

# Dagstuhl seminar: Challenges in Document Mining

## Preliminary Program

Day	Time	
Tu	9-12:15	Welcome, Talks
Tu	14-17:30	Talks
Tu	20-21:00	Demos
Wed	9-12:15	Talks
Wed	14-17:30	Working groups
Thu	9-12:15	Talks
Thu	14-17:30	Working groups
Fri	9-10:30	WG reports
Fri	10:45-12:15	Wrapup

Day	Time	Name	Title
Tu	09:15	Stefan M. Rueger	Knowledge Discovery in the web: potential, automation and limits
Tu	10:00	Hamish Cunningham	Full lifecycle information extraction and multiparadigm indexing with GATE
Tu	10:45	John A. Carroll	Unsupervised and Semi-Supervised Approaches to Cross-Domain Sentiment Classification
Tu	11:30	Marc Lechtenfeld	Determining the Polarity of Postings for Discussion Search
Tu	11:50	Thomas Gottron	Challenges in Mining Social Media: Sparsity and Quality
Tu	14:00	Steven Burrows	Data Mining for Civil Engineering Bridge Simulation: A Research Proposal
Tu	14:30	Ted Briscoe	Unsupervised entailment detection for IE query expansion
Tu	14:50	M. Ciaramita/H. Schütze	Piggyback: Using Search Engines for Robust Cross-Domain Named Entity Recognition
Tu	15:10	Michael Granitzer	Feedback Methods in Large Scale Visual Document Analysis
Tu	16:00	Norbert Fuhr	The Optimum Clustering Framework
Tu	16:45	Ingo Frommholz	A Quantum-inspired Polyrepresentation Framework
Tu	17:05	Tim Gollub	Data Mining Experiments with TIRA
We	09:00	Andreas Henrich	LFRP-Search: Multi-Layer Faceted Search with Ranking and Parallel Coordinates...
We	09:45	Alexander Loeser	Beyond Search: Interactive Web Analytics
We	10:45	Leif Azzopardi	The Retrievability of Documents
We	11:05	Gerhard Heyer	Looking at document collections from a bird's eye view: Exploratory search...
We	11:25	Harald Reiterer	Facet-Streams and Search-Tokens - Tangible User Interfaces for Information Seeking
We	11:45	Wolf Siberski	What do you mean? - Determining the Intent of Keyword Queries on Structured Data
Thu	09:00	Oliver Niggemann	A Probabilistic MajorClust Variant
Thu	09:20	Peter Prettenhofer	Learning cross-lingual representations from unlabeled data
Thu	09:40	Benno M. Stein	Exploiting unlabeled data for classification purposes
Thu	10:30	C.J. Keith van Rijsbergen	Document clustering revisited
Thu	10:50	Arjen P. de Vries	Search by Strategy
Thu	11:10	Dennis Hoppe	Challenges in Patent Retrieval and Mining