In the era of Big Data, it is more important than ever to adequately manage data such that they can adequately be leveraged in various domains and applications. One common use is for instance the analysis of data with the goal to gain new insights. Indeed, business analytics, data mining, or machine learning rely on a large data basis, and their result quality highly depends on the quality of the underlying data. Therefore, it is crucial that these data are correct, meaningful, accessible, up to date, unbiased, etc.

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

Topics of Interest

- **Data integration**: How can data from multiple 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?
- **Meta-data**: How can data and the data engineering process be documented?
- **Systems**: What system architectures support data engineering?
- **Applications**: What data engineering techniques need to be devised to support applications with varying requirements and of different domains?

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.

Organizers

Prof. Melanie Herschel (examiner) and team
IPVS / Data Engineering

Course Details

- The seminar kicks off during the first two weeks of the lecture time. All registered students should be present for the introductory session to secure their spot. No shows risk to lose it to students on the waiting list, encouraged to attend as well.
- Further details will be communicated via Campus/ILIAS approximately two weeks before lectures start.