A Hybrid Imputation Approach Using Web Tables
Ahmad Ahmadov, Maik Thiele, Wolfgang Lehner, Robert Wrembel

What can be done to improve data quality?

1. Entity Augmentation
2. Tuple Augmentation
3. Entity Verification
4. Data Imputation
5. Table Augmentation
6. Data Standardisation

Web Based Imputation Approach

Model Based Imputation Approach

Motivating Example Dataset

Companies Dataset – Impute Marketvalue

Countries Dataset – Impute Poverty

Using Hybrid Approach:

Companies Dataset:
Significant increase of imputation accuracy by increasing coverage with the help of web data

Countries Dataset:
Slight increase in overall imputation accuracy using the hybrid approach.