

BDASEM: Big Data Analytics Seminar (2 SEM)

0434 L 477

Content:

Both the sciences and industry are currently undergoing a profound transformation: largescale, diversedata sets - derived from sensors, the web, or via crowd sourcing - present a huge opportunity for datadriven decision making. This data poses new challenges in a variety of dimensions: in its unprecedented volume, in the speed at which it is generated (its velocity) and in the variety of data sources that need to be integrated. A whole new breed of systems and paradigms is currently developed to be able to cope with that these challenges. The field of Big Data Analytics deals with the technological means of gaining insights from huge amounts of data. In this seminar, students will review the current state of the art in this field. At the beginning of the semester students will receive a set of primary literature, which consists of basic reading material for every participant. Students will review this material under the guidance of a mentor. At the same time, students will learn about presentation techniques, techniques for critical reading of scientific papers, as well as scientific writing. Students will use secondary literature to research the assigned topic assigned, including research journals and articles published at information management conferences such as WWW, VLDB, or SIGMOD.

Participants of this seminar will acquire knowledge about recent research results and trends in the analysis of web-scale data. Through the work in this seminar, students will learn the comprehensive preparation and presentation of a research topic in this field. In order to achieve this, students will get to read and categorise a scientific paper, conduct background literature research and present as well as discuss their findings. After the course, students will be able to critically read and evaluate scientific publications, and to conduct background research. They will be capable of preparing for and giving oral presentations on research topics for an expert audience, of analyzing the state of the art of a research topic, and of summarizing it in a scientific paper. They should also understand techniques used in the scientific community like peer reviews, conference presentations, and defenses of the findings after their presentation, as well as they should understand methods for large-scale data analytics.

The course is principally designed to impart:

Technical skills: 50x, Methodological skills: 20x, System skills: 10x, Social skills: 20x

At the beginning of the semester you will receive a set of primary literature. Then you should find and use secondary sources to investigate the topic assigned to you in the seminar, which should go beyond the supplied primary literature. A few weeks after receiving the primary literature you will have to give a short presentation of approx. 10 minutes. The presented topic will be discussed by the other seminar participants in order to give you constructive feedback about the content of your talk. In the further course of the semester you will have to give a long final presentation of approx. 30 minutes. In addition to the talk you will also have to hand in a 10-15 page technical report describing and evaluating your topic. Details may vary and will be announced in the respective semester.

Target group:

This course is aimed at **master students** with a focus on database systems and information management and should be chosen after the 2nd semester.

Prerequisite:

To be able to participate you should have successfully completed 'Database Technology DBT' and one of the 'Advanced Information Management AIM' (1,2,3) courses. In the ideal case this seminar should be taken by students who are directly in front of their master thesis, or have already completed it and are interested in a doctorate in information management.

It is vital that you have a sound understanding of written and spoken English.

Registration:

Students are required to register via the DIMA course registration tool before the start of the first lecture

(<http://www.dima.tu-berlin.de>). Within the first six weeks after commencement of the lecture, students will have to register for the course at **QISPOS (university examination protocol tool)** and **ISIS (course organization tool)** in addition to the registration at the DIMA course registration tool.

Contributions:

PORTFOLIOPRÜFUNG

The exam will be done as a 'portfolio examination', including two deliverable assessments, totaling for 100 portfolio points:

- seminar presentation (50 portfolio points) (about 60 Min.)
- written seminar report (50 portfolio points) (20 - 25 pages)

The final grade according to § 47 (2) AllgStuPO will be calculated with the faculty grading table 2.

(Die Gesamtnote gemäß § 47 (2) AllgStuPO wird nach dem Notenschlüssel 2 der Fakultät IV ermittelt.)

Short Comment:

Das Modul kann in 1 Semester(n) abgeschlossen werden.

Das Modul ist auf 16 Teilnehmer begrenzt.

The lab capacity limits this course to max. 16 participants.

Contact persons:

Gábor Gévay, Prof. Dr. Volker Markl