

Annotation Guide

Ink Action Item Definition

Introduction

The goal of this section is to define what we mean by ink action items in this project. This will help disambiguate how we are choosing to label whether a sentence is actionable or not.

In general, an action item is a single task that can be assigned to an individual or small team to accomplish. In this project, we are particularly interested in the subset of action items that we expect to be written on a whiteboard, in a collaborative setting such as drafting notes in a meeting.

While there may be ambiguity in some samples, the guiding principle for our scenario is that *an action item is a discrete task that can be assigned to an individual and that is intended to be completed after it was drafted.*

There are a few characteristics that we expect are common in action items drafted on a whiteboard:

- Action items may appear as bulleted items as part of a ToDo list. Example:
 - Send email to Amy by EOD.
- Inked action items may not be a perfect sentence with polished grammar, spelling or punctuation.
 - will send out next steps document tmrw morning
 - pls schedule the team sync asap
- Regardless of phrasing style, they can naturally be assigned to an owner. Though, the method of assignment may differ. Examples:
 - John: Create document to define action items.
 - run experiment with latest dataset. @Ryan
 - [Sam] Can you create the repo by Monday?
 - Jenny – Gather feedback on design.
 - We should transplant model to C++ codebase -> John Doe

It is important that we do not simply rely on the writing style to identify an action item. People may write an action item with the same task content in various styles:

- Alex to schedule a meeting with the PAWS team on Monday.
- alex: will you please schedule a meeting with paws on Monday?
- PAWS expects a call from us on Monday. @Alex
- @Alex - We should schedule a call with the PAWS folks on Monday.
- Alex, please schedule some time with the PAWS team on Monday.

In the sections below, we provide a non-exhaustive list of cases to label as tasks/non-tasks based.

Examples of Sentences to Label as Action Items

- ToDo List Style Item:
 - Create experimentation repo.
 - Discuss next steps with Rebecca and her team.
 - John to send out secret sauce recipe.
- Request or delegation:
 - Can you please prepare the slides for the engineering design presentation on Thursday?
 - Jack, please send the project documentation to my email.
 - Could you reach out to Mark to see if he is on board with the plan?
- Commitment:

- I will prepare the engineering presentation slides for Thursday.
- Jack will send out the project documentation by email.
- Will check with Mark to see if he is on board with the plan.
- Something to be done:
 - We need the engineering design presentation slides ready for Thursday.
 - [Jack] Project documentation should be sent out by email.
 - Let's make sure Mark is on board with the plan.
- Something to be followed-up on:
 - Follow-up with Ryan to understand the engineering design.
- Incomplete task sentences with owners clearly assigned:
 - Engineering design presentation slides @Rebecca
 - Alex – Exploratory data analysis

Examples of Sentences to *NOT* Label as Action Items

- Sentences that are clearly not tasks.
 - I am feeling hungry today.
- Actions to be completed instantaneously, rather than in the future.
 - Calculate 1 + 1.
- Sentences which represent inaction:
 - Mike will wait for the experiment.
 - Wait for manager approval.
- Negated actions:
 - Mike will not run the experiment.
- Tasks completed in the past:
 - Mike ran the experiment.
- Conditional tasks that may or may not need to be fulfilled:
 - If we win the contract, Mike will run the experiment.
- Past-for-future:
 - Mike said he would run the experiment.

Special Cases

- Generic Action Items: Sentences that could be action items but lack any details of the task since they are too generic. Examples:
 - I will do this.
 - Please do it by next week.
 - Could you rush this task?

We would like to label these separately from the case the classes we discussed above (action item/not an action item). We will create a third category for these generic items when labelling.

- Action Items due to Context: Sentences or phrases that would not be action items on their own but are action items due to context. Example:

To do list:

- *Exploratory Data Analysis*
- *Model Integration*
- *Customer feedback*

In the above example, none of the bulleted items are action items on their own. However, the fact that they are part of a To Do list makes them action items. We would like to label such items (those that are action items only due to context) separately from our binary categorization (action item/not an action item) and will create a special label for these items.

Due to errors propagating through the HWR we have spelling errors and also grammatical errors which could change the meaning of a sentence but we still want to surface such cases if we know for instance if they are part of a todo list.

To do list:

- *Meet my friend in the evening*
- *Buy notebooks for class*
- *See I up the team roster*

Here "*Setup my team roster*" is wrongly parsed as "*See I up my team roster*" but we still want to surface such text and let the user correct the errors in the wrongly parsed text as we know this is part of a todo list

- Non-action item due to Context: Sentences which seem actionable but are not due to the context they occur in. Examples:

Butternut Squash Soup:

- *Cut the carrots*
- *Boil water*
- *Grind the vegetables*

All the above seem like action items but are actually instructions for a recipe regarding how to make butternut squash soup and hence are not action items due to context. Another common such scenario is driving instructions.

Dataset Annotation Guide

Introduction

In this section we will go over the instructions to manually annotate the data. It will be helpful to read the Ink Action Item Definition section of this document before starting the labeling process. The labeling process is as follows:

1. You will be assigned some rows in a csv file to annotate.
2. In the csv file you only need to annotate the column named "annotate"
3. You may find some of the other columns in the csv helpful in making decisions while annotating.

In the cells of the "annotate" column of the csv, you will be given some whitespace delimited text corresponding to a writing region in a whiteboard. Your task is to identify the beginning and end of each sentence in the text and also the type of sentence it is according to categories defined below. For each sentence you identify, use the tags defined below to demarcate the bounds of the sentence based on its category.

Tags for Each Category

1. Not an action item: <s>sentence<a0>
2. Action item: <s>sentence<a1>
3. Generic Action item: <s>sentence<a2>
4. Action item due to context: <s>sentence<a3>

5. Not an action item due to context: <s>sentence<a4>

Please refer to the Ink Action Item Definition section of this document for the definition of each of these categories.

Example Annotation

Below, we have an example of what the before and after annotation looks like for one of the cells in the csv. Please use this as a reference for the syntax of how to annotate. Also, while most of the sentences are clearly action items in this example, note the use of the <a3> tag for the “action item due to context” category.

Before Annotation:

To Do list Meet w/ Ruth cut the grass read the case studies finish the reading project set up a meeting w/ Alley make copies for the seminar contact the leaders make an essay list contact Meg about the fundraising call Chef John about the banquet call Pele and set up meeting meet w/ Pete see I up the team roster update the website w/ new pics email the class about the assignment email Liz about the club meetings call the phone tree to make sure everyone shows up on Sunday have coffee with Todd

After Annotation:

<s>To Do list<a0> <s>Meet w/ Ruth<a1> <s>cut the grass<a1> <s>read the case studies<a1> <s>finish the reading project<a1> <s>set up a meeting w/ Alley<a1> <s>make copies for the seminar<a1> <s>contact the leaders<a1> <s>make an essay list<a1> <s>contact Meg about the fundraising<a1> <s>call Chef John about the banquet<a1> <s>call Pele and set up meeting meet w/ Pete<a1> <s>see I up the team roster<a3> <s>update the website w/ new pics<a1> <s>email the class about the assignment<a1> <s>email Liz about the club meetings<a1> <s>call the phone tree to make sure everyone shows up on Sunday<a1> <s>have coffee with Todd<a1>

Annotation CSV Schema

Below are the descriptions of each of the columns in the annotation csv.

- `annotate`: The cells to annotate.
- `wr_lines`: The list of document lines for this writing region, detected by the IA engine.
- `line_list_item_mask`: Binary list where a 1 represents that the corresponding line in `wr_lines` is a list item.
- `wr_sentence`: List of sentences for this writing region obtained by passing the writing region text through a sentence segmentation model. Note that this segmentation is often inaccurate, and you should only use this as a reference rather than thinking of it as ground truth.
- `wh_id`: Unique whiteboard id.
- `wr_id`: id of the writing region within the whiteboard.