JACOBS UNIVERSITY

Study Program Handbook

Data Engineering

Master of Science

Valid for all students who start their studies in fall 2017
Date: August 18, 2017
Program Coordinator: Dr. Mathias Bode
http://www.jacobs-university.de/data-engineering
dataengineering@jacobs-university.de

Table of Contents

1. Program Overview

1.1. Concept

Today we are "drowning in data and starving for information" while acknowledging that "data is the new gold". However, deriving value from all the data now available requires a transformation in data analysis, in how we see, maintain, share and understand data. Data Engineering is an emerging profession concerned with the task of acquiring large collections of data and extracting insight from them. It is driving the next generation of technological innovation and scientific discovery, as this will be data-driven.

The graduate program in Data Engineering offers a fascinating and profound insight into the methods and technologies of this rapidly growing area. The program is more broadly conceived than competing programs: it combines big data aspects of "Data Analytics" and "Data Science" with the technological aspects of data acquisition, curation and management. Thus the program provides the essentials for paving a successful career: computer skills and mathematical understanding paired with practical experience in selected application fields.

The program is embedded into the "Mobility" focus area at Jacobs University. This focus area investigates the mobility of people, goods and information. Even though the Data Engineering program is centered in "Mobility", it includes contributions from and supports applications in the two other research foci: Health (bioactive substances), and Diversity (in modern societies).

The graduate program in Data Engineering is tailored to a diverse student body (see also Section 1.3) with a wide variety of interests, academic backgrounds, and previous experiences. Small group sizes, low teacher-to-student ratios, and personalized supervision/advising make it possible to keep the graduation requirements flexible so that the program can cater to the 21-year-old full-time student as well as the long-time employee of a data-intensive company who wants to keep up with current data engineering practices.

1.2. Qualification Aims

The Data Engineering program aims to provide an in-depth understanding of the essential aspects of data based decision making and the skills required to apply and implement these powerful methods in a successful and responsible manner.

Apart from the necessary programming skills, this comprises:

- Methods of data acquisition both from the internet and from sensors

- Methods to efficiently store and access data in large and distributed data bases
- Statistical model building including a wide range of data mining methods, signal processing, and machine learning techniques
- Visualization of relevant information
- Construction and use of confidence intervals, hypothesis testing, and sensitivity analyses
- The legal foundations of Data Engineering

1.3. Target Audience

The target audience of the Data Engineering graduate program comprises students who have completed their BSc in areas like computer science, physics, applied mathematics, statistics, electrical engineering, communications engineering or related disciplines and who want to deepen their knowledge and proceed to research oriented work towards a master's or ultimately a PhD degree. Typical examples are:

- A computer science BSc who wants to acquire skills in data analysis and micro/macro-economics for a career in computational finances;
- A business major with a solid statistics and analysis foundation but without any programming experience;
- A geology student who wants to become a data scientist and needs to obtain the mathematical and statistical foundations;
- A student with a bachelor's or master's degree in one of the natural sciences who wishes to boost his/her career in empirical research or industrial research and development, where professional handling of very large-scale data collections has become a prime bottleneck for success;
- A graduate in mathematics or theoretical physics who wants to capitalize on his/her theoretical knowledge of modeling methods by learning about the hands-on side of data analysis, interesting fields for applications, and options for employment.

1.4. Career Options

Demand for Data Engineers is massive. Typical fields of work encompass the finance sector, the automotive and health industry as well as retail and telecommunications. Companies and institutions in almost every domain need:

- **experts for data acquisition** who find out how to collect the data needed;
- **experts for data management** who know how to make the optimum use of the obtained data;
- **experts for data analysis** who evaluate and interpret the collected data correctly and are able to visualize the findings clearly.

Graduates of the program work as data analysts, data managers, data architects, business consultants, software and web developers, or system administrators.

An MSc degree in Data Engineering also allows students to move on to a PhD and a career in academia and research institutions.

1.5. Admission Requirements

In general, applicants need to submit the following documents in order to be considered for admission:

- Letter of motivation
- Curriculum vitae (CV)
- University transcript in English or German
- Bachelor's degree certificate
- Two letters of recommendation
- English language proficiency test with a minimum score of 90 (TOEFL) or 6.5 (IELTS). Alternatively, students may submit a confirmation from their previous university that their education was conducted in English.

Please visit http://www.jacobs-university.de/study/graduate/application-information for more details on the application process.

2. The Curriculum

2.1. The Curriculum at a Glance

The Data Engineering graduate program is composed of a mixture of foundational lectures, specialized courses, industry seminars and applied project work, leading to a master's thesis that will usually be conducted in close collaboration with an industry partner or even at a company site. The program takes four semesters (two years). The following table shows an overview of the modular structure of the program. All credit points (CP) are ECTS (European Credit Transfer System) credit points.

<table>
<tr><td>Semester 4</td><td colspan="5">RESEARCH-MT
Master's Thesis (30 CP)</td></tr>
<tr><td>Semester 3</td><td rowspan="3">CORE-ERC
Electives and Remedial Courses (15 CP)</td><td rowspan="2">CORE-AMA
Advanced Methods and Applications (20 CP)</td><td rowspan="2">RESEARCH-IRP
Industry and Research Projects (20 CP)</td><td></td><td rowspan="3">CAREER-SL
Skills and Languages (15 CP)</td></tr>
<tr><td>Semester 2</td><td rowspan="2">CORE-FDE
Foundations of Data Engineering (20 CP)</td></tr>
<tr><td>Semester 1</td><td></td><td></td></tr>
</table>

In order to graduate, students need to obtain 120 ECTS credit points. The ECTS (European Credit Transfer System) is a system defining the student workload required to achieve the objectives of a study program. At Jacobs University, 1 credit point is equivalent to 25 hours of student workload. In each module, students must obtain a minimum amount of credit points (with an average of 30 credit points per semester).

Each module consists of mandatory and/or elective components as outlined below. Detailed course descriptions are available in the course catalogue (https://campusnet.jacobs-university.de).

2.2. Modules

2.2.1. CORE-FDE - Foundations of Data Engineering

Amount of credit points to be obtained in this module: 20

This module covers general methods of data engineering and constitutes the foundations for further, more advanced instruction and applied projects by introducing the fundamental concepts, methods and technologies used in data engineering. The module consists of intensive courses accompanied by hands-on tutorials and labs in which they are applied and consolidated.

Course Title	Course No.	Semester	Mandatory	Credits
The Big Data Challenge: Topics, Applications, Perspectives	340111	1	yes	5
Data Analytics	340131	1	yes	5
Big Data Bases and Cloud Services	340151	2	no	5
Principles of Statistical Modeling	340101	2	no	5
Web Analytics	340161	1/3	no	5

2.2.2. CORE-ERC - Electives and Remedial Courses

Amount of credit points to be obtained in this module: 15

The Data Engineering graduate program attracts students with diverse career goals, backgrounds, and prior work experience (see chapter 1.3). This module allows students to either make up for missing academic prerequisites or to strengthen their knowledge in application areas of data engineering by taking dedicated preparatory lectures, courses from the undergraduate curriculum or other graduate programs at Jacobs University. All third-year courses in the undergraduate majors in the mobility focus are directly admissible as electives and remedial courses. All other courses require the approval by the program coordinator. Please see CampusNet (https://campusnet.jacobs-university.de) for current course offerings. The list below gives a number of examples:

Course Title – Preparatory Lectures	Course No.	Semester	Mandatory	Credits
Calculus and Linear Algebra for Graduate Students	340181	1	no	5
Data Management for Graduate Students	340163	1	no	5
Probabilities for Graduate Students	340171	1	no	5
Data Bases and Web Services	CO19-320302	1/3	no	5
Geospatial Data Analysis	210222	1/3	no	2.5
Probability and Random Signal Processing	CO26-300321	1/3	no	5
Econometrics	JTME-990222	1/3	no	2.5
Earth System Monitoring and Remote Sensing	CO12-210213	2	no	2.5
Process Modelling and Simulation	CO29-050212	2	no	5

To enhance flexibility, the credits in this module can also be earned in part or in total by a qualifying internship in an external research organization or a company. 5 credit points are awarded for a full-time (40 hours/week) internship of 1 month.

The study plan for this module is fixed by the student individually in close consultation with his/her academic advisor. It has to be approved by the program coordinator of the Data Engineering graduate program.

Generally, the focus in this module will be on remedial courses in the first two semesters – courses may only be given in the spring or fall semesters – and on electives in the second/third semester. The distribution between these categories will be decided on a case-by-case basis.

2.2.3. CORE-AMA - Advanced Methods and Applications

Amount of credit points to be obtained in this module: 20

This module is the centerpiece of the Data Engineering graduate program. Building on the CORE-FDE module and first remedial courses, it introduces advanced concepts, methods and technologies of data engineering with a view towards industrial applications. The components of this module range from full, semester-long courses for general methods, via short block seminars that introduce special techniques and applications, to application-driven, problem-solving workshops in which industry partners come together with data engineering students and faculty to study and solve real-world problems. Courses in this module may, for example, include:

Course Title	Course No.	Semester	Mandatory	Credits
Semantic Web and Internet of Things	340152	2 - 3	no	5
Document Analysis	340241	2 - 3	no	5
Data Acquisition Technologies and Sensor Networks	340112	2 - 3	no	5
Big Data Management	340221	2 - 3	no	5
Data Visualization and Image Processing	340231	2 - 3	no	5
Statistical Modeling and Predictive Analytics	340172	2 - 3	no	5
Internet Security and Privacy	340251	2 - 3	no	5
Modern Signal Processing	340153	2 - 3	no	5
Data Mining	tba	2 – 3	no	5

Data Compression, Compressed Sensing, and Modern Coding	340143	2 – 3	no	5
Biomedical Signal Processing	340113	2 - 3	no	2.5

Please check the course catalogue (https://campusnet.jacobs-university.de) for up-to-date course offerings.

Courses on other topics can be offered depending on availability of teaching resources and student interest. Topics of special block courses may include, for example:
a) Numerics, b) Parallel Computing c) R course, d) SAP Software e) Computational Statistics f) Hadoop, g) Novel Data Bases for Big Data, h) Semantic Web, RDF, Linked Open Data, i) Design Thinking.

On individual request, and upon approval by the program coordinator, credit points in this module may also be obtained by taking suitable courses from other graduate programs or via independent study courses. Independent study courses are tailor-made courses comprising the in-depth study of specific subjects or guided projects. Projects for independent study courses include the preparation of academic publications and the preparation of presentations for academic conferences. Independent study courses need to be approved by the course instructor, the program coordinator and the academic advisor of the student. Each independent study course has a syllabus outlining the structure of the course and its learning outcomes. The amount of credit points to be obtained in an independent study course is based on the workload of the student (1 credit point is equivalent to 25 hours of workload). Data Engineering graduate students may obtain a maximum amount of 10 credit points via independent study courses.

Additional credit points in this module may be obtained through transfer credits. Transfer credits may be awarded for coursework completed at other academic institutions, e.g. at summer schools. See § 4.12 of the Policies for Master Studies for details (http://www.jacobs-university.de/academic-policies).

2.2.4. CAREER-SL - Skills and Languages

Amount of credit points to be obtained in this module: 15

In this module students acquire skills preparing them for a career as data engineers in industry. Apart from mandatory courses on “Ethics in Science and Technology” and “Legal Foundations of Data Engineering”, students take language courses and other courses from the “Jacobs Track”, a special offer of Jacobs University fostering a broad range of career competencies. Please see http://language-program.user.jacobs-university.de and https://campusnet.jacobs-university.de for current course offerings and more information.

Course Title	Course No.	Semester	Mandatory	Credits
Ethics in Science and Technology	JTSC-040121	1 - 3	yes	2.5
Legal Foundations of Data Engineering	340121	1/3 intersession	yes	2.5
Jacobs Track courses including language courses I	depending on course	1 - 3	no	2.5
Jacobs Track courses including language courses II	depending on course	1 - 3	no	2.5
Jacobs Track courses including language courses III	depending on course	1 - 3	no	2.5
Jacobs Track courses including language courses IV	depending on course	1 - 3	no	2.5

2.2.5. RESEARCH-IRP - Industry and Research Projects

Amount of credit points to be obtained in this module: 20

This module features two advanced projects in data engineering in semesters 2 and 3, each worth 10 credit points. These projects are supervised in close collaboration between Jacobs University faculty and partner companies. Projects in the research groups at Jacobs University are also admissible. These projects are guided research or development projects that result in a report with a target size of ca. 20 pages and a presentation to the Data Engineering program members. Both components contribute to the grade of the project.

Title	Course No.	Semester	Mandatory	Credits
Data Engineering Advanced Project I	340182	2	yes	10
Data Engineering Advanced Project II	340282	3	yes	10

2.2.6. RESEARCH-MT - Master's Thesis

Amount of credit points to be obtained in this module: 30

In the fourth semester students conduct research and write a master's thesis guided and supported by their academic supervisor. The thesis has a target size of ca. 60 pages and presents the research of the student. It will be jointly judged by a thesis committee which consists of the thesis supervisor and at least one other member. The other member(s) can be Jacobs University faculty members or external members. The

thesis will be graded using the Jacobs University grading system ranging from 1 (excellent) to 5 (fail).

Title	Course No.	Semester	Mandatory	Credits
Master's Thesis	no number	4	yes	30

3. Data Engineering Graduate Program Regulations

3.1. Scope of these Regulations

The regulations in this handbook are valid for all students who entered the Data Engineering graduate program at Jacobs University in fall 2017. The general Policies for Master Studies at Jacobs University apply to this program. In cases of conflict between the program regulations and the general policies, the general policies apply (see: http://www.jacobs-university.de/academic-policies).

3.2. Degree

Upon successful completion of the program, students are awarded a Master of Science (MSc) degree in Data Engineering.

3.3. Graduation Requirements

In order to graduate, students need to obtain 120 credit points. In addition, the following graduation requirements apply:

- In each module, students need to obtain a minimum amount of credit points as indicated in chapter 2 of this handbook.
- Students need to complete all mandatory components of the program as indicated in chapter 2 of this handbook.

Jacobs University Bremen reserves the right to substitute courses by replacements and /or reduce the number of mandatory/mandatory elective courses offered.