Data structure in OMERO and organization with tags

Workshop: **FAIR data handling for microscopy: Structured metadata annotation in OMERO**

April 29th & 30th, 2024

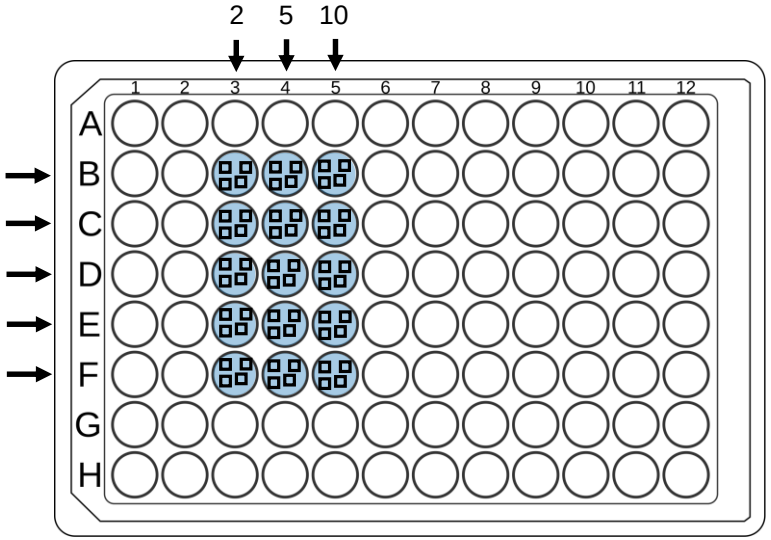
Day 1 Session 3

Trainer: Vanessa Fuchs, **Tom Boissonnet**, Christian Schmidt

29-04-2024

Example data for this session

- A plate acquired at week intervals
- One compound per row
  - Compound B, Compound C, …
- Tested at different concentration (columns):
  - 2 µM, 5 µM, 10 µM
- Multiple images are taken from each well for an accurate sampling.



Not so serious analogy

datasets

images

annotations

By Tom Rossini, CC-BY

By Gordon Jolly, https://flickr.com/photos/loopzilla/8132242517/, CC-BY-SA, https://creativecommons.org/licenses/by-sa/2.0/

Not so serious analogy

Heinrich Heine Universität Düsseldorf

datasets

images

Tags

10 µM – week 1 – compound A –

Key-value pairs

compound	A
concentration	10 µM
week	1

Week1_150607_B02_s1_c1-DAPI_compoundA_10µM.tif

By Tom Rossini, CC-BY

By Gordon Jolly, https://flickr.com/photos/loopzilla/8132242517/, CC-BY-SA, https://creativecommons.org/licenses/by-sa/2.0/

I3D:bio

Searching the optimal organization

Heinrich Heine
Universität Düsseldorf

- What should our "dataset unit" be ?
  - the whole plate, of every week ?
  - the whole plate, one dataset per week ?
  - the compound of every week ?
  - the compound of a single week ?
  - …
- What's a convenient way to describe the data content ?

I3D:bio

Introducing OMERO at HHU

Heinrich Heine
Universität Düsseldorf

OMERO has only a two level hierarchy. But you can use tags

Tags: just like hashtags in Twitter or Instagram

You can use annotations to add metadata

metadata?

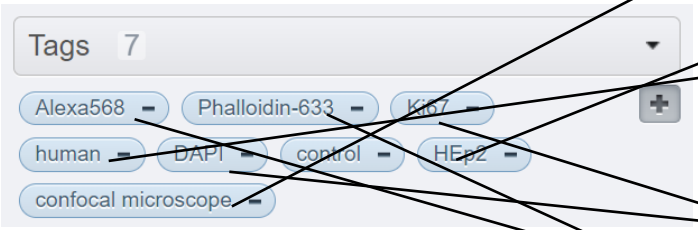
I'll think about how to organize my data later

I3D:bio

Overlapping concepts

Heinrich Heine Universität Düsseldorf

Tags



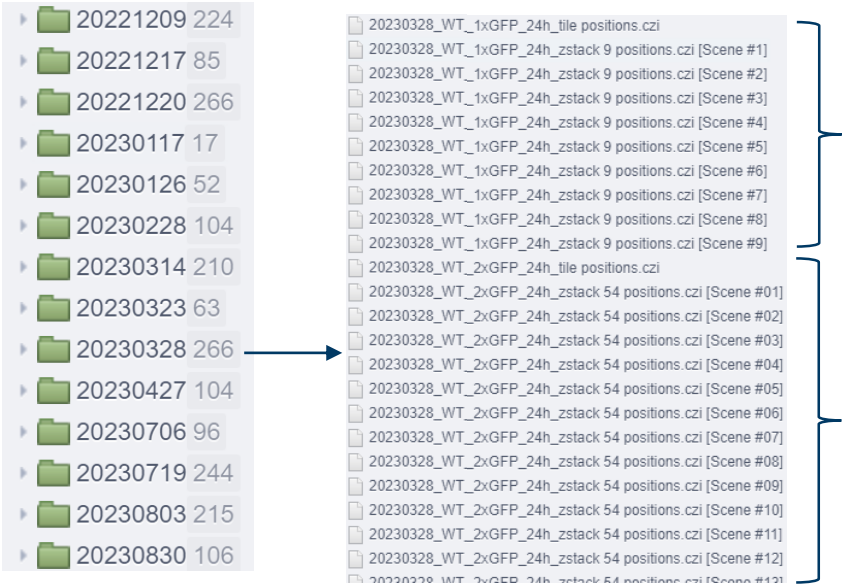
Study component	
imaging method	confocal laser scanning microscopy
imaging method term accession number	http://purl.obolibrary.org/obo/CHMO_0000089
imaging method term accession number source REF	chemical methods ontology
Biosample	
biological entity	HEp2 cells
biological entity term accession number	http://purl.obolibrary.org/obo/BTO_0000976
biological entity term accession number source REF	BRENDA tissue / enzyme source
species	human
species term accession number	http://purl.obolibrary.org/obo/NCBITaxon_9606
species term accession number source REF	NCBI organismal classification
Specimen	
preparation method	#EXP00114_20230615_CAI_Test_Staining_for_Practicals
staining	
staining	#EXP00114_20230615_CAI_Test_Staining_for_Practicals
channel1 - content	DAPI
channel1 - biological entity	DNA, mainly nucleus
channel2 - content	antiKi67(rabbit), antiRabbit-AlexaFluorophore488
channel2 - biological entity	Ki67, proliferation marker
channel3 - content	antiLaminA(mouse), antiMouse-AlexaFluorophore568
channel3 - biological entity	LaminA, nuclear membrane
channel4 - content	Phalloidin-633
channel4 - biological entity	filamentous Actin

Key-value pairs

I3D:bio

Using project & dataset as folders

- Folder habits are intuitively matched to projects and datasets
- Thinking dataset as folders leads to a confusing structure in OMERO:

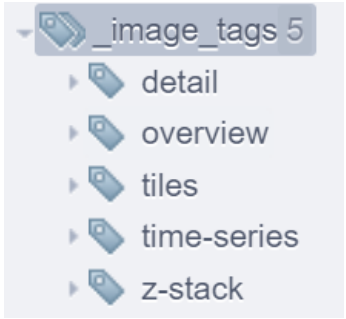


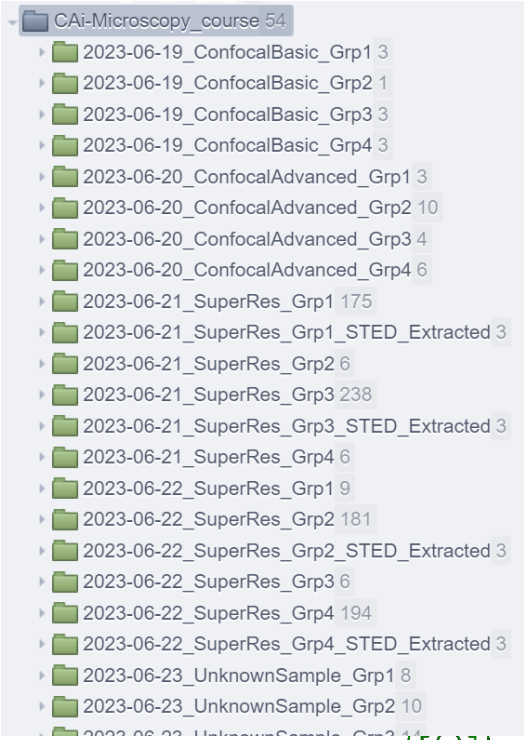
- How should I annotate this?
  - Annotate dataset: unclear to know which images it applies to
  - Annotate first image of every set: other images need reference to the annotated image
  - Annotate every image: tedious and annotation duplication
- What if someone wants to look at the data differently?

Need some organization for data?

- We can use tags like folders (not possible with datasets)
- Don't overdo it: less tag categories -> less work
  (and less oversight)
- Tags are flexible and should be used for individual preferences

Need to group images across datasets?
Tags can do that





The scope of tags

- KV-pairs takes care of describing the data, so no need to redo it with tags
- Tags are like folders, but better (multiple tags per object, descriptions)
- Tags can be categorized with Tagsets

-> Good for organizing and filtering

- Tags could emulate my datasets (don't try)

- So why should we even bother with datasets ?

The scope of datasets

Heinrich Heine
Universität Düsseldorf

- Datasets can be annotated !
  - tags
  - key-value pairs
  - attachements
  - ratings
  - comments

- By annotating a dataset, I can implicitely annotate the images it contains
  - group images of a same experiment in dataset

  -> less dupliction of annotations
  - images can be annotated to give more details

I3D:bio

The scope of Key-value pairs

- The annotation that describes the data in details:

species	human
species term accession number	http://purl.obolibrary.org/obo/NCBITaxon_9606
species term accession number source REF	NCBI organismal classification
Specimen	
preparation method	#EXP00114_20230615_CAI_Test_Staining_for_Practicals
staining	DAPI
channel1 - biological entity	DNA, mainly nucleus
channel2 - content	phalloidin - 488

Precise terms description with ontologies (next session)

link to protocol

Back to the plate data

- How do I describe the data content, with the lesser effort?
- What is the "dataset unit"?

- Practice time:
- Split down the data so that a dataset contain only images of the exact same condition:
  - Per well (one compound, one concentration)
  - Per plate (one repetition of the experiment)
  - Per week (one sampling)

- Annotate condition with key value pairs on the dataset
- Tag the datasets with what you estimate is necessary

Annotations

- Projects, Datasets and Images can be annotated

Tags

arabidopsis

Comments

Tom Boissonnet at 2024-04-22 10:09:33

noice

Key-value pairs

temperature	21 °C
CO_2	10 %

Ratings

Attachments

Table

image	roi	area	avg_inten_C0	avg_inten_C1	avg
13278	65584	1205677.0	38701.0	42363.0	760
13278	65585	1470548.0	28573.0	42220.0	947
13278	65586	1200417.0	35192.0	44121.0	747
13278	65587	1447938.0	32550.0	45276.0	933
13278	65588	1274438.0	32099.0	36876.0	799
13278	65589	1421678.0	58407.0	60823.0	884
13278	65590	1220604.0	41654.0	47156.0	792
13278	65591	1184172.0	39427.0	46411.0	756
13278	65591	1224946.0	24840.0	33459.0	768
13279	65592	1424809.0	17905.0	33864.0	941
13279	65593	995294.0	25585.0	38470.0	616
13279	65594	1204997.0	15061.0	28537.0	775
13279	65595	1081951.0	49136.0	86084.0	129
13279	65596	1475485.0	37292.0	48017.0	950
13279	65597	1526607.0	27268.0	41203.0	985
13279	65598	1134129.0	36584.0	42564.0	706

Treasure hunt

Follow the clues left on OMERO
to find a hidden reward.

Your search starts in the tags