@normanrz
normanrz@mastodon.social

scalableminds

Interoperability for large images with OME-Zarr and WEBKNOSSOS

29 May 2024

Norman Rzepka ~ OME Meeting

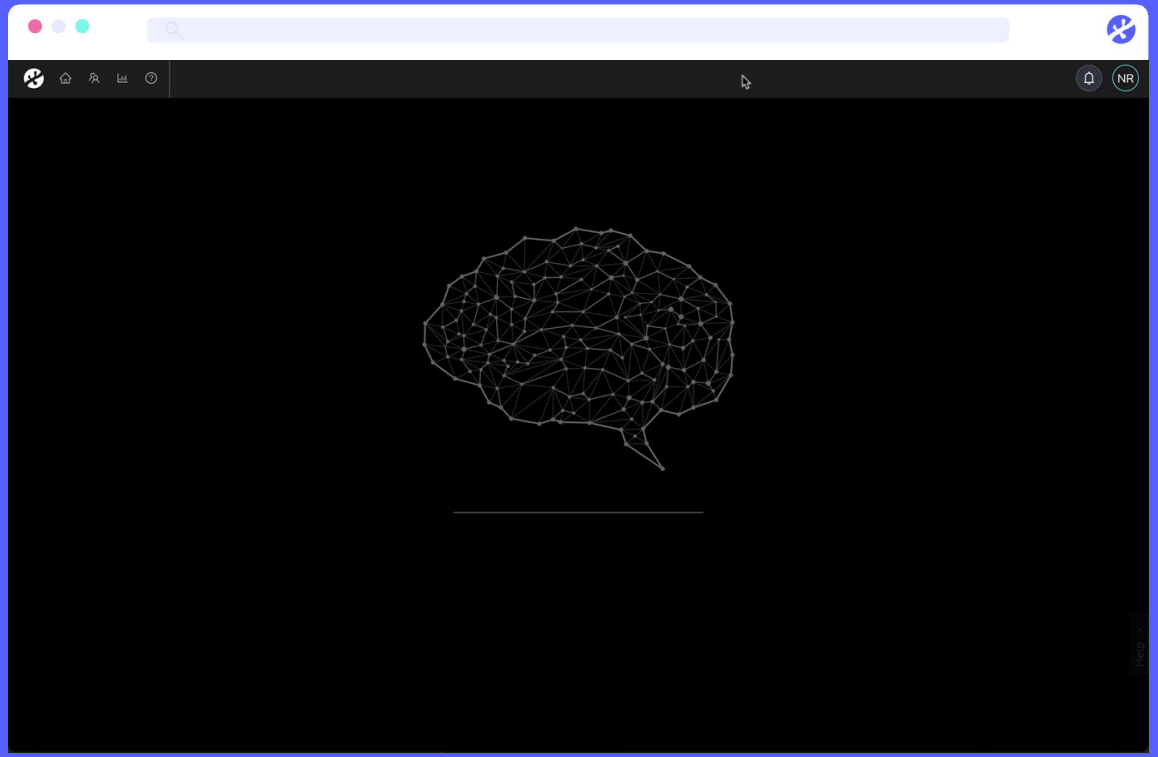
scalableminds

WEBKNOSSOS

A large-image data platform

Easily browse through terabytes of data

over the web

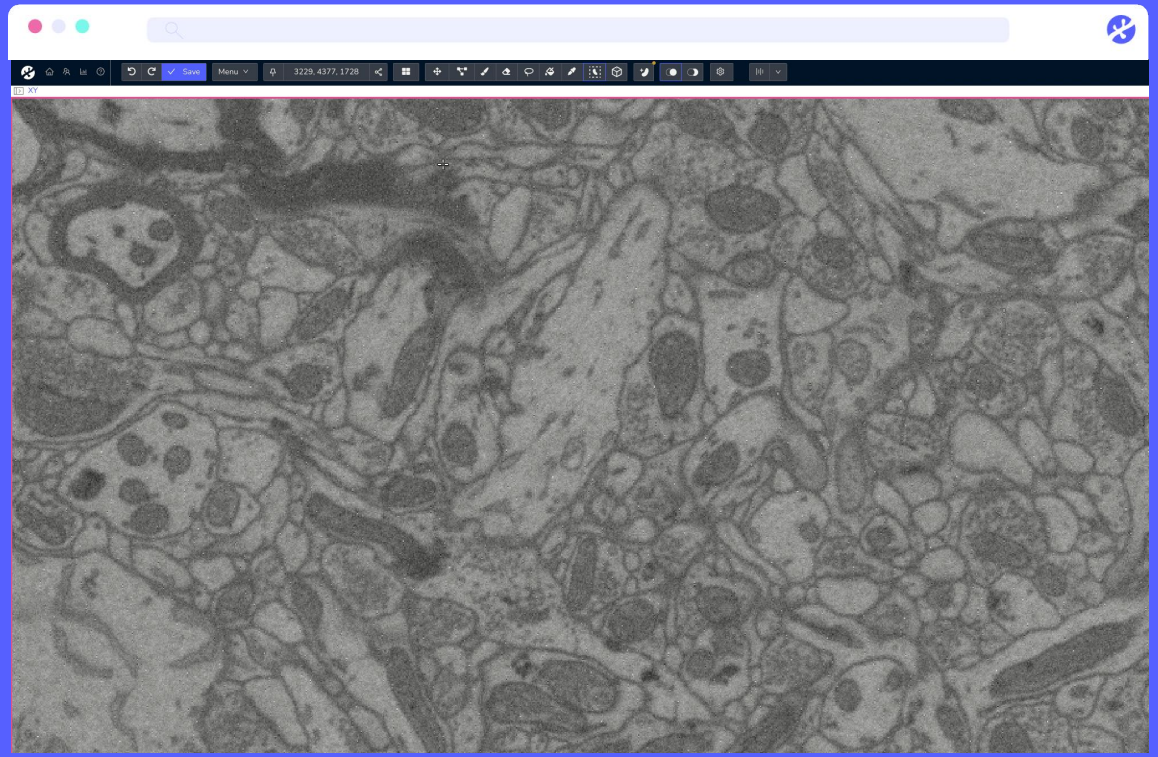


WEBKNOSSOS

EM data of drosophila melanogaster by Zhang et al., *Cell* 2017

Annotate data

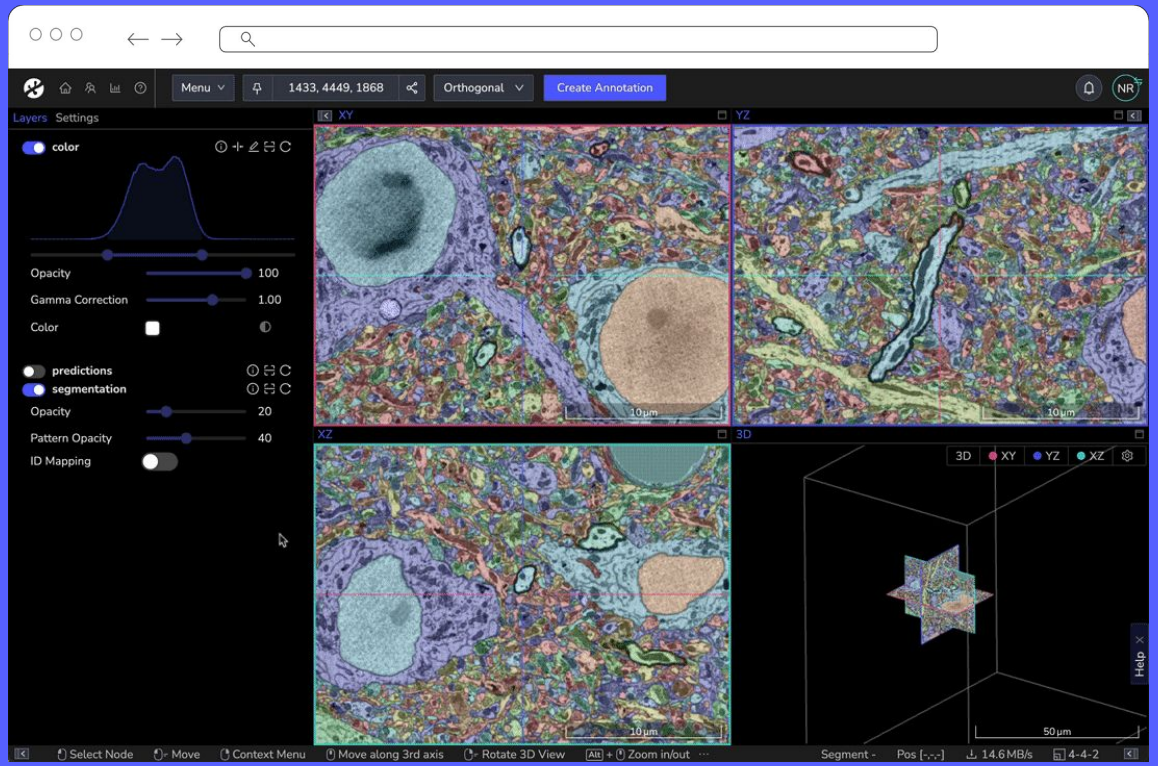
Quickly and collaboratively



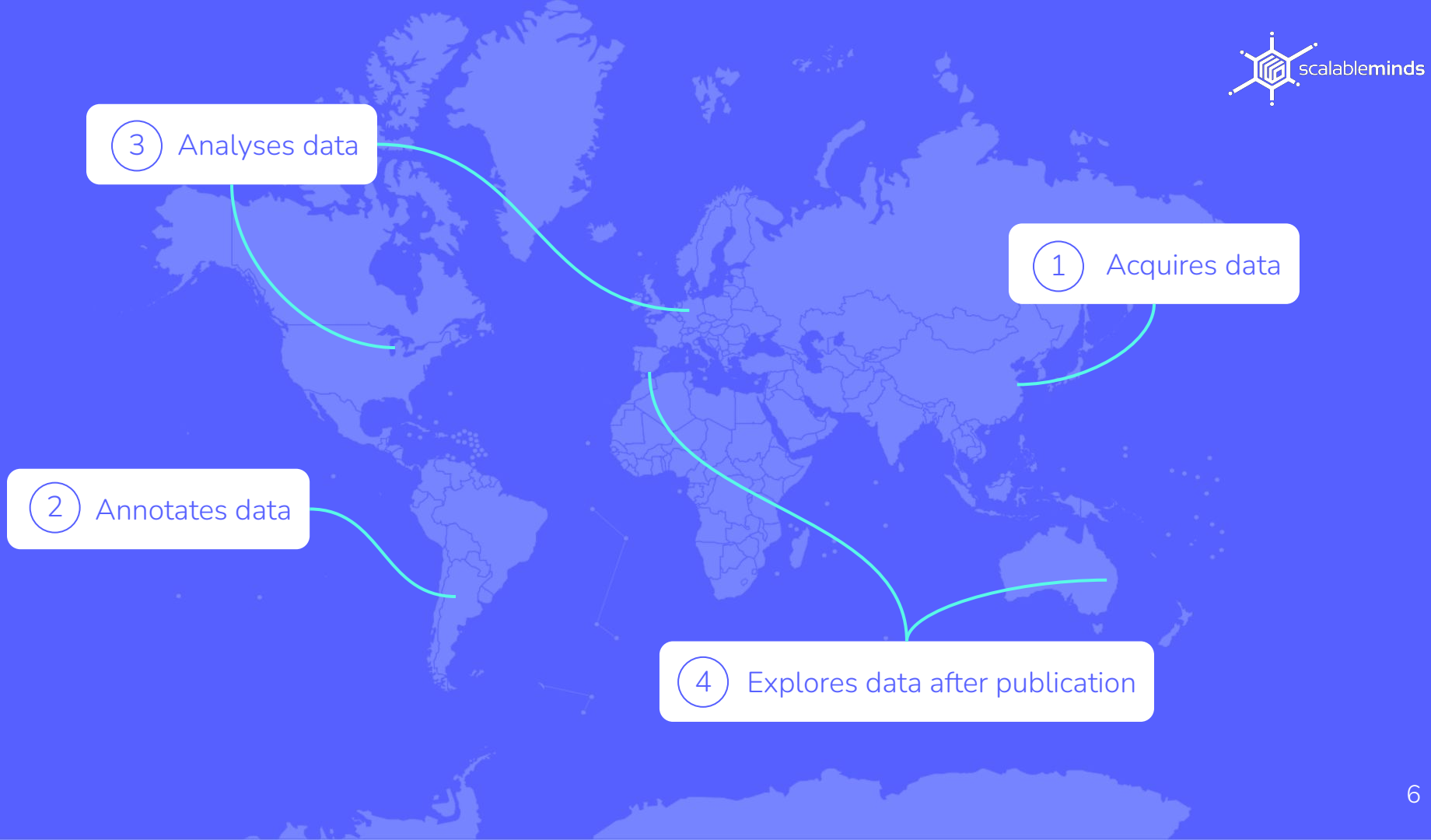
WEBKNOSSOS

EM data of mouse cortex by
Motta et al., *Science* 2019
Segmentation by scalable minds

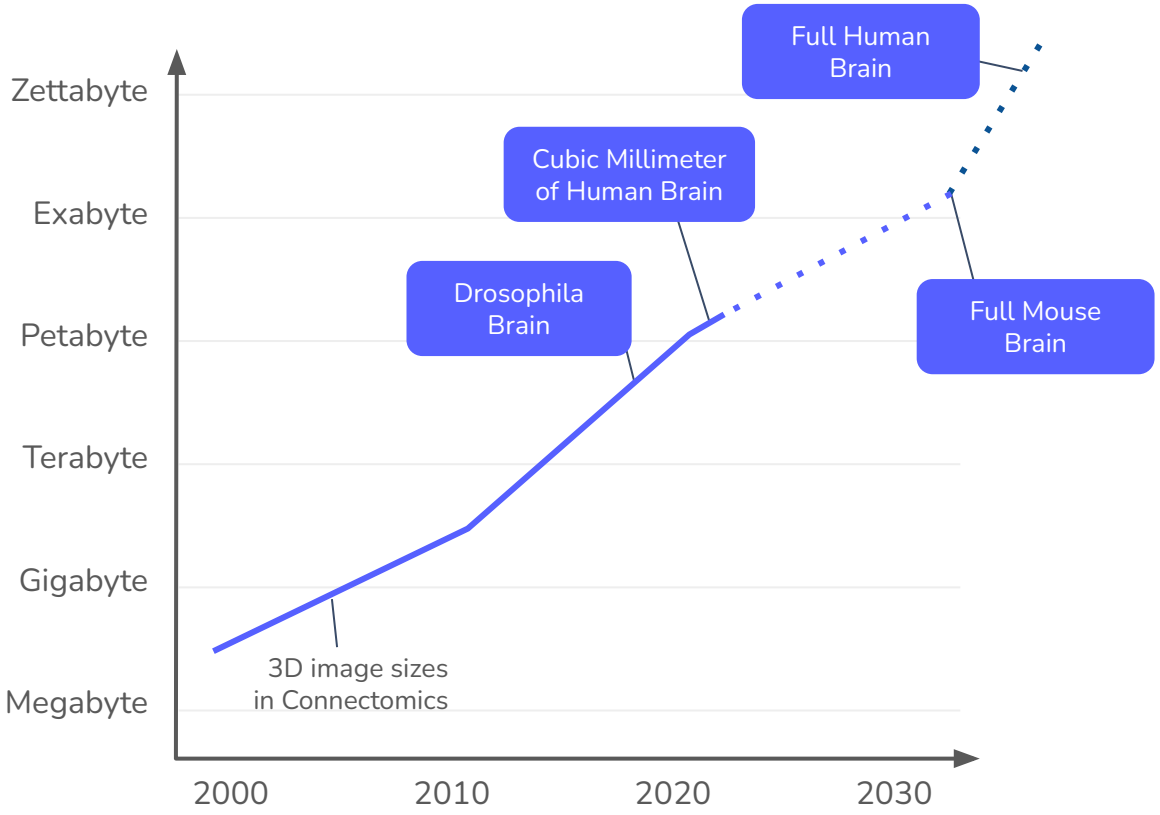
Train and run
AI models
(paid feature)



EM data of mouse cortex by
Motta et al., *Science* 2019
Segmentation by scalable minds

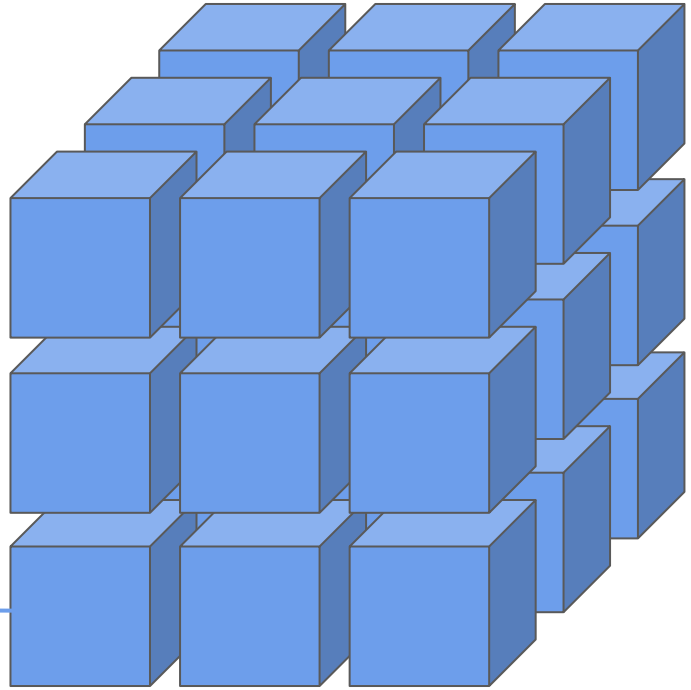

3 Analyses data
1 Acquires data
2 Annotates data
4 Explores data after publication

How can we store images that are 100s of TBs?



scalableminds

Key enabler: Chunked 3D data streaming and storage



Chunks

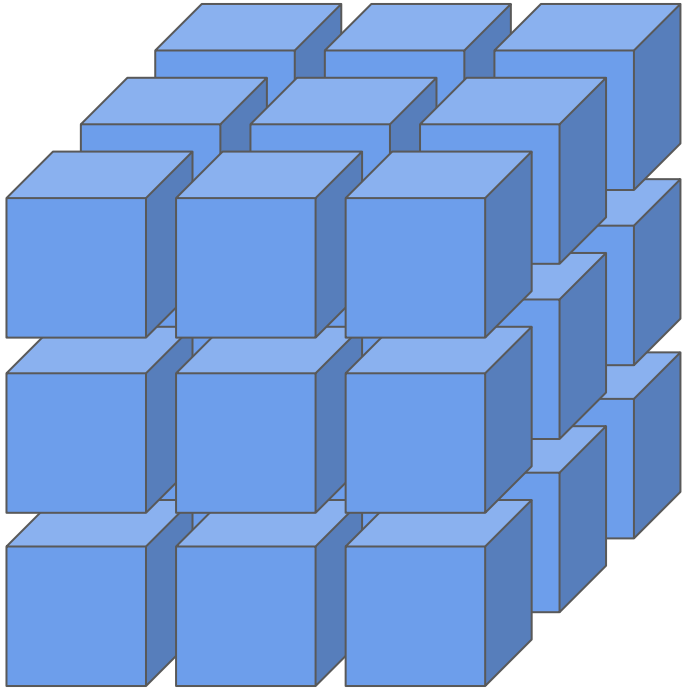
32 x 32 x 32 vx

= 32kB for uint8

scalableminds

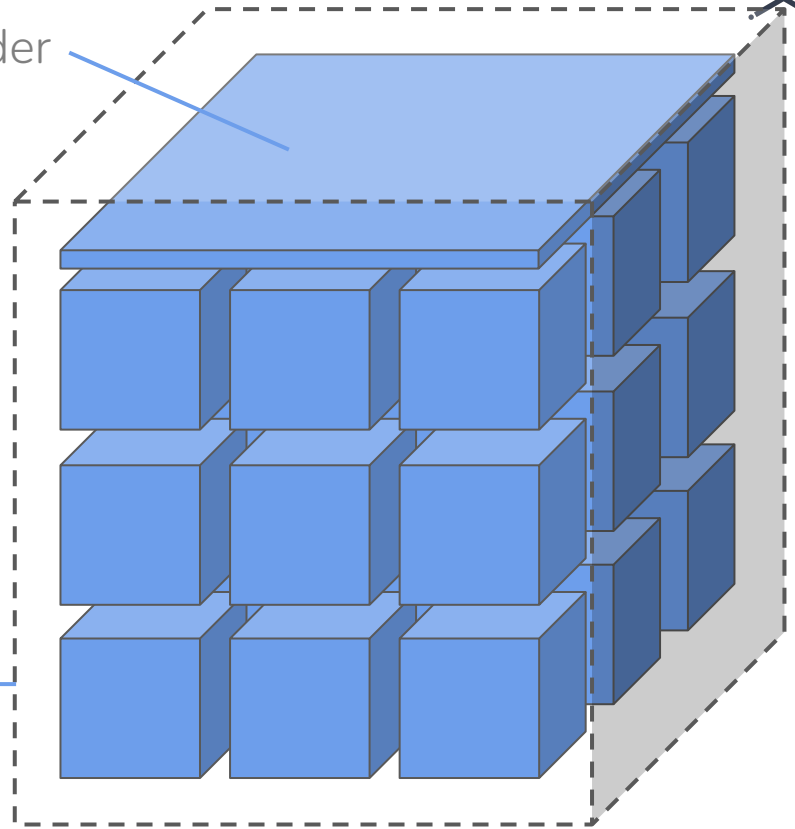
Many 32kB files don’t play well with current file systems

1TB = 33.554.432 files



scalableminds

Idea: Put many chunks in a larger file (shard)



32 x 32 x 32 chunks

= 1GB for uint8

What about interoperability?

scalableminds

WEBKNOSSOS Wrap

N5

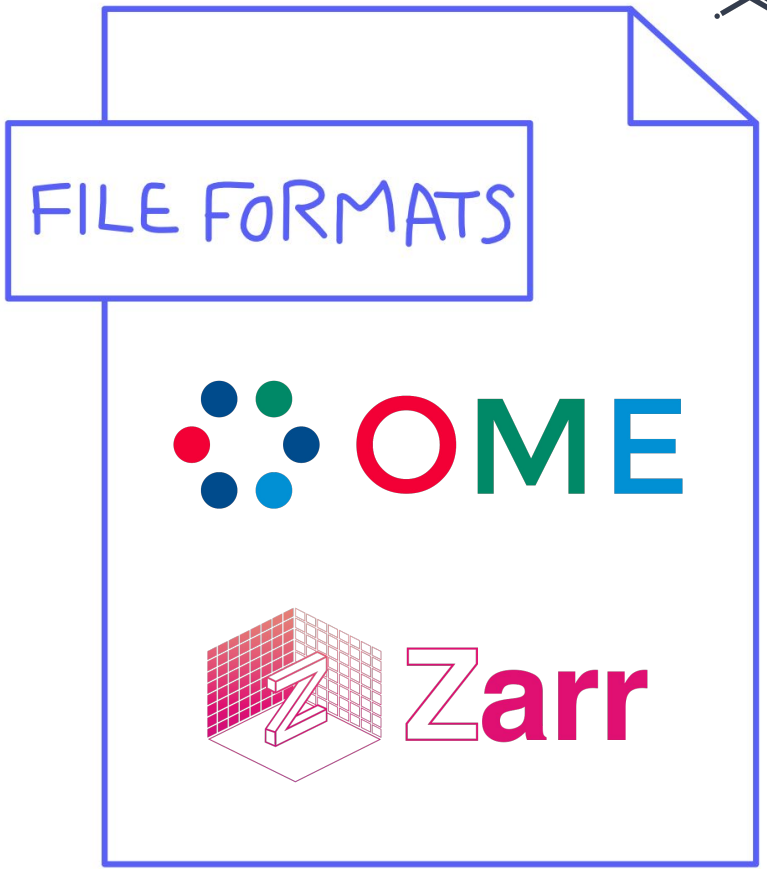
Neuroglancer precomputed

BossDB

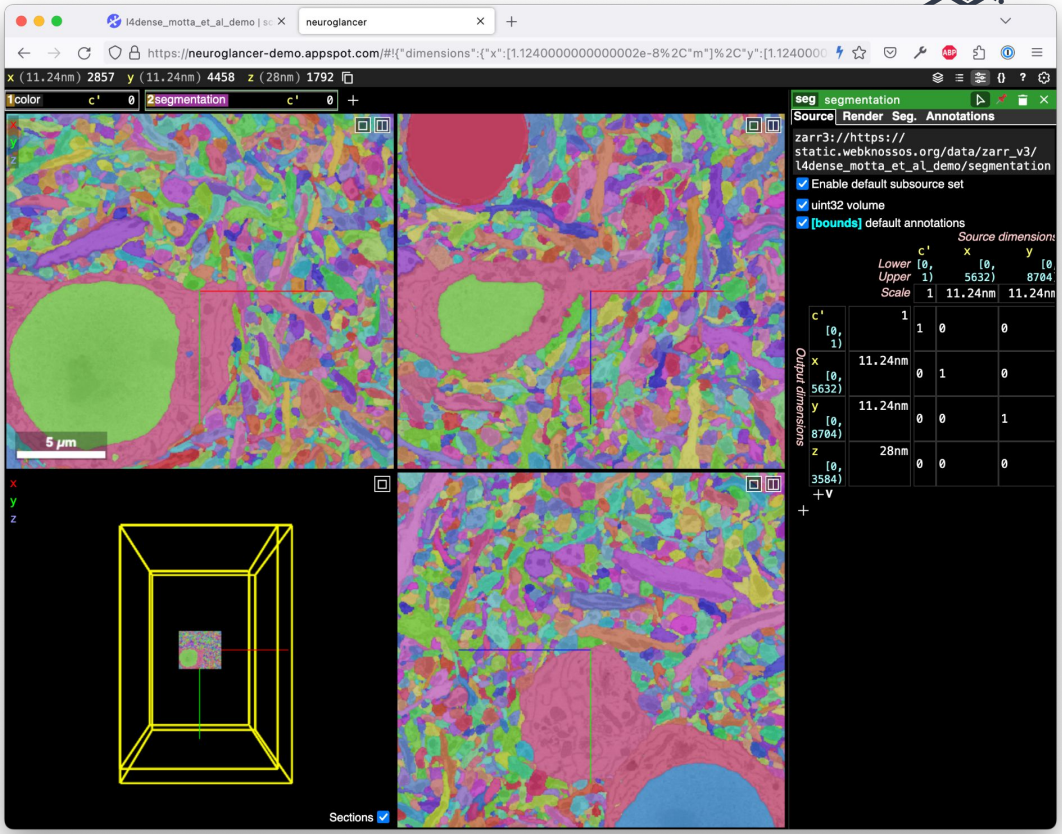
Zarr

HDF5

A next-generation file format (NGFF)

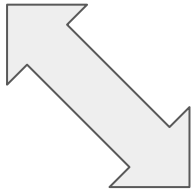


Broad support from applications



EM Data my Motta et al., segmentation by scalable minds
Neuroglancer by Jeremy Maitin-Shepard et al.

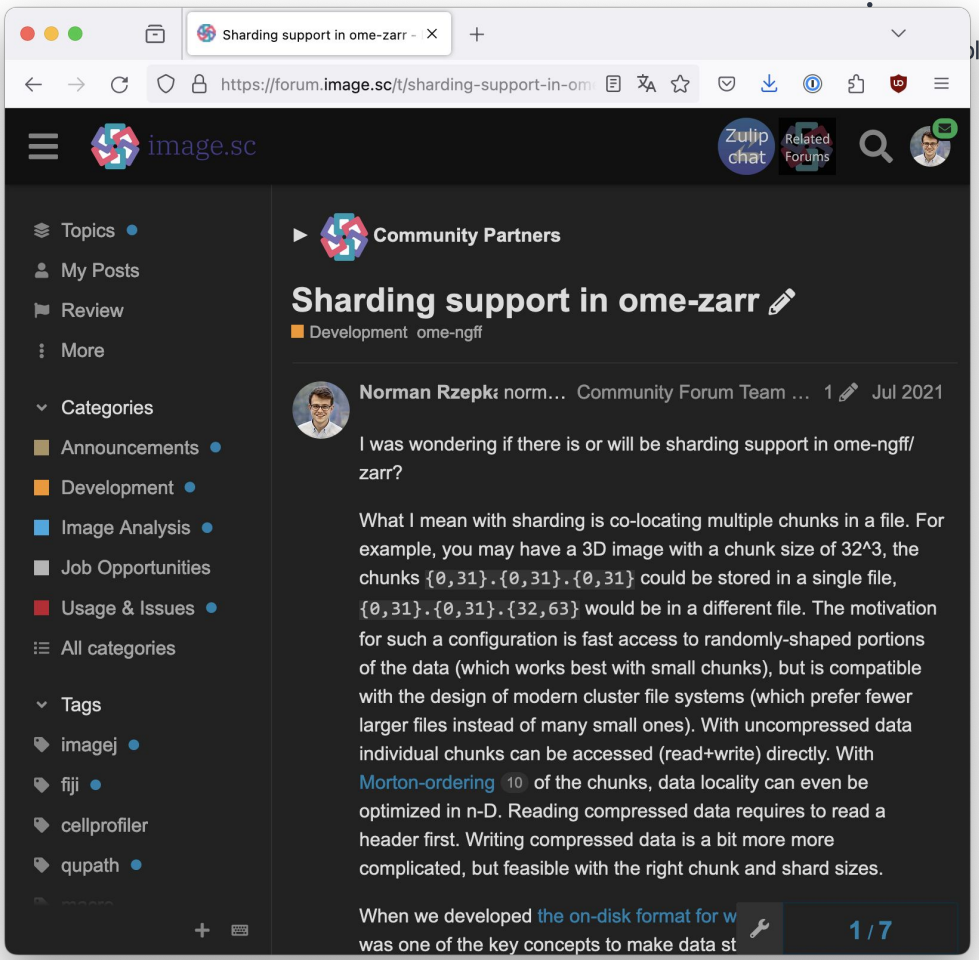
Integration with OMERO



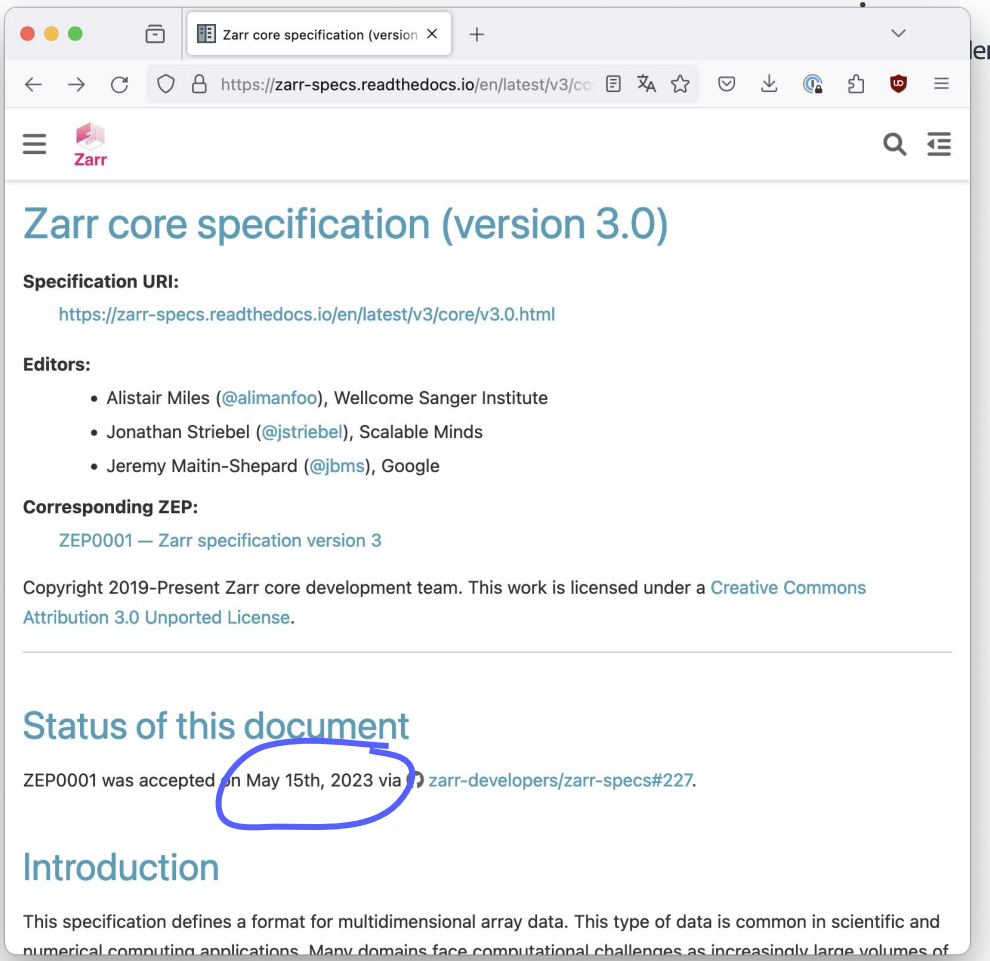
Data streaming for scripting

```python
import webknossos as wk


with wk.webknossos_context(token="..."):
    dataset = wk.Dataset.remote_open("4496763.zarr")
    layer = dataset.get_color_layer()
    mag = layer.get_mag(1)

    data = mag.read(
        absolute_offset=(0, 0, 0),
        size=(1024, 1024, 25),
    )

run_analysis(data)
```

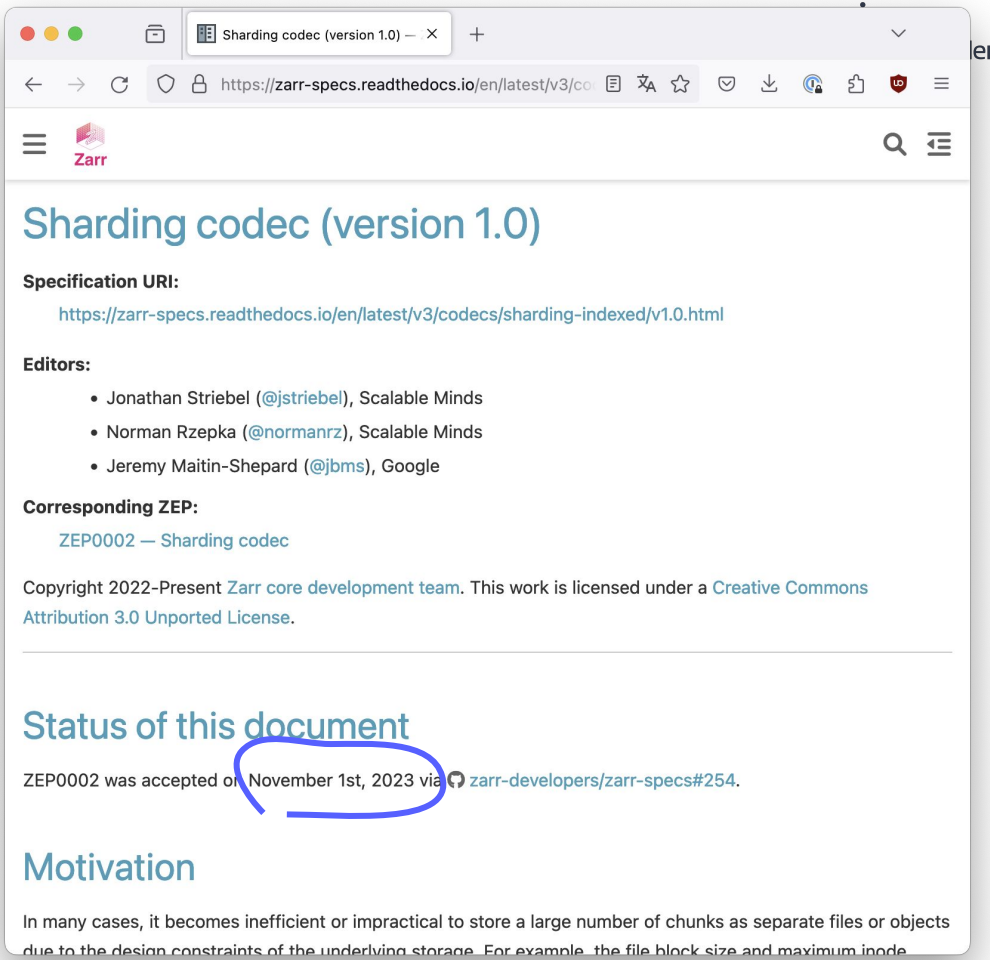
But, OME-Zarr doesn't have sharding



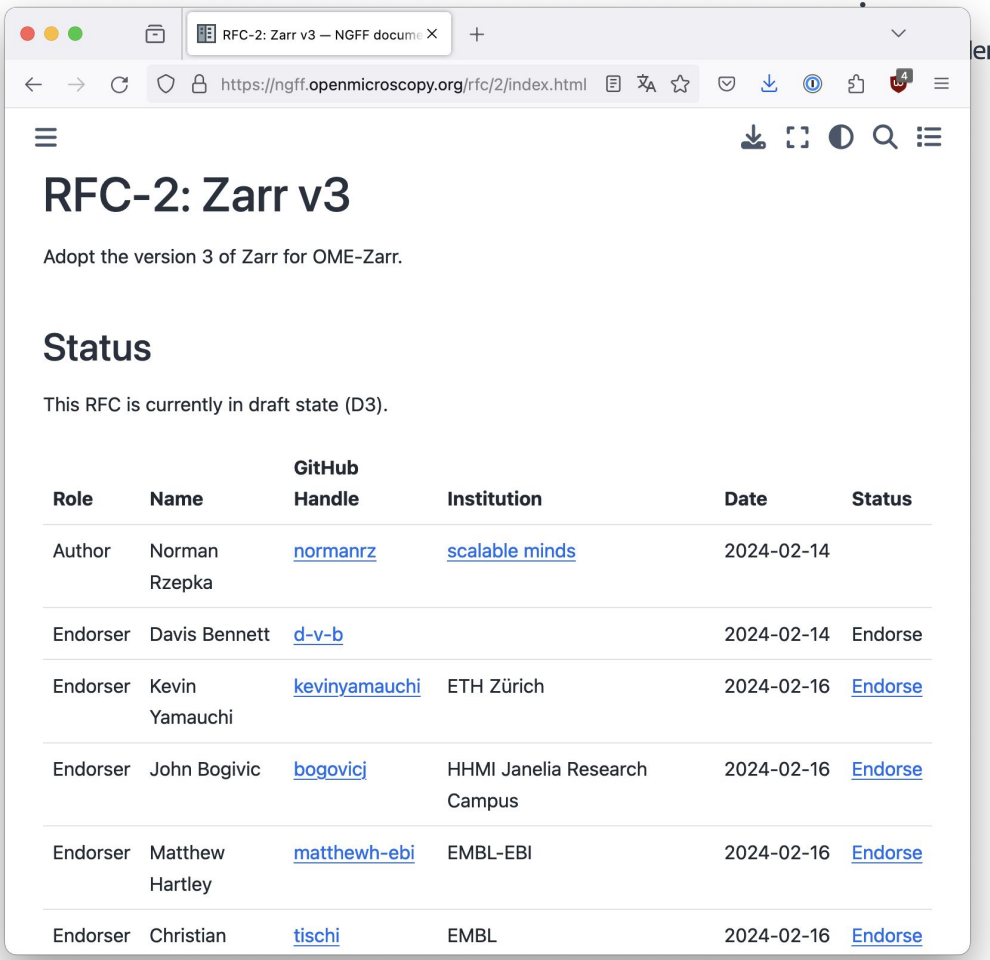
The ZEP process didn't exist and Zarr v3 was still in under development



Adding sharding as a codec to Zarr

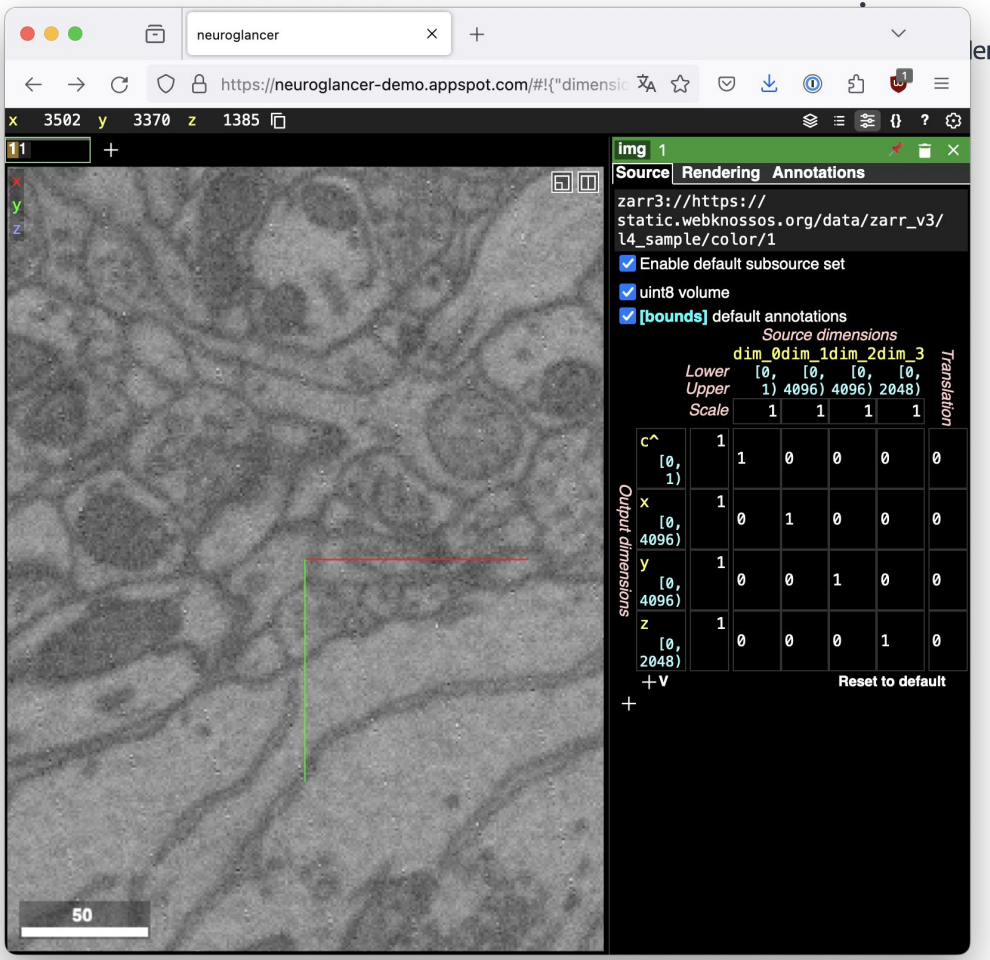


Adopt sharding and Zarr v3 in OME-Zarr



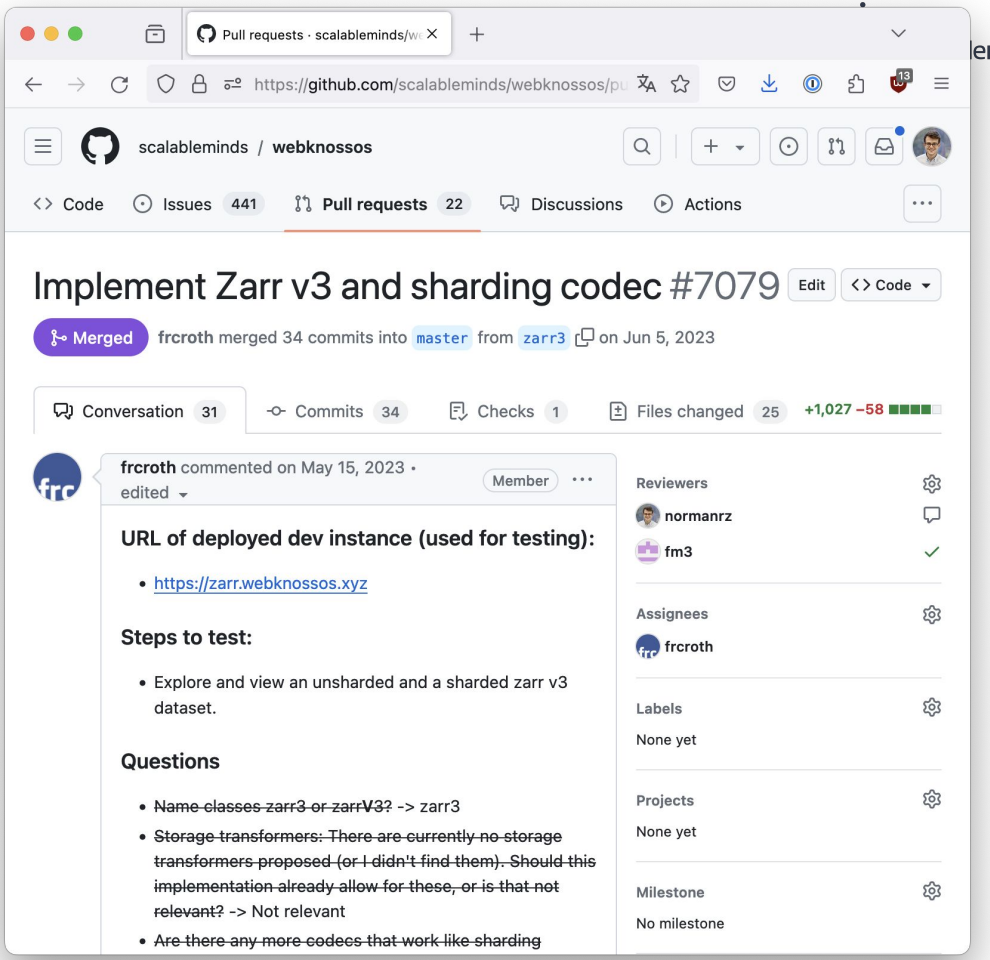
- [x] Specification
- [ ] Implementation

Visualize with neuroglancer

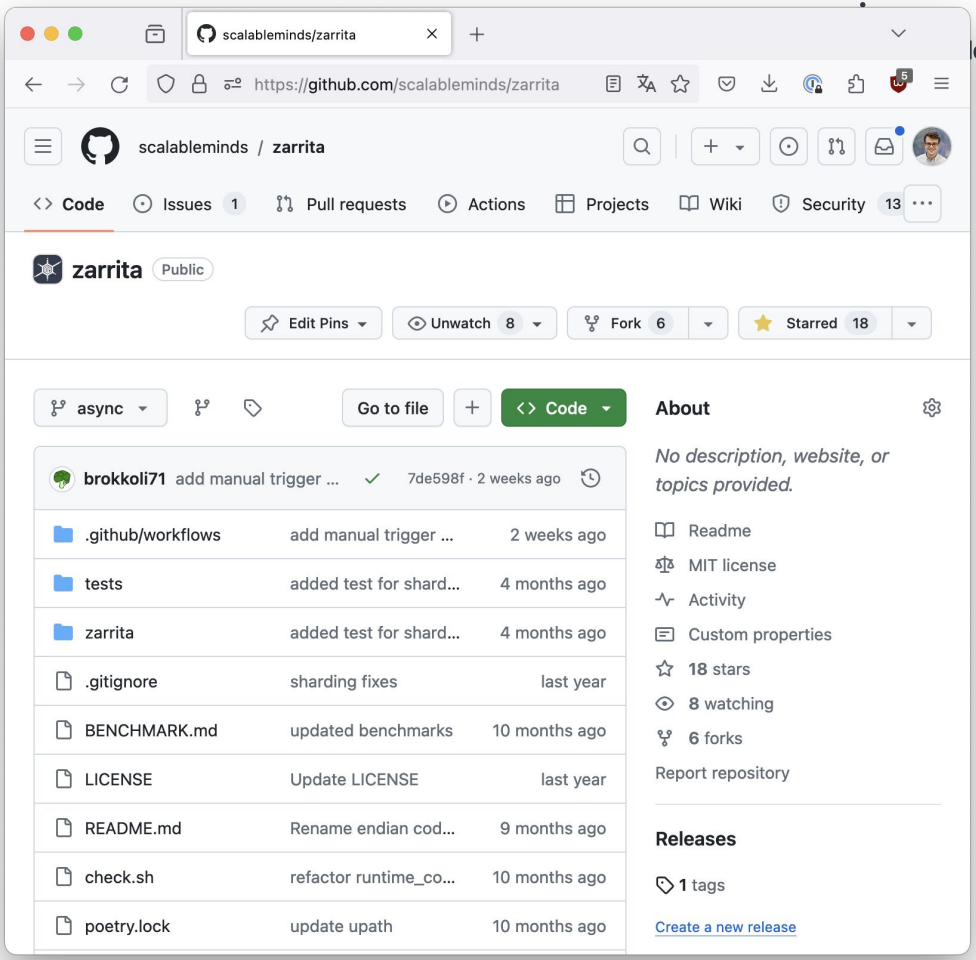


EM Data my Motta et al.
Neuroglancer by Jeremy Maitin-Shepard et al.

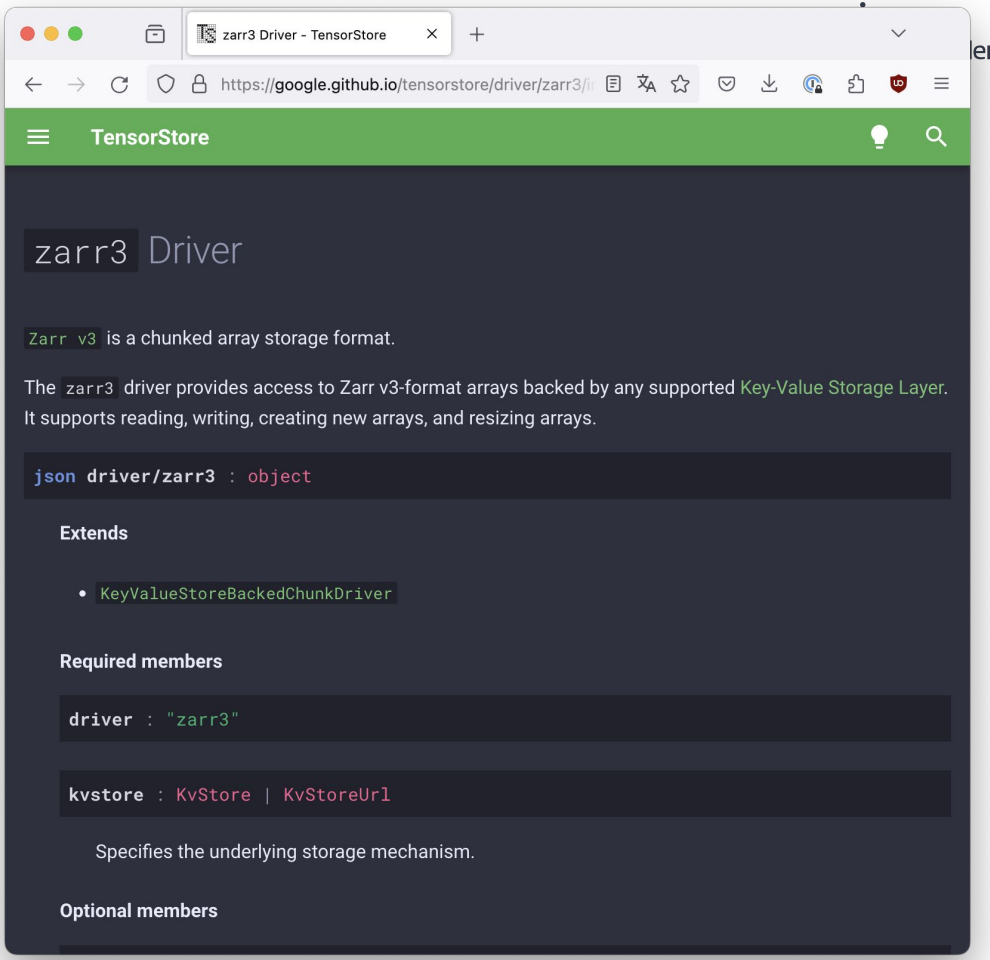
Visualize and annotate with WEBKNOSSOS



Read/write with zarrita



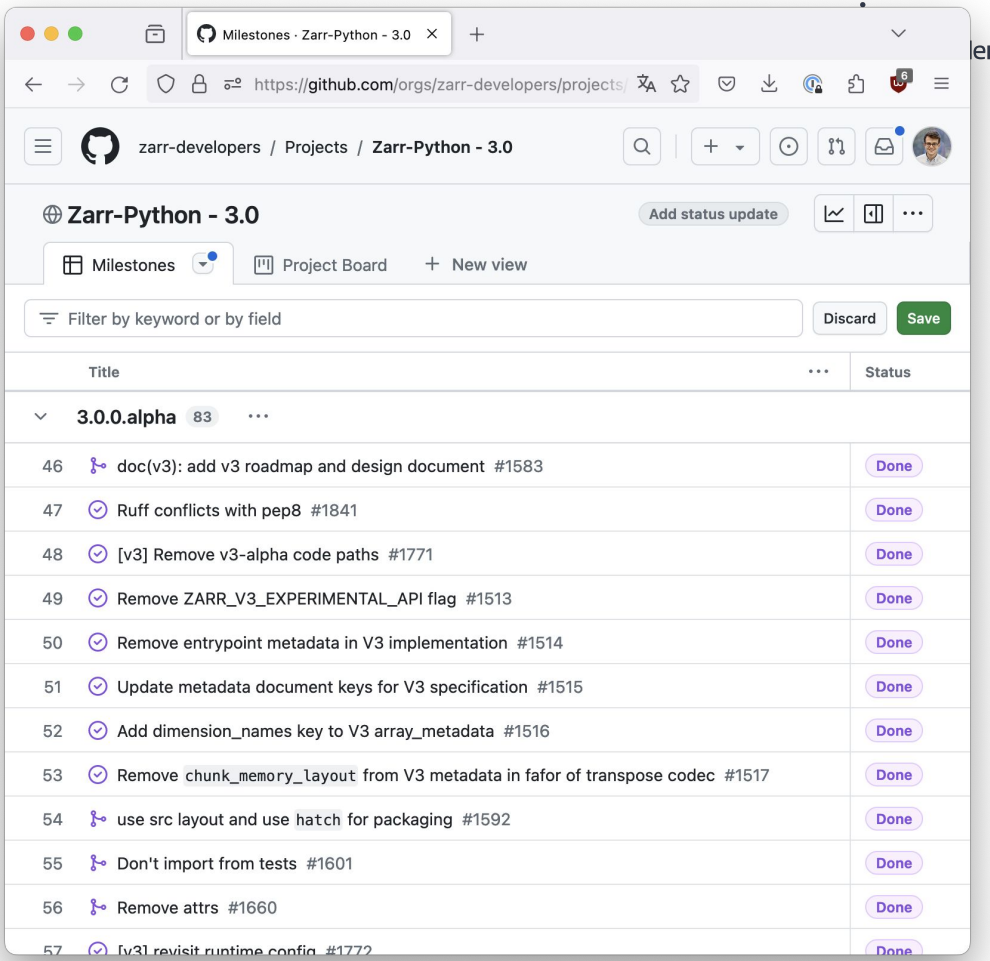
Read/write with tensorstore



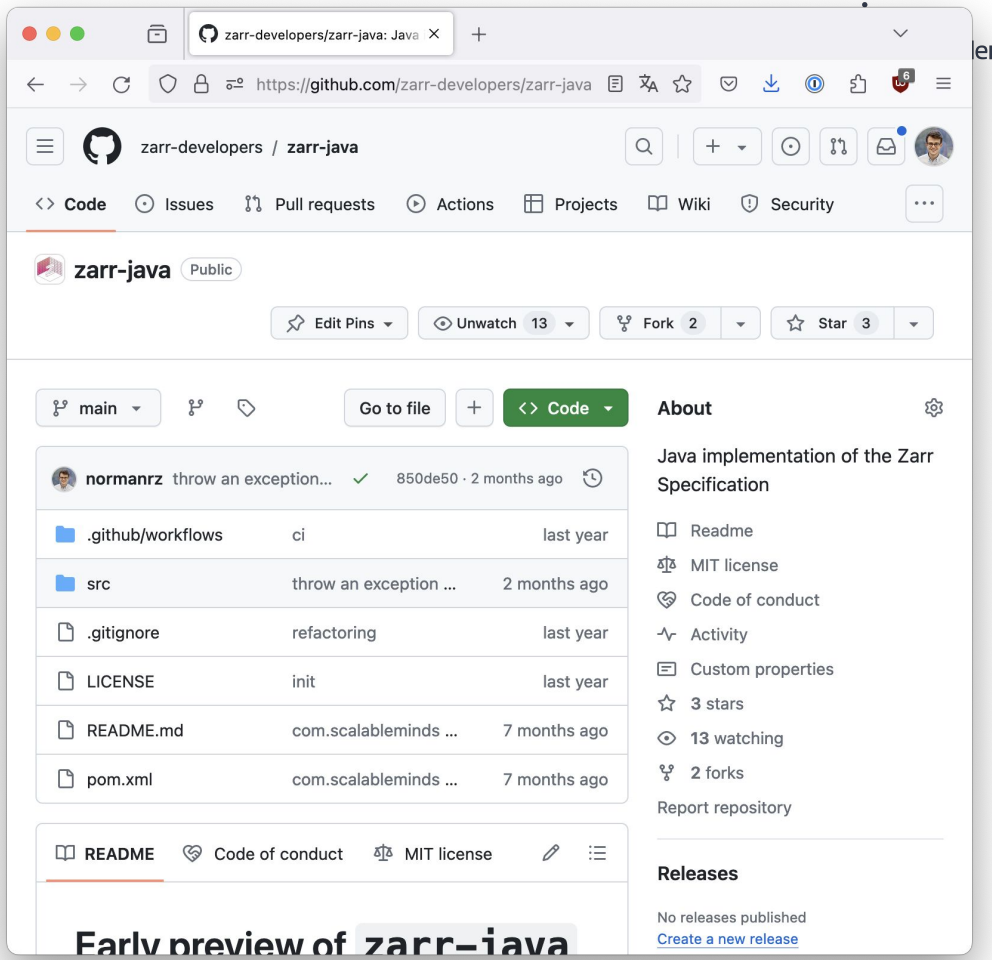
By Jeremy Maitin-Shepard et al.

zarr-python 3.0

Release scheduled for June

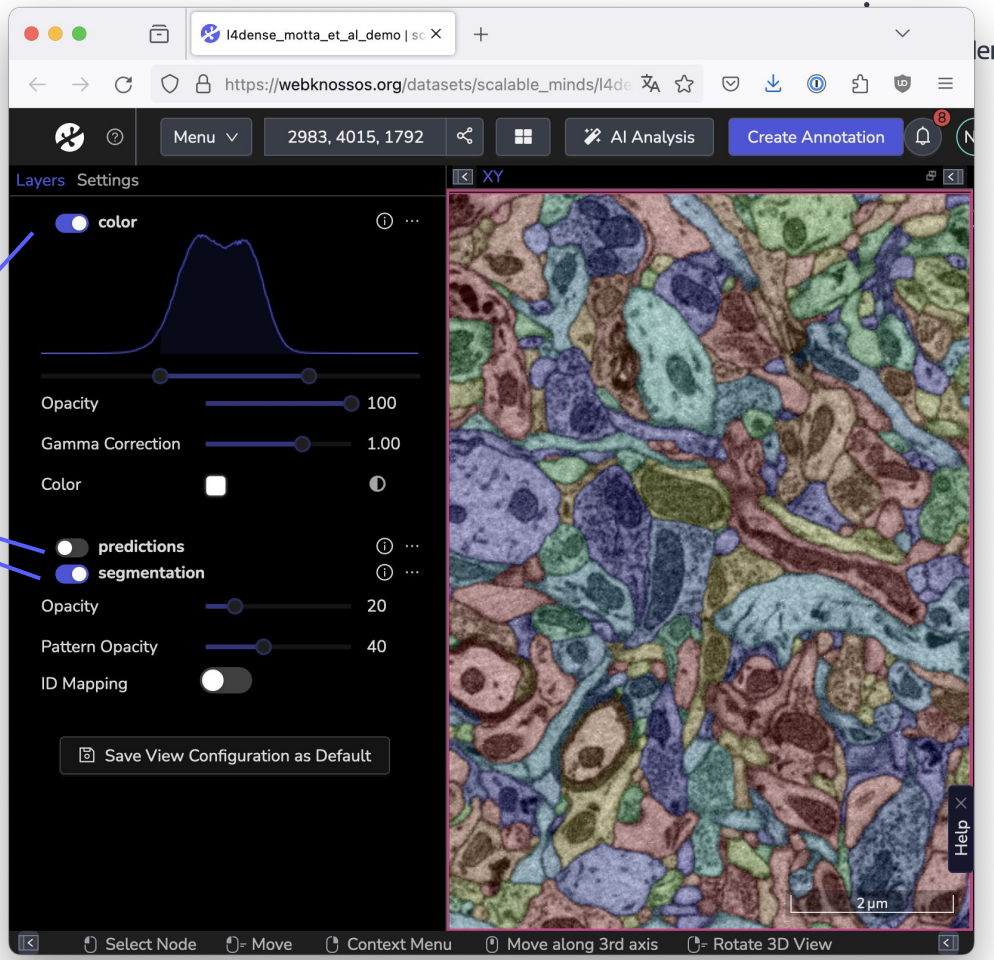


zarr-java and
n5-zarr for
Fiji, OMERO,
bf2raw etc.



What is next?

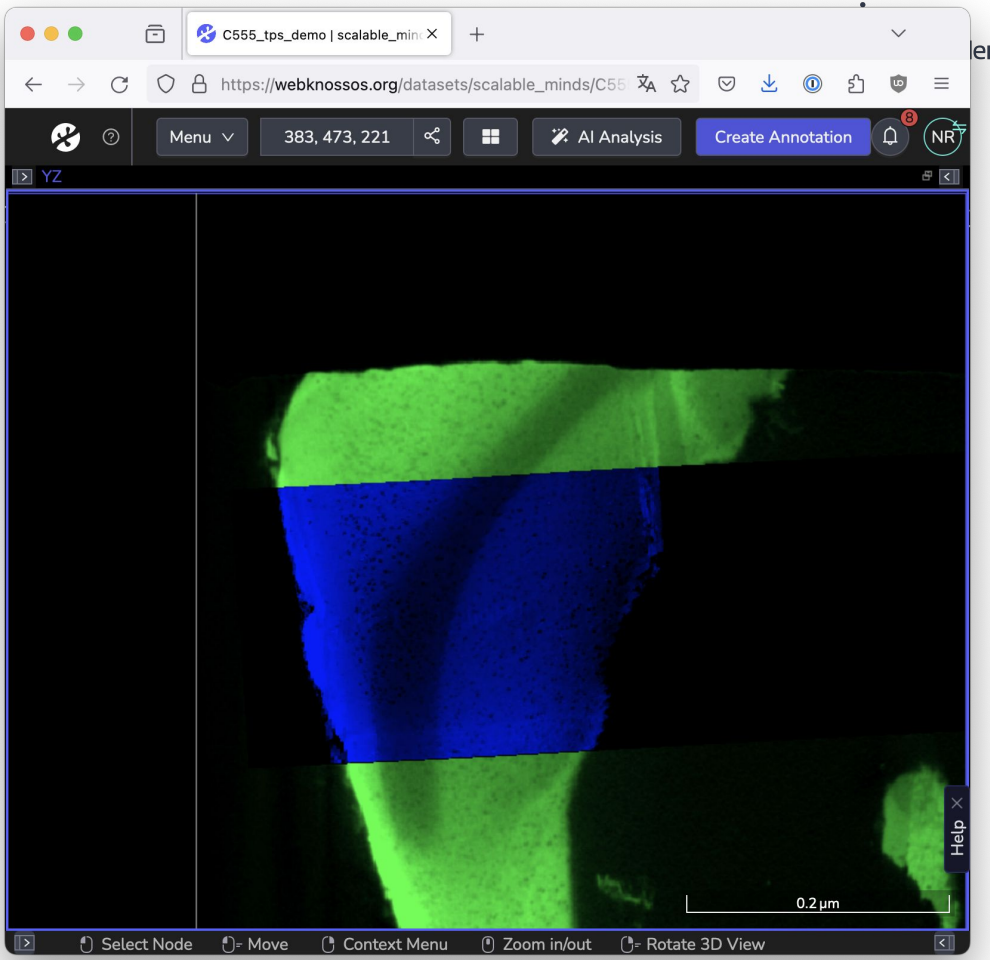
Collections of images



Collaboration with Christian Tischer, Davis Bennett, Joel Lüthi and many more from the OME-Zarr community

Collections of images

Correlative visualization and annotation



Proposal by John Bogovic
Data by Bosch et al.

~~OME-NGFF~~

OME-Zarr 1.0

scalableminds

Thank you!

@webknossos

webknossos@mstdn.science

https://webknossos.org

https://scalableminds.com