

# WEHI's Spatial Omics Data Analytics (SODA) Hub

An environment for spatial analysis from data generation to analysis and beyond

Marek Cmero, Michael Milton, Pradeep Rajasekhar, Lachlan Whitehead and Julie Iskander

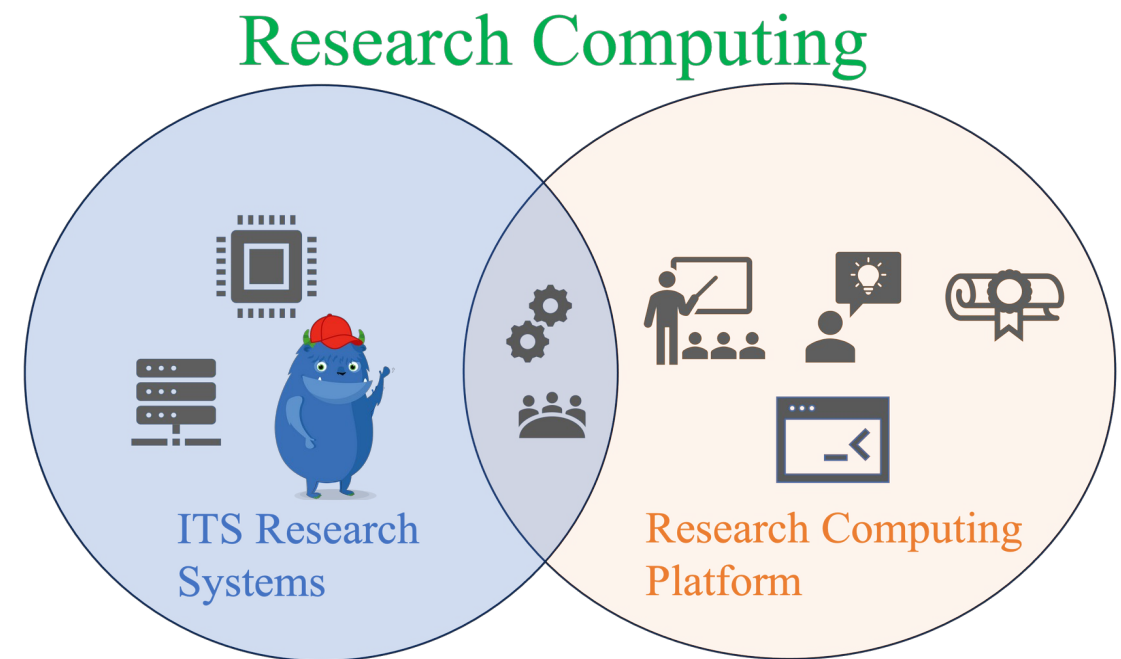
SODA Technical Team

Walter and Eliza Hall Institute of Medical Research

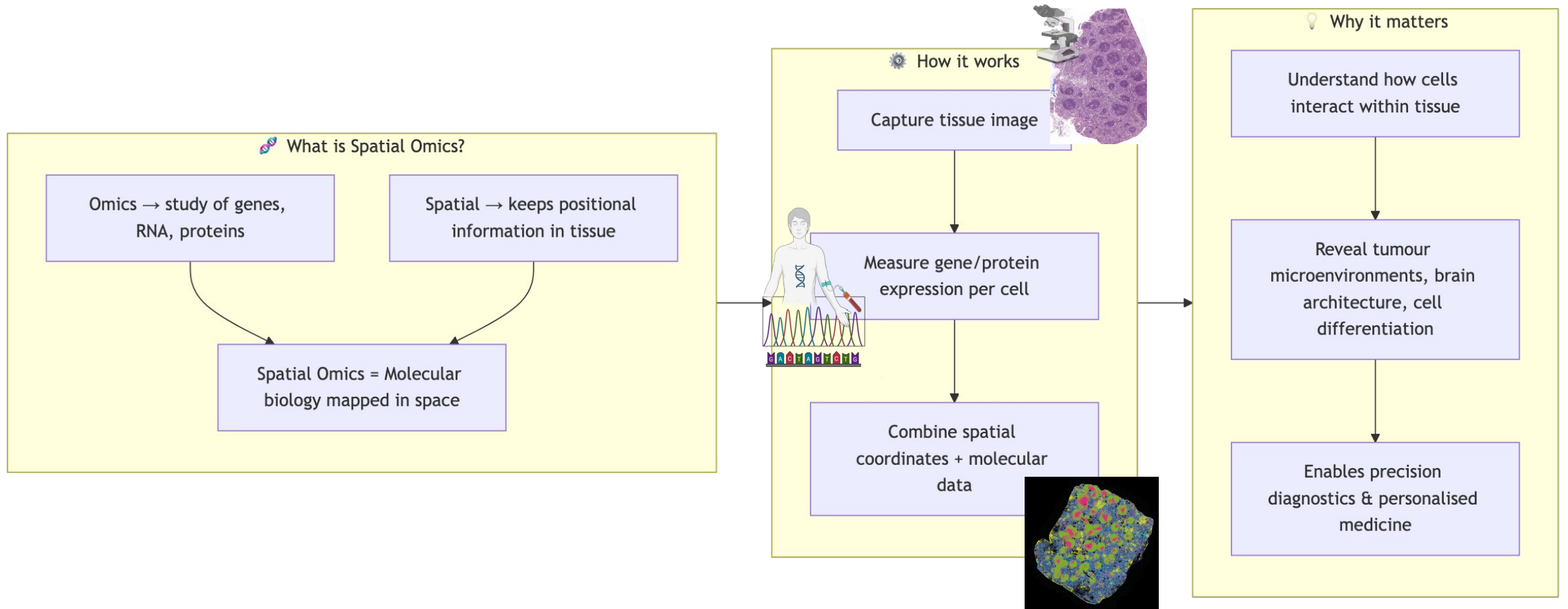
eResearch Australiasia, 21<sup>st</sup> October 2025

## Introducing Research Computing Platform

- Part of WEHI Research Computing
- A team of RSEs with varied backgrounds that include computer science, data science, bioinformatics and biomechanics.
- The main support categories:
  - HPC Support
  - Research Data Management
  - Specialised Support for Divisions/Initiatives



# Spatial Omics





**Spatial omics is where biology meets data intensity.**

**Spatial data is rich, complex and  
presents many challenges**

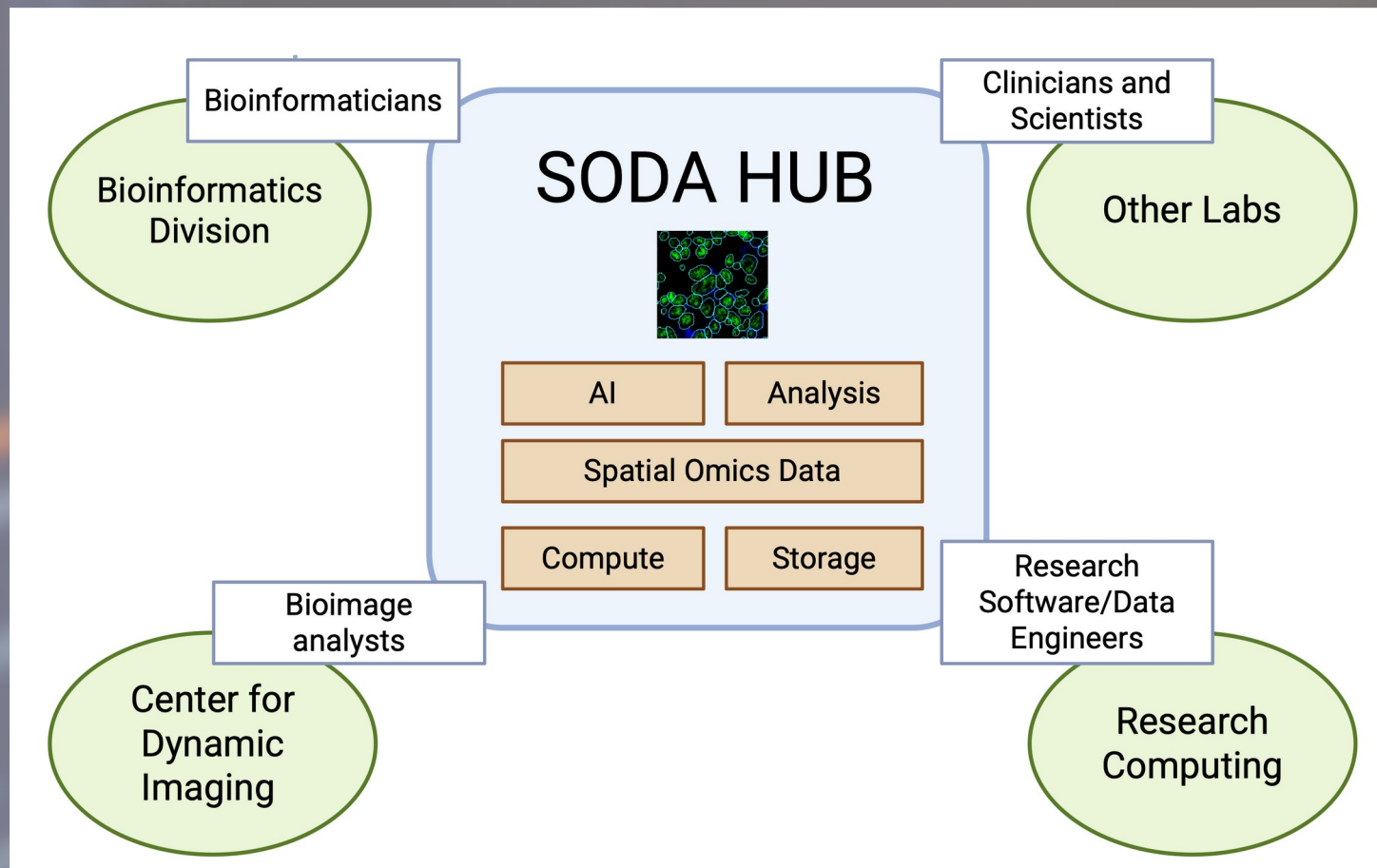
## Why Spatial Omics Matters

- Spatial omics pushes the limits of **data management and computation**:
  - **Terabytes per experiment** — large multiplexed imaging and sequencing outputs.
  - **Heterogeneous data types** — images, RNA counts, spatial coordinates, metadata.
  - **Cross-disciplinary workflows** — microscopy, bioinformatics, HPC, and cloud.
- Requires:
  - **High-performance storage and compute**
  - **Standardised pipelines**
  - **Metadata management**
  - **Interdisciplinary collaboration**

**Spatial omics is the kind of science that only works when biologists, data engineers, and computational scientists come together.**

# WEHI Approach: SODA Hub

## Spatial Omics Data Analytics Hub



## Aims of the SODA Hub

- Make spatial data Findable, Accessible, Interoperable and Reusable
- Develop workflows and new tools to streamline spatial analysis
- Foster collaboration, contribute open-source software, provide educational resources
- **Accelerate spatial omics research, leading to clinical application**

# Challenges

## Instruments



## Infrastructure



## Data management

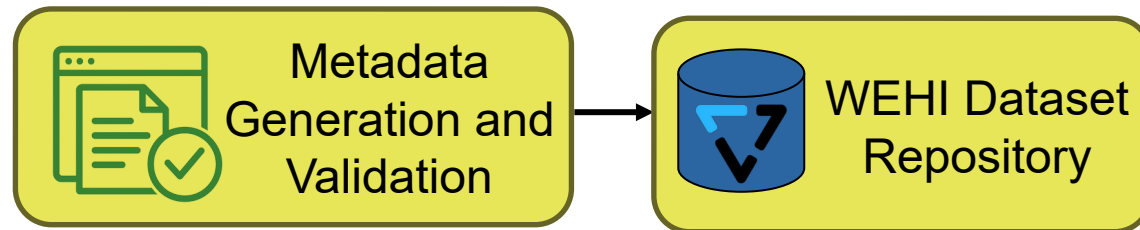


## SODA Architecture



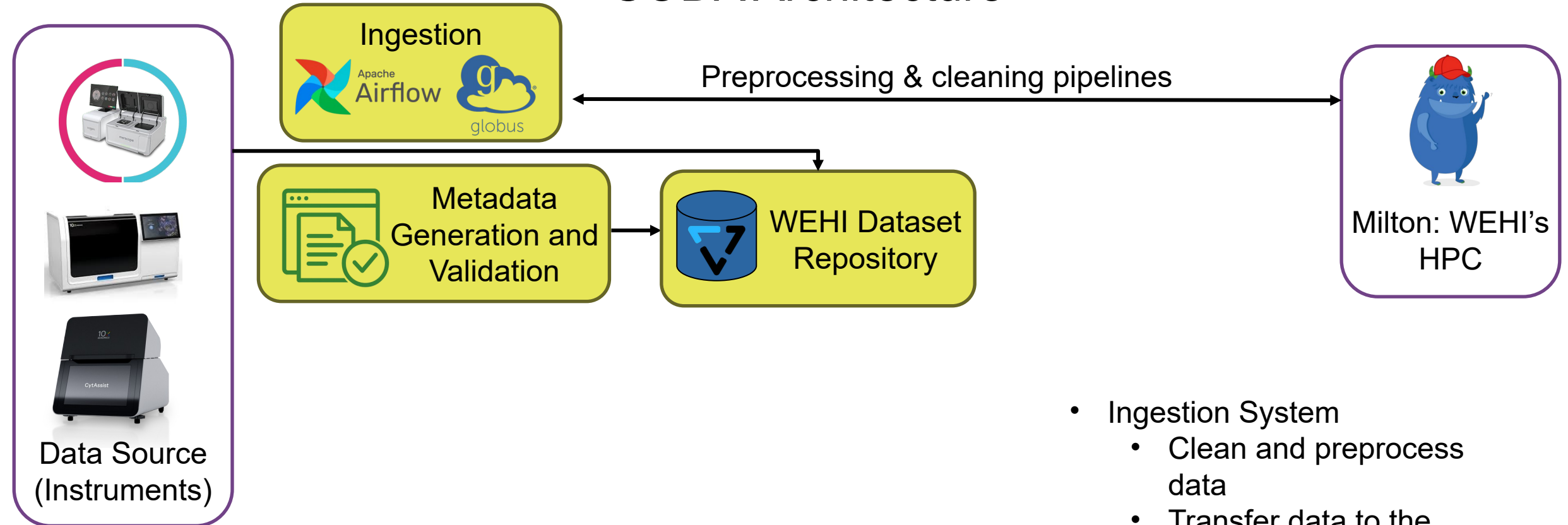
- A central repository for datasets that require long term retention and disaster recovery.
- Stores all types of data, such as, raw data, processed data, or summarised data, private data or shared data, in different formats and structures
- Data is immutable and version-controlled when required
- Access is limited to read-only and a streamlined transfer process

## SODA Architecture



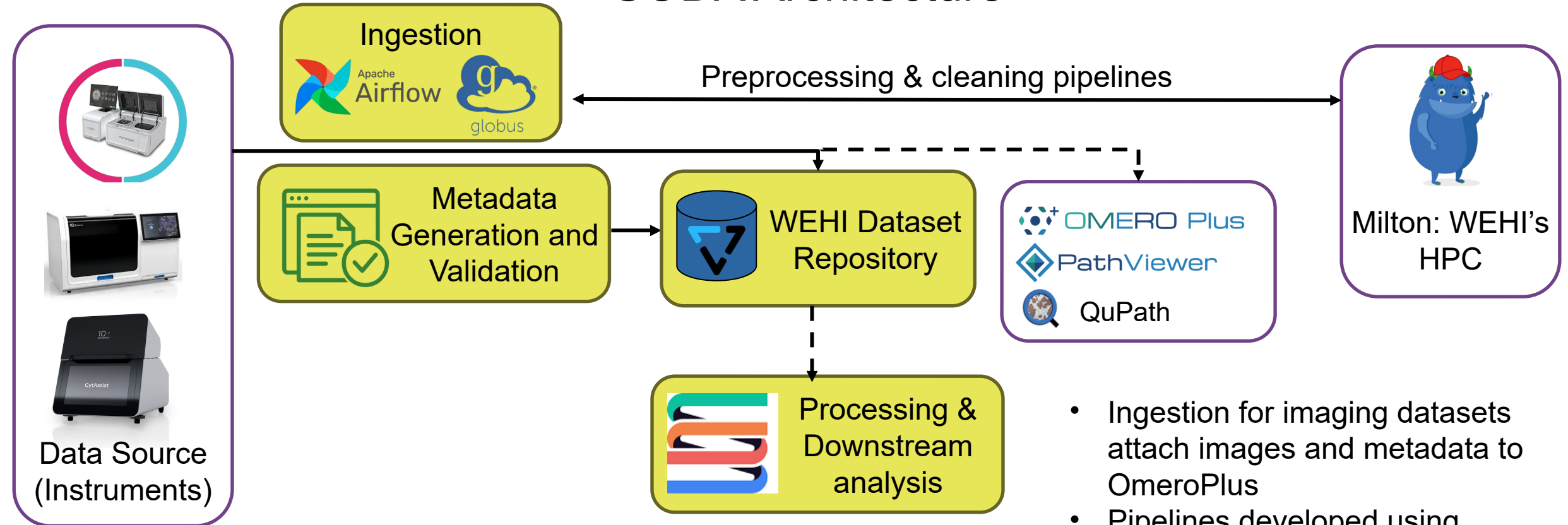
- Each Dataset must have
  - A record in the Governance and Compliance DB
  - Attached research metadata file (ro-crate.json)

# SODA Architecture



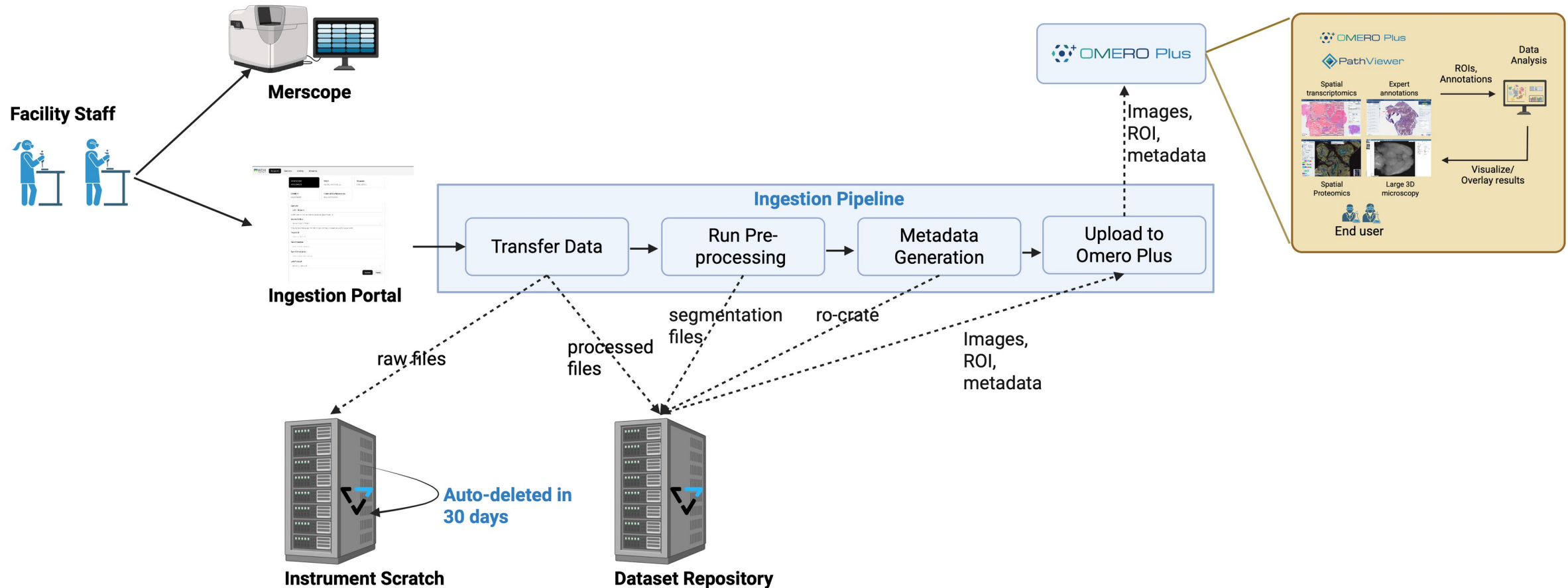
- Ingestion System
  - Clean and preprocess data
  - Transfer data to the repository
  - Add Metadata

# SODA Architecture

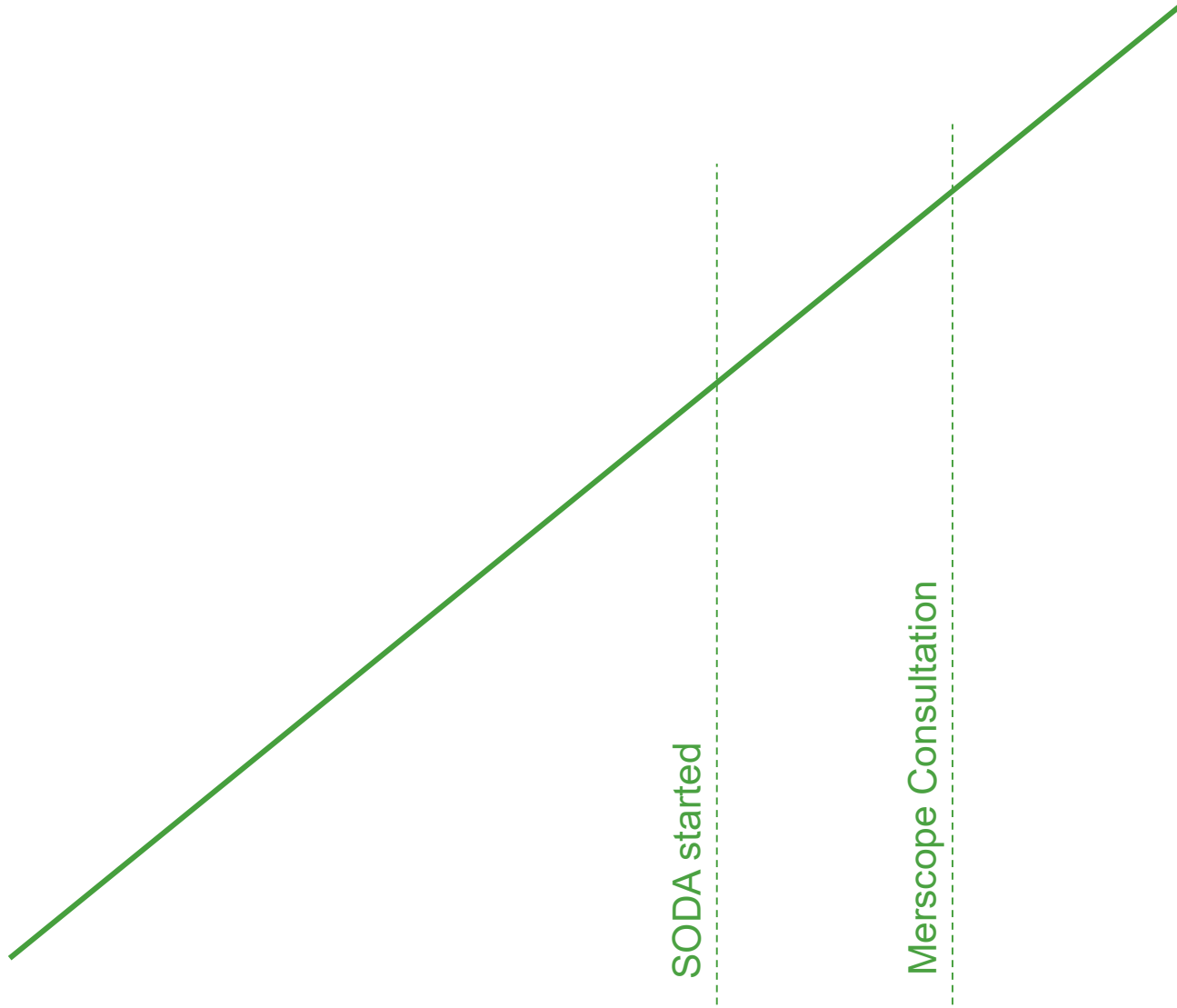


- Ingestion for imaging datasets attach images and metadata to OmeroPlus
- Pipelines developed using Nextflow and launched through Seqera Platform

# Merscope Data Ingestion Workflow

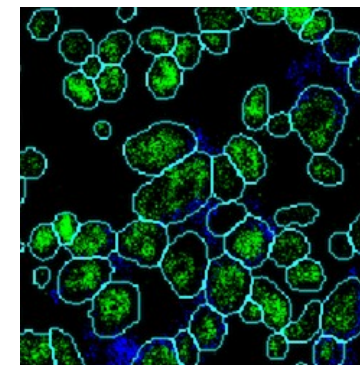
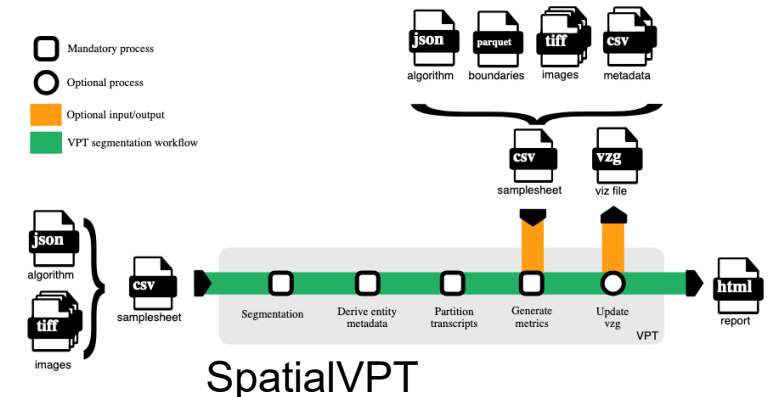


# Data Growth

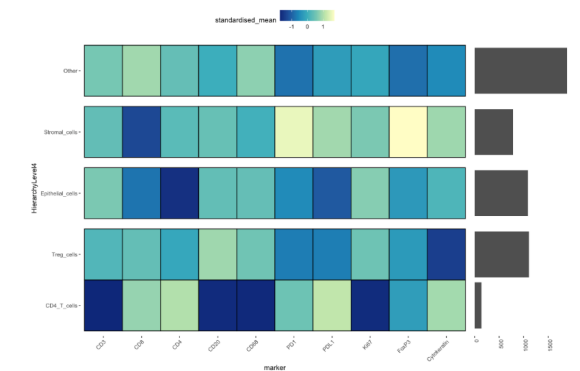


# Pipelines & Tools

- The ingestion runs preprocessing pipelines
  - SODA’s [spatialVPT](#) for Merscope cell segmentation
  - SODA’s [spatialproteomics](#) for COMET cell segmentation
  - For Visium and Xenium, nf-core pipeline
    - **nf-core/spatialvi**
    - **nf-core/sopa**
    - **nf-core/spatialxe**
- Developed other open source tools like
  - [spatialVis R package](#)
- Downstream analysis like cell phenotyping



Cell Segmentation




R Analysis - SpatialVis

# Running Spatialvpt via Seqera Platform

## WEHI-SODA-Hub/spatialvpt pipeline parameters

Upload params file

 Workflow run name

Workflow run name \*  
irreverent\_lavoisier


A unique name randomly assigned to this workflow run. Customize this with a name of your choice (optional).


Labels


MERSCOPE ×


A label must contain at least 2 alphanumeric characters.


>\_ Input/output options

 Segmentation options

 Report options

 Report only options

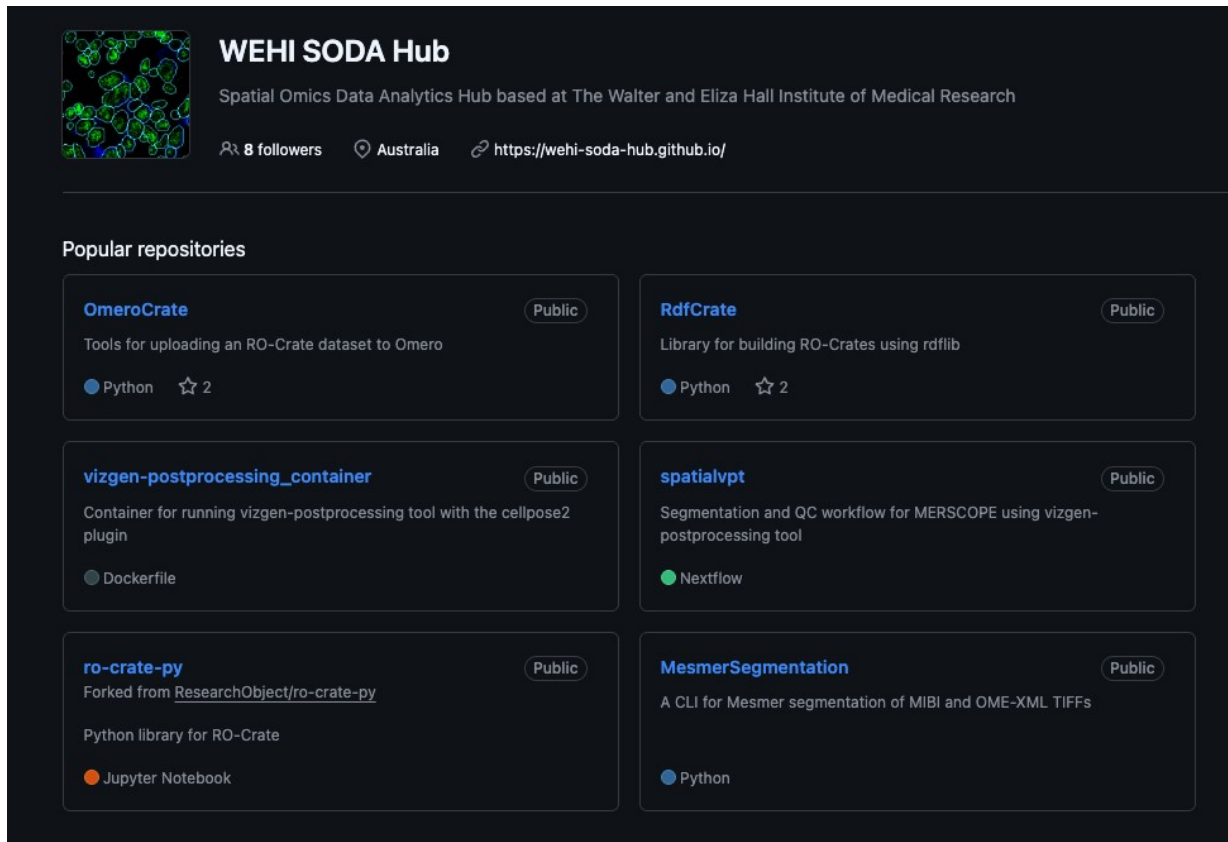
 Generic options

 Show hidden params

Launch settings

Launch

# Know more about the WEHI SODA Hub

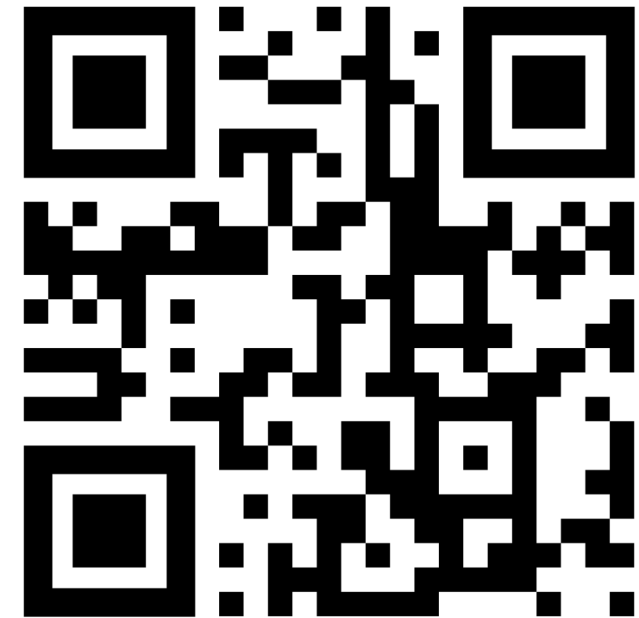


**WEHI SODA Hub**  
Spatial Omics Data Analytics Hub based at The Walter and Eliza Hall Institute of Medical Research

8 followers Australia <https://wehi-soda-hub.github.io/>

### Popular repositories

<b>OmeroCrate</b> Public Tools for uploading an RO-Crate dataset to Omero Python ☆ 2	<b>RdfCrate</b> Public Library for building RO-Crates using rdflib Python ☆ 2
<b>vizgen-postprocessing_container</b> Public Container for running vizgen-postprocessing tool with the cellpose2 plugin Dockerfile	<b>spatialvpt</b> Public Segmentation and QC workflow for MERSCOPE using vizgen-postprocessing tool Nextflow
<b>ro-crate-py</b> Public Forked from <a href="#">ResearchObject/ro-crate-py</a> Python library for RO-Crate Jupyter Notebook	<b>MesmerSegmentation</b> Public A CLI for Mesmer segmentation of MIBI and OME-XML TIFFs Python



<https://github.com/WEHI-SODA-Hub>

# SODA Team

## SODA Chief Investigators



**Prof Marie-Liesse Labat**  
Joint Division Head  
Personalised Oncology Division  
Chair, Spatial Omics Leadership Group



**A/Prof Kelly Rogers, OAM**  
Division Head,  
Advanced Technology & Biology Division  
Head, Centre for Dynamic Imaging  
Co-Chair, Technology Governance Committee



**Prof Tony Papenfuss**  
Deputy Director  
Head, Papenfus Lab, Bioinformatics and Computation  
Biology Division  
Chair, IT Governance  
Chair, Research Data Governance & Management

## SODA Technical Team



**Dr. Julie Iskander**  
Head  
Research Computing Platform  
Bioinformatics Division



**Mr. Michael Milton**  
Research Computing Engineer  
Research Computing Platform  
Bioinformatics Division



**Dr. Marek Cmero**  
Senior Research Officer  
Genomics Lab



**Dr. Lachlan Whitehead**  
Senior Research Officer  
Imaging Lab



**Dr. Pradeep Rajeseckhar**  
Senior Research Officer  
Imaging Lab

# Acknowledgement



**Generously funded by the Kinghorn Foundation**

## **Research Computing Platform**

Mr. Alex Ringeri  
Mr. James U

## **ITS Research systems**

Ms. Miffy Edwards  
Mr. Miguel Esteva  
Mr. Tim Martin  
Ms. Stephanie Gibbs

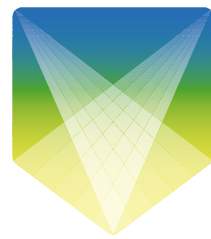
## **Imaging Lab**

Dr. Raymond Yip  
Dr. Claire Marceaux

## **Australian Biocommons**

Dr. Ziad Al Bkhetan





**WEHI**  
brighter together

Questions?

 [WEHI\\_research](#)

 [WEHIresearch](#)

 [WEHImovies](#)

 [WEHI\\_research](#)

 [Walter and Eliza Hall Institute](#)