**Speaker**: Fabian Theis

Institute of Computational Biology, Helmholtz Munich - http://comp.bio

**Title**: Latent space learning in single cell genomics

**Abstract**:

Modeling cellular state as well as dynamics e.g. during differentiation or in response to perturbations is a central goal of computational biology. Single-cell technologies now give us easy and large-scale access to state observations on the transcriptomic, epigenomic and more recently also spatial level. In particular, they allow resolving potential heterogeneities due to asynchronicity of differentiating or responding cells, and profiles across multiple conditions such as time points, space and replicates are being generated, with a series of implications across biology and medicine.

Most computational methods for single cell genomics are operating on an intermediate often nonlinear representation of the high-dimensional data such as a cell-cell knn graph or some more general latent space. Interpretation of these led already in early days towards models of cellular differentiation for example by pseudotemporal ordering or mapping time information. Hence latent space modeling and manifold learning have become a popular tool to learn overall variation in single cell gene expression, more recently also across data sets and modalities.

After a short review of these approaches, I will discuss how latent space learning can be achieved using variants of autoencoders, with applications from denoising, imputation to learning perturbations. I will then show how it can be used to integrate single cell RNA-seq data sets across multiple labs in a privacy-aware manner, and demonstrate mapping disease variation by querying COVID-19 patients ontop of a healthy immune reference atlas. I will present our recent resource Sfaira of data loaders and shared latent spaces across tissues, and finish with short outlook towards spatial modeling and interpretability of latent projections under perturbations.