

# Advanced Data Engineering

## Topics of Interest

- **Data collection:** how do we find relevant data sources?
- **Data integration:** Given the properties of Big Data, how can data from multiple data sources be combined to get a more global perspective on a subject to be analyzed?
- **Data cleaning:** How can important properties and errors of data be assessed and corrected?
- **Provenance:** How can the whole data engineering process be documented and using so-called provenance meta-data?
- **Systems:** What system architectures support data engineering?

In the era of Big Data, it is more important than ever to manage the available data such that it can be reasonably used. One use case for instance is analysis of the data with the overall goal to gain new insights. Indeed, business analytics, data mining, or machine learning rely on a large data basis. However, the results of an analysis typically depend on the data they are based on, so it is crucial that these data are correct, meaningful, accessible, up to date, etc.

The goal of data engineering is to ensure these qualities of the data. To this end, data engineering focuses on technologies, software, algorithms, and tools to support elevating raw data to information useful for further analysis. This seminar covers selected research on foundations, algorithms, and systems for data engineering.

## Prerequisites

- The language of this seminar is **English**, i.e., to participate you have to be able to write a document and give a presentation in English!
- Knowledge of the subjects addressed in the **Data Engineering or Information Integration lecture** (or equivalent) are required.

## Additional Information

### Lecturer

Prof. Melanie Herschel  
and team

IPVS / Data Engineering

### Course Website

[http://www.ipvs.uni-stuttgart.de/abteilungen/as/abteilung/mitarbeiter/Melanie.Herschel\\_infos/teaching/seminar\\_data\\_engineering\\_ws1718.html](http://www.ipvs.uni-stuttgart.de/abteilungen/as/abteilung/mitarbeiter/Melanie.Herschel_infos/teaching/seminar_data_engineering_ws1718.html)