



Good Instructions Checklist for Data Enrichment Projects

The choices made by AI practitioners have an important influence on the working conditions of data enrichment professionals who perform tasks such as data annotation, data cleaning, and human review of algorithmic outputs.

This checklist lists what should be included in a set of task instructions to make sure they are as clear as possible for data enrichment workers.

To learn more about how this resource was developed, visit PAI's [responsible data enrichment sourcing library](#).

Data enrichment task instructions should always include:

Step-by-step instructions and/or training materials for how to successfully complete a task with the given application interface, which may include:

- Screenshots as needed
- Recorded videos (screenshare) of someone completing tasks correctly with voiceover as needed
- For data labeling tasks, list the range of labels you want the worker to consider while completing tasks for your project

Instructions for how to use the application interface for any additional activities related to the task

- Example: If there is a reference or guide built into the application that may or may not be needed for every task but can support workers in completing the task

Clearly specified criteria for task acceptance/rejection

A mechanism for workers to get in touch if instructions are unclear, they are experiencing a problem with the task, or need to contest a rejected task

Context of how this data will be used

For any project where tasks submitted by data enrichment workers have the potential of being rejected, include:

Example of well completed tasks, both common and edge cases

- Screenshots and/or Videos

Examples of common mistakes that might be made and why they are wrong

- Screenshots and/or Videos

Sandbox mode

- An environment which allows workers to practice completing a few tasks where they are shown the answer key upon completing each task (*Note: Workers should be paid for time spent in the 'sandbox'*)

Depending on the task, you may also need:

Training Module (and/or Videos) covering any specialized background knowledge needed for the task

Links to reference materials needed that cover the necessary domain knowledge (but please avoid implicitly requesting workers to invest unpaid time into reviewing background and training materials)