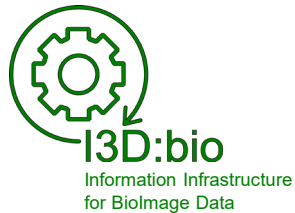


Research Data Management for Bioimage Data at the HHU

Metadata curation 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](#).

What is „metadata“?

In general, metadata is often called „data about the data“.

In bioimaging, metadata accompanies the actual image (pixel) data:

- Technical metadata *(automatically recorded or added manually)*
 - Information about the instrument, including hardware components, filter settings, etc.
- Sample metadata *(researcher's documentation or published protocol)*
 - Information about the specimen, organ, cell type, sample type, test group, etc., and the experimental procedures during sample preparation, e.g., sample fixation, staining, use of antibodies, etc.
- Analysis metadata

Where to find the metadata?

Image file header Sidecar / additional file Electronic labbook Paper notebook In the data organization ...

Why does metadata matter? – the societal perspective

Good scientific practice principles (DFG Code of Conduct: doi:[10.5281/zenodo.6472827](https://doi.org/10.5281/zenodo.6472827))

- Data is fully and correctly described
- Results are reproducible

Open Science principles ([OECD recommendations](#), adopted 2006, amended 2021)

- Publicly funded research output should be available open access

FAIR data principles (Wilkinson et al., 2016, *Sci Data*, doi: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18))

- Data should be Findable, Accessible, Interoperable, and Reusable

→ **Rich metadata enhances reproducibility, trust, and openness, contributing to sustainable research**

Why does metadata matter? – the personal perspective

Funding agency demands

- (Minimum) metadata standards become mandatory for third-party funding, e.g.:
[DFG guidelines](#), [ERC guidelines](#)

Career Advancement

- Publications with links to deposited data correlate with more citations, see Colavizza et al., 2020, *PLoS ONE*, doi: [10.1371/journal.pone.0230416](https://doi.org/10.1371/journal.pone.0230416)
- Sharing (reusable) data can foster new collaboration

Future-Me

- Metadata will help your colleagues and yourself to re-use your own data in the future

→ Rich metadata annotation adds value to the research data

Metadata „standards“ in bioimaging?

Which metadata should be collected?

- Research-discipline-specific recommendations
- Bioimaging-specific recommendations:
 - Sarkans *et al.* (2021) REMBI: Recommended Metadata for Biological Images – enabling reuse of microscopy images in biology. *Nat Methods*, Dec;18(12). doi: [10.1038/s41592-021-01166-8](https://doi.org/10.1038/s41592-021-01166-8).
 - Hammer *et al.* (2021) Towards community-driven metadata standards for light microscopy: tiered specifications extending the OME model. *Nat Methods*, Dec;18(12). doi: [10.1038/s41592-021-01327-9](https://doi.org/10.1038/s41592-021-01327-9).
- Repository-specific requirements
 - e.g., the Image Data Resource (IDR): <https://idr.openmicroscopy.org/about/submission.html>
(see also: Ellenberg *et al.*, 2018, *Nat Methods*, doi: [10.1038/s41592-018-0195-8](https://doi.org/10.1038/s41592-018-0195-8))

What is the “correct“ or “best“ way to work with metadata in OMERO?

There is not “one correct“ way!



OMERO offers possibilities for structured metadata for bioimaging data.

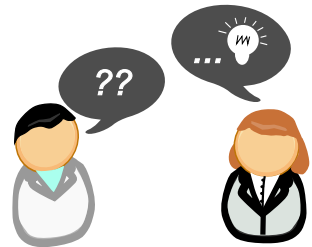
As a researcher or research group, familiarize yourself with OMERO’s possibilities.



Like with good experimental design, the suitable metadata enrichment for your data should be part of your (group’s) considerations when performing research.

I3D:bio recommends consulting with core facility staff and colleagues who have previously used OMERO for their work and attending OMERO-specific workshops.

(<https://www.i3dbio.de>)



How is metadata organized in OMERO?

Metadata is contained:

1. In the data organization tree (implicitly)

2. Under the General tab

- Image Details
- Tags
- Key-Value Pairs
- Attachments
- Ratings

3. Under the Acquisition tab:

- Original metadata

The screenshot shows the OMERO web client interface. A red box labeled 'example' is overlaid on the file tree. Three orange arrows labeled 1, 2, and 3 point to different parts of the interface: arrow 1 points to the file tree, arrow 2 points to the 'General' tab, and arrow 3 points to the 'Acquisition' tab. A red box labeled 'Tags and Key-Value Pairs' points to the 'Key-Value Pairs' section in the 'General' tab. The 'Key-Value Pairs' section contains a table with the following data:

Key	Value
CellType	CD4+ T cell
Organism	Mus musculus
Organ	Spleen
IsolationMethod	MACS negative sort
ActivationMethod	anti-CD3/anti-CD28

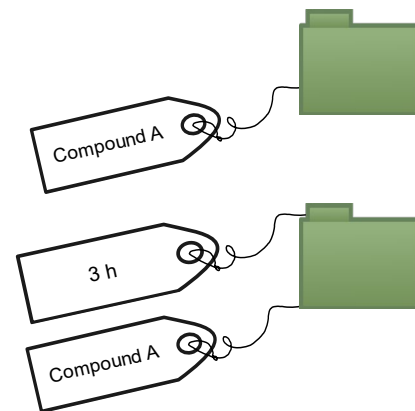
Tags

Tags denote a specific property of an entity (like a price tag in the supermarket)

- Tags allow a dynamic representation of the data tree
 - Data tree is re-arranged in Tag-based search according to the search interest
- Tags can define similarities and relationships
- Tags can be used to quickly access and process data of interest (see also Chapter 6.1)

Examples (Tags):

- „show all data for samples treated with **compound A**“
- „show all data for samples incubated for **3 h**“
- „show all data recorded with **instrument A**“
- etc...



Tags in OMERO

Which tags you should use depends on your research and your data

Tags can be added to:

- Images
- Groups of Images
- Datasets
- Projects

You can define **Tag Categories**

- To represent nested tags
- E.g., in a research collaboration
- *Example:* <https://neuroinformatics.icm-institute.org/knowledge-base/omero-tags-dictionary/>

Tags can be

- based on *ontology* terms
- based on a collaboration agreement

I3D:bio recommends:

- Use Tags for data organization across datasets and projects (categorization) (instead of deep folder hierarchies)
- Use Key-Value Pairs for metadata enrichment

→ **Discuss suitable Tags within your group / with the data manager**

Key-Value Pairs

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

Consist of

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

Examples:

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

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

Key-Value Pairs in OMERO

- Use Key-Value Pairs for metadata details about the experiment
- If available, review metadata recommendations in your field of research
- Recommended: Use *ontology terms* in Tags and Key-Value Pairs (see subchapter)

Example:

- Using the REMBI guidelines, which suggest specific Keys, see:
 - *Sarkans et al., Nature Methods*, doi: [10.1038/s41592-021-01166-8](https://doi.org/10.1038/s41592-021-01166-8).
 - The RDMbites (Elixir UK) YouTube series about REMBI, incl. a suggestion on how to use REMBI in OMERO:
<https://www.youtube.com/playlist?list=PLyCNTVs-UBvttFwfV6RuRKZH6G8NGlodA>

Practical metadata annotation in OMERO

- → Subchapter 07.1: Tag annotation
- → Subchapter 07.2: Key-Value Pair annotation
- → Subchapter 07.3: Ontology terms for metadata annotation