

The background features a complex network of red lines connecting green dots, overlaid on a light blue grid. A white banner with a subtle geometric pattern is positioned across the center, containing the title text. To the left of the banner, there is a small inset image showing a cluster of orange and red dots.

Session 6: An Overview of Clustering Different Types of Data

Question

- ❑ Consider a heterogeneous network of publications that contains nodes of type *author*, *paper*, and *venue*, where *papers* connect to *authors* via the *written-by* relation, *papers* connect to *venues* via the *published-in* relation, and *papers* connect to each other via the *cited-by* relation. What are some feasible ways to transform this network into a homogeneous network?
- ❑ We can create a citation network by only keeping *papers* and the *cited-by* links (induced subgraph on nodes of the type *paper*).
- ❑ We can create a co-authorship network by keeping *authors* and creating links between *authors* that are connected to the same *paper*.
- ❑ Answer: Both. Here we give two examples of projecting a heterogeneous network into a homogeneous network by leaving out certain information about the complete dataset for the sake of homogeneity.