

Alegion How to Get High-Quality Training Data for Computer Vision from Your Data Labeling Platform

# How to Get High-Quality Training Data for Computer Vision from Your Data Labeling Platform?

## Introduction

There is nothing more important to the success of your machine learning initiatives than acquiring contextually labeled, high quality training data.

You need expertise to arrive at quality training data, in addition to the right tools and technology. Developing the capacity to annotate massive volumes of data while maintaining quality is a function of the model development lifecycle that enterprises often underestimate. It's resource intensive and requires specialized expertise. Your data science team needs partners and platforms it can trust to deliver the data quality you need.

At Alegion, we've established quality management best practices based on our experience labeling tens of millions of images, video, text, and audio records. In this guide, we explain the fundamental steps to quality data and how the Alegion team partners with you at each stage to ensure the quality you need.



## The Four Steps to Quality Data

Training data quality is an evaluation of a dataset's fitness to serve its purpose in a given ML use case. There is no one definition of quality; "quality data" is completely contingent on your specific project. The path to quality data can be broken down into four prioritized phases.

## Set Quality Requirements and Annotation Criteria

## **Annotation Criteria**

Before you can measure quality, you need to establish an unambiguous set of rules that describe what "quality" means in the context of your project. Annotation criteria are the collection of rules that define which objects to annotate, how to annotate them correctly, and what your quality targets are.

#### Accuracy/Quality Targets

Accuracy or quality targets define the lowest acceptable result for evaluation metrics like accuracy, recall, precision, F1 score, etc. Typically, your team will have quality targets for how accurately objects of interest were classified, how accurately objects were localized, and how accurately relationships between objects were identified.

#### **OBJECT CLASSIFICATION**

Object classification criteria describe how to apply labels to qualified objects, using predefined taxonomies.

#### **OBJECT LOCALIZATION**

Localization rules define how objects should be identified within the frame by bounding boxes or keypoints.

#### **OBJECT RELATIONSHIP**

Object relationship criteria describe the connections between objects in the frame that are of interest and how they should be annotated.

# The Four Steps to Quality Data – Setting Requirements

## **Quality Evaluation Metrics**

#### Individual or Difference Metrics

In a large, annotated dataset, there are thousands and thousands of specific annotations. Each of these annotations must be compared to and measured against ground truth data (data that is known to be labeled correctly) in order to evaluate the quality of the dataset as a whole. The difference between the ground truth annotation and the worker generated annotation is the basis for all other quality metrics.

#### **Batch Metrics**

Batch metrics are the metrics used to evaluate a labeled dataset as a whole, and they rely on counts of the individual metrics. The most common ones are accuracy, recall, precision, and F1 Score. Note that these are the same evaluation metrics typically used to evaluate model performance, which makes sense given that the goal of the training data is to represent the answers the models need to predict.

Some examples of metrics used to evaluate the quality of a specific annotation:



Intersection Over Union (IOU)

Pixel Tolerance

Object Keypoint Similarity (OKS)



Root Mean Squared Error



# The Four Steps to Quality Data – Setting Requirements

Accuracy is a measure of correctness —out of all of the objects in the dataset, both labeled and unlabeled, did we annotate the objects we were supposed to and not annotate the objects that didn't fit the criteria?

**Precision** is a measure of exactness—did we annotate only the objects we were supposed to? Did we capture only the objects that satisfy the annotation criteria, the true positives (TP), or did we incorrectly capture some false positives (FP), objects that did not satisfy our annotation criteria or weren't desired at all?

**Recall** is a measure of completeness—did we annotate all of the objects that met the annotation criteria, or have we failed to capture some of the qualified objects in frame?

**F1 Score** measures the balance between recall and precision in a dataset—how well are we capturing all AND only the objects we are supposed to annotate? We define F1 Score in terms of recall and precision by calculating their harmonic mean.







# Four Steps to Quality Data - Managing Workforce & Platform

Manage Workforce Training and Platform Configuration



## **Platform Configuration**

Task design and workflow setup require time and expertise, and accurate annotation requires task-specific tools. At this stage, you need a partner with expertise to help you determine how best to configure your labeling tools, classification taxonomies, and annotation interfaces for accuracy and throughput.



## Worker Testing and Scoring

To accurately label your data, annotators need a welldesigned training curriculum so they fully understand your annotation criteria and domain context. The platform ensures accuracy by actively tracking annotator proficiency against gold data tasks and/or when a judgement is modified by a higher-skilled worker or admin.

## **Ground Truth or Gold Data**

Your ground truth data is crucial at this stage of the process as the baseline to score workers and measure output quality. Many data science teams are working with a ground truth data set, however, if you are not there yet, Alegion's Customer Success (CS) team and trained workforce can get you there as part of the process of building and scaling your data pipeline.



# The Four Steps to Quality Data – Establishing Quality

Establish Quality Assurance and Conduct Quality Control

There is no one-size-fits-all QA approach that will meet the quality standards of all ML use cases. Your specific business objectives, as well as the risk associated with an under-performing model, will drive your quality requirements. Some projects reach target quality using multiple annotators. Others require complex reviews against ground truth data or escalation workflows with verification from a subject matter expert.

There are three primary sources of authority that can be used to measure the quality of annotations and that are used to score workers.

#### **GOLD DATA**

The gold data or ground truth set of records from your team can be used both as a qualification tool for testing and scoring workers at the outset of the process and also as the measure for output quality. When you use gold data to measure quality, you compare worker annotations to your own annotations for the same dataset, and the difference between these two independent, blind answers can be used to produce quantitative measurements like accuracy, recall, precision, and F1 scores.

#### **EXPERT REVIEW**

This method of quality assurance relies on expert review from a highly skilled worker, an admin, or from an expert on the customer side, sometimes all three. It can be used in conjunction with gold data QA. The expert reviewer looks at the answer given by the qualified worker and either approves it or makes corrections as needed, producing a new correct answer. Initially, an expert review may take place for every single instance of labeled data, but over time, as worker quality improves, expert review can utilize random sampling for ongoing quality control.

#### **VERIFYING OUTPUT QUALITY**

After your training data project has been configured in the platform and put through several small batch rounds of annotation, testing, and scoring, you will typically need to verify that the output quality meets your expectations before data production can get under way at scale.



# The Four Steps to Quality Data – Iterate & Scale

Iterate and Scale the Process

Congratulations! You have successfully launched your high quality training data pipeline. The last and ongoing stage in the journey to perfectly labeled training data is all about optimization and quality control. In this stage, you will need ongoing support and expertise from your platform partner in order to:

#### **TRACK VELOCITY**

In order to scale effectively, it's good to measure annotation throughput. How long is it taking data to move through the process? Is the process getting faster?

#### **TUNE WORKER TRAINING**

As the project scales, your labeling and quality requirements may evolve. This necessitates ongoing workforce training and scoring.

#### **TRAIN ON EDGE CASES**

Over time, your training data should include more and more edge cases in order to make your model as accurate and robust as possible.



# How Alegion Ensures Your Data Quality

At Alegion, quality data is our top priority. The Customer Success (CS) team is your primary point of contact at all four stages to make sure you and your data are shepherded through the annotation process in such a way that your resultant labeled datasets are high quality, and you can train a high-performing model. Because of our commitment to quality, our CS team routinely conducts admin reviews regardless of whether reviews are formally included in the project parameters for scoring workers or measuring quality.

### How We Partner with You

#### Set Quality Requirements and Annotation Criteria

Our CS experts can advise on your quality management approach with expectations around project budget, timeline, quality targets, and risk tolerance.

#### Manage Workforce Training and Platform Configuration

The Alegion CS team uses their nuanced knowledge of the platform and the demands of your project requirements to design tasks and workflows and identify, train, and score a workforce.

#### Establish Quality Assurance and Conduct Quality Control

This is where our CS team really shines. We take ownership of quality assurance and conduct ongoing quality reviews regardless of your project parameters to verify that workers are improving and annotation quality stays high.

#### Iterate and Scale the Process

Once you are thrilled with the results of small batch testing, the CS team identifies how to maximize efficiency and throughput in your workflows and how to finetune performance on edge cases in order to increase the speed and scale of data production.

# **Quality for the Full Range of Alegion Services**

Our process is designed to improve labeling accuracy and throughput, while minimizing risk and offsetting the cost of in-house data labeling.



# Quality for the Full Range of Alegion Services

At Alegion, our powerful platform and subject matter expertise allows us to support quality data labeling for your team no matter how large or small your needs.

- For fully managed labeling projects, our CS team completely owns quality from beginning to end.
- For teams who prefer a managed platform approach, our CS team provides platform configuration and support so that your annotators can do quality work.
- For teams who prefer our self-service platform option, Alegion Control, the CS team is available for training and support, helping your team use the platform more effectively to meet your quality goals.



## Conclusion

Alegion guarantees the quality of our data, and we have proven our ability to deliver. We have defined annotation industry best practices while labeling tens of millions of records for enterprise AI and ML teams.

With an ML integrated platform built to produce the highest quality annotations at scale, dedicated specialists on our CS team focused on understanding your problem space, and management of the workforce to ensure the annotations delivered will optimize model performance, Alegion offers you the high quality training data pipeline you need.

Alegion is the proven partner to build the training data pipeline that will fuel your model throughout its lifecycle. Contact us at solutions@alegion.com

