



**ifis**

Institut für Informationssysteme  
Technische Universität Braunschweig

**Information Retrieval &  
Web Search Engines**

---

## **HW: Language Models and Evaluation of IR Systems**

**Please note:** even though the homework assignments are optional, you are encouraged to answer them, as they will help you prepare for your final exam. Solutions can be dropped off at the institute's homework mailbox located on Informatikzentrum 2<sup>nd</sup> floor, next to room 238. Please do not forget your **Matrikelnummer** and your **full name** on your solutions. Another alternative is to send your homework by email to: [pinto@ifis.cs.tu-bs.de](mailto:pinto@ifis.cs.tu-bs.de). Please write your solutions in English. **Due date:** before next lecture (18.05.2017).

1. Explain what is the idea behind statistical language models?
2. What is the pooling method?
3. In language models one can estimate a model's probability parameters by the relative term frequencies as they appear in the respective document. What is the problem with this approach? And how can it be fixed?