

Arena3D: A visualization tool for high level relationship and clustering analysis in large scale networks

Georgios Pavlopoulos, PhD Student

EMBL Heidelberg, Meyerhofstraße 1, 69117 Heidelberg, Germany
pavlopou@embl.de

The analysis and interpretation of relationships between biological molecules, networks and concepts is becoming a major bottleneck in systems biology. Very often the pure amount of data and their heterogeneity provides a challenge for the visualization of the data. There exist a wide variety of graph representations, which most often map the data on 2D graphs. These methods are applicable to a wide range of problems, nevertheless many of them reach a limit in terms of user friendliness and visualization when a large number of nodes and connections have to be analyzed and visualized.

Here we present a new visualization tool called Arena3D which introduces a new, staggered multi layer concept that allows the analysis of big networks in a three dimensional space representation. The different layers in the representation correspond to different data types respective concepts like sequences, structures, chemicals diseases, pathways etc. The data entries for one specific data type, like sequences, can be ordered or clustered on their respective layer by applying a data focused similarity measurement like sequence similarity. Several clustering methods are also supported to cluster several data types in different layers. The different nodes of the layers can be connected according to known or predicted relationships between the nodes. The relationships are typically extracted from available databases (SRS, web services) and available text mining machineries (Bioalma, ArenaText) and are predicted by various data mining methods or originate from experimentally generated data. Indirect connections that may hide some additional information can also be explored.