# Introduction

* What is the problem?
* Motivation: why the chosen topic is interesting and relevant?
* Goals: what is it the paper strives to achieve? What is the contribution?
* Outline:
  + short presentation of the approach and structure of the paper / welche Gedanken gemacht / wie genau vorgehen
  + KDD (Source: The KDD Process for Extracting Useful Knowledge from Volumes of Data; U s a m a F a y y a d , G r e g o r y P i a t e t s k y - S h a p i r o , a n d P a d h r a i c S m y t h)
    - Framework for getting from data to knowledge
    - Keep in mind that it is not a binding sequence, but includes numerous iterations (for optimizations etc .)
    - Selection: select out of data the target data which should be discovered/analyzed
    - Preprocessing: exploratory data analysis and data cleaning (handle noise, outliers, missing values, transform 99 values)
    - Transformation: transform preprocessed data into transformed data (Data reduction: PCA, feature selection), encoding and projection (scaling…)
    - Data mining: choose DM function (clustering), algorithms (which to use for pattern search in data, models and parameters)
    - Interpretation: interpret results from data mining and possibly connections to previous steps: what can be improved?
  + 1.Iteration,
    - Focus on rather simple and fast methods
    - Decide with methods work well and which should be replaced by more complex ones
  + 2. Iteration
    - Try more complex methods
    - E.g. multidimensional outlier detection