Welcome to the Mannheim Master in Data Science

Introductory Session

HWS 2018
Agenda

1. Motivation and Goals of the Program
2. Program Structure
   - Fundamentals
   - Data Management
   - Data Analytics
   - Projects & Seminars
   - Master’s Thesis
3. The people you will deal with
4. Scholarships
5. MMDS Communication Channels
6. Your Questions
Motivation of the Program: The Data Deluge

More and more data is generated.
“We are Drowning in Data...”

Sloan Digital Sky Survey
≈ 200 GB/day
≈ 73 TB/year
≈ 12 Wikipedias/year

Analyze
• Type of sky object:
  Star or galaxy?
“We are Drowning in Data...”

Hundreds of biomedical data sources available; growing rapidly!
“We are Drowning in Data...”

US Library of Congress
≈ 235 TB archived
≈ 40 Wikipedias

Analyze
• Topic distributions
• Citation networks
• Historic trends*.  

---

Facebook
≈ 12 TB/day added
(as of Mar. 2010)
≈ 2 Wikipedias/day
≈ 782 Wikipedias/year

Analyze
• Current interests and behavior of over one billion people.

“We are Drowning in Data...”
“We are Drowning in Data...”

Google
\[ \approx 20 \text{ PB/day processed} \]
\[ (\text{Jan. 2010}) \]
\[ \approx 3,389 \text{ Wikipedias/day} \]
\[ \approx 7,300,000 \text{ Wikipedias/year} \]

Analyze
• Browsing behavior and interests of users.
“We are Drowning in Data...”

Analyze
• Current behavior and interests of mankind.
“We are Drowning in Data...”

NSA
Unknown amount of communication data from all over the world.

Analyze
• Identify suspects and terrorists.
“We are Drowning in Data...”

← Amount of data that is produced.

← Amount of data that can be looked at by humans.
Data Science in a Nutshell

- Take advantage of the data deluge.

- Challenges
  - Data collection
  - Data integration
  - Storage
  - Processing
  - Analysis
  - Validity of results
  - Communication of results
  - Application of results
Becoming a „Data Scientist“

- Competencies
  - Mathematics
  - Computer Science
  - Statistics

- Job Areas
  - Business (Recommender systems, market research, pricing, …)
  - Industry 4.0 (Production planning, predictive maintenance, early warning systems, …)
  - Research (Social Sciences, Computer Science, Digital Humanities, …)
  - Administration (Traffic planning, intelligent energy networks, …)
  - Medical Science (Research studies, personalized medication, …)
LinkedIn analyzed the skills of members who started new jobs or received interest from recruiters in 2016.

### The Top Skills of 2016 on LinkedIn Germany

<table>
<thead>
<tr>
<th>Rank</th>
<th>Skill</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cloud and Distributed Computing</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Statistical Analysis and Data Mining</td>
<td>+2</td>
</tr>
<tr>
<td>3</td>
<td>SEO/SEM Marketing</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Marketing Campaign Management</td>
<td>NR</td>
</tr>
<tr>
<td>5</td>
<td>Data Engineering and Data Warehousing</td>
<td>+4</td>
</tr>
<tr>
<td>6</td>
<td>Database Management and Software</td>
<td>+1</td>
</tr>
<tr>
<td>7</td>
<td>Software QA and User Testing</td>
<td>-2</td>
</tr>
<tr>
<td>8</td>
<td>Retail Store Operations</td>
<td>+2</td>
</tr>
<tr>
<td>9</td>
<td>Electronic and Electrical Engineering</td>
<td>-7</td>
</tr>
<tr>
<td>10</td>
<td>Channel Marketing</td>
<td>-2</td>
</tr>
</tbody>
</table>

Mannheim Master in Data Science

- Interdisciplinary
  - Computer Science
  - Mathematics and Statistics
  - Social Science

- Research-Oriented

- Applied
  - Hands-on experience
  - Close cooperation with industry partners
Programme Structure

- The programme consists of different blocks:
  - Fundamentals (0-14 ECTS)
  - Data Management (24-36 ECTS)
  - Data Analytics (30-54 ECTS)
  - Projects & Seminars (12-16 ECTS)
  - Master‘s Thesis (30 ECTS)
Fundamentals

- The MMDS is an interdisciplinary study programme
  - Hence, most of you have different backgrounds: you are computer scientists, social scientists, mathematicians, …
- Fundamentals help you catch up with each other

- Offerings:
  - Data acquisition
  - Multivariate analysis
  - Database technology
  - Programming
Fundamentals

- You decide which fundamentals you need!
- You can visit up to two fundamental courses
- Take a look at descriptions of fundamentals in the module catalogue
  - Do the contents sound familiar or unfamiliar?
- Take a look at the requirements of other courses in the module catalogue
  - Do they sound familiar or unfamiliar?
Data Management (24-36 ECTS)

- Storing, processing, integrating, and searching for data
  - Formats, standards, and methodologies

- Lecture offerings include
  - Large Scale Data Management
  - Information Retrieval and Web Search
  - Query Optimization
  - Semantic Web Technologies
  - Model-driven Development
  - Algorithmics
  - Web Data Integration
  - Record Linkage
  - Data Security
Data Analytics (30-54 ECTS)

- Analysis, subsequent processing, and visualization
  - Methods and applications
- Lecture offerings include
  - Data Mining I + II
  - Text Analytics
  - Hot Topics in Machine Learning
  - Mathematical Visualization
  - Cross Sectional Data Analysis
  - Longitudinal Data Analysis
  - Baysian Statistics
  - Lineare Optimierung
  - Decision Support
  - Big Data and Social Sciences
Seminars and Individual Projects

- Seminar (4 ECTS)
  - Mostly theoretical work (e.g., literature reviews on certain topics)
  - Offered every semester by various chairs
  - Watch for the announcements of different chairs

- Individual Projects (8 ECTS)
  - Practical work on a data science problem
  - Approach chairs directly for topics
Team Projects (12 ECTS)

- Work on a data science project as a team
  - Acquire team organization and management skills on top!
- Run over one or two semesters
- Topics are announced every semester in a special session
  - This semester: 07.09., starting at 10:15
- Registration via portal2 after the announcement
Industry Partners

- We are constantly extending our network of industry partners
  - Joint projects
  - Internships (can be combined with individual project)
  - Master’s theses
Programme Structure

1st Semester
- Fundamentals
  - Up to two out of
    - Data Acquisition
    - Database Technology
    - Multivariate Analysis
    - Programming

- Data Management
  - e.g., one out of
    - Algorithms
    - Database Systems II
    - Data Security
    - Information Retrieval and Web Search

- Data Analytics
  - e.g., two out of
    - Advanced Quantitative Methods
    - Computational Finance
    - Data Mining and Matrices
    - Data Mining II
    - Web Mining
    - Non-linear Optimization
    - Longitudinal Data Analysis

2nd Semester
- Research
  - - Team or Individual Project or Seminar

3rd Semester
- Data Management
  - e.g., one or two out of
    - Query Optimization
    - Advanced Software Engineering
    - Semantic Web Technologies

- Data Analytics
  - e.g., one or two out of
    - Algebraic Statistics
    - Hot Topics in Machine Learning
    - Mathematics and Information
    - Mathematical Visualization

4th Semester
- Master’s Thesis

Research
- - Team or Individual Project
- - Seminar

(approx. 30 ECTS / Semester)
Module Catalogue

- See online at http://www.fmi.uni-mannheim.de/de/studium/modulkatalog-uebersicht/
Module Overview

- See online at http://mmds.wim.uni-mannheim.de/

<table>
<thead>
<tr>
<th>Fundamentals (0 – 14 ECTS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
</tr>
<tr>
<td>Data Acquisition</td>
</tr>
<tr>
<td>Database Technology</td>
</tr>
<tr>
<td>Multivariate Analysis</td>
</tr>
<tr>
<td>Programming Course</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Management (24 – 36 ECTS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
</tr>
<tr>
<td>Advanced Software Engineering</td>
</tr>
<tr>
<td>Algorithmik</td>
</tr>
<tr>
<td>Anfrageoptimierung</td>
</tr>
<tr>
<td>Data Security</td>
</tr>
<tr>
<td>Datenbanksysteme II</td>
</tr>
<tr>
<td>Information Retrieval and Web Search</td>
</tr>
<tr>
<td>Large Scale Data Management</td>
</tr>
<tr>
<td>Model-driven Development</td>
</tr>
<tr>
<td>Semantic Web Technologies</td>
</tr>
<tr>
<td>Web Data Integration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Analytics (30 – 54 ECTS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
</tr>
<tr>
<td>Advanced Quantitative Methods</td>
</tr>
<tr>
<td>Algebraische Statistik</td>
</tr>
<tr>
<td>Computational Finance</td>
</tr>
<tr>
<td>Cross sectional Data Analysis</td>
</tr>
<tr>
<td>Data Mining I</td>
</tr>
</tbody>
</table>
Detailed Information about Courses

- You will find additional information about the courses on the chairs’ web sites

- **Information session** about master courses and chairs: Monday, 3rd September, 3:30 pm, B6 A1.01

- **Course registration**
  - Central registration for many courses (closed Wednesday)
  - If a course is booked, write an e-mail to instructor
    - Sometimes, there are leftover places
People you will deal with: Professors and Lecturers

Institute of Computer Science and Business Informatics
- Frederik Armknecht
- Colin Atkinson
- Christian Bizer
- Rainer Gemulla
- Goran Glavaš
- Margret Keuper
- Matthias Krause
- Guido Moerkotte

Institute of Mathematics
- Heiko Paulheim
- Simone Ponzetto
- Ursula Rost
- Heiner Stuckenschmidt
- Simone Göttlich
- Andreas Neuenkirch
- Daniel Roggenkamp
- Wolfgang K. Seiler

Department of Politics
- Sean Carey
- Thomas Gschwend

Department of Sociology
- Thomas Gautschi
- Florian Keusch
- Frauke Kreuter
People you will deal with: MMDS Program Management

Program Director
Prof. Dr. Heiko Paulheim
B6, 26, Room B1.16
heiko@informatik.uni-mannheim.de

Program Manager
Lisa Wessa, M.A.
B6, 26, Room B 1.04
wessa@wim.uni-mannheim.de
People you will deal with: Student Advisors

Academic Advisor
Samuel Broscheit
A5, 6, Room B2.27
studienberatung@wim.uni-mannheim.de

Study Coach
Birgit Czanderle, M.A.
A5, 6, Room B 2.20
czanderle@wim.uni-mannheim.de
Scholarships and Funding Opportunities

- Various funding opportunities are available
  - e.g., Deutschlandstipendium (10 scholarships only for MMDS students 2018/2019)
  - see https://www.uni-mannheim.de/stipendium/english/
MMDS Mailing List

- You will be added to the mmds emailing list.

- Please note: All email correspondence from now on will go to your Uni Mannheim email account!

→ Check your university mail account regularly!
Students’ Communication Channels

- There are also communication channels run by your fellow students
- Slack Channel
  - currently 29 students
  - mmds-students.slack.com
  - Use Firstname_Lastname as displayed username
- WhatsApp group
  - currently 36 students
  - ask your fellow students to get in
- Contact person
  - Jan Wagner (3rd semester student)
  janwwagn@mail.uni-mannheim.de
And now for Something Completely Different

- The Data and Web Science Group hosts a BBQ at the beginning of each semester
- Drinks and grill are provided, please bring your own food
- Thursday, September 6th, 6pm, basement of B6 26
Questions?