



NFDI4
BIOIMAGE

BioImage Data Handling and Analysis with OMERO

Workshop in 4 Modules

May 13th, 2024

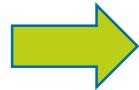
Trainers: Michele Bortolomeazzi, Riccardo Massei, Christian Schmidt

Support: Lena Krämer & Tom Boissonnet



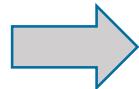
With the exception of logos and cited third-party content, the content
of these slides is shared under the terms of the [Creative Commons
Attribution License \(CC-BY 4.0\)](#) unless the content is marked otherwise.

Programme



Module 1 (9 am - 10.15 am);
Basics of OMERO, data structuring and annotation
(Christian)

Time to go to Lunch



Module 3 (1.45 pm - 3.45 pm);
OMERO and Jupyter Notebooks
(Riccardo)

Module 4 (4.15 pm - 6. pm);

Publication-ready figures and data with OMERO
(Christian , Riccardo , Michele)

Who are the trainers?



Dr. Riccardo Massei (Helmholtz-Center for Environmental Research, UFZ, Leipzig) -
Data Steward for Bioimaging Data in NFDI4BIOIMAGE

Dr. Michele Bortolomeazzi (DKFZ, Single cell Open Lab, bioimage data specialist,
bioinformatician, staff scientist in the NFDI4BIOIMAGE project)

Dr. Christian Schmidt (Science Manager for Research Data Management in
Bioimaging, DKFZ, Heidelberg, Project Coordinator of the NFDI4BIOIMAGE
project)

Introduction and expectations



- Join your nearest neighbors into groups of 3 – 5 people
- Find two things you have all in common
- Collect your expectations for this workshop (excalidraw)

5 min only!





NFDI4
BIOIMAGE

Introduction to OMERO and structured annotations

Workshop: **Bioimage data management and analysis with OMERO**

May 13th, 2024, Module 1

Trainers: Michele Bortolomeazzi, Riccardo Massei, **Christian Schmidt**



With the exception of logos and cited third-party content, the content
of these slides is shared under the terms of the [Creative Commons
Attribution License \(CC-BY 4.0\)](#) unless the content is marked otherwise.

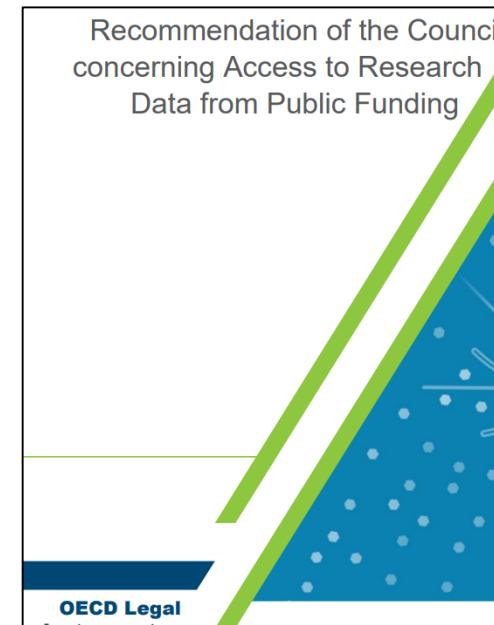
Significance of research data management



- **Intrinsic motivation**
 - Impact and efficiency of my research
- **Extrinsic factors**
 - Funding agency demands
 - Good Scientific Practice
- **Ethical aspect**
 - Make science more sustainable
 - Promote *Open Science* and data sharing

The screenshot shows the DFG (Deutsche Forschungsgemeinschaft) website. The top navigation bar includes links for Research Funding, Funded Projects, and DFG in Profile. Below the navigation, a breadcrumb trail shows Home > Funding > Principles of DFG Funding > Handling of Research Data. The main content area is titled "Handling of Research Data" and discusses the importance of appropriate handling of data based on subject-specific discipline. A sidebar on the right contains news items and a link to "Specification of Requirements Relating to the Handling of Research Data in Funding Proposals".

The screenshot shows the European Research Council (ERC) website. The header features the ERC logo and the text "European Research Council Scientific Council". Below the header, it says "Established by the European Commission". At the bottom, there is a link to "Open Research Data and Data Management Plans".



The screenshot shows a tweet from the BMBF (@BMBF_Bund) dated September 19, 2023. The tweet discusses the importance of making research data accessible and datenschutzkonform, mentioning Bettina Stark-Watzinger, the Federal Minister for Education and Research. It includes a quote in German: "„Daten sind der Rohstoff der Zukunft. Je mehr Daten etwa die Forschung zur Verfügung gestellt bekommt, umso besser kann sie arbeiten.“" Below the tweet is a snippet of code and a small photo of Bettina Stark-Watzinger.

- https://www.dfg.de/en/research_funding/principles_dfg_funding/research_data/index.html
- https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf
- https://twitter.com/BMBF_Bund/status/1571801906074337280?s=20&t=krDcwOPMuPwjs-VisYBgVg
- <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>



Reused from: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>, CC-BY 4.0

The FAIR principles for data and data stewardship



Findable

Accessible

Interoperable

Reusable

Box 2 | The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
- A1.1 the protocol is open, free, and universally implementable
- A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
- R1.1. (meta)data are released with a clear and accessible data usage license
- R1.2. (meta)data are associated with detailed provenance
- R1.3. (meta)data meet domain-relevant community standards

- Wilkinson et al., 2016, Scientific Data, DOI: 10.1038/sdata.2016.18, CC-BY 4.0 (<http://creativecommons.org/licenses/by/4.0>)
- Jacobsen et al., 2020, FAIR Principles: Interpretations and Implementation Considerations. Data Intelligence, DOI: 10.1162/dint_r_00024

Guiding principles

(not a dictation of technical solutions)

Goal is machine-readability

(not only human ability to reuse)

Guidelines for data producers and publishers

Applies to data and their provenance

FAIR is not

- the same as „open“
- a standard

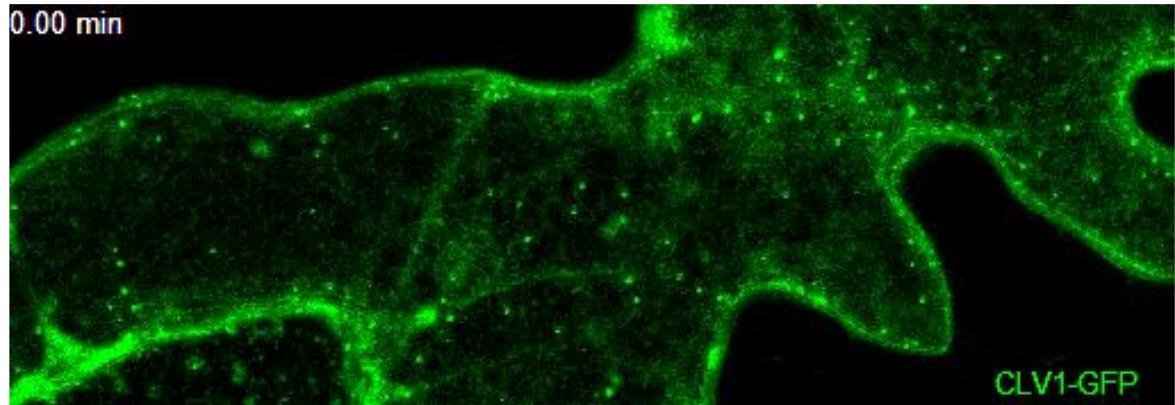


Reused from: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>, CC-BY 4.0

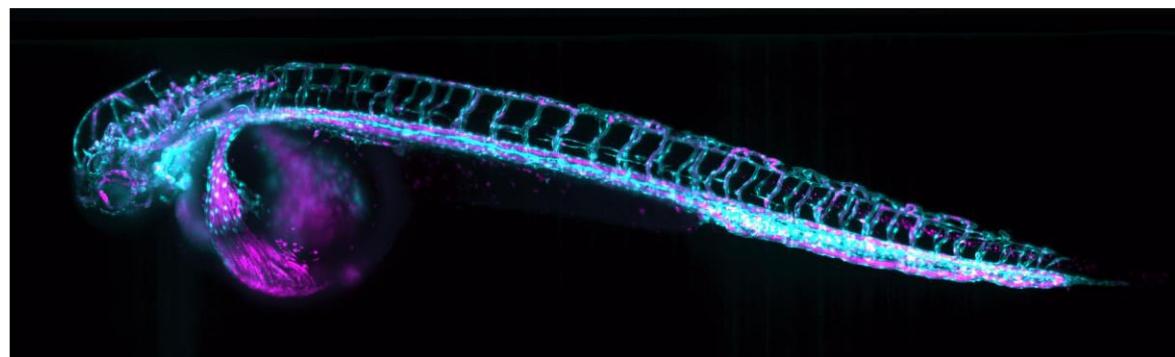
The bioimage data type

Microscopy data is (often):

- high-dimensional (X, Y, Z, Channel, Time, ...)
- saved in proprietary file formats
- of large file size
(often in GB-, sometimes in TB-range)
- produced with complex experimental setups
- used for quantitative analysis → derived data
- ... i.e. cumbersome to handle, store, and share?



Courtesy of: Y. Stahl, S. Weidtkamp-Peters, HHU Düsseldorf

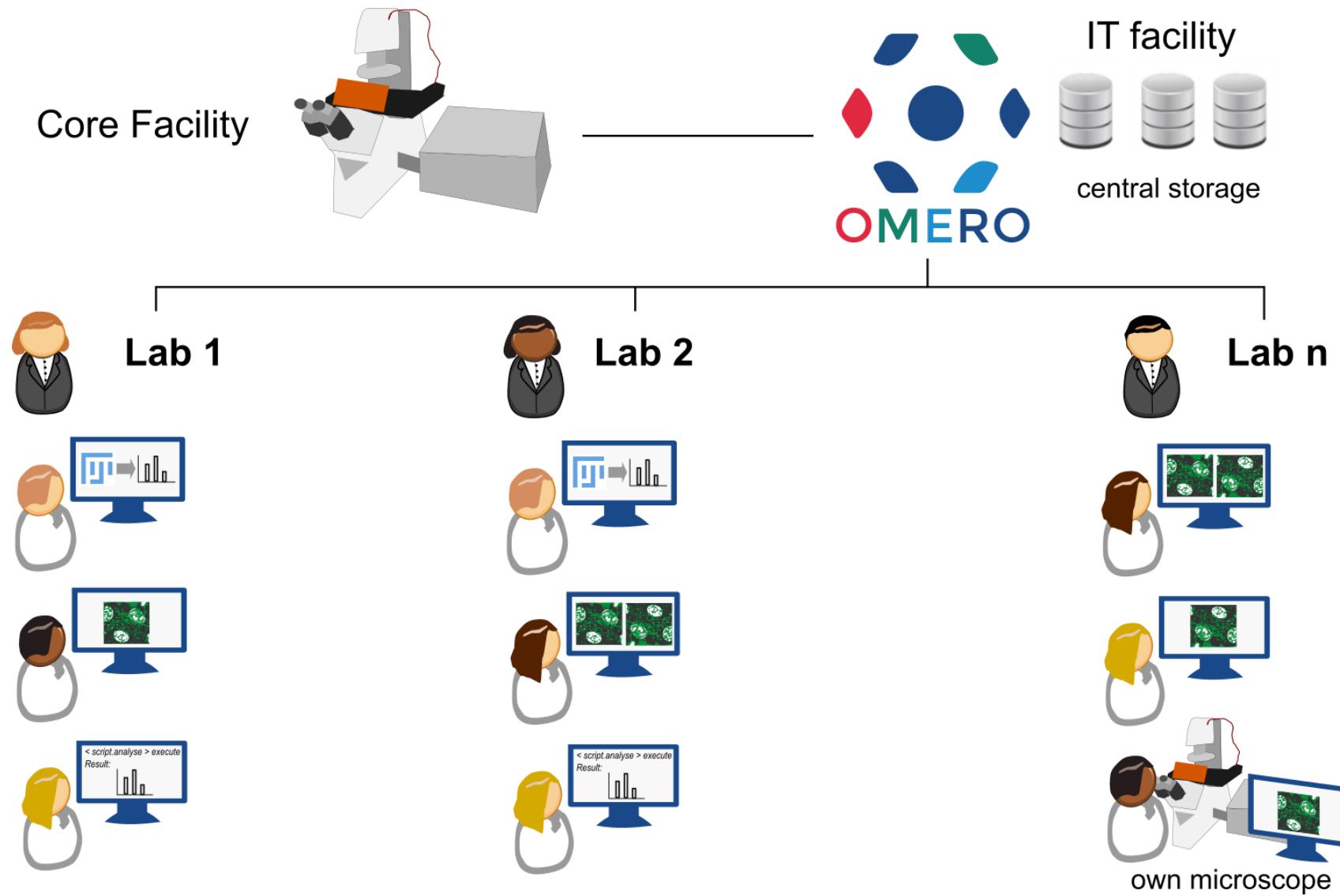


Courtesy of: Jan Huisken, University of Göttingen



Reused from: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>, CC-BY 4.0

OMERO – What is it?



Reused from: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>, CC-BY 4.0

OMERO – What is it?

OMERO: An open-source software for image data management

OMERO = „OME Remote Objects“

Created by the **Open Microscopy Environment Consortium (OME)**



- User computer
 - Microscope computer
 - Processing / analysis server
-
- Installed on an institute or central IT server
 - Storage and handling of imaging data
 - Accessible from outside via „clients“

OMERO.web – the main user interface

The screenshot illustrates the OMERO.web interface with several key features highlighted:

- User-Project-Dataset-Image file:** A red box highlights the navigation menu on the left.
- Figure preparation:** A red box highlights the thumbnail view of microscopy images.
- Preview of the images in the dataset:** A red box highlights the thumbnail view of microscopy images.
- Tags for flexible organization:** A red box highlights the tags section on the right.
- Key-Value pairs for metadata annotation:** A red box highlights the key-value pairs section on the right.
- Technical metadata:** A red box highlights the detailed technical metadata panel on the right.

Technical metadata (Right Panel):

Import Date:	2023-10-26 22:17:26
Dimensions (XY):	476 x 476
Pixels Type:	uint8
Pixels Size (XYZ) (µm):	0.09 x 0.09 x -
Z-sections/Timepoints:	1 x 1
Channels:	Alexa Fluor 647, Alexa Fluor 568, Alexa Fluor 488, DAPI
ROI Count:	0
Tags	0
Key-Value Pairs	0

Added on Project Microscope_Training

Study	master course
study type	master course
study name	CAI-Module_WS-2023/24
study description	practical microscopy course for master students

Added on Dataset 2023-10-25_LSM780

Study component	confocal laser scanning microscopy
imaging method	confocal laser scanning microscopy
imaging method term accession number	http://purl.obolibrary.org/obo/CHM_O_0000089
imaging method term accession number source REF	chemical methods ontology
Biosample	HeLa cells
biological entity	HeLa cells
biological entity term accession number	http://www.ebi.ac.uk/efo/EFO_0001185
biological entity term accession number source REF	Experimental Factor Ontology
species	human
species term accession number	http://purl.obolibrary.org/obo/NCBI_Taxon_9606
species term accession number source REF	NCBI Taxonomy

Demonstration and Exercise

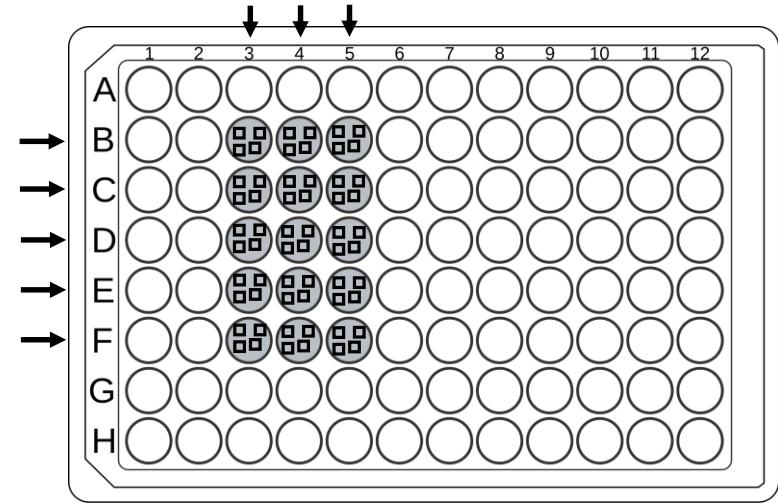


Make yourself familiar with OMERO.web



Example Data for this workshop

- 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.
- Two independent samples are tested (with the sample IDs: 18746 and 22123)



How is this structured correctly?



- What should our "dataset unit" be ?
 - the whole week, both plates with all images?
 - One whole plate, one dataset per week ?
 - each compound of every week ?
 - each compound of a single week ?
 - ...
- What's a convenient way to describe the data content ?

A little analogy (not so serious)

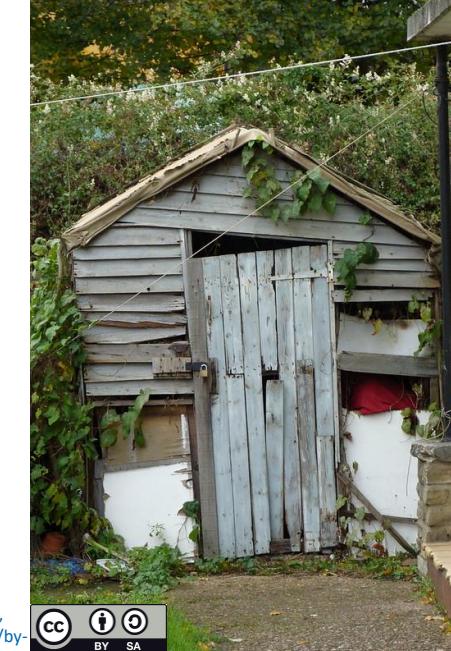
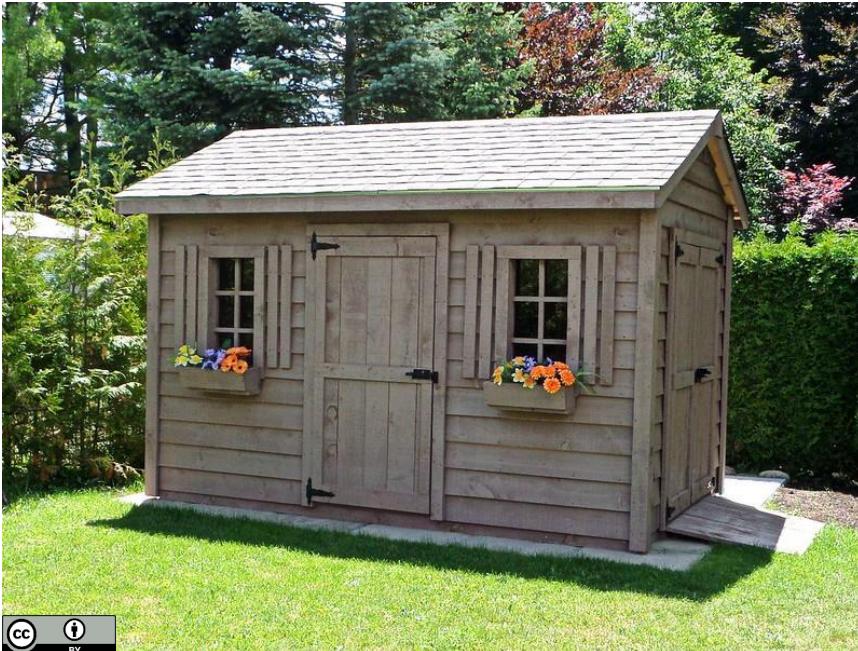
datasets



images



annotations



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



Reused from Tom Boissonnet, available at: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>, CC-BY 4.0

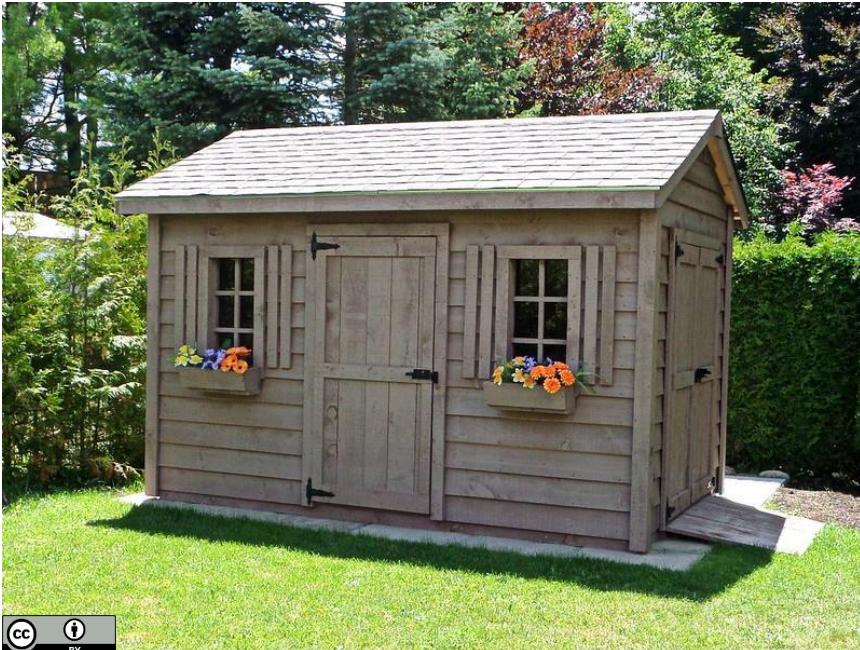
A little analogy (not so serious)



datasets

Tags	Key-value
10 µM -	compound A
week 1 -	concentration 10 µM
compound A -	week 1

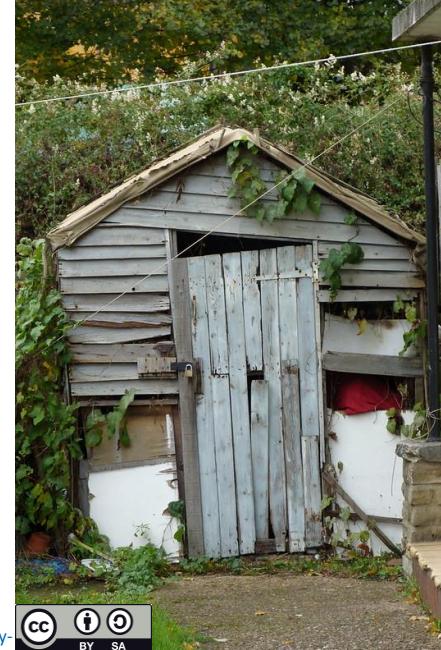
Week1_150607_B02_s1_c1-DAPI_compoundA_10µM.tif



By Tom Rossini,
CC-BY



Reused from Tom Boissonnet, available at: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>, CC-BY 4.0

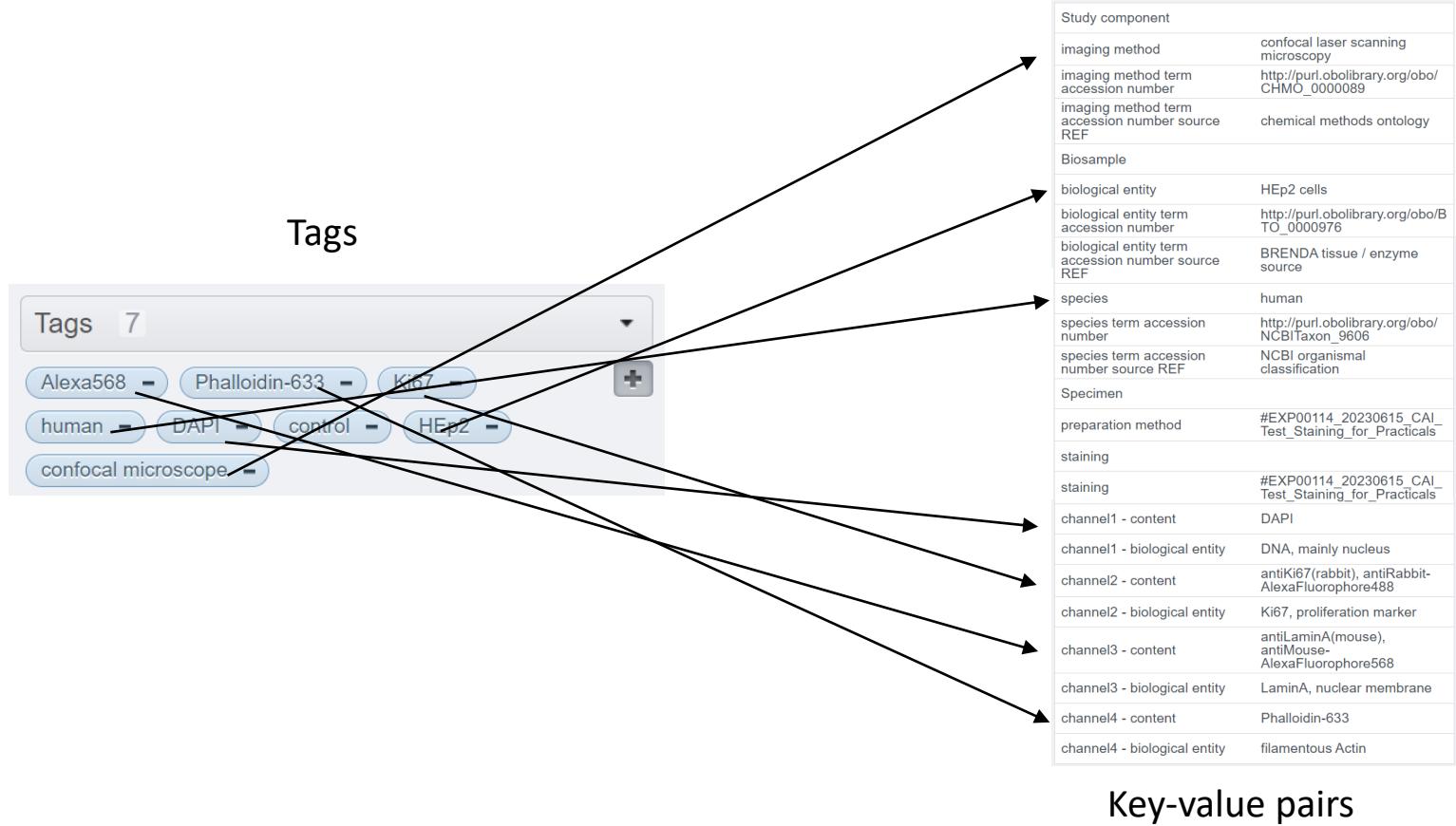


By Gordon Jolly,
<https://flickr.com/photos/loopzilla/8132242517/>,
CC-BY-SA, <https://creativecommons.org/licenses/by-sa/2.0/>



Redundant concepts?

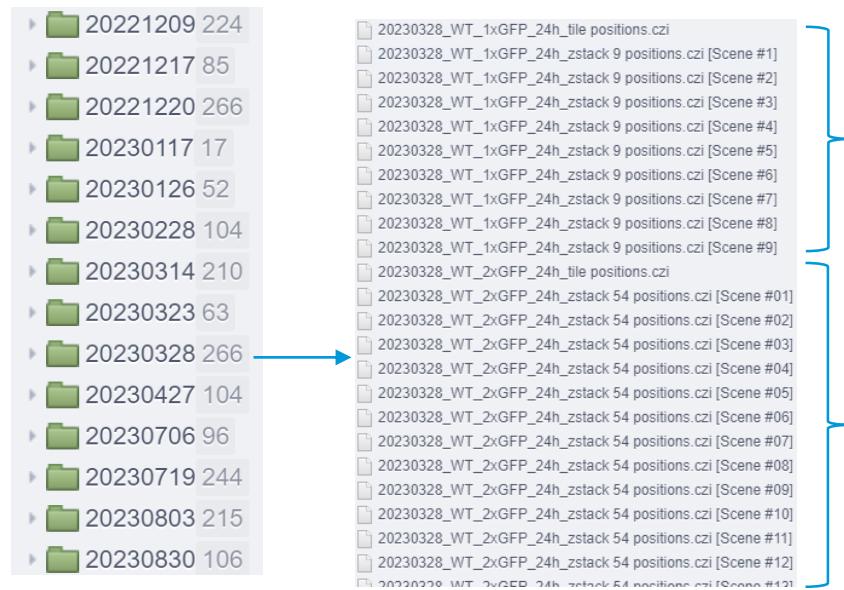
(REMBI checklist)
→ Module 4



Folder hierarchies are what we are used to



- 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?



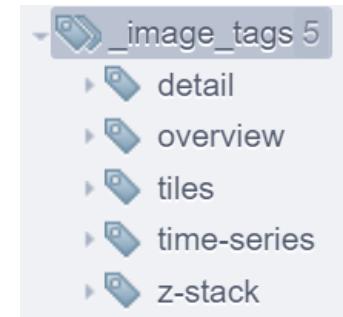
Tags are like folders, but better

- 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



Are datasets even needed?



- 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 ?



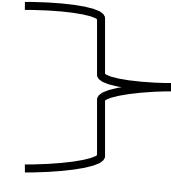
Reused from Tom Boissonnet, available at: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>. CC-BY 4.0

Advantage of metadata annotation



- Datasets can be annotated and used as structure in analysis!

- Tags
- Key-value pairs
- Attachments
- Ratings
- Comments



These will be captured in the Structured Annotation metadata in the OME.XML header of an OME-TIFF

- By annotating a dataset, I can implicitly annotate the images it contains
 - group images of a same experiment in the dataset
-> less duplication of annotations
 - images can be annotated to give more details



Adapted after Tom Boissonnet, available at: Fuchs, V. A. F., Schmidt, C., & Boissonnet, T. (2024, Mai 6). [Workshop] FAIR data handling for microscopy: Structured metadata annotation in OMERO. Zenodo. <https://doi.org/10.5281/zenodo.11109616>, CC-BY 4.0

Structured metadata

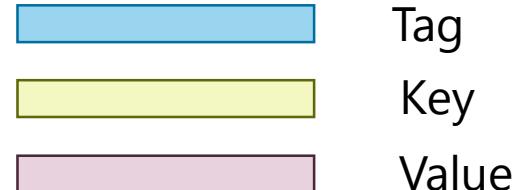
```

<structuredAnnotations>
    <TagAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbd-769b-474a-8e7b-93bcd175df65_47706:2998580">
        <Description>tim of 5 µg/ml anti-CD3/anti-CD28 stimulation before staining (in h)</Description>
        <Value>0 h anti-CD3/CD38</Value>
    </TagAnnotation>

    <TagAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbd-769b-474a-8e7b-93bcd175df65_19802:622773">
        <Description>Experiments with staining for phosphorylated ERK protein</Description>
        <Value>phospho-ERK</Value>
    </TagAnnotation>

    <TagAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbd-769b-474a-8e7b-93bcd175df65_19545:561233">
        <Value>DMSO control</Value>
    </TagAnnotation>

    <MapAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbd-769b-474a-8e7b-93bcd175df65_46089:1280547"
Namespace="openmicroscopy.org/omero/client/mapAnnotation">
        <Value>
            <M K="Organism">Mus musculus</M>
            <M K="Strain">C57BL/6</M>
            <M K="Organ">Spleen</M>
            <M K="Cell Type">CD4-positive, alpha-beta T cell</M>
            <M K="Manipulation 1">unstimulated</M>
            <M K="Treatment 1">0.3% DMSO (control)</M>
            <M K="Treatment 1 time">0 h</M>
            <M K="Target 1">phospho-ERK</M>
            <M K="Staining Method">immunofluorescence indirect</M>
            <M K="Imaging Method">CLSM</M>
            <M K="Lab Notebook Ref">CSchm#4 page 112</M>
        </Value>
    </MapAnnotation>
</structuredAnnotations></OME>
```



Visualizing the OME.XML metadata in Fiji



```
<StructuredAnnotations>
  <TagAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbcd-769b-474a-8e7b-93bcd175df65_47706:2998580">
    <Description>tim of 5 µg/ml anti-CD3/anti-CD28 stimulation before staining (in h)</Description>
    <Value>0 h anti-CD3/CD38</Value>
  </TagAnnotation>
  <TagAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbcd-769b-474a-8e7b-93bcd175df65_19802:622773">
    <Description>Experiments with staining for phosphorylated ERK protein</Description>
    <Value>phospho-ERK</Value>
  </TagAnnotation>
  <TagAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbcd-769b-474a-8e7b-93bcd175df65_19545:561233">
    <Value>DMSO control</Value>
  </TagAnnotation>
  <MapAnnotation ID="urn:lsid:export.openmicroscopy.org:Annotation:a99ddfbcd-769b-474a-8e7b-93bcd175df65_46089:1280547" Namespace="openmicroscopy.org/omero/client/mapAnnotation">
    <Value>
      <M K="Organism">Mus musculus</M>
      <M K="Strain">C57BL/6</M>
      <M K="Organ">Spleen</M>
      <M K="Cell Type">CD4-positive, alpha-beta T cell</M>
      <M K="Manipulation 1">unstimulated</M>
      <M K="Treatment 1">0.3% DMSO (control)</M>
      <M K="Treatment 1 time">0 h</M>
      <M K="Target 1">phospho-ERK</M>
      <M K="Staining Method">immunofluorescence indirect</M>
      <M K="Imaging Method">CLSM</M>
      <M K="Lab Notebook Ref">CSchm#4 page 112</M>
    </Value>
  </MapAnnotation>
```



Exercise – Bring structure to the data



Enrich the training dataset (structured according to your preference) with Tags (and KV pairs)



Acknowledgments



In cooperation with

Information Infrastructure for BioImage Data (I3D:bio)

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



German Cancer Research Center (DKFZ), Heidelberg
Department Enabling Technology

Dr. Christian Schmidt, Project Coordinator

office@nfdi4bioimage.de

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the National Research Data Infrastructure – NFDI 46/1 – 501864659



In cooperation with



The NFDI4BIOIMAGE consortium comprises legally independent partners and does not act autonomously towards third parties. The authors represent the contributions from their respective affiliated institutions and work together for the project.