

Highlight

Note

Dataset Versioning

Datasets can be *versioned*, enabling you to track historical versions of datasets that were used in experiments, and reproduce those experiments with data in the same state.

Creating a New Version of a Dataset

You can create a new version of a dataset by registering it with the same name as a previously registered dataset and specifying the **create_new_version** property:

```
img_paths = [(blob_ds, 'data/files/images/*.jpg'),  
              (blob_ds, 'data/files/images/*.png')]  
file_ds = Dataset.File.from_files(path=img_paths)  
file_ds = file_ds.register(workspace=ws, name='img_files', create_new_version=True)
```

In this example, the .png files in the **images** folder have been added to the definition of the **img_paths** dataset example used in the previous topic.

Retrieving a Specific Dataset version

You can retrieve a specific version of a dataset by specifying the **version** parameter in the **get_by_name** method of the **Dataset** class:

```
img_ds = Dataset.get_by_name(workspace=ws, name='img_files', version=2)
```