Web Vijayi

SERVICE BROCHURE

Data annotation

Your images, video, audio and text labelled by people, to rules you approve.





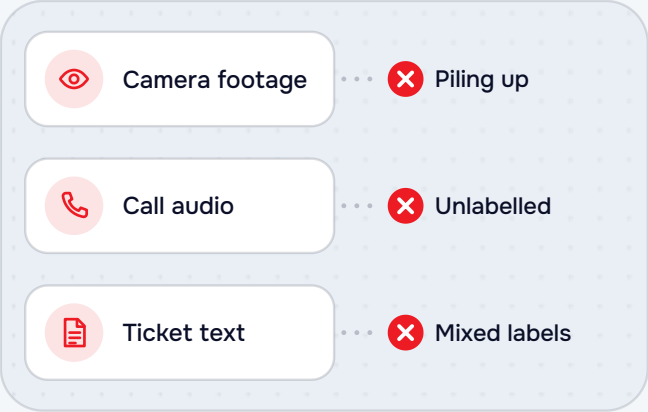
Who it is for

Medium and large companies building their own models, with more images, video, audio or text than their own people can label.

- COMPUTER VISION TEAMS
- SPEECH AND CALL TEAMS
- DATA SCIENCE GROUPS
- TEAMS FIXING MIXED LABELS

When the data outgrows your team

- Your team has far more images, recordings or records than its own people have time to label.
- Labels from several suppliers or crowd work don't agree and need checking against one set of rules.
- A model learns from a wrong label just as readily as from a right one.



BEFORE: DATA PILING UP, LABELS THAT DISAGREE

What we take on

Our annotators label your images, video, audio and text to guidelines you approve. Every label has a definition, so two people should give the same item the same answer. Hidden test items and a second review check the labels before they reach you.

Guidelines you approve

Every label gets a definition, when to use it and when not, with right and wrong examples.

Checked before you get it

Hidden test items check the work, and a second person reviews every item or an agreed sample.

A report with every batch

Figures for agreement and errors, the most common mistakes, and how each figure was worked out.

How it works

WE DO THIS · TOGETHER

1 TOGETHER

You send a sample and the goal

A batch of your real data, awkward items included, and what the model should learn.

2 TOGETHER

We write the guidelines with you

We agree each label, the hard cases and how deep the checks go. You approve the guidelines before anyone labels for real.

3 US

We label in batches and check the work

Automated checks catch mechanical errors, hidden test items test each annotator, and a second person reviews every item or an agreed sample.

4 TOGETHER

You check each batch and its report

Labels arrive in the agreed format with figures for agreement and errors. Send back any items you disagree with.

What you get

- ✓ Labelling guidelines written with you, with right and wrong examples and a version number
- ✓ Labelled data in your export format
- ✓ A quality report with every batch: agreement, errors found and guideline changes
- ✓ A log of every question annotators raised and how it was decided
- ✓ Quality targets set with you for your project

What we need from you

- A sample of real data, and what the model should do
- Someone who can rule on hard cases while the guidelines settle
- Items with answers your experts have checked, or time to build them
- The export format you need
- Your data protection terms, with unneeded personal details removed

Timeline We agree the timeline with you before we start.

How you stay in charge

You set the rules, see every batch and decide what changes.

Guidelines you approve

You approve the guidelines before labelling starts, and no rule changes without your sign-off.

Batches and reports you can check

Each batch comes with a quality report. Open any batch, check any item and read the log of every question raised.

A named person on each side

A named contact on our side runs your project. Anything the guidelines miss goes to the person you name.

Change, pause or stop it

Between batches, change the guidelines, the volume or what comes first. Pause or stop with the notice in your contract.

Questions people ask

Is annotation the same as labelling?

In practice, yes. US teams spell it labeling. Both mean adding labels to raw data so a model has examples to learn from.

How do you check the labels are right?

Hidden test items with checked answers, agreement between annotators on the same items, and a check of a random sample. Each report shows all three.

Which formats and languages do you cover?

We agree the export format, the label types and the languages with you before work starts, once we have seen your data.

How is our data kept safe?

Where your data is kept, who can open it and what happens to it at the end are agreed in the contract before work starts.

Send us a sample of your data

Tell us what the model should learn. In a free consultation we'll show you how we'd start the guidelines.

Every project is quoted in writing after a free consultation, with the timeline agreed in the same quote.

webvijayi.com

care@webvijayi.com

Book a free consultation: **webvijayi.com/contact**