# Hands-on Network Biology Analysis using NetworkX and yEd

## Objective:

Analyze a biological network representing protein-protein interactions.

Introduce fundamentals of graph theory, centrality measurements, and graph models.

Tools:

* Python (with NetworkX library installed)
* yEd Graph Editor

Biological Network: Human PPi network

We will use a protein-protein interaction network based on data from a hypothetical study.

## Exercises:

### E1 Degree Distribution Analysis:

Calculate the degree distribution of the protein-protein interaction network.

Plot a histogram showing the degree distribution.

**Centrality Comparison:**

Calculate and compare different centrality measures:degree centrality, betweenness centrality) for the network.

Identify the top 10 central nodes in the network based on each centrality measure.

### E2 Graph Model Simulation:

Create and visualize an Erdos-Renyi graph and a Barabasi-Albert graph with similar characteristics (number of nodes, average degree) as the protein-protein interaction network.

Compare the structural properties (e.g., degree distribution, clustering coefficient) of the simulated graphs with the protein-protein interaction network.

Path Analysis:

Find the shortest path between randomly selected pairs of nodes in the protein-protein interaction network.

Calculate the average shortest path length for the network.

Community Detection:

Use a community detection algorithm (e.g., Louvain algorithm) to identify communities in the protein-protein interaction network.

Visualize the identified communities in the network.

Node Removal Impact:

Simulate the impact of removing high centrality nodes from the network.

Measure changes in network connectivity and identify critical nodes for network integrity.

Graph Comparison:

Compare the structural properties of the protein-protein interaction network with a random network (Erdos-Renyi) and a scale-free network (Barabasi-Albert).

Discuss the relevance of each model in representing biological networks.

Visualization Enhancement:

Enhance the visualization of the protein-protein interaction network in yEd by customizing node colors, sizes, and edge styles based on node attributes (e.g., centrality measures).

Real Data Analysis:

(Optional) Obtain a real protein-protein interaction network dataset from resources like STRING or BioGRID.

Analyze the real dataset using the same techniques and compare the results with the hypothetical network.

Network Comparison:

Compare the properties of the protein-protein interaction network with other types of biological networks (e.g., metabolic networks, gene regulatory networks).

Discuss the similarities and differences in network properties across different biological systems.