Hallo

- **Prof. Dr. Christian Bizer**
- Professor for Information Systems V
- Research Interests:
  - Web Data Integration
  - Data and Web Mining
  - Data Web Technologies
- Room: B6 - B1.15
- eMail: chris@informatik.uni-mannheim.de
- Consultation: Wednesday, 13:30-14:30
Hallo

- Anna Primpeli
- Graduate Research Associate
- Research Interests:
  - Data Extraction
  - Web Data Integration
  - Active Learning
  - Structured Data on the Web
- Room: B6, 26, C 1.04
- eMail: anna@informatik.uni-mannheim.de
Hallo

- **Ralph Peeters**
- Graduate Research Associate
- Research Interests:
  - Entity Matching using Deep Learning
  - Product Data Integration
- Room: B6, 26, C 1.04
- eMail: ralph@informatik.uni-mannheim.de
Hallo

- **M. Sc. Wi-Inf. Alexander Brinkmann**
- Graduate Research Associate
- Research Interests:
  - Data Search using Deep Learning
  - Product Data Categorization
- Room: B6, 26, C 1.03
- eMail: alex.brinkmann@informatik.uni-mannheim.de
Agenda of Today’s Kickoff Meeting

1. Seminar organization
2. Seminar topics
3. How to structure your seminar paper / presentation?
4. Questions and guidance
1. Organization
Learning Targets

- Writing a seminar thesis as an exercise for your master thesis
- Understanding and presenting state-of-the-art scientific work
- Searching and citing scientific papers / journal articles
- How to structure your thesis and presentation
- How to argue, how to explain, how to write!
- How to write a nicely formatted paper using LaTeX
## Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunday, 7.03.2021</td>
<td>Send list of preferred topics via eMail</td>
</tr>
<tr>
<td>Wednesday, 17.03.2021</td>
<td>Kick-off meeting and topic assignment</td>
</tr>
<tr>
<td></td>
<td>Read papers about your topic and search for additional literature</td>
</tr>
<tr>
<td></td>
<td>Prepare outline and argumentation line for the presentation</td>
</tr>
<tr>
<td>Until Friday 23.04.2021</td>
<td>Meet with your mentor to discuss your presentation</td>
</tr>
<tr>
<td></td>
<td>Prepare draft of your presentation</td>
</tr>
<tr>
<td>Until Sun. 17.05.2021</td>
<td>Send draft presentation to your mentor</td>
</tr>
<tr>
<td></td>
<td>Finalize your presentation</td>
</tr>
<tr>
<td>Friday, 28.05.2021</td>
<td>Presentation and discussion of your topic</td>
</tr>
<tr>
<td></td>
<td>(30 % of your final grade)</td>
</tr>
<tr>
<td></td>
<td>Write seminar thesis</td>
</tr>
<tr>
<td>Sunday, 28.06.2021</td>
<td>Submission of your seminar thesis (70 % of your final grade)</td>
</tr>
</tbody>
</table>
Formal Requirements

- **Presentation**
  - 15 minutes + 10 minutes discussion
  - should be 100% understandable for all participants

- **Written report (paper)**
  - 10-12 pages single column
    - including abstract and appendixes
    - not including bibliography
    - every additional page reduces your grade by 0.3
  - written in English language
  - use latex template of Springer Computer Science Proceedings

- **Final grade**
  - 70% written report
  - 30% presentation
Which template to use?

http://www.springer.com/de/it-informatik/lncs/conference-proceedings-guidelines
2. Seminar Topics
Topics

1. Table Search using Deep Learning (Giang, mentor: Alex)


Topics

2. Entity Matching using Deep Learning (Keti, mentor: Ralph)
3. Schema Matching using Deep Learning (Xuehui, mentor: Ralph)


Topics

4. Data Imputation using Deep Learning (Yifan, mentor: Ralph)


5. Deep Active Learning (Andreas, mentor: Anna)


3. How to Structure Your Paper / Presentation
Goal of Seminar Paper

- A seminar paper differs significantly from a master thesis
  - The topic is already defined
  - No need to implement or develop algorithms
  - No need to perform experiments
  - Primarily: reproduction and re-organization of content that is already available

- Goal of seminar paper
  - Describe the problem, describing several existing methods for handling the problem, comparing the methods and their evaluation using a systematic comparison schema
How to Structure Your Paper?

1. Introduction and Problem Statement
   • Which problem is addressed? Why is the problem important?
   • Structure of your paper

2. Description of Existing Approaches
   • Overview of existing methods and features used by the methods
   • Detailed description of two selected methods
   • Comparison of the selected methods using a set of comparison criteria

3. Evaluation
   • Comparison and discussion of the evaluation tasks, metrics
   • Comparison of the evaluation results

4. Conclusion
   • What did the comparison of the methods and evaluation results show?
   • Can something be concluded for future work?

5. Bibliography
Learn from Examples

- Read survey articles and identify the structure from the previous slide
  - Why can this paragraph be found at that position?
  - What is the purpose of some section / subsection?

- Important
  - Read survey articles!
  - Read conference or journal papers.

- Textbook on how to write a thesis

- University Library: Academic Writing Consultancy
  - https://www.bib.uni-mannheim.de/en/writing-consultancy/
  - Open consulting hour: every Wednesday 10 am - noon
Citing different Types of Publications

- Journal article
  - Good to cite, current research results
  - Survey articles (very good for an overview)
- Conference and workshop paper
  - Good to cite, current research results
- Books (sometimes cited)
  - Textbooks
  - Collections of articles/papers => Cite specific paper in book
- Websites
  - better not cited, exceptions are, e.g., W3C Specifications
  - Wikipedia is not an exception!!! Do not cite Wikipedia, ever!
- Slide sets
  - Never cite!
How to Find Relevant Publications?

- Use Standard Search Engines
- **Use Google Scholar**
  - we use it a lot ourselves
- Search Engines of the University’s library
  - see slides from the library course
- **Exploit references**: Given a relevant document $x$
  - Follow references in the past: papers $y$ that $x$ has cited
  - Follow references in the future: papers $y$ that cited $x$
    ("cited by" functionality in Google scholar)
4. Questions?