ABSTRACT: In a lot of the lectures the speakers described a similar workflow and software stack for data mining processes. Lennart Heuckendorf gave an overview over the software stack typically used. This poster generalizes the typical data mining workflow seen in the lectures and shows typical software used in each step.

1. Data Integration & Data Cleaning
Remove inconsistencies and bring data from heterogeneous sources into an uniform format. Therefore one needs to deal e. g. with missing values, mapping attribute names and duplicates.

2. Data Mining
Depending on the task, techniques are chosen and implemented to find new information in the data. The techniques reach from simple statistics, clustering, outlier mining and classification techniques to using neural networks.

3. Data Visualization
Visualize the patterns one found, or try to find new information through visual data mining.