What is this package about?

This elective package is for students in the bachelor program Data Science and other programs (except the bachelor program Computer Science), who are interested in the master program Data Science in Engineering (DSiE). The courses in this package are required prior knowledge for enrollment in this master program. Students are only admissible in the master program DSiE if they have successfully completed both the packages (1) Data modeling foundations and (2) Computer Science for Data Science.

Course descriptions

Data management for data analytics

As we enter an era of big data and data science, core knowledge and skills in data modeling and data management are now recognized as essential in many disciplines. The primary goal of JBI050 is to master the core best-practices of data management systems, applied towards using contemporary tools to support effective data analytics. In particular, this course focuses on preparing students to meet contemporary data modeling and data management challenges which arise in applications in their own fields of study.

Linear algebra and applications

This course offers a wide range of interesting linear algebra techniques and very nice applications, including Big Data applications. These techniques are almost indispensable for the current Big Data era. You will learn how to solve linear systems, arguably the most common and important scientific problem. Least squares methods can be used to determine an approximate line (or polynomial, or spline) through a number of points (used in Computer Graphics). Rotations and reflections are useful for Computer Graphics and Robotics. Angles between vectors can be used to compare tastes of movies, music, or books, and to predict
such tastes for the future. You will learn the idea behind Google PageRank and its connection with eigenvalues, how it is computed, and you can also apply similar techniques to finally determine which of Ajax, Feyenoord, or PSV is the best team. We can use text mining methods to determine the most important tweets or keywords from a set of tweets; useful for effective communication. The course is designed to be interesting for many students, including students from Computer Science, Data Science, and Robotics. The subjects are fascinating on their own, but also form an ideal preparation for subsequent courses such as Data Mining and Machine Learning and Computer Graphics. The course offers a nice combination of theoretical (math) and practical (implementing) work, partly done in groups. You can use a language you already know (for instance Python or Matlab) or learn a new language.

Data intensive systems & applications

The database and data engineering market is going through unprecedented shifts and challenges demanding new approaches for processing heterogeneous data at massive scale. As a result, a new breed of data-intensive systems which aims to handle smart data-driven applications is emerging. This course prepares students to meet the new challenges of contemporary data engineering in which traditional assumptions break, where new data models, query languages and programming interfaces are required.