Wikipedia Contradiction Annotation Guideline

Table of Contents

1. Setup the Label Studio environment 2

2. Task annotation 6

Step1: Check whether the inconsistent tag is valid. 6

Step2: find the inconsistent passages. 7

Step3: Annotate the contradiction reason. 8

Step4: Annotate the contradiction types. 9

Step5: Annotate the question and answers. 14

3. Export annotations 15

1. Setup the Label Studio environment

1) install Label Studio locally
   Github: https://github.com/HumanSignal/label-studio
   Here's the command I used to install it in my laptop (make sure python>=3.8):

```
conda create --name label-studio
conda activate label-studio
conda install psycopg2
pip install label-studio
```

   About LS version: I'm using Label Studio 1.8.2. Normally all new versions later than 1.8.2 should work, I recommend using google Chrome to do annotation.

2) run "label-studio" to start the server at http://localhost:8080
3) create an account and log in
4) create a new project called "WikipediaContradict" or any other names
5)

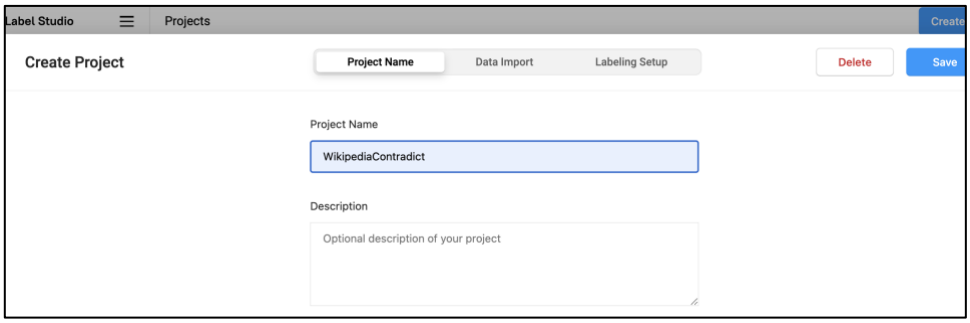

Fig 1: Create a new project

6) import the annotation tasks into the platform: first, click "upload files" and choose "inconsistentArticleTags_all.json"; next, click "save" to finish importing the 170 annotation tasks

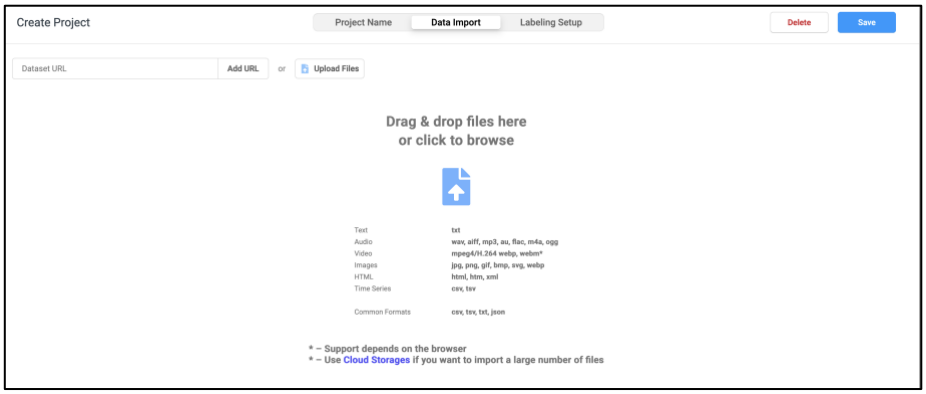

Fig 2: Import annotated data

7) setup the annotation UI:
   a) click "settings" on the upper right corner:

Fig 3: Setup the annotation UI

b) click “labelling interface”, copy-paste the following UI code into the box (or the code from the “WikipediaContradiction_LabelSstudioUI”), then click “save”. After this, you can start annotating by clicking “Label All Tasks” or any task in the panel.

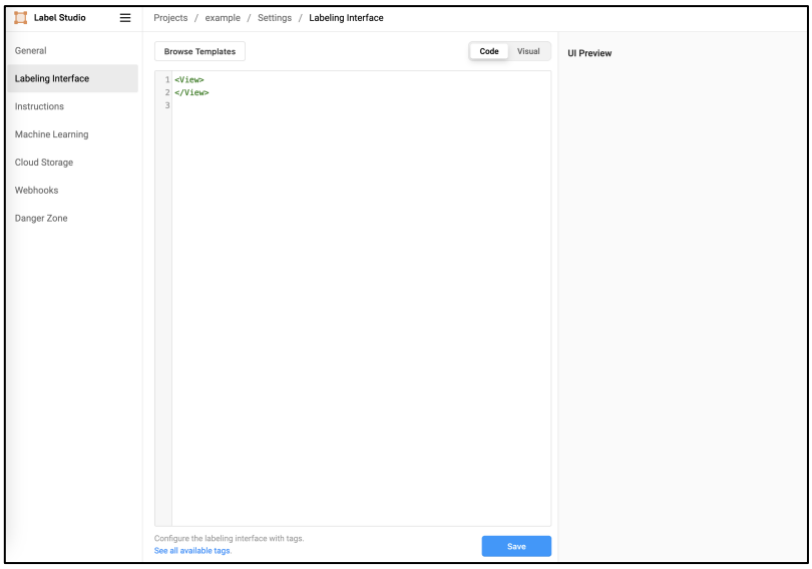

Fig 4: Setup the annotation UI - Continue

```
<View>
<!--  <header name="articletitle" value="Wikipedia article: $title"/> -->
  <HyperText clickableLinks="true" name="articlelink" inline="true" target="_blank" value="">
    <h2><a target="_blank" href="$url">Wikipedia article: $title</a></h2>
  </HyperText>

  <View style="box-shadow: 2px 2px 5px #999;          padding: 20px; margin-top: 2em;          border-radius: 5px;">
  <Header value="Inconsistence or contradictory tag"/>
  <Text name="wikitag" value="$paragraph_A"/>
  </View>

  <View style="box-shadow: 2px 2px 5px #999;          padding: 20px; margin-top: 2em;          border-radius: 5px;">
  <Header value="* Step1: Is this tag valid?"/>
  <Choices name="wikitag_label_valid" toName="wikitag" choice="single" showInLine="true">
     <Choice value="Valid"/>
     <Choice value="Invalid"/>
  </Choices>
 <Header value="Additional comment" />
    <TextArea name="valid_comment" toName="wikitag"
          showSubmitButton="true" maxSubmissions="1" editable="true"
          required="false" />

</View>

<View visibleWhen="choice-selected" whenTagName="wikitag_label_valid" whenChoiceValue="Valid" style="box-shadow: 2px
2px 5px #999;          padding: 20px; margin-top: 2em;  border-radius: 5px; ">
<Header value="* Step 2: Copy the inconsistent sentences, paragraphs, or information (e.g., table or infobox rows) from the
wikipedia article(s) into the following boxes"/>
<Header value="* Article A title" />
<TextArea name="paragraphA_article" toName="wikitag"
          showSubmitButton="true" maxSubmissions="1" editable="true"
```

```
        required="true" />

<Header value="* Article A relevant information (Passage 1)" />
   <TextArea name="paragraphA_information" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="true" />
<Header value="Article A relevant information_stand-alone (Passage 1)" />
   <TextArea name="paragraphA_information_standalone" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="false" />


<Header value="* Article B title" />

  <TextArea name="paragraphB_article" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="true" />

<Header value="* Article B relevant information (Passage 2)" />
   <TextArea name="paragraphB_information" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="true" />
<Header value="Article B relevant information_stand-alone (Passage 2)" />
   <TextArea name="paragraphB_information_standalone" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="false" />

</View>

<View visibleWhen="choice-selected" whenTagName="wikitag_label_valid" whenChoiceValue="Valid" style="box-shadow: 2px
2px 5px #999;              padding: 20px; margin-top: 2em;  border-radius: 5px; ">
<Header value="* Step 3: Contradiction reason" />
<Header size = "8">* Passage 1 states that:  </Header> <TextArea name="relevantInfo_comment_A" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="true" />
<Header size = "8">* However, passage 2 states that:  </Header>    <TextArea name="relevantInfo_comment_B"
toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="true" />
<Header size = "8">If possible, copy the contradicted span from passage 1:  </Header> <TextArea
name="relevantInfo_comment_A_Span" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="false" />

<Header size = "8">If possible, copy the contradicted span from passage 2:  </Header>    <TextArea
name="relevantInfo_comment_B_Span" toName="wikitag"
        showSubmitButton="true" maxSubmissions="1" editable="true"
        required="false" />

</View>

<style>
  .center-text {
    text-align: center;
  }
</style>


<View visibleWhen="choice-selected" whenTagName="wikitag_label_valid" whenChoiceValue="Valid" style="box-shadow: 2px
2px 5px #999;              padding: 20px; margin-top: 2em;  border-radius: 5px; ">
<Header value="* Step 4: Choose all options that can describe the above contradictory information"/>
  <Taxonomy name="taxonomy" toName="wikitag">
    <Choice value="Contradict type I">
      <Choice value="(PhraseLevel) Entity - Date/time" />
      <Choice value="(PhraseLevel) Entity - Location/GPE (Non-GPE locations, mountain ranges, bodies of water, and Countries,
cities, states)" />
      <Choice value="(PhraseLevel) Entity - Number" />
      <Choice value="(PhraseLevel) Entity - Organization (Companies, agencies, institutions, etc.)" />
      <Choice value="(PhraseLevel) Entity - Person" />
      <Choice value="(PhraseLevel) Entity - NORP (Nationalities or religious or political groups)" />
      <Choice value="(PhraseLevel) Entity - FAC (Buildings, airports, highways, bridges, etc.)" />
      <Choice value="(PhraseLevel) Entity - Work-of-Art (Titles of books, songs, etc.)" />
      <Choice value="(PhraseLevel) Entity - Product (Titles of books, songs, etc.)" />
      <Choice value="(PhraseLevel) Entity - Law (Named documents made into laws)" />
      <Choice value="(PhraseLevel) Entity - Language (Any named language)" />
      <Choice value="(PhraseLevel) Entity - Event (Named hurricanes, battles, wars, sports events, etc.)" />
      <Choice value="(PhraseLevel) Entity - Other" />
      <Choice value="(PhraseLevel) NP-related (non-entity)" />
      <Choice value="(PhraseLevel) Event/Relation (e.g., verb)" />
      <Choice value="(DiscourseLevel) NP-related " />
```

```
    <Choice value="(DiscourseLevel) Event/Relation-related " />
   </Choice>
   <Choice value="Contradict type II">
    <Choice value="Text - Text" />
    <Choice value="Text - Infobox/table" />
    <Choice value="Infobox/table - Infobox/table" />
    <Choice value="Other" />
   </Choice>
   <Choice value="Contradict type III">
    <Choice value="Within the same article" />
    <Choice value="Across different articles" />
   </Choice>
   <Choice value="Contradict type IV">
    <Choice value="Explicit" />
    <Choice value="Implicit (reasoning required)" />
   </Choice>
  </Taxonomy>
<Header value="Additional comment" />
   <TextArea name="contradict_comment" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="false" />

</View>

<View visibleWhen="choice-selected" whenTagName="wikitag_label_valid" whenChoiceValue="Valid" style="box-shadow: 2px
2px 5px #999;              padding: 20px; margin-top: 2em;  border-radius: 5px; ">
<Header value="Step 5: Write at least one question that highlights the contradictions between Passage 1 and Passage 2, eliciting
different responses based on each passage."/>
<Header value="* Question 1" />
   <TextArea name="question1" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="true" />
<Header value="* Answers for Question 1" />
   <TextArea name="question1_answer1" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="true" />
   <TextArea name="question1_answer2" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="true" />

<Header value="Question 2" />
   <TextArea name="question2" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="false" />
<Header value="Answers for Question 2" />
   <TextArea name="question2_answer1" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="false" />
   <TextArea name="question2_answer2" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="false" />


<Header value="Additional comment" />
   <TextArea name="qa_comment" toName="wikitag"
         showSubmitButton="true" maxSubmissions="1" editable="true"
         required="false" />

  </View>



</View>
```

Another way is to choose the assigned tasks (e.g., 31 - 40) and click “Label 10 Tasks”, this will open the annotation window for these 10 tasks.

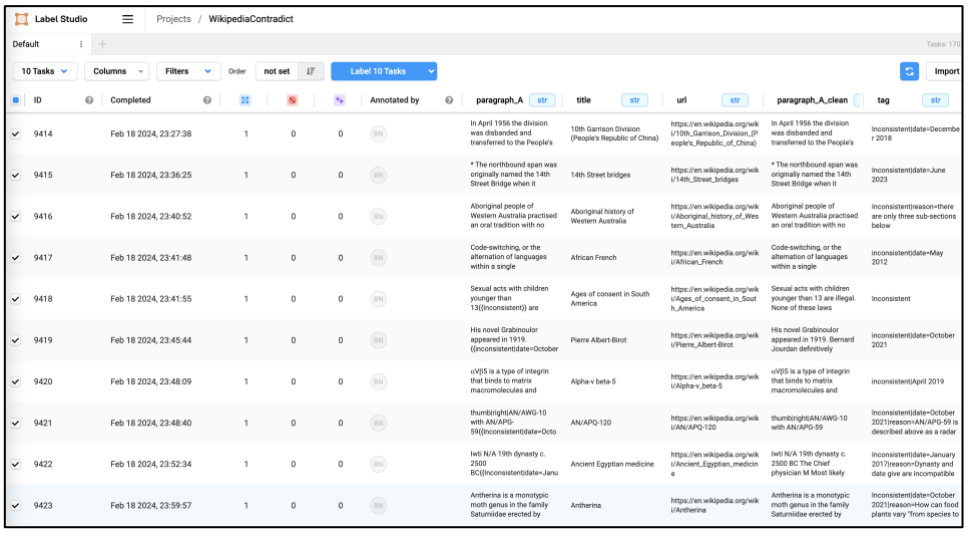

Fig 5: Choosing the annotation tasks

2. Task annotation

Step1: Check whether the inconsistent tag is valid.

1) Open the Wikipedia article by clicking the corresponding link (Fig 6), identify the paragraph tagged with the inconsistent tag by searching “inconsistent” (Fig 7).

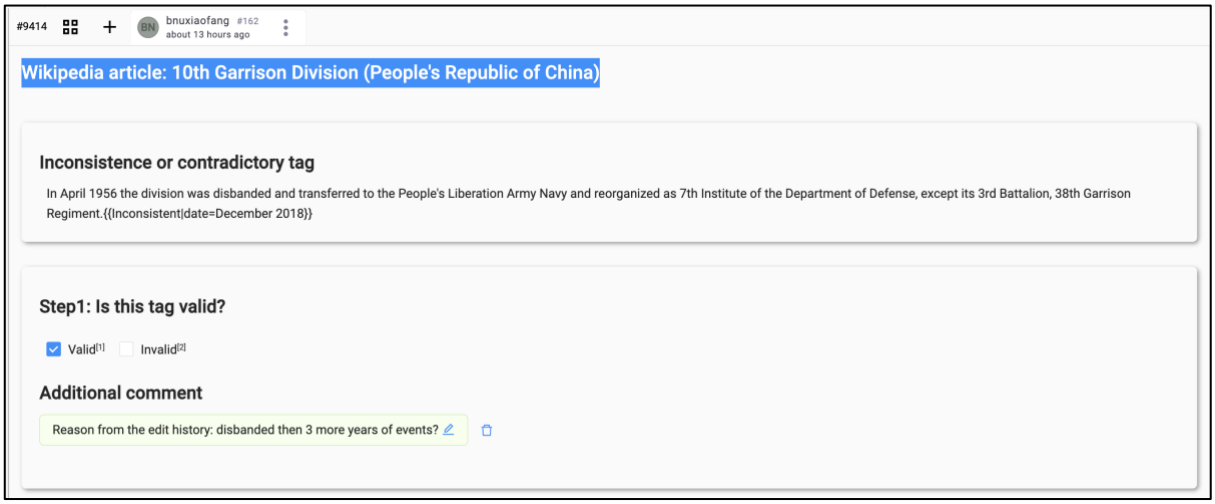

Fig 6: Annotation – step 1

Fig 7: Read the Wikipedia article

2) Read the Wikipedia article to check whether the tag makes sense. In the above example, the Wikipedia editor who added this tag didn’t specify the reason as an attribute of the tag. Although this is the recommended template, they simply didn’t follow the rules, so we need to investigate further. We know that this tag was added in December 2018, so we can check

the editing history of this article to see under what condition this tag was added. In the revision history (see the figure below) we see that the editor put the reason in the edit comment “disbanded then 3 more years of events?” (last line in Fig 8). After checking this reason, if we agree that this inconsistent tag is valid, we go back to label studio and choose “valid” and put the reason in the additional comment box, as shown in Fig 1.

Fig 8: Checking the revision history of a Wikipedia article

Step2: find the inconsistent passages.

Copy the inconsistent sentences, paragraphs, or information (e.g., table or infobox rows) from the Wikipedia article(s) into the following boxes:

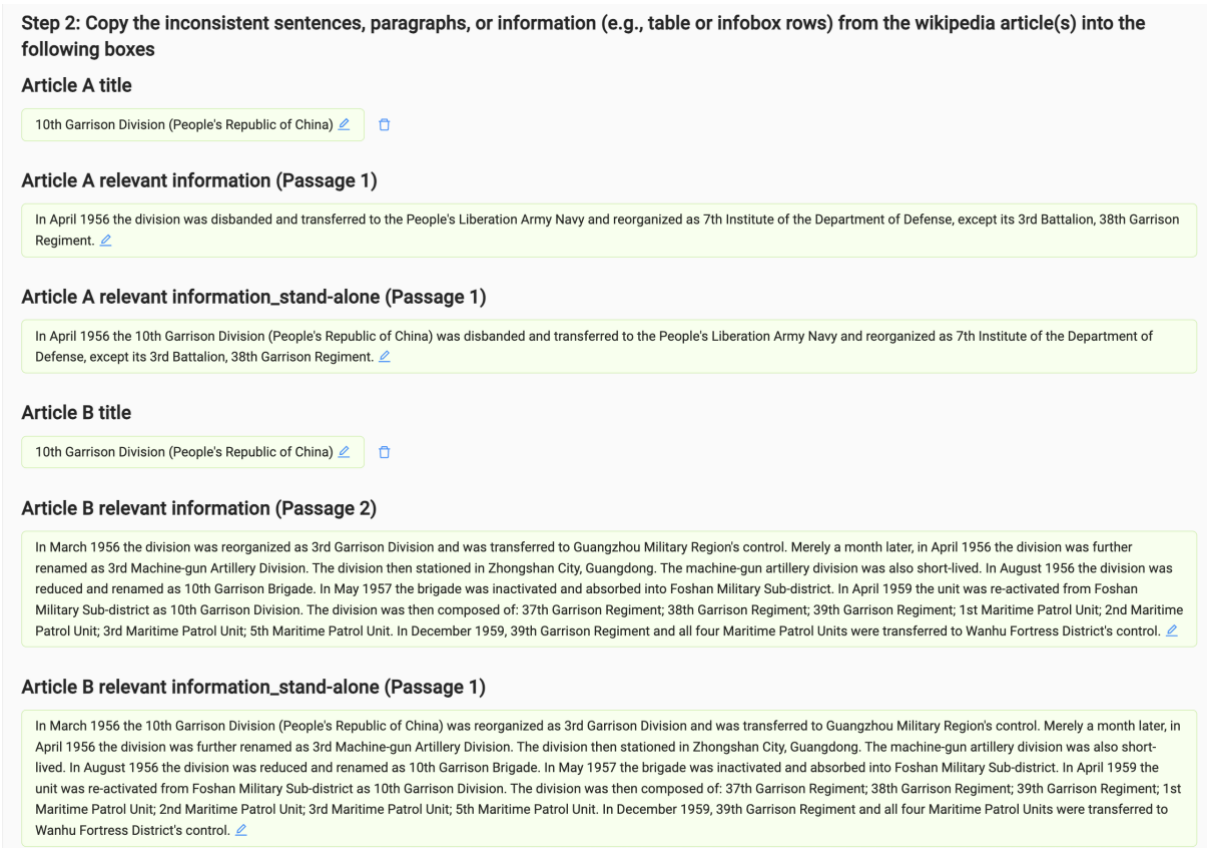

Fig 9: Annotation – step 2

For this annotation task, often "Article A title" and "Article B title" are the same. Note that "Article A relevant information (Passage 1)" and "Article B relevant information (Passage 2)" contain the original passage information from the Wikipedia, which means that you should copy-paste the original information into these boxes without modifying them. When copying the original passages into these boxes, please remove the citation marks and the inconsistent tags ({{inconsistent …}}). For "Article A relevant information_stand-alone (Passage 1)" and "Article B relevant information_stand-alone (Passage 2)", you are required to slightly modify the original passages to make them stand-alone (decontextualization). Normally, this requires you to resolve the coreference anaphors or the bridging anaphors in the first sentence. In Wikipedia, oftentimes the antecedents for these anaphors are the article titles themselves.

Example of resolving coreference anaphors:
In the example shown in Fig 9, we replace "the division" in the first sentence of both passages as "the 10th Garrison Division (People's Republic of China)"

Below is another example of resolving coreference anaphors:

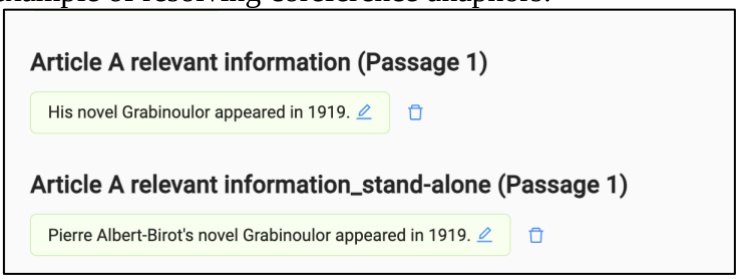

Fig 10: resolving coreference anaphors

Example of resolving bridging anaphors (one of my favorite topics ☺):
In the following example as shown in Fig 11, we replace "The larvae" in the first sentence as "The larvae of Antherina"

Fig 11: resolving bridging anaphors

Step3: Annotate the contradiction reason.

In this step, we use the template "passage 1 states that …, however, passage 2 states that …" to annotate the contradiction reason (see Fig 12). If possible, please copy-paste the exact contradicted spans (short phrases within a sentence) from both passages, such as "was disbanded" from passage 1 as shown in Fig 12. In this example, we leave the contradicted span from passage 2 empty since it involves a series of events across multiple sentences. Fig 13 shows an example in which we can easily identify the contradicted spans in both passages.

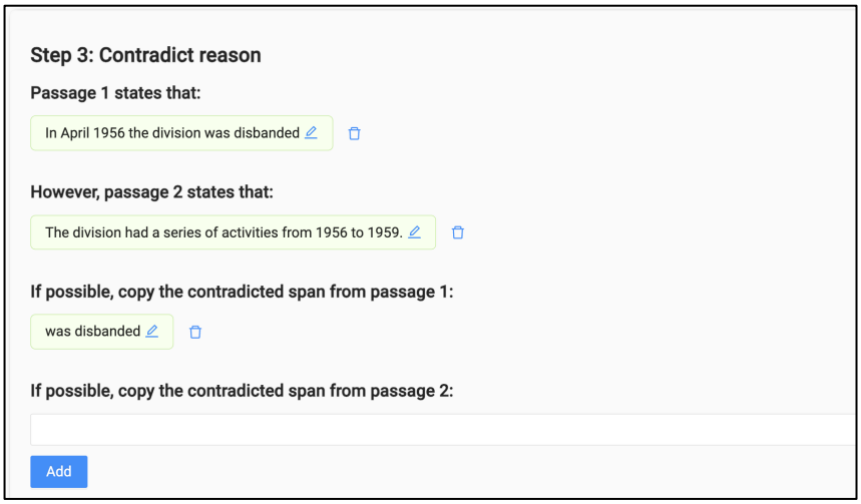

Fig 12: Annotate the contradiction reason

Fig 13: Annotate the contradiction reason – another example

Step4: Annotate the contradiction types.

In this step, we assign the identified contradictory information to the appropriate types according to a pre-defined taxonomy. Back to our first example as shown in Fig 6/9/12, we assign it to the following contradiction types as shown in Fig 14

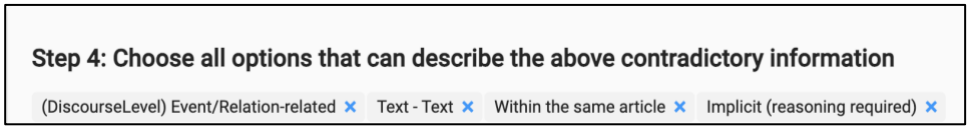

Fig 14: Annotate the contradiction types

Below we provide more details four the pre-defined four contradiction types.

1) Contradiction type I: As shown in Fig 15, contradiction type I focuses on the fine-grained semantics of the contradiction.

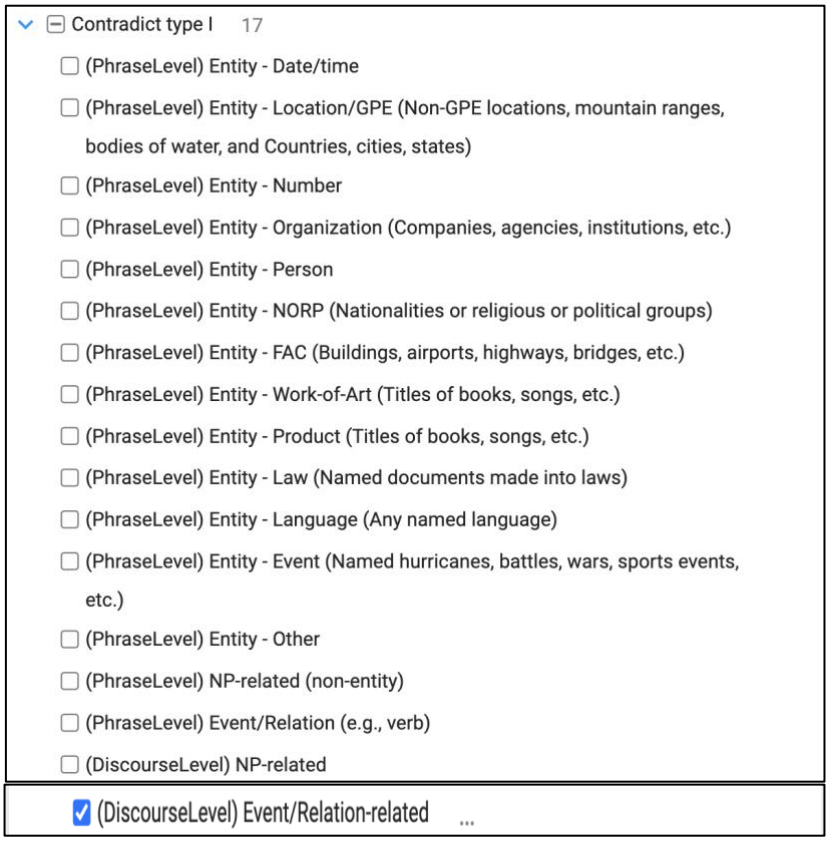

Fig 15: Contradiction type I

<u>Contradiction Type 1 – Phrase level – Entity</u>: The contradictory information is around two named entities in passage 1 and passage 2. Normally phrase level contradiction can be easily fixed by changing one of the named entities (if we know which one is factually correct). We adapt OntoNotes named entity type definitions to describe the different types of contradicted named entities. For each named entity type, its and the explanation included in the parenthesis should provide a clear definition of the corresponding named entity type. The following example shown in Fig 16 was assigned to "(phraseLevel) – Entity-Date/time".

Fig 16: Contradiction type example

Contradiction Type 1 – Phrase level – Non-entity NP: The contradictory information is around two noun phrases that are not named entities in passage 1 and passage 2. In the following example as shown in Fig 17, the contradicted information are around two common nouns: monotypic (passage 1) and species to species (passage 2).

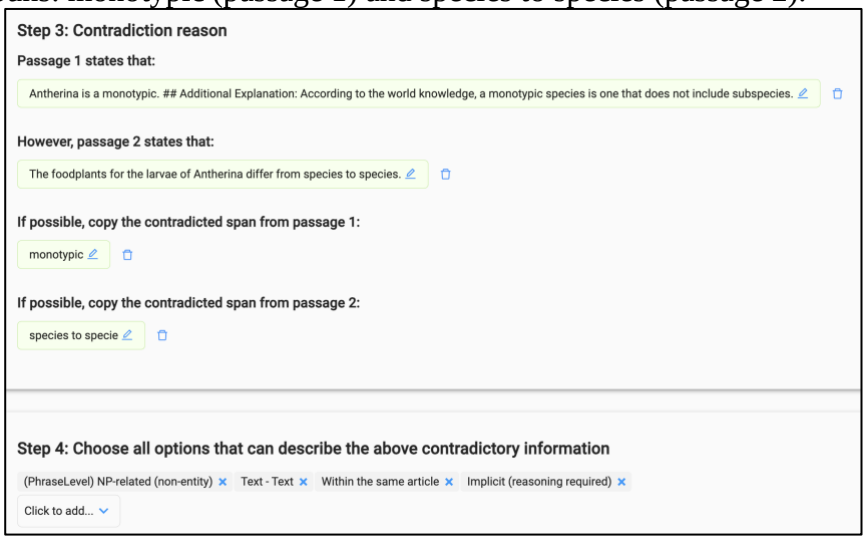

Fig 17: Contradiction type example

Contradiction Type 1 – Phrase level – Event/relation: The contradictory information is around two verb phrases that are describe two contradicted events or relations in passage 1 and passage 2. In the following example as shown in Fig 18, the contradicted information are around two verbs: "stimulates" (passage 1) and "inhibits" (passage 2).

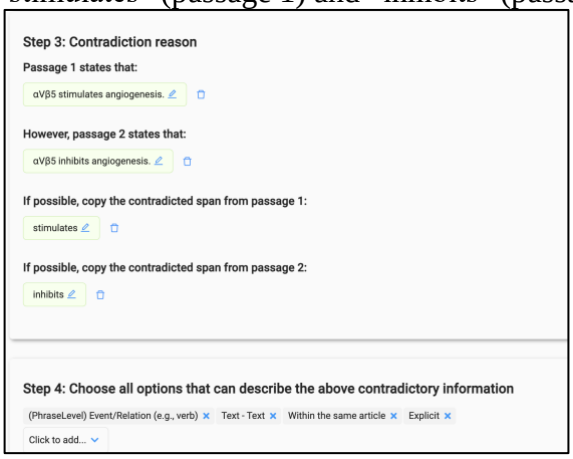

Fig 18: Contradiction type example

Contradiction Type 1 – Discourse level – NP-related: The contradictory information is beyond the single sentence level from passage 1 and passage 2. The contradicted information from passage 1 can be anchored to a NP span, and the contradicted information from passage 2 is across multiple sentences. In the following example as shown in Fig 19, the contradicted information from passage 1 can be anchored to an NP: "five periods of time", and the contradicted information from passage 2 contains a few paragraphs that contains three sub-section headers indicating three time periods.

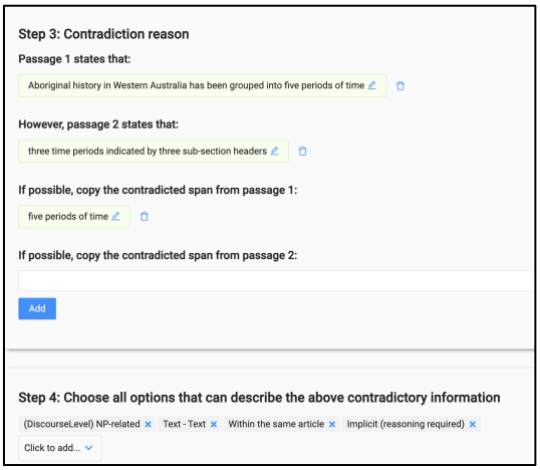

Fig 19: Contradiction type example

Contradiction Type 1 – Discourse level – Event/relation-related: The contradictory information is beyond the single sentence level from passage 1 and passage 2. The contradicted information from passage 1 can be anchored to a verb phrase, and the contradicted information from passage 2 is across multiple sentences. In the following example as shown in Fig 20, the contradicted information from passage 1 can be anchored to a VP: "was disbanded", and the contradicted information from passage 2 contains a few paragraphs that describes a series of events.

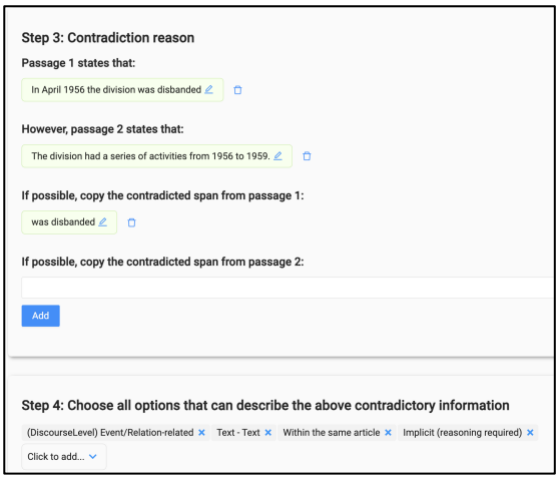

Fig 20: Contradiction type example

2) Contradiction type II: As shown in Fig 21, contradiction type II focuses on the modality the contradiction. It describes the modality of passage 1 and passage 2, whether the information is from a piece of text, or from a row an infobox or a table. Fig 22 shows an example of "contradict type II – infobox/table – Infobox/table".

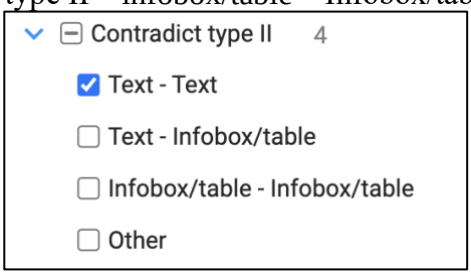

Fig 21: Contradiction type II

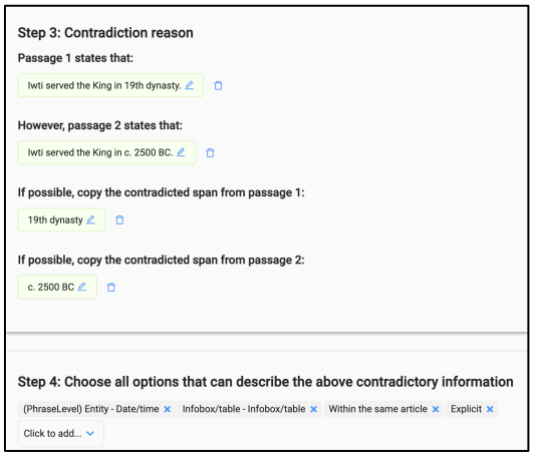

Fig 22: Contradiction type example

3) Contradiction type III: As shown in Fig 23, contradiction type III focuses on the source the contradiction. It describes whether passage 1 and passage 2 are from the same article or not. For the inconsistent tags, most of contradictions are from the same article. In a few rare cases, the contradiction is from different articles. Fig 24 illustrates such an example: passage 1 is from the English version of the Wikipedia article "Pierre Albert-Birot" and passage 2 is from the French version of the article.

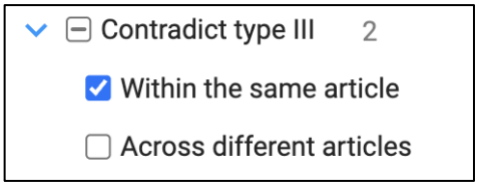

Fig 23: Contradiction type III

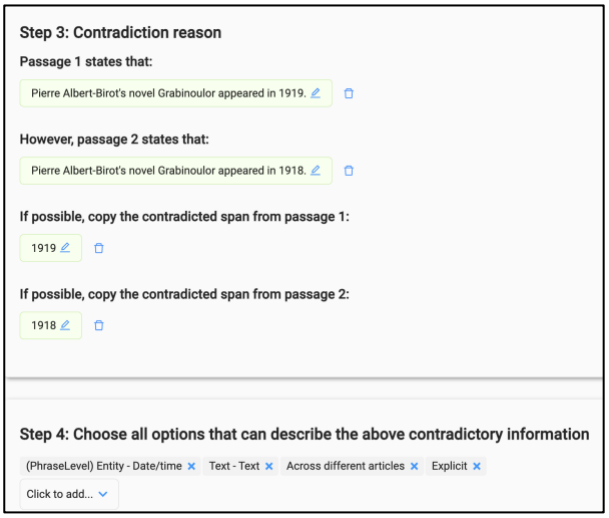

Fig 24: Contradiction type example

4) Contradiction type IV: As shown in Fig 25, contradiction type IV focuses on the reasoning aspect. It describes whether the contraction is explicit or implicit. Implicit contradiction requires us do to some reasoning to understand why passage 1 and passage 2 are contradicted. In the example shown in Fig 17, to understand that "monotypic (passage 1)" and "species to species (passage 2)" are contradicted, we need to carry out additional reasoning steps, i.e., first we need to know that according to commonsense knowledge, a

monotypic species is one that does not include subspecies; then “species to species” entails that there are more than one specie. Therefore, we can conclude that passage 1 is contradicted with passage 2.

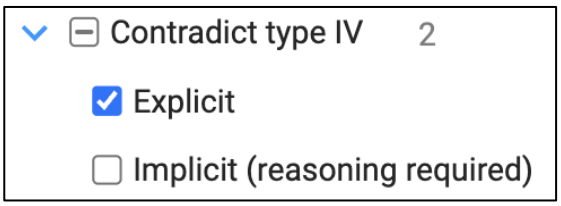

Fig 25: Contradiction type IV

Note: If the contradictory information exhibits other important attributes that are not covered by the existing taxonomy, using the additional comment box to describe it. Figure 26 shows such an example: passage 1 is from the English Wikipedia article, and passage 2 is from the French Wikipedia article, therefore in the additional comment for Step 4 we put “across different languages”

Fig 26: Additional contradiction Type

Step5: Annotate the question and answers.

In this step, we formulate at least one question that highlights the contradictions between Passage 1 and Passage 2, eliciting different responses based on each passage. Fig 27 shows a question and answers example about the contradicted information in Fig 16.

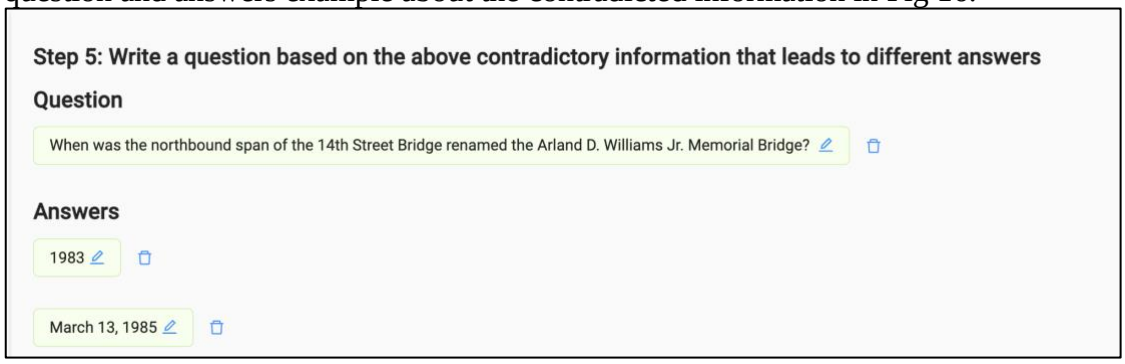

Fig 27: write a question and different answers

3. Export annotations

After finishing the assigned annotation tasks, go back to the task pool panel, click “Export” and choose “JSON” as the format (Fig 28). Please rename the exported file using the following template: wikipediaContradict_10_<your name>.json

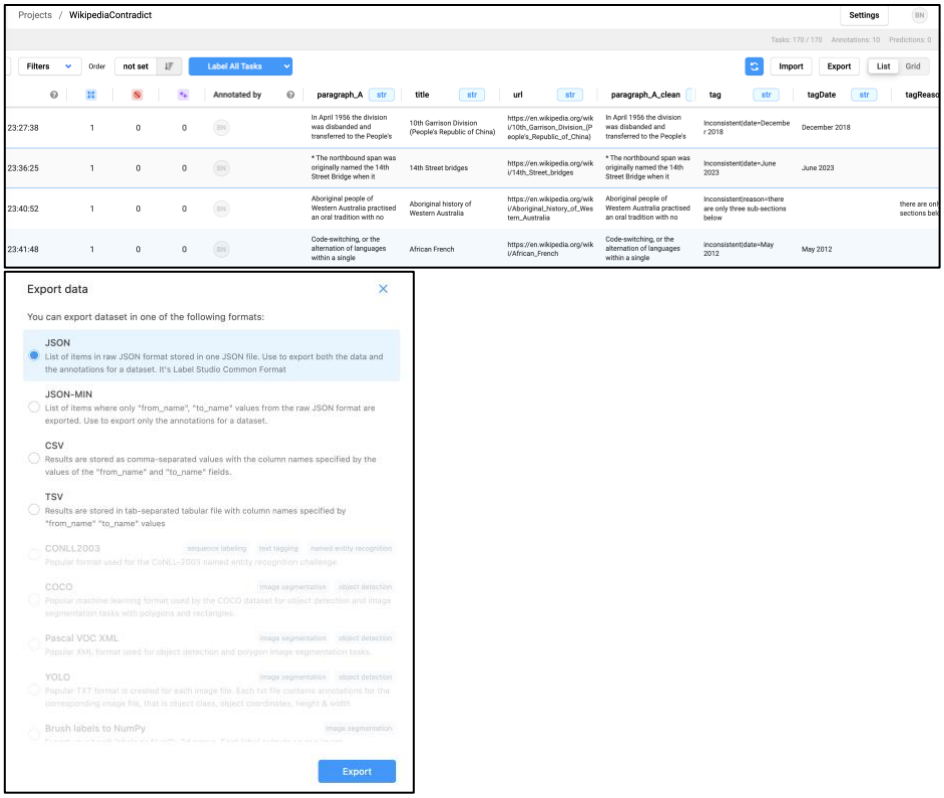

Fig 28: Export annotations