Cloud AutoML Vision

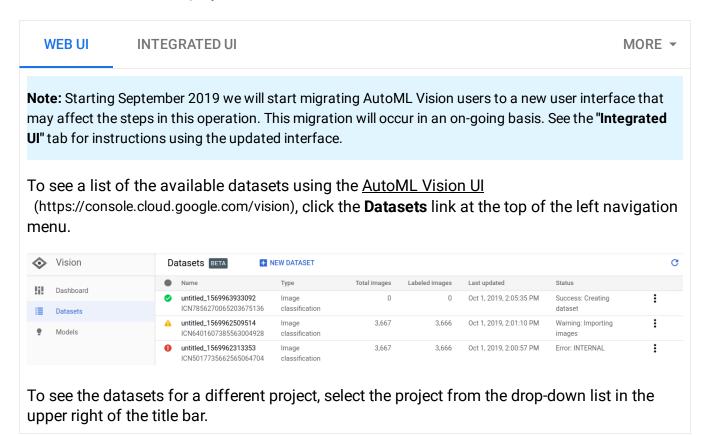
Managing datasets

A project can have multiple datasets, each used to train a separate model. You can get a <u>list of the available datasets</u> (#list-datasets), <u>get a specific dataset</u> (#get-dataset), <u>export a dataset</u> (#export-dataset) and can <u>delete a dataset</u> (#delete-dataset) you no longer need.

Use labels with user-defined metadata to organize your datasets. See <u>Using labels to organize Google Cloud Platform resources</u> (https://goo.gl/xmQnxf) for more information on labels including common uses.

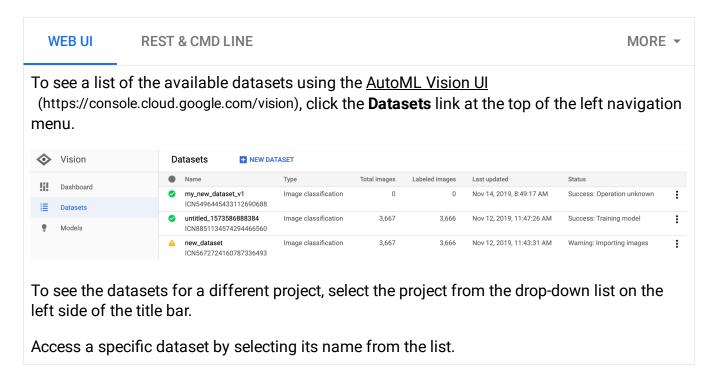
Listing datasets

A project can include numerous datasets. This section describes how to retrieve a list of the available datasets for a project.



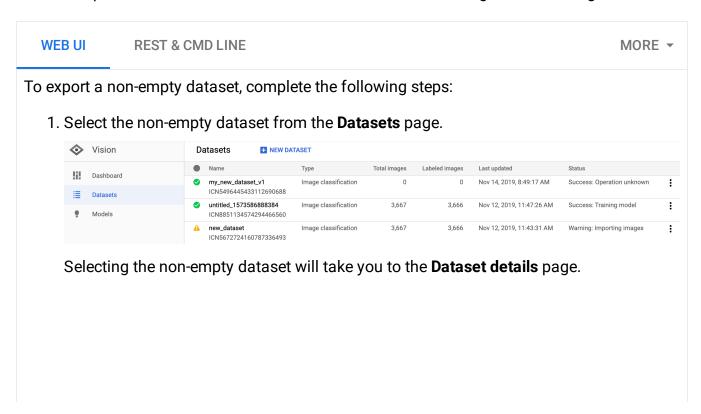
Get a dataset

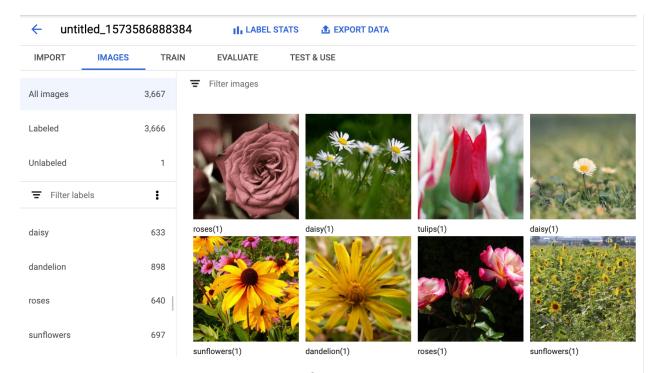
You can also get a specific dataset using a dataset ID.



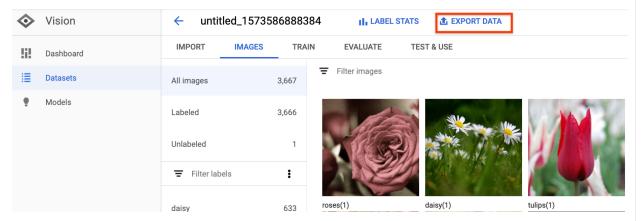
Export a dataset

You can export a CSV file with all a dataset's information to a Google Cloud Storage bucket.





2. Select the Export data option at the top of the Dataset details page.



This opens a window where you can choose a Google Cloud Storage bucket location, or create a new bucket and designate it as the location to store the CSV file.

Export data

Your dataset will be exported as a list of images in CSV format. Each row contains a Cloud Storage path, any label(s) assigned to that image, bounding boxes, and a flag that indicates whether that image is in the training, validation, or test sets. You will be emailed once exporting has completed.



BROWSE

Destination folder is required.

CANCEL EXPORT

3. Select **Export CSV** after you have selected a new or existing Google Cloud Storage bucket location.

Export data

Your dataset will be exported as a list of images in CSV format. Each row contains a Cloud Storage path, any label(s) assigned to that image, bounding boxes, and a flag that indicates whether that image is in the training, validation, or test sets. You will be emailed once exporting has completed.



You will be notified when exporting finishes.



4. You will receive an email when the data export process has completed.

Exported CSV format

The exported CSV file contains the same formatting as the <u>training data import CSV</u> (https://cloud.google.com/vision/automl/docs/prepare):

```
set,path,label0[,label1,label2,...]
```

This CSV file is saved in a created export folder that is distinguished by a unique timestamp. Below are some sample lines from an exported CSV file:

```
my-storage-bucket/export_data-my_dataset_name-2019-11-
08T22:28:13.081Z/image_classification_1.csv:
```

TRAIN, gs://my-storage-bucket/export_data-my_dataset_name-2019-11-08T22:28:13.081Z/fivalIDATION, gs://my-storage-bucket/export_data-my_dataset_name-2019-11-08T22:28:13.08 TRAIN, gs://my-storage-bucket/export_data-my_dataset_name-2019-11-08T22:28:13.081Z/fivalIDATION, gs://my-storage-bucket/export_data-my_dataset_name-2019-11-08T22:28:13.08 TRAIN, gs://my-storage-bucket/export_data-my_dataset_name-2019-11-08T22:28:13.081Z/fivalIDATION.

Deleting a dataset

You can delete a dataset you no longer need in the UI, or with the dataset's ID via the following code samples.

WEB UI INTEGRATED UI MORE ▼

Note: Starting September 2019 we will start migrating AutoML Vision users to a new user interface that may affect the steps in this operation. This migration will occur in an on-going basis. See the "**Integrated UI**" tab for instructions using the updated interface.

- 1. In the <u>AutoML Vision UI</u> (https://console.cloud.google.com/vision), click the **Datasets** link at the top of the left navigation menu to display the list of available datasets.
- Click the three-dot menu at the far right of the row you want to delete and select **Delete** dataset.
- 3. Click **Delete** in the confirmation dialog box.



Except as otherwise noted, the content of this page is licensed under the <u>Creative Commons Attribution 4.0 License</u> (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the <u>Apache 2.0 License</u> (https://www.apache.org/licenses/LICENSE-2.0). For details, see our <u>Site Policies</u> (https://developers.google.com/terms/site-policies). Java is a registered trademark of Oracle and/or its affiliates.

Last updated January 22, 2020.