

# Master Projects

## Winter Term 2018/2019

Friday 19<sup>th</sup> of October 2018

Albert-Ludwigs-Universität Freiburg

Prof. Dr. Georg Lausen  
Anas Alzogbi  
Victor Anthony Arrascue Ayala  
Michael Färber  
Io Taxidou



**UNI  
FREIBURG**

# Agenda



- Organization
  - Regulations
  - Prerequisites
  - Goals
  - Project management & Supervision
  - Infrastructure
  - Assessment
  
- Topics



- Study regulations (Studienordnung)
  - 16 ECTS → 480 hours (17-1 weeks WS1819 ⇒ 30 hours/week  $\cong$  6 hours/day from Mo. to Fr.)
  - Curriculum: Master of Science, 3rd Semester
  
- Apply via HisInOne
  1. Master Project Veranstaltungsart: Projekt
  2. After Kick-off meeting: choose a topic, set priority → group
  3. At the end: Masterprojekt - Prüfungsleistung (Prüfung)

# Prerequisites



- **Finish before end of March 2019**
- Each project has preconditions
  - Background knowledge, ex: Spark, Machine Learning, Semantic Web, Recommender Systems
  - Programming languages: Java, Scala, Python
  - **Compulsory task (optional)**
- Highly recommended:
  - Data Analysis and Query Language
  - Advanced Databases and Information Systems
  - Lab: Distributed Computing Using Spark

# Goals – Perspectives



- CS Program
  - Preparation for the Master Thesis
  - Gain experience in research and development methods
- Professional profile
  - Data Science, Data Analyst, Machine Learning specialist
  - Application areas: Semantic Web, Recommender systems
- Gain soft skills
  - Collaborative work on a project
  - Improve individual programming skills
  - Management of mid-sized projects

# Project management & Supervision



- Team size: ~2 students (depends on project)
  
- Supervision model depends on supervisor, typically
  - Regular meetings every 2 weeks (depends on the needs)
  - At supervisor's office or Room 051-01-029



- VMs running at physical servers (ex: lima, hobbes)
  - Up to 100 GB of RAM
- 2 Cloudera Clusters
  - Student cluster, 10 Machines
  - Worker cluster, if required
- Private github repository
  - Alternatively SVN repository
- Datasets

# Assessment



- Source code
- Project report: ~15-20 pages for each group member
- Final presentation: 25 min



# Assessment criteria (tangible)



- The scope and difficulty of the work / implementation
- Individual contribution
- Autonomous work
- Quality of code (design, documentation, tests)
- Quality of written work (project report)
- Impact of the results
- Quality of presentation

# Assessment criteria (less tangible)



- Workload of every student must be clearly distinguishable
- Team performance: a successful project has a positive effect
- Role and participation in the team (coordination, etc.)
- Presentations (intermediate presentations)

# Topics & Supervisors



- T1: Leveraging Embeddings for Cross-Domain Recommendations (**Arrascue**)
- T2: Extending P<sub>Ro</sub>ST (**Arrascue**)
- T3: Error prediction in automotive industry (**Alzogbi**)
- T4: Trust and Distrust Assessment in Online User Activity (**Taxidou**)
- T5: News Bias Detection (**Färber**)
- T6: How Fair are Researchers? Analyzing Biases of Paper Citations (**Färber**)
- T7: A Platform for Meta-Learning of Entity Linking (**Färber**)



*Any questions about organization?*