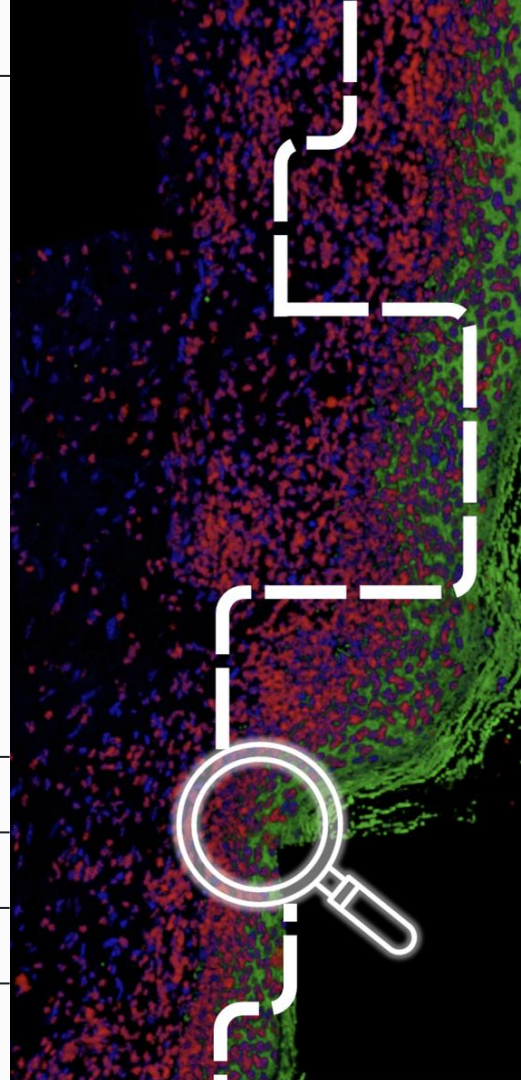


Guided Discovery in Multiplexed Volumetric Data

Chahat Kalsi

Visualization, Imaging and Data Analysis (VIDA) Center

New York University

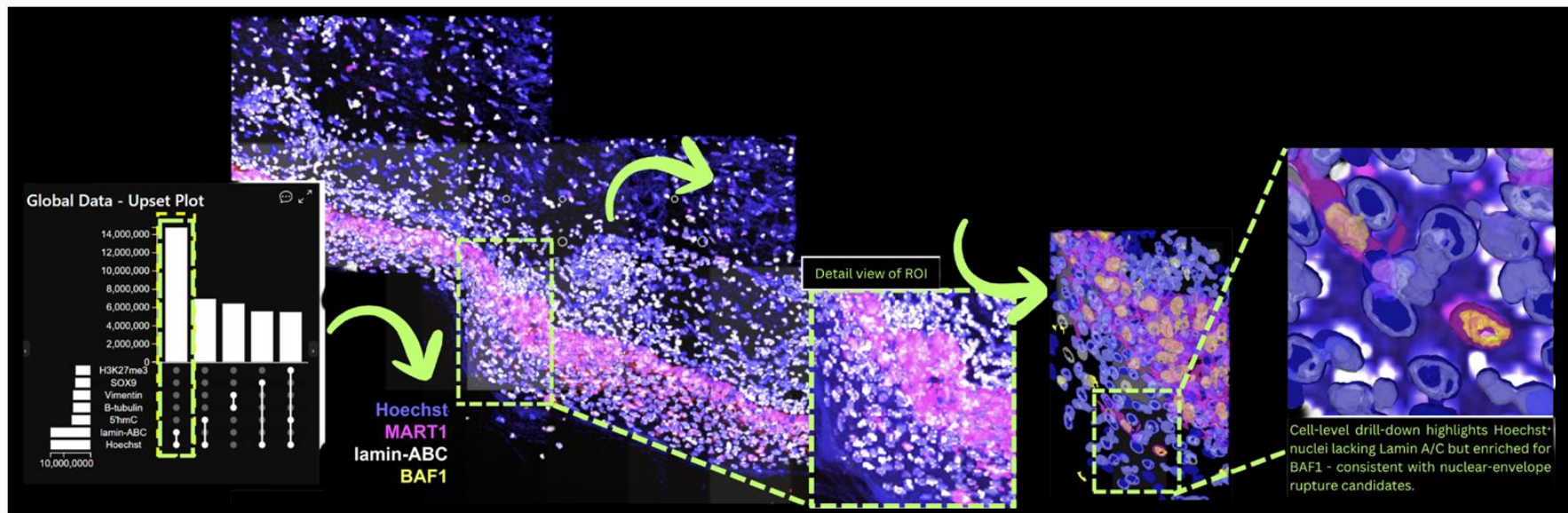


BioSET - Biomarker-based Spatial co-Expression analysis in Tumor environments



Award of Excellence

Chahat Kalsi, Yuancheng Shen, Sophia Gaupp, Luca Reichmann, Meri Rogava, Michael Krone, Saeed Boorboor, Robert Krüger





BioSET
VISUALIZER

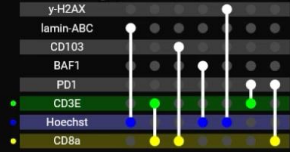
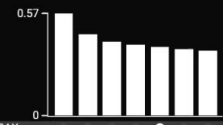
BioSET web tool

Analysis ✕

MARKER COMBINATIONS

Global Local All Selected ⌵ ⌵ ⌵

1 - 7

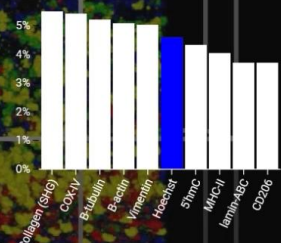


GitHub Link

MARKER COVERAGE

Global Local All Selected ⌵ ⌵ ⌵

1 - 10



DILATION CURVES

Global Local Single Multiple ⌵

DATA SOURCES

Zarr URL
<https://lsp-public-data.s3.amazonaws.com>
Metadata URL
<https://lsp-public-data.s3.amazonaws.com>

CLEAR DATA

Analysis results (bioset)
melanoma_1_ttu_bioset

melanoma_in_situ_bioset

ANALYSIS PARAMETERS

SETTINGS

TOOLS

CHANNELS 70

Hoechst



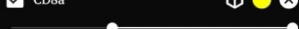
SOX10



CD3E



CD8a

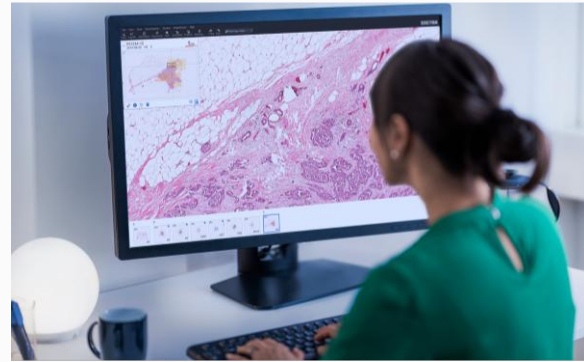
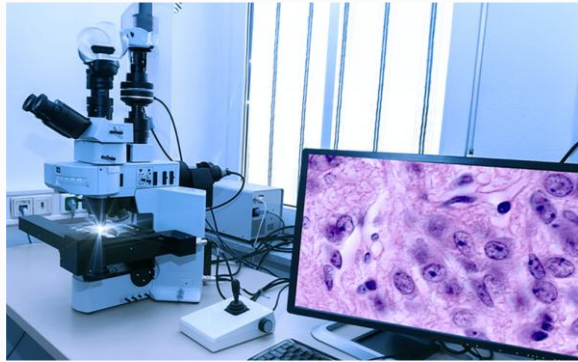
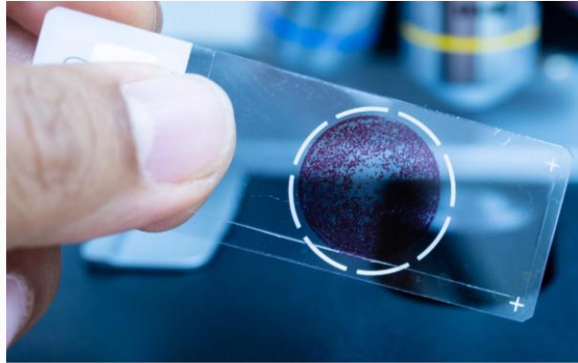


ADD CHANNEL

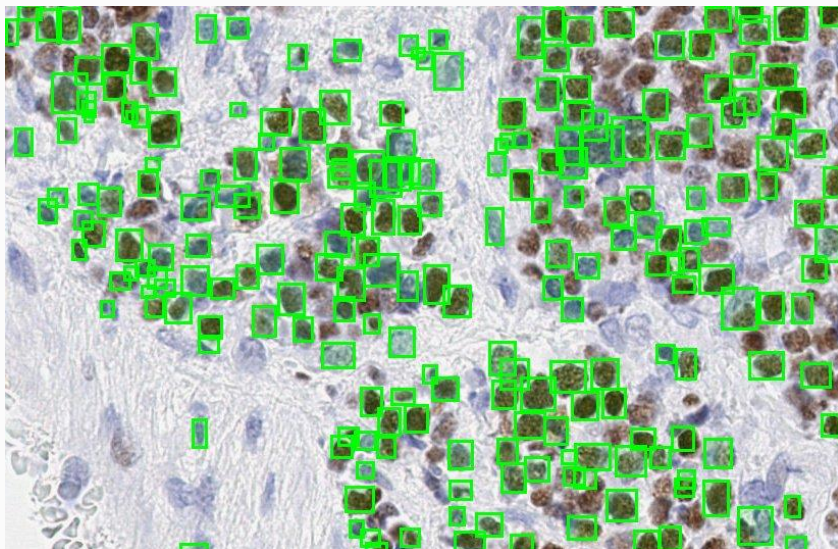
112 μ m



Digital pathology



Semi-automated workflows

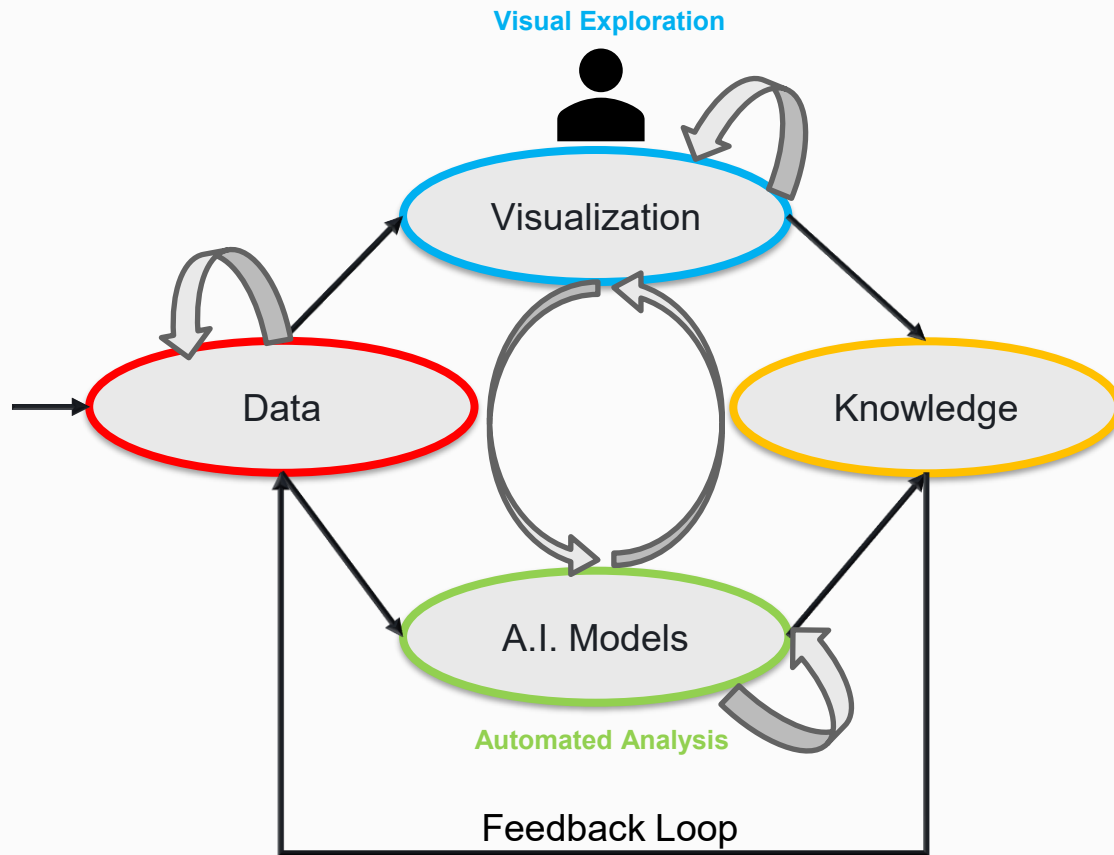


Visual analytics

- Human-AI Collaboration
- Expert in the loop analysis



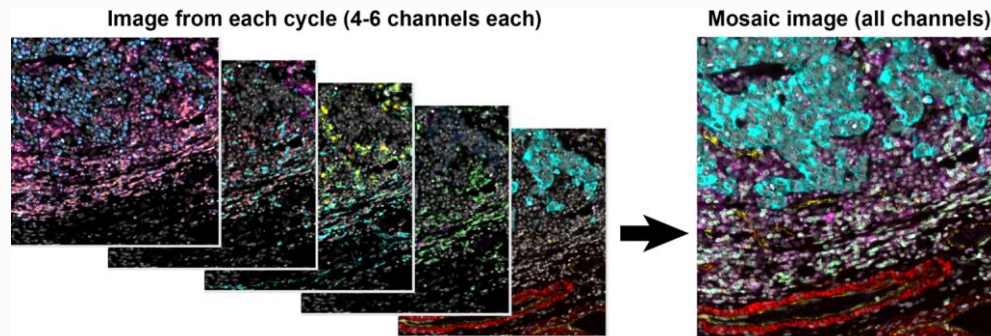
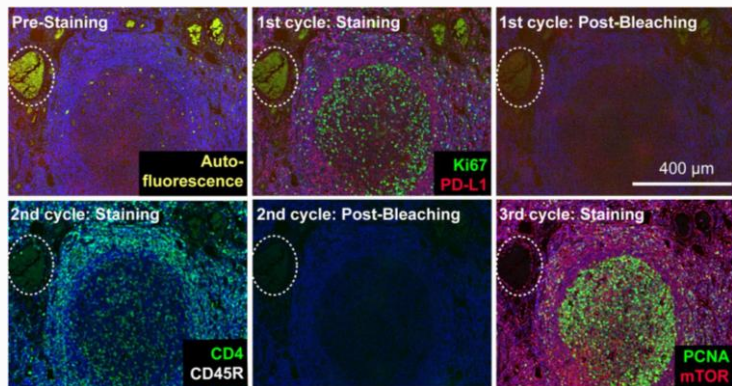
D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. Visual analytics: Definition, process, and challenges. In A. Kerren, J. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization*, volume 4950 of *Lecture Notes in Computer Science*, pages 154–175. Springer Berlin Heidelberg, 2008.



Multiplexed Cyclic Immunofluorescence (CyCIF)

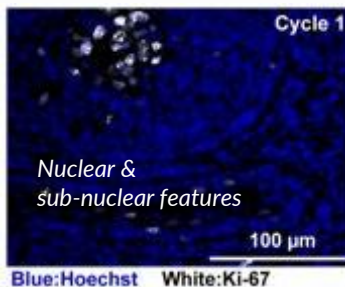


Lin, J.R., et al., 2018. Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife*



Note: The mosaic image shown here contains data from only a small portion of a slide - a whole slide image would be made up of many image 'tiles' of small regions, which must be assembled into the full whole slide image.

- **Multiple image channels** (up to 100)
- **High-resolution** (100k x 100k px)
- Making proteins visible to reveal cell functions.

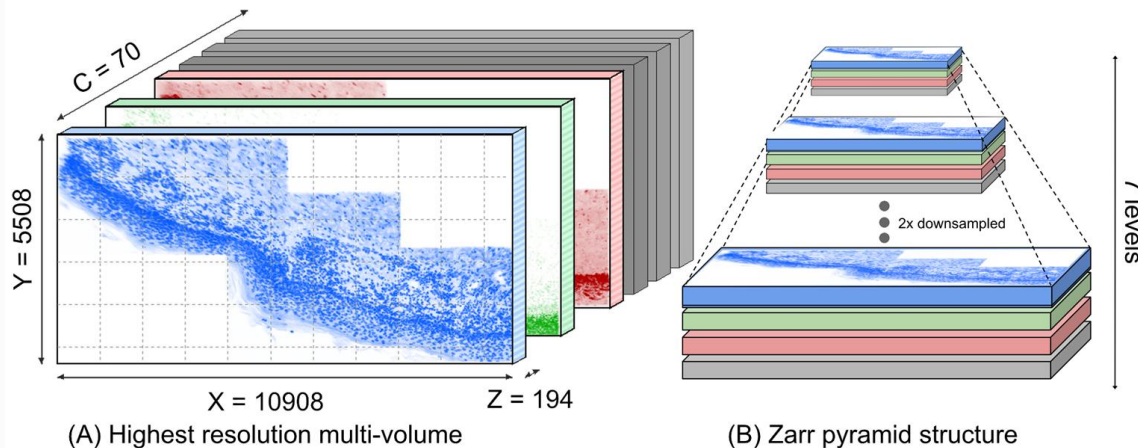


...Multiplexed CycIF... in 3D!

3D
30-50 μm Thick
(upto 2 layers of intact cells)

Multi-Channel
8-18 imaging cycles
producing 50+ channels

High Resolution
 $\sim 500 \text{ GB mm}^{-2}$
(1-2 TB datasets)



Dataset: Highly multiplexed 3D profiling of cell states and immune niches in human tumors, Yapp et al., *Nature Methods*, 2025

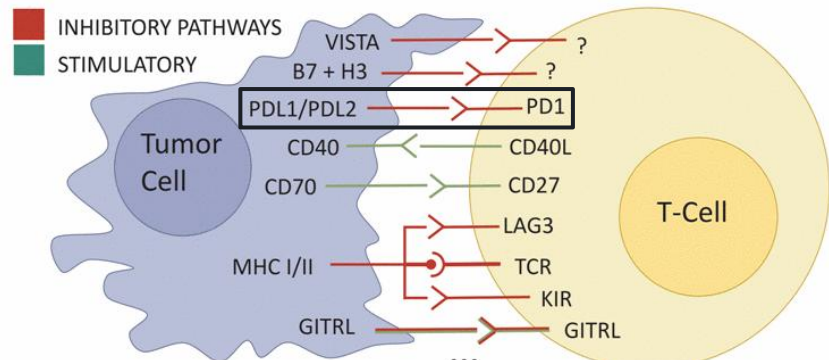
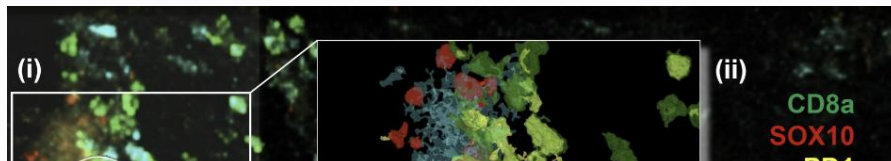
What are we discovering?

Cell Phenotype

Cell identity can be inferred from the biomarkers it exhibits

Cell-Cell Interactions

Cells interact with each other through biological pathways which can exhibit specific biomarkers



Cell2Cell: Exploratory Cell-Cell Interaction Analysis in Multi-Volumetric Tissue Data



BioSET: Biomarker based Spatial co-Expression analysis in Tumor environments, Kalsi et al., Bio+MedVis Challenge 2025

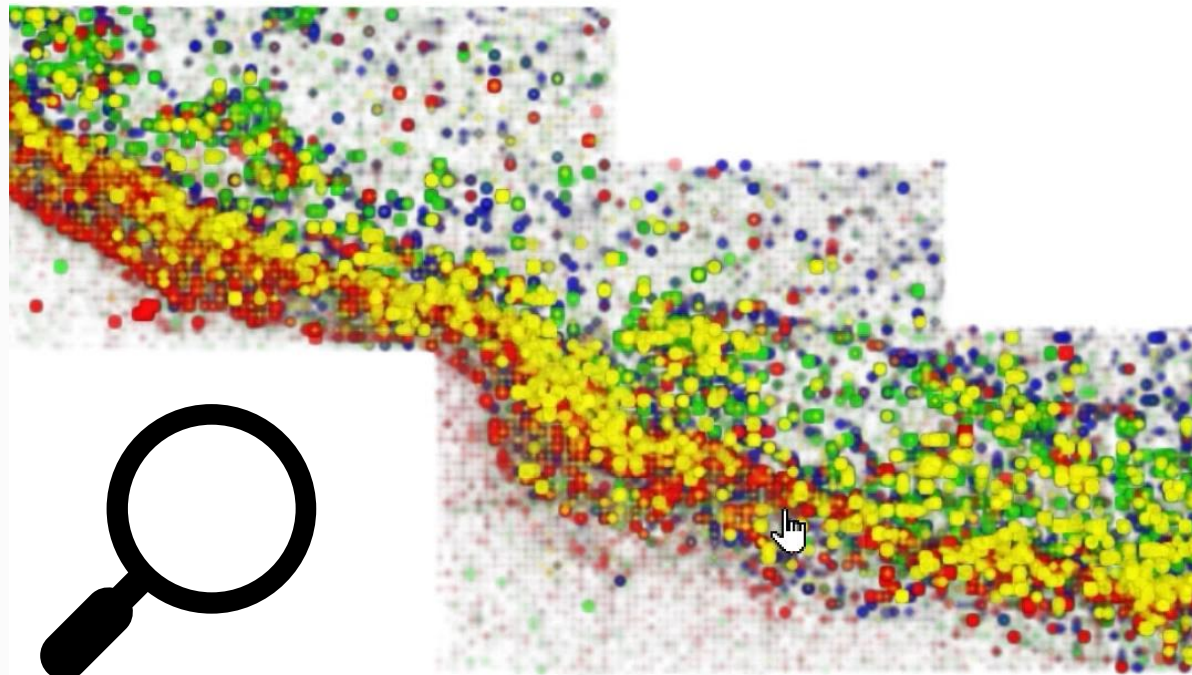
Manual exploration driving discovery

Tissues are

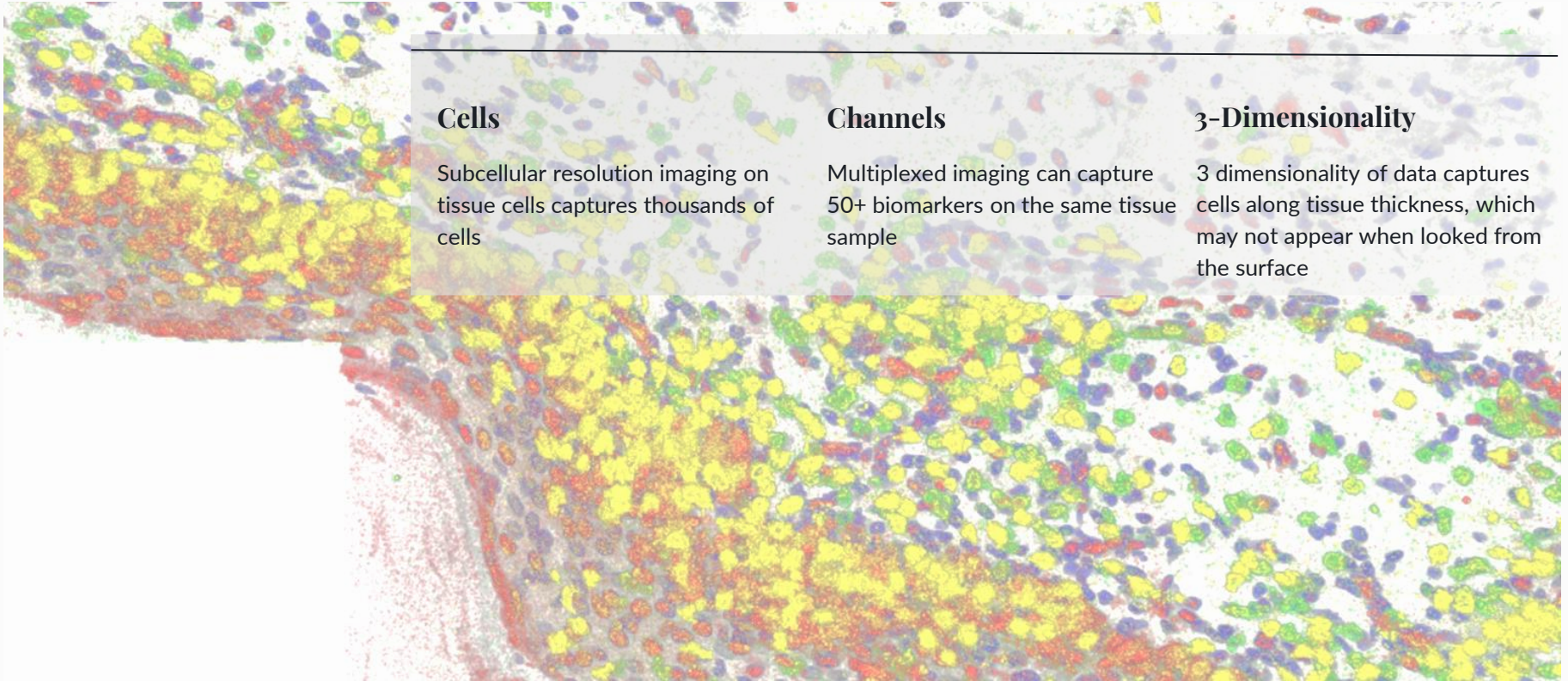
- **Large**
- **High-res**, and,
- **High-plex**

Cells/cellular interactions are

- **Small**, and,
- **Channel-dependent**



Challenges posed by the scale of the data



Cells

Subcellular resolution imaging on tissue cells captures thousands of cells

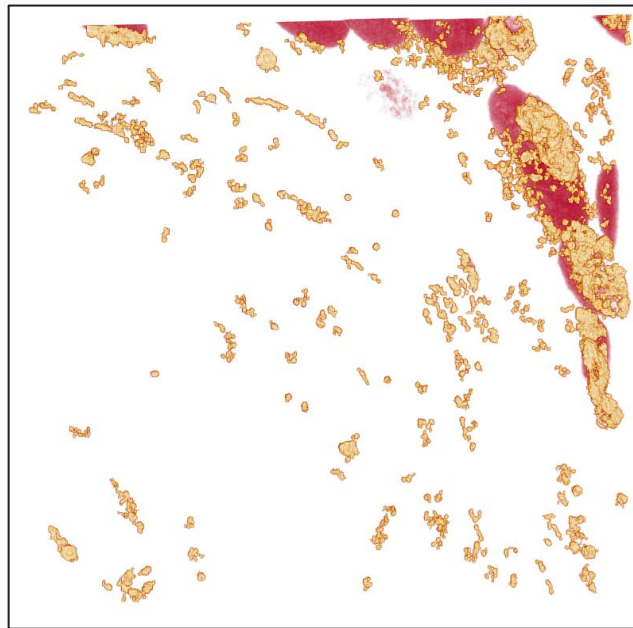
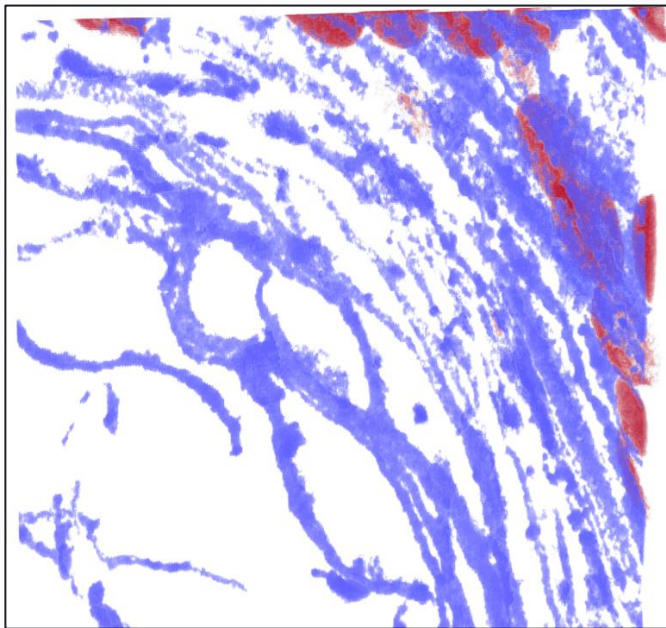
Channels

Multiplexed imaging can capture 50+ biomarkers on the same tissue sample




3-Dimensionality

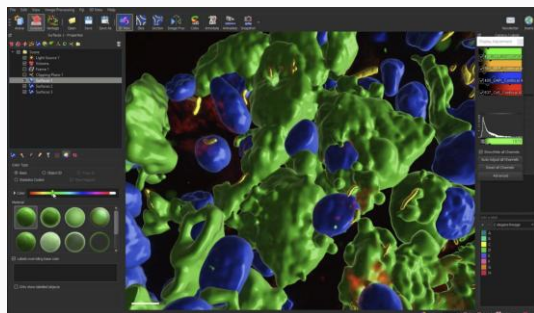
3 dimensionality of data captures cells along tissue thickness, which may not appear when looked from the surface

3-Dimensionality as a challenge to discovery

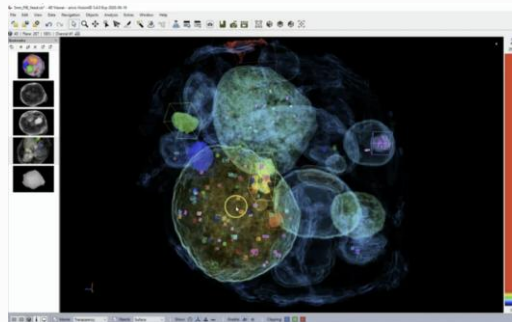


Multi-volume visualization tools

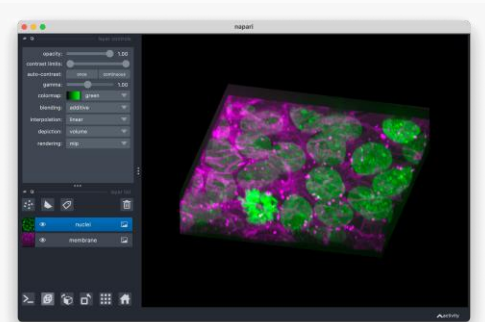
-  Commercial
-  Open sourced
-  Web-based



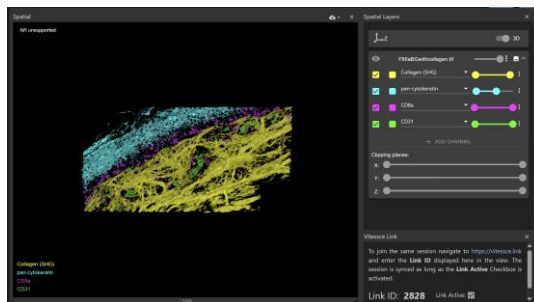
Imaris 



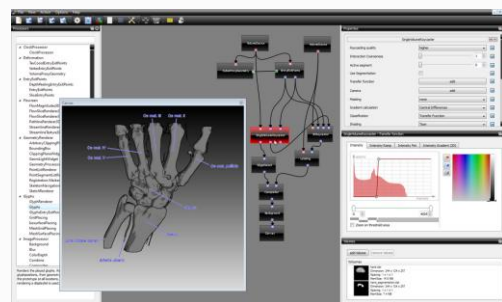
Arivis 



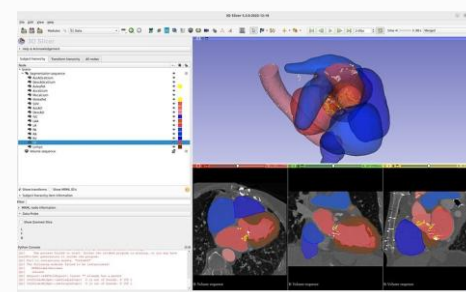
Napari 



Vitesse  



Voreen 



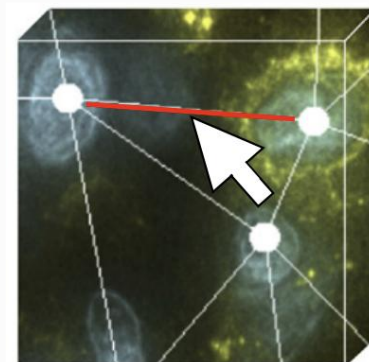
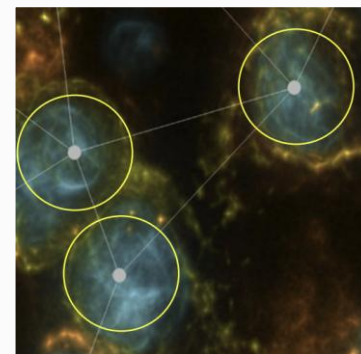
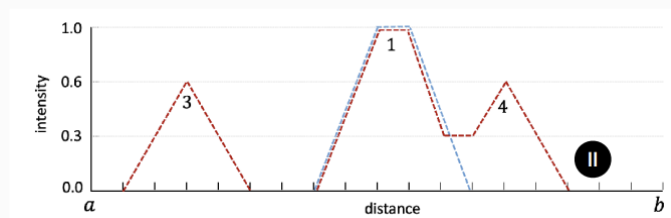
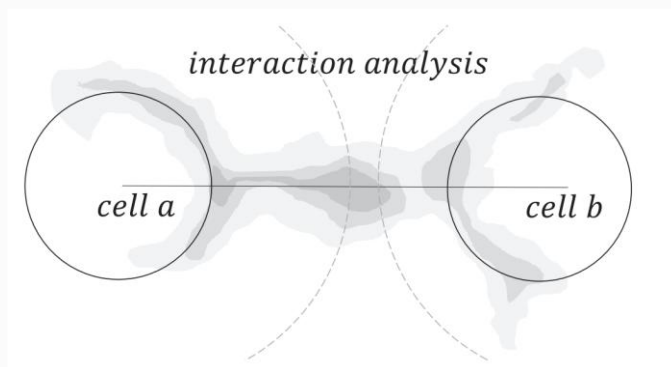
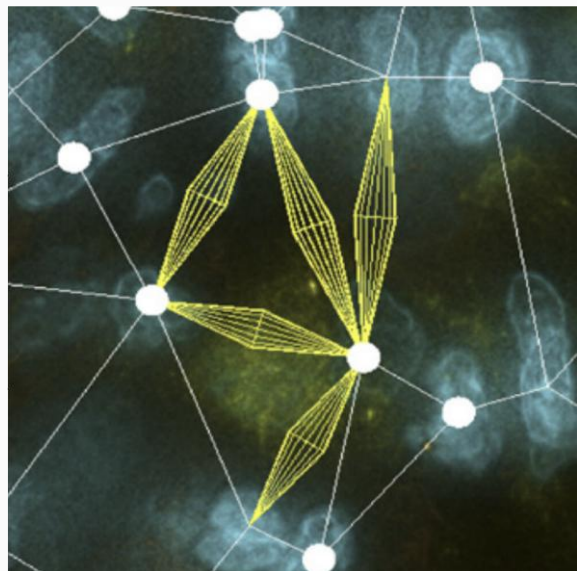
Paraview 

And more....

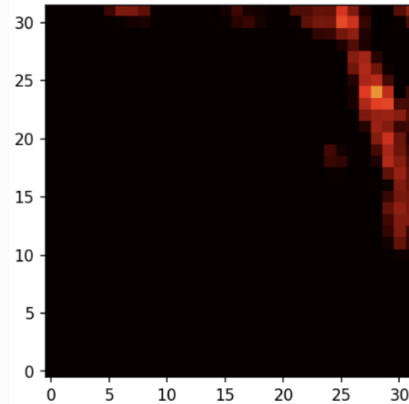
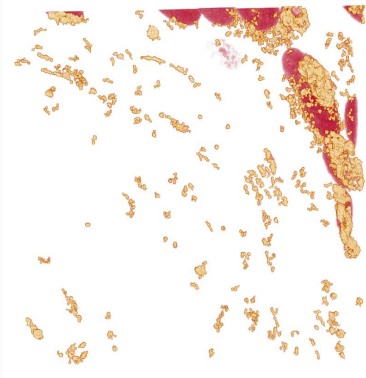
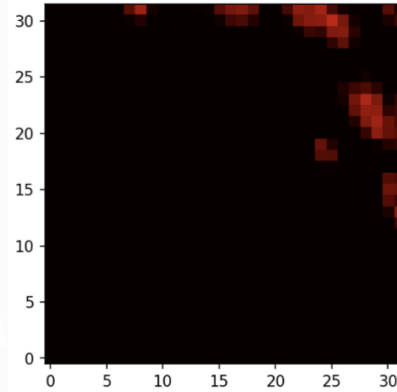
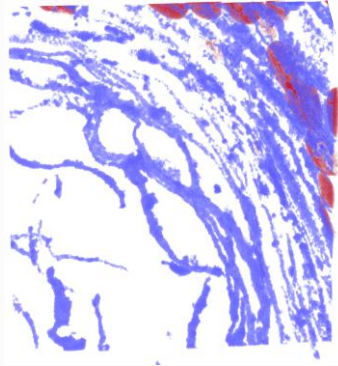
Expert-in-the-loop visual analysis



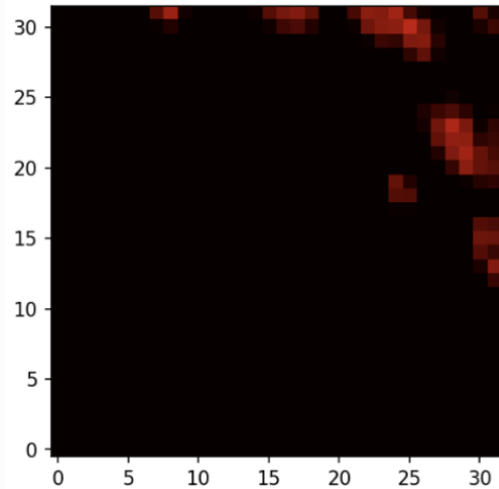
Cell2Cell: Explorative Cell Interaction Analysis in Multi-Volumetric Tissue Data, Morth et al., TVCG 2024



BioSET: Spatial binning of co-expression



2D spatial binning of 3D tissue data helps makes sense of the 3-Dimensionality of it



Challenges posed by the scale of the data

Cells

Subcellular resolution imaging on tissue cells captures thousands of cells

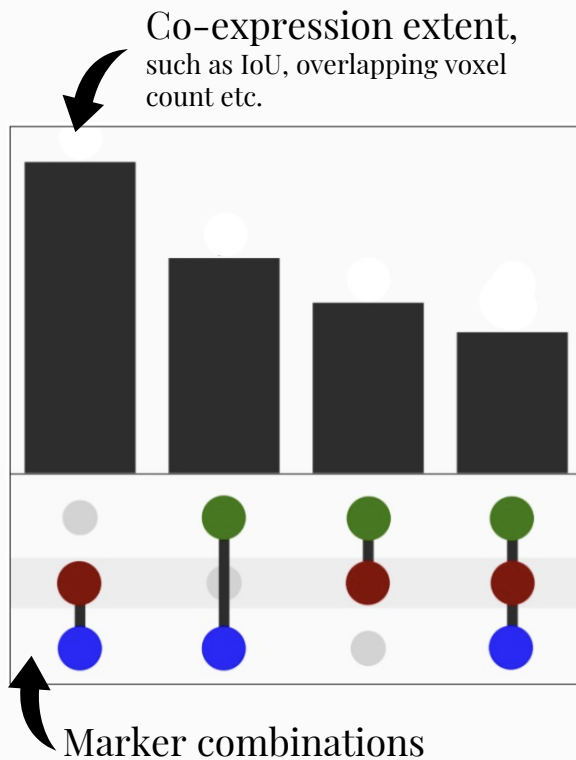
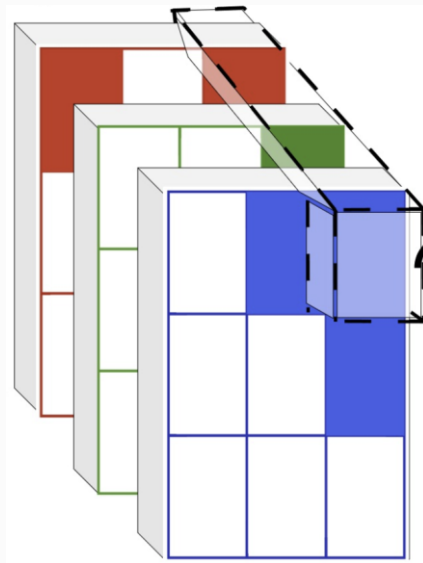
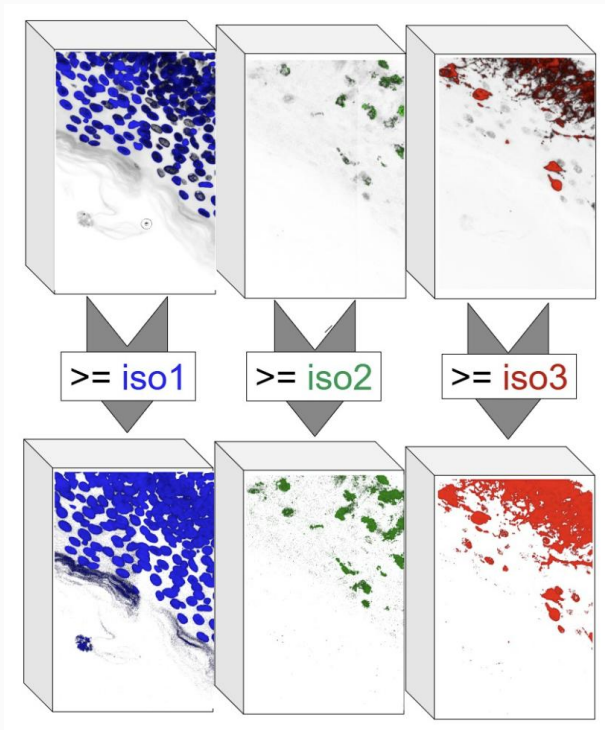
Channels

Multiplexed imaging can capture 50+ biomarkers on the same tissue sample

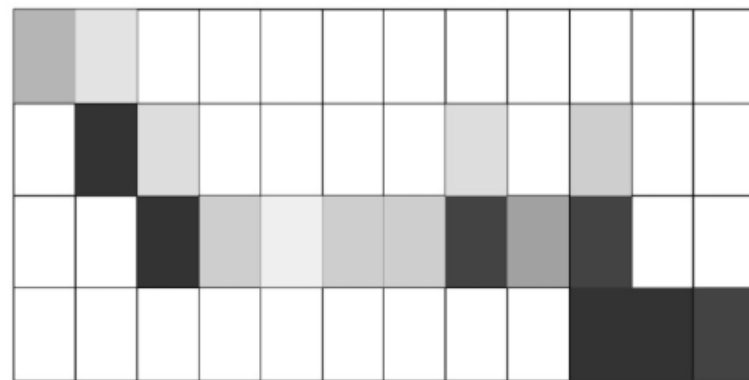
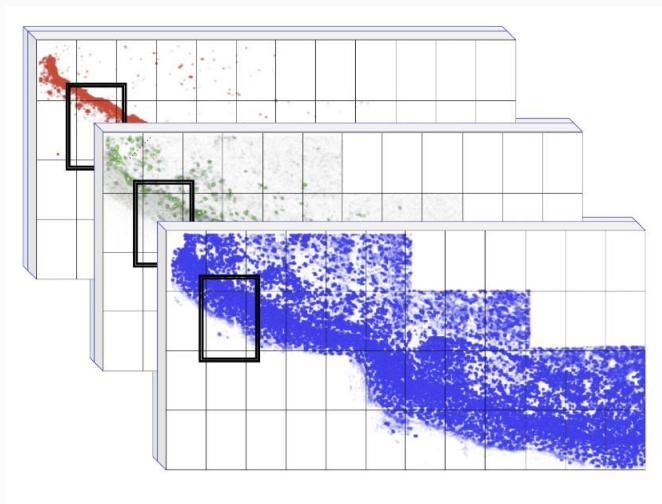
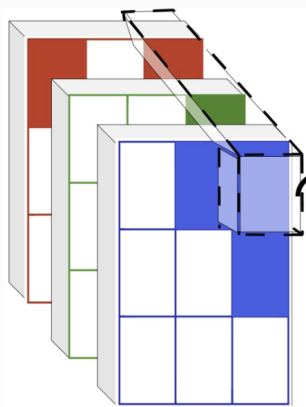
3-Dimensionality

3 dimensionality of data captures cells along tissue thickness, which may not appear when looked from the surface

Detecting co-expression



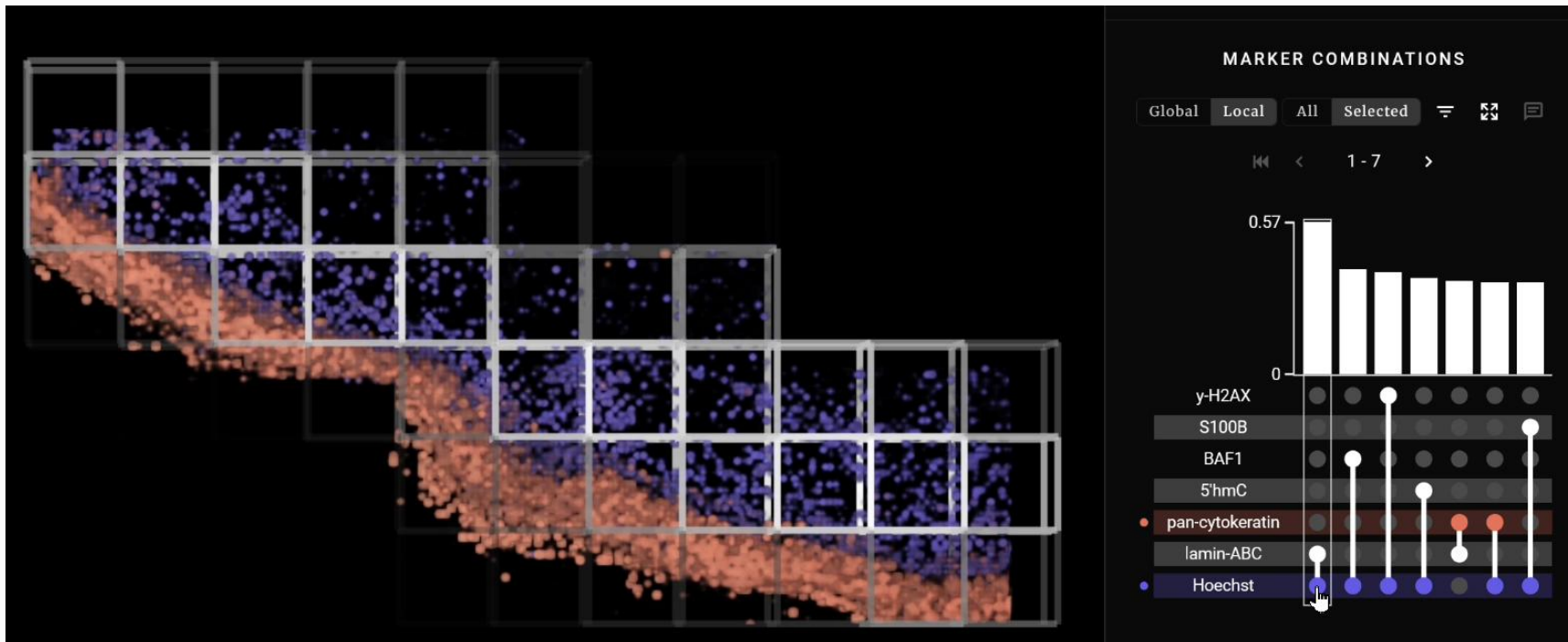
Spatial co-expression extent



Linking quantification and distribution

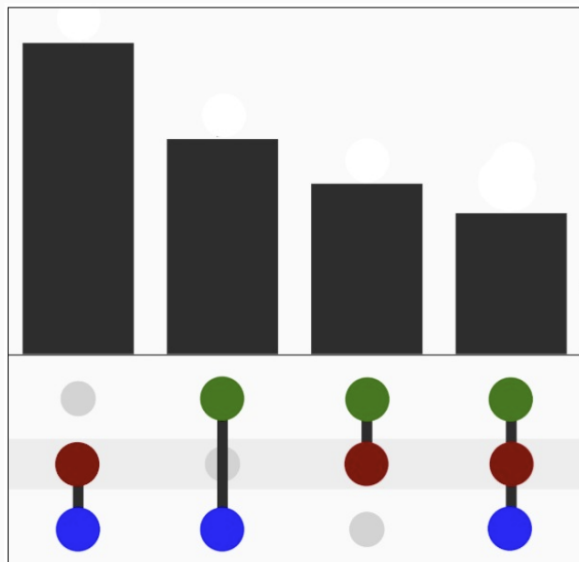
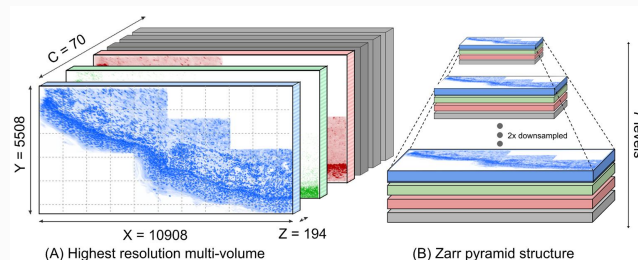


UpSet: Visualization of Intersecting Sets, Lex et al., TVCG 2009



Which channel combinations to consider?

Number of channels, $C = 49$



All

$$2^C - 1$$

563 trillion combinations

Up to size
 k_{\max}

$$\sum_{k=1}^{\min(C, k_{\max})} \binom{C}{k}$$

$$\mathcal{O}(C^{k_{\max}})$$

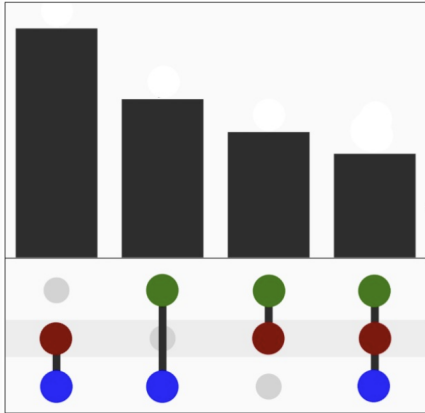
$k_{\max} = 4$
231k combinations

Higher-order
set pruning

Evaluate pairwise support.

Only grow sets from pairs with enough support.

Pre-determining channel combinations helps make sense of multiplexed data



Challenges posed by the scale of the data

Cells

Subcellular resolution imaging on tissue cells captures thousands of cells

Channels

Multiplexed imaging can capture 50+ biomarkers on the same tissue sample

3-Dimensionality

3 dimensionality of data captures cells along tissue thickness, which may not appear when looked from the surface

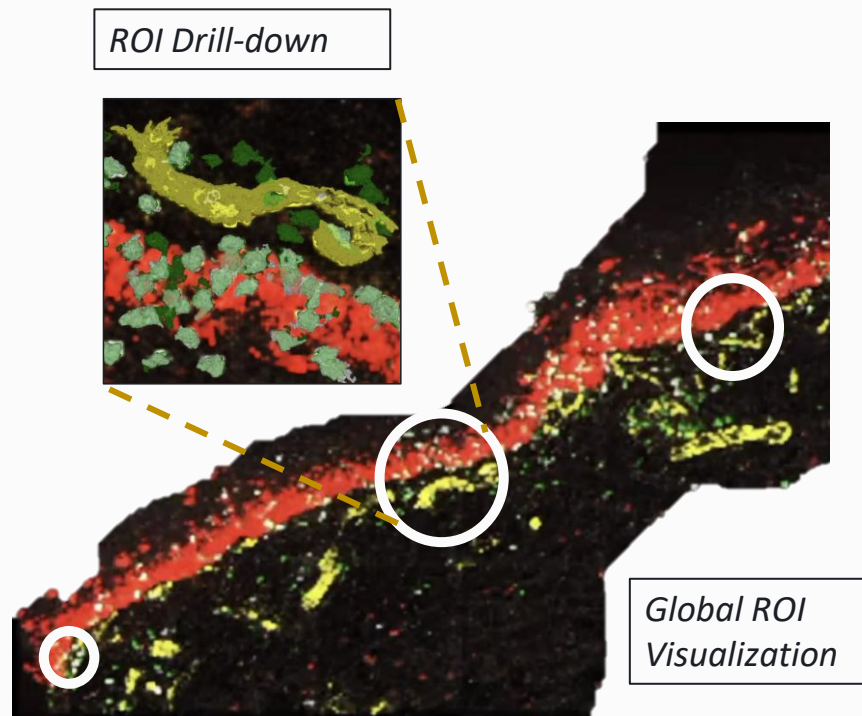
Guidance mechanism

Global placement of ROI

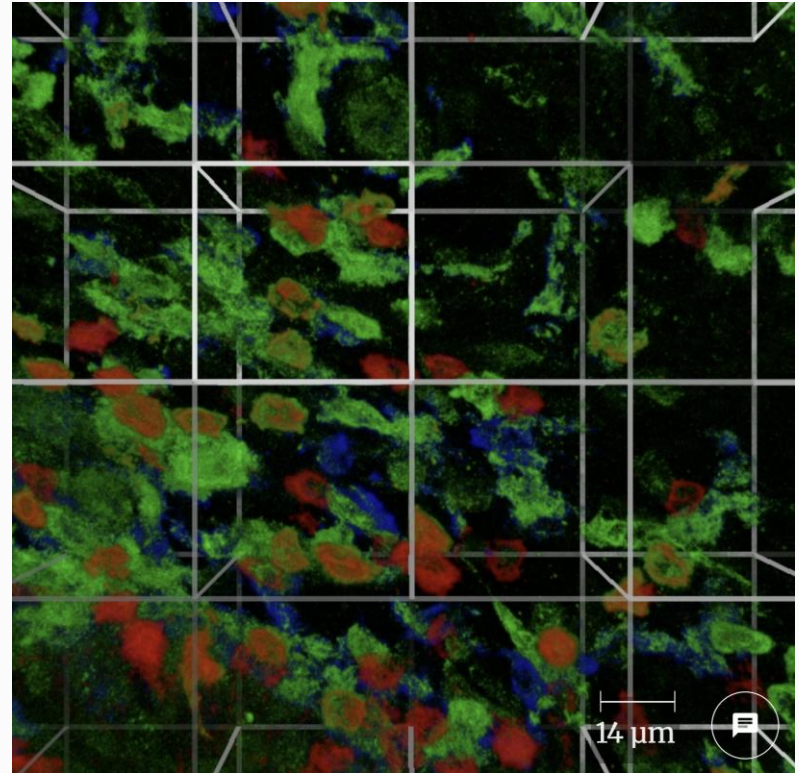
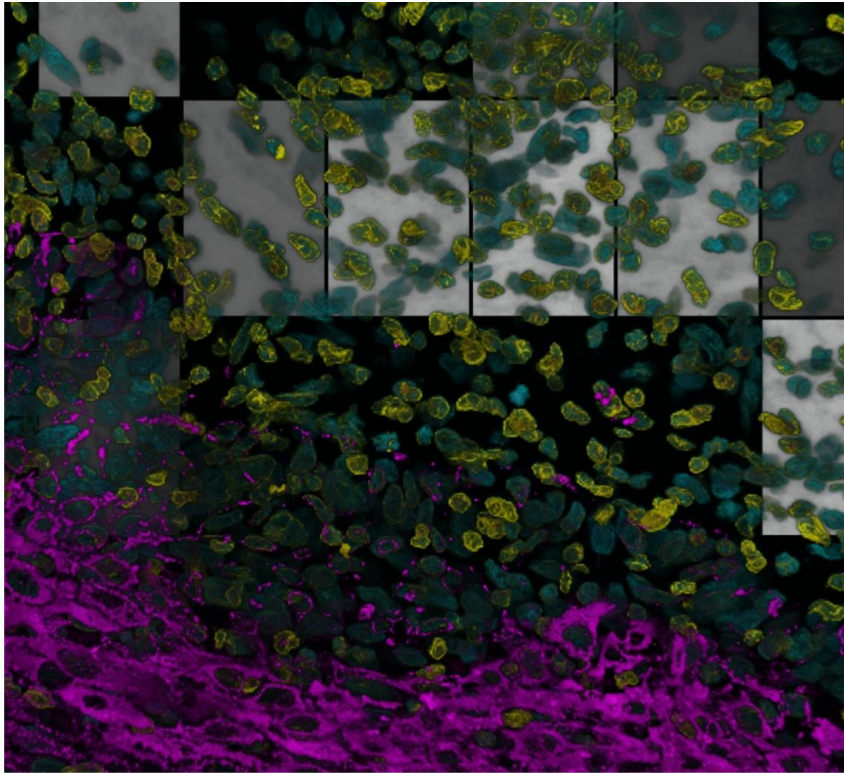
Need to see the ROI in
tissue context

Drill-down into an ROI

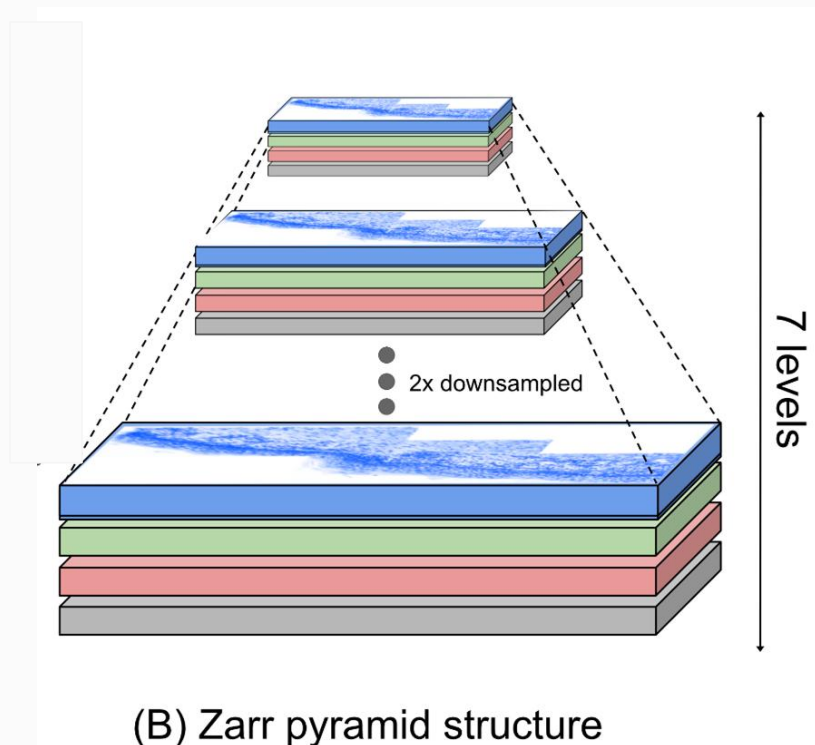
Need to be able to
investigate the ROI in
detail



Guidance embedded in the image space



Small regions of interest in large data

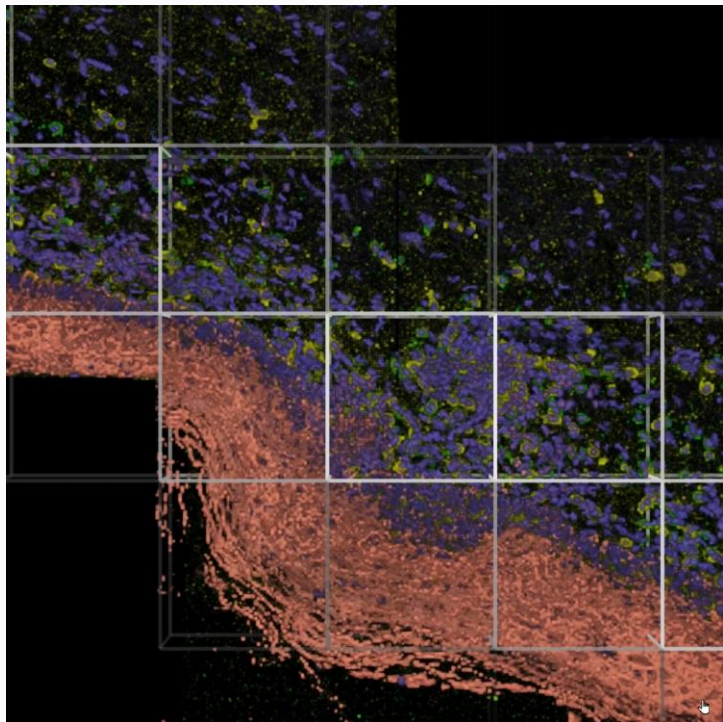


Coarse-to-fine scale navigation

Multi-Resolution Volume
Rendering

Hierarchically Resolving
Heatmap

Small regions of interest in large data

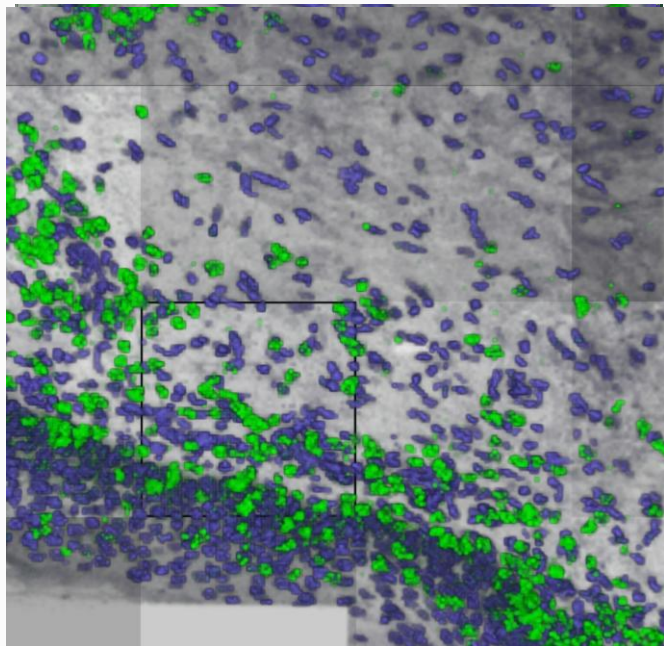


Coarse-to-fine scale navigation

Multi-Resolution Volume
Rendering

Hierarchically Resolving
Heatmap

Small regions of interest in large data

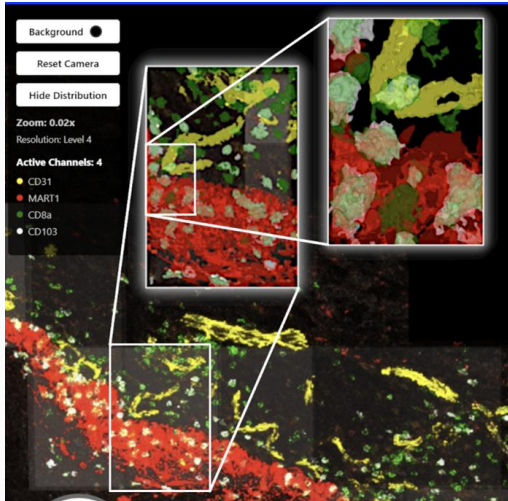


Coarse-to-fine scale navigation

Multi-Resolution Volume
Rendering

Hierarchically Resolving
Heatmap

Hierarchical navigation to individual cells helps make sense of large tissues



Challenges posed by the scale of the data

Cells

Subcellular resolution imaging on tissue cells captures thousands of cells

Channels

Multiplexed imaging can capture 50+ biomarkers on the same tissue sample

3-Dimensionality

3 dimensionality of data captures cells along tissue thickness, which may not appear when looked from the surface

Supporting discovery

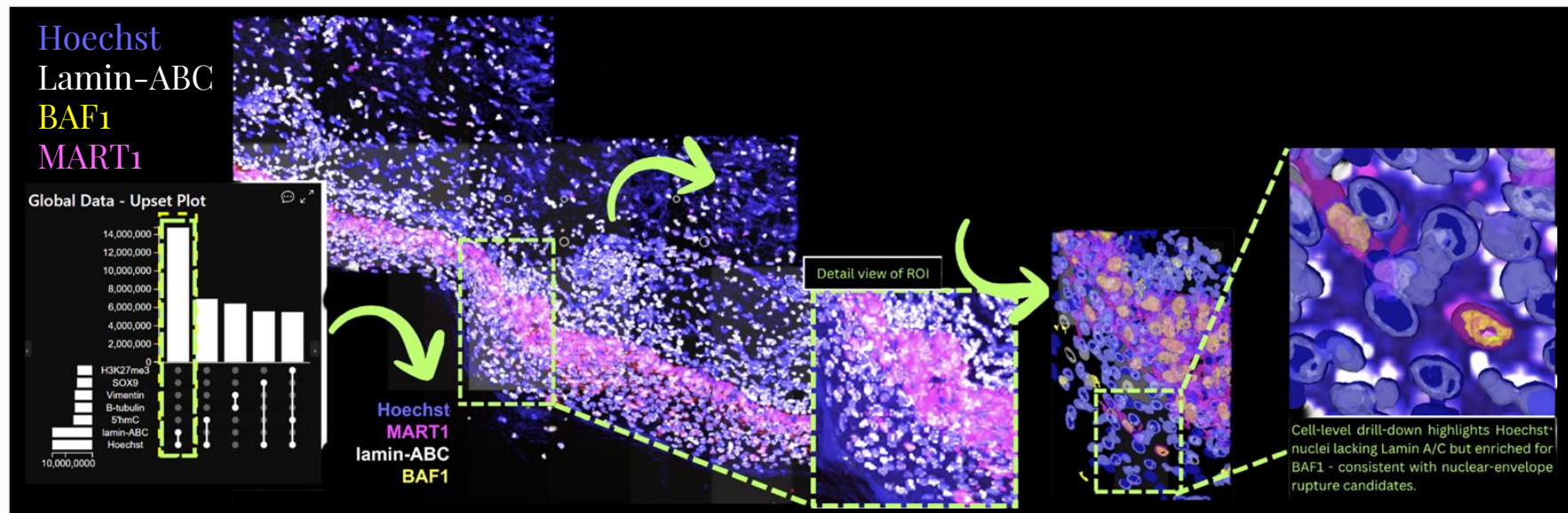
Case study

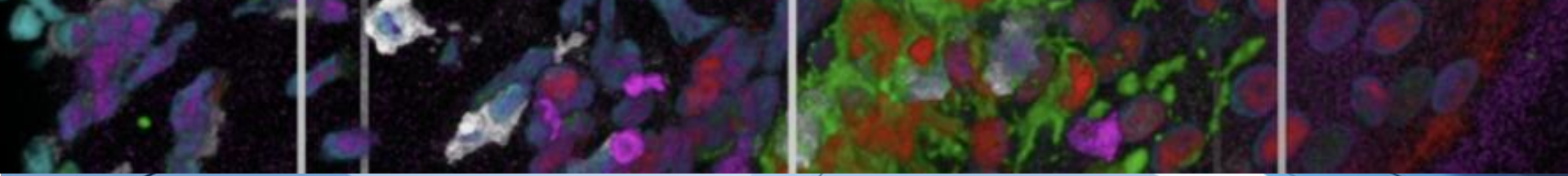
Finding a ruptured nucleus



Meri Rogava

Assistant Professor
Zucker School of Medicine at
Hofstra/Northwell





Thank You!



Chahat Kalsi

Visualization, Imaging and Data Analysis (VIDA) Center
New York University