JACOBS UNIVERSITY

Data Engineering (MSc)
Graduate Program Handbook

Version: Fall 2015
Date: October 26, 2015
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 Jacobs University 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 graduate program in Data Engineering 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 will be tightly integrated with the accompanying executive program "Jacobs Academy: Data Engineering" (JA:DE) which is currently under development. The courses in the central module "CORE-AMA - Advanced Methods and Applications" (see section 2.2.3) will be shared with the JA:DE program. This puts master's students in direct contact with the industry in the and facilitates interaction, networking and recruiting, boosting student's employability and career options.

The graduate program in Data Engineering caters 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, 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 computer science 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 a 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 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 – in industry, commerce and the public sector. From IT to finance, from automotive to oil and gas, from health to retail: companies

and institutions in almost every domain need experts for data acquisition, data management and data analysis. Likewise, an MSc degree in Data Engineering allows students to move on to a PhD and to a career in academia and research institutions.

1.5. Admission Requirements

Applicants need to submit the following documents in order to be considered for admission:

- Essay (letter of motivation)
- Curriculum vitae (CV)
- University transcript in English or German
- Bachelor's degree certificate or equivalent
- 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) credits. Students acquire 30 credit points per semester on average.

<table>
<tr><td>Semester 4</td><td colspan="4">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>CORE-AMA
Advanced Methods and Applications
(20 CP)</td><td>RESEARCH-IRP
Industry and Research Projects
(20 CP)</td><td rowspan="3">CAREER-SL
Skills and Languages
(15 CP)</td></tr>
<tr><td>Semester 2</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>Semester 1</td></tr>
<tr><td colspan="3">CORE-FDE
Foundations of Data Engineering (20 CP)</td></tr>
</table>

In order to graduate, students need to obtain 120 credit points. In each module, students need to obtain a minimum amount of credit points:

- Module CORE-FDE - Foundations of Data Engineering: 20 credit points
- Module CORE-ERC - Electives and Remedial Courses: 15 credit points
- Module CORE-AMA - Advanced Methods and Applications: 20 credit points
- Module CAREER-SL - Skills and Languages: 15 credit points
- Module RESEARCH-IRP - Industry and Research Projects: 20 credit points
- Module RESEARCH-MT - Master's Thesis: 30 credit points

Each module consists of mandatory and/or elective components as outlined below. Detailed course descriptions are available in the course catalogue on CampusNet (http://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 lays 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 that introduce concepts, methods and technologies. The courses are accompanied by hands-on tutorials and labs in which those are applied and consolidated.

Course Title	Course No.	Semester	Mandatory	Credits
The Big Data Challenge: Topics, Applications, Perspectives	340111	1	yes	5
Big Data Bases and Cloud Services	340151	tba	yes	5
Principles of Statistical Modeling	340101	1	yes	5
Data Analytics	340131	1	yes	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 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 (http://campusnet.jacobs-university.de) for current course offerings.

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 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 semesters. 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 modules 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.

Full courses in this module may, for example, include (please see http://campusnet.jacobs-university.de for up-to-date course offerings):

Course Title	Course No.	Semester	Mandatory	Credits
Semantic Web and Internet of Things	tba	2 or 3	no	5
Document Analysis	tba	2 or 3	no	5
Machine Learning	tba	2 or 3	no	5
Data Acquisition Technologies	tba	2 or 3	no	5
Big Data Management	tba	2 or 3	no	5
Data Visualization and Image Processing	tba	2 or 3	no	5
Statistical Modeling and Predictive Analytics	tba	2 or 3	no	5
Internet Security and Privacy	tba	2 or 3	no	5

Courses on other topics can be offered depending on the 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.

On individual request, credits points in this module may also be obtained 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 needs 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 Academic 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 courses on Data Ethics and the Legal Foundations of Data Engineering, students take language courses and other courses from the "Jacobs Track". Please see http://language-program.user.jacobs-university.de and http://campusnet.jacobs-university.de for current course offerings and more information.

Course Title	Course No.	Semester	Mandatory	Credits
Data Ethics	tba	tba	yes	2.5
Legal Foundations of Data Engineering	tba	tba	yes	2.5
German language courses or other language courses	tba	tba	no	2.5 each
Further courses from the "Jacobs Track"	tba	tba	no	2.5 each

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
Advanced Project 1	tba	2	yes	10
Advanced Project 2	tba	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	none	4	yes	30

3. Data Engineering Graduate Program Policies

3.1. Scope of these Policies

The policies in this handbook are valid for all students who entered the Data Engineering graduate program at Jacobs University in fall 2015. Besides the Data Engineering Graduate Program Policies, the general Policies for Academic Master Studies at Jacobs University (see http://www.jacobs-university.de/academic-policies) apply to this program. In cases of conflict between the program policies and the general policies, the general policies apply.

3.2. Degree

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

3.3. Graduation Requirements

In order to graduate, students need to obtain at least 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.

3.4. Application for PhD Program

Students who want to continue their education with a PhD and find a supervisor who also supplies PhD funding may apply after their third semester. If successful, a PhD committee is formed immediately, and students write an extended master's thesis, that – in addition to preliminary results – also contains a well-developed work plan for PhD research so that it can serve as a research proposal for the PhD program in Data Engineering (see the handbook of the PhD program in Data Engineering).