

Universität des Saarlandes,  
Sonderforschungsbereich 1102

## Workshop

### “Data mining and its use and usability for linguistic analysis”

**Friday, 13<sup>th</sup> 2015, building A.2.2, R. 1.20**  
**(conference room)**

At this workshop, we aim to bring together people from **data mining, language modeling, linguistics** and **digital humanities** to further enhance collaborations between these fields.

#### PROGRAM

- 09:00-09:20 **Elke Teich and Stefania Degaetano-Ortlieb**, *Welcome and Introduction*
- 09:20-10:00 **Jon Dehdari**, *An Overview of Language Modeling and its Applications*
- 10:00-10:40 **Dietrich Klakow**, *Practical applications for language models*
- 10:40-11:10 *Coffee break*
- 11:10-11:50 **Peter Fankhauser**, *Observing Surprisal through the blurry lense of Language Models*
- 11:50-12:30 **Jilles Vreeken**, *Mining Sequential Patterns*
- 12:30-13:30 *Lunch*
- 13:30-14:10 **Steffen Koch**, *Trends and Topics: How visual approaches foster synergetic effects by combining linguistic analyses and interactive exploration*
- 14:10-14:50 **Stefan Evert**, *A multivariate approach to linguistic variation and distribution*
- 14:50-15:20 *Coffee break*
- 15:20-15:40 **Project B1**, *Information Density and Scientific Literacy in English*
- 15:40-16:00 **Project B3**, *Information Density and Fragments in German*
- 16:00-16:20 **Project C4**, *Modeling mutual intelligibility between Slavic languages*
- 16:20-17:30 **Discussion**

Please distribute to potentially interested colleagues.