

5<sup>th</sup> International Conference on

# Big Data Analysis and Data Mining

June 20-21, 2018 | Rome, Italy

## High dimensional single cell gene expression data and batch effect corrections

Laleh Haghverdi

EMBL Heidelberg, Germany

Emerging about a decade ago, single cell genes expression measurement technologies have facilitated the study of heterogeneous populations of cells such as in development and cell differentiation. Single cell ribonucleic acid sequencing (scRNA-seq) techniques can measure the expression level of several thousand genes at the single cell level for millions of cells. Increasingly used by several laboratories, the technique provides a big amount of data which opens new opportunities for knowledge extraction using new machine learning and computational methods. The author will discuss the properties of high-dimensional data which needs to be taken care of when dealing with such big expression data, and discuss in an instance on how high-dimension properties allowed us to develop a new method for batch effects correction and data integration across several laboratories.

[laleh@ebi.ac.uk](mailto:laleh@ebi.ac.uk)