Enabling large, N-dimensional array data from imaging in EOSC with OME-Zarr

Christian Schmidt,
0000-0001-9568-895X
Enabling Technology Department,
German Cancer Research Center (DKFZ),
Heidelberg, Germany

Aastha Mathur,
0000-0001-9734-9767
Euro-BioImaging ERIC, Bio-Hub
European Molecular Biology Laboratory,
Heidelberg, Germany

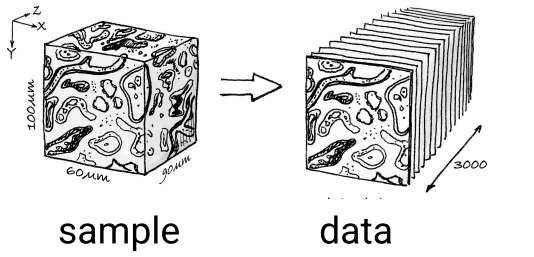
*Josh Moore**
0000-0003-4028-811X
German BioImaging e.V
Society for Microscopy and Image Analysis
Konstanz, Germany

** contact:* ***joshua.moore@gerbi-gmb.de***

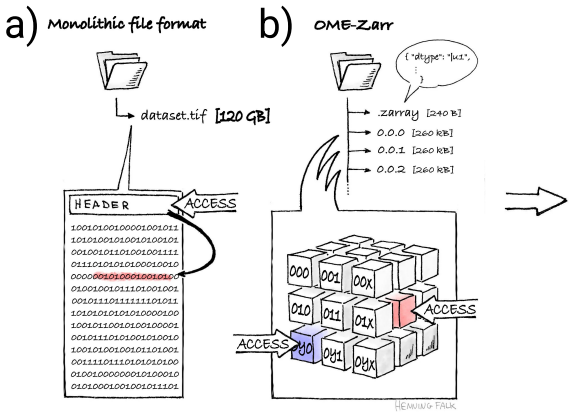
Advanced imaging modalities are used across scientific domains which create large, multidimensional, information-rich datasets. Being able to exploit such data in full via cross-dimensional data processing and analysis demands critical file format and data structure capabilities, pivotal to enabling efficient collaboration, sharing and reuse.

OME-Zarr: a specification for bioimaging data

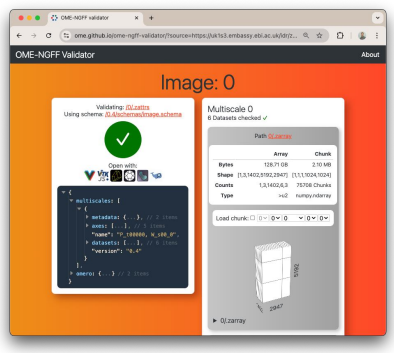
Example:
Volume Electron Microscopy



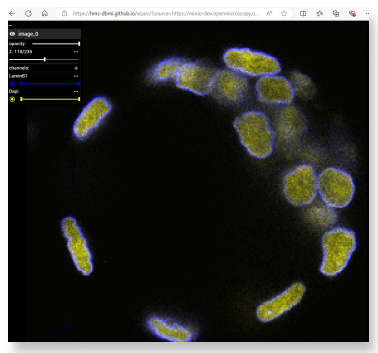
Data structure



Suite of tools



Validation, ...



Visualization, etc...

The technology: Zarr - a cross-domain solution for large array-typed data

bioimaging and medical imaging

geospatial

data science

Zarr

distributed and **cloud** storage

Chunked structure for efficient I/O in parallel processing

built-in **compression**

Moore & Kunis, CoRDI 2023

https://zarr.dev/

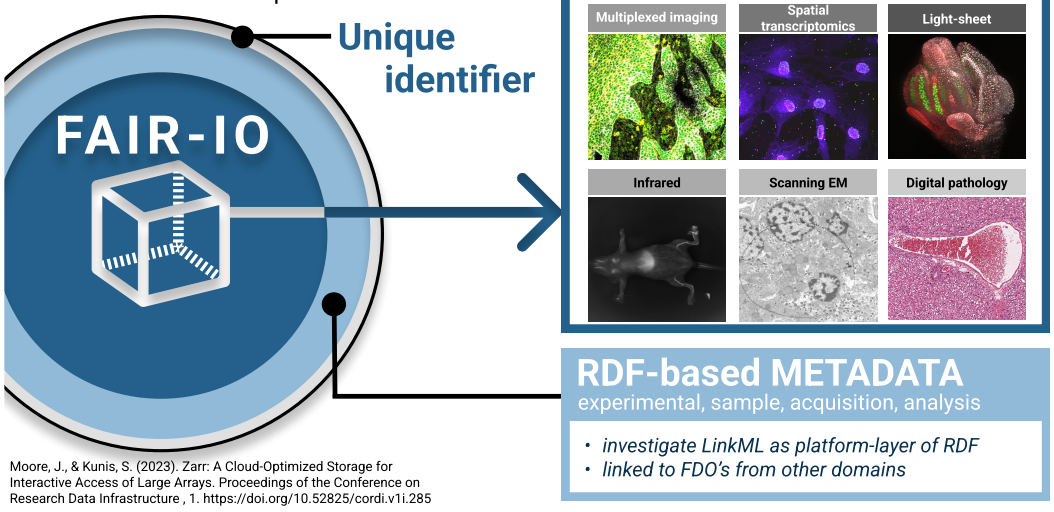
Wide-spread adoption and community-approved metadata support:

A common file format becomes a pillar of FAIR imaging data in Europe and beyond

NFDI4 BIOIMAGE

NFDI4BIOMAGE is a consortium within Germany's National Research Data Infrastructure (NFDI) framework, focusing on FAIR research data management for microscopy and image analysis.

FAIR Image Objects:
an FDO implementation

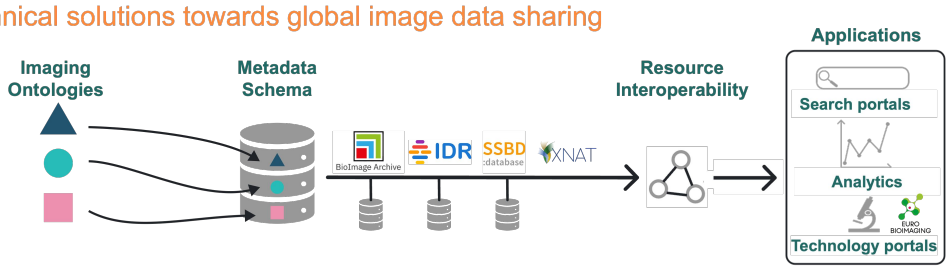


Moore, J., & Kunis, S. (2023). Zarr: A Cloud-Optimized Storage for Interactive Access of Large Arrays. Proceedings of the Conference on Research Data Infrastructure , 1. https://doi.org/10.52825/cordi.v1i.285

foundingGIDE

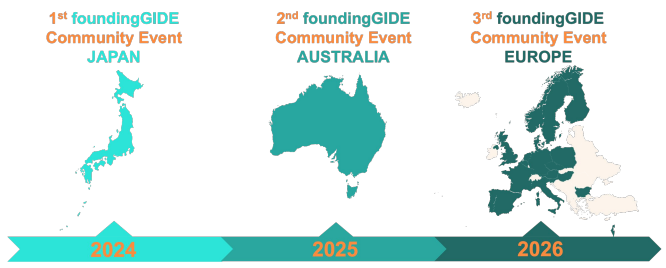
Founding a Global Image Data Ecosystem (GIDE) for bioimage data exchange based on global coordination of technical developments among data infrastructures and communities

Technical solutions towards global image data sharing



Engagement with global community

- Global coordination events
- Technical events for common Ontologies and Metadata
- Interaction with image data communities through liaisons
- Training on FAIR image data management practices
- Tools for metadata recording and image search



https://nfdi4bioimage.de
Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the National Research Data Infrastructure –**NFDI 46/1 – 501864659**

https://founding-gide.eurobioimaging.eu
FoundingGIDE is coordinated by Euro-BioImaging ERIC and has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement number **101130216**

References:

Moore J, Allan C, Besson S, et al (2021) OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies. Nat Methods 18:1496–1498. https://doi.org/10.1038/s41592-021-01326-w

Falk, H. zarr-developers/zarr-illustrations-falk-2022 | Zenodo [WWW Document], 2022. URL https://doi.org/10.5281/zenodo.7037367 (accessed 8.31.22).