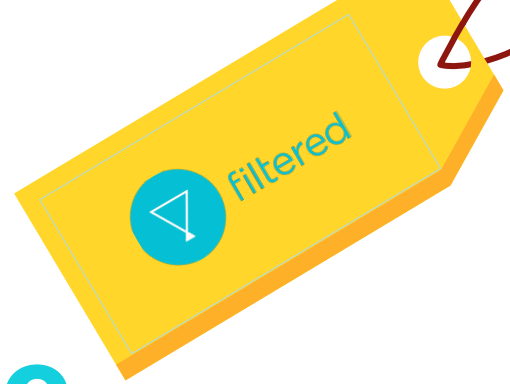
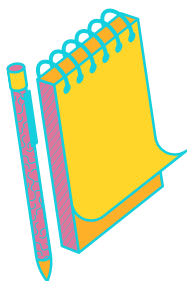


# The Ultimate Content Tagging & Validation Guide



This is your guide to the best content tagging practices and the suggested approach to its validation.



## Best practices

### Skills

It's crucial that skills are well-defined and well-thought-out as part of the content learning tagging process.

### Process

Process matters. Think through how your tagging process will work. Choose what's most efficient for curators.

### Native tags

Pre-filled tags can make the process go faster. But, much of this metadata is often surprisingly partial and inconsistent.

### Extra eyes

For the content tagging that's most essential to your organisation, make sure curators' work is marked independently. This allows for a second (or third) set of eyes.

### Size

How much content needs tagging? Before diving into a tagging process, it's valuable to review some high-level summary statistics on the content you want to tag.

### Data

Measure your tagging. Generate data on the usefulness of your content libraries, measure its provision against the your skills frameworks and find skills gaps.



## What next?

Pair solid content tagging with well-thought-out content curation, and you have a winning formula — one that favours personalising content for each learner, running engagement campaigns through integrations, building pathways, and catering for social learning.

## Filtered's optimisation & validation

In preparation for tagging, we capture feedback on their performance which we use to optimise and validate using these steps.

### Internal validation metrics

Such as the frequency of a keyword in our databases of assets. This tells us the influence of the keyword on results and how powerful it is in terms of differentiating content which should/shouldn't attract a tag.

### Training-set results review

During configuration, we review results dynamically (predictions as to which assets should have the tag applied) and make subjective assessments of their quality.

### Post-hoc precision on the training set

We calculate a metric that we call 'post-hoc precision'. This involves taking a random sample of assets that the model has tagged, and calculating the percentage of those decisions which seem sensible to a human subject matter expert.

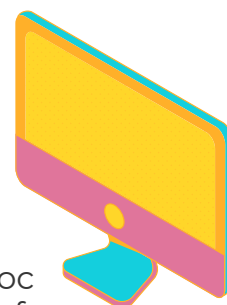
### Post-hoc precision on the prediction set

This metric is calculated in exactly the same way as post-hoc precision on the training set, but using a sample of the assets which we are ultimately trying to tag for the client.

## Client-side evaluation recipe

We recommend that clients evaluate tagging results using the 'post-hoc precision' metric. This is a good indicator of the quality of tagging results for onward use. The recipe to calculate post-hoc precision of the supplied assets is:

- Select a random sample of the tagged assets (start with around 50 assets in the sample).
- Share the sample with a human curator.
- Ask them to review the sample and mark those tagging decisions they agree with.
- Calculate the % of the sample where they agree with the tag applied.



## Better way to tag

On average, our customers reduce their content spend by 30% and produce a library of impeccably tagged content at 10x the speed of human curators.

Filtered's **Content Intelligence** accurately and efficiently tags learning content by turning the text of content itself into data.