Image from idr0100, Capar et al.

OME-NGFF: cloud-optimized format

CRG Barcelona 2023 🍌

OME Team

HORIZ N 2020























What is OME-NGFF ?

Brief Communication Open Access Published: 29 November 2021

OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies

Josh Moore, Chris Allan, Sébastien Besson, Jean-Marie Burel, Erin Diel, David Gault, Kevin Kozlowski, Dominik Lindner, Melissa Linkert, Trevor Manz, Will Moore, Constantin Pape, Christian Tischer & Jason R. Swedlow 🖂

Nature Methods 18, 1496–1498 (2021) Cite this article

6045 Accesses | 4 Citations | 80 Altmetric | Metrics

https://ngff.openmicroscopy.org https://ngff.openmicroscopy.org/data

NEUBIAS 2023

The Image Data Resource (IDR)

- Public access
- **Reference datasets** complete datasets containing molecular and functional annotations, associated with an existing or upcoming publication.
- **Study integration** integrating studies or datasets with other datasets via **genes**, **compounds** or **phenotypes**.
- Curated metadata
- Cloud re-analysis

OME-NGFF uses Zarr



Binary TIFF files contain a metadata block and a linear sequence of 2D tiles of varying size. Binary HDF5 files contain a hierarchy of ND arrays with associated metadata. NGFFs folders store the same information as HDF5s spread across many, smaller files. NEUBIAS workshop

Zarr



chunk = file = object

Remote cloud storage



NEUBIAS 2023

Access possibilities of images

image from idr0044, McDole et al.	Download	IDR API Access	OME-Zarr Access via S3
Load image subregion, e.g., single chunk or tile	<mark>No</mark> , only per file	Yes	Yes
Lazy loading	No	No	Yes. Use Dask collections: da.from_zarr(endpoi nt_url)
Easily analyze in parallel	No, depends on file format which may require a translation library.	Difficult due to the transfer protocol used (zeroc-ice)	Yes. Use Dask schedulers: dask.delayed(analyz e)(t, c, z)

Adapted from Table 3, Josh Moore et. al.: OME-Zarr: a cloud-optimized bioimaging file format with international community support, bioRxiv 2023.02.17.528834



Thank you









