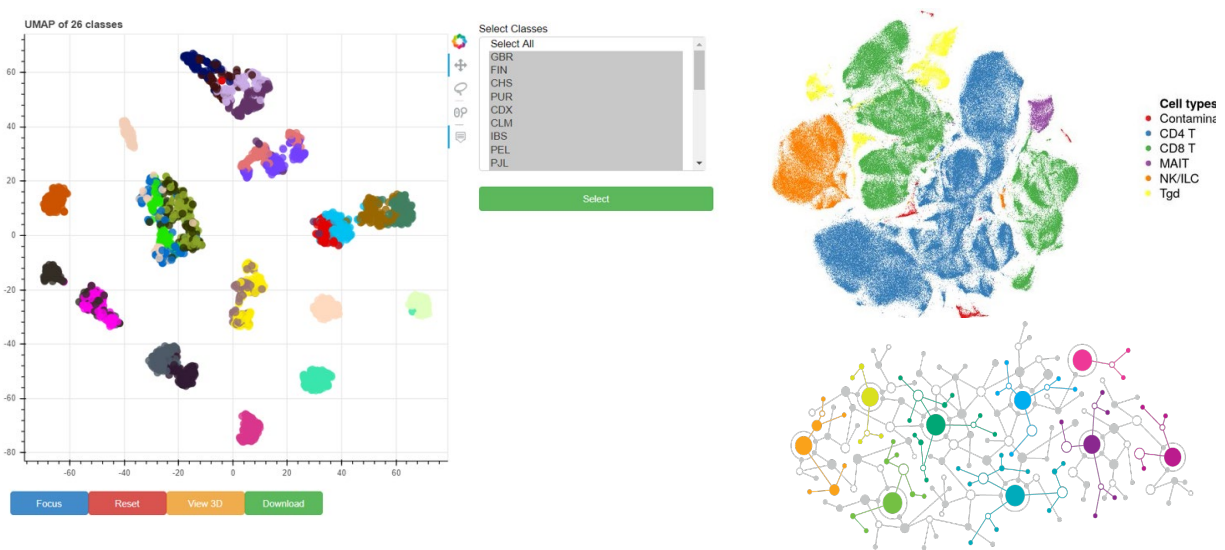


GuiDR: an interface to visualize and cluster high dimensional data



Project Description:

In the big data era, analysts need to understand complex phenomena in various domain of science like genomics, climate, cyber-security through their associated data. These data are high-dimensional and the challenge is to visualize them in a trustworthy way so cluster and outlier patterns visible in their 2D graphical representation correspond to actual clusters and outliers of the data.

We develop GuiDR, an interactive visualization tool to discover real clusters in high-dimensional data through their distorted visualizations. We want to add new functionalities to this tool, enabling to store, and visualize as a network, the current status of the exploration conducted by the analyst. We also want to navigate the history of the user interactions to get an overview of the exploration paths.

Duties/Activities: Implementing the network view and the history view of the user exploration path with Bokeh and Python 3, on top of the existing tool.

Required Skills: Base knowledge of Python 3 and Bokeh, base knowledge of projection techniques.

Preferred Intern Academic Level: BSc 3rd year or BSc 4th year

Learning Opportunities: Advanced visualization and interactions with Bokeh, analysis of high-dimensional data

Expected Team Size: 2

Mentors

Name: Michaël Aupetit email: maupetit@hbku.edu.qa

Abdelkader Baggag abaggag@hbku.edu.qa