


# ome-ngff: scalable format strategies for interoperable bioimaging data

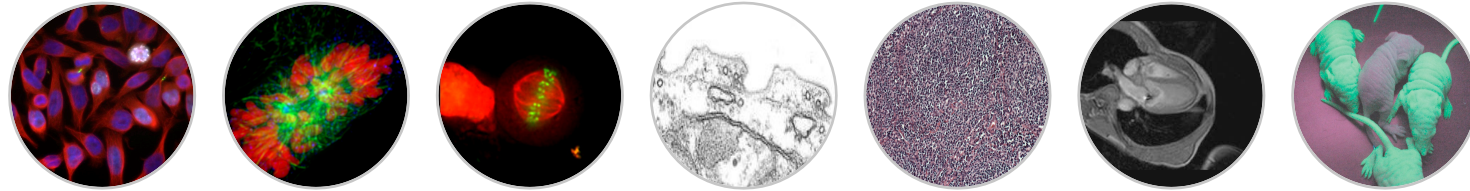
<https://j.mp/ome-ngff-biorxiv> (10.1101/2021.03.31.437929)



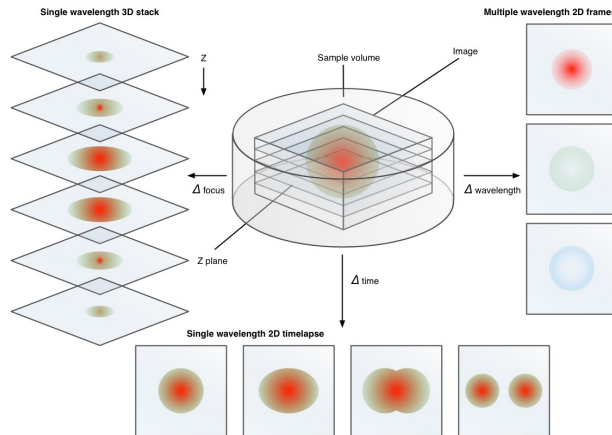
 **OME**'s mission is to facilitate the storage,  
exchange and re-use of bioimaging data.

# Bioimaging Domain

Imaging Modalities



5D Image Data Model



3D, multi-color, movies, or any combination thereof

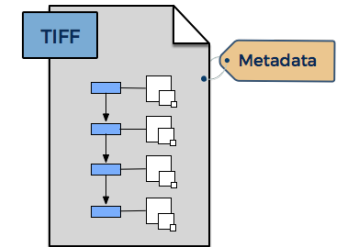
No Standard Format



150+ Proprietary File Formats (PFFs)

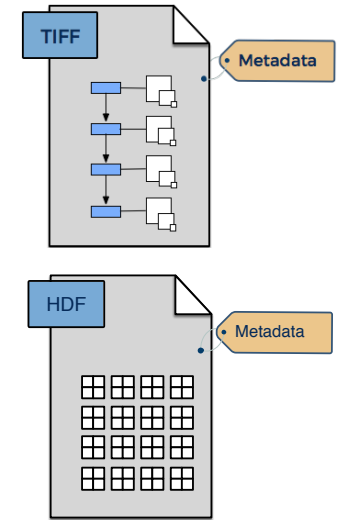
# Search for a format

	Maturity	Dimensionality	Scalability
TIFF	Ubiquitous	2D tiles	Linear slowdown



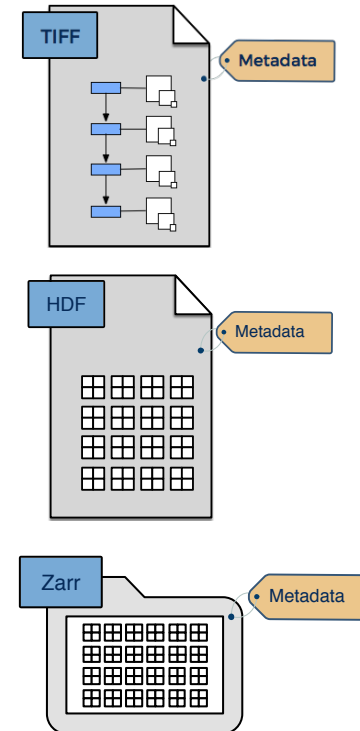
# Search for a format

	Maturity	Dimensionality	Scalability
TIFF	Ubiquitous	2D tiles	Linear slowdown
HDF5	Well-supported	3D+ chunks	Serial writes *



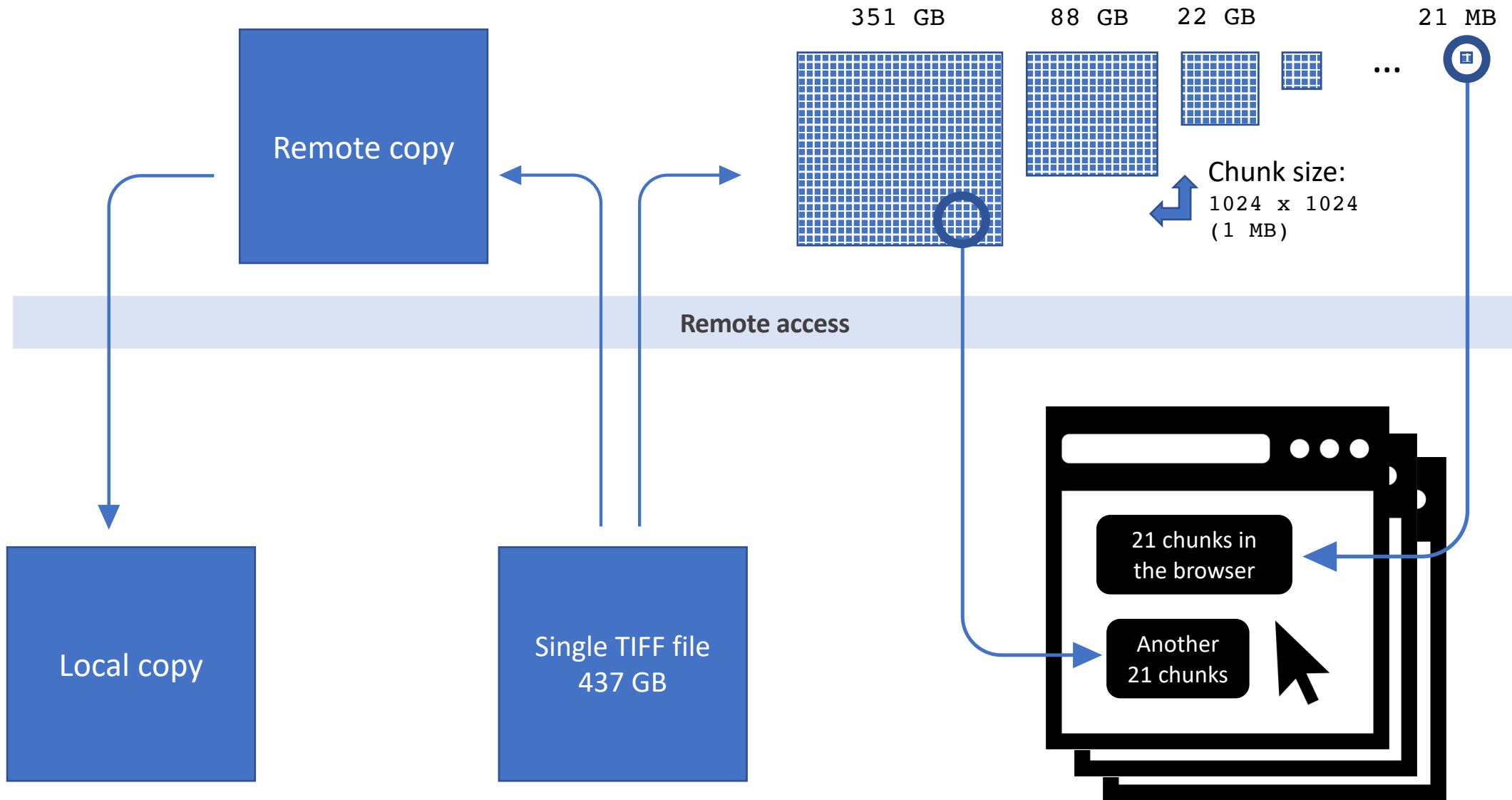
# Search for a format

	Maturity	Dimensionality	Scalability
TIFF	Ubiquitous	2D tiles	Linear slowdown
HDF5	Well-supported	3D+ chunks	Serial writes *
Zarr/N5	Next-generation	3D+ chunks	Cloud-native

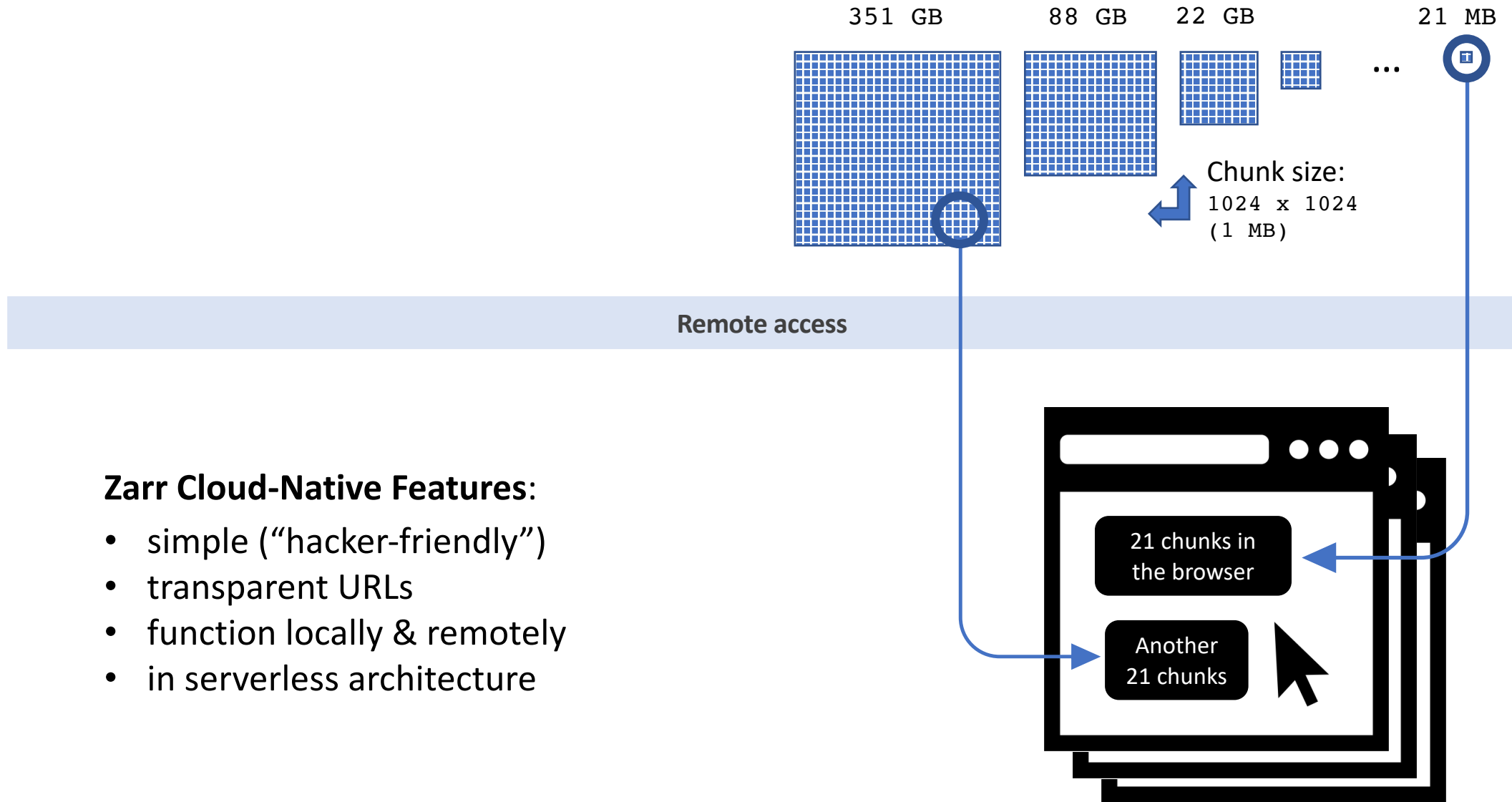


# Monolithic format

# Next-generation file format (NGFF)

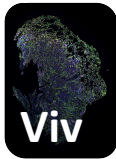


# Next-generation file format (NGFF)

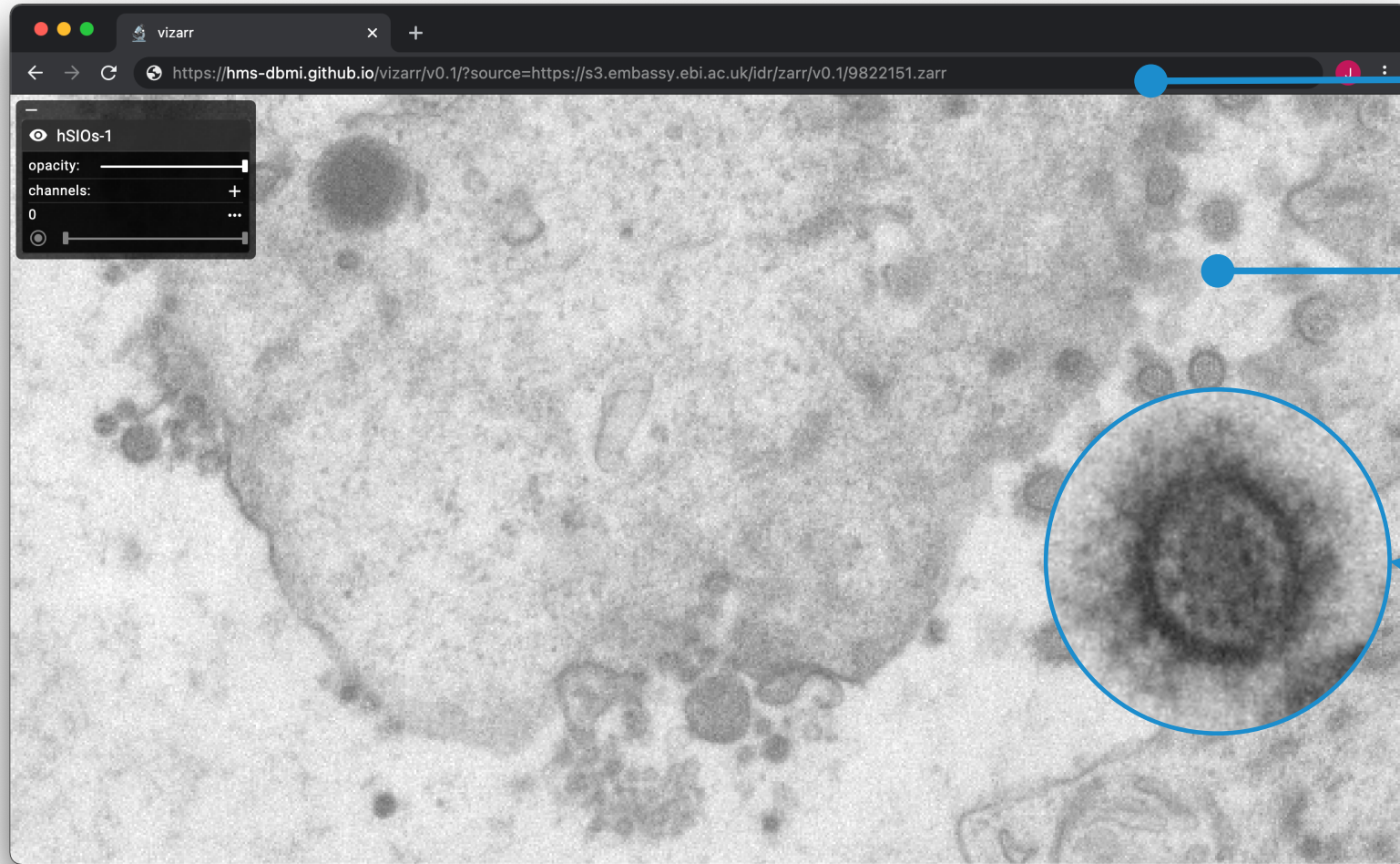


## Zarr Cloud-Native Features:

- simple (“hacker-friendly”)
- transparent URLs
- function locally & remotely
- in serverless architecture



Gehlenborg  
Lab (HMS)

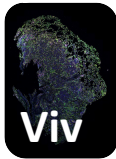


**Data:** S3 in EBI's  
Embassy Cloud

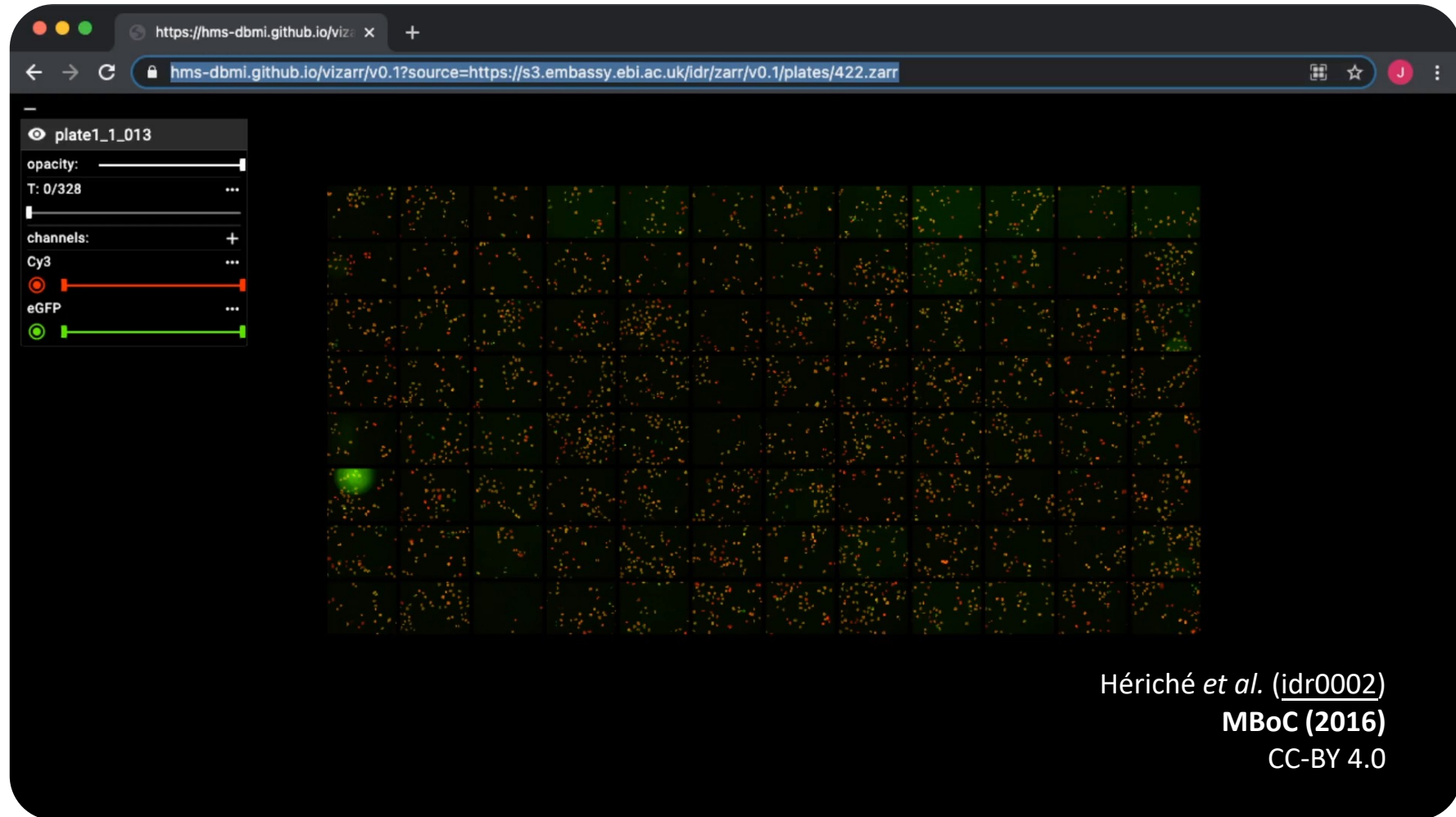
Lamers *et al.* ([idr0083](#))  
**Science (2020)**  
CC BY 4.0

30 GB from EBI's S3 to your browser  
<https://bit.ly/ngff-covid-zarr>  
(<http://bit.ly/ngff-covid-source>)

Highest resolution image  
of SARS-CoV-2 in May



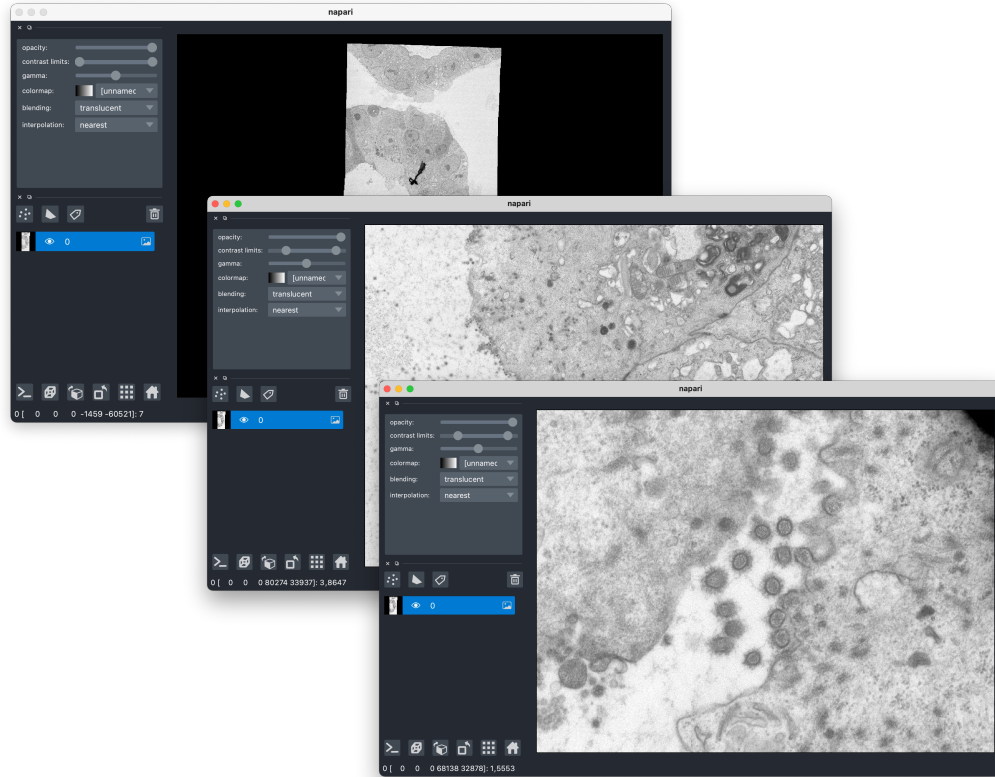
Gehlenborg  
Lab (HMS)



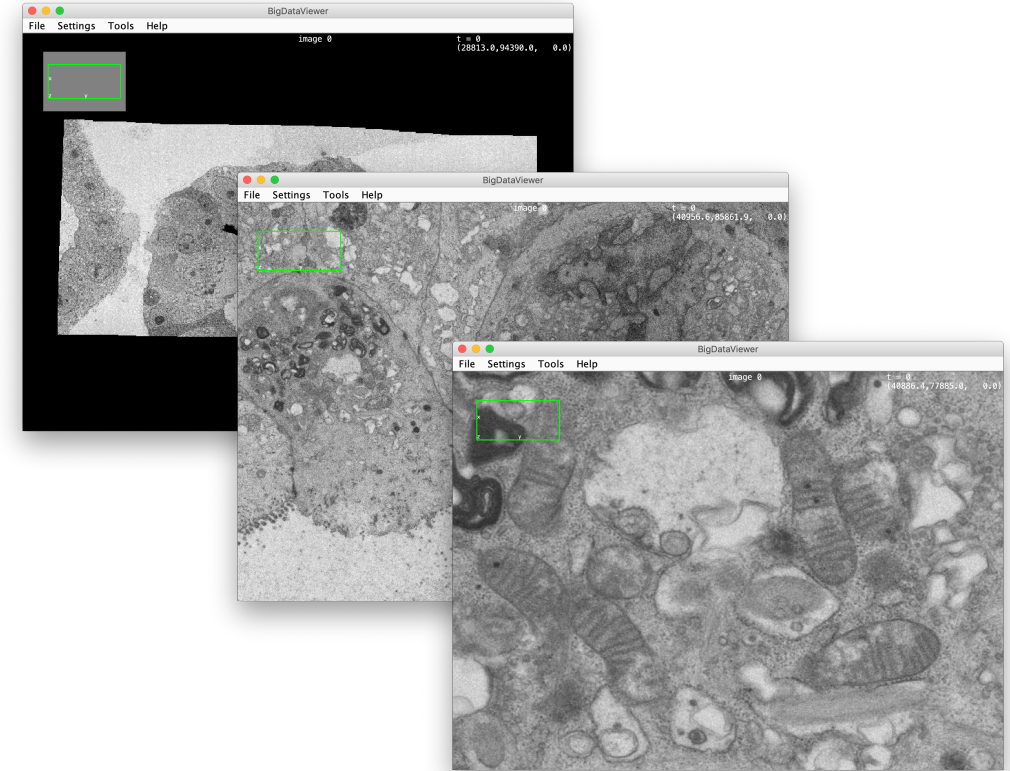
Hériché *et al.* ([idr0002](#))  
**MBoC (2016)**  
CC-BY 4.0

<http://bit.ly/idr0002-vizarr>

# Cross-platform

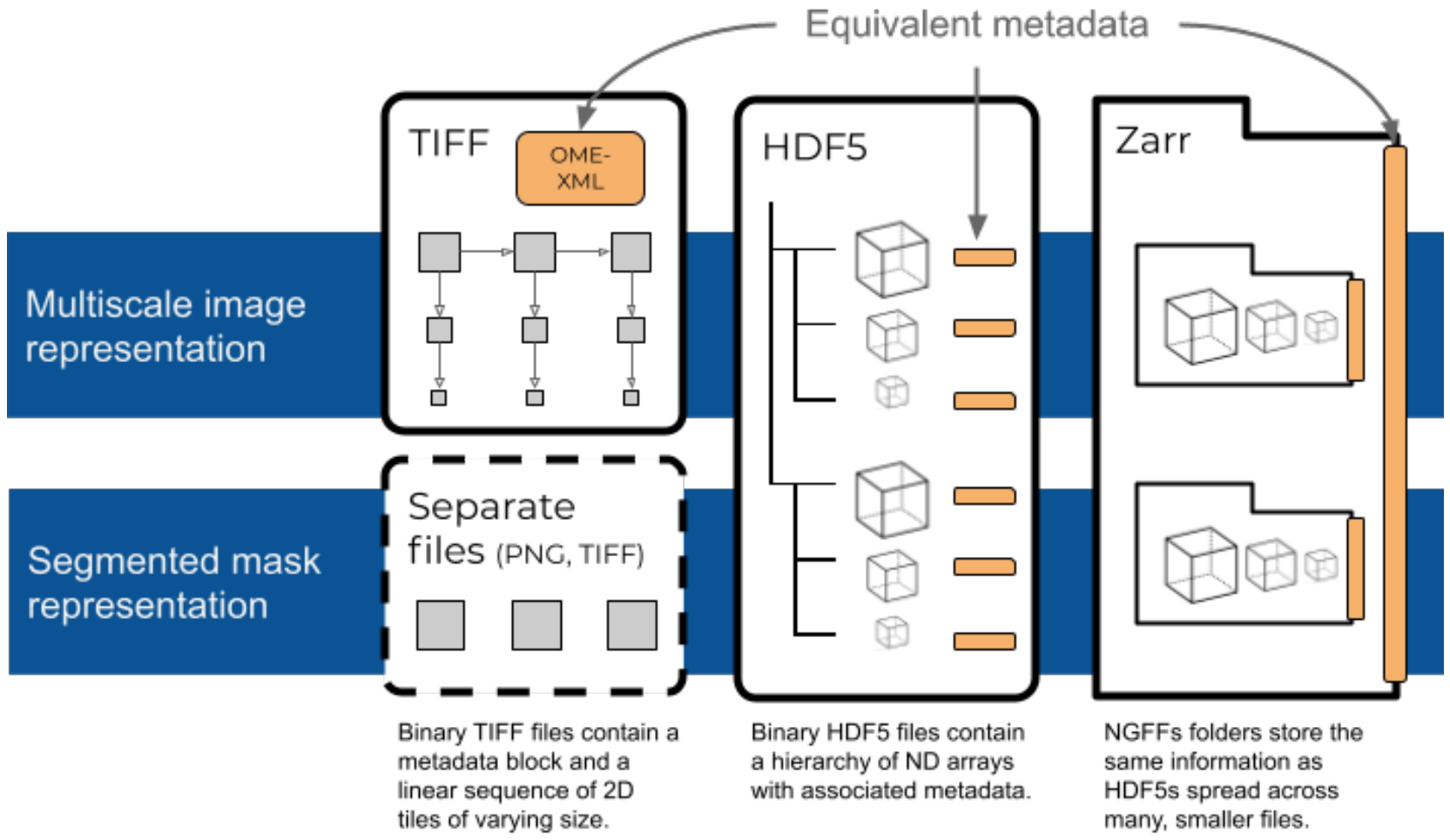


**ome-zarr-py plugin**  
Will Moore, U. Dundee *et al.*



**MOBIE Plugin**  
Christian Tischer, EMBL Heidelberg





<http://j.mp/ome-ngff-biorxiv> (Fig 4.)

```
$ bioformats2raw retina_large.ims retina_large.ome.zarr
```

```
$ find retina_large.ome.zarr -name ".z*" -o -name 0.0.0.0.0
```

```
retina_large.ome.zarr/.zattrs  
retina_large.ome.zarr/.zgroup
```

```
retina_large.ome.zarr/0/.zarray  
retina_large.ome.zarr/0/0.0.0.0.0
```

...

```
retina_large.ome.zarr/1/.zarray  
retina_large.ome.zarr/1/0.0.0.0.0
```

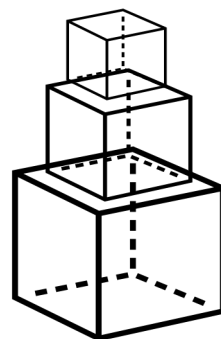
...

```
retina_large.ome.zarr/2/.zarray  
retina_large.ome.zarr/2/0.0.0.0.0
```

...

```
retina_large.ome.zarr/3/.zarray  
retina_large.ome.zarr/3/0.0.0.0.0
```

...



**Zarr Group: *group\_name***

.zgroup

.zattrs

**Zarr Array: *array\_name***

.zarray

.zattrs

0.0

0.1

1.0

1.1

2.0

2.1

```

intake/fsspec-reference-maker x +
github.com/intake/fsspec-reference-maker
README.md
Version 1
Metadata structure in JSON. We note, for future possible binary storage, that "version",
"gen" and "templates" should be considered attributes, and "refs" as the data that
ought to dominate the storage size. The previous definition, Version 0, is compatible
with the "refs" entry, but here we add features. It will also be possible to expand this
new enhanced spec into Version 0 format.
{
  "version": (required, must be equal to) 1,
  "templates": (optional, zero or more arbitrary keys) {
    "template_name": jinja-str
  },
  "gen": (optional, zero or more items) [
    "key": (required) jinja-str,
    "url": (required) jinja-str,
    "offset": (optional, required with "length") jinja-str,
    "length": (optional, required with "offset") jinja-str,
    "dimensions": (required, one or more arbitrary keys) {
      "variable_name": (required)
      {"start": (optional) int, "stop": (required) int, "step": (optional) int
      OR
      [int, ...]
    }
  ],
  "refs": (optional, zero or more arbitrary keys) {
    "key_name": (required) str OR [url(jinja-str)] OR [url(jinja-str), offset(in
  }
}

```

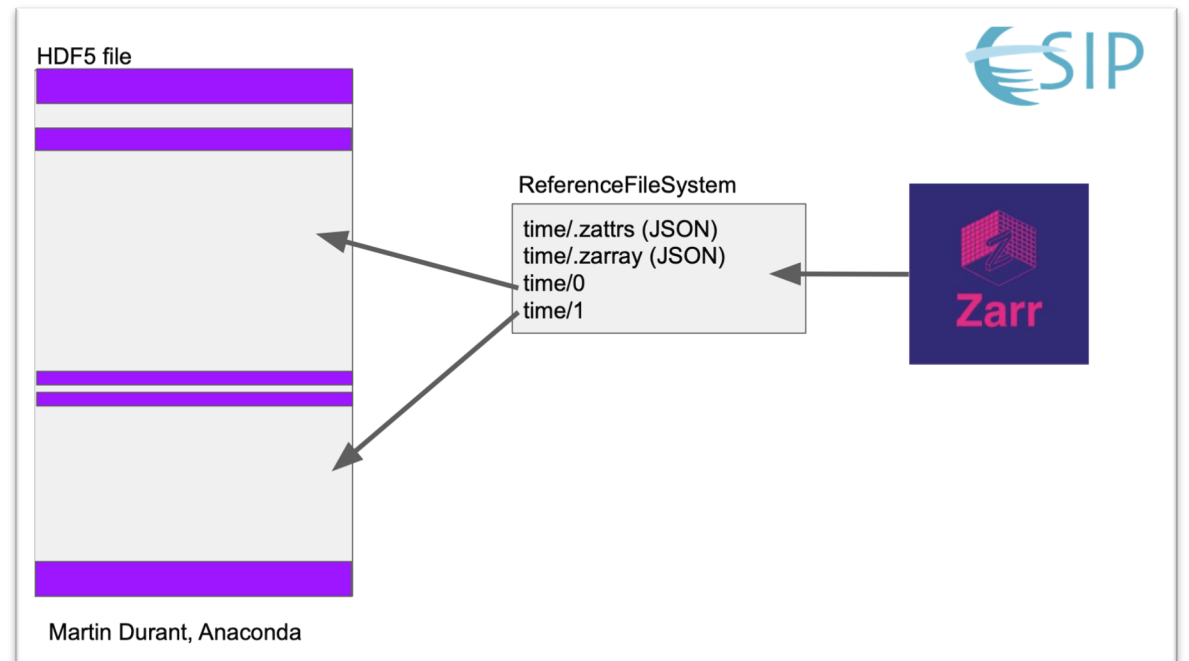


# intake/fsspec-reference-maker

```

with fsspec.open(INPUT,) as i:
  with fsspec.open(OUTPUT, mode="w") as o:
    h5 = SingleHdf5ToZarr(i, INPUT)
    o.write(json.dumps(h5.translate()))

```



```

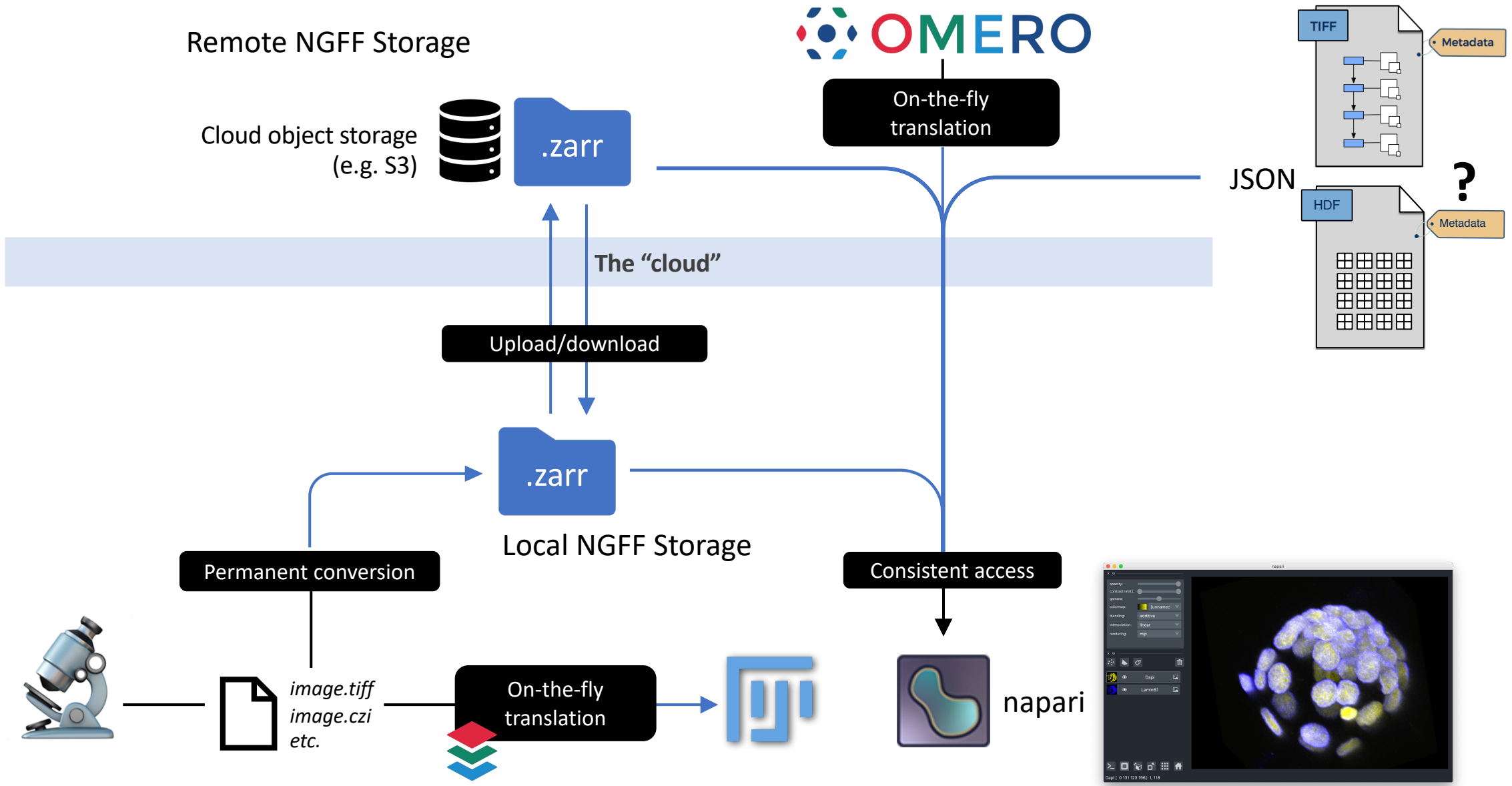
{
  "version": 1,
  "templates": {
    "u": "retina_large.ims"
  },
  "refs": {
    ".zgroup": "{\n  \"zarr_format\": 2\n}",
    ".zattrs": "{\n  \"DataSetDirectoryName\": [\n    \"D\", \n    \"a\", \n    \"t\", \n    ], \n  \"DataSetInfoDirectoryName\": [\n    \"D\", \n    \"a\", \n    \"t\", \n    ], \n  \"I\", \n    \"n\", \n    \"f\", \n    \"o\", \n    ], \n  \"ImarisDataSet\": [\n    \"a\", \n    \"r\", \n    \"i\", \n    \"s\", \n    \"D\", \n    \"a\", \n    \"t\", \n    ], \n  \"ImarisVersion\": [\n    \"5\", \n    \".\", \n    \"5\", \n    ], \n  \"ThumbnailDirectoryName\": [\n    \"T\", \n    \"h\", \n    \"u\", \n    \"m\", \n    \"a\", \n    \"i\", \n    \"l\", \n    ]\n}",
    "DataSet/.zgroup": "{\n  \"zarr_format\": 2\n}",
    "DataSet/ResolutionLevel 0/.zgroup": "{\n  \"zarr_format\": 2\n}",
    "DataSet/ResolutionLevel 0/TimePoint 0/.zgroup": "{\n  \"zarr_format\": 2\n}",
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/.zgroup": "{\n  \"zarr_format\": 2\n}",
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/.zattrs": "{\n  \"HistogramMax\": [\n    \"2\", \n    \".\", \n    \"0\", \n    ], \n  \"HistogramMin\": [\n    \"0\", \n    \".\", \n    \"0\", \n    ], \n  \"ImageSizeX\": [\n    \"2\", \n    \"4\", \n    \"8\", \n    \"1\", \n    \"5\", \n    \"6\", \n    \"7\", \n    ], \n  \"ImageSizeZ\": [\n    \"6\", \n    ], \n  \"DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/.zarray": "{\n  \"chunks\": [\n    16, \n    ], \n  \"id\": \"zlib\", \n  \"level\": 2, \n  \"dtype\": \"lu1\", \n  \"fill_value\": \"C\", \n  \"shape\": [\n    64, \n    1792, \n    2048, \n    ], \n  \"zarr_format\": 2\n}",
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/.zattrs": "{\n  \"_ARRAY_DIMENSIONS\": [\n    ], \n  \"DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0\": [...],
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0.1": [...],
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0.2": [...],
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0.3": [...],
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0.4": [
      "{u}",
      339421,
      39879
    ],
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0.5": [...],
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0.6": [...],
    "DataSet/ResolutionLevel 0/TimePoint 0/Channel 0/Data/0.0.7": [...],
  },
}

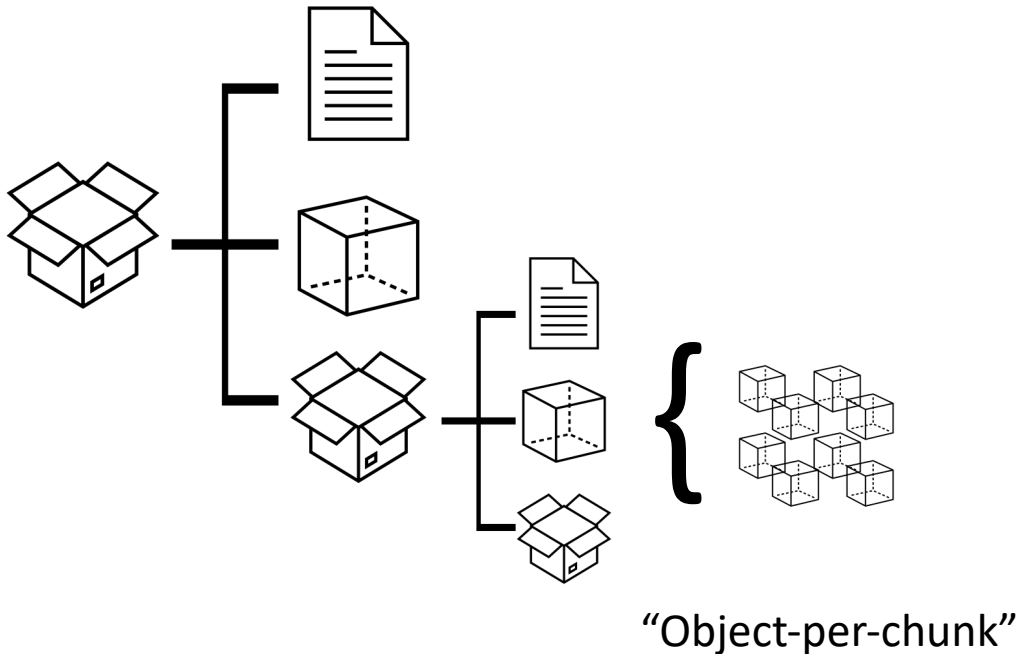
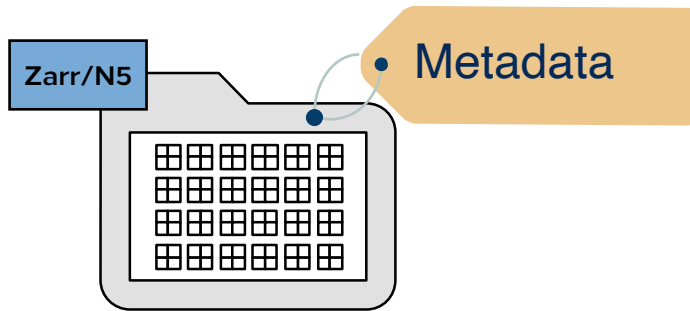
```

Reduce repetition

Optionally inline metadata

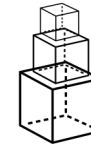
Offset & length





## Specifications:

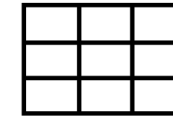
- Multiscales



- Labels



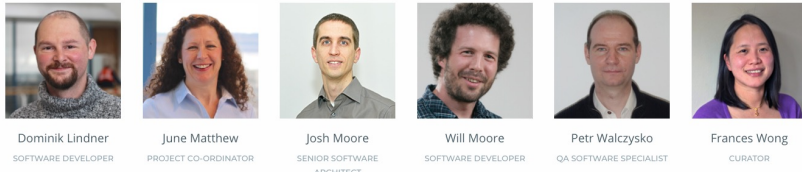
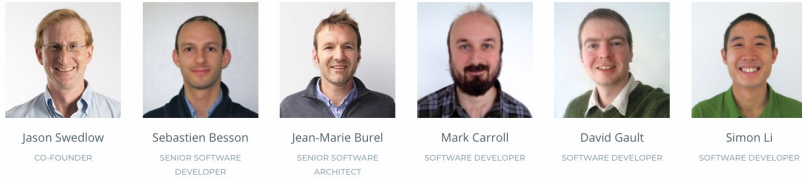
- HCS Plates



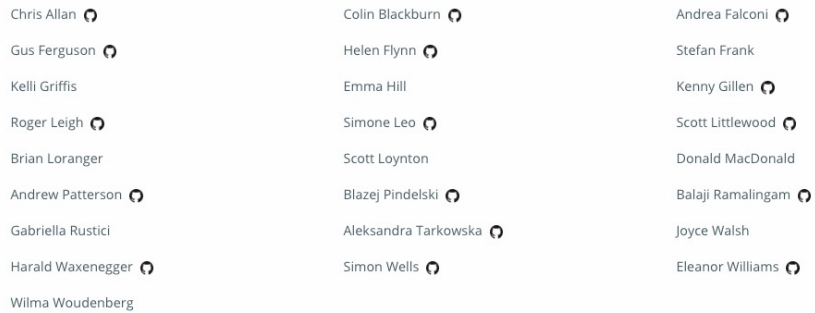
## Process:

- Discussions: <https://image.sc/tag/ome-ngff>
- Publication: <https://ngff.openmicroscopy.org>
- Samples: <https://s3.embassy.ebi.ac.uk/idr/zarr>

# Acknowledgements



Former members of the OME team in Dundee

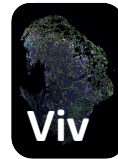


## Development Teams

Other teams are also working on developing or integrating OME tools.



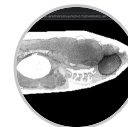
<https://www.openmicroscopy.org/teams>



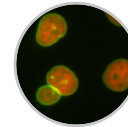
Gehlenborg Lab (HMS)



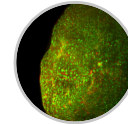
napari



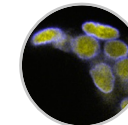
Faas *et al.* ([idr0053](#), CC BY-NC-SA 3.0)  
**J Cell Biol (2012)**



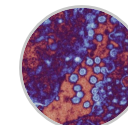
Hériché *et al.* ([idr0002](#), CC-BY 4.0)  
**MBoC (2016)**



McDole *et al.* ([idr0044](#), CC BY 4.0)  
**Cell (2018)**



Blin *et al.* ([idr0062](#), CC BY 4.0)  
**PLOS Biology (2019)**



Lamers *et al.* ([idr0083](#), CC BY 4.0)  
**Science (2020)**

**BDV**

Tomancak Lab  
(MPI-CBG)

**KLB**

Keller Lab  
(Janelia)

**N5**

Saalfeld Lab  
(Janelia)

**Zarr**

Alistair Miles  
(Oxford)



WORLD

