

Data Formats and OMERO

by Dominic Waithe
10th December 2019

IAFIG-RMS - Bioimage Analysis With Python
Cambridge Bioinformatics Training Centre

TODAYS TALK:

- Data formats.
- Big-Data Formats
- OMERO
- OMERO practical

UK Research
and Innovation

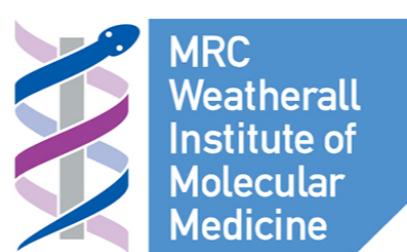
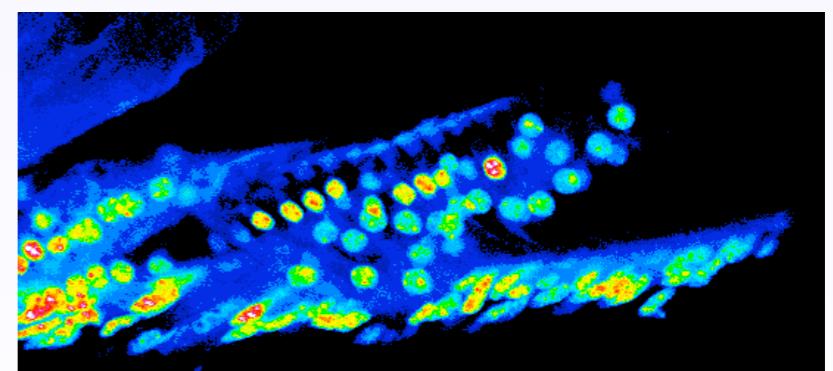
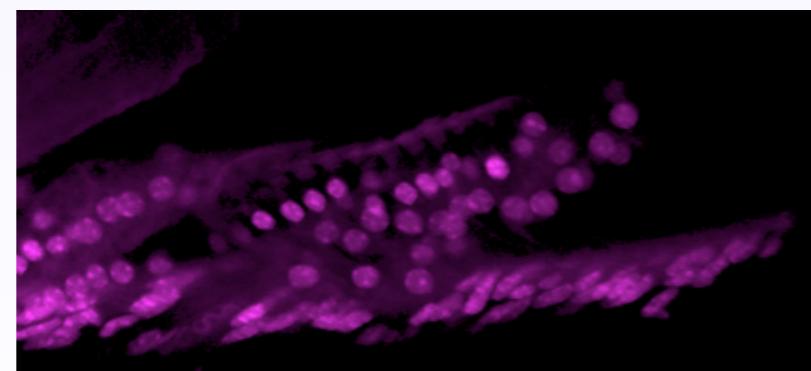
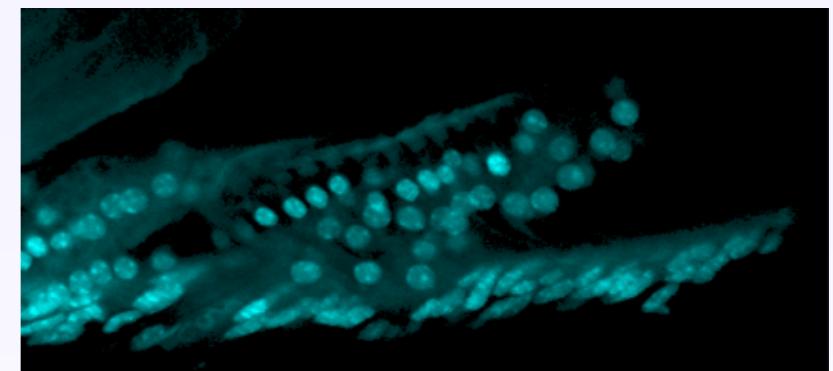
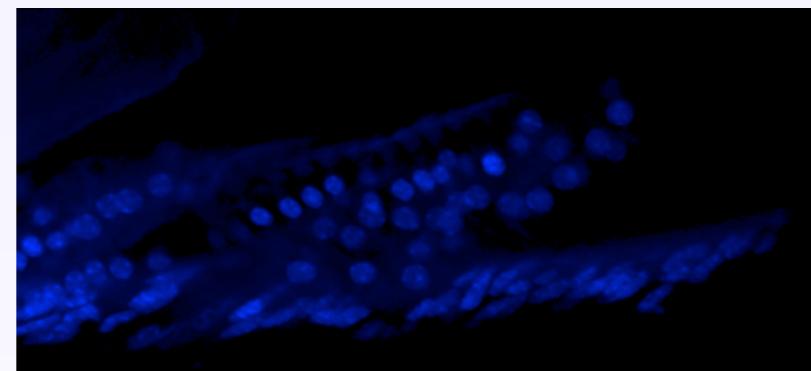
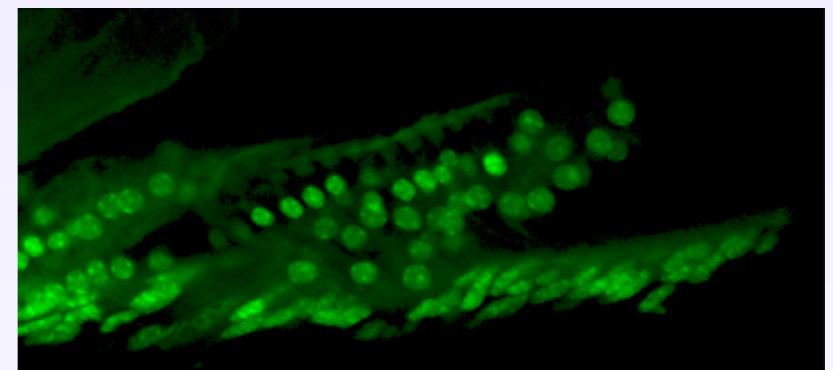
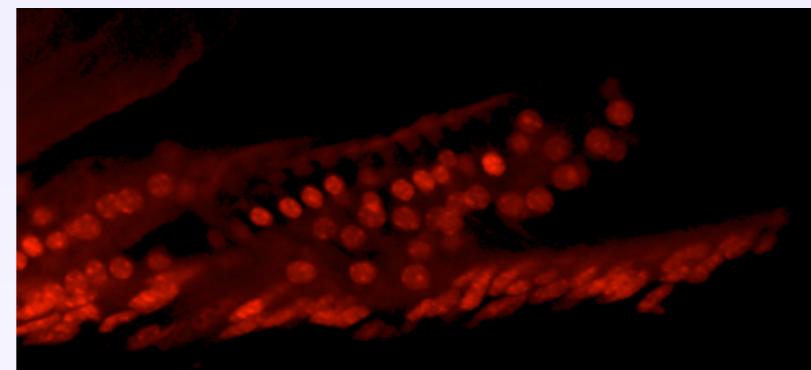
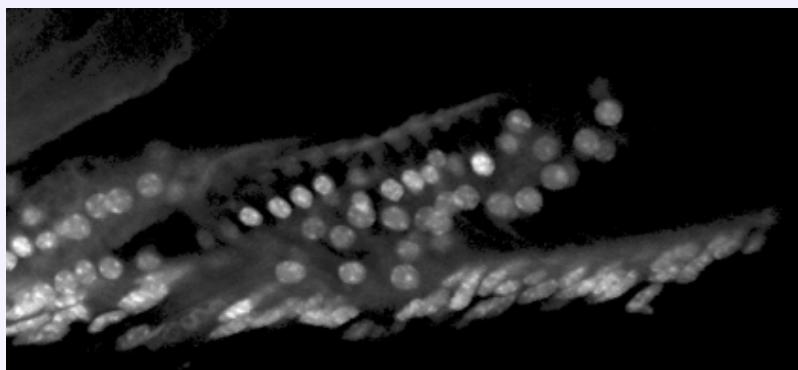


Image intensity

Pixel values when visualised on a screen can be coloured in a number of different ways, independent of the actual pixels values.



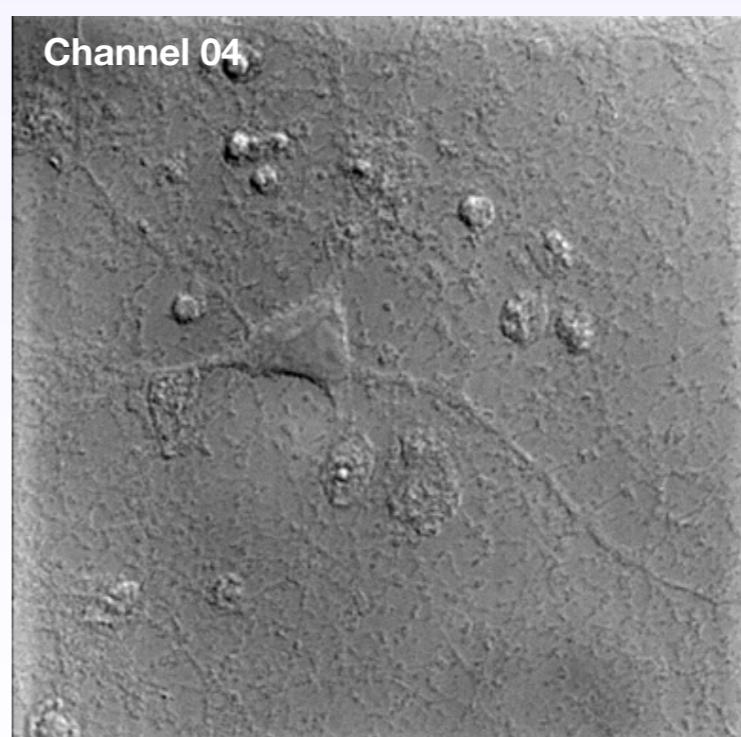
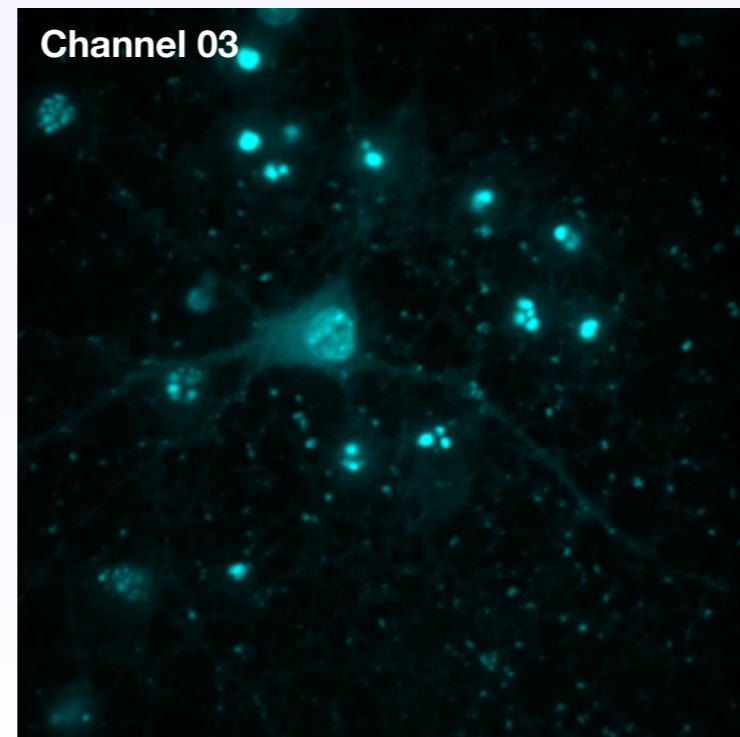
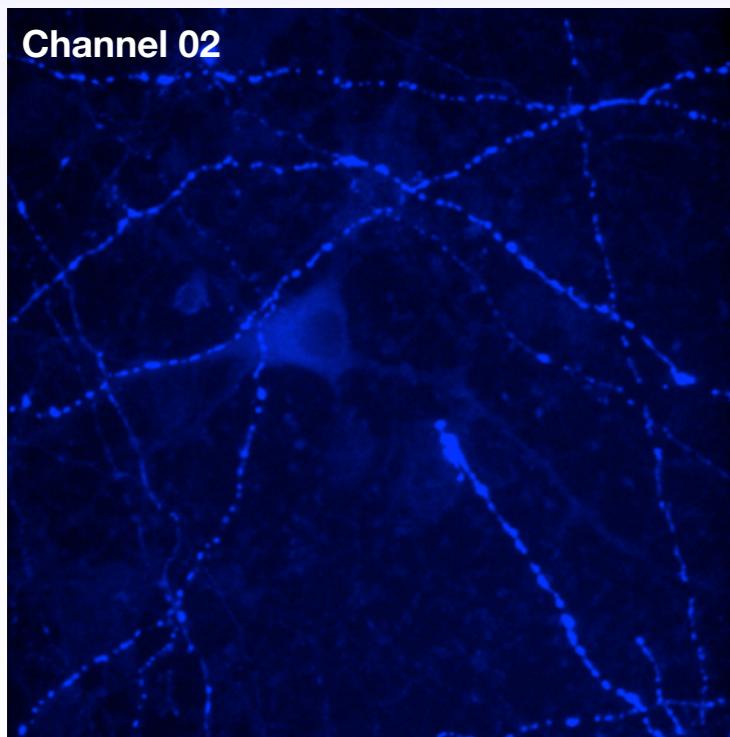
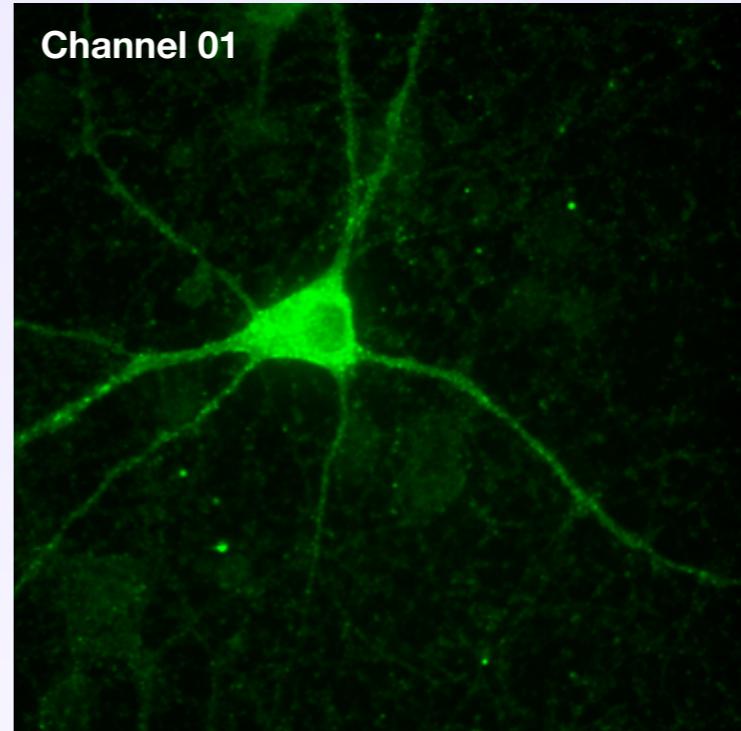
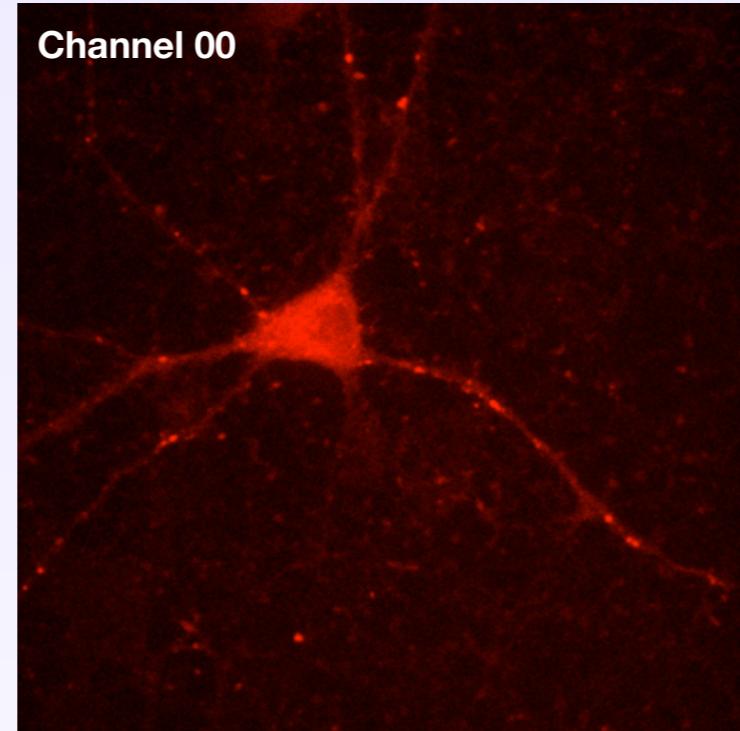
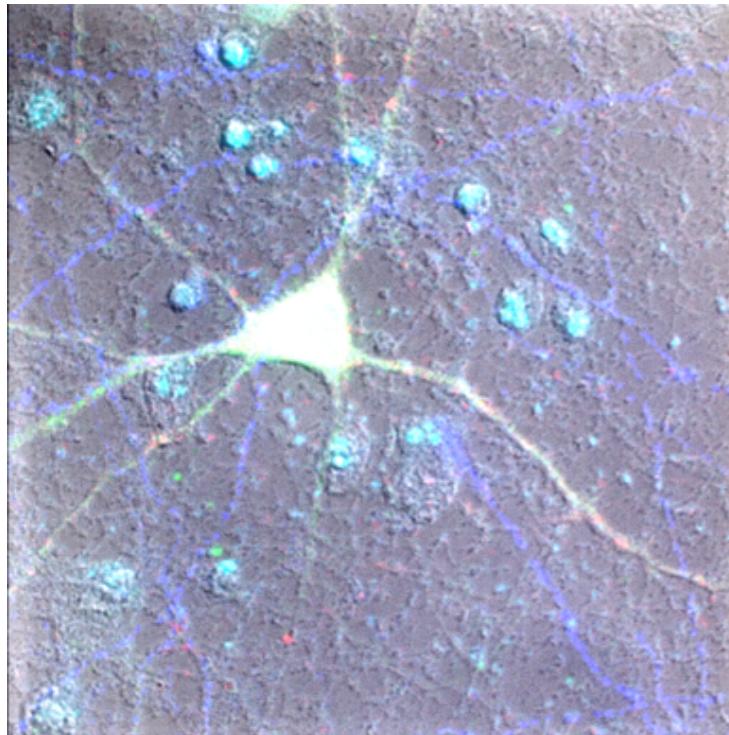
The colour pixels appear on a screen is defined by the look-up-table.

Any grayscale value can be mapped to any tone.

Source: Fiji, organ-of-corti.tif

Image representation

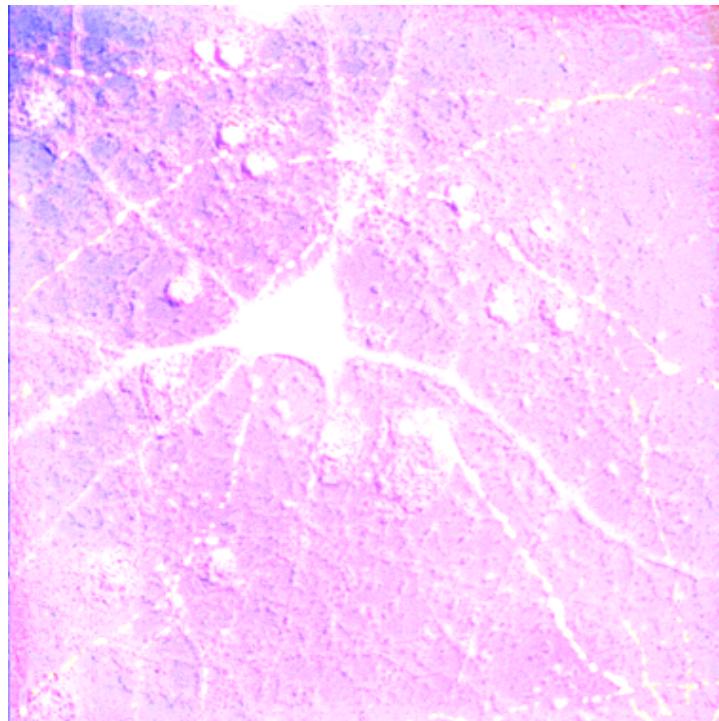
Composite representation



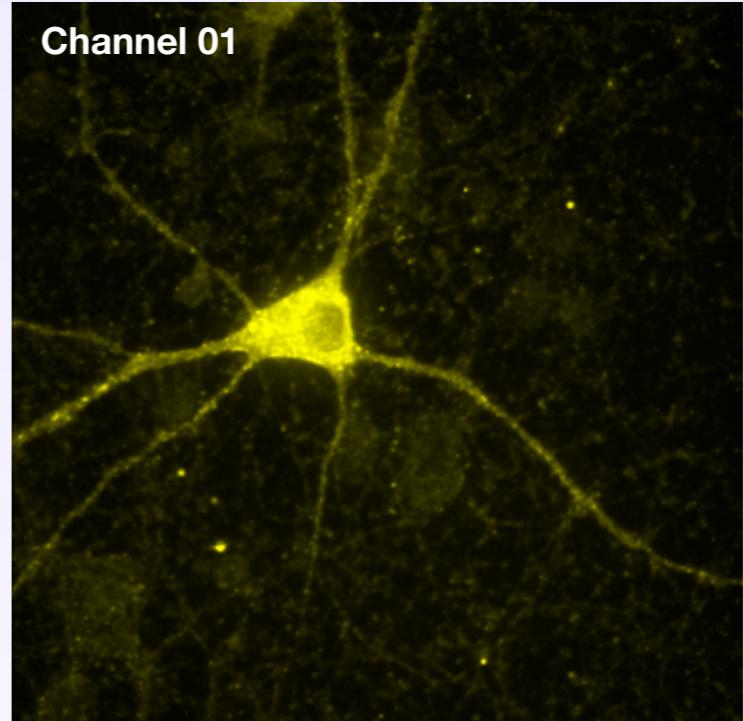
Source: ImageJ Rat_Hippocampal_Neuron.tif.

Image representation

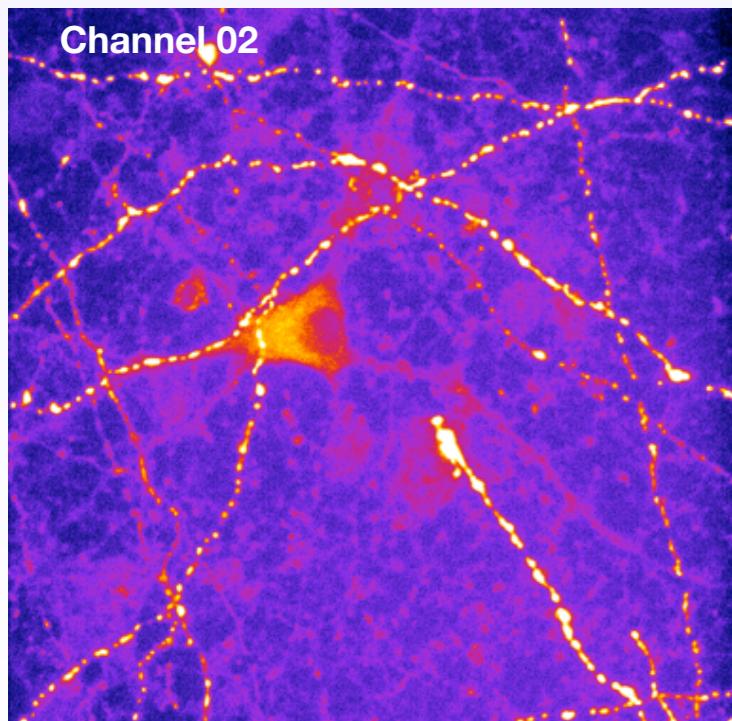
Composite representation



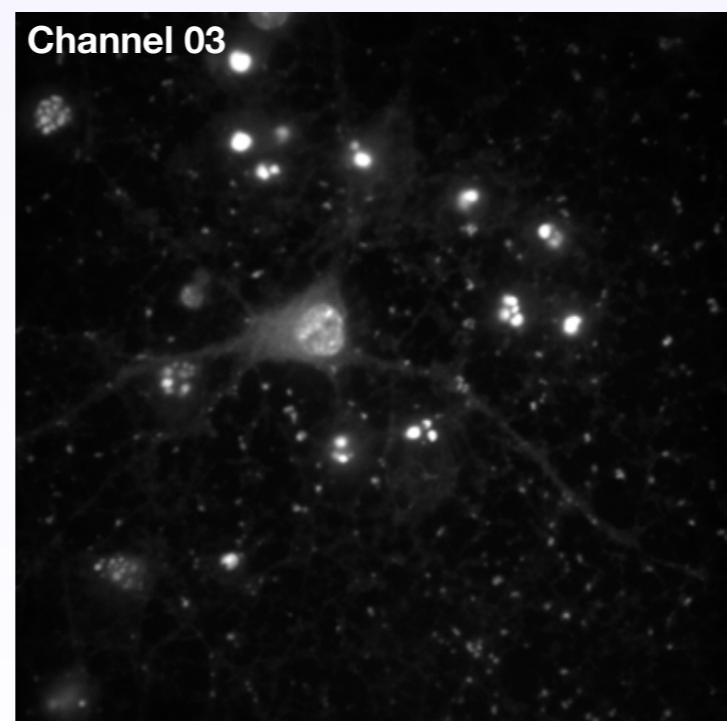
Channel 00



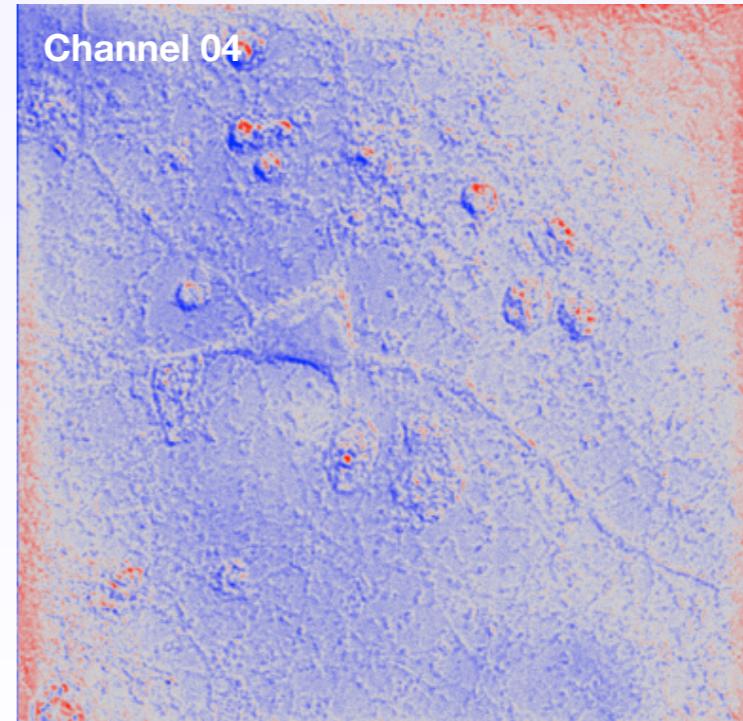
Channel 01



Channel 02



Channel 03

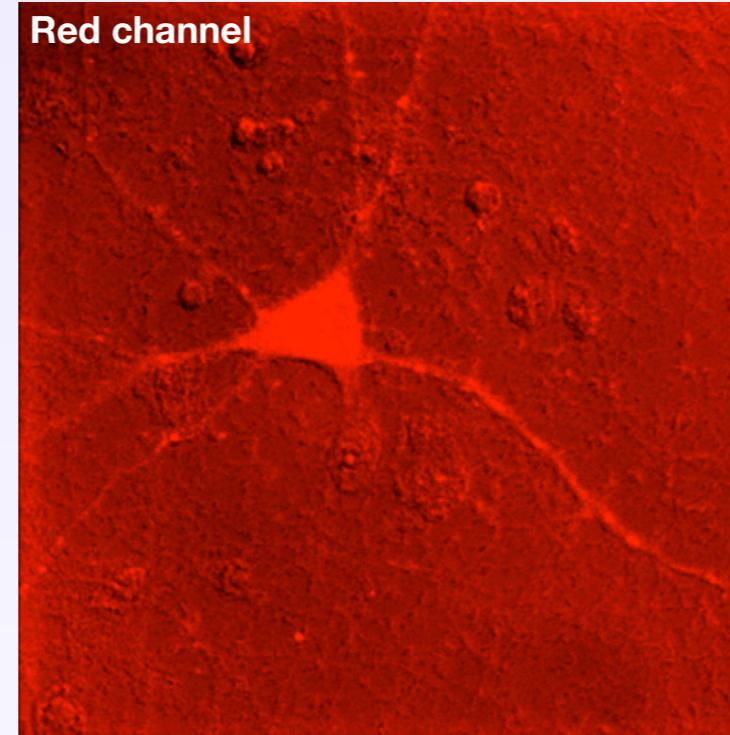
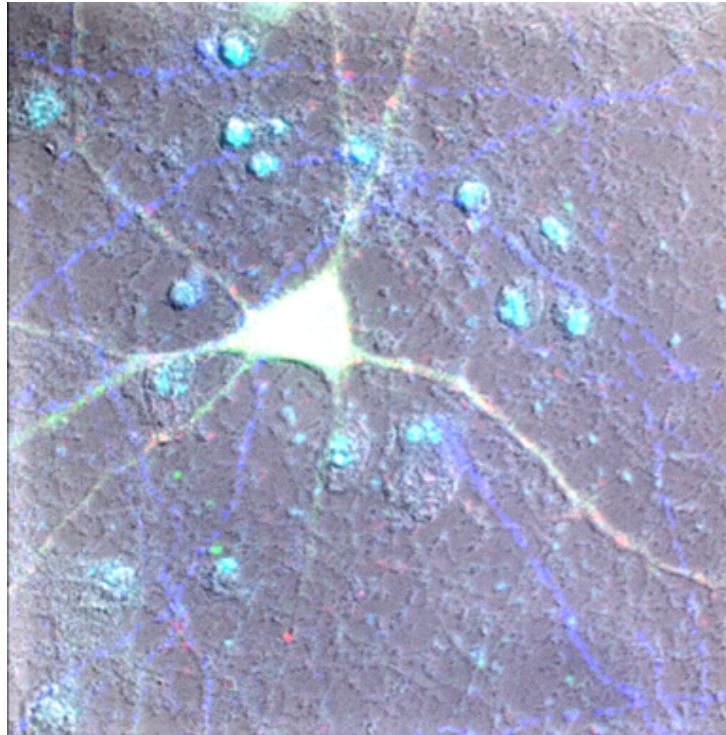


Channel 04

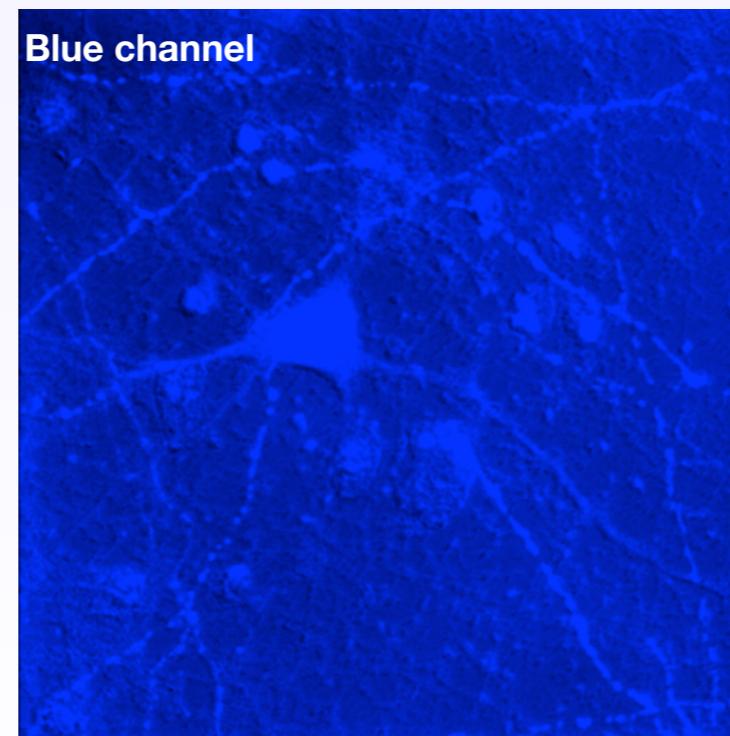
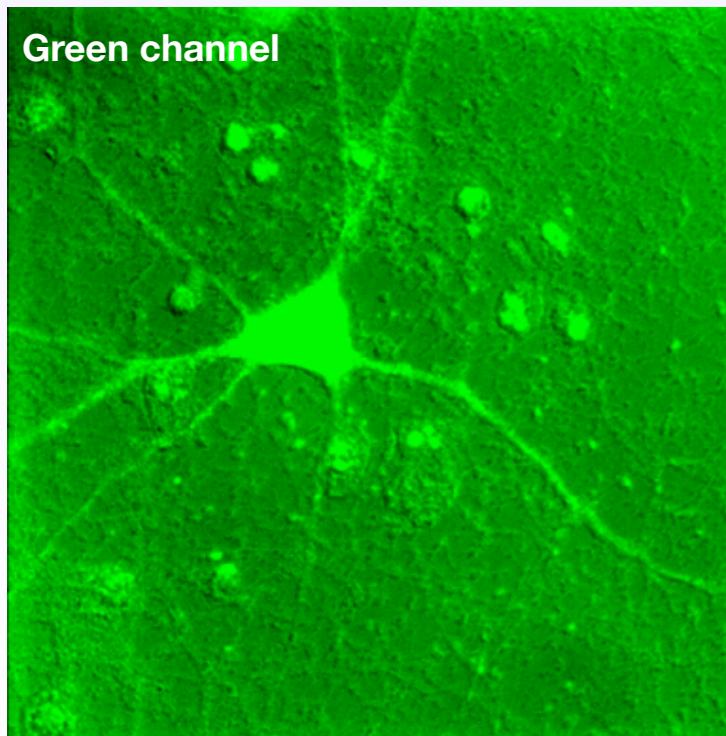
Source: ImageJ Rat_Hippocampal_Neuron.tif.

Image representation

RGB representation



If saved as a conventional image file (e.g. png or jpeg). The image intensity values are distributed amongst the red green and blue channels.



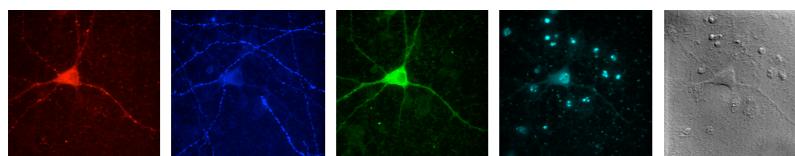
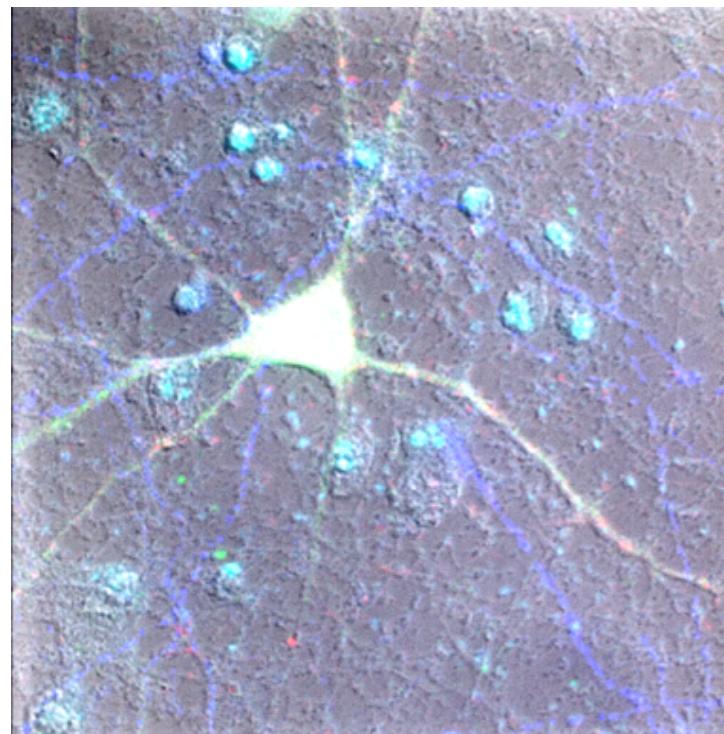
The intensity attributed to each channel is defined by the look-up-table. This is fine for making a picture or figure but bad for analysis as information is mixed and lost.

Source: ImageJ Rat_Hippocampal_Neuron.tif.

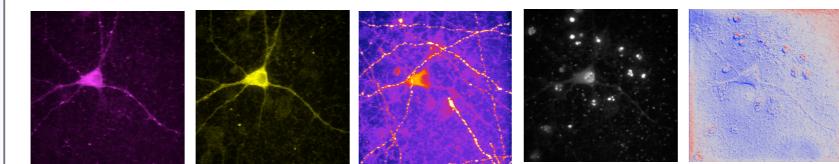
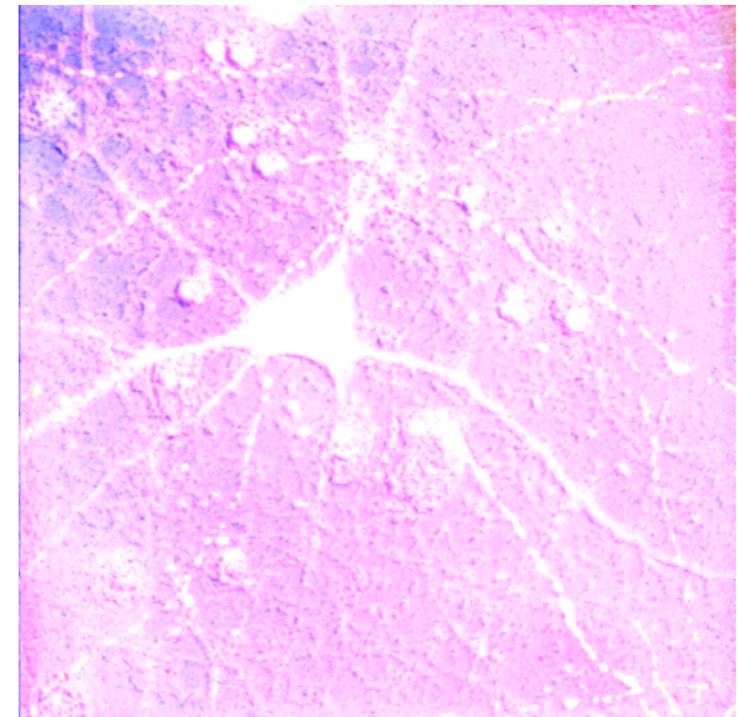
Image inspection

Same information just represented with different colours

Composite representation



Composite representation



Remember to save in TIFF format if you have many channels (multi-page format).
Don't be tricked by the colour of something. Use Fiji to inspect.

Source: <http://fiji.sc/Fiji>

Image -> color -> Channels Tool

Imaging pitfalls: Compression

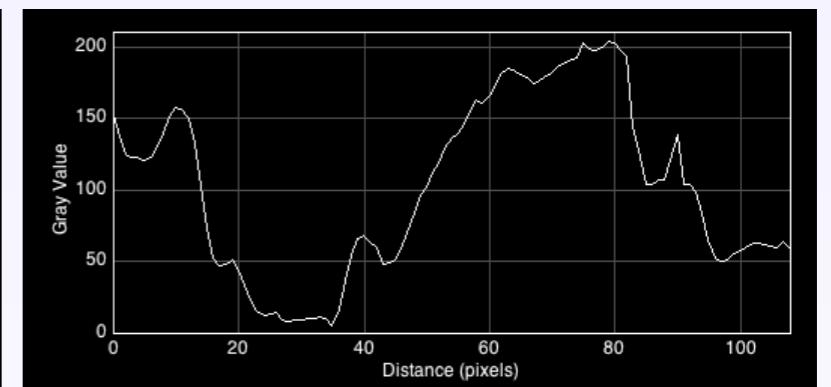
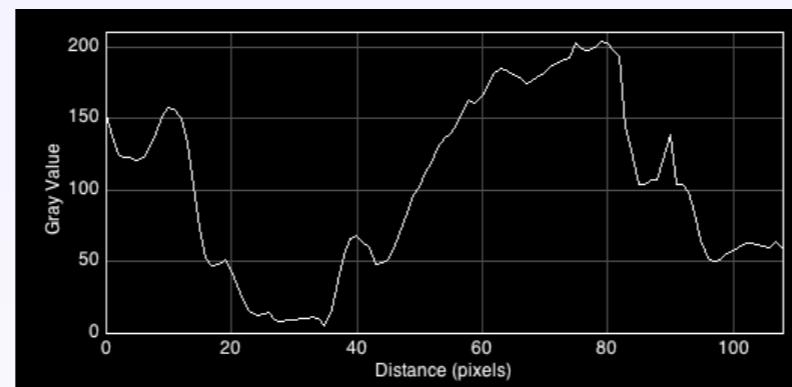
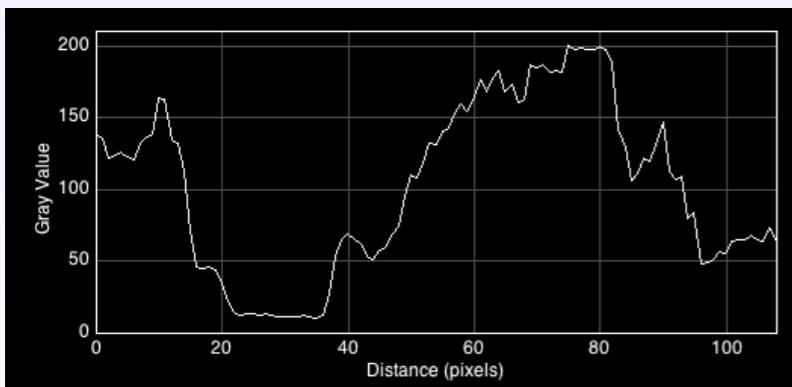
no Compression



medium Compression



max Compression



JPEG, movie formats = LOSSY & BAD (BMP, GIF, PNG also not great)
Compression also reduces certain colours more than others (e.g. blue tones).
TIFF is good, especially OME-TIFF

Source: http://upload.wikimedia.org/wikipedia/commons/c/ce/Quality_comparison_jpg_vs_saveforweb.jpg

Drinks and how to drink them (i.e. data import speeds)



In the old days, the straw was a good size for the drink, the drink could be drank quickly. E.g. How Fiji/ImageJ would load a TIFF/CZI/LIF... file.

Source: image.sc

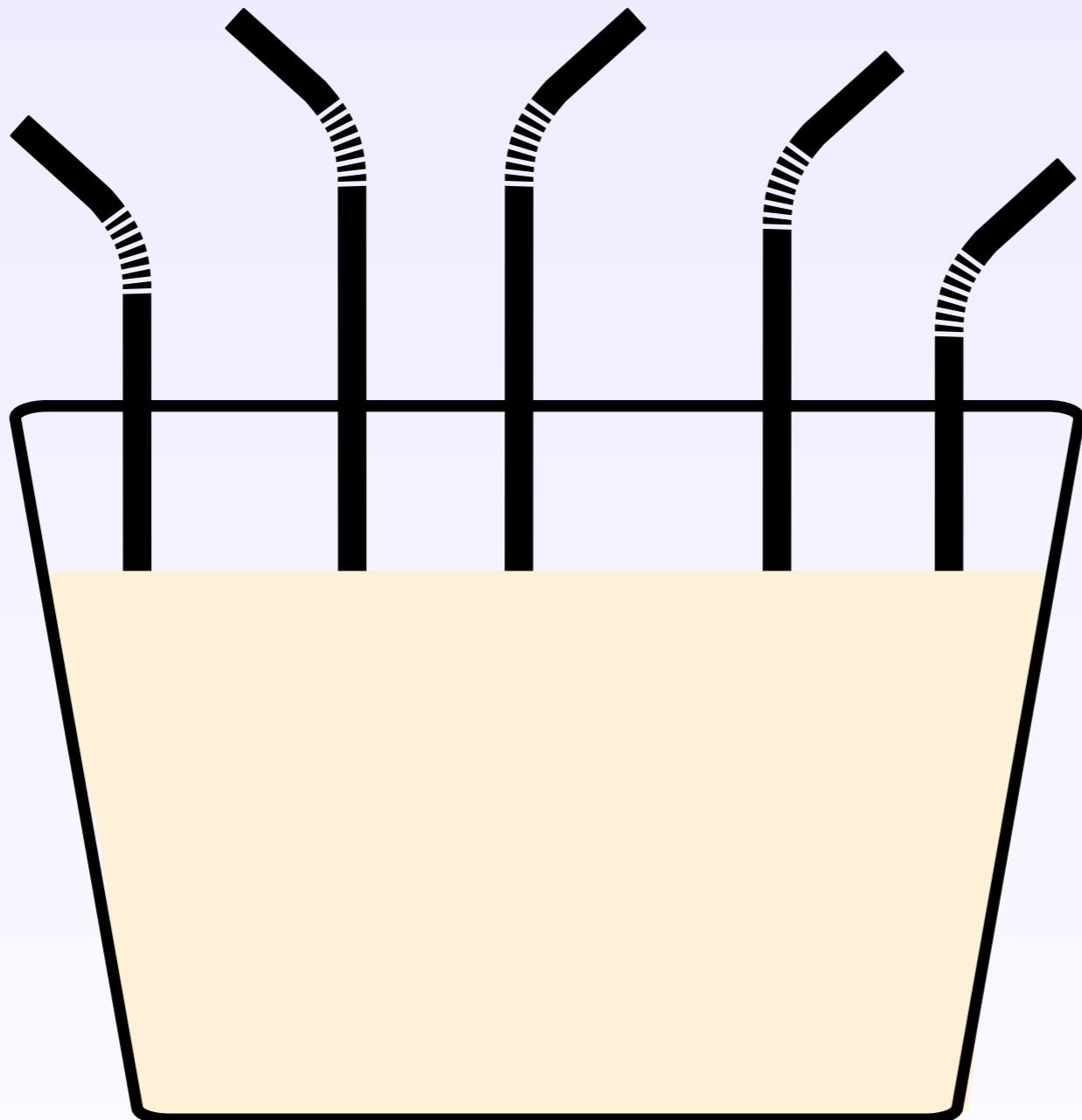
Drinks and how to drink them.



Drinks got bigger. The straw has remained the same. Microscopes will create very large files 1GB-1TB and still use the same file format. Takes long time to drink in all the data.

Source: image.sc

Drinks and how to drink them.



With many straws and many people you can drink the drink very quickly.
Computers can handle multiple straws, but the drinks containers aren't really there

Source: image.sc

HDF5 a solution to the large file format?

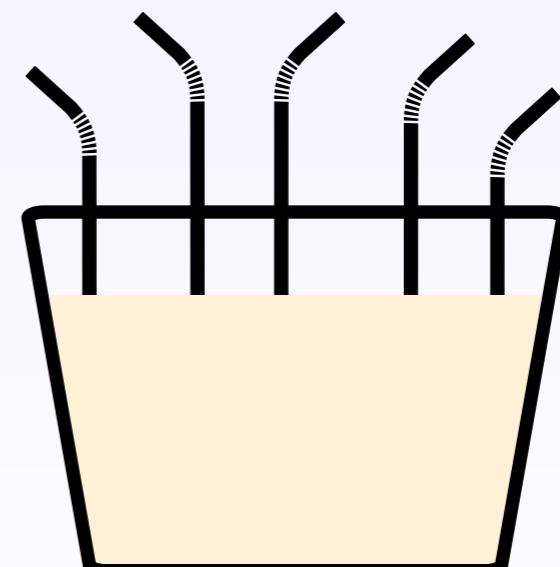
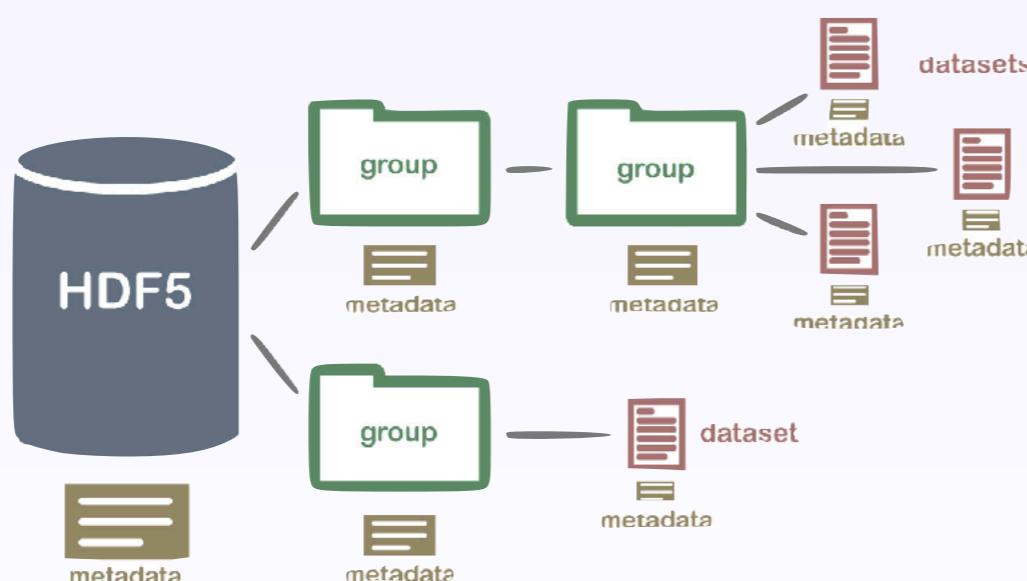
Conventional binary formats (e.g. TIFF file).

- Reads data sequentially, generally on one CPU thread at time.



File format which is useful for big-data e.g. (HDF5) Hierarchical Data Format v5:

- Random Access
- Parallel execution (CPUs have multiple cores and threads).



Source: <https://stackoverflow.com/questions/22125778/how-is-hdf5-different-from-a-folder-with-files>

HDF5 - several solutions as present

Unfortunately, like TIFF there are multiple types of HDF file. Each with their own metadata:

- SVI Huygens (deconvolution software) have a type <https://svi.nl/HDF5>
- Electron Microscopy Datasets <https://emdatasets.com/format/>
- BigDataViewer (BDV) format (https://imagej.net/BigDataViewer#Multi-view_dataConverted_to_XML)
- Screening HDF format (<https://github.com/CellH5>)

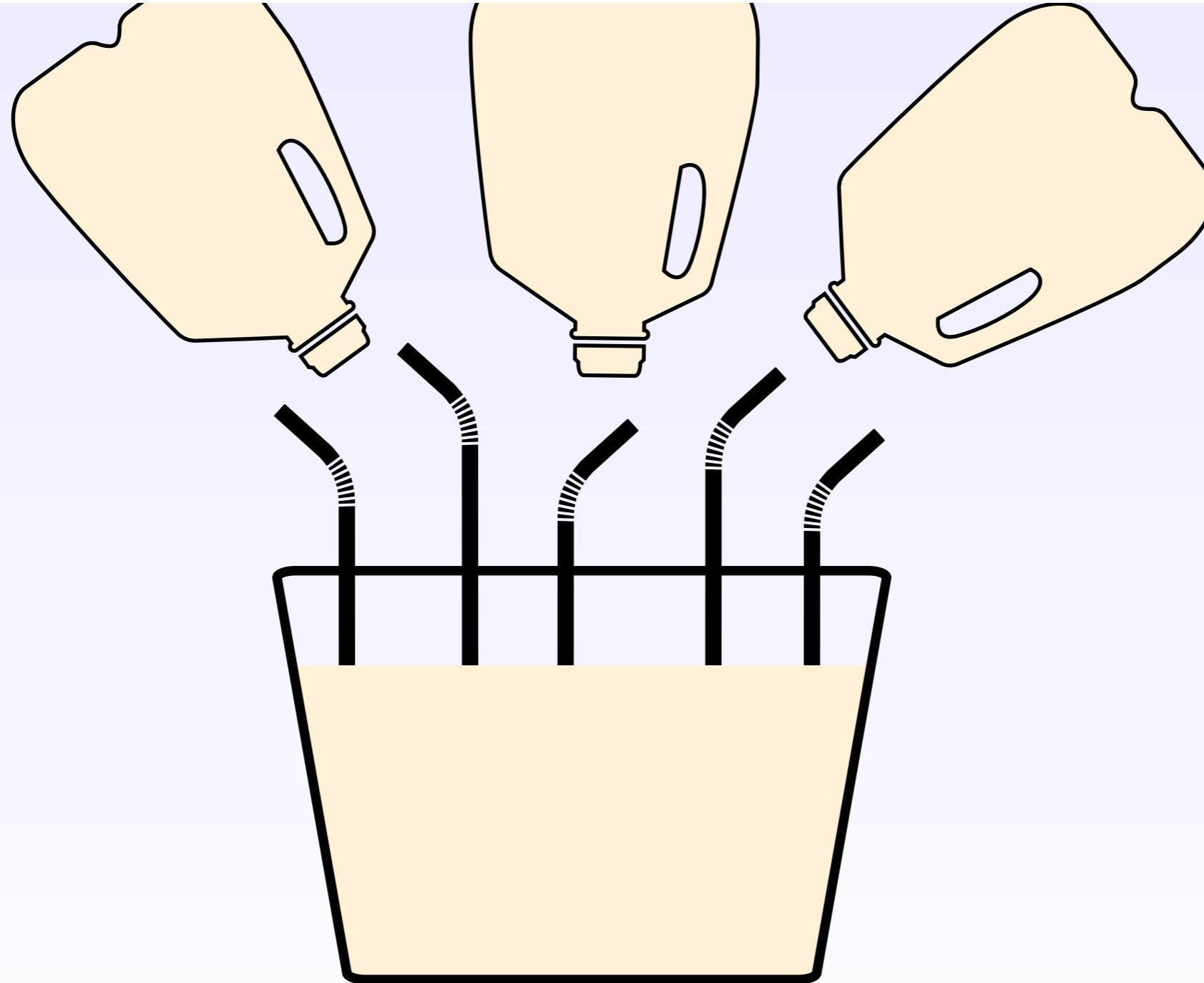
Attempts to unify:

- BigDataViewer (BDV) format supported by Bioformats (<https://forum.image.sc/t/ome-s-position-regarding-file-formats/26952>)
- Unifying Biological Image Formats with HDF5 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3016045/>)
- Bioformats H5 Reader (<https://docs.openmicroscopy.org/bio-formats/5.8.2/metadata/CellH5Reader.html>)

OME (<https://forum.image.sc/t/ome-s-position-regarding-file-formats/26952>)

Source:

It is not just about drinking, but also pouring.



Next Generation formats work on fast parallel saving and loading and also cloud based solutions. Zarr, N5 formats.

Source: <https://zarr-developers.github.io/> <https://github.com/saalfeldlab/n5>

HDF5 for Python

HDF5 for Python

[Downloads](#) [Documentation](#) [GitHub Project](#)



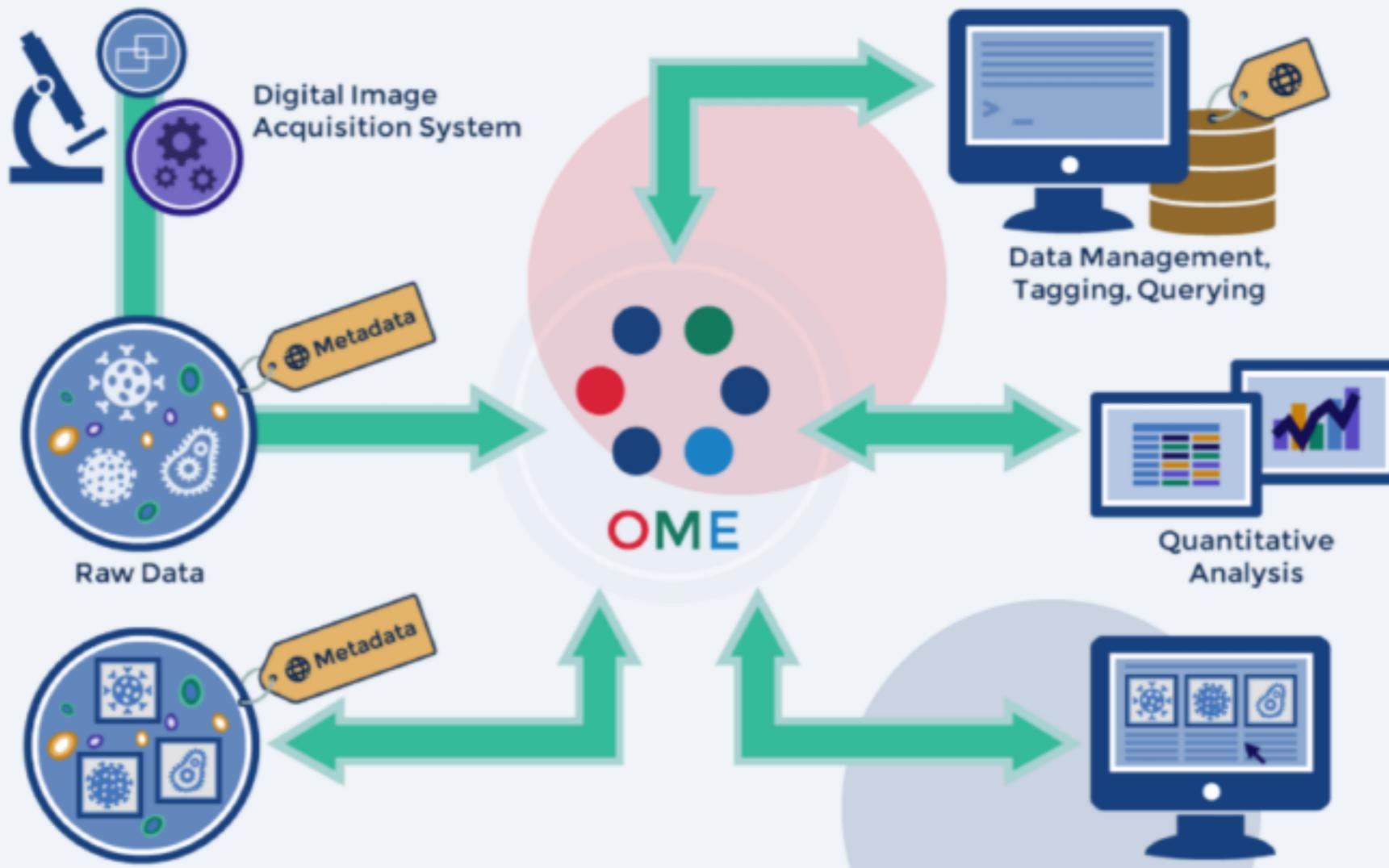
We recommend you check out this resource for using and managing hdf5 in Python.

Source: <https://www.h5py.org/>

Introduction to OME and OMERO

OMERO: Introduction: OME

OME: Interoperability

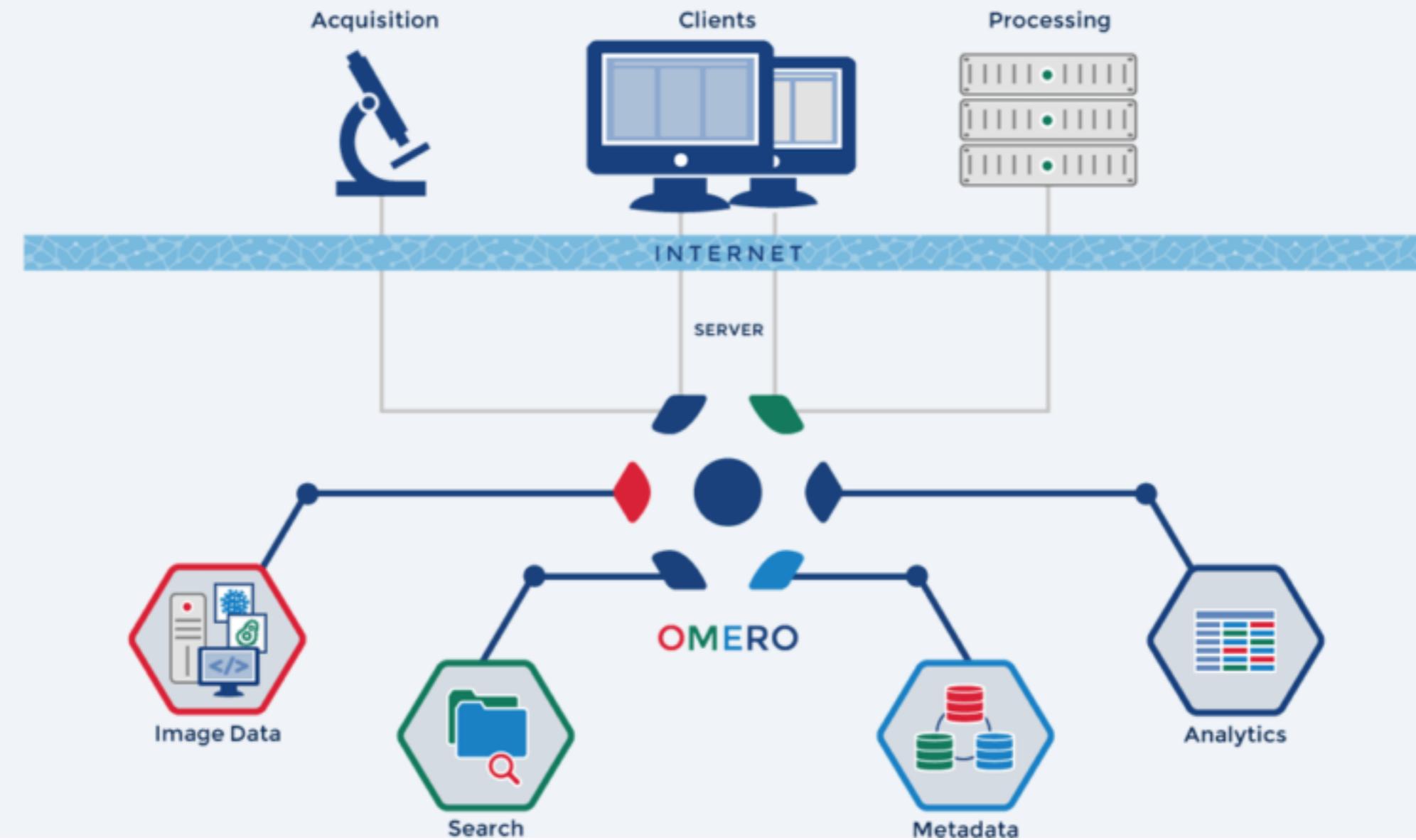


OME traditionally write software/formats which allows enables interchangeability between microscopes and analysis software.

Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

OMERO is a server with clients

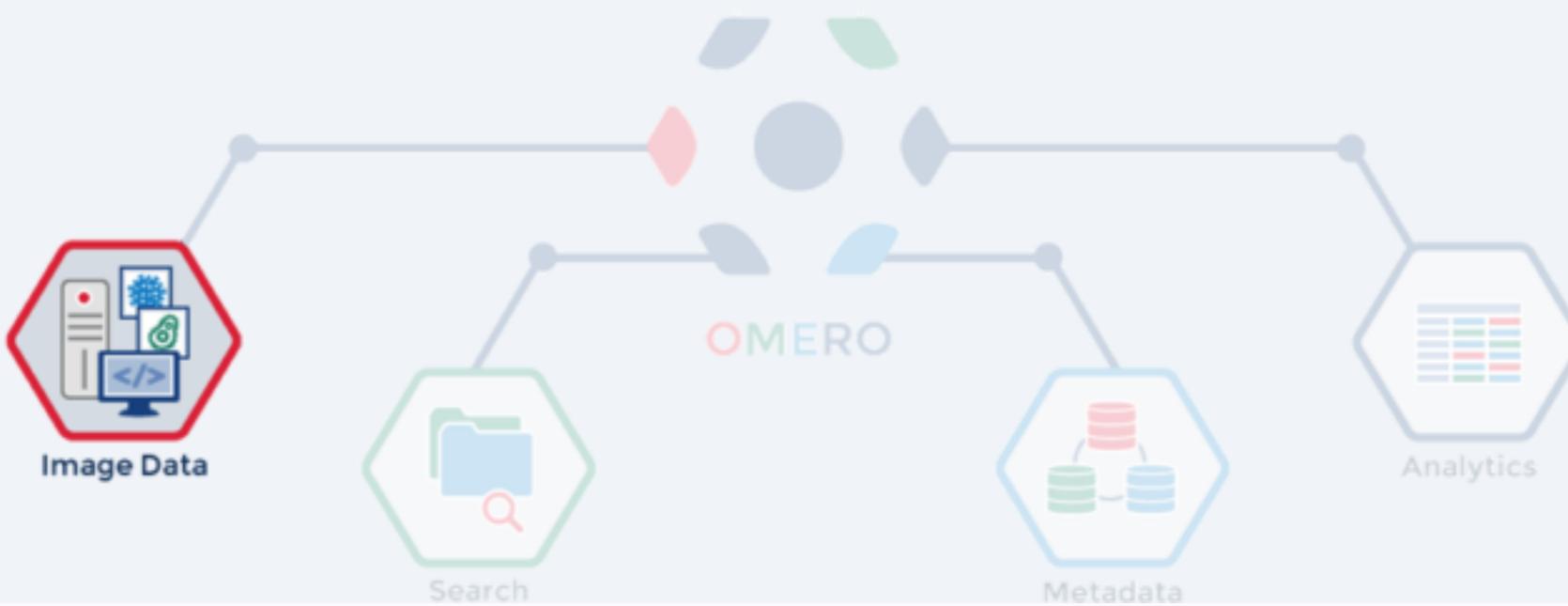


Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

Image Data - Read by Bio-Formats

- Support for reading > **150 image formats**
- Read pixel data and metadata
- Includes 5D images, HCS data, Tiled WSI images
- Domains: biological, medical, general (tiff, png, etc.)



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

Metadata

- Stored in a Relational Database
- Acquisition Metadata
- User-added Annotations
- ROIs / Segmentations

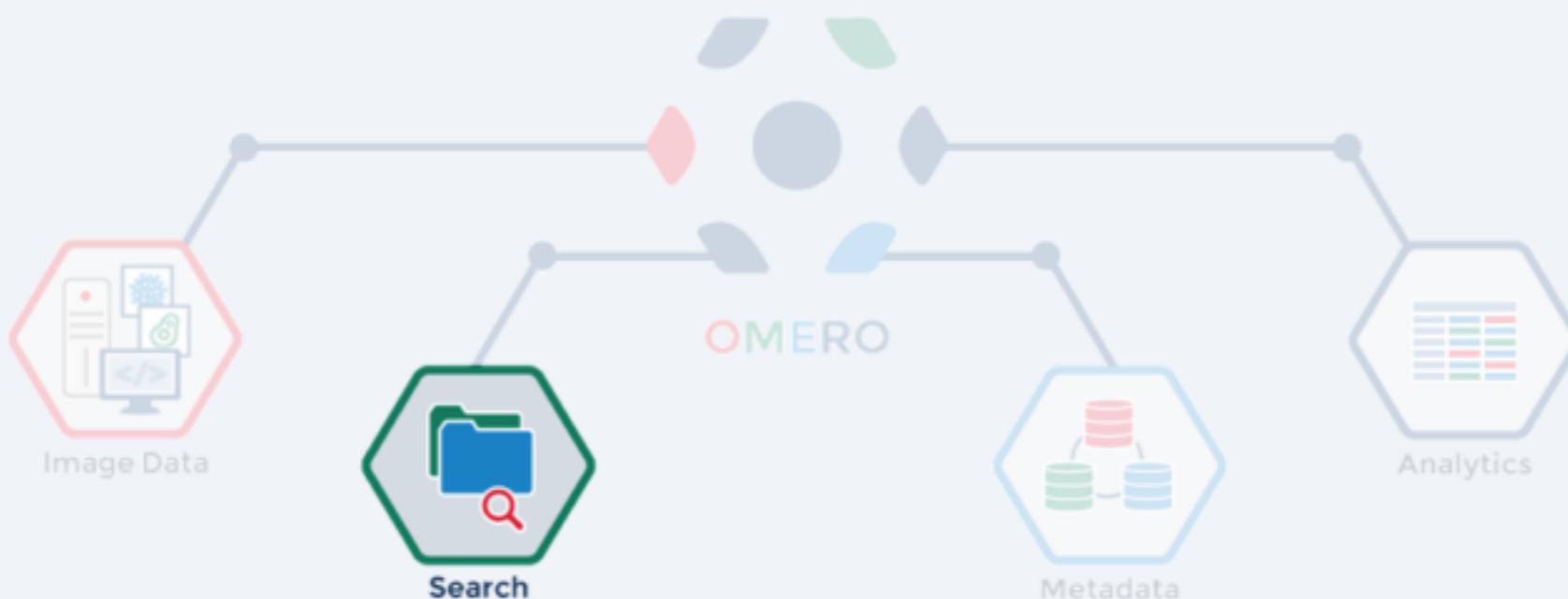


Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

Search

- Text and Annotations indexed with Lucene
- Allows keyword searches



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

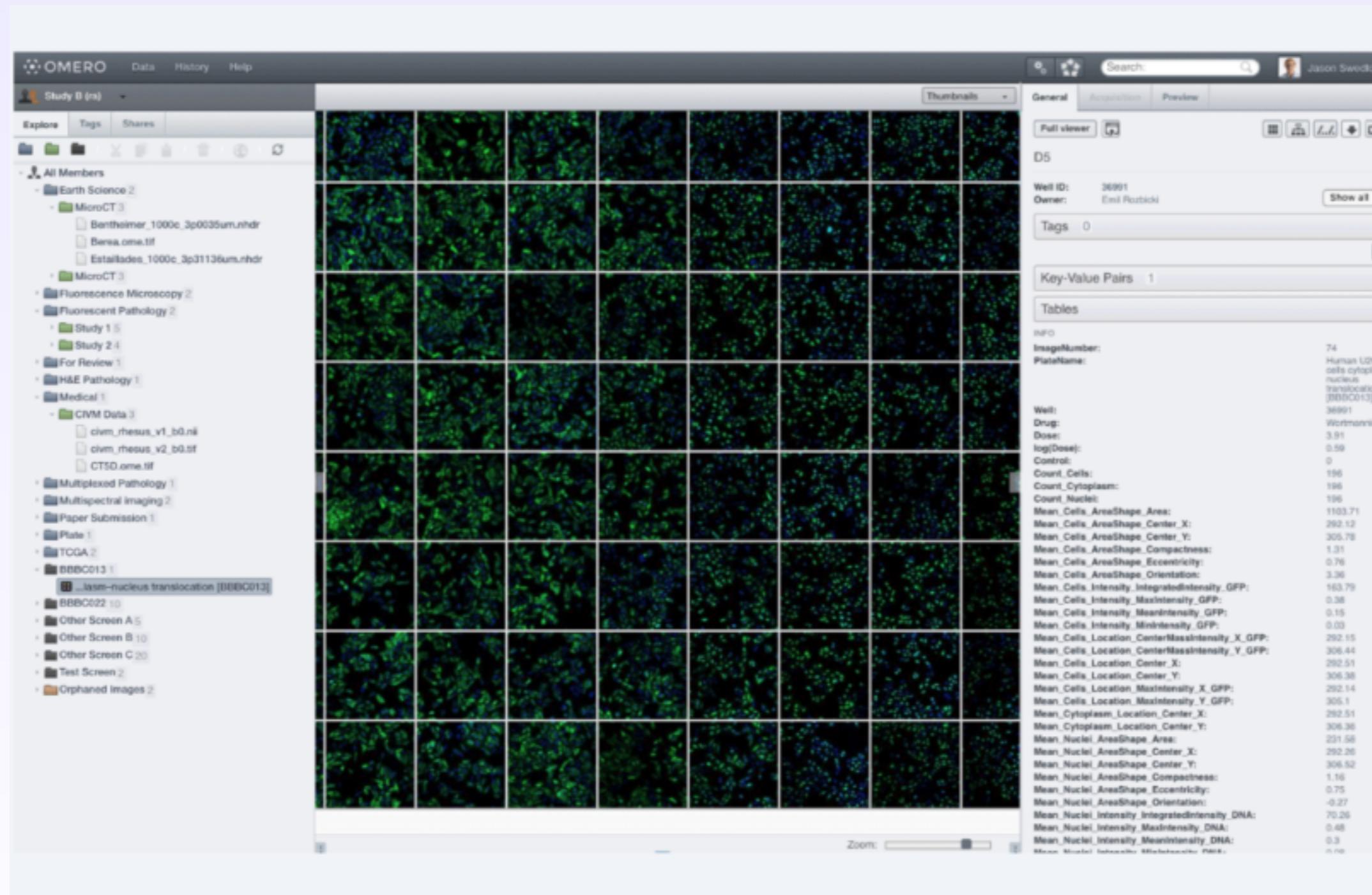
OMERO.tables

- HDF5 tables for analysis results
- Flexible columns



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

The screenshot shows the OMERO web interface. The top navigation bar includes links for OMERO, Data, History, and Help, along with a search bar and user profile for Jason. The main area displays a dataset named "Study B (ra)". On the left, a sidebar lists study members and a folder named "ACC 231" containing numerous image files. The central area shows a grid of thumbnail images labeled "ACC". To the right, a detailed panel titled "Image Details" provides technical information about the images, such as acquisition date (2013-01-22), import date (2015-11-24), dimensions (113287 x 40455 pixels), and various counts and areas related to the images.

General Acquisition Preview

Image Details

Aperio Image Library vFS90 a asd

Acquisition Date: 2013-01-22 02:23:31
Import Date: 2015-11-24 05:40:14
Dimensions (XY): 113287 x 40455
Pixels Type: uint8
Pixels Size (XYZ) (µm): 0.25 x 0.25 x -
Z-sections/Timepoints: 1 x 1
Channels: Channel1, Channel2, Channel3
ROI Count: 0

Tags 3

Key-Value Pairs 0

Tables

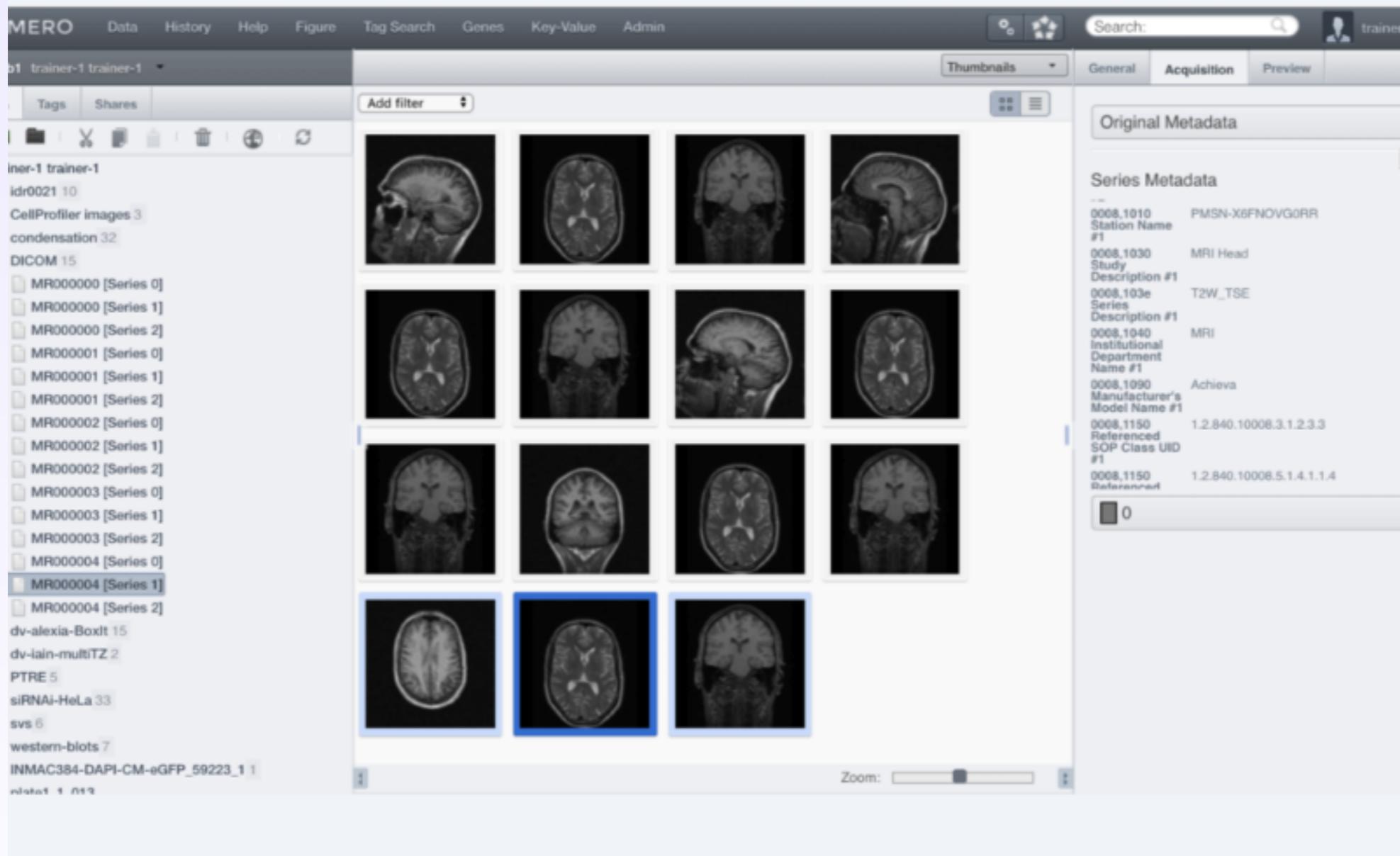
INFO

Dataset:	2601
Image:	156028
Nuclear Count:	214782
Stromal Count:	118372
Epithelial Count:	96410
Stromal Area:	0.55
Epithelial Area:	0.45
Image Name:	TCGA-OR-A5J2-01A-01-TS1.F951E65D-4231-83AB-D17520D1AC95.svs

Attachments 0

Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

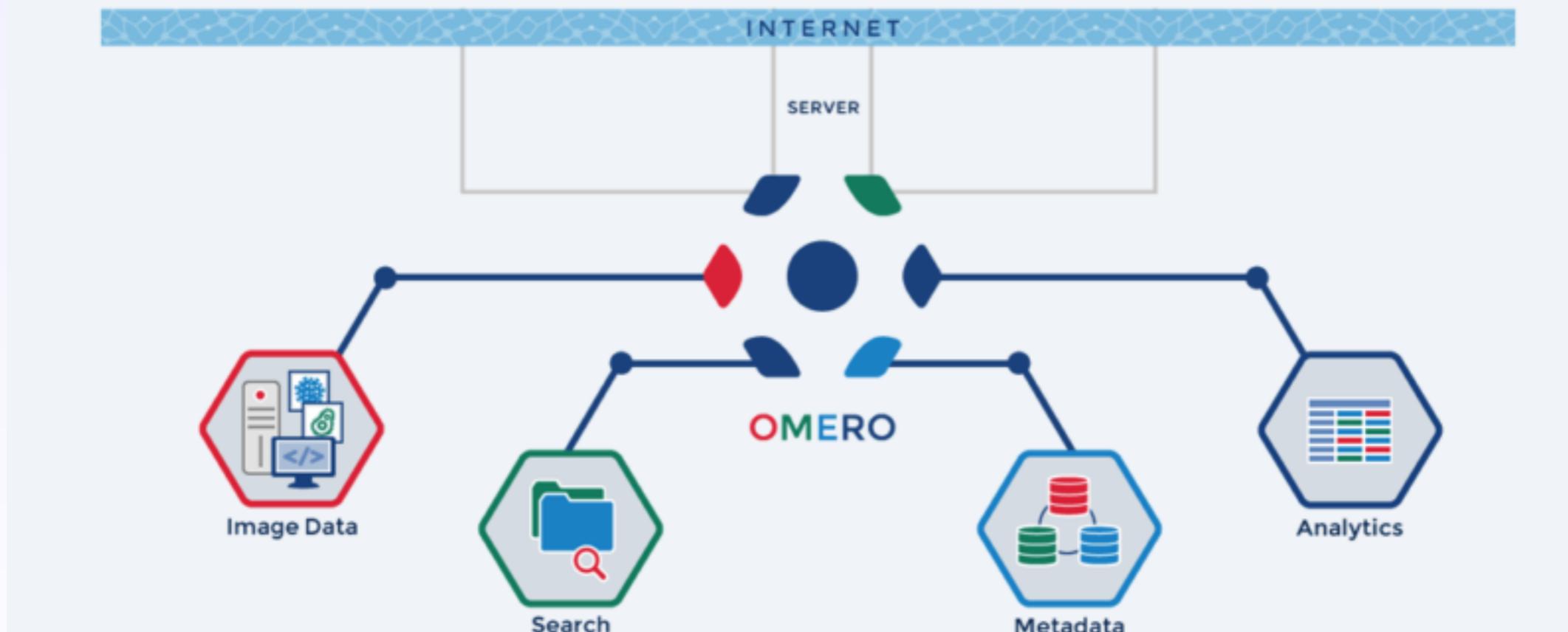


Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

OMERO API

Python | Java | Matlab | R | C++

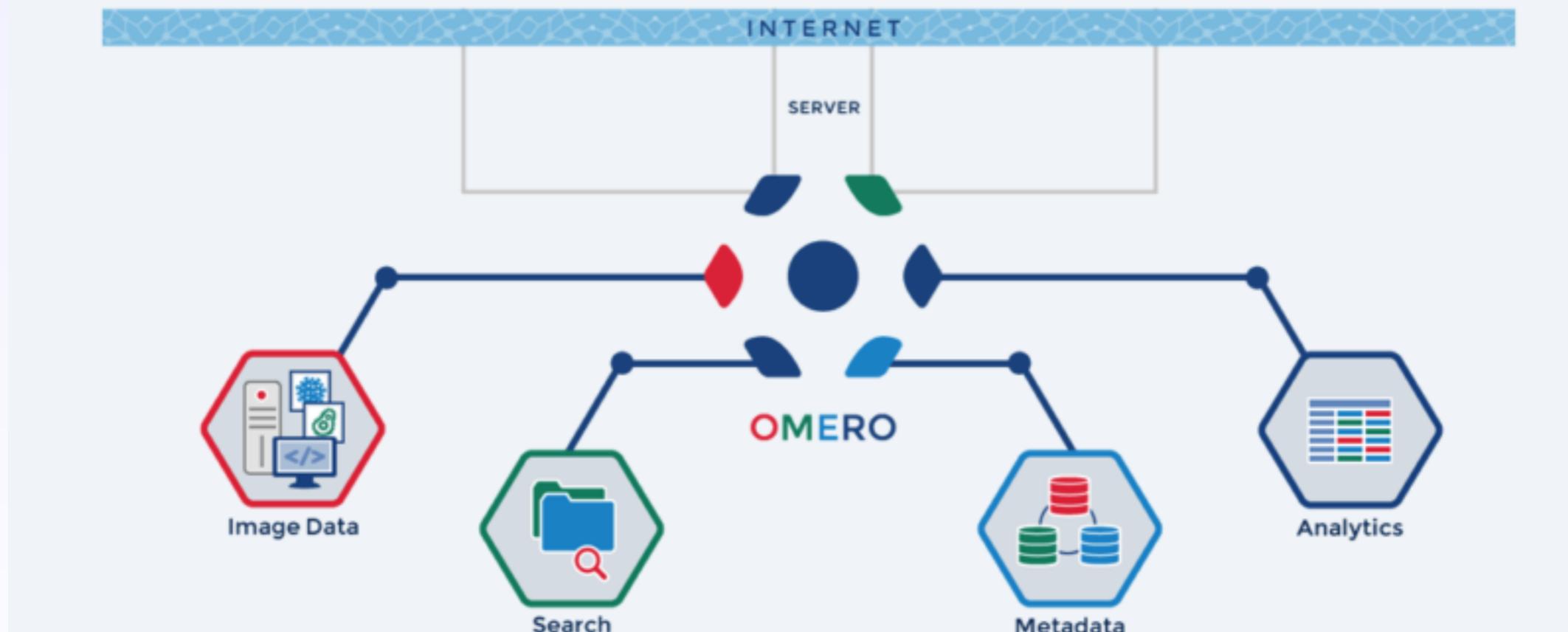


Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko:

OMERO: Introduction

OMERO API

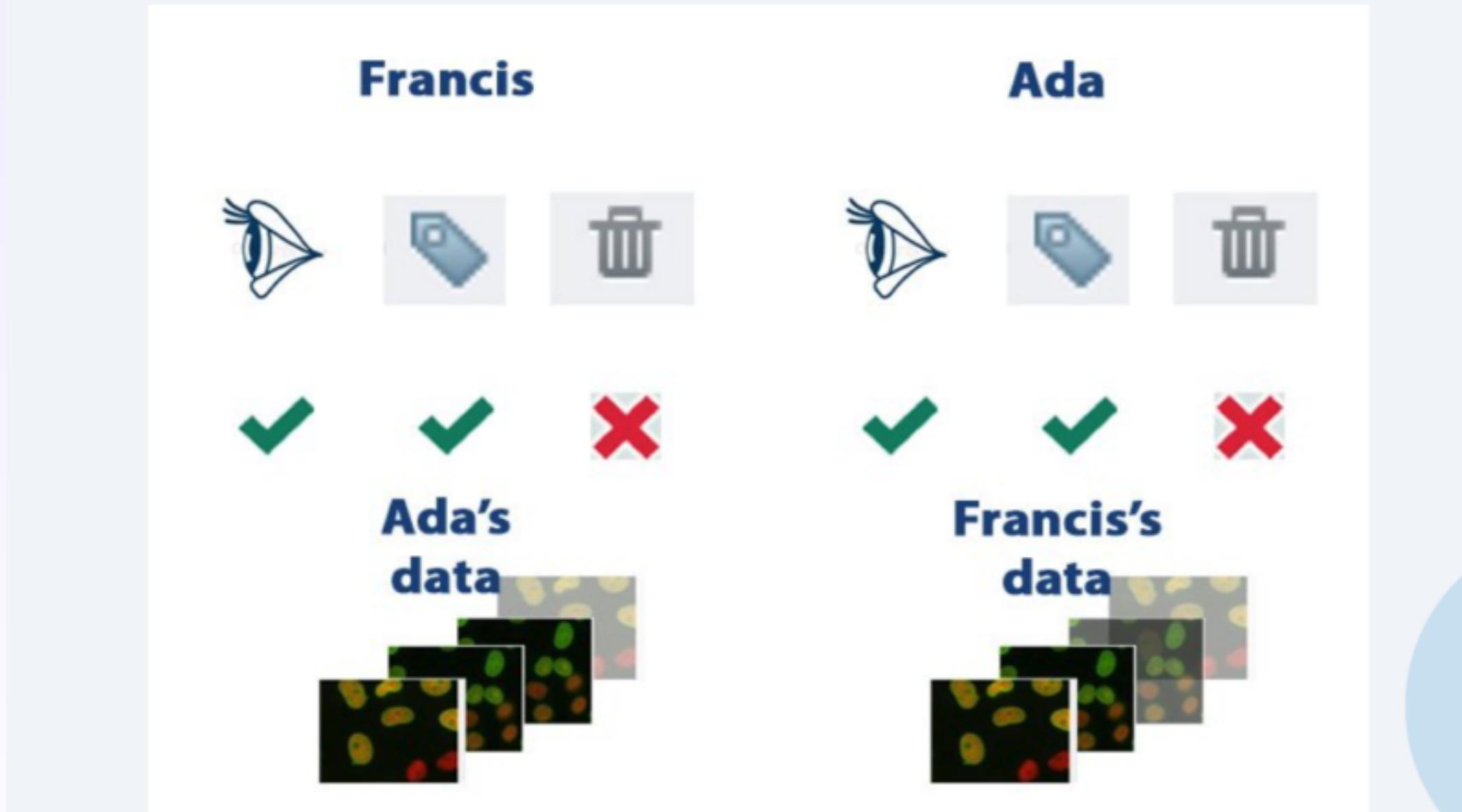
Python | Java | Matlab | R | C++



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

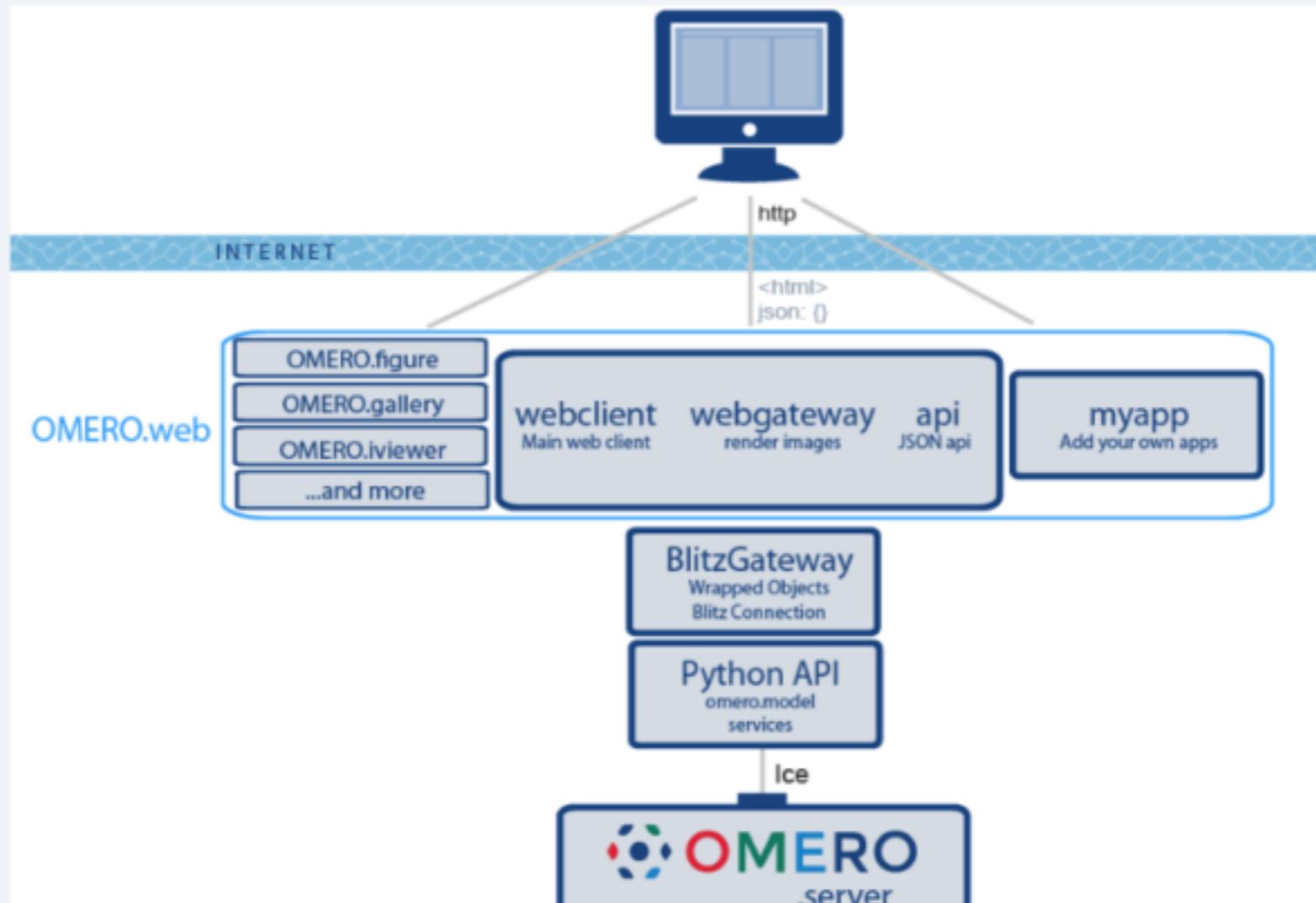
Each other's data – read-annotate setup



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

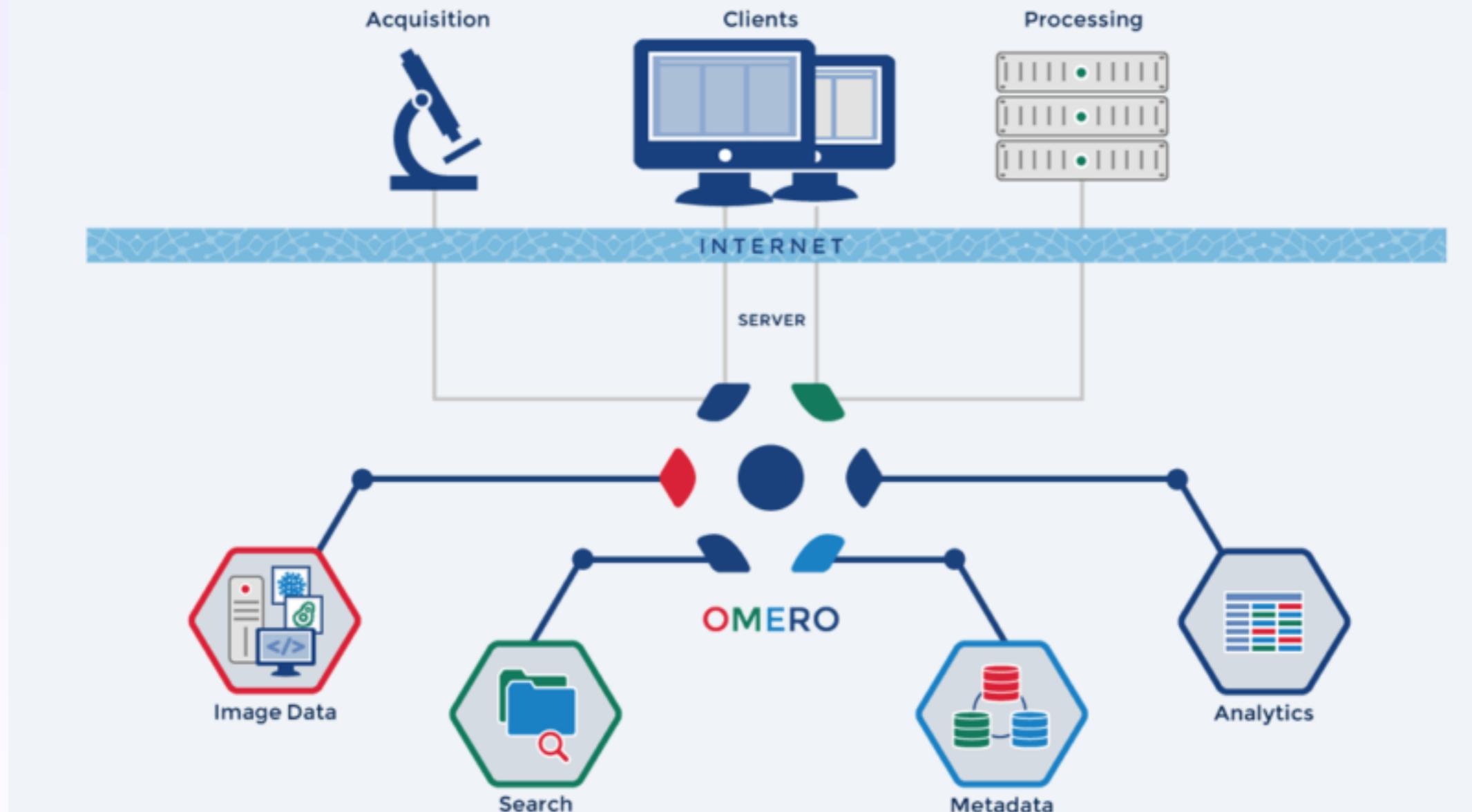
OMERO.web framework is extensible



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

Different clients for different jobs



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

OMERO clients

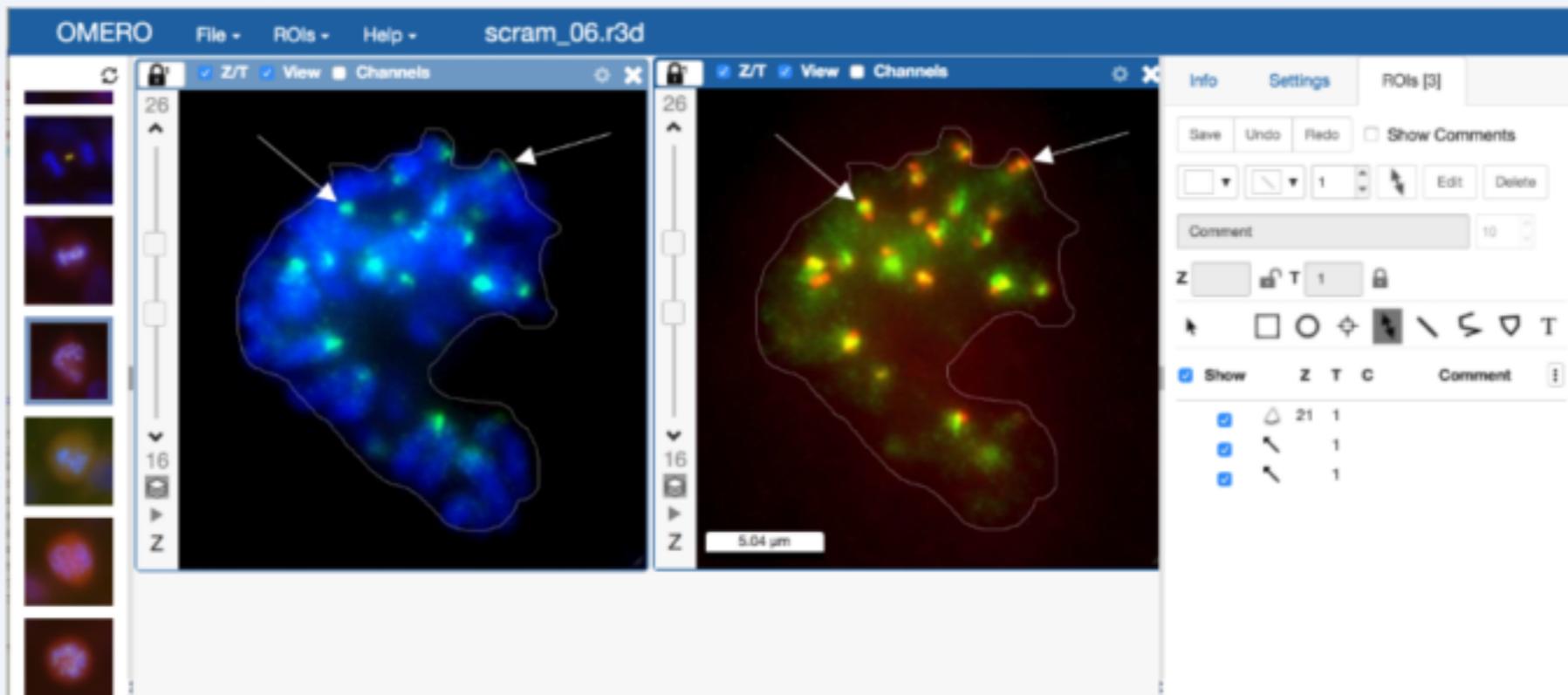


- OMERO.insight
- OMERO.webclient
 - Browsing images
 - Viewing metadata
 - Data management
 - Searching
- OMERO web apps
 - OMERO.iviewer
 - OMERO.figure
 - OMERO.parade
 - ...
- ...

Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

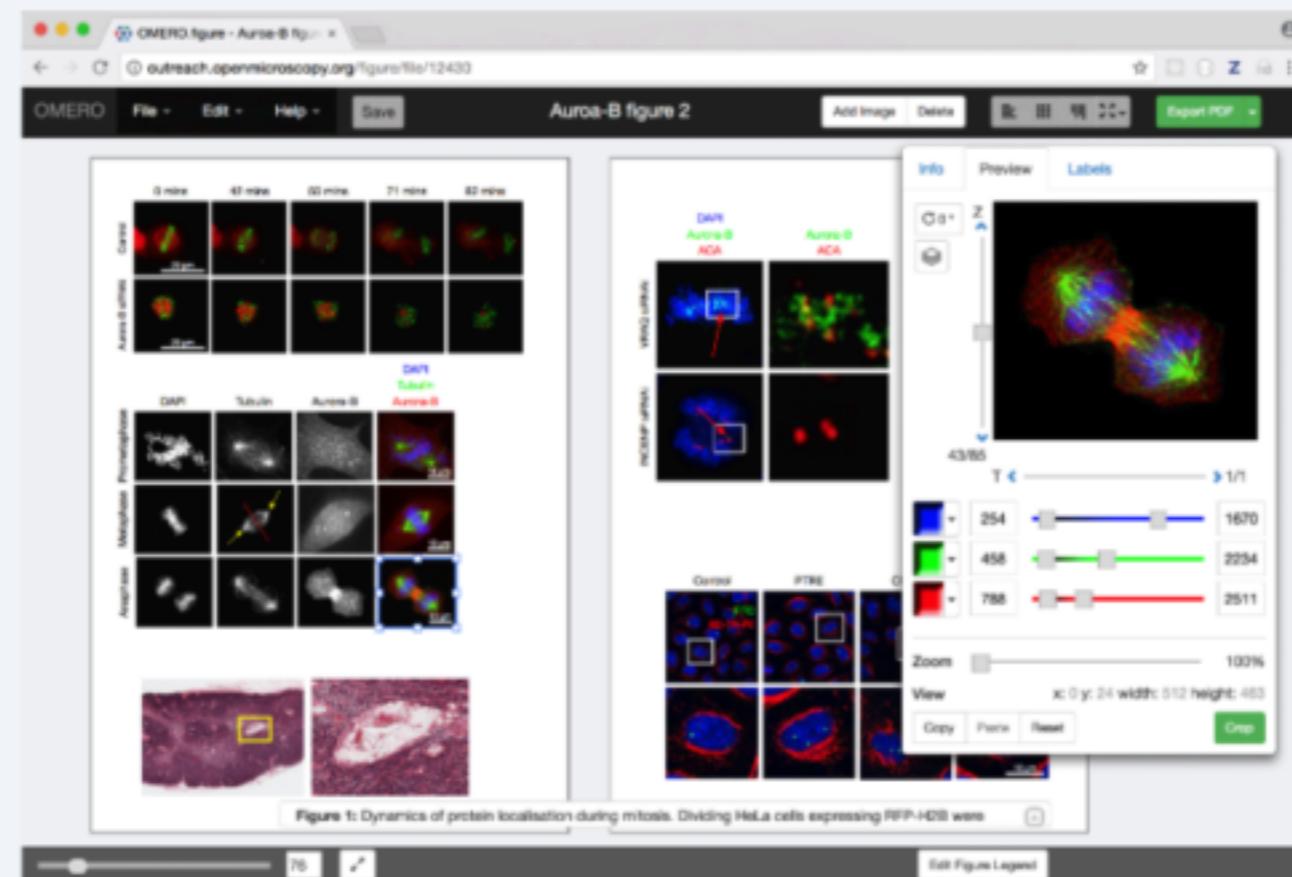
OMERO.iviewer



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction

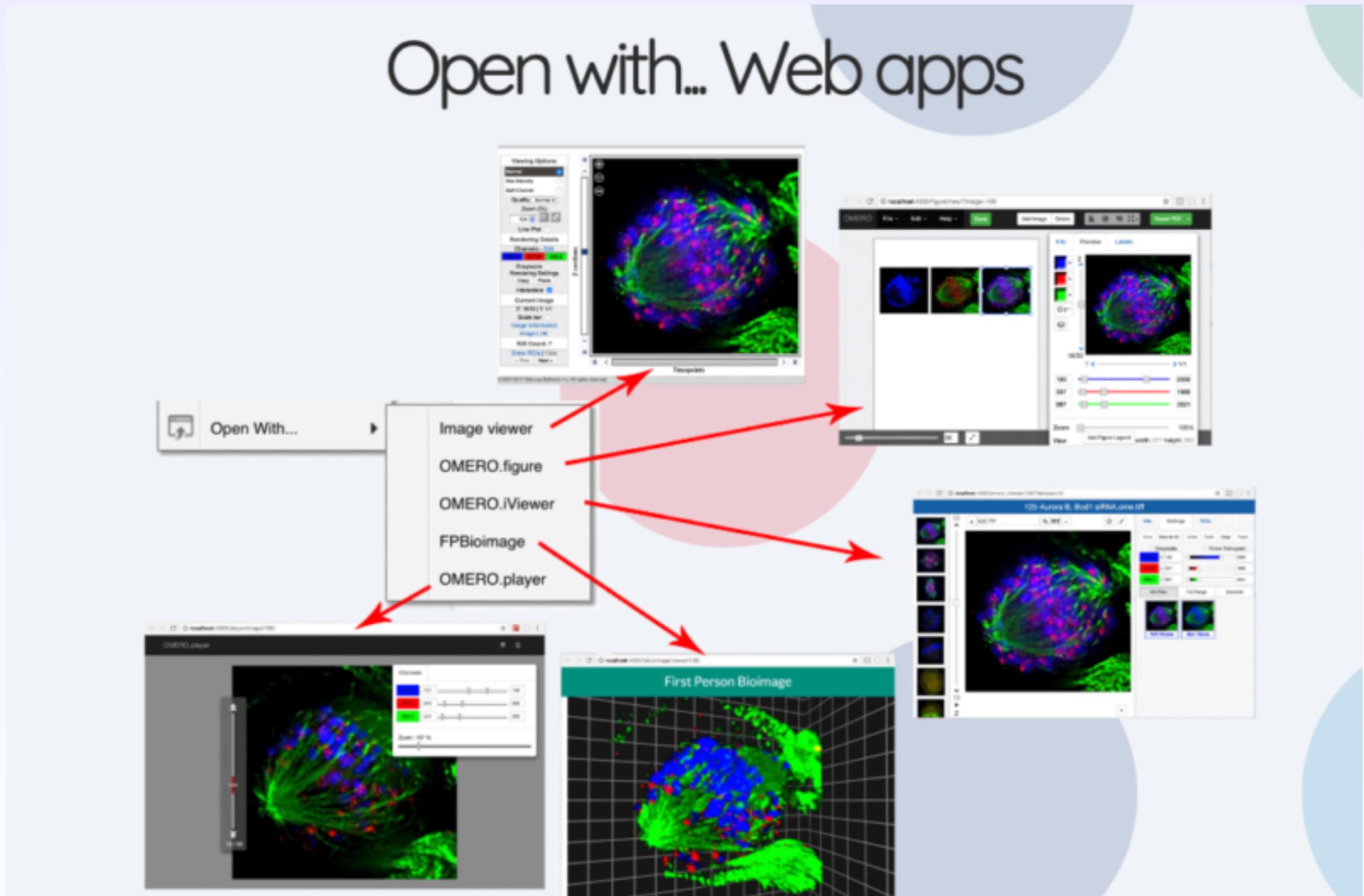
OMERO.figure



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

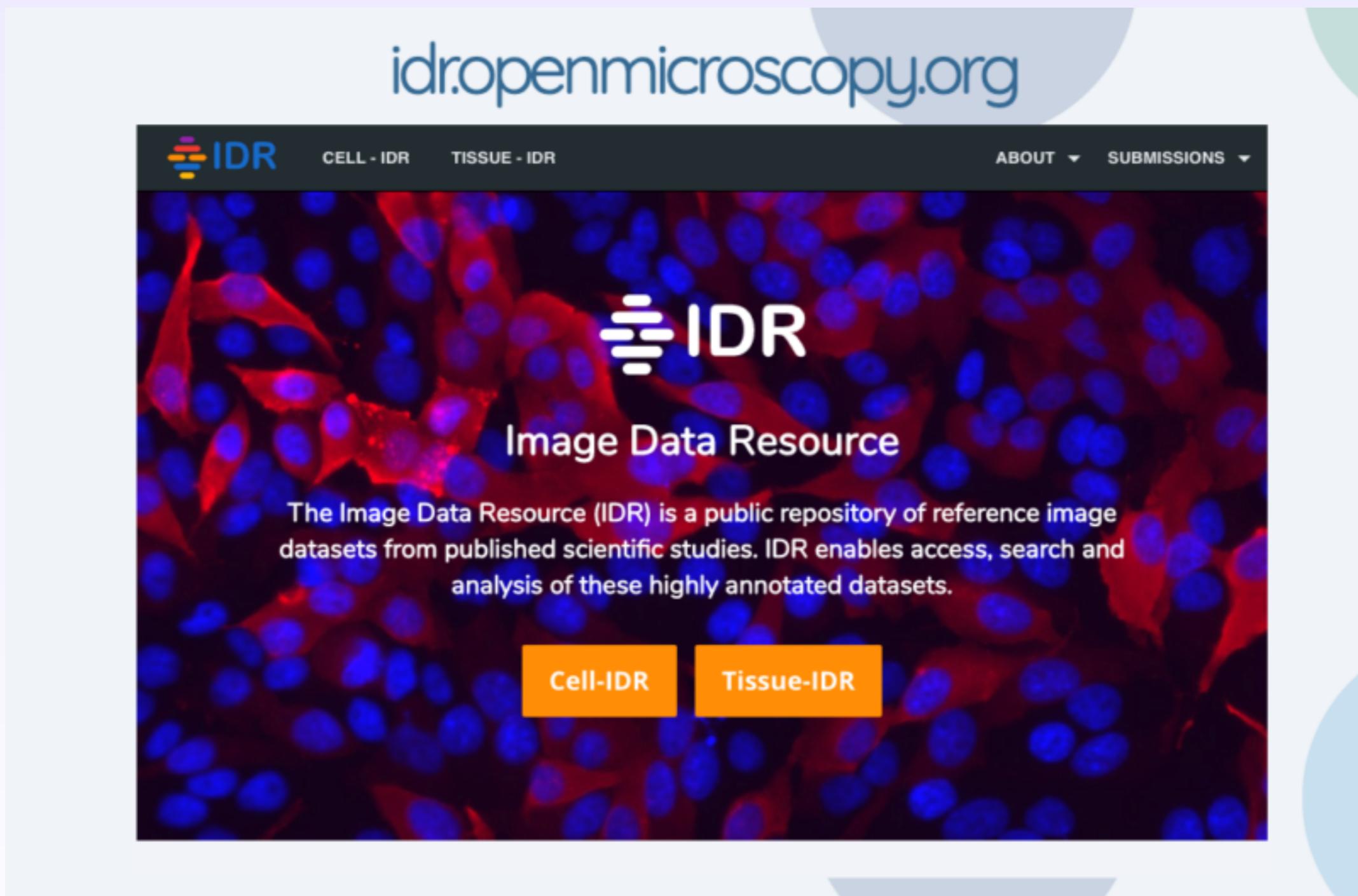
OMERO: Introduction

Open with... Web apps



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO: Introduction



Source: <https://downloads.openmicroscopy.org/> Dominik Lindner, Petr Walczysko

OMERO DEMO with Python

Source:

Last slide, a chance to promote your things

For this content and more:

<https://github.com/IAFIG-RMS/Bioimage-training>

<https://twitter.com/dwaithe>



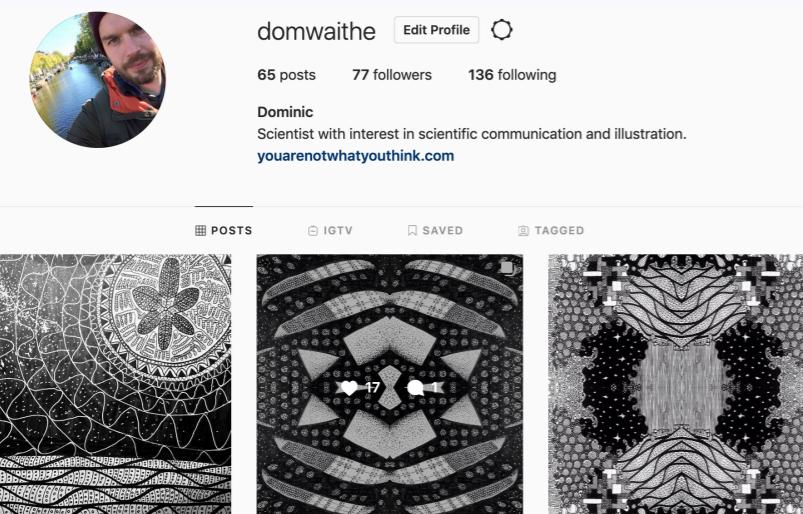
A screenshot of a Twitter profile for 'Dominic Waithe' (@dwaithe). The profile picture shows a man with a beard. The bio reads: 'Dominic Waithe @dwaithe'. Below the bio, it says 'Tweets 401', 'Following 157', and 'Followers 196'. The background of the profile page is a photograph of a modern building in a field at sunset.



<https://github.com/dwaithe>



<https://instagram.com/dwaithe>



A screenshot of an Instagram profile for 'domwaithe'. The profile picture shows a man in a red jacket. The bio reads: 'domwaithe Edit Profile'. Below the bio, it says '65 posts', '77 followers', and '136 following'. The bio continues: 'Dominic Scientist with interest in scientific communication and illustration. youarenotwhatyouthink.com'. Below the bio, there are three thumbnail images showing intricate scientific illustrations of biological structures.

dominic.waithe@imm.ox.ac.uk



UK Research
and Innovation



Sources