Next Generation File Formats and Platforms

Joshua Moore[1] and Nicholas Sofroniew[2]

[1]University of Dundee, Scotland, UK, Walluf, Germany, [2]Chan Zuckerberg Initiative, United States

Quite simply, the volume and complexity of microscopy data that can be acquired today, though astounding, are also troublesome. The introductions of most recent bioimaging papers touch on this to some degree. Gifted researchers naturally find ways to overcome their immediate problems but often must do so alone without time or funds to maintain their creations. The Chan Zuckerberg Initiative (CZI), the Open Microscopy Environment (OME) and similarly minded groups of research scientific engineers (RSE) are working to bring together friendly, generalizable, scalable but also maintainable infrastructures to ease the individual burden on labs. Collaboration of this type drives ecosystems of interoperable tools and leads to FAIRer, more open data.

First, we will present recent work on data formats culminating in the development of OME-Zarr, a next generation format with the aim of facilitating the sharing of images on highly-parallel object storage. We hope having a common, cloud optimized, extensible format can avoid the proliferation of data formats which has been a hallmark of bioimaging. Examples of multi-resolution images stored together with labeled annotations in the Image Data Resource (IDR) have already expedited the sharing of recent SARS-CoV2 imaging data.

Then we will present napari, a fast, interactive, multi-dimensional image viewer for Python that can load next generation file formats like OME-Zarr. napari is built on top of and aims to integrate well with foundational open-source scientific Python libraries like numpy, Dask, vispy, and scikit-image. napari also aims to provide graphical user interface (GUI) access to a plugin ecosystem of image analysis tools for scientists to use in their daily work.

We propose that such next-generation tools can revolutionize the sharing and utility of bioimaging data by expediting access and driving discovery. We look forward to feedback and involvement towards these ends from participants, including vendors, researchers, and developers.

References

Besson S, *et al*. (2019) Bringing Open Data to Whole Slide Imaging. Digital Pathology ECDP 2019. Lecture Notes in Computer Science, vol 11435. doi:10.1007/978-3-030-23937-4_1

Eleanor Williams, *et al.* (2017) The Image Data Resource: A Bioimage Data Integration and Publication Platform. Nature Methods 14(8), 775-781. Published 19 June 2017 doi:10.1038/nmeth.4326

napari contributors (2019). napari: a multi-dimensional image viewer for python. doi:10.5281/zenodo.3555620