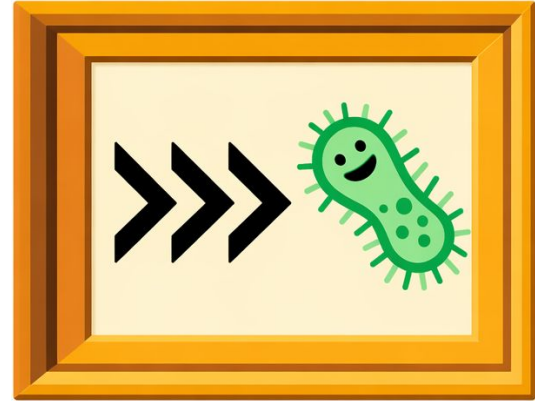


# OME-Arrow

Building data relationships  
for bioimage analysis

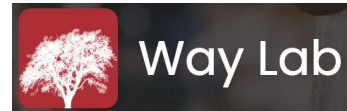


**Dave Buntten**

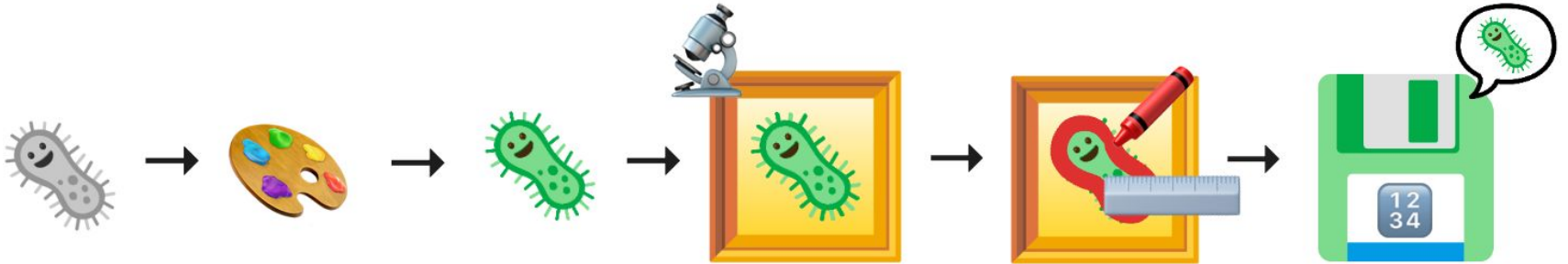
Principal Research Software Engineer  
University of Colorado Anschutz  
Department of Biomedical Informatics



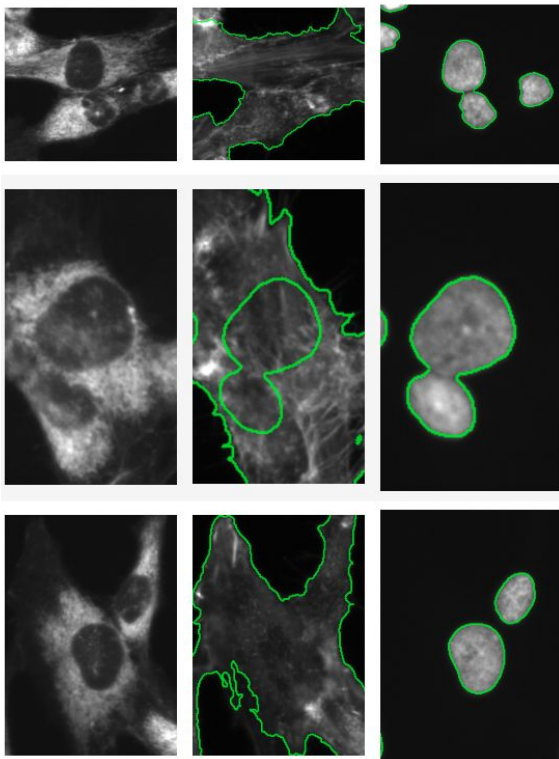
Anschutz



# Leveraging morphology for biological insight



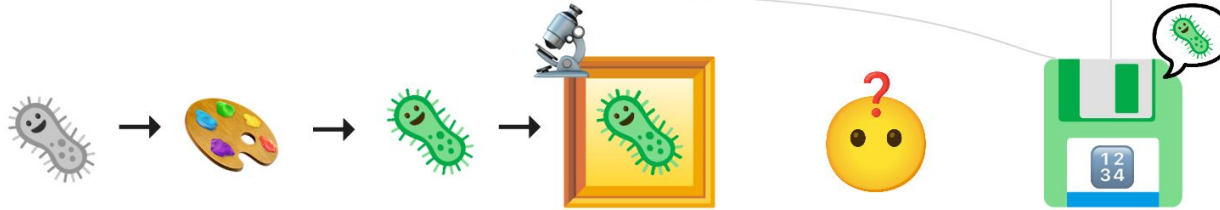
# Quantitative features benefit from image context



Metadata_ImageNumber	Metadata_Cells_Number_Object_Number	cqc.large_nuclei.is_outlier	
699	50	2	True
1557	113	10	True
568	45	9	False

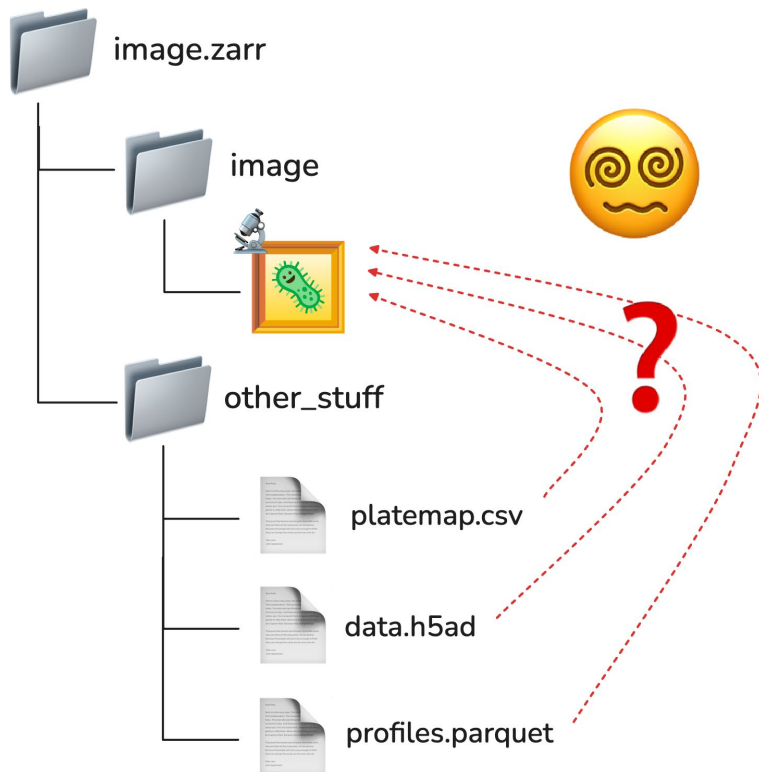
# Relating images to other data is difficult!

	A	B	C
1	ImageNumber	ObjectNumber	Metadata_FileLocation
2		1	1file:/app/inputdata/nuclei1_out_c00_dr90_image.tif
3		1	2file:/app/inputdata/nuclei1_out_c00_dr90_image.tif
4		1	3file:/app/inputdata/nuclei1_out_c00_dr90_image.tif



Using filenames means  
**lost data connection.**

# Bioimaging lacks a complete relational data model



Imaging implementations accumulate external data references out of necessity.

These conventions embed assumptions that complicate reproducibility.

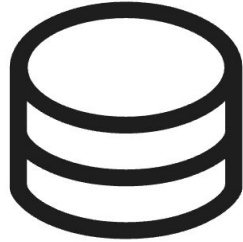
# It doesn't have to be like this!



Images

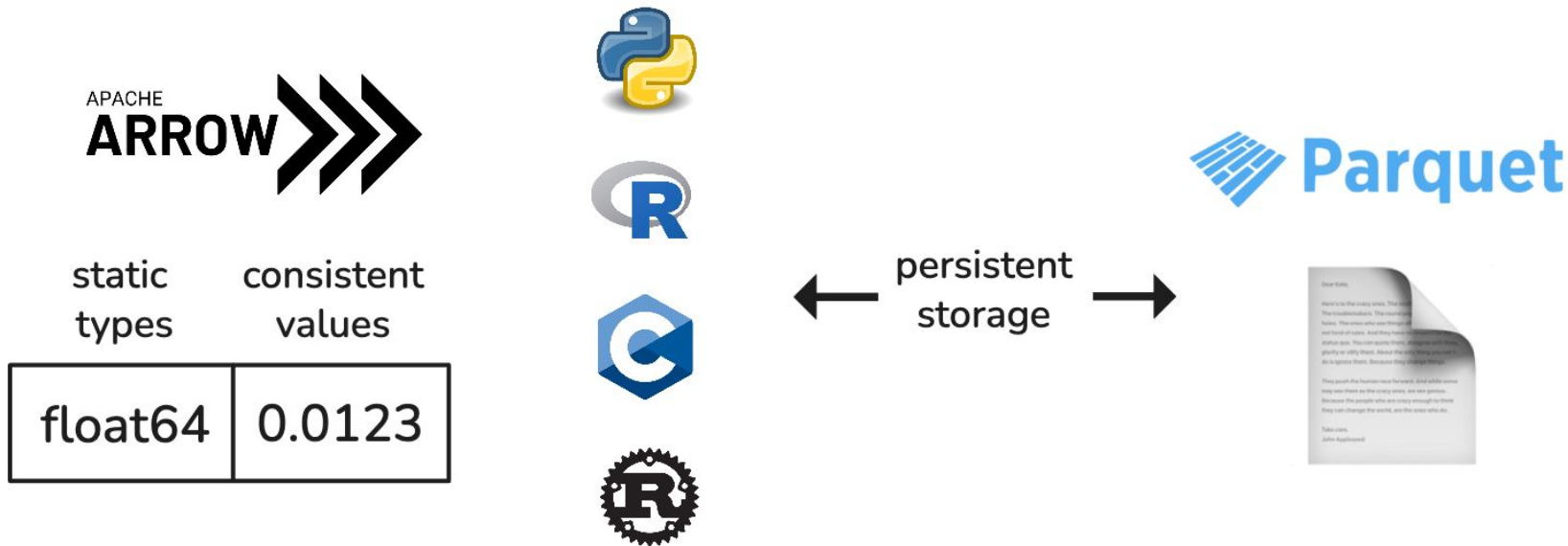


Other data

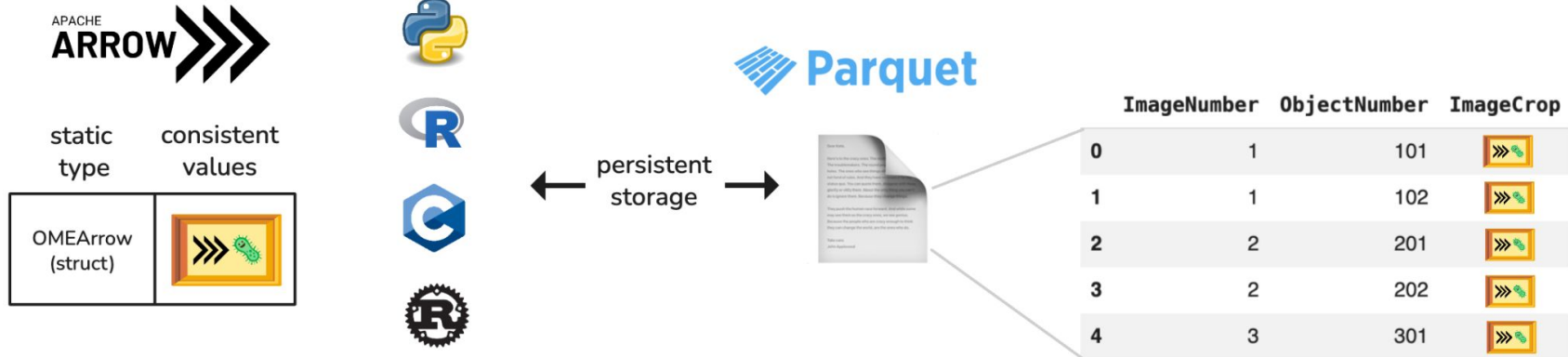


Modern data technologies enable a true relational data model without external file references.

# Apache Arrow is built for scale and interoperability



# OME-Arrow is a relational bioimage data type



<https://github.com/WayScience/ome-arrow>

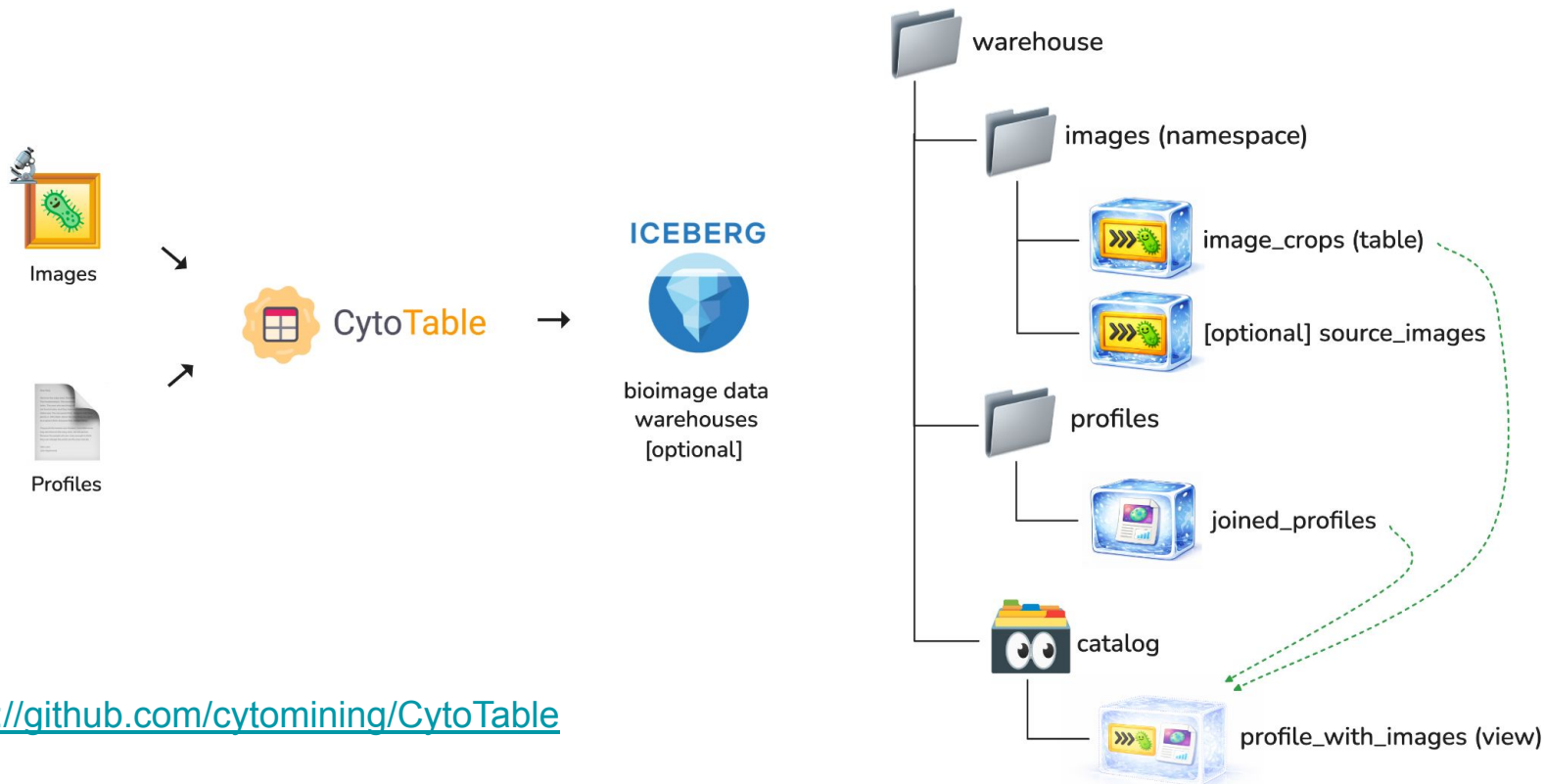


# OME-Arrow helps uncover patterns in bioimage cycles



Metadata_Cells_Number_Object_Number	cqc.large_nuclei.is_outlier	Image_FileName_GFP	Image_FileName_RFP	Image_FileName_DAPI
2	True			
10	True			
9	False			

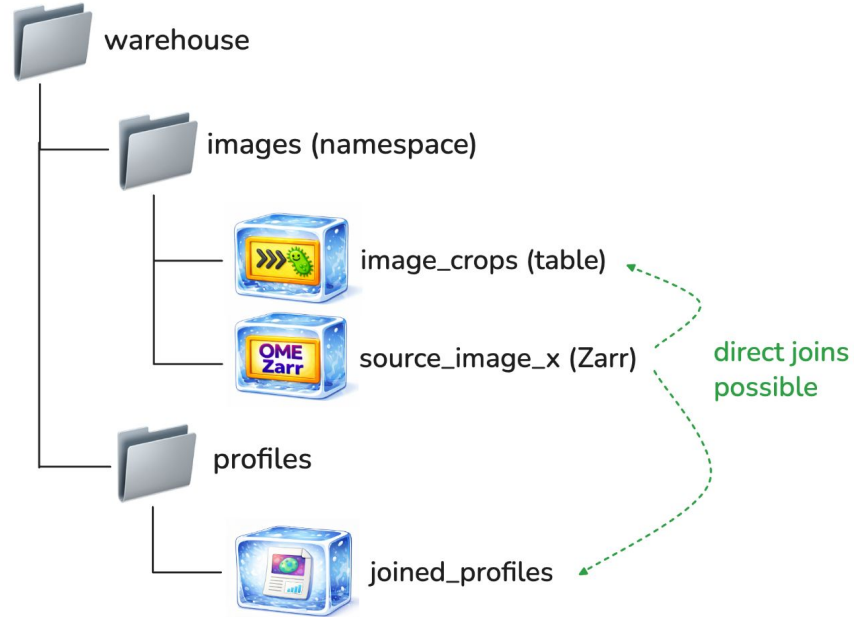
# CytoTable for OME-Arrow warehouse creation



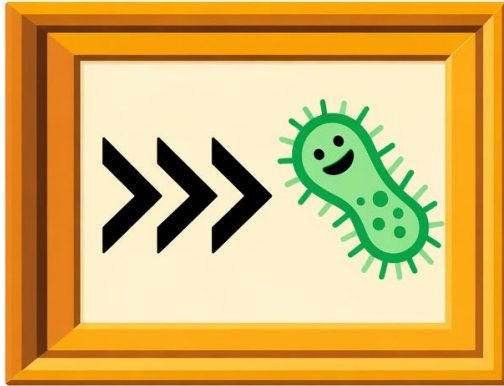
<https://github.com/cytomining/CytoTable>

# Iceberg-bioimage for viaducts to existing formats

ICEBERG  
BIOIMAGE



<https://github.com/WayScience/iceberg-bioimage>



# Thanks!

Please visit us at [poster #3](#) !