Overview

1. Find out why software engineering is important
2. Get acquainted with
   - the Chair of Software Engineering
   - the research
   - the people
   - the teaching

Code Search Evaluation Framework (CSEF)

Master Team Project FSS 2017

Marcus Kessel (kessel@informatik.uni-mannheim.de)

Chair of Software Engineering
Context – Merobase Code Search

- **Goal**
  - Software Code Reuse

- **merobase**
  - Keyword search
  - Interface-driven code search
  - Test-driven search

- **SOCORA**
  - Ranking engine for candidate search results
    - Functional/Non-functional properties
  - Ranking criteria
    - Static analysis / dynamic analysis
    - Metrics
The Challenge – Evaluate Code Search

Source Code Repositories

1. Fetch
2. Analyze
3. Index

Crawler

Source Code Analyzer

Software Component Index (Solr)

How to evaluate results/rankings?

Millions of components – „Blackbox“
**Code Search Evaluation Framework**

**Exploration & Characterization of Component Index**
- Faceting, clustering, collapsing (+ other suitable data mining techniques)
- Visualization options (e.g., Voronoi maps)
- ...

**Analysis / Knowledge Discovery**
- Construction of realistic code search scenarios
  - Component index, internet platforms, questionnaires ..
- Linking (establishing relationships)
  - Object-orientation, clones etc.
- Ranking Criteria – New Insights from Code Analysis?

**Corpus curation / Evaluation**
- Human judgment (“ground truth”)
- Enable (ranking) evaluations, comparisons etc.
Conditions

- Duration
  - 12 months

- Required skills
  - Programming (Java)

- Participants:
  - 4-8 students

- Language:
  - English

- MMDS (Mannheim Master in Data Science) – Suitable?
  - Yes!

- Note: Requirements, goals and timetable defined by agreement with the supervisors
Supervisors

- **Main Supervisor**
  - Marcus Kessel (Lehrstuhl Softwaretechnik)
  - kessel@informatik.uni-mannheim.de

- **Professor (Lehrstuhl Softwaretechnik)**
  - Colin Atkinson
  - atkinson@informatik.uni-mannheim.de
Thanks for your attention!