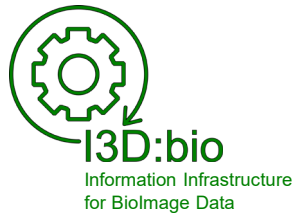


Research Data Management for Bioimage Data at the HHU

Data Organization in OMERO



Tom Boissonnet

Adapted from: Schmidt C., Bortolomeazzi M., Boissonnet T., Fortmann-Grote C. *et al.* (2023). I3D:bio's OMERO training material: Re-usable, adjustable, multi-purpose slides for local user training. Zenodo. DOI: 10.5281/zenodo.8323588
If not stated otherwise, the content of this material (except for logos and the slide design) is published under a [Creative Commons Attribution 4.0 license](#).

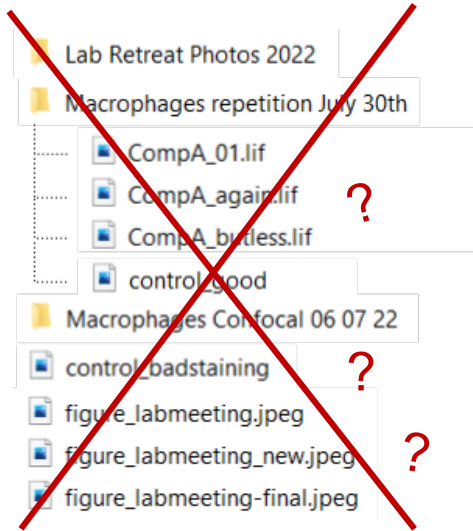
Questions/considerations around data organization and management

- Local computer vs. shared drives (Where?)
- Backed-up network drives vs. safety copies on hard drives (Where? Who?)
- Version control software vs. manual versioning (file names) (How?)
- Arbitrary file naming vs. (any level) of standardization (How? Who?)
- Management software vs. file folder hierarchies (How?)
- Documentation in paper notebook vs. electronic lab notebook (How? Who?)
- Original data vs. derived data (What? How?)
- Automatic recording, sample barcoding, etc.
- ...



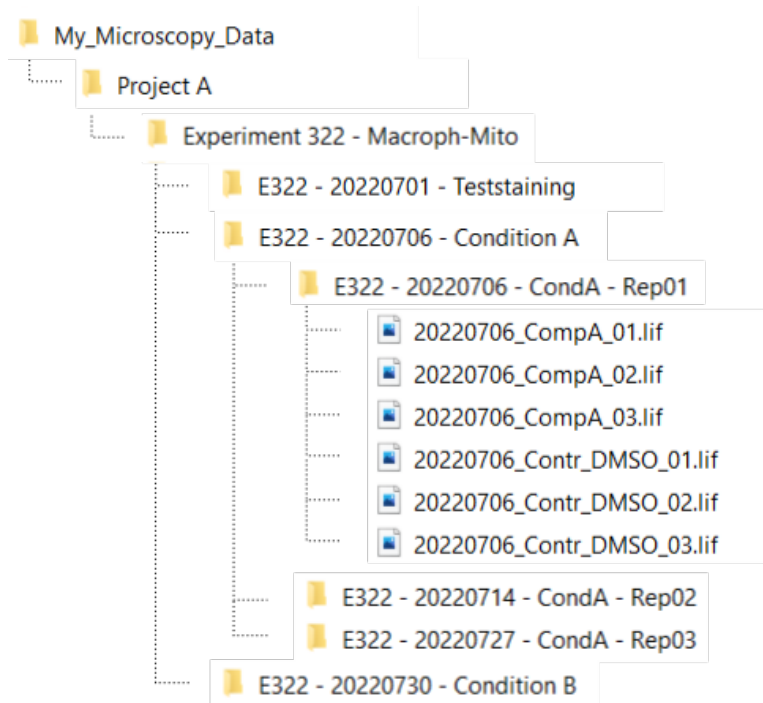
Managing data in classical file folder hierarchies

Not managed



Understanding the data is only possible for the data producer (if at all)

Managed in a file system hierarchy (arbitrary example)



Understanding the data is possible based on the researcher's documentation

Hierarchy structure standardization?

- At the individual's level
- At the group's level
- At the collaboration level
- Discipline-specific standards

No preview and limited direct access to metadata

A data management system helps to organize data (here: OMERO)

Example:
Access with
OMERO.web

Example: A file with a „multi-scene image“

- Managed data
- Preview available
- Access to metadata
- User-friendly but machine-accessible

The screenshot shows the OMERO web client interface. On the left, a file explorer shows a directory structure under 'Mary Mayperson' with a file '14 08943 8_A1APOBEC3A_1.czi [0]' selected. In the center, three image thumbnails are displayed: a blue-tinted microscopy image, a label image with handwritten text '14 08943 8', 'APOBEC 3A', and 'H15 844', and a green-tinted microscopy image. On the right, a metadata panel for the selected file shows details such as 'Image ID: 15201', 'Owner: Mary Mayperson', 'Acquisition Date: 2017-05-17 16:33:20', and 'Import Date: 2022-07-20 12:52:05'. The interface includes a search bar, navigation tabs, and various tool icons.

Preview thumbnails

Metadata

Data organization in OMERO – part 1

OMERO.web offers a tree-view data hierarchy in the **Explore** tab



Projects



Datasets



Screens (multi-well plates)

user/group

Projects (blue)

Datasets (green)

Image(s) (grey)

GROUP Mary Mayperson

Explore Tags Shares

- Mary Mayperson
 - 2010_CMBE_ExpressionTest 2
 - 2011_VTK_ZyxinConfocal 1
 - 2012_MPIMolBio_CrossSec_Lymphangio 1
 - 2017_Nuc-Intens_Tcells 6
 - 20171115_M33_pERK 25
 - 20171115_0h_DMSO_pERK_01_D1.czi
 - 20171115_0h_DMSO_pERK_02_D1.czi
 - 20171115_0h_MG_pERK_01_D1.czi
 - 20171115_0h_MG_pERK_02_D1.czi
 - 20171115_0h_X_pERK_01_D1.czi
 - 20171115_0h_X_pERK_02_D1.czi
 - 20171115_3h_DMSO_pERK_01_D1.czi
 - 20171115_3h_DMSO_pERK_01_D2.czi
 - 20171115_3h_DMSO_pERK_02_D1.czi
 - 20171115_3h_DMSO_pERK_02_D2.czi

Dataset within project

Non-nested Dataset

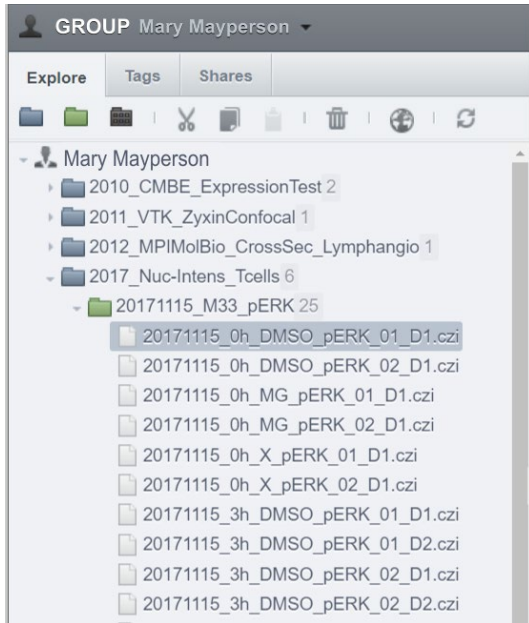
Scenes from multi-scene image (e.g., slide scanner)

Explore Tags Shares

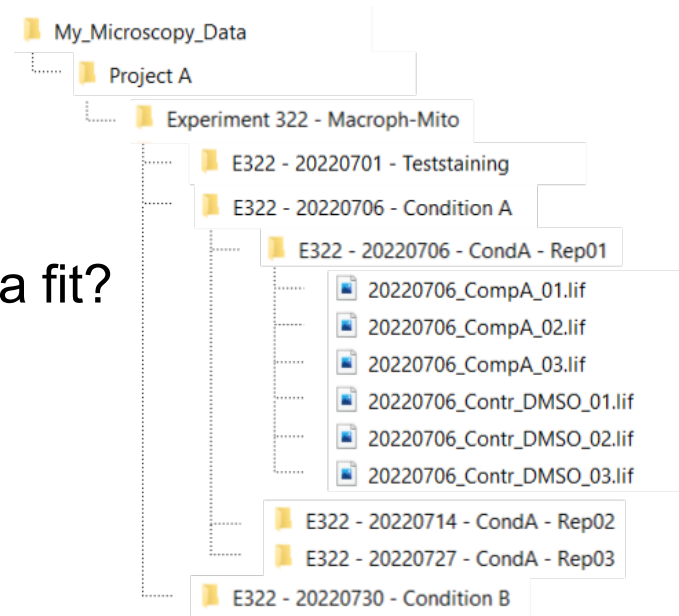
- Mary Mayperson
 - ExampleData_BiolImageArchive_WSI 1
 - BIA_Accession_S-BSST110 3
 - 14 08943 8_A1APOBEC3A_1.czi [0]
 - 14 08943 8_A1APOBEC3A_1.czi [label image]
 - 14 08943 8_A1APOBEC3A_1.czi [macro image]
 - Orphaned Images

Re-think data organization: File folder hierarchy vs. object-oriented data structure in OMERO?

If OMERO offers only a two-folder deep hierarchy...



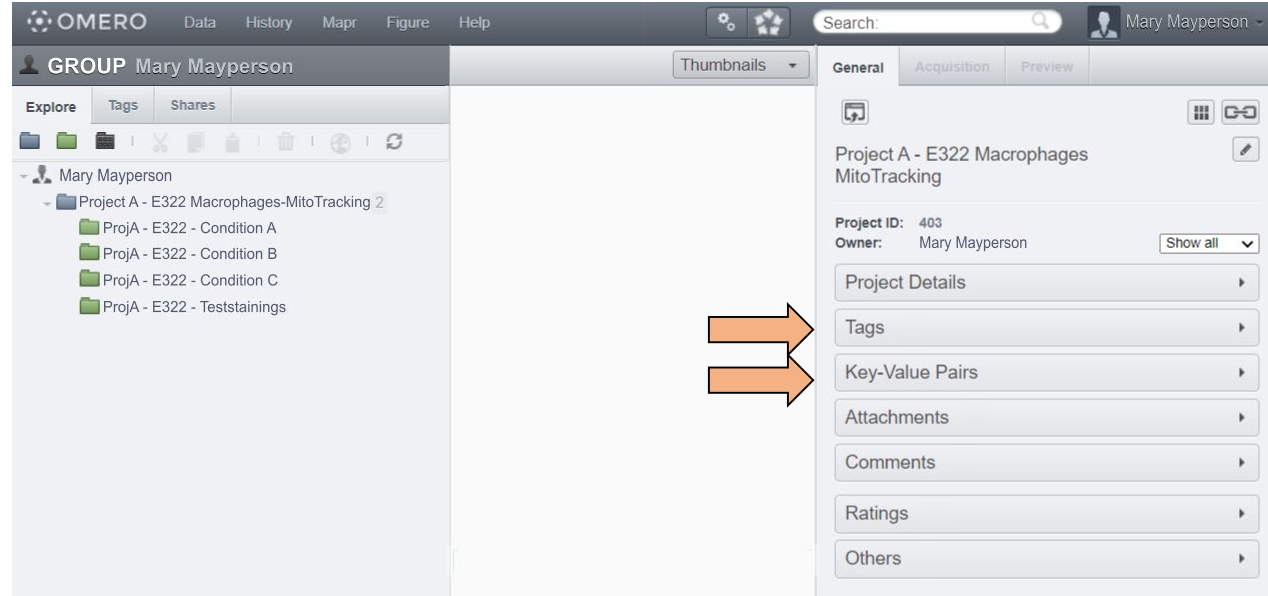
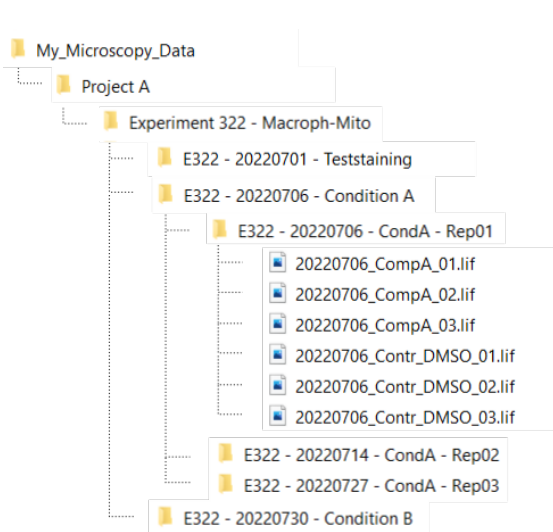
... how does my data fit?



File folder hierarchy in explorer vs object-based data structure

A file folder hierarchy is itself a form of metadata

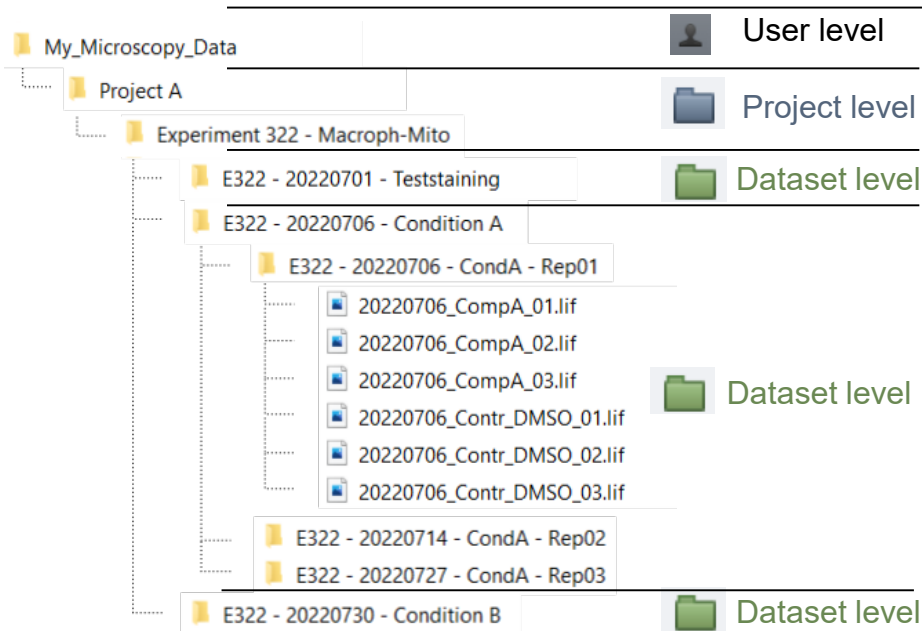
OMERO uses *structured metadata*, e.g., with **Tags** and **Key-Value Pairs**



→ Annotate data with Tags and Key-Value Pairs

Re-think data organization: Object-based data organization

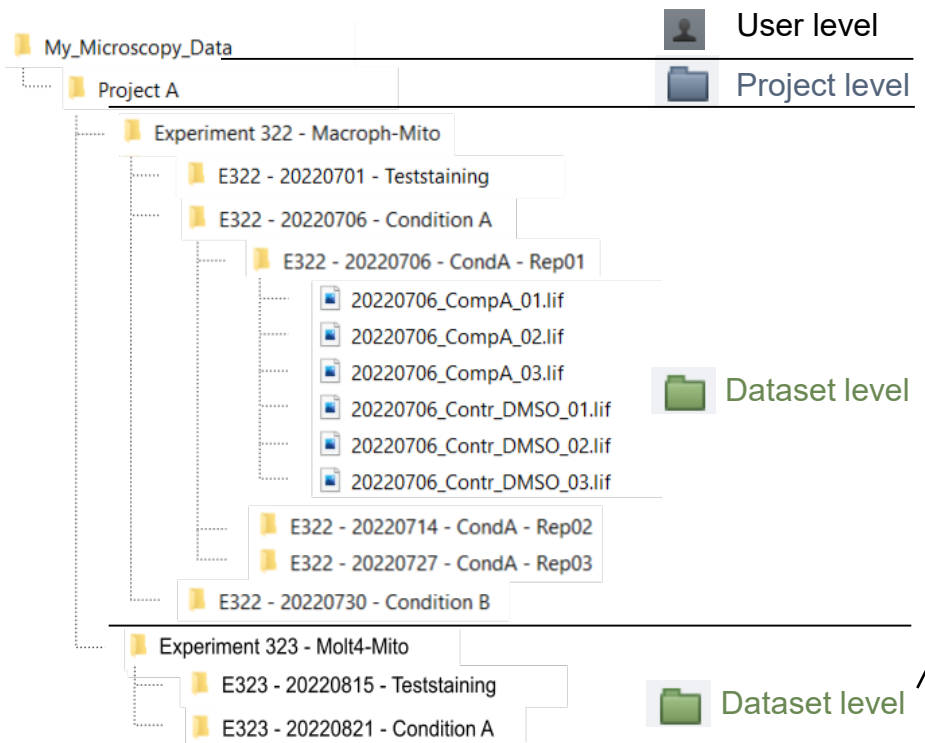
Organization example 1



Use Tags for information across datasets and to *substitute for deep folder hierarchies*
(→ How? See the following slides and chapters!)

Re-think data organization: Object-based data organization

Organization example 2



Use **Tags** for information across datasets (e.g., „Compound A“, „DMSO control“, etc.)

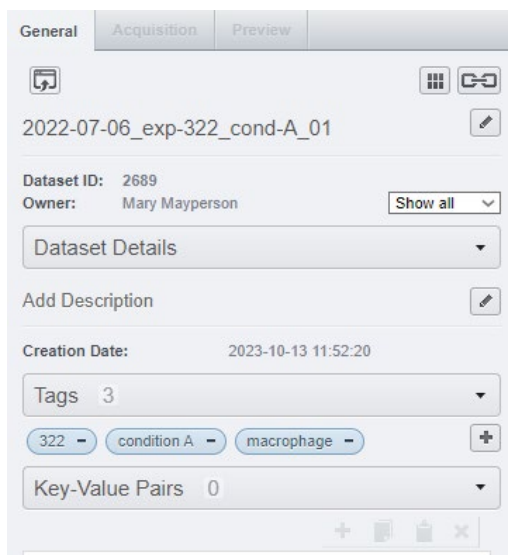
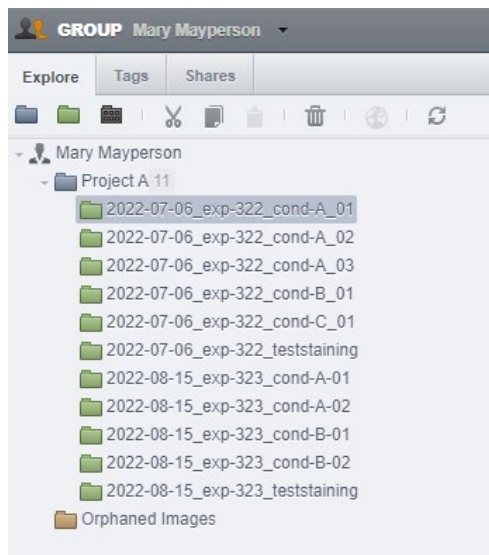
Use **Key-Value Pairs** to enrich with metadata details

(→ How? See the following slides and chapters!)

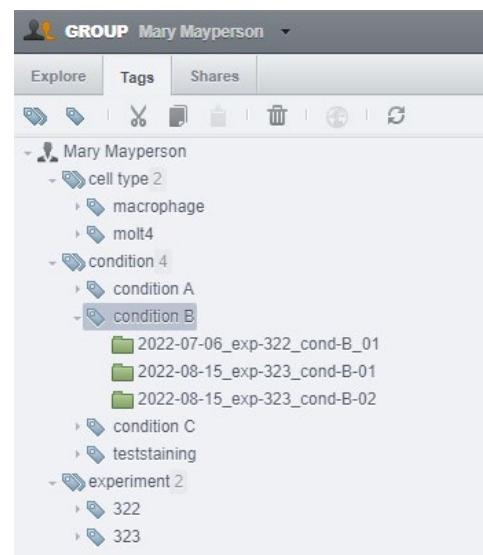
Re-think data organization: Object-based data organization

Organization example 3

Datasets on a flat list act as a "data library"



All folders are mapped to tags



- Images of an experiment are divided to obtain datasets of the same sample + condition + experiment + ...
- As the "data library" grows, the tags grant a flexible and efficient filtering
- Because images have the same "origin", they are implicitly annotated all at once (by annotating the dataset)

Re-think data organization!

This is an important concept:

OMERO is *not* intended for use as a file hierarchy system

OMERO is object-oriented

How to leverage the potential of object-oriented data organization using Tags and Key-Value-Pairs will be shown throughout the following chapters

Re-think data organization: Object-based data organization

How to organize data in OMERO?

- Structure data *according to your (group's) needs*
 - Make use of **Tags** (and/or Key-Value Pairs) for data organization (instead of deep folder hierarchies)
 - Explore ways of data organization and discuss them with your research group

→ What are Tags and Key-Value Pairs?

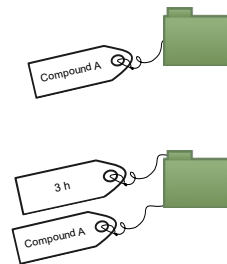
What are Tags and their advantages?

Tags denote a property of an entity, similar to a price tag in the supermarket

- Tags allow a dynamic re-representation of the data tree (Tag-based search)
- Tags can help organize data across datasets and projects (e.g., similarities, relationships)

Examples (Tags):

- „show all data treated with **compound A**“
- „show all data recorded with **instrument A**“
- etc...



Note: Tags are associated with users and groups!

- In a shared group, we recommend discussing which tags to use and potentially assigning a user who curates and manages the tags
- For private data, tags can be used based on user preference alone

What are Key-Value Pairs and their advantages?

Key-Value Pairs allow (standardized) annotation of detailed metadata

Consists of

- **Key:** Denotes a real-world object or an abstract concept that has a specific value (out of several or many possible values)
- **Value:** Number or text-string that describes the object denoted under „Key“

Examples:

Key: „cell type“ **Value:** „CD4+ T cell“

Key: „disease model“ **Value:** „experimental autoimmune encephalomyelitis“

→ Allows structured and standardized metadata details curation

Making use of the data organisation

→ see sub-chapter 6: Search Data in OMERO

Data annotation with Key-Value Pairs and Tags

→ see Chapter 07 on Metadata Curation

→ see official OMERO documentation

<https://omero-guides.readthedocs.io/en/latest/introduction/docs/index.html>