

**HEIBRiDS Lecture Series – Wednesday 22<sup>nd</sup> May, 16.00 - 17.00**  
**@ Einstein Center Digital Future, RKF, Wilhelmstrasse 67**

**Location:** Room 104/105/106

**Speaker:** Helena Mihaljević, HTW-Berlin

### **Machine Learning approaches to authorship disambiguation**

Author Name Disambiguation (AND), and the more general problem of word sense disambiguation, is an open issue in computational linguistics and information retrieval, with particular relevance to digital libraries. The extraction of author entities is of interest in many cases, including optimization of search engines, the analysis of social networks or the evaluation of author impacts. Challenges such as name homography (author entities sharing the same name), name variability, incomplete information and misspellings, and the bare size of the current corpora make AND a difficult algorithmic problem that is actively researched through various approaches. In this talk, I will describe the challenges in more detail, showcase approaches based on supervised and unsupervised Machine Learning algorithms and discuss some open challenges.

**Next Lecture Series:** [Wednesday, June 5<sup>th</sup>](#)