

## **Neue Wege in der Nutzung von Korpora – Data-Mining für die textorientierten Geisteswissenschaften**

**Zeit und Ort:** Fr., 30.10.2015, Berlin Brandenburgische Akademie der Wissenschaften (BBAW), Einsteinsaal, Jägerstraße 22/23, 10117 Berlin

Tagungsprogramm (Stand 22.10.2015):

- 9:30 – 9:40 **Begrüßung und Eröffnung**
- 9:40 – 10:05 **Korpusbasierte Recherche und Analyse mit Hilfe von Data-Mining: Das BMBF-Projekt KobRA** (Prof. Dr. Angelika Storrer, Universität Mannheim)
- 10:05 – 10:30 **RapidMiner als Werkzeug für die textorientierten Geisteswissenschaften** (Prof. Dr. Katharina Morik, TU Dortmund)
- 10:30 – 11:00 Kaffeepause
- 11:00 – 11:30 **Routineaufgaben bei der Nutzung von Korpora: Disambiguieren, Klassifizieren, Annotieren mit KobRA-Verfahren** (Thomas Bartz/Christian Pölitz, TU Dortmund)
- 11:30 – 12:00 **Extraktion guter Belege aus Textkorpora durch Kombination eines regelbasierten Verfahrens mit maschinellem Lernen** (Dr. Alexander Geyken/Dr. Lothar Lemnitzer, BBAW)
- 12:00 – 12:30 **Analyse von sozialen Medien und internetbasierter Kommunikation** (PD Dr. Michael Beißwenger, TU Dortmund/Dr. Harald Lungen, IDS, Christian Pölitz, TU Dortmund)
- 12:30 – 14:00 Mittagspause
- 14:00 – 14:25 **Neue Entwicklungen und Wege bei der Erstellung, Erweiterung und Nutzung von Korpora am Zentrum Sprache** (Dr. Alexander Geyken/Bryan Jurish/Kay Michael Würzner, BBAW)
- 14:25 – 14:50 **Schnittstellen zur Nutzung der Korpusanalyseplattform KorAP** (Nils Diewald/Dr. Marc Kupietz, IDS)
- 14:50 – 15:15 **WebLicht Integration** (Prof. Dr. Erhard Hinrichs/Marie Hinrichs, SfS Universität Tübingen)
- 15:15 - 16:00 Kaffeepause, z.T. parallel zur Postersession
- 15:30 – 16:50 **Postersession und Systemdemonstrationen** (Liste siehe folgende Seite)
- 17:00 - 19:00 **Podiumsdiskussion: Digitale Textkorpora in den Geisteswissenschaften: Standortbestimmung und Perspektiven.**  
*Moderation:* Prof. Dr. Angelika Storrer  
*Podiumsgäste:* Prof. Dr. Thomas Gloning (Gießen), Prof. Dr. Erhard Hinrichs (Tübingen), Prof. Dr. Katharina Morik (Dortmund), Prof. Dr. Caroline Sporleder (Göttingen) und Prof. Dr. Holger Wormer (Dortmund)

**Postersession und Systemdemonstrationen** (Stand: 22.10.2015)

Thomas Bartz/Christian Pölitz/Angelika Storrer: *Korpus-basierte linguistische Recherche und Analyse mit Hilfe von Data-Mining („KobRA“). Erprobung innovativer Data-Mining-Verfahren für die empirische Arbeit mit strukturierten Sprachressourcen*

Michael Beißwenger/Thomas Bartz/Axel Herold/Marc Kupietz/Lothar Lemnitzer/Harald Lungen/Angelika Storrer: *Aufbau von Social-Media-Korpora für die Digital Humanities: Standards und Perspektiven*

Michael Beißwenger/Christian Pölitz/Lena Meyer: *KobRA @ Classroom: Ein interdisziplinäres Projektseminar zum Einsatz von Data-Mining bei der korpusgestützten Analyse internetbasierter Kommunikation*

Tassja Weber/Laura Herzberg: *Studienarbeiten zur internetbasierten Kommunikation mit KobRA-Verfahren an der Universität Mannheim*

Jörg Didakowski/Nadja Radtke: *Korpusbasierte Analyse von deutschen Stützverbgefügen*

Sebastian Buschjäger/Lukas Pfahler: *Discovering Subtle Word Relations in Large German Corpora*

Maria Ermakova: *Baby vs. Säugling: Wer gewinnt im XX. Jahrhundert? Eine korpusbasierte Studie zur Semantik und Pragmatik des Fremdwortgebrauchs*

Erik-Lân Do Dinh/Richard Eckart de Castilho/Iryna Gurevych: *In-Tool Learning for Selective Manual Annotation in Large Corpora* (Poster und Systemdemonstration)

Nils Diewald/Marc Kupietz: *Korpusanalyseplattform (KorAP)* (Poster und Systemdemonstration)

Erhard Hinrichs/Marie Hinrichs: *WebLicht* (Systemdemonstration)

Bryan Jurish: *Deutsches Textarchiv (DTA)* (Poster und Systemdemonstration)

Christian Pölitz: *RapidMiner als Werkzeug für die korpusbasierte Sprachanalyse* (Systemdemonstration)