scverse

Inaugural conference
Technical University, Munich
September 10-12, 2024

Conference booklet

Table of Contents

Day 1
Short talks

Scaling immune-cell receptor analysis in scirpy to millions of single cells
Presenter: **Felix Petschko**

Location: Main conference room
Date / Time: 10th September / 11:00 - 11:15

Abstract:
Scirpy is a scverse core package for single-cell analysis of adaptive immune receptor repertoires (AIRR) in Python. A key functionality is the "clonotype clustering" step that aims to identify immune cells likely to recognize the same antigen or have a common ancestor. As dataset sizes increased to millions of cells thanks to improvements in single-cell technologies, clonotype clustering can become a bottleneck in the analysis pipeline. This project aimed at making it feasible to analyze large-scale, single-cell AIRR datasets with scirpy.

****

User-friendly exploration of (epi)genomic data in single cells using sincei
Presenter: **Vivek Bhardwaj**

Location: Main conference room
Date / Time: 10th September / 11:15 - 11:30

Abstract:
In the past few years, we have witnessed the exciting development of several single-cell epigenomics methods such as scATAC-seq, scBS-seq or scCUTnRUN/CUTnTag. These new technologies have the potential to enable a quantitative analysis of the epigenetic landscape of cells and resolve tissue heterogeneity. However, despite the rising popularity of single-cell epigenomic protocols, issues related to data analysis could limit their broad adoption among biologists: 1) lack of quality control, normalization, and downstream analysis methods tailored to epigenomic data 2) lack of tools that extend beyond droplet-based scATAC-seq methods, 3) need of significant scripting/programming knowledge. We introduce the Single-Cell Informatics (sincei) toolkit, that tackles the above challenges. sincei provides an easy-to-use, command-line interface for the exploration of data from a wide range of single-cell epigenomics and transcriptomics protocols directly from BAM files. It adopts bulk-epigenomics analysis standards into single-cell genomics workflows and simplifies data integration. sincei is available open-source at https://github.com/bhardwaj-lab/sincei. I will introduce the novel features of the sincei toolkit, discuss challenges related to epigenomics, and welcome contributions from the community to help build a scalable interaction between sincei and the anndata/mudata structures.

****

Unleashing the potential of multiplexed imaging experiments
Presenter: **Pierre Bost**

Location: Main conference room

Date / Time: 10th September / 11:30 - 11:45

Abstract:
The last decade has been the witness of a meteoric development of multiplexed imaging (MI) technologies, allowing in theory the in-depth spatial profiling of biological tissues. Here we show that the potential of MI can be unlocked by using tailored experimental sampling and data analysis approaches that embody the intrinsically spatial nature of MI. First, by coupling in-depth sampling simulation with a large scale spatial transcriptomic atlas, the rules governing MI experimental design were elucidated [1], and more efficient sampling strategies were established. Second, inspired by spatial ecology, we developed statistical models to analyze cell count derived from MI data, resulting in a drastic increase of statistical power for differential abundance analysis, as illustrated by the reanalysis of MI COVID-19 data [2]. Finally, we developed a new approach to model the distribution of cells through space in a quantitative manner, thus providing an interpretable feature extraction method to describe the shape of a group of cells.

****

anndataR: easily interact with anndata in R

Presenter: **Louise Deconinck**

Location: Main conference room
Date / Time: 10th September / 11:45 - 12:00

Abstract:
anndataR aims to make it easy to work with anndata files in R, either by converting it to a SingleCellExperiment or Seurat object, or by interacting with them directly. Existing single-cell tools and packages most often live in one of three pre-existing ecosystems. The underlying data structures and on-disk formats differ between these frameworks, making smooth interoperability difficult. Community-based efforts exist to make this conversion process easier, but they rely on using reticulate and while they can make the conversion process easier, it is far from foolproof: due to the differences between the data structures and philosophies underlying the frameworks, lossless conversion is not possible in all cases. That is where anndataR comes in: it aims to both make anndata a first-class citizen in R, and to provide conversion functions to and from SingleCellExperiment and Seurat classes. Using anndataR, you can read and write h5ad files in R without using reticulate and access anndata slots.

Day 1
Poster flash talks

	Title	Presenter	Time	Location
1	High-resolution single-cell atlas reveals diversity and plasticity of tissue-resident neutrophils in colorectal cancer	Valentin Marteau	14:45 - 14:48	Main conference room
2	Deciphering Intra-Tumoral Heterogeneity and Metastatic Processes in Pancreatic Ductal Adenocarcinoma Using In Situ Sequencing	Johannes Wirth	14:48 - 14:51	Main conference room
3	Identifying retinal ganglion cell types	Paulina Paiz	14:51 - 14:54	Main conference room
4	Integrative Analysis of Tumor Microenvironment Dynamics in Response to Immune Checkpoint Inhibition	Zhongyang Lin	14:54 - 14:57	Main conference room
5	Novae: a graph-based foundation model for spatial transcriptomics data	Quentin	14:57 - 15:00	Main conference room
6	SCleeStacks: tool and modality-organized images for containerized development, execution, and analysis in the scverse	Chris Tastad	15:00 - 15:03	Main conference room
7	Spatialproteomics - Streamlining spatial proteomics analysis	Matthias Meyer-Bender	15:03 - 15:06	Main conference room
8	BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis	Vipul Singhal	15:06 - 15:09	Main conference room
9	Heartbreaking - Resolving Myocardial Infarction at Subcellular Resolution	Paul Kiessling	15:15 - 15:18	Main conference room
10	Combinator: A Semi-Supervised Trajectory Inference Algorithm to Track Rare Cell State Transitions in scRNAseq Data	Shyam	15:18 - 15:21	Main conference room

Day 2
Short talks

Higher throughput and fidelity screens with higher accuracy barcode decoding
Presenter: **John Hawkins**

Location: Main conference room
Date / Time: 11th September / 09:15 - 09:30

Abstract:
DNA barcodes and their decoding are at the core of high-throughput molecular biology. Large-scale assays such as pooled CRISPR screens and their single-cell manifestations are pushing the boundaries of what is accurately feasible with current barcode decoding technology, as we ever demand the capability to identify more perturbations, more cells, more omics, more insights in the same assay. Despite the critical nature of the decoding task, few methods exist that permit accurate, flexible and versatile decoding in the more demanding of these applications. Current state-of-the-art is many tools for individual tasks, often with severe limitations in their applicability, and often with inflexible vendor-specific solutions. Here, we propose a principled information theory-based barcode decoding software solution that addresses a wide range of decoding tasks in a unified framework, including demanding applications. Our approach can handle flexible error models, including those encountered in long-read sequencing such as Nanopore, and supports decoding arbitrary and non-trivial constructions of multiple barcodes, linkers, UMIs, and logical combinations thereof. Finally, we present a graphical user interface for clear, user-friendly specification of each barcode decoding task.

****

Characterizing cell state heterogeneity in a primary acute myeloid leukemia hierarchy using Single-Cell Proteomics by Mass Spectrometry
Presenter: **Pedro Aragón Fernandez**

Location: Main conference room
Date / Time: 11th September / 09:30 - 09:45

Abstract:
In the era of single-cell screenings, RNA-based technologies have been the gold standard for the systematic study of cell state heterogeneity and differentiation dynamics of complex cellular systems. Recently, single-cell proteomics by mass spectrometry (scp-MS) has reached a maturity level where it can reliably measure the proteomes of thousands of individual cells in an unbiased manner. Here, we leverage these technological advances for the characterization of a primary acute myeloid leukemia (AML) sample obtained from a relapsed patient. Our analysis not only recapitulates the known hierarchical cellular composition observed in AML but suggest a new non-canonical differentiation avenue through which quiescent leukemic stem cells are able to reconstitute the terminal CD15+CD14+ blast pool, recognized as the main effectors of the disease and targets of commonly used therapeutics. Through trajectory analysis, we show several key proteins exhibiting an asynergic expression profile across the developmental bifurcation leading to a heterogeneous blast reservoir not previously observed by measurements of mRNA transcripts alone.In summary, in this work we demonstrate the capability of scp-MS to interrogate complex biological systems at single cell resolution. We hypothesize a new developmental model in AML

awaiting further experimental validation and underline the potential towards integrative analysis employing different modalities of biological information in combination with scpMS.

****

NeuroVelo: interpretable learning of temporal cellular dynamics from single-cell data
Presenter: **Idris Kouadri Boudjelthia**

Location: Main conference room
Date / Time: 11th September / 09:45 - 10:00

Abstract:
Reconstructing temporal cellular dynamics from static single-cell transcriptomics remains a major challenge. Methods based on RNA velocity are useful, but interpreting their results to learn new biology remains difficult. Here we show NeuroVelo, a method that couples learning of an optimal linear projection with non-linear Neural Ordinary Differential Equations. Unlike current methods, it uses dynamical systems theory to model biological processes over time, hence NeuroVelo can identify what genes and mechanisms drive the temporal cellular dynamics. We benchmark NeuroVelo against several state-of-the-art methods using single-cell datasets, demonstrating that NeuroVelo has high predictive power but is superior to competing methods in identifying the mechanisms that drive cellular dynamics over time. We also show how we can use this method to infer gene regulatory networks that drive cell fate directly from the data.

****

Comprehensive Analysis Integrating Single-Cell Spatial Transcriptomics, Genomics, and Histology in Gastric Cancer
Presenter: **Sano Kyohei**

Location: Main conference room
Date / Time: 11th September / 11:30 – 11:45

Abstract:
Gastric cancer is a highly heterogeneous tumor, not only in its clinical and biological behavior but also histologically. Numerous classification attempts have been made in the past, but these have relied on histological analysis and bulk sequencing, potentially failing to fully capture the complexity of this disease. Therefore, we collected single-cell level spatial transcriptomics (scST) data from large number of surgical cases. This gastric cancer atlas is unprecedented in scale for scST analysis, incorporating genomic mutation data and H&E-stained images, enabling more comprehensive analysis. We used a deep neural network model to extract morphological features from pathological images and quantitatively integrated the three modalities of scST, genomic mutation information, and pathological images. This allowed for a comprehensive interpretation of how tumor cells with specific gene variation affect surrounding cells, form niches, and how these are reflected in histological variation. Using this atlas, we identified characteristic niches associated with drug response and elucidated the interactions between stromal cells and tumor cells as well as distinctive histology. This study provides a new standard for detailed tumor characterization and potentially paves the way for advanced precision medicine in gastric cancer.

****

Analysis of tumor infiltrating lymphocytes in the AML bone marrow
Presenter: **Nicholas Ceglia**

Location: Main conference room
Date / Time: 11th September / 11:45 – 12:00

Abstract:
While immune-based therapies have transformed the treatment of many hematologic malignancies, T cell-based immunotherapies for acute myeloid leukemia (AML) have had limited success. However, the curative potential of donor T cells in allogeneic hematopoietic cell transplantation underscores the need to understand the mechanisms behind ineffective anti-leukemic T cell activity to advance these therapies. To investigate, we performed longitudinal analysis of single cell RNA-seq on pre- and serial post-treatment samples with paired CITE-seq and VDJ sequencing for 38 patients with newly diagnosed AML. These analyses reveal the immunosuppressive nature of T cells within the AML BM, the dynamic evolution of T cell clones across activation and exhaustion states, and the immunomodulatory impact of malignant AML cells on T cells. Finally, we identified key T cell intrinsic features affecting the relationship between T cell phenotype, repertoire, and disease status in the AML BM. We hypothesize that both malignant and non-malignant cells in the AML BM contribute to impaired T cell immunity, resulting in distinct T cell compositions across different disease states.

****

DeepSpaCE: A deep learning framework to detect spatial single cell domains
Presenter: **Behnam Yousefi**

Location: Main conference room
Date / Time: 11th September / 14:15 – 14:30

Abstract:
The rapidly growing field of spatial transcriptomics has provided new means to discover cellular states within their local microenvironments, thereby revealing the spatial structures that form the basis of organ function, development, and disease pathology (Sudupe, Laura, et al., 2023). Cells that are spatially close to each other are often deemed to be functionally related, emphasizing the need to consider spatial relationships to understand cellular roles in the tissue. However, despite significant advances in spatial molecular imaging technologies, such as MERFISH and 10X Xenium, current methodologies face considerable challenges in accurately deciphering spatial domains and cell niches. In this work, we introduce DeepSpaCE, a Deep learning-based Spatial Cell Explorer model, that employs graph neural networks to address the challenge of identifying spatial domains and cell niches with high accuracy and robustness. We summarize the significance and novelties of our work in the following. Leveraging a GNN-based encoder-decoder architecture and bi-objective function, DeepSpaCE takes into account both cell transcriptomic data and spatial relationships among cells. DeepSpaCE requires only three hyperparameters. We either demonstrate the robustness of our pipeline to the choice of a hyperparameter or provide algorithmic solutions to find the optimum hyperparameter. We applied DeepSpaCE on spatial single-cell data and showed the robustness of our method in accurately detecting the spatial domains and its superiority over existing methods. Additionally, we utilized DeepSpaCE to identify spatial domains in in-house kidney biopsies obtained using 10X Xenium technology. The results demonstrated the potential of DeepSpaCE to significantly enhance the accuracy of spatial domain detection in diverse tissue types, thereby providing a means to more precise and insightful spatial transcriptomic analyses.

****

SPARCSpy: Building single cell representations based on microscopy images
Presenter: **Sophia Mädler**

Location: Main conference room
Date / Time: 11th September / 14:30 – 14:45

Abstract:
Microscopy imaging routinely produces large datasets that capture the spatial composition and arrangement of millions of cells. Gaining biological insights from these data requires transforming images into descriptive features and integrating single cell information across datasets. Here we introduce SPARCSpy, a computational framework to generate single cell representations from raw microscopy images. SPARCSpy solves several challenges that come with scaling image operations such as stitching and segmentation to millions of cells. Out-of-core computation enables SPARCSpy to efficiently handle datasets in which individual images, for example covering whole microscopy slides, exceed available memory. By introducing an open file format that interfaces with OME-NGFF and scverse spatialData, SPARCSpy facilitates the integration of newly recorded and publicly available datasets. To generate meaningful single cell representations, SPARCSpy's standardized data format directly enables training and applying the latest deep learning-based computer vision models. We demonstrate the utility of SPARCSpy on several biological use cases including phenotype identification in image-based genetic screening and representation learning on highly multiplexed microscopy datasets.

Workshops

Time slot 15:45 – 16:45

Avoiding common pitfalls in single-cell analysis – hands-on guide through preprocessing, annotation, and differential expression.
Presenter: **Mark Sanborn**

Location: Main conference room
Abstract:
This live coding workshop is an introduction to preprocessing and the core analyses of single-cell sequencing. We will focus on the best practices while avoiding pitfalls that are commonly missed or leftover from earlier single-cell recommendations. We will briefly cover background correction, outlier filtering, transformation, and integration. The emphasis of this workshop will be annotation, differential expression, and pathway analysis. Each attendee will have access to a dataset and optional analysis environment.

Projections, interactive plots, and interactive online documentation of your scRNA project
Presenter: **Samuele Soraggi**

Location: Workshop room I
Abstract:
A small workshop on existing projection methods for high-dimensional data such as scRNA-seq, interactive plots in python and how to create a documentation of your research that can be hosted online (including the interactive plots). Associated to this there will be a google colab notebook, github repository and webpage for the participants as active exercise and documentation for future use. Optional/Reading: The PacMap paper is a good technical intro to dimensionality reduction https://www.jmlr.org/papers/volume22/20-1061/20-1061.pdf

---

Interactive AnnData & SpatialData analysis
Presenter: **Luca Marconato**

Location: Workshop room II
Abstract:
Introduction to interactive AnnData analysis with cellxgene, and interactive SpatialData analysis with Napari, vitessce and more.

Workshops

Time slot 17:00 – 18:00

****

Single-cell analysis - dos and don'ts in trajectory inference
Presenter: **Roshan Sharma**

Location: Main conference room
Abstract:
We will use Scanpy to discuss some of the basic pre-processing of single-cell data, and transition to discuss common trajectory methods. Major emphasis of the workshop will be on guidelines for trajectory detection - things a user is recommended to keep in mind when doing such analyses. We will provide example code (with Palantir and CellRank) and illustrations. All code used will be made publicly available.

Marsilea: Declarative creation of composable visualization for Omics data
Presenter: **Yimin Zheng**

Location: Workshop room I
Abstract:
In this workshop, we will learn to use Marsilea, a newly developed Python package to transform your omics data into composable visualization. Agenda: (1) Introduction of "cross-layout" and Marsilea - The talk will start by discussing the limitations of traditional single-plot methods in visualizing complex scientific data and highlighting the advantages of composable visualization techniques. Next, we will provide an in-depth overview of Marsilea by introducing the concept of "cross-layout" paradigm to create the composable visualization. After that, we will see how "cross-layout" is applied in Marsilea. We will then talk about the design and advantages of Marsilea over other plotting libraries. A minimum example will be used to illustrate how a composable visualization can be built with Marsilea incrementally in a declarative way. (2) Interactive Hands-on Session - We like to engage attendees to recreate three composable visualizations in single-cell omics with provided datasets and Jupyter Notebook in a preconfigured environment (etc, Google Colab). They will get hands-on creating a complex heatmap visualizing the expression profile of single cells, a dot heatmap to visualize the cell-cell communications, and a track plot to visualize ATAC-seq data. When attendees are familiar with the usage of Marsilea, they will be asked to use a provided dataset and maximize their creativity to build a visualization using Marsilea. All masterpieces will be showcased on the Marsilea website. (3) Conclusion and Wrap-Up - In the end, the talk will be wrapped up with a cheat sheet to review the "cross-layout" concept and Marsilea's APIs, and more examples of the broad applications of Marsilea beyond omics data will be showcased.

Benchmarking Open Problems in Single Cell Analysis
Presenters: **Malte Luecken, Robrecht Cannoodt**

Location: Workshop room II
Abstract:

The "Benchmarking Open Problems in Single Cell Analysis" workshop aims to address critical challenges in the field by fostering community engagement in the development of robust benchmarks. This 90-minute session will introduce participants to the core mission of the Open Problems in Single-Cell Analysis, followed by an interactive session where we will build a new benchmark from scratch. As part of this tutorial, participants not only learn about the technical aspects of setting up a benchmark within the Open Problems framework, but will also learn about essential best practices in benchmarking computational methods.

Day 3
Short talks

Fast and accurate cell segmentation of highly multiplexed spatial omics using graph neural networks with segger

Presenter: **Elyas Heidari**

Location: Main conference room
Date / Time: 12th September / 12:00 - 12:15

Abstract:
Imaging-based spatial omics datasets present challenges in reliably segmenting single cells. Achieving accurate segmentation at single-cell resolution is crucial to unravelling multicellular mechanisms and understanding cell-cell communications in spatial omics studies. Despite the considerable progress and the variety of methods available for cell segmentation, challenges persist, including issues of over-segmentation, under-segmentation, and contamination from neighbouring cells. While combining multiple segmentation methods with distinct advantages has been proposed, it does not completely resolve these issues. Additionally, scalability remains an obstacle, particularly when applying these methods to larger tissues and gene panels in targeted studies.
Here we introduce Segger, a cell segmentation model designed for single-molecule resolved datasets, leveraging the co-occurrence of nucleic and cytoplasmic molecules (e.g., transcripts). It employs a heterogeneous graph structure on molecules and nuclei, integrating fixed-radius nearest neighbor graphs for nuclei and molecules, along with edges connecting transcripts to nuclei based on spatial proximity. A heterogeneous graph neural network (GNN) is then used to propagate information across these edges to learn the association of molecules with nuclei. Post-training, the model refines cell borders by regrouping transcripts based on confidence levels, overcoming issues like nucleus-less cells or overlapping cells. Benchmarks on 10X Xenium and MERSCOPE technologies reveal Segger's superiority in accuracy and efficiency over other segmentation methods, such as Baysor, Cellpose, BIDcell, and simple nuclei-expansion. Segger can be pre-trained on one or more datasets and fine-tuned with new data, even acquired via different technologies. Its highly parallelizable nature allows for efficient training across multiple GPU machines, facilitated by recent graph learning techniques. Compared to other model-based methods like Baysor and BIDCell Segger's training is orders of magnitude faster, and more accurate making it ideal for integration into preprocessing pipelines for comprehensive spatial omics atlases.

Workshops

Time slot 09:00 – 10:00

Training models on atlas-scale single-cell datasets
Presenter: **Ryan Williams**

Location: Main conference room
Abstract:
As the field of single-cell RNA sequencing continues to evolve, researchers are increasingly interested in using these datasets to train foundational models for a wide range of applications. While training models on smaller datasets that fit into memory is relatively straightforward, scaling up beyond single machines presents significant technical challenges. In this workshop co-presented by TileDB and the Chan Zuckerberg Initiative, participants will receive hands-on experience training models on a large dataset comprising 70 million cells and learn about the key technologies and resources that make this possible.

Good first contributions
Presenter: **Mikaela Koutrouli**

Location: Workshop room I
Abstract:
An overview on how do to your first contribution to scverse community, even if you are not a developer including: replying on discourse queries, contributing to documentation, best practices book, posting and using issues even for problems you have solved, where to find the contributing guidelines.

Data Version Control: the missing link in team science
Presenter: **Chris Tastad**

Location: Workshop room II
Abstract:
Advancements in tools and data management ecosystems have led to a mature state for collaboration across research compute efforts. The refinement of code hosting, dependency management, DevOps, documentation, and automation have established a range of workflows that remove barriers and accelerate scientific endeavors across teams and communities. Still, many of these systems place an emphasis on the management and organization of code. The complexity and importance of code base management cannot be overstated, but there remains a comparable set of challenges around handling the evolution of a shared dataset. Much like shared code, shared data can be complex, iterative, transitory, and distributed. Collaborative efforts with shared data also retain the same needs for data traceability, reproducibility, and portability. Among all available collaborative resources, a common problem stands in the way of addressing these needs - data are large.

To confront this, our group has implemented an approach to using Data Version Control (DVC) for team science with bioinformatics applications and single cell data. DVC is a data science toolset created by Iterative.ai for change tracking inputs in machine learning workflows. At its core, this resource offers a generic framework for codified

data versioning and provenance which sits on top of well-established tools. This key element of the DVC paradigm allows for the separation of data from metadata in change tracking. Object-linked metadata are managed through git while objects can be stored in a separate remote storage. This work offers a guide of the DVC framework that is intended to service requirements specific to the needs of bioinformatics and single cell data. We show how to tailor tracking of data in a manner that can efficiently address challenges around the diverse set of intermediary single cell outputs. We also place a particular emphasis on the need to bridge gaps in data hand-offs that may exist between hybrid collaborations of dry and wet lab scientists. Altogether, we hope this application of a lesser-known toolset offers an on ramp to the adoption of stronger data management practices, enabling collaboration across single cell biology and beyond.

Workshops

Time slot 14:00 – 15:00

GPU accelerated single-cell analysis
Presenter: **Severin Dicks**

Location: Main conference room
Abstract:
Learn how to incorporate NVIDIA GPU acceleration in your day-to-day data analysis. This workshop will demonstrate how to use rapids-singlecell to accelerate traditional Scanpy and other scverse workflows. Attendees will gain practical insights into leveraging GPUs outside of deep learning, exploring the capabilities of RAPIDS and CuPy. Discover how these tools can enhance your package and improve your single-cell data analysis efficiency.

Get your package ready for the scverse ecosystem + cookiecutter template
Presenters: **Lukas Heumos, Philipp Angerer**

Location: Workshop room I
Abstract:
Introducing the cookiecutter template and how to use it. We also introduce the scverse ecosystem and how to contribute to it.

Zoom & Enhance: Enabling Subcellular Analysis of Spatial Transcriptomics with Bento
Presenter: **Clarence Mah**

Location: Workshop room II
Abstract:
This goal of this workshop is to demonstrate how spatial transcriptomics data can be used to extract subcellular insights about cell biology, primarily emphasizing analysis with Bento – a package in the Scverse ecosystem – that utilizes the SpatialData framework enabling interoperability with Scanpy and Squidpy. The workshop will include a brief intro to spatial transcriptomics technologies capable of subcellular measurements. Then we will introduce hands-on exercises (Jupyter notebooks) for spatial analysis, such as annotating subcellular spatial patterns and domains, gene colocalization, and measuring cell morphology. We will conclude with a short Q&A session for more open-ended discussion.

Workshops

Time slot 15:15 – 16:15

Big Data in AnnData
Presenter: **Ilan Gold**

Location: Main conference room
Abstract:
As both individual datasets and atlases grow, so to should the capacities of our data structures. Listen in to find out how AnnData is addressing big data, and preparing for the future.

Polyglot programming for single-cell analysis
Presenters: **Louise Deconinck, Benjamin Rombaut, Robrecht Cannoodt**

Location: Workshop room I
Abstract:
Any bioinformatician that has analysed a single-cell dataset knows that using methods developed for different ecosystems or programming languages is necessary but painful. Any package developer has asked themselves the question on how to best provide access to their tool or method. We will give an overview of the interoperability tools you can use when analysing a single-cell dataset: do you want to convert your data to a different data format, or is just calling one R function in your Jupyter notebook sufficient? Do you want fine-grained control over each step in the analysis pipeline or do you run a series of scripts that you really should convert to a workflow system? We will give information on different options for package developers to provide better interoperability. Should you reimplement your package in a new language? How do you ensure that the results are the same? In order to follow this workshop, we expect the participants to have some Python or R programming knowledge.