

## Example

**Query:** What are some benefits of taking Omega 3,6,9?

**Rules of interpretation:**

1: The answer must be about Omega 3,6,9

2: Answers should be about the benefits of taking Omega 3,6,9

3: Omega 3,6,9 can also be called Omega 3

**Citation:**

Who have higher levels of Omega 3 was learning difficulties

**Source text:**

الذين لديهم مستويات أعلى من أوميغا 3 كانت صعوبات التعلم لديهم أقل.

*Partial human translation: Those with higher levels of Omega 3 have fewer learning difficulties.*

Does the English citation contain any relevant information?

YES, it addresses the query.

NO, it does not address the query.

Discussion: While the source text is relevant to the query, the English citation does not contain the key piece of information that points to a benefit of Omega-3: the English citation leaves out the key word “fewer”, and thus appears to be describing a

drawback rather than a benefit of taking Omega-3. Since the citation does not discuss benefits, we must conclude that it does not address the query.

## 9. DECISION: Determine if the citation contains extraneous material

[Note that as stated above, this question will NOT be presented for Arabic and Chinese kits.]

If you are working on an English kit and have judged the English citation to meet the rules of interpretation (when the source text is not viewed), or you have judged that the English citation contains relevant information (when the source text is viewed), then you must make one additional decision.

Does the English citation contain large amounts of extraneous material?

- YES, it contains a lot of extra junk.
- NO, it does not contain a lot of extra junk.

If you answer “YES” to this question, you will be prompted to copy/paste the relevant span (relspan) of the citation into a text box. You should swipe your mouse over the string of text that is relevant and paste that text into the box.

→ Do not select individual words to paste in – swipe a string and paste the full string

→ Do not edit the text you paste into the text box – just swipe, copy and paste, and move on to the next citation

By “extraneous material”, we mean a significant amount of text that is completely irrelevant to the query. The goal of this question is not to do precision editing of a good citation. Instead, the goal is to flag citations that potentially waste the user’s time because they contain too much extra material.

Do not answer YES if the citation contains a bit of extra content that could have been left out without affecting the citation’s relevance. Only answer YES if the citation contains a great deal of extra material unrelated to the query.

The vast majority of citations should be marked as “NO”, and require no relspan editing.

### **RULE OF THUMB**

**If you can’t make up your mind between YES and NO, choose NO.**

Let us consider some examples that illustrate how to answer this question.

Note that if you answer NO to this question, you will immediately proceed to the next citation. If you answer YES to this question, you will proceed to the next citation after you copy paste a relspan string into the text box.

## 9.1 Examples that do not contain large amounts of extraneous material

In all of the following examples the correct answer is “NO, it does not contain a lot of extra junk”. These citations may contain some content that does not directly address the query, but the citation is primarily on topic and should not be edited at all.

All of the following examples are taken from the Dry Run queries, and were incorrectly edited for relspan. The original incorrect answers are shown for contrast, with the text marked by the annotator as being outside the relspan represented with a ~~strike through~~.

### Example

**Query:** What happens if you get addicted to the internet?

**Citation:** It is believed that between 5 to 10% of internet users have become addicted to it in the sense that they are no longer able to control its use

**Does the English citation contain large amounts of extraneous material?**

YES, it contains a lot of extra junk.

NO, it does not contain a lot of extra junk.

It is incorrect to do any relspan editing on this citation:

**\*Incorrect:** ~~It is believed that between 5 to 10% of internet users have become addicted to it in the sense~~ that they are no longer able to control its use

### Example

**Query:** Is it safe to travel to Ethiopia?

**Citation:** Transport is expensive in Ethiopia, and you definitely need a four-wheel drive for safety and comfort, with a second vehicle for back up. On a recent visit we only got stuck twice but met others digging out for the seventh time...

**Does the English citation contain large amounts of extraneous material?**

YES, it contains a lot of extra junk.

NO, it does not contain a lot of extra junk.

It is incorrect to do any relspan editing on this citation:

**\*Incorrect:** Transport is expensive in Ethiopia, and you definitely need a four-wheel drive for safety and comfort, with a second vehicle for back up. ~~On a recent visit we only got stuck twice but met others digging out for the seventh time...~~

## Example

**Query:** Describe Russia's relations with the nation of Georgia.

**Citation:** Deputy Foreign Minister Sergei Ryabkov's remarks appeared to reflect alarm over the idea that Western-leaning neighbors such as Ukraine or Georgia, Russia's foe in a war last year could potentially host U.S. missile defense facilities.

**Does the English citation contain large amounts of extraneous material?**

- YES, it contains a lot of extra junk.  
 NO, it does not contain a lot of extra junk.

It is incorrect to do any relspan editing on this citation:

~~**\*Incorrect:** Deputy Foreign Minister Sergei Ryabkov's remarks appeared to reflect alarm over the idea that Western-leaning neighbors such as Ukraine or Georgia, Russia's foe in a war last year, could potentially host U.S. missile defense facilities.~~

## Example

**Query:** Describe Russia's relations with the nation of Georgia.

**Citation:** well i mean that georgia does not deserve such agression from Russia, becoz we don't have any terrorist bases or anti russian elements in our country!

**Does the English citation contain large amounts of extraneous material?**

- YES, it contains a lot of extra junk.  
 NO, it does not contain a lot of extra junk.

It is incorrect to do any relspan editing on this citation:

~~**\*Incorrect:** well i mean that georgia does not deserve such agression from russia, becoz we don't have any terrorist bases or anti russian elements in our country!~~

## Example

**Query:** Is abortion ethical?

**Citation:** While I detest abortion from a ethical standpoint, from a legal standpoint the most likely outcome is that abortion will remain legal in the United States but will be heavily restricted; late term abortions would be banned and things like parental

**Does the English citation contain large amounts of extraneous**

**material?**

- YES, it contains a lot of extra junk.  
 NO, it does not contain a lot of extra junk.

It is incorrect to do any relspan editing on this citation:

**\*Incorrect:** ~~While I detest abortion from a ethical standpoint, from a legal standpoint the most likely outcome is that abortion will remain legal in the United States but will be heavily restricted; late term abortions would be banned and things like parental~~

## 9.2 Examples that do contain large amounts of extraneous material

In the following examples the correct answer is “YES, it does contain a lot of extra junk”. These citations contain significant content that does not directly address the query and is completely irrelevant to the topic.

### Example

**Query:** What are the effects of outsourcing jobs?

**Citation:** How Outsourcing Creates Jobs for Americans | Publications | National Center for Policy Analysis | NCPA

**Does the English citation contain large amounts of extraneous material?**

- YES, it contains a lot of extra junk.  
 NO, it does not contain a lot of extra junk.

Because you answered YES, you should copy and paste the relevant span (relspan) of text into the relspan box in the annotation tool:

**RelSpan:** How Outsourcing Creates Jobs for Americans

### Example

**Query:** Why is climate change happening?

**Citation:** Changes in the Sun’s Surface to Bring Next Climate Change 2008 01 13 < /SPAN > From : spaceandscience.net Today , the Space and Science Research Center , ( SSRC ) in Orlando , Florida announces that it has confirmed the recent web announcement of .

**Does the English citation contain large amounts of extraneous material?**

- YES, it contains a lot of extra junk.  
 NO, it does not contain a lot of extra junk.

Because you answered YES, you should copy and paste the relevant span (relspan) of text into the relspan box in the annotation tool:



**RelSpan:** Changes in the Sun's Surface to Bring Next Climate Change

## **10. Conclusion**

To summarize, you will be assigned a kit containing a single query and multiple citations to judge. The assessment process involves answering a series of questions. The outcome for each question will determine what question(s) you see next. You will answer all questions for a citation before moving to the next citation.

On the next page is a checklist summarizing the annotation decisions you need to make for each citation. You may wish to print this out for easy reference.

## BOLT IR Assessment Procedure Checklist

**Step 1: Read the query, rules and citation carefully.**

**Step 2. Can you answer relevance questions based on this English citation alone, without looking at the source text?**

- YES (Go to “A. No Source View” Questions)
  - NO because the translation is incomprehensible. (Go to next citation)
  - NO because I need to see the source text to resolve pronouns, get more context and/or clarify the translation. (Go to “B. Source View” Questions)
- **If in doubt, choose “NO because I need to see the source text...”**

### A. No Source View Questions

**Step 3A. Does the English citation satisfy the rules of interpretation?**

- YES, it meets all the rules. (Go to Step 4A)
  - NO, it fails to meet one or more rules. (Go to next citation)
- **If in doubt, choose “YES, it meets all the rules”.**

**Step 4A. Is the English citation useful?**

- YES, it adds information beyond restating the query. (Go to Step 5A if source language is English, otherwise go to next citation)
  - NO, it does not add information. (Go to Step 5A if source language is English, otherwise go to next citation)
- **If in doubt, choose “YES, it adds information beyond restating the query”.**

**Step 5A. Does the English citation contain large amounts of extraneous material?**

- YES, it contains a lot of extra junk. (Fill in RelSpan box then go to next citation)
  - NO, it does not contain a lot of extra junk. (Go to next citation)
- **If in doubt, choose “NO, it does not contain a lot of extra junk”.**

### B. Source View Questions

**Step 3B. Does the red/nearby source text satisfy the rules of interpretation?**

- YES, it meets all the rules. (Go to Step 4B)
  - NO, it fails to meet one or more rules. (Go to next citation)
- **If in doubt, choose “YES, it meets all the rules”.**

**Step 4B. Is the red/nearby source text useful?**

- YES, it adds information beyond restating the query. (Go to Step 6B if source language is English, otherwise go to Step 5B)
  - NO, it does not add information. (Go to Step 6B if source language is English, otherwise go to Step 5B)
- **If in doubt, choose “YES, it adds information beyond restating the query”.**

**Step 5B. Re-read English citation. Does it contain any relevant information?**

- YES, it addresses the query. (Go to next citation)
  - NO, it does not address the query. (Go to next citation)
- **If in doubt, choose “YES, it addresses the query”.**

**Step 6B. Does the English citation contain large amounts of extraneous material?**

- YES, it contains a lot of extra junk. (Fill in RelSpan box then go to next citation)
  - NO, it does not contain a lot of extra junk. (Go to next citation)
- **If in doubt, choose “NO, it does not contain a lot of extra junk”.**

<b>Remember: Reasonable! Be Generous!</b>
---